Computational Screening of Putative Catalyst Transition Metal Complexes as Guests in a $Ga_4L_6^{-12}$ Nanocage

Clorice R. Reinhardt,[1,2] Melissa T. Manetsch,[1,2] Wan-Lu Li,[3,4,5,#] Teresa Head-Gordon[3,4,5,6,7] and Heather J. Kulik[1,2,*]

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2]*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[3]*Kenneth S. Pitzer Center for Theoretical Chemistry, Berkeley, CA 94720, USA*
[4]*Department of Chemistry, University of California, Berkeley, CA 94720, USA*
[5]*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[6]*Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, United States*
[7]*Department of Bioengineering, University of California, Berkeley, CA 94720, United States*
[#]Current address: *Department of NanoEngineering, University of California San Diego, CA 92093, United States*
*email: hjkulik@mit.edu

ABSTRACT: Metal-organic cages form well-defined microenvironments that can enhance the catalytic proficiency of encapsulated transition metal catalysts (TMCs). We introduce a screening protocol to efficiently identify TMCs that are promising candidates for encapsulation in the $Ga_4L_6^{-12}$ nanocage. We obtain TMCs from the Cambridge Structural Database with geometric and electronic characteristics amenable to encapsulation and mine the text of associated manuscripts to curate TMCs with documented catalytic functionality. By docking candidate TMCs inside the nanocage cavity and carrying out electronic structure calculations, we identify a subset of successfully optimized candidates (TMC-34) and observe that encapsulated guests occupy an average of 60% of the cavity volume, in line with previous observations. Notably, highly charged guests occupy as much as 72% of the cavity as a result of linker rotation. Encapsulation has a universal effect on the electrostatic potential, systematically decreasing the electrostatic potential at the metal center of each TMC in the TMC-34 dataset, while minimally altering TMC metal partial charges. Collectively these observations support geometry-based screening of potential guests and suggest that encapsulation in Ga_4L_6 cages could electrostatically stabilize diverse cationic or electropositive intermediates. We highlight candidate guests with associated known reactivity and solubility most amenable for encapsulation in experimental follow-up studies.

1. Introduction

Synthetic chemists have long sought to mimic the catalytic effect of the enzyme environment around an active site. One promising direction involves the use of a molecular capsule or cage, acting as a "host" to bind a "guest" reactant or smaller catalyst, in the same way an enzyme binds a small molecule substrate or utilizes a catalytic residue in the active site. Since the initial discovery of crown-ethers[1], cryptands,[2,3] and other molecular capsules, the field of supramolecular chemistry has expanded significantly as larger and more elaborate hosts are designed, including metal-organic cages.[4-6] Metal-organic cages are composed of metal vertices and organic, often aromatic linkers. Due to the organic nature of the linkers, metal-organic cages are often stable in organic solvents, although some can also be stable in polar solvents such as water.[7] The identity of the coordinating metal and linker molecules that form the cage walls can be modified to change the solubility or electronic properties of the cage.[7,8] Particularly well-studied cages include the $[Pd_6L_4]^{+12}$,[9] $[Co_4L_2]^{-8}$,[10] $[Pd_2L_4]^{4+}$,[11] $[In_4L_6]^{-12}$,[8] and $[Ga_4L_6]^{-12}$ cages.[12] Numerous other examples of cages with octahedral,[13] cubic,[14] and tetrahedral geometries exist,[15] with increasing diversity in other cage topologies.[16]

As more examples of cages emerge in the literature, computational studies have become central to understand the physical principles behind host-guest recognition,[17-20] predict the assembly of novel cage architectures,[16,21,22] and identify opportunities for cage-mediated catalytic rate enhancement.[8,23] Experimental and computational efforts to understand the physics driving host-guest recognition have identified electrostatics,[24] guest size,[25] properties of counterions[26], and solvent effects[20] as factors that strongly influence host-guess affinity. Once a catalytic guest is bound by the host, researchers have identified that the local pH inside the cage,[27] transition state stabilization,[8,23,24] and electrostatic effects[24,27] can all contribute to enhanced catalytic rates and

selectivity observed within many cage environments. Because the physical features that most strongly impact binding or catalysis can vary with cage architecture, guest characteristics, and solvent type, it is challenging to identify promising guests according to a universal chemical or physical descriptor of the guest.

While numerous computational studies have focused on cage design or encapsulation effects,[28] few studies have systematically identified promising guest candidates for existing metal-organic cages. Notable exceptions include work by Ward and colleagues,[29,30] in which they applied a docking protocol to screen organic guests for encapsulation inside a $[Co_8L_{12}]$ octahedral cage. While this approach worked with multiple guests, it necessitated refitting the weights of specific terms in the docking energy expression against a training set of experimental data, limiting its applicability to cages where experimental data is available. More recent work demonstrated the use of a machine learning model to identify promising small organic guests for encapsulation in cucurbituril CB[6] and Pd_2L_4 using spatial and electrostatic host-guest compatibility.[31] This effort leveraged electrostatic and non-covalent interactions to predict new small organic guests but is not suited in its current form for large cage cavities or more diverse guests because of the size of molecules in its training data. Notably, these screening approaches have not been extended to diverse transition metal-containing guests, where additional complexities arise from variable oxidation and spin states of transition metal centers. While transition metal complexes (TMCs) introduce computational modeling challenges, these guests can have desirable reactivity, such as cross-coupling[32], for which additional rate enhancement by encapsulation would be beneficial.

Here we identify promising TMC guests by systematically screening the chemical literature for catalytic TMCs that are suited for encapsulation in a representative $[Ga_4L_6]^{-12}$ cage. The $[Ga_4L_6]^{-12}$ nanocage is a well-characterized anionic cage formed of N,N'-bis(2,3-dihydroxy

benzoyl)-1,5-diaminonapthalene linkers and gallium vertices and has demonstrated promise in enhancing reductive elimination by TMCs (Figure 1).[33] We screened the Cambridge structural database (CSD)[34] for candidate TMCs with desirable size and catalytic properties for potential guest molecules. Over a subset of 34 TMCs successfully docked into the nanocage (TMC-34), we identified universal effects of encapsulation on electrostatic properties but variable enlargement of the cage depending on the size and charge of the guest. We leverage these insights to propose catalyst and reactant complexes from the TMC-34 set that would be most promising for follow-up encapsulation-accelerated catalysis study in experiments.

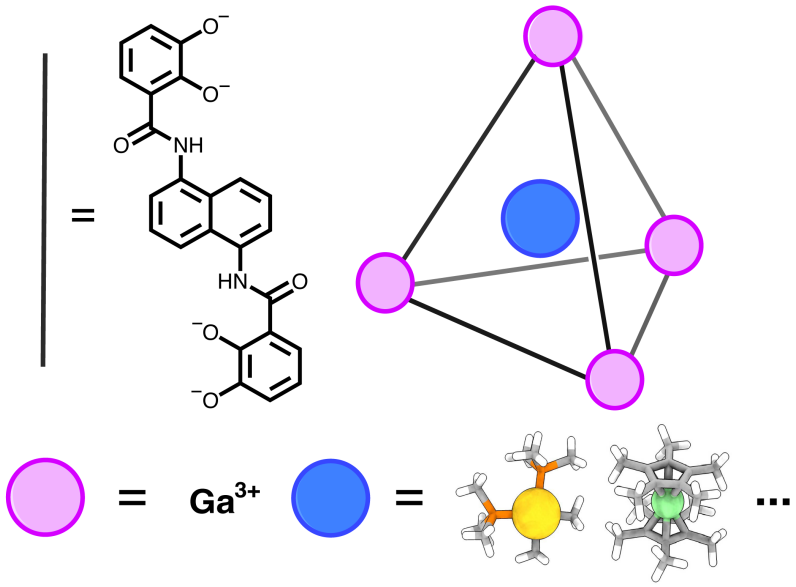


Figure 1. Architecture of the $[Ga_4L_6]^{-12}$ nanocages with two examples of transition metal complexes previously characterized as guests and crystalized shown in licorice representation. The cage linker (N,N'-bis(2,3-dihydroxy benzoyl)-1,5-diaminonapthalene) is represented by black lines that correspond to the provided chemical structure. TMC guests are the bis(trimethylphosphine) gold cation and $Cp^*_2Co^+$ as described in refs [35] and [36] respectively. For structures, gray is used for carbon atoms, white for hydrogen atoms, yellow for gold atoms, and light green for cobalt atoms.

2. Computational Details

2a. Transition Metal Complex Screening

We started from a previously curated dataset of mononuclear transition metal complexes[37] from the Cambridge Structural Database (CSD). Briefly, this set contains 204,981 unique complexes consisting of mononuclear 3d, 4d, or 5d transition metal complexes and was curated from CSD version 5.41 (Nov. 2019) with March and May 2020 updates included. In this previous work, [37] a corpus was constructed from downloadable manuscripts mapped from CSD IDs using the ArticleDownloader package[38] following protocols previously detailed by our group.[39-41] Consistent with the cationic character of TMC guests in previous experimental studies of Ga_4L_6 guests[35,42,43] and the anionic nature of the cage, we filtered this subset of the CSD to obtain TMCs with an overall positive charge. We constrained the dataset to compounds that either contained non-disordered crystallographic hydrogens or where any added hydrogens were deemed "trustworthy" by the CSD API. Because square planar complexes are likely to feature open catalytic sites, we further curated the dataset for these coordination geometries.

We also computed the vdW volumes of all TMCs as a prerequisite for encapsulation. Molecular volumes were evaluated using the standalone program MoloVol v1.0.0[44] with a single-probe method and a grid size of 0.1 Å. Following literature precedent, we employed the standard radii for elements and set the probe size to 1.1 Å to mimic hydrogen.[45] We enforced a TMC volume cutoff of < 300 $Å^3$, motivated by the estimate of cage cavity size and known guest size reported in a prior study[36] or calculated in this study (Supporting Information Table S1). Prior work indicated that the cage encapsulated guests with volumes ranging from 138 $Å^3$ for tetraethyl ammonium cation to 291 $Å^3$ for cobalt metallocene cation[36], and the volume of the cage cavity ranged from 253 to 434 $Å^3$ for these same guests. We further down-selected complexes to only include those that had an even number of electrons, as determined by counting the number of electrons based on the molecular formula and the net charge present in the CSD entry (see Sec. 3).

We next identified complexes with manuscripts that studied catalysis following a similar procedure to our previous work.[40,41,46] Briefly, the DOI associated with each candidate structure was searched against the aforementioned corpus, and complexes with inaccessible manuscripts were excluded. With the accessible manuscripts, the main text was mined for keywords such as "catalysis, kinetic studies, reaction rate", along with capitalization permutations of those phrases (Supporting Information Table S2). We disregarded the first 25% of the paper to avoid false positive matches from the introductory text, which was previously determined to be a good threshold for metal-organic framework literature.[46] As we wanted to avoid discarding potentially catalytically relevant compounds, we carried out further review on any paper that had at least one catalysis keyword. We then manually inspected these papers to ensure that they did not only contain synthetic methods or structural characterization and that the keywords were associated with the relevant TMC (Supporting Information Table S3).

For all complexes that passed the catalytic-relevance test, we prepared them for encapsulation through a simple geometric docking approach. We computed the centroid of each guest without mass weighting along with the centroid of the cage. We then applied a coordinate translation to the guest's coordinates to superimpose the two centroids. The in-house script that was used to carry out this transformation is available in the Zenodo repository.[47] This avoided the need for docking with an energy function, as had been employed previously for a different cage[34] but which would be difficult to implement for TMCs due to the lack of available parameters.

2b. Electronic Structure Calculations

We performed density functional theory (DFT) calculations with a developer version of Terachem (v1.9-2018)[48,49]. We optimized all TMCs and nanocages with the B3LYP functional[50-52] and

applied a semi-empirical D3 dispersion correction with default Becke–Johnson damping.[53] All DFT calculations in Sec. 3 were carried out with an LACVP* composite basis set, corresponding to an LANL2DZ effective core potential[54,55] for heavier elements (i.e., Pd, Pt, Ni, Rh, Au, Ir, Cu, Hg, Ag, and Ga) and a 6-31G* basis set[56-58] for the lighter elements. To confirm limited sensitivity of the cage geometry to the electronic structure method, we performed additional calculations with PBE[59] and ωB97X[60] in TeraChem as well as GFN2-xTB in the standalone CREST[61,62] program (Supporting Information Tables S4–S5 and Figure S1). We observed minimal DFT functional dependence, with all optimized geometries within 0.5 Å RMSD of the X-ray crystal structure[35] of encapsulated bis(trimethylphosphine) gold cation. We performed xTB optimizations with the standalone xTB program version 6.6.0,[63] using tight optimization criteria and analytical linearized Poisson–Boltzmann (ALPB) implicit solvation.[64] Prior to geometry optimization with DFT, we carried out conformer searching using GFN-2xTB and CREST[61,62] to identify the lowest-energy conformer (Supporting Information Text S1 and Figure S2).

To aid self-consistent field (SCF) convergence of DFT calculations in TeraChem, we used the hybrid direct inversion in the iterative subspace (DIIS) combined with the augmented form (ADIIS) method[65], and level-shifting of 0.25 hartree was applied to the virtual orbitals.[66,67] A denser grid (dftgrid = 3) than the default (dftgrid = 1) was used for all calculations to improve convergence of the geometry optimizations and reduce numerical noise. The translation-rotation-internal coordinate (TRIC) algorithm for assigning coordinates for geometry optimization was used with the L-BFGS algorithm for optimization.[68] Complexes were solvated implicitly with the conductor-like polarizable continuum model (C-PCM) as implemented in TeraChem,[69-71] with water ($\varepsilon = 78.39$) and methanol ($\varepsilon = 32.70$) dielectric constants. In all cases, the solute atomic radii used for forming the cavity were set as the Bondi[72] radii scaled by 1.2, following standard

practice.[69] Default radii provided by TeraChem were used for all elements except for gold, which we instead determined from the original Bondi[72] radii publication.

To ensure that the ground state electron configurations of transition metals in the guest complexes were closed-shell singlets, we optimized structures in implicit solvent using restricted DFT in singlet states or unrestricted DFT in triplet and quintet states and compared the electronic energies. For the unrestricted calculations, we used level-shifting of 0.25 hartree on the majority- and minority-spin orbitals, and no significant spin contamination was observed with all complexes having $\langle S^2 \rangle$ that deviated from the expected $S(S+1)$ value by less than 1.0. Because the energetic difference between spin states may be sensitive to the amount of Hartree–Fock exchange as well as functional family or basis set choice,[73-79] we excluded any guest complexes with singlet-triplet spin splitting energies less than 20 kcal/mol when evaluated with B3LYP-D3(BJ)/LACVP* upon optimization in C-PCM water implicit solvent. Results were consistent within 0.1–0.2 kcal/mol if methanol implicit solvent was considered instead with only a few exceptions where geometry optimizations differed (Supporting Information Table S6 and Figures S3–S4).

To obtain qualitative trends in the host-guest affinity between the cage and docked TMCs, we computed the rigid binding electronic energies for each optimized host-guest complex. In this manuscript, we defined binding as encapsulation into the cages central cavity as seeded by the docking procedure and not including other interactions that could contribute experimentally, such as the association of the TMC to the nanocage surface[80,81], due to the large number of structures being optimized and the difficulty of obtaining quantitatively accurate binding free energies for host-guest systems.[82-84]

2c. Electrostatic Potential Calculations and Analysis

We computed the electrostatic potential (ESP) at the metal center of each transition metal complex guest following our previous work.[85,86] Briefly, we computed the partial charges on each atom in a complex using iterative Hirshfeld (Hirshfeld-I) charges,[87] as implemented in Multiwfn version 3.7 (Supporting Information Text S2).[88] We selected the Hirshfeld-I scheme since previous reports indicate that it accurately approximates the electrostatic potential[89] and is relatively insensitive to the level of theory[90] or basis set. Then, we applied the partial charges to compute the electrostatic potential at the atom of interest as follows:

$$V_a = \frac{1}{4\pi\varepsilon_0\kappa}\sum_{i\neq a}\frac{q_i}{r_{ia}} \quad (1)$$

Consistent with previous work, we estimated the effective dielectric constant inside the cage, κ, as 10% of the dielectric constant of the bulk solvent outside of the cage.[20] Thus, we used effective dielectric constants of $\kappa = 7.839$ (water) and $\kappa = 3.24$ (methanol) to account for intra-cage electrostatic screening in systems optimized with water and methanol implicit solvent environments respectively. The in-house script that was used to calculate the ESP is available in the Zenodo repository.[47]

2d. Calculation of Cavity Size

To quantify the volume of cage cavities after encapsulating each guest, we first removed the encapsulated guest and evaluated the molecular volumes with a two-probe (i.e., using a large and small probe) mode in MoloVol, as recommended[44] to avoid difficulties in delineating the boundary of the enclosed region adjacent to cavity openings.[44,91-93] If the large probe is too small, it will not block off the central cavity from the outside, which will result in small or discontinuous cavities. We calculated the largest cavity volume across the various host-guest combinations with large probe sizes ranging from 3.0 to 5.0 in 0.1 Å increments, selecting a radius of 3.6 Å because it led

to consistent behavior across all TMC-34 guests without discontinuous cavities (Supporting Information Figures S5 and S6). For final calculations, we also used the default[44] grid size of 0.2 Å and a small probe size of 1.1 Å to be consistent with our molecular volume calculations (sec. 2a).

3. Results and Discussion

3a. Screening Workflow and Generated Datasets

To identify synthesizable guest complexes that could be encapsulated in the Ga_4L_6 anionic nanocages, we developed a procedure to screen the Cambridge Structural Database (CSD) and mine the text of associated papers. We focused our search on square planar TMCs, which have previously demonstrated catalytic activity in Ga_4L_6 nanocages.[35,94] Leveraging a previously curated[37] dataset of 242,829 mononuclear TMCs (see Computational Details) from the CSD, we identified 49,643 square planar complexes. As described in Sec. 2, additional quality control metrics were applied to ensure that the hydrogens added by the CSD python API were reliable, yielding a remaining total of 35,874 complexes (Figure 2).

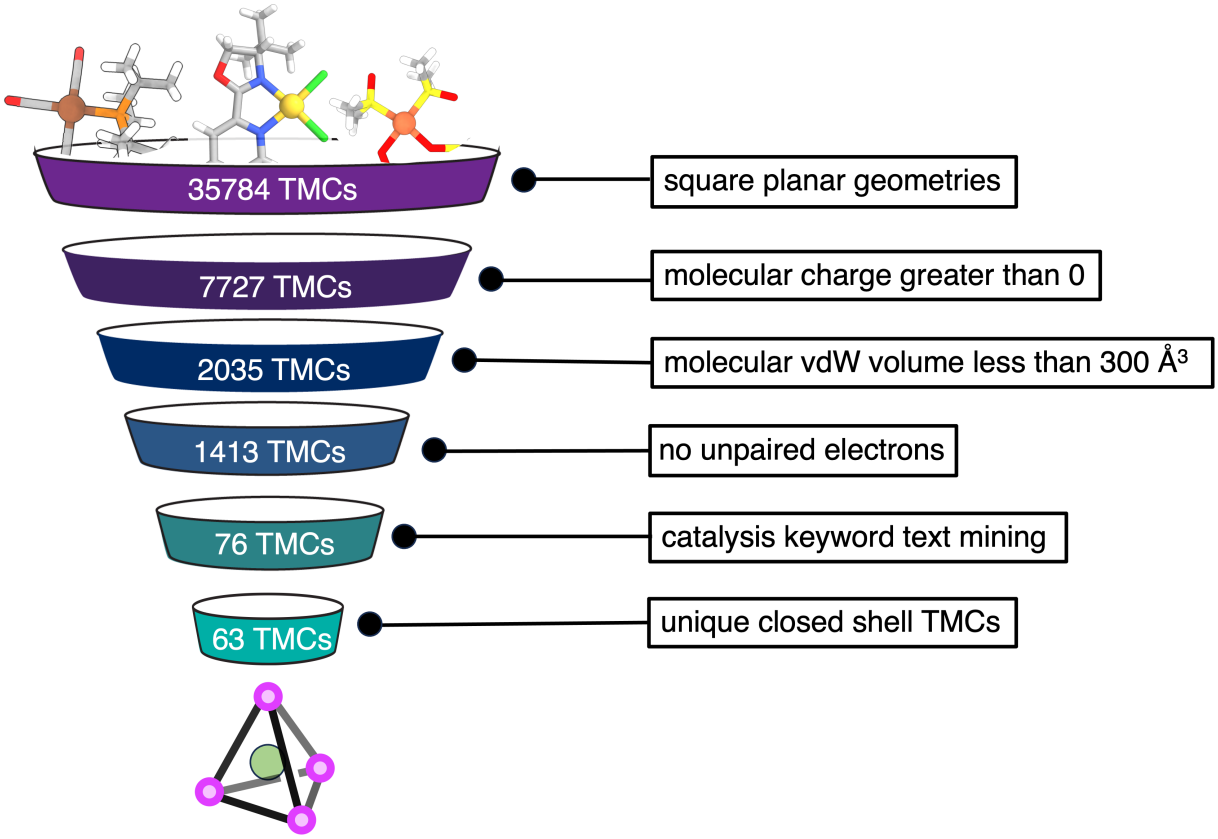


Figure 2. A funnel summarizing the screening workflow we apply to identify transition metal complex guests for encapsulation into the Ga_4L_6 nanocage. Each text box lists a filtering criterion applied to downselect the dataset to the number of structures labeled in the corresponding funnel segment. The 63 TMCs remaining after these filtering stages were studied further with electronic structure methods as candidates for encapsulation in the cage.

Because positively charged guests bind electrostatically to the highly anionic cage, we further filtered the dataset to include only cationic TMCs, eliminating a large portion of the dataset and retaining 7,727 complexes (Figure 2). We computed the vdW volume of the remaining compounds using MoloVol[44] to ensure that candidate TMCs were small enough to fit in the cage. Here, we chose a volume cutoff of < 300 Å^3 for potential guests based on a previous literature study[36] (see also Sec. 2). After imposing these size constraints, we eliminated nearly two-thirds of cationic TMCs from the dataset, retaining 2,035 small, cationic, square planar TMCs (Supporting Information Figure S7). We next identified complexes with even numbers of electrons (i.e., that

could form closed-shell species), reducing this set to 1,413 TMCs. Of this set, about half (i.e., 708 of 1,413) of all associated manuscripts were downloadable, and we identified 129 to have matches for catalysis keywords (see Sec. 2).

Manual review of these papers led to identification of 76 final TMCs. To identify if any duplicates had been missed at earlier stages of filtering, we both visually inspected the complexes and computed their molecular graph determinant.[39] We identified two pairs of duplicates and eliminated the structure from each pair with the higher R factor value (Supporting Information Table S7). We validated the closed-shell assumption on these 74 TMCs, eliminating 11 structures with low-lying triplet spin states (see Sec. 2), and achieving an initial dataset of candidate TMCs for encapsulation of 63 complexes, which we refer to as TMC-63 (Supporting Information Table S6). This set of selection criteria enabled us to select *a priori*, the crystal structure[35] of a gold cation (refcode: OVUHUK), which was experimentally encapsulated in the Ga_4L_6 cage, demonstrating the utility of this workflow to identify viable candidates for encapsulation into the cage.

After eliminating a large majority of the initial TMC candidates, we characterized the final selection of structures in terms of the transition metal identity and ligand denticity. The initial CSD set of 35,784 square planar structures are largely composed of Pd, Pt, Ni, Cu and Rh metal centers, which is consistent with d^8 electronic configurations. In the TMC-63 dataset, we observe many of the most common metal centers, with some notable changes. Because we eliminated TMCs with predicted odd numbers of electrons, Cu metal centers are underrepresented, as predominantly Cu(II) metal centers with d^9 configurations have likely been removed (Figure 3 and Supporting Information Table S8 and Figure S8). All types of ligand denticities present in the initial set are preserved after filtering in the TMC-63 set, but the complexes consisting of two bidentate ligands

are dramatically decreased, likely due to their limited association with catalysis keywords (Supporting Information Figure S9). We observe monodentate, bidentate, and tetradentate ligands such as macrocycles and other fused ring structures in both the original 35,784 structures and in TMC-63 (Supporting Information Figure S9). However, the volume filtering step removes a large proportion of TMCs with four nitrogen coordinating atoms, a configuration that is characteristic of nitrogen-containing macrocycles such as porphyrins that are likely to be too bulky for encapsulation (Supporting Information Figure S10). Amongst TMC-63, the most prevalent motif is a bidentate ligand coordinating through nitrogen atoms paired with two monodentate ligands, such as CO or a methyl group (Supporting Information Figure S10).

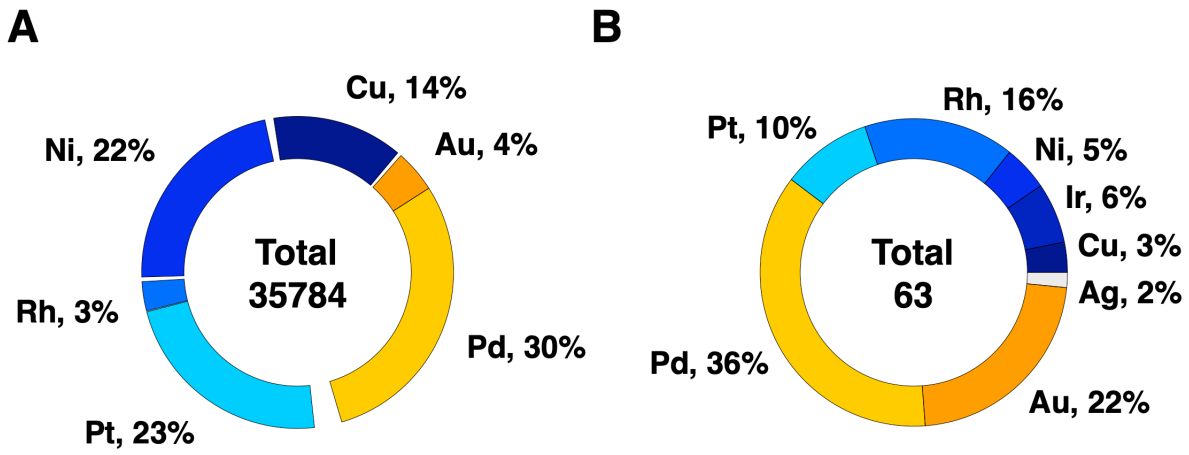


Figure 3. Pie charts showing the distribution of TMC metal center identity among A) all square planar complexes in comparison to B) the structures in TMC-63 identified to be potential guests of the Ga_4L_6 nanocage. Only unique TMCs pursued for docking into the cage are present in B and metals that make up less than 1% of the total are omitted for clarity in A.

3b. Geometric Changes to TMCs Upon Encapsulation

To establish a baseline for properties of the TMCs in TMC-63 prior to encapsulation, we optimized each geometry in C-PCM implicit methanol and water solvents to mimic solvent environments that are experimentally compatible with this nanocage.[12] Among optimized structures in the TMC-

63 dataset, we observe a median heavy-atom root-mean-square deviation (RMSD) of 0.21 Å, indicating that DFT optimization does not substantially alter the geometry of most TMCs from the crystal structure. In both solvents (i.e., water or methanol), the average RMSD of 0.41 Å is expectedly larger, while the minimum RMSD value of 0.04 Å indicates some structures are nearly unchanged during optimization. The large difference between the median and average is driven by greater than 1.0 Å RMSD values for six complexes that have ligands with rotatable bonds, such as methyl or phenyl groups (Supporting Information Figures S11–S12). Some high RMSDs could also be attributed to the conformational sampling performed (see Sec. 2) prior to DFT optimization. Overall, our analysis indicates most of TMC-63 consists of relatively small and rigid transition metal complexes that change little from their crystal structure, with a few outliers.

Next, we docked the optimized, isolated TMC structures into the nanocages by aligning the centroid of the TMC and the centroid of the cage. From this docked structure we optimized the full TMC-cage system (see Sec. 2). Of the initial set of 63 compounds, only slightly more than half (i.e., 34) were successfully optimized, and we refer this set of compounds retained for further analysis as TMC-34. Optimized geometries are available in the Zenodo repository.[47] The failed optimizations were predominantly cases in which TMC atoms were too close to neighboring cage atoms. The initial shortest distance between an atom in the TMC and a cage atom of an unsuccessfully docked complex averaged 0.55 Å, whereas successfully optimized complexes typically had larger separation from the cage, and all complexes initially distant by at least 1 Å completed successfully (Figure 4 and Supporting Information Figure S13). This suggests that in future work, success rates could likely be improved by a docking procedure in which rigid rotation of the TMCs was used to attempt to avoid close contacts with the cage before optimization.

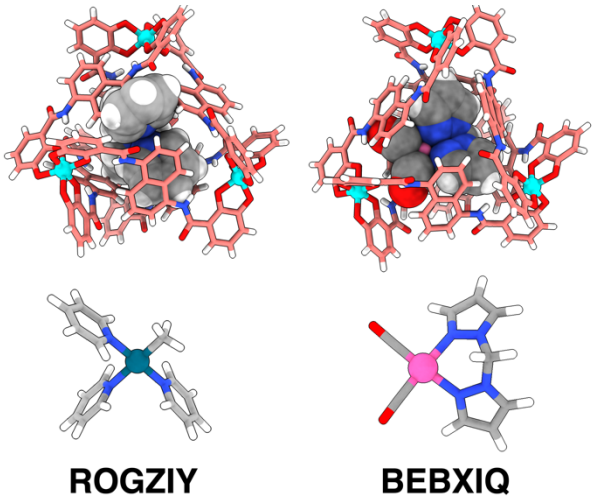


Figure 4. Representative docked structures with small (left) and large (right) interatomic distances. Space-filling models are used for the docked TMCs in the top panel to visualize the short interatomic distances, and the TMC structure with its CSD refcode is given in the bottom panel. For cage structures, pink coloring is used for carbon atoms, white for hydrogen atoms, blue for nitrogen atoms, cyan for gallium metal ions, and red for oxygen atoms. For the TMC guest structures, carbon atoms are grey, hydrogen atoms are white, and oxygen atoms are red. Dark blue-green and pink are used for the palladium and the iridium metal ions, respectively.

For each guest that successfully optimized in the cage, we quantified both global and local geometric changes. Using RMSD as a global metric, we observe an average change of only 0.13 Å between the solvated and encapsulated TMC structures (Supporting Information Figure S14). To characterize the local metal structure we defined a scaled metal–ligand bond length, d_{rel}(M–X) evaluated relative to the sum of the covalent radii[95] of each ligand element, X, with transition metal M:

$$d_{\mathrm{rel}}(\mathrm{M}-\mathrm{X}) = \frac{d(\mathrm{M}-\mathrm{X})_{\mathrm{opt}} - d(\mathrm{M}-\mathrm{X})_{\mathrm{crystal}}}{r_{\mathrm{M}} + r_{\mathrm{X}}} \quad (1)$$

and we averaged this quantity over all metal-ligand bonds in the complex. Consistent with the global RMSD metric, we observe that encapsulation minimally alters metal-local TMC structure, with average changes in scaled metal–ligand bond lengths of only 0.019 from the crystal structure reference (Figure 5). Nevertheless, three outliers are observed: MUPHIP, MUGJOQ, and YIDLOP, which are the only TMCs in TMC-34 with Au–Cl bonds (Figure 5). The TMC with the

largest RMSD, MUPHIP, which is 1-(2-(amino)ethyl)-3-methyl-1H-imidazol-3-ium)-trichloro-gold, has an average Au–Cl scaled bond length change of 9% upon encapsulation with respect to the crystal structure. In unscaled bond length terms, this corresponds to an increase from 2.28 Å in the crystal structure to 2.37 Å in the DFT optimized structure to 2.48 Å after encapsulation, potentially indicating that the Au–Cl bond is uniquely weakened and thus more susceptible to cleavage inside the cage (Supporting Information Table S9). A halide pre-dissociation step has been proposed for mechanisms involving reductive eliminations from Au(III) halides in the Ga_4L_6 nanocage,[35,96] and it is an intriguing possibility that the dissociation could be more favorable inside the cage as a result of host-guest interactions.

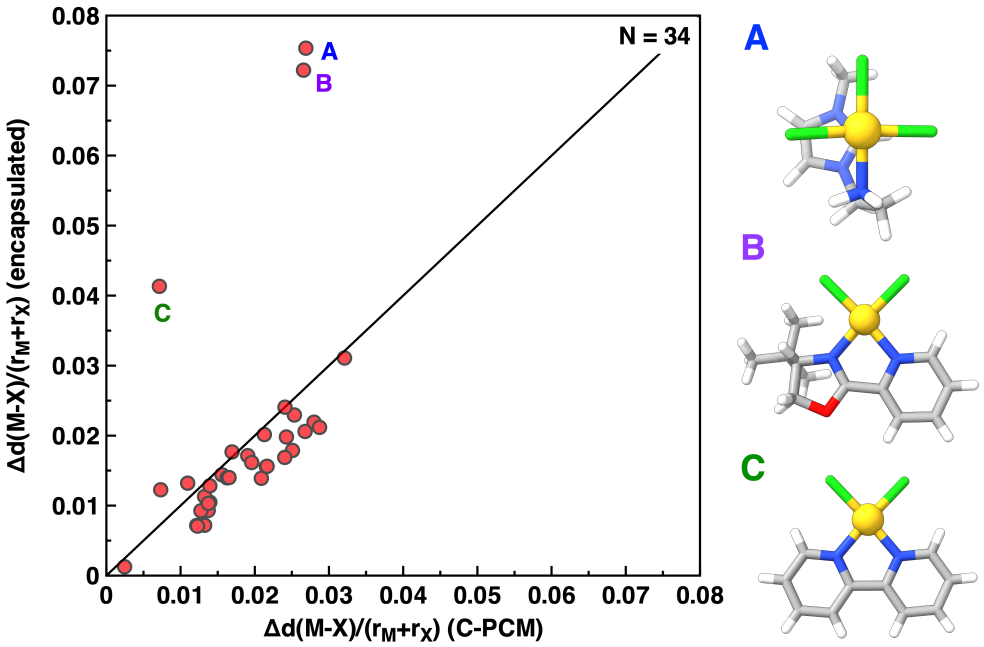


Figure 5. Average change in the scaled bond length ratio relative to the crystal structure value for optimized TMCs. The covalent radii for all element are defined in Ref. [95], and the x-axis corresponds to TMCs optimized in water C-PCM solvent, where the y-axis corresponds to the same TMC optimized in the cage environment. The four metal–ligand bonds are averaged for each complex in this plot to produce one data point per TMC. A) Structure of MUPHIP, B) structure of MUGJOQ, and C) structure of YIDLOP. For structures, gray coloring is used for carbon atoms, white for hydrogen atoms, blue for nitrogen atoms, green for chlorine atoms, yellow for gold atoms, and red for oxygen atoms.

3c. Energetics and Conformational Changes in the Cage Upon Encapsulation

While most TMC guest geometry is not substantially altered by encapsulation, we observe pronounced differences in nanocage geometry after binding diverse guests. The average heavy-atom RMSD of cage atoms with respect to a reference X-ray crystal structure is 0.89 Å, which is more than double the average TMC-atom RMSD (Supporting Information Figures S11 and S15). It is worth noting that the reference cage structure we chose for computing the RMSD is the X-ray crystal structure[35] of the Ga_4L_6 nanocage with an encapsulated bis(trimethylphosphine) gold cation, as there is no empty Ga_4L_6 nanocage structure available from experiment. Nevertheless, the large RMSD observed in cage atoms can largely be attributed to the rotation of the cage linkers to accommodate encapsulated guests. To characterize this behavior, we examined the atom-wise contributions to the cage RMSD (e.g., on individual linkers) for cages with the largest and smallest deviations. The contrast between the low RMSD values in some linkers (< 1.1 Å) and the high RMSDs in others (> 1.9 Å) demonstrates that some linkers are relatively fixed, while the other linkers rotate to encapsulate the guest (Figure 6 and Supporting Information Table S10). Consistent with linker rotation, higher atom-wise RMSDs are generally observed on the edge of the linker naphthalene rings, which undergo greater displacement during rotation. While this trend is generally observed across our dataset of square planar TMCs, further investigation is required to determine if other TMC geometries (e.g., tetrahedral structures) would similarly perturb the cage structure.

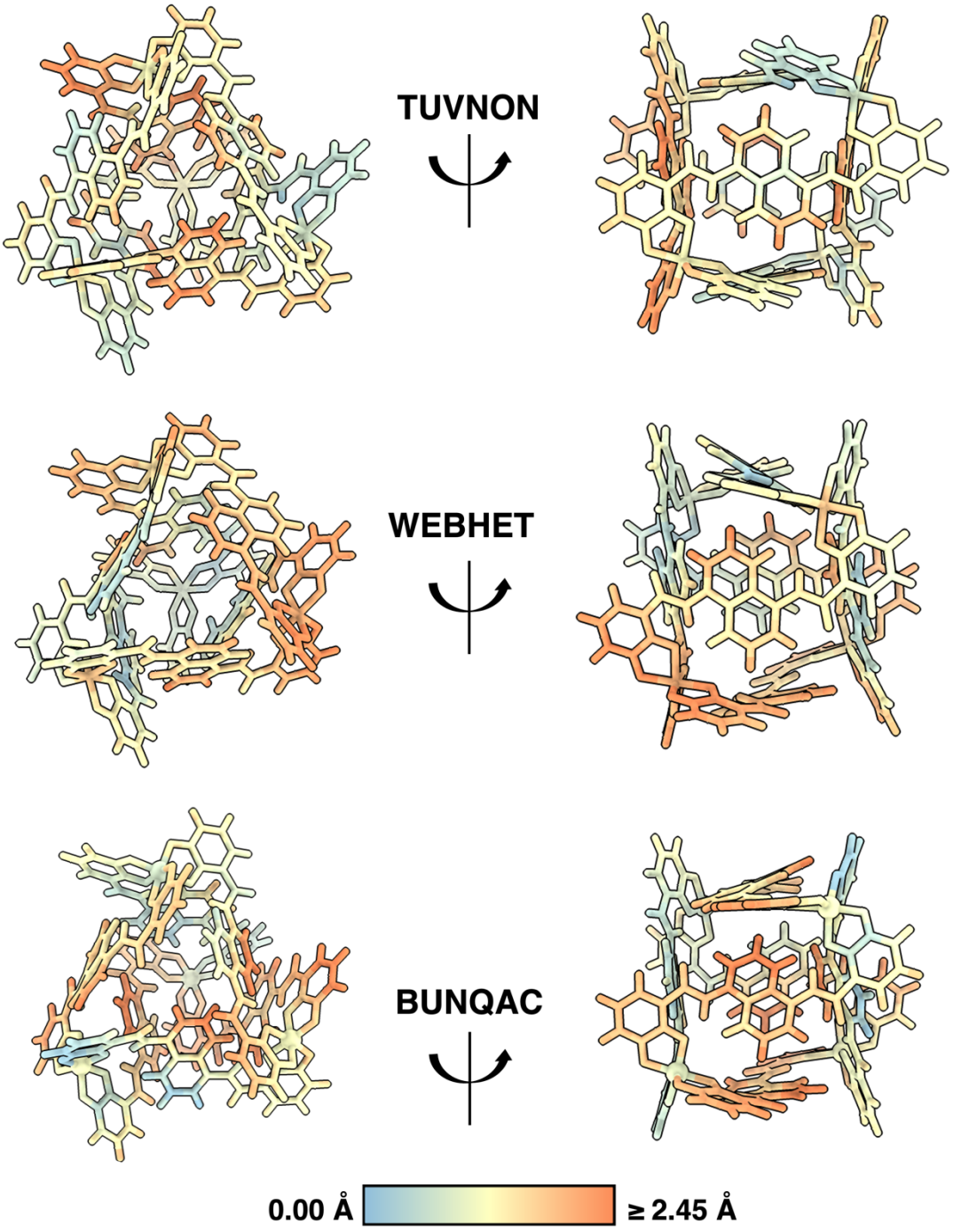


Figure 6. Root-mean-square deviations (RMSDs) per cage atom from the optimized encapsulated TMC geometry. RMSDs were computed with respect to the cage geometry in the X-ray crystal structure[35] of the encapsulated bis(trimethylphosphine) gold cation. The same color map is used for both species with the smaller range of values spanned by TUVNON used as bounds, thus all values greater than or equal to 2.45 Å are given in the same color for WEBHET and BUNQAC.

As linkers rotate to accommodate a guest, the cage volume necessarily changes. The final cage volumes range from 356 Å^3 to 535 Å^3, with an average of 435 Å^3 that is 5.6% smaller than the volume of the reference crystal structure (460 Å^3) computed with the same approach. Because the TMC encapsulated in the reference crystal structure was also obtained *a priori* by our screening procedure, it is possible to compare the cavity size of the optimized cage obtained by DFT to the experimental structure. This comparison yields a cavity size of 405 Å^3 for the DFT optimized structure vs. 460 Å^3 from the crystal structure. While cavity analysis can be sensitive to aperture volume and pore choice, this difference suggests that cavities may be smaller in DFT-optimized structures. The reason for this decrease in cavity size could be due to lack of explicit treatment of solvent or counterions both inside and outside the cage, but these effects are expected to be systematic across the entire dataset and thus not impact the relative comparisons we make.

As could be expected, we observe a positive correlation between the volume of the occupied cage and the volume of the guest, which highlights the flexibility of the cage to expand with increasing guest size (Supporting Information Figure S16). Over our range of optimized guest sizes (211–310 Å^3), the encapsulated guests have a wide range of packing fractions from 0.55 to 0.72 (average: 0.60, Supporting Information Figure S17). While factors such as shape, charge, and ligand identity influence the upper bound of the realizable packing fraction, we note that the calculated ratios are consistent with those previously quantified for guests studied in this system.[25,36] Notably, the packing fraction exhibits no correlation with the average heavy-atom RMSD of the cage, indicating that the particular rotation of the cage linkers we observe during each DFT geometry optimization can change the shape of the cavity but does not correlate with the volume guests can occupy (Supporting Information Figure S18). The average packing ratio we observed across this expanded set of diverse TMC guests corroborates Rebek's 55% rule[97] but also

highlights opportunities for larger packing fractions when the chemistry of the TMC is taken into account. The largest packing fraction in our dataset was 0.72, which corresponds to NICGEO (5,5'-difluoro-2,2'-bipyridine)-(4-fluorophenyl)-iodo-gold), the only complex in our dataset with fluorinated ligands and iodide coordinating to the gold center. The hydrophobic nature[98] of fluorine and the larger ionic radii of iodide (i.e., relative to chloride) could explain this large packing fraction value.

We next computed the binding electronic energies of the structures in TMC-34 to interrogate the geometric and electronic properties associated with favorable binding energetics. We computed the binding energies from electronic energies of the isolated and complexed cage-TMC pair using two dielectric values in an implicit solvent model (see Sec. 2). These binding energies ranged from –92.5 to –33.2 kcal/mol in water, with an average of –49.9 kcal/mol (Supporting Information Figures S19). Binding energies computed in methanol span a larger range from –131 kcal/mol to –44 kcal/mol and are more favorable than those observed in water (Supporting Information Figures S19). The combination of ligand and metal identity has nuanced effects, with distinct identities giving rise to similar packing fractions and binding energies (Figure 7). Electrostatics play an important role in binding, as illustrated by highly positively charged molecules binding most favorably with the anionic cage (e.g., refcode BUNQAC with a net charge of +4, –92.5 kcal/mol, Figure 7). This effect is more pronounced in implicit methanol solvent, in which a lower dielectric constant leads to less electrostatic screening and thus stronger electrostatic interactions that promote TMC encapsulation (Supporting Information Figures S19). The role of solvent screening is also clear when comparing a TMC in our dataset (refcode OVUHUK) to previous computational efforts[24] where similar TMCs were studied. In this prior work, a TMC possessing one fewer trimethylphosphine ligand than OVUHUK was studied as an intermediate in

the same nanocage using a hybrid GGA functional and a larger basis set with neutralizing cations and water molecules present. The authors obtained a binding energy of –121.9 kcal/mol,[24] which is much larger than our computed value of -38.7 kcal/mol, likely due to differences in solvent treatment and inclusion of counterions. While the results from these two studies are not directly comparable due to different chemical models and levels of theory, these results show that the electronic contribution to binding is very large for these charged systems, as expected and previously noted.[27]

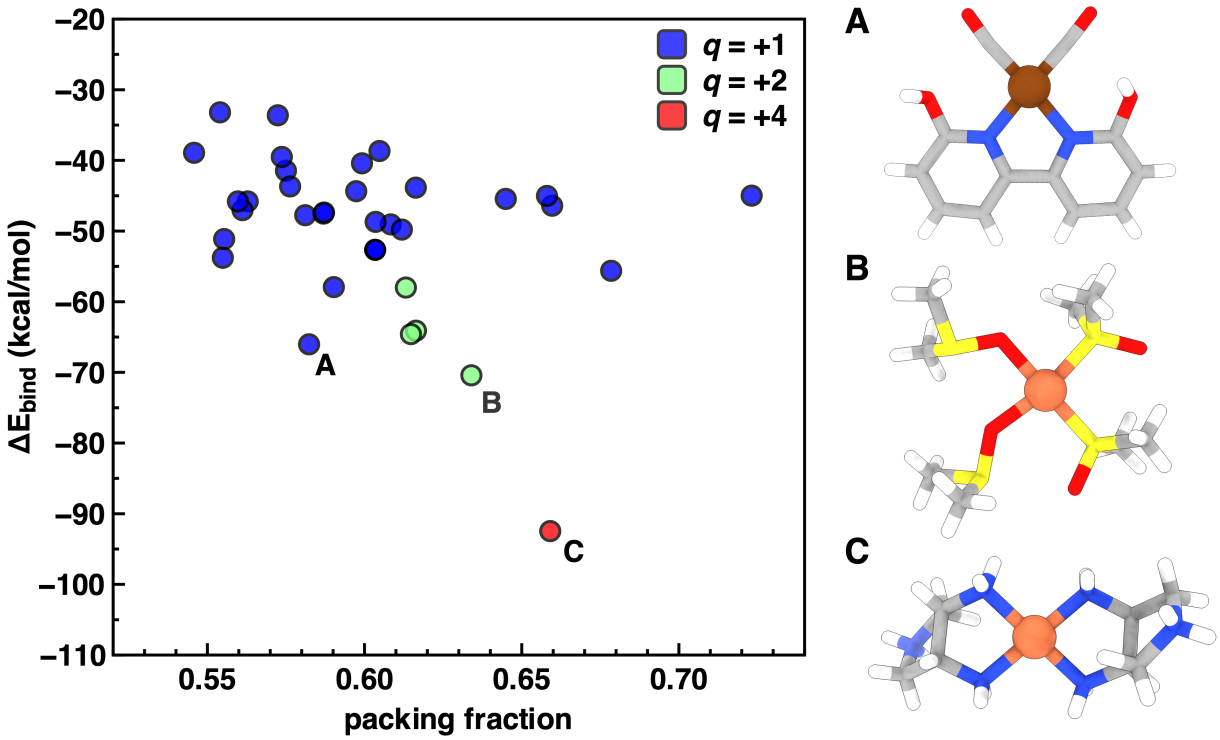


Figure 7. Binding energies as a function of packing fraction for the 34 encapsulated TMC guests. TMCs are colored by their total molecular charge. Packing fraction is a ratio of volumes and is thus unitless. The binding energies as computed in C-PCM solvent of water (ε = 78.39), A) Structure of IMIFEQ, B) structure of MEZHUW, and C) structure of BUNQAC. For structures, gray coloring is used for carbon atoms, white for hydrogen atoms, blue for nitrogen atoms, and red for oxygen atoms. Dark brown and burnt orange are used for the rhodium and palladium metal centers, respectively.

One might expect that the packing fraction of the TMC could be related to binding energy, where the complexes most likely to violate Rebek's 55% rule[97] are those for which electrostatic stabilization with the cage is most favorable. Nevertheless, we observe only a weak correlation between packing fraction and electronic binding energy, with slightly higher binding energies for the TMCs with larger packing fractions (Figure 7). To examine if the individual components of the packing fraction ratio have stronger correlations with the binding energy, we examined the van der Waals volume of the encapsulated TMC and the cavity volume of the optimized cage (Supporting Information Figure S20 and S21). We observed no significant trends between these geometric measures and binding energy, which indicates that fit of the TMC in the cage is not correlated with the electronic contribution to binding. This indicates that a computationally demanding binding energy calculation is also not required, and it should instead be more straightforward to apply a relaxed (i.e., 0.55–0.72) interpretation of Rebek's rule[97] for identifying potential TMC guests based on packing fraction.

3d. Electronic Changes to TMC-34 Upon Encapsulation

Motivated by studies that have attributed encapsulation-mediated rate enhancements to the electrostatic environment provided by the cage,[27,99,100] we aimed to identify if encapsulation in the Ga_4L_6 cage alters the electrostatic characteristics of TMCs. Specifically, we compared the electrostatic potential and partial charges at each TMC metal center in solvated and cage environments (see Sec. 2). Overall, we observe that the partial charge at TMC metal centers is weakly affected by encapsulation, with nearly all TMCs retaining unchanged partial charges (Figure 8). This suggests that much greater differences in partial charges are determined by TMC chemistry, such as the metal identity (e.g., Au charges range from +0.94 to +1.48 versus Rh partial

charges range from -0.27 to +0.86 in the encapsulated state), than the effect of encapsulation (Figure 6). In line with observations on changes in the Au–Cl bond upon encapsulation, we observe some Au complexes to exhibit a slight decrease in positive charge upon encapsulation (Figure 8). Conversely, for several of the rhodium-containing TMCs, we do observe a slight shift in partial charge to more positive values (Figure 8). While the observed partial charges are a combined effect of the metal and ligand chemistry, the smaller partial charge on the rhodium metal center is attributed to its electron-rich, nucleophilic nature. On the contrary, gold is more electronegative and thus carries a more positive partial charge. To examine the chemical diversity that could also be contributing, we inspected representative TMCs spanning the range of metal partial charges: the maximum (refcode: IPISAD), the minimum (refcode: SEBBIL), and a structure in the midpoint (refcode: MODKEY), finding that these complexes preserved our observation of differences between Au and Rh, with the Ni complex (refcode: MODKEY) residing intermediate between the two. Interpreting these trends further is complicated by the diversity of coordinating ligands, which can have differing effects on reactivity across metal series.

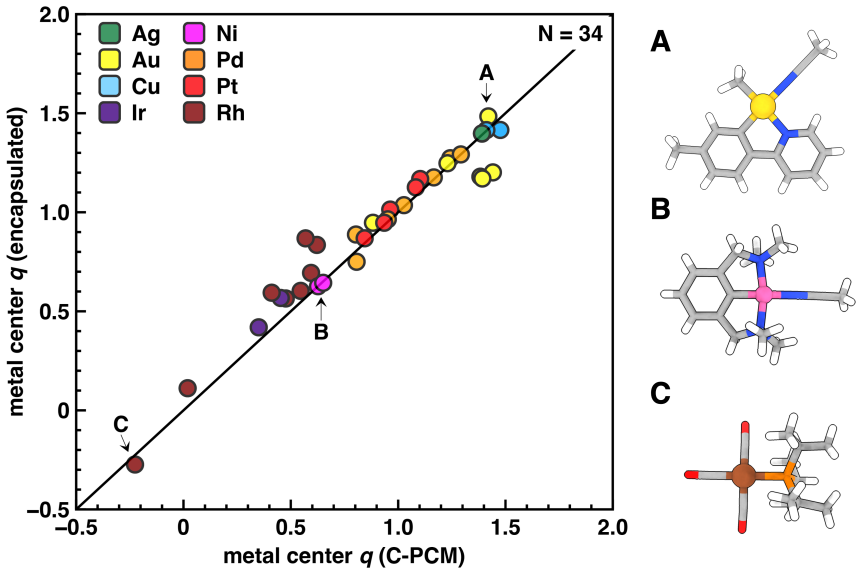


Figure 8. Iterative Hirshfeld partial charge, *q*, of the metal center while in implicit solvent versus when encapsulated for each TMC successfully optimized in the cage. A) Structure of IPISAD, B) structure of MODKEY, and C) structure of SEBBIL. For structures, gray coloring is used for

carbon atoms, white for hydrogen atoms, blue for nitrogen atoms, orange for phosphorous atoms, and red for oxygen atoms. The metal centers are colored consistently with the plot legend.

We next considered the relationship between the net TMC charge and partial charge at the metal center. Although the number of highly charged complexes is limited (i.e., four complexes with charges above +1) and complexes vary in ligand and metal identity, encapsulation leads to a negligible charge shift regardless of overall TMC molecular charge except in some cases of Rh and Au previously noted (Figure 8 and Supporting Information Figure S22). These conclusions are largely insensitive to the partial charge scheme (Supporting Information Figure S23). To understand the weak impact of the cage environment on the metal center partial charge, we analyzed the structure of each encapsulated TMC. Generally, we observe that TMCs closely approach cage linkers, with an average distance of 3.8 Å between the metal center and the nearest cage linker atom, while maintaining an average of 7.8 Å from the gallium cage vertices. The metal center of the TMC with refcode HUVNIW most closely approaches a gallium vertex at 5.5 Å, albeit still at a greater distance than the closest linker atoms (Supporting Information Figure S24). This distance is too far for the positively charged gallium centers to strongly influence the charge distribution on the metal, and non-polar atoms in the cage linkers would be expected to have an even more limited effect.

In contrast to metal partial charge, we observe that encapsulation strongly alters the magnitude of the electrostatic potential (ESP) at the TMC metal centers. ESPs uniformly decrease upon encapsulation in the strongly anionic cage, with the change in ESP upon encapsulation, ΔESP, ranging from -417.9 to -918.9 kcal/mol•e (Figure 9). Differences in net charge of the TMC have limited impact on the ESP shift. The ESP at the Rh center of a TMC with a net charge of +4 (refcode: BUNQAC) exhibits a similar ΔESP to the Rh center in the +1 net charge complexes with

refcodes AFAZUE or HUVNIW. Moreover, the relative ESP of these TMCs (i.e., AFAZUE < HUVNIW < BUNQAC) is preserved upon encapsulation, emphasizing the relatively uniform impact of encapsulation on this diverse set of TMCs. These trends are largely maintained when the implicit solvent dielectric is changed from methanol to water, although small differences arise when the change in solvent impacts the optimized geometry (Supporting Information Figures S25–S26).

To probe small variations in the nearly uniform ΔESP across TMC-34, we investigated the proximity of the TMC to the gallium ions at the cage vertices and to the cage linkers. Quantifying the relationship between these geometry proximity measures and the ESP shift, we observe only a weak correlation (Figures 9 and Supporting Information Figures S27–S28). Although the TMCs that are most distant from the cage have smaller ESP shifts than those that are closest to the cage, there is no monotonic trend between ΔESP and the distance from the TMC metal to the cage (i.e., either Ga ions or cage linkers). Nevertheless, the range of distances between TMCs and cage linkers or Ga vertices is modest, leaving the possibility that incorporating values from smaller TMCs or larger cages could lead to a stronger trend. Overall, this indicates that a rapid, purely volumetric screening of TMCs inside the cage (e.g., with a force field) should be sufficient because the ΔESP can be rapidly ascertained on approximate structures or inferred from other TMCs.

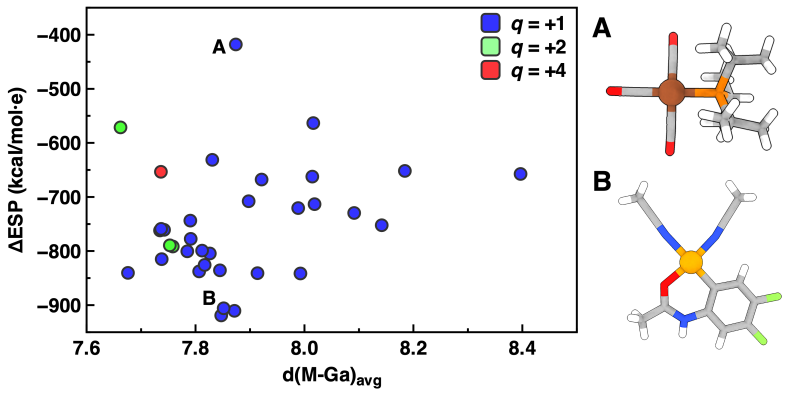


Figure 9. Encapsulation-mediated shift in electrostatic potential at the TMC metal center plotted against the average distance between the TMC metal and the gallium cage vertices. Data are colored according to the overall charge of the TMC. A) Structure of SEBBIL and B) structure of KUDJEZ. For structures, gray coloring is used for carbon atoms, white for hydrogen atoms, blue for nitrogen atoms, dark orange for phosphorous atoms, light green for fluorine atoms, and red for oxygen atoms. Brown is used to denote the rhodium metal center in SEBBIL, while light orange is used to denote the palladium metal center in KUDJEZ.

Next, we evaluated the cumulative ESP, which we obtained by starting at the metal center and incrementally incorporating atoms at increasing radial distances into the calculation of the ESP. We carried out this analysis for three representative TMCs, refcodes IPISAD, MODKEY, and SEBBIL, that span the range of metal partial charges. For each of these TMCs, we observe an initial local maximum associated with ESP contributions from the primary and secondary coordination shell atoms (Figure 10). The cumulative ESP then decreases upon inclusion of the cage atoms, with oscillations in ESP due to local TMC and cage geometry as illustrated by SEBBIL (Figure 11). These oscillations are in contrast with the radially decaying ESP one would expect if the electron density were uniformly distributed inside a sphere spanning the distance between the Rh center of SEBBIL and the closest Ga atom (Supporting Information Figure S29). The location of the nearest cage atom varies slightly among the three complexes, with distances of 3.28 Å (IPISAD), 4.15 Å (MODKEY), and 3.65 Å (SEBBIL). At larger distances (~4.5 to ~8 Å), both cage and TMC atoms contribute to the ESP, making it difficult in some cases to deconvolute the two contributions. For the relatively compact MODKEY (i.e., the largest radial distance spanned by the TMC complex is 5.7 Å with respect to Ni), the sudden decrease in the electrostatic potential observed between 6.0 and 6.5 Å can be entirely attributed to the nearest cage linker. Importantly, we observe that the full integrated limit of the ESP is consistent with the ESP obtained only at the metal center (Figure 8). The relative values of the ESP are also correlated with the charge on the metal atom. Of the three complexes, IPISAD contains both the most negative metal center and

most negative final ESP amongst the dataset, SEBBIL contains the least negative metal center and least negative ESP, and MODKEY lies between the two. Collectively, we find that cage encapsulation shifts the electrostatic potential at TMC centers in a uniformly negative fashion. These ESP shifts are remarkably similar in magnitude across the diverse dataset and show only weak dependence on TMC location in the cage.

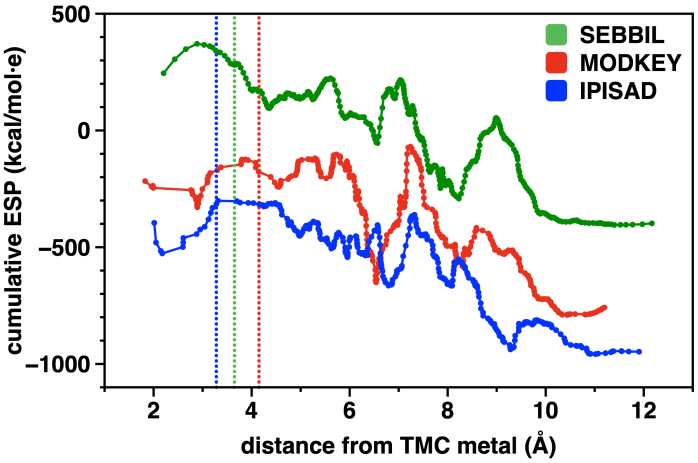


Figure 10. Cumulative electrostatic potential for encapsulated TMCs accounting for contributions from atoms up to different radial cutoffs from TMC center. Cage atoms contribute to the ESP starting at radial distances of 3.28 Å (refcode: IPISAD), 4.15 Å (refcode: MODKEY), and 3.65 Å (refcode: SEBBIL), which are indicated with dashed lines.

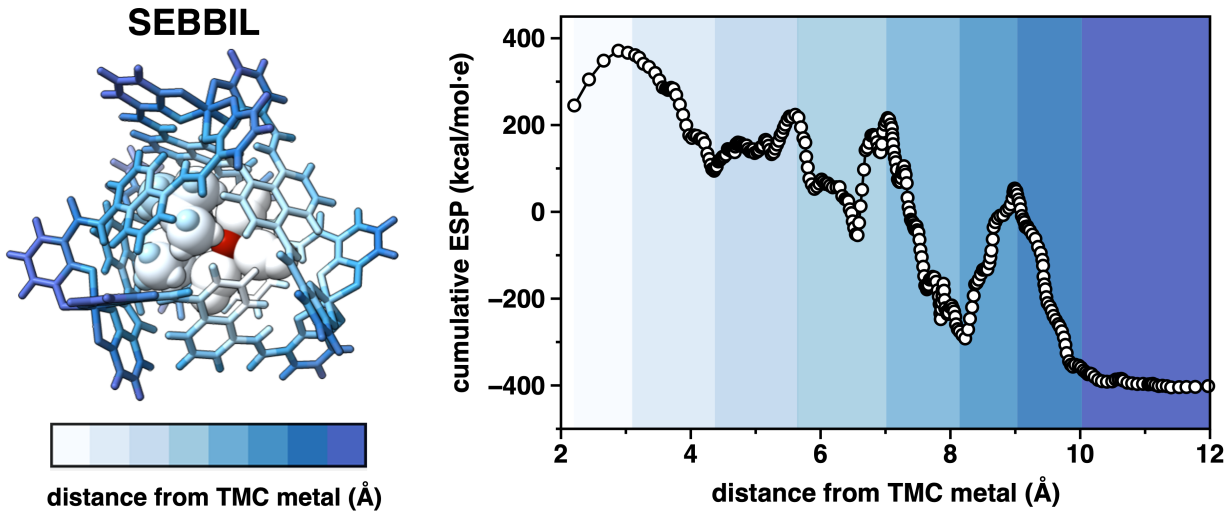


Figure 11. Cumulative electrostatic potential for the encapsulated TMC refcode: SEBBIL colored by radial distance from the TMC metal. Local minima and maxima are used to delineate boundaries between color blocks. The values used to delineate the minima and maxima are 2.2 Å,

3.1 Å, 4.4 Å, 5.6 Å, 7.0 Å, 8.1 Å, 9.0 Å, 10.0 Å, and 12.2 Å. All atoms are colored by their radial distance except the Rh metal center, which is shown in red.

3e. Identifying Guests for Catalytic Studies

We now examine the size of candidate complexes for catalysis (i.e., both the TMC and reacting molecules) to identify the species with the greatest potential for rate acceleration via encapsulation. We apply two main criteria: whether the catalyst is small enough that it can be encapsulated, and whether associated reacting molecules are small enough to be encapsulated as well. We also take into consideration the solvent used experimentally and note that for follow up study. From TMC-34, we select five TMCs with reactant volumes of less than 150 $Å^3$, such that the total volume of a reactant and TMC would be < 450 $Å^3$ (i.e., all TMCs are < 300 $Å^3$ in size, see Sec. 2, Table 1). This choice is also experimentally motivated by the fact that the largest Ga_4L_6 cage volume measured in the literature for this system[36] is 434 $Å^3$ and our average cage cavity volume across the 34 encapsulated guests was 435 $Å^3$. Assuming the cavity can expand to the maximum value of 535 $Å^3$ seen in our study, this would result in a packing fraction of 0.84 for the largest combined sizes of TMCs and reactants. While this exceeds the packing fraction we observed to be favorable for our isolated TMCs, we err on the side of including more complexes as it is difficult to predict if reactants need to be completely encapsulated or how the cage volume will change for different guests.

Table 1. Experimental information from literature and computed volumes for five promising TMCs investigated in this study based on experimental properties and reactant size including CSD refcode, chemical name, details of the solvent and relevant reaction, and the volume of the reactant. Chemical names are taken directly from the CSD, where the solubility and reaction type were gleaned from the associated reference.

CSD refcode	CSD Chemical Name	Solvent[a]	Type of Reaction	Reactant, TMC, vdW volume[b]	Combined volume of TMC + reactant[c]	Ref.
AXEDEN	Dicarbonyl-(N-((1-methyl-1H-imidazol-2-yl)methylene)aniline)-rhodium	MeCN; AcOH/ H_2O mixture	carbonylation	Methyl acetate, 74.8 Å^3	316.2 Å^3	[101]
BEBXIQ	Dicarbonyl-(bis(1-pyrazolyl)methane-N,N')-iridium	THF	triethylsilane alcoholysis	Triethylsilane, 147.2 Å^3	352.4 Å^3	[102]
IMIFEQ	Dicarbonyl-(6,6'-dihydroxy-2,2'-bipyridine)-rhodium	AcOH, H_2O	carbonylation	Methyl acetate, 74.8 Å^3	308.4 Å^3	[103]
INIJUK	Dicarbonyl-(bis(3,5-dimethylpyrazolyl)methane-N,N')-rhodium	THF	alkene hydroformylation and hydroamination	1-oct-ene, 89.2 Å^3	361.2 Å^3	[104]
MEZHUW	bis((methylsulfinyl)methane-S)-bis((methylsulfinyl)methane-O)-palladium	DMSO; DCE	C–C bond forming	Cyclohex-2-enone, 101.0 Å^3	394.9 Å^3	[105]

[a]Full chemical names of solvents and utilized abbreviations are acetonitrile (MeCN), acetic acid (AcOH), tetrahydrofuran (THF), Dimethyl sulfoxide (DMSO), 1,2-Dichloroethane (DCE)
[b]Volumes were computed from 3D structures obtained from PubChem entry for the listed reactant. Reactant size was computed by MoloVol v1.0.0 using the standard radii set for the elements, with a grid size of 0.1 Å.
[c]Combined volumes are computed from the CSD structure of the TMC of interest plus the reactant listed in the column immediately to the left.

Applying this criterion for cage compatibility, we see a particular enrichment of rhodium-containing compounds and compounds containing carbonyl ligands, which can be expected due to their widespread catalytic use (Table 1). These complexes: dicarbonyl-(N-((1-methyl-1H-imidazol-2-yl)methylene)aniline)-rhodium (refcode: AXEDEN), dicarbonyl-(bis(1-pyrazolyl)methane-N,N')-iridium (refcode: BEBXIQ), dicarbonyl-(6,6'-dihydroxy-2,2'-bipyridine)-rhodium (refcode: IMIFEQ), dicarbonyl-(bis(3,5-dimethylpyrazolyl)methane-N,N')-rhodium (refcode: INIJUK), and bis((methylsulfinyl)methane-S)-bis((methylsulfinyl)methane-O)-palladium (refcode: MEZHUW), are known to catalyze a wide range of reactions such as carbonylation, hydroamination, C–C bond formation, and polymerization. For one example (refcode: INIJUK), the catalytic efficiency of the alkene hydroformylation and hydroamination

reaction was noted to be poor due to migratory insertion not being as rapid as β-elimination reforming an alkene from the alkyl rhodium intermediate.[104] Upon encapsulation in the Ga_4L_6 cage, this TMC becomes 10% more electropositive with a shift in partial charge from +0.546 to +0.602 a.u. (Figure 6). More electropositive metal centers are expected to accelerate migratory insertion reactions, and the increased Lewis acidity inside the cage cavity could promote the required ligand binding for the migratory insertion step. Nevertheless, catalysis by this TMC has yet to be explored in a more polar solvent. Across our entire TMC-34 set, almost all our TMCs have reactants that may fit inside the cage and fall below a total volume of 450 $Å^3$, and many have been studied in a polar solvent like acetonitrile. A full analysis of all TMC-34 reaction conditions is provided in the Zenodo repository.[47] This analysis suggests that many of our TMC-34 complexes possess properties that would make them ideal starting points for further experimental investigation.

4. Conclusions

We conducted a screen of the CSD using a combination of physical properties and text mining to identify catalytic transition metal complexes (TMCs) as potential binders in the anionic $Ga_4L_6^{12-}$ nanocage. We demonstrated a workflow that that is general to other metal-organic cage systems. From thousands of possible square-planar complexes, we obtained the TMC-63 set of 63 unique TMCs encompassing eight different metal centers and a diverse set of coordination ligands.

We retained a subset of 34 TMCs that were successfully optimized inside the cage (TMC-34), and analyzed properties of these TMCs. We characterized the encapsulation-mediated electronic and geometric changes in each TMC and found that they were relatively uniform and modest across most complexes. Nevertheless, we observed that for the only TMCs in TMC-34 with Au–Cl bonds, substantial elongation of the Au–Cl bond occurred in encapsulation. Across

all complexes, the cage demonstrated higher flexibility, with linker rotations corresponding to the largest deviation from a reference cage structure. Individual linkers exhibited RMSDs as high as ~2.0 Å, often in combination with less mobile linkers with RMSDs of less than <1.0 Å, showing cooperative motion. The TMCs that were successfully encapsulated exhibited a range of packing fractions with a mean of 0.60, which bears similarity to prior heuristics (i.e., Rebek's 55% rule). However, notable exhibitions included complexes (e.g., refcode: NICGEO) with a packing fraction substantially greater than the 55% rule of thumb (i.e., up to 0.72). Taken together, we have shown that encapsulation impacts small cationic TMCs in a similar manner, while the cage itself has the larger changes with shifts in linker geometry.

Across TMC-34, the electrostatic potential (ESP) at the metal center consistently decreased upon encapsulation while the partial charge remained nearly constant, with some TMC metal centers (e.g. Rh) showing a slight increase and some gold complexes showing a slight decrease. Examining the atom-wise contributions to the ESP, we observed that the TMC ligands strongly screen the metal center from the anionic environment in the cage, minimizing the negative shift in electrostatic potential and maintaining the TMC partial charge. TMCs with larger ESPs before encapsulation maintain the largest ESPs after binding in the cage, illustrating that the cage has a uniform effect across TMC-34. This observation highlighted how purely geometric screening of TMC size could be used to rapidly assess cage compatibility for charged TMCs.

Finally, we identified the TMC-34 structures that would be most amenable to follow-up study. We identified five TMCs associated with reactions with small reactants that have been previously studied in polar solvents that could be simultaneously encapsulated alongside the TMC. Complexes could then also be pursued experimentally by using our approach to select for complexes predisposed to rate acceleration upon encapsulation. Future work could also

computationally study elementary reaction steps inside the cage to allow further fine-tuning of our approach.

AUTHOR INFORMATION

Corresponding Author

*email:hjkulik@mit.edu

Notes

The authors declare no competing financial interest.

Supporting Information.

Computed volumes of cationic guests; keywords used for text-mining; number of complexes and reasons text-mined manuscripts were excluded after manual paper review; RMSDs of optimized structures with different levels of theory; comparison of optimized structures with different levels of theory; comparison of calculated and experimental binding energies; structures of cationic guests used for benchmarking; CREST optimization criteria; comparison of CREST and crystal structure geometries; Singlet-triplet splitting electronic energies; singlet-triplet splitting energies in water and methanol; additional details for the calculation of electrostatic potential; effects of large probe size of computed cavity volume and visualization of cage cavities; van der Waals volumes for cationic square complexes; CSD structure codes of removed duplicate complexes; metal center identities present in dataset across filtering steps; metal center ligand denticities and metal center coordinating atoms in dataset; heavy atom RMSDs for transition metal complexes, highest RMSD structures compared to crystal structures; smallest host-guest interatomic distances upon docking; Au(III)-Cl distances in selected CSD entries before and after optimization; metal center partial charge with various charge schemes; average and smallest metal to cage atom distance; shifts in electrostatic potential upon encapsulation in methanol and water implicit solvent DFT optimizations; shifts in electrostatic potential as a function of distance to gallium vertices or cage atoms; cage heavy atom RMSD, van der Waals volume of TMC guests and nanocage cavities; computed packing fractions; heavy atom RMSD vs. packing fraction; binding energies in different implicit solvents; binding energies as a function of TMC and cavity volume. (PDF)

This material is available free of charge via the Internet at http://pubs.acs.org.

Acknowledgments

This work was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing, Office of Basic Energy Sciences, via the Scientific Discovery through Advanced Computing (SciDAC) program (to C.R.R.; M.T.M.; W.L.; T.H.G.; and H.J.K.) C.R.R. was partially supported by an Arnold O. Beckman Postdoctoral Fellowship. H.J.K. holds an Alfred P. Sloan Fellowship in Chemistry and a Simon Family Faculty Research Innovation Fund grant which supported this research. The authors thank Adam H. Steeves for providing a critical reading of the manuscript and Ilia Kevlishvili for scientific discussion.

References

(1) Pedersen, C. Cyclic Polyethers and Their Complexes with Metal Salts. *J. Am. Chem. Soc.* **1967,** *89*, 7071.
(2) Cram, D. J.; Cram, J. M. Host-Guest Chemistry. *Science* **1974,** *183*, 803-809.
(3) Lehn, J. M. Cryptates: Inclusion Complexes of Macropolycyclic Receptor Molecules. *Pure Appl. Chem.* **1978,** *50*, 871-892.
(4) Brown, C. J.; Toste, F. D.; Bergman, R. G.; Raymond, K. N. Supramolecular Catalysis in Metal–Ligand Cluster Hosts. *Chem. Rev.* **2015,** *115*, 3012-3035.
(5) Yoshizawa, M.; Klosterman, J. K.; Fujita, M. Functional Molecular Flasks: New Properties and Reactions within Discrete, Self-Assembled Hosts. *Angew. Chem. Int. Ed.* **2009,** *48*, 3418-3438.
(6) Zhao, L.; Jing, X.; Li, X.; Guo, X.; Zeng, L.; He, C.; Duan, C. Catalytic Properties of Chemical Transformation within the Confined Pockets of Werner-Type Capsules. *Coord. Chem. Rev.* **2019,** *378*, 151-187.
(7) Percástegui, E. G.; Ronson, T. K.; Nitschke, J. R. Design and Applications of Water-Soluble Coordination Cages. *Chem. Rev.* **2020,** *120*, 13480-13544.
(8) Li, W.-L.; Hao, H.; Head-Gordon, T. Optimizing the Solvent Reorganization Free Energy by Metal Substitution for Nanocage Catalysis. *ACS Catalysis* **2022,** *12*, 3782-3788.
(9) Fujita, M.; Oguro, D.; Miyazawa, M.; Oka, H.; Yamaguchi, K.; Ogura, K. Self-Assembly of Ten Molecules into Nanometre-Sized Organic Host Frameworks. *Nature* **1995,** *378*, 469-471.
(10) Fox, O. D.; Dalley, N. K.; Harrison, R. G. A Metal-Assembled, Ph-Dependent, Resorcinarene-Based Cage Molecule. *J. Am. Chem. Soc.* **1998,** *120*, 7111-7112.
(11) Martí-Centelles, V.; Lawrence, A. L.; Lusby, P. J. High Activity and Efficient Turnover by a Simple, Self-Assembled "Artificial Diels–Alderase". *J. Am. Chem. Soc.* **2018,** *140*, 2862-2868.

(12) Caulder, D. L.; Powers, R. E.; Parac, T. N.; Raymond, K. N. The Self-Assembly of a Predesigned Tetrahedral M4l6 Supramolecular Cluster. *Angew. Chem. Int. Ed.* **1998,** *37*, 1840-1843.
(13) Moon, D.; Kang, S.; Park, J.; Lee, K.; John, R. P.; Won, H.; Seong, G. H.; Kim, Y. S.; Kim, G. H.; Rhee, H.; Lah, M. S. Face-Driven Corner-Linked Octahedral Nanocages: M6l8 Cages Formed by C3-Symmetric Triangular Facial Ligands Linked Via C4-Symmetric Square Tetratopic Pdii Ions at Truncated Octahedron Corners. *J. Am. Chem. Soc.* **2006,** *128*, 3530-3531.
(14) Tidmarsh, I. S.; Faust, T. B.; Adams, H.; Harding, L. P.; Russo, L.; Clegg, W.; Ward, M. D. Octanuclear Cubic Coordination Cages. *J. Am. Chem. Soc.* **2008,** *130*, 15167-15175.
(15) Pugh, C. J.; Santolini, V.; Greenaway, R. L.; Little, M. A.; Briggs, M. E.; Jelfs, K. E.; Cooper, A. I. Cage Doubling: Solvent-Mediated Re-Equilibration of a [3 + 6] Prismatic Organic Cage to a Large [6 + 12] Truncated Tetrahedron. *Cryst. Growth Des.* **2018,** *18*, 2759-2764.
(16) Santolini, V.; Miklitz, M.; Berardo, E.; Jelfs, K. E. Topological Landscapes of Porous Organic Cages. *Nanoscale* **2017,** *9*, 5280-5298.
(17) Mao, X.-R.; Wang, Q.; Zhuo, S.-P.; Xu, L.-P. Reactivity and Selectivity of the Diels–Alder Reaction of Anthracene in [Pd6l4]12+ Supramolecular Cages: A Computational Study. *Inorg. Chem.* **2023,** *62*, 4330-4340.
(18) 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*, 1300-1310.
(19) Nurttila, S. S.; Brenner, W.; Mosquera, J.; van Vliet, K. M.; Nitschke, J. R.; Reek, J. N. H. Size-Selective Hydroformylation by a Rhodium Catalyst Confined in a Supramolecular Cage. *Chem. Eur. J.* **2019,** *25*, 609-620.
(20) Sebastiani, F.; Bender, T. A.; Pezzotti, S.; Li, W.-L.; Schwaab, G.; Bergman, R. G.; Raymond, K. N.; Toste, F. D.; Head-Gordon, T.; Havenith, M. An Isolated Water Droplet in the Aqueous Solution of a Supramolecular Tetrahedral Cage. *Proc. Natl. Acad. Sci. U.S.A.* **2020,** *117*, 32954-32961.
(21) Tarzia, A.; Jelfs, K. E. Unlocking the Computational Design of Metal-Organic Cages. *Chem Commun (Camb)* **2022,** *58*, 3717-3730.
(22) Zhou, J.; Mroz, A.; Jelfs, K. E. Deep Generative Design of Porous Organic Cages Via a Variational Autoencoder. *Digital Discovery* **2023,** *2*, 1925-1936.
(23) Welborn, V. V.; Li, W.-L.; Head-Gordon, T. Interplay of Water and a Supramolecular Capsule for Catalysis of Reductive Elimination Reaction from Gold. *Nat. Commun.* **2020,** *11*, 415.
(24) Vaissier Welborn, V.; Head-Gordon, T. Electrostatics Generated by a Supramolecular Capsule Stabilizes the Transition State for Carbon–Carbon Reductive Elimination from Gold(Iii) Complex. *J. Phys. Chem. Lett.* **2018,** *9*, 3814-3818.
(25) 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*, 4370-4381.
(26) 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*, 15022-15026.

(27) Frushicheva, M. P.; Mukherjee, S.; Warshel, A. Electrostatic Origin of the Catalytic Effect of a Supramolecular Host Catalyst. *J. Phys. Chem. B* **2012,** *116*, 13353-13360.
(28) Sciortino, G.; Norjmaa, G.; Maréchal, J. D.; Ujaque, G. In *Supramolecular Catalysis*, 2022,271-285
(29) Cullen, W.; Turega, S.; Hunter, C. A.; Ward, M. D. Virtual Screening for High Affinity Guests for Synthetic Supramolecular Receptors. *Chem. Sci.* **2015,** *6*, 2790-2794.
(30) Taylor, C. G.; Cullen, W.; Collier, O. M.; Ward, M. D. A Quantitative Study of the Effects of Guest Flexibility on Binding inside a Coordination Cage Host. *Chemistry* **2017,** *23*, 206-213.
(31) Parrilla-Gutiérrez, J. M.; Granda, J. M.; Ayme, J.-F.; Bajczyk, M. D.; Wilbraham, L.; Cronin, L. Electron Density-Based Gpt for Optimization and Suggestion of Host–Guest Binders. *Nat. Comput. Sci.* **2024,** *4*, 200-209.
(32) Cherney, A. H.; Kadunce, N. T.; Reisman, S. E. Enantioselective and Enantiospecific Transition-Metal-Catalyzed Cross-Coupling Reactions of Organometallic Reagents to Construct C–C Bonds. *Chem. Rev.* **2015,** *115*, 9587-9652.
(33) Hong, C. M.; Bergman, R. G.; Raymond, K. N.; Toste, F. D. Self-Assembled Tetrahedral Hosts as Supramolecular Catalysts. *Acc. Chem. Res.* **2018,** *51*, 2447-2455.
(34) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Cryst. B* **2016,** *72*, 171-179.
(35) Levin, M. D.; Kaphan, D. M.; Hong, C. M.; Bergman, R. G.; Raymond, K. N.; Toste, F. D. Scope and Mechanism of Cooperativity at the Intersection of Organometallic and Supramolecular Catalysis. *J. Am. Chem. Soc.* **2016,** *138*, 9682-9693.
(36) 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*, 111-120.
(37) Nandy, A.; Taylor, M. G.; Kulik, H. J. Identifying Underexplored and Untapped Regions in the Chemical Space of Transition Metal Complexes. *J. Phys. Chem. Lett.* **2023,** *14*, 5798-5804.
(38) Kim, E.; Huang, K.; Jegelka, S.; Olivetti, E. Virtual Screening of Inorganic Materials Synthesis Parameters with Deep Learning. *npj Computational Materials* **2017,** *3*, 53.
(39) Taylor, M. G.; Yang, T.; Lin, S.; Nandy, A.; Janet, J. P.; Duan, C.; Kulik, H. J. Seeing Is Believing: Experimental Spin States from Machine Learning Model Structure Predictions. *J. Phys. Chem. A* **2020,** *124*, 3286-3299.
(40) Nandy, A.; Duan, C.; Kulik, H. J. Using Machine Learning and Data Mining to Leverage Community Knowledge for the Engineering of Stable Metal–Organic Frameworks. *J. Am. Chem. Soc.* **2021,** *143*, 17535-17547.
(41) Nandy, A.; Terrones, G.; Arunachalam, N.; Duan, C.; Kastner, D. W.; Kulik, H. J. Mofsimplify, Machine Learning Models with Extracted Stability Data of Three Thousand Metal–Organic Frameworks. *Scientific Data* **2022,** *9*, 74.
(42) Kaphan, D. M.; Levin, M. D.; Bergman, R. G.; Raymond, K. N.; Toste, F. D. A Supramolecular Microenvironment Strategy for Transition Metal Catalysis. *Science* **2015,** *350*, 1235-1238.
(43) Leung, D. H.; Bergman, R. G.; Raymond, K. N. Scope and Mechanism of the C−H Bond Activation Reactivity within a Supramolecular Host by an Iridium Guest: A Stepwise Ion Pair Guest Dissociation Mechanism. *J. Am. Chem. Soc.* **2006,** *128*, 9781-9797.

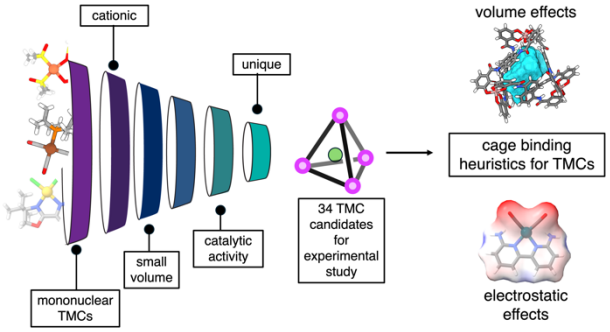
(44) Maglic, J. B.; Lavendomme, R. Molovol: An Easy-to-Use Program for Analyzing Cavities, Volumes and Surface Areas of Chemical Structures. *J. Appl. Crystalogr.* **2022,** *55*, 1033-1044.
(45) Rowland, R. S.; Taylor, R. Intermolecular Nonbonded Contact Distances in Organic Crystal Structures: Comparison with Distances Expected from Van Der Waals Radii. *J. Phys. Chem.* **1996,** *100*, 7384-7391.
(46) Adamji, H.; Nandy, A.; Kevlishvili, I.; Román-Leshkov, Y.; Kulik, H. J. Computational Discovery of Stable Metal–Organic Frameworks for Methane-to-Methanol Catalysis. *J. Am. Chem. Soc.* **2023,** *145*, 14365-14378.
(47) Reinhardt, C. R.; Manetsch, M. T.; Li, W.-L.; Head-Gordon, T.; Kulik, H. J. Computational Screening of Transition Metal Complexes as Guests in the Ga4l6-12 Nanocage. **2024**, DOI:10.5281/zenodo.11163686
(48) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 2. Direct Self-Consistent-Field Implementation. *J. Chem. Theory Comput.* **2009,** *5*, 1004-1015.
(49) Ufimtsev, I. S.; Martínez, T. J. Quantum Chemistry on Graphical Processing Units. 1. Strategies for Two-Electron Integral Evaluation. *J. Chem. Theory Comput.* **2008,** *4*, 222-231.
(50) Becke, A. Density-Functional Thermochemistry. Iii. The Role of Exact Exchange. *J. Chem. Phys* **1993,** *98*, 5648.
(51) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys Rev B Condens Matter* **1988,** *37*, 785-789.
(52) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994,** *98*, 11623-11627.
(53) 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*, 154104.
(54) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for the Transition Metal Atoms Sc to Hg. *J. Chem. Phys.* **1985,** *82*, 270-283.
(55) Wadt, W. R.; Hay, P. J. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for Main Group Elements Na to Bi. *J. Chem. Phys.* **1985,** *82*, 284-298.
(56) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. Ix. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972,** *54*, 724-728.
(57) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theoret. Chim. Acta* **1973,** *28*, 213-222.
(58) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor. Chim. Acta* **1977,** *44*, 129-138.
(59) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996,** *77*, 3865-3868.
(60) Chai, J.-D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008,** *128*, 084106.

(61) Pracht, P.; Bohle, F.; Grimme, S. Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. *Phys. Chem. Chem. Phys.* **2020,** *22*, 7169-7192.
(62) Grimme, S. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *J. Chem. Theory Comput.* **2019,** *15*, 2847-2862.
(63) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended Tight-Binding Quantum Chemistry Methods. *WIREs Computational Molecular Science* **2021,** *11*, e1493.
(64) Ehlert, S.; Stahn, M.; Spicher, S.; Grimme, S. Robust and Efficient Implicit Solvation Model for Fast Semiempirical Methods. *J. Chem. Theory Comput.* **2021,** *17*, 4250-4261.
(65) Hu, X.; Yang, W. Accelerating Self-Consistent Field Convergence with the Augmented Roothaan–Hall Energy Function. *J. Chem. Phys.* **2010,** *132*.
(66) Saunders, V. R.; Hillier, I. H. A "Level–Shifting" Method for Converging Closed Shell Hartree–Fock Wave Functions. *Int. J. Quantum Chem.* **1973,** *7*, 699-705.
(67) Guest, M. F.; Saunders, V. R. On Methods for Converging Open-Shell Hartree-Fock Wave-Functions. *Mol. Phys.* **1974,** *28*, 819-828.
(68) Wang, L.-P.; Song, C. Geometry Optimization Made Simple with Translation and Rotation Coordinates. *J. Chem. Phys.* **2016,** *144*.
(69) Liu, F.; Luehr, N.; Kulik, H. J.; Martínez, T. J. Quantum Chemistry for Solvated Molecules on Graphical Processing Units Using Polarizable Continuum Models. *J. Chem. Theory Comput.* **2015,** *11*, 3131-3144.
(70) York, D. M.; Karplus, M. A Smooth Solvation Potential Based on the Conductor-Like Screening Model. *J. Phys. Chem. A* **1999,** *103*, 11060-11079.
(71) Lange, A. W.; Herbert, J. M. A Smooth, Nonsingular, and Faithful Discretization Scheme for Polarizable Continuum Models: The Switching/Gaussian Approach. *J. Chem. Phys.* **2010,** *133*.
(72) Bondi, A. Van Der Waals Volumes and Radii. *J. Phys. Chem.* **1964,** *68*, 441-451.
(73) Paulsen, H.; Duelund, L.; Winkler, H.; Toftlund, H.; Trautwein, A. X. Free Energy of Spin-Crossover Complexes Calculated with Density Functional Methods. *Inorg. Chem.* **2001,** *40*, 2201-2203.
(74) Reiher, M.; Salomon, O.; Artur Hess, B. Reparameterization of Hybrid Functionals Based on Energy Differences of States of Different Multiplicity. *Theor. Chem. Acc.* **2001,** *107*, 48-55.
(75) Reiher, M. Theoretical Study of the Fe(Phen)2(Ncs)2 Spin-Crossover Complex with Reparametrized Density Functionals. *Inorg. Chem.* **2002,** *41*, 6928-6935.
(76) Bowman, D. N.; Jakubikova, E. Low-Spin Versus High-Spin Ground State in Pseudo-Octahedral Iron Complexes. *Inorg. Chem.* **2012,** *51*, 6011-6019.
(77) Ioannidis, E. I.; Kulik, H. J. Towards Quantifying the Role of Exact Exchange in Predictions of Transition Metal Complex Properties. *J. Chem. Phys.* **2015,** *143*.
(78) Nandy, A.; Chu, D. B. K.; Harper, D. R.; Duan, C.; Arunachalam, N.; Cytter, Y.; Kulik, H. J. Large-Scale Comparison of 3d and 4d Transition Metal Complexes Illuminates the Reduced Effect of Exchange on Second-Row Spin-State Energetics. *Phys. Chem. Chem. Phys.* **2020,** *22*, 19326-19341.

(79) Zhang, D.; Truhlar, D. G. Spin Splitting Energy of Transition Metals: A New, More Affordable Wave Function Benchmark Method and Its Use to Test Density Functional Theory. *J. Chem. Theory Comput.* **2020,** *16*, 4416-4428.
(80) Sgarlata, C.; Mugridge, J. S.; Pluth, M. D.; Tiedemann, B. E. F.; Zito, V.; Arena, G.; Raymond, K. N. External and Internal Guest Binding of a Highly Charged Supramolecular Host in Water: Deconvoluting the Very Different Thermodynamics. *J. Am. Chem. Soc.* **2010,** *132*, 1005-1009.
(81) Ward, M. D.; Hunter, C. A.; Williams, N. H. Coordination Cages Based on Bis(Pyrazolylpyridine) Ligands: Structures, Dynamic Behavior, Guest Binding, and Catalysis. *Acc. Chem. Res.* **2018,** *51*, 2073-2082.
(82) Sure, R.; Grimme, S. Comprehensive Benchmark of Association (Free) Energies of Realistic Host–Guest Complexes. *J. Chem. Theory Comput.* **2015,** *11*, 3785-3801.
(83) Yin, J.; Henriksen, N. M.; Slochower, D. R.; Shirts, M. R.; Chiu, M. W.; Mobley, D. L.; Gilson, M. K. Overview of the Sampl5 Host–Guest Challenge: Are We Doing Better? *J. Comput. Aided Mol. Des.* **2017,** *31*, 1-19.
(84) Xu, P.; Sattasathuchana, T.; Guidez, E.; Webb, S. P.; Montgomery, K.; Yasini, H.; Pedreira, I. F. M.; Gordon, M. S. Computation of Host–Guest Binding Free Energies with a New Quantum Mechanics Based Mining Minima Algorithm. *J. Chem. Phys.* **2021,** *154*.
(85) Nazemi, A.; Steeves, A. H.; Kastner, D. W.; Kulik, H. J. Influence of the Greater Protein Environment on the Electrostatic Potential in Metalloenzyme Active Sites: The Case of Formate Dehydrogenase. *J. Phys. Chem. B* **2022,** *126*, 4069-4079.
(86) Mehmood, R.; Kulik, H. J. Both Configuration and Qm Region Size Matter: Zinc Stability in Qm/Mm Models of DNA Methyltransferase. *J. Chem. Theory Comput.* **2020,** *16*, 3121-3134.
(87) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbo-Dorca, R. Critical Analysis and Extension of the Hirshfeld Atoms in Molecules. *J Chem Phys* **2007,** *126*, 144111.
(88) Lu, T.; Chen, F. Multiwfn: A Multifunctional Wavefunction Analyzer. *J. Comp. Chem.* **2012,** *33*, 580-592.
(89) Van Damme, S.; Bultinck, P.; Fias, S. Electrostatic Potentials from Self-Consistent Hirshfeld Atomic Charges. *J. Chem. Theory Comput.* **2009,** *5*, 334-340.
(90) Bultinck, P.; Ayers, P. W.; Fias, S.; Tiels, K.; Van Alsenoy, C. Uniqueness and Basis Set Dependence of Iterative Hirshfeld Charges. *Chem. Phys. Lett.* **2007,** *444*, 205-208.
(91) Guerra, J. V. S.; Alves, L. F. G.; Bourissou, D.; Lopes-de-Oliveira, P. S.; Szalóki, G. Cavity Characterization in Supramolecular Cages. *J. Chem. Inf. Model.* **2023,** *63*, 3772-3785.
(92) Miklitz, M.; Jelfs, K. E. Pywindow: Automated Structural Analysis of Molecular Pores. *J. Chem. Inf. Model.* **2018,** *58*, 2387-2391.
(93) Martí-Centelles, V.; Piskorz, T. K.; Duarte, F. Cagecavitycalc (C3): A Computational Tool for Calculating and Visualizing Cavities in Molecular Cages. *ChemRxiv* **2024**.
(94) Leung, D. H.; Bergman, R. G.; Raymond, K. N. Highly Selective Supramolecular Catalyzed Allylic Alcohol Isomerization. *J. Am. Chem. Soc.* **2007,** *129*, 2746-2747.
(95) Cordero, B.; Gómez, V.; Platero-Prats, A. E.; Revés, M.; Echeverría, J.; Cremades, E.; Barragán, F.; Alvarez, S. Covalent Radii Revisited. *Dalton Trans.* **2008**, 2832-2838.
(96) Winston, M. S.; Wolf, W. J.; Toste, F. D. Halide-Dependent Mechanisms of Reductive Elimination from Gold(Iii). *J. Am. Chem. Soc.* **2015,** *137*, 7921-7928.

(97) Mecozzi, S.; Rebek, J., Julius. The 55 % Solution: A Formula for Molecular Recognition in the Liquid State. *Chem. Eur. J.* **1998,** *4*, 1016-1022.
(98) Dalvi, V. H.; Rossky, P. J. Molecular Origins of Fluorocarbon Hydrophobicity. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107*, 13603-13607.
(99) Siddiqui, S. A.; Shaik, S.; Kalita, S.; Dubey, K. D. A Porphyrin-Based Molecular Cage Guided by Designed Local-Electric Field Is Highly Selective and Efficient. *Chem. Sci.* **2023,** *14*, 10329-10339.
(100) Nguyen, Q. N. N.; Xia, K. T.; Zhang, Y.; Chen, N.; Morimoto, M.; Pei, X.; Ha, Y.; Guo, J.; Yang, W.; Wang, L.-P.; Bergman, R. G.; Raymond, K. N.; Toste, F. D.; Tantillo, D. J. Source of Rate Acceleration for Carbocation Cyclization in Biomimetic Supramolecular Cages. *J. Am. Chem. Soc.* **2022,** *144*, 11413-11424.
(101) Conifer, C. M.; Law, D. J.; Sunley, G. J.; Haynes, A.; Wells, J. R.; White, A. J. P.; Britovsek, G. J. P. Dicarbonylrhodium(I) Complexes of Bipyridine Ligands with Proximate H-Bonding Substituents and Their Application in Methyl Acetate Carbonylation. *Eur. J. Inorg. Chem.* **2011,** *2011*, 3511-3522.
(102) Field, L. D.; Messerle, B. A.; Rehr, M.; Soler, L. P.; Hambley, T. W. Cationic Iridium(I) Complexes as Catalysts for the Alcoholysis of Silanes. *Organometallics* **2003,** *22*, 2387-2395.
(103) Conifer, C. M.; Taylor, R. A.; Law, D. J.; Sunley, G. J.; White, A. J. P.; Britovsek, G. J. P. First Metal Complexes of 6,6′-Dihydroxy-2,2′-Bipyridine: From Molecular Wires to Applications in Carbonylation Catalysis. *Dalton Trans.* **2011,** *40*, 1031-1033.
(104) Teuma, E.; Loy, M.; Le Berre, C.; Etienne, M.; Daran, J.-C.; Kalck, P. Tandem Carbonylation Reactions: Hydroformylation and Hydroaminomethylation of Alkenes Catalyzed by Cationic [(H2c(3,5-Me2pz)2)Rh(Co)L]+ Complexes. *Organometallics* **2003,** *22*, 5261-5267.
(105) Walker, S. E.; Boehnke, J.; Glen, P. E.; Levey, S.; Patrick, L.; Jordan-Hore, J. A.; Lee, A.-L. Ligand- and Base-Free Pd(Ii)-Catalyzed Controlled Switching between Oxidative Heck and Conjugate Addition Reactions. *Org. Lett.* **2013,** *15*, 1886-1889.

For Table of Contents Use Only



Synopsis

Nanocages are structures that are known to enable rate enhancement of select reaction steps via encapsulation by creating a confined environment like enzymes. To understand the cage's effect on transition metal complexes with identified catalytic activity, we identify complexes from the Cambridge Structural Database with complementary properties as potential guests for the $Ga_4L_6^{-12}$ nanocage. We find that the cage undergoes the most geometric changes and electronic changes to the encapsulated transition metal complexes are relatively uniform.