THE JOURNAL OF PHYSICAL CHEMISTRY C
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCC

Article

Locating Guest Molecules inside Metal–Organic Framework Pores with a Multilevel Computational Approach

Published as part of The Journal of Physical Chemistry virtual special issue "Early-Career and Emerging Researchers in Physical Chemistry Volume 2".

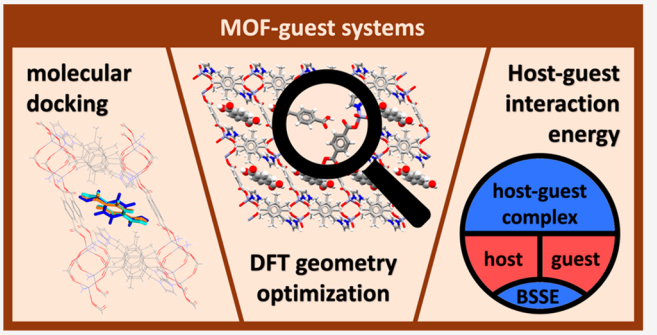
Michelle Ernst,* Tomasz Poręba, Lars Gnägi, and Ganna Gryn'ova*

Cite This: *J. Phys. Chem. C* 2023, 127, 523–531

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Molecular docking has traditionally mostly been employed in the field of protein–ligand binding. Here, we extend this method, in combination with DFT-level geometry optimizations, to locate guest molecules inside the pores of metal–organic frameworks. The position and nature of the guest molecules tune the physicochemical properties of the host–guest systems. Therefore, it is essential to be able to reliably locate them to rationally enhance the performance of the known metal–organic frameworks and facilitate new material discovery. The results obtained with this approach are compared to experimental data. We show that the presented method can, in general, accurately locate adsorption sites and structures of the host–guest complexes. We therefore propose our approach as a computational alternative when no experimental structures of guest-loaded MOFs are available. Additional information on the adsorption strength in the studied host–guest systems emerges from the computed interaction energies. Our findings provide the basis for other computational studies on MOF–guest systems and contribute to a better understanding of the structure–interaction–property interplay associated with them.

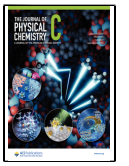


INTRODUCTION

Key features of metal–organic frameworks (MOFs), resulting from their porous structure and enormous internal surface, include the capture, transport, and release of guest molecules. Their architecture, consisting of easily combinable metal nodes and organic linkers, allows for remarkable synthetic and structural tunability. Hence, they are employed in a plethora of applications in gas storage,[1] separation,[2] sensing,[3] catalysis,[4] biomedicine,[5] and many other fields. Most of these applications rely on noncovalent interactions between the guest molecules adsorbed on the pore walls and the framework. Therefore, the location of the adsorption site and the strength of the interaction are critical parameters, and a profound understanding of these features is indispensable for a rational MOF design.[6] A single experiment alone can often only partially unravel these features: X-ray and neutron diffraction experiments, although widely used in MOF research, rely on the guests being adsorbed in a regular and ordered manner and only provide insights about long-range order. Spectroscopic experiments, on the one hand, allow for a characterization of the functional groups and the change in symmetry and force constants due to guest–framework interactions but not a precise structure elucidation. On the other hand, adsorption isotherm measurements and differential scanning calorimetry can shed light on the stoichiometry and adsorption strength but not on the site geometry. None of these techniques alone provide local atomic-scale information simultaneously on site and strength of the host–guest interaction. *In silico* modeling offers, arguably, the only means to obtain both the structural and the energetic description of the MOF–guest complexes simultaneously, with an added benefit of elucidating the structure–property relationships.[7,8]

The use of computational modeling to identify potential adsorption sites has a long tradition in the field of protein–ligand interactions, where molecular docking is used to predict the binding sites and conformations of the ligands. Broadly speaking, the method consists of two parts: (1) sampling over various positions and conformations of the ligand and (2) evaluation of the free energies of the sampled conformations via a scoring function. The stronger the interaction between

Received: August 4, 2022
Revised: November 8, 2022
Published: December 26, 2022



https://doi.org/10.1021/acs.jpcc.2c05561
J. Phys. Chem. C 2023, 127, 523–531

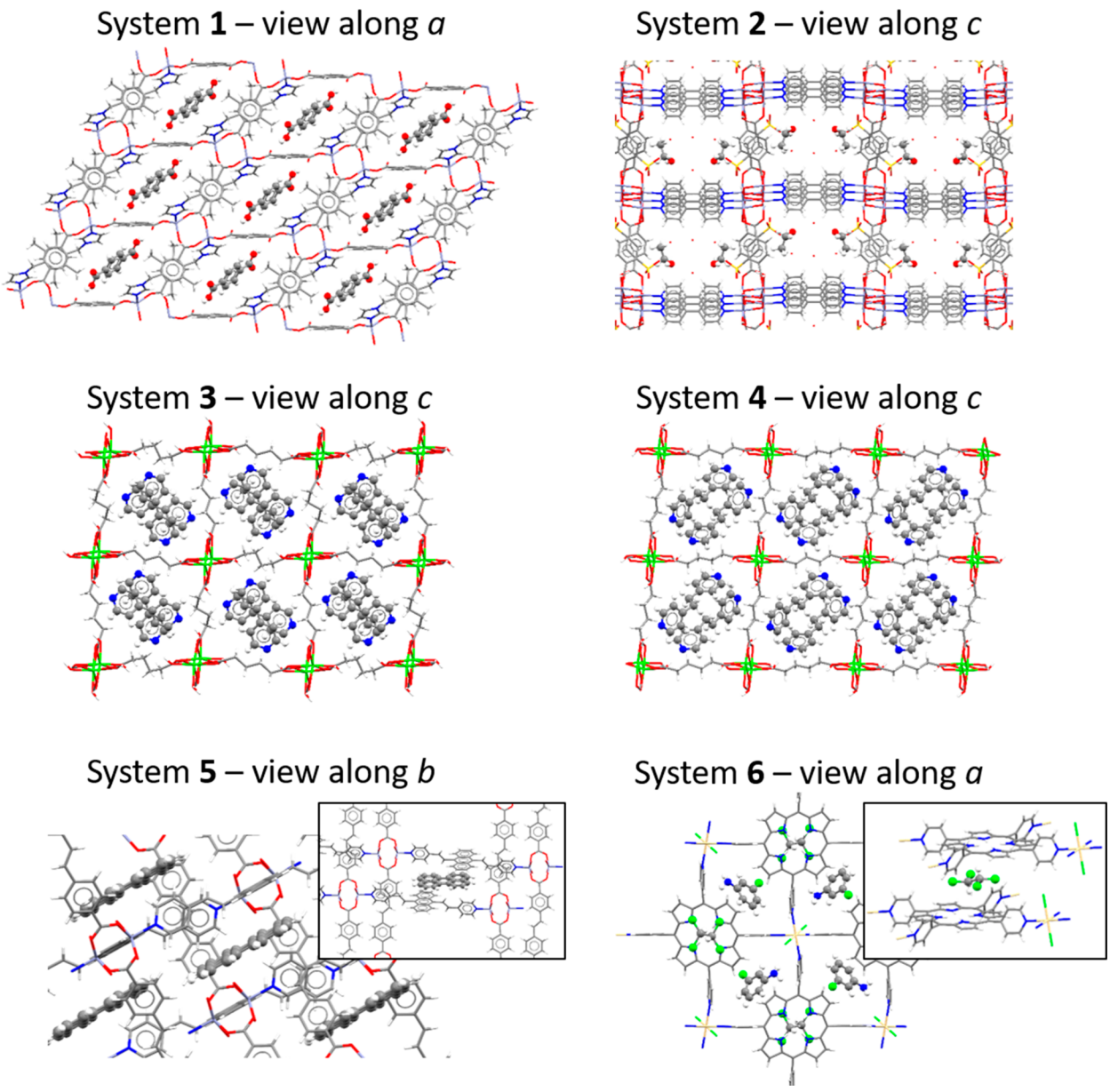


Figure 1. Crystal structures of the six examined MOF–guest systems. The guests are shown with balls and sticks, and the frameworks are shown with sticks. The insets for systems **5** and **6** highlight the guest position.

protein and ligand, the lower (more stabilizing) is the free energy associated with a specific conformation. In applying molecular docking to find adsorption sites in MOFs, the framework takes the role of the protein, and the guest takes the role of the ligand. Literature examples on the docking of small molecules to MOFs include drugs in MOF-5,[9] anionic dye reactive blue-4 in ZIF-8,[10] ethanol in Zn-MOF,[11] 5-nitroimidazole in MIL-101(Cr),[12] metolachlor in MIL-53(Al),[13] and polycyclic aromatic hydrocarbons in UiO-66.[14] The vast majority of MOF–guest research focuses on gas molecules in MOF pores, where advanced techniques combining simulated annealing and periodic density functional theory (DFT) have been put forward.[15] In contrast, these studies, as well as the approach presented in this article, primarily concentrate on medium-sized organic guest molecules. However, unlike our proposed methodology, in these literature examples optimization at the DFT level was not performed; thus the results of molecular docking were not evaluated against quantum-chemical data.

Weakly interacting guests show a very dynamic behavior in MOF pores, which often results in an orientational disorder. In this work we focus on guests, which interact with the MOF strongly and thus become ordered within the crystalline framework. Our goals are to find their preferred adsorption sites and geometries and to determine their interaction energies *in silico*. Although we consider a periodic framework, these features are sufficiently local and can be properly represented with a finite cluster model, which allows for a much broader set of computational tools to be employed. We have recently compared cluster and periodic results for two exemplary MOF–guest systems and found that the cluster model is qualitatively able to reproduce the periodic results and is useful to analyze the host–guest interactions.[16]

In this work, we applied the method (and implementation) of molecular docking, which is well-established in protein–

https://doi.org/10.1021/acs.jpcc.2c05561
J. Phys. Chem. C 2023, 127, 523–531

ligand binding, to the field of MOF−guest complexes. The software AutoDock v4.2.6, originally designed for protein−ligand binding, is used in the initial screening of various adsorption sites of guest molecules inside MOF pores. We employ a genetic sampling algorithm in combination with additional force field parameters, needed to describe MOFs (see Methods). Subsequently, the so-obtained structures are used as inputs for DFT-based geometry optimization. With these structures, in turn, host−guest interaction energies are computed.

This new approach is validated by comparing the structures, obtained here via multilevel simulations, to five experimentally determined MOF−guest structures deposited in the Cambridge Structural Database (CSD). We discuss how computational modeling can be used as an efficient tool to elucidate where, in which conformation, and how strongly the guest molecule is adsorbed. Finally, we successfully apply this method to a newly synthesized MOF containing two different halogenated organic guests at two different adsorption sites.

METHODS

Materials. We have selected five MOF−guest systems from the Cambridge Structural Database as examples for which the structures of the framework and the guest have been determined. Additionally, a new MOF with two guests was synthesized and characterized experimentally and computationally. The six investigated systems are the following (see Figure 1).

System 1. MOF $[Zn_2(bdc)_2(bimx)]$ (H_2bdc = 1,4-benzenedicarboxylic acid, bimx = 1,4-bis(imidazol-1-ylmethyl)-2,3,5,6-tetramethylbenzene) with guest 1,4-benzenedicarboxylic acid (CSD refcode: DOGZIJ).[17] The guest molecules reside in channels along the *a*-axis.

System 2. $\{[Zn_3(\mu 3\text{-}OH)_3(2\text{-}stp)(bpy)_{1.5}(H_2O)](EtOH)(2H_2O)\}_n$ (2-stp = 2-sulfonylterephthalate, bpy = 4,4′-bipyridyl) (CSD refcode: JITPOS).[18] This three-dimensional MOF consists of zinc−oxygen secondary building units forming chains, which are in turn connected by bipyridine and 2-sulfonatoterephthalate in the other two directions (cf. Figure 1). The ethanol guest molecules trapped in the pores are close to the sulfonate groups of the organic linkers.

System 3. GW-MOF-7 (CSD refcode: PARHAS).[19] This MOF consists of a calcium-adipate framework and the 4,4′-dipyridyl guest in the channels.

System 4. GW-MOF-8 (CSD refcode: PARHEW).[19] This is the same MOF as GW-MOF-7, but it contains a different guest, 1,2-bis(4-pyridyl)ethane, in the channels.

System 5. This MOF (CSD refcode: FUNLEH)[20] consists of a 4-fold interpenetrating network with zinc paddlewheel nodes connected via the two organic linkers. The perylene guests are π-stacked in between the aromatic linkers. In contrast to systems **1**−**4**, in which the guests form strong hydrogen bonds with the framework, in system **5** the host−guest interactions are dominated by $\pi-\pi$ interactions.

System 6. A newly synthesized MOF, named "UB-MOF-1", consists of octahedral cadmium nodes with two coordinated chlorine atoms and four nitrogen atoms in the equatorial plane per cadmium. The organic linkers consist of pyridine-substituted porphyrin rings. This MOF hosts 3-chloroaniline (3-CA) in the pore channels and 1,1,2,2-tetrachloroethane (TCE) between the porphyrin rings (cf. Figure 1). The synthesis and structure determination are described in the experimental part.

These systems show high coordination flexibility and form a variety of framework geometries. System **1** contains a bulky polar guest molecule, whereas system **2** contains a very small polar guest. Comparison of systems **3** and **4** allows evaluating the effect of a different guest alone, as the framework is the same in both cases. In contrast to systems **1**−**4**, in which the MOFs bind the guest primarily via hydrogen bonding, in system **5** the host−guest interactions are dominated by $\pi-\pi$ stacking. System **6** contains both a polar (chloroaniline) and a nonpolar (tetrachloroethane) guest occupying two positions.

Computational Methods. *Preparation of the Pore Model.* To define a proper pore model for molecular docking, we start from the experimental crystal structure of the guest-occluded MOFs. Since MOFs expand in one, two, or three dimensions in which they are linked by covalent or coordination bonds, some of these bonds need to be cut to terminate the periodic extension. The removed atoms were replaced with hydrogen atoms to obtain a chemically reasonable and charge-balanced pore model. The resulting cluster should include all possible docking sites available for a guest inside the MOF framework. For certain topologies, in particular the interpenetrated networks, this can be a more delicate task because the choice of a cluster is less straightforward. One must pay attention to include all possible docking sites while not giving the guest too much space by omitting interpenetrated parts of the framework. The pore model should have the charge of the framework (usually, zero). In case of a relevant disorder, either two different pore models or a larger one including different disordered parts must be prepared. In this study, we have only included nondisordered MOFs. The selection of the cluster model for every system is described in more detail in the Supporting Information.

Molecular Docking. We used AutoDock 4.2.6 to perform the molecular docking and AutoDockTools 1.5.6 for the preparation of the inputs.[21] By default, AutoDock does not have force field parameters for many metals. Therefore, the parameters from the universal force field (UFF),[22] which is the original source of parameters in AutoDock, were manually added.

The automatic assignment of Kollman or Gasteiger atomic charges in AutoDockTools, in general, cannot be applied to MOFs since it is constructed for proteins and is not parametrized for metals. Computing accurate atomic charges at higher *ab initio* levels of theory is suboptimal due to the large system sizes and high computational cost. Therefore, we used Mulliken charges of the pore models at the B3LYP/6-31G level of theory.

The grid box, defining where the guest molecules can go during the generation of the adsorption conformations with molecular docking, was chosen such that all possible symmetry-independent adsorption sites are included. The spacing between two grid points was always kept at the default value of 0.375 Å, and the size of the box was adjusted to the system size. To perform the conformational search, the genetic algorithm (GA) was used, where the fitness is the binding free energy of a given conformation. The lower (more negative) the binding energy, the more fit is the respective individual (in GA terms). The selected search parameters are as follows: GA runs, 20; population size, 150; maximum number of evaluations, 2 500 000 (medium); maximum number of generations, 27 000; maximum number of top individuals that automatically survive, 1; rate of gene mutation, 0.02; rate of crossover, 0.8; GA window size (i.e., number of generations

https://doi.org/10.1021/acs.jpcc.2c05561
J. Phys. Chem. C 2023, 127, 523−531

considered when ranking the individuals in the current population), 10. If not otherwise specified, the ligands were fully flexible while the framework was kept rigid. The number of GA runs corresponds to the number of dockings executed, and each of them results in one optimum conformation; i.e., 20 runs lead to 20 conformations. If the root-mean-square deviation (RMSD) between conformations is less than 2.0 Å, they are grouped into the same cluster such that in the end, one or several (max 20) clusters of conformations are generated.

Partial Geometry Optimization of Guest with Density Functional Theory (DFT). The coordinates from the best conformation of every cluster were used as starting point for a subsequent DFT geometry optimization. Only the coordinates of the guest were optimized at the DFT level, while the atoms of the framework were kept frozen. Partial geometry optimizations were performed at the B3LYP-D3/6-31G level of theory except for system **6** with cadmium atoms, for which the LANL2DZ basis set with an effective core potential was used.

Computation of Interaction Energy. The geometry from the partial optimization of the guest was used for a single point computation of the host−guest interaction energy at the B3LYP-D3/cc-pVTZ level of theory with counterpoise correction for the basis set superposition error with the software Gaussian 16.[23] The clusters usually contain several metals and several hundred atoms; thus we recommend a lower level of theory for the geometry optimization, while a reasonable basis set is needed for accurate interaction energies. The interaction energy (E^{int}) was calculated as follows:

$$E^{int} = E_{hg} - [E_h + E_g] + E^{BSSE} \quad (1)$$

where the electronic energies of both the host (E_h) and the guest (E_g) are computed in the unrelaxed geometry of the complex, and E^{BSSE} is the basis set superposition error correction.

Visualization. Discovery Studio 2021 was used for the visualization of the guests at the adsorbed sites.

Experimental Section. *Synthesis of UB-MOF-1.* All chemicals are commercially available and were used without further purification. The synthesis was performed in a custom-made glass-tube (5.0 mm internal diameter, 30 cm long), essentially following the procedure described by DeVries et al.[24]

The layering of the following three components was performed using long needles for precise and slow placement and to avoid contamination of the upper layers by residual droplets of the lower solutions at the tube wall.

A dark blue solution of *meso*-tetra-4-pyridylporphyrin (TPyP, 99%, 6.0 mg, 0.0097 mmol, 1.0 equiv) in a mixture of 1,1,2,2-tetrachloroethane (TCE, 98.5%, 0.75 mL) and methanol (HPLC grade, 0.25 mL) was placed at the bottom of the vertically fixed tube. Then, 3-chloroaniline (99%, 1.0 mL) was carefully layered on top of the previous solution. Lastly, a solution of cadmium(II) iodide (99%, 7.2 mg, 0.02 mmol, 2.0 equiv) in *N*,*N*-dimethylacetamide (1.0 mL) was added on top. The tube was then sealed with a septum, wrapped in aluminum foil, and allowed to sit undisturbed for 6 weeks at ambient temperature. The content of the tube was poured into a Petri dish, and the dark-red prism crystals suitable for XRD analysis were removed from the solution shortly before measurement.

The incorporation of chloride is most likely due to the Finkelstein reaction between TCE and CdI_2. Filtration of TCE as the potential source of HCl over basic aluminum oxide and repetition of the crystallization yielded the same MOF, ruling out the influence of any acidic substrate impurities.

Crystallographic Section. A single crystal of UB-MOF-1 with approximate dimensions of 0.080 mm × 0.045 mm × 0.030 mm was chosen for the diffraction experiments. The experiment was carried out at ID15B beamline at ESRF, using a convergent monochromatic beam (30 keV). Data were collected at room temperature by an Eiger2 9M CdTe detector. The sample-to-detector distance (180 mm) was calibrated using a Si powder standard and a vanadinite $[Pb_5(VO_4)_3Cl]$ single crystal. The data collection strategy consisted of a ω-scan ($-70° \leq \omega \leq +70°$), with 0.5° step width and 0.5 s exposure. The data reduction has been carried out in the CrysAli Pro software. The structure was solved with the dual-space algorithm, as implemented in SHELXT. All non-hydrogen atoms were refined anisotropically using the full-matrix, least-squares method on F^2 by the SHELXL software. All hydrogen atoms were refined using the riding model. Isotropic displacement factors of hydrogen atoms were equal to 1.2 times the value of an equivalent displacement factor of the parent atoms. Measurement details are presented in Table S1 of the Supporting Information. CCDC 2157164 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/structures.

RESULTS

System 1: MOF $[Zn_2(bdc)_2(bimx)]$. The pore model was chosen such that the 1,4-benzenedicarboxylic acid guest molecule is in a cage delimitated by the ligands and the zinc−oxygen SBUs at the corners of the pore (Figure 2 and Figure S1). The 20 GA runs of the molecular docking resulted in only one cluster of conformations, i.e., all 20 resulting conformations are the same within the root-mean-square deviation (RMSD tolerance) of 2.0 Å. The estimated free energy of binding for the most stable docked conformation is −8.37 kJ mol^{-1} (Table 1). The simulated position of the 1,4-benzenedicarboxylic acid guest is in excellent agreement with the experimentally determined position (Figure 2). The hydrogen atoms of both carboxylic acid groups of the guest form hydrogen bonds with the oxygen atoms of the carboxylate of the linker (1.83 Å in the experimental crystal structure and 1.60 Å in the optimized structure), while several longer contacts (2.5−3.0 Å) are established between the C═O groups of the guest and the hydrogen atoms of the framework. The geometry optimization did not substantially alter the position and geometry of the guest molecule (Figure 2). The interaction energy computed with DFT is −259.87 kJ mol^{-1}. We also performed full periodic geometry relaxation of system **1** with and without the guest molecule at the B3LYP-D3/pob_TZVP_rev2 level of theory using the software Crystal17.[25] The structures of the guest-occluded and guest-free MOF (Figure 3) are very similar, with only small deviations in atomic coordinates and unit cell parameters (see Table S2 in Supporting Information).

System 2: MOF $\{[Zn_3(\mu_3\text{-}OH)_3(2\text{-}stp)(bpy)_{1.5}(H_2O)](EtOH)(2H_2O)\}_n$. The docking with 20 GA runs resulted in only one cluster of conformations, since the RMSD for the ethanol docked at the same functional group is low regardless

https://doi.org/10.1021/acs.jpcc.2c05561
J. Phys. Chem. C 2023, 127, 523−531

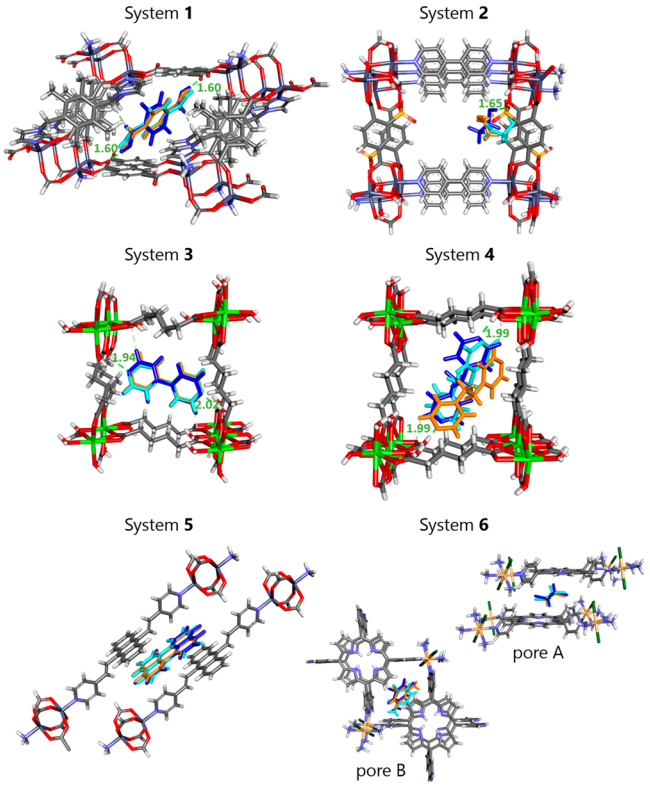


Figure 2. Experimental and computed geometries of the host–guest complexes in pore models of the studied MOFs. The guest molecules are shown in different colors: orange for experimental, light blue for docked, and dark blue for the geometry optimized structure. Hydrogen bonds between geometry optimized guest and framework are shown with dashed green lines. Hydrogen atoms are depicted in white, carbon in gray, nitrogen in violet, oxygen in red, sulfur in yellow, chlorine in dark green, calcium in green, zinc in silver, and cadmium in ochre color.

Table 1. Host–Guest Interaction Energies in kJ mol^{-1}

	system				
	1	2	3	4[a]	5
energy from docking	−8.37	−4.02	6.82	13.85	−35.35
energy from DFT	−259.87	−58.66	−166.9	−145.14	−158.11

[a]Guest geometry is frozen (no change of internal coordinates).

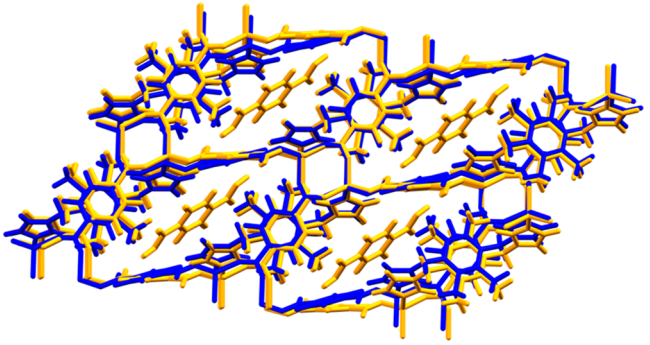


Figure 3. Optimized geometries of guest-free (blue) and guest-occluded (orange) MOF [Zn_2(bdc)$_2$(bimx)].

of the orientation. The lowest free energy of binding is −4.02 kJ mol^{-1}. The DFT geometry optimization led to a reorientation of the ethanol molecule, but the main hydrogen bond between the hydrogen of the ethanol's hydroxyl group and the sulfonate group of the framework remained. The cluster model used in molecular docking and geometry optimization is large (24 metal atoms and 362 atoms in total), impeding the use of a relatively large cc-pVTZ basis set in the subsequent interaction energy evaluation. Therefore, a smaller cluster model was used in the latter (see Figure S2). The idea behind this multilevel approach is precisely that the appropriate model and level of theory can be used for each problem set. The higher the level of theory, the smaller is the structural model and *vice versa*. Since in this system the guest interacts only with the linker, we built a smaller model including only these fragments to compute the interaction energy of −58.66 kJ mol^{-1}.

System 3: GW-MOF-7. For GW-MOF-7, all 20 GA runs of the molecular docking simulation produced only one cluster of conformations, which is in excellent agreement with the experimental guest position. The best conformation from docking has a positive binding energy of +6.82 kJ mol^{-1}. The subsequent geometry optimization of the guest did not substantially alter the position and structure. The interaction energy between framework and guest, computed in this geometry at the B3LYP-D3/cc-pVTZ level of theory, is clearly stabilizing (−166.9 kJ mol^{-1}).

System 4: GW-MOF-8. Initially, all torsions of the guest were kept active (movable) during the molecular docking runs. Three distinct clusters of conformations were found, with free binding energies of +10.59 kJ mol^{-1}, +11.30 kJ mol^{-1}, and +12.64 kJ mol^{-1}, respectively. All three types of conformations differ from the experimental position of the guest. The DFT host–guest interaction energies in these three conformations are −126.78 kJ mol^{-1}, −147.57 kJ mol^{-1}, and −87.49 kJ mol^{-1}, respectively. Figure S3 in the Supporting Information shows the experimental guest position and conformation, as well as the three distinct conformations resulting from docking.

In the next step, the torsions were kept inactive (frozen) in an attempt to recover the experimental guest geometry. One cluster of conformations was found with a free binding energy of +13.85 kJ mol^{-1}. However, the docked site was still in disagreement with the experimental one (Figure 2). A smaller-spaced grid map for the docking (0.275 Å instead of 0.375 Å) did not improve the outcome. The intermolecular interaction distances in the six examined systems are generally overestimated by docking. The 1,2-bis(4-pyridyl)ethane guest occupies a large portion of the pore space since it is present during the synthesis. Ultimately, molecular docking is not able to locate the correct adsorption site in this tight space and prefers another adsorption geometry, where the guest has more space and the intermolecular interaction distances are longer.

System 5: [Zn_2(SDC)$_2$(An2Py)]·perylene. In contrast to the previous system, dominated by hydrogen bonding, in system **5** the π–π interactions are responsible for the formation of the host–guest complex. A simplified pore model was constructed from two π-conjugated linkers and their adjacent zinc-paddlewheel nodes (see Figure 2). Molecular docking was able to correctly locate the perylene guest between the π-conjugated linker systems. The free binding energy from docking is −35.35 kJ mol^{-1}. The DFT geometry optimization did not significantly alter the geometry, and the DFT interaction energy is −158.11 kJ mol^{-1}. Importantly, without dispersion correction the geometry

https://doi.org/10.1021/acs.jpcc.2c05561
J. Phys. Chem. C 2023, 127, 523−531

optimization erroneously leads to guest molecule rotation (see Figure S4 in the Supporting Information).

System 6: UB-MOF-1. A new TPyP-based MOF was synthesized (i) to achieve a MOF with pores large enough to incorporate chlorinated solvent molecules used in crystallization and (ii) to maximize chemical affinity toward them. A number of structures with TCE and $CHCl_3$ interacting with linkers were reported for MOFs comprising closed-shell metal nodes and TPyP linkers.[26−28] Among them, cadmium analogues showed rich coordination chemistry, resulting in a multitude of three-dimensional networks.[29] The largest pores in this group of networks were obtained by keeping the linker in the H_2TPyP form, which allowed incorporating both the 1,1,2,2-tetrachloroethane (TCE) and 3-chloroaniline (3-CA).

The newly synthesized UB-MOF-1 consists of Cd(II) colinear with two coordinated chlorine atoms, which are in turn connected to four meso-tetra-4-pyridylporphyrin linkers, forming octahedral nodes. It presents channels in the crystallographic *a* direction. Such channels are common adsorption sites for guest molecules. In addition, there is another adsorption site between the porphyrin rings (distance between the porphyrin planes: 7.127(4) Å) where guests can be adsorbed in an "atop" arrangement. This adsorption mode has been observed in other systems.[30,31] The two adsorption sites differ in chemical affinity, and indeed, UB-MOF-1 contains two guests, halogenated organic solvents TCE and 3-CA, clearly located at these distinct sites. The TCE molecule is disordered by an inversion center so that the carbon atoms have 0.5 occupancy each.

We constructed two separate pore models (cf. Figure 2). Pore model A consists of two π-stacked porphyrin rings and the adjacent SBUs and therefore contains the experimentally found adsorption site of TCE. Pore model B contains the channel, in which the 3-CA is located, as well as four adjacent porphyrin rings. To ensure a chemically reasonable charge-neutral pore model, two diagonal nitrogen atoms per porphyrin ring are bonded to a hydrogen atom, whereas the other two are not.

Next, we performed four docking simulations: using pore A and pore B (with the simulation box constrained to the channel) with the two guests (TCE and 3-CA), respectively. The molecular docking was able to locate both experimentally found guest positions. The position and orientation of TCE in pore A and the position of the phenyl ring of 3-CA in pore B are in excellent agreement between experiment and simulation, while the orientation of the amine functional group in 3-CA is not (cf. Figure 2 and Figure S1). However, the symmetry-equivalent (but rotated) representations of pore B with the guest position taken from the experimental crystal structure (Figure 4, left) and the optimized structure (Figure 4, right) clearly show that the orientation identified by simulation is very similar to the experimental one.

The docking energies in Table 2 indicate that both guest molecules are more favored at the adsorption site between the two porphyrin rings (pore A). When we use the entire cluster model of pore B to perform the docking simulation, likewise, both guest molecules prefer the adsorption site between the porphyrin rings. On account of this, the best adsorption conformations from all four scenarios (TCE and 3-CA in pores A and B, respectively) were used to perform partial geometry optimizations of the guests and subsequent computations of the host−guest interaction energies (Table 2). The DFT interaction energies are in line with the results of the X-ray diffraction measurements: TCE is preferentially bound at adsorption site A, and 3-CA is preferentially bound at adsorption site B.

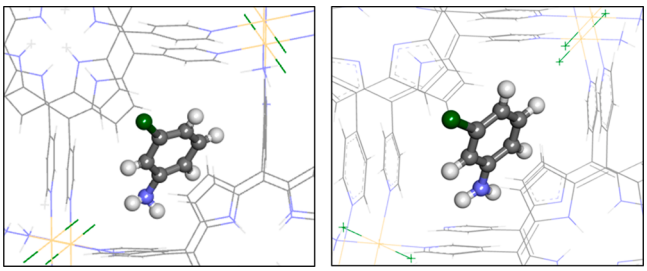


Figure 4. Symmetry-equivalent pore B with experimentally determined guest (left) and geometry optimized guest in pore B (right). In this representation, the orientation of the chlorine atoms in the experimental structure differs with respect to pore model B, while the orientation of the guest is similar.

Table 2. Interaction Energies from Molecular Docking and DFT in kJ mol^{-1} [a]

	1,1,2,2-tetrachloroethane (TCE)	3-chloroaniline (3-CA)
	Free Binding Energy from Molecular Docking	
pore A	***−14.56***	**−12.51**
pore B	−7.20	*−4.94*
	Interaction Energy from DFT	
pore A	***−114.45***	−102.41
pore B	−74.78	***−124.56***

[a]The computationally preferred site is in bold. Italic font indicates where the molecule is located in the experimental structure in system **6**.

DISCUSSION

Locating guest molecules inside MOFs is challenging both computationally and experimentally, since both approaches entail uncertainties.[32,33] When the experimental structure of the host−guest complex is not available, the starting geometries can be created based on chemical intuition,[34] electrostatic complementarity,[35] or other force-field-based methods.[36] Here, we introduce and evaluate an alternative multilevel approach combining molecular docking with density functional theory. Specifically, molecular docking is employed to locate possible adsorption sites. The guest positions in them are subsequently reoptimized at a DFT level with a small basis set. Finally, DFT computations with a larger basis set, dispersion, and counterpoise corrections are used to refine the interaction energies.

Adsorption Geometry. For all systems except system **4**, molecular docking correctly located the adsorption site. For systems **2** and **6** (pore B), the simulated adsorption geometry deviates from the experimentally determined one, while for systems **1**, **3**, **5**, and **6** (pore A) modeling is in very good agreement. Success of the molecular docking approach is therefore evident in systems, in which the guest assumes a well-defined adsorption position, i.e., when there is strong complementarity in the sizes of the pore and the guest, as well as in the noncovalent interactions, established between the host framework and the guest molecule.

Accurate prediction of the adsorption sites with molecular docking becomes challenging in systems lacking such complementarity. For example, in system **2**, the pores are relatively large while the guest molecule, ethanol, is small and

https://doi.org/10.1021/acs.jpcc.2c05561
J. Phys. Chem. C 2023, 127, 523−531

can establish only a single relatively strong hydrogen bond with the framework. This high conformational flexibility is further complicated by the residual electron density in the pores in the experimental structure. Moreover, the computed interaction energy is only insignificantly affected by the orientation of the ethanol molecule (see Figure S5 in the Supporting Information), hampering a distinction between different conformations based on energy alone. Nevertheless, with molecular docking, a hydrogen bond between the guest and the sulfonate group of the linker was identified correctly. Our findings for system **2** resemble those of Korb et al.[11] Using the software GOLD,[37] which, like AutoDock, is optimized for protein−ligand binding, the authors also located a hydrogen bond between ethanol and the framework's sulfonate group but could not correctly predict the XRD position.

System **4** is illustrative of an alternative, yet equally challenging scenario: the 1,2-bis(4-pyridyl)ethane guest occupies a large portion of the pore space. As a result, molecular docking is not able to locate the correct adsorption site in this tight space and prefers another adsorption geometry, where the guest has more space and the intermolecular interaction distances are longer. Accordingly, the interaction energy, computed using the experimental geometry of the host−guest complex, is lower (more stabilizing) than that in the structure from molecular docking (see Table S3 in the Supporting Information).

Upon insertion or removal of guests, MOFs can undergo deformations, phase transitions, and/or chemical reactions. These phenomena, critical in highly flexible and breathing MOFs, are not readily captured by molecular docking; thus the scope of applicability of this approach is limited to relatively rigid MOFs. In the latter, the difference between the guest-free and guest-occluded geometry is minimal (see Figure 3 for system **1**); thus experimental structures of the guest-free MOFs, which are much more readily available than the MOF−guest complex structures, provide a reliable starting point for molecular docking.

The results of molecular docking are likely to be improved when a force field, explicitly designed for MOFs, is used. A systematic assessment of atomic charge schemes could further enhance these results. To include all possible adsorption sites, pore models in molecular docking must be sufficiently large; in this work, they consist of 200−450 atoms. However, such system sizes preclude further refinement of geometries and energies above the accuracy of the force-field-based methods. This issue is addressed by turning to smaller, finite cluster models of the adsorption sites, identified in molecular docking, and treating them at higher levels of theory. While such finite cluster models neglect the periodicity of the MOFs and the guest−guest interactions, they are generally able to accurately capture the local host−guest interactions.[16]

Host−Guest Interaction Energy. Interaction energies, computed using density functional theory electronic energies, cannot be directly compared in absolute terms to the free energies of binding from molecular docking. However, according to the results in Table 1, the latter does not provide a good estimation of the interaction energy, not only in comparison to DFT but more importantly in relation to experiment. For example, in systems **3** and **4** (which have been isolated and characterized experimentally), molecular docking predicts repulsive interactions (positive interaction energies), in contrast to the appreciably negative interaction energies at the B3LYP-D3/cc-pVTZ level. Moreover, in system **6** featuring several guests and several adsorption sites that are possible, molecular docking incorrectly predicts the preferred adsorption site (Table 2). Only subsequent DFT optimization allows us to discriminate the correct binding position. This emphasizes the need in energy refinement at higher *ab initio* levels of theory following the preliminary sampling of adsorption sites with molecular docking.

CONCLUSIONS

Many practical applications of metal−organic frameworks involve the incorporation of guests in their pores. Detailed information on the structures of and interactions in the host−guest complexes are a prerequisite for an optimal choice of a MOF for a specific guest and application. Moreover, the infinite variability of chemical composition and topology of the frameworks calls for reliable yet facile modeling approaches in place of time-consuming and often trial-and-error experimentation. The multilevel approach (Figure 5) addresses this need.

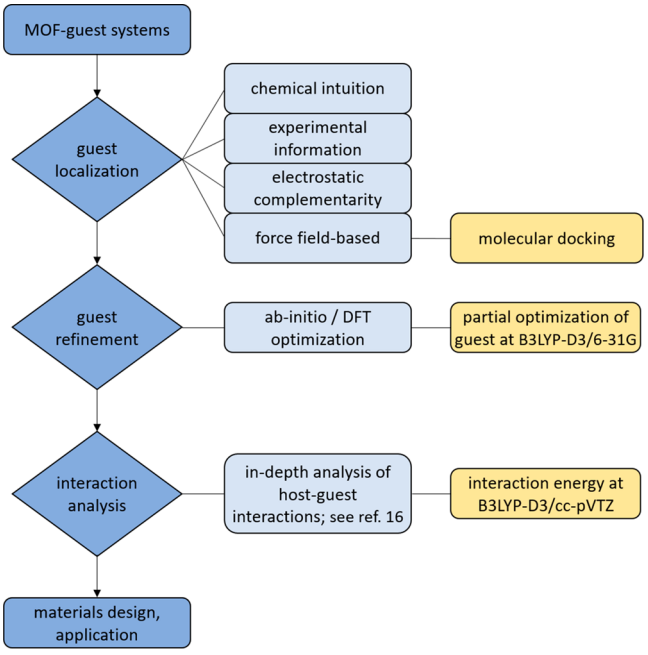


Figure 5. Flowchart illustrating the computational location and analysis of guests in MOFs. In yellow is the approach adopted within this work.

Six types of MOF−guest complexes, for which the experimental structures have already been reported, were used to evaluate this methodology. These systems feature various types of host−guest interactions, including hydrogen bonds and dispersion-dominated interactions, pores, and guests of different sizes, as well as several guests and adsorption sites within a single system. Furthermore, we considered systems both with and without guests templating the framework during the synthesis, as well as interpenetrated networks. Our multilevel approach combining molecular docking for conformational sampling with subsequent geometry and energy refinement using density functional theory and finite cluster models correctly located guest molecules in most of the studied cases.

To date, the vast majority of computational studies focused on gas molecules in MOF pores and typically used a single level of theory, with some exceptions.[15,38] The methodology,

https://doi.org/10.1021/acs.jpcc.2c05561
J. Phys. Chem. C 2023, 127, 523−531

presented here, affords rapid prescreening of adsorption geometries and accurate prediction of the interaction energies for medium-sized organic molecule guests. Advantageously, the accuracy of the computed properties can be systematically improved by using more and more advanced levels of electronic structure theory. While the force fields, common in molecular docking for the sampling of possible adsorption/binding sites in biomolecules or on surfaces,[39,40] are lacking in their description of metals, here we have demonstrated that by simply adding the necessary parameters they can be successfully applied to MOF−guest complexes. Admittedly, the manual selection of the finite cluster models hampers high-throughput automatization, necessitating further development of the automatic definition of pore cluster models. Once such tools are tested, the presented methodology can be incorporated into a computational workflow for a high-throughput screening of new frameworks and facile assessment of the host−guest complementarity.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcc.2c05561.

Details of X-ray diffraction (CCDC 2157164), molecular docking, and coordinates for all pore models (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Michelle Ernst − *Computational Carbon Chemistry Group, Heidelberg Institute for Theoretical Studies (HITS gGmbH), 69118 Heidelberg, Germany; Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany*; Email: michelle.ernst@h-its.org

Ganna Gryn'ova − *Computational Carbon Chemistry Group, Heidelberg Institute for Theoretical Studies (HITS gGmbH), 69118 Heidelberg, Germany; Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany*; orcid.org/0000-0003-4229-939X; Email: ganna.grynova@h-its.org

Authors

Tomasz Poręba − *European Synchrotron Radiation Facility, 38000 Grenoble, France*; orcid.org/0000-0002-0383-4639

Lars Gnägi − *Institute of Organic Chemistry, RWTH Aachen University, 52074 Aachen, Germany*; orcid.org/0000-0003-4610-7252

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcc.2c05561

Author Contributions

M.E.: Conceptualization, formal analysis, methodology, visualization, funding acquisition, writing of original draft, review and editing of manuscript. T.P.: Crystallographic investigation, writing of original draft, review and editing of manuscript. L.G.: Synthesis, writing of original draft, review and editing of manuscript. G.G.: conceptualization, funding acquisition, resources, supervision, review and editing of manuscript.

Funding

M.E. and G.G. acknowledge support from the Klaus Tschira Foundation. M.E acknowledges support from the Swiss National Science Foundation (Project P2BEP2_195236). The authors thank the state of Baden-Württemberg through bwHPC (bwForCluster JUSTUS2) for granting access to computational resources.

Notes

The authors declare no competing financial interest.

All input and output files were deposited on the MaterialsCloud archive (10.24435/materialscloud:zq-3t).

■ ACKNOWLEDGMENTS

M.E. gratefully acknowledges discussions with Dr. Ariane Nunes Alves and Manuel Glaser.

■ REFERENCES

(1) Connolly, B. M.; Madden, D. G.; Wheatley, A. E. H.; Fairen-Jimenez, D. Shaping the Future of Fuel: Monolithic Metal-Organic Frameworks for High-Density Gas Storage. *J. Am. Chem. Soc.* **2020**, *142*, 8541−8549.

(2) Li, J. R.; Sculley, J.; Zhou, H. C. Metal-Organic Frameworks for Separations. *Chem. Rev.* **2012**, *112*, 869−932.

(3) Yi, F. Y.; Chen, D.; Wu, M. K.; Han, L.; Jiang, H. L. Chemical Sensors Based on Metal−Organic Frameworks. *ChemPlusChem* **2016**, *81*, 675−690.

(4) Dhakshinamoorthy, A.; Li, Z.; Garcia, H. Catalysis and Photocatalysis by Metal Organic Frameworks. *Chem. Soc. Rev.* **2018**, *47*, 8134−8172.

(5) Horcajada, P.; Gref, R.; Baati, T.; Allan, P. K.; Maurin, G.; Couvreur, P.; Férey, G.; Morris, R. E.; Serre, C. Metal−Organic Frameworks in Biomedicine. *Chem. Rev.* **2012**, *112*, 1232−1268.

(6) Li, Q.; Zhang, W.; Miljanić, O. Š.; Sue, C.-H.; Zhao, Y.-L.; Liu, L.; Knobler, C. B.; Stoddart, J. F.; Yaghi, O. M. Docking in Metal-Organic Frameworks. *Science* **2009**, *325* (5942), 855−859.

(7) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Identifying Promising Metal−Organic Frameworks for Heterogeneous Catalysis via High-Throughput Periodic Density Functional Theory. *J. Comput. Chem.* **2019**, *40*, 1305−1318.

(8) Mancuso, J. L.; Mroz, A. M.; Le, K. N.; Hendon, C. H. Electronic Structure Modeling of Metal-Organic Frameworks. *Chem. Rev.* **2020**, *120*, 8641−8715.

(9) Rodrigues, M. O.; de Paula, M. V.; Wanderley, K. A.; Vasconcelos, I. B.; Alves, S.; Soares, T. A. Metal Organic Frameworks for Drug Delivery and Environmental Remediation: A Molecular Docking Approach. *Int. J. Quantum Chem.* **2012**, *112*, 3346−3355.

(10) Panda, J.; Sahoo, J. K.; Panda, P. K.; Sahu, S. N.; Samal, M.; Pattanayak, S. K.; Sahu, R. Adsorptive Behavior of Zeolitic Imidazolate Framework-8 towards Anionic Dye in Aqueous Media: Combined Experimental and Molecular Docking Study. *J. Mol. Liq.* **2019**, *278*, 536−545.

(11) Korb, O.; Wood, P. A. Prediction of Framework−Guest Systems Using Molecular Docking. *Chem. Commun.* **2010**, *46*, 3318−3320.

(12) Lu, N.; Wang, T.; Zhao, P.; Zhang, L.; Lun, X.; Zhang, X.; Hou, X. Experimental and Molecular Docking Investigation on Metal-Organic Framework MIL-101(Cr) as a Sorbent for Vortex Assisted Dispersive Micro-Solid-Phase Extraction of Trace 5-Nitroimidazole Residues in Environmental Water Samples Prior to UPLC-MS/MS Analysis. *Anal. Bioanal. Chem.* **2016**, *408*, 8515−8528.

(13) Ahmad Isiyaka, H.; Jumbri, K.; Soraya Sambudi, N.; Uba Zango, Z.; Ain Fathihah Binti Abdullah, N.; Saad, B. Effective Adsorption of Metolachlor Herbicide by MIL-53(Al) Metal-Organic Framework: Optimization, Validation and Molecular Docking Simulation Studies. *Environ. Nanotechnol., Monit. Manage.* **2022**, *18*, 100663.

(14) Zango, Z. U.; Sambudi, N. S.; Jumbri, K.; Abu Bakar, N. H. H.; Abdullah, N. A. F.; Negim, E. S. M.; Saad, B. Experimental and Molecular Docking Model Studies for the Adsorption of Polycyclic Aromatic Hydrocarbons onto UiO-66(Zr) and NH2-UiO-66(Zr) Metal-Organic Frameworks. *Chem. Eng. Sci.* **2020**, *220*, 115608.

(15) Cui, P.; Ma, Y. G.; Li, H. H.; Zhao, B.; Li, J. R.; Cheng, P.; Balbuena, P. B.; Zhou, H. C. Multipoint Interactions Enhanced CO2 Uptake: A Zeolite-like Zinc-Tetrazole Framework with 24-Nuclear Zinc Cages. *J. Am. Chem. Soc.* **2012**, *134*, 18892−18895.
(16) Ernst, M.; Gryn'ova, G. Strength and Nature of Host-Guest Interactions in Metal-Organic Frameworks from a Quantum-Chemical Perspective. *ChemPhysChem* **2022**, *23*, 202200098.
(17) Schaate, A.; Klingelhöfer, S.; Behrens, P.; Wiebcke, M. Two Zinc(II) Coordination Polymers Constructed with Rigid 1,4-Benzenedicarboxylate and Flexible 1,4-Bis(Imidazol-1-Ylmethyl)-2,3,5,6-Tetramethylbenzene Linkers: From Interpenetrating Layers to Templated 3D Frameworks. *Cryst. Growth Des.* **2008**, *8*, 3200−3205.
(18) Horike, S.; Bureekaew, S.; Kitagawa, S. Coordination Pillared-Layer Type Compounds Having Pore Surface Functionalization by Anionic Sulfonate Groups. *Chem. Commun.* **2008**, *8*, 471−473.
(19) de Lill, D. T.; Bozzuto, D. J.; Cahill, C. L. Templated Metal−Organic Frameworks: Synthesis, Structures, Thermal Properties and Solid-State Transformation of Two Novel Calcium−Adipate Frameworks. *Dalt. Trans.* **2005**, *12*, 2111−2115.
(20) Quah, H. S.; Chen, W.; Schreyer, M. K.; Yang, H.; Wong, M. W.; Ji, W.; Vittal, J. J. Multiphoton Harvesting Metal−Organic Frameworks. *Nat. Commun.* **2015**, *6*, 7954.
(21) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785−2791.
(22) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(23) 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.; et al.*Gaussian 16*, revision C.01.; Gaussian, Inc.: Wallingford, CT, 2016.
(24) DeVries, L. D.; Barron, P. M.; Hurley, E. P.; Hu, C.; Choe, W. "Nanoscale Lattice Fence" in a Metal−Organic Framework: Interplay between Hinged Topology and Highly Anisotropic Thermal Response. *J. Am. Chem. Soc.* **2011**, *133*, 14848−14851.
(25) Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalleri, B.; Maschio, L.; Rérat, M.; Casassa, S.; Baima, J.; Salustro, S.; Kirtman, B. Quantum-Mechanical Condensed Matter Simulations with CRYSTAL. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1360.
(26) Seidel, R. W.; Oppel, I. M. 1D and 2D Solid-State Metallosupramolecular Arrays of Freebase 5,10,15,20-Tetra(4-Pyridyl)Porphyrin, Peripherally Linked by Zinc and Manganese Ions. *Struct. Chem.* **2009**, *20*, 121−128.
(27) Krishnamohan Sharma, C. V.; Broker, G. A.; Rogers, R. D. Polymorphous One-Dimensional Tetrapyridylporphyrin Coordination Polymers Which Structurally Mimic Aryl Stacking Interactions. *J. Solid State Chem.* **2000**, *152*, 253−260.
(28) Sharma, C. V. K.; Broker, G. A.; Huddleston, J. G.; Baldwin, J. W.; Metzger, R. M.; Rogers, R. D. Design Strategies for Solid-State Supramolecular Arrays Containing Both Mixed-Metalated and Freebase Porphyrins. *J. Am. Chem. Soc.* **1999**, *121*, 1137−1144.
(29) Zheng, N.; Zhang, J.; Bu, X.; Feng, P. Cadmium−Porphyrin Coordination Networks: Rich Coordination Modes and Three-Dimensional Four-Connected CdSO 4 and (3,5)-Connected Hms Nets. *Cryst. Growth Des.* **2007**, *7*, 2576−2581.
(30) Flanagan, K. J.; Twamley, B.; Senge, M. O. Investigating the Impact of Conformational Molecular Engineering on the Crystal Packing of Cavity Forming Porphyrins. *Inorg. Chem.* **2019**, *58*, 15769−15787.
(31) Ochsenbein, P.; Ayougou, K.; Mandon, D.; Fischer, J.; Weiss, R.; Austin, R. N.; Jayaraj, K.; Gold, A.; Terner, J.; Fajer, J. Conformational Effects on the Redox Potentials of Tetraarylporphyrins Halogenated at The*β*-Pyrrole Positions. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 348−350.
(32) Inokuma, Y.; Yoshioka, S.; Ariyoshi, J.; Arai, T.; Hitora, Y.; Takada, K.; Matsunaga, S.; Rissanen, K.; Fujita, M. X-Ray Analysis on the Nanogram to Microgram Scale Using Porous Complexes. *Nature* **2013**, *495*, 461−466.
(33) Ramadhar, T. R.; Zheng, S.-L.; Chen, Y.-S.; Clardy, J. Analysis of Rapidly Synthesized Guest-Filled Porous Complexes with Synchrotron Radiation: Practical Guidelines for the Crystalline Sponge Method. *Acta Crystallogr. Sect. A Found. Adv.* **2015**, *71*, 46−58.
(34) Horcajada, P.; Serre, C.; Maurin, G.; Ramsahye, N. A.; Balas, F.; Vallet-Regí, M.; Sebban, M.; Taulelle, F.; Férey, G. Flexible Porous Metal-Organic Frameworks for a Controlled Drug Delivery. *J. Am. Chem. Soc.* **2008**, *130*, 6774−6780.
(35) Delle Piane, M.; Corno, M.; Pedone, A.; Dovesi, R.; Ugliengo, P. Large-Scale B3LYP Simulations of Ibuprofen Adsorbed in MCM-41 Mesoporous Silica as Drug Delivery System. *J. Phys. Chem. C* **2014**, *118*, 26737−26749.
(36) Vieira Soares, C.; Maurin, G.; Leitaõ, A. A. Computational Exploration of the Catalytic Degradation of Sarin and Its Simulants by a Titanium Metal-Organic Framework. *J. Phys. Chem. C* **2019**, *123*, 19077−19086.
(37) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(38) Sillar, K.; Kundu, A.; Sauer, J. Ab Initio Adsorption Isotherms for Molecules with Lateral Interactions: CO2 in Metal-Organic Frameworks. *J. Phys. Chem. C* **2017**, *121*, 12789−12799.
(39) Nunes-Alves, A.; Ormersbach, F.; Wade, R. C. Prediction of the Drug−Target Binding Kinetics for Flexible Proteins by Comparative Binding Energy Analysis. *J. Chem. Inf. Model.* **2021**, *61*, 3708−3721.
(40) Li, M.; Li, Q.; Xu, M.; Liu, B.; Calatayud, D. G.; Wang, L.; Hu, Z.; James, T. D.; Mao, B. Amphiphilic Engineering of Reduced Graphene Oxides Using a Carbon Nitride Coating for Superior Removal of Organic Pollutants from Wastewater. *Carbon N. Y.* **2021**, *184*, 479−491.

https://doi.org/10.1021/acs.jpcc.2c05561
J. Phys. Chem. C 2023, 127, 523−531