High-throughput screening of metal–organic frameworks for macroscale heteroepitaxial alignment

Andrew Tarzia,† Masahide Takahashi,‡ Paolo Falcaro,¶ Aaron W. Thornton,§ Christian J. Doonan,*,† and David M. Huang*,†

†Department of Chemistry and Centre for Advanced Nanomaterials, The University of Adelaide, South Australia 5005, Australia
‡International Institute for Nano/Meso Materials Science, Osaka Prefecture University, Sakai, Osaka 599-8531, Japan
¶Institute of Physical and Theoretical Chemistry, Graz University of Technology, Stremayrgasse 9, 8010 Graz, Austria
§CSIRO Future Industries, Clayton South, Victoria 3169, Australia

E-mail: david.huang@adelaide.edu.au; christian.doonan@adelaide.edu.au

Abstract

The ability to align porous metal–organic frameworks (MOFs) on substrate surfaces on a macroscopic scale is a vital step towards integrating MOFs into functional devices. But macroscale surface alignment of MOF crystals has only been demonstrated in a few cases. To accelerate the materials discovery process, we have developed a high-throughput computational screening algorithm to identify MOFs that are likely to undergo macroscale aligned heteroepitaxial growth on a substrate. Screening of thousands of MOF structures by this process can be achieved in a few days on a desktop
workstation. The algorithm filters MOFs based on surface chemical compatibility, lattice matching with the substrate, and interfacial bonding. Our method presents a simple and computationally efficient measure of the interfacial energy that considers both bond and defect formation at the interface. Further, we show that this novel approach is a better predictor of aligned heteroepitaxial growth than other, established, interface descriptors by testing our screening algorithm on a sample set of copper MOFs that have been grown heteroepitaxially on a copper hydroxide surface. Application of the screening process to several MOF databases, revealed that the top candidates for aligned growth on copper hydroxide comprise mostly MOFs with rectangular lattice symmetry in the plane of the substrate. This result indicates a substrate-directing effect that could be exploited in targeted synthetic strategies. We also identified that MOFs likely to form aligned heterostructures have broad distributions of in-plane pore sizes and anisotropies. Accordingly, this suggests that aligned MOF thin films with a wide range of properties may be experimentally accessible.

Keywords
metal–organic framework, porous material, heteroepitaxy, high-throughput screening, lattice matching

1 Introduction

Metal–organic frameworks (MOFs) are a class of materials formed by connecting metal nodes and organic links into an extended network. During the past two decades, a focus of MOF research has been to develop synthetic strategies for the precise control of pore structure and functionality on the single-crystal scale. However, more recently, extending the design principles of MOF chemistry from single-crystal to the macroscopic dimensions (nm to cm) has garnered significant interest. Approaches to MOF thin-film fabrication include layer-
by-layer synthesis, liquid-phase epitaxy and chemical vapor or atomic layer deposition.\textsuperscript{5,6} However, these techniques have yielded partially oriented thin films with out-of-plane alignment only.\textsuperscript{2,7–9} To fully realize the potential of crystalline MOF thin films in applications such as optoelectronics, gas separation, and sensing, alignment in both in-plane and out-of-plane directions is required.\textsuperscript{3,6} Recently, Falcaro and co-workers developed a one-pot methodology for the fabrication of centimeter-scale MOF films that achieved both in-plane and out-of-plane alignment.\textsuperscript{10} In this case, pre-aligned crystalline copper (II) hydroxide (Cu(OH)\textsubscript{2}) nanobelts were used as a sacrificial substrate for the epitaxial growth\textsuperscript{11} of copper-based MOFs and precise pore alignment was confirmed by comprehensive diffraction experiments. It was found that a necessary condition for epitaxial growth was a close matching between the MOF and Cu(OH)\textsubscript{2} lattice parameters. Such macroscale alignment was demonstrated only for seven MOFs, which represents a very small subset of known MOF structure types and functionalities. To fully exploit this fabrication strategy, an efficient procedure for selecting MOFs that possess lattice parameters that match those of the Cu(OH)\textsubscript{2} substrate from the vast database of known structures is essential.

Materials design through high-throughput computational screening offers a low-cost strategy to efficiently guide experimental endeavors. In recent years, advancements in computer hardware and the development of open databases of experimental and theoretical material properties, which are continually growing, has led to rapid growth in this research area.\textsuperscript{12–14} Indeed, high-throughput screening of candidate materials for aligned heteroepitaxial growth of MOFs would avoid an arduous and inefficient experimental approach. To this end, we have developed a high-throughput screening process to select and rank MOFs based on their likelihood to form aligned crystallites on a Cu(OH)\textsubscript{2} substrate. Although an atomic-level description of the growth mechanism and interfacial chemistry is necessary to fully characterize the kinetics and thermodynamics of these MOF films, such considerations are not amenable to rapid screening.\textsuperscript{4,15–19} Instead, we have extended established methods for evaluating interfacial lattice matching\textsuperscript{17,20–24} and chemical compatibility\textsuperscript{25} to assess MOF–substrate in-
teractions in a manner that accounts for framework porosity and allows for high-throughput screening. Geometrical lattice-matching algorithms\textsuperscript{20,21} have been employed in many screening studies as they offer an efficient way to exclude incompatible interfacial structures based on lattice mismatches.\textsuperscript{17,22–25} However, chemical compatibility at the interface is expected to be a more discriminating measure of interfacial stability than lattice matching.\textsuperscript{24,25} For example, Walsh and co-workers have reported an efficient screening process for matching contact layers in hybrid perovskites that considered the chemical compatibility at the interface by calculating the overlap of atoms on either side of the interface.\textsuperscript{25} Although this approach was demonstrated to be effective for non-porous solids, in this present work we show that the atomic site overlap is not the most suitable measure of interfacial binding for highly porous materials such as MOFs. We identify a simple and easily computed proxy for the interfacial energy that considers the difference between the number of bonds formed and broken as a result of interface formation, which accurately predicts macroscale alignment of MOFs on a Cu(OH)$_2$ substrate. Application of our screening process to the CORE MOF,\textsuperscript{26} hMOF,\textsuperscript{27} and TOBACCO\textsuperscript{28,29} databases shows that specific MOF topologies are favored for aligned growth from Cu(OH)$_2$, but that a wide range of pore network properties can nevertheless be obtained. Such knowledge will underpin design principles that will lead to a significant decrease in the experimental workload to develop precisely aligned MOF films.

2 Computational Approach

We have applied a three-step screening process: (1) selection of candidate MOFs based on chemistry (in this study, MOF structures were searched for that possess copper-bound carboxylate functionalities); (2) refinement of the data set based on lattice matching; and (3) ranking of MOFs by interfacial bonding (see Figure 1). Step 3 is the most computationally intensive: for example, analysis of 20 crystal structures took approximately 10 CPU hours on two Intel i5-6200U processors. Nevertheless, high-throughput screening can be achieved with
Figure 1: Flow chart of the screening process, in which a database of MOF crystal structures and a target substrate crystal structure act as inputs. The MOF database is filtered toward a smaller list of candidate structures for aligned heteroepitaxial growth. Each step is illustrated schematically on the right. Step 1 filters the database for MOFs that contain a carboxylate functionality bound to copper atoms and do not contain any metal other than copper. Step 2 checks if matching supercells (e.g., those outlined in solid lines, with dashed lines the corresponding unit cells; in this example the red supercell is the same size and shape as the unit cell) exist for a given MOF–substrate pair. Step 3 checks all possible supercells for interfacial bonding, allowing for translations along the supercell vectors.
this process: for example, \( \sim 5000 \) structures were screened in \( \sim 25 \) hours on six processors.

In general, crystallographic information files (CIFs) are used as the input for the MOF layer (film) and a single crystal structure (also from a CIF) is used as input for the substrate layer (\( \text{Cu(OH)}_2 \) in the following work). A definition of atom–atom binding across the interface between the two materials is also required. This work focuses on the epitaxial growth of MOFs from a metal-hydroxide surface (specifically carboxylate-bound copper MOFs on \( \text{Cu(OH)}_2 \)); however, the screening algorithm is not fundamentally limited to these specific systems.

### 2.1 Identifying Structures with Appropriate Chemistry

Step 1 of the screening process is designed to filter out all MOFs that do not possess the appropriate chemistry to bind to the substrate. This step can be generalized to any binding chemistry. The algorithm uses the Atomic Simulation Environment (ASE)\(^{30}\) and pymatgen\(^{31}\) Python libraries (description provided in Supporting Information Section S1). In this study, the search was limited to MOFs containing copper nodes and carboxylate functional groups with both oxygens bound to a copper atom (see Figure 1). This constraint was guided by experiments by Falcaro and co-workers, in which \( \text{Cu(OH)}_2 \) acted as a sacrificial substrate and metal source for the MOF. However, there is no experimental evidence that heteroepitaxial growth of MOFs is limited to sacrificial substrates or to MOFs of the same metal as the substrate. Walsh and co-workers\(^{22}\) have suggested that the chemical similarity of both sides of the interface (defined in their work by the difference in Pearson hardness between the metals in the two environments) is a driving force for stability at the interface. According to this criterion, copper-based MOFs on a copper-based substrate are expected to produce stable interfaces.
2.2 Identifying Structures with Matching Lattices

Step 2 of the screening process determines whether at least one Miller plane of a given MOF matches the lattice dimensions of the target substrate. In this work, the (010) Miller plane of Cu(OH)$_2$ was selected as it is known from experiments that this plane is exposed at the surface of this material. Zur and McGill have developed an efficient geometrical lattice-matching algorithm that has been widely adopted and shown to predict heteroepitaxial interfaces consistent with experimental data. The lattice-matching algorithm requires that the translational symmetry of each lattice matches at the interface within a specified tolerance. Although lattice matching is not a sufficient condition for growth, poor lattice matching can impact interfacial growth and can be used to rule out candidate structures for aligned MOF films. For example, large lattice mismatches have been found to anisotropically limit the size of hybrid perovskite crystals/grains in the direction of high mismatch. Given a set of primitive vectors for a specified lattice on one side of the interface (the unit cell), the algorithm (implemented using the pymatgen Python library) produces a series of supercells made up of an integer multiple of contiguous unit cells. A supercell is defined by two lattice vectors, which are reduced such that the supercell is unique. A pair of supercells on either side of the interface, with match areas $m$ times the MOF unit cell area and $n$ times the substrate unit cell area, are deemed to match based on four criteria: (1) the supercell match areas are below a defined threshold (set to the smaller of 90× the substrate unit cell area or 9× the film unit cell area), (2) the area ratio, $|A_{\text{MOF}}/A_{\text{sub}} - \frac{n}{m}|$, where $A_{\text{MOF}}$ and $A_{\text{sub}}$ are the MOF and substrate unit cell areas, respectively, is below a defined threshold (set to 15%), (3) corresponding lattice vectors are equal in length within a defined tolerance (set to 10%), and (4) the angles between the lattice vectors on either side of the interface are equal within a defined tolerance (set to 2%). (The choice of the parameters used in the matching criteria are discussed further in Section 3.1 and Supporting Information Section S2). In the screening process, a MOF structure is eliminated if no Miller planes with indices $-1 \leq h, k, l \leq 1$ are found to have a matching lattice with the substrate. If at
least one pair of supercells for a MOF–substrate pair passes all four tests, then that MOF continues to step 3 of the screening process.

### 2.3 Identifying Interfaces with Favorable Binding

Even if MOF and substrate lattices are found to match, chemical compatibility at the interface is expected to be the major factor in determining interfacial stability. Hence, the purpose of the final step of the screening algorithm is to identify interfaces with favorable binding by employing a geometric definition of interfacial binding. To do this, all possible binding planes of each MOF Miller plane that pass step 2 of the screening process are determined. We define MOF binding planes as those that include oxygens from carboxylate groups and Cu(OH)$_2$ binding planes as those that have a layer of copper atoms in the (010) Miller plane. When building the MOF binding plane, a 3D slab of the crystal structure with one face corresponding to the interface Miller plane is selected in which the axis bisecting the O–C–O angle of each carboxylate functionality at the surface is approximately orthogonal to the interface (see Supporting Information Section S3). We use this definition because density functional theory calculations show that monodentate binding (one oxygen bound to one metal) of deprotonated carboxylates with this axis orthogonal to a metal-oxide surface is most stable for the benzene dicarboxylate (BDC) ligand. To test the chemical compatibility between a selected Miller plane of a MOF and the (010) Miller plane of Cu(OH)$_2$, the degree of binding at the interface of the MOF and Cu(OH)$_2$ binding plane unit cells is calculated once the unit cells are mapped onto each of their corresponding supercells (identified in the lattice-matching step (step 2)). By iterating over all prescribed supercells and searching interfacial configurations, interfaces with the most favorable binding for each MOF–substrate pair are identified (see Supporting Information Section S3.4).

Several different geometric measures of interfacial binding were used to quantify the chemical compatibility of an interface. Each of these measures can take values between 0 (no binding) and 1 (perfect binding). One previously applied geometric definition of the
degree of bonding at the interface is the atomic site overlap (ASO), defined as

$$\text{ASO} = \frac{2S_C}{S_A + S_B},$$

(1)

where $S_A$ and $S_B$ are the number of binding atoms in the substrate and film binding planes, respectively, and $S_C$ is the number of coincident atom pairs at the interface. Throughout this work, two atoms on either side of the interface were defined as coincident (bonded) if their separation in the 2D binding plane was less than the mean of their ionic radii (0.96 Å for copper (0.57 Å) and oxygen (1.35 Å)). This definition for the ASO has previously been applied to determine the chemical stability of contact layers in hybrid perovskites; however, for interactions involving porous structures, such as MOFs, it yields low ASOs in general due to the disparity between the number of binding sites per unit area in the MOF and substrate layers as a result of the porous MOF structure. Therefore, we have redefined the ASO to be a function of the coordination of the MOF binding sites only as

$$\text{ASO} = \frac{S'_C}{S_B},$$

(2)

where $S'_C$ is the number of MOF (film) binding atoms that form bonds with the substrate and $S_B$ remains the number of film binding atoms, as defined above. Thus, the ASO describes the proportion of MOF binding atoms that can form bonds with the substrate.

The atom positions in a given supercell are not uniquely specified, but can be varied by rigid translations of the lattice along the supercell vectors without changing the supercell itself. We use a Metropolis Monte Carlo (MC) algorithm to apply random rigid translations along the MOF supercell vectors with respect to the fixed substrate to find the maximum ASO (see Supporting Information Section S4 for details). For computational simplicity, we assumed rigid MOF and substrate structures (we also did not consider the possibility of surface reconstruction); however, flexibility is implicitly taken into account by the lattice-matching and binding tolerances used in screening steps 2 and 3. Rigid frameworks could lead
to false negatives in the screening process as structural relaxation could enhance interfacial binding. Nevertheless, we found that the rigid approximation yields sufficient agreement with experimental results while allowing efficient screening (see Section 3.1). Structural distortions of the MOF framework\textsuperscript{10,16} and in-plane rotations that afford improved epitaxial matching\textsuperscript{39} have previously been shown to relieve interfacial stress in MOF heterostructures. Accordingly, our screening process accounts—either explicitly or implicitly—for mechanisms by which interfacial stress can be reduced by considering multiple interface orientations and inherent MOF flexibility.

As MOFs are formed from metal nodes and organic linkers, slicing their crystal structures without breaking strong covalent bonds (that may be part of the organic linkers) is non-trivial.\textsuperscript{22} We use a bond network approach to determine appropriate binding planes when slicing a MOF crystal at a particular Miller plane (see Supporting Information Section S3), which introduces missing linker defects (i.e. removes the entire linker) when covalent bonds of the linker are broken by the slicing process. This approach is similar to that used in recent work that applied a graph network analysis to determine the most stable way to slice a crystal structure based on the assumption that the most stable slice minimizes the number of broken bonds.\textsuperscript{40} A missing linker defect results in coordinatively unsaturated metal centers in the MOF near the binding plane, which is likely to be energetically unfavorable with respect to the pristine structure.\textsuperscript{41} However, we note that a number of MOFs possessing defect sites are known to retain their structural integrity.\textsuperscript{16,41–43} As the ASO does not account for the number of dangling or non-coordinated MOF atoms that are not carboxylate oxygens, thermodynamically unstable interfaces with a large proportion of dangling bonds can potentially have high ASOs and thus falsely predict macroscale aligned heteroepitaxial growth. Therefore, we propose a novel geometric measure of interfacial binding that is similar to the ASO, but also considers the number of dangling MOF bonds at the interface and thus is expected to be a better proxy for the interfacial energy. We call this quantity the change in
interfacial bonds or $\Delta IB$ and define it to be

$$\Delta IB = \frac{S'_C}{S_B + S_D} = \frac{n_{\text{formed}}}{n_{\text{formed}} + n_{\text{broken}}},$$

(3)

where $n_{\text{formed}}$ is the number of MOF binding atoms that make bonds with substrate binding atoms and $n_{\text{broken}}$ is the number of bonds broken when building the interface that do not form bonds at the interface, $S'_C$ and $S_B$ are as defined earlier, and $S_D$ is the number of MOF non-binding atoms with dangling bonds. $\Delta IB$ is identical to the ASO if no missing linker defects (broken bonds) are introduced when building the binding interface, but decreases relative to the ASO with increasing numbers of missing linker defects. We contend that considering the energetic cost of forming the interface using $\Delta IB$ provides a more accurate measure of interface stability than previously used geometric descriptors. Indeed, in Section 3.1.2 we demonstrate the superiority of the $\Delta IB$ over the ASO as a predictor for aligned heteroepitaxial growth of MOFs.

2.4 Structural Characterization

We measured crystal structure porosity using Zeo++,\textsuperscript{44} which applies a Voronoi decomposition to model the void space in a crystal structure and, due to its efficiency, has been widely employed for high-throughput screening processes.\textsuperscript{12,45–47} Where applicable, all Zeo++ calculations used a probe radius of 1.82 Å to represent $N_2$. The pore-limiting diameter (or largest free sphere) was calculated along each of the three crystallographic axes using the resex command in Zeo++. All structures were rendered with OVITO.\textsuperscript{48}
3 Results and Discussion

3.1 Validation of Screening Methodology

We validated our screening algorithm and determined an optimal set of parameters for each step and substep by comparing the results to experimental data obtained for a series of aligned MOFs heteroepitaxially grown through a one-pot synthetic method on a Cu(OH)$_2$ (010) surface. The experimental MOF data set (given in Supporting Information Table S2) comprised six ditopic carboxylate-linked MOFs and CuBTC (also known as HKUST-1), a well-studied 3D MOF with a tritopic carboxylate linker. Two of the dicarboxylate-linked MOFs were 3D pillared frameworks (containing 1,4-biphenyl dicarboxylate (BDC) and either 4,4’-bipyridine (BPY) or 1,4-diazabicyclo[2.2.2]octane (DABCO) coordinated to the axial positions of the copper paddle-wheel), while the remaining four were 2D eclipsed sheet structures. Each of the MOFs possess copper paddle-wheels as their metal nodes. Aside from the 2D MOF formed from the 1,4-terphenyl dicarboxylate (TDC) linker, all ditopic MOFs were shown to grow with in-plane alignment on the Cu(OH)$_2$ substrate. CuBTC did not form crystals with in-plane alignment, but out-of-plane alignment was reported. Note that our algorithm specifically identifies structures that are likely to grow with in-plane alignment on Cu(OH)$_2$.

A length mismatch tolerance between supercell vectors of 10%, an angle mismatch tolerance of 2%, a maximum match area of the smaller of 90× the substrate unit cell area (∼1350 Å$^2$ for Cu(OH)$_2$ (010)) or 9× the film unit cell area, and a maximum area ratio tolerance of 15% were used. A larger length mismatch tolerance was used than generally applied in the lattice-matching literature due to the flexible nature of MOFs. This choice is supported by experimental data for synthesized MOF-on-MOF heterostructures, in which large lattice mismatches are alleviated by MOF flexibility. To allow testing of MOFs with large unit cells and to ensure that all MOFs in the experimental data set were found to have at least one Miller plane with a substrate-matching lattice, the maximum match area
used was much larger than values previously used for metal-oxide interfaces ($\sim 600 \AA^2$)\textsuperscript{20,21} or for MOF-on-metal-oxide or MOF-on-metal interfaces ($20 \times$ the unit cell area for the MOF or metal/metal-oxide).\textsuperscript{22} A limit on the number of unit cells in the supercell was imposed because large-area interfaces are unlikely to form without defects. Thus, comparison of computed results for very large match areas with experiment is not expected to be meaningful.

An analysis of the effect of varying the tolerances and maximum area constraints on the final results for a test database found that the chosen parameters provide robust results (see Supporting Information Section S2).

Figure 2: Categorical scatter plots of all crystal structures in the experimental data set (listed in Supporting Information Tables S2 and S3) showing categorization based on the (a) minimum coincident match area, (b) minimum ratio of MOF and substrate unit-cell areas, (c) maximum ASO, and (d) maximum $\Delta IB$. The shaded region indicates the approximate threshold $\Delta IB$ above which in-plane heteroepitaxial alignment occurs. The red line indicates the $\Delta IB$ threshold (0.67) applied throughout the remainder of this work. Separated nets of CuBDC-BPY are shown as CuBDC-BPY-i-1 and CuBDC-BPY-i-2, while the interpenetrated structure is not shown (see Supporting Information Section S7). Categorization is based on whether the associated Miller plane grows experimentally from Cu(OH)$_2$ with in-plane alignment.
Figure 2 shows the experimental data set\textsuperscript{10} categorized by four different descriptors: (1) the minimum coincident match area, (2) the minimum ratio of MOF and substrate unit cell areas, (3) the maximum ASO, and (4) the maximum $\Delta IB$ obtained for each crystal structure in the experimental data set (Table S3). Screening of materials, including MOFs, for heteroepitaxial growth using the lattice-matching algorithm\textsuperscript{20,21} has previously shown that top candidates can be determined by ranking matching lattices by their minimal coincident match area (or the smallest unit-cell multiple that yields a matching lattice).\textsuperscript{17,22,23} This is because interfaces with smaller match areas (and smaller mismatch) are expected to have fewer defects.\textsuperscript{20,21,25} Figure 2a shows that the minimum coincident match area does not distinguish between MOFs that undergo aligned heteroepitaxial growth and those that do not; furthermore, in some cases for which such growth occurs, the MOF interface that grows on the substrate is not the one with the minimum match area. On the other hand, Figure 2b shows that the minimum ratio of the MOF and substrate unit-cell areas for all Miller planes successfully categorizes almost all the MOFs in the experimental data set and suggests that a small ratio of MOF and substrate unit-cell areas is correlated with aligned heteroepitaxial growth.

Both the minimum coincident match area and unit-cell area ratio are geometrical descriptors that do not consider the chemistry at the interface and are significantly faster to calculate than the ASO or $\Delta IB$. But the neglect of interface chemistry misses crucial features of interface formation. For example, in each of the cases in Figure 2a for which the lowest match-area interface is not one that is known to exhibit heteroepitaxial growth, the MOF binding interface comprises pairs of carboxylate groups bound to the same copper paddlewheel that are each at approximately a $45^\circ$ angle to the plane of the interface. Such binding configurations would require significant distortion of the MOF and/or substrate structures to avoid steric clashes and form stable interfaces, which is exemplified by their low ASOs and $\Delta IB$ (not shown in Figure 2). Figure 2b suggests that aligned heteroepitaxial growth occurs for a unit-cell area ratio below about 10. However, several MOFs in the validation data set
have Miller planes that satisfy this criterion but are not observed to undergo aligned growth on Cu(OH)$_2$. By contrast, the $\Delta$IB descriptor, which does account for bonding between atoms at the interface, correctly categorizes almost all of the MOF interfaces that grow aligned on Cu(OH)$_2$, as shown in Figure 2d, without predicting aligned growth for Miller planes that have not been observed. On the other hand, the ASO, which also considers interfacial binding between atoms and has previously been used to categorize heteroepitaxial growth,\textsuperscript{25} does not appear to be as good a predictor of aligned heteroepitaxy as $\Delta$IB, as indicated by the lack of any obvious trend in Figure 2c.

Figure 2 shows that CuTDC is an outlier by both the unit-cell area-ratio and $\Delta$IB descriptors; i.e. by these measures, it would be predicted to undergo aligned heteroepitaxial growth that is not observed experimentally. We suggest that experimental factors associated with CuTDC that are not accounted for in the screening algorithm, such as the low solubility of the TDC linker, may inhibit aligned MOF growth. Out of the MOFs that experimentally show aligned heteroepitaxial growth on Cu(OH)$_2$, only CuNDC has a lower ASO or $\Delta$IB than CuTDC. CuNDC has been found to deform at MOF–MOF interfaces with high lattice mismatch,\textsuperscript{16} which our screening algorithm does not consider, and so this MOF could also potentially be miscategorized by the metrics in Figure 2. Neglecting CuTDC, Figure 2 suggests that MOFs with a $\Delta$IB above a threshold somewhere between 0.60 and 0.77 undergo aligned heteroepitaxial growth, while those with $\Delta$IB below this threshold do not. We have applied a threshold value around the midpoint of this range ($\Delta$IB = 0.67) in further analysis. For all MOFs for which aligned heteroepitaxial growth has been shown to occur experimentally, only one in-plane orientation of the MOF crystallites was found. Our screening results are consistent with this observation: for example, the MOFs in Figure 2d with $\Delta$IB above the threshold for aligned heteroepitaxial growth ($\geq$ 0.67) have no other in-plane orientations or Miller planes with a $\Delta$IB above this threshold. (The method used to identify unique interfaces is described in Supporting Information Section S5.1). The following two sections discuss the robustness of the screening algorithm and highlight the
importance of considering the interfacial energy (via $\Delta IB$) to avoid false positives.

### 3.1.1 Effect of Crystal Structure

In many cases, multiple published crystal structures exist for a given set of MOF building blocks (i.e. organic linker and metal node). For example, there are at least 17 entries of CuBTC in the CORE MOF database, which contains approximately 5000 experimental MOF structures. The lattice parameters of these structures vary in many cases by a small amount ($\sim 0.1 \text{ Å}$). Step 3 of our screening process uses the structures of matching MOF and substrate supercells identified in step 2 to calculate the ASO and $\Delta IB$. Each supercell has a particular relative orientation of the atoms on either side of the interface, with only certain relative orientations aligning the binding sites such that the ASO (or $\Delta IB$) is high (see Supporting Information Section S3). Both the lattice mismatch and the set of matching supercells depend on the crystal lattice parameters, and these affect the ASO and $\Delta IB$. We have used a set of crystal structures (Supporting Information Table S3) for the MOFs CuBDC, CuBDC-DABCO, and CuBTC in the experimental data set to investigate the impact of different crystal structures of the same MOF on the results of our screening process.

For both CuBDC and CuBDC-DABCO, slight differences between unit-cell vectors for different crystal structures yield MOF supercells with the same highest ASO that have the same orientation but different area. This is because small changes in the unit-cell vectors can lead to certain supercells failing the lattice-matching step of the screening. However, these differences do not impact the screening process, as the same ASO or $\Delta IB$ is obtained for all crystal structures. On the other hand, in the case of CuBTC, Figure 3a shows that the particular crystal structure can affect the maximum ASO and, more importantly, the Miller plane with the highest ASO. Figure 3b–e shows the four different binding interfaces found for all CuBTC crystal structures. This difference between the maximum ASO for different CuBTC structures is a consequence of two factors. First, small changes to the unit-cell vectors can result in a given supercell orientation of a particular MOF Miller plane...
failing the lattice-matching criteria, i.e. different crystal structures yield different supercell orientations and ASOs. This explains the dramatic difference between the maximum ASO of the binding interface (Figure 3b) for CuBTC-g compared with the other crystal structures (Figure 3a). Second, small changes to the unit-cell vectors modify the spacing between atoms in the supercell, which can impact the coincidence of atoms at the interface. Figure 3f shows the orientation and lattice spacing that gives rise to the ASO of 1 for CuBTC-g and an example interface with a maximum ASO \(\approx 0.33\) for the other CuBTC crystal structures. These issues could potentially be partially resolved by using less stringent interface bonding criteria, but large tolerances required to yield comparable ASOs for all crystal structures for the interface in Figure 3b would likely lead to spurious lattice and binding matches for other MOFs. On the other hand, the maximum \(\Delta IB\) is comparable and low for all CuBTC crystal structures (Figure 2d). Thus, \(\Delta IB\) does not show the same sensitivity to the crystal structure lattice parameters as the ASO and so appears to be a more robust descriptor for aligned heteroepitaxial growth.
3.1.2 Cost of Interface Formation on Chemical Compatibility

Figure 3b shows the binding interface of the (111) Miller plane of CuBTC and highlights the carboxylate-oxygen binding sites and coordinatively unsaturated copper missing linker defects. This interface gives the maximum possible ASO of 1 for the CuBTC-g crystal.
structure. However, this surface of CuBTC has not been shown to exist experimentally on the facets of CuBTC crystals,\textsuperscript{49} nor has it been shown to be the most energetically stable.\textsuperscript{50} Furthermore, it has not been shown to bind to surfaces in any examples of surface-anchored CuBTC structures.\textsuperscript{51,52} This contradiction between the computed ASO and experimental observations indicates that the ASO is not always a good predictor of favorable interface formation. In particular, the ASO ignores energetically unfavorable broken or dangling bonds. In the case of the (11\bar{1}) Miller plane of CuBTC, 16 bonds are broken per unit cell to form the interface in Figure 3b.

On the other hand, the $\Delta$IB metric that we have introduced considers both the bonds that are formed and those that are broken to form the interface, resulting in a low $\Delta$IB of 0.125 for this interface. In contrast, the (111) and (100) Miller planes, which have been experimentally observed at CuBTC surfaces, each have lower ASOs but no missing linker defects are required to form interfaces with Cu(OH)$_2$ (Figure 3d and e); thus, they have higher $\Delta$IBs ($\approx$0.36 and $\approx$0.28, respectively) than the (11\bar{1}) Miller plane. This result is consistent with previous theory and experiment, in which the (111) and (100) Miller planes have been found to be the most stable surfaces\textsuperscript{49,50} and have been shown to form interfaces with other materials.\textsuperscript{51,52} In contrast to the ASO, $\Delta$IB is similar and low for all Miller planes of CuBTC, indicating that in-plane heteroepitaxial alignment of this MOF on Cu(OH)$_2$ is not favored. This agrees with experimental results that show crystal growth of CuBTC with out-of-plane orientation only.\textsuperscript{10} It is also possible that the presence of multiple interfaces with similar $\Delta$IB values could lead to competitive growth of multiple distinct orientations, which could be the cause of diminished in-plane alignment on the macroscale for CuBTC. In either case, this example highlights the superiority of the $\Delta$IB over the ASO in explaining the experimental observations.
3.2 Screening CORE MOF Database

Figure 4: Nitrogen-accessible void fraction versus pore-limiting diameter for all crystal structures in the CORE MOF database. Points are colored by how far the MOF progressed through the screening process.

We have applied our screening algorithm to the Computation-Ready, Experimental (CORE) MOF database. The CORE MOF database comprises approximately 5000 MOF structures with pore-limiting diameters greater than 2.4 Å that have been processed by removing solvent from the pores, fixing atom position disorder, and converting each structure to its primitive unit cell. Figure 4 shows the nitrogen-accessible void fraction as a function of the pore-limiting diameter (also known as the largest free sphere) of all structures in the CORE MOF database. For each crystal structure, the screening process was applied with the validated parameters determined from the analysis of the experimental data set in Section 3.1. Only 8% of the CORE MOF database passed the chemistry test and 89% of those MOFs passed the lattice-matching step, resulting in 7% of the database or 336 structures whose interfacial binding was tested. 20 of these structures were interpenetrated, which led to 40 additional structures representing the separated nets of those 20 structures that were screened. (The processing of interpenetrated structures is described in Supporting Information Section S7.) To emphasize the efficiency of the screening process, the first two steps took minutes on six Intel i7-4790k CPU cores for the entire CORE MOF database, while the final step for the remaining 376 structures took approximately 25 hours on six CPUs.
Figure 5 shows the maximum $\Delta IB$ of each of the 376 structures that passed the first two screening steps as a function of the ratio of the MOF and substrate unit-cell areas. Because macroscale aligned heteroepitaxial growth relies on the preferential growth of only one MOF crystal orientation with respect to the substrate, the points in Figure 5 are colored according to the second highest $\Delta IB$ ($\Delta IB_2$), which corresponds to a different Miller plane or a different in-plane orientation the same Miller plane. Macroscale aligned heteroepitaxial growth is expected for MOFs for which the maximum $\Delta IB$ is above the threshold determined in Section 3.1, with the second highest $\Delta IB$ below this threshold. The symbol shape in Figure 5 indicates the MOF topology, which was obtained from the TTO ToposPro topological database.\textsuperscript{53,54} The standard cluster representation of 3D MOFs was used (this database uses a mixture of the Reticular Chemistry Structure Resource (RCSR) three-letter codes\textsuperscript{55} and Topos-specific codes).\textsuperscript{56} The cluster representation separates the atomic net into building blocks (such as the metal clusters and organic linkers), which are then treated as nodes to determine the underlying network topology. Structures listed as unknown did not have cluster representation topology entries in the database. Similar to what was found for the limited experimental data set in Figure 2, there is no clear correlation between the maximum $\Delta IB$ and the ratio of the unit-cell areas for the MOF structures in the CORE MOF database. Although all structures with high $\Delta IB$ values correspond to low unit-cell area ratios, the converse is not true, with structures with low $\Delta IB$ exhibiting a wide range of unit-cell area ratios. Therefore, the data suggests that a low ratio of unit cell areas does lead to favorable interfaces, as has previously been suggested,\textsuperscript{22} but the interfacial energy (through $\Delta IB$) should be considered to identify MOFs likely to show aligned heteroepitaxial growth.
Figure 5: Maximum $\Delta IB$ of crystal structures in the CORE MOF database that passed screening steps 1 and 2 as a function of the ratio of the MOF and substrate unit-cell areas for the Miller plane with the maximum $\Delta IB$. Coloring is by the value of the second highest $\Delta IB$ ($\Delta IB_2$). The symbol shape indicates the 3D MOF topology (structures with no reported topology are given as ‘unknown’). The percentage of MOFs that passed screening step 2 of each topology is given in the legend. The shaded region indicates the approximate $\Delta IB$ threshold for heteroepitaxial in-plane alignment determined in Section 3.1. The red line indicates the value of $\Delta IB$ used as a threshold for top candidates.

From Figure 5, the top candidates in the MOF database for aligned heteroepitaxial growth on Cu(OH)$_2$ can be identified as those with maximum $\Delta IB$ above the previously defined threshold of approximately 0.67 (see Section 3.1) and second highest $\Delta IB$ ($\Delta IB_2$) below this threshold. Figure 6 shows the interface structures of the top candidates (MOFs with similar structures have been grouped together). Three of the top candidates (NEJRUR, NEJSAY and NEJSEC) correspond to different crystal structures (all of which are interpenetrated MOFs with pcu topology) of the MOF CuBDC-BPY (Figure 6a), which has been shown experimentally to grow aligned from Cu(OH)$_2$ (one of these structures, NEJRUR, was in the parameterization data set). Top candidate CEHPIP (Figure 6b) corresponds to an interpenetrated MOF that is isoreticular to CuBDC-BPY, in which the BDC linker has been replaced with 1,4-cyclohexanedicarboxylic acid. CEHPIP also has an underlying pcu topology. Top candidate ZAZBUZ (tfc topology) is made up of BDC linkers bound to chain-like copper nodes, rather than the paddle-wheel in CuBDC (Figure 6e) and the top candidate ZECKID has a dia topology (Figure 6d). Two binding planes (equivalent within a 90° rotation) for the same Miller plane of ZECKID were found with the same maximum
ΔIB, one of which was excluded manually upon visual inspection of all top candidates. Finally, one of the top candidates corresponds to an interpenetrated MOF (UNABUH, Figure 6e), which has an unknown topology by the cluster representation (using the standard unsimplified representation, it has a tfs topology).

Figure 5 along with experimental data show that the Cu(OH)$_2$ (010) substrate favors the growth of MOFs with the pcu topology (many of the high-ranking points with pcu topology are overlapping in Figure 5a). Furthermore, the Miller planes that meet the ΔIB criteria for top candidates for aligned heteroepitaxial growth have approximate rectangular symmetry, as indicated by their in-plane unit-cell lattice vectors (Supporting Information Figure S10a). These findings suggest that the rectangular symmetry of the (010) plane of Cu(OH)$_2$ selects out top candidates for aligned heteroepitaxial growth with similar symmetry. Note that MOFs with pcu and other similar topologies that commonly have at least one lattice plane with rectangular symmetry have been shown previously to exhibit in-plane alignment when grown via heteroepitaxy;$^{39,57}$ importantly, the pcu topology has been shown to produce superior internal surface areas compared with other MOF topologies, suggesting that the ability to target pcu MOFs with heteroepitaxy could be beneficial for developing useful MOF heterostructures.$^{58}$ The top candidate structures comprise mostly interfaces with one carboxylate binding site per unit cell (or two oxygen binding sites per unit cell), similar to the experimental data set (Figure 6). Note that MOFs with topologies other than pcu have high maximum ASOs, while their maximum ΔIB is significantly diminished due to the number of missing linker defects required to form the interface. This result suggests that certain topologies may not grow heteroepitaxially due to the difficulty of cleaving the crystal structure at a planar interface without breaking a large number of bonds.
3.3 Selective MOF Growth by Topology

Experimental data\(^{10}\) as well as the top candidates found using our screening process on the CORE MOF database suggest that the rectangular symmetry of the (010) Miller plane of Cu(OH)$_2$ preferentially selects MOFs with at least one Miller plane with rectangular symmetry. The majority of selected MOFs were found to also have pcu topology, a topology that often corresponds to structures having lattice planes of rectangular symmetry. The ability to select a specific MOF topology could greatly aid the development of MOF heterostructures and provides a design principle for further experimental study. To investigate whether selectivity by MOF topology is expected to occur generally for Cu(OH)$_2$, we applied our screening process to two other MOF databases, the TOBACCO database\(^{28,29}\) and
the hMOF database,\textsuperscript{27} which contain hypothetical MOF structures that were built using \textit{in silico} structure generation. The TOBACCO database contains approximately 13 000 non-interpenetrated MOF structures constructed from a diverse range of metal nodes, organic linkers, and underlying network topologies.\textsuperscript{28,29} Figure 7a shows the maximum ∆IB as a function of the ratio of the MOF and substrate unit-cell areas for all crystal structures that passed screening steps 1 and 2 (10.3% of the full database). The points are also colored by the value of the second highest ∆IB. Note that none of the MOFs in this database with \textbf{pcu} topology passed the chemistry test of the screening process, as they were all zinc MOFs. Nonetheless, the screening process picks out a small subset of topologies that are predicted to yield aligned heteroepitaxial growth on Cu(OH)\textsubscript{2}, namely \textbf{lvt-b} and \textbf{nbo-b}. The \textbf{lvt-b} and \textbf{nbo-b} topologies have networks similar to those with the \textbf{pcu} topology. Screening of the TOBACCO database predicts again that the Cu(OH)\textsubscript{2} substrate selects MOFs with only a small set of topologies as well as Miller planes with in-plane rectangular symmetry (see Supporting Information Figure S10b). Note that there are MOFs in the TOBACCO database with topologies other than \textbf{lvt-b} and \textbf{nbo-b} with high maximum ASO and low maximum ∆IB, which suggests that the ability to cleave 3D MOF structures at a 2D planar interface with minimal broken bonds also drives their selection. In addition, Figure 7a shows that many binding planes in the TOBACCO database have large unit-cell areas compared with the Cu(OH)\textsubscript{2} substrate, which are not expected to be good candidates for aligned heteroepitaxial growth because large, pristine substrate surfaces would be required, which is unlikely to occur in real systems.
Figure 7: Maximum ΔIB of all crystal structures that passed screening steps 1 and 2 as a function of the ratio of the MOF and substrate unit-cell areas for the (a) TOBACCO database\textsuperscript{28,29} and (b) hMOF database.\textsuperscript{27} Points are colored by the value of the second highest ΔIB (ΔIB\textsubscript{2}). The symbol shape indicates MOF topology in (a) (the percentage of MOFs in the entire TOBACCO database with the specified topologies are given in the legend). The shaded region indicates the approximate ΔIB threshold for heteroepitaxial in-plane alignment determined in Section 3.1. The red line indicates the value of ΔIB used as a threshold for top-candidates.

To confirm the favorability of the pcu topology for heteroepitaxial alignment on the Cu(OH)\textsubscript{2} substrate, we screened the hMOF database, which contains approximately 150,000 hypothetical MOF structures, approximately 90% of which have underlying pcu nets.\textsuperscript{27} Only the unfunctionalized and non-interpenetrated structures were considered, resulting in only 2570 structures that were screened, 27% or 693 of which made it to the final step of the screening process. Figure 7b shows that out of the 693 structures, 266 have only orientation of one Miller plane with ΔIB greater than the 0.67 threshold, all of which have pcu topologies, which supports the assertion that the pcu topology is favorable for aligned growth on the Cu(OH)\textsubscript{2} substrate. As observed for the other two MOF databases, these top candidates
for aligned growth also have binding planes with approximate rectangular symmetry (see Supporting Information Figure S10c). Figure 7b also shows that many MOFs in this database (145 in total) have multiple Miller planes with \( \Delta IB \geq 0.67 \), unlike the experimental data set, in which only one Miller plane was found to have a high \( \Delta IB \) for each MOF. This is because each hMOF structure is built from up to three unique (but sometimes structurally similar) ligands from a ligand database, leading to more inequivalent Miller planes per MOF and often to multiple Miller planes with similar \( \Delta IB \) values. Many MOFs in the hMOF database also have similar organic linkers, which results in similar structures and interfaces and helps to explain why many MOFs have a maximum \( \Delta IB \) above the threshold. Both the hMOF and TOBACCO databases were compiled from mostly hypothetical structures, which limits their applicability in an experimental setting. Nevertheless, screening of the three MOF databases has identified a small number of MOF topologies (which are all similar) as useful starting points for future experimental investigation of MOF heteroepitaxy.

### 3.4 Templating pore architecture through epitaxial growth

Intuitively, aligned heteroepitaxially grown MOF crystals could lead to anisotropy in the pores along the different directions with respect to the substrate. To test this hypothesis, we calculated the pore-limiting diameters in the \( a \), \( b \), and \( c \) crystallographic dimensions for all pcu MOFs in the hMOF database whose maximum \( \Delta IB \) occurs for a Miller plane with \( h + k + l = 1 \) (a subset of those in Figure 7b). This subset of MOFs was chosen because, for the pcu topology, it can reasonably be assumed that the pore windows align with the three crystallographic axes of the MOF for these Miller planes, for which the pore-limiting diameter was calculated. Supporting Information Figures S12 and S13 show that the distribution of the maximum in-plane pore-limiting diameters does not differ substantially from the distribution of the maximum overall pore-limiting diameters of their crystals. High \( \Delta IB \) MOFs that are predicted to have aligned heteroepitaxial growth have a similar distribution of in-plane pore-limiting diameters to low \( \Delta IB \) MOFs, which suggests that epitaxial growth
does not constrain the in-plane pore-limiting diameters as a result of the registry with the substrate lattice. This suggests that a broad range of possible pore sizes can be achieved for a given substrate and, therefore, MOFs with precise pore sizes can be chosen for target applications.

Falcaro and co-workers\textsuperscript{10} showed that macroscale alignment of the CuBPDC MOF on a Cu(OH)\textsubscript{2} substrate could be employed to align fluorescent molecules within the MOF and yield anisotropic fluorescence from the MOF film. It follows that anisotropy of the pore network is an important design aspect for MOF thin films that could be controlled through in-plane heteroepitaxial alignment. Furthermore, the directionality of the pores with respect to the surface due to a preferred crystal orientation has been shown to impact adsorption properties of porous-coordination polymers,\textsuperscript{59} and such properties can be predicted once the in-plane orientation is determined. To investigate this concept further, we calculated the pore network anisotropies from the pore-limiting diameters along the three crystallographic axes (described above) for each \textbf{pcu} MOF in the hMOF database with a maximum $\Delta IB \geq 0.67$ associated with a Miller plane with $h + k + l = 1$. Pore network anisotropies were calculated as either the ratio or the absolute difference between the maximum in-plane pore-limiting diameter ($PLD_{\text{in-plane, max}}$) and the pore-limiting diameters in the other two crystallographic dimensions (the minimum in-plane pore-limiting diameter ($PLD_{\text{in-plane, min}}$) or the out-of-plane pore-limiting diameter ($PLD_{\text{out-of-plane}}$)). Figure 8 shows that the top candidates from our screening process for the hMOF database cover a broad range of anisotropies and pore shapes (using either definition), including some very anisotropic pore networks. While this allows for the selection of the desired pore properties for a target application, we note that the definition of anisotropy that matters will be different for certain applications. For example, a MOF in which all three pores are large may have high anisotropy with respect to the difference between the pore sizes, but low anisotropy with respect to their ratios. This type of pore network would not provide any noticeable anisotropic effect for the adsorption of relatively small molecules, but may be useful for some other application, such as alignment.
of large guests.

Figure 8: Pore network anisotropy for all MOF structures in the hMOF database predicted to grow aligned on Cu(OH)$_2$ whose maximum $\Delta IB$ is associated with a Miller plane where $h+k+l=1$: (a) ratio and (b) absolute difference of the largest in-plane pore-limiting diameter ($PLD_{in-plane,max}$) and minimum in-plane pore-limiting diameter ($PLD_{in-plane,min}$) versus that of the largest in-plane pore-limiting diameter and out-of-plane pore-limiting diameter ($PLD_{out-of-plane}$). Points are colored by the value of the maximum $\Delta IB$ and only those with a second highest $\Delta IB < 0.67$ are shown.

4 Conclusions

We have developed a robust and efficient screening process of metal–organic frameworks (MOFs) for aligned heteroepitaxial growth on metal-hydroxide substrates, which we have validated by comparison with experimental measurements. This process applies a geomet-
tical lattice-matching algorithm\textsuperscript{20,21} and extends previous methodologies for determining chemical compatibility at an interface\textsuperscript{25} to filter and categorize MOF structures. We have proposed a novel metric for the favorability of interfacial binding, $\Delta \text{IB}$, which is simple and efficient to compute and accounts for both the energetic gain of bond formation and the energetic cost of broken and dangling bonds. Importantly, our screening process can be generalized to cases other than those studied in this work. The MOF in the experimental data set made from the TDC linker was found to be an outlier in the validation of the screening process, which we suggest could be due to experimental difficulties associated with the solubility of that linker. This finding highlights the need to consider the chemistry of the MOF building blocks of any top candidates generated by computational screening before further experimental evaluation. Our screening process does not consider the kinetics or thermodynamics of MOF growth from the substrate or the possibility of competing binding modes of the ligands used to form the MOF (for example, the nitrogen donors of bipyridine ligands may coordinate to copper species and compete with carboxylate groups). More computationally intensive simulations that quantitatively model different interfacial interactions could be used after initial high-throughput screening in the final selection of MOF candidates. Nonetheless, the criteria we have implemented allows for an efficient screening process that considers the necessary conditions for aligned heteroepitaxial growth of MOF crystallites on the macroscale.

We used the screening process to compile a list of top candidates for aligned heteroepitaxial growth from the CORE MOF,\textsuperscript{26} hMOF,\textsuperscript{27} and TOBACCO\textsuperscript{28,29} databases, which equated to screening approximately 20,000 structures in a few days on a desktop workstation. The screening results showed that MOFs that have at least one lattice plane with rectangular symmetry have the greatest propensity for aligned heteroepitaxial growth on Cu(OH)$_2$, a condition that is commonly satisfied only for certain MOF topologies. (The \texttt{pcu}, \texttt{nbo-b} and \texttt{lvt-b} topologies were found to make up the majority of top candidates). This finding indicates a substrate-directing effect, whereby the symmetry of the substrate surface determines
the preferential symmetry of MOF binding sites at the interface and, hence, the symmetry of the MOF binding plane. Knowledge of substrate-dependent design principles like this allows for topologically targeted synthetic strategies, which offers an efficient route to viable experimental targets.\textsuperscript{58,60,61} We identified six, experimentally realized MOF structures (REFCODES: ZAZBUZ, CEHPIP, NEJRUR, NEJSAY, NEJSEC, and UNABUH) that are likely to form aligned heteroepitaxial structures on Cu(OH)$_2$.

Furthermore, we used the hypothetical MOF database (hMOF), which comprises mostly pcu MOF structures, to show that the dimensions of the substrate lattice, which determine the registry and alignment of the MOF, do not constrain the available range of pore diameters and the anisotropy of the pore networks. A broad range of possible pore architectures means that the design of MOF thin films can be tailored toward target applications. The screening algorithm we have described offers a fast and robust way to find MOF candidates for further experimental analysis in the growing field of MOF thin films and device development. In addition, the categorization of vast MOF databases afforded the determination of design principles and structure–property relationships.

All code used in this work is freely available at https://bitbucket.org/andrewtarzia/epitmof/src/master/.

**Supporting Information Available**

Detailed description of screening algorithm, parameter selection, list of MOFs in experimental data set and corresponding crystallographic information files (CIFs), and additional screening results.

This material is available free of charge via the Internet at http://pubs.acs.org/.
Acknowledgement

A. T. was supported by an Australian Government RTP Scholarship and a PhD top-up scholarship from CSIRO Division of Materials Science and Engineering. CJD would like to acknowledge the JSPS for a short-term visiting professorship.

References

(1) Furukawa, H.; Cordova, K. E.; O’Keeffe, M.; Yaghi, O. M. The Chemistry and Applications of Metal-Organic Frameworks. Science 2013, 341, 1230444.

(2) Stassen, I.; Burtch, N.; Talin, A.; Falcaro, P.; Allendorf, M.; Ameloot, R. An Updated Roadmap for the Integration of Metal–Organic Frameworks with Electronic Devices and Chemical Sensors. Chem. Soc. Rev. 2017, 46, 3185–3241.

(3) Shekhah, O.; Liu, J.; Fischer, R. A.; Woll, C. MOF Thin Films:Existing and Future Applications. Chem. Soc. Rev. 2011, 40, 1081–1106.

(4) Haraguchi, T.; Otsubo, K.; Kitagawa, H. Emergence of Surface- and Interface-Induced Structures and Properties in Metal–Organic Framework Thin Films. Eur. J. Inorg. Chem. 2018, 2018, 1697–1706.

(5) Stassen, I.; Campagnol, N.; Fransaer, J.; Vereecken, P.; De Vos, D.; Ameloot, R. Solvent-Free Synthesis of Supported ZIF-8 Films and Patterns through Transformation of Deposited Zinc Oxide Precursors. CrystEngComm 2013, 15, 9308–9311.

(6) Falcaro, P.; Ricco, R.; Doherty, C. M.; Liang, K.; Hill, A. J.; Styles, M. J. MOF Positioning Technology and Device Fabrication. Chem. Soc. Rev. 2014, 43, 5513–5560.

(7) Liu, J.; Wöll, C. Surface-Supported Metal–Organic Framework Thin Films: Fabrication Methods, Applications, and Challenges. Chem. Soc. Rev. 2017, 46, 5730–5770.
(8) Oldenburg, M.; Turshatov, A.; Busko, D.; Jakoby, M.; Haldar, R.; Chen, K.; Emandi, G.; Senge, M. O.; Wöll, C.; Hodgkiss, J. M.; Richards, B. S.; Howard, I. A. Enhancing the Photoluminescence of Surface Anchored Metal–Organic Frameworks: Mixed Linkers and Efficient Acceptors. Phys. Chem. Chem. Phys. 2018, 20, 11564–11576.

(9) Liu Bo.; Shekhah Osama.; Arslan Hasan K.;; Liu Jinxuan.; Wöll Christof.; Fischer Roland A., Enantiopure Metal–Organic Framework Thin Films: Oriented SURMOF Growth and Enantioselective Adsorption. Angew. Chem. Int. Ed. 2011, 51, 807–810.

(10) Falcaro, P.; Okada, K.; Hara, T.; Ikigaki, K.; Tokudome, Y.; Thornton, A. W.; Hill, A. J.; Williams, T.; Doonan, C.; Takahashi, M. Centimetre-Scale Micropore Alignment in Oriented Polycrystalline Metal-Organic Framework Films via Heteroepitaxial Growth. Nat. Mater. 2017, 16, 342–348.

(11) Reboul, J.; Furukawa, S.; Horike, N.; Tsotsalas, M.; Hirai, K.; Uehara, H.; Kondo, M.; Louvain, N.; Sakata, O.; Kitagawa, S. Mesoscopic Architectures of Porous Coordination Polymers Fabricated by Pseudomorphic Replication. Nat. Mater. 2012, 11, 717–723.

(12) Thornton, A. W.; Simon, C. M.; Kim, J.; Kwon, O.; Deeg, K. S.; Konstas, K.; Pas, S. J.; Hill, M. R.; Winkler, D. A.; Haranczyk, M.; Smit, B. Materials Genome in Action: Identifying the Performance Limits of Physical Hydrogen Storage. Chem. Mater. 2017, 29, 2844–2854.

(13) Simon, C. M.; Kim, J.; Gomez-Gualdron, D. A.; Camp, J. S.; Chung, Y. G.; Martin, R. L.; Mercado, R.; Deem, M. W.; Gunter, D.; Haranczyk, M.; Sholl, D. S.; Snurr, R. Q.; Smit, B. The Materials Genome in Action: Identifying the Performance Limits for Methane Storage. Energy Env. Sci 2015, 8, 1190–1199.

(14) Coudert, F.-X.; Fuchs, A. H. Computational Characterization and Prediction of
Metal–Organic Framework Properties. *Coord. Chem. Rev.* 2016, 307, Part 2, 211–236, Chemistry and Applications of Metal Organic Frameworks.

(15) Stavila, V.; Volponi, J.; Katzenmeyer, A. M.; Dixon, M. C.; Allendorf, M. D. Kinetics and Mechanism of Metal–Organic Framework Thin Film Growth: Systematic Investigation of HKUST-1 Deposition on QCM Electrodes. *Chem. Sci.* 2012, 3, 1531–1540.

(16) Wang, Z.; Liu, J.; Lukose, B.; Gu, Z.; Weidler, P. G.; Gliemann, H.; Heine, T.; Wöll, C. Nanoporous Designer Solids with Huge Lattice Constant Gradients: Multiheteroepitaxy of Metal–Organic Frameworks. *Nano Lett.* 2014, 14, 1526–1529.

(17) Bristow, J. K.; Butler, K. T.; Svane, K. L.; Gale, J. D.; Walsh, A. Chemical Bonding at the Metal–Organic Framework/Metal Oxide Interface: Simulated Epitaxial Growth of MOF-5 on Rutile TiO2. *J. Mater. Chem. A* 2017, 5, 6226–6232.

(18) Shekhah, O.; Wang, H.; Zacher, D.; Fischer, R. A.; Wöll, C. Growth Mechanism of Metal–Organic Frameworks: Insights into the Nucleation by Employing a Step-by-Step Route. *Angew. Chem. Int. Ed.* 2009, 48, 5038–5041.

(19) Wang, Z.; Weidler, P. G.; Azucena, C.; Heinke, L.; Wöll, C. Negative, Anisotropic Thermal Expansion in Monolithic Thin Films of Crystalline Metal-Organic Frameworks. *Microporous Mesoporous Mater.* 2016, 222, 241–246.

(20) Zur, A.; McGill, T. C. Lattice Match: An Application to Heteroepitaxy. *J. Appl. Phys.* 1984, 55, 378–386.

(21) Zur, A.; McGill, T. C.; Nicolet, M.-A. Transition-metal Silicides Lattice-matched to Silicon. *J. Appl. Phys.* 1985, 57, 600–603.

(22) Butler, K. T.; Hendon, C. H.; Walsh, A. Designing Porous Electronic Thin-Film Devices: Band Offsets and Heteroepitaxy. *Faraday Discuss.* 2017, 201, 207–219.
(23) Ding, H.; Dwaraknath, S. S.; Garten, L.; Ndione, P.; Ginley, D.; Persson, K. A. Computational Approach for Epitaxial Polymorph Stabilization through Substrate Selection. *ACS Appl. Mater. Interfaces* **2016**, *8*, 13086–13093.

(24) Raclariu, A.-M.; Deshpande, S.; Bruggemann, J.; Zhuge, W.; Yu, T.; Ratsch, C.; Shankar, S. A Fast Method for Predicting the Formation of Crystal Interfaces and Heterocrystals. *Comput. Mater. Sci.* **2015**, *108, Part A*, 88–93.

(25) Butler, K. T.; Kumagai, Y.; Oba, F.; Walsh, A. Screening Procedure for Structurally and Electronically Matched Contact Layers for High-Performance Solar Cells: Hybrid Perovskites. *J. Mater. Chem. C* **2016**, *4*, 1149–1158.

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

(27) Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q. Large-Scale Screening of Hypothetical Metal-Organic Frameworks. *Nat. Chem.* **2012**, *4*, 83–89.

(28) Gómez-Gualdrón, D. A.; Colón, Y. J.; Zhang, X.; Wang, T. C.; Chen, Y.-S.; Hupp, J. T.; Yildirim, T.; Farha, O. K.; Zhang, J.; Snurr, R. Q. Evaluating Topologically Diverse Metal–Organic Frameworks for Cryo-Adsorbed Hydrogen Storage. *Energy Environ. Sci* **2016**, *9*, 3279–3289.

(29) Colón, Y. J.; Gómez-Gualdrón, D. A.; Snurr, R. Q. Topologically Guided, Automated Construction of Metal–Organic Frameworks and Their Evaluation for Energy-Related Applications. *Cryst. Growth Des.* **2017**, *17*, 5801–5810.

(30) Larsen, A. H. et al. The Atomic Simulation Environment—a Python Library for Working with Atoms. *J. Phys. Condens. Matter* **2017**, *29*, 273002.
(31) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* 2013, *68*, 314–319.

(32) Oswald, H. R.; Reller, A.; Schmalle, H. W.; Dubler, E. Structure of Copper(II) Hydroxide, Cu(OH)$_2$. *Acta Cryst. C* 1990, *46*, 2279–2284.

(33) Chen, J.; Fu, Y.; Samad, L.; Dang, L.; Zhao, Y.; Shen, S.; Guo, L.; Jin, S. Vapor-Phase Epitaxial Growth of Aligned Nanowire Networks of Cesium Lead Halide Perovskites (CsPbX$_3$, X = Cl, Br, I). *Nano Lett.* 2017, *17*, 460–466.

(34) Sun, W.; Ceder, G. Efficient Creation and Convergence of Surface Slabs. *Surf. Sci.* 2013, *617*, 53–59.

(35) Domenico, J.; Foster, M. E.; Spoerke, E. D.; Allendorf, M. D.; Sohlberg, K. Effect of Solvent and Substrate on the Surface Binding Mode of Carboxylate-Functionalized Aromatic Molecules. *J. Phys. Chem. C* 2018, *122*, 10846–10856.

(36) Mathew, K.; Singh, A. K.; Gabriel, J. J.; Choudhary, K.; Sinnott, S. B.; Davydov, A. V.; Tavazza, F.; Hennig, R. G. MPInterfaces: A Materials Project Based Python Tool for High-Throughput Computational Screening of Interfacial Systems. *Comput. Mater. Sci.* 2016, *122*, 183–190.

(37) Ahrens, L. H. The Use of Ionization Potentials Part 1. Ionic Radii of the Elements. *Geochim. Cosmochim. Acta* 1952, *2*, 155–169.

(38) Shannon, R. D. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Cryst. A* 1976, *32*, 751–767.

(39) Furukawa, S.; Hirai, K.; Nakagawa, K.; Takashima, Y.; Matsuda, R.; Tsuruoka, T.; Kondo, M.; Haruki, R.; Tanaka, D.; Sakamoto, H.; Shimomura, S.; Sakata, O.; Kita-
gawa, S. Heterogeneously Hybridized Porous Coordination Polymer Crystals: Fabrication of Heterometallic Core–Shell Single Crystals with an In-Plane Rotational Epitaxial Relationship. *Angew. Chem. Int. Ed.* **2009**, *48*, 1766–1770.

(40) Witman, M.; Ling, S.; Boyd, P.; Barthel, S.; Haranczyk, M.; Slater, B.; Smit, B. Cutting Materials in Half: A Graph Theory Approach for Generating Crystal Surfaces and Its Prediction of 2D Zeolites. *ACS Cent. Sci.* **2018**, *4*, 235–245.

(41) Bristow, J. K.; Svane, K. L.; Tiana, D.; Skelton, J. M.; Gale, J. D.; Walsh, A. Free Energy of Ligand Removal in the Metal–Organic Framework UiO-66. *J. Phys. Chem. C* **2016**, *120*, 9276–9281.

(42) Fang, Z.; Bueken, B.; De Vos, D. E.; Fischer, R. A. Defect-Engineered Metal–Organic Frameworks. *Angew. Chem. Int. Ed.* **2015**, *54*, 7234–7254.

(43) St. Petkov, P.; Vayssilov, G. N.; Liu, J.; Shekhah, O.; Wang, Y.; Wöll, C.; Heine, T. Defects in MOFs: A Thorough Characterization. *ChemPhysChem* **2012**, *13*, 2025–2029.

(44) Willems, T. F.; Rycroft, C. H.; Kazi, M.; Meza, J. C.; Haranczyk, M. Algorithms and Tools for High-Throughput Geometry-Based Analysis of Crystalline Porous Materials. *Microporous Mesoporous Mater.* **2012**, *149*, 134–141.

(45) Altintas, C.; Avci, G.; Daglar, H.; Nemati Vesali Azar, A.; Velioglu, S.; Eruçar, I.; Keskin, S. Database for CO2 Separation Performances of MOFs Based on Computational Materials Screening. *ACS Appl. Mater. Interfaces* **2018**, *10*, 17257–17268.

(46) Qiao, Z.; Xu, Q.; Jiang, J. High-Throughput Computational Screening of Metal-Organic Framework Membranes for Upgrading of Natural Gas. *J. Memb. Sci.* **2018**, *551*, 47–54.

(47) Evans, J. D.; Huang, D. M.; Haranczyk, M.; Thornton, A. W.; Sumby, C. J.; Doonan, C. J. Computational Identification of Organic Porous Molecular Crystals. *CrystEngComm* **2016**, *18*, 4133–4141.
(48) Stukowski, A. Visualization and Analysis of Atomistic Simulation Data with OVITO—the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.* 2010, 18, 015012.

(49) Mao, Y.; Su, B.; Cao, W.; Li, J.; Ying, Y.; Ying, W.; Hou, Y.; Sun, L.; Peng, X. Specific Oriented Metal–Organic Framework Membranes and Their Facet-Tuned Separation Performance. *ACS Appl. Mater. Interfaces* 2014, 6, 15676–15685.

(50) Amirjalayer, S.; Tafipolsky, M.; Schmid, R. Surface Termination of the Metal-Organic Framework HKUST-1: A Theoretical Investigation. *J. Phys. Chem. Lett.* 2014, 5, 3206–3210.

(51) Liu, J.; Shekhah, O.; Stammer, X.; Arslan, H. K.; Liu, B.; Schüpbach, B.; Terfort, A.; Wöll, C. Deposition of Metal-Organic Frameworks by Liquid-Phase Epitaxy: The Influence of Substrate Functional Group Density on Film Orientation. *Materials* 2012, 5, 1581–1592.

(52) Biemmi, E.; Scherb, C.; Bein, T. Oriented Growth of the Metal Organic Framework Cu₃(BTC)₂(H₂O)₃⋅xH₂O Tunable with Functionalized Self-Assembled Monolayers. *J. Am. Chem. Soc.* 2007, 129, 8054–8055.

(53) Blatov, V. A.; Shevchenko, A. P.; Proserpio, D. M. Applied Topological Analysis of Crystal Structures with the Program Package ToposPro. *Cryst. Growth Des.* 2014, 14, 3576–3586.

(54) Alexandrov, E. V.; Blatov, V. A.; Kochetkov, A. V.; Proserpio, D. M. Underlying Nets in Three-Periodic Coordination Polymers: Topology, Taxonomy and Prediction from a Computer-Aided Analysis of the Cambridge Structural Database. *CrystEngComm* 2011, 13, 3947–3958.

(55) O’Keeffe, M.; Peskov, M. A.; Ramsden, S. J.; Yaghi, O. M. The Reticular Chemistry Structure Resource (RCSR) Database of, and Symbols for, Crystal Nets. *Acc. Chem. Res.* 2008, 41, 1782–1789.
(56) Bonneau, C.; O’Keeffe, M.; Proserpio, D. M.; Blatov, V. A.; Batten, S. R.; Bourne, S. A.; Lah, M. S.; Eon, J.-G.; Hyde, S. T.; Wiggin, S. B.; Öhrström, L. Deconstruction of Crystalline Networks into Underlying Nets: Relevance for Terminology Guidelines and Crystallographic Databases. *Cryst. Growth Des.* 2018, 18, 3411–3418.

(57) Otsubo, K.; Haraguchi, T.; Sakata, O.; Fujiwara, A.; Kitagawa, H. Step-by-Step Fabrication of a Highly Oriented Crystalline Three-Dimensional Pillared-Layer-Type Metal–Organic Framework Thin Film Confirmed by Synchrotron X-Ray Diffraction. *J. Am. Chem. Soc.* 2012, 134, 9605–9608.

(58) Martin, R. L.; Haranczyk, M. Insights into Multi-Objective Design of Metal–Organic Frameworks. *Cryst. Growth Des.* 2013, 13, 4208–4212.

(59) Hirai, K.; Sumida, K.; Meilikhov, M.; Louvain, N.; Nakahama, M.; Uehara, H.; Kitagawa, S.; Furukawa, S. Impact of Crystal Orientation on the Adsorption Kinetics of a Porous Coordination Polymer–Quartz Crystal Microbalance Hybrid Sensor. *J. Mater. Chem. C* 2014, 2, 3336–3344.

(60) Boyd, P. G.; Woo, T. K. A Generalized Method for Constructing Hypothetical Nanoporous Materials of Any Net Topology from Graph Theory. *CrystEngComm* 2016, 18, 3777–3792.

(61) Addicoat, M. A.; Coupry, D. E.; Heine, T. AuToGraFS: Automatic Topological Generator for Framework Structures. *J. Phys. Chem. A* 2014, 118, 9607–9614.
Graphical TOC Entry

1. MOF database
2. Graphical representation of MOF on target substrate
3. Aligned MOF on target substrate