Molecular level understanding of the free energy landscape in early stages of MOF nucleation.

Loukas Kollias[1], David C. Cantu[2,3], Marcus A. Tubbs[2], Roger Rousseau[2], Vassiliki-Alexandra Glezakou[2]*, Matteo Salvalaglio[1]*

[1]Thomas Young Centre and Department of Chemical Engineering, University College London, London WC1E 7JE, UK; [2]Basic and Applied Molecular Foundations, Pacific Northwest National Laboratory, Richland WA, 99352 USA; 3Chemical and Materials Engineering Department, University of Nevada, Reno, Reno NV, 89557 USA

ABSTRACT: The assembly mechanism of hierarchical materials controlled by the choice of solvent and presence of spectator ions. In this paper, we use enhanced sampling molecular dynamics methods to investigate these effects on the configurational landscape of metal-linker interactions in the early stages of synthesis, using MIL-101(Cr) as a prototypical example. Microsecond-long well-tempered metadynamics (WTmetaD) simulations uncover a complex free energy structural landscape, with distinct crystal (C) and non-crystal (NC) like configurations and their equilibrium population. In presence of ions (Na+, F-), we observe a complex effect on the crystallinity of secondary building units (SBUs), by encouraging/suppressing salt bridges between C configurations and consequently controlling the percentage of defects. Solvent effects are assessed by introducing N, N dimethylformamide (DMF) instead of water, where SBU adducts are appreciably more stable and compact. These results shed light on how solvent and ionic strength impact the free energy of assembly phenomena that ultimately control material synthesis

INTRODUCTION

Tunable materials offer numerous possibilities for synthesis leading to desirable attributes like high porosity, surface area and separations capability. Porous materials of this kind include metal-organic frameworks (MOFs) that attract increasing interest due to their potential applicability in numerous industrial processes that include CO_2 capture and sequestration[1-3] to reduce environmental pollution, CH_4 storage[4-5] for natural gas applications, catalysis[6-7], drug delivery[8-9], separations[10-11] and water production from the air in dry climates[12]. MOF properties can be significantly changed upon the presence of defects[13-14]; for example, defects are known to compromise mechanical stability[15-16], or they can offer novel ways of functionalization, without loss of stability[17]. Understanding the formation of defects during MOF synthesis is essential[18] to devising strategies that regulate their occurrence.

In recent years, MOF synthesis research has focused on material discovery by employing various experimental techniques to develop new MOFs. Nevertheless, a thorough investigation involves understanding the mechanism of synthesis that takes place. In this respect, theoretical models should be validated with experimental findings. Molecular simulations can intrinsically access atomic level information and provide mechanistic details that are not always possible to identify from experimental studies. In this work, we provide insights into MOF synthesis by assessing interactions between its building blocks during the early stages of nucleation through molecular simulations. Furthermore, we investigate solvent and ionic effects on these interactions in order to elucidate their role in MOF assembly.

Due to the metal cations in MOFs, their synthesis is regularly carried out in presence of counterions. Therefore, studying the effect that ionic strength has in this process is meaningful in the modelling of MOF synthesis. More precisely, metal precursors are introduced in the form of metal nitrate hydrates (e.g. $Cr(NO_3)_3 \cdot 9H_2O$ in the case of chromium)[19] and fluoride has been used by experimentalists in the synthesis of MIL-101 (Cr) as it promotes rapid crystal growth of the material[20-21]. Fluoride anions have an affinity to the chromium at the metal center of MOF building units and their interaction assists the stability of the crystal structure[22]. Nevertheless, recent studies[23-26] have used alternative synthetic methods by replacing fluoride with hydroxide anions using deionized water[23-25].

In this work, the self-assembly of MIL-101 (Cr) SBU from its constituent half-SBUs is studied as a prototypical example to highlight the most important stages of nucleation. Existing data on synthesis kinetics[27] allows for validation of our models. The half-SBUs used in this research work are structural isomers of the MIL-101 (Cr) half structural building unit (half-SBU).

MIL-101 (Cr) half-SBUs are comprised of the tri-metal center with three coordinatively bound organic linkers, and three "open" sites available to interact with other half-SBUs during nucleation. These have been found to be relevant in the study by Cantu et al.[28], where they emerge from minimum energy pathways in the formation of a MIL-101 (Cr) SBU. Furthermore, in situ spectroscopy and NMR studies provide concrete evidence that supports the hypothesis that nucleation proceeds by the formation and assembly of half-SBUs[29-31].

Overall, MOF formation is widely believed to follow a nucleation and crystal growth mechanism. Our hypothesis is that nucleation of the MOF lattice is the result of the assembly of these half-SBUs, and is therefore directly impacted by their configuration complexity. In this work we sample the configurational space of SBUs resulting from the dimerization of half-SBUs and estimate the equilibrium distribution of their conformational variants. With this analysis, we ultimately assess the propensity to develop defects in the early stages of MOF crystallization. Our analysis finds its context within a number of recent computational studies on the assembly of complex materials such as metal nanoparticles[32-33], metal-organic frameworks with single-metal centers[34-35] and "metal-free" covalent-organic frameworks[36]. Implicit solvation and united-atom force fields are often employed in these works[32-34] to qualitatively analyze assembly processes. In this work we make a leap forward in complexity, by tackling a system that presents a complex metal-center topology with all-atom, explicit solvent simulations

aimed at exploring and quantifying the free energy landscape in the early stages of MOF assembly, which has not been comprehensively understood yet[37].

To achieve this goal, several microsecond-long metadynamics simulations have been conducted, unveiling the complexity of a configurational landscape characterized by hundreds of possible SBU configurations. Furthermore, the addition of an increasing number of ions to the system allows to uncover the role of solution composition in the formation of SBUs, and the probability of assembling defective units from half-SBU assembly.

METHODS

In the following sections, we present the methodology applied to carry out molecular simulations and processing of the data to compute equilibrium distributions.

Half-SBUs For the sake of clarity, the MIL-101 (Cr) half-SBU structural isomers are referred to in this work as MLA, MLB, MLC: in metal linker arrangement A (MLA), all linkers are on the same side of the metal center plane; in arrangement B (MLB), two linkers are on one side and one on the other side of an open site; in arrangement C (MLC) two linkers are on one side and one on the other side of an occupied site. Furthermore, we should note that MLA and MLB are also stereo isomers. A representation of their chemical structure is illustrated in Fig. 1[28]. We carried out an extensive analysis of the conformational space of both isolated half-SBUs and all their possible combinations into SBUs at 298K and 1 bar.
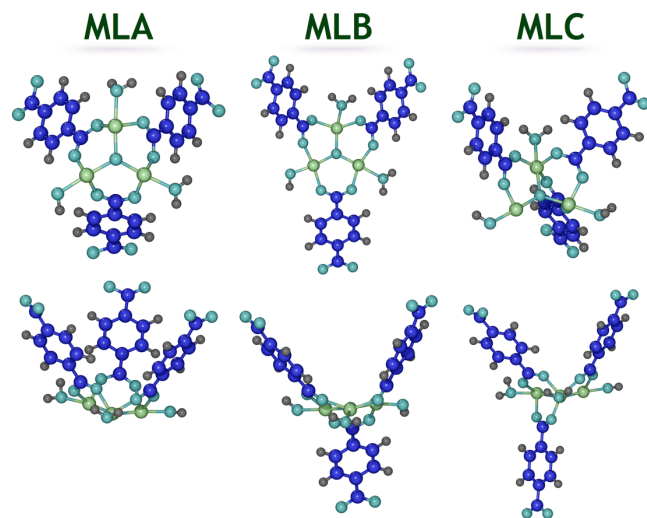

Figure 1. Structural isomers of the MIL-101 (Cr) half-SBU. Color code: Chromium - Lime, Oxygen - Cyan, Hydrogen – Grey, Carbon – Blue. Top: top view, bottom: side view. Considering as a visual reference the plane defined by the Cr atoms in the metal cluster, MLA is characterized by having all terephthalate linkers pointing away from the metal cluster. MLB and MLC instead have only two linker pointing in the same direction. MLB and MLC display a different connectivity between the metal cluster and the linkers, whereby the same two Cr atoms are simultaneously coordinating two linkers. An animated version of this figure is provided as Supplementary Animation. Chemical structures were visualized with VMD[38]

Force fields To achieve an extensive sampling of conformational isomers of both half-SBUs and SBUs in explicit solution, we resorted to a classical representation of molecular interactions. In all simulations we used the OPLS-AA force field[39] to model interactions in the terephthalate linkers. To model non-bonded interactions between Cr atoms and linkers in different half-SBUs, we tailored the pair-wise interactions between chromium atoms and linkers from a different half-SBU using a Lennard-Jones potential following the work of Cantu et al.[28]. The overall procedure is presented in the SI, section I. Regarding internal half-SBU bonded interactions, the values of the harmonic oscillator constants of bonds and angles involving Cr were set to high values to keep the internal half-SBU coordinates close to equilibrium, since the focus of the simulations is to understand how different half-SBUs assemble. Bond and angle equilibrium values are based on SBU structures optimized with density functional theory (DFT). Dihedral angles, instead, are estimated from geometry optimizations of the isolated metal center.

All simulations have been carried out in explicit solvent. We investigate both aqueous and N, N-dimethylformamide (DMF) solutions, with the explicit inclusion of counterions (Na^+, F^-). Water is simulated with the TIP3P model[40], while ions and DMF are modelled using the OPLS-AA[39] force field. For counterions we used the OPLS-AA[39] parameters included in GROMACS[41]. DMF force field parameters are obtained from the *virtualchemistry.org* database[42-43]. The dipole moment of this classical model of DMF is 3.52 D, in good agreement with the experimental value of 3.86 D[44-45].

Conformational sampling with well-tempered metadynamics. Well-tempered metadynamics (WTmetaD)[46] was used to extensively characterize the conformational space of both half-SBUs and their dimers (SBUs). In all simulations, temperature and pressure conditions are maintained constant (298 K and 1 bar) by using the velocity rescale thermostat[47] and the Berendsen barostat[48], respectively. A 10 Å cut-off was used for non-bonded interactions. Long range electrostatic interactions were treated using a particle-mesh Ewald scheme[49]. Dynamics were propagated with a leapfrog integrator using a time step of 2 fs, while preserving bond lengths using LINCS[50]. A cubic simulation box with 6500 atoms and an edge of 40 Å for isolated half-SBUs and 12000 atoms and 50 Å edge for couples forming an SBU is used in all cases, respectively, while imposing three-dimensional periodic boundary conditions. All simulations were carried out using GROMACS 5.1[41] with PLUMED 2.2[51].

In the case of isolated half-SBUs in solution, we investigate the structural flexibility of all three isomers as conformational isomerism is found to regulate the flexibility of the resulting framework[52]. In this context, we sample their conformational space performing WTmetaD[46] simulations using the mass-weighted radius of gyration (R_{gyr}) as a collective variable (CV).

The assembly of half-SBUs into SBUs was examined via enhanced biasing of two CVs, defined as the smallest among all distances between any two chromium atoms (labelled M-M) and the smallest among all distances between any terminal carboxylic carbon in a terephthalate linker and any metal center (labelled M-T). These CVs are invariant with respect the specific identity of the chromium or carbon atoms, and are schematically depicted in the SI, Fig. S2.

All simulations were performed updating the WTmetaD[46] bias every ps, with an initial Gaussian height of 0.598 kcal/mol and width of 0.1 Å. The isolated half-SBUs simulations were carried out with a bias factor of 10. In the assembly of half-SBUs into SBUs instead, the bias factor is chosen as follows: 30 for MLA-MLA (AA), 20 for MLA-MLB (AB), MLA-MLC (AC), MLB-MLB (BB) and 10 for MLB-MLC (BC), MLC-

MLC (CC). The choice of different bias factors was dictated by differences in the characteristic energy barriers associated to each case. Nevertheless, the final free energy estimates are independent of this choice. The same WTmetaD[46] setup was implemented in all aqueous simulations, both the absence and presence of ions. For simulations in DMF, the bias factors used were: 50 for AA, 40 for AC, 30 for AB, BB, BC and CC.

Each one of the 6 combinations of half-SBUs, was simulated in the absence of ions for an overall simulation time of 7 μs. To investigate the effect of the ionic concentration in water all 6 half-SBUs combinations were studied at 6 different ionic concentrations, for a total of 32 independent simulations, extended to a simulation time of 1 μs each. In DMF, the simulation time is 4 μs for AA, 2.5 μs for AC and 2 μs for the remaining four couples of half-SBUs.

Free energy calculations allow us to study the formation of crystal-like and non-crystal-like SBUs from detached half-SBUs. All possible couples of configurational isomers MLA, MLB, MLC are investigated. The SBU dimers (AA, AB, AC, BB, BC, CC) are distinguished into two categories as D_C for crystal-like and D_{NC} for non-crystal-like SBUs, where D indicates one of the aforementioned dimers e.g., AA.

DFT calculations. We further examine the relative stability of the conformational isomers of half-SBU dimers by computing their relative energy at the DFT level. Calculations are conducted in Gaussian 16[53] with the hybrid, gradient-corrected Perdew, Burke and Ernzerhof (PBE0)[54-55] functional for exchange correlation. Single point energy calculations were carried out in the gas phase. We used the 6-31G**[56] basis set for carbon, hydrogen and oxygen and the "Stuttgart RSC 1997"[57] along with its corresponding effective core potential for chromium. The latter were obtained from the EMSL basis set exchange[57-58].

Equilibrium distribution of species. Free energy differences obtained from sampling the conformational landscape of SBUs allow us to compute the equilibrium distribution of SBU conformers in absence and presence of ions. The equilibrium probability of forming crystal-like or non-crystal-like SBUs from any of the six possible half-SBU dimers is estimated from free energy differences under the assumption that half-SBUs can only form SBUs without considering their further assembly into larger structures. It has been reported in the literature that MOF crystallization quickly follows after SBU formation[29]. Two types of dimerization processes are considered:

$$M1 + M2 \rightleftharpoons D_C, \qquad M1 + M2 \rightleftharpoons D_{NC} \qquad (1)$$
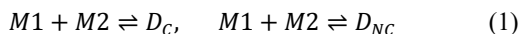
where M1 and M2 are any two half-SBUs among the MLA, MLB and MLC types, and D_C and D_{NC} are crystal-like (C) and non-crystal-like (NC) dimers (SBUs) respectively. This leads to considering 12 reactions, resulting in the formation of 12 SBUs, and 15 species (3 half-SBUs and 12 SBUs). Each species is indicated with the index i. For each individual, isolated reaction, denoted with index j, we compute the free energy of dimerization ΔG_j^0, with respect to an ideal reference state common to all reactions, in which half-SBUs are isolated in solution. Transitions between this reference state and different SBUs are sampled through WTmetaD[46], and the convergence of ΔG_j^0 values is reported in Fig. S12-S14, Supplementary Information (SI).

With this information at hand, we computed the equilibrium composition of a system in which all reactions are occurring simultaneously. The equilibrium probability for species i, P_i, is given by:

$$P_i = P_{i,0} + v_{ij} \cdot \lambda_j \qquad (2)$$

where v_{ij} is the stoichiometric coefficient of species i in reaction j, $P_{i,0}$ is the initial probability of species i, and λ_j is the conversion associated to the equilibration of reaction j. A system, in which all 12 reactions are active, is in equilibrium when the condition $e^{-\beta \Delta G_j^0} = \prod_i P_i^{v_{ij}}$ holds for every reaction.

Values of P_i that satisfy this condition can therefore be computed from Eq. 2, by numerically evaluating the equilibrium conversion λ_j by solving the following system of non-linear equations:

$$\beta \Delta G_j^0 + \sum_i v_{ij} \cdot log(P_i) = 0, \quad 1 \le j \le 12 \qquad (3)$$

where β is the inverse temperature ($[k_B T]^{-1}$, where k_B is the Boltzmann's constant). A numerical solution is computed using a trust-region approach with a dogleg step calculation[59-60].

In order to assess how the uncertainty in the estimates of ΔG_j^0 propagates to the distribution we carry out a bootstrap analysis, extracting random values for ΔG_j^0 within the confidence interval estimated for every SBU conformer (see SI).

RESULTS AND DISCUSSION

Insight into SBU assembly. WTmetaD[46] simulations exploring the conformational space of half-SBU couples were performed at this stage for all combinations of the three half-SBUs – MLA, MLB, and MLC. Our analysis leads to 48 FES and approximately 300 dimer structures in total. The resulting free energy surfaces (FES), reported in Fig. 2 for the AA SBU, exhibits 7 local minima corresponding to (meta) stable SBU configurations. Low values for both CVs correspond to crystal-like SBUs (states 1, 2C) and high to detached, non-interacting, half-SBUs (state 5). When M-M is low and M-T is high, non-crystal-like SBUs are projected to form (states 4, 6, 7). Interactions between chromium atoms of a metal center and terminal oxygen atoms of the adjacent terephthalate linker with a planar T-shape represent structures found in the crystal lattice of MIL-101 (Cr). At last, we note that crystal-like SBUs resemble the relative orientation found in the crystal structure, shown in Supporting Information (SI), Fig. S17; hence they are in accordance with experimental observations. All FES are provided in SI, Fig. S4.

Interactions between MLA half-SBUs result in the most stable SBUs. The associated FES, in the case of a pure water solution, is presented in Fig. 2. States 1-3 are characterized by to interactions between the metal center and a terephthalate linker, while states 4, 6 and 7 are dominated by interactions between metal centers. The nomenclature followed to further describe specific atomic interactions is OMetal: oxygen of the metal center, OTerminal: terminal oxygen of the terephthalate linker. Consequently, we identify the following interactions present, in state 1: [2 OTerminal – 1 Cr], state 2C: [2 OTerminal – 2 Cr], state 6: [1 OMetal – 1 Cr] and state 7: [2 OMetal – 2 Cr].

It should be noted that state 2 dominates the equilibrium distribution of species and states 6, 7 correspond to non-crystal-like SBUs with highly positive values that exceed the current scale.

In state 2, we identify a crystal-like structure (labelled 2C, corresponding to state 2 in Fig. 2), with two OTerminal – Cr inter-

actions. In the same region of the M-M, M-T CV space a quadruple oxygen – chromium interaction (labelled 2Q) with four OTerminal – Cr interactions is sampled. This configuration, which belongs to the set of non-crystal-like SBUs, can be seen in Fig. S3 – 2a, SI. In order to quantitatively assess the contribution of 2Q and 2C to the species distribution we independently compute their free energy difference through an independent WTmetaD[46] simulation (for the details see SI, section III). The latter shows that $\Delta G_{2C \to 2Q}$ is of the order of -10 kcal/mol, hence indicating that in the absence of counterions 2Q

is the most stable configuration for an SBU. We note that this is also the case in DMF, as discussed in detail in the following, while in the presence of excess ions 2Q is not sampled at all and 2C dominates state 2. In states 3 and 4, oxygen atoms of one half-SBU and hydrogen atoms of its counterpart interact as shown in the insets of Fig. 2.
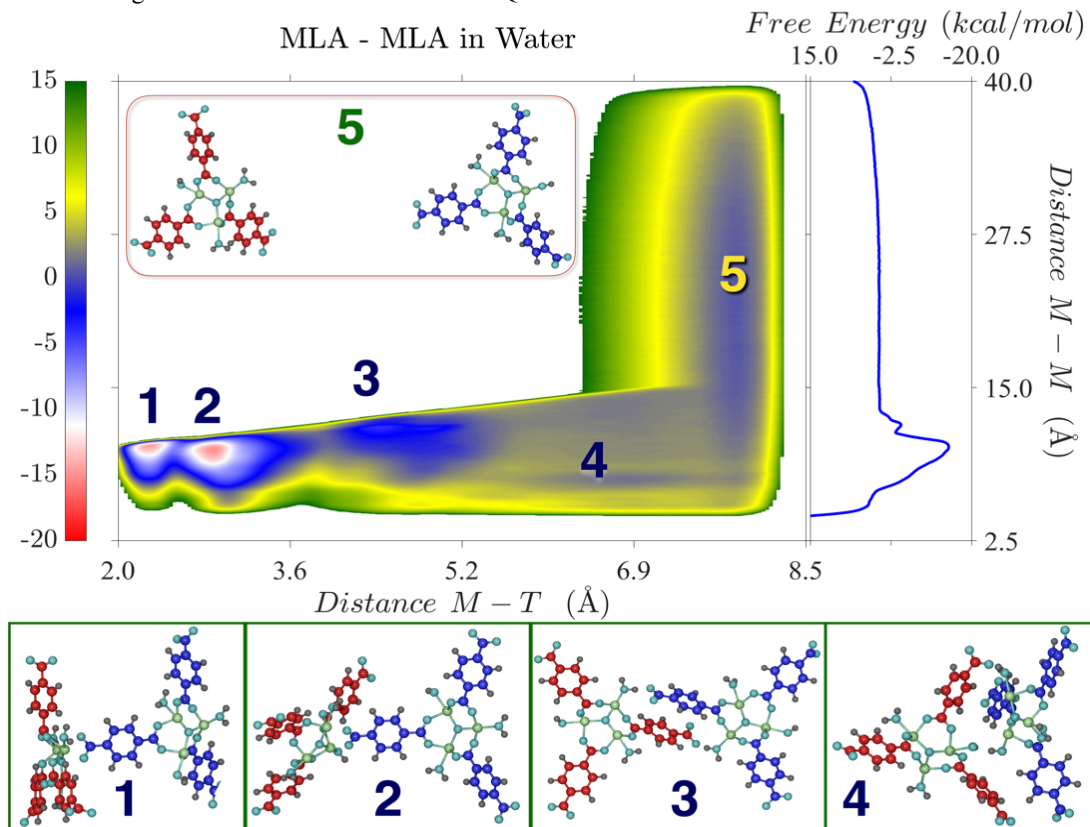

Figure 2. Free energy surface projected on the 2 CVs biased during the simulation of 2 MLA half-SBUs in an aqueous solution at 298K, 1 bar. The 4 states of interactions and their respective configurations are shown. A structure representing 2C is displayed in state 2. State 5 corresponds to detached half-SBUs. The color code for all atoms except from carbon is consistent with Fig. 1. In this representation, carbon atoms are colored blue and red in order to distinguish between different half-SBUs.

Characterization of half-SBU flexibility. We examine the flexibility of isolated half-SBUs in solution via WTmetaD[46] simulations performed using the mass – weighted radius of gyration (R_{gyr}) of each isomer in solution as a collective variable for a total simulation time of $0.2\ \mu s$. The resulting probability densities are reported in Fig. 3a. We note that MLA displays a rather rigid structure, showing a narrow distribution of R_{gyr}. MLB exhibits a wider distribution with structures characterized

by the largest R_{gyr}. MLC shows a configuration with two of the ligands almost parallel to each other at the lowest R_{gyr} of all half-SBUs. The latter observation shows that the most probable configuration for MLC diverges from the orientation found in the crystal lattice, and supports the fact that MLC, as emerged from dimerization simulation, is unlikely to produce crystal-like SBUs; especially in the absence of MLA.
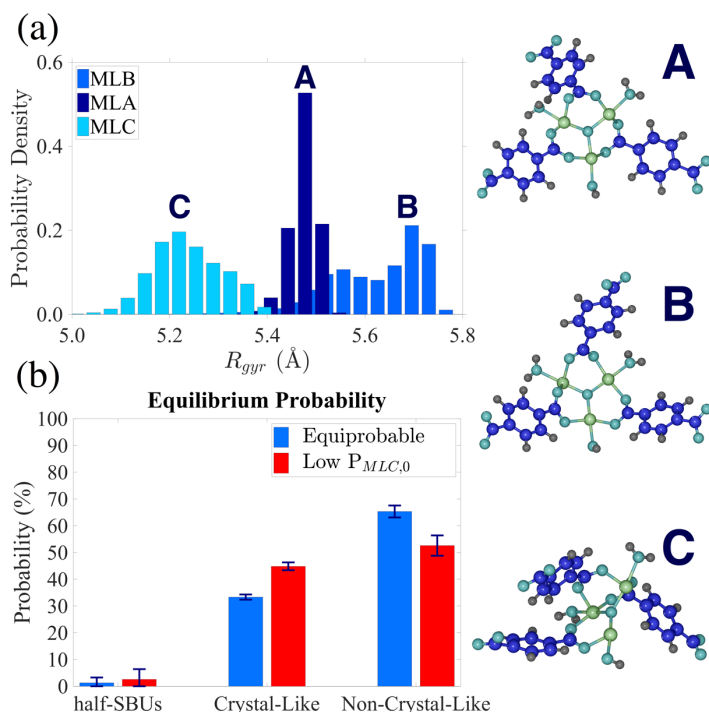
4

Figure 3. (a) Probability density of isolated half-SBUs in terms of R_{gyr}. The color code is consistent with Fig. 1. All structures refer to (a). (b) Equilibrium distribution of species in water with an ionic concentration of 0.025M. Blue: Equiprobable initial distribution of half-SBUs, red: low initial probability for MLC.

MLA-MLA SBU energies. Following the analysis of the FES conformational landscape, we have conducted single point energy calculations on the AA SBUs in the gas phase. A qualitative agreement with estimates from the classical force field is observed. Therefore, we report that the rank of the structures in terms of energy stays the same as at a higher order of theory. Nevertheless, the potential energy difference (ΔE) is larger at the higher level of theory. The reference structure for these calculations is state 2C. All estimates from DFT calculations are provided in table S1, SI.

Equilibrium probability of crystal-like vs non-crystal-like SBUs. The probabilities of different species at equilibrium were calculated to provide an estimate for the likelihood of forming crystal-like (C) vs non-crystal-like (NC) SBUs starting from an initial distribution of half-SBUs. Since NC structures dominate in absence of ions as discussed, the analysis was conducted for the aqueous simulations in presence of ions, in order to quantitatively assess the effect of the concentration of ions on the population of SBUs conformers.

We investigate two conditions for the initial distribution of half-SBUs: (1) Equiprobable, (2) $P_{MLA,0} = 45\%, P_{MLB,0} = 45\%, P_{MLC,0} = 10\%$ (Low initial probability of MLC). The reason for the latter is that the energy barrier is larger in the formation process of MLC than in MLA, MLB[28]. We shall evaluate the case of water at very low ionic strength (0.025M). An Equiprobable initial distribution results in $\sim 33\%$ probability of formation of crystal-like SBUs, which is increased to $\sim 45\%$ when the initial probability of MLC is lower than MLA and MLB. Detached half-SBUs have a probability of less than $\sim 5\%$ at equilibrium for both (1) and (2). The latter follows after an appreciable decrease in probability of non-crystal-like SBUs ($\sim$ 52%) in (2) in contrast to (1) ($\sim 65\%$). The resulting probability distributions, considering ΔG in the calculations, are shown in Fig. 3b. The equilibrium probabilities of all SBUs are available in Fig. S15-S16, SI.

Effect of solution composition. The presence of spectator ions and choice of solvent changes the FES of dimerization and impacts the probability of formation of crystal-/non-crystal-like structures.

We investigate the impact of a solvent different than water on the half-SBU association process. DMF is frequently used in MOF synthesis[30, 61-62]. Compared with those obtained in water, SBUs are much more stable and compact in DMF: the values for the ΔG (in kcal/mol) are -115 for AA, -55 for AB, AC, -35 for CC and -20 for the rest. In most cases, only one SBU configuration dominates the FES. The configurations that corresponds to this state display 4 OMetal – Cr interactions and the half-SBUs are parallel to each other (Fig. S3 – 2a) or slightly bent (Fig. S3 – 2b). The former is the most stable structure in DMF and it has been formed in water, in absence of ions, as well, however it is not as stable in the latter. Moreover, AA again dominates all SBUs from a free energy of association standpoint. Another interesting structure that is found in DMF presents the terminal oxygen atoms from 2 linkers of one half-SBU bound to all 3 Cr atoms of its counterpart (Fig. S3 – 2c). It should be noted that $\pi - \pi$ stacking of the benzene rings characterizes these structures. All of the aforementioned conformers are provided along with the corresponding FES in Fig. S3, SI.
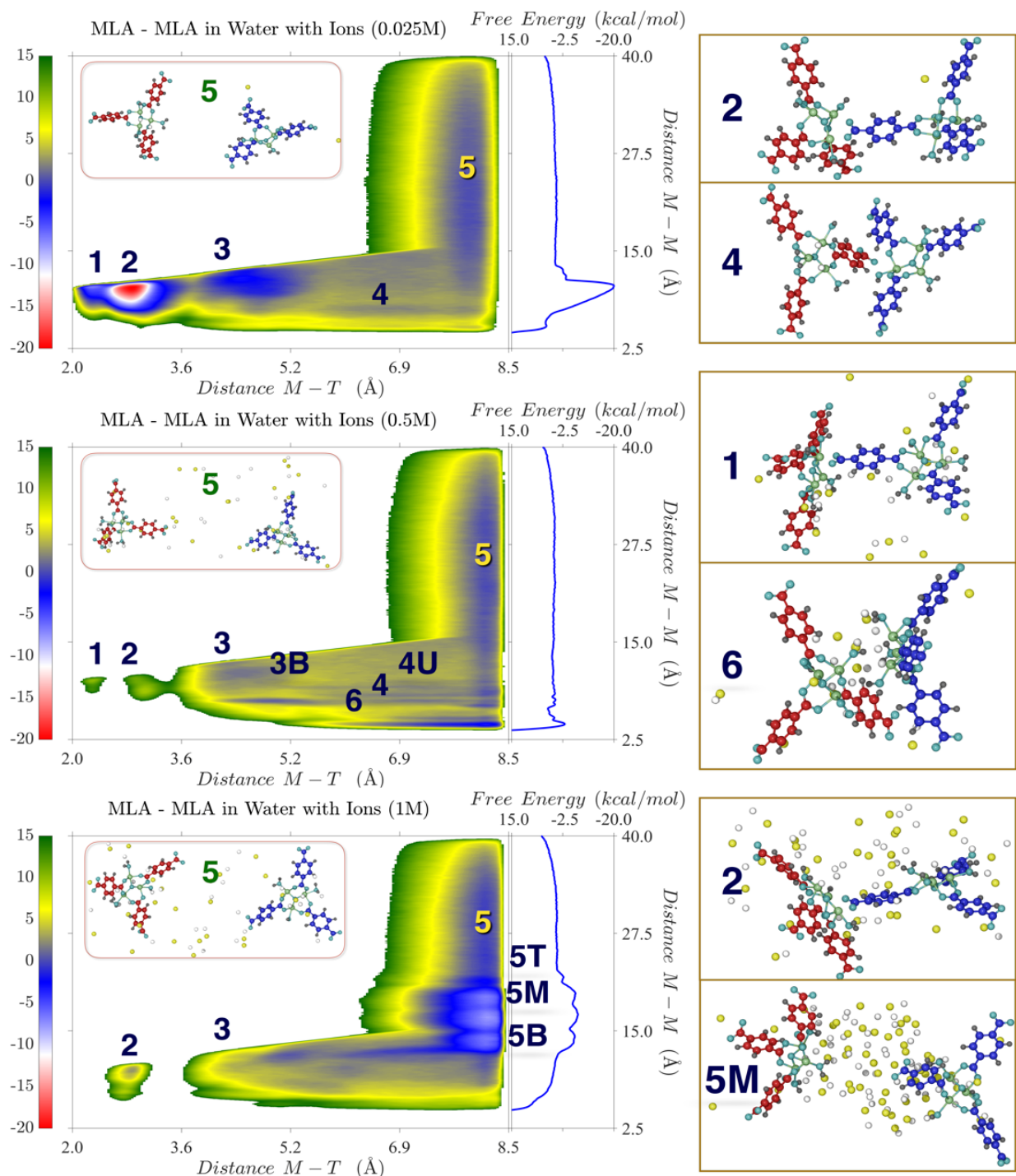
5

Figure 4. Free Energy Surfaces of MLA – MLA interactions in the presence of ions in water at 298K, 1 bar. The structure of the dominant non-crystal-like state is shown for every case (bottom configuration at the side of each FES) along with the crystal-like conformer (top configuration at the side of each FES). Reference state (5) is also shown for each concentration of ions. The color code is consistent with Fig. 2.

Impact of ions on the SBUs configurational landscape. The same approach is followed after the introduction of spectator ions to the system of half-SBUs in water, both explicitly represented. Ionic species are Na^+ and F^- at concentrations of 0.025, 0.125, 0.25, 0.5, 0.75 and 1M. In order to characterize the binding of half-SBUs in the presence of ions, we discuss AA as it is the most energetically favorable SBU. The FES for AA SBUs in water at ionic concentrations 0.025M, 0.5M and 1M along with the corresponding structures are available in Fig. 4, while the intermediate concentrations are available in Fig. S5, SI.

We note that a small concentration of ions (0.025M) results in a free energy surface comparable to that obtained in absence of ions (Fig. 4, top panel). Nevertheless, ions inhibit the spontaneous formation of the 2Q configuration, which instead dominates the pure water case. In these conditions the probability of observing the crystal-like 2C configuration is the highest as shown

6

in Fig. 5. A further increase in the ionic concentration to 0.125M renders crystal-like structures improbable because the terephthalate-chromium interactions, dominating state 2 are hindered by ions. On the contrary states 6, 7 characterized by M-M interactions are stabilized by fluoride anions and dominate the FES. Additional results are reported in the SI, Fig. S5-S10. An analogous behavior is observed for ionic concentrations of 0.25M and 0.5M, where state 6 is the most stable, while states 1, 2 and 7 are completely inhibited ($\Delta G_{5\rightarrow1,2,7} > +15$ kcal/mol, see Fig 5). Additional interactions between metal centers are observed for 0.5M, where F^- linked to the metal centers interact with Na^+ forming salt bridges between the half-SBUs (state 4U, see Fig. 4).

Beyond 0.5 M, ions have a two-fold effect. On the one hand ions decrease the total amount of SBUs formed, on the other hand however the introduction of salt bridges favors the stabilization of crystal-like SBUs. This behavior is apparent observing state 2 (Fig. 4) at 1M.

Impact of spectator ions on the equilibrium distribution of SBU conformers. Equilibrium probability distributions allow us to assess the effect of ionic strength on the relative abundance of conformational isomers of SBUs. Following the procedure outlined in the previous paragraphs, the equilibrium probability of detached half-SBUs, crystal like SBUs and non-crystal-like SBUs has been computed for all the concentrations of ions investigated. The resulting probability distributions are available in SI, Fig. S15-S16.

In Fig. 5 we observe that the probability of crystal-like SBUs presents a non-monotonous trend as a function of the ionic strength. In particular, we see that in absence of ions, non-crystal-like SBUs dominate the distribution. This is due to the fact that in these conditions even in aqueous solution, the NC configuration corresponding to state 2Q, discussed in the previous sections, is the most stable.

For small ionic concentrations (0.025M) crystal-like SBU reach a maximum in their population, contributing for more than 30% of the species in solution, as shown in detail in the histogram reported in Fig. 3b. Increasing the concentration of ions however, long lived interactions between counterions and metal clusters are observed. These interactions hinder the formation of bonds between terephthalate linkers metal clusters, and lead to a negligible probability of observing crystal-like SBU configurations, such as those corresponding to states 1 and 2 in the FES reported in Fig. 2 and 4. In these intermediate conditions non-crystal-like SBUs, characterized by specific interactions (states 4, 6 and 7) are favored. Nevertheless, when the ionic concentration exceeds 0.5 M, salt bridges are formed between half-SBUs. Such ion-mediated interactions hinder rotation of the linkers and promote crystal-like SBU configurations.

These findings suggest a complex role played by ions during MOF synthesis:

i. In absence of ions, non-crystal-like SBUs dominate the early stages of MOF nucleation, hence potentially leading to the formation of highly defective material. This provides a mechanistic rationale for the need of spectator ions in MOF synthesis.

ii. At low ionic concentrations (0.025M) crystal-like SBU configurations represent a large proportion (30-40% see Fig. 3b and Fig. 5). In these conditions these SBUs are likely to provide the building blocks for the nucleation of crystalline MOFs.

iii. At slightly higher concentrations, 0.2 - 0.5 M in our simulations - spectator ions completely inhibit the formation of crystal-like SBUs.

iv. Above 0.5 M, instead the formation of salt bridges between with half-SBUs has the opposite effect, favoring the formation of crystal-like units. In these conditions salting out effects take place, whereby half-SBUs act as nucleation centers, and are sequestrated by large ionic clusters. In these conditions the overall probability of forming SBUs detached from said clusters remains low, however the SBUs formed are crystal-like, stabilized by salt bridges.

v. The dissociation rate of SBUs with increasing ionic concentration plateaus at values four to seven orders of magnitude larger hen compared to the case studied in absence of ions. This observation indicates that with increasing ionic concentrations the clusters formed by the assembly of SBUs become more dynamic and potentially more prone at self-healing defects formed in the early stages of the crystallization process.
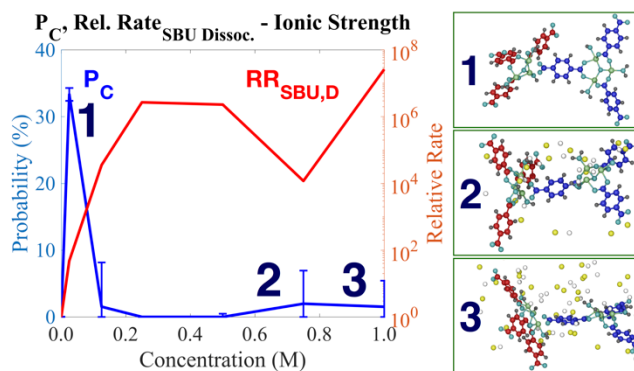

Figure 5. Blue line: Average probability of formation of crystal-like SBUs at equilibrium with increasing concentration of ions and associated standard deviation (error-bars). Red plot: SBU relative dissociation rate with respect to the system in absence of ions. A qualitative upper bound estimate of the dissociation rate is provided by considering it proportional to $e^{\beta\langle\Delta G\rangle}$, where $\langle\Delta G\rangle$ is the ensemble average free energy of formation of SBUs. The configurations present at different instances correspond to the AA SBU and the color code is consistent with Fig. 2.

CONCLUSIONS AND OUTLOOK

This study documents in detail how the complexity of the configurational landscape controls MOF synthesis. Choice of solvents can be used to direct speciation towards desired isomers, as shown in the case of DMF vs water. The implications are that both thermodynamic and kinetic factors are at play, indicating how solvent mixtures can be used to control self-assembly and defect population. We find that structural rigidity of the SBUs is strongly correlated, with ability to form crystalline units. This is controlled partially through the molecular identity of the SBUs, as is dramatically demonstrated in Fig. S4, SI. Similarly, spectator ions and solution ionic strength can also be used to selectively shut down or open up pathways and control defect concentrations. At low ionic strength, higher populations of crystal-like structures are facilitated by ions (F-) attached to the metal center which obstruct the occurrence of double metal center - linker interactions (2Q), whereas higher ionic strengths salting out effects and blocking of linkers become prevalent. While our analysis does not provide an exhaustively quantitative estimate of reaction kinetics, it suggests that the composition of the liquid phase affects significantly the dynamical rearrangement of SBUs, and hence is key to control the self-healing of defects formed in the early stages of nucleation.

The simulations reveal a remarkable conformational ensemble of 300+ structures that are implicated in the early stages of nucleation of MIL-101. Given that these effects arise from both enthalpic and entropic drivers, it is not a priori obvious which will dominate based on simple intuition alone. The identification of these factors demands knowledge of the free energy landscapes and their dependence on synthesis conditions. Modern molecular simulations methods can be used to discriminate these critical drivers, and provide a hypothesis-based approach to synthesis that departs from traditional Edisonian practices. Although this study is specific to MIL-101, we postulate that these observations are generalizable and suggest that systematic theory/experimental studies of solvent mixtures, and varying ionic strength are likely to illuminate the complex nature of MOF synthesis.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Supplementary information (SI) linked with the main article (PDF)
Half-SBUs atomic coordinate files (XYZ)
Half-SBUs topology files (ITP)
Half-SBUs movie (MPG)

AUTHOR INFORMATION

Corresponding Author

*m.salvalaglio@ucl.ac.uk

*Vanda.Glezakou@pnnl.gov

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

1.Yu, J.; Xie, L. H.; Li, J. R.; Ma, Y.; Seminario, J. M.; Balbuena, P. B., CO_2 Capture and Separations Using MOFs: Computational and Experimental Studies. *Chem Rev* **2017,** *117* (14), 9674-9754.
2.Trickett, C. A.; Helal, A.; Al-Maythalony, B. A.; Yamani, Z. H.; Cordova, K. E.; Yaghi, O. M., The chemistry of metal-organic frameworks for CO2 capture, regeneration and conversion. *Nat. Rev. Mater.* **2017,** *2* (8), 17045-17045.
3.Fracaroli, A. M.; Furukawa, H.; Suzuki, M.; Dodd, M.; Okajima, S.; Gandara, F.; Reimer, J. A.; Yaghi, O. M., Metal-organic frameworks with precisely designed interior for carbon dioxide capture in the presence of water. *J. Am. Chem. Soc.* **2014,** *136* (25), 8863-6.
4.Duren, T.; Sarkisov, L.; Yaghi, O. M.; Snurr, R. Q., Design of new materials for methane storage. *Langmuir* **2004,** *20* (7), 2683-9.
5.Jiang, J.; Furukawa, H.; Zhang, Y. B.; Yaghi, O. M., High Methane Storage Working Capacity in Metal-Organic Frameworks with Acrylate Links. *J. Am. Chem. Soc.* **2016,** *138* (32), 10244-51.
6.Lee, J.; Farha, O. K.; Roberts, J.; Scheidt, K. A.; Nguyen, S. T.; Hupp, J. T., Metal-organic framework materials as catalysts. *Chem. Soc. Rev.* **2009,** *38* (5), 1450-9.
7.Park, H. D.; Dinca, M.; Roman-Leshkov, Y., Continuous-Flow Production of Succinic Anhydrides via Catalytic beta-Lactone Carbonylation by Co(CO)4 subsetCr-MIL-101. *J. Am. Chem. Soc.* **2018,** 140(34), 10669-10672.
8.Horcajada, P.; Gref, R.; Baati, T.; Allan, P. K.; Maurin, G.; Couvreur, P.; Ferey, G.; Morris, R. E.; Serre, C., Metal-organic frameworks in biomedicine. *Chem Rev* **2012,** *112* (2), 1232-68.
9.Horcajada, P.; Chalati, T.; Serre, C.; Gillet, B.; Sebrie, C.; Baati, T.; Eubank, J. F.; Heurtaux, D.; Clayette, P.; Kreuz, C.; Chang, J. S.; Hwang, Y. K.; Marsaud, V.; Bories, P. N.; Cynober, L.; Gil, S.; Ferey, G.; Couvreur, P.; Gref, R., Porous metal-organic-framework nanoscale carriers as a potential platform for drug delivery and imaging. *Nat Mater* **2010,** *9* (2), 172-8.
10.Van Vleet, M. J.; Weng, T.; Li, X.; Schmidt, J. R., In Situ, Time-Resolved, and Mechanistic Studies of Metal-Organic Framework Nucleation and Growth. *Chem Rev* **2018,** *118* (7), 3681-3721.
11.Li, J. R.; Sculley, J.; Zhou, H. C., Metal-organic frameworks for separations. *Chem Rev* **2012,** *112* (2), 869-932.
12.Fathieh, F.; Kalmutzki, M. J.; Kapustin, E. A.; Waller, P. J.; Yang, J.; Yaghi, O. M., Practical water production from desert air. *Sci Adv* **2018,** *4* (6), eaat3198.
13.Webber, T. E.; Liu, W.-G.; Desai, S. P.; Lu, C. C.; Truhlar, D. G.; Penn, R. L., Role of a Modulator in the Synthesis of Phase-Pure NU-1000. *ACS Applied Materials & Interfaces* **2017,** *9* (45), 39342-39346.
14.Liu, W.-G.; Truhlar, D. G., Computational Linker Design for Highly Crystalline Metal–Organic Framework NU-1000. *Chem. Mater.* **2017,** *29* (19), 8073-8081.
15.Howarth, A. J.; Liu, Y. Y.; Li, P.; Li, Z. Y.; Wang, T. C.; Hupp, J.; Farha, O. K., Chemical, thermal and mechanical stabilities of metal-organic frameworks. *Nat. Rev. Mater.* **2016,** *1* (3).
16.Howarth, A. J.; Peters, A. W.; Vermeulen, N. A.; Wang, T. C.; Hupp, J. T.; Farha, O. K., Best Practices for the Synthesis, Activation, and Characterization of Metal-Organic Frameworks. *Chem. Mater.* **2017,** *29* (1), 26-39.
17.Taddei, M., When defects turn into virtues: The curious case of zirconium-based metal-organic frameworks. *Coord. Chem. Rev.* **2017,** *343*, 1-24.
18.Gutov, O. V.; Gonzalez Hevia, M.; Escudero-Adan, E. C.; Shafir, A., Metal-Organic Framework (MOF) Defects under Control: Insights into the Missing Linker Sites and Their Implication in the Reactivity of Zirconium-Based Frameworks. *Inorg Chem* **2015,** *54* (17), 8396-400.
19.Ferey, G.; Mellot-Draznieks, C.; Serre, C.; Millange, F.; Dutour, J.; Surble, S.; Margiolaki, I., A chromium terephthalate-based solid with unusually large pore volumes and surface area. *Science* **2005,** *309* (5743), 2040-2.
20.Loiseau, T.; Ferey, G., Crystalline oxyfluorinated open-framework compounds: Silicates, metal phosphates, metal fluorides and metal-organic frameworks (MOF). *J. Fluor. Chem.* **2007,** *128* (4), 413-422.
21.Leng, K. Y.; Sun, Y. Y.; Li, X. L.; Sun, S.; Xu, W., Rapid Synthesis of Metal-Organic Frameworks MIL-101(Cr) Without the Addition of Solvent and Hydrofluoric Acid. *Crystal Growth & Design* **2016,** *16* (3), 1168-1171.
22.Mishra, S.; Daniele, S., Metal-Organic Derivatives with Fluorinated Ligands as Precursors for Inorganic Nanomaterials. *Chem Rev* **2015,** *115* (16), 8379-448.
23.Bromberg, L.; Diao, Y.; Wu, H. M.; Speakman, S. A.; Hatton, T. A., Chromium(III) Terephthalate Metal Organic Framework (MIL-101): HF-Free Synthesis, Structure, Polyoxometalate Composites, and Catalytic Properties. *Chem. Mater.* **2012,** *24* (9), 1664-1675.
24.Yang, J. F.; Zhao, Q.; Li, J. P.; Dong, J. X., Synthesis of metal-organic framework MIL-101 in TMAOH-Cr(NO3)(3)-H2BDC-H2O and its hydrogen-storage behavior. *Microporous Mesoporous Mater.* **2010,** *130* (1-3), 174-179.
25.Rallapalli, P. B. S.; Raj, M. C.; Senthilkumar, S.; Somani, R. S.; Bajaj, H. C., HF-free synthesis of MIL-101(Cr) and its hydrogen adsorption studies. *Environ. Prog. Sustain. Energy* **2016,** *35* (2), 461-468.

26.Liang, Z. J.; Marshall, M.; Ng, C. H.; Chaffee, A. L., Comparison of Conventional and HF-Free-Synthesized MIL-101 for CO2 Adsorption Separation and Their Water Stabilities. *Energy & Fuels* **2013,** *27* (12), 7612-7618.

27.Low, J. J.; Benin, A. I.; Jakubczak, P.; Abrahamian, J. F.; Faheem, S. A.; Willis, R. R., Virtual high throughput screening confirmed experimentally: porous coordination polymer hydration. *J. Am. Chem. Soc.* **2009,** *131* (43), 15834-42.

28.Cantu, D. C.; McGrail, B. P.; Glezakou, V. A., Formation Mechanism of the Secondary Building Unit in a Chromium Terephthalate Metal-Organic Framework. *Chem. Mater.* **2014,** *26* (22), 6401-6409.

29.Férey, G.; Haouas, M.; Loiseau, T.; Taulelle, F., Nanoporous Solids: How Do They Form? An In Situ Approach. *Chem. Mater.* **2013,** *26* (1), 299-309.

30.Goesten, M. G.; Stavitski, E.; Juan-Alcaniz, J.; Martinez-Joaristi, A.; Petukhov, A. V.; Kapteijn, F.; Gascon, J., Small-angle X-ray scattering documents the growth of metal-organic frameworks. *Catal. Today* **2013,** *205*, 120-127.

31.Embrechts, H.; Kriesten, M.; Hoffmann, K.; Peukert, W.; Hartmann, M.; Distaso, M., Elucidation of the Formation Mechanism of Metal-Organic Frameworks via in-Situ Raman and FTIR Spectroscopy under Solvothermal Conditions. *J. Phys. Chem. C* **2018,** *122* (23), 12267-12278.

32.Yoneya, M.; Tsuzuki, S.; Yamaguchi, T.; Sato, S.; Fujita, M., Coordination-Directed Self-Assembly of M12L24 Nanocage: Effects of Kinetic Trapping on the Assembly Process. *ACS Nano* **2014,** *8* (2), 1290-1296.

33.Yoneya, M.; Yamaguchi, T.; Sato, S.; Fujita, M., Simulation of Metal–Ligand Self-Assembly into Spherical Complex M6L8. *J. Am. Chem. Soc.* **2012,** *134* (35), 14401-14407.

34.Yoneya, M.; Tsuzuki, S.; Aoyagi, M., Simulation of metal–organic framework self-assembly. *Physical Chemistry Chemical Physics* **2015,** *17* (14), 8649-8652.

35.Biswal, D.; Kusalik, P. G., Probing Molecular Mechanisms of Self-Assembly in Metal–Organic Frameworks. *ACS Nano* **2016,** *11* (1), 258-268.

36.Nguyen, V.; Grünwald, M., Microscopic Origins of Poor Crystallinity in the Synthesis of Covalent Organic Framework COF-5. *J. Am. Chem. Soc.* **2018,** *140* (9), 3306-3311.

37.Cheetham, A. K.; Kieslich, G.; Yeung, H. H. M., Thermodynamic and Kinetic Effects in the Crystallization of Metal–Organic Frameworks. *Accounts of Chemical Research* **2018,** *51* (3), 659-667.

38.Humphrey, W.; Dalke, A.; Schulten, K., VMD - Visual Molecular Dynamics. *J. Mol. Graph.* **1996,** *14*, 33-38.

39.Jorgensen, W. L.; Tirado-Rives, J., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988,** *110* (6), 1657-66.

40.Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983,** *79* (2), 926-935.

41.Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015,** *1-2*, 19-25.

42.Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D., Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J Chem Theory Comput* **2012,** *8* (1), 61-74.

43.van der Spoel, D.; van Maaren, P. J.; Caleman, C., GROMACS molecule & liquid database. *Bioinformatics* **2012,** *28* (5), 752-753.

44.Hutchinson, M. H.; Sutton, L. E., The electric dipole moments of a vinylenic series. *J. Chem. Soc.,* **1958,** 0, 4382-4386.

45.Reichardt, C.; Welton, T., *Solvents and Solvent Effects in Organic Chemistry*. Wiley: 2011.

46.Barducci, A.; Bussi, G.; Parrinello, M., Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett* **2008,** *100* (2), 020603.

47.Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J Chem Phys* **2007,** *126* (1), 014101.

48.Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.

49.Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.

50.Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18* (12), 1463-1472.

51.Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G., PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014,** *185* (2), 604-613.

52.Petkov, P.; Bon, V.; Hobday, C.; Kuc, A.; Melix, P.; Kaskel, S.; Düren, T.; Heine, T., Conformational isomerism controls collective flexibility in metal-organic framework DUT-8(Ni). *Phys. Chem. Chem. Phys.,* **2019,** *21, 674-680.*

53.Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16, Revision A.03*, Gaussian Inc.: Wallingford CT, 2016.

54.Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys Rev Lett* **1996,** *77* (18), 3865-3868.

55.Adamo, C.; Barone, V., Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999,** *110* (13), 6158-6170.

56.Hariharan, P. C.; Pople, J. A., The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973,** *28* (3), 213-222.

57.Feller, D., The role of databases in support of computational chemistry calculations. *J. Comput. Chem.* **1996,** *17* (13), 1571-1586.

58.Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L., Basis set exchange: a community database for computational sciences. *J Chem Inf Model* **2007,** *47* (3), 1045-52.

59.Dennis, J. E.; Schnabel, R. B., *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics: 1996.

60.Powell, M. J. D., *Numerical methods for nonlinear algebraic equations*. Gordon and Breach Science Publishers: 1970.

61.Goesten, M. G.; Magusin, P. C.; Pidko, E. A.; Mezari, B.; Hensen, E. J.; Kapteijn, F.; Gascon, J., Molecular promoting of aluminum metal-organic framework topology MIL-101 by N,N-dimethylformamide. *Inorg Chem* **2014,** *53* (2), 882-7.

62.Stylianou, K. C.; Heck, R.; Chong, S. Y.; Bacsa, J.; Jones, J. T.; Khimyak, Y. Z.; Bradshaw, D.; Rosseinsky, M. J., A guest-responsive fluorescent 3D microporous metal-organic framework derived from a long-lifetime pyrene core. *J. Am. Chem. Soc.* **2010,** *132* (12), 4119-30.

Graphic entry for the Table of Contents (TOC)