

High Throughput Methods in the Synthesis, Characterization, and Optimization of Porous Materials

Ivan G. Clayson, Daniel Hewitt, Martin Hutereau, Tom Pope, and Ben Slater*

Porous materials are widely employed in a large range of applications, in particular, for storage, separation, and catalysis of fine chemicals. Synthesis, characterization, and pre- and post-synthetic computer simulations are mostly carried out in a piecemeal and ad hoc manner. Whilst high throughput approaches have been used for more than 30 years in the porous material fields, routine integration of experimental and computational processes is only now becoming more established. Herein, important developments are highlighted and emerging challenges for the community identified, including the need to work toward more integrated workflows.

1. Introduction

One of the greatest contemporary challenges to targeted materials science is accelerating the time to realize functional materials with desired properties. The obstacles to achieving this goal are multifarious and include factors such as:

- Composition space is vast and hence full exploration of combinatorial phase space is intractable whilst blind exploration is inefficient.
- The financial and temporal costs of performing physical experiments limit what can be achieved on a practical time-scale.
- Analysis and characterization of physical experiments is onerous and potentially nonlinear with respect to the number of components.
- Identifying descriptors that capture the leading factors that influence the desired property can be elusive.
- Computational experiments that could complement laboratory experiments may not predict outcomes correctly.
- Establishing feedback loops between the outcome of physical experiments and computational predictions can be non-trivial.

I. G. Clayson, D. Hewitt, M. Hutereau, T. Pope, Prof. B. Slater
Department of Chemistry
University College London
20 Gower Street, London WC1E 6BT, UK
E-mail: b.slater@ucl.ac.uk

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adma.202002780>.

© 2020 The Authors. Published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adma.202002780

Notwithstanding these impediments, in many fields of materials science, solutions are being designed to mitigate hindrances to the efficient sampling of chemical space and improving the robustness of computational screening models through a combination of high throughput synthesis (HTS), characterization, and machine learning.

In this review, we highlight key developments in high throughput approaches pertinent to porous materials. Specifically, we focus on developments in the field of zeolitic materials, metal-organic frameworks (MOFs), and touch upon covalent organic frameworks (COFs). Although superficially these classes of materials have little in common, there are structural links such as topology which allow for the consideration of how high throughput methods contrast with one another. By contrasting developments in different fields, we identify some underutilized approaches which could be leveraged in the future. We target structurally ordered materials because the exploration of chemical and topological space is more straightforward to enumerate, though similar approaches could be applied to disordered materials.

The scale of the problem faced in materials science of targeted structure and function^[1] is by no means unique; in drug discovery, HTS and chemoinformatic approaches have been used for more than 40 years in an attempt to accelerate the discovery of druggable molecules.^[2,3] Unlike the drug discovery process, where Lipinski's "rule of 5"^[4] provides a reliable top level sift for viable targets, identifying the most promising material for a target application requires very particular properties that may be intertwined in complex and contradictory ways. For example, thermoelectrics require high electrical conductivity and low thermal conductivity and yet these properties are typically correlated unless they can be separated.^[5] Given the large scope of potential applications, the advent of the Materials Genome Initiative (MGI)^[6,7] in the United States has undoubtedly helped to promote ways to solve the aforementioned obstacles and numerous others. Similarly there are major initiatives within the EU (e.g., NOMAD^[8] and BIGmax^[9]) and Switzerland (MARVEL^[10]). In the MGI, nanoporous materials, including zeolites and MOFs, have been specifically targeted^[11] and in this review, we seek to highlight particular challenges in the field of porous materials and how researchers have sought to overcome these challenges. We focus primarily on the methods used in HTS and rapid throughput/screening computational approaches. Aspects of high throughput computation applied to porous materials have been reviewed before,^[12–14] as has high throughput experimentation (HTE)^[15,16] but here we focus on their combination within an integrated workflow.

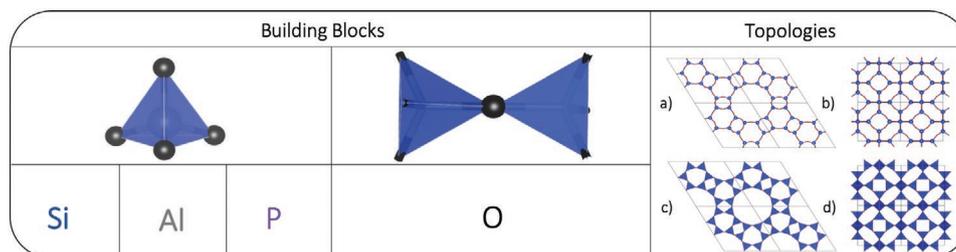


Figure 1. Building blocks for zeolites, separated into their constituent parts: tetrahedral atoms which can be Si, Al, or P, and their linker, a bridging oxygen atom. Two examples of topological nets for zeolites, GME and SOD, are shown in the Topologies column as a) and b), respectively, with their polyhedral representations depicted in c) and d). MOFs with the same topology, and their corresponding building blocks, are shown in Figure 9.

In particular, we highlight the growing trend for the virtuous circle where: 1) computer simulation screens are carried out to identify potential candidate materials for synthesis, 2) experiments are performed in a rapid manner, 3) the outcome of the experiments are relayed back to a simulation algorithm, to decide how to modify the experiment or stoichiometry and to improve the property or properties of choice and the cycle continues until some target condition is met. We feel this review is timely as robotic synthesis is becoming more routine, high performance desktop computing and supercomputing are more widely deployed, and the rise and extensive adoption of efficient machine learning approaches to screen materials, identify descriptors or principal components, and set synthetic targets is ever improving. Our aim is to highlight exemplars of major developments in approaches toward the realization of the virtuous circle that exploits the synergy between physical experiment, computational experiment, and the analysis and exploitation of data.

2. Experimental High Throughput Methods for Zeolite Synthesis and Characterization

Zeolites are most commonly identified as microporous framework silicates, composition SiO_2 , consisting of edge-sharing tetrahedra, where the silicon atom sits at the centre of a tetrahedron and oxygen atoms are at the vertices of the tetrahedra, as depicted in **Figure 1**. These materials typically contain channels and/or cages in one, two, or three dimensions with a pore size ranging from $\approx 4\text{--}15$ Å, allowing for the transport of small molecules within the structure. By far the greatest uses of zeolites are in the fields of petrochemistry and fine chemicals where they are used to separate oil fractions and catalyze the formation of chemical feedstocks. Although zeolites are most commonly realized as framework silicates there are also germanosilicates, aluminosilicates, aluminophosphates (AlPO) or silica-aluminophosphates (SAPO), borosilicates, and other variants and the composition and stoichiometry of these materials can vary dramatically. For a thorough introduction to the field, the reader is referred to Wright et al.^[17] In catalysis, zeolites typically have aluminosilicate composition with protons acting to supplement the charge of Al^{3+} to be equivalent to $\text{Si}^{4+}-(\text{Si}_{1-x}\text{Al}_x\text{H}_x)$. Zeolite synthesis is relatively difficult and can be time intensive as well as expensive^[18] due to:

- A tailored organic template also known as a structure directing agent (SDA) often being required to achieve a particular topology.

- The synthesis often being carried out at quite aggressive pH, initially often 10–11 rising to >14.
- The timescale of synthesis ranging from hours to several days.

Moreover, bespoke laboratory ware is used—bombs—to make use of autogenous pressure during the synthesis. Another complicating factor is that zeolites are kinetic products rather than thermodynamic products (e.g., all pure framework silicates are metastable with respect to quartz^[19,20]) which makes isolating phase pure samples particularly challenging. Indeed, the reproducibility of syntheses and optimization of synthetic conditions have been longstanding targets for the zeolite community.^[21] These complex synthetic challenges are generally not manifest in MOF chemistry, though there is some overlap in zeolitic imidazolate framework (ZIF) chemistry.^[22] Some zeolite phases have very narrow regions of stability and can only be made at particular ratio of Si:Al (also often reported as Si/Al). Exploration of zeolite phase space has, for more than 70 years, involved methodically traversing ternary phase maps, scanning the ratio of Si:Al:M (where M is an alkali metal) to identify stability fields for different zeolites. Naturally, this type of approach to synthesis lends itself to HTS methods and we now focus on major developments in this field.

2.1. Synthesis and Preparation

The array of samples, referred to as the sample library, in HTE should be preferably produced on as large a scale and as rapidly as possible in order to allow for the most expeditious exploration of phase space. To achieve this and minimize the time required to produce each individual sample, HTS must be employed; this section will discuss the implementation of HTS, examples of zeolite discovery through HTS, and finally the robotization of HTS.

2.1.1. Initial Method Development

Zeolitic sample libraries are formed through a systematic survey of either the gel compositional space or the SDA-phase relationship, with the former first being achieved with zeolitic materials by Akporiaye et al.^[23] Although HT methods for parallel protein synthesis had been used before,^[24] Akporiaye and co-workers explored the $\text{Na}_2\text{O}\text{--}\text{SiO}_2\text{--}\text{Al}_2\text{O}_3$ phase space for a $\text{Na}_2\text{O}\text{--}\text{SiO}_2\text{--}\text{Al}_2\text{O}_3\text{--}\text{H}_2\text{O}$ zeolite gel with a specially developed

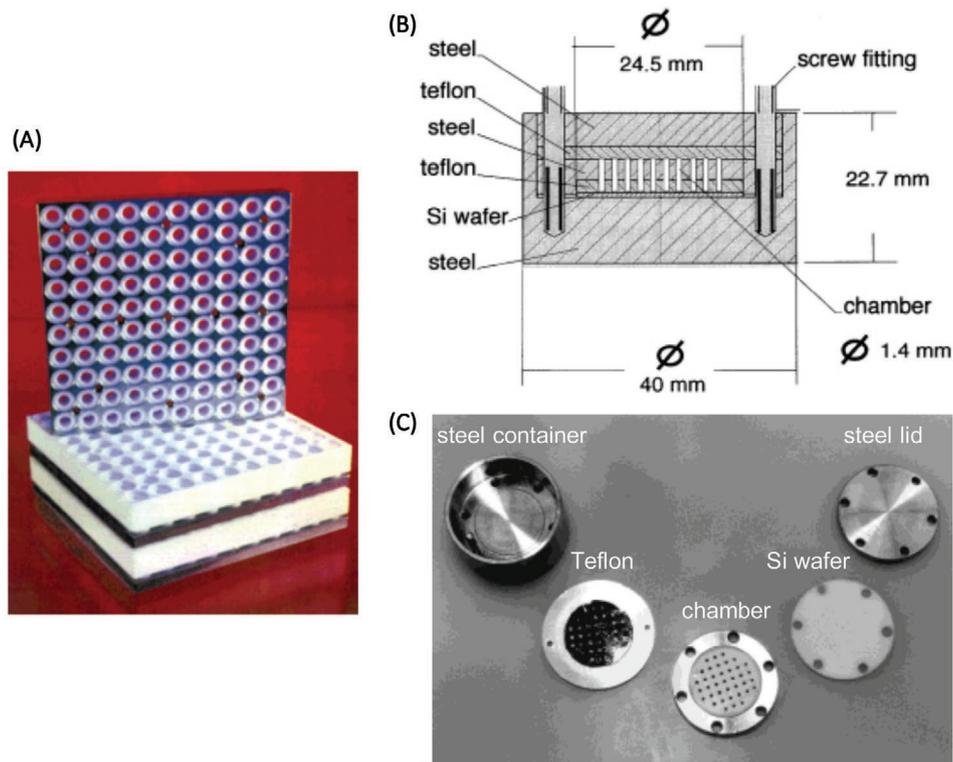


Figure 2. A) The HTS autoclaves produced by Akporiaye et al. in 1998 to explore the compositional space of an aluminosilicate gel.^[23] B) The cross-section diagram of the multi-autoclave employed by Klein et al.^[29] C) The related components prior to HTS. A) Adapted with permission.^[23] Copyright 1998, Wiley-VCH. B,C) Adapted with permission.^[29] Copyright 1998, Wiley-VCH.

multi-autoclave apparatus (Figure 2A).^[25,26] The phase diagram produced via this HT hydrothermal synthesis was in partial agreement with the previously reported phase diagram produced by Breck,^[27] with the authors attributing slight differences to variations in the water content in the gels formed by Breck. Notwithstanding the differences between the phase boundaries and the lack of formation of chabazite (CHA type)^[28] crystals observed by Akporiaye, the general mapping of a particular composition to a particular zeolite topology appeared to agree, implying that HTS with a multi-autoclave was valid. The authors' multi-autoclave enabled up to 1000 different gel compositions to be tested, saving several orders of magnitude of time with the time necessary for just hydrothermal synthesis being reduced from one and a half years to only 6 days. By exploiting the parallel synthesis technique, Akporiaye et al. explored the phase diagram of a more complex gel system containing not only the aforementioned compounds, but additionally some proportion of MeNH_3^+ , Li^+ , and/or Cs^+ . Thus Akporiaye et al. demonstrated that multi-autoclaves could be used to sample vast regions of compositional space with the authors noting that the development of such HT methods would enable novel catalyst identification with greater efficiency and speed.

Akporiaye et al.^[30] and others^[31,32] further demonstrated the transferability of this mass hydrothermal synthesis technique to AlPOs and other heteroatom zeolites. The Maier group notably developed an alternative multi-autoclave that enabled coarse X-ray diffraction (XRD) spectra of the samples to be subsequently taken, thus combining synthesis and characterization

in a zeolite HT workflow for the first time in the academic literature (Figure 2B,C where further details are discussed in Section 2.2.1).^[29] The microgram amounts of Ti-containing silicate TS-1 (MFI type) synthesized and then examined by Klein et al.^[29] additionally demonstrated the possible prototyping compatibilities of HTS to exhaustively sample regions of interest without prohibitive time or monetary investment. Other HTSs besides hydrothermal synthesis were later devised with Zhang et al. developing an HT vapor-phase transport synthesis method for SAPOs where the dry gel composition-SAPO topology relationship was subsequently studied.^[33]

Following the pioneering work by these groups and others,^[23,29,31,34–36] Newsam et al.^[37] noted the particular difficulty with zeolites in specifying which region of the compositional space to investigate. Even nearly two decades on, the translation between in silico zeolite design to a rigorous synthesis procedure is still difficult^[38,39] as a reproducible one-to-one mapping of gel composition to a particular zeolite topology and composition remains elusive. Slight experimental variation can lead to significant and non-trivial effects on the resultant zeolite topology and composition, with the relationship between the contents of the gel and the final crystallized product being difficult to establish. To further complicate matters, not all points on the compositional phase diagram relate to crystalline products with amorphous, non-porous, and/or mixed-phase products populating many regions. Thus, faster explorations of unknown phase space are necessary and accordingly automation has been employed to help streamline the HTS process.

Choi et al.^[31] were the first authors to apply robotics to support this multi-autoclave approach in order to increase the HT capabilities, achieved through the automation of liquid reagent dispensing. These authors examined the resultant AlPO topology while varying the proportion of the Co(III) dicyclopentadienyl and (*i*-Pr)₂NH SDAs present in the synthesis gel; AlPO-5 (AFI type) was found to only crystallize if at least 40% of the SDA (*i*-Pr)₂NH was present, demonstrating the importance of SDAs as variables in HT phase explorations. Reagent dispensing robots were subsequently used to study other zeolites^[40–44] and crucially for zeolite discovery where various members of the UZM^[45,46] and ITQ^[47–52] families are notable examples.

2.1.2. Zeolite Discovery through High Throughput Synthesis

The methodology for these HT discovery studies was succinctly summarized by Stock^[53] as: i) define the phase space to be explored and design experiments accordingly, ii) generate a selection of reaction mixtures produced systematically via robotic arms, iii) attempt synthesis with the caveat that not all elements of the library will produce single-phase crystalline products, iv) isolate and separate, and finally v) characterize and/or analyze samples. The work by Corma and colleagues particularly illustrate this workflow with the ITQ zeolites where the synthesis was largely automated through the use of a robotic arm that composed the gel solutions which formed the sample library.

When targeting novel ITQ-24 (IWR type) compositions,^[47] the authors first noted that the larger Si–O–Ge bond angles found in silicogermanate ITQ-24 were of a similar size to that of Si–O–B bonds. Therefore they hypothesized that beyond a given B₂O₃ gel composition, B should be able to occupy all the Ge sites in silicogermanate ITQ-24 and thus enable a Ge-free borosilicate ITQ-24 zeolite to crystallize. A HT search of various SiO₂–GeO₂–B₂O₃ gel compositions was performed in order to determine the proportion of B₂O₃ necessary for the authors to successfully synthesize borosilicate ITQ-24.^[47] As trivalent B will impart a negative charge onto the framework, a pure silica ITQ-24 analog would require a lower charge density in the SDA to reduce the number of charged defects being included during crystallization. Therefore, a variety of possible organic SDA (OSDA) candidates were trialled and one was found that produced ITQ-24 with Si/Ge ratios from 2 to infinity (i.e., pure silica), where the a pure silica variant of ITQ-24 had previously been unobtainable synthetically.

Screening a variety of possible OSDA candidates with HTS has been performed by many authors^[54–56] as the SDA is an additional degree of freedom when exploring new possible topologies or compositions. A novel example of this is the discovery of the exceptionally low density ITQ-37 zeolite (ITV type).^[50] This silicogermanate has an unusual topology due to the large elliptical 30-member rings and periodic framework interruptions. The framework interruption sites, preferentially occupied by Ge, are terminated with hydroxyls, resulting in T–OH bonds being inherent to the internal structure as opposed to predominantly arising from post-synthetic treatments and defects. Additionally, the novel zeotype, crystallizable as both

the silicogermanate and aluminosilicogermanate form, is chiral due to the gyroidal channel formed from the 30-membered rings. Whilst the Corma group had previously selected a particular enantiomer of zeolite beta^[57] through a chiral OSDA,^[58] the study by Sun et al.^[50] is the first example of zeolite discovery by considering a particular OSDA optical isomer, implying that the chirality of an OSDA is a further degree of freedom available for interrogation. ITQ-37 was found to be stable after calcination at 813 K for up to 2 weeks, where crystallinity is comprised afterward, and had triple the initial activity toward bulky aldehydes for catalytic acetalization when compared with the common zeolite H-beta, demonstrating the possible utility for near-mesoporous and chiral zeolites for industry.

Moreover, a novel large-pore silicogermanate zeolite ITQ-33 (ITT type) with 10- and 18-membered ring channels was discovered by the Corma group through varying the gel composition as well as the concentration of mineralizing F[−] present.^[49] Whilst the process for discovering this zeolite is similar to prior examples,^[47,50] the distinctive feature of this synthesis is the successful targeting of a large-pore zeolite by exploiting the ability of F[−] to stabilize double 4-membered ring motifs common in large-pore frameworks.^[47,59] The corresponding large-pore ITQ-33 found by Corma and colleagues had a large Brunauer–Emmett–Teller (BET) surface area and superior hydrocarbon cracking capabilities than other zeolites. This study further exemplifies how HTS, directed by chemical intuition and prior knowledge, either a posteriori or a priori, can be employed to target specific types of catalysts or physical features.

2.1.3. Implementation of Robotics and Automation

The next major innovation in HTS was proposed by Caremans et al.^[60] who developed an “all-in-one” automated synthesis and analysis system, with the only manual input being the supply of reagents to the robotic dispenser (Figure 3A) and the moving of filtrate to the ovens. The all-encompassing automated process described by Caremans et al. demonstrated no contamination and was successfully employed in synthesizing zeolite-2,^[61] a microporous analog of the amorphous-walled mesoporous MCM-48 material,^[62–66] and clathrasil phases, pure silica framework materials where only small molecules are able to pass between cages.^[67]

Whilst the samples produced by Caremans et al.^[60] were less crystalline than the reference material,^[61] evidenced by the broader peaks in the XRD spectrum, they nonetheless demonstrated that porous silica materials could be produced and analyzed with minimal human input. Janssen et al.^[68] further developed the “all-in-one” automated HT research system proposed by Caremans and colleagues,^[60] with Janssen et al. developing an automatic workflow for the: i) dispensing of pre-formed zeolite (Figure 3C), ii) pre-treatment of zeolites (Figure 3D), and iii) catalytic tests on modified zeolites. Whilst an impressive feature of engineering, the workflow deployed by Janssen et al. did not perform any zeolite synthesis and used pre-made commercial zeolites, falling short of a completely autonomous “start-to-end” research station. In the subsequent year, Serna et al.^[69] produced a more advanced complete workflow for synthesizing, testing and examining Ti-grafted MCM-41 and ITQ-2 for large olefins

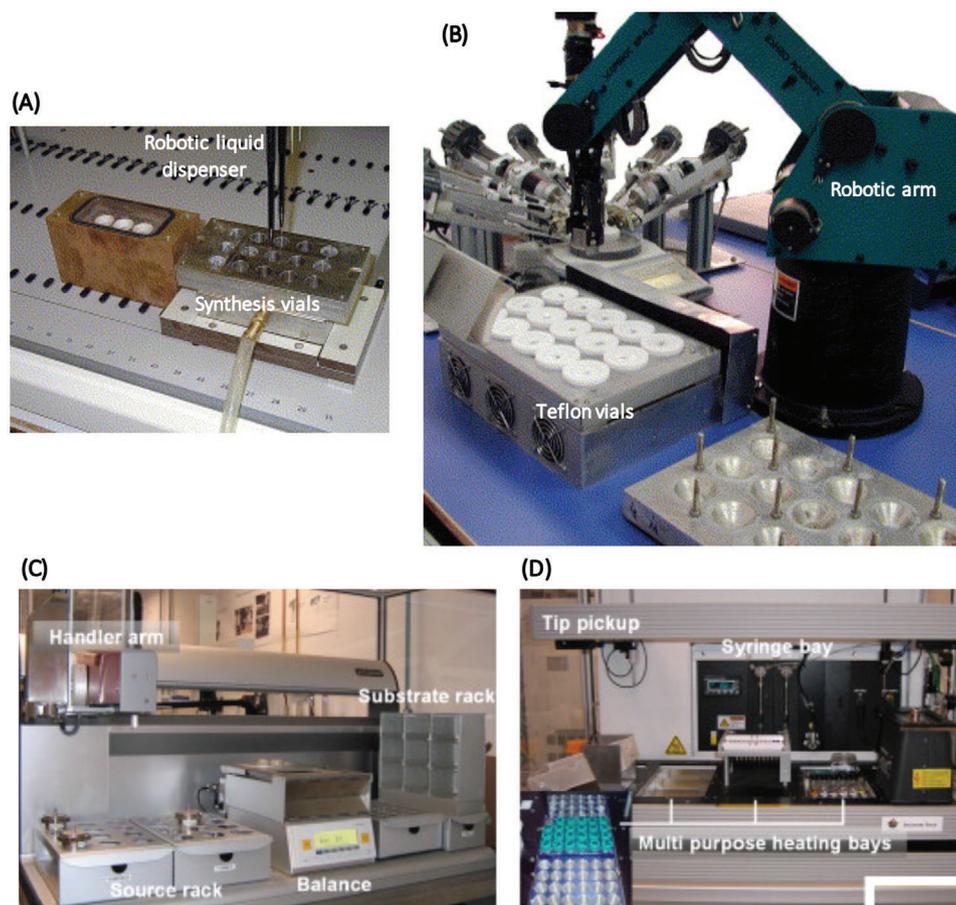


Figure 3. The robotic apparatus employed in the literature: A) the liquid dispensing arm utilized by Caremans et al. to dispense zeolite precursor solution into vials prior to hydrothermal synthesis,^[60] B) the robotic arm employed by Moliner et al. to weigh solid reagents and dispense zeolite precursor into the Teflon vials for hydrothermal synthesis,^[42] C) the solid zeolite dispensing station, and D) the automated ion-exchange solution dispensing used by Janssen et al.^[68] A) Adapted with permission.^[60] Copyright 2006, Elsevier. B) Adapted with permission.^[42] Copyright 2005, Elsevier. C,D) Adapted with permission.^[68] Copyright 2007, Elsevier.

and methyl oleate epoxidation. The HT platform developed by these authors: i) dispersed the reagents for the post-synthetic Ti-grafting and silylation, ii) performed HT XRD, and iii) measured catalytic activity with gas chromatography. However as with the work by Janssen and colleagues,^[68] the HT process developed by Serna et al.^[69] lacked any HT initial porous material synthesis and required human input to move samples from one HT station to another. These studies demonstrate the difficulty in creating a single HT experimental station and thus starting in the mid-2000s, the community moved toward a combined HTS and data-mining approach with the authors Moliner et al.^[42] constructing an artificial neural network machine learning (ML) model based on experimental data to aid in the analysis. Others subsequently employed various ML models to analyze synthesis data^[70–72] though initially the use of these approaches was limited to examining the pre-collected experimental data and the models were not used for predictions outside of the range of the training set. Later ML models have been successfully employed for novel predictions that have been experimentally verified where further details can be seen in Section 3.4. Whilst other alternative high throughput production methods were developed,^[73–75] the majority of new innovations are limited to HT structural characterization and activity or HT post-synthetic treatments.^[44,76–83]

Notably, there is a lack of development of “full” HT library generation workflows, with current state-of-the-art systems only automating a few portions of the synthesis process or lack robotization in important areas, such as zeolite synthesis or post-synthetic treatment. Also, the automated workstations are often isolated from other stations with manual, as opposed to more efficient robotic, sample transport being required.

2.2. Characterization

The characteristics of a zeolite with respect to its structure, in terms of crystallinity and topological features, and chemistry, in terms of the environments present on and within the material, determine its efficacy and viability. This section will describe both the physical and chemical descriptors that have been employed to investigate zeolitic samples in a HT manner.

2.2.1. Structural Proprieties

Since zeolites are crystalline materials, XRD was the first method applied to characterize the crystallinity of the

zeolites produced from HTS by the pioneering authors Akporiaye et al.^[23] and Klein et al.^[29] Due to the design of the parallel multi-autoclave, Akporiaye et al.^[23] were required to transfer the products to another container before performing analysis. However, the multi-reactor developed by Klein et al.^[29] (Figure 2B,C) contained segregated chambers in the autoclave with a Si wafer placed on the base. Porous rods were then inserted into the microchambers after the hydrothermal synthesis treatment to separate the liquid and solid phases, where further drying was done to the zeolite solid left on the wafer. The zeolite crystals were then sequentially scanned by a focused X-ray beam. Due to the set-up devised by Klein et al., the final products from the HTS were already present on an XRD stage and therefore could be taken for XRD analysis without further transfer or modification. This enabled the spatial component during the initial gel formation to be preserved and thus allow for simple mapping of diffraction spectrum to gel composition.

The use of a multi-sample XRD stage, often automated with movement in at least the *xy* plane or similar, became standard into the new millennium.^[31,42,45,46,48,60,69,72] Baumes et al. further innovated the XRD analysis procedure in 2009 through the development of the adaptable time warping (ATW) technique,^[84] an ML spectral assignment method which was significantly more reliable in assigning unknown and multi-phase spectra than traditional and other ML methods. The advent of this new method by the Corma group was crucial for screening large numbers of XRD patterns from HTS when scanning compositional space, accelerating the zeolite discovery process.

Compositional imaging of zeolitic materials in a non-destructive manner was first introduced by Koster et al.^[85] where the authors imaged a AgNa-Y zeolite (FAU type) using 3D transmission electron microscopy (TEM). By rotating the sample 143° and recording the TEM image, a reconstructed 3D representation of the crystal could be made whilst recording the composition of the surface. However, this approach is not appropriate for HT and would become a bottleneck process since the sample would need to be carefully rotated through a given angle interval. Despite developments in automated 3D-TEM imaging^[86] and the predictions that this automation could be exploited for HT,^[87] 3D-TEM for porous materials is still to be used in an effective HT workflow. If a non-tomographic imaging method is used, then compositional imaging can be employed in a HT manner through micro X-ray fluorescence (MXRF). MXRF involves focusing an X-ray beam to a small cross-section on the order of μm^2 and illuminating a small section of a multi-sample matrix. The emitted fluorescence is then recorded for a particular sample element of the array and after a given time, the sample array is translated in the *xy* direction to have a different sample exposed to the fluorescing beam, resulting in the presence of $Z \geq 11$ elements within the library being recorded.^[88] Minogue and colleagues utilized this MXRF to perform an HT screening of zeolites for their propensity and selectivity to Cs uptake from a solution of Cs and other metals in order to determine the optimal sieve for removing Cs from nuclear waste and other solutions containing a large number of cations.^[89] Through the use of MXRF, 11 zeolites were screened for Cs⁺ uptake with the CHA type SAPO-34 and other small-pore zeolites having the greatest

selectivity to Cs. The authors particularly noted that the HT screening took only a few hours as opposed to weeks, with such a speed-up being crucial for exploring a wider range of samples and reducing labor for each experiment.

The utility of porous materials relates to their high surface area, where screening various topologies and treated zeolites for their accessible surface area (ASA) is crucial for determining their efficacy. Groen et al. developed a complete HT post-synthetic workflow followed by a HT BET surface analysis for various ZSM-5 zeolites (MFI type) in order to investigate post-synthetic treatments that maximize mesoporosity without framework degradation.^[76] The authors employed a parallel post-synthetic approach with a library of pre-formed zeolites with the desilication treatment variables being the: i) concentration of HF acid or hydroxide base, ii) counter-cation for the hydroxide base, iii) temperature, iv) duration, v) stirring speed, and vi) initial framework Si/Al ratio. The library of samples was manually generated and then placed in a workstation containing isolated cells that allowed for separate temperature and stirring speed control. The novel aspect of this work was the characterization of the samples via N₂ adsorption isotherms conducted in parallel on porous materials for the first time in the academic literature. Si was found to be more readily extracted from the framework compared with Al, with the mesoporosity of the samples increasing as the Si/Al increased. However, the zeolites with the greatest Si content, Si/Al > 50, had a lower mesoporous surface area than the samples with an Si/Al between 25 and 50. Groen et al. attributed this counter-intuitive observation to the assistance of Al in the desilication mechanism reported by previous authors.^[90] Another notable observation was how the extent of desilication varied with the identity of the alkali metal present in the base; a comparison of Li, Na, and K, showed NaOH yielded the greatest mesoporosity. These authors further improved their HT post-synthetic treatment method in subsequent papers through the use of robotic syringes to disperse the reagents,^[78,91] greatly improving the possible speed-up afforded by post-synthetic HT methods.

2.2.2. Chemical Properties

A powerful technique to determine the internal chemistry of zeolites is Fourier-transform infrared (FTIR) spectroscopy, which can be used to probe the character of bonds present in the system. Snively et al.^[92–95] first developed a 7-chamber parallel FTIR spectrometer that, whilst predominantly employed for metal-alumina supported catalysts,^[93–95] was shown to be transferable to zeolites.^[92,93] However, the main appeal of the device was the increased high-quality spectrum acquisition speed as the apparatus collected seven spectra simultaneously (averaged to produce a final single spectrum) from the same sample, rather than collecting multiple spectra stepwise.^[92] The sample chambers could contain different samples,^[93] but often the chamber's contents were duplicated so as to reduce the measurement time.

The Schüth group further developed this parallelized FTIR method shortly afterward^[96,97] but whilst Lauterbach and colleagues initially focused on increased acquisition speed, Schüth and co-workers focused on producing a device that

enabled the analysis of a larger number of samples. This was achieved through a widening of the beam to generate a parallel beam that simultaneously shone on multiple samples. After transmitting through a sample, a portion of the resultant beam was re-focused on to a designated block of the detector array assigned to that particular sample. This resulted in 8 IR spectra being measured in parallel with the authors further improving the apparatus to record 49 spectra.^[96] Kubanek et al.^[98] employed this new HT characterization method to investigate six Pt-ZSM-5 catalysts with three different metal loadings and two different post-calcination thermal treatments at different temperatures. CO was passed over the metal zeolites and the behavior of the Pt-CO adsorption peaks was used to determine the Pt speciation, with Kubanek et al.^[98] stating that high loading and temperature treatment causes Pt to aggregate within the zeolite crystal into larger nanoparticles.

Since zeolites are solid acid catalysts, it is desirable to characterize them in terms of acidity. Temperature-programmed desorption of NH₃ (NH₃-TPD) enables a variety of differing acid sites within a sample to be enumerated through the number of NH₃ desorption peaks in a temperature spectrum where the strength of an acid site is related to the temperature required for bound NH₃ to be released. Wang et al.^[99] used a multi-stream mass spectrometer (MS) developed earlier^[100] to record the NH₃ released after saturation by following the intensity of the $m/e = 16$ mass fragment as a function of temperature. The multi-stream MS involved a parallel 80-channel reactor system where gas feeds were sequentially selected to pass through a quadrupole MS in rapid succession, enabling the NH₃-TPD of various zeolites of differing compositions and metal loadings to be measured in a HT manner.^[100] Woltz et al.^[101] produced a similar 6-channel reactor system with a MS attached where, though the use of sequential mass spectroscopy, a HT NH₃-TPD was taken for Pt-loaded H-beta zeolites that were post-synthetically sulfated with H₂S followed by oxidation. As with Wang et al.,^[99] Woltz and colleagues sampled the mass peaks of NH₃ from the post-catalyst gas streams consecutively where the intensity of these peaks was employed to determine the acidity. The Pt/H-beta catalysts were found to be bifunctional catalysts with Brønsted zeolite and sulfate acid sites that enhanced the hydroisomerization of pentane. Another HT method to determine the acidity of a zeolite system was developed by Fischer and colleagues where propene was introduced into various H-ZSM-5 aluminosilicate zeolites of differing Si/Al ratios and the near-edge X-ray absorption fine-structure (NEXAFS) spectra were recorded at a synchrotron.^[102] Whilst characterizing a zeolite with synchrotron radiation was not unprecedented, Fischer et al. expanded traditional NEXAFS spectroscopy into a combinatorial form through the collection of X-ray absorption spectra sequentially with the sample array being translated in the xz plane, granting an additional spatial dimension to the NEXAFS spectrum and thus enabling different samples to be measured without superimposing the spectra. The authors utilized the intensity of the carbon K-edge peak as an indicator for the amount of propylene retained, following evacuation, on the zeolite after saturation. Unsurprisingly, increasing the Al content resulted in a greater amount of hydrocarbon absorbed. Therefore, Fischer et al. additionally recorded the normalized C core 1s $\rightarrow \pi^*$ C=C bond excitation intensity for the chemisorbed

propylene. This absorption intensity is proportional to the amount of reactive propene available on the zeolite and thus the authors used this as a measurement of the acidity for a single site; it was found that the acidity for a single Brønsted site increased with decreasing Al content until Si/Al = 150, when it roughly plateaus. Recording specific excitation intensities can allow for the probing of particular interactions, such as the formation of active species as the authors discussed, that other methods cannot indicate in an unambiguous way. Therefore the development of such a method, whilst dependent on substantive knowledge of the system in question, into a HT workflow could allow for the rapid screen of catalysts with respect to reactive intermediate formation and thus indirectly screen based on mechanistic characteristics.

The Kaskel group developed an alternative acidity measurement procedure, initially for MOFs (see Section 4.2) and then for zeolites in 2012, using the instruments InfraSORB^[103,104] and later InfraSORP.^[105] These two pieces of apparatus were based on measuring the thermal response of a gaseous adsorbate on a sample with an IR sensor and relating the magnitude of the heat released from the adsorption to the acidity. Multiple sample cells with separate sensors were then linked together in an array via a shared gas inlet and outlet channel to enable parallel acidity measurements.

The authors Bae et al.^[106] examined the absorption of CO₂ in alkali-metal-exchanged LTA and FAU typed zeolites through the use of a multi-channel gas absorption analyzer where each sample chamber contained an independent pressure sensor. Bae et al. reported that Ca-A had an affinity for CO₂ over N₂. Further gas separation and adsorption work in the academic literature regarding zeolites has been performed primarily with HT computational screens of experimentally realized and hypothetical frameworks (see Section 3.3 for further details) as opposed to physical experiments. Additionally, academic research is now concerned more with larger-pore and less dense materials for gas sorption, most notably MOFs (see Section 5.4.5). Some authors have utilized gas adsorption for other types of characterization indirectly though, as with Fischer et al. where they studied the change in water pressure of separate isolated reservoirs connected to a set of zeolites as a function of temperature;^[107,108] the resultant isochores were used to determine the hydrothermal stability of the zeolites where the parallel nature of the measurements enabled the authors to record the stability in a HT manner. Experimental HT academic studies such as these are more common to gas adsorption studies performed on MOFs, where Section 4.2 contains further details. Zeolites have been employed industrially as gas separators for decades;^[109,110] however, the majority of recent academic research has been primarily concerned with other materials with greater void volumes.

2.3. Catalytic Activity

As mentioned in Section 1, zeolites are frequently employed as catalysts in the fine chemical and petrochemical industry where their efficacy and stability are of vital importance. In order to maximize the impact of HTS, the largest possible array of samples should be tested contemporaneously so as

to minimize the temporal costs of catalyst evaluation. Therefore, HT catalytic testing is desirable for evaluating the sample library with this speed-up being particularly achieved when these HT activity measurements are coupled with the methods described in Sections 2.1 and 2.2. This section will introduce and discuss the HT variants of the traditional gas chromatography/ mass spectroscopy methods use to measure activity as well as non-destructive HT evaluation procedures based on photon absorption.

2.3.1. Early Methods of Catalytic Testing

Microflow reactors were used far before zeolite HTS became prevalent,^[16,111,112] where the first zeolite HT academic study was published by Creer et al.^[112] in 1986. Six different catalysts were tested concurrently with up to four different gaseous reagents being available for the catalytic testing during a single run. In the six separate systems, one could vary the catalysts present and maintain the same gas flow or have differing gas composition streams with the catalyst remaining fixed, enabling a vast number of degrees of freedom for exploring the catalytic search space. The products from the reaction were then analyzed with gas chromatography (GC). Whilst there were issues with regards to reproducibility, attributed to non-uniform temperature between the individual reactors, the work by Creer et al.^[112] was clearly pioneering as microflow reactors became the standard method by which catalytic activity from zeolites was recorded. Using the system developed by Creer et al.,^[112] Bessell and Seddon examined the formation of larger hydrocarbons from olefins over H-ZSM-5.^[113] The authors utilized an improved version of the parallel microreactor-GC system where a quadrupole MS was attached to increase the analytical capacity of the system, as the hydrocarbons had similar GC retention times.^[113] Bessell et al. found that as the temperature increased, the formation of aromatic products became more favorable and dominated at the expense of paraffin formation; increasing the Al content increased the cracking activity, rationalized as due to an increase in acidity. The use of GC in HT workflows to evaluate activity is ubiquitous and used for a variety of applications (Table 1).

Mass spectroscopy has been employed to measure zeolite activity since 1987.^[113] However, this was only for a small number of samples with authors preferring to initially analyze products from zeolite catalysis with GC. Whilst non-porous supports with metals loaded on them had been employing mass spectroscopy to measure activity in a HT manner since 1990,^[161,162] it was not until 2003 that Wang et al.^[100] developed an HT catalytic testing system incorporating a mass spectrometer (MS) which was used for zeolites. Wang and colleagues developed an 80-channel reactor system where the reaction effluent was sequentially but rapidly sampled. Whilst the MS itself was not parallel, the multi-stream MS could be used for recording the activity of a large number of catalysts in the same run with the throughput of the system reaching 400 catalysts per week. The authors examined the efficacy of zeolite, alumina, and silica for the catalytic condensation of acetone, where measuring the relative intensity of the product mass peak compared to Ar allowed Wang et al. to record the conversion

Table 1. The wide range of possible zeolite applications that have been successfully investigated through the use of high throughput parallelized gas chromatography.

High throughput gas chromatography application		References
Petrochemical	Cracking	[112,114–130]
	Isomerization	[101,112,122–124,130–134]
	Small molecule formation	[68,113,135–138]
Oxidation	Ammoxidation	[139]
	CO oxidation	[140,141]
	Controlled Combustion	[142]
	Epoxidation	[69,143,144]
Other industrially relevant processes	Adsorption	[83,145]
	N ₂ O abatement	[146–154]
	Coupling reactions	[155–158]
	Lactate formation	[159,160]

of acetone and the selectivity toward particular products for each catalyst (Figure 4); the H-beta and H-Y zeolites were noted to consume large amounts of acetone but had little selectivity toward any particular product. The multi-stream MS was later employed for further catalysis measurements of various metal-docked H-ZSM-5 zeolites to the conversion of CO and methane to ethene and aromatic compounds^[163] as well for a HT NH₃-TPD method (see Section 2.2.2).^[99] Other authors later developed similar reactors that allowed for rapid sequential analysis^[120,164–166] and beyond the porous materials field, the authors Casacuberta et al.^[167] recently employed a miniaturized ¹⁴C mass spectrometry-based analyzer system^[168] to sequentially determine the concentration of CO₂ present in seven seawater samples, where this latter example is of particular note due to the smaller size of the MS enabling a multi-instrument HT workflow. However a parallel, concurrent MS for academic use is yet to materialize and be used for zeolite catalysts. This inevitably results in the number of samples and the time intervals that products can be analyzed being limited. When using an HT MS, this inverse relationship causes an elongation of the time between two products analysis points as the number of effluent streams increases, adversely affecting the temporal component of the activity measurement. Even if the temporal component is not desired, the shorter timescale for spectral acquisition would result in a greater uncertainty in the intensity measurements and would require more careful experimental planning to ensure the reagent exposure time before measurement is consistent.

In 2002, Ozturk and Senkan performed an extensive HT search of catalytic metals on alumina and zeolitic supports with over 1500 different samples being screened for their NO reduction capacities under fuel-lean, or air-rich, conditions.^[169] This wide survey of possible catalysts involved evaluating the resultant gas products from the microreactors via MS.^[170] This exhaustive evaluation not only enabled directed comparisons between a large range of catalysts, but the systematic approach taken discovered a new bimetallic CuOs-13X (FAU type) catalyst that was superior to the state-of-the-art catalyst at the time, Cu-ZSM-5.^[171,172] The non-obvious CuOs-13X catalyst was

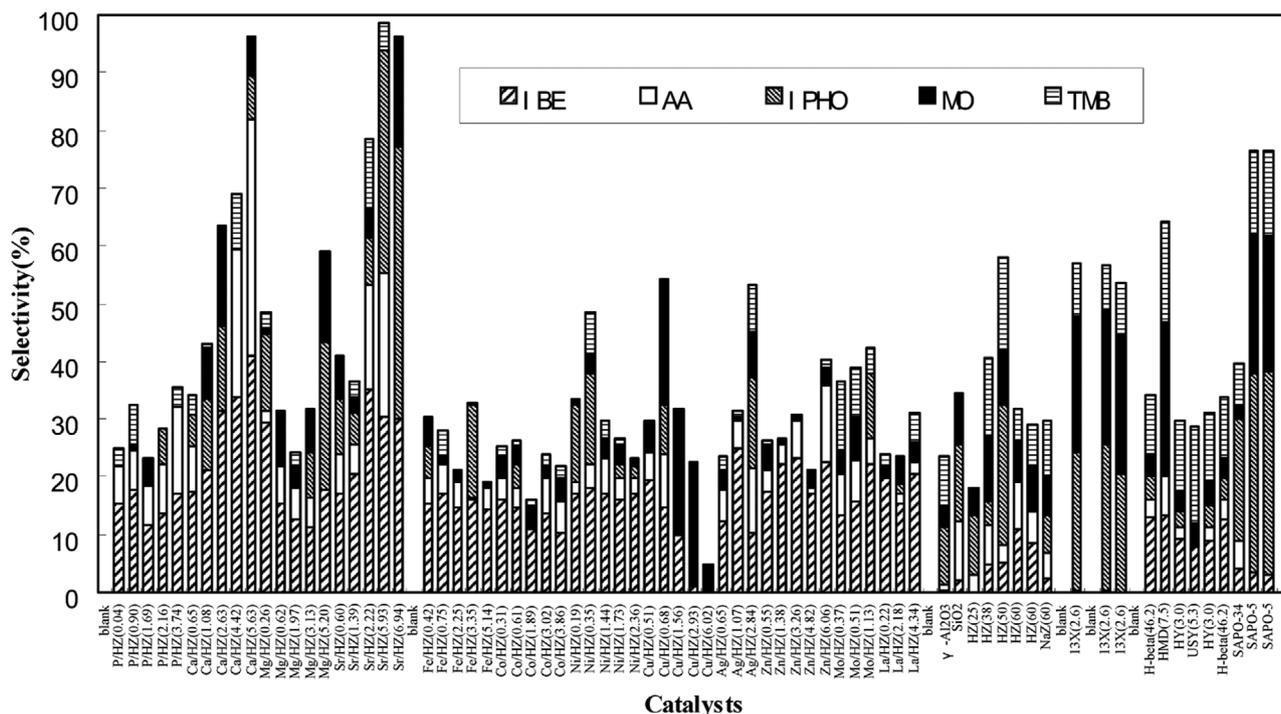


Figure 4. The product distribution of recorded products for the catalytic condensation of acetone over the sample library investigated by Wang et al.^[100] using the HT multi-stream MS developed by the authors. The products recorded were: isobutene (I BE), acetic acid (AA), isophorone (I PHO), mesityl oxide (MO), and mesitylene (TMB) (detailed reaction scheme and product structures can be found within ref. [100]). Reproduced with permission.^[100] Copyright 2003, American Chemical Society.

formed through sequential ion-exchange procedures, first with a $\text{Cu}(\text{NO}_3)_2$ solution and then with a $(\text{NH}_4)_2\text{OsCl}_6$ solution for 2 days each where CuOs-13X displayed greater activity to NO abatement for all temperatures and maintained significant activity over a wider temperature region than Cu-ZSM-5 . Moreover, CuOs-13X retained a majority of its activity after steam aging for 36 h at 450 °C. The discovery of this more effective catalyst clearly demonstrated the value of HT methods as the mixed-metal nature of the catalyst, with 10% and 1% exchanged Cu and Os, would have resulted in its discovery only through serendipity or a search of the composition phase space, with the latter being possible through HT testing due to the large number of candidates that can be evaluated. Many authors following Ozturk and Senkan applied a similar HT catalytic testing procedure to find optimal catalysts for a variety of predominantly petrochemical reactions.^[173–176]

Whilst the implementations of GC and mass spectroscopy above record the final products of a reaction, an alternative way of interrogating a catalyst and particularly the mechanism is through the use of temporal analysis of products (TAP), developed by Gleaves and colleagues in 1988.^[177] This method involves a very short gas pulse being injected into the reactor and then rapidly evacuated on the order of milliseconds. The post-reactor gas stream is then sampled by an MS, for example, to evaluate the intermediate products formed just after contact with the surface. The utility of TAP arises from its capability to analyze intermediate products formed during the initial contact stage where the formation of such products can give insight in the mechanism and pathway by which catalysis occurs. Van

Veen et al. were the first to produce a parallel reactor with 12 channels that allow for concurrent TAP experiments where the authors investigated the absorption enthalpy of ethane on 11 different zeolite samples simultaneously.^[178] The reported values were acknowledged by the authors to be an underestimation of the previous literature values, though they justify the discrepancy by stating that their rapid screen measurements could be partially affected by the differing diffusion rate of adsorbates within the crystals, where these difference would be significant given the small time window of the measurements. This does suggest that careful calibration of pulse length is necessary to ensure that the results are not governed by mass transport effects, which will be of greater importance when comparing porous materials with differing channel and pore sizes.

2.3.2. Photoluminescence and Photon-Based Spectroscopy

As opposed to GC and mass spectroscopy methods, where the final desorbed product is physically recorded, photon-based analysis offers an alternative non-destructive way of interrogate the products or intermediates. Due to the samples and products not being consumed during the measurement, other analysis procedures can be done following the methods outlined below. Since these techniques allow for the products to be subject to multiple measurements, they are ideal for HT systems as a single effluent stream can yield various spectra detailing different behavior, maximizing the data produce from each sample in the library and thus improve throughput.

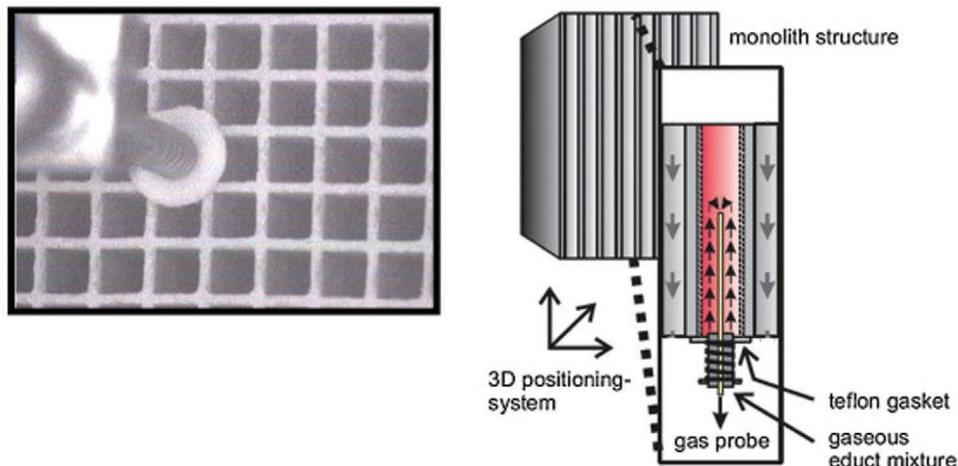


Figure 5. The insertion of the gas probe into a channel of the honeycomb monolith (left) and a schematic cross-section of the screening of an individual zeolite within the channel (right) where an FTIR spectrometer analyzed the reaction products downstream.^[79] Reproduced with permission.^[79] Copyright 2010, Elsevier.

A novel method of activity measurements on zeolites was proposed by Su and Yeung^[179,180] named laser-induced fluorescence imaging (LIFI) where this measures activity with fluorescence. Whilst initially applied to V_2O_5 catalysts, the procedure involved the samples being deposited separately on a steel plate and then placed in the reactor. The reagents then are passed over the samples with the system being simultaneously illuminated with a laser. The fluorescence of the product, naphthoquinone for the studies by Su et al.,^[179,180] is captured by a camera, enabling both spatial and temporal measurements to be performed. Whilst this is a relatively simple experiment to perform and analyze in a HT manner, this method requires the product to fluoresce and thus has limited application outside a few specific reactions. Nevertheless, the use of HT photoluminescence to record activity has been used by authors such as Atienzar et al.^[181] to screen various alkali metal zeolites for activity toward the formation of the phenylenevinylene oligomer, demonstrating the use of such an approach for certain systems.

The use of UV–vis to determine the efficacy of a zeolite was first performed by Gao et al.^[182] where the authors developed a multi-flow reactor with a UV adsorption cell that the effluent gas passed through after catalysis. Through the use of a unidirectional translation stage to move the UV cells, the products from up to 80 catalysts could be sequentially analyzed in quick succession and thus result in a rapid and HT testing of the catalysts. The authors employed the set-up to investigate the NO abatement of a series of mixed-metal exchanged H-ZSM-5 samples by recording the absorbance at 214.5 nm, the highest wavelength absorbed by NO, in the post-catalyst gas flow. Whilst no combination of Co/Ce or Ce/In ZSM-5 catalyst outperformed the industrial standard at the time, Pd/Al₂O₃, Gao and colleagues nevertheless demonstrated an alternative manner for measuring activity. However, this approach is inherently hindered by the translating stage, forcing the samples to be slightly delayed or the measurement points to be shifted by some amount, reducing the comparability of the resultant spectra. Moreover, the use of UV requires the reagents or products to have a non-overlapping peak that is easy to measure and relate

to the concentration, which becomes increasingly difficult as the number of species in the effluent flow increases due to multiple products being produced. Furthermore, such a technique would be impossible to use if no absorption even occurs in the UV–vis window. HT UV–vis spectroscopy is not limited to exclusively recording products from catalysis and has been successfully employed in 2010 by Horcajada et al.^[183] to indirectly measure the toxicity of several Fe-containing MOFs in parallel with a UV–vis active dye after exposure to said MOFs (see Section 4.2 for further details and discussion related to use of UV–vis spectroscopy to characterize and evaluate MOFs).

FTIR had initially been used in a HT manner to only characterize zeolite catalysts (see Section 2.2.2) but in 2003, the Lauterbach group applied their HT FTIR spectrometer^[92] to record the nitrous oxide storage capabilities for metals on γ -alumina supports using a 16-channel system.^[184,185] Fickel et al. later employed this set-up to interrogate Cu-zeolites for NO_x abatement in 2011 where the Cu small-pore zeolites were seen to be more robust and maintained greater activity after ageing than the comparative Cu-ZSM-5 zeolite.^[186] The Claus group developed an alternative HT gas probe that sequentially screened first metal-alumina catalysts and then poisoned Fe-zeolites for NO_x reduction activity in 2003 and 2010, respectively.^[79,187] This procedure involved a gas probe being translated over a honeycomb monolith, where the channels contained different catalysts, and then lowered into the desired channel with a gasket being used to prevent the gas escaping to another channel (Figure 5). The products were then subsequently analyzed by an on-line FTIR spectrometer. Kern et al.^[79] examined the effect of post-synthetic treatments of the various Fe-zeolites where it was found that the introduction of Cr or Cu resulted in an increased production of N₂ but concurrently increased the amount of the undesirable greenhouse gas N₂O produced, first only to a mild degree at low temperatures but then doubling at the higher temperature of 450 °C. This is a particular concern as the exhaust gas that the catalyst operates on is projected to be at or above this temperature for the majority of the vehicle's journey. Moreover, the addition of these transition metals resulted in a greater amount of ammonia slippage, the

amount of NH_3 reducing agent which does not react and would be expelled from the vehicle exhaust, where these metals result in a less efficient catalyst and increases the toxicity of the automotive emission. Additionally, the authors compared the deactivating effects of alkali and alkali earth metals on the Fe-MFI and the commonly used $\text{V}_2\text{O}_5\text{-WO}_3/\text{TiO}_2$ catalysts where it was found that the MFI type zeolites were more resistant to this poisoning. Further evaluation with UV-vis spectroscopy determined that rather than causing a change in Fe speciation or affecting the Fe's reducibility, the alkali cations simply blocked some of the channels and pores.

This study therefore demonstrates the efficacy of HTE as the effects of a wide range of different post-synthetic ageing treatments could be evaluated in a relatively short amount of time, about 20 h for 128 samples, and therefore allow for the durability and resistance of catalysts to be reported along with their effectiveness. Prior to the work by Kern et al.,^[79] thermography was employed for screening various catalysts. Cypes et al.,^[188] for example, investigated the activity of various catalysts, including mordenite, toward CO oxidation. The activity was recorded through an IR camera that imaged a 16×16 array as CO in air was passed over the samples and registered the change in temperature associated with the catalyst oxidizing CO. Whilst thermography can rapidly screen catalysts, with the Maier group commonly employing this,^[189] it is unable to give any information with regard to selectivity or prevalence of side reactions as with full FTIR spectrum.

2.4. Critical Remarks

Despite HTS being able to traverse vast regions of phase space, phase space is effectively infinite and thus a small region must be specified for investigation. However, identification of compositional ranges that are expected to contain novel topologies or contain samples with desirable features is difficult.^[38] Therefore, contemporary HT zeolite work focuses predominantly on a targeting or screening via computational methods prior to synthesis (see Section 3) with novel zeolite discovery through HTS, for example, becoming increasingly infrequent besides a few instances.^[39,56] Beyond HTS, HTE focuses on optimization or the benchmarking of catalysts with this avenue similarly involving simulations to pre-screen and flag samples for further investigation. Such developments are to be expected, and indeed welcomed, due to continual increase in computational capabilities coupled with experimental unexplored phase regions increasing in complexity and dimensionality as time increases.

With regard to automation and mechanization, synthesis, characterization, and catalytic evaluation either have only some sections automated^[49,69,145] or if all desired procedures are robotized, then there is never a system of transferring the components from one station to another without manual intervention.^[60,68] This results in a major bottleneck inhibiting massive scale HT and prohibits self-contained virtuous circles from being formed. Whilst many recent developments have increased the speed of analysis or enable in situ analysis,^[190–194] these methods are still done on samples sequentially rather than concurrently where parallelizing these new methods would enable a substantial speed-up of current HT workflows.

Despite the ubiquity of robots employed for experimental interrogations of zeolites, linking HT synthesis to characterization and activity measurements still remains a significant challenge, demonstrated by the use of pre-formed zeolites in characterization and activity workflows, preventing a “full-cycle” synthesis–treatment–characterization–test workflow for a defined region of parameter space with minimal human input.

Experimental HT work with zeolites is a relatively mature field with HT catalytic testing and HTS being ≈ 34 and ≈ 22 years old, respectively, whilst comparable work with other porous materials such as MOFs has only developed in the last 15 years (see Section 4). As a consequence of the developed nature of zeolite HTE and its pervasive use in industry (where HT autoclaves, XRD systems, and reactors are available to purchase), simulation and work with other newer porous material has dominated the literature for the past half-decade save for notable examples highlighted here. Nevertheless despite the plethora of work relating to zeolite HTE, issues regarding automation mentioned here continue to plague both particularly zeolite HT as well as HT in general.

3. Computational Screening of Zeolites

Zeolites are valued for their exceptional shape selectivity, making them highly applicable to many industrial applications. High throughput computational screening (HTCS) of these materials allows for the rapid identification of potentially industrially relevant structures, bypassing the expensive and time-consuming nature of experiment. Structure–property relationships vital to the understanding of thermodynamic and transport properties can be determined through computational work, which can help to identify optimal compositions. These screening approaches are facilitated by both experimentally determined and computer-generated structures, available in different forms as described in Section 2.

3.1. Databases

To assist the high throughput screening of siliceous zeolites, there are a number of structural databases available, the most well known of which is the IZA Structure Commission database (<http://www.iza-structure.org/databases/>);^[195] synthetically reported structures undergo a refinement process through the fixing of unit cell lengths and adjustment of atomic coordinates through least squares refinement. As of April 2020 there were 248 distinct topologies reported in the database, each of which has been given a three-letter code derived from the name of the corresponding synthesized material. The three-letter codes refer only to the connectivity of reported structures, and not to any particular material or composition. All topologies have an idealized siliceous form associated with them, despite many of these structures being currently inaccessible synthetically in that form; Crystallographic Information Files (CIFs) are readily available from their website for these idealized structures. Typically this database is the first port-of-call in a high throughput screen, as materials with the described connectivity have known synthetic pathways.

The Deem database is an example of a hypothetical set of siliceous structures;^[196,197] other such databases exist, with different methods for generating the structures, such as that of Foster and Treacy.^[198] The structures in the Deem database are generated via a process of simulated annealing followed by structure optimizations using both the Sanders–Leslie–Catlow (SLC)^[199] and the van Beest–Kramer–van Santen (BKS)^[200] force fields. In total there are 2.6 million distinct structures present in this database, all of which are available in CIF form containing their topological information, as well as associated energies for the structures calculated using the aforementioned force fields. Although millions of structures have been generated using this method, a subset of only $\approx 330\,000$ are within $+30\text{ kJ mol}^{-1}/\text{Si}$ of alpha-quartz (the most thermodynamically stable phase at ambient temperature, see also Henson et al.^[19]) when using the SLC force field, which defines the upper bound of energy for known zeolites. It should be noted that here the generation of structures is done in such a way as to maximize the number of energetically favorable structures by use of the ZEFSAII program,^[201–203] and thus this database contains by far the largest number of such favorable structures. Owing to its extensive quantity of thermodynamically favorable structures, and the availability of CIFs through the Predicted Crystallography Open Database (PCOD)^[204] and the Atlas of Prospective Zeolite Structures,^[205] the Deem database remains one of the most widely explored hypothetical database of zeolitic structures for use in high throughput computation.

AlPO—zeolites where the tetrahedral (T) sites are occupied by aluminum or phosphorous atoms in an alternating fashion—have also been explored in a high throughput manner. Innovative work by Li et al. generated over 80 000 hypothetical AlPO structures by variation of the stacking sequence of six rings, resulting in two newly synthesized materials;^[206] this work led on to further high throughput studies by the authors, examining the necessity for heteroatoms in their database of structures.^[207] Although useful, this database is far less extensive than the Deem database owing to its method of construction, as it does not explore the varied composite building units that zeolites can be made from.

3.1.1. Treatment of Aluminosilicate Structures

Currently there is no database that allows for the rapid screening of aluminosilicate zeolites (although some 43 000 structures were assessed but not made publicly accessible in a study by Muruoka et al.^[208]). Databases that do contain aluminosilicate structures, such as the IZA, typically use partial occupancies to describe the aluminum and charge compensating cation positions, which creates difficulty when attempting to use these structures in a screening approach. A commonly used method when studying aluminosilicate structures is to apply Löwenstein's rule of aluminum avoidance^[209] to describe the aluminum distribution, which states that Al–O–Al linkages are unfavorable; another such assumption is to place aluminum atoms in the structure according to Dempsey's rule, ensuring maximum separation of aluminum atoms within the structure.^[210] These methods are employed in order to capture the chemistry of functional zeolite catalysts, which remains one of the greatest challenges in zeolite science to date. Significant

modeling difficulties arise from the effects of SDAs, hexacoordinate aluminum, and extra-framework cation and aluminum disorder; the combinatorial nature of these hindrances makes the structural elucidation of zeolite catalysts costly.

Toward this goal, Fletcher et al.^[211] performed the first exhaustive periodic DFT study on a fully enumerated data set where each unique structure associated with a particular aluminum content was evaluated. Surprisingly, it transpired that when the charge compensating cation is a proton, the lowest energy configuration is associated with a non-Löwenstein structure. Conversely, when the counter-cation is a common alkali metal, Löwenstein structures are the most stable form. Recent computational studies by Heard et al.^[212] have shown that the presence of water suppresses the driving force for the formation of non-Löwenstein Al–O–Al linkages. Whilst these purely DFT studies are insightful, it is currently not practical to carry out a full DFT ranking of important catalyst structures like ZSM-5 due to the combinatorial explosion and huge compute cost (for ZSM-5 with the MFI structure type there are 24 distinct T sites in the monoclinic form). However, recent work by Evans et al.^[213] has reported a new machine learnt aluminosilicate model that is similar in cost to a force field but presents accuracy comparable to DFT. Remarkably, this machine learnt model, trained on ≈ 1600 DFT optimized crystal bulk configurations was able to blind predict the correct ranking of protonated mordenite geometries and surface slabs of chabazite—neither of which featured in the training set. The advent of such models heralds the possibility of identifying a manifold of plausible and low energy structures that are likely to be redolent of real catalysts for the first time.

3.2. Screening Techniques

For the screening of microporous materials such as zeolites, there are a number of possible techniques that can be used with varying degrees of accuracy and computational cost. These range from purely structural and geometric considerations to higher-level techniques which involve calculating the precise energetics of the system to rank their stability.

3.2.1. Textural and Geometric Characterization

Nanoporous materials such as zeolites and MOFs excel at the adsorption and separation of molecular species, and so a useful first step is to calculate the relevant textural/geometric properties that would impact these processes, such as pore limiting diameter (PLD) and largest cavity diameter along the free path (LCD). This can be done relatively cheaply in computational terms, and several programs are available that make evaluating these measures very rapid: Zeo++ is a widely cited program, and is able to calculate these properties as well as others using Voronoi decomposition, which produces a graph representation of the free space, allowing the resulting network to be analyzed for features of interest.^[214] Poreblazer splits the structure into cubelets and identifies which of these overlap with framework atoms, allowing the free space to be assigned as those cubelets with no overlap;^[215] the Hoshen–Kopelman cluster labeling algorithm is then used to analyze the connectivity.^[216]

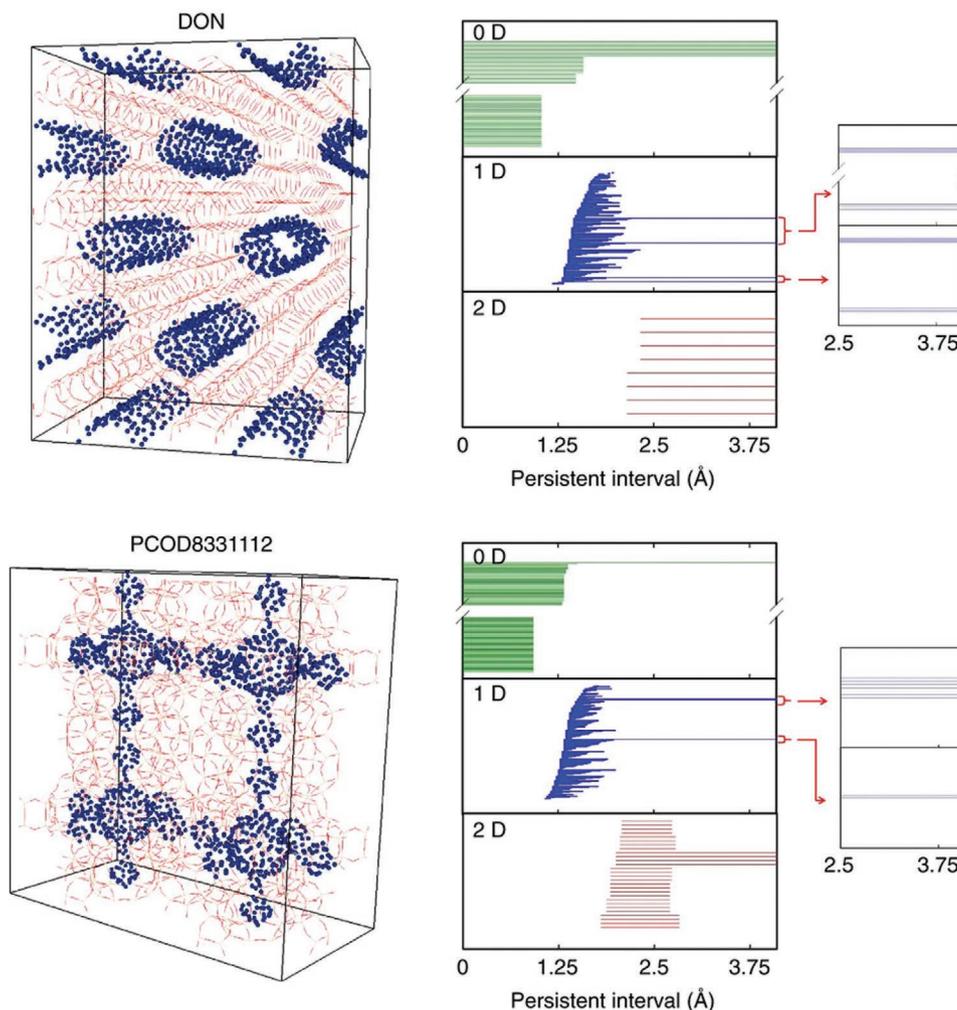


Figure 6. Image showing the structure and fingerprints of the zeolite DON (top), and the hypothetical zeolite PCOD8331112 (bottom). The figures on the left show the zeolite structure in red, and the randomly sampled points to identify the pore structure in blue. The 3D pore network of PCOD8331112 can be seen by its singular long interval in the 0D barcode, suggesting that the pore system is connected. The disjointed nature of the pore system in DON can be seen by its eight long intervals in the 0D barcode, suggesting that eight separate channel systems exist. These conclusions are evident by inspection of the structures on the left. Reproduced under the terms of the CC-BY Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0/>).^[217] Copyright 2017, The Authors, published by Springer Nature.

A novel approach to examining pore geometry was developed by Lee et al. whereby unique fingerprints of nanoporous materials are generated based on their channel and pore systems.^[217] Example fingerprints are shown for two zeolites in **Figure 6**. Comparison of these fingerprints between different structures allows identification of structures with similar pore geometries, which allows comparison between different zeolites and also allows MOFs or other nanoporous materials with similar pore geometries to be identified. Techniques such as this could prove very revealing in correlating the features of known data sets to a wider range of nanoporous materials, allowing for the identification of high performance porous networks of a different class.

Similarly, a geometry-based descriptor has been developed by Martin et al. to capture the shape and size of a material's accessible void space in a 3D vector, termed a Voronoi hologram.^[218] Combining this representation with a modified similarity coefficient and dissimilarity-based selection algorithms allows for diversity selection on a given set of porous materials; use

of such techniques can allow for systematic reduction in the number of structures investigated in a high throughput study without resorting to random selection.

3.2.2. Computational Methods

The use of molecular simulations to predict thermodynamic and transport properties is valuable for the characterization of nanoporous materials. Highly accurate approaches such as those based on quantum mechanical calculations can be used to compute electronic structure properties of the system. However, when employed in HT workflows, these methods are generally restricted to relatively small sample sizes due to their high computational cost.

Simulations based on classical mechanics^[219] are typically used in high throughput workflows due to their ability to determine properties of interest while striking a balance between accuracy

and execution time. A variety of techniques exist for the calculation of adsorption properties; two commonly used schemes are grand canonical Monte Carlo (GCMC) and its extension, configurational-bias GCMC (CB-GCMC),^[220] which allow for the prediction of thermodynamic equilibrium properties such as heats of adsorption, adsorption isotherms and isobars, and the selectivities of mixtures.^[221] GCMC simulations are typically used for the modeling of rigid adsorbates, while CB-GCMC can be employed for the treatment of flexible molecules in order to generate distributions of adsorbate conformers. Molecular dynamics (MD) simulations are commonly used to calculate transport properties such as diffusion coefficients, although other approximations can be made to determine these characteristics as mentioned later in Section 3.3.2.^[222,223] We note that free energies of adsorption can be calculated from molecular dynamics simulations, however, the typically slow convergence of averaged ensemble properties such as adsorption energy means GCMC and CB-GCMC are usually more efficient.

3.2.3. Framework Flexibility

Modeling framework flexibility can have a considerable impact on the calculation of properties, however lattices are often modeled as rigid to reduce computational cost. This has been shown to be a reasonable assumption under certain conditions; a study by García-Sánchez et al.^[224] highlighted that the effects of framework flexibility on adsorption properties were small, whereas their influence on transport properties was more pronounced. These findings were investigated further by Krishna et al.^[225] who determined that framework flexibility had no influence on self-diffusivities on an examined subset of cage-type zeolites interconnected by 8-membered rings. More recently the separation of ethane/ethene was investigated by Bereciartua et al.^[226] using ab initio MD, which showed that framework flexibility is key to this separation process. Similar consideration is needed when treating MOFs; a recent study by Witman et al.^[227] revealed significant changes in the separation ability of MOFs when framework flexibility was taken into account, leading to the conclusion that its employment is necessary to accurately rank the performance of structures. Based on these findings the necessity of modeling framework flexibility is dependent on the structures and the sensitivity of the processes under investigation, and so neglecting flexibility carries a potential significant error.

3.3. Applications

This section highlights high throughput approaches applied to the identification of ideal zeolite structures for a variety of purposes, as well as the unique descriptors discovered in these studies that allow for the rapid screening and optimization of structures.

3.3.1. Adsorption

The storage of gases such as hydrogen or methane, for use as fuels, is an active area of research. However, a major drawback to these technologies is the relatively low volumetric energy

density while in gaseous form. Carbon capture is another topic garnering significant attention currently, due to the ever increasing impact of the greenhouse effect;^[228] capture and sequestration of greenhouse gases such as carbon dioxide is thus of urgent interest. Nanoporous materials offer a potential solution to these problems by adsorbing the gases within their pores, permitting a higher density of the stored gas at ambient conditions. Zeolites, being one of the most widely studied classes of nanoporous materials, have been proposed as promising candidates for this application.

Carbon Sequestration: A study by Kim et al. on the uptake of CO₂ in siliceous and aluminosilicate zeolites showed interesting trends between free volume, largest included sphere, and uptake;^[229] a range of 8 < LCD < 13 Å for the largest included sphere was shown to be optimal to promote higher uptakes. Beyond this point the free volume inside the pores increased, but a lack of available adsorption sites caused an overall decrease in uptake. The IZA^[195] and Deem^[197] databases were both used in this study, and the challenge of screening aluminosilicates was overcome by the generation of structures with the random replacement of silicon with aluminum while obeying Löwenstein's rule. The force field developed by Garcia-Sánchez et al.^[230] was employed for the simulations, which has been shown to reproduce experimental isotherms across many topologies, and improves on previous models by allowing for the free movement of extra-framework cations. The results of the study allowed interesting correlations to be uncovered between adsorption of CO₂ in siliceous zeolites to that in aluminosilicate structures; the two main descriptors found to increase CO₂ uptake were topologies with a large free volume, and those with the greatest fraction of framework–framework atom distances between 3 and 4.5 Å. Incorporation of descriptors like these into a high throughput workflow allows for a drastic reduction in compute cost; promising aluminosilicate structures can be identified rapidly without the need for the calculation of properties at varied Si:Al, which drastically increases the number of possible structures.

More recently a study by Fang et al. sought to develop a transferable force field fit to energies from density functional theory and coupled cluster calculations.^[231–233] By fitting to high-level quantum mechanical data, a broader description of the energy landscape can be obtained for areas that may not be sufficiently sampled by adsorption experiments, making the force field more transferable to a wider range of applications. The data set contained a combination of siliceous and aluminosilicate structures with both sodium and/or potassium as the extra-framework cation, allowing for high throughput identification of promising structures and identification of the optimal Si:Al ratio for adsorption across a range of topologies. Similarly to the study by Kim et al.,^[229] aluminosilicate structures were generated by random substitution of Al in place of Si while obeying Löwenstein's rule, although the only topologies investigated were those from the IZA database with verified experimental aluminosilicate analogs. Zeolites with a range of ring sizes were investigated with a focus on 10-membered ring structures and those with large pore volumes in order to avoid possible pore blocking by guest molecules. Data obtained from these simulations showed that aluminosilicate structures adsorbed greater quantities of CO₂ than siliceous structures

due to the presence of strong interactions between the CO₂ and the extra-framework ions. However, siliceous structures displayed a higher working capacity due to these interactions preventing the adsorbed molecules from desorbing and leaving the framework.

Methane Storage: A wide range of nanoporous materials were investigated in a high throughput study by Simon et al. in order to identify their performance limits for storage of methane.^[234] The goal was to identify structures with the highest deliverable capacity of methane at different pressure swings: 65 to 5.8 bar, and 35 to 1 bar. In this study over 135 000 hypothetical structures from the Deem database^[197] were tested, as well as 187 structures from the IZA database.^[195] Pores that could adsorb methane but were not accessible to the gas phase were blocked, giving a more realistic description of the maximum loading.^[235] The two most promising candidate structures identified from this study were both hypothetical structures, highlighting the value of screening such databases. The two structures exhibited very high working capacities of 200 and 172 volumes of gas per volume of material, at the storage pressure of 35 bar, and discharge pressure of 1 bar. Key properties were identified to guide the search for new promising materials: a high density of adsorption sites, a moderate heat of adsorption, and adsorption pockets positioned optimally for guest–guest adsorption.

Further to this, Krishnapriyan et al. recently published a study using this data set to train a machine learning model for the prediction of adsorption properties using different descriptors.^[236] Commonly used descriptors were used as a “baseline” for comparison: LCD, PLD, ASA, accessible volume (AV), and crystal density; topological descriptors were then computed by use of persistent homology, allowing vector representations of channels and voids in the materials to be determined. The use of random forest regression to predict the methane adsorption isotherms allowed for importance values to be calculated for each descriptor, distinguishing this work from others by demonstrating how the topological features complemented the baseline descriptors for an overall more accurate model of adsorption. The results of this study showed that at low pressures, the topological descriptors calculated through persistent homology, specifically the 2D topology, performed significantly better than baseline features, with these baseline features dominating at higher pressure, as shown in Figure 7. Machine learning approaches such as this can provide in-depth knowledge of structure–property relationships that govern the capabilities of nanoporous materials to perform well industrially; use of descriptors identified in such studies can aid in the rapid screening of zeolites for many applications.

3.3.2. Separation

Although zeolites have been successfully screened for storage applications as highlighted above, MOFs generally perform better for these tasks owing to their relatively low crystal density in comparison to zeolites, affording guest molecules more free space.^[237] This naturally restricts zeolites to the adsorption and separation of smaller molecules, and so they are instead extensively used in the shape-selective separation of isomers and

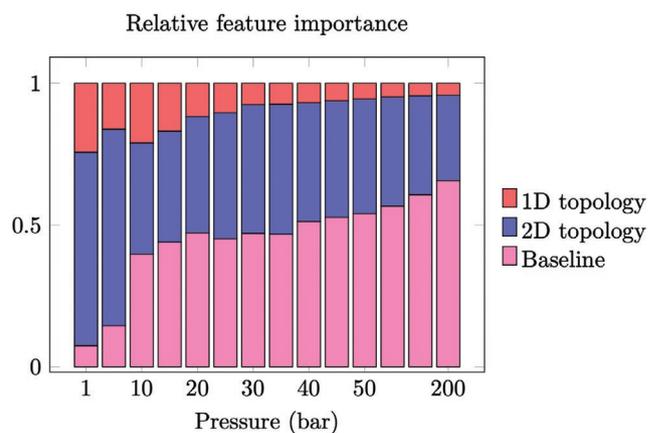


Figure 7. Image depicting the relative importance of topological descriptors (1D and 2D topology) over the baseline features described above, illustrating the need for the development of tools to extract more sophisticated descriptors. Reproduced with permission.^[236] Copyright 2020, American Chemical Society.

mixtures of low molecular weight adsorbates; the ability to tune their pore and channel systems allows them to be optimized for a variety of such applications.

High Throughput Separations of Binary Mixtures: High throughput screening of zeolites and MOFs was undertaken for the separation of a range of binary mixtures at different temperatures by First et al.^[223] The methodology employed was an interesting intersection between structural and molecular dynamics approaches, allowing for the rapid screening of structures for a range of purposes; first, energy barriers were calculated for the movement of the molecule under consideration through each pore, allowing for a pathway to be constructed through the periodic lattice. The energy of a pathway was then taken to be the maximum energy of any pore along this pathway, with the assumption being that transport is limited solely by this bottleneck. Boltzmann factors were then calculated for the movement of a molecule through the pathways of a framework, with the difference between two different molecule’s Boltzmann factors representing a framework’s separation ability. Using this methodology, 196 siliceous zeolite structures from the IZA database, and 1690 MOF structures were screened for their separation ability for eight different processes of interest, allowing for the identification of many high performing topologies that had not before been considered. Trends in separation ability with respect to temperature were also investigated, revealing important failings in the ability to discriminate between similar shaped molecules as temperature was increased.

The benefits of large-scale computational screening are illustrated succinctly in the work of Kim et al. on the separation of ethane/ethene mixtures.^[238] Here a typical HT approach was employed to screen both the IZA and Deem databases, using textural characterization as an initial screen followed by random selection from the filtered set in order to reduce computational cost, while still exploring a representative portion of the data set. A performance metric was determined to be the product of selectivity and working capacity, in order to find a structure able to refine the mixture to a given purity in fewer cycles.

The best performing structure based on this metric was further examined, and the structural features responsible for this performance were determined; by identifying other topologies with a similar framework geometry to the leading candidate structure (see Section 3.2.1), 90% of the optimum separation candidates could be determined without the need for simulation. This illustrates the power of high throughput computation in elucidating structure–property relationships, which in turn allows for the identification of metrics for the rapid screening of large databases.

Further to this work, CO₂ membrane separations were studied^[222] using a similar work flow, although with different selection criteria: for a given separation a minimum selectivity was identified, and from the pool of structures which surpassed this boundary, those with the highest permeability were deemed to be the best performing candidates. Permeability was defined here as the product of the solubility and the diffusion coefficients of the gas molecules. Diffusion coefficients were calculated rapidly without the need for expensive MD calculations, by using transition state theory; peaks and troughs in the free-energy profiles of channel systems allowed diffusion coefficients for individual channels to be calculated, which in turn allowed for diffusion coefficients in a given direction, as well as self-diffusion coefficients, to be calculated. The use of such techniques allows for the exhaustive search of a manifold of materials, unveiling key descriptors for the process under consideration; here it was found that a relatively low heat of adsorption and an intermediate Henry constant were optimal for CO₂/CH₄ separation.

Application-Specific Separations: Complementary to the carbon capture work referenced earlier,^[229] a high throughput study into the separation of CO₂ from flue gas was undertaken by Lin et al.^[239] Flue gas is the gas produced by power plants, and is approximated as a mixture of N₂ and CO₂; separating and sequestering the CO₂ from this is crucial to limiting the ever-increasing amounts of greenhouse gases in the atmosphere. Siliceous and aluminosilicate zeolites, as well as ZIFs, were examined for their separation capabilities; siliceous structures were taken from the IZA and Deem databases. Aluminosilicate forms of these structures were generated by obeying Löwenstein's rule, yielding one structure with an Si:Al ratio of 1, and aluminum distributions between this and the siliceous analog generated by random replacement of silicon with aluminum to yield ten structures; calculated properties for these ten structures were averaged to give a representative value for a particular Si:Al ratio. ZIF structures were generated by taking a siliceous zeolite structure and scaling the unit cell by 1.95, then replacing Si with Zn, and O with the imidazole group.

A remarkably complete picture of the flue gas separation process was presented, with a novel metric, the parasitic energy, introduced to discriminate between candidate structures. The parasitic energy metric was calculated to be the energy required to: heat the material, supply the heat of desorption, and pressurize CO₂ to 150 bar. The process investigated was a temperature–pressure swing whereby the temperature could be increased, the pressure decreased, or both, in order to allow desorption of the CO₂. By calculating the parasitic energy, key trends were extrapolated from the data, showing a nonlinear relationship between Henry constants and parasitic

energy. This work is crucial in illustrating the importance of modeling an entire process as opposed to a single property, as other metrics may not take the energy required for desorption into account. Another key observation from this work is the abundance of hypothetical structures in the basin of optimal structures, many of which were identified to have low densities, suggesting that these may be key targets for synthetic work.

The separations of methane from coal mine ventilation air and low-quality natural gas mixture feeds were investigated by Kim et al. in order to identify new materials for these processes.^[240] A significant challenge in the capture of methane through separation processes, is finding a material with a preference for adsorption of the non-polar methane molecule over other constituent gas molecules in the mixture feed, for example, CO₂. Aluminosilicates have been seen to not lend themselves well to these challenges, as the presence of cations in the structures creates strong binding sites for CO₂. Thus, 87 000 hypothetical^[197] and 190 IZA siliceous zeolite structures^[195] were explored for their use in these processes using the force field of García-Pérez et al.,^[241] with the D2FF force field of Sholl et al.^[233] for further validation of leading candidate isotherms.

For the purpose of capturing methane from low-quality natural gas feeds, the zeolite SBN was identified as the most promising candidate due to its topological disposition of allowing optimal CH₄–CH₄ separations of 4 to 4.6 Å between the molecules at high pressures, facilitated by the shape and size of the pores. The adsorption isotherm shows that the separation ability of SBN for this process is pressure sensitive, as at lower pressures the higher binding energy of the more polar CO₂ molecule contributes more significantly to uptake. Similarly for the dilute concentration source examined, coal mine ventilation air, a strong correlation between Henry's constants and uptake values was seen due to the process occurring at 1 bar. It was found that leading materials possessed a narrow channel system rather than cage-like channels, as identified through calculations of the largest included sphere along the free path. These narrow channels allowed an increased number of oxygen atoms within close vicinity to the centre of binding sites. Two known zeolites, ZON and FER, were proposed as candidates for this process due to their large Henry constants and high CH₄/CO₂ selectivity, although the greater performance of some hypothetical structures again illustrates the need for the screening of these predicted materials.

A study by Bai et al.^[221] into the purification of ethanol from fermentation broths found FER to also be promising for this separation, outperforming MFI which has seen extensive exploration in the literature. Use of a high throughput multi-tiered workflow allowed for efficient screening of leading structures to undergo further calculations; for the purification of ethanol from an ethanol/water solution, calculations at the lowest solution-phase concentration were used as the initial screen, with the performance metric being the product of selectivity and loading. Analysis of leading structures showed that the increased performance of FER for this separation was due to its ability to disfavor formation of hydrogen bonds from ethanol to water, preventing their co-adsorption; similarly the ATN structure is able to facilitate high selectivity at increased concentrations due to its narrow windows separating well-spaced ethanol adsorption sites. Although only IZA structures were screened

for this process, application of high throughput analysis techniques to a wider data set, such as use of the similarity measures described earlier, could help identify promising predicted structures; moreover, techniques such as persistent homology could aid in the elucidation of structural features responsible for these properties.

In the same study, the group devised a model for finding structures useful in the refining of oil: high affinities for long linear alkanes correlates to a high concentration of these molecules near active sites, while low affinities for branched chain alkanes promotes their desorption to prevent further cracking. A further performance measure was low selectivity between linear alkanes of different lengths to allow conversion of a broad range of different length chains. A similar tiered approach was taken, with simulations at the infinite-dilution limit and 573 K on all IZA structures, as well as the 330 000 thermodynamically accessible structures from the Deem database; the top candidates then underwent longer simulations under these same conditions, as well as at a fixed pressure of 3 MPa. Data from the IZA database set of structures readily identified current leading structures as the most promising, with a selection of hypothetical structures outperforming these by several orders of magnitude. Many structures with favorable adsorption characteristics were found to be in the range of $4.5 \text{ \AA} < \text{LCD} < 7.5 \text{ \AA}$, with high selectivity for linear versus branched alkanes being best correlated to “pore bumpiness.” Although some structural features were seen to have an impact on selectivity, the lack of clear correlation between simple geometric features and selectivity shows the need for the development and implementation of more high throughput analysis tools which are able to help identify structure–property relationships.

3.4. Machine Learning

We now briefly discuss the emerging role of machine learning (ML), distinguishing between applications focused on using simulation derived data and those where experimental data is used.

3.4.1. Computational Data

The identification of promising hypothetical nanoporous materials (e.g., see Section 3.3.1, Simon et al.^[234]) prompted a focus on efficient methods of evaluating the viability of synthesizing these structures and their characterization. The parallel evolution of ML and HT data acquisition techniques has facilitated the marriage of these methods to enable the screening of such databases for the massively combinatorial problems arising in nanoporous materials.

Use of quantum mechanical data for the training of highly accurate predictive algorithms is a powerful tool offered by ML, allowing these models to reproduce DFT data to a high degree of accuracy on structures not in the training set.^[213] Previously, DFT-derived data sets have been used for the derivation of force fields,^[231–233] however the application of ML to these challenges offers huge benefits, due to the relative ease of training the models, as well as the ability for machine learning models to

identify nonlinear trends. A recent highlight of the application of machine learning techniques trained on computational data includes the design of organic SDAs for the synthesis of zeolite beta by Daeyaert et al.,^[242] determination of the most thermodynamically favorable aluminum distribution in a range of zeolite topologies by Evans et al.,^[213] and the reliable calculation of anisotropic properties toward the discovery of auxetic zeolite frameworks by Gaillac et al.^[243] Modeling of these features demonstrates the ability of machine learning to accelerate data-driven predictions that can help inform experiment, cementing their importance in the virtuous circle.

3.4.2. Experimental Data

The abundance of experimental data available for zeolites, most notably for synthesis, has given rise to a surge of interest in applying ML techniques to these data sets; recent developments in natural language processing have allowed for extraction of this experimental data from the literature, leading to insightful predictions that can help inform synthesis. The power of such techniques is highlighted in the work of Jensen et al.^[244] who developed an automated data extraction pipeline which gathers data on synthesis and topology from both tables and text. This data set was then used to make predictions on the density of the zeolite that would be formed under different synthesis conditions. An account of the group’s work was summarized by Moliner et al.^[245] who demonstrate the capabilities and limitations of applying machine learning to such tasks; in their work the group illustrate the virtuous circle of continuous feedback between computation and experiment (see **Figure 8**), and importantly emphasize the need for standardization of data in computer-readable forms. In this work, the group applied machine learning techniques to address the “missing link” to a complete workflow for zeolite synthesis: tools that can extract and process large amounts of data from the literature, predict the formation of stable hypothetical frameworks, and the corresponding prediction of OSDA’s that can facilitate their synthesis.

A similar approach to Jensen et al.^[244] was taken by Muraoka et al.^[246] who collated experimental data in order to train an ML algorithm with structural and synthetic descriptors. Evaluating the importance of these descriptors allowed for the construction of a similarity network, which provided new insights into structural similarity between zeolite topologies. The synthesis of a family of zeolites which are structurally related (EEI-EUONES) was attempted using the SDA employed in the formation of IHW, a topology revealed by the similarity network to be structurally similar, but previously not considered part of this family. Successful synthesis of EUO using this SDA under the synthesis conditions for EEI confirmed the similarity between the structures that was identified in this study. Incorporation of hypothetical structures into this workflow could allow for key insights into their synthesis, facilitating the discovery of new materials.

Although a powerful tool, data mining of published work relies immensely on the reproducibility of the data in question. Recent work in the domain of life sciences research has shown prevalence of irreproducible preclinical data,^[247] calling in to

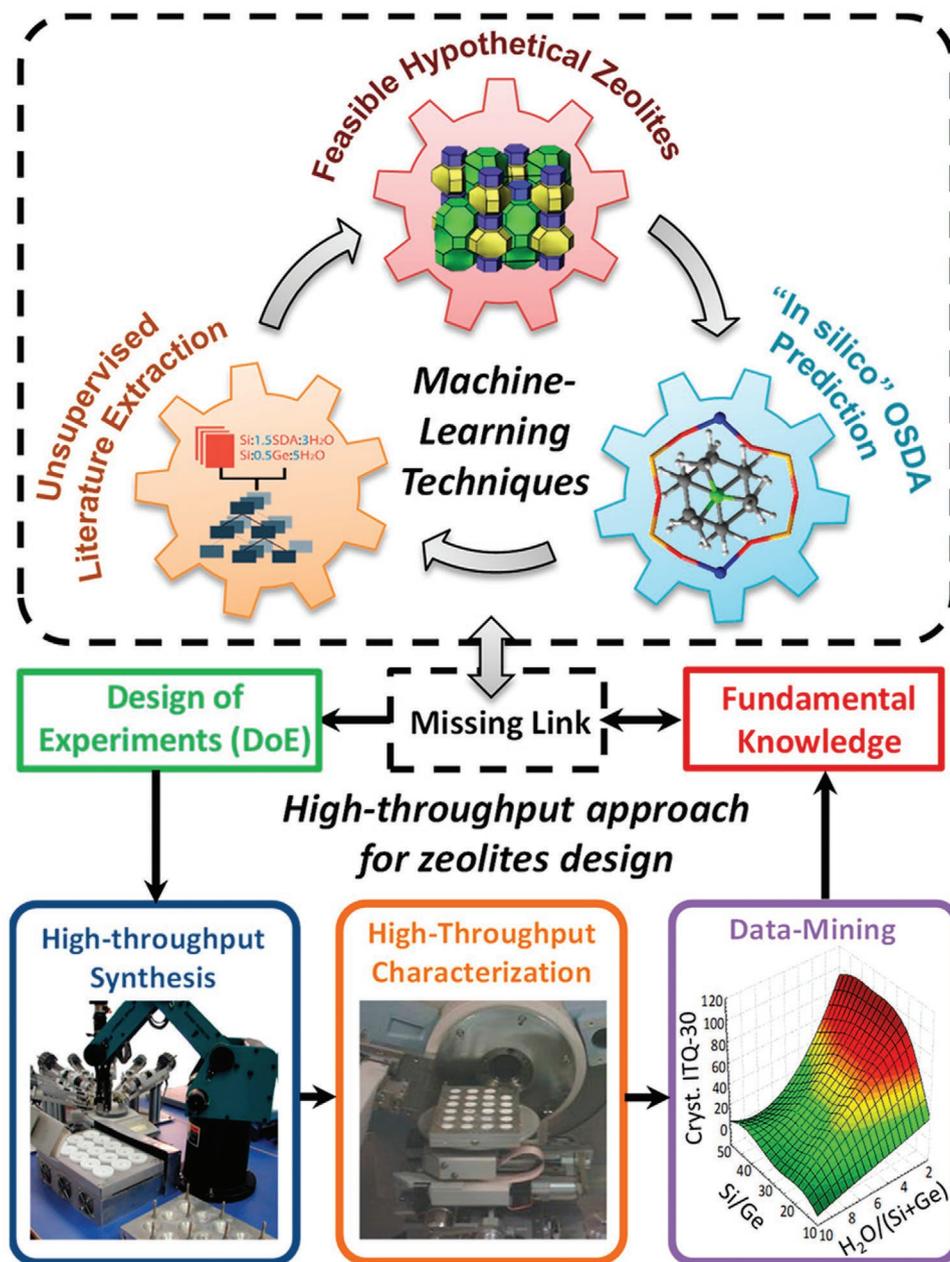


Figure 8. High throughput workflow for the computer-guided synthesis of nanoporous materials. Reproduced with permission.^[245] Copyright 2017, American Chemical Society.

question the legitimacy of extracting data from the published literature, although we note that ML could be potentially used to identify anomalous studies.

3.5. Critical Remarks

Many high throughput computational studies have been performed for the assessment of zeolites as media for separation and adsorption, and as catalysts. By their nature these studies cannot completely capture all physical factors that govern the performance of the material, and so assumptions are made

to allow for their efficient analysis while minimizing the loss of accuracy. Common assumptions made are the neglect of framework flexibility, the random ordering of aluminum within aluminosilicate structures, and the derivation of transport properties through calculation of energetic bottlenecks. Through the development of more efficient algorithms and increased availability of computational resources, more realistic calculations will be possible, leading to large repositories of high-quality data.

Continued development and use of the analytical tools available will lead to the identification of powerful descriptors, allowing for in-depth understanding of

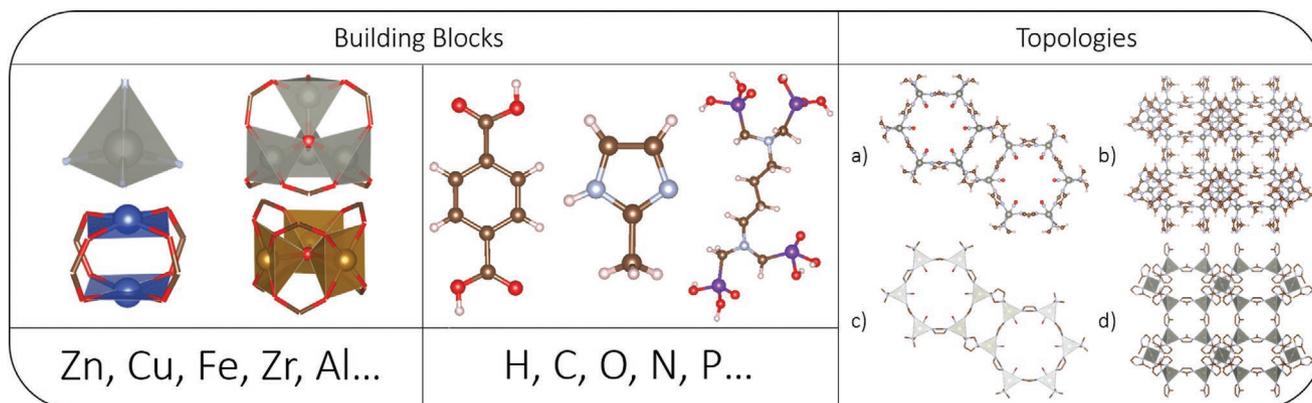


Figure 9. Some common MOF building blocks, separated into SBUs and linkers. Structural fragments of ZIF-70 and ZIF-8 are shown in a) and b) respectively, with polyhedral representations also depicted in c) and d). These systems have topologies gme and sod, respectively, which are also shown for zeolites in Figure 1. Zn, Cu, Fe, C, H, O, N, and P atoms have been respectively colored grey, dark blue, gold, brown, white, red, light blue, and purple.

structure–property relationships; this is exemplified in the work of Krishnapriyan et al.^[236] where topological descriptors were computed by use of persistent homology, and their importance in governing the adsorption process investigated was assessed through use of machine learning. The combination of such analytical tools with machine learning will be key to unveiling the nonlinear relationships between structure and property.

As the tools and architecture for high throughput computation develop, so must our abilities to effectively model real applications; in order to capture the full chemistry of a given process, proposed models must explicitly take into account all operating conditions. Exemplars of this include the studies described in Section 3.3.2, particularly the work of Lin et al.^[239] in their determination of a novel metric, the parasitic energy, for describing the energy penalty imposed on a power plant by conducting carbon capture and sequestration.

Cooperation between groups working on different classes of microporous materials is key to the continued development of this field; currently this is facilitated by initiatives such as the MGI, NOMAD, BIGmax, and MARVEL as mentioned in Section 1. Many tools and methodologies employed for the modeling and analysis of zeolites are directly applicable to MOFs and vice versa, hence cross-disciplinary collaboration will prove auspicious for the identification of new industry-leading materials.

4. High Throughput Experimental Work on MOFs

MOFs are another important category of porous materials. Unlike zeolites, the general structure of a MOF can be broken down into distinct inorganic and organic building blocks (respectively, termed the secondary building unit, or SBU, and the linker). There exist myriads of combinations of such components, imbuing the set of possible frameworks with incredible compositional and topological diversity. Some example systems are shown in **Figure 9**, and the reader is referred to reviews by Furukawa et al. and Yuan et al. for a broader introduction.^[248,249] High porosity, well-defined metal centers, and almost endless possibilities for functionalization have attracted attention to these materials for commercial and industrial applications, ranging from gas storage and separation to

catalysis.^[250–252] It is worth noting that the field of MOFs is not as mature as that of zeolites, with the former having grown over only the last ≈ 20 years. Therefore, many of the HT techniques which have been used in MOF discovery and screening have previously been developed for and applied to zeolites, as described earlier in Sections 2 and 3. However, due to differences in chemistry, synthetic protocols, and target applications between the two classes of materials, these existing procedures have sometimes been applied in a different manner and been supplemented by new models and approaches, as will be discussed here and in Section 5.

In experimental work involving MOFs, high throughput methods can be deployed at two stages. Synthesis can be accelerated by carrying out the dispensing of reagents and subsequent heating in a parallel manner, yielding a large number of samples in minimal time and with reduced human effort. Structural and other characterization may also be parallelized to rapidly identify crystalline products or evaluate properties of interest across multiple materials. In studies focused on the discovery of new MOFs, streamlining synthesis is more fruitful, while for work aiming to evaluate material performance at multiple condition–framework combinations, it is important to deploy suitable parallel instrumentation. In both cases, HT methods may need to be applied at only one stage to relax the workflow's bottleneck and accelerate the development of MOFs.

This section introduces high throughput techniques used in MOF synthesis along with key studies which have made use of them. The number of HT articles published to date is not very extensive, but there have been enough successful experiments to affirm the utility of such methods in discovering new frameworks and tuning their properties. The following sections examine the synthetic and characterization stages, while also touching on the development of new parallel tools, the importance of feedback loops, and issues surrounding reliability and consistency.

4.1. Preparation and Synthesis

A key concern when synthesizing MOFs is sensitivity to reaction parameters, as even minor changes in conditions can lead to amorphous products or different phases; like zeolites, they

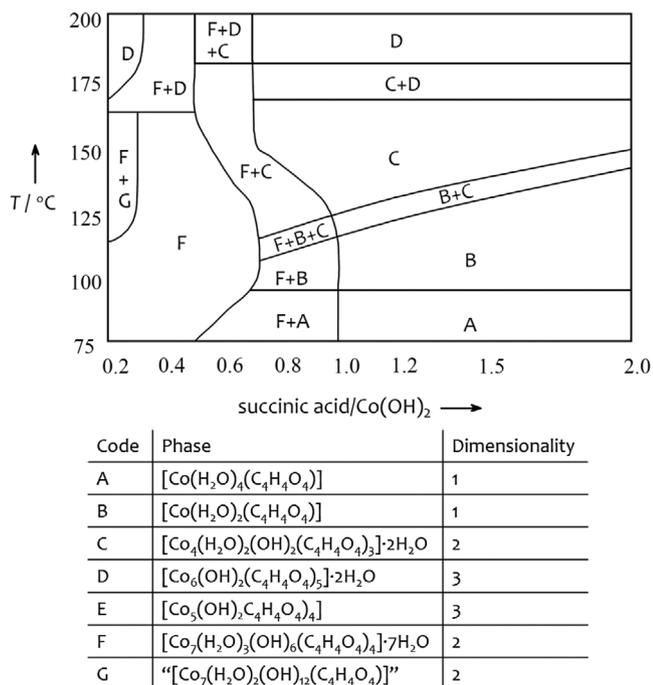


Figure 10. Temperature–composition phase diagram of cobalt succinate and a description of each phase. Reproduced with permission.^[253] Copyright 2005, Wiley-VCH.

are kinetic products. This is well exemplified by the cobalt succinate system, with a phase diagram shown in **Figure 10**.^[253] Taking into account just two synthesis variables, the system exhibits 15 different product mixtures, with some regions of the diagram existing over relatively narrow bands. The complexity of phase space here makes it evident that control and consistency are paramount in achieving reproducible results, and any attempt to increase throughput should not neglect these key factors. This can be done by using automatic dispensers and multi-compartment reaction vessels, both of which were first developed for research on small pharmaceuticals over 40 years ago, though they have since been applied to zeolite synthesis (see Section 2.1).^[3] In this manner, consistency can be maintained whilst minimizing manual involvement, in terms of both time and human effort.

A simple and reliable way to streamline the dispensing of reagents is to use a programmable liquid handler. This approach is particularly powerful if the metal salts or precursors can reliably be prepared as stock solutions or are readily available in this form. This allows for the rapid testing of combinations of common linkers, solvents, and additives such as acidic modulators. Sonnauer et al. made use of a liquid dispenser to synthesize a chromium MOF with the MIL-101 structure in the presence of a dozen different solvents and liquid additives, with the presence of acetic acid proving important to product crystallinity.^[254] More recently, Kelty et al. investigated the influence of reaction parameters on the synthesis of porphyrinic zirconium MOFs.^[255] 1027 combinations were tested, encompassing seven different acidic modulators and solvents, ultimately yielding nanocrystalline forms of three existing frameworks. In the case that most reactants are readily

available in liquid forms from suppliers, as was the case in both of these studies, programmable liquid dispensers allow phase space to be explored with limited human input beyond experimental design and characterization.

The dosing of solids can also be done in an automated manner with solid dispensers. This allows for the simultaneous use of more metal precursors, low-solubility reagents, and linkers directly in powder form in a high throughput manner. For some synthetic routes, such as hydrothermal synthesis, the use of solids is necessary. Stock and Bein have developed an HT methodology which includes solid dispensing based on an automated dosing station.^[256] This workflow has been adopted in the high throughput study of metal phosphonates synthesized via a hydrothermal route. Bauer et al. used this approach to carry out a synthetic screen of cobalt phosphonates using four different salts, which resulted in the identification of four new frameworks.^[257] Bauer and Stock later examined combinations of a new phosphonate linker with six different divalent cations, from which they isolated four novel metal phosphonates.^[258]

Some research groups have also made use of more advanced robotic workstations that can handle both liquid and solid reagents without manual involvement. Sumida et al. have made use of an automated robotic dispensing module in research on an iron MOF for gas storage, although their synthesis was limited to two powder reactants.^[259] The group of Martí-Gastaldo has also made use of a robotic dispenser to optimize the syntheses of new titanium frameworks. Padial et al. synthesized a new Ti MOF, MUV-11, using conventional methods, then refined the synthesis parameters via the high throughput use of a robotic workstation.^[260] Crucially, this allowed them to screen numerous Ti precursors, both allowing them to establish the use of previously neglected Ti sources and optimizing sample crystallinity. This optimization is particularly important for titanium frameworks, given the tendency for amorphous by-products to form during the syntheses of these systems.^[261] The same group used this methodology to make a titanium analog of MIL-101 in a study by Castells-Gils et al.^[262] Using a robot allowed the authors to first find a viable synthesis route for this MOF before screening multiple precursor–solvent combinations to increase sample crystallinity. In light of this experimental work, automated solid dispensing seems particularly well suited to studies where the precursor is varied to tune the synthetic route.

When carrying out HT synthesis, it is necessary to choose appropriate vessels so as to allow for carrying out multiple reactions in parallel. The most common format is the multi-well plate, which typically contains 24, 48, or 96 individual compartments, and is suitable for hydrothermal or solvothermal synthesis if made of Teflon-lined steel. The use of multi-well plates for porous materials is well established with their application in zeolite synthesis described in Section 2.1. These were first employed in the context of MOFs by the groups of Bein and Stock to study the compositional phase diagrams of metal phosphonates through parallel hydrothermal synthesis.^[256,263,264] These early studies often used a single plate and kept the number of microreactions to a minimum of 48 or 96. However, later work has often explored reaction parameters more broadly, using either larger or simply more plates, with the work of Banerjee et al. on ZIFs remaining the largest synthetic screen to

date in this field.^[22] In their study, 9600 different combinations of linker, solvent, metal source, temperature, stoichiometric ratios, and reaction times were evaluated using a hundred 96-well plates. Here the combination of automated liquid dispensers, only two metal precursors, and highly parallel solvothermal synthesis allowed for a fully systematic exploration of compositional phase space and the discovery of 16 new ZIFs.

Once all reagents have been prepared and combined appropriately, heating and aging are required to complete the MOF synthesis process. Conventional heating can be used, but this can require long aging times and subjects all samples in the multi-well plate or reaction vessel to the same conditions. Temperature gradients can be effected across a single multi-well plate using the methodology developed by Bauer and Stock, which allows for a maximum differential of 40 K.^[265] The group of Stock has also shown that microwave-assisted heating can be used in a high throughput manner.^[266–268] The benefits of this method were made clear in Maniam and Stock's study on Ni-paddlewheel MOFs, where samples synthesized with microwave heating required less aging and showed higher purity than those resulting from conventional heating, both of which are ideal in a high throughput workflow. A further technique which can be used to decrease synthesis time and improve crystallinity in a parallel manner is ultrasonication. Schilling and Stock employed this approach to isolate four new metal phosphonocarboxylates, again noting the drastically reduced aging needed to yield crystalline products.^[269] Note that for all these methods, a magnetic multi-stirrer can be used to ensure that the reaction mixtures remain homogenized. A downside of these alternate heating approaches is that they may not be compatible with standard multi-well plates; microwave vials and cell culture plates were used in the studies described above.

Following heating and aging, it is necessary to isolate the products from solution (where microreactions have been successful). The two ways this is commonly done are filtering and centrifuging.^[53] Individually doing this for hundreds of samples can be very time consuming, so tools have been developed to parallelize this process. Bauer et al. have used a custom-built 48-tube apparatus that can both dry and wash products in a single step to examine cadmium phosphonates.^[270] Plabst et al. have also described using a custom instrument to filter and wash products in an investigation of lanthanide phosphonates.^[271] Some robotic platforms can also isolate solid products from multi-well plates in an automated manner; Sumida et al. did this to recrystallize samples for powder XRD (PXRD) in their study of an iron MOF.^[259]

4.2. Characterization and Properties

High throughput synthetic screens result in large numbers of specimens which require further examination. Sometimes, a number of experiments can be discounted on the grounds of obviously non-crystalline products or failed reactions and the remainder can be characterized using conventional methods and instruments. In such cases, time-consuming techniques such as single-crystal X-ray diffraction (SCXRD) or electron microscopy (for which no straightforward parallelization schemes are currently known) can be used without significantly

throttling the study. However, if this is not the case, then more streamlined approaches with higher throughput must be relied on at least for identifying samples for further analysis, with routines such as powder PXRD being more suitable. The last 10 years have also seen the development of new tools and the acceleration of previously slow methodologies, particularly with regard to adsorption properties.

PXRD is currently the most widely applied characterization technique in the context of high throughput experiments on MOFs.^[53] It is mainly used to determine the crystallinity of specimens and the presence of possible by-products, making it suitable for rapidly pruning the results of an initial screen. The method is applied in a fast serial manner with an automated instrument containing a movable *xy* stage capable of holding multiple samples. Issues can arise when mixed or unknown products are present, as this can lead to patterns that overlap or which cannot be automatically assigned, complicating phase identification. This has also been a concern in zeolite HT synthesis (see Section 2.2),^[84] and remains an active area of research for MOFs.^[272] An alternative approach is to use an optical microscope to visually determine when reactions have been successful, as was done by Banerjee et al. for their 9200-sample study of ZIFs.^[22] A representative set of microscopy images taken this way is shown in **Figure 11**. The different panels show that crystalline products can be found using this method with little ambiguity, given that well-separated millimetre-scale crystals can be identified in each. Sumida et al. also used this method in their study on iron MOFs, where in this case, optical capabilities were integrated in the high throughput robotic workstation.^[259]

There have been attempts to supplement PXRD and optical microscopy with new methods for rapidly identifying crystalline porous materials. The Kaskel group has developed a tool which can rapidly measure the heat of adsorption of butane in samples via optical methods to ascertain porosity.^[103] They further showed that, by comparison with results from well-calibrated reference materials, the sensor readings could be used to quickly determine gravimetric butane capacities and BET surface areas. In later work, they extended this to also work with other gases such as cyclohexane, CO₂, and H₂O.^[104] This approach requires materials which have already been activated, a process which is typically time consuming and leads to poor properties if done improperly.^[273] The Long group attempted to bypass this problem by formulating a new heating protocol for use with existing thermogravimetric analysis (TGA) machines.^[274] This approach, compatible with small samples from HT synthetic screens, can rapidly identify porous samples and identify optimal activation conditions using multiple short heating–gas-adsorption cycles. However, note that some frameworks require more sophisticated approaches, such as solvent exchange or supercritical CO₂ drying, to reach optimal porosity.^[275] Nuclear magnetic resonance (NMR) has also been proposed by Chen et al. for the purpose of swiftly identifying porous products, as the sample isolation and activation steps are replaced by solvent exchange.^[276] The authors have demonstrated a strong correlation between BET surface area and transverse relaxation of spin in solvent atoms for a series of established frameworks and zeolites, noting also that by using autosampling hardware NMR can be deployed in a HT fashion.

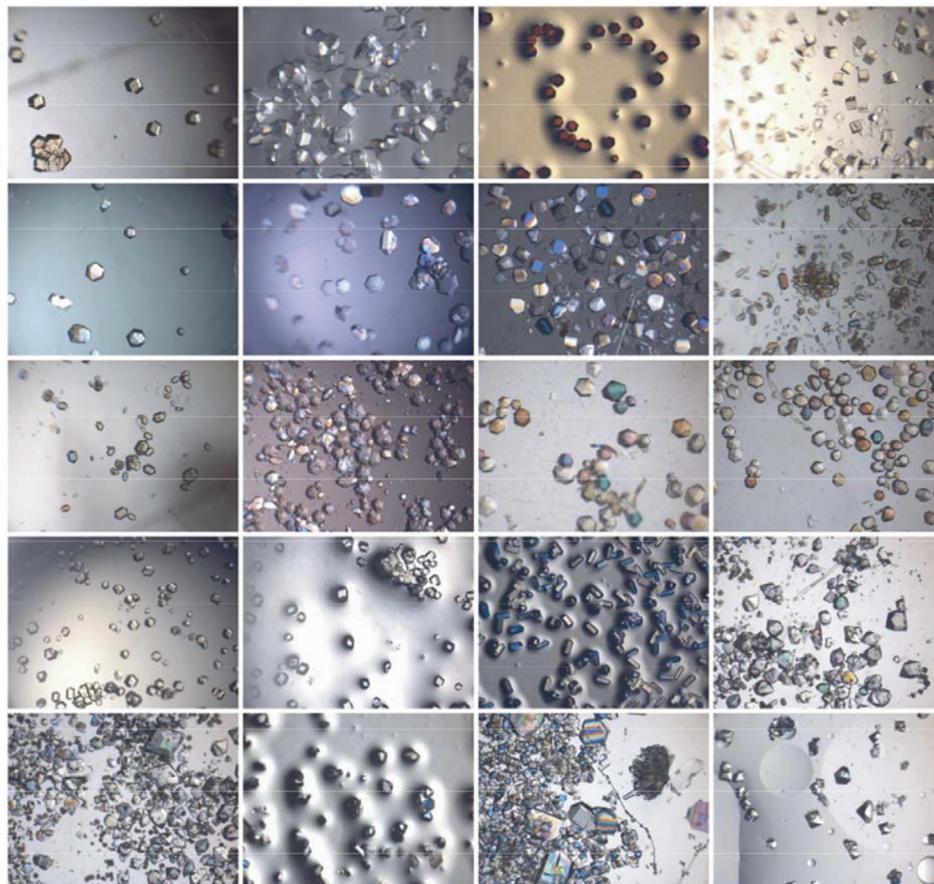


Figure 11. Optical microscopy images of successful ZIF microreactions; crystal sizes are of the order of 0.1–1 mm. Reproduced with permission.^[22] Copyright 2008, AAAS.

A set of measurements that is routinely collected for MOFs is those related to gas uptake and selective adsorption. However, standard protocols are too time consuming to be used in a high throughput fashion. To remedy this, Han et al. developed a piece of apparatus capable of recording the gas uptake of 36 samples concurrently.^[277,278] This allowed them to evaluate 16 frameworks over two studies for potential in carbon capture, from which they found that structures with nicotinate linkers showed a combination of good CO₂ selectivity and uptake along with stability toward moisture and acidic gases. Wiersum et al. also developed a new instrument capable of measuring adsorption properties in parallel.^[279] Their approach is better suited to generating full isotherms, which they did for a few simple gases in functionalized MIL-100 and CAU-10 to examine the impacts of metal centre and linker moiety on selectivity and uptake. A different approach was used by Mason et al. to rank frameworks for carbon capture, as they instead developed an instrument capable of estimating multi-component adsorption for multiple samples in parallel at equilibrium.^[166] This is faster than conventional breakthrough experiments and also avoids some of the shortcomings of this latter technique associated with non-equilibrium, so the authors argue that this is more chemically sound. With this machine, they tested 15 porous materials under flue gas conditions, finding that alkylamine moieties provide good selectivity and uptake for CO₂ while maintaining stability toward moisture.

Important features to consider for MOFs in the scope of industrial or commercial applications are thermal and water stability. While a good fraction of the reported MOFs are reasonably thermally stable, the number of water stable systems is remarkably small. Stability can be evaluated through gas adsorption experiments or by boiling samples, as reviewed by Burtch et al.^[280] However, for high throughput screens which yield many promising hits, measuring thermal and water stability is costly and time consuming. Hydrothermal stability was estimated in the studies by Han, Wiersum, and Mason through gas uptake measurements both prior to and after exposure to moisture,^[166,277,279] so methodologies exist for evaluating this property in a high throughput manner. However, note that these experiments were limited to flue gas conditions, where temperature is relatively low (20–40 °C) and water content is fixed. A more systematic approach was developed by Low et al. for mapping hydrothermal stability.^[281] In a combined experimental and computational study, the authors built a custom 48-chamber steaming instrument to examine ten different frameworks at temperatures up to 300 °C and under gas mixtures with up to 50% steam. This allowed them to build a comprehensive “steam stability map” and infer that the strength of the metal-linker bond is an important variable for hydrothermal stability. Most of the subsequent high throughput water and thermal stability work on MOFs has been computational in

nature, but more recent experimental work by Fischer et al. on zeolites (see Section 2.2) has relied on another multi-compartment piece of apparatus with similar capabilities to that of Low et al.^[107,108] A wider mapping of hydrothermal stability in MOFs would be invaluable to their development and potentially be a crucial element of the screening process.

Much research has been devoted to MOFs as drug delivery vehicles and biocompatible platforms.^[282–284] In order for such applications to be realized, it must be shown that promising frameworks are non-toxic to human cells. To this end, there have been numerous studies which have made use of dye-based assays combined with high throughput UV–vis spectroscopy to evaluate MOF cytotoxicity. The scheme used in such experiments is similar to that used for zeolites (see Section 2.3): some substance is exposed to a sample of MOF, time is allowed for potential catalytic or degradation processes to occur, and the subsequent response to UV–vis irradiation is recorded. An early study by Horcajada et al. examined 6 iron MIL-type frameworks for drug delivery and medical imaging.^[183] In order to evaluate toxicity to cell lines *in vitro*, the authors ran assays in which mouse macrophages were exposed to varying concentrations of MOFs. They subsequently used the UV–vis active MTT dye to measure the impact on the mouse cells,^[285] finding from this that their target materials were indeed safe for biological uses. A later study by Ruyra et al., using the XTT dye,^[286] focused on 16 frameworks representing a range of metals.^[287] From this, they found that cytotoxicity was often associated with bare metal ions from framework degradation.

One further use for UV–vis spectroscopy in high throughput experiments on MOFs has been for evaluating catalysis. The method has been deployed by the Cohen group to measure the catalytic breakdown of nerve gas simulants by porous materials. Palomba et al. developed a statistically robust assay system for evaluating the degradation of dimethyl 4-nitrophenol via a parallel UV–vis spectrometer.^[288] In their screen of 96 porous materials, they found that Zr-MOFs such as UiO-66 outperformed the rest. This led them to repeat their experiments using 26 different multivariate UiO-66 samples, from which they found that mixed ligands improved the catalytic breakdown of nerve gas simulants.^[289] Further work on a different simulant, involving a screen of 117 frameworks and zeolites, yielded concordant results regarding the high activity of UiO-66 and its multivariate forms.^[290] The authors have also noted that this screening method for catalytic activity could be applied for other reactions. Given the potential of MOFs as catalysts in various reaction archetypes, this approach may see more use going forward for high throughput catalysis measurements.^[291]

4.3. Exploiting Feedback Loops

In an HT screen, the marginal costs of testing additional reaction conditions are low. This can encourage the use of HT techniques as a brute-force method for exploring phase space, and as noted by Plabst and Bein, this is especially powerful when a synthetic strategy or chemical knowledge of the system is unavailable.^[292] However, it can remain fruitful to carry out syntheses in batches, deriving insight into what reaction parameters most influence the system, before adjusting and

proceeding with the next batch in a feedback process. In a study on metal phosphonates, Maniam et al. used feedback-informed multi-stage screening to discover new copper frameworks and refine their formation fields.^[293] Synthesis was carried out in three batches, shown in **Figure 12**, initially encompassing a large range of conditions. Following batch A, progressively more focused batches B and C were explored, eventually leading to very tight stability regions for three of the compounds discovered. In this case, feedback from earlier screens has allowed for a focused study of target MOFs without the time-consuming and wasteful testing of phase space which yields no products.

Recent work by the Smit group has taken this approach a step further using a computational model that captures chemical inferences from successful and failed reactions. In a proof-of-concept study, Moosavi et al. showed that by using an appropriate measure of reaction success, a genetic algorithm (GA) could be trained to refine synthesis in the same way that a chemist would using chemical knowledge from the results of experiments.^[294] The authors tested this technique on HKUST-1 using BET surface area as a measure of success, and parametrized the reaction conditions in terms of nine variables. By using the MaxMin method,^[295] the 1000 most diverse combinations of synthesis parameters were generated, the top 30 of which were used in an HT synthetic screen. This is outlined in **Figure 13**.

Following the first round of synthesis, the sample surface areas were measured, and the results were fed into the GA to generate 30 new sets of parameters (G-2). This process was repeated once more to reach a third generation (G-3), after which the authors elected to stop the procedure as a resulting sample was found to have the highest BET surface area reported to date for this material. It is clear from **Figure 13** that, as the GA proceeds, synthesis conditions are converging and reaching some optimum point. Follow-up analysis of the data using machine learning allowed weights, reflecting impact on surface area, to be assigned to each synthesis variable. The screening was repeated from scratch for the zinc analog of HKUST-1, but using the parameter weights determined from Cu-HKUST-1, and this yielded crystalline samples within a single generation. This powerful approach illustrates not only the role of feedback from prior screens to inform subsequent screens, but also the value of failed or partially successful reactions. There have been other attempts in materials chemistry to guide synthesis using automated equipment and algorithms capable of learning, such as the Cronin group's "chemputer" approach or the work of Raccuglia et al.^[296,297] More recently, Chen et al. used post hoc machine learning to analyze synthesis and characterization data from experiments on UiO-67, from which they were able to infer the relative importance of reaction parameters for crystal growth.^[298] Nevertheless, Moosavi et al.'s work remains exemplary, having elegantly combined a model which learns on-the-fly with high throughput techniques for exploring and optimizing MOF synthesis. Given that this approach can determine synthetic optima without prior chemical knowledge and with minimal human involvement, it may be a powerful potential partner to the computational methods discussed in Section 5; screening identifies target frameworks and the synthetic GA realizes them in an almost wholly automated and integrated process.

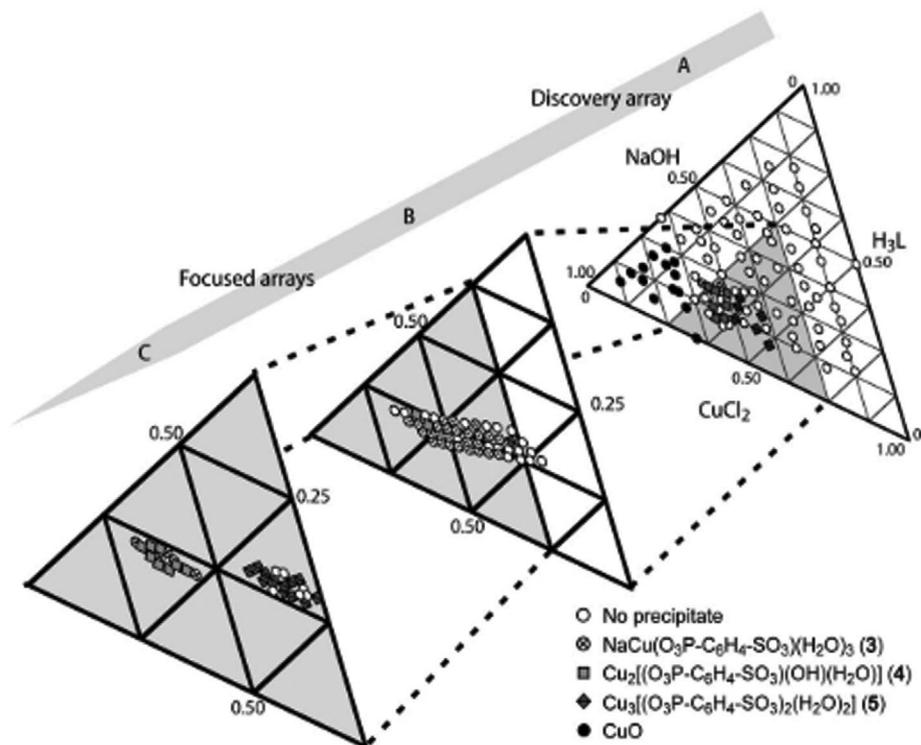


Figure 12. Copper phosphonate discovery and focused synthesis arrays. The process of synthesis focusing on a narrow part of phase space in going from array A to array C is made clear in this diagram. H₃L refers to 4-phosphonobenzenesulfonic acid; molar ratio values have been normalized to 1, and products have been identified with PXRD. Reproduced with permission.^[293] Copyright 2010, Wiley-VCH.

4.4. Reliability and Consistency

The use of automated dispensers, robots, and high throughput instruments can increase the consistency of experiments, the

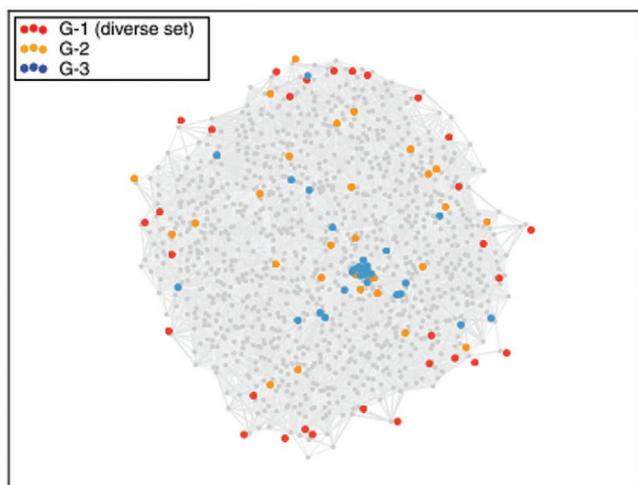


Figure 13. Library of synthesis conditions considered for the ML-assisted synthesis of HKUST-1. This representation is a 2D projection of the 9D reaction parameter space. The three generations of samples are labeled G-1, G-2, and G-3 in red, orange, and blue, respectively. The grey points represent other diverse points found with the MaxMin method. Reproduced under the terms of the CC-BY Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0>).^[294] Copyright 2019, The Authors, published by Springer Nature.

former two having been found to be more reliable than manual methods.^[299] Nevertheless, it remains important to evaluate the consistency of methodologies and collect repeats when possible, as this can partially fulfil the role of validation and allow for the estimation of errors. Given MOF sensitivity to reaction conditions, work centred on synthesis could benefit greatly from repeat runs. Despite the lower relative cost of doing this in high throughput studies, this is seldom done. Work by Biemmi et al. on reaction parameters for MOF-5 and HKUST-1 remains the main example of where this has been done.^[300] The authors repeated some of their syntheses for both frameworks, also using different positions in their multi-well plates to rule out spatial differences across their reaction vessels. This works as both an assessment of the HT method's reliability and as verification that the same conditions lead to the same products. The latter point is particularly important, as it has been found that downscaling large-scale reactions to the micro-scale sometimes leads to amorphous results or no reaction, the reasons for which are not well understood.^[256,301]

It is more common to repeat runs when measuring sample properties, namely those related to gas adsorption or catalysis. In their study on the sorption of CO₂ from flue gas, the Sholl group made use of their new parallel instrument to record uptakes in triplicate. This has also allowed them to estimate the errors in their measurements, explaining most of the variation seen in the results for adsorption prior to and after exposure to moisture and acid gases. Given that experimental gas adsorption isotherms for MOFs often lack reproducibility, this represents a best practice that both simplifies data analysis and

increases trust in the data.^[302] Moreover, such concerns present a strong case for establishing reference MOF samples, in order that adsorption measurements can be verified, as has recently been achieved in the zeolite field.^[303] The work of Palomba et al. on nerve gas degradation using porous materials also excels in its thoroughness with repeat measurements. In this case the authors first assessed the viability of their assay of catalytic activity by means of a statistical method known as the Z-factor on a subset of their target MOFs.^[304] After having determined that their level of noise was acceptable, they proceeded to take every measurement in triplicate, and were thereafter able to make robust conclusions on key framework features for this catalytic process. HT methods make repeat measurements relatively low cost and, as exemplified by these studies, this duplication of results allows for a quantification of noise and increases reliability.

4.5. Critical Remarks

To date, there has only been a relatively limited number of high throughput experimental studies focused on MOFs. There are several reasons for this, all of which reflect the weakness of high-volume approaches in this part of MOF chemistry. A prohibitively large range of different metal precursors, solvents, and linkers are currently in use, not all of which can be explored in a single screen. Although liquid dispensers can be used when all reagents are available in solution form, this still requires preparing a different stock solution for each precursor–solvent combination. Hence, experiments tend to be targeted at optimizing a single framework (e.g., the work of Sonnauer et al. on MIL-101) or toward material discovery using limited metal sources (Banerjee et al.'s 9200-sample screen used only two precursors). Using an automated solid dispenser or robot may help alleviate this problem (when investigating the synthesis of MUV-11, the group of Martí-Gastaldo examined seven different precursors), but concerns remain regarding inconsistent performance amongst powder types, non-negligible errors at small masses, and rates of machine stall.^[305] For both solid and liquid automated dispensers, high equipment cost may also remain a barrier to entry. Additionally, if bespoke starting materials (such as an exotic linker or precursor) are required, these need to be synthesized and functionalized prior to any experiment. Further concerns can include long reaction times and the lengthy procedures of washing and activation to remove solvent and impurities from the products, though this is highly dependent on the method of synthesis.^[273,306] Some groups have developed ways of accelerating this step, but these practices have not yet become widespread as they often involve custom-made equipment.

The same issues are also pervasive in property-based screening studies. In order to measure adsorption or catalytic properties across a large number of porous networks, they must first be synthesized. This again limits the sample set to either a limited number of frameworks, or variations of a handful of syntheses at once. Following this, instruments for high throughput characterization must also be available. Some techniques are currently not amenable to parallelization, such as SCXRD and electron diffraction, though ongoing efforts

may eventually streamline such methods sufficiently for HT use.^[191,194] In some cases authors must develop entirely new equipment in order to carry out their screening, which further requires initial benchmarking and validation (as discussed in Section 4.2). The synthesis stage will be limited by the number of precursors, the presence of solid reagents, and the reaction method, but will be fast if these parameters are carefully chosen. Subsequent washing and activation (particularly important if adsorption properties are of interest) are also likely bottlenecks. The results of an initial screen, if unsatisfactory, may also prompt secondary screens. Characterization is rapid if the methods chosen can make use of small samples (on the milligram scale), otherwise scale-up and instrument set-up will require manual involvement. The commercial availability of porous materials, such as activated carbon or certain zeolites, can accelerate a study, though these are usually used as reference materials or for calibration purposes.

Despite these weaknesses, it has been shown that high throughput methodologies are effective in MOF experimental work. Parallel synthesis is a powerful approach for studying a single framework at a time, particularly when the aim is to optimize conditions so as to tune a given material property. Automated dispensing techniques and robotics may also pair very well with computational screens, as the use of hypothetical databases can lead to the identification of promising networks which have not yet been synthesized. In such cases, as shown by the work in Section 4.3, even without a chemical starting point from which a viable synthesis could be derived, a model-backed approach could quickly yield crystalline samples. Even when materials discovery is the objective, the isorecticular chemistry of many MOFs means that the combination of a small number of precursors and numerous linkers can still lead to a substantial array of new frameworks. Machine learning and GA methods have been sparingly used in MOF synthesis, but the success of Moosavi et al.'s HT study, along with similar developments in other fields, may drive their more widespread adoption for these materials. Indeed, the most successful high throughput MOF studies of the future are likely to be those involving parallelization and streamlining at every stage, from computational screening to rapid synthesis and characterization, powered by feedback loops that enable more targeted secondary screens.

5. Computational Screening of Metal–Organic Frameworks

The combinatorial nature of MOFs, constructed by self-assembly of inorganic nodes and organic linkers, makes them exciting materials due to the tunability of their chemical composition and structure. However, it also poses a significant challenge to modeling approaches as the search space of all possible frameworks is intractably large. The ever-increasing power of high performance computing and the creation of large repositories of experimentally determined and/or computer-generated structures has led to a paradigm shift of *in silico* materials science over the last decade. HTCS has emerged as an invaluable asset to the scientific community, allowing fast and accurate property prediction of up to hundreds of

thousands of structures, expediting the time frame between hypothesis and discovery. We note that due to the diversity of potential applications of MOFs, HTCS can occur at different stages of a workflow. We also note that we have to take a flexible definition of “high throughput” in the context of computer simulation. Screening based on geometric properties of MOFs can take fractions of a second for a single MOF (or zeolite) to be examined; classical approaches (FF) involving geometry optimization, Monte Carlo/molecular dynamics would typically be minutes to hours timescale per MOF and *ab initio*/DFT geometry optimization, Monte Carlo/molecular dynamics could be tens to hundreds of hours. Such computational costs imply a practical limitation to the number of calculations that can be performed on a tractable timescale. Depending on the application, potentially all three methods may be used as part of a hierarchical sift, but each one has associated limitations. In this section we highlight some of the most critical aspects in the HTCS of MOFs: different parts of the screening process, including the construction and mining of structural databases, and the limitations of classical force fields (transferability and assignment of charges). These issues are then discussed in the context of adsorption and separation applications, the most heavily studied potential uses of MOFs.

5.1. Databases

HTCS of MOFs usually necessitates parsing thousands of crystal structures to identify frameworks for use in a particular application; this requires large repositories of MOF structures to be available. There are a plethora of available databases to choose from, comprising either experimentally determined or computer-generated frameworks. Each of these repositories come with their own set of limitations, such as: differences in the algorithms used for solvent removal and treatment of charge compensating ions in the experimentally reported structures, and the synthetic feasibility and restricted topological diversity of hypothetical frameworks. This poses the question of how to identify the relevant database for a certain application, which we now discuss.

5.1.1. CSD-Derived Databases

Newly synthesized frameworks typically have their structures deposited in the Cambridge structural database (CSD) (<https://www.ccdc.cam.ac.uk>), amongst hundreds of thousands of crystal structures, with no label identifying them as MOFs, leaving them as needles in a haystack. To remedy this, methods have been developed for identifying, extracting, and readying frameworks from the CSD for visualization or simulation. Search criteria are employed to identify MOFs and post-extraction approaches focus on whether to remove bound or unbound solvent and algorithms to treat disorder. Watanabe and Sholl sought to identify structures for applications in the high throughput screening of CO₂/N₂ separation. To this end, they extracted and screened ≈30 000 MOFs from the CSD, leading to 1163 frameworks being examined for this separation process.^[307] Further to this, Goldsmith et al. published an

automated approach for screening 20 000 frameworks from the Cambridge repository for their use in hydrogen storage applications.^[308] In 2014, and more recently 2019, Chung and co-workers published a database of MOFs tailored to contain those for use in adsorption processes.^[309,310] The construction process of these Computation Ready, Experimental Metal–Organic Framework (CoREMOF) databases consisted of extracting only porous networks with 3D channel structures and PLDs greater than 2.4 Å; the most current CoREMOF database includes ≈14 000 structures and more than 350 unique topologies. Additional computational data has been added to the database, such as the work by Nazarian et al. who published *ab initio* derived point charges for ≈2900 structures in the CoREMOF database.^[311]

Whilst development of these methods aids researchers in the identification of structures for study, they are external to the CSD and require manual updating for identification of new porous structures as they are deposited. Recent work by Moghadamm et al. sought to develop an integrated MOF subset of the CSD; this was done by establishing a set of seven “look for MOF” search criteria implemented in a custom CSD Python Application Programming Interface (API) workflow.^[312] Analysis showed that disorder was present in a number of the identified structures, leading to the creation of the CSD non-disordered MOF subset. These two lists are integrated into the CSD’s structure search software, ConQuest,^[313] and currently contain over 90 000 entries; where the subsets are automatically updated with newly deposited structures.

These studies all parse the Cambridge repository with search criteria based on atomic types and bonds present, and remove or manipulate structures that contain disorder intrinsic to the frameworks; a disparity between them is the treatment of charge compensating ions and the removal of solvent bound to the structure (seen in **Table 2**). The approximations made are valid on a case-by-case basis and choosing the appropriate database, or automated workflow, is dependant on the application under study. Examples include the CoREMOF^[309] treatment of bound solvent, where all identified solvent molecules are removed even if this results in undercoordinated metal sites. This is a reasonable approximation if the process that the simulation seeks to model occurs at high temperature or under non-humid conditions. Alternatively, processes at low temperature and/or under humid conditions are likely to see water vapor bind within the MOF pore, altering the adsorption properties of the framework and consequently discrepancies between simulation and experiment are likely to arise.

5.1.2. Computer-Generated Databases

Aside from the CSD-derived databases, there is a set of repositories containing computer-generated frameworks. In 2011, Wilmer et al. constructed 137 953 hypothetical MOFs (hMOFs) using crystal enumeration algorithms and a library of 102 building units (BU) extrapolated from existing crystallographic data;^[314] this has been dubbed the hMOF database. The BUs consisted of 5 metal clusters, 42 linkers (terminated with either nitrogen atoms or carboxylic acid groups), and 13 functional groups. Each structure was generated with no more

Table 2. Techniques for solvent removal, charge compensating ion (CCI) treatment and removing/fixing disorder used in the development of the CSD-derived databases (or automated workflow). In every study tabulated here free solvent was removed.

Study	Bound solvent	Charge compensating ions	Disorder present in framework
Goldsmith et al. ^[308]	Removed	Structures with CCIs removed	Structures removed
Chung et al. ^[309] (2014)	Removed	CCIs retained for charge neutrality	Structures fixed if possible or removed
Chung et al. ^[310] (2019)	Two subsets: one with bound solvent removed, the other without	CCIs retained for charge neutrality	Structures fixed if possible or removed
Moghadamm et al. ^[312]	Removed only if bound to Cu-paddlewheel or MOF-74 type clusters	Not currently considered	Removed in the case of the non-disordered subset

than four unique BUs and no consideration was taken for post-construction optimization. In order to validate the hypothetical frameworks, methane storage performance was computed and compared between a subset of the constructed MOFs, their energetically relaxed counterparts and their experimentally reported analogs; once validated the entire database was screened for high-pressure room-temperature methane adsorption.

In a similar vein, Aghaji and co-workers generated 324 500 frameworks from a library of 90 BUs,^[315] comprising 70 inorganic/organic building units and 20 functional groups. The authors excluded interpenetrating structures from their construction process and functionalized linkers only in symmetric hydrogen positions to enhance their synthetic viability. The generated structures were then energetically relaxed with the universal force field^[316] and screened for their CO₂/CH₄ sorption selectivity.

Colón and co-workers implemented a different approach for constructing hypothetical frameworks. As opposed to the bottom-up crystal enumeration used in the aforementioned studies, the authors used a reverse topological approach (RTA)^[317] to generate over 13 500 hypothetical MOF structures with 41 different edge-transitive topologies. Their implementation of this approach, where constituent building units are mapped onto topological blueprints in a “top-down” fashion, has been published in the ToBaCCo code;^[318] the structures’ use has been demonstrated by assessing them for their hydrogen and methane storage capabilities, as well as their Xe/Kr sorption selectivity.

5.2. Computational Methods

The calculation of textural, topological, and adsorption properties underpins HTCS procedures. For an overview of the methods used, we refer the reader to Sections 3.2.1, 3.2.2, and 3.2.3 where geometric and topological descriptors, the hierarchy of simulation methods, and the validity of some common approximations are discussed. The important input parameters when computing adsorption properties with classical approaches (such as GCMC and MD) are the force fields that describe both the framework atoms and the guest molecules. In standard molecular mechanics type force fields, there are bonding terms and non-bonding terms; the latter are typically partitioned into the electrostatic terms and van der Waals terms. It is common practice to simulate MOFs as rigid,^[314,315,319] taking the van der Waals parameters from generic force fields, such as the Universal^[316] and Dreiding^[320] force field. The issue

of assigning partial charges is more complex and is discussed in Section 5.3. Adsorbate molecules are often described by empirical force fields fit to match the vapor/liquid coexistence properties of the molecule;^[321] however, when these schemes prove to be inadequate, first principle-derived force fields^[322,323] or corrections to account for quantum effects may be deployed.^[324,325] Notably, Verploegh et al.^[326] have derived an enhanced force field for assessing the diffusion of small molecules in MOFs (the specific study was on ZIF structures). The force field was fit to DFT data and gives similar quality of forces to those extracted from periodic ab initio MD studies. Hence, the reported force field lends itself to HTCS of either single or multi-component studies of diffusion and breakthrough measurements, as well as competitive diffusion which is relevant to membrane separation applications.

5.3. Charge Assignment

One of the most significant challenges to developing transferable force fields for MOFs is the issue of charge. MOFs can accommodate, in practice, almost all elements of the periodic table. In zeolites, the framework nature of the materials coupled with their semi-ionic nature and relatively limited atom types means that the parameterization is intrinsically quite tightly constrained. In MOFs there is the possibility of under-coordinated transition metals (exhibiting unusual Jahn-Teller distortions) in varied oxidation states, within the same framework. Added to the mix is the organic linker which can display varied degrees of charge transfer depending on the oxidation state of the metal. This complexity clearly pushes the limitations of fixed charge force fields, although adaptive force fields such as ReaxFF^[327] might be promising if adequately trained on ab initio data, for example.

The difficulty of finding a method that can accurately model charge distributions in a wide variety of chemical and structural compositions means that assigning these charges is an onerous and hazardous task. Whilst charges can be assigned through DFT-derived partial charges, for example,^[328–330] transferability is important and these methods require compute intensive calculations for each framework being considered. It becomes unfeasible to perform calculations on samples consisting of tens or hundreds of thousands of structures. To remedy this, Zhong and co-workers developed a connectivity-based atom contribution method (CBAC) for fast assignment of partial atomic charges; in this approach it is assumed that atoms with the same bonding connectivity have identical charges across different MOFs.^[331,332]

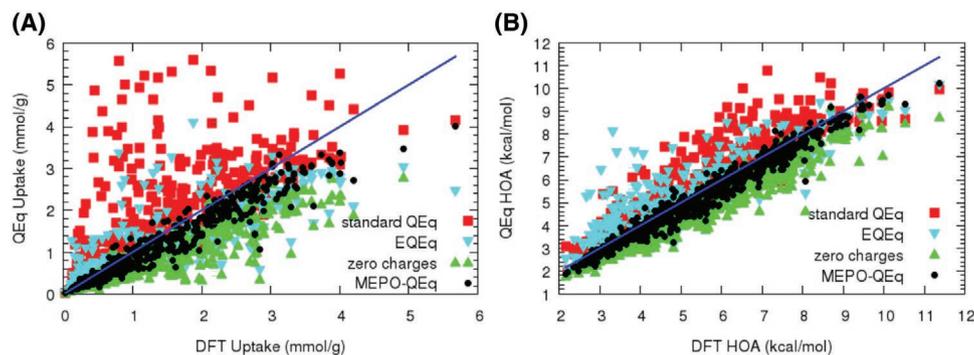


Figure 14. A) CO_2 uptake and B) heats of adsorption at 0.15 bar, 298 K for the test set of MOFs obtained from the GCMC simulations using standard QEq, EQEq and MEPO-QEq charges, zero charges, and DFT-derived charges. The line of perfect correlation is shown in dark blue. A,B) Reproduced with permission.^[335] Copyright 2013, American Chemical Society.

The authors computed CO_2 adsorption isotherms using the CBAC method on a small test set of frameworks and good correlation to those calculated with ab initio derived charges was observed. Other methods have been developed by building on the work of Rappe et al., who developed a charge equilibration scheme (QEq) for predicting charge distributions in molecules.^[333] QEq assigns charges that are based on the molecular geometry of the system and the experimentally determined atomic ionization potential, electron affinities, and atomic radii. Several groups, such as that of Wilmer and co-workers, have built upon the work of Rappe et al. by developing an extended charge equilibration method (EQEq), taking into account all measured ionization energies for every atom in the periodic table.^[334] Moreover, Kadantsev et al. parameterized the QEq method (MEPO-QEq) to reproduce ab initio derived electrostatic potentials in a diverse training set of 543 MOFs.^[335] Using a test set of 693 structures, the parameterization was validated by means of cross-comparison between CO_2 uptake and heats of adsorption calculated with charges from the MEPO-QEq method and those derived from DFT.^[329] The influence of charge partitioning schemes on simulation results can be seen in **Figure 14**, where the data produced using different variants of the QEq method is compared to data computed with ab initio derived charges.

5.4. Applications

In this section we highlight HTCS procedures applied to either computer generated or experimentally reported MOFs for the identification of high performance candidates. In particular, we emphasize studies that broaden our knowledge of combinatorial MOF space, restrict the search space necessary for candidate identification using well-informed filters, and present key descriptors for optimizing material properties. Moreover, we will discuss limitations of some of the current methodology and summarize the rate determining steps in these screening procedures.

5.4.1. DFT Virtual Screening

High throughput DFT is not commonplace on MOF structures, owing to their structural complexity, diverse compositions, and

large unit cells, often containing hundreds of atoms. Hence, DFT is yet to be routinely performed over sample sizes consisting of tens or hundreds of thousands of MOF structures. Early work has been conducted in this field on two pertinent aspects of MOFs, hydrothermal stability and the competitive adsorption of water. Low et al. tackled the issues of MOF hydrothermal stability by devising a cluster model approach, where linking ligands were replaced with capping species containing the same functional group bound to the metal.^[281] This model was applied to eight experimentally realized frameworks and DFT was used to compute energies associated with hydrolysis and ligand displacement. When this data was compared with experiment data it was found that the activation energies of ligand displacement served as a useful approximation of the relative water stability of the small sample of MOFs used in their study. Further to this, Capena et al. investigated the effects of searching composition space, with respect to MOF constituent metals, on the adsorption of H_2O , CO_2 , CH_4 , and H_2 .^[336] The authors focused on MOF-74 as it is known to have a high density of unsaturated metal sites, which have been shown both experimentally and computationally to interact strongly with various adsorbates.^[337,338] The group proceeded to substitute Zn-MOF-74 with 25 different metals and subsequently optimized the MOF-74 analogs. The outcome of the DFT calculations led to the discovery of five M-MOF-74 structures, where M=Rh, Pd, Os, Ir, or Pt, that showed preferential binding of CO_2 over H_2O . In a more recent study, Rosen and co-workers published a fully automated periodic DFT workflow for assessing optimal MOF candidates for use in catalytic processes.^[339] As a proof-of-concept the authors applied their procedure to screen MOFs containing unsaturated metal sites from the CoREMOF^[309] database for their use in the oxidative C–H bond activation of methane.

5.4.2. Hydrogen Storage

Alternatively, classical molecular simulation has proven to be an invaluable tool for property prediction, and when used with structure databases it allows HT investigation of the physicochemical properties that influence adsorption in porous materials. Inspired by the prospect of MOFs as hydrogen fuel delivery media, one of the most widely studied phenomena

is the adsorption of molecular hydrogen.^[340] In attempts to optimize hydrogen storage and deliverable capacity by increasing the MOF-H₂ heat of adsorption through functionalization, Colón and co-workers used a bottom-up approach to generate over 18 000 MOFs and porous aromatic frameworks (PAFs) composed of linkers functionalized with various numbers of magnesium alkoxide sites.^[341] GCMC was deployed to screen the structures for their hydrogen deliverable capacities at 243 K and a pressure swing between 100 and 2 bar. Since generic force fields are not sufficiently accurate to model the strong interaction of H₂ with Mg, a first principles derived force field was employed to model this phenomenon. The authors demonstrated that there is a fundamental limitation that prevents a structure from having large volumetric and gravimetric deliverable capacities simultaneously. Optimal gravimetric deliverable capacity was exhibited in structures with low framework density and insertion of relatively heavy Mg sites did not improve this property. On the other hand, optimal volumetric deliverable capacity was exhibited in structures with a balance between void fraction and material density.

Furthermore, work by Sikora et al. in analyzing the hMOF database^[314] showed that the chosen building units of this bottom-up approach produced only six topologies, where the majority were the primitive cubic unit (pcu) net (over 90% of MOFs in the database).^[342] Gómez-Gualdrón, Snurr, and co-workers took note of this work and examined geometric dependencies of hydrogen storage in a topologically diverse sample of over 13 500 hMOF structures, generated using an RTA,^[317] and assessing them for their deliverable capacity between 100 bar/77 K and 5 bar/160 K.^[343] The authors found that volumetric and gravimetric deliverable capacities were inversely related, where this was linked to the topologically dependant trade-off between volumetric and gravimetric surface areas.^[344] It was discovered that different topologies reach a maximum volumetric deliverable capacity at different linker lengths. Since topology inherently captures additional spatial information on the relationship between the local geometric features given by textural properties, it is desirable to use framework topology in tandem with textural descriptors as design variables

for novel high performance frameworks. Further to this, the group validated their automated MOF construction process by successfully synthesizing four “she” topology frameworks; comparison of the empirically determined and simulated PXRD patterns of these structures showed the computational predictions were consistent with experimental observations.

In more recent work, Siegel and co-workers conducted the largest screening of MOFs for hydrogen storage to date.^[345] Half a million frameworks were collated from 11 published databases, including CSD-derived and computer-generated structures. In an effort to offset the temporal cost associated with brute-force screening of their entire sample, the semi-empirical Chahine rule^[308] was used to estimate total gravimetric and volumetric deliverable capacities of each structure, refining the search space of their aggregated repository to 43 777 frameworks. Subsequently, the frameworks underwent further evaluation by GCMC simulation to compute deliverable capacities of each framework at cryogenic operating conditions. The pseudo-Feynman–Hibbs model for H₂ was used to account for quantum effects that are expected to be significant at low temperature. The authors used the record holder for hydrogen balanced storage capacity, IRMOF-20, as a benchmark; this yielded 102 CSD-derived frameworks and 5957 hypothetical structures that exceeded their benchmark in usable pressure swing deliverable capacities. In order to empirically verify their screening procedure, two real MOFs and one hypothetical structure were synthesized and tested. The resulting experimental data was in good accord with the simulated adsorption isotherms. Moreover, analysis of the full data set showed a theoretical usable volumetric capacity ceiling at ≈ 40 g-H₂ L⁻¹, highlighting a need for the design of new frameworks with respect to high volumetric usable capacity.

The difference in hydrogen deliverable capacities between experimentally determined and computer-generated MOFs can be seen on inspection of **Figure 15**. Functional groups are explicitly considered as building units in the construction of some hypothetical databases,^[314,315] meaning their may be higher numbers of functionalized hypothetical structures than pristine frameworks. This possibility coupled with the limited

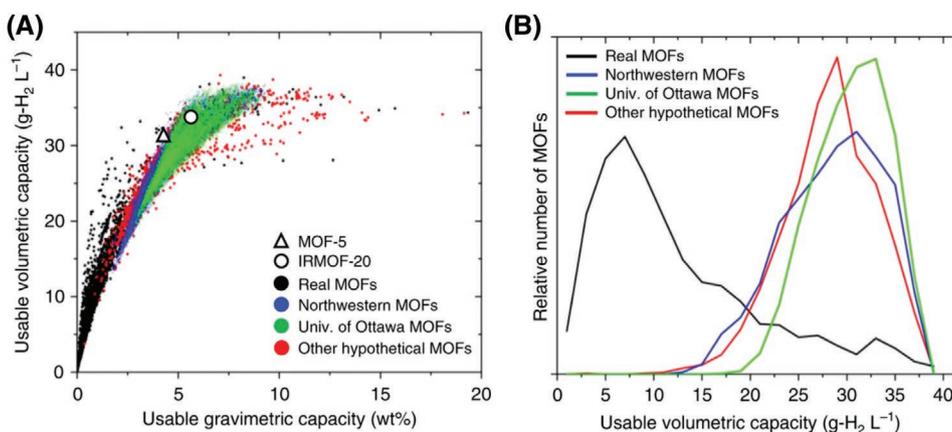


Figure 15. For MOFs in the 11 databases used by Siegel et al.,^[345] (A) shows their hydrogen volumetric and gravimetric deliverable capacities and (B) displays the probability distribution of usable volumetric deliverable capacities of the real and hypothetical frameworks. A,B) Reproduced under the terms of the CC-BY Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0>).^[345] Copyright 2019, The Authors, published by Springer Nature.

number of framework topologies^[342] and metal clusters present in the hMOF repositories is seen to be beneficial for this application. However, the restricted topological diversity in particular hinders the transferability of these hypothetical frameworks to shape-selective applications, such as catalysis. Further work on top-down approaches to increase the size, topological, and structural distributions of these databases will allow HT structure–property information to be produced for a wider range of applications over larger sample sizes.^[318] Moreover, the data produced by studies such as this could benefit from scrupulous analysis by means of machine learning algorithms (discussed in Section 5.5). This could aid in the identification of optimal framework property combinations in unexplored regimes, which could constitute design criteria for new high performance structures. However, depending on the repository used in data generation the statistics produced may be biased toward the dominant topology present.

5.4.3. Methane Storage

Methane is another important gas fuel and was heavily studied at the beginning of the last decade, motivated by applications such as natural gas vehicles.^[346] The pioneering work by Wilmer et al. in the automated construction of hypothetical MOF structures (and the hMOF database), which were subsequently analyzed for their methane storage performance,^[314] has been further built upon by Gómez-Gualdrón and co-workers.^[347] Approximately 120 pcu MOFs were taken from the hMOF database, supplemented by 39 idealized carbon-based porous materials, and using GCMC the authors explored the limits of methane deliverable capacity. An investigation into the effects of manipulating the well depth, ϵ , of their Lennard-Jones parameters describing the framework was conducted. This was done by multiplying ϵ by 2 and 4 to homogeneously increase MOF–methane interaction strength in order to approximate functionalization that maintained structural characteristics. From this, several structure–properties relationships were identified, it was found that at higher values of ϵ : i) the lower bound of the range of volumetric surface area necessary for optimal deliverable capacity decreased and ii) the optimal pore size range shifted toward larger pores, suggesting that functionalization only increases deliverable capacity past a minimum pore size threshold. Further to this, the authors derived an analytical equation, based on MOF and methane properties, that could successfully predict GCMC-simulated methane deliverable capacities for 95% of their repository with an error of $50 \text{ cm}^3(\text{STP})\text{cm}^{-3}$.^[348] Whilst not a limitation-free metric, analytical predictions such as these are effective descriptors for reducing the search space needed for the identification of high performance candidates. The methods of database construction mentioned in the studies above are based on known building units extrapolated from crystallographic data;^[314,315] these techniques rely on exploration of MOF space using libraries of pre-existing linker species. This is a limitation of the existing databases as sampling of composition space is restricted by the input of pre-defined building units (often commodity chemicals). This has been addressed by Bao et al. who developed a de novo evolutionary algorithm to explore the combinatorial

space of linker molecules in order to optimize methane deliverable capacity in predicted MOFs.^[349] The method explicitly considers the synthetic viability of the linker species by using known chemical transformations and a precursor library of commercially available molecules for an in silico search of linker space. The algorithm is initiated with a population of 100 linker molecules and in each generation a linker molecule is subject to several filters; if the linker species passes all filters an MOF is built with a selected topological net. The newly constructed MOF is evaluated for methane deliverable capacity with a pressure swing of 65–5.8 bar, at 298 K, and the linker is inserted into the population in rank order, whilst maintaining a population size of 100; hence, the lowest rank is discarded. Using this approach and MOF-5 as a benchmark, the authors found 48 predicted MOFs in four nets, amongst the nine used, having higher deliverable capacity than their benchmark material. Instead of exhaustively screening large databases of porous materials, the authors have evolved MOFs in composition space whilst taking into account the known constraints of chemical synthesis; providing a rare connection between HTCS and synthetic chemistry.

5.4.4. Other Adsorption Applications

Genetic algorithms (GA) have also been deployed by Collins and co-workers who developed a GA that makes use of experimentally realized MOFs.^[350] Their algorithm searches materials space with respect to the functionalization of these structures in order to optimize adsorption properties of interest. The GA was applied in an effort to maximize CO_2 uptake in 141 experimentally characterized frameworks; the myriad of combinations possible due to the functionalization of the linker species led to a total search space of 1.65 trillion structures. Thirteen GA parameters were optimized using three properties: CO_2 uptake, surface area, and parasitic energy^[239] (which been discussed in Section 3.3.2); where the gas adsorption properties were determined from GCMC simulations. A unique mutation algorithm was employed that replaced a chosen functional group with a chemically similar analog, determined by electrostatic and van der Waals potentials, and local steric availability. Remarkably, CO_2 uptake was explicitly calculated for only $\approx 580\,000$ frameworks to screen the entire search space; leading to the identification of 1035 derivatives of 23 parent structures that displayed exceptional CO_2 uptake.

In a recent study, Li et al. evaluated ethanol as a working fluid for alcohol-based adsorption driven heat pumps.^[351] This was done by using GCMC integrated with standard thermodynamic equations to screen ≈ 2900 MOFs with high-quality ab initio derived point charges from the CoREMOF^[311] database. A three-tiered screening process was deployed, systematically increasing the number of MC steps in each simulation round, and simulating: evaporation, condensation, and desorption. From this physicochemical properties were extricated that influenced the coefficient of performance for cooling (COP_c). Analysis showed frameworks with LCDs between 10 and 15 Å correspond to a high ethanol working capacity and relatively low enthalpy of adsorption, which maximizes the COP_c value. Moreover, principal component analysis and decision tree modeling were employed to determine the dominant features

Table 3. Adsorbent evaluation criteria used by Bae and Snurr to assess the effectiveness of porous materials for CO₂ separation and capture.^[352] N, γ and the superscripts “ads” and “des” refer to number of molecules, the mole fraction in the gas phase, adsorption, and desorption conditions, respectively. Reproduced from Bae and Snurr.^[352] Copyright 2011, Wiley-VCH.

Criterion	Definition
CO ₂ uptake [mol kg ⁻¹]	N_1^{ads}
Working capacity [mol kg ⁻¹]	$\Delta N_1 = N_1^{\text{ads}} - N_1^{\text{des}}$
Regenerability [%]	$R = (\Delta N_1 / N_1^{\text{ads}}) \times 100$
Selectivity	$\alpha_{12}^{\text{ads}} = (N_1^{\text{ads}} / N_2^{\text{ads}}) / (\gamma_2 / \gamma_1)$
Sorbent selection parameter	$S = (\alpha_{12}^{\text{ads}})^2 / (\alpha_{12}^{\text{des}}) (\Delta N_1 / \Delta N_2)$

affecting performance, revealing that LCD, working capacity, and enthalpy of adsorption were key descriptors for COP_C. In particular pore size played a dominant role in the COP_C value for MOFs with a low working capacity, whilst enthalpy of adsorption dominated in those with high working capacity.

5.4.5. Carbon Capture and Separation

Modeling the separation of gaseous species is a more complex task than pure component adsorption, as either multi-component simulations must be conducted or approximations using pure component data must be employed to assess a material's capacity. Diffusion effects can also play a more dominant role in separation processes and the metrics used to rank promising candidates must be defined on a case-by-case basis. One of the most intensely studied separations is that involving CO₂ capture, as the development of carbon capture and separation technologies is of high importance to mitigate greenhouse-gas emissions.^[352] Using their previously constructed database of over 130 000 hypothetical structures, Wilmer and co-workers computed pure component adsorption data for CO₂, CH₄, and N₂; the results were used to calculate five adsorbent evaluation criteria (shown in Table 3) for four different separation cases, based on pressure swing adsorption (PSA) and vacuum swing adsorption (VSA) processes.^[353] The resulting data was scrupulously analyzed to identify structure–performance relationships, which revealed trends that were not apparent from studies with fewer samples. In the case of landfill gas separation via the PSA process, optimal working capacity was achieved in structures with heats of adsorption ≈ 21 kJ mol⁻¹ and optimal selectivity was achieved in MOFs with void fractions ≈ 0.6 to 0.8 . On the other hand, for the sequestration of CO₂ from flue gas via the VSA process, where CO₂ has a lower partial pressure, selectivity was optimized in structures with void fractions in the range ≈ 0.3 to 0.4 .

The competitive adsorption of water in MOFs is an important consideration when modeling processes that occur under humid, or even ambient, conditions. This aspect is often neglected in simulations, probably due to the complexity of accurately modeling water adsorption. However, Li et al. tackled

this issue by screening ≈ 5000 CoREMOF^[309] structures for their CO₂/N₂/H₂O sorption selectivity.^[354] The authors' screening procedure consisted of: i) efficiently assigning framework charges with the EQEq method, ii) employing Widom particle insertion to calculate Henry constants for each species, iii) refining the search space through selectivity-based ratios of CO₂/H₂O Henry constants, iv) recalculating framework charges with the higher accuracy REPEAT method^[329] for 15 top performing MOFs, v) deploying GCMC to compute binary CO₂/H₂O and ternary CO₂/N₂/H₂O sorption selectivity. By cross-comparison of the data produced by the two charge schemes, it was found that Henry constants for H₂O were more sensitive to the choice of method for estimating partial charges than CO₂ and N₂. Additionally, the MOF–guest interaction energies showed that Coulombic interactions contributed a higher fraction of the adsorption energy of H₂O than the other two non-polar adsorbates, again emphasizing the importance of selecting an internally consistent approach for calculating framework charges when modelling competitive adsorption.

The inorganic/organic hybrid nature of MOFs provides the possibility of highly tunable structures. Searching the combinatorial space of these frameworks by introduction of multiple linker types and functionalities can have synergistic effects. Multivariate MOFs (MTV-MOFs), those containing multiple linker types within one structure, have been experimentally realized for their selective sorption capabilities.^[355] Later, Li et al. built on this idea by conducting a systematic large-scale screening of ≈ 10 000 computer-generated MTV-MOFs to probe the structures for their CO₂/N₂ sorption selectivity and CO₂ capacity.^[356] The construction process consisted of generating pcu topology frameworks, with copper paddlewheel nodes and 20 organic linkers; each MTV-MOF consisted of three linker types, functionalized with –F, –NH₂, and –OCH₃. 560 unfunctionalized parent frameworks and 10 monolinker MOFs were constructed for cross-comparison. CO₂ uptake and CO₂/N₂ sorption selectivity were computed with single component and multi-component GCMC, respectively. By grouping MTV-MOFs by families of unfunctionalized parent structures and taking the average of the adsorption properties per family, the authors found that the functionalized derivatives exhibited better CO₂/N₂ selectivity and higher CO₂ capacity than their parent counterparts; except for seven frameworks all with largest cavity diameters between 5 and 6.5 Å. The enhancement in CO₂ selectivity and capacity caused by functionalization was maximized in MOFs exhibiting small pore geometries. However, if the pores are too small then functionalization blocks accessibility and causes a reduction in gas adsorption, revealing important descriptors for the design of new high performance frameworks.

Another important application involving the separation of gas mixtures is the pre-combustion processing of high pressure streams of CO₂/H₂ mixtures.^[357] Chung et al. sought to identify top MOF candidates for the selective uptake and sequestration of CO₂ from the aforementioned mixture.^[358] A novel aspect of this HTCS approach was the development of a GA for the in silico discovery of high performance porous networks for this application. Starting with an initial population of 100 hMOFs, each generation was evolved by implementing elitism and applying genetic operations to hMOF pairs to form a subsequent generation (maintaining a population of 100 frameworks

Table 4. Comparison of relative computational expense for brute-force search versus GA. ΔN_1 is the CO₂ working capacity, α_{12}^{ads} is the CO₂/H₂ selectivity, and APS is the adsorbent performance score. Reproduced from Chung et al.^[358] Copyright 2016, AAAS.

Method	Fitness measure	Number of GCMC simulations	Relative compute time [%]
Brute force	—	51 163	100
	ΔN_1	340	0.66
GA	α_{12}^{ads}	322	0.63
	APS	268	0.52

at each step). Three independent GA runs were performed to optimize the three fitness measures, namely CO₂ working capacity, CO₂/H₂ selectivity, and adsorbent performance score (which is the product of the former two fitness measures) computed by “on the fly” GCMC simulations. The GA runs produced data for the fitness criteria of 730 unique hMOFs, of which ~50 frameworks showing high performance for this application were extracted. The remarkable aspect of this GA-guided search was the reduction in computational cost relative to a brute-force screening approach, which is shown in Table 4.

5.4.6. Alternative Separation Materials and Processes

Whilst MOFs have shown great promise as adsorption-based carbon capture technologies, recently, an area of interest for HT researchers has been mixed matrix membranes (MMMs) as potential candidates for CO₂ capture. MMMs consist of polymer membranes with inorganic particles (in this case MOFs) dispersed throughout the polymer matrix. Budhathoki and co-workers conducted a multi-scale study, computing the selectivity of CO₂/N₂ mixtures, CO₂ permeability, and carrying out techno-economic evaluations, presenting a novel connection between atomistic MOF structures and the cost of carbon capture (CCC).^[359] The authors screened both CoREMOF^[309] and Wilmer’s database,^[314] and in order to take into account whether MOFs in the MMMs were CO₂/H₂O sorption selective or not, the selectivity data from Li et al.^[354] was used to rank real MOFs for their selectivity. By calculating the permselectivity of each framework and employing the Maxwell model,^[360] along with experimental data for nine polymers, the gas permeabilities of over a million MMMs were computed. The authors went on to use process modeling to assign a predicted CCC to each hypothetical MMM in their repository. Their analysis showed MOFs with LCD in the range 4–10 Å and PLD in the range 4–5 Å had superior adsorption and diffusion selectivity, respectively. Furthermore, techno-economic evaluation found 1153 MMMs were predicted to yield a low CCC, 16 of which were based on CoREMOF structures with favorable CO₂/H₂O sorption selectivity.

The final application we have chosen to highlight relates to the industrially important process of separating linear and monobranched hexane isomers from their dibranched counterparts to enhance the research octane number. Chung and co-workers conducted a screening of both MOFs, from the CoREMOF database,^[309] and zeolites, taken from the IZA,^[195] in

order to identify optimal adsorbent materials for this separation process.^[361] Unlike the previous HT separation studies, which focus on small gaseous molecules modeled as rigid bodies, hexane isomers require full flexibility for accurate property prediction. Therefore, the authors employed CB-GCMC (see Section 3.2.2) in order to produce distributions of adsorbate conformers and used the Widom particle insertion approach to calculate Henry constants. A pore size cutoff and selectivity-based ratios of Henry constants were used as an initial screen to reduce the sample size to 501 structures. Selectivities for an equimolar five-component mixture of hexane isomers was computed for the refined set of structures, which were then ranked by the structure’s affinity to adsorb linear and monobranched hexane isomers. The outcome of the screening procedure was the identification of 22 high performance candidates for this separation process. The authors went on to further assess three structures’ viability for this industrial process by conducting column breakthrough simulations. Moreover, by assessing the role of channel shape and conducting thermodynamic analysis, important molecular-level insights of this separation process were discovered.

The studies covered in this section give a broad overview of HTCS methodologies that can be implemented to study particular applications and seeks to highlight their limitations and rate determining steps. GCMC or MD simulations are not individually expensive in the case of adsorption and separation of small gaseous molecules, but the cumulative cost of running these calculations across the vast number of structures available can lead to intractable simulation time frames. In these cases, development of structural screens and semi-empirical methods to restrict the search space are key to the efficient identification and evaluation of promising candidates. This has been highlighted through the use and development of analytical equations to approximate methane^[347] and hydrogen^[308] storage in MOFs, the identification of key descriptors for materials, and GA-guided searches through combinatorial MOF space.^[349,350,358] In the case of flexible guest molecules, further problems arise as the computational cost increases dramatically when CB-GCMC is employed. These issues can be addressed by implementing initial filters that assess a framework’s validity in terms of both its structural characteristics and chemical composition, taking the form of pore size cutoffs and ratios of Henry constants.^[361] This allows further restriction of the search space to only the most promising structures, necessary for offsetting the cost of these simulations. Key information produced by these studies includes the synergistic combinations of structural and chemical descriptors that can be extricated from high performance candidates. However, these relationships are not always linear and machine learning algorithms may need to be deployed both to identify these dominant features and to identify feature combinations or frameworks not included in the original data set.

5.5. Applications of Machine Learning

The combinatorial nature of MOF structures lends itself to HT in silico investigation. However, depending on the time investment associated with assessing these structures’ properties,

simulation time frames can become unfeasible when deployed across vast databases of frameworks. ML proves to be an invaluable asset in this regard as employing these technologies can make virtual screening of colossal search spaces feasible. The data produced by HTCS can be used to train such algorithms, which can be deployed to: identify nonlinear structure–property relationships that were previously not discernible, further search MOF space to identify high performance candidates not present in the given data, and to guide the *in silico* synthesis of new frameworks for application. The field of HT-assisted machine learning is an ever-growing research area and has been recently reviewed;^[362] we give here a brief overview of the descriptors and methodologies that can be implemented for data analysis and acquisition, and materials discovery.

Fernandez et al. employed quantitative structure–property relationship (QSPR) models using multiple linear regression (MLR) analysis, decision tree regression (DT), and nonlinear support vector machines (SVM) to systematically correlate MOF structural features with their methane storage performance.^[363] Using six structural descriptors and methane storage data calculated for $\approx 130\,000$ hMOFs, the authors identified the framework void fraction and dominant pore diameter as the key features affecting methane storage. It was also found that SVMs outperformed MLR and DTs in the prediction of methane storage. Response surfaces of methane uptake produced by the SVM showed that the MOF database contained a limited distribution of void fraction and dominant pore diameter combinations and identified a maximum corresponding to combinations of these features not present in the original data set. The same group went on to do preliminary QSPR analysis on the same library of MOFs to examine their CO₂ capture properties and observed poor correlation between geometrical features and CO₂ capture at pressures relevant for gas separation applications. This led to the development of a novel atomic property weighted radial distribution function (AP-RDF) descriptor tailored for large-scale QSPR predictions of gas adsorption in MOFs.^[364] The group trained an SVM with AP-RDF scores to categorize MOFs as having high or low CO₂ capacity and, when applied to a test set of $\approx 290\,000$ structures, the QSPR classifier could recover 945 of the top 1000 MOFs whilst flagging only 10% of the repository for compute intensive screening.^[365] Using these machine learning classifiers to supplement high throughput workflows could lead to orders of magnitude reduction in the computational expense associated with HTCS.

Furthermore, Simon et al. implemented a hybrid ML/molecular simulation workflow to a database of over 600 000 experimentally realized and hypothetical porous materials for their Xe/Kr sorption selectivity.^[366] The authors trained a random forest of decision tree regressors using a training set of 15 000 structures, six structural descriptors, and the Voronoi energy (this novel descriptor takes into account both structural features and the energetics associated with guest–host interaction).^[367] The model was applied to their entire repository of porous networks and molecular simulation was used to assess only the most promising candidates. Of the $\approx 600\,000$, the random forest predicted 20 000 structures as promising candidates; these subsequently underwent further evaluation by GCMC simulation. This demonstrates further the use of this screening paradigm

in refining the necessary search space for high performance candidate identification.

Recent work conducted by Zhang et al. sought to develop a generative model for the *in silico* synthesis of high performance frameworks.^[368] By selecting ten different combinations of metal nodes and topologies, and deploying an algorithm utilizing Monte Carlo tree search combined with a recurrent neural network, the authors were able to search composition space and tailor novel MOFs to target applications. The algorithm's use was demonstrated by applying it to the case study of methane storage and carbon capture (estimating the reward function by GCMC simulation), where it successfully and efficiently designed high performance frameworks for these applications. Moreover, topological data analysis was employed to assess whether the set of novel MOFs generated were sufficiently diverse in their composition. As a similarity measure, organic linkers were represented using a topological fingerprint^[369] and pairwise comparisons of each framework was conducted. This demonstrated that the algorithm was able to generate a diverse set of high performing structures, as opposed to being constrained to derivatives of a particular composition. In a similar vein, Lee et al. sought to identify top MOF candidates for methane deliverable capacity;^[370] the authors developed an advanced MOF construction algorithm with the capability to generate ≈ 247 trillion structures by utilizing 1775 topologies and a large variety of structural building units. By combining an evolutionary algorithm with an artificial neural network, the authors were able to efficiently parse a search space of over 100 trillion structures; this led to the identification of 96 frameworks that exceed the current world record for methane deliverable capacity.

As highlighted in this section, ML can be a powerful tool for structure and property prediction; however, the Achilles heel of these algorithms is the data they are trained on. Whilst importance weighting can be used in order to produce algorithms that are unbiased by anomalous data, large quantities of erroneous data produced by HTCS procedures may skew ML predictions, hindering the rate of materials discovery.

5.6. Critical Remarks and Further Studies

HTCS, as seen in the studies covered in this section, has proven to be an invaluable tool for the fast identification of porous materials for particular applications, helping to expedite the time between hypothesis and discovery. However, non-uniformity of methodology for database construction and simulation parameters can lead to hazardous ramifications such as: highly skewed simulation statistics, unrealistic representation of guest–host interactions and computational sampling of MOF space that is not necessarily representative of the topological and structural diversity of real MOFs.

The methodological differences in constructing the CSD-derived databases present a challenge for HTCS, as different operating conditions require different treatment of solvent and automating this process can be onerous when deployed over large samples; bringing into question whether the simulated crystal structure is representative of the true nature of the framework. To address the latter, work has been conducted to

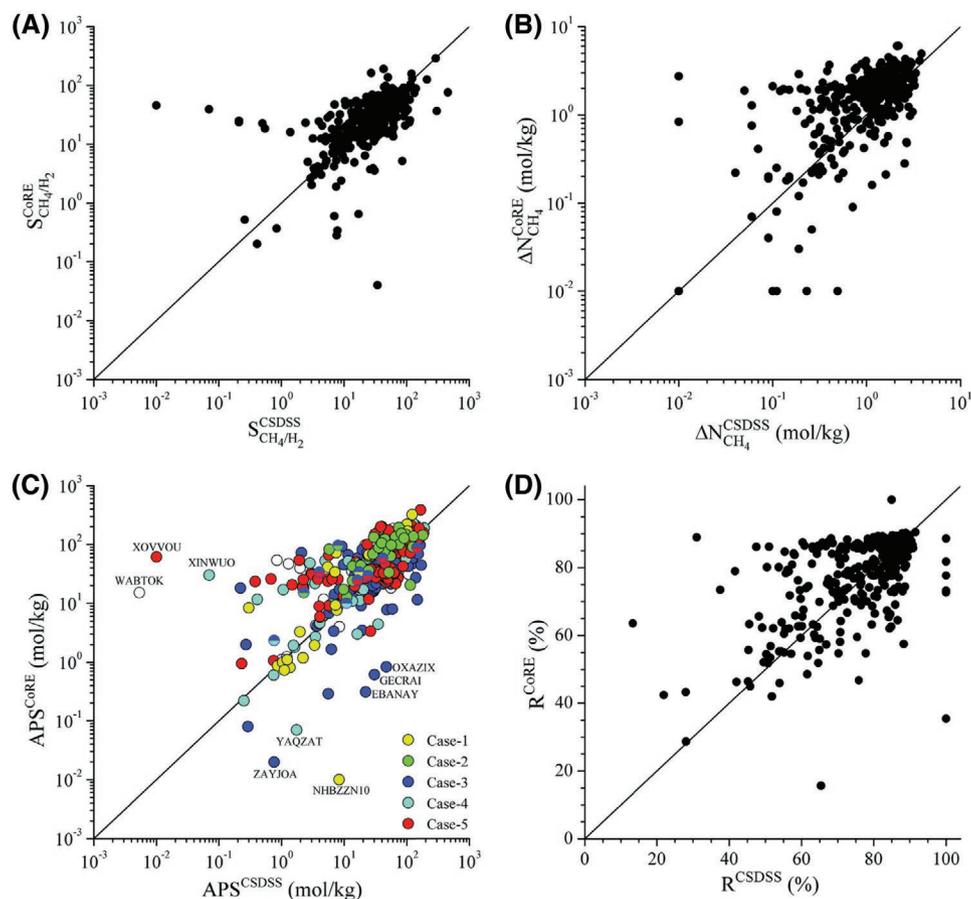


Figure 16. A–D) Comparison of CH_4/H_2 sorption selectivities (A), working capacities (B), adsorption performance score (C), and regenerability (D) of the 387 identified MOFs. The five cases are: 1) missing hydrogen atoms, 2) removal of unbound solvents, 3) removal of bound solvents, 4) retaining charge balancing ions and 5) missing parts in MOFs (ligand, metal, guest, etc.). MOFs in (C) are color-coded based on the cases they were categorized in, where uncolored data points represent the MOFs that were corrected by the authors during the study and double-colored data points represent the MOFs where two cases apply. A–D) Reproduced with permission.^[371] Copyright 2019, Royal Society of Chemistry.

identify erroneous structures labeled as identical between the CSD non-disordered MOF subset^[312] and the CoREMOF^[309] database.^[371] By computing pure component and binary mixture data for CH_4 and H_2 , and assessing four adsorption performance metrics for 3490 identically labeled structures, it was determined that 387 frameworks produced significantly different simulation statistics between the two databases. The cause for this was the difference in the structural information present for each framework in the repositories; the five cases outlining how the structures information differed between the CSD non-disordered MOF subset and the CoREMOF database are shown in **Figure 16**.

Further to this, different synthetic conditions coupled with the solvent removal techniques used in database construction and/or geometric relaxation of the structures can lead to duplicate frameworks being present within databases. Barthel et al.^[372] sought to systematically identify topological duplicates in a set of 502 DFT optimized structures from the CoREMOF database. By assessing the invariance of a set of descriptors such as the atom type, number, and those derived from the graph describing the frameworks bonding network, it was found that only 72.5% (364) of the 502 structures were reliable, with the other 27.5% having incorrect structural information or

being redundant duplicates. The findings of these two studies highlight the need for caution when conducting HTCS procedures, as if these erroneous structures are left untreated the statistics produced can be skewed and may even lead to the identification of candidate materials that could be fictitious.

The building block nature of MOFs has been heavily exploited by the computational community, indeed the number of hMOFs far surpasses the number of known experimentally reported frameworks. Many studies make use of the hMOF repositories to produce large quantities of structure–property information that constitute design criteria for new structures. However, the topological diversity in hypothetical framework databases can be limited and therefore structure–property information produced can be bias toward the dominant topology present in the data set; this limits the transferability of these structures to shape-selective processes like catalysis. However, work is being conducted to mitigate this hindrance, such as implementation of the RTA^[317] in the ToBaCCo code^[318] and further developments of this code in extending the number of nets to include non-edge-transitive topologies.^[373] One pressing question is the synthetic feasibility of the frameworks. Work by Anderson and Gómez-Gualdrón addressed this issue; they devised a computational approach to assess the

synthetic likelihood of computer-generated MOFs.^[374] Their study provides evidence that crystal free energies could be key to understanding the synthetic likelihood of hypothetical structures. Establishing *in silico* procedures for assessing synthetic viability, structural, and hydrothermal stability of MOFs will play a key role in restricting the ever-growing search space of these materials to only those that could be viable candidates for industrial processes.

In regard to the force fields used in HTCS, generic force fields available for framework atoms and those that are optimized to reproduce properties of the adsorbates differ by their treatment of the van der Waals and electrostatic components. Efforts have been made to systematically quantify the discrepancies of simulation data produced between different force fields. McDaniel et al. compared the validity of the EMP2 force field for CO₂^[375] and the TraPPE force field for methane^[376] with a benchmark *ab initio* derived force field.^[377] It was found that whilst gas uptake is relatively insensitive to force field choice at high pressures (assuming accurate adsorbate–adsorbate potential), MOF–guest interaction is of higher significance at low pressures and the accuracy of standard force fields is dependant on functional and topological features of the structures as well as the adsorbate type. In the case of CO₂, polarization and van der Waals interactions have distinctly different effects on the adsorption site distributions in MOFs; however, good correlation between the ranking of the force fields was found in general. For HTCS studies, where the goal is to identify the top percentage of MOFs for a given application, employing generic force fields is a reasonable approximation. However, caution is still warranted, especially where performance metrics that rely on the sampling of configurational space are used.

Despite these shortcomings HTCS, as seen in this section, is a powerful and invaluable tool for efficient property prediction. The data produced from these studies helps both computational and experimental scientists to refine the search space of whichever property they are looking to optimize. However, data is not always made publicly available and when it is, it may not be in a form that is easily intelligible or parsed. Notably, Moghadam and co-workers screened ≈2900 frameworks with high-quality *ab initio* charges, developed by Nazarian et al.,^[311] for their oxygen deliverable capacities.^[378] The arising data was then made available in an open-access interactive 5D visualization and data-mining tool, allowing 1000 unique structure–property relationships to be generated according to the user's interests. Open-access tools such as these are a desirable commodity that allow experimental or simulation effort to be focused on only the most promising candidates for given applications.

6. Covalent Organic Frameworks

We now very briefly highlight selected developments from the COFs to illustrate the use of techniques and approaches which have only partially been adopted or exploited in the zeolite and MOF community. The aim of this section is not to review these fields in toto but to pick out some novel studies that could inspire different approaches within the MOF and zeolite fields.

COFs are 2D or 3D porous crystalline polymeric materials composed of the light elements H, B, C, N, O, and S or P. Like

zeolites and MOFs, secondary building units are linked to form well-ordered, crystalline materials where the SBUs in this case are organic molecules. COFs are attractive gas storage materials owing to the strength of covalent bonds, and their intrinsic light weight frameworks due to their chemical composition. Despite the almost unlimited potential for forming frameworks, the number of experimentally reported COF structures is still relatively small, less than 500, and the number of computational studies in this area is also relatively small. For a comprehensive introduction and overview of COFs and their uses, the reader is referred to several excellent reviews on this topic.^[379–381] We note that the structure of COFs is usually less ambiguous than that of MOFs, for example; disorder is less common, and solvent is typically weakly coordinated to the framework rather than completing the coordination shell of an atom such as a metal. However, partial interpenetration of the networks is a feature of some COFs, leading to an occupancy of a framework of less than unity for the averaged crystal structure.

There are relatively few reports of experimental HTS of COFs but a notable early paper is that of Dogru et al.^[382] who synthesized the mesoporous material BTP-COF. The structure is notable for the large pore sizes of 4 nm and also for the synthetic approach. The synthetic route and parameters were reported to have been optimized through a robotized dosing system, although the number of permutations was not reported. A recent study by Wang et al.^[194] reported the use of divergent synthesis strategies to generate pools of modified reagents. These were then combined to generate eight hitherto unreported COFs using a multi-step synthesis route. Such a systematic approach would appear to be very amenable to robotic synthesis, potentially informed by successful and unsuccessful experiments in the manner reported by Moosavi et al.^[294] for MOFs.

An early study due to Bureekaew and Schmid^[383] used a small pool of reagents to generate hypothetical structures in a variety of topologies. The structures were fully optimized using the *ab initio* derived MOF-FF^[384] force field which yielded predictions of the lowest energy structure out of the possible topologies, although the overall feasibility of these frameworks was not reported. A notable feature of the work, apart from being one of the earliest studies of hypothetical COFs, was the use of a GA to identify the optimal orientation of ligands within a given topology.

Martin et al.^[385] reported an early attempt to shrink the synthesis space of COFs according to known synthetic routes and using commodity/commercially available reagents for the bridging linkers. Since COFs lack metal constituents, these structures lend themselves to energetic ranking by more sophisticated approaches than simply force fields. In this study, Martin et al. used the PM6-DH2 method within MOPAC12 (applying periodic boundary conditions) to optimize 620 non-interpenetrated frameworks and these structures were then assessed for their capability to interpenetrate using Zeo++.^[214] Combining all permutations of the 620 frameworks to form interpenetrated structures, a total of 4147 structures were found to be geometrically well matched. These 4147 structures were finally optimized using DFT to identify structures for potential methane storage applications. By explicitly considering the economy of synthesis, simulation starts to push the technology

readiness level up for the winning candidates because the raw cost of the reagents is, in effect, a screening parameter. While this approach could be more widely adopted in MOF screening, there is a small danger that disruptive materials are missed because the constituents are not currently easily available. Of course, new routes to synthesis and demand for reagents leading to an economy of scale, can lead to enormous savings in reagent cost. A similar approach, using 666 known reagents and established synthetic routes was reported by Mercado et al.^[386] who reported a database of 69480 COF structures^[387] that was used to discover a methane storage framework based on the tbd topology.^[388] The best material found had a deliverable methane capacity of 216 v(STP)v⁻¹,^[348] the highest reported figure at the time of publication.

A computation ready CoRe COF database was reported by Tong et al.^[14,389] which now contains 451 (experimentally realized) structures at the time of writing. However, there are now much larger databases available featuring structures not yet realized synthetically. A novel use of GA approaches to COFs was reported by Lam et al.,^[75] who developed a database of 471 990 COFs^[390] constructed through an elegant computational algorithm. SBUs are treated as genetic structural units and reacted together according to consideration of their functional groups in a manner reminiscent of Wilmer et al.,^[314] with the distinction that 2D materials were explicitly targeted in this work as well as 3D materials. 2D MOFs have become a recent area of great interest (see e.g., Cliffe et al.^[391] and Firth et al.^[392]) for their potential use in catalysis and/or as membranes for example, hence it would be interesting to see the targeting of 2D MOFs. Although the generated structures were optimized using the DREIDING force field and/or Forcite, the structures were not ranked according to thermodynamic stability, which would be a logical way to assess the synthetic viability, as was done in a landmark study by Lukose et al.^[393] in 2010 using tight binding DFT approaches implemented in the deMON2k package.^[394]

It is also becoming more widespread to see machine learning impacting the identification of useful COFs. Li et al.^[395] reported a recent study to identify COFs as potential heat pumps as is discussed in greater detail in Section 5.4.4. Descriptors were identified to predict their performance as heating or cooling agents. One potential extension to this work would be to use the descriptor approach to assess the virtual databases available, to examine whether far more efficient materials exist in the databases, which could provide an additional incentive to target synthetic effort.

7. Outlook

In this review, we have sought to highlight developments in the field of porous materials toward the realization of a wholly integrated workflow between experiment and computer simulation and analysis. While synthetic approaches are relatively mature and the new vanguard is automating synthesis, new approaches are being developed that promise to dramatically improve the efficacy of synthesis, such as the “chemputer”.^[396,397]

For simulation and analysis, there is greater scope for improving interoperability and veracity. Any computer model is necessarily a simplification of the experimental conditions

and so integration of experimental results provides invaluable data to assess the robustness of the current model and refine it. Physical experiments are the ultimate test of a model's predictive power; on one level, the experiment may reveal inadequacies in the parameterization of the model. These deficiencies may lead to an incorrect prediction of the thermodynamic ground state, but with adequate training, a useful model should be trustworthy and accurately predictive. However, a far more subtle and complex aspect of the physical experiment is that it probes kinetic aspects of the reaction. Typical syntheses of the materials discussed in this review require several hours and this timescale is not accessible to the models that are typically used for this type of prediction (DFT-based simulations, FF simulations, and DFT trained/machine learned models). In terms of improving models, the advent of machine learning approaches, taught from increasingly large and diverse databases of ever more reliable DFT data, should ensure that the models have the potential to become ever more robust. The inexorable growth of computing power means that an adequate training pool for geometries and their associated energies will be achieved more readily.

There are cross-cutting and interdisciplinary challenges to developing more automated materials discovery processes. It is noteworthy that often techniques are blooded in the metal oxide field before finding application in the porous materials area, so more closely monitoring that literature could help expedite new methods in the porous materials field. In addition, here, we list areas where there are opportunities to enhance the interoperability of techniques and to reduce the time overhead in some rate limiting steps within the virtuous circle of computation, experiment and analysis:

1. Reducing the initial cost of robots and platforms. More widespread adoption of automated experimental workflows, such as the “chemputer” approach,^[396] to design and print bespoke vessels for the delivery of molecules and solids, and robotic approaches^[398] will drive down the economic cost of the physical components. Beyond this initial outlay, there is the cost of developing the software to enable the interoperability of the components which may be both significant and highly specialized.
2. Harnessing of interdisciplinary skills. Programming of robots and writing codes capable of automatically interpreting the outcomes of experiments, capitalizing on ML approaches and acting upon them to design new experiments, additional characterization, or new virtual screening experiments are currently rare skills. Aside from stepping outside the traditional recruitment and collaboration field to tap into the knowledge of engineers, mathematicians, and computer scientists, there is an impetus to train up physical scientists to be conversant and comfortable with using these new integrated approaches in research. There is a growing need for interdisciplinary methods to be taught at the undergraduate level, in order to enhance their training and ensure that the start-of-the-art approaches, robotics, ML, etc., as described, are adopted more widely and become ingrained in the design of experiments.
3. Repurposing of synthetic and characterization methods from other classes of materials. For example, physical and chemical vapor deposition (CVD) are very well-established techniques

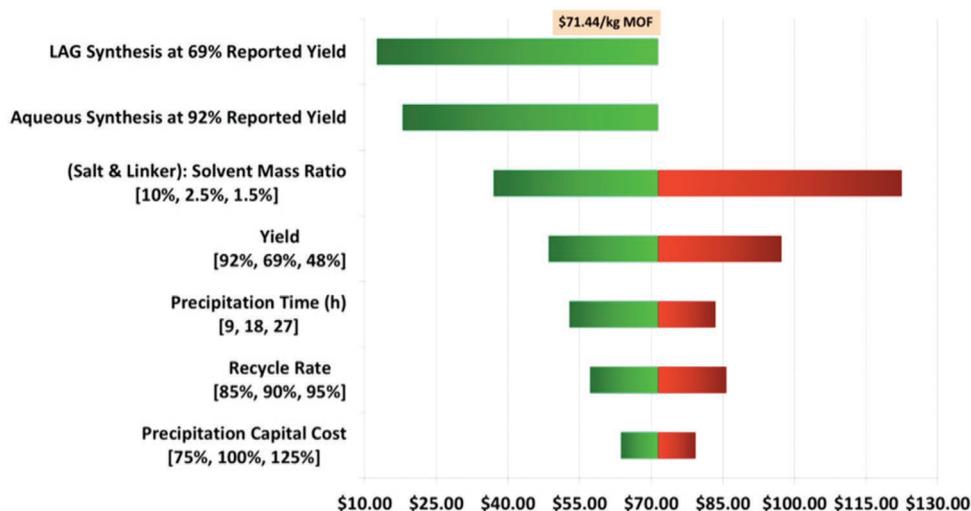


Figure 17. Sensitivity assessment of factors affecting the economic cost of synthesizing Mg₂(dobc) (MOF-74). LAG is liquid-assisted grinding. Switching from solvothermal to aqueous or LAG syntheses has strongest affect on reducing the total cost of synthesis, followed by mass ratio of salt and linker to solvent. Reproduced with permission.^[404] Copyright 2017, American Chemical Society.

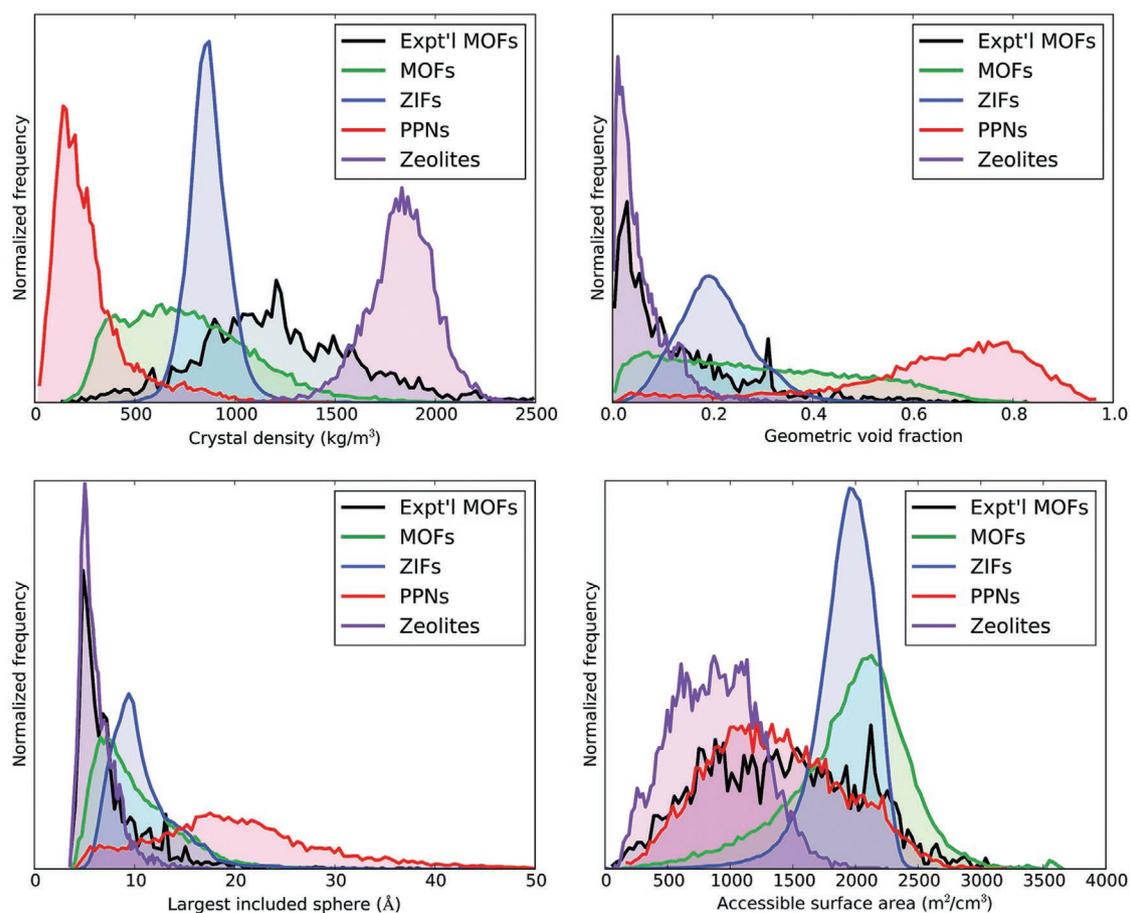


Figure 18. An example of the assessment of a greater diversity of class of nanoporous materials, enabled by consistent computational settings. In the search for outstanding materials, the facility to search across completely different chemical compositions and distributions of porosity could lead to important discoveries, especially when combined with ML approaches. The particular databases mined here actually show strong overlap of the properties, which could indicate that looking to other classes of materials (e.g., amorphous materials) could be fruitful for certain applications. Reproduced with permission.^[237] Copyright 2015, Royal Society of Chemistry.

that have been used in materials discovery for decades and have been successfully used for the high throughput phase screening of metal oxides.^[399,400] Stassen et al.^[401] reported how MOFs can be synthesized using CVD, offering new routes to controlling the surface chemistry of MOFs. However, taking inspiration from other fields, CVD offers an intuitive way to exploring, for example, multivariate MOFs on a single sheet, that is, to produce a printable phase diagram. Whilst it is now possible to perform many characterizations (e.g., XRD) in a high throughput fashion, some analysis is still incredibly time consuming and manual, for example, HRTEM.^[402] Nevertheless, disruptive techniques such as machine learning are beginning to reduce the time overhead to this most onerous of techniques.^[403] Similarly, strides in automating time consuming characterization methods such as electron diffraction have started to address these experimental bottlenecks.^[191]

- Incorporation of economic factors and considerations (including potentially upstream costs, sequestration, disposal, or processing of harmful by-products) in the initial synthetic and computational screening process. There are relatively few examples in the literature but the study by DeSantis et al. (see **Figure 17**)^[404] focused on identifying factors that could drive down MOF synthesis costs to <10 \$ kg⁻¹ provides an excellent example of the potential value of these studies. Similarly, assessing reagent costs as part of the screening process, as described in Martin et al.,^[385] is clearly a powerful approach toward identifying commercially viable materials and this could be employed more routinely in truly integrated workflows.
- Coordination between materials disciplines to improve the transferability and consistency in data. Harmonizing of electronic databases allows for scanning of multiple materials types, as depicted in **Figure 18**,^[237] to resolve the issue of inconsistent computational settings between databases. With a sufficiently powerful machine learning approach, structure/energy relationships from one database could be “corrected” to place them within the same potential energy landscape of another, obviating the need to re-evaluate all structures at the same level of theory.
- Improving the veracity of computer simulations. Density functional theory has advanced tremendously over the last decade or two in terms of accuracy and transferability. There are still problematic cases, such as highly correlated systems and excited states, but the advent of quantum Monte Carlo and potentially quantum computing approaches means the chemical accuracy and range of chemical states in databases will only increase. Ever-growing databases could be used to develop machine learnt potentials, in the spirit of Gaussian approximation potentials^[405] with the ability to describe reaction chemistry, like reaxFF.^[327] Another opportunity is to use ab initio or DFT data to further improve tight binding (such as GFN2-xTB),^[406,407] or semi-empirical DFT approaches (such as HSEsol-3c, PBEh-3c, and HF-3c^[408]) in order to reduce the mean error associated with these approaches.
- Machine learning to harvest and collate existing data while rejecting anomalies. It has been long proposed that machine learning methods could be used to extract data from disparate sources to produce a far more comprehensive virtual compendium or lab book. Indeed, Jensen et al.^[244]

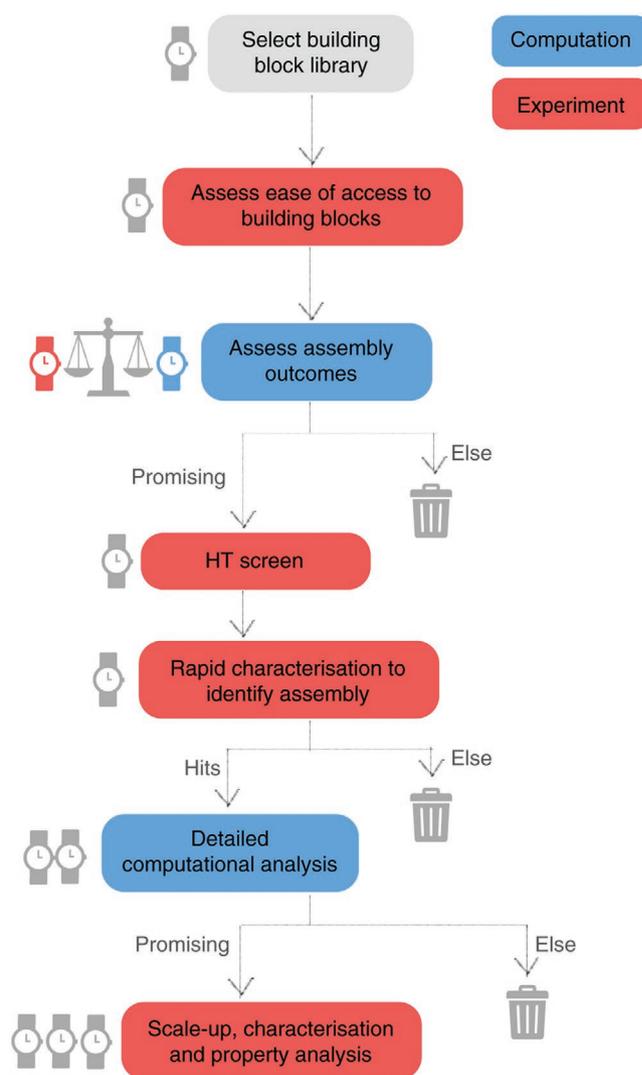


Figure 19. Proposed integrated workflow reproduced from Greenaway et al.^[398] The rate determining steps are easily identified by the wrist-watches in the figure. Reproduced under the terms of the CC-BY Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0>).^[398] Copyright 2018, The Authors, published by Springer Nature.

have recently applied this technology to germanosilicates and advocated that the approach could be used for other problems in zeolite synthesis and to gain greater insight into the factors affecting crystal habit. One potential hazard in this approach is the veracity of the data in the literature. Similar approaches have been used in the pharmaceutical industry where it is chastening to learn that “An analysis of past studies indicates that the cumulative (total) prevalence of irreproducible preclinical research exceeds 50%, resulting in \approx US\$28 000 000 000 (US\$28B)/year spent on preclinical research that is not reproducible—in the United States alone”.^[247] However, we can anticipate that machine learning models could be trained to detect suspect or identify substantially outlying data. The emergence of studies such as Moosavi et al.^[294] and Porwol et al.^[409] that generate valuable data from

failed as well as successful experiments will help to empower existing and future GA/ML-guided synthesis approaches.

It is also instructive to look to other fields for inspiration of how to improve HT approaches for zeolites, MOFs, and COFs. Robotic synthesis is playing an increasingly important role in performing HTS. For instance, Greenaway et al.^[398] used a combination of computational screening and robotic synthesis to expedite exploration of 78 precursor combinations resulting in the identification of 32 new porous organic cage molecules. **Figure 19** shows the entire workflow used in this study. Interestingly, consideration is given to the cost of reagents tensioned against the cost of the computation. In this particular study, the computational role varied depending on whether the organic building blocks were commercially available and inexpensive, or required substantial synthetic effort. For time-consuming and expensive synthesis to be justified, a complete assessment of the potential stability and properties of the cage materials was performed to establish the potential high value of the likely products of synthesis. In other cases where the building blocks were commodity chemicals, the computational screening was sidestepped. In future work it was noted that elimination of low value targets would be a desirable aspect of screening, presumably, when the compute cost of ranking and screening properties becomes even more tractable through advances in computer software, hardware, and machine learning approaches.^[410] Additionally, adding a feedback loop to learn from successful and unsuccessful experiments,^[294] as previously discussed in Section 4.3 could greatly accelerate the discovery process.

Notwithstanding the opportunities listed above, by harnessing the full gamut of techniques surveyed in this review and emerging technologies, it is hoped that fully integrated, self-learning, and self-guided identification of promising materials for real-world applications will become increasingly tractable, to avoid making poor pore choices.

Acknowledgements

B.S. would like to thank EPSRC for funding the programme grant EP/N004884/1 and a CASE studentship and a doctoral training centre studentship (with Johnson Matthey) which have helped with this review. B.S. would also like to thank the Royal Society for an Industry Fellowship IF160062 and a studentship INF\PHD\180018 which also contributed to this review.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

high throughput methods, high throughput synthesis, machine learning, metal–organic frameworks, porous materials, zeolites

Received: April 24, 2020

Revised: June 2, 2020

Published online:

- [1] W. F. Maier, K. Stöwe, S. Sieg, *Angew. Chem., Int. Ed.* **2007**, *46*, 6016.
- [2] K. H. Bleicher, H.-J. Böhm, K. Müller, A. I. Alanine, *Nat. Rev. Drug Discovery* **2003**, *2*, 369.
- [3] R. Bogue, *Ind. Robot* **2012**, *39*, 113.
- [4] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* **2001**, *46*, 3.
- [5] L. M. Daniels, S. Ling, S. N. Savvin, M. J. Pitcher, M. S. Dyer, J. B. Claridge, B. Slater, F. Corà, J. Alaria, M. J. Rosseinsky, *Energy Environ. Sci.* **2017**, *10*, 1917.
- [6] M. L. Green, C. L. Choi, J. R. Hatrick-Simpers, A. M. Joshi, I. Takeuchi, S. C. Barron, E. Campo, T. Chiang, S. Empedocles, J. M. Gregoire, A. G. Kusne, J. Martin, A. Mehta, K. Persson, Z. Trautt, J. Van Duren, A. Zakutayev, *Appl. Phys. Rev.* **2017**, *4*, 011105.
- [7] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Mater.* **2013**, *1*, 011002.
- [8] Website of NOMAD Laboratory, <https://nomad-coe.eu> (accessed: January 2020).
- [9] Webpage for BiGmax Max Planck Research Network on big-data-driven materials science, <https://www.bigmax.mpg.de> (accessed: January 2020).
- [10] Website for MARVEL National Centre of Competence in Research, <https://nccr-marvel.ch/> (accessed: January 2020).
- [11] Webpage for Nanoporous Materials Genome Center at University of Minnesota, <http://www1.chem.umn.edu/nmgc/> (accessed: January 2020).
- [12] P. G. Boyd, Y. Lee, B. Smit, *Nat. Rev. Mater.* **2017**, *2*, 17037.
- [13] G. Fraux, S. Chibani, F.-X. Coudert, *Philos. Trans. R. Soc., A* **2019**, *377*, 20180220.
- [14] M. Tong, Y. Lan, Q. Yang, C. Zhong, *Chem. Eng. Sci.* **2017**, *168*, 456.
- [15] J. D. Evans, B. Garai, H. Reinsch, W. Li, S. Dissegna, V. Bon, I. Senkowska, R. A. Fischer, S. Kaskel, C. Janiak, N. Stock, D. Volkmer, *Coord. Chem. Rev.* **2019**, *380*, 378.
- [16] W. F. Maier, *ACS Comb. Sci.* **2019**, *21*, 437.
- [17] P. A. Wright, *Microporous Framework Solids*, RSC Materials Monographs. The Royal Society of Chemistry, London, UK **2008**.
- [18] S. M. Auerbach, K. A. Carrado, P. K. Dutta, *Handbook of Zeolite Science and Technology*, CRC Press, Boca Raton, FL, USA **2003**.
- [19] N. J. Henson, A. K. Cheetham, J. D. Gale, *Chem. Mater.* **1994**, *6*, 1647.
- [20] P. M. Piccione, C. Laberty, S. Yang, M. A. Cambor, A. Navrotsky, M. E. Davis, *J. Phys. Chem. B* **2000**, *104*, 10001.
- [21] *Verified Syntheses of Zeolitic Materials*, 3rd ed. (Ed: S. Mintova), International Zeolite Association, **2016**.
- [22] R. Banerjee, A. Phan, B. Wang, C. Knobler, H. Furukawa, M. O'Keeffe, O. M. Yaghi, *Science* **2008**, *319*, 939.
- [23] D. E. Akporiaye, I. M. Dahl, A. Karlsson, R. Wendelbo, *Angew. Chem., Int. Ed.* **1998**, *37*, 609.
- [24] S. P. Fodor, J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu, D. Solas, *Science* **1991**, *251*, 767.
- [25] R. Wendelbo, A. Karlsson, D. E. Akporiaye, I. M. Dahl (Sinvent AS) *Norway* **304,355**, **1997**.
- [26] D. Akporiaye, I. M. Dahl, A. Karlsson, R. Wendelbo (Sinvent AS) *USA* **6,821,486**, **2004**.
- [27] D. W. Breck, *Zeolite Molecular Sieves: Structure, Chemistry, and Use*, John Wiley & Sons, New York **1974**.
- [28] CHA denotes the three letter structure code of the Chabazite type topology—see www.iza-structure.org for a more comprehensive introduction and a current list of experimentally realised zeolite topologies and Section 3.1 for further discussion.
- [29] J. Klein, C. W. Lehmann, H. W. Schmidt, W. F. Maier, *Angew. Chem., Int. Ed.* **1998**, *37*, 3369.
- [30] D. Akporiaye, I. Dahl, A. Karlsson, M. Plassen, R. Wendelbo, D. S. Bem, R. W. Broach, G. J. Lewis, M. Miller, J. Moscoso, *Microporous Mesoporous Mater.* **2001**, *48*, 367.

- [31] K. Choi, D. Gardner, N. Hilbrandt, T. Bein, *Angew. Chem., Int. Ed.* **1999**, *38*, 2891.
- [32] J. Holmgren, D. Bem, M. Bricker, R. Gillespie, G. Lewis, D. Akporiaye, I. Dahl, A. Karlsson, M. Plassen, R. Wendelbo, *Stud. Surf. Sci. Catal.* **2001**, *135*, 113.
- [33] L. Zhang, J. Yao, C. Zeng, N. Xu, *Chem. Commun.* **2003**, *17*, 2232.
- [34] P. G. Schultz, X. Xiang, I. Goldwasser (Symyx Technologies and the Regents of the University of California) *US 6,326,090*, **2001**.
- [35] W. F. Maier, J. Klein, C. Lehmann, H.-W. Schmidt (The Aktiengesellschaft, the high throughput experimentation company) *US 6,825,048*, **2004**.
- [36] P. G. Schultz, X. Xiang, I. Goldwasser (Lawrence Berkeley National Laboratory and Symyx Technologies, Inc.) *US 6,686,205*, **2004**.
- [37] J. M. Newsam, T. Bein, J. Klein, W. F. Maier, W. Stichert, *Microporous Mesoporous Mater.* **2001**, *48*, 355.
- [38] Y. Li, H. Cao, J. Yu, *ACS Nano* **2018**, *12*, 4096.
- [39] C.-R. Boruntea, G. Sastre, L. F. Lundegaard, A. Corma, P. N. R. Vennestrøm, *Chem. Mater.* **2019**, *31*, 9268.
- [40] P. P. Pescarmona, J. J. T. Rops, J. C. van der Waal, J. C. Jansen, T. Maschmeyer, *J. Mol. Catal. A: Chem.* **2002**, *182–183*, 319.
- [41] M. Tagliabue, L. C. Carluccio, D. Ghisletti, C. Perego, *Catal. Today* **2003**, *81*, 405.
- [42] M. Moliner, J. M. Serra, A. Corma, E. Argente, S. Valero, V. Botti, *Microporous Mesoporous Mater.* **2005**, *78*, 73.
- [43] T. J. Daou, J. Dhainaut, A. Chappaz, N. Bats, B. Harbuzaru, H. Chaumeil, A. Defoin, L. Rouleau, J. Patarin, *Oil Gas Sci. Technol.* **2015**, *70*, 447.
- [44] T. Willhammar, J. Su, Y. Yun, X. Zou, M. Afeworki, S. C. Weston, H. B. Vroman, W. W. Lonergan, K. G. Strohmaier, *Inorg. Chem.* **2017**, *56*, 8856.
- [45] L. M. Knight, G. J. Lewis, in *Recent Advances in the Science and Technology of Zeolites and Related Materials*, Vol. 154A (Eds: E. van Steen, I. M. Claeys, L. H. Callanan), Elsevier, Amsterdam, The Netherlands **2004**, pp. 171–179.
- [46] L. M. Knight, M. A. Miller, S. C. Koster, M. G. Gatter, A. I. Benin, R. R. Willis, G. J. Lewis, R. W. Broach, in *From Zeolites to Porous MOF Materials: The 40th Anniversary of International Zeolite Conference*, Vol. 170 (Eds: R. Xu, Z. Gao, J. Chen, W. Yan), Elsevier, Amsterdam, The Netherlands **2007**, pp. 338–346.
- [47] A. Cantín, A. Corma, M. J. Díaz-Cabanas, J. L. Jordá, M. Moliner, *J. Am. Chem. Soc.* **2006**, *128*, 4216.
- [48] A. Corma, M. J. Díaz-Cabanas, M. Moliner, C. Martínez, *J. Catal.* **2006**, *241*, 312.
- [49] A. Corma, M. J. Díaz-Cabanas, J. L. Jordá, C. Martínez, M. Moliner, *Nature* **2006**, *443*, 842.
- [50] J. Sun, C. Bonneau, Á. Cantín, A. Corma, M. J. Díaz-Cabanas, M. Moliner, D. Zhang, M. Li, X. Zou, *Nature* **2009**, *458*, 1154.
- [51] J. Jiang, J. L. Jorda, J. Yu, L. A. Baumes, E. Mugnaioli, M. J. Diaz-Cabanas, U. Kolb, A. Corma, *Science* **2011**, *333*, 1131.
- [52] J. Jiang, Y. Yun, X. Zou, J. L. Jorda, A. Corma, *Chem. Sci.* **2015**, *6*, 480.
- [53] N. Stock, *Microporous Mesoporous Mater.* **2010**, *129*, 287.
- [54] Y. Song, J. Li, J. Yu, K. Wang, R. Xu, *Top. Catal.* **2005**, *35*, 3.
- [55] J. Jiang, Y. Xu, P. Cheng, Q. Sun, J. Yu, A. Corma, R. Xu, *Chem. Mater.* **2011**, *23*, 4709.
- [56] S. C. Weston, B. K. Peterson, J. E. Gatt, W. W. Lonergan, H. B. Vroman, M. Afeworki, G. J. Kennedy, D. L. Dorset, M. D. Shannon, K. G. Strohmaier, *J. Am. Chem. Soc.* **2019**, *141*, 15910.
- [57] Zeolite beta is considered to be a partially disordered material due to the crystal consisting of 2 to 3 intergrown polymorphs.^[41] Therefore whilst the framework type BEA is assigned to the zeolite beta structure, this is with the caveat that zeolite beta itself has a disordered topology due to these intergrowths and thus is a polytype rather than a singular framework.
- [58] Á. Cantín, A. Corma, M. J. Díaz-Cabanas, J. L. Jordá, M. Moliner, F. Rey, *Angew. Chem., Int. Ed.* **2006**, *45*, 8013.
- [59] J. L. Guth, H. Kessler, P. Caullet, J. Hazm, A. Merrouche, J. Patarin, in *Proc. Ninth Int. Zeolite Conf.* (Eds: R. von Ballmoos, J. B. Higgins, M. M. J. Treacy), Butterworth-Heinemann, Oxford, UK **1993**, pp. 215–222.
- [60] T. P. Caremans, C. E. A. Kirschhock, P. Verlooy, J. S. Paul, P. A. Jacobs, J. A. Martens, *Microporous Mesoporous Mater.* **2006**, *90*, 62.
- [61] S. P. B. Kremer, C. E. A. Kirschhock, A. Aerts, C. A. Aerts, K. J. Houthoofd, P. J. Grobet, P. A. Jacobs, O. I. Lebedev, G. van Tendeloo, J. A. Martens, *Solid State Sci.* **2005**, *7*, 861.
- [62] C. T. Kresge, M. E. Leonowicz, W. J. Roth, J. C. Vartuli, J. S. Beck, *Nature* **1992**, *359*, 710.
- [63] J. S. Beck, J. C. Vartuli, W. J. Roth, M. E. Leonowicz, C. T. Kresge, K. D. Schmitt, C. T.-W. Chu, D. H. Olson, E. W. Sheppard, S. B. McCullen, J. B. Higgins, J. L. Schlenker, *J. Am. Chem. Soc.* **1992**, *114*, 10834.
- [64] C. T. Chu, C. T. Kresge, K. G. Simmons (Mobil Oil Corp.) *US 5,215,737*, **1993**.
- [65] V. Alfreðsson, M. W. Anderson, *Chem. Mater.* **1996**, *8*, 1141.
- [66] A. Carlsson, M. Kaneda, Y. Sakamoto, O. Terasaki, R. Ryoo, S. H. Joo, *J. Electron Microsc.* **1999**, *48*, 795.
- [67] F. Liebau, *Zeolites* **1983**, *3*, 191.
- [68] K. P. F. Janssen, J. S. Paul, B. F. Sels, P. A. Jacobs, *Appl. Surf. Sci.* **2007**, *254*, 699.
- [69] P. Serna, L. A. Baumes, M. Moliner, A. Corma, *J. Catal.* **2008**, *258*, 25.
- [70] A. Corma, M. Moliner, J. M. Serra, P. Serna, M. J. Díaz-Cabanas, L. A. Baumes, *Chem. Mater.* **2006**, *18*, 3287.
- [71] J. M. Serra, L. A. Baumes, M. Moliner, P. Serna, A. Corma, *Comb. Chem. High Throughput Screening* **2007**, *10*, 13.
- [72] L. A. Baumes, M. Moliner, A. Corma, *QSAR Comb. Sci.* **2007**, *26*, 255.
- [73] R. Lai, B. S. Kang, G. R. Gavalas, *Angew. Chem., Int. Ed.* **2001**, *40*, 408.
- [74] T. Vandermeersch, T. R. C. van Assche, J. F. M. Denayer, W. D. Malsche, *Microporous Mesoporous Mater.* **2016**, *226*, 133.
- [75] C.-H. Lam, W.-J. Hsu, H.-Y. Chi, Y.-H. Kang, J.-J. Chen, D.-Y. Kang, *Microporous Mesoporous Mater.* **2018**, *267*, 171.
- [76] J. C. Groen, J. A. Moulijn, J. Pérez-Ramírez, *Ind. Eng. Chem. Res.* **2007**, *46*, 4193.
- [77] K. P. F. Janssen, J. S. Paul, B. F. Sels, P. A. Jacobs, in *From Zeolites to Porous MOF Materials: The 40th Anniversary of International Zeolite Conference*, Vol. 170 (Eds: R. Xu, Z. Gao, J. Chen, W. Yan), Elsevier, Amsterdam, The Netherlands **2007**, pp. 1222–1227.
- [78] S. Abelló, S. Dhir, G. Colet, J. Pérez-Ramírez, *Appl. Catal., A* **2007**, *325*, 121.
- [79] P. Kern, M. Klimczak, T. Heinzlmann, M. Lucas, P. Claus, *Appl. Catal., B* **2010**, *95*, 48.
- [80] A. Nanoti, S. Dasgupta, V. Agnihotri, P. Gupta, A. N. Goswami, M. O. Garg, E. Tangstad, M. Stöcker, A. Karlsson, Ø. B. Vistad, *Microporous Mesoporous Mater.* **2011**, *146*, 158.
- [81] P. P. Pescarmona, B. C. Gagea, P. van der Aa, P. A. Jacobs, J. A. Martens, *Catal. Today* **2011**, *159*, 120.
- [82] A. Sundermann, O. Gerlach, *Catalysts* **2016**, *6*, 23.
- [83] Y. Khabzina, C. Laroche, J. Perez-Pellitero, D. Farrusseng, *Microporous Mesoporous Mater.* **2017**, *247*, 52.
- [84] L. A. Baumes, M. Moliner, A. Corma, *Chem. - Eur. J.* **2009**, *15*, 4258.
- [85] A. J. Koster, U. Ziese, A. J. Verkleij, A. H. Janssen, K. P. de Jong, *J. Phys. Chem. B* **2000**, *104*, 9368.
- [86] A. H. Janssen, A. J. Koster, K. P. de Jong, *J. Phys. Chem. B* **2002**, *106*, 11905.
- [87] U. Ziese, A. H. Janssen, J. L. Murk, W. J. C. Geerts, T. van der Krift, A. J. Verkleij, A. J. Koster, *J. Microsc.* **2002**, *205*, 187.
- [88] T. C. Miller, G. Mann, G. J. Havrilla, C. A. Wells, B. P. Warner, R. T. Baker, *J. Comb. Chem.* **2003**, *5*, 245.
- [89] E. M. Minogue, T. P. Taylor, A. K. Burrell, G. J. Havrilla, B. P. Warner, M. T. Janicke, *Chem. Commun.* **2005**, *33*, 4167.
- [90] A. Saito, H. C. Foley, *Microporous Mater.* **1995**, *3*, 531.

- [91] J. C. Groen, S. Abelló, L. A. Villaescusa, J. Pérez-Ramírez, *Microporous Mesoporous Mater.* **2008**, *114*, 93.
- [92] C. M. Snively, S. Katzenberger, G. Oskarsdottir, J. Lauterbach, *Opt. Lett.* **1999**, *24*, 1841.
- [93] C. M. Snively, G. Oskarsdottir, J. Lauterbach, *Catal. Today* **2001**, *67*, 357.
- [94] C. M. Snively, G. Oskarsdottir, J. Lauterbach, *Angew. Chem.* **2001**, *113*, 3117.
- [95] C. M. Snively, G. Oskarsdottir, J. Lauterbach, *Angew. Chem., Int. Ed.* **2001**, *40*, 3028.
- [96] P. Kubanek, O. Busch, S. Thomson, H. W. Schmidt, F. Schüth, *J. Comb. Chem.* **2004**, *6*, 420.
- [97] O. M. Busch, W. Brijioux, S. Thomson, F. Schüth, *J. Catal.* **2004**, *222*, 174.
- [98] P. Kubanek, H.-W. Schmidt, B. Spliethoff, F. Schüth, *Microporous Mesoporous Mater.* **2005**, *77*, 89.
- [99] H. Wang, Z. Liu, J. Shen, H. Liu, *Catal. Commun.* **2004**, *5*, 55.
- [100] H. Wang, Z. Liu, J. Shen, *J. Comb. Chem.* **2003**, *5*, 802.
- [101] C. Woltz, A. Jentys, J. A. Lercher, *J. Catal.* **2006**, *237*, 337.
- [102] D. A. Fischer, K. Efimenko, R. R. Bhat, S. Sambasivan, J. Genzer, *Macromol. Rapid Commun.* **2004**, *25*, 141.
- [103] P. Wollmann, M. Leistner, U. Stoeck, R. Grünker, K. Gedrich, N. Klein, O. Throl, W. Grählert, I. Senkovska, F. Dreisbach, S. Kaskel, *Chem. Commun.* **2011**, *47*, 5151.
- [104] P. Wollmann, M. Leistner, W. Grählert, O. Throl, F. Dreisbach, S. Kaskel, *Microporous Mesoporous Mater.* **2012**, *149*, 86.
- [105] A. Werner, M. Wöllner, P. Bludovsky, M. Leistner, C. Selzer, S. Kaskel, *Microporous Mesoporous Mater.* **2018**, *268*, 46.
- [106] T.-H. Bae, M. R. Hudson, J. A. Mason, W. L. Queen, J. J. Dutton, K. Sumida, K. J. Micklash, S. S. Kaye, C. M. Brown, J. R. Long, *Energy Environ. Sci.* **2013**, *6*, 128.
- [107] F. Fischer, E. Laevemann, *Meas. Sci. Technol.* **2015**, *26*, 065603.
- [108] F. Fischer, W. Lutz, J.-C. Buhl, E. Laevemann, *Microporous Mesoporous Mater.* **2018**, *262*, 258.
- [109] J. T. Mullhaupt, P. C. Stephenson (Praxair Technology, Inc.) *USA 5,554,208*, **1996**.
- [110] Z. Tang, L.-F. Li, H. Jiang (Bettinger Corp.) *US 9,126,830*, **2015**.
- [111] R. Thomas, J. A. Moulijn, V. H. J. de Beer, J. Medema, *J. Mol. Catal.* **1980**, *8*, 161.
- [112] J. G. Creer, P. Jackson, G. Pandey, G. G. Percival, D. Seddon, *Appl. Catal.* **1986**, *22*, 85.
- [113] S. Bessell, D. Seddon, *J. Catal.* **1987**, *105*, 270.
- [114] W. Huybrechts, J. Mijoin, P. A. Jacobs, J. A. Martens, *Appl. Catal., A* **2003**, *243*, 1.
- [115] J. F. M. Denayer, R. A. Ocakoglu, W. Huybrechts, B. Dejonckheere, P. Jacobs, S. Calero, R. Krishna, B. Smit, G. V. Baron, J. A. Martens, *J. Catal.* **2003**, *220*, 66.
- [116] B. Xu, F. Rotunno, S. Bordiga, R. Prins, J. A. van Bokhoven, *J. Catal.* **2006**, *241*, 66.
- [117] M. A. Miller, G. J. Lewis, J. G. Moscoco, S. Koster, F. Modica, M. G. Gatter, L. T. Nemeth, in *From Zeolites to Porous MOF Materials: The 40th Anniversary of International Zeolite Conference*, Vol. 170 (Eds: R. Xu, Z. Gao, J. Chen, W. Yan), Elsevier, Amsterdam, The Netherlands **2007**, pp. 487–492.
- [118] K. Hayasaka, D. Liang, W. Huybrechts, B. R. de Waele, K. J. Houthoofd, P. Eloy, E. M. Gaigneaux, G. van Tendeloo, J. W. Thybaut, G. B. Marin, J. F. M. Denayer, G. V. Baron, P. A. Jacobs, C. E. A. Kirschhock, J. A. Martens, *Chem. - Eur. J.* **2007**, *13*, 10070.
- [119] B. Xu, S. Bordiga, R. Prins, J. A. van Bokhoven, *Appl. Catal., A* **2007**, *333*, 245.
- [120] G. Morra, A. Desmartin-Chomel, C. Daniel, U. Ravon, D. Farrusseng, R. Cowan, M. Krusche, C. Mirodatos, *Chem. Eng. J.* **2008**, *138*, 379.
- [121] F. Ocampo, H. S. Yun, M. M. Pereira, J. P. Tessonnier, B. Louis, *Cryst. Growth Des.* **2009**, *9*, 3721.
- [122] N. Kasian, T. I. Koranyi, G. Vanbutsele, K. Houthoofd, J. A. Martens, C. E. A. Kirschhock, *Top. Catal.* **2010**, *53*, 1374.
- [123] N. Kasian, G. Vanbutsele, K. Houthoofd, T. I. Koranyi, J. A. Martens, C. E. A. Kirschhock, *Catal. Sci. Technol.* **2011**, *1*, 246.
- [124] F. Marques Mota, C. Bouchy, E. Guillon, A. Fécant, N. Bats, J. A. Martens, *J. Catal.* **2013**, *301*, 20.
- [125] J. E. Bedenbaugh, S. Kim, E. Sasmaz, J. Lauterbach, *ACS Comb. Sci.* **2013**, *15*, 491.
- [126] K. Arve, P. Mäki-Arvela, K. Eränen, M. Tiitta, T. Salmi, D. Y. Murzin, *Chem. Eng. J.* **2014**, *238*, 3.
- [127] A. D. F. Ferreira, A. J. Maia, B. Guatiguaba, M. H. Herbst, P. T. L. Rocha, M. M. Pereira, B. Louis, *Catal. Today* **2014**, *226*, 67.
- [128] J. Zečević, G. Vanbutsele, K. P. de Jong, J. A. Martens, *Nature* **2015**, *528*, 245.
- [129] M. M. Pereira, E. S. Gomes, A. V. Silva, A. B. Pinar, M.-G. Willinger, S. Shanmugam, C. Chizallet, G. Laugel, P. Losch, B. Louis, *Chem. Sci.* **2018**, *9*, 6532.
- [130] L. Geerts, R. K. Ramachandran, J. Dendooven, S. Radhakrishnan, J. W. Seo, C. Detavernier, J. Martens, S. P. Sree, *Catal. Sci. Technol.* **2020**, *10*, 1778.
- [131] A. Aerts, A. van Isacker, W. Huybrechts, S. P. B. Kremer, C. E. A. Kirschhock, F. Collignon, K. Houthoofd, J. F. M. Denayer, G. V. Baron, G. B. Marin, P. A. Jacobs, J. A. Martens, *Appl. Catal., A* **2004**, *257*, 7.
- [132] W. Huybrechts, J. W. Thybaut, B. R. de Waele, G. Vanbutsele, K. J. Houthoofd, F. Bertinchamps, J. F. M. Denayer, E. M. Gaigneaux, G. B. Marin, G. V. Baron, P. A. Jacobs, J. A. Martens, *J. Catal.* **2006**, *239*, 451.
- [133] L. Ha, J. Mao, J. Zhou, Z. C. Zhang, S. Zhang, *Appl. Catal., A* **2009**, *356*, 52.
- [134] N. Kasian, E. Verheyen, G. Vanbutsele, K. Houthoofd, T. I. Koranyi, J. A. Martens, C. E. A. Kirschhock, *Microporous Mesoporous Mater.* **2013**, *166*, 153.
- [135] P. Mertens, F. Verpoort, A.-N. Parvulescu, D. de Vos, *J. Catal.* **2006**, *243*, 7.
- [136] V. Subramanian, V. L. Zholobenko, K. Cheng, C. Lancelot, S. Heyte, J. Thuriot, S. Paul, V. V. Ordonsky, A. Y. Khodakov, *ChemCatChem* **2016**, *8*, 380.
- [137] T. N. Nguyen, T. T. P. Nhat, K. Takimoto, A. Thakur, S. Nishimura, J. Ohyama, I. Miyazato, L. Takahashi, J. Fujima, K. Takahashi, T. Taniike, *ACS Catal.* **2020**, *10*, 921.
- [138] J. L. Weber, N. A. Krans, J. P. Hofmann, E. J. M. Hensen, J. Zečević, P. E. de Jongh, K. P. de Jong, *Catal. Today* **2020**, *342*, 161.
- [139] M. J. M. Mies, E. V. Rebrov, C. J. B. U. Schiepers, M. H. J. M. de Croon, J. C. Schouten, *Chem. Eng. Sci.* **2007**, *62*, 5097.
- [140] C. Hoffmann, A. Wolf, F. Schüth, *Angew. Chem., Int. Ed.* **1999**, *38*, 2800.
- [141] U. Rodemerck, P. Ignaszewski, M. Lucas, P. Claus, *Chem. Eng. Technol.* **2000**, *23*, 413.
- [142] M. J. M. Mies, E. V. Rebrov, L. Deutz, C. R. Kleijn, M. H. J. M. de Croon, J. C. Schouten, *Ind. Eng. Chem. Res.* **2007**, *46*, 3922.
- [143] P. P. Pescarmona, J. van Noyen, P. A. Jacobs, *J. Catal.* **2007**, *251*, 307.
- [144] P. P. Pescarmona, P. A. Jacobs, *Catal. Today* **2008**, *137*, 52.
- [145] T. Duerinck, P. Leflaive, P. Martin, G. D. Pirngruber, J. F. M. Denayer, *Adsorption* **2011**, *17*, 347.
- [146] J. Pérez-Ramírez, F. Kapteijn, G. Mul, J. A. Moulijn, *Chem. Commun.* **2001**, *8*, 693.
- [147] J. Pérez-Ramírez, F. Kapteijn, G. Mul, J. A. Moulijn, *J. Catal.* **2002**, *208*, 211.
- [148] J. Pérez-Ramírez, F. Kapteijn, G. Mul, J. A. Moulijn, *Appl. Catal., B* **2002**, *35*, 227.
- [149] J. Pérez-Ramírez, F. Kapteijn, G. Mul, J. A. Moulijn, *Catal. Commun.* **2002**, *3*, 19.
- [150] R. J. Berger, J. Pérez-Ramírez, F. Kapteijn, J. A. Moulijn, *Appl. Catal., A* **2002**, *227*, 321.

- [151] J. Pérez-Ramírez, F. Kapteijn, A. Brückner, *J. Catal.* **2003**, *218*, 234.
- [152] I. Melián-Cabrera, F. Kapteijn, J. A. Moulijn, *Catal. Today* **2005**, *110*, 255.
- [153] P. J. Smeets, M. H. Groothaert, R. M. van Teeffelen, H. Leeman, E. J. M. Hensen, R. A. Schoonheydt, *J. Catal.* **2007**, *245*, 358.
- [154] P. J. Smeets, B. F. Sels, R. M. van Teeffelen, H. Leeman, E. J. M. Hensen, R. A. Schoonheydt, *J. Catal.* **2008**, *256*, 183.
- [155] S. van der Beken, E. Dejaegere, K. A. Tehrani, J. S. Paul, P. A. Jacobs, G. V. Baron, J. F. M. Denayer, *J. Catal.* **2005**, *235*, 128.
- [156] S. van der Beken, E. Dejaegere, H. Verelst, G. V. Baron, J. F. M. Denayer, *Int. J. Chem. React. Eng.* **2005**, *3*, <https://doi.org/10.2202/1542-6580.1271>.
- [157] K. Mantri, E. Dejaegere, G. V. Baron, J. F. M. Denayer, *Appl. Catal., A* **2007**, *318*, 95.
- [158] M. Salzinger, M. B. Fichtl, J. A. Lercher, *Appl. Catal., A* **2011**, *393*, 189.
- [159] P. P. Pescarmona, K. P. F. Janssen, C. Stroobants, B. Molle, J. S. Paul, P. A. Jacobs, B. F. Sels, *Top. Catal.* **2010**, *53*, 77.
- [160] P. P. Pescarmona, K. P. F. Janssen, C. Delaet, C. Stroobants, K. Houthoofd, A. Philippaerts, C. de Jonghe, J. S. Paul, P. A. Jacobs, B. F. Sels, *Green Chem.* **2010**, *12*, 1083.
- [161] L. Singoredjo, M. Slagt, J. van Wees, F. Kapteijn, J. A. Moulijn, *Catal. Today* **1990**, *7*, 157.
- [162] J. Pérez-Ramírez, R. J. Berger, G. Mul, F. Kapteijn, J. A. Moulijn, *Catal. Today* **2000**, *60*, 93.
- [163] H. Wang, Z. Liu, J. Shen, H. Liu, J. Zhang, *Catal. Commun.* **2005**, *6*, 343.
- [164] C. Mukarakate, X. Zhang, A. R. Stanton, D. J. Robichaud, P. N. Ciesielski, K. Malhotra, B. S. Donohoe, E. Gjersing, R. J. Evans, D. S. Heroux, R. Richards, K. Iisa, M. R. Nimlos, *Green Chem.* **2014**, *16*, 1444.
- [165] C. Mukarakate, M. J. Watson, J. ten Dam, X. Baucherel, S. Budhi, M. M. Yung, H. Ben, K. Iisa, R. M. Baldwin, M. R. Nimlos, *Green Chem.* **2014**, *16*, 4891.
- [166] J. A. Mason, T. M. McDonald, T.-H. Bae, J. E. Bachman, K. Sumida, J. J. Dutton, S. S. Kaye, J. R. Long, *J. Am. Chem. Soc.* **2015**, *137*, 4787.
- [167] N. Casacuberta, M. Castrillejo, A.-M. Wefing, S. Bollhalder, L. Wacker, *Radiocarbon* **2020**, *62*, 13.
- [168] H.-A. Synal, M. Stocker, M. Suter, *Nucl. Instrum. Methods Phys. Res., Sect. B* **2007**, *259*, 7.
- [169] S. Ozturk, S. Senkan, *Appl. Catal., B* **2002**, *38*, 243.
- [170] S. Senkan, K. Krantz, S. Ozturk, V. Zengin, I. Onal, *Angew. Chem., Int. Ed.* **1999**, *38*, 2794.
- [171] M. Iwamoto, H. Furukawa, Y. Mine, F. Uemura, S. Mikuriya, S. Kagawa, *J. Chem. Soc., Chem. Commun.* **1986**, *16*, 1272.
- [172] M. Iwamoto, H. Yahiro, Y. Mine, S. Kagawa, *Chem. Lett.* **1989**, *18*, 213.
- [173] J. M. Serra, A. Corma, D. Farrusseng, L. Baumes, C. Mirodatos, C. Flego, C. Perego, *Catal. Today* **2003**, *81*, 425.
- [174] J. M. Serra, E. Guillon, A. Corma, *J. Catal.* **2004**, *227*, 459.
- [175] J. M. Serra, E. Guillon, A. Corma, *J. Catal.* **2005**, *232*, 342.
- [176] K. S. Oh, Y. K. Park, S. I. Woo, *Rev. Sci. Instrum.* **2005**, *76*, 062219.
- [177] J. T. Gleaves, J. R. Ebner, T. C. Kuechler, *Catal. Rev.: Sci. Eng.* **1988**, *30*, 49.
- [178] A. C. van Veen, D. Farrusseng, M. Rebeilleau, T. Decamp, A. Holzwarth, Y. Schuurman, C. Mirodatos, *J. Catal.* **2003**, *216*, 135.
- [179] H. Su, E. S. Yeung, *J. Am. Chem. Soc.* **2000**, *122*, 7422.
- [180] H. Su, E. S. Yeung, *Appl. Spectrosc.* **2002**, *56*, 1044.
- [181] P. Atienzar, A. Corma, H. García, J. M. Serra, *Chem. - Eur. J.* **2004**, *10*, 6043.
- [182] K. Gao, L. Yuan, L. Wang, *J. Comb. Chem.* **2006**, *8*, 247.
- [183] P. Horcajada, T. Chalati, C. Serre, B. Gillet, C. Sebrie, T. Baati, J. F. Eubank, D. Heurtaux, P. Clayette, C. Kreuz, J.-S. Chang, Y. K. Hwang, V. Marsaud, P.-N. Bories, L. Cynober, S. Gil, G. Férey, P. Couvreur, R. Gref, *Nat. Mater.* **2010**, *9*, 172.
- [184] R. J. Hendershot, P. T. Fanson, C. M. Snively, J. A. Lauterbach, *Angew. Chem., Int. Ed.* **2003**, *42*, 1152.
- [185] R. J. Hendershot, S. S. Lasko, M.-F. Fellmann, G. Oskarsdottir, W. N. Delgass, C. M. Snively, J. Lauterbach, *Appl. Catal., A* **2003**, *254*, 107.
- [186] D. W. Fickel, E. D'Addio, J. A. Lauterbach, R. F. Lobo, *Appl. Catal., B* **2011**, *102*, 441.
- [187] M. Lucas, P. Claus, *Appl. Catal., A* **2003**, *254*, 35.
- [188] S. Cypes, A. Hagemeyer, Z. Hogan, A. Lesik, G. Streukens, A. F. Volpe Jr., W. H. Weinberg, K. Yaccato, *Comb. Chem. High Throughput Screening* **2007**, *10*, 25.
- [189] J. Loskyll, K. Stoewe, W. F. Maier, *ACS Comb. Sci.* **2012**, *14*, 295.
- [190] B. Pinho, K. Minsariya, A. Yen, N. Joshi, R. L. Hartman, *Energy Fuels* **2017**, *31*, 11640.
- [191] M. O. Cichocka, J. Ångström, B. Wang, X. Zou, S. Smeets, *J. Appl. Crystallogr.* **2018**, *51*, 1652.
- [192] W. Chen, P. Vashistha, A. Yen, N. Joshi, Y. Kapoor, R. L. Hartman, *Energy Fuels* **2018**, *32*, 12205.
- [193] P. M. Dietrich, S. Bahr, T. Yamamoto, M. Meyer, A. Thissen, *J. Electron Spectrosc. Relat. Phenom.* **2019**, *231*, 118.
- [194] L.-K. Wang, J.-J. Zhou, Y.-B. Lan, S.-Y. Ding, W. Yu, W. Wang, *Angew. Chem., Int. Ed.* **2019**, *58*, 9443.
- [195] C. Baerlocher, L. McCusker, Database of zeolite structures, <http://www.iza-structure.org/databases/> (accessed: January 2020).
- [196] D. J. Earl, M. W. Deem, *Ind. Eng. Chem. Res.* **2006**, *45*, 5449.
- [197] R. Pophale, P. A. Cheeseman, M. W. Deem, *Phys. Chem. Chem. Phys.* **2011**, *13*, 12407.
- [198] M. M. J. Treacy, I. Rivin, E. Balkovsky, K. H. Randall, M. D. Foster, *Microporous Mesoporous Mater.* **2004**, *74*, 121.
- [199] M. J. Sanders, M. Leslie, C. R. A. Catlow, *J. Chem. Soc., Chem. Commun.* **1984**, *19*, 1271.
- [200] B. W. H. Van Beest, G. J. Kramer, R. A. Van Santen, *Phys. Rev. Lett.* **1990**, *64*, 1955.
- [201] M. Falcioni, M. W. Deem, *J. Chem. Phys.* **1999**, *110*, 1754.
- [202] M. W. Deem, J. M. Newsam, *Nature* **1989**, *342*, 260.
- [203] M. W. Deem, J. M. Newsam, *J. Am. Chem. Soc.* **1992**, *114*, 7189.
- [204] Crystallography Open database, <http://www.crystallography.net/pcod/>, <http://www.crystallography.net/cod/> (accessed: January 2020).
- [205] M. D. Foster, M. M. J. Treacy, Atlas of prospective zeolite structures, <http://www.hypotheticalzeolites.net/NEWDATABASE/index.html> (accessed: January 2020).
- [206] Y. Li, X. Li, J. Liu, F. Duan, J. Yu, *Nat. Commun.* **2015**, *6*, 8328.
- [207] L. Li, B. Slater, Y. Yan, C. Wang, Y. Li, J. Yu, *J. Phys. Chem. Lett.* **2019**, *10*, 1411.
- [208] K. Muraoka, W. Chaikittisilp, T. Okubo, *J. Am. Chem. Soc.* **2016**, *138*, 6184.
- [209] W. Loewenstein, *Am. Mineral.* **1954**, *39*, 92.
- [210] E. Dempsey, G. H. Kuehl, D. H. Olson, *J. Phys. Chem.* **1969**, *73*, 387.
- [211] R. E. Fletcher, S. Ling, B. Slater, *Chem. Sci.* **2017**, *8*, 7483.
- [212] C. J. Heard, L. Grajciar, P. Nachtigall, *Chem. Sci.* **2019**, *10*, 5705.
- [213] J. D. Evans, R. E. Fletcher, D. Hewitt, I. G. Clayson, B. Slater, unpublished.
- [214] T. F. Willems, C. H. Rycroft, M. Kazi, J. C. Meza, M. Haranczyk, *Microporous Mesoporous Mater.* **2012**, *149*, 134.
- [215] L. Sarkisov, A. Harrison, *Mol. Simul.* **2011**, *37*, 1248.
- [216] J. Hoshen, R. Kopelman, *Phys. Rev. B* **1976**, *14*, 3438.
- [217] Y. Lee, S. D. Barthel, P. Dlotko, S. M. Moosavi, K. Hess, B. Smit, *Nat. Commun.* **2017**, *8*, 15396.
- [218] R. L. Martin, B. Smit, M. Haranczyk, *J. Chem. Inf. Model.* **2012**, *52*, 308.
- [219] An important factor when using classical mechanics based methods is the choice of force field, which is a set of analytic equations describing the interactions between atoms. The form of the FF can vary, with non-bonded terms commonly described by Lennard-Jones and Coulomb interactions and harmonic or

anharmonic terms describing bonding interactions. Parameters for these equations also vary between force fields, and can be derived from empirical and/or ab initio data.^[199,231]

- [220] J. I. Siepmann, D. Frenkel, *Mol. Phys.* **1992**, *75*, 59.
- [221] P. Bai, M. Y. Jeon, L. Ren, C. Knight, M. W. Deem, M. Tsapatsis, J. I. Siepmann, *Nat. Commun.* **2015**, *6*, 5912.
- [222] J. Kim, M. Abouelnasr, L. C. Lin, B. Smit, *J. Am. Chem. Soc.* **2013**, *135*, 7545.
- [223] E. L. First, C. E. Gounaris, C. A. Floudas, *Langmuir* **2013**, *29*, 5599.
- [224] A. García-Sánchez, D. Dubbeldam, S. Calero, *J. Phys. Chem. C* **2010**, *114*, 15068.
- [225] R. Krishna, J. M. Van Baten, *J. Phys. Chem. C* **2010**, *114*, 18017.
- [226] P. J. Bereciartua, Á. Cantín, A. Corma, J. L. Jordá, M. Palomino, F. Rey, S. Valencia, E. W. Corcoran Jr., P. Kortunov, P. I. Ravikovitch, A. Burton, C. Yoon, Y. Wang, C. Paur, J. Guzman, A. R. Bishop, G. L. Casty, *Science* **2017**, *358*, 1068.
- [227] M. Witman, S. Ling, S. Jawahery, P. G. Boyd, M. Haranczyk, B. Slater, B. Smit, *J. Am. Chem. Soc.* **2017**, *139*, 5547.
- [228] The Intergovernmental Panel on Climate Change (IPCC) 6th assessment report, <https://www.ipcc.ch/assessment-report/ar6/> (accessed: January 2020).
- [229] J. Kim, L. C. Lin, J. A. Swisher, M. Haranczyk, B. Smit, *J. Am. Chem. Soc.* **2012**, *134*, 18940.
- [230] A. García-Sánchez, C. O. Ania, J. B. Parra, D. Dubbeldam, T. J. H. Vlugt, R. Krishna, S. Calero, *J. Phys. Chem. C* **2009**, *113*, 8814.
- [231] H. Fang, P. Kamakoti, J. Zang, S. Cundy, C. Paur, P. I. Ravikovitch, D. S. Sholl, *J. Phys. Chem. C* **2012**, *116*, 10692.
- [232] H. Fang, P. Kamakoti, P. I. Ravikovitch, M. Aronson, C. Paur, D. S. Sholl, *Phys. Chem. Chem. Phys.* **2013**, *15*, 12882.
- [233] H. Fang, A. Kulkarni, P. Kamakoti, R. Awati, P. I. Ravikovitch, D. S. Sholl, *Chem. Mater.* **2016**, *28*, 3887.
- [234] C. M. Simon, J. Kim, L. C. Lin, R. L. Martin, M. Haranczyk, B. Smit, *Phys. Chem. Chem. Phys.* **2014**, *16*, 5499.
- [235] P. Gómez-Álvarez, A. R. Ruiz-Salvador, S. Hamad, S. Calero, *J. Phys. Chem. C* **2017**, *121*, 4462.
- [236] A. S. Krishnapriyan, M. Haranczyk, D. Morozov, *J. Phys. Chem. C* **2020**, *124*, 9360.
- [237] C. M. Simon, J. Kim, D. A. Gomez-Gualdrón, J. S. Camp, Y. G. Chung, R. L. Martin, R. Mercado, M. W. Deem, D. Gunter, M. Haranczyk, D. S. Sholl, R. Q. Snurr, B. Smit, *Energy Environ. Sci.* **2015**, *8*, 1190.
- [238] J. Kim, L. C. Lin, R. L. Martin, J. A. Swisher, M. Haranczyk, B. Smit, *Langmuir* **2012**, *28*, 11914.
- [239] L. C. Lin, A. H. Berger, R. L. Martin, J. Kim, J. A. Swisher, K. Jariwala, C. H. Rycroft, A. S. Bhowan, M. W. Deem, M. Haranczyk, B. Smit, *Nat. Mater.* **2012**, *11*, 633.
- [240] J. Kim, A. Maiti, L. C. Lin, J. K. Stolaroff, B. Smit, R. D. Aines, *Nat. Commun.* **2013**, *4*, 1694.
- [241] E. García-Pérez, J. B. Parra, C. O. Ania, A. García-Sánchez, J. M. van Baten, R. Krishna, D. Dubbeldam, S. Calero, *Adsorption* **2007**, *13*, 469.
- [242] F. Daeyaert, F. Ye, M. W. Deem, *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 3413.
- [243] R. Gaillac, S. Chibani, F.-X. Coudert, *Chem. Mater.* **2020**, *32*, 2653.
- [244] Z. Jensen, E. Kim, S. Kwon, T. Z. H. Gani, Y. Román-Leshkov, M. Moliner, A. Corma, E. Olivetti, *ACS Cent. Sci.* **2019**, *5*, 892.
- [245] M. Moliner, Y. Román-Leshkov, A. Corma, *Acc. Chem. Res.* **2019**, *52*, 2971.
- [246] K. Muraoka, Y. Sada, D. Miyazaki, W. Chaikittisilp, T. Okubo, *Nat. Commun.* **2019**, *10*, 4459.
- [247] L. P. Freedman, I. M. Cockburn, T. S. Simcoe, *PLoS Biol.* **2015**, *13*, e1002165.
- [248] H. Furukawa, K. Cordova, M. O’Keeffe, O. Yaghi, *Science* **2013**, *341*, 974.
- [249] S. Yuan, L. Feng, K. Wang, J. Pang, M. Bosch, C. Lollar, Y. Sun, J. Qin, X. Yang, P. Zhang, Q. Wang, L. Zou, Y. Zhang, L. Zhang, Y. Fang, J. Li, H. C. Zhou, *Adv. Mater.* **2018**, *30*, 1704303.
- [250] K. Sumida, D. L. Rogow, J. a. Mason, T. M. McDonald, E. D. Bloch, Z. R. Herm, T.-H. Bae, J. R. Long, *Chem. Rev.* **2012**, *112*, 724.
- [251] J. Lee, O. K. Farha, J. Roberts, K. A. Scheidt, S. T. Nguyen, J. T. Hupp, *Chem. Soc. Rev.* **2009**, *38*, 1450.
- [252] S. Ma, H. C. Zhou, *Chem. Commun.* **2010**, *46*, 44.
- [253] P. M. Forster, N. Stock, A. K. Cheetham, *Angew. Chem., Int. Ed.* **2005**, *44*, 7608.
- [254] A. Sonnauer, F. Hoffmann, M. Fröba, L. Kienle, V. Duppel, M. Thommes, C. Serre, G. Férey, N. Stock, *Angew. Chem., Int. Ed.* **2009**, *48*, 3791.
- [255] M. L. Kelty, W. Morris, A. T. Gallagher, J. S. Anderson, K. A. Brown, C. A. Mirkin, T. D. Harris, *Chem. Commun.* **2016**, *52*, 7854.
- [256] N. Stock, T. Bein, *Angew. Chem., Int. Ed.* **2004**, *43*, 749.
- [257] S. Bauer, T. Bein, N. Stock, *Inorg. Chem.* **2005**, *44*, 5882.
- [258] S. Bauer, N. Stock, *J. Solid State Chem.* **2007**, *180*, 3111.
- [259] K. Sumida, S. Horike, S. S. Kaye, Z. R. Herm, W. L. Queen, C. M. Brown, F. Grandjean, G. J. Long, A. Dailly, J. R. Long, *Chem. Sci.* **2010**, *1*, 184.
- [260] N. M. Padiál, J. Castells-Gil, N. Almora-Barrios, M. Romero-Angel, I. Da Silva, M. Barawi, A. García-Sánchez, V. A. De La Peña O’Shea, C. Martí-Gastaldo, *J. Am. Chem. Soc.* **2019**, *141*, 13124.
- [261] H. L. Nguyen, *New J. Chem.* **2017**, *41*, 14030.
- [262] J. Castells-Gil, N. M. Padiál, N. Almora-Barrios, I. Da Silva, D. Mateo, J. Albero, H. García, C. Martí-Gastaldo, *Chem. Sci.* **2019**, *10*, 4313.
- [263] N. Stock, T. Bein, *J. Mater. Chem.* **2005**, *15*, 1384.
- [264] A. Sonnauer, C. Näther, H. A. Höpfe, J. Senker, N. Stock, *Inorg. Chem.* **2007**, *46*, 9968.
- [265] S. Bauer, N. Stock, *Angew. Chem., Int. Ed.* **2007**, *46*, 6857.
- [266] P. Maniam, N. Stock, *Inorg. Chem.* **2011**, *50*, 5085.
- [267] M. Feyand, C. F. Seidler, C. Deiter, A. Rothkirch, A. Lieb, M. Wark, N. Stock, *Dalton Trans.* **2013**, *42*, 8761.
- [268] H. Reinsch, N. Stock, *Microporous Mesoporous Mater.* **2013**, *171*, 156.
- [269] L. H. Schilling, N. Stock, *Dalton Trans.* **2014**, *43*, 414.
- [270] S. Bauer, T. Bein, N. Stock, *Solid State Sci.* **2008**, *10*, 837.
- [271] M. Plabst, R. Köhn, T. Bein, *CrystEngComm* **2010**, *12*, 1920.
- [272] F. Gándara, T. D. Bennett, *IUCrJ* **2014**, *1*, 563.
- [273] A. J. Howarth, A. W. Peters, N. A. Vermeulen, T. C. Wang, J. T. Hupp, O. K. Farha, *Chem. Mater.* **2017**, *29*, 26.
- [274] T. M. McDonald, E. D. Bloch, J. R. Long, *Chem. Commun.* **2015**, *51*, 4985.
- [275] J. E. Mondloch, O. Karagiari, O. K. Farha, J. T. Hupp, *CrystEngComm* **2013**, *15*, 9258.
- [276] J. J. Chen, X. Kong, K. Sumida, M. A. Manumpil, J. R. Long, J. A. Reimer, *Angew. Chem., Int. Ed.* **2013**, *52*, 12043.
- [277] S. Han, Y. Huang, T. Watanabe, Y. Dai, K. S. Walton, S. Nair, D. S. Sholl, J. C. Meredith, *ACS Comb. Sci.* **2012**, *14*, 263.
- [278] S. Han, Y. Huang, T. Watanabe, S. Nair, K. S. Walton, D. S. Sholl, J. C. Meredith, *Microporous Mesoporous Mater.* **2013**, *173*, 86.
- [279] A. D. Wiersum, C. Giovannangeli, D. Vincent, E. Bloch, H. Reinsch, N. Stock, J. S. Lee, J. S. Chang, P. L. Llewellyn, *ACS Comb. Sci.* **2013**, *15*, 111.
- [280] N. C. Burtch, H. Jasuja, K. S. Walton, *Chem. Rev.* **2014**, *114*, 10575.
- [281] J. J. Low, A. I. Benin, P. Jakubczak, J. F. Abrahamian, S. A. Faheem, R. R. Willis, *J. Am. Chem. Soc.* **2009**, *131*, 15834.
- [282] J. Della Rocca, D. Liu, W. Lin, *Acc. Chem. Res.* **2011**, *44*, 957.
- [283] P. Horcajada, R. Gref, T. Baati, P. K. Allan, G. Maurin, P. Couvreur, G. Férey, R. E. Morris, C. Serre, *Chem. Rev.* **2012**, *112*, 1232.
- [284] Y. Bai, Y. Dou, L. H. Xie, W. Rutledge, J. R. Li, H. C. Zhou, *Chem. Soc. Rev.* **2016**, *45*, 2327.
- [285] T. Mosmann, *J. Immunol. Methods* **1983**, *65*, 55.

- [286] K. D. Paull, R. H. Shoemaker, M. R. Boyd, J. L. Parsons, P. A. Risbood, W. A. Barbera, M. N. Sharma, D. C. Baker, E. Hand, D. A. Scudiero, A. Monks, M. C. Alley, M. Grote, *J. Heterocycl. Chem.* **1988**, 25, 911.
- [287] A. Ruyra, A. Yazdi, J. Espín, A. Carné-Sánchez, N. Roher, J. Lorenzo, I. Imaz, D. MasPOCH, *Chem. - Eur. J.* **2015**, 21, 2508.
- [288] J. M. Palomba, C. V. Credille, M. Kalaj, J. B. Decoste, G. W. Peterson, T. M. Tovar, S. M. Cohen, *Chem. Commun.* **2018**, 54, 5768.
- [289] M. Kalaj, J. M. Palomba, K. C. Bentz, S. M. Cohen, *Chem. Commun.* **2019**, 55, 5367.
- [290] J. M. Palomba, S. P. Harvey, M. Kalaj, B. R. Pimentel, J. B. Decoste, G. W. Peterson, S. M. Cohen, *ACS Appl. Mater. Interfaces* **2020**, 12, 14672.
- [291] D. Yang, B. C. Gates, *ACS Catal.* **2019**, 9, 1779.
- [292] M. Plabst, T. Bein, *Inorg. Chem.* **2009**, 48, 4331.
- [293] P. Maniam, C. Näther, N. Stock, *Eur. J. Inorg. Chem.* **2010**, 3866.
- [294] S. M. Moosavi, A. Chidambaram, L. Talirz, M. Haranczyk, K. C. Stylianou, B. Smit, *Nat. Commun.* **2019**, 10, 539.
- [295] A. Polinsky, R. Feinstein, S. Shi, A. Kuki, in *Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery* (Eds: I. M. Chaiken, K. D. Janda), American Chemical Society, Washington, DC, USA **1996**, pp. 219–232.
- [296] J. M. Granda, L. Donina, V. Dragone, D. L. Long, L. Cronin, *Nature* **2018**, 559, 377.
- [297] P. Raccuglia, K. C. Elbert, P. D. F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, A. J. Norquist, *Nature* **2016**, 533, 73.
- [298] P. Chen, Z. Tang, Z. Zeng, D. Zhou, C. Wang, W. Lin, P. Chen, Z. Tang, Z. Zeng, X. Hu, L. Xiao, Y. Liu, X. Qian, *Matter* **2020**, 2, 1651.
- [299] F. Kong, L. Yuan, Y. F. Zheng, W. Chen, *J. Lab. Autom.* **2012**, 17, 169.
- [300] E. Biemmi, S. Christian, N. Stock, T. Bein, *Microporous Mesoporous Mater.* **2009**, 117, 111.
- [301] M. Lammert, S. Bernt, F. Vermoortele, D. E. De Vos, N. Stock, *Inorg. Chem.* **2013**, 52, 8521.
- [302] J. Park, J. D. Howe, D. S. Sholl, *Chem. Mater.* **2017**, 29, 10487.
- [303] H. G. T. Nguyen, L. Espinal, R. D. van Zee, M. Thommes, B. Toman, M. S. L. Hudson, E. Mangano, S. Brandani, D. P. Broom, M. J. Benham, K. Cychoz, P. Bertier, F. Yang, B. M. Krooss, R. L. Siegelman, M. Hakuman, K. Nakai, A. D. Ebner, L. Erden, J. A. Ritter, A. Moran, O. Talu, Y. Huang, K. S. Walton, P. Billefont, G. De Weireld, *Adsorption* **2018**, 24, 531.
- [304] J.-H. Zhang, T. D. Y. Chung, K. R. Oldenburg, *J. Biomol. Screening* **1999**, 4, 67.
- [305] M. N. Bahr, D. B. Damon, S. D. Yates, A. S. Chin, J. D. Christopher, S. Cromer, N. Perrotto, J. Quiroz, V. Rosso, *Org. Process Res. Dev.* **2018**, 22, 1500.
- [306] M. Rubio-Martinez, C. Avci-Camur, A. W. Thornton, I. Imaz, D. MasPOCH, M. R. Hill, *Chem. Soc. Rev.* **2017**, 46, 3453.
- [307] T. Watanabe, D. S. Sholl, *Langmuir* **2012**, 28, 14114.
- [308] J. Goldsmith, A. G. Wong-Foy, M. J. Cafarella, D. J. Siegel, *Chem. Mater.* **2013**, 25, 3373.
- [309] Y. G. Chung, C. Jeffrey, M. Haranczyk, B. J. Sikora, W. Bury, V. Krungleviciute, T. Yildirim, O. K. Farha, D. S. Sholl, R. Q. Snurr, *Chem. Mater.* **2014**, 26, 6185.
- [310] 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, R. Q. Snurr, *J. Chem. Eng. Data* **2019**, 64, 5985.
- [311] D. Nazarian, J. S. Camp, D. S. Sholl, Y. G. Chung, S. Q. Snurr, *Chem. Mater.* **2016**, 28, 785.
- [312] P. Z. Moghadam, A. Li, S. B. Wiggin, A. Tao, A. G. Maloney, P. A. Wood, S. C. Ward, D. Fairen-Jimenez, *Chem. Mater.* **2017**, 29, 2618.
- [313] I. J. Bruno, J. C. Cole, P. R. Edgington, C. F. Macrae, J. Pearson, R. Taylor, *Acta Crystallogr., Sect. B* **2002**, B58, 389.
- [314] C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp, R. Q. Snurr, *Nat. Chem.* **2012**, 4, 83.
- [315] M. Z. Aghaji, M. Fernandez, P. G. Boyd, T. D. Daff, T. K. Woo, *Eur. J. Inorg. Chem.* **2016**, 2016, 4505.
- [316] A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. G. Goddard III, W. M. Skif, *J. Am. Chem. Soc.* **1992**, 114, 10024.
- [317] S. Bureekaew, V. Balwani, S. Amirjalayer, R. Schmid, *CrystEngComm* **2015**, 17, 344.
- [318] Y. J. Colon, D. A. Gomez-Gualdrón, R. Q. Snurr, *Cryst. Growth Des.* **2017**, 17, 5801.
- [319] R. Q. Snurr, Y. J. Colón, *Chem. Soc. Rev.* **2014**, 43, 5735.
- [320] S. L. Mayo, B. D. Olafson, W. A. Goddard III, *J. Phys. Chem.* **1990**, 94, 8897.
- [321] B. L. Eggimann, A. J. Sunnarborg, H. D. Stern, A. P. Bliss, J. I. Siepmann, *Mol. Simul.* **2014**, 40, 101.
- [322] J. Fu, H. Sun, *J. Phys. Chem. C* **2009**, 113, 21815.
- [323] E. Haldoupis, J. Borycz, H. Shi, K. D. Vogiatzis, P. Bai, W. L. Queen, L. Gagliardi, J. I. Siepmann, *J. Phys. Chem. C* **2015**, 119, 16058.
- [324] G. Garberoglio, A. I. Skoulidas, J. K. Johnson, *J. Phys. Chem. B* **2005**, 109, 13094.
- [325] S. E. Wenzel, M. Fischer, F. Hoffmann, M. Fröba, *Inorg. Chem.* **2009**, 48, 6559.
- [326] R. J. Verploegh, A. Kulkarni, S. E. Boulfelfel, J. C. Haydak, D. Tang, D. S. Sholl, *J. Phys. Chem. C* **2019**, 123, 9153.
- [327] T. P. Senthle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, A. C. T. van Duin, *NPJ Comput. Mater.* **2016**, 2, 15011.
- [328] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, 11, 361.
- [329] C. Campana, B. Mussard, T. K. Woo, *J. Chem. Theory Comput.* **2009**, 5, 2866.
- [330] T. A. Manz, D. S. Sholl, *J. Chem. Theory Comput.* **2010**, 6, 2455.
- [331] C. Zheng, C. Zhong, *J. Phys. Chem. C* **2010**, 114, 9945.
- [332] Q. Xu, C. Zhong, *J. Phys. Chem. C* **2010**, 114, 5035.
- [333] A. K. Rappe, W. A. Goddard III, *J. Phys. Chem.* **1991**, 95, 3358.
- [334] C. E. Wilmer, K. C. Kim, R. Q. Snurr, *J. Phys. Chem. Lett.* **2012**, 3, 2506.
- [335] E. S. Kadantsev, P. G. Boyd, T. D. Daff, T. K. Woo, *J. Phys. Chem. Lett.* **2013**, 4, 3056.
- [336] P. Canepa, C. A. Arter, E. M. Conwill, D. H. Johnson, B. A. Shoemaker, K. Z. Soliman, T. Thonhauser, *J. Mater. Chem. A* **2013**, 1, 13597.
- [337] S. R. Caskey, A. G. Wong-Foy, A. J. Matzger, *J. Am. Chem. Soc.* **2008**, 130, 10870.
- [338] Y. Liu, H. Kabbour, C. M. Brown, D. A. Neumann, C. C. Ahn, *Langmuir* **2008**, 24, 4772.
- [339] A. S. Rosen, J. M. Notestein, R. Q. Snurr, *J. Comput. Chem.* **2019**, 40, 1305.
- [340] Y. Basdogan, S. Keskin, *CrystEngComm* **2015**, 17, 261.
- [341] Y. J. Colón, D. Fairen-Jimenez, C. E. Wilmer, R. Q. Snurr, *J. Phys. Chem. C* **2014**, 118, 5383.
- [342] B. J. Sikora, R. Winnegar, D. M. Proserpio, R. Q. Snurr, *Microporous Mesoporous Mater.* **2014**, 186, 207.
- [343] 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, R. Q. Snurr, *Energy Environ. Sci.* **2016**, 9, 3279.
- [344] Volumetric surface area (SA) relates to the ratio of the SA of a material and its volume and gravimetric SA is the ratio of the materials SA and its mass.
- [345] A. Ahmed, A. J. Matzger, S. Seth, J. Purewal, A. G. Wong-Foy, M. Veenstra, D. J. Siegel, *Nat. Commun.* **2019**, 10, 1568.
- [346] H. Yabing, W. Zhou, Q. Guodong, *C. Banglin Chem. Soc. Rev.* **2014**, 43, 5657.
- [347] D. A. Gómez-Gualdrón, C. E. Wilmer, O. K. Farha, J. T. Hupp, R. Q. Snurr, *J. Phys. Chem. C* **2014**, 118, 6941.
- [348] STP refers to standard temperature and pressure.

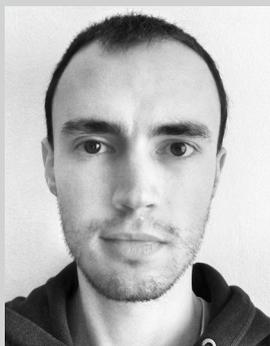
- [349] Y. Bao, R. L. Martin, C. M. Simon, M. Haranczyk, B. Smit, M. W. Deem, *J. Phys. Chem. C* **2015**, *119*, 186.
- [350] S. P. Collins, T. D. Daff, S. S. Piotrowski, T. K. Woo, *Sci. Adv.* **2016**, *2*, e1600954.
- [351] W. Li, X. Xia, M. Cao, S. Li, *J. Mater. Chem. A* **2019**, *7*, 7470.
- [352] Y.-S. Bae, R. Q. Snurr, *Angew. Chem., Int. Ed.* **2011**, *50*, 11586.
- [353] C. E. Wilmer, O. K. Farha, Y.-S. Bae, J. T. Hupp, R. Q. Snurr, *Energy Environ. Sci.* **2012**, *5*, 9849.
- [354] S. Li, Y. G. Chung, R. Q. Snurr, *Langmuir* **2016**, *32*, 10368.
- [355] H. Deng, C. J. Doonan, H. Furukawa, R. B. Ferreira, J. Towne, C. B. Knobler, B. Wang, O. M. Yaghi, *Science* **2010**, *327*, 846.
- [356] S. Li, Y. G. Chung, C. M. Simon, R. Q. Snurr, *J. Phys. Chem. Lett.* **2017**, *8*, 6135.
- [357] S. Sircar, T. C. Golden, *Sep. Sci. Technol.* **2000**, *35*, 667.
- [358] Y. G. Chung, D. A. Gómez-Gualdrón, P. Li, K. T. Leperi, P. Deria, H. Zhang, N. A. Vermeulen, J. F. Stoddart, F. You, J. T. Hupp, O. K. Farha, R. Q. Snurr, *Sci. Adv.* **2016**, *2*, 1600909.
- [359] J. A. Steckel, A. Olukayode, J. A. Steckel, C. E. Wilmer, *Energy Environ. Sci.* **2019**, *12*, 1255.
- [360] J. C. Maxwell, *Treatise on Electricity and Magnetism*, Oxford University Press, London, UK **1873**.
- [361] Y. G. Chung, P. Bai, M. Haranczyk, K. T. Leperi, P. Li, H. Zhang, T. C. Wang, T. Duerinck, F. You, J. T. Hupp, O. K. Farha, J. I. Siepmann, R. Q. Snurr, *Chem. Mater.* **2017**, *29*, 6315.
- [362] Z. Shi, W. Yang, X. Deng, C. Cai, Y. Yan, H. Liang, Z. Liu, Z. Qiao, *Mol. Syst. Des. Eng.* **2020**, *5*, 725.
- [363] M. Fernandez, T. K. Woo, C. E. Wilmer, R. Q. Snurr, *J. Phys. Chem. C* **2013**, *117*, 7681.
- [364] M. Fernandez, N. R. Tre, T. K. Woo, *J. Phys. Chem. C* **2013**, *117*, 14095.
- [365] M. Fernandez, P. G. Boyd, T. D. Daff, M. Z. Aghaji, T. K. Woo, *J. Phys. Chem. Lett.* **2014**, *5*, 3056.
- [366] C. M. Simon, R. Mercado, S. K. Schnell, B. Smit, M. Haranczyk, *Chem. Mater.* **2015**, *27*, 4459.
- [367] The Voronoi energy is the average energy of a xenon atom at the accessible Voronoi nodes that represent the pore topology.
- [368] X. Zhang, K. Zhang, Y. Lee, *ACS Appl. Mater. Interfaces* **2020**, *12*, 734.
- [369] RDKit, Open-source cheminformatics, <http://rdkit.org/> (accessed: January 2020).
- [370] S. Lee, B. Kim, J. Kim, *Discovery of Record-Breaking Metal-Organic Frameworks for Methane Storage using Evolutionary Algorithm and Machine Learning*, <https://doi.org/10.26434/chemrxiv.12333077.v1>.
- [371] C. Altintas, G. Avci, H. Daglar, A. Nemati Vesali Azar, I. Erucar, S. Velioglu, S. Keskin, *J. Mater. Chem. A* **2019**, *7*, 9593.
- [372] S. Barthel, E. V. Alexandrov, D. M. Proserpio, B. Smit, *Cryst. Growth Des.* **2018**, *18*, 1738.
- [373] R. Anderson, D. A. Gómez-Gualdrón, *CrystEngComm* **2019**, *21*, 1653.
- [374] R. Anderson, D. A. Gómez-Gualdrón, *Large-Scale Free Energy Calculations on a Computational MOF Database: Toward Synthetic Likelihood Predictions*, <https://doi.org/10.26434/chemrxiv.11874816.v1>.
- [375] J. G. Harris, K. H. Yung, *J. Phys. Chem.* **1995**, *99*, 12021.
- [376] M. G. Martin, J. I. Siepmann, *J. Phys. Chem. B* **1998**, *102*, 2569.
- [377] J. G. Mcdaniel, S. Li, E. Tylianakis, R. Q. Snurr, J. R. Schmidt, *J. Phys. Chem. C* **2015**, *119*, 3143.
- [378] P. Z. Moghadam, T. Islamoglu, S. Goswami, J. Exley, M. Fantham, C. F. Kaminski, R. Q. Snurr, O. K. Farha, D. Fairen-Jimenez, *Nat. Commun.* **2018**, *9*, 1378.
- [379] M. S. Lohse, T. Bein, *Adv. Funct. Mater.* **2018**, *28*, 1705553.
- [380] K. Geng, T. He, R. Liu, S. Dalapati, K. T. Tan, Z. Li, S. Tao, Y. Gong, Q. Jiang, D. Jiang, *Chem. Rev.* **2020**, *120*, 8814.
- [381] X. Guan, F. Chen, Q. Fang, S. Qiu, *Chem. Soc. Rev.* **2020**, *49*, 1357.
- [382] M. Dogru, A. Sonnauer, A. Gavryushin, P. Knochel, T. Bein, *Chem. Commun.* **2011**, *47*, 1707.
- [383] S. Bureekaew, R. Schmid, *CrystEngComm* **2013**, *15*, 1551.
- [384] S. Bureekaew, S. Amirjalayer, M. Tafipolsky, C. Spickermann, T. K. Roy, R. Schmid, *Phys. Status Solidi B* **2013**, *250*, 1128.
- [385] R. L. Martin, C. M. Simon, B. Medasani, D. K. Britt, B. Smit, M. Haranczyk, *J. Phys. Chem. C* **2014**, *118*, 23790.
- [386] R. Mercado, R.-S. Fu, A. V. Yakutovich, L. Talirz, M. Haranczyk, B. Smit, *Chem. Mater.* **2018**, *30*, 5069.
- [387] R. Mercado, R.-S. Fu, A. V. Yakutovich, L. Talirz, M. Haranczyk, B. Smit, In *Silico Design of 2D and 3D Covalent Organic Frameworks for Methane Storage Applications*, *Materials Cloud Archive* **2018.0003/v2**, **2018**, <https://doi.org/10.24435/materialscloud:2018.0003/v2>.
- [388] For more information about the tbd topology, the reader is directed to the reticular chemistry structure resource at <http://rcsr.anu.edu.au> (accessed: January 2020).
- [389] The CoRE-COF-Database, <https://core-cof.github.io/CoRE-COF-Database/> (accessed: February 2020).
- [390] Database of genetic structural units, <https://figshare.com/s/f8377d33bb9e9a73935a> (accessed: February 2020).
- [391] M. J. Cliffe, E. Castillo-Martinez, Y. Wu, J. Lee, A. C. Forse, F. C. N. Firth, P. Z. Moghadam, D. Fairen-Jimenez, M. W. Gaultois, J. A. Hill, O. V. Magdysyuk, B. Slater, A. L. Goodwin, C. P. Grey, *J. Am. Chem. Soc.* **2017**, *139*, 5397.
- [392] F. C. N. Firth, M. J. Cliffe, D. Vulpe, M. Aragones-Anglada, P. Z. Moghadam, D. Fairen-Jimenez, B. Slater, C. P. Grey, *J. Mater. Chem. A* **2019**, *7*, 7459.
- [393] B. Lukose, A. Kuc, J. Frenzel, T. Heine, *Beilstein J. Nanotechnol.* **2010**, *1*, 60.
- [394] The deMon2k Home Page, http://www.demon-software.com/public_html/index.htm (accessed: January 2020).
- [395] W. Li, X. Xia, S. Li, *ACS Appl. Mater. Interfaces* **2020**, *12*, 3265.
- [396] P. S. Gromski, J. M. Granda, L. Cronin, *Trends Chem.* **2020**, *2*, 4.
- [397] S. S. Zaleskiy, P. J. Kitson, P. Frei, A. Bubliskas, L. Cronin, *Nat. Commun.* **2019**, *10*, 5496.
- [398] R. L. Greenaway, V. Santolini, M. J. Bennison, B. M. Alston, C. J. Pugh, M. A. Little, M. Miklitz, E. G. Eden-Rump, R. Clowes, A. Shakil, H. J. Cuthbertson, H. Armstrong, M. E. Briggs, K. E. Jelfs, A. I. Cooper, *Nat. Commun.* **2018**, *9*, 2849.
- [399] X. D. Xiang, X. Sun, G. Briceno, Y. Lou, K.-A. Wang, H. Chang, W. G. Wallace-Freedman, S.-W. Chen, P. G. Schultz, *Science* **1995**, *268*, 1738.
- [400] S. Guerin, B. E. Hayden, *J. Comb. Chem.* **2006**, *8*, 66.
- [401] I. Stassen, M. Styles, G. Greci, H. V. Gorp, W. Vanderlinden, S. D. Feyter, P. Falcaro, D. D. Vos, P. Vereecken, R. Ameloot, *Nat. Mater.* **2016**, *15*, 304.
- [402] L. Liu, Z. Chen, J. Wang, D. Zhang, Y. Zhu, S. Ling, K.-W. Huang, Y. Belmabkhout, K. Adil, Y. Zhang, B. Slater, M. Eddaoudi, Y. Han, *Nat. Chem.* **2019**, *11*, 622.
- [403] J. Madsen, P. Liu, J. Kling, J. B. Wagner, T. W. Hansen, O. Winther, J. Schiøtz, *Adv. Theory Simul.* **2018**, *1*, 1800037.
- [404] D. DeSantis, J. A. Mason, B. D. James, C. Houchins, J. R. Long, M. Veenstra, *Energy Fuels* **2017**, *31*, 2024.
- [405] A. P. Bartók, G. Csányi, *Int. J. Quantum Chem.* **2015**, *115*, 1051.
- [406] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652.
- [407] Semiempirical Extended Tight-Binding Program Package, <https://github.com/grimme-lab/xtb> (accessed: January 2020).
- [408] L. Doná, J. G. Brandenburg, B. Civalieri, *J. Chem. Phys.* **2019**, *151*, 144103.
- [409] L. Porwol, D. J. Kowalski, A. Henson, D.-L. Long, N. Bell, L. Cronin, *Angew. Chem. Int. Ed.* **2020**, *59*, 11256.
- [410] L. Turcani, R. L. Greenaway, K. E. Jelfs, *Chem. Mater.* **2019**, *31*, 714.
- [411] T. Lu, W. Yan, R. Xu, *Inorg. Chem. Front.* **2019**, *6*, 1938.



Ivan G. Clayson is an Eng.D. student who has worked with the Slater group since 2017 at University College London. His work focuses on the transport of molecules under solvation in aluminosilicate zeolites through a combination of both first-principle simulation and experimental work. His research interests span both zeolite atomic structure and catalytic ability as well as software development and the application of quantum computing to chemical systems.



Daniel Hewitt is a Ph.D. student in the Slater group at the University College London Chemistry Department. His work focuses on the high throughput modeling of aluminosilicate zeolites for catalytic applications, and the identification of descriptors pertinent to their structure–property relationships. His research interests also include the structural elucidation of aluminosilicate zeolites through a combination of quantum mechanical and machine learning approaches; this research is facilitated by his development of custom software in Python.



Martin Hutereau is in the midst of a Ph.D. with the Slater group. His project revolves around the defect behavior of UiO-66 and other zirconium metal–organic frameworks, which he studies using quantum mechanical methods. His interests also extend to coding, and he has written a suite of Python utilities to aid his research.



Tom Pope is currently undertaking a Ph.D. in the Slater group at the University College London Chemistry Department. His work focuses on the high throughput modeling of metal–organic frameworks, for applications in fine chemical capture and transport. His research interests also include software design to aid in high throughput calculations, and development of molecular-level descriptions of organic molecules through a combination of force field and quantum mechanical approaches.



Ben Slater is a professor of computational and materials chemistry at University College London. He is currently a Royal Society industry fellow in conjunction with Johnson Matthey. He has a broad interest in simulation and materials chemistry and had published extensively on the topic of water ice and a variety of different porous materials including zeolites and MOFs.