JCTC Journal of Chemical Theory and Computation

Open Access



pubs.acs.org/JCTC

Article

# *Metallicious*: Automated Force-Field Parameterization of Covalently Bound Metals for Supramolecular Structures

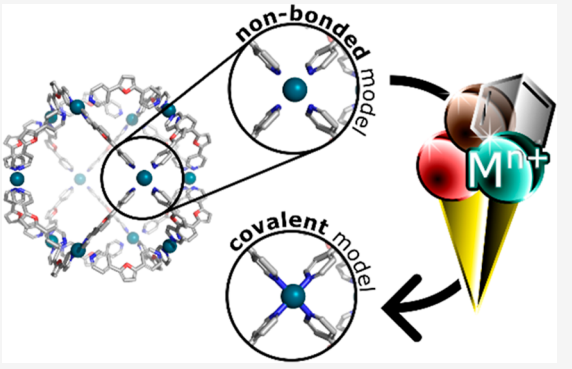
Tomasz K. Piskorz, Bernadette Lee, Shaoqi Zhan, and Fernanda Duarte*

Cite This: *J. Chem. Theory Comput.* 2024, 20, 9060−9071

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**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 parameterization, especially for supramolecular systems containing repetitive units. To streamline this process, we introduce *metallicious*, a Python tool designed for efficient force-field parameterization of supramolecular structures. *Metallicious* has been tested on diverse systems including supramolecular cages, knots, and MOFs. Our benchmarks demonstrate that parameters accurately reproduce the reference properties obtained from quantum calculations and crystal structures. Molecular dynamics simulations of the generated structures consistently yield stable simulations in explicit solvent, in contrast to similar simulations performed with nonbonded and cationic dummy models. Overall, *metallicious* facilitates the atomistic modeling of supramolecular systems, key for understanding their dynamic properties and host−guest interactions. The tool is freely available on GitHub (https://github.com/duartegroup/metallicious).






## 1. INTRODUCTION

Metal ions play significant roles 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,[12,13] and hybrid QM/MM.[11,12] These techniques have been applied to model metalloproteins,[14,15] metallo-organic cages,[16,17] and MOFs.[18−21] 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 nonbonded,[22−34] cationic dummy,[35−44] and covalent models (Figure 1b).[45−49] Nonbonded models describe metal centers as van der Waals spheres with integer charge, describing ligand−metal interactions through electrostatic Coulombic and Lennard−Jones (L−J) terms. Li and Merz have reported parameter sets for 56 metals, aiming to reproduce hydration-free energy (HFE), coordination number (CN), and ion-oxygen distances 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.[22,23,27−30] Zhang sampled a larger L−J parameter space to identify parameters capable of reproducing HFE and CN for a set of 47 ions.[31−34]

The dummy model, originally developed by Åqvist and Warshel,[24,50] 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 nonbonded models by simultaneously reproducing HFE and CN. However, its predefined configuration reduces flexibility, and only a limited number of metal ions have been parameterized.[35−42,44] While both nonbonded and cationic dummy models can describe

Received: July 2, 2024
Revised: September 20, 2024
Accepted: September 25, 2024
Published: October 7, 2024

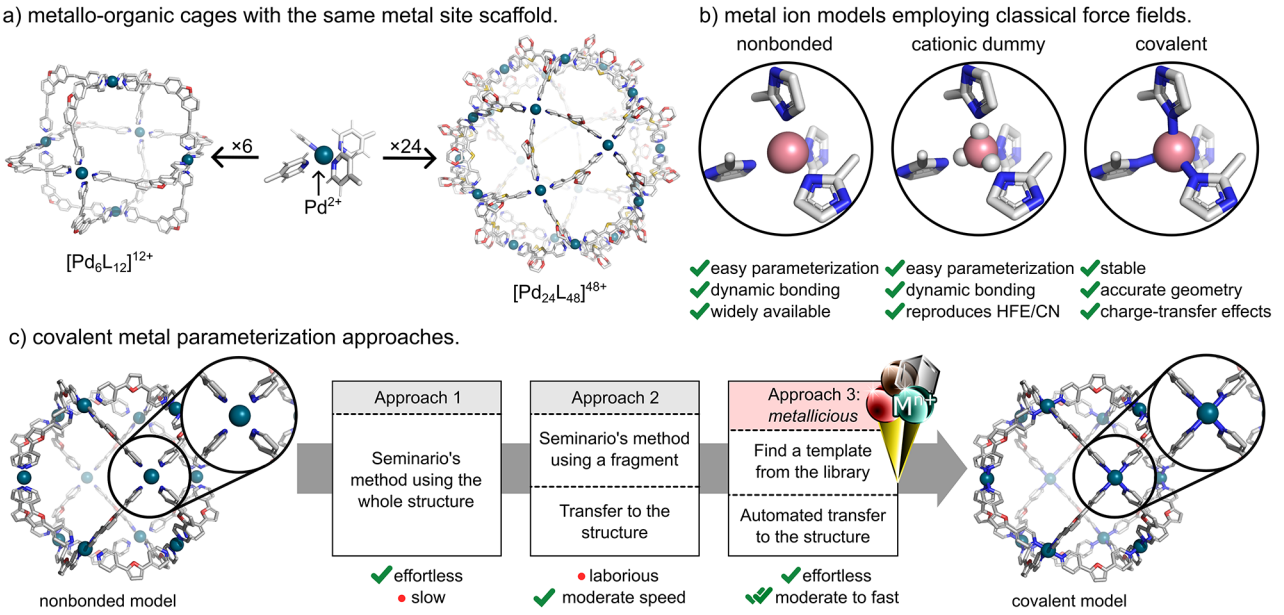



**Figure 1.** Force-field parameterization of supramolecular systems. (a) Supramolecular metallocages often exhibit repeated metal binding.[91,92] (b) Approaches for deriving metal force-field parameters include nonbonded, cationic dummy, and covalent models. (c) Methodologies for parameterization of a covalent metal model. *Metallicious* logo appears courtesy of Tomasz K. Piskorz.

bond-forming and -breaking processes,[42,43,51,52] they often exhibit unexpected behavior when transferred to different systems[35,53] and fail to account for charge transfer.[43,54,55]

The covalent bond model explicitly describes bonds, angles, torsions, and 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. Parameterizing 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[49,56−59] or employing energy-fitting or/and force-matching techniques.[60−64] Among Hessian-based methods, the Seminario method is the most widely used. It involves the projection of the Hessian matrix elements onto relevant bonded parameters using the harmonic approximation (Figure 1c; approach 1).[56] In this approach, dihedral parameters involving metal ions are omitted as they usually have minimal impact on the structure,[65−68] although exceptions including these terms exist.[69] This method has been successfully applied to model small metallo-organic molecules,[70−72] metalloproteins,[67,73] MOFs,[74−76] and supramolecular cages.[77−79] Cole and co-workers have improved the Seminario method by addressing the problem of double-counting bending interactions.[57] Alternatively, the energy- and force-matching techniques fit energies and forces to reproduce the reference data.[64,80,81] These methods have also been successfully applied to metallo-organic molecules[82] and cages.[83] Hu and Ryde have compared these two approaches for $Zn^{2+}$ complexes, finding that while energy- and force-matching techniques more accurately reproduced QM-optimized geometry of the complexes, they were also more laborious, requiring significant human intervention.[84] The L−J parameters are typically taken from nonbonded model parameters or the Universal Force Field, which covers 126 elements,[85] while partial charges are obtained using the restrained electrostatic potential (RESP) method,[86] available in various software including AmberTools (via Gaussian, GAMESS-US),[87] R.E.D. server (Gaussian, GAMESS-US, Firefly),[88] psi4/resp (psi4),[89] and psiRESP (psi4).[90]

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 parameterization (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,[93] offering a semiautomatic 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 parameterization (e.g., *pdb*2gmx 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 modeling of supramolecular structures containing repeating metal site units (Figure 1a).

It is important to note that more advanced force-field potentials for specific metal ions have been developed, for example, using a polarizable force field, such as the classical Drude oscillator model,[94] or more advanced models, including atomic multipole moments[95] or electrostatic penetration.[96] Another promising avenue has been the development of a generic force field by Grimme et al.,[97] which includes parameters for most metals and has been used for gas-phase optimization of MOFs and metallo-organic cages.[21] While such approaches are usually more accurate than standard force fields, they remain computationally too expensive for routine all-atoms condensed-phase MD simulations.

In this work, we introduce *metallicious* (Figure 1c; approach 3), a tool designed to streamline the parameterization of metal



https://doi.org/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071

centers using the covalent metal model based on a *template library*. It enables the parameterization of new templates using Seminario's method and RESP charges. As a result, *metallicious* has 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,[98] MDAnalysis,[99] RDKit,[100] ParmEd,[101] and NetworkX[102] for core functionalities. Optionally, the parameterization of new templates requires the QM engine ORCA,[103,104] autodE (used as an interface for the QM engine),[105] and psiRESP,[90] and parameterization of the organic molecules requires AmberTools.[106] 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 2a.

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 a suitable template from the database, and execution of template parameterization. For each metal site identified in the input, a *metal_site* class is generated, storing the metal index and the corresponding template. Preparameterized 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 parameterize the new metal site (truncated coordinate file, metal name, charge, multiplicity, L−J parameters, and ligand charges). Lastly, the *patcher* class combines all of the *metal_site* classes with the input structure and topology to produce the output files.

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

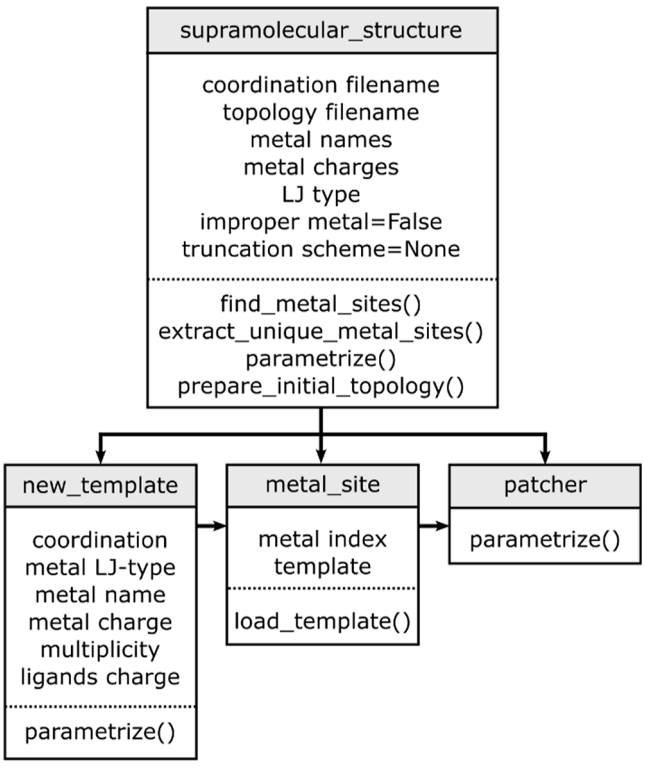
(a) *Coordinate file (f).* Accepted formats include .xyz, .pdb, .gro, and others supported by MDAnalysis. The coordination file might be taken directly from the Cambridge Crystallographic Data Centre (CCDC) (after removing solvent and additives), generated using our tool for automated metallocage construction, *cgbind*,[107] or Avogadro and optimized with the Universal Force Field.[108]

(b) *Topology file (p)* or specify *prepare_initial_topology* argument. Topology file containing force-field parameters for the input coordination file. Accepted formats include .top, .itp, .prmtop, and other files supported by ParmEd.[101] Note that the L−J parameters and charge of metal atoms present in the topology files are not used, only their indices. Such parameters can be obtained from tools such as AmberTools,[106] ATB,[70] or CHARMM-GUI.[109] Instead of providing a topology file, the user might request them by specifying the *prepare_initial_topology* argument, which interfaces with the antechamber to obtain general AMBER force field parameters for the ligand.[106]

(c) *LJ type (LJ_type).* Specify the name of the library used to extract the metal L−J 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 parameterization of new templates.

a) class structure




b) usage example

```
metallicious -f cage.pdb -p cage.top -LJ_type uff
                    -metal_charges Pd 2
```

```
from metallicious import supramolecular_structure
cage = supramolecular_structure(f='cage.pdb',
                        p='cage.top', LJ_type='uff',
                        metal_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 via either a command line or Python script. Input requirements include coordinate (.pdb), topology (.top), L−J type, and identity and charge of the metal ion.

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

Upon completion of the parameterization, two output files are generated: a coordination file with reordered atoms (default: out.pdb) and a topology file saved in GROMACS format



https://doi.org/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071

(default: out.top), although other formats are supported via ParmEd. The topology file contains modified bonds, angles, dihedrals, and partial charges of the metal and its surrounding atoms and modified L−J parameters of metals.

**2.3. Limitations.** The code currently supports organometallic structures with metals separated by at least two nonmetal atoms; metal clusters are not currently supported.

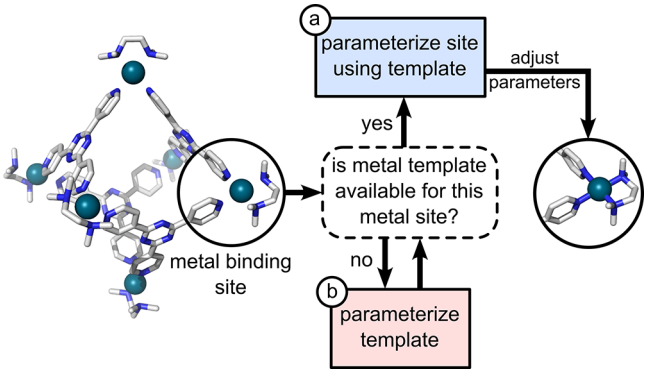



**Figure 3.** Overview of metal parameterization in *metallicious*, illustrating the communication between the two primary program subroutines: parameterization site using a template and (new) template parameterization.

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

(a) Parameterization using a template library. Selects a suitable template from a repository of predefined templates and adjusts the input topology accordingly.
(b) Parameterization of a new template. If a suitable template is unavailable, it performs template parameterization.

*2.4.1. Parameterization Using a 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):

*2.4.1.1. Template Selection and (Optionally) Truncation.* This search aims to identify file names that match the user-specified metal name, charge, and L−J type (Figure 4a.i). The first matching template is used for the next step. If unavailable, users can truncate the system and perform the search again or parameterize a new template. Truncation involves applying a distance cutoff 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 Section 3.3.

*2.4.1.2. Reduction of the Metal Site.* The original metal site is aligned with a selected template, chosen after matching the name, charge, and L−J type through an iterative process. A distance cutoff equal to the template's radius plus 2 Å is used [Figure 4a.ii, blue dashed line]. The template's radius is defined as the distance between the metal and its furthest atom [Figure 4a.ii, 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

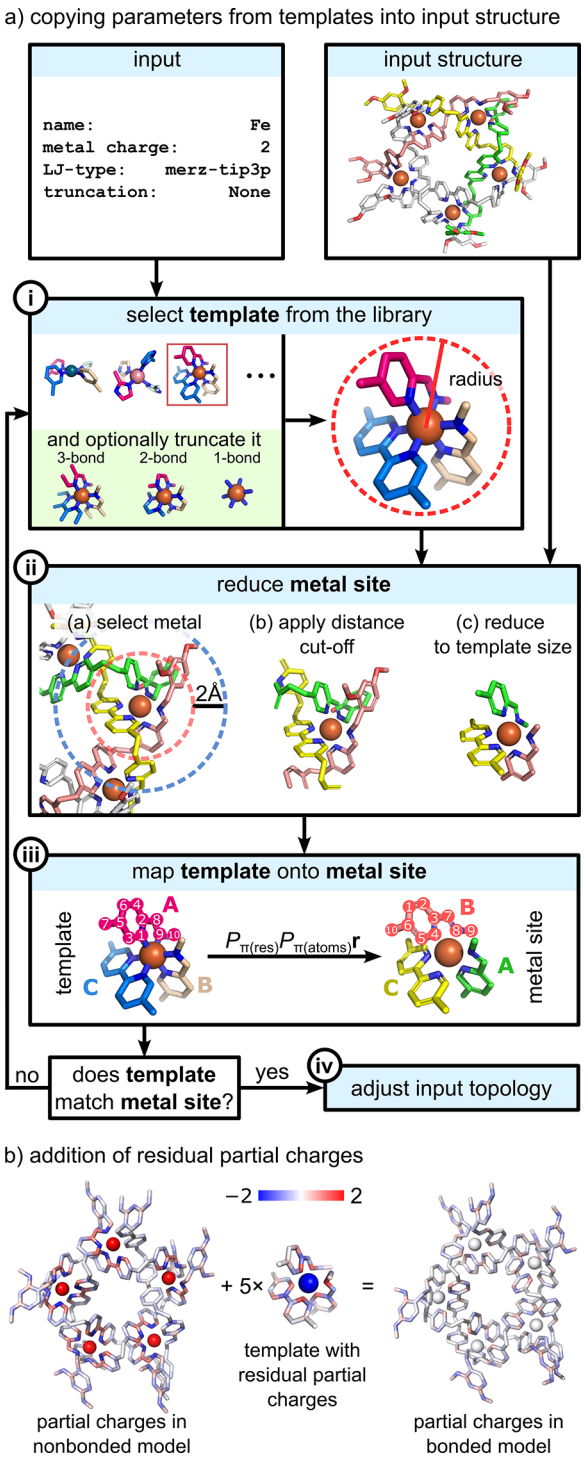



**Figure 4.** Parameterization workflow using templates: (a) Copying parameters from a template into the input coordination and topology files. $P_{\pi(res)}$ and $P_{\pi(atoms)}$ denote permutation matrix over residues and residues' atoms, respectively, and **r** denotes atom positions. (b) Partial charges accounting for charge transfer are obtained by summing partial charges from the nonbonded model and template's residual partial charges.

of the metal site's graphs. The metal site's linkers are then pruned to match template graphs, resulting in identical structures that only differ in the order of the atoms.

*2.4.1.3. Mapping of the Template onto the Metal Site.* Achieving the correct atom order involves an exhaustive


https://doi.org/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071

exploration of all possible permutations of ligand and atom numbering, the most time-consuming part of the process (Figure 4a.iii). For efficiency, the process is divided into two stages. First, the order of all non-hydrogen atoms is determined by exhaustively exploring permutations of the pruned 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 linker graphs, constraining the order of heavy atoms. Successful mapping is defined as RMSD <2 Å, although this criterion can be modified. If RMSD >2 Å, the next template from the library is evaluated; if unavailable, the procedure for parameterizing a new template is followed (Figure 5).

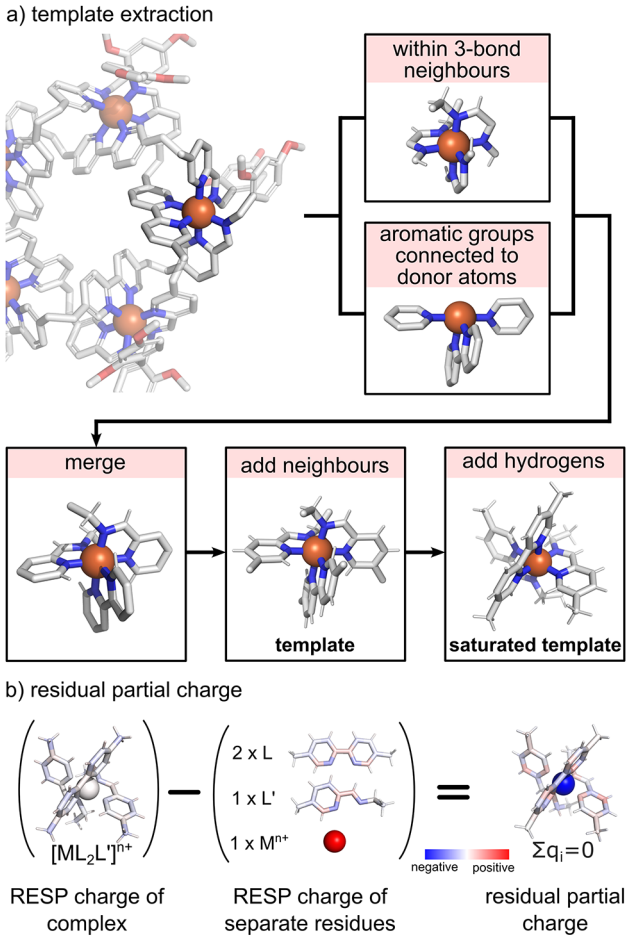



**Figure 5.** New template parameterization. (a) Extracting the metal template from the supramolecular complex. (b) Calculating the residual partial charge of the template.

*2.4.1.4. 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 (Figure 4a.iv). This 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.

*2.4.2. Parameterization of a New Template.* When no template is found in the library, *metallicious* performs parameterization using the following steps:

*2.4.2.1. Extraction of a New Template.* To extract the new template, *metallicious* iterates through metal atoms in the input file (Figure 5a). 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 parameterization using the template library. Hydrogen atoms are added to atoms with unfulfilled valence, saturating the template, which is required for QM calculations (*saturated template*).

*2.4.2.2. Parameterization of Bonded Parameters.* Bonds and angles are parameterized using Seminario's method, adapted from Cole and co-workers.[57] 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 S4). Similar to the work of others, proper dihedrals are not parameterized.[65−68] 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).

*2.4.2.3. Residual Partial Charge.* Partial charges are calculated using the RESP method.[86] 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.[110−113] 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.

*2.4.2.4. 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.

## 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 ($[Pd_2L_4]^{4+}$,[114,115] $[Ga_4L_6]^{12-}$,[116] $[Fe_4L_6]^{4-}$,[117] $[Pd_6L_4]^{12+}$,[118] $[Co_8L_{12}]^{16+}$,[119] $[Pd_6Ru_8L_{24}]^{28+}$,[120] and $[Pd_{48}L_{96}]^{96+}$),[121] two knots ($[Fe_5L_5]^{10+}$,[122] $[Zn_3L_3]^{6+}$),[123] and two MOFs (ZIF-8[124] and ZIF-67;[125] Table S2). Among these systems, $[Pd_2L_4]^{4+}$,[126] $[Ga_4L_6]^{12-}$,[77,127] $[Pd_6L_4]^{12+}$,[78] ZIF-


https://doi.org/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071

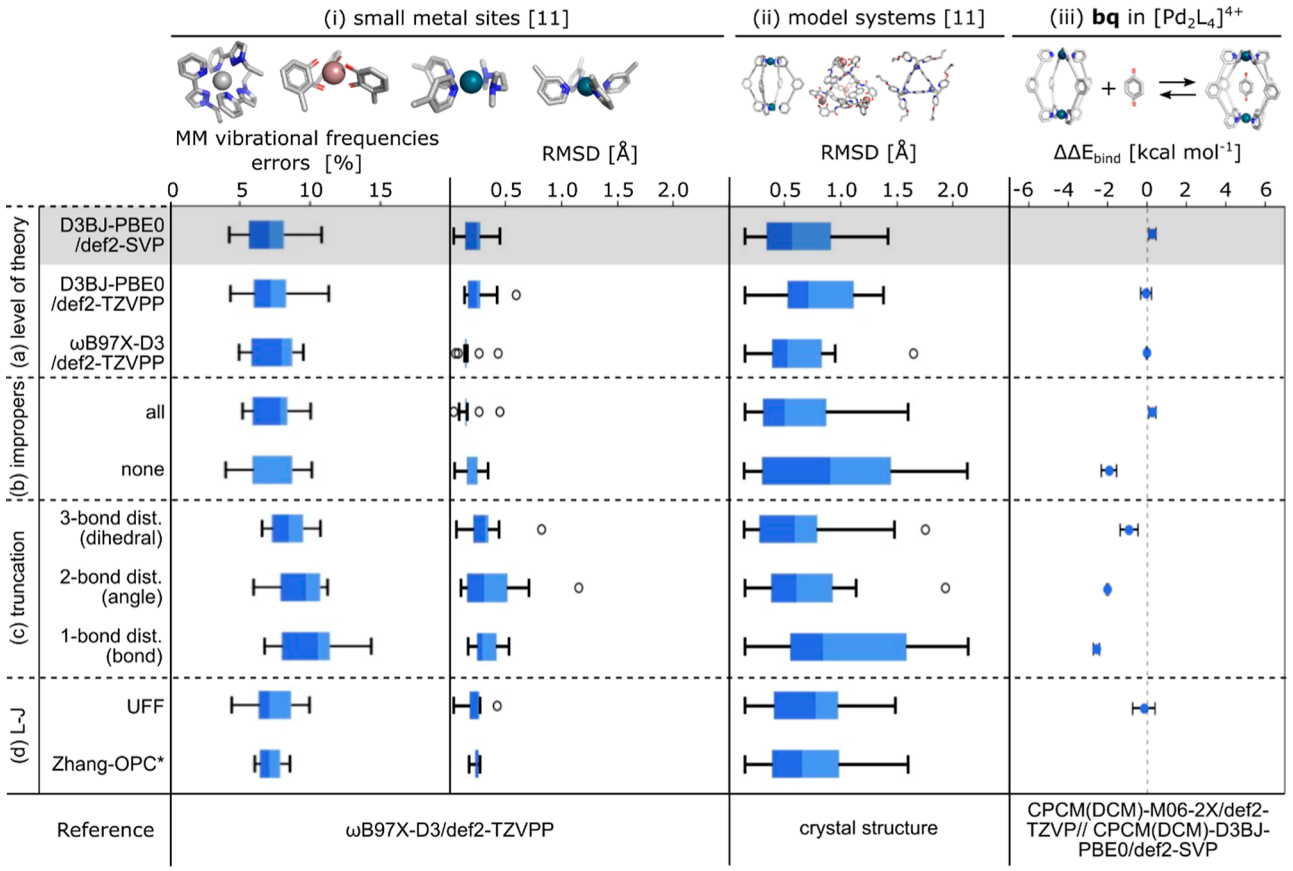



**Figure 6.** Benchmark of *metallicious* considering (i) 11 small metal sites, including a comparison of MM- and QM-computed vibrational frequencies and RMSD values between MM- and QM-optimized structures, (ii) 11 representative supramolecular systems, including comparison between MM-optimized and crystal structures, and (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 parameterization, (b) inclusion of improper parameters, (c) truncation scheme, and (d) type of L−J parameters (Zhang-OPC parameters are available only for nonpalladium systems). Except where specified by (a−d), the template was parameterized at the D3BJ-PBE0/def2-SVP level of theory; "none" for the truncation scheme; Merz-OPC L−J parameters and improper parameters included only for systems with square-planar complexes (gray shaded area). The dark and light blue colors represent lower and upper interquartile ranges, respectively.

8,[74] and ZIF-67[75] have been previously modeled 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 the 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 6i; see Figure S5 for all structures), RMSD values between MM-optimized and crystal structures of 11 representative supramolecular systems (Figure 6ii), and MM and QM binding energies for the benzoquinone (**bq**)-$[Pd_2L_4]^{4+}$ complex (Table S3), with the MM value obtained from five different starting structures (Figure 6iii).

By default, the templates were parameterized (Section 2.4.2) at the D3BJ-PBE0/def2-SVP level of theory without truncation using the Merz-OPC L−J parameters; improper parameters were included only for systems with square-planar complexes. Using the default parameters shows good performance across the different metrics evaluated (Figure 6). For example, when comparing vibrational frequencies for the small metal sites, the mean percentage error is only 6.7%, similar to the 6.4% obtained by Cole and co-workers for 70 small molecules using the modified Seminario method.[57] 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 parameterization obtained through *metallicious* is robust enough to obtain not only structural energetics but also host−guest energetics.

**3.1. Influence of the 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-TZVPP) 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-TZVPP level of theory, the 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


https://doi.org/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071

the loss of accuracy during parameterization originates from the inadequate depiction of interactions by conventional force fields. For example, harmonic bonds might poorly approximate metal−ligand interactions.[128] Since D3BJ-PBE0/def2-SVP is computationally efficient, it is used as a default methodology to parameterize templates in *metallicious*.

**3.2. Importance of Improper Dihedral Parameters Involving Metals.** Proper and improper dihedral parameters involving metal centers are rarely considered as they are generally thought to have little impact on geometry.[65−68] However, exceptions have been reported for square-planar $Ni^{2+}$ complexes, where including the out-of-plane bending term was necessary to reproduce crystal structures.[69] Similarly, during MD simulations of the $[Pd_2L_4]^{4+}$ cage in implicit solvent, we observed deviation from the expected square-planar geometry when improper dihedrals centered on the donor atom were not included (Figure 7a). Including improper dihedrals centered on

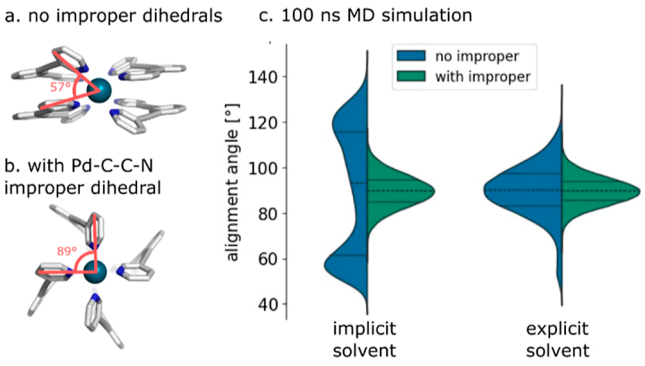



**Figure 7.** Importance of improper dihedral parameters involving metals. Energy-minimized structures using L-BFGS with GBSA implicit solvent (a) with and (b) without the improper dihedral. (c) Histogram of the angle between two ligands from 100 ns MD simulations of the $[Pd_2L_4]^{4+}$ cage with implicit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

the donor atom yielded structures similar to those obtained from QM in an implicit solvent (Figure 7b). Interestingly, these parameters can be omitted in explicit solvent, likely due to the dominant effect of solvent−solute interactions over solute−solute stacking interactions. For example, during 100 ns MD simulation in explicit dimethyl sulfoxide (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 (Figures 6b and S6). Otherwise, including improper parameters has no impact on the geometries. For this reason, and considering the computational cost associated with obtaining these parameters, their parameterization is disabled by default in *metallicious*, except for $Rh^+$, $Ir^+$, $Pd^{2+}$, $Pt^{2+}$, and $Au^{3+}$, which often have square-planar configurations.

**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)[129]

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

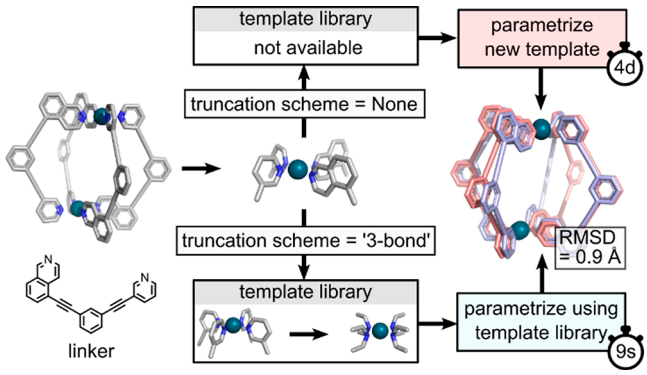



**Figure 8.** Truncation scheme. This scheme speeds up calculations by replacing missing templates with substructures from related templates. Timing is shown for a single CPU@2.5 GHz.

lacks a template, *metallicious* parameterizes a new template in around ∼100 CPUhs compared to only a few seconds when it uses a 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 parameterization with the least truncated template.

**3.4. L−J 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 the bonded parameters.

**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 parameterized and performed each MD simulation for 100 ns of 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. We compared the results with simulations using the nonbonded[29,85] and dummy metal models,[35,42,44] 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 Table S4 and Figures S7−S17 for RMSD, coordination sphere analysis, and snapshots of individual trajectories), with a median RMSD = 1.1 Å relative to the crystal structure. In contrast, most of the simulations utilizing nonbonded 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 counterions 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−counterion interactions.


https://doi.org/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071

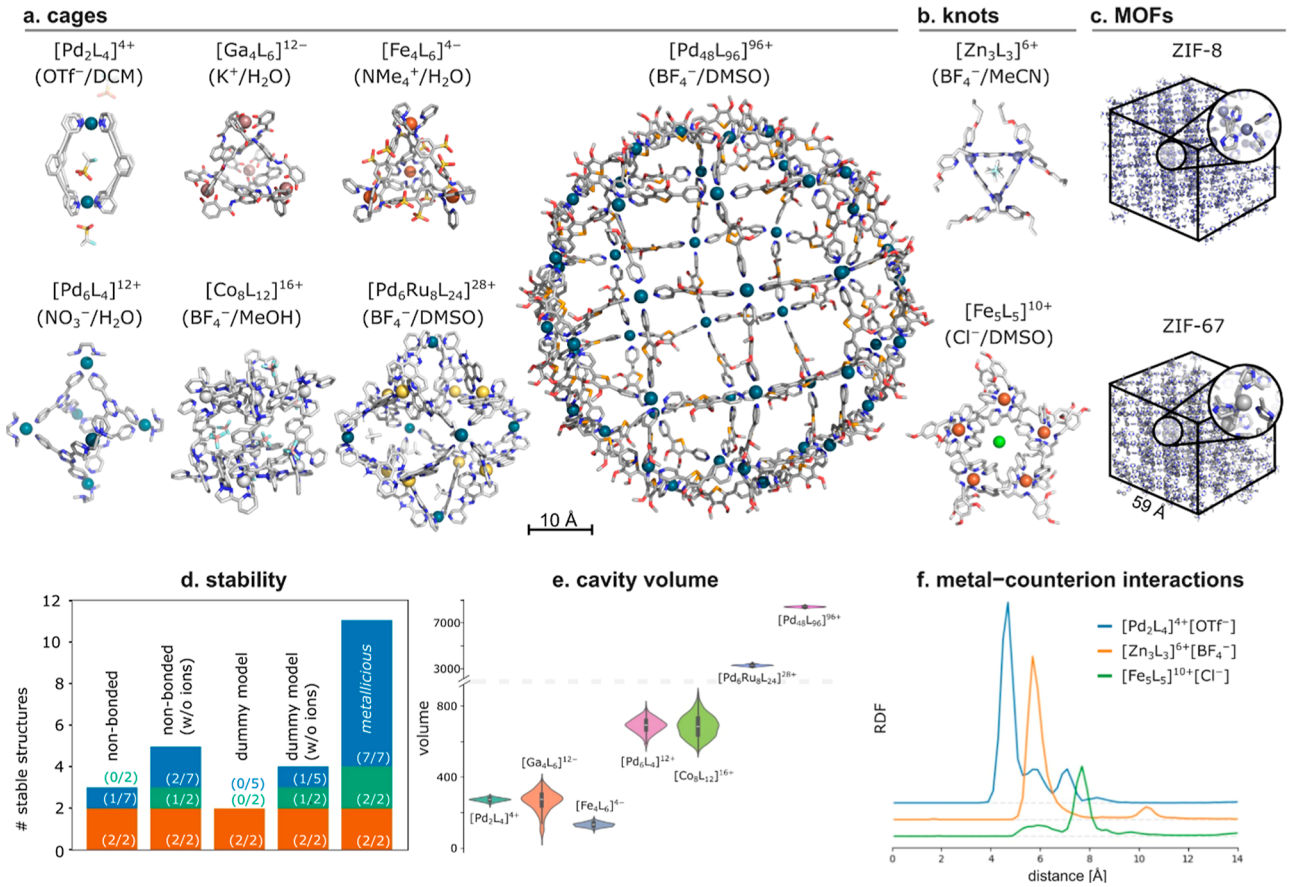



**Figure 9.** Analysis from 100 ns MD simulations for systems parameterized by *metallicious*. Final snapshots for (a) cages, (b) knots, and (c) MOFs (solvent and counterions not shown for clarity, except for $[Pd_2L_4]^{4+}$, $[Zn_3L_3]^{6+}$, and $[Fe_5L_5]^{10+}$). (d) Histogram of stable structures after simulation (blue—metallo-organic cages, green—supramolecular knots, and orange—MOFs). (e) Distribution of cavity volumes calculated using C3.[132] (f) RDFs calculated for metal centers and counterions (see Figure S20 for further details). DCM = dichloromethane; DMSO = dimethyl sulfoxide; MeOH = methanol; MeCN = acetonitrile; $OTf^-$ = triflate; $NMe_4^+$ = tetramethylammonium; $NO_3^-$ = nitrate; $BF_4^-$ = tetrafluoroborate.

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$^3$ (no grid point can be placed inside) to 267 Å$^3$ (Figure 9e; Figure S19). 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 to 160 Å$^3$.[127,130]

Furthermore, an analysis of the interactions between the supramolecular systems and counterions can be done by computing the corresponding radial distribution functions (RDFs; Figure S20). For example, the $[Pd_2L_4]^{4+}$ cage reveals 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.[115] Similar results were observed for the $[Zn_3L_3]^{6+}$ and $[Fe_5L_5]^{10+}$ knots, indicating a strong binding interaction at 5.7 and 7.7 Å, respectively. Indeed, Leigh and co-workers showed that the $[Fe_5L_5]^{10+}$ knot can extract chloride traces from solvent and glassware.[131]

## 4. CONCLUSIONS

We developed an automated tool called *metallicious* for parameterizing covalent metal models in supramolecular structures. Our method leverages the repetitive patterns of binding metal motifs in these structures. Several standard templates were parameterized and stored in a template library, which *metallicious* uses to parameterize input structures. Once *metallicious* identifies the template that matches the metal site in the structure, it copies the bonded parameters from the template and performs charge redistribution to account for charge transfer. To broaden the scope of the template library and increase efficiency, *metallicious* provides convenient truncation schemes, allowing for the recycling of available templates. In cases where no suitable template is found, *metallicious* automatically performs parameterization. The results of the benchmarks show good agreement with reference data obtained from QM and the crystal structure. While it can be argued that two other popular models, the nonbonded and cationic dummy atom model, offer more flexibility, the MD simulations conclusively show that the covalent metal model is the only robust option enabling simulations of these systems. Overall, *metallicious* provides a valuable resource for researchers working with metal-containing systems, facilitating their atomistic modeling in explicit solvent.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

The source code and associated Python files are freely available at https://github.com/duartegroup/metallicious. Documenta-



https://doi.org/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071

tion and tutorials are available at https://metallicious.readthedocs.io.

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.4c00850.

Implementation details, benchmark methods, and individual results (PDF)

MD topologies and input files (ZIP)

100 ns MD trajectories of the 11 supramolecular systems parameterized with *metallicious* and 3 supramolecular systems with the cationic dummy model (ZIP)

## AUTHOR INFORMATION


### Corresponding Author

**Fernanda Duarte** − *Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.;* orcid.org/0000-0002-6062-8209; Email: fernanda.duartegonzalez@chem.ox.ac.uk

### Authors

**Tomasz K. Piskorz** − *Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.;* orcid.org/0000-0003-0716-6874

**Bernadette Lee** − *Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.;* orcid.org/0000-0001-8548-6756

**Shaoqi Zhan** − *Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.; Department of Chemistry—Ångström, Ångströmlaboratoriet, Uppsala S-751 20, Sweden*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.4c00850


### Author Contributions

T.K.P. and F.D. conceptualized the study. T.K.P. carried out the calculations. All authors participated in data analyses and writing of the manuscript. B.L. and S.Z. performed tests. T.K.P. and F.D. wrote the first draft. F.D. supervised the study.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors thank V. Martí-Centelles and V. Jusraskova for helpful comments. 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 (EP/L015838/1) for a studentship, generously supported by GSK, MSD, Syngenta, and Vertex. S.Z. thanks the Swedish Research Council (Vetenskapsrådet) for an international postdoctoral fellowship (2021-00366).

## 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.

(2) Vardhan, H.; Verpoort, F. Metal-Organic Polyhedra: Catalysis and Reactive Intermediates. *Adv. Synth. Catal.* **2015**, *357* (7), 1351−1368.

(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.

(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.

(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.

(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.

(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.

(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.

(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.

(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.

(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.

(12) Li, P.; Merz, K. M. Metal Ion Modeling Using Classical Mechanics. *Chem. Rev.* **2017**, *117* (3), 1564−1686.

(13) Norrby, P. O.; Brandt, P. Deriving Force Field Parameters for Coordination Complexes. *Coord. Chem. Rev.* **2001**, *212* (1), 79−109.

(14) Li, W.; Wang, J.; Zhang, J.; Wang, W. Molecular Simulations of Metal-Coupled Protein Folding. *Curr. Opin. Struct. Biol.* **2015**, *30*, 25−31.

(15) Riccardi, L.; Genna, V.; De Vivo, M. Metal-Ligand Interactions in Drug Design. *Nat. Rev. Chem* **2018**, *2* (7), 100−112.

(16) 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 Catal.* **2022**, *12*, 5806−5826.

(17) Tarzia, A.; Jelfs, K. E. Unlocking the Computational Design of Metal-Organic Cages. *Chem. Commun.* **2022**, *58* (23), 3717−3730.

(18) 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.

(19) Coudert, F. X.; Fuchs, A. H. Computational Characterization and Prediction of Metal-Organic Framework Properties. *Coord. Chem. Rev.* **2016**, *307*, 211−236.

(20) 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), 75.

(21) Spicher, S.; Bursch, M.; Grimme, S. Efficient Calculation of Small Molecule Binding in Metal-Organic Frameworks and Porous Organic Cages. *J. Phys. Chem. C* **2020**, *124* (50), 27529−27541.

(22) 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.

(23) 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.

(24) 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.
(25) 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.
(26) Stote, R. H.; Karplus, M. Zinc Binding in Proteins and Solution: A Simple but Accurate Nonbonded Representation. *Proteins: Struct., Funct., Bioinf.* **1995**, *23* (1), 12−31.
(27) 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.
(28) 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.
(29) 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.
(30) 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.
(31) 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**, *61*, 4613−4629.
(32) 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.
(33) 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.
(34) 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.
(35) 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.
(36) 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.
(37) 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.
(38) 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.
(39) 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.
(40) 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.
(41) 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.
(42) 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.
(43) 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.
(44) 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.
(45) 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.
(46) 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.
(47) 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.
(48) 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.
(49) 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.
(50) 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.
(51) Biswal, D.; Kusalik, P. G. Probing Molecular Mechanisms of Self-Assembly in Metal-Organic Frameworks. *ACS Nano* **2017**, *11* (1), 258−268.
(52) Yoneya, M.; Tsuzuki, S.; Aoyagi, M. Simulation of Metal-Organic Framework Self-Assembly. *Phys. Chem. Chem. Phys.* **2015**, *17* (14), 8649−8652.
(53) Marrone, T. J.; Merz, K. Transferability of Ion Models. *J. Phys. Chem.* **1993**, *97*, 6524−6529.
(54) Sakharov, D. V.; Lim, C. Zn Protein Simulations Including Charge Transfer and Local Polarization Effects. *J. Am. Chem. Soc.* **2005**, *127* (13), 4921−4929.
(55) Sakharov, D.; Lim, C. Force Fields Including Charge Transfer and Local Polarization Effects: Application to Proteins Containing Multi/Heavy Metal Ions. *J. Comput. Chem.* **2008**, *30* (2), 191−202.
(56) Seminario, J. M. Calculation of Intramolecular Force Fields from Second-Derivative Tensors. *Int. J. Quantum Chem.* **1996**, *60* (7), 1271−1277.
(57) 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.
(58) Wang, R.; Ozhgibesov, M.; Hirao, H. Partial Hessian Fitting for Determining Force Constant Parameters in Molecular Mechanics. *J. Comput. Chem.* **2016**, *37*, 2349−2359.
(59) 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.
(60) Ercolessi, F.; Adams, J. B. Interatomic Potentials from First-Principles Calculations. *Mater. Res. Soc. Symp. Proc.* **1992**, *291*, 31−36.
(61) 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.
(62) 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.
(63) 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), 064108.
(64) Betz, R. M.; Walker, R. C. Paramfit: Automated Optimization of Force Field Parameters for Molecular Dynamics Simulations. *J. Comput. Chem.* **2015**, *36* (2), 79−87.
(65) Hoops, S. C.; Anderson, K. W.; Merz, K. M. Force field design for metalloproteins. *J. Am. Chem. Soc.* **1991**, *113* (22), 8262−8270.
(66) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Effective Tailor-Made Force Field Parameterization of the Several Zn Coordination


https://doi.org/10.1021/acs.jctc.4c00850
*J. Chem. Theory Comput.* 2024, 20, 9060−9071

Environments in the Puzzling FTase Enzyme: Opening the Door to the Full Understanding of Its Elusive Catalytic Mechanism. *Theor. Chem. Acc.* **2006**, *117* (1), 171−181.
(67) 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.
(68) 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.
(69) Adam, K. R.; Antolovich, M.; Brigden, L. G.; Lindoy, L. F. Comparative Molecular Mechanics Study of the Low-Spin Nickel(II) Complexes of an Extended Series of Tetraaza Macrocycles. *J. Am. Chem. Soc.* **1991**, *113*, 3346−3351.
(70) 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.
(71) 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.
(72) 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.
(73) 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.
(74) 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.
(75) 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.
(76) 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.
(77) 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.
(78) 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.
(79) 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.
(80) Wang, L. P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theory Comput.* **2013**, *9* (1), 452−460.
(81) Norrby, P.; Liljefors, T. Automated molecular mechanics parameterization with simultaneous utilization of experimental and quantum mechanical data. *J. Comput. Chem.* **1998**, *19* (10), 1146−1166.
(82) Rydberg, P.; Olsen, L.; Norrby, P. O.; Ryde, U. General Transition-State Force Field for Cytochrome P450 Hydroxylation. *J. Chem. Theory Comput.* **2007**, *3* (5), 1765−1773.
(83) Poole, D. A.; Bobylev, E. O.; Mathew, S.; Reek, J. N. H. Topological Prediction of Palladium Coordination Cages. *Chem. Sci.* **2020**, *11* (45), 12350−12357.
(84) Hu, L.; Ryde, U. Comparison of Methods to Obtain Force-Field Parameters for Metal Sites. *J. Chem. Theory Comput.* **2011**, *7* (8), 2452−2463.
(85) Rappe, 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.
(86) 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.
(87) 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.
(88) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F. Y. R. E. D. 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.
(89) 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.
(90) Wang, L.; O'Mara, M. L. PsiRESP: Calculating RESP Charges with Psi4. *J. Open Source Softw.* **2022**, *7* (73), 4100.
(91) Suzuki, K.; Tominaga, M.; Kawano, M.; Fujita, M. Self-Assembly of an M6L12 Coordination Cube. *Chem. Commun.* **2009**, *13*, 1638−1640.
(92) Sun, Q. F.; Iwasa, J.; Ogawa, D.; Ishido, Y.; Sato, S.; Ozeki, T.; Sei, Y.; Yamaguchi, K.; Fujita, M. Self-Assembled $M_{24}L_{48}$ Polyhedra and Their Sharp Structural Switch upon Subtle Ligand Variation. *Science* **2010**, *328* (5982), 1144−1147.
(93) Li, P.; Merz, K. M. MCPB.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, *56* (4), 599−604.
(94) Yu, H.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; MacKerell, A. D.; Roux, B. Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.* **2010**, *6* (3), 774−786.
(95) Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. Simulation of Ca2+ and Mg2+ Solvation Using Polarizable Atomic Multipole Potential. *J. Phys. Chem. B* **2006**, *110* (37), 18553−18559.
(96) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J. P. Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and Intramolecular Interactions and Ligand Macromolecule Complexes. A Bottom-up Strategy. *J. Chem. Theory Comput.* **2007**, *3* (6), 1960−1986.
(97) Spicher, S.; Grimme, S. Robust Atomistic Modeling of Materials, Organometallic, and Biochemical Systems. *Angew. Chem., Int. Ed.* **2020**, *59* (36), 15665−15673.
(98) 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.
(99) 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 *Proceedings of the 15th Python in Science Conference*; Benthall, S.; Rostrup, S., Eds.; Los Alamos National Laboratory, 2016, pp 98−105.
(100) RDKit: Open-source cheminformatics. http://www.rdkit.org (accessed Sep 20, 2024).
(101) ParmEd. https://parmed.github.io/ParmEd/html/index.html (accessed Sep 20, 2024).
(102) Hagberg, A. A.; Schult, D. A.; Swart, P. J. Exploring Network Structure, Dynamics, and Function Using NetworkX *Proceedings of the 7th Python in Science Conference*; Varoquaux, G., Vaught, T.; Millman, J., Eds.; SciPy Conferences: Pasadena, CA USA, 2008, pp 11−15.
(103) 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/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071

(104) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152* (22), 224108.
(105) 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.
(106) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25* (2), 247−260.
(107) Young, T. A.; Gheorghe, R.; Duarte, F. Cgbind: A Python Module and Web App for Automated Metallocage Construction and Host-Guest Characterization. *J. Chem. Inf. Model.* **2020**, *60* (7), 3546−3557.
(108) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminf.* **2012**, *4* (1), 17.
(109) 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.
(110) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0Model. *J. Chem. Phys.* **1999**, *110* (13), 6158−6170.
(111) 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.
(112) 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.
(113) 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.
(114) 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.
(115) 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.
(116) 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.
(117) 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.
(118) Takezawa, H.; Murase, T.; Resnati, G.; Metrangolo, P.; Fujita, M. Halogen-Bond-Assisted Guest Inclusion in a Synthetic Cavity. *Angew. Chem., Int. Ed.* **2015**, *54* (29), 8411−8414.
(119) 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. *Chem.—Eur. J.* **2018**, *24*, 1554−1560.
(120) 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.
(121) 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.
(122) 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.
(123) 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.
(124) 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.
(125) 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.
(126) 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.
(127) 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.
(128) Burton, V. J.; Deeth, R. J.; Kemp, C. M.; Gilbert, P. J. Molecular Mechanics for Coordination Complexes: The Impact of Adding d-Electron Stabilization Energies. *J. Am. Chem. Soc.* **1995**, *117* (32), 8407−8415.
(129) 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.
(130) 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.
(131) August, D. P.; Borsley, S.; Cockroft, S. L.; della Sala, F.; Leigh, D. A.; Webb, S. J. Transmembrane Ion Channels Formed by a Star of David [2]Catenane and a Molecular Pentafoil Knot. *J. Am. Chem. Soc.* **2020**, *142*, 18859−18865.
(132) Martí-Centelles, V.; Piskorz, T. K.; Duarte, F. CageCavityCalc (C3): A Computational Tool for Calculating and Visualizing Cavities in Molecular Cages. *J. Chem. Inf. Model.* **2024**, *64*, 5604.


https://doi.org/10.1021/acs.jctc.4c00850
J. Chem. Theory Comput. 2024, 20, 9060−9071