metallicious: Automated force-field parametrization of covalently bound metals for supramolecular structures

Tomasz K. Piskorz,[1] Bernadette Lee,[1] Shaoqi Zhan,[1,2] Fernanda Duarte[1]*

[1]Department of Chemistry, University of Oxford, Oxford OX1 3QZ, UK.

[2]Department of Chemistry - Ångström, Ångströmlaboratoriet Box 523, S-751 20 Uppsala, Sweden

E-mail: fernanda.duartegonzalez@chem.ox.ac.uk

KEYWORDS metal ions; force-field parametrization; molecular dynamics; supramolecular chemistry; metal-organic frameworks

ABSTRACT

Metal ions play a central functional and structural role in many molecular structures, from small catalysts to metal-organic frameworks (MOFs) and proteins. Computational studies of these systems typically employ classical or quantum mechanical approaches, or a combination of both. Among classical models, only the covalent metal model reproduces both geometries and charge transfer effects but requires time-consuming parametrization, especially for supramolecular systems containing repetitive units. To streamline this process, we introduce *metallicious*, a Python tool designed for efficient force-field parametrization of supramolecular structures. *metallicious* has been tested on diverse systems, including supramolecular cages, knots, and MOFs. Our benchmarks demonstrate that parameters obtained from *metallicious* accurately reproduce the reference properties obtained from quantum calculations and crystal structures. MD simulations of the generated structures consistently yield stable simulations in explicit solvent, in contrast to similar simulations performed with non-bonded and cationic dummy models. Overall, *metallicious* facilitates the setup of molecular dynamics (MD) of supramolecular systems simulations, providing insights into their dynamic properties and host-guest interactions. The tool is freely available on GitHub (https://github.com/duartegroup/metallicious)

1 Introduction

Metal ions play a significant role in chemistry, biology, and material science. Approximately one-third of the proteins in the Protein Data Bank contain metals, serving essential structural and catalytic functions.[1] Metal ions are also key building blocks in supramolecular chemistry, enabling the formation of complex metallo-organic cage structures[2–5] and metal-organic frameworks (MOFs).[6,7] Their unique properties stem from their strong directional interactions and coordination patterns, which are unavailable in carbon-based chemistry. Supramolecular structures often feature repeated metal binding sites within their architecture and even across different systems (**Figure 1a**). For instance, the majority of metal sites in palladium-based supramolecular cages are derivatives of the tetrakis(pyridine)palladium(II) building block.[8]

Various methodologies have been used to model metal-containing systems, including quantum mechanics (QM),[9–12] molecular mechanics (MM) such as molecular dynamics (MD) and Monte Carlo (MC) simulations,[12] and hybrid QM/MM.[11,12] These techniques have been applied to model metalloproteins,[13,14] metallo-organic cages,[15,16] and MOFs.[17–19] MM approaches, in particular, offer a significant advantage over QM methods in terms of computational costs.[12] MM based approaches for deriving force-field parameters for metal ions, include the non-bonded,[20,21,30–32,22–29] cationic dummy,[33–42] and covalent models (**Figure 1b**).[43–47] Non-bonded models describe metal centers as van der Waals spheres with integer charge, describing ligand-metal interactions through electrostatic Coulombic and Lennard-Jones (LJ) terms. Li and Merz have reported parameter sets for 56 metals, aiming to reproduce hydration-free energy (HFE),

coordination number (CN), ion-oxygen distances (IOD) of aqua complexes. However, achieving this often involves identifying parameters that compromise accuracy for these specific observables, since no parameter set can reproduce all of them.[20–25] Zhang sampled a larger LJ parameter space to identify parameters capable of reproducing HFE and CN for a set of 47 ions.[26–29]

The dummy model, originally developed by Åqvist and Warshel,[30,48] places cationic dummy particles bonded to metal in a predefined coordination geometry. These particles exclusively interact with ligands via electrostatic interactions. This model provides an improvement over non-bonded models by simultaneously reproducing HFE and CN. However, its predefined configuration reduces flexibility, and only a limited number of metal ions have been parametrized.[33–40,42] While both non-bonded and cationic dummy models can describe bond forming and breaking processes,[40,41,49,50] they often exhibit unexpected behavior when transferred to different systems[33,51] and fail to account for charge transfer.[41,52,53]

The covalent bond model explicitly describes bonds, angles, torsions, van der Waals and Coulombic interactions between coordinating ligands and metal ions, allowing it to account for charge transfer effects but unable to account for ligand exchange. Parametrizing the different terms is time-consuming, often relying on equilibrium bonded parameters and force constants obtained from QM calculations. The two most common approaches for obtaining bond force constants involve calculating the Hessian matrix[47,54–57] or employing forces-matching techniques.[58–61] Among Hessian-based methods, the Seminario method is the most

widely used, involving the projection of the Hessian Matrix elements onto relevant bonded parameters (**Figure 1c**; Approach 1).[54] This procedure has been successfully applied to model small metallo-organic molecules,[62–64] metalloproteins,[65,66] MOFs,[67–69] and supramolecular cages.[70–72] In this approach, dihedral parameters involving metal ions are omitted as they usually have minimal impact on the structure.[66,73–75] Cole and co-workers have improved the Seminario method by addressing the problem of double counting bending interactions.[55] The LJ parameters are typically taken from non-bonded model parameters or the Universal Force-Field, which covers 126 elements,[76] while partial charges are obtained using the restrained electrostatic potential (RESP) method,[77] available in various software including AmberTools (via Gaussian, GAMESS-US),[78] R.E.D. server (Gaussian, GAMESS-US, Firefly),[79] psi4/resp (psi4),[80] and psiRESP (psi4).[81]

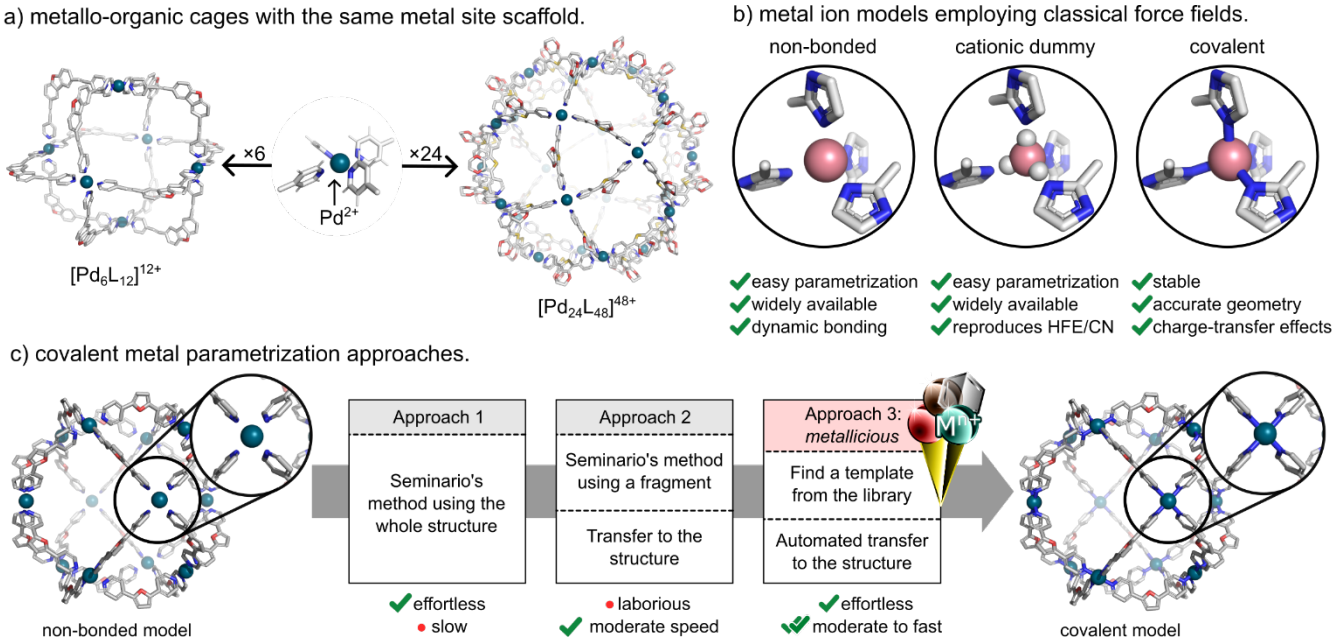


Figure 1. Force-field parametrization of supramolecular systems. (a) Supramolecular metallocages often exhibit repeated metal binding.[82–85] (b) Approaches for deriving metal force-field parameters include non-bonded, cationic dummy, and covalent models. (c) Methodologies for parametrization of a covalent metal model.

The Seminario method and RESP calculations are computationally costly, restricting their use to small systems. For larger structures, such as metalloproteins, a cluster representing the metal binding site is often used for parametrization (**Figure 1c**; Approach 2). These parameters are then transferred to the larger system, usually involving significant manual intervention. To speed up this process, Li and Merz developed MCPB.py,[86] offering a semi-automatic approach to derive bonded parameters and charges. MCPB.py is compatible with the AMBER force-field, and while it is suitable for proteins with a few distinct metal sites, its use becomes laborious for supramolecular structures, as they often consist of multiple identical metal sites. Due to differences in atom order, each site would require separate QM calculations, resulting in a time-consuming procedure. Ideally, an approach similar to those available for protein parametrization (*e.g., pdb2gmx* in GROMACS and *LeAP* in AMBER), where residue parameters are tabulated and automatically assigned, tailored to metal sites and neighboring residues would significantly streamline the modelling of supramolecular structures containing repeating metal site units (**Figure 1a**).

In this work, we introduce *metallicious* (**Figure 1c**; Approach 3), a tool designed to streamline the parametrization of metal centers using the covalent metal model based on a *template library*. Similar to MCPB.py, it enables the parametrization of new templates using Seminario's

method and RESP charges but with increased efficiency and applicability to systems with repetitive metal units, including supramolecular cages, MOFs and knots.

2. Methodology

In the following paragraphs, we explain the structure and functionalities of *metallicious*.

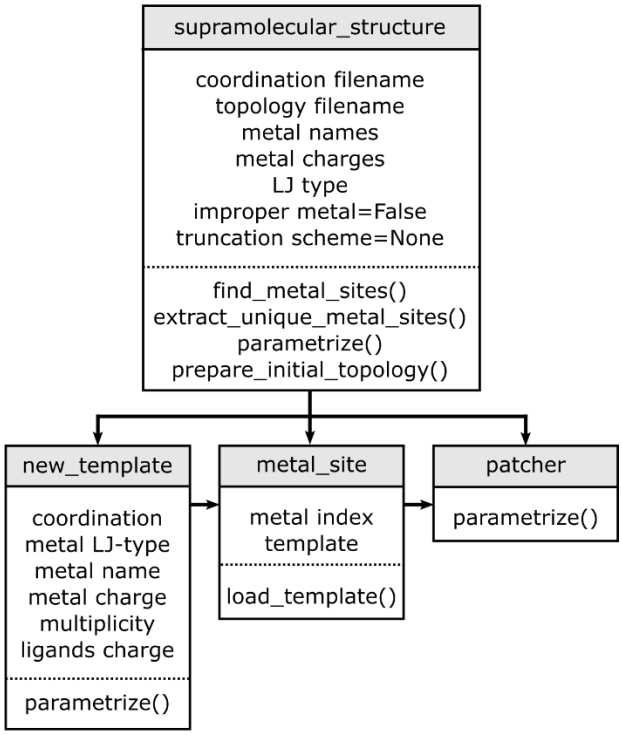
2.1 Code Structure.

metallicious is written in object-oriented Python and uses scientific programming modules, including Numpy,[87] MDAnalysis,[88] RDKit,[89] ParmEd[90] and NetworkX[91] for core functionalities. Optionally, the parametrization of new templates requires the QM engine ORCA,[92,93] autodE (used as an interface for QM engine),[94] and psiRESP,[81] and parametrization of the organic molecules requires AmberTools.[95] The code is built around four key objects (*supramolecular_structure, new_template, metal_site, and patcher*). The relationship between these objects is outlined in **Figure 2b**.

The *supramolecular structure* class serves as the central element within the module, holding information about the overall code structure. It encompasses the subclasses *new_template*, *metal_site*, and *patcher*, as well as their associated coordination and topology files. Core functionalities include extraction of metal sites from the *coordinate file*, identification of suitable template from the database and execution of template parametrization. For each metal site identified in the input, a *metal_site* class is generated, storing the metal index and the corresponding template. Pre-parametrized templates are stored in a dedicated subdirectory

within *metallicious* (in Gromacs topology format). In cases where the template is absent, a *new_template* class is created, containing all information necessary to parametrize the new metal site (truncated coordinate file, metal name, charge, multiplicity, LJ parameters, and ligands charges). Lastly, the *patcher* class combines all the *metal_site* classes with input structure and topology to produce the two output files.

a) class structure



b) usage example

```
metallicious -f cage.pdb -p cage.top -lj_type uff
                  -metal_and_charges Pd 2
```

```
from metallicious import supramolecular_structure
cage = supramolecular_structure(f='cage.pdb',
                        p='cage.top', lj_type='uff',
                        metal_and_charges = {'Pd':2})
cage.parametrize()
```

Figure 2. ***metallicious* code structure**. (a) Class structure of *metallicious*, including its four main objects and their relationship. (b) *metallicious* can be executed either via a command line or Python script. Input requirements include coordinate (.pdb), topology (.top), LJ type, and identity and charge of the metal ion.

2.2 Code usage.

metallicious can be executed either *via* a command line or imported as a Python script (**Figure 2a**). The tool requires the following input information:

(a) *Coordinate file (f).* Accepted formats include .xyz, .pdb, .gro, and others supported by MDAnalysis.

(b) *Topology file* *(p).* Containing parametrized organic molecules and metal. Accepted formats include .top, .itp, .prmtop and other files supported by ParmEd.[90] Note that the LJ parameters and charge present in the topology files are not used, only their indices. Users can either input the force-field parameters for the organic ligands (obtained from tools such as ambertools,[95] ATB,[62] or charmm-gui[96]), or request them by specifying the *prepare_initial_topology* argument, which interfaces with antechamber to obtain General AMBER Force Field (GAFF) parameters.

(c) *LJ type (LJ_type).* Specify the name of the library used to extract the metal LJ parameters. Options include Merz, Zhang or UFF libraries.

(d) *Metal name(s) and charge(s) (metal_and_charge,* and *optionally multiplicity.* This information is used to identify the metal ions in the input structure and the template to be used. Multiplicity is required only for the parametrization of new templates.

For new templates, users can also include parametrization of improper dihedral formed by metal and bound aromatic molecules, which is turned off by default (See SI §S4.1). This can be done by specifying the *improper_metal* option. Additionally, several truncation schemes of templates are available to simplify parametrization, which are also turned off by default. A full description of the input is provided in **Table S1**.

Upon completion of the parametrization, two output files are generated: a coordination file with reordered atoms (default: out.pdb) and a topology file saved in GROMACS format (default: out.top), although other formats are supported *via* ParmEd. The topology file contains modified bonds, angles, dihedrals, and partial charges of metal and its surrounding atoms and modified LJ parameters of metals.

2.3 Code functionality

As illustrated in **Figure 3**, *metallicious* iterates through each metal center in the input structure, executing one of the following actions:

(a) **Parametrization using template library.** Selects a suitable template from a repository of predefined templates and adjust the input topology accordingly.

(b) **Parametrization of new template.** If a suitable template is unavailable, it performs template parametrization.

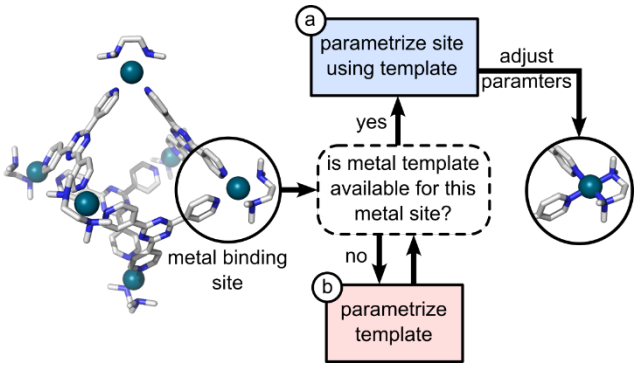


Figure 3. Overview of metal parametrization in *metallicious*, illustrating the communication between the two primary program subroutines: parametrization site using a template and a (new) template parametrization.

(a) Parametrization using template library. The most efficient approach is to use readily available templates. This subroutine is always performed for each site and can be summarized in the following key steps (Figure 4a):

(a.i) Template selection and (optionally) truncation: This search aims to identify filenames that match the user-specified metal name, charge, and LJ type. The first matching template is used for the next step. If unavailable, users can truncate the system and perform the search again, or parametrize a new template. Truncation involves applying a distance cut-off based on the user-defined truncation_scheme variable (default: None for which the whole template structure is considered). Depending on the scheme, atoms more than 1, 2 or 3-bonds away from the metal center are removed, and their charges are evenly redistributed among the remaining atoms. The influence of different truncation schemes on the final parameters is discussed in § **3.3**.

(a.ii) Reduction of the metal site: The original metal site is aligned with a selected template, chosen after matching the name, charge, and LJ type through an iterative process. A distance cut-off equal to the template's radius plus 2Å is used (**Figure 4**, step (ii), blue dashed line). The template's radius is defined as the distance between the metal and its furthest atom (**Figure 4 (i)**, red dashed line). Matching ligands in both structures are identified by constructing their molecular graphs and locating combinations where the template's linker graphs are subgraphs of the metal site's graphs. The metal site's linkers are then reduced to match template graphs, resulting in identical structures that only differ in the order of the atoms.

(a.iii) Mapping of the template onto the metal site: Achieving the correct atom order involves an exhaustive exploration of all possible permutations of ligand and atom numbering, the most time-consuming part of the process. For efficiency, the process is divided into two stages. Firstly, the order of all non-hydrogen atoms is determined by exhaustively exploring permutations of the reduced metal site and selecting the one with the lowest root mean square displacement (RMSD) with the template with reordered atoms (**Figure S1**). Subsequently, hydrogen atoms are reconstructed by finding an isomorphism between the template's and the metal site's linkers graphs, constraining the order of heavy atoms. Successful mapping is defined as RMSD < 2Å. If RMSD > 2 Å, the next template from the library is evaluated; if unavailable, the procedure for parametrizing new template is followed (**Figure 5**).

(a.iv) Adjusting input topology: If the steps above are successful, the structure's topology is updated by substituting its bonded parameters with those from the reordered template. This

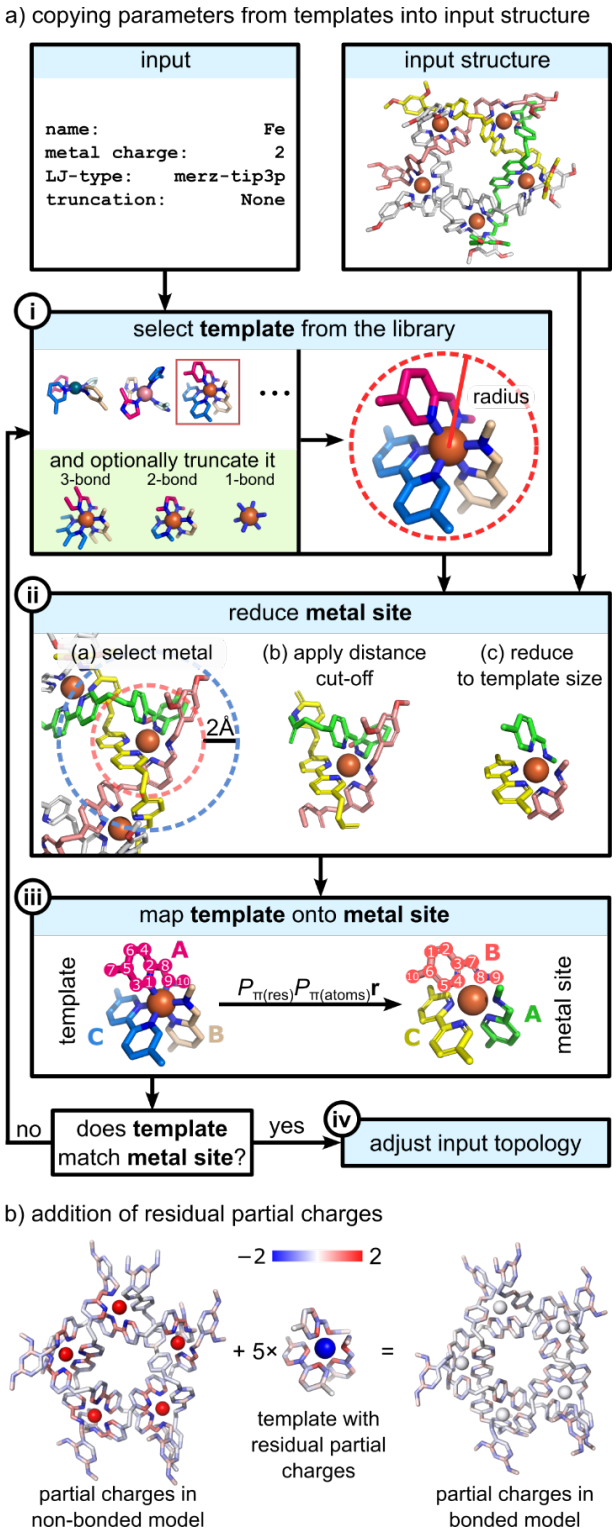


Figure 4. Parametrization workflow using templates (a) Copying parameters from a template into the input coordination and topology files. (b) Partial charges accounting for charge transfer are obtained by summing partial charges from non-bonded model and template's residual partial charges.

topology is further updated by adding residual partial charges from the template, which account for charge transfer effects (**Figure 4b**; for comparison with RESP charges see **Figure S2**). These partial charges are obtained from the differences in charges between the complex and separate species (*vide infra*, **Figure 5b**). Since the sum of residual charges of the template is zero, the result of its addition to the new topology only results in the redistribution of partial charges.

(b) Parametrization of a new template. When no template is found in the library, *metallicious* performs parametrization using the following steps (**Figure 5a**):

(b.i) Extraction of a new template: To extract the new template, *metallicious* iterates through metal atoms in the input file. It selects atoms that are either (a) within a distance of 3 bonds from the metal or (b) part of a connected aromatic group(s) coordinated with the metal. The selected atoms from both categories are combined to form the template's backbone. Atoms that are not part of the template's backbone but directly connected to it by one bond distance are added to the structure, resulting in the *template*, which is the main structure for parametrization using the template library. Hydrogen atoms are added to atoms with unfulfilled valance, saturate the template needed for QM calculations (*saturate template*).

(b.ii) Parametrization of bonded parameters. Bonds and angles are parametrized using Seminario's method, adapted from Cole and co-workers.[55] This uses the Hessian matrix computed via ORCA/autodE for the *saturated template*. Equivalent bonds and angles are symmetrized by comparing molecular graphs of metal-ligand pairs (**Figure S3**). Similarly to

the work of others, proper dihedrals are not parametrized.[66,73–75] Parameters for improper dihedral are obtained by performing a 1D scan involving the metal ion and the donor atoms. We found that including improper parameters improves the geometry of some metal sites (*vide infra*, Section **3.2**).

(b.iii) Residual partial charge. Partial charges are calculated using the RESP method.[77] During this process, the electrostatic potential (ESP) is computed at the D3BJ-PBE0/def2-SVP level of theory using ORCA for the individual ligands and the saturated template separately. This methodology was selected based on the wide availability of the basis set for most elements of the periodic table and the robustness and efficiency of the hybrid PBE0 functional.[97–100] Moreover, this level of theory was found to provide similar results to the popularly used B3LYP/6-31+G*. **Figure S2**). The total charge of the saturated template, required for initiating QM calculations, is determined by summing up the user-specified metal charge and RDKIT-generated ligand charges. Subsequently, the partial charges are computed using psiRESP. During this process, charges of linking atoms are constrained to zero. From these calculations, a residual partial charge is computed by subtracting the partial charges of the *saturated template* from the partial charges of individual ligands and the metal (**Figure 5b**). Lastly, the obtained charges are symmetrized for identical ligands, determined using the isomorphism of their molecular graphs.

(b.iv) Map parameters onto the template: The additional atoms added to form *the saturated template* are removed, resulting in the final *template,* which is added to the template library.

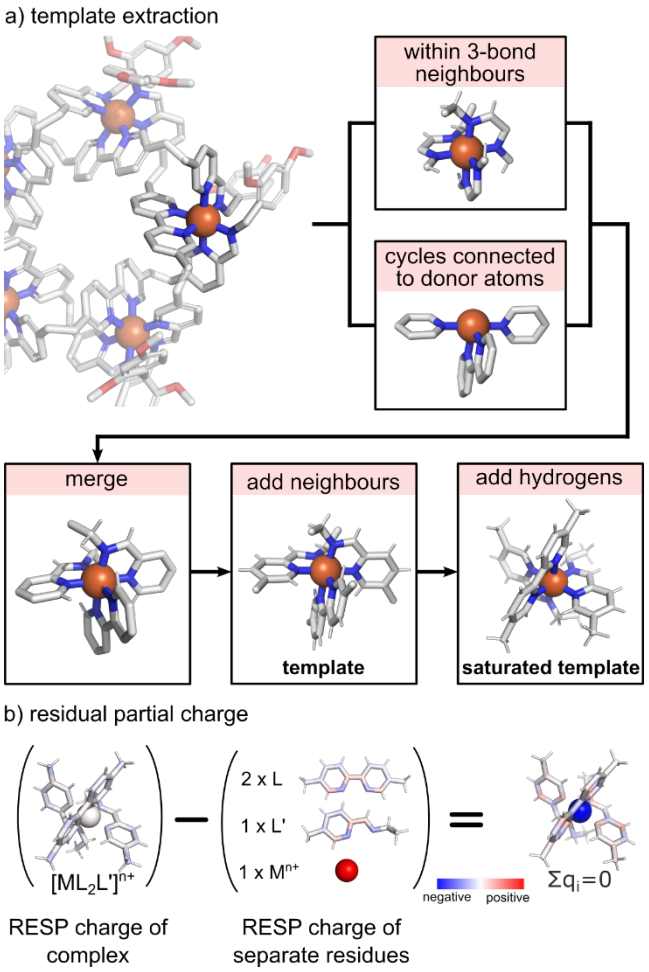


Figure 5. New template parametrization. (a) Extracting the metal template from the supramolecular complex. (b) Calculating the residual partial charge on the template.

3. Results

To assess the capabilities of *metallicious*, we considered 11 supramolecular systems whose structures were obtained from the Cambridge Crystallographic Data Centre (CCDC), including seven cages ($[Pd_2L_4]^{4+}$,[101,102] $[Ga_4L_6]^{12-}$,[103] $[Fe_4L_6]^{4-}$,[104] $[Pd_6L_4]^{12+}$,[105] $[Co_8L_{12}]^{16+}$,[106] $[Pd_6Ru_8L_{24}]^{28+}$,[107] $[Pd_{48}L_{96}]^{96+}$)[108], two knots ($[Fe_5L_5]^{10+}$,[109] $[Zn_3L_3]^{6+}$)[110] and two MOFs (ZIF-8,[111] ZIF-67[112]); **Table S2**). Among these systems, $[Pd_2L_4]^{4+}$,[113] $[Ga_4L_6]^{12-}$,[70,114] $[Pd_6L_4]^{12+}$,[71] ZIF-8[67] and ZIF-67[68] have

been previously modelled using MD simulations. The others were selected due to their technical challenges for automation. For example, $[Co_8L_{12}]^{16+}$contains eight metal sites, four of which are stereoisomers, differing only by orientation of the ligands around metal; $[Pd_6L_4]^{12+}$ comprises two different ligands, pyridine and diamine; $[Pd_6Ru_8L_{24}]^{28+}$ requires routines to handle two distinct metals and ligands. Finally, $[Pd_{48}L_{96}]^{96+}$ was chosen as it is the largest synthesized supramolecular cage to date.

The metric used to assess the quality of the parameters includes: comparing the MM and QM-computed vibrational frequencies and optimized structures for 11 small metal sites (**Figure 6(i)**; see **Figure S4** for all structures), RMSD values between MM optimized and crystal structures of 11 representative supramolecular systems (**Figure 6(ii)**), and comparison between MM and QM binding energies for benzoquinone (**bq**)-$[Pd_2L_4]^{4+}$ complex, with the MM value obtained from five different starting structures (**Figure 6(iii)**).

By default, the templates were parametrized (**§ 2.3(b)**) at the D3BJ-PBE0/def2-SVP level of theory without truncation using the Merz-OPC Lennard-Jones parameters; improper parameters were only included for systems with square planar complexes. Using the default parameters shows a good performance across the different metrics evaluated (**Figure 6**). For example, when comparing vibrational frequencies for the small metal sites, the mean percentage errors only 6.7%, similar to the 6.4% obtained by Cole and co-workers for 70 small molecules using the modified Seminario method.[55] The MM-optimized small metal sites also result in structures similar to those optimized by QM (RMSD < 0.5 Å). Similarly, good

agreement is obtained when comparing the MM-optimized and crystal structures of supramolecular systems (RMSD < 1.5 Å). Lastly, the binding energy of (**bq**)-$[Pd_2L_4]^{4+}$ shows excellent agreement with QM computed energies (<0.5 kcal mol^{-1}). These results indicate that the parametrization obtained through *metallicious* is robust enough to obtain not only structural but also host-guest energetics.

3.1 influence of level of theory on quality of parameters. We first assessed the impact of different functionals (D3BJ-PBE0 and ωB97X-D3) and basis sets (def2-SVP and def2-TZVP) on the quality of the parameters obtained from the Seminario method (**Figure 6a**). The results obtained for the different benchmarks were similar regardless of the functional and basis set used. Surprisingly, despite the reference values being derived from calculations at the ωB97X-D3/def2-TZVP level of theory, parameters obtained at this level of theory were not better than those obtained at the D3BJ-PBE0/def2-SVP level of theory. This observation suggests that the loss of accuracy during parametrization originates from the inadequate depiction of interactions by conventional force fields. Since D3BJ-PBE0/def2-SVP is computationally efficient, it is used as a default methodology to parametrize templates in *metallicious*.

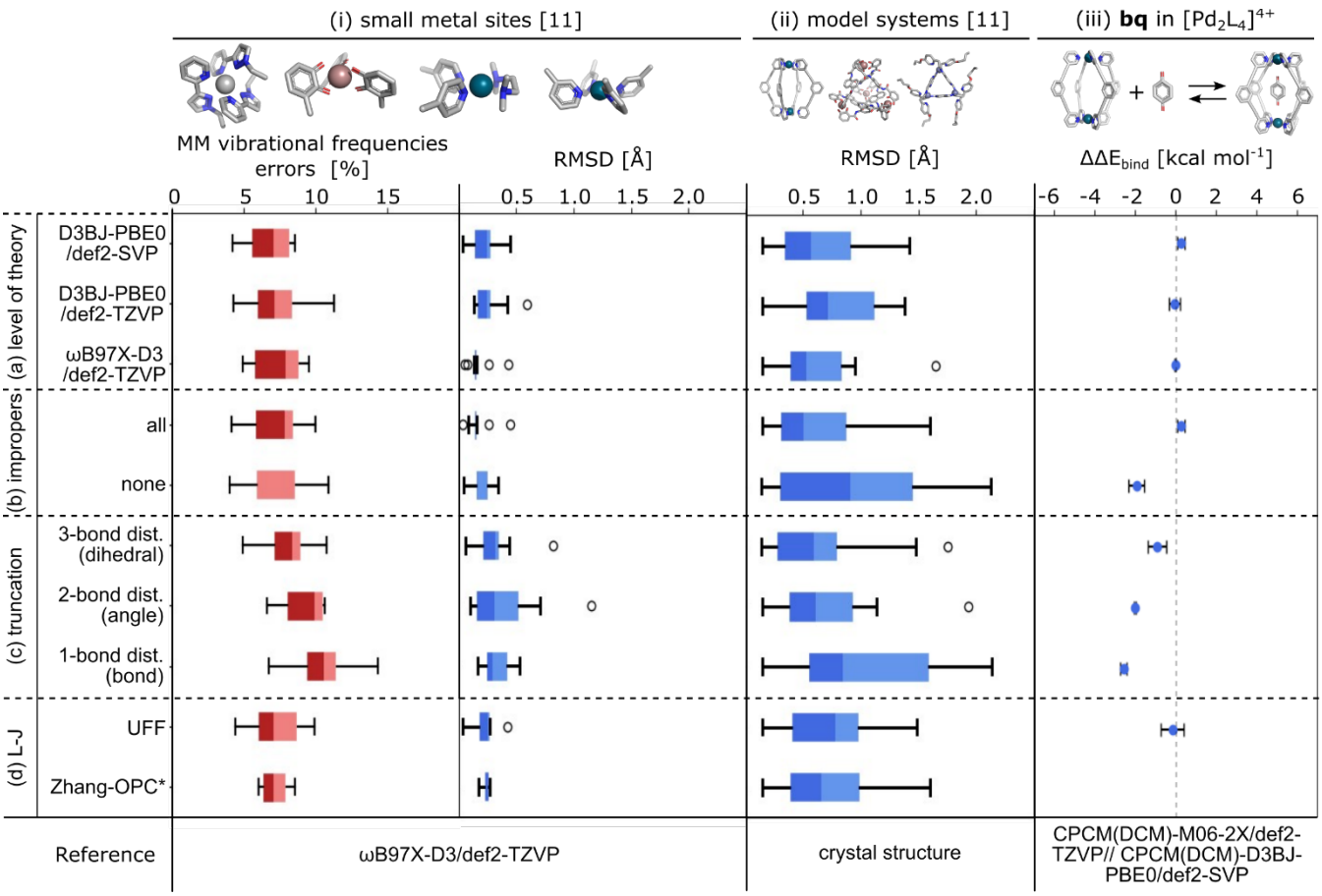


Figure 6. Benchmark of *metallicious* considering: (i) 11 small metal sites, including a comparison of MM and QM-computed vibrational frequencies and RMSD values between MM and QM optimized structures; (ii) 11 representative supramolecular systems, including comparison between MM-optimized and crystal structures; (iii) benzoquinone (BQ)-$[Pd_2L_4]^{4+}$ complex, comparing MM and QM binding energies. The benchmark considered the influence of: (a) different levels of theory on parametrization, (b) inclusion of improper parameters, (c) truncation scheme, (d) type of LJ parameters (Zhang-OPC parameters are available only for non-palladium systems). Except where specified by (a-d), the template was parametrized at the D3BJ-PBE0/def2-SVP level of theory, none truncation scheme, Merz-OPC Lennard-Jones parameters and improper parameters included only for systems with square planar complexes.

3.2. Importance of improper dihedral parameters involving metals. Proper and improper dihedral parameters involving metal centers are rarely considered as they are expected to have a negligible influence on geometry.[66,73–75] However, while testing the $[Pd_2L_4]^{4+}$ cage in implicit solvent, we observed during the MD simulation structures where linkers were stacked on top of each other (**Figure 7a**). As a result, the donor atom connected to the metal deviates from the expected planar position. This made us wonder if such an artefact was due to the omission of dihedral parameters around the metal. Indeed, including improper dihedrals centred on the donor atom yielded structures similar to those obtained from QM in implicit solvent (**Figure 7b**). Interestingly, these parameters can be omitted in explicit solvent, likely due to the dominant effect of solvent-solute interaction over solute-solute stacking interactions. For example, during 100 ns MD simulation in explicit dimethylsulfoxide (DMSO), when no metal-involved improper dihedrals are included, less than 10% of the configurations adopted a staggered configuration compared to 50% in implicit solvent (**Figure 7c**).

Optimization of all other structures with and without improper dihedrals in implicit solvent demonstrated that this issue is specific to systems featuring a square planar configuration (**Figure 6b; Figure S5**). Otherwise, including improper parameters has no impact on the geometries. For this reason, and considering the computational cost associated with obtaining these parameters, their parametrization is disabled by default in *metallicious*.

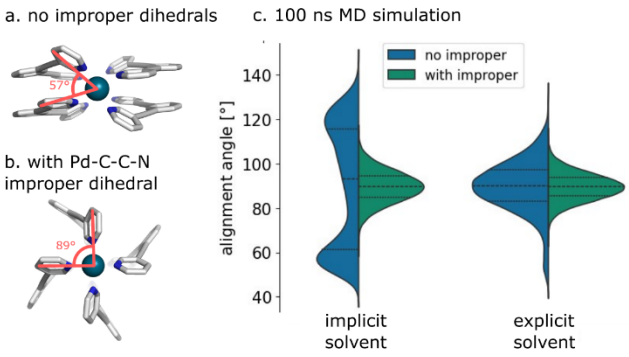


Figure 7. Energy-minimized structures using L-BFGS with GBSA implicit solvent (a) with and (b) without the improper dihedral. (c) Histogram of the angle between two ligands from 100 ns MD simulations of the $[Pd_2L_4]^{4+}$ cage with implicit and explicit DMSO solvent.

3.3. Truncation schemes. Deriving new templates in *metallicious* is fully automated but computationally intensive, especially for systems with unique metal sites like the asymmetric cage reported by Lewis *et al.* (**Figure 8**).[115]

To expedite this process, we created truncation schemes that reuse existing templates. The degree of truncation depends on the selected scheme: none, 3-, 2- and 1-bond distance from the metal center. For the example mentioned above, which currently lacks a template, *metallicious* parametrizes a new template in around ~100 CPUhs compared to only a few seconds when it uses template library by applying a 3-bond distance truncation scheme to an existing template. While truncation schemes expand the utility of the existing template library, they sacrifice accuracy and, therefore, must be used with caution. The benchmarks reveal that the accuracy loss is proportional to the extent of truncation (**Figure 6c**). It is

therefore recommended, when possible, to start parametrization with the least truncated template.

3.4. Lennard-Jones parameters. The effect of the use of different L-J parameters (Merz-OPC, UFF and Zhang-OPC) on the evaluated metrics was found to be minimal (**Figure 6d**). This is likely due to the dominance of bonded parameters.

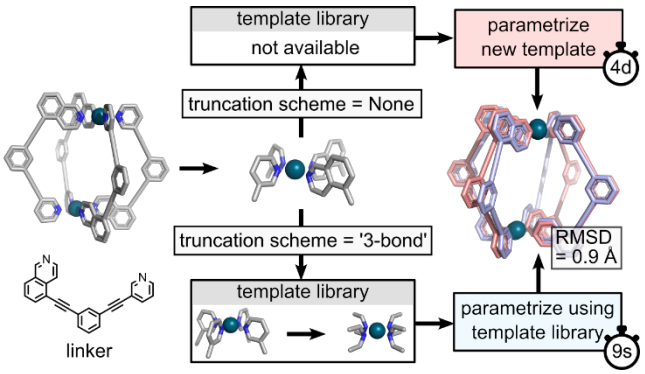


Figure 8. Truncation scheme. This scheme speeds up calculations by replacing missing templates with substructures from related templates.

3.5 Application of *metallicious*.

MD simulations provide access to the dynamics properties of supramolecular systems that cannot be obtained from crystal or QM-optimized structures. *metallicious* simplifies the setup of such simulations, facilitating the analysis of dynamic properties and their implications for host-guest interactions. To illustrate this, we parametrized and performed 100 ns each MD simulations f the 11 supramolecular systems described above (**Figure 9a-c**). Except for ZIF-8 and ZIF-67, which lack counterions and solvent molecules in their crystals, simulations were

conducted in explicit solvent and included counterions, which often strongly interact with metals.[40] We compared the results with simulations using the non-bonded[24,76] and dummy metal models,[33,40,42] evaluating stability of the simulations through changes in the metal coordination sphere.

As anticipated, structures using covalent metal models remained stable over the simulation time (**Figure 9d;** see **Figure S6-S16** for RMSD, coordination sphere analysis and snapshots of individual trajectories), with a median RMSD = 1.9 Å relative to the crystal structure. In contrast, most of the simulations utilizing non-bonded and dummy models resulted in disassembly. Only ZIF-8 and ZIF-67 MOFs produced stable simulations with these models, possibly due to the absence of competing interactions with counter ions and solvent. Removing counterions in the other systems resulted in a marginal improvement in stability. While the stability of the covalent model is expected as dissociation is not allowed, the poor performance of other models for almost all systems was surprising. This is likely the result of parameters being overfitted to reproduce aqueous complexes rather than interactions with other heteroatoms containing molecules, leading to imbalanced metal-ligand and metal-counterions interactions.

Analyzing the flexibility of the cages provides a more realistic picture of how much a cage cavity changes and adapts to different guests. Among the simulated cages, the $[Ga_4L_6]^{12-}$ cage shows the largest relative change in volume, from < 1 $Å^3$ (not grid point can be placed inside) to 267 $Å^3$ (**Figure 9e; Figure S17**). This flexibility is due to the rotatable central naphthalene

linker. Indeed, the $[Ga_4L_6]^{12-}$ cage has been found to bind a broad range of substrates, including quaternary ammonium cations with volumes ranging from 80–160 $Å^3$.[114,116]

Furthermore, analysis of the interactions between the supramolecular systems and counterions can be done by computing the corresponding radial distribution functions (RDFs; **Figure S18**). For example, the $[Pd_2L_4]^{4+}$ cage reveal a prominent peak at 4.7 Å, indicative of a conserved interaction between the cage and the triflate counterion (**Figure 9f**). Notably, Lusby and co-workers showed that this counterion exhibits strong binding affinity to the $[Pd_2L_4]^{4+}$ cavity.[102] Similar results were observed for the $[Zn_3L_3]^{6+}$ and $[Fe_5L_5]^{10+}$ knots, indicating a strong binding interaction at 5.7 Å and 7.7 Å, respectively. Indeed, the authors showed that the $[Fe_5L_5]^{10+}$ knot is capable of extracting chloride traces from solvent and glassware.[117]

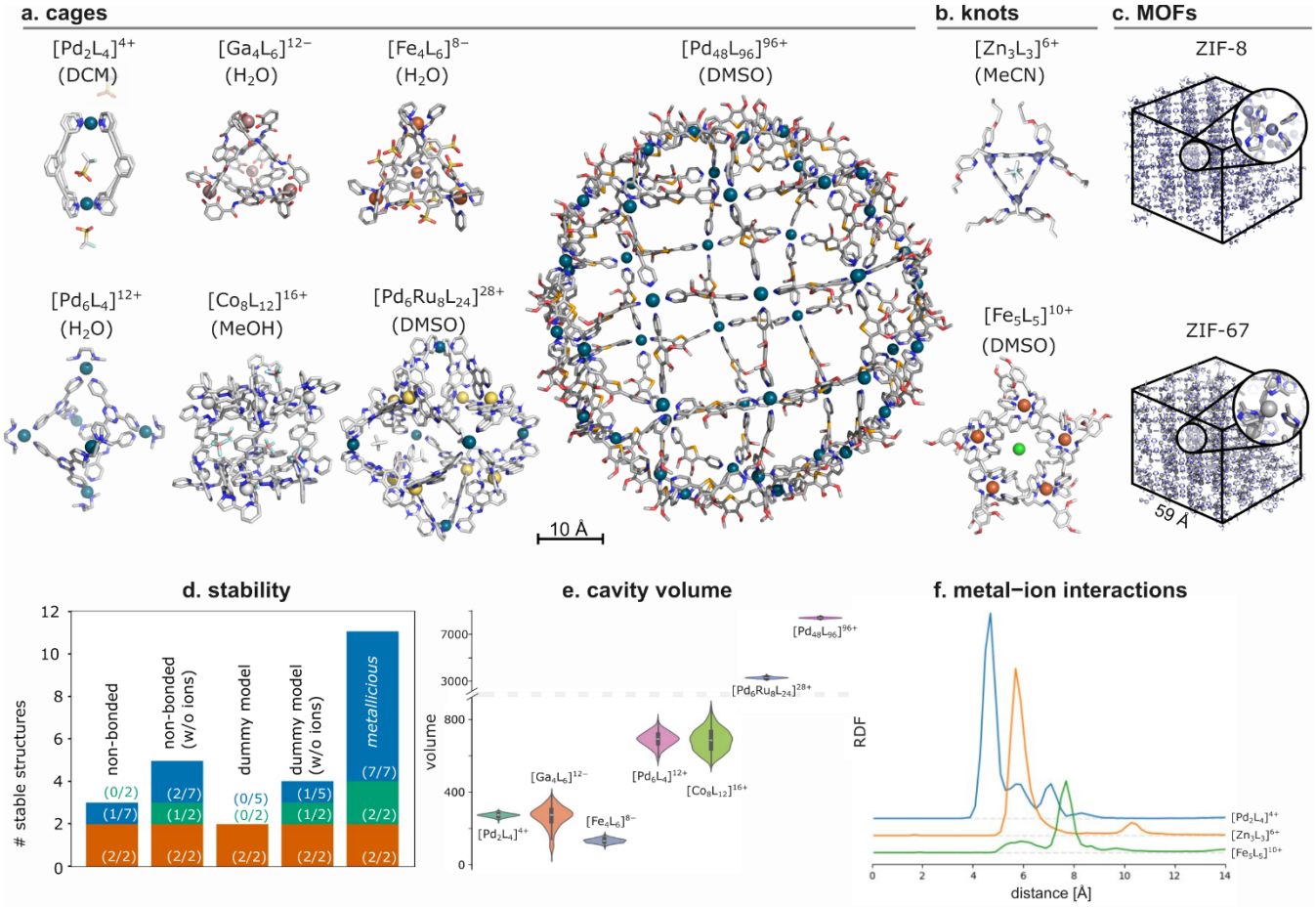


Figure 9. **Analysis from 100 ns MD simulations for systems parametrized by *metallicious*.** Final snapshots for (a) cages, (b) knots and (c) MOFs (solvent and counterions not shown for clarity, except for $[Pd_2L_4]^{4+}$, $[Zn_3L_3]^{6+}$ and $[Fe_5L_5]^{10+}$). (d) Histogram of stable structures after simulation. (e) Distribution of cavity volumes calculated using *C3*.[118] (f) RDFs calculated for metal centers and counterions (see **Table S2** and **Figure S18** for further details). DCM= dichloromethane; DMSO: dimethylsulfoxide; MeOH=methanol; MeCN=acetonitrile.

4. Conclusions

We developed an automated tool called *metallicious* for parametrizing covalent metal models in supramolecular structures. Our method leverages the repetitive patterns of binding metal motifs in these structures. Several standard templates for such structures were parametrized and stored in the library, which *metallicious* utilize to parametrize input structures by

matching a suitable template and copying its parameters. To broaden the scope of the template library and increase efficiency, *metallicious* provides convenient truncation schemes, allowing the recycling of available templates. Lastly, *metallicious* can automatically parametrize new templates if no matching one is available.

The results of the benchmarks show good agreement with reference data obtained from QM and crystal structure. While it can be argued that two other popular models, the non-bonded and cationic dummy atom model, offer more flexibility and can also account for ligand exchange, the MD simulation benchmark conclusively shows that the covalent metal model is the only robust option enabling simulations of these systems.

Supporting Information.

Implementation details. Benchmark methods and individual results.

AUTHOR INFORMATION

Corresponding Author

Fernanda Duarte – Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford OX1 3TA, United Kingdom; ORCID: https://orcid.org/0000-0002-6062-8209; Email: fernanda.duartegonzalez@chem.ox.ac.uk

Author Contributions

TKP and FD conceptualized the study. TKP carried out the calculations. All authors participated in data analyses and writing of the manuscript. BL and SZ performed tests. TKP and FD wrote the first draft. FD supervised the study.

ACKNOWLEDGMENTS

T.K.P and F.D. acknowledge the financial support from EPSRC (EP/W010666/1 and EP/W009803/1) and the John Fell Fund (ref 0006752). This work used the Cirrus UK National Tier-2 HPC Service at EPCC (http://www.cirrus.ac.uk) funded by the University of Edinburgh and EPSRC (EP/P020267/1). B.L. is grateful to the Agency for Science, Technology and Research (A*STAR) and the Centre for Doctoral Training in Synthesis for Biology and Medicine for a studentship, generously supported by GSK, MSD, Syngenta, and Vertex. SZ thanks the Swedish Research Council (the Vetenskapsrådet international postdoc grant). Authors thank V. Martí-Centelles for helpful comments.

REFERENCES

(1) Holm, R. H.; Kennepohl, P.; Solomon, E. I. Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* **1996**, *96* (7), 2239–2314. https://doi.org/10.1021/cr9500390.

(2) Vardhan, H.; Verpoort, F. Metal-Organic Polyhedra: Catalysis and Reactive Intermediates. *Adv. Synth. Catal.* **2015**, *357* (7), 1351–1368. https://doi.org/10.1002/adsc.201400778.

(3) Brown, C. J.; Toste, F. D.; Bergman, R. G.; Raymond, K. N. Supramolecular Catalysis in Metal-Ligand Cluster Hosts. *Chem. Rev.* **2015**, *115* (9), 3012–3035. https://doi.org/10.1021/cr4001226.

(4) Ayme, J. F.; Beves, J. E.; Campbell, C. J.; Leigh, D. A. Template Synthesis of Molecular Knots. *Chem. Soc. Rev.* **2013**, *42* (4), 1700–1712. https://doi.org/10.1039/c2cs35229j.

(5) Cook, T. R.; Stang, P. J. Recent Developments in the Preparation and Chemistry of Metallacycles and Metallacages via Coordination. *Chem. Rev.* **2015**, *115* (15), 7001–7045. https://doi.org/10.1021/cr5005666.

(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), 1450–1459. https://doi.org/10.1039/b807080f.

(7) Cook, T. R.; Zheng, Y. R.; Stang, P. J. Metal-Organic Frameworks and Self-Assembled Supramolecular Coordination Complexes: Comparing and Contrasting the Design, Synthesis, and Functionality of Metal-Organic Materials. *Chem. Rev.* **2013**, *113* (1), 734–777. https://doi.org/10.1021/cr3002824.

(8) Debata, N. B.; Tripathy, D.; Chand, D. K. Self-Assembled Coordination Complexes from Various Palladium(II) Components and Bidentate or Polydentate Ligands. *Coord. Chem. Rev.* **2012**, *256* (17–18), 1831–1945. https://doi.org/10.1016/j.ccr.2012.04.001.

(9) Cramer, C. J.; Truhlar, D. G. Density Functional Theory for Transition Metals and Transition Metal Chemistry. *Phys. Chem. Chem. Phys.* **2009**, *11* (46), 10757–10816. https://doi.org/10.1039/b907148b.

(10) Vogiatzis, K. D.; Polynski, M. V.; Kirkland, J. K.; Townsend, J.; Hashemi, A.; Liu, C.; Pidko, E. A. Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities. *Chem. Rev.* **2019**, *119* (4), 2453–2523. https://doi.org/10.1021/acs.chemrev.8b00361.

(11) Mancuso, J. L.; Mroz, A. M.; Le, K. N.; Hendon, C. H. Electronic Structure Modeling of Metal-Organic Frameworks. *Chem. Rev.* **2020**, *120* (16), 8641–8715. https://doi.org/10.1021/acs.chemrev.0c00148.

(12) Li, P.; Merz, K. M. Metal Ion Modeling Using Classical Mechanics. *Chem. Rev.* **2017**, *117* (3), 1564–1686. https://doi.org/10.1021/acs.chemrev.6b00440.

(13) Li, W.; Wang, J.; Zhang, J.; Wang, W. Molecular Simulations of Metal-Coupled Protein Folding. *Curr. Opin. Struct. Biol.* **2015**, *30*, 25–31. https://doi.org/10.1016/j.sbi.2014.11.006.

(14) Riccardi, L.; Genna, V.; De Vivo, M. Metal–Ligand Interactions in Drug Design. *Nat. Rev. Chem.* **2018**, *2* (7), 100–112. https://doi.org/10.1038/s41570-018-0018-6.

(15) Piskorz, T. K.; Martí-Centelles, V.; Young, T. A.; Lusby, P. J.; Duarte, F. Computational Modeling of Supramolecular Metallo-Organic Cages-Challenges and Opportunities. *ACS Catalysis*. 2022, pp 5806–5826. https://doi.org/10.1021/acscatal.2c00837.

(16) Tarzia, A.; Jelfs, K. E. Unlocking the Computational Design of Metal-Organic Cages. *Chem. Commun.* **2022**, *58* (23), 3717–3730. https://doi.org/10.1039/d2cc00532h.

(17) Daglar, H.; Keskin, S. Recent Advances, Opportunities, and Challenges in High-Throughput Computational Screening of MOFs for Gas Separations. *Coord. Chem. Rev.* **2020**, *422*, 213470. https://doi.org/10.1016/j.ccr.2020.213470.

(18) Coudert, F. X.; Fuchs, A. H. Computational Characterization and Prediction of Metal-Organic Framework Properties. *Coord. Chem. Rev.* **2016**, *307*, 211–236. https://doi.org/10.1016/j.ccr.2015.08.001.

(19) Mashhadzadeh, A. H.; Taghizadeh, A.; Taghizadeh, M.; Munir, M. T.; Habibzadeh, S.; Salmankhani, A.; Stadler, F. J.; Saeb, M. R. Metal–Organic Framework (Mof) through the Lens of Molecular Dynamics Simulation: Current Status and Future Perspective. *J. Compos. Sci.* **2020**, *4* (2), 1–13. https://doi.org/10.3390/jcs4020075.

(20) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, *9* (6), 2733–2748. https://doi.org/10.1021/ct400146w.

(21) Li, P.; Song, L. F.; Merz, K. M. Parameterization of Highly Charged Metal Ions Using the 12-6-4 LJ-Type Nonbonded Model in Explicit Water. *J. Phys. Chem. B* **2015**, *119* (3), 883–895. https://doi.org/10.1021/jp505875v.

(22) Li, P.; Song, L. F.; Merz, K. M. Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model. *J. Chem. Theory Comput.* **2015**, *11* (4), 1645–1657. https://doi.org/10.1021/ct500918t.

(23) Sengupta, A.; Li, Z.; Song, L. F.; Li, P.; Merz, K. M. Parameterization of Monovalent Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *J. Chem. Inf. Model.* **2021**, *61* (2), 869–880. https://doi.org/10.1021/acs.jcim.0c01390.

(24) Li, Z.; Song, L. F.; Li, P.; Merz, K. M. Systematic Parametrization of Divalent Metal Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *J. Chem. Theory Comput.* **2020**, *16* (7), 4429–4442. https://doi.org/10.1021/acs.jctc.0c00194.

(25) Li, Z.; Song, L. F.; Li, P.; Merz, K. M. Parametrization of Trivalent and Tetravalent Metal Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *J. Chem. Theory Comput.* **2021**, *17* (4), 2342–2354. https://doi.org/10.1021/acs.jctc.0c01320.

(26) Zhang, Y.; Jiang, Y.; Qiu, Y.; Zhang, H. Rational Design of Nonbonded Point Charge Models for Highly Charged Metal Cations with Lennard-Jones 12-6 Potential. *J. Chem. Inf. Model.* **2021**. https://doi.org/10.1021/acs.jcim.1c00723.

(27) Zhang, Y.; Jiang, Y.; Peng, J.; Zhang, H. Rational Design of Nonbonded Point Charge Models for Divalent Metal Cations with Lennard-Jones 12-6 Potential. *J. Chem. Inf. Model.* **2021**, *61* (8), 4031–4044. https://doi.org/10.1021/acs.jcim.1c00580.

(28) Fan, K.; Zhang, Y.; Qiu, Y.; Zhang, H. Impacts of Targeting Different Hydration Free Energy References on the Development of Ion Potentials. *Phys. Chem. Chem. Phys.* **2022**, *24* (26), 16244–16262. https://doi.org/10.1039/d2cp01237e.

(29) Qiu, Y.; Jiang, Y.; Zhang, Y.; Zhang, H. Rational Design of Nonbonded Point Charge Models for Monovalent Ions with Lennard-Jones 12–6 Potential. *J. Phys. Chem. B* **2021**, *125* (49), 13502–13518. https://doi.org/10.1021/acs.jpcb.1c09103.

(30) Åqvist, J.; Warshel, A. Free Energy Relationships in Metalloenzyme-Catalyzed Reactions. Calculations of the Effects of Metal Ion Substitutions in Staphylococcal Nuclease. *J. Am. Chem. Soc.* **1990**, *112* (8), 2860–2868. https://doi.org/10.1021/ja00164a003.

(31) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112* (30), 9020–9041. https://doi.org/10.1021/jp8001614.

(32) Stote, R. H.; Karplus, M. Zinc Binding in Proteins and Solution: A Simple but Accurate Nonbonded Representation. *Proteins Struct. Funct. Bioinforma.* **1995**, *23* (1), 12–31. https://doi.org/10.1002/prot.340230104.

(33) Duarte, F.; Bauer, P.; Barrozo, A.; Amrein, B. A.; Purg, M.; Åqvist, J.; Kamerlin, S. C. L. Force Field Independent Metal Parameters Using a Nonbonded Dummy Model. *J. Phys. Chem. B* **2014**, *118* (16), 4351–4362. https://doi.org/10.1021/jp501737x.

(34) Jiang, Y.; Zhang, H.; Feng, W.; Tan, T. Refined Dummy Atom Model of Mg2+ by Simple Parameter

Screening Strategy with Revised Experimental Solvation Free Energy. *J. Chem. Inf. Model.* **2015**, *55* (12), 2575–2586. https://doi.org/10.1021/acs.jcim.5b00286.

(35) Liao, Q.; Kamerlin, S. C. L.; Strodel, B. Development and Application of a Nonbonded Cu2+ Model That Includes the Jahn-Teller Effect. *J. Phys. Chem. Lett.* **2015**, *6* (13), 2657–2662. https://doi.org/10.1021/acs.jpclett.5b01122.

(36) Jiang, Y.; Zhang, H.; Tan, T. Rational Design of Methodology-Independent Metal Parameters Using a Nonbonded Dummy Model. *J. Chem. Theory Comput.* **2016**, *12* (7), 3250–3260. https://doi.org/10.1021/acs.jctc.6b00223.

(37) Saxena, A.; Sept, D. Multisite Ion Models That Improve Coordination and Free Energy Calculations in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9* (8), 3538–3542. https://doi.org/10.1021/ct400177g.

(38) Liao, Q.; Pabis, A.; Strodel, B.; Kamerlin, S. C. L. Extending the Nonbonded Cationic Dummy Model to Account for Ion-Induced Dipole Interactions. *J. Phys. Chem. Lett.* **2017**, *8* (21), 5408–5414. https://doi.org/10.1021/acs.jpclett.7b02358.

(39) Jiang, Y.; Zhang, H.; Cui, Z.; Tan, T. Modeling Coordination-Directed Self-Assembly of M 2 L 4 Nanocapsule Featuring Competitive Guest Encapsulation. *J. Phys. Chem. Lett.* **2017**, *8* (9), 2082–2086. https://doi.org/10.1021/acs.jpclett.7b00773.

(40) 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. https://doi.org/10.1021/ja303542r.

(41) 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. https://doi.org/10.1021/nn404595j.

(42) Pang, Y. P. Successful Molecular Dynamics Simulation of Two Zinc Complexes Bridged by a Hydroxide in Phosphotriesterase Using the Cationic Dummy Atom Method. *Proteins Struct. Funct. Genet.* **2001**, *45* (3), 183–189. https://doi.org/10.1002/prot.1138.

(43) Talmazan, R. A.; Podewitz, M. PyConSolv: A Python Package for Conformer Generation of (Metal-Containing) Systems in Explicit Solvent. *J. Chem. Inf. Model.* **2023**, *63* (17), 5400–5407. https://doi.org/10.1021/acs.jcim.3c00798.

(44) Zheng, S.; Tang, Q.; He, J.; Du, S.; Xu, S.; Wang, C.; Xu, Y.; Lin, F. VFFDT: A New Software for Preparing AMBER Force Field Parameters for Metal-Containing Molecular Systems. *J. Chem. Inf. Model.* **2016**, *56* (4), 811–818. https://doi.org/10.1021/acs.jcim.5b00687.

(45) Vanduyfhuys, L.; Vandenbrande, S.; Verstraelen, T.; Schmid, R.; Waroquier, M.; Van Speybroeck, V. QuickFF: A Program for a Quick and Easy Derivation of Force Fields for Metal-Organic Frameworks from Ab Initio Input. *J. Comput. Chem.* **2015**, *36* (13), 1015–1027. https://doi.org/10.1002/jcc.23877.

(46) Nilsson, K.; Lecerof, D.; Sigfridsson, E.; Ryde, U. An Automatic Method to Generate Force-Field Parameters for Hetero-Compounds. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **2003**, *59* (2), 274–289. https://doi.org/10.1107/S0907444902021431.

(47) Burger, S. K.; Lacasse, M.; Verstraelen, T.; Drewry, J.; Gunning, P.; Ayers, P. W. Automated Parametrization of AMBER Force Field Terms from Vibrational Analysis with a Focus on Functionalizing

Dinuclear Zinc(II) Scaffolds. *J. Chem. Theory Comput.* **2012**, *8* (2), 554–562. https://doi.org/10.1021/ct2007742.

(48) Åqvist, J.; Warshel, A. Computer Simulation of the Initial Proton Transfer Step in Human Carbonic Anhydrase I. *J. Mol. Biol.* **1992**, *224* (1), 7–14. https://doi.org/10.1016/0022-2836(92)90572-2.

(49) Biswal, D.; Kusalik, P. G. Probing Molecular Mechanisms of Self-Assembly in Metal-Organic Frameworks. *ACS Nano* **2017**, *11* (1), 258–268. https://doi.org/10.1021/acsnano.6b05444.

(50) Yoneya, M.; Tsuzuki, S.; Aoyagi, M. Simulation of Metal-Organic Framework Self-Assembly. *Phys. Chem. Chem. Phys.* **2015**, *17* (14), 8649–8652. https://doi.org/10.1039/c5cp00379b.

(51) Marrone, T. J.; Merz, K. Transferability of Ion Models. *J. Phys. Chem.* **1993**, *97*, 6524–6529.

(52) Sakharov, D. V.; Lim, C. Zn Protein Simulations Including Charge Transfer and Local Polarization Effects. *J. Am. Chem. Soc.* **2005**, *127* (13), 4921–4929. https://doi.org/10.1021/ja0429115.

(53) SAKHAROV, D.; LIM, C. Force Fields Including Charge Transfer and Local Polarization Effects: Application to Proteins Containing Multi/Heavy Metal Ions. *J. Comput. Chem.* **208AD**, *30* (2), 191–201. https://doi.org/10.1002/jcc.

(54) Seminario, J. M. Calculation of Intramolecular Force Fields from Second-Derivative Tensors. *Int. J. Quantum Chem.* **1996**, *60* (7), 1271–1277. https://doi.org/10.1002/(SICI)1097-461X(1996)60:7<1271::AID-QUA8>3.0.CO;2-W.

(55) Allen, A. E. A.; Payne, M. C.; Cole, D. J. Harmonic Force Constants for Molecular Mechanics Force Fields via Hessian Matrix Projection. *J. Chem. Theory Comput.* **2018**, *14* (1), 274–281. https://doi.org/10.1021/acs.jctc.7b00785.

(56) Wang, R.; Ozhgibesov, M.; Hirao, H. Partial Hessian Fitting for Determining Force Constant Parameters in Molecular Mechanics. *J. Comput. Chem.* **2016**, 2349–2359. https://doi.org/10.1002/jcc.24457.

(57) Wang, R.; Ozhgibesov, M.; Hirao, H. Analytical Hessian Fitting Schemes for Efficient Determination of Force-Constant Parameters in Molecular Mechanics. *J. Comput. Chem.* **2018**, *39* (6), 307–318. https://doi.org/10.1002/jcc.25100.

(58) Ercolessi, F.; Adams, J. B. Interatomic Potentials from First-Principles Calculations. *Mater. Res. Soc. Symp. Proc.* **1993**, *291*, 31–36. https://doi.org/10.1557/proc-291-31.

(59) Masia, M.; Guàrdia, E.; Nicolini, P. The Force Matching Approach to Multiscale Simulations: Merits, Shortcomings, and Future Perspectives. *Int. J. Quantum Chem.* **2014**, *114* (16), 1036–1040. https://doi.org/10.1002/qua.24621.

(60) Wang, L. P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5* (11), 1885–1891. https://doi.org/10.1021/jz500737m.

(61) Akin-Ojo, O.; Song, Y.; Wang, F. Developing Ab Initio Quality Force Fields from Condensed Phase Quantum-Mechanics/Molecular-Mechanics Calculations through the Adaptive Force Matching Method. *J. Chem. Phys.* **2008**, *129* (6). https://doi.org/10.1063/1.2965882.

(62) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J. Chem. Theory Comput.*

2011, *7* (12), 4026–4037. https://doi.org/10.1021/ct200196m.

(63) Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. Rapid Parameterization of Small Molecules Using the Force Field Toolkit. *J. Comput. Chem.* **2013**, *34* (32), 2757–2770. https://doi.org/10.1002/jcc.23422.

(64) Horton, J. T.; Allen, A. E. A.; Dodda, L. S.; Cole, D. J. QUBEKit: Automating the Derivation of Force Field Parameters from Quantum Mechanics. *J. Chem. Inf. Model.* **2019**, *59* (4), 1366–1381. https://doi.org/10.1021/acs.jcim.8b00767.

(65) Peters, M. B.; Yang, Y.; Wang, B.; Füsti-Molnár, L.; Weaver, M. N.; Merz, K. M. Structural Survey of Zinc-Containing Proteins and Development of the Zinc AMBER Force Field (ZAFF). *J. Chem. Theory Comput.* **2010**, *6* (9), 2935–2947. https://doi.org/10.1021/ct1002626.

(66) Lin, F.; Wang, R. Systematic Derivation of AMBER Force Field Parameters Applicable to Zinc-Containing Systems. *J. Chem. Theory Comput.* **2010**, *6* (6), 1852–1870. https://doi.org/10.1021/ct900454q.

(67) Krokidas, P.; Castier, M.; Moncho, S.; Brothers, E.; Economou, I. G. Molecular Simulation Studies of the Diffusion of Methane, Ethane, Propane, and Propylene in ZIF-8. *J. Phys. Chem. C* **2015**, *119* (48), 27028–27037. https://doi.org/10.1021/acs.jpcc.5b08554.

(68) Krokidas, P.; Castier, M.; Moncho, S.; Sredojevic, D. N.; Brothers, E. N.; Kwon, H. T.; Jeong, H. K.; Lee, J. S.; Economou, I. G. ZIF-67 Framework: A Promising New Candidate for Propylene/Propane Separation. Experimental Data and Molecular Simulations. *J. Phys. Chem. C* **2016**, *120* (15), 8116–8124. https://doi.org/10.1021/acs.jpcc.6b00305.

(69) Verploegh, R. J.; Kulkarni, A.; Boulfelfel, S. E.; Haydak, J. C.; Tang, D.; Sholl, D. S. Screening Diffusion of Small Molecules in Flexible Zeolitic Imidazolate Frameworks Using a DFT-Parameterized Force Field. *J. Phys. Chem. C* **2019**, *123* (14), 9153–9167. https://doi.org/10.1021/acs.jpcc.9b00733.

(70) Norjmaa, G.; Maréchal, J. D.; Ujaque, G. Microsolvation and Encapsulation Effects on Supramolecular Catalysis: C-C Reductive Elimination inside [Ga4L6]12- Metallocage. *J. Am. Chem. Soc.* **2019**, *141* (33), 13114–13123. https://doi.org/10.1021/jacs.9b04909.

(71) Pesce, L.; Perego, C.; Grommet, A. B.; Klajn, R.; Pavan, G. M. Molecular Factors Controlling the Isomerization of Azobenzenes in the Cavity of a Flexible Coordination Cage. *J. Am. Chem. Soc.* **2020**, *142* (21), 9792–9802. https://doi.org/https://dx.doi.org/10.1021/jacs.0c03444.

(72) Xue, W.; Pesce, L.; Bellamkonda, A.; Ronson, T. K.; Wu, K.; Zhang, D.; Vanthuyne, N.; Brotin, T.; Martinez, A.; Pavan, G. M.; et al. Subtle Stereochemical Effects Influence Binding and Purification Abilities of an FeII4L4 Cage. *J. Am. Chem. Soc.* **2023**, *145* (9), 5570–5577. https://doi.org/10.1021/jacs.3c00294.

(73) Hoops, S. C.; Anderson, K. W.; Merz, K. M. Field Force Field Design for Metalloproteins Stephen. *J. Am. Chem. Soc.* **1991**, *113* (20), 8262–8270.

(74) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Effective Tailor-Made Force Field Parameterization of the Several Zn Coordination Environments in the Puzzling FTase Enzyme: Opening the Door to the Full Understanding of Its Elusive Catalytic Mechanism. *Theor. Chem. Acc.* **2007**, *117* (1), 171–181. https://doi.org/10.1007/s00214-006-0170-9.

(75) Neves, R. P. P.; Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Parameters for Molecular Dynamics

Simulations of Manganese-Containing Metalloproteins. *J. Chem. Theory Comput.* **2013**, *9* (6), 2718–2732. https://doi.org/10.1021/ct400055v.

(76) 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* (25), 10024–10035. https://doi.org/10.1021/ja00051a040.

(77) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280. https://doi.org/10.1021/j100142a004.

(78) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174. https://doi.org/10.1002/jcc.20035.

(79) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F. Y. R.E.D. Server: A Web Service for Deriving RESP and ESP Charges and Building Force Field Libraries for New Molecules and Molecular Fragments. *Nucleic Acids Res.* **2011**, *39* (SUPPL. 2), 511–517. https://doi.org/10.1093/nar/gkr288.

(80) Alenaizan, A.; Burns, L. A.; Sherrill, C. D. Python Implementation of the Restrained Electrostatic Potential Charge Model. *Int. J. Quantum Chem.* **2020**, *120* (2), 1–7. https://doi.org/10.1002/qua.26035.

(81) Wang, L.; O'Mara, M. L. PsiRESP: Calculating RESP Charges with Psi4. *J. Open Source Softw.* **2022**, *7* (73), 4100. https://doi.org/10.21105/joss.04100.

(82) Suzuki, K.; Tominaga, M.; Kawano, M.; Fujita, M. Self-Assembly of an M6L12 Coordination Cube. *Chem. Commun.* **2009**, No. 13, 1638–1640. https://doi.org/10.1039/b822311d.

(83) Sun, Q. F.; Iwasa, J.; Ogawa, D.; Ishido, Y.; Sato, S.; Ozeki, T.; Sei, Y.; Yamaguchi, K.; Fujita, M. Self-Assembled M24L48 Polyhedra and Their Sharp Structural Switch upon Subtle Liqand Variation. *Science* **2010**, *328* (5982), 1144–1147. https://doi.org/10.1126/science.1188605.

(84) Zhu, J. L.; Zhang, D.; Ronson, T. K.; Wang, W.; Xu, L.; Yang, H. B.; Nitschke, J. R. A Cavity-Tailored Metal-Organic Cage Entraps Gases Selectively in Solution and the Amorphous Solid State. *Angew. Chem. Int. Ed.* **2021**, *60* (21), 11789–11792. https://doi.org/10.1002/anie.202102095.

(85) Davies, J. A.; Ronson, T. K.; Nitschke, J. R. Twisted Rectangular Subunits Self-Assemble into a Ferritin-like Capsule. *Chem* **2022**, *8* (4), 1099–1106. https://doi.org/10.1016/j.chempr.2022.01.003.

(86) Li, P.; Merz, K. M. MCPB.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, *56* (4), 599–604. https://doi.org/10.1021/acs.jcim.5b00674.

(87) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; et al. Array Programming with NumPy. *Nature* **2020**, *585* (7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2.

(88) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; et al. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. In *Proceedings of the 15th Python in Science Conference*; Benthall, S., Rostrup, S., Eds.; 2016; pp 98–105. https://doi.org/10.25080/Majora-629e541a-00e.

(89) RDKit: Open-source cheminformatics http://www.rdkit.org (accessed Feb 14, 2022).

(90) ParmEd https://parmed.github.io/ParmEd/html/index.html (accessed Feb 15, 2022).

(91) Hagberg, A. A.; Schult, D. A.; Swart, P. J. Exploring Network Structure, Dynamics, and Function Using NetworkX. In *Proceedings of the 7th Python in Science Conference*; Varoquaux, G., Vaught, T., Millman, J., Eds.; Pasadena, CA USA, 2008; pp 11–15.

(92) Neese, F. Software Update: The ORCA Program System—Version 5.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2022**, *12* (5), 1–15. https://doi.org/10.1002/wcms.1606.

(93) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152* (22). https://doi.org/10.1063/5.0004608.

(94) Young, T. A.; Silcock, J. J.; Sterling, A. J.; Duarte, F. AutodE: Automated Calculation of Reaction Energy Profiles— Application to Organic and Organometallic Reactions. *Angew. Chem. Int. Ed.* **2021**, *60* (8), 4266–4274. https://doi.org/https://doi.org/10.1002/anie.202011941.

(95) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25* (2), 247–260. https://doi.org/10.1016/j.jmgm.2005.12.005.

(96) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2012**, *32*, 174–182. https://doi.org/10.1002/jcc.

(97) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170. https://doi.org/10.1063/1.478522.

(98) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–3305. https://doi.org/10.1039/b508541a.

(99) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 1–19. https://doi.org/10.1063/1.3382344.

(100) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465. https://doi.org/10.1002/jcc.

(101) Liao, P.; Langloss, B. W.; Johnson, A. M.; Knudsen, E. R.; Tham, F. S.; Julian, R. R.; Hooley, R. J. Two-Component Control of Guest Binding in a Self-Assembled Cage Molecule. *Chem. Commun.* **2010**, *46* (27), 4932–4934. https://doi.org/10.1039/c0cc00234h.

(102) August, D. P.; Nichol, G. S.; Lusby, P. J. Maximizing Coordination Capsule–Guest Polar Interactions in Apolar Solvents Reveals Significant Binding. *Angew. Chem. Int. Ed.* **2016**, *55* (48), 15022–15026. https://doi.org/10.1002/anie.201608229.

(103) Pluth, M. D.; Johnson, D. W.; Szigethy, G.; Davis, A. V.; Teat, S. J.; Oliver, A. G.; Bergman, R. G.; Raymond, K. N. Structural Consequences of Anionic Host-Cationic Guest Interactions in a Supramolecular Assembly. *Inorg. Chem.* **2009**, *48* (1), 111–120. https://doi.org/10.1021/ic8012848.

(104) Mal, P.; Breiner, B.; Rissanen, K.; Nitschke, J. R. White Phosphorus Is Air-Stable within a Self-Assembled

Tetrahedral Capsule. *Science* **2009**, *324* (5935), 1697–1699. https://doi.org/10.1126/science.1175313.

(105) Takezawa, H.; Murase, T.; Resnati, G.; Metrangolo, P.; Fujita, M. Halogen-Bond-Assisted Guest Inclusion in a Synthetic Cavity ** Angewandte. **2015**, *1* (3), 8411–8414. https://doi.org/10.1002/anie.201500994.

(106) Metherell, A. J.; Cullen, W.; Williams, N. H.; Ward, M. D. Binding of Hydrophobic Guests in a Coordination Cage Cavity Is Driven by Liberation of " High-Energy " Water. **2018**, 1554–1560. https://doi.org/10.1002/chem.201704163.

(107) Wu, K.; Li, K.; Hou, Y. J.; Pan, M.; Zhang, L. Y.; Chen, L.; Su, C. Y. Homochiral D4-Symmetric Metal–Organic Cages from Stereogenic Ru(II) Metalloligands for Effective Enantioseparation of Atropisomeric Molecules. *Nat. Commun.* **2016**, *7* (1), 10487. https://doi.org/10.1038/ncomms10487.

(108) Fujita, D.; Ueda, Y.; Sato, S.; Mizuno, N.; Kumasaka, T.; Fujita, M. Self-Assembly of Tetravalent Goldberg Polyhedra from 144 Small Components. *Nature* **2016**, *540* (7634), 563–566. https://doi.org/10.1038/nature20771.

(109) Ayme, J. F.; Beves, J. E.; Leigh, D. A.; McBurney, R. T.; Rissanen, K.; Schultz, D. Pentameric Circular Iron(II) Double Helicates and a Molecular Pentafoil Knot. *J. Am. Chem. Soc.* **2012**, *134* (22), 9488–9497. https://doi.org/10.1021/ja303355v.

(110) Zhang, L.; August, D. P.; Zhong, J.; Whitehead, G. F. S.; Vitorica-Yrezabal, I. J.; Leigh, D. A. Molecular Trefoil Knot from a Trimeric Circular Helicate. *J. Am. Chem. Soc.* **2018**, *140* (15), 4982–4985. https://doi.org/10.1021/jacs.8b00738.

(111) Morris, W.; Stevens, C. J.; Taylor, R. E.; Dybowski, C.; Yaghi, O. M.; Garcia-Garibay, M. A. NMR and X-Ray Study Revealing the Rigidity of Zeolitic Imidazolate Frameworks. *J. Phys. Chem. C* **2012**, *116* (24), 13307–13312. https://doi.org/10.1021/jp303907p.

(112) Banerjee, R.; Phan, A.; Wang, B.; Knobler, C.; Furukawa, H.; O'Keeffe, M.; Yaghi, O. M. High-Throughput Synthesis of Zeolitic Imidazolate Frameworks and Application to CO2 Capture. *Science* **2008**, *319*, 939–943. https://doi.org/10.1126/science.1152516.

(113) Young, T. A.; Martí-Centelles, V.; Wang, J.; Lusby, P. J.; Duarte, F. Rationalizing the Activity of an "Artificial Diels-Alderase": Establishing Efficient and Accurate Protocols for Calculating Supramolecular Catalysis. *J. Am. Chem. Soc.* **2020**, *142* (3), 1300–1310. https://doi.org/10.1021/jacs.9b10302.

(114) Norjmaa, G.; Vidossich, P.; Maréchal, J.-D.; Ujaque, G. Modeling Kinetics and Thermodynamics of Guest Encapsulation into the [M4L6]12– Supramolecular Organometallic Cage. *J. Chem. Inf. Model.* **2021**, *61* (9), 4370–4381. https://doi.org/10.1021/acs.jcim.1c00348.

(115) Lewis, J. E. M.; Tarzia, A.; White, A. J. P.; Jelfs, K. E. Conformational Control of Pd2L4 Assemblies with Unsymmetrical Ligands. *Chem. Sci.* **2020**, *11* (3), 677–683. https://doi.org/10.1039/c9sc05534g.

(116) Davis, A. V.; Fiedler, D.; Seeber, G.; Zahl, A.; Van Eldik, R.; Raymond, K. N. Guest Exchange Dynamics in an M4L6 Tetrahedral Host. *J. Am. Chem. Soc.* **2006**, *128* (4), 1324–1333. https://doi.org/10.1021/ja056556+.

(117) August, D. P.; Borsley, S.; Cockroft, S. L.; Sala, F.; Leigh, D. A.; Webb, S. J. Transmembrane Ion Channels Formed by a Star of David [2]Catenane and a Molecular Pentafoil Knot. **2020**. https://doi.org/10.1021/jacs.0c07977.

(118) Martí-Centelles, V.; Piskorz, T. K.; Duarte, F. *CageCavityCalc (C3): A Computational Tool for Calculating and Visualizing Cavities in Molecular Cages.*; 2024. https://doi.org/10.26434/chemrxiv-2024-fmlx0.