Two-Dimensional Metal–Organic Framework Self-Assembly and Defect Engineering Studied *via* Coarse-Grained Simulations

Reum N. Scott,*,† Claire E. Frank,† Maya M. Martirossyan,† Phillip J. Milner,*,‡ and Julia Dshemuchadse*,†

†*Department of Materials Science and Engineering, Cornell University, Ithaca, NY 14853, USA*

‡*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA*

E-mail: rns76@cornell.edu; pjm347@cornell.edu; jd732@cornell.edu

Abstract

Metal–organic frameworks (MOFs) are crystalline materials that self-assemble from inorganic nodes and organic linkers, and isoreticular chemistry allows for modular and synthetic reagents of various sizes. In this study, a MOF's components — metal nodes and organic linkers — are constructed in a coarse-grained model from isotropic beads, retaining the basic symmetries of the molecular components. Lennard-Jones and Weeks–Chandler–Andersen pair potentials are used to model attractive and repulsive particle interactions, respectively. We analyze the crystallinity of the self-assembled products and explore the role of modulators — molecules that compete with the organic linkers in binding to the metal nodes, and which we construct analogously — during the self-assembly process of defect-engineered MOFs. Coarse-grained simulation allows for the uncoupling of experimentally interdependent variables to broadly map and determine essential MOF self-assembly conditions, among which are properties of the modulator: binding strength, size (steric hindrance), and concentration. Of these, the simulated modulator's binding strength has the most pronounced effect on the resulting MOF's crystal size.

Introduction

The ever-expanding field of metal–organic frameworks (MOFs) is propelled by isoreticular chemistry, *i.e.*, the ability to vary the MOF components while retaining the same topological net.[1–4] Metal-containing secondary building units (SBUs) and organic linkers are joined by strong bonds in a periodic repetition of the building units in a crystalline MOF.[5] MOFs have vast applications—in chemical separations[6] and gas storage,[2,7,8] fuel cells, super-capacitors, and catalysis[6,9–11]—as a result of the variability of possible MOF components and the porous crystalline structure in which they grow. Producing molecules such that they form a given ordered, crystalline structure[12] with defined and ideally tunable materials properties is at the heart of materials design, for which MOFs therefore provide an ideal medium. Furthermore, structural errors can become desirable features[13] if their location and concentration can be controlled. Control of structural defects, *i.e.*, defect engineering, along with control over the size and shape of crystalline materials, offers an opportunity to further tailor a synthesized material's properties.[14–20] One approach to introduce well-defined defects into MOFs is to incorporate modulators—monodentate ligands that compete with linkers in binding to node sites, forming modulator–linker substitution defects.[21,22] Modulator-induced defect engineering has been used to enhance the catalytic and gas sorption properties of MOFs and control their crystallite sizes and morphologies.[23]

With a vast selection of SBUs, linkers, modulators, and synthesis conditions, experimental sampling to determine desirable MOF structures and properties from the resulting combination of parameters becomes infeasible. We can use coarse-grained simulations to build models of MOFs and understand their self-assembly *via* a systematic reduction of molecular structures to representative "rigid bodies" composed of beads that capture the general intermolecular forces and molecular geometry, allowing for the self-assembly of ordered MOF structures. The isoreticular property of MOFs lends itself to using simple structural constructs of beads in simulation to represent the building units and functional groups with pair potentials used to mimic the intermolecular interactions between SBU and linker in solution. Previously, coarse-grained molecular dynamics have been used to simulate the synthesis of the covalent–organic frameworks (COFs), MOFs' metal-free relatives, with the results suggesting experimental strategies for growing COFs with enhanced crystallinity.[24]

The Zn-carboxylate MOF PPF-1[25] (see Fig. 1a) is used as an instructive example system to establish design rules for the coarse-grained model for the simulation of MOF self-assembly. The coarseness of the simulated model, *i.e.*, the degree to which the molecules' structural and chemical properties are represented, impacts the degree of crystallinity of the assembled MOF (see Fig. 1d–i). Simulated modulator molecules—reagents used to control the assembly process and allowing for greater reversibility of MOF formation, *i.e.*, to correct structural errors[21,26–29]—are introduced to understand defect formation along with tuning MOF reactivity and crystalline domain size. Two-dimensional MOF assembly not only reduces the degrees of freedom of this computational study, but also finds application in MOF synthesis on surfaces.[30]

In building a coarse-grained model, we are guided by the symmetry of the MOF, and we add a level of tunability to the simulated synthesis by introducing a modulator molecule. By simulating the MOF self-assembly *via* coarse-grained molecular dynamics simulations, we explore the influence of modulation on the crystallinity, resulting in the tuning of MOF properties through defect engineering. We also validate the fidelity of the model by relating our results to observed trends in MOF chemical synthesis.

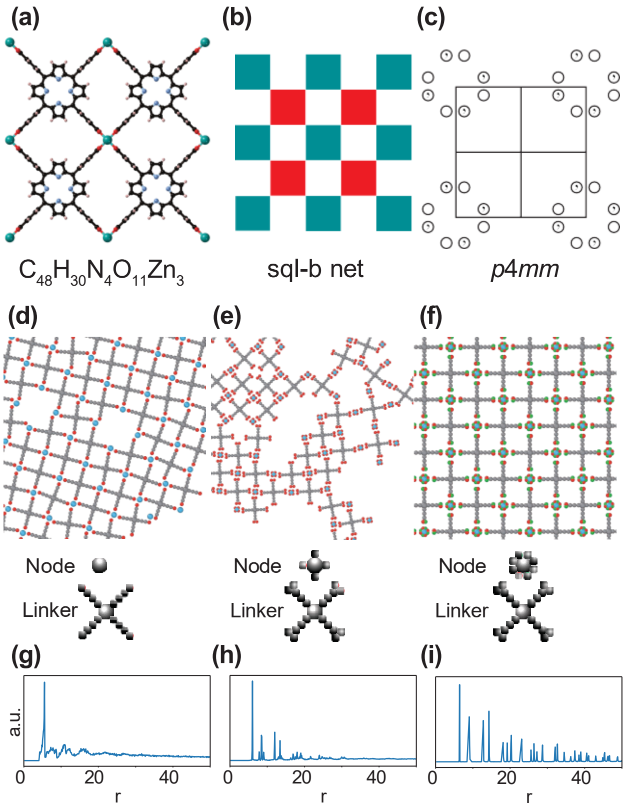


Figure 1: (**a**) The structure of the two-dimensional (2D) Zn-MOF, (**b**) the corresponding topological net (sql-b), (**c**) the corresponding plane group (*p*4*mm*), as well as snapshots of self-assembly simulations of the 2D MOF with decreasing degrees of coarseness (**d–f**) and their respective radial distribution functions (**g–i**).

Self-Assembly Simulations

Simulations were performed using the molecular dynamics algorithm implemented in the *HOOMD-blue* software package. [31] An NVT canonical ensemble was simulated, with a constant total system size of $N = 882$ particles (nodes, linkers, and modulators), employing a Langevin integrator and a Nosé–Hoover thermostat. The 2D simulation box is initialized with periodic boundary conditions at a high starting temperature ($T_{\text{start}} = 0.3$ for the majority of the simulations; $T_{\text{start}} = 0.2$ for the coarser-grained models) and is successively cooled to a final temperature of 0.01 over 10^8 molecular dynamics steps, with a step size of $\delta\tau = 0.005$. The particles are initialized on a dilute square grid, and the identity of the particles in the initialization stage is assigned at random based on the node-to-linker-to-modulator ratio of $\left[1 + \frac{x}{4}\right] : 1 : x$. After the original particles are assigned to be the central beads of nodes, linkers, or modulators, additional bead particles are added to define the intended shape and interactions of each building block.

The *signac* data management program was used for simulation production and simulation data management.[32,33] The final MOF structures were analyzed using the *freud* package[34,35] and visualizations were created using *OVITO*.[36,37]

The final simulation frame of each self-assembly simulation was used to compute radial distribution functions (RDFs) ($r = 0.1$–50 with 500 bins). Coordination numbers (CNs) were determined by counting the nearest neighbors around each node particle within the final simulation frame. A radial cutoff value was determined as the position of the minimum after the first maximum of the RDF, which corresponds to the maximum distance between a pair of nearest neighbors (a node and a linker). The reported average coordination number is computed by taking the mean of all particles in a simulation frame.

The porphyrin paddle-wheel framework, PPF-1 ($C_{48}H_{30}N_4O_{11}Zn_3$),[25] was chosen as an example system to develop the coarse-grained two-dimensional (2D) self-assembly model (see Fig. 1a), as its three-dimensional structure is the product of a staggered stacking of 2D MOF layers. Meso–tetra(4–carboxyl–phenyl) porphyrin ($TCPP^{2-}$) is used as a planar tetratopic linker to allow for the formation of 2D rigid frameworks due to the carboxylate functional group's ability to aggregate metal ions. [38] The space group symmetry of the three-dimensional MOF is $I4/mmm$,[39] and the symmetry of the 2D sql-b net simulated in the plane—representing one layer of said MOF—corresponds to the wallpaper group $p4mm$ (see Fig. 1b).

The Lennard-Jones (LJ) potential[40,41] is chosen to model attractive interactions:

$$V_{\mathrm{LJ}} = 4\varepsilon_{\mathrm{LJ}}\left[\left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{6}\right],$$

and the Weeks–Chandler–Andersen (WCA) potential[42,43]—a LJ potential that is truncated and shifted to zero at its minimum—is chosen to model repulsive interactions between coarse-grained beads that represent components of the MOF building blocks and is employed to prevent bead overlap:

$$V_{\mathrm{WCA}} = 4\varepsilon_{\mathrm{WCA}} \quad \frac{\sigma_{\mathrm{WCA}}}{r}^{12} - \frac{\sigma_{\mathrm{WCA}}}{r}^{6} + \varepsilon_0.$$

Results and Discussion

Coarse-Grained Model Development

Initially, we simulated the coarse-grained MOF self-assembly by only considering the four-fold symmetry (4*mm*) of the linker to produce an assembly of beads to mimic the tetratopic linker with four arms, with the node reduced to a single bead (see Fig. 1d). Only the terminal beads of the linker arms, representative of the carboxylate functional groups around the porphyrin ring, are attracted to the node. The number of beads composing the linker arms along with the values of the free parameters in the LJ and WCA pair potentials shown in Tab. 1 for the various bead types (node and linker end) were optimized for the assembly of the square MOF lattice. The cutoff radius for the node–node repulsion is increased to prevent the trapping of nodes between linker arms (see Tab. 1). Upon increasing the linker arm length in test simulations to ensure the correct coordination of nodes with 4 linkers, a linker arm length of 4 beads was found to be optimal (see Fig. 1d), ensuring sufficient steric hindrance to prevent over-coordination of nodes with linkers while minimizing the number of required bead particles and therefore computational cost. The radial distribution function (RDF), used here to qualitatively evaluate long-range order, showed only a defined peak for nearest neighbors (see Fig. 1g). The lack of well-defined crystallites necessitated the addition of greater specificity of the intermolecular interactions to the coarse-grained model.

To improve the definition of the MOF grains, the orientation of the local bonding environments of the node and linker, *i.e.*, the four-fold coordination symmetry of the node and

Table 1: Values for the pair-potential constants determined for the initial coarse-grained simulation.

	Pair Potential Parameters			
	$\varepsilon_{\mathrm{LJ}}$	σ_{LJ}	$\varepsilon_{\mathrm{WCA}}$	σ_{WCA}
Node	1.0	1.4	1.0	4.5
Linker				1.6

linker in the sql-b net (see Fig. 1b), are encoded with specific binding points at 90° angles added around the node in a finer-grained model. Additionally, two equivalent binding points on the linker arms are added to emulate the chelating behavior of the carboxylate functional groups. The crystallinity of the MOF does increase (see Fig. 1e) with more distinguishable RDF peaks at a greater radial distance along with overall narrower peaks (see Fig. 1h), corresponding to better-defined relative particle positions on the crystal lattice and crystals growing to larger grain sizes. Similarly, Monte Carlo simulation for the packing of porous species had previously shown that the directionality and orientation of the bonds formed impacts the phase behaviour, and that strong direct molecular interactions allow for only one ordered phase to form.[44]

Further improvement of the crystallinity is achieved by encoding the anisotropy of the local environments in the MOF crystal *via* two different beads with selective attraction at each of the four binding sites around the node (shown in red and green color in Fig. 1f). This amendment of the model now also integrates the handedness of the node and linker molecule (see Fig. 2), by having a red–green bead pair on the node be attracted to a mirrored pair on the linker. The anisotropy of the simulated node and linker beads can be seen as a reflection of the symmetry present in the $p4mm$ wallpaper group (see Fig. 1c), where the positions equivalent to red *vs.* green beads are not identical but mirror images of each other. The greater apparent definition of the MOF grains is observed qualitatively *via* the RDF, which has several narrow peaks that propagate beyond 50 a.u. (see Fig. 1i). Incorporating directional bonding and saturability (coordination number) for reagents of MOF simulations produces an assembled framework reflecting the symmetry of the topological net and is in agreement with simulations of surface-confined metal–organic nanostructures (SMONs).[45] With the finer coarse-grained model allowing for successful self-assembly of the desired MOF structure, further study into how reaction parameters impact the crystallization process, including modulation, is possible.

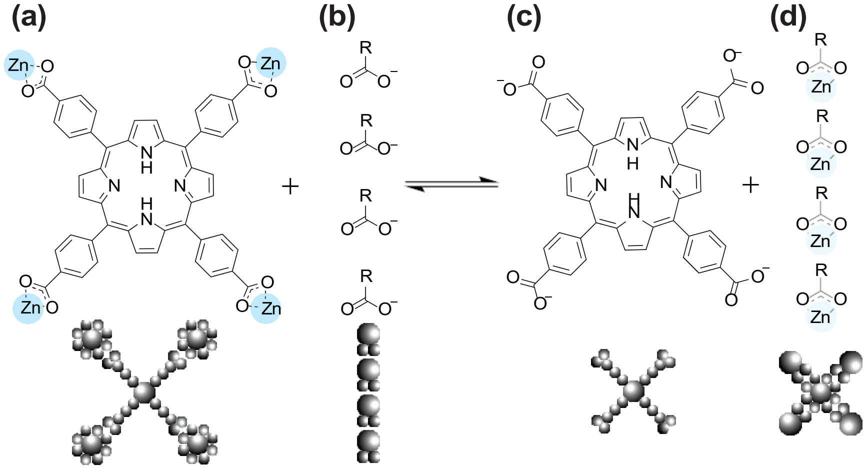


Figure 2: The mechanism for competitive coordination modulation: the deprotonated conjugate base of the modulator competes with the linker for coordination to the node. (**a**) The tetradentate porphyrin linker bound to four Zn-nodes through the carboxylate, (**b**) the carboxylate modulator in solution, (**c**) the unbound tetradentate porphyrin linker, (**d**) the carboxylate modulators bound to a node

Modulation

Modulators, such as monocarboxylic acids,[28,46,47] are used to improve MOF crystallinity by enhancing the reversibility of MOF assembly, and they can also introduce defects during the assembly process.[19] The influence of acid modulators on the domain size and reactivity of crystalline MOFs have been widely investigated through experimental work, using hydrochloric,[48] formic,[49] acetic,[28,49] trifluoroacetic,[49] benzoic acids,[28,50] as well as methanesulfonic and pivalic acids, phenol, and benzenesulfonamide, [47] among many others. The pK_a ($-\log_{10}K_a$) of these acids is best reflected by the binding strength of the simulated modulator and the reversibility of the binding process, as modulators with more basic conjugate bases should interact more strongly with the nodes. The same coarse-grained two-bead binding pattern used for the linker arms is employed to simulate the modulator, and the relative attraction of node-to-modulator is varied through the parameter ε_{n-m} in the LJ potential, as the atoms neighboring the carboxylate or Lewis base can be electron donating or withdrawing and thus can either raise or lower the binding attraction strength of the modulator to the node. The competitive coordination mechanism for modulation uses a conjugate base (carboxylate or general anion) of the modulator to compete with the linker for coordination with the node (see Fig. 2). A single functional group on a modulator enables it to act as a monodentate linker to produce missing-cluster defects or missing-linkers, where the modulator building units incorporated into the assembled MOF structure act as truncated linkers. [49,51–53] The modulator's presence in the assembled MOF is reflected in the RDF by the appearance of a new peak at a shorter radial distance—due to the modulator binding site's direct connection to its central bead while 5 beads (node binding site, linker binding site, linker arm beads) separate the central beads of the node and linker.

The effective size of the modulator molecule can have steric effects[47] and is represented by additional beads to form arms at two sites perpendicular to the modulator binding site; we vary the modulator arm length from 0 to 3 beads (see Fig. 3a). Modulator concentration is the parameter that has been most thoroughly studied in previous experimental studies, as it can be readily adjusted without changing any other modulator properties. In general, higher concentrations of modulator relative to the linker concentration lead to more missing-linker defects and more effective modulation.[54] Modulator concentration can be readily studied using coarse-grained models by increasing the modulator-to-linker ratio while keeping the modulator size and binding strength constant across a series of simulations. A balance in the concentration of "modulator added" with respect to the number of nodes is employed, as experimental studies have shown that crystallization can be halted in modulator excess.[55] To maintain charge balance—represented as parity in binding sites—the ratio of node-to-linker-to-modulator ($n : l : m$) is set as $\left\lceil 1 + \frac{x}{4} \right\rceil : 1 : x$, while increasing equivalents of modulator x in simulation from 0 to 10 in increments of 2. This means that four bidentate modulators replace a single tetratopic linker. The coarse-grained modeling approach allows for three properties—size, binding strength to the node, and concentration—of the simulated modulator to be varied independently throughout the simulations of MOF self-assembly in 2D.

Modulator size

Modulator incorporation into the lattice is observed for smaller simulated modulators, with arms consisting of fewer than 3 beads (see Fig. 3), when the node–modulator interaction strength is equal to the strength of the node–linker interaction (at a constant $\varepsilon_{n\text{-}m} = 1.0$, $\varepsilon_{n\text{-}l} = 1.0$) and at a modulator quantity of 4 equivalents; $n : l : m = 2 : 1 : 4$. Increasing the effective size of the modulator shows fewer modulator particles incorporated into the final lattice and more located on the surfaces of crystallites (see Fig. 3e), while the final MOF grain size remains relatively constant. The incorporated modulator building blocks cause distortions of the lattice. As the modulator arm length increases, the larger modulators are increasingly excluded from the MOF and the peaks of the RDF become increasingly defined (see Fig. 3b–e). The lattice distortions observed at smaller modulator sizes result in greater variance in the expected positions of neighbors, *i.e.*, broader RDF peaks, and increasingly so at larger distances. As larger modulators are more excluded from the lattice due to steric effects, there is less distortion around the nodes: modulators with an arm length of 3 are completely excluded from pores of the final lattice and decorate the crystal surface (see Fig. 3e), such that the modulators do not displace linkers in the interior of the crystallites anymore. For the largest modulator, the size of the modulator impacts the crystalline order of the MOF but does not considerably impact the size of the assembled MOF crystal grains. This finding supports that sterically small modulators (relative to the linker) are optimal if missing-linker defects are desired. This work is consistent with experimental findings, which showed that smaller carboxylic acids tend to be more effective modulators than chemically similar but larger modulators.[47]

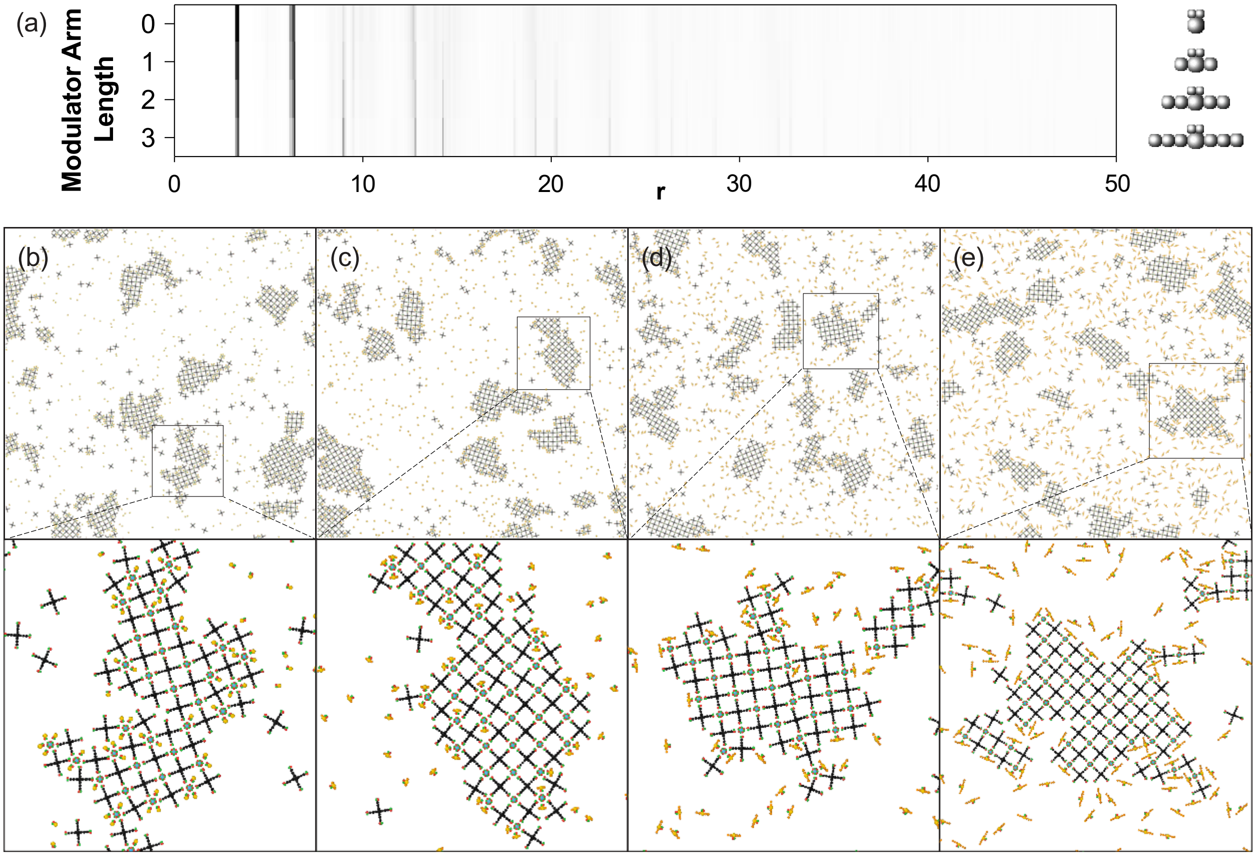


Figure 3: The steric effect of modulator arm length at a constant ratio of equivalents ($n : l : m = 2 : 1 : 4$) and constant node–modulator interaction strength $\varepsilon_{n\text{-}m} = 1.0$: (**a**) radial distribution functions of the self-assembled final simulation product, as well as final simulation frames with a modulator arm length of (**b**) 0, (**c**) 1, (**d**) 2, and (**e**) 3.

Modulator binding strength

Increasing the node-to-modulator binding strength $\varepsilon_{n\text{-}m}$ at constant modulator concentration for the smallest modulator size leads to a decrease in crystal domain size (see Fig. 4). As the modulator binds to the node more tightly, the probability of the bond being broken decreases as the system is cooled, and the modulator acts as a capping agent. The current simulated MOF self-assembly with modulation captures the molecule's function as a capping agent and not as an acid — or H^+ source — that protonates the linker functional groups. Consistently, the MOF MIL-101(Cr) exhibited decreasing crystallite size as the pK_a-equivalent binding strength of the modulator increased in experiments.[56] Similarly, simulated modulator capping is minimally observed for $\varepsilon_{n\text{-}m} = 0.5$: the modulators are still present in the solution, and the largest MOF crystallites are observed (see Fig. 4b). For $\varepsilon_{n\text{–}l} = \varepsilon_{n\text{–}m} = 1.0$, the effect of the direct competition between linker and modulator binding to the node is evident with a stronger RDF peak corresponding to the modulator (see Fig. 4a), fewer modulators in solution, and greater numbers of modulator on the MOF surface and within the pores of the MOF (see Fig. 4c). No individual modulators are observed in solution when $\varepsilon_{n\text{–}l} < \varepsilon_{n\text{–}m}$, and some linker molecules remain uncoordinated to nodes (see Fig. 4d–i): as node–modulator binding strength increases, the modulators can fully exclude the linkers from binding to the nodes with as many as six modulators bound to a single node (see Fig. 4f–i). The simulated modulator's behavior would suggest that the modulator's nucleophilicity needs to be within the same general range or of the same functional group as the linker, which matches the observation that carboxylate modulators used with carboxylate linkers for MOF assembly have been shown to be effective in size and crystallinity control.[57]

Insight into the collective behavior of the simulated modulator is gleaned by comparing the RDFs of the final simulation frames with varying modulator size, concentration (0 to 10 equivalents), and binding strength ($\varepsilon_{n\text{–}m}$ = 0.5, 1.0, 2.0, and 4.0), as shown in Fig. 5 at a constant linker-to-node binding strength, $\varepsilon_{n\text{–}l} = 1.0$. The first column, at the lowest modulator binding strength, shows that exclusively varying the modulator size does not prevent MOF self-assembly when the modulator binds to the node more weakly than the linker. At low $\varepsilon_{n\text{–}m}$, the continuous horizontal lines, corresponding to RDF peaks, extend to the greatest values of r. The weak binding of the modulator to the surface of the final MOF crystal is presented by only faint RDF peaks at very low value of r (corresponding to the small size of the modulators), with increasing equivalents of modulator. Parity between the binding strengths between node and modulator *vs.* node and linker, *i.e.*, $\varepsilon_{n\text{–}m} = \varepsilon_{n\text{–}l} = 1.0$, demonstrates the size-dependent effects of the modulator on the final crystalline product: the smallest modulators hinder crystal growth and act as a capping agent when using 2 equivalents or greater, as evident by the decreasing intensity of the horizontal lines with greater modulator equivalents with the exception of the first observed RDF peak. The RDF

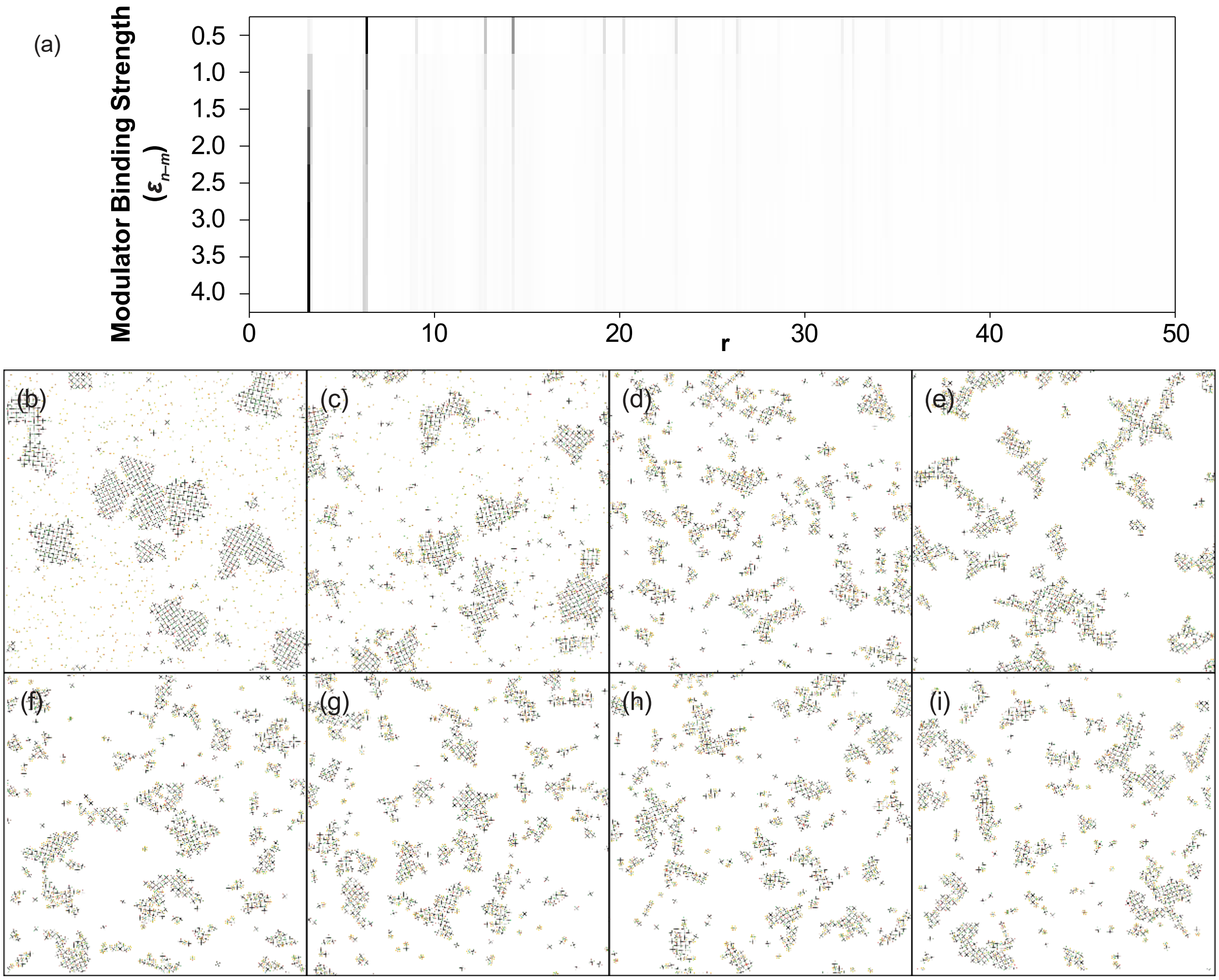


Figure 4: The effect of node-to-modulator binding strength $\varepsilon_{n\text{-}m}$ at a constant ratio of equivalents ($n : l : m = 2 : 1 : 4$) and constant arm length (of 0 beads): (**a**) radial distribution functions of the self-assembled final simulation product at varying $\varepsilon_{n\text{-}m}$, and the final simulation frames with increasing node-to-modulator binding strength at $\varepsilon_{n\text{-}m}$ values of (**b**) 0.5, (**c**) 1.0, (**d**) 1.5, (**e**) 2.0, (**f**) 2.5, (**g**) 3.0, (**h**) 3.5, and (**i**) 4.0.

peak at the lowest value of r, a result of the bonds between nodes and modulators, becomes more intense with greater modulator quantities as the modulators act as capping agents and permanently bind to the nodes. As the modulator size increases at $\varepsilon_{n\text{-}m} = \varepsilon_{n\text{-}l}$, the intensity of the RDF peak corresponding to the modulator decreases as larger modulators are excluded from the pores of the final assembled MOF but are still present at the surface (as previously observed in Fig. 3). For $\varepsilon_{n\text{-}m} > \varepsilon_{n\text{-}l}$, the modulator interferes with the MOF assembly process: prominent RDF peaks that correspond to the MOF—*i.e.*, excluding the first peak representing the capping modulator—are only seen for nearest and second-nearest neighbors. The modulator capping behavior, due to $\varepsilon_{n-m} > \varepsilon_{n-l}$, reaches saturation at $\varepsilon_{n-m} \approx 2.0$: the RDFs for $\varepsilon_{n-m} = 2.0$ through 4.0 are strikingly similar.

The effect of the relative quantity of modulator is tested at a constant system size (see Fig. 5). Experimentally testing the effect of changing the node-to-modulator binding interaction is challenging due to the interdependence of variables such as steric size, pK_a, and nucleophilicity. However, this is easy to interrogate using coarse-grained modeling, as the node–linker interaction strength can be kept constant while the node–modulator interaction strength is varied. As previously seen (in Fig. 4), for only 4 equivalents of modulator relative to linker the MOF self-assembles without issue and the modulator is largely excluded from the lattice when the modulator binds more weakly to the node than the linker does. In Fig. 5, this effect is shown for a range of modulator concentrations. As the interaction strengths between node and modulator *vs.* node and linker reach parity, modulator incorporation into the lattice is observed (see Figs. 3 and 4c, and in the SI, Fig. S6). Further increasing the node-to-modulator binding strength ε_{n-m} at constant modulator concentration and size leads to a decrease in crystal domain size (see Fig. 4d–i). As the modulator binds more strongly (and thus irreversibly) to the node, MOF self-assembly is greatly suppressed, leading to poorly ordered and smaller crystallites. As such, this finding suggests that a close match in the interaction strength between the node and linker *vs.* modulator is desirable to achieve defect engineering without completely suppressing MOF formation. Lastly, the modulator dilutes the solution, thus obstructing potential bonds between nodes and linkers, and the presence of modulator prevents further growth of the MOF crystal once node and linker bond and an initial MOF grain nucleates. At a constant reaction volume (here, area), increased modulator concentration reduces the probability of MOF formation and growth. This is consistent with experimental findings that excess modulators can completely suppress MOF formation.[55]

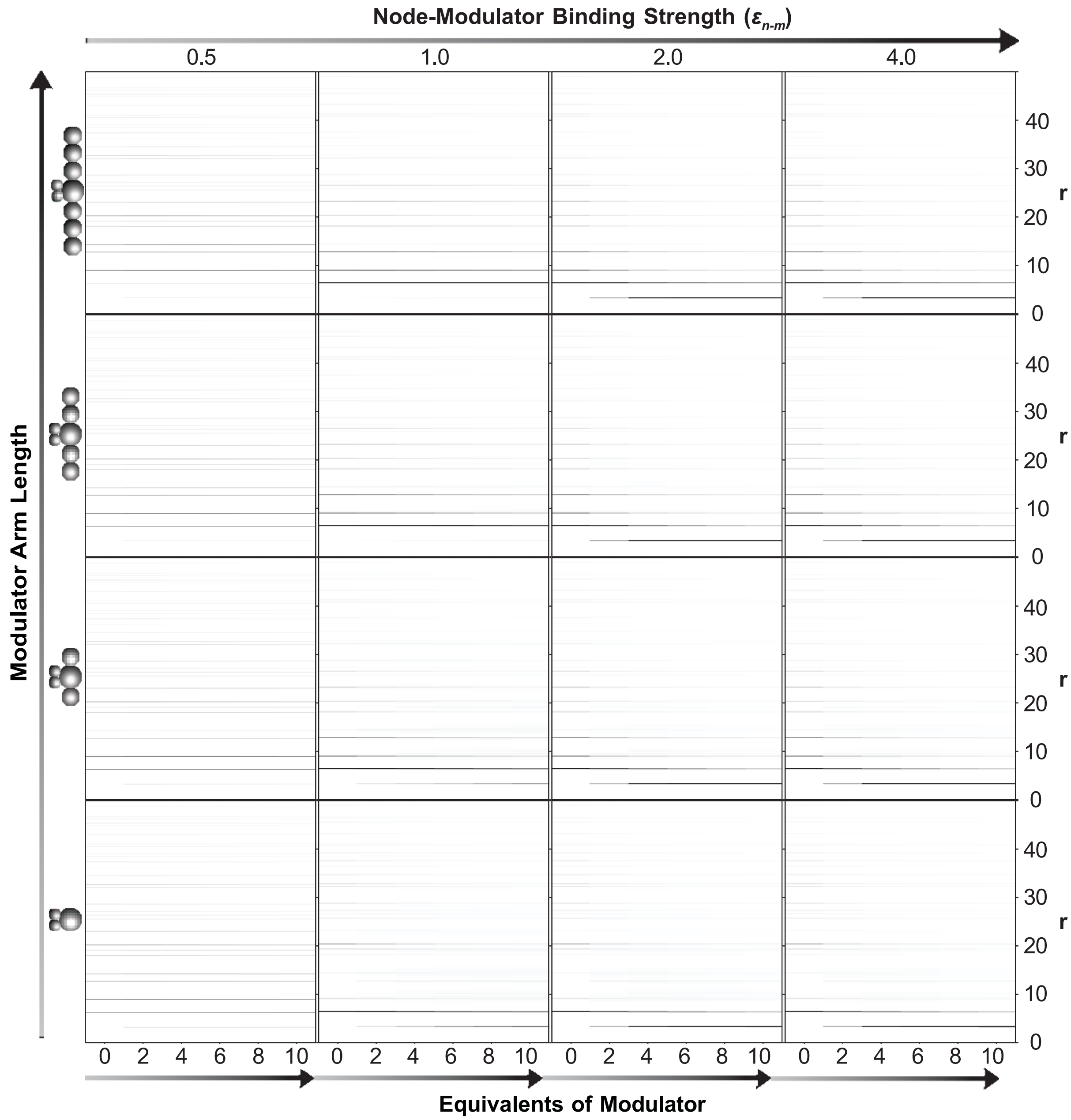


Figure 5: Radial distribution functions of the self-assembled final simulation product showing the effect of node-to-modulator binding strength (ε_{n-m}) with increasing equivalents of modulator ($n : l : m = \lceil 1 + \frac{x}{4} \rceil : 1 : x$). The modulator arm length ranges from 0 to 3 (increasing bottom to top, across subplots), ε_{n-m} values of 0.5, 1.0, 2.0, 4.0 (increasing left to right, across subplots), and the modulator concentration ranges from 0 to 10 equivalents (increasing left to right in each subplot).

Coordination numbers

Thus far we have evaluated MOF crystallinity by analyzing the RDFs of the self-assembled arrangements, but this measure does not provide immediate information about the amount of crystal defects. We therefore calculate the average coordination number (CN) for all nodes in the final assembled MOF in order to quantify the degree to which defects are present. By excluding the modulator from this calculation, we use the CN value as a metric to approximate how many linkers are correctly bound to nodes in the simulated MOF. In a perfect, infinite 2D MOF single crystal, each node is surrounded by 4 linkers, *i.e.*, CN = 4. As surface particles are under-coordinated, measuring the CN can be used as a proxy for crystal size, *i.e.*, a crystal with CN $\approx$ 4 is large. Reduction of the CN is not only a result of crystal surface termination but also of internal node and linker vacancy defects.

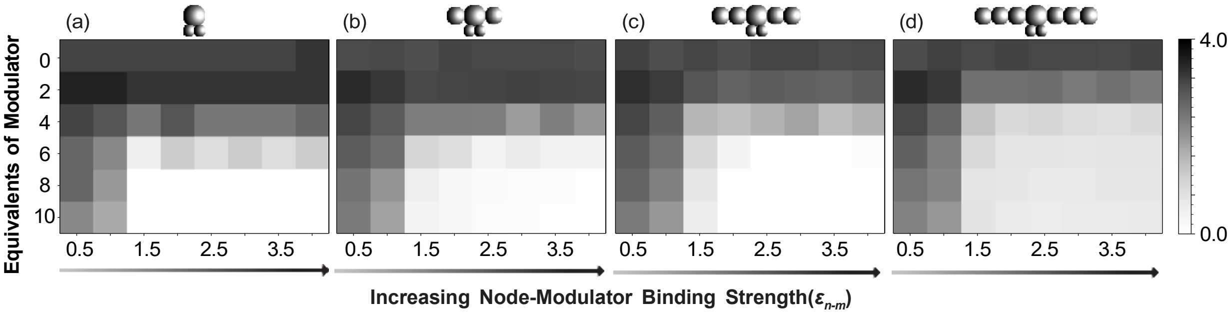


Figure 6: Average coordination number of the assembled MOFs. The average number of nearest neighbors across each system at the final assembled state is plotted for different equivalents of modulator (increasing top to bottom), node-to-modulator binding strength ε_{n-m} (increasing left to right in each subplot), and modulator size (increasing left to right, across subplots).

The CNs across the simulated modulator arm lengths, binding strengths, and numbers of equivalents are shown in Fig. 6 and are listed in detail in the Supporting Information (in Tab. S1). The CNs for the MOFs at constant $n : l : m = 2 : 1 : 4$ and $\varepsilon_{n-m} = \varepsilon_{n-l} = 1.0$ with varying modulator arm length (see Fig. 3) are constant at approximately 2.8, corroborating the above-discussed observation that the crystal size remains constant within these simulation parameters. The simulations with $\varepsilon_{n-m} \leq \varepsilon_{n-l} = 1.0$ and 2 or fewer modulator equivalents ($m \leq 2$ in $n : l : m$) display the same general trend in CN for all modulator sizes. A decrease in the average CN occurs, moving from left to right with increasing ε_{n-m}, in each subplot in Fig. 6, showing very low CNs with strongly binding modulators at high concentrations. Increasing the modulator concentration—from top to bottom in the subplots in Fig. 6—does indeed lead to an overall decrease in CN, regardless of modulator binding strength or size. The lowest CNs, and therefore smallest numbers of correctly bonded nodes and linkers, are thus generally observed when the modulator is small, strongly binding, and present at a high concentration relative to the MOF building blocks. In this case, the CN approaches 0, indicating that MOF self-assembly is completely disrupted. The decrease in CN becomes more gradual as the modulator increases in size. As the modulator size increases and binds more strongly to the node than the linker ($\varepsilon_{n-m} > \varepsilon_{n-l}$), the modulator's ability to impede crystal growth declines and what is a perceivable white region for modulators without arms becomes a visibly gray region for the largest modulator. Thus, smaller modulators with stronger binding to the node are better capping agents to control and restrict crystal growth.

The plotted heat maps of the average coordination in Fig. 6 illustrate both the subtle and drastic effects of various modulator properties. The greatest effect on the MOF crystalline domain size comes from the modulators' binding strength, while modulator concentration and sterics only have a significant effect when coupled with strong binding. At low binding strengths, the CN increases upon incorporation of a small amount of modulator (going from 0 to 2 equivalents), which we attribute to an "annealing" effect that adds reversibility to node-to-linker bond formation. Together, these findings support that there is a delicate balance between modulator size, concentration, and binding strength to the node that allows us to achieve defect engineering without completely suppressing MOF self-assembly.

Summary and Conclusions

A coarse-grained model for MOF synthesis is possible with careful consideration of the molecular components and their symmetries. Incorporating specific interactions, *i.e.*, using a finer-grained model that represents the anisotropy and handedness of the molecular orbitals responsible for bonding between node and linker, allows for the development of increase MOF grain size and definition.

Systematic testing of the various modulator properties (binding strength, size, and concentration) as independent simulation variables reveals a general hierarchy in which these factors influence MOF self-assembly: a modulator's binding strength has the greatest effect on the MOF crystal size. If the node-to-modulator bond exceeds the strength of the node-to-linker bond, the MOF crystal size decreases at modulator concentrations as low as 4 equivalents across varying modulator sizes. Currently, the simulated modulator only captures the molecule's function as a capping agent and not as an acid (*i.e.*, H^+ source) that protonates the linker functional groups. Others have observed a seesaw mechanism for modulation where there is both a reduction of MOF crystallite size with initial modulator addition and then an increase in crystallite size with greater modulator quantity.[57]

Ranked behind the effect of the modulator binding strength is the steric effect of modulators with different sizes. The modulator's steric effect is most prominent in the system when the node-to-modulator bond is stronger than the node-to-linker bond. Strongly binding modulators act as capping agents and thus inhibit crystal growth. Strongly bound modulators with greater steric hindrance seem to prevent further modulator binding based on a greater minimum coordination number for the largest modulator at the highest concentration. As the modulator's dimensions are comparable to the MOF's pore size, the modulator becomes excluded from the crystal pores and acts as a surface capping agent.

Finally, the effect of the modulator concentration depends on their binding strength and steric hindrance. Weakly binding modulators—with node-to-modulator bonds weaker than the node-to-linker bond—do not significantly impact the MOF crystal size, even at high modulator concentration and irrespective of the modulator size.

The process to develop the 2D MOF self-assembly model can be applied to the vast space of MOF compounds. In the future, the collapse of MOF structures as a result of missing-linker defects caused by synthesis conditions including modulation[23] can be investigated within a similar coarse-grained setting. Beyond sampling the crystallization process for effects of the modulator properties studied here, the effect of functional groups bonded to linkers—*i.e.*, the sterics of the linkers themselves—can be investigated, as the activity of MOFs like UiO-66 can be significantly increased by using electron-withdrawing groups on the linkers.[49]

Acknowledgement

We thank Rachael S. Skye and Andreia L. Fenley for helpful discussions. This material is based upon work supported by the National Science Foundation MPS-Ascend Postdoctoral Research Fellowship under Grant No. DMR-2139237. Acknowledgement is made to the donors of the American Chemical Society Petroleum Research Fund for support of this research. P.J.M. acknowledges support from the National Science Foundation (CBET-2047627).

Supporting Information Available

- Additional details on coarse-grained model development, including radial distribution functions for varying levels of model coarseness, as well as node-to-modulator binding strength, modulator size, and modulator concentration
- Additional details on the crystallinity of simulated self-assemblies, including coordination numbers for systems with varying node-to-modulator binding strength, modulator size, and modulator concentration, and simulation snapshots at varying modulator concentration

References

(1) Winarta, J.; Shan, B.; Mcintyre, S. M.; Ye, L.; Wang, C.; Liu, J.; Mu, B. A Decade of UiO-66 Research: A Historic Review of Dynamic Structure, Synthesis Mechanisms, and Characterization Techniques of an Archetypal Metal–Organic Framework. *Crystal Growth & Design* **2020**, *20*, 1347–1362.

(2) Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M. Systematic Design of Pore Size and Functionality in Isoreticular MOFs and Their Application in Methane Storage. *Science* **2002**, *295*, 469–472.

(3) Ockwig, N. W.; Delgado-Friedrichs, O.; O'Keeffe, M.; Yaghi, O. M. Reticular Chemistry: Occurrence and Taxonomy of Nets and Grammar for the Design of Frameworks. *Accounts of Chemical Research* **2005**, *38*, 176–182.

(4) O'Keeffe, M.; Yaghi, O. M. Deconstructing the Crystal Structures of Metal–Organic Frameworks and Related Materials into Their Underlying Nets. *Chemical Reviews* **2012**, *112*, 675–702.

(5) Yuan, S. et al. Stable Metal–Organic Frameworks: Design, Synthesis, and Applications. *Advanced Materials* **2018**, *30*, 1704303.

(6) Seo, J. S.; Whang, D.; Lee, H.; Jun, S. I.; Oh, J.; Jeon, Y. J.; Kim, K. A homochiral metal–organic porous material for enantioselective separation and catalysis. *Nature* **2000**, *404*, 982–986.

(7) Noro, S.-i.; Kitagawa, S.; Kondo, M.; Seki, K. A New, Methane Adsorbent, Porous Coordination Polymer [CuSiF6(4,4'-bipyridine)2n]. *Angewandte Chemie International Edition* **2000**, *39*, 2081–2084.

(8) Banerjee, R.; Furukawa, H.; Britt, D.; Knobler, C.; O'Keeffe, M.; Yaghi, O. M. Control of Pore Size and Functionality in Isoreticular Zeolitic Imidazolate Frameworks and

their Carbon Dioxide Selective Capture Properties. *Journal of the American Chemical Society* **2009**, *131*, 3875–3877.

(9) Corma, A. From Microporous to Mesoporous Molecular Sieve Materials and Their Use in Catalysis. *Chemical Reviews* **1997**, *97*, 2373–2420.

(10) Vermoortele, F.; Vandichel, M.; Van de Voorde, B.; Ameloot, R.; Waroquier, M.; Van Speybroeck, V.; De Vos, D. E. Electronic Effects of Linker Substitution on Lewis Acid Catalysis with Metal–Organic Frameworks. *Angewandte Chemie International Edition* **2012**, *51*, 4887–4890.

(11) Vermoortele, F.; Ameloot, R.; Alaerts, L.; Matthessen, R.; Carlier, B.; Fernandez, E. V. R.; Gascon, J.; Kapteijn, F.; De Vos, D. E. Tuning the catalytic performance of metal–organic frameworks in fine chemistry by active site engineering. *J. Mater. Chem.* **2012**, *22*, 10313–10321.

(12) Whitesides, G. M.; Boncheva, M. Beyond molecules: Self-assembly of mesoscopic and macroscopic components. *Proceedings of the National Academy of Sciences* **2002**, *99*, 4769–4774.

(13) Wu, H.; Chua, Y. S.; Krungleviciute, V.; Tyagi, M.; Chen, P.; Yildirim, T.; Zhou, W. Unusual and Highly Tunable Missing-Linker Defects in Zirconium Metal–Organic Framework UiO-66 and Their Important Effects on Gas Adsorption. *Journal of the American Chemical Society* **2013**, *135*, 10525–10532.

(14) Cliffe, M. J.; Hill, J. A.; Murray, C. A.; Coudert, F.-X.; Goodwin, A. L. Defect-dependent colossal negative thermal expansion in UiO-66(Hf) metal–organic framework. *Phys. Chem. Chem. Phys.* **2015**, *17*, 11586–11592.

(15) Dissegna, S.; Epp, K.; Heinz, W. R.; Kieslich, G.; Fischer, R. A. Defective Metal-Organic Frameworks. *Advanced Materials* **2018**, *30*, 1704501.

(16) Fang, Z.; Bueken, B.; De Vos, D. E.; Fischer, R. A. Defect-Engineered Metal–Organic Frameworks. *Angewandte Chemie International Edition* **2015**, *54*, 7234–7254.

(17) Sholl, D. S.; Lively, R. P. Defects in Metal–Organic Frameworks: Challenge or Opportunity? *The Journal of Physical Chemistry Letters* **2015**, *6*, 3437–3444.

(18) Huang, Y.; Jiao, Y.; Chen, T.; Gong, Y.; Wang, S.; Liu, Y.; Sholl, D. S.; Walton, K. S. Tuning the Wettability of Metal–Organic Frameworks via Defect Engineering for Efficient Oil/Water Separation, url = https://doi.org/10.1021/acsami.0c08803. *ACS Applied Materials & Interfaces* **2020**, *12*, 34413–34422.

(19) Vermoortele, F.; Bueken, B.; Le Bars, G.; Van de Voorde, B.; Vandichel, M.; Houthoofd, K.; Vimont, A.; Daturi, M.; Waroquier, M.; Van Speybroeck, V.; Kirschhock, C.; De Vos, D. E. Synthesis Modulation as a Tool To Increase the Catalytic Activity of Metal–Organic Frameworks: The Unique Case of UiO-66(Zr). *Journal of the American Chemical Society* **2013**, *135*, 11465–11468.

(20) Simonov, A.; Goodwin, A. L. Designing disorder into crystalline materials. *Nature Reviews Chemistry* **2020**, *4*, 657–673.

(21) Forgan, R. S. Modulated self-assembly of metal–organic frameworks. *Chem. Sci.* **2020**, *11*, 4546–4562.

(22) Cai, G.; Jiang, H.-L. A Modulator-Induced Defect-Formation Strategy to Hierarchically Porous Metal–Organic Frameworks with High Stability. *Angewandte Chemie International Edition* **2017**, *56*, 563–567.

(23) Bueken, B.; Van Velthoven, N.; Krajnc, A.; Smolders, S.; Taulelle, F.; Mellot-Draznieks, C.; Mali, G.; Bennett, T. D.; De Vos, D. Tackling the Defect Conundrum in UiO-66: A Mixed-Linker Approach to Engineering Missing Linker Defects. *Chemistry of Materials* **2017**, *29*, 10478–10486.

(24) Nguyen, V.; Grünwald, M. Microscopic Origins of Poor Crystallinity in the Synthesis of Covalent Organic Framework COF-5. *Journal of the American Chemical Society* **2018**, *140*, 3306–3311.

(25) Choi, E.-Y.; Wray, C. A.; Hu, C.; Choe, W. Highly tunable metal–organic frameworks with open metal centers. *CrystEngComm* **2009**, *11*, 553–555.

(26) Feng, L.; Wang, K.-Y.; Powell, J.; Zhou, H.-C. Controllable Synthesis of Metal-Organic Frameworks and Their Hierarchical Assemblies. *Matter* **2019**, *1*, 801–824.

(27) Liu, B.; Vellingiri, K.; Jo, S.-H.; Kumar, P.; Ok, Y. S.; Kim, K.-H. Recent advances in controlled modification of the size and morphology of metal-organic frameworks. *Nano Research* **2018**, *11*, 4441–4467.

(28) Schaate, A.; Roy, P.; Godt, A.; Lippke, J.; Waltz, F.; Wiebcke, M.; Behrens, P. Modulated Synthesis of Zr-Based Metal–Organic Frameworks: From Nano to Single Crystals. *Chemistry – A European Journal* **2011**, *17*, 6643–6651.

(29) Zhu, L.; Zhang, Y.-B. Crystallization of Covalent Organic Frameworks for Gas Storage Applications. *Molecules* **2017**, *22*, 1149.

(30) Zou, H.; Liu, L.; Zhang, S.; Miao, X.; Ying, L.; Deng, W.; Cao, Y. Different Stepwise Growth Mechanism of AIE-Active Tetraphenylethylene-Functionalized Metal–Organic Frameworks on Au(111) and Cu(111) Surfaces. *The Journal of Physical Chemistry Letters* **2023**, *14*, 489–498.

(31) Anderson, J. A.; Glaser, J.; Glotzer, S. C. HOOMD-blue: A Python package for high-performance molecular dynamics and hard particle Monte Carlo simulations. *Computational Materials Science* **2020**, *173*, 109363.

(32) Adorf, C. S.; Dodd, P. M.; Ramasubramani, V.; Glotzer, S. C. Simple data and workflow

management with the signac framework. *Computational Materials Science* **2018**, *146*, 220–229.

(33) Ramasubramani, V.; Adorf, C. S.; Dodd, P. M.; Dice, B. D.; Glotzer, S. C. signac: A Python framework for data and workflow management . Proceedings of the 17th Python in Science Conference. 2018; pp 152–159.

(34) Ramasubramani, V.; Dice, B. D.; Harper, E. S.; Spellings, M. P.; Anderson, J. A.; Glotzer, S. C. freud: A Software Suite for High Throughput Analysis of Particle Simulation Data. *Computer Physics Communications* **2020**, *254*, 107275.

(35) Dice, B. D.; Ramasubramani, V.; Harper, E. S.; Spellings, M. P.; Anderson, J. A.; Glotzer, S. C. Analyzing Particle Systems for Machine Learning and Data Visualization with freud. Proceedings of the 18th Python in Science Conference. 2019; pp 27–33.

(36) Stukowski, A. OVITO. https://www.ovito.org, accessed 2023-06-19.

(37) Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. *Modelling and Simulation in Materials Science and Engineering* **2009**, *18*, 015012.

(38) Eddaoudi, M.; Moler, D. B.; Li, H.; Chen, B.; Reineke, T. M.; O'Keeffe, M.; Yaghi, O. M. Modular Chemistry: Secondary Building Units as a Basis for the Design of Highly Porous and Robust Metal—Organic Carboxylate Frameworks. *Accounts of Chemical Research* **2001**, *34*, 319–330.

(39) 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. *Accounts of Chemical Research* **2008**, *41*, 1782–1789.

(40) Jones, J. E. On the determination of molecular fields.—I. From the variation of the

viscosity of a gas with temperature. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1924**, *106*, 441–462.

(41) Jones, J. E. On the determination of molecular fields. —II. From the equation of state of a gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1924**, *106*, 463–477.

(42) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids. *The Journal of Chemical Physics* **1971**, *54*, 5237–5247.

(43) Chandler, D.; Weeks, J. D.; Andersen, H. C. Van der Waals Picture of Liquids, Solids, and Phase Transformations. *Science* **1983**, *220*, 787–794.

(44) Wolpert, E. H.; Jelfs, K. E. Coarse-grained modelling to predict the packing of porous organic cages. *Chem. Sci.* **2022**, *13*, 13588–13599.

(45) Gorbunov, V. A.; Uliankina, A. I.; Myshlyavtsev, A. V. Off-Lattice Coarse-Grained Model of Surface-Confined Metal–Organic Architectures. *The Journal of Physical Chemistry C* **2023**, *127*, 8281–8293.

(46) Morris, W.; Wang, S.; Cho, D.; Auyeung, E.; Li, P.; Farha, O. K.; Mirkin, C. A. Role of Modulators in Controlling the Colloidal Stability and Polydispersity of the UiO-66 Metal–Organic Framework. *ACS Applied Materials & Interfaces* **2017**, *9*, 33413–33418.

(47) Chen, F. E.; Pitt, T. A.; Okong'o, D. J.; Wetherbee, L. G.; Fuentes-Rivera, J. J.; Milner, P. J. A Structure–Activity Study of Aromatic Acid Modulators for the Synthesis of Zirconium-Based Metal–Organic Frameworks. *Chemistry of Materials* **2022**, *34*, 3383–3394.

(48) Katz, M. J.; Brown, Z. J.; Colón, Y. J.; Siu, P. W.; Scheidt, K. A.; Snurr, R. Q.;

Hupp, J. T.; Farha, O. K. A facile synthesis of UiO-66, UiO-67 and their derivatives. *Chem. Commun.* **2013**, *49*, 9449–9451.

(49) Shearer, G. C.; Chavan, S.; Bordiga, S.; Svelle, S.; Olsbye, U.; Lillerud, K. P. Defect Engineering: Tuning the Porosity and Composition of the Metal–Organic Framework UiO-66 via Modulated Synthesis. *Chemistry of Materials* **2016**, *28*, 3749–3761.

(50) Atzori, C.; Shearer, G. C.; Maschio, L.; Civalleri, B.; Bonino, F.; Lamberti, C.; Svelle, S.; Lillerud, K. P.; Bordiga, S. Effect of Benzoic Acid as a Modulator in the Structure of UiO-66: An Experimental and Computational Study. *The Journal of Physical Chemistry C* **2017**, *121*, 9312–9324.

(51) Taddei, M. When defects turn into virtues: The curious case of zirconium-based metal-organic frameworks. *Coordination Chemistry Reviews* **2017**, *343*, 1–24.

(52) Vandichel, M.; Hajek, J.; Vermoortele, F.; Waroquier, M.; De Vos, D. E.; Van Speybroeck, V. Active site engineering in UiO-66 type metal–organic frameworks by intentional creation of defects: a theoretical rationalization. *CrystEngComm* **2015**, *17*, 395–406.

(53) 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. *The Journal of Physical Chemistry C* **2016**, *120*, 9276–9281.

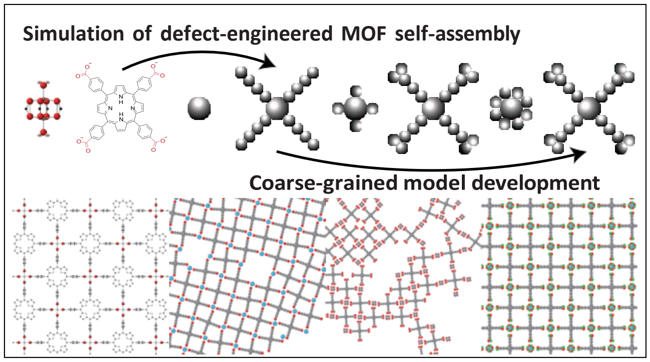
(54) Yuan, L.; Tian, M.; Lan, J.; Cao, X.; Wang, X.; Chai, Z.; Gibson, J. K.; Shi, W. Defect engineering in metal–organic frameworks: a new strategy to develop applicable actinide sorbents. *Chem. Commun.* **2018**, *54*, 370–373.

(55) Ren, J.; Langmi, H. W.; North, B. C.; Mathe, M.; Bessarabov, D. Modulated synthesis of zirconium-metal organic framework (Zr-MOF) for hydrogen storage applications. *International Journal of Hydrogen Energy* **2014**, *39*, 890–895.

(56) Jiang, D.; Burrows, A. D.; Edler, K. J. Size-controlled synthesis of MIL-101(Cr) nanoparticles with enhanced selectivity for CO_2 over N_2. *CrystEngComm* **2011**, *13*, 6916–6919.

(57) Marshall, C. R.; Timmel, E. E.; Staudhammer, S. A.; Brozek, C. K. Experimental evidence for a general model of modulated MOF nanoparticle growth. *Chem. Sci.* **2020**, *11*, 11539–11547.

TOC Graphic


Simulation of defect-engineered MOF self-assembly
Coarse-grained model development