JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

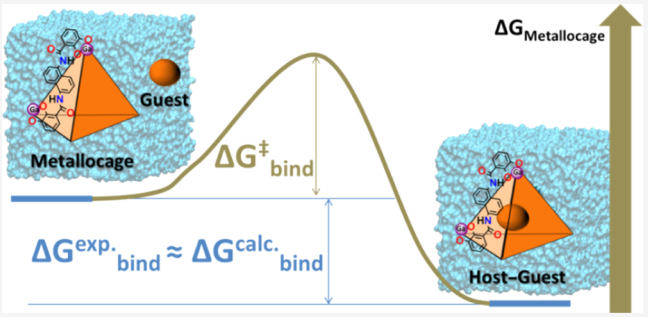
Modeling Kinetics and Thermodynamics of Guest Encapsulation into the $[M_4L_6]^{12-}$ Supramolecular Organometallic Cage

Gantulga Norjmaa, Pietro Vidossich, Jean-Didier Maréchal,* and Gregori Ujaque*

Cite This: *J. Chem. Inf. Model.* 2021, 61, 4370−4381 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The encapsulation of molecular guests into supramolecular hosts is a complex molecular recognition process in which the guest displaces the solvent from the host cavity, while the host deforms to let the guest in. An atomistic description of the association would provide valuable insights on the physicochemical properties that guide it. This understanding may be used to design novel host assemblies with improved properties (i.e., affinities) toward a given class of guests. Molecular simulations may be conveniently used to model the association processes. It is thus of interest to establish efficient protocols to trace the encapsulation process and to predict the associated magnitudes ΔG_{bind} and $\Delta G^{\ddagger}_{bind}$. Here, we report the calculation of the Gibbs energy barrier and Gibbs binding energy by means of explicit solvent molecular simulations for the $[Ga_4L_6]^{12-}$ metallocage encapsulating a series of cationic molecules. The $\Delta G^{\ddagger}_{bind}$ for encapsulation was estimated by means of umbrella sampling simulations. The steps involved were identified, including ion-pair formation and naphthalene rotation (from L ligands of the metallocage) during the guest's entrance. The ΔG_{bind} values were computed using the attach−pull−release method. The results reveal the sensitivity of the estimates on the force field parameters, in particular on atomic charges, showing that higher accuracy is obtained when charges are derived from implicit solvent quantum chemical calculations. Correlation analysis identified some indicators for the binding affinity trends. All computed magnitudes are in very good agreement with experimental observations. This work provides, on one side, a benchmarked way to computationally model a highly charged metallocage encapsulation process. This includes a nonstandard parameterization and charge derivation procedure. On the other hand, it gives specific mechanistic information on the binding processes of $[Ga_4L_6]^{12-}$ at the molecular level where key motions are depicted. Taken together, the study provides an interesting option for the future design of metal−organic cages.



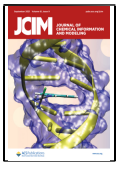
1. INTRODUCTION

Host−guest interactions play an essential role in supramolecular chemistry, a rapidly growing field of research.[1] Among the hosts developed, supramolecular organometallic cages (SOCs) play a prominent role.[2,3] SOCs, also named metal−organic cages (MOCs) or metallocages, are supermolecules generated by self-assembly of metal ions or clusters with organic ligands.[4−8] They have been extensively studied because of their interesting structures that account with a variety of topologies and well-defined shapes at the nanoscale. The capability for tailoring the cavities of the SOCs has led to develop a rich host−guest chemistry with many applications such as recognition of specific molecular targets,[9] the purification of product mixtures,[10] gas sorption,[11,12] biomedicine,[13] and catalysis.[5]

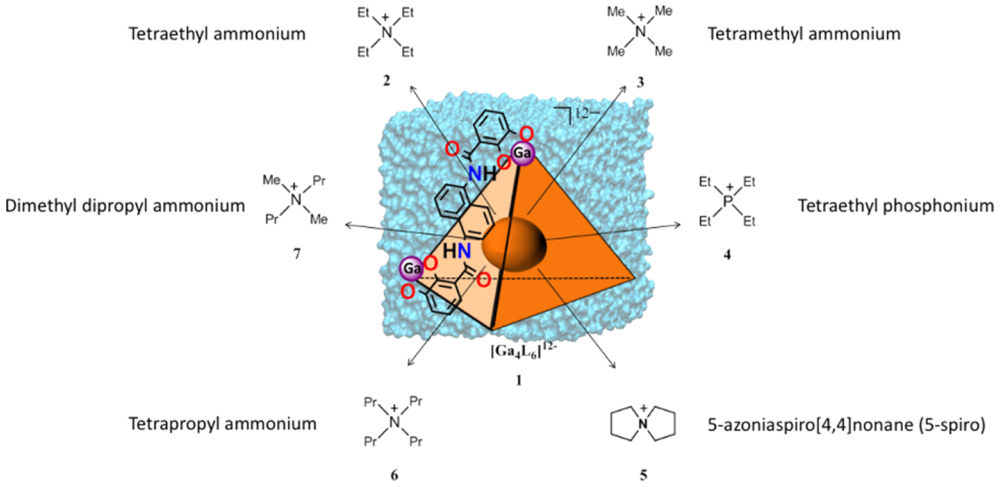
The rules governing the specificity and strength of molecular binding events for these systems are complex. Among key questions is the exact contribution of the different physicochemical variables involved in the process of encapsulation. From the thermodynamic point of view, the process is governed by the Gibbs energy associated with the encapsulating process. However, a molecular description of the process considering the flexibility of the cage, size, and shape complementarities, in between others, is highly desired.

Among the families of SOCs that are receiving considerable attention, the self-assembled tetrahedral metallocage, $[Ga_4L_6]^{12-}$ (**1**, Scheme 1), developed by Raymond and coworkers has shown versatility to encapsulate cationic guests and perform host−guest catalysis.[5] Recent studies by our group and others have demonstrated that computational chemistry can have major implications on understanding the mechanism of metallocage-supported catalysis.[14−19] So far, the works performed were based on quantum chemistry approaches combined with limited conformational sampling

Received: April 7, 2021
Published: September 10, 2021



Scheme 1. Chemical Structures of the $[Ga_4L_6]^{12-}$ Metallocage (the Host) and Cationic Ligands (the Guests) Used in this Work



performed via molecular dynamics (MD) (either based on force fields or quantum approaches).[14,15,17−20] These works were focused at decoding the particularities of the host−guest catalytic mechanism compared to the reaction taking place in solution. Among the most striking results was the identification that the few solvent molecules that access the host cavity are fundamental for the catalytic effect provided by the supramolecular assembly. In these studies, however, another fundamental aspect in host−guest chemistry—the encapsulation process—was not addressed.

The experiments on guest-exchange dynamics showed that the $[Ga_4L_6]^{12-}$ SOC deforms to enlarge an entrance for the passage of the guest, instead of transiently breaking a metal−organic ligand bond.[21,22] Previous MM3 calculations in the gas phase in combination with experiments of the guest displacement from the cage have been reported, although these calculations are based on constrained searches on the potential energy surface that do not account for the full flexibility of the system.[23] However, uncovering the dynamics and structural fluctuations of these supramolecular assemblies in solution is fundamental to properly describe the process and for the rational design of novel cages.[24]

When it comes to simulate binding processes by computational means, techniques allowing extensive conformational samplings are necessary, something unsustainable in a quantum mechanics-based framework whose computational cost is still demanding. In many fields, classical MD simulations, in combination with enhanced sampling techniques, proved successful for the calculation of binding Gibbs energies.[19,25−29] In principle, classical MDs offers an excellent opportunity to study host−guest binding processes at the atomic scale and evaluate their thermodynamics profiles. However, the application of classical MD to SOCs is not free of challenges. One of the most important points is to establish an appropriate protocol in terms of atomic parameterization and conformational sampling. Derivation of metal-compatible force fields is still one of the most complicated steps in parameterization of atomic force fields.[30−33] Regarding the encapsulation process, there are only very recent works addressing this issue.[34,35]

Here, we report a study aimed at both, establishing a computational framework valid to calculate both Gibbs binding energies ($\Delta G_{binding}$) and Gibbs energy barriers ($\Delta G^{\ddagger}_{binding}$) associated with the host−guest encapsulation process, as well as providing an atomistic description for the encapsulation process itself. To do so, we take advantage of the experimental data available on the $[Ga_4L_6]^{12-}$ metallocage (**1**) (Scheme 1), which we used as a benchmarking system. We computed the binding Gibbs energies of several cationic guests into **1** and described the encapsulation process at the molecular level in aqueous solution using MD simulations and subsequent statistical analysis. Furthermore, the potential of mean force (PMF) for the encapsulating process was also investigated. The atomistic description of this process revealed very interesting motions in the metallocage and rearrangements in the system.

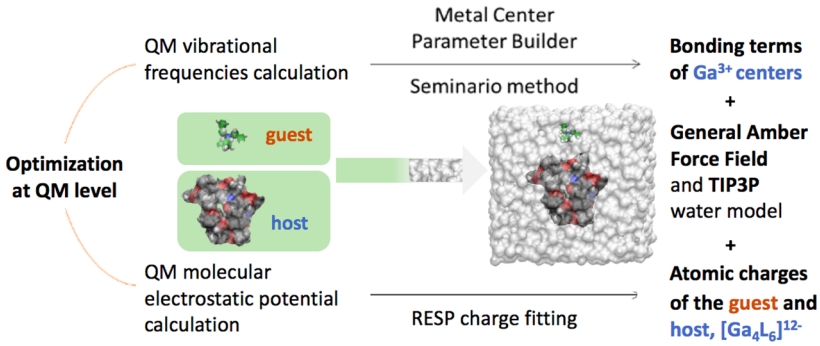
2. COMPUTATIONAL DETAILS

2.1. Derivation of Force Field Parameters. The supramolecular metallocage, **1**, contains a Ga^{3+} metal center on each vertex of the tetrahedron. A bonded scheme was employed to describe these centers, according to which the interactions between the metal and its ligands were modeled using harmonic potentials for bonds and angles (the AMBER force field functional form), while rotations around metal−ligand bonds were not explicitly included (as is customarily done for organometallic systems). The parameters were derived using the Python-based metal center parameter builder (MCPB)[36] developed by Merz and co-workers and based on Seminario's method.[37] The parameters of the bonding terms for the organic moieties of the metallocage and for the guests were taken from the general AMBER force field (gaff).[38,39] Van der Waals parameters were obtained from the optimized potentials for liquid simulations force field (OPLS),[40,41] except those of the metal (Ga^{3+}) taken from the universal force field (UFF).[42] Atomic charges were computed by fitting the molecular electrostatic potential computed at the quantum mechanics (QM) level (either in the gas phase or including an implicit solvent representation—see the Results and Discussion section) according to the restrained electrostatic potential (RESP) method.[43] To this aim, we optimized the full metallocage and each ligand at the density functional theory (DFT) level, as described below. The antechamber[44] program was used to assign atom types and to derive atomic charges (Scheme 2).

2.2. Quantum Chemical Calculations. DFT geometry optimizations were performed with the B3LYP-D3 functional[45−47] (including three body correction terms) using the

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

Scheme 2. General Force Field Parameterization Procedure for Host and Guest Molecules Investigated in This Study



Gaussian 09 program.[48] The SDD pseudopotential[49] and related basis set complementing with a set of d polarization functions[50] were used for gallium, while the 6-31G(d) basis set[51] was used for all other atoms. Optimizations were performed separately for the metallocage and each ligand, both in the gas phase and using the SMD solvation model[50] to account for solvation effects. Binding energy estimates at the QM level account for thermal effects via the quasi-rigid-rotor-harmonic-oscillator (quasi-RRHO) approach.[52] The standard state correction[53] (1.9 kcal/mol to each of the compounds) was also included.

The cavity volumes for the metallocage were obtained using the CAVER Analyst 1.0 software tool[54] employing the default probe radii. The molecular surface areas of the metallocage were calculated using UCSF Chimera.[55]

2.3. Molecular Simulations. MD simulations were performed using the CUDA version of the pmemd program from the AMBER 16 package.[56] The MD simulation box, of size 46 × 49 × 47 Å and treated under periodic boundary conditions, contains metallocage **1**, the guest, 11 potassium counterions, and ∼2500 water molecules (TIP3P).[57] The simulations were performed at constant temperature (298.15 K, using a Langevin thermostat) and pressure (1 bar, using a Monte Carlo barostat[58]). A cutoff of 9 Å was used for nonbonded interactions, with long-range electrostatics interactions accounted by the PME method.[59] A time step of 2 fs was used in plain MD simulations, whereas for Gibbs energy calculations, different time steps were used according to recommendations from the literature.

The absolute Gibbs energies of binding (ΔG_{bind}) of each ligand were computed using the attach−pull−release (APR) method.[60] This method allows for precise calculations of supramolecular host−guest binding, and may thus also be used to evaluate and optimize force fields.[60] Three dummy atoms were used to setup the restraints required by the APR method, together with two atoms of the guest and three of the host (see Figure S12). The distance force constant was set to 5.0 kcal/(mol Å^2), whereas the angle and torsional force constants were set to 100.0 kcal/(mol rad^2). The distance between one atom of the guest and one dummy atom placed at the bottom of the metallocage was increased by 0.4 Å from 6 Å (guest inside the cage) to 24 Å (guest outside) during the simulation. For each simulated window, equilibration and accumulation included 500,000 minimization cycles, 1 ps NVT at 10 K, 100 ps heating to 298.15 K, 250 ps NPT equilibration, and 2.5−25 ns NPT production (depending on the standard error of the mean of the restraint forces). Hydrogen mass repartitioning allowed for a time step of 4 fs. Snapshots from APR simulation are shown in Figure S7.

The relative Gibbs energies of binding ($\Delta\Delta G_{bind}$) were computed using alchemical transformations[61] of guest **2** into guests **3**, **4**, **6**, and **7** (ligand **5** was not considered here because of the spiro substituents). The transformations were performed for the bound and unbound states using 12 values of the alchemical coupling parameter (0.00922, 0.04794, 0.11505, 0.20634, 0.31608, 0.43738, 0.56262, 0.68392, 0.79366, 0.88495, 0.95206, and 0.99078). The windows were sampled sequentially (starting from λ = 0.00922), performing 10 ns NPT simulations each. Since SHAKE was not applied during the alchemical transformation, a time step of 1 fs was used. $\Delta\Delta G_{bind}$ was estimated by thermodynamic integration. Averages of $\partial H/\partial\lambda$ for each window were computed on uncorrelated data after removing the initial nonequilibrated part of the time series.[62]

The PMF for the binding of guest **2** was computed via umbrella sampling.[63] The distance between the center of mass (COM) of the host (defined by the four Ga atoms) and the COM of the guest was selected as the reaction coordinate. A total of 38 windows (spaced by 0.4 Å) were used to cover the range between 0.7 and 15.1 Å (one further window was added in the transition-state region). A harmonic potential with a force constant of 5.0 kcal/mol was applied to sample configurations in each window. The production time was 2.5 ns for each window, and before the production run, every window was minimized and equilibrated as described above for the APR method.[60] A time step of 4 fs was used. The weighted histogram analysis (WHAM) method[64] was used to reconstruct the Gibbs energy profile.

3. RESULTS AND DISCUSSION

The supramolecular tetrahedral metallocage **1** (Scheme 1) offers a particularly interesting system for benchmarking computational protocols for the encapsulation of ligands into metallocages. Indeed, ^{1}H NMR spectroscopy experiments led to valuable observations in binding trends. The values for binding Gibbs energies show that metallocage **1** preferentially encapsulates NEt_4^+, **2** (−6.2 kcal/mol[9] and/or −6.3 kcal/mol),[65] over its smaller NMe_4^+, **3** (−3.4 kcal/mol)[10] and larger NPr_4^+, **6** (−2.8 kcal/mol[9] and/or −4.3 kcal/mol)[65] counterparts.

The first step in the computational study of guest encapsulation in **1** was centered on NEt_4^+ and consisted in assessing an optimal parameterization procedure for the different molecular entities compatible with the AMBER force field. Once the benchmark on this system was performed,

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

the same protocol was applied for the calculation of binding Gibbs energies of other cationic guests, which we compared to experimental data. Next, the Gibbs energy profile for the whole encapsulation process was investigated.

3.1. Parameterization for Binding Gibbs Energy Calculations: $NEt_4^+ \subset [Ga_4L_6]^{12-}$. The metallocage under study is highly charged, $[Ga_4L_6]^{12-}$. One of the first questions in terms of modeling is to assess how the presence of the solvent could affect the physicochemical properties of the metallocage and how this could influence the step of force field parameterization. We started by optimizing the geometry of the metallocage at the QM level. We observed significant differences between vacuum and continuum aqueous solvent conditions in terms of size and shape (Figure 1). For instance,

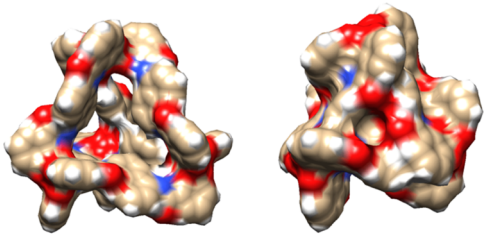


Figure 1. DFT-optimized geometries of the metallocage, **1**, (a) in the gas phase and (b) in the implicit water solvent.

the surface area of the metallocage is 1995 Å^2 in the gas phase, whereas it is 1351 Å^2 in the implicit water solvent. Thus, accounting for solvent effects provokes a shrinking of the metallocage, accompanied by a decrease of the host cavity.

This observation prompted the question on the influence of the medium on the derived parameters for the Ga^{3+} centers. To answer it, we systematically derived the force field parameters from both the gas phase and implicit solvent calculations and checked the effect of the host and the guest parameterizations on the computed Gibbs binding energy (Table 1). As an informative additional calculation, we also provide the binding Gibbs energy obtained with the AMBER-BCC charges for the guest, whose calculated binding Gibbs energy is −8.2 ± 0.6 kcal/mol (entry 9 of Table 1). The binding Gibbs energies are calculated to be between −2.1 and −11.2 kcal/mol depending on the atomic charges of host and guest when the bonding terms for the Ga^{3+} centers are derived in the gas phase. These values are in a narrower range of −2.5 to −6.3 kcal/mol when the bonding terms for the Ga^{3+} centers are derived in the implicit aqueous solvent.

The experimental Gibbs energy of binding of NEt_4^+ **2** into **1** is −6.21 kcal/mol, as estimated from the equilibrium constants measured at 25 °C in water. Table 1 shows that the calculation of both the charges and bonding parameters from continuum solvent calculations is the one with the better agreement with the experimental value (−6.3 ± 0.6 kcal/mol; entry 5). All other protocols lead to values that differ between 2 and 5 kcal/mol from the experiment (except for entry 6 with less than 0.5 kcal/mol difference). These results confirm the importance of accounting for solvent effects during the parameterization.

The values shown in Table 1 illustrate that for Ga^{3+} bonding parameters derived in vacuum, the effect of host charges derived in the gas phase and solvent in the binding Gibbs energy ($\Delta\Delta G_{calc.}^{\circ}$) is 5.8 kcal/mol for guest charges derived in the solvent (entry 4−entry 1), whereas this effect accounts for 9.1 kcal/mol when guest charges are derived in the gas phase (entry 3−entry 2), respectively. The effect of guest charges, in turn, is 1.4 kcal/mol for host charges derived in the solvent (entry 1−entry 2) and 1.9 kcal/mol for host charges derived in the gas phase (entry 3−entry 4). The effect of host charges on the $\Delta G_{calc.}^{\circ}$ is clearly higher than the effect of guest charges.

The analogous analysis taking Ga^{3+} bonding parameters derived in the solvent shows that the effect of host charges derived in the gas phase and solvent ($\Delta\Delta G_{calc.}^{\circ}$) is 3.8 kcal/mol for guest charges derived in the solvent (entry 8−entry 5), whereas this effect is 3.2 kcal/mol when guest charges are derived in the gas phase (entry 7−entry 6). The effect of guest charges, in turn, is null for host charges derived in the gas phase (entry 8−entry 7) and 0.6 kcal/mol for host charges derived in the solvent (entry 8−entry 5). As in the previous case, the effect of host charges on the $\Delta G_{calc.}^{\circ}$ is larger than the effect of guest charges, but the differences obtained are reduced.

As far as the bonding terms ($\Delta G_{calc.}^{\circ}$) are concerned, regarding bonding terms for Ga^{3+} calculated in vacuum or the solvent, those obtained in vacuum tend to give lower binding energies (more negative) than their counterparts derived in the solvent (except for the comparison of entry 3 with entry 7). When all parameters are derived in vacuum (entry 3 of Table 1), the computed Gibbs binding energy is −2.1 ± 0.5 kcal/mol, providing an estimate much worse than that obtained when all parameters are derived in the solvent (entry 5, −6.3 ± 0.6 kcal/mol). Indeed, the best agreement with experiment is obtained when all parameters are derived in the solvent. Thus, this is the selected parameterization protocol used in the following part of the study.

At this point, it is worth comparing the above results from simulations with the binding estimate based on quantum

Table 1. Binding Gibbs Energies for 2 ⊂ 1 Computed with the APR Approach Using the Force Field Parameters Derived from the Gas Phase or Continuum Water Solvent Calculations

	bonding terms of Ga^{3+} centers derived in	host charges calculated in	guest charges calculated in	$\Delta G_{calc.}^{\circ}$	$\Delta G_{exp.}^{\circ}$
entry 1	vacuum	solvent	solvent	−9.8 ± 0.8	
entry 2		solvent	gas phase	−11.2 ± 0.8	
entry 3		gas phase	gas phase	−2.1 ± 0.5	
entry 4		gas phase	solvent	−4.0 ± 0.6	
entry 5	**solvent**	**solvent**	**solvent**	**−6.3±0.6**	**−6.2**
entry 6		solvent	gas phase	−5.7 ± 0.6	
entry 7		gas phase	gas phase	−2.5 ± 0.6	
entry 8		gas phase	solvent	−2.5 ± 0.5	
entry 9		solvent	AMBER-BCC	−8.2 ± 0.6	

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

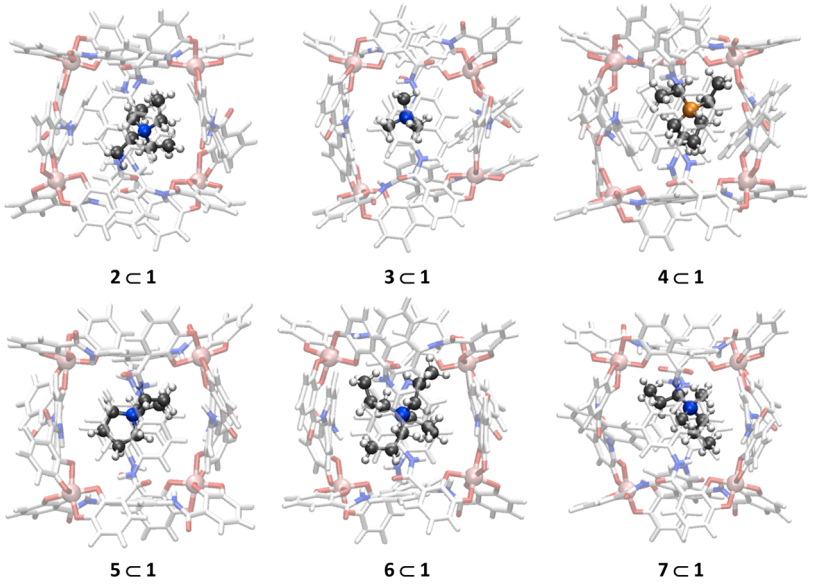


Figure 2. Representative structures of the host−guest complexes considered in this study.

chemical calculations. The $\Delta G_{binding}$ obtained for this system with an implicit SMD solvent at the B3LYP-D3 level is −19.6 kcal/mol, overestimating the strength of binding and suggesting that accounting for an explicit solvent and flexibility are important ingredients for accurate modeling. The DFT-optimized structures are shown in Figure S1. It is worth mentioning that a mixed implicit−explicit solvent method was shown to work properly for complexation of self-assembled capsules in water, but a thorough study to determine the number of solvent molecules to be included is needed.[66]

Overall, based on the results of this part of the study, the APR protocol based on force field parameters (both bonding terms and atomic charges) derived from implicit solvent DFT calculations provides a very good agreement with experimental Gibbs binding energies.

3.2. Computing Binding Energies for a Set of Guests.

3.2.1. Robustness of the Method. To further assess the suitability of the parameterization procedure characterized in the previous part of the study for a general purpose in supramolecular chemistry, the binding Gibbs energies of five additional cationic guests (**3**, **4**, **5**, **6**, and **7**) with known experimental values were calculated (Figure 2). The same protocol based on the parameters derived according to the scheme described in the previous section was employed.

Calculated and experimental energies of binding are in very good agreement for all guests (Table 2). For complexes **4** ⊂ **1**, **5** ⊂ **1**, and **7** ⊂ **1**, differences are ca. 1.0 kcal/mol, while for **3** ⊂ **1** and **6** ⊂ **1**, they are ca. 2 kcal/mol. It is quite remarkable to reach such quality, indicating that the combination between force field and the sampling protocol is quite robust. Plotting the experimental against the computed values shows that there is good linear relationship between both quantities, although the experimental range is narrower, leading to a slope <1 (Figure S4).

Table 2. Binding Gibbs Energies of the Host−Guest Complexes Considered in This Study

host−guest complex	$\Delta G_{calc.}{}^{\circ}$	$\Delta G_{exp.}{}^{\circ}$[22]	$\Delta\Delta G$
2 ⊂ 1	−6.28 ± 0.56	−6.21 ± 0.01 (−6.29)[65]	0.07
3 ⊂ 1	−1.15 ± 0.82	−3.42 ± 0.001[65]	2.27
4 ⊂ 1	−8.32 ± 0.63	−6.85 ± 0.03	1.47
5 ⊂ 1	−4.36 ± 0.60	−5.59 ± 0.01[65]	1.23
6 ⊂ 1	−0.62 ± 0.73	−2.75 ± 0.03 (−4.25)[65]	2.13
7 ⊂ 1	−5.86 ± 0.97	−4.80 ± 0.03	1.06

The computation of relative free energies of binding using alchemical transformations is a convenient technique to compare the affinity of ligands for a given receptor.[67−70] Converting guest **2** into each of the remaining guests, both in solution and encapsulated in the metallocage, we estimated $\Delta\Delta G_{2 \rightarrow j}$ (**j** = **3**, **4**, **6**, and **7**; see Table S2). The ranking of guests based on the $\Delta\Delta G$ values is in line with the experimental trend. For guests **3** and **6**, the estimates are consistent with the APR results (within statistical errors), whereas for compounds **4** and **7**, the difference in the computed $\Delta\Delta G$ is significant. For guest **4**, the FEP estimate is closer to the experimental data, whereas for guest **7**, the APR estimate is closer to the experimental result. The relative potential energies of binding were calculated, but they fail to reproduce the experimental trend (see Table S2).

Importantly, calculated binding Gibbs energies with APR reflect the same trend between encapsulation of compounds than those found experimentally (Table 2). Thus, the largest binding energies correspond to those species bearing four Et substituents (compounds **2** and **4**), whereas the lowest binding energies correspond to species bearing four Me or Pr substituents (compounds **3** and **6**; the origin for the similar behavior, however, is different for both species, vide infra). The binding energies of compounds **5** and **7**, bearing two cyclic butyl substituents and either two Me and two Pr substituents, respectively, lie in between. Moreover, guests with linear alkyl chains, **2** and **4** (−6.3 and −8.3 kcal/mol) have a higher binding affinity than their cyclic analogue **5** (−4.4 kcal/mol), which is consistent with the experimental observations on relative binding affinities between *n*-alkanes (C5−C8) and their cyclic isomers.[71] This agreement further validates the computational protocol and suggests that molecular simulations may provide access to the molecular features that control the binding of guests to the SOC host. This is extremely relevant because the energetic behaviors between

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

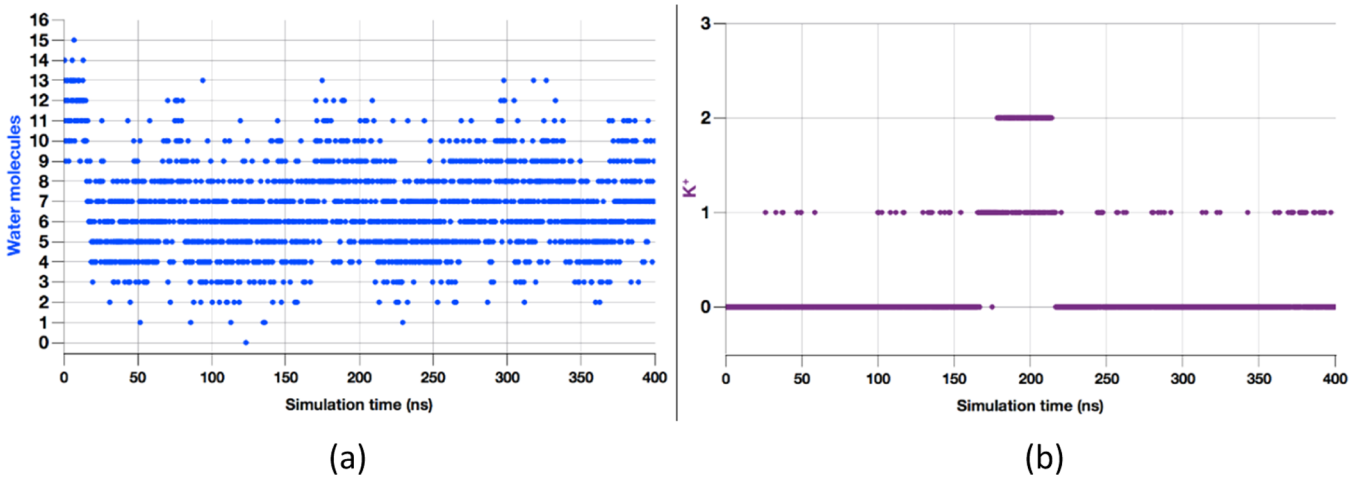


Figure 3. (a) Number of solvent water molecules in the metallocage during classical MD simulation. (b) Number of K^+ located at a distance < 5 Å from the COM of the metallocage during classical MD simulation.

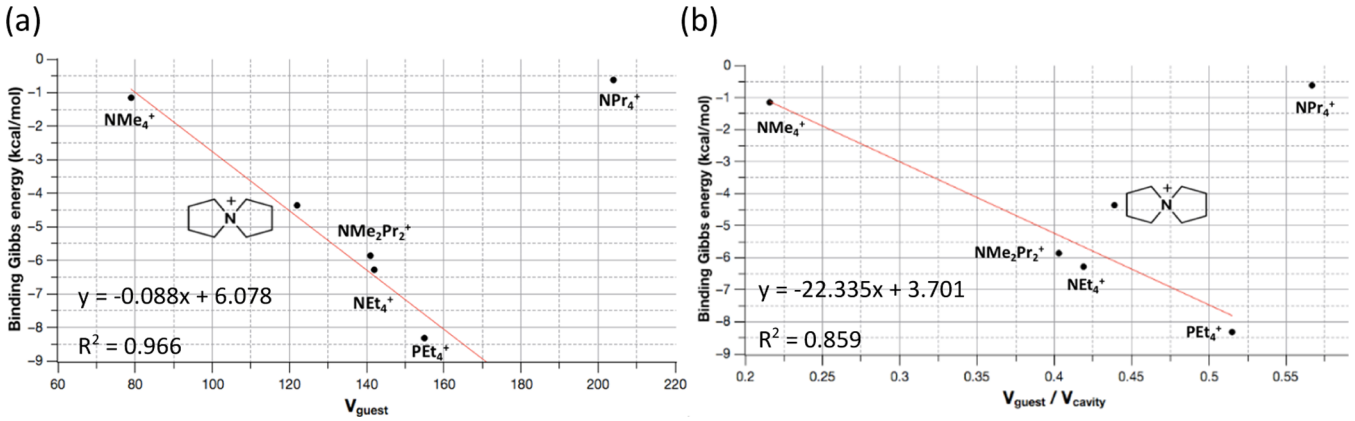


Figure 4. (a) Correlation between binding Gibbs energy and V_{guest}. (b) Correlation between binding Gibbs energy and V_{guest}/V_{cavity}. In both cases, regression excluding compound **6**, NPr_4^+.

the different host−guest partners are somehow counterintuitive. Indeed, one would expect some correlation between the size and shape of the substituent such as methyl < ethyl < propyl. The simulations offer therefore a unique opportunity to better understand the molecular grounds of the energetics of encapsulation.

3.2.2. Molecular Grounds for Binding Energy Trends. Owning a computational protocol that could help in defining the relative affinities of different guests to the same metallocage is crucial in supramolecular chemistry and constitutes one of the objectives of this work. Computing reliable binding affinities can be a precious tool for designing supramolecular hosts. Based on previous studies on supramolecular systems,[14,15,18,32,33,65,72] several variables could relate to such prediction: (1) the cavity volume of the host alone, (2) the volume of the guest, (3) the packing values, that is, the ratio between the volume of the guest in front of the volume of the cavity of the host, and (4) the possible involvement of solvent molecules. It is likely although that the real relationship lies as a combination of these different factors.

We started analyzing the metallocage **1** in solution without any guest. The volume for the free state of the cage ranges from void (meaning not even one solvent molecule could fit) to 480 $Å^3$. The number of water molecules inside the cavity varies from 0 to 15 (Figure 3a), with an average of 6−9; this result is in agreement with that obtained by AIMD calculations of 9 ± 1 water molecules inside the metallocage.[73] The number of K^+ counterions in the close environment of the cage (up to 11 Å from the COM of the cage) mostly fluctuates between 3 and 6 (Figure S1), with up to two K^+ inside the cavity (Figure 3b). For comparison, in the case of the encapsulation of a substrate, that is, **2** ⊂ **1**, the volume of the cavity fluctuates from ca. 180 to 530 $Å^3$, with a mean of 318 ± 51 $Å^3$ (Figure S3).

The large variability of the size of the cavity of the cage clearly indicates a high flexibility of the host. Importantly, these MD simulations suggest that the flexibility of the cage enables adaptation of its cavity to the guest to be encapsulated. Our purpose was to look for a correlation, if any, between some properties of the host and guest and the binding process. Interestingly, no correlation is observed between the binding Gibbs energy and the size of the cavity (the volume of the cavity is calculated for the most populated structure along the dynamics) for the six systems under study in this work (Figure S5). The same occurs when plotting the volume of the cavity and the volume of the guest, meaning that the structuring of the cage was not only a product of the size of the guest alone (Figure S6). The correlation between binding Gibbs energy and volume of the guest is surprisingly good, especially if guest **6** is not considered (R^2 = 0.96 for theoretical values (Figure

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

4a) and R^2 = 0.79 for experimental ones). The larger the V_{guest}, the greater (more negative) the Gibbs binding energy. This is accomplished until the volume of the guest overcomes a certain threshold: all the guests fitting the correlation have a volume in the range 80−160 Å^3, whereas **6** has a volume of ∼200 Å^3. Further analysis of the **6** ⊂ **1** complex reveals that this guest tends to be partially outside the cavity (Figure S10), thus explaining the reason why the binding energy decreases dramatically despite having a larger volume. Overall, for those guests with a volume below the threshold, V_{guest} is an excellent predictive indicator for binding affinity trends. Those with larger volumes are expected to lie partially outside the cavity, consequently having lower binding energies.

The complexing properties were also evaluated as a function of the V_{guest}/V_{cavity} ratio (Figure 4b). Close inspection of the values obtained shows that compound **6** acts also as a clear outlier. Its removal from the set allows an acceptable correlation (R^2 = 0.86, Figure 4b). The best binding affinity is obtained for the encapsulation of **4**, whereas for compound **6**, displaying a similar V_{guest}/V_{cavity} ratio, presents the worse binding affinity. The same reasoning for this guest explained before can be applied here: **6** tends to be partially outside the cavity of the metallocage. This explains why it is the worse binder among those considered for this metallocage. Overall, for those systems below the threshold, the V_{guest}/V_{cavity} ratio, ranging from 0.22 to 0.51, it is also a good predictive indicator of binding affinity trends.

Next, we analyzed the packing coefficient of the guest in the internal cavity of the molecular cage. In order to be able to calculate packing coefficients, the presence of solvent molecules inside the cavity, if any, must be taken also into account. For that, we performed additional 200 ns MD simulation for each of the guests. The analysis of the structures along the MD could identify the average solvent molecules inside the cavity (Table S1): for **3**, there are four water molecules (V_{guest} ∼ 80 Å^3), for **5**, there are three water molecules (V_{guest} ∼ 120 Å^3), for **2**, **4**, and **7**, there are two water molecules (V_{guest} ∼ 150 Å^3), and for **6**, there are no water molecules (V_{guest} ∼ 200 Å^3). The V_{cavity}, in turn, remains around 300 Å^3 for all the guests, except for **6**, where the volume rises up to ∼450 Å^3 (close to the highest volume of 480 Å^3 found when encapsulating the largest amount of water, vide supra). A comparison among the guests indicates that the larger the V_{guest}, the lower the number of water molecules inside the cavity. This allows keeping the packing values ranging from 0.55 to 0.63 (Table S1). These values are close to the theoretical value of 0.55 proposed by Rebek,[74] constituting a rule of the thumb in supramolecular chemistry: the closer the packing to 55%, the better the binding affinity. For the case of **6**, the packing coefficient is 0.45, in line with this guest having the least binding energy.

The results obtained up to here clearly indicate that the APR method provides with very relevant results for the prediction of $\Delta G_{binding}$. Moreover, among the parameters computed, both V_{guest} and V_{guest}/V_{cavity} produce good correlations with binding energies (for guests with a volume lower than a threshold). Thus, these two parameters are potentially useful for virtual screening of the ability of this metallocage to bind a guest. To satisfy Rebek's rule on the packing coefficient, the solvent molecules present inside the cavity need to be also considered, which can conveniently be obtained by performing explicit solvent MD simulations. Overall, pure volumetric analysis of the guests could be already indicative of the binding affinity with the metallocage. This also highlights that computational approaches, such as the one we tested here, could be a good ally for both metallocage and guest selection and useful for supramolecular design.

3.3. Encapsulation of NEt_4^+ in the Metallocage: Molecular Insights on Encapsulation Profiles. In the view to develop a computational framework able to describe the encapsulation of guest molecules into metallocages, we extended our study to the calculation of the Gibbs energy barriers of the encapsulation process. With this purpose, we used an umbrella sampling approach[63] and calculated the PMF associated with the entrance of a guest in the metallocage **1**. Such kind of simulations are computationally much more demanding than those to obtain binding Gibbs energies, and we thus decided to focus on the encapsulation of NEt_4^+, **2**, in the metallocage (system **2** ⊂ **1**; Figure 5). This approach

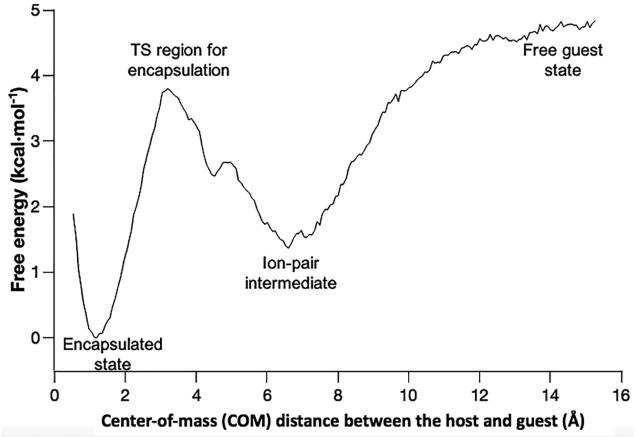


Figure 5. Free energy profile (PMF) of encapsulation process of NEt_4^+ by metallocage **1**.

allows a thorough exploration of the mechanism of guest encapsulation under realistic conditions. The variable selected as the reaction coordinated to analyze the encapsulation process was the distance between the centers of mass of the host and guest (see computational details).

The encapsulation process of guest **2** in a macromolecular host **1** can be segmented in two steps. The first event consists in the formation of an ion-pair complex (Figure 5: ion-pair intermediate) by the association of the cationic guest, NEt_4^+, with the metallocage from the two isolated species in solution (Figure 5: free guest state). This event was experimentally detailed by analyzing the host−guest encapsulation equilibria.[75] In this state, the guest remains on the outer surface of the highly anionic host, $[Ga_4L_6]^{12-}$. This process is exothermic due to strong attractive interactions (as observed in the Gibbs energy profile). The second event consists in the penetration of the guest from exterior into the internal cavity of the metallocage (Figure 5: from the ion-pair to encapsulated state). This step is favorable, although with a barrier associated with the entry of the cationic guest into the metallocage.

The flexibility of the metallocage has been demonstrated by the large variability of the size of the cavity shown during the MD simulations, as well as by encapsulating guests with different sizes and shapes inside the cavity. During the guest release and encapsulation, the deformation of the metallocage by movements of six naphthalene-based bisbidentate catechol ligands has been proposed, but the process has not been yet

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

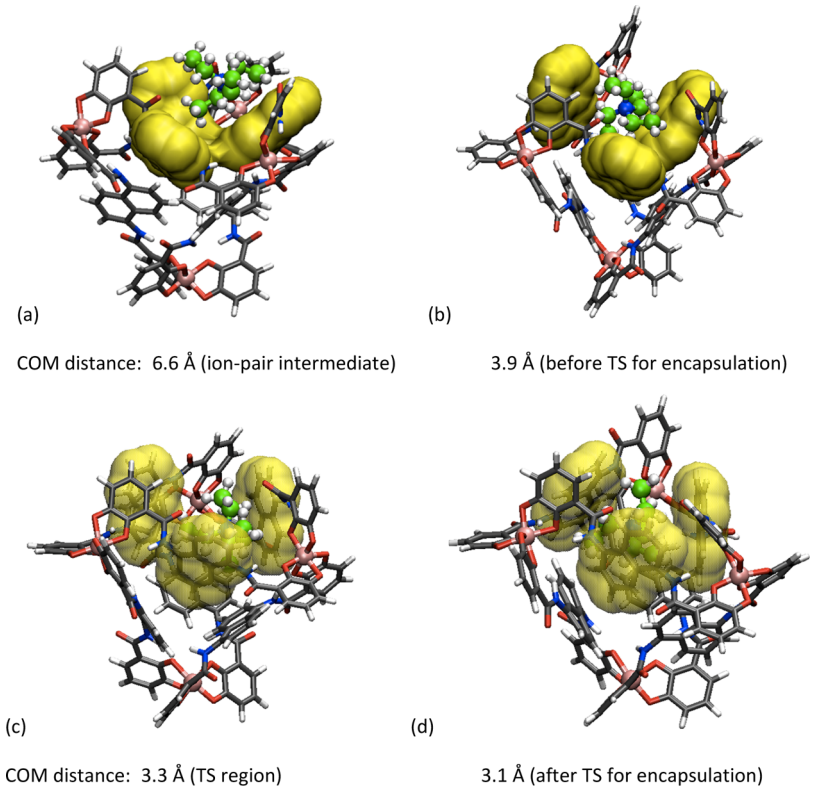


Figure 6. Representative molecular structures of the encapsulation of the guest into the metallocage. The solvent molecules are not shown for clarity. COM = center of mass.

described at the atomic level in explicit solvents for any guest. Therefore, we intended to discuss this process in an illustrative way from MD with an explicit solvent and counter ions. The general representation of the process is shown by the most populated molecular structures in Figure 6 (a different view is given in Figure S8).

In the free guest state (Figure 5), at a distance between the COM of the metallocage and the COM of the guest around 15.0 Å, the metallocage and the guest are both fully solvated. The ion-pair intermediate ($NEt_4^+ \cdot [Ga_4L_6]^{12-}$) forms at a distance of ∼6.6 Å between their respective centers of masses (Figure 5). A representative geometry of this intermediate is shown in Figure 6a. In this intermediate, the guest is positioned close to one of the four triangular faces of the tetrahedron host, which is formed by three naphthalene rings of the metallocage and shaped as a "concave cup" (Figure 6a). This is the most stable place for NEt_4^+ outside the metallocage in aqueous solution. The interactions between the host and the guest can be described as C−H···π interactions between the C−H bonds of the guest and the naphthalene rings of the metallocage and C−H···O(amide) interactions between the C−H bonds of the guest and the amide oxygen of the metallocage. Moreover, the overall charge is also at play. Interestingly, the internal cavity of the metallocage at this stage is surprisingly small, with a volume lower than 30 $Å^3$ (average of the cavity volumes of the three most populated structures). This is due to the fact that the naphthalene rings rotate to accommodate the guest, provoking that the inside cavity shrinks considerably (with no other molecules inside the cavity). Nevertheless, as previously shown, the system is dynamic and this volume fluctuates. The calculated Gibbs energy associated with this ion-pair intermediate state is −3.4 kcal/mol relative to the free guest state (Figure 5); this is in excellent agreement with the estimated value of −3.5 kcal/mol from the experimental observations.[75]

The following step is the encapsulation of the guest from the ion-pair state into the internal cavity of the metallocage (Figure 5: from ion-pair to encapsulated state). When the guest passes through the ion-pair intermediate state to the transition state for encapsulation, the "concave-cup" structure of the face of the metallocage changes in shape by rotational movements of the naphthalene rings of the metallocage. To illustrate this, we selected a point where the distance between the COM of the metallocage and the COM of the guest, NEt_4^+, is 3.9 Å (0.6 Å prior to the transition state for encapsulation; Figure 6b). At this point, the cavity of the metallocage is calculated to be 233 $Å^3$ (calculated for a representative structure of the most populated conformation; it includes the zone comprised among the three naphthalene rings; Figure S9). This shows that the rotation of the naphthalene rings significantly modifies the accessible cavity of the metallocage.

The transition-state region for the encapsulation is calculated to be at the distance of around 3.3 Å between the COM of the metallocage and the COM of NEt_4^+ (a representative snapshot is shown in Figure 6c). This separation is about half of the distance observed for the ion-pair intermediate. The three naphthalene rings of the metallocage are positioned nearly parallel to each other. They still keep C−H···π interactions and the C−H···O(amide) interactions with NEt_4^+ as in the ion-pair intermediate state. The cavity volume of the metallocage is ∼350 $Å^3$, indicating that the cavity increases at the transition state. For comparison, we also calculated the volume of the cavity at a distance 3.1 Å after the transition state (Figure 6d), which turned out to be 238 $Å^3$,

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

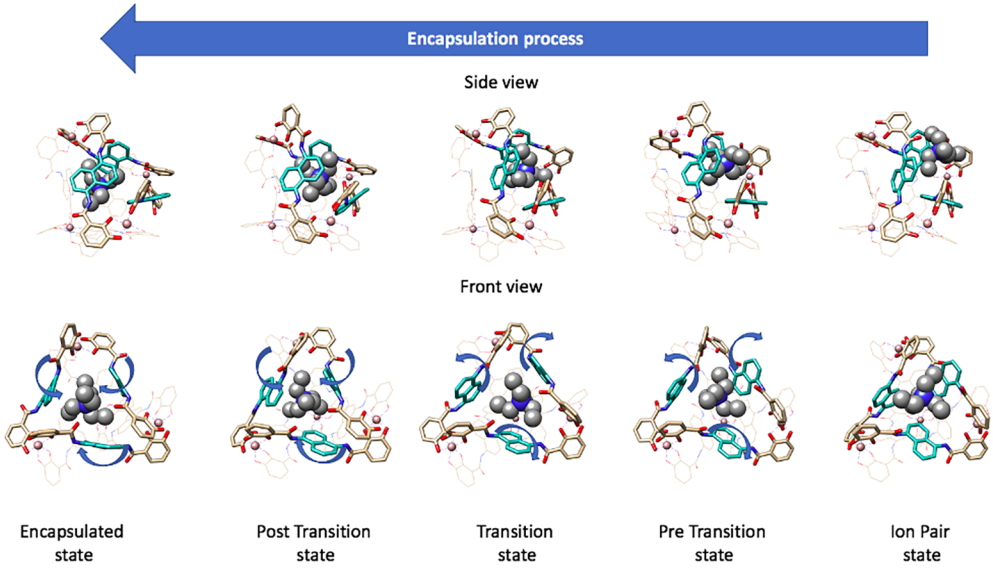


Figure 7. Schematic view of the molecular motions of the encapsulation process from the ion-pair state (right) to encapsulated state (left). Branches of the metallocages involved in controlling guest entrance are represented in stick with the naphthalene groups in cyan. Branches from the opposite side of the cage and with little implication in the transitional motion are represented in wires. The guest molecule is represented in spheres.

showing that the cavity shrinks significantly after the transition state (Figure S9). At this point, we observed that one of the naphthalene rings of the metallocage rotates in the closing direction of the guest-entering face of the tetrahedron host (Figure 6d). A representation of the whole motion for the encapsulation process is shown in Figure 7.

After the guest passes through the transition state for encapsulation, the encapsulated state is observed at a distance of 1.1 Å between the COM of the metallocage and the COM of NEt_4^+. This indicates that the guest does not sit at the center of the cavity, and it relays closer to one of the host walls. In the encapsulated state, the volume of the cavity of the metallocage is 272 $Å^3$ (calculated for a representative structure of the most populated conformation). There is a solvent water molecule found inside the cavity of the metallocage along with NEt_4^+; the calculated packing coefficient is 58%, which is quite close to Rebek's 55% rule[74] and the classical plain MD result (56%). This shows that the cage is able to vary its volume, trying to keep the packing coefficient close to the Rebek rule.

A closer look on the encapsulation process shows a transient state where the metallocage generates two cavities separated by one of the naphthalene-based ligands. The distance between COMs of the host and guest is 5.1 Å, which is a state between the ion-pair intermediate (6.6 Å) and the transition state (3.3 Å). The two cavities, with volumes of 129 and 188 $Å^3$ are shown in Figure 8, in cyan and blue, respectively. Interestingly, one of the cavities (in cyan in Figure 8) includes K^+ with three solvent water molecules, while the other (in blue in Figure 8) is occupied by the guest, NEt_4^+.

The presence of the potassium at this particular step of the encapsulation process prompted us to hypothesize a mechanistic role of this ion and a possible exchange between K^+ (inside the cage) with the guest (outside the cage). Nevertheless, the analysis of the distribution of the K^+ cations shows that during most of the simulation time, there are no K^+ cations inside the cavity (Figure S13) and that the presence of K^+ in one of the cavities formed during migration of the ligand is more likely to be due to the proximity of the ions from the

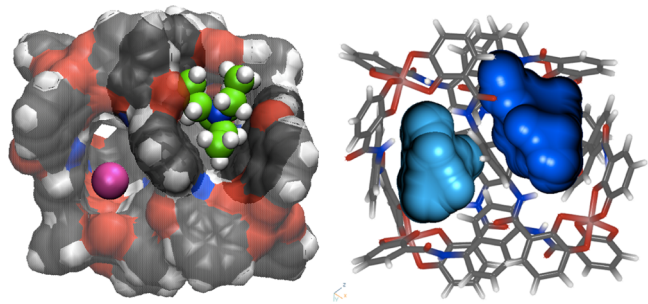


Figure 8. Two different representations of the cavities generated during the encapsulation of NEt^{4+}. They are taken from a snapshot of the most populated structure of MD simulations at a COM distance of 5.1 Å. K^+ in purple.

cage during the simulation. Our work suggests therefore that the mechanism proceeds by a simple encapsulation of the guest to the metallocage more than an ion−ligand exchange process.

The Gibbs energy barrier for encapsulation is calculated to be 2.4 kcal/mol from the ion-pair intermediate (Figure 5), whereas it is 3.8 kcal/mol for the guest releasing from the encapsulated state. The encapsulated state is more stable than the ion-pair intermediate by 1.4 kcal/mol, which is in very good agreement with 2.7 kcal/mol obtained from the experiment.[75] Moreover, the encapsulation of NEt_4^+ within the metallocage is found to be a spontaneous process with a rather low barrier. This is in line with the displacement experiment: encapsulated NMe_4^+ in the metallocage **1** is readily displaced by NEt_4^+, and the process is too fast to be monitored by ^1H NMR spectroscopy.[9]

From a computational point of view, the binding Gibbs energy estimated from the PMF (ca. −5.0 kcal/mol) is consistent with the APR result (−6.3 kcal/mol). The small difference is likely due to the different sampling strategies, probably limited in the umbrella sampling simulation at large host−guest separation.

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

4. CONCLUSIONS

This study shows that classical MD in an explicit solvent with proper parameters provides the level of accuracy required for analyzing the encapsulation of cationic species by a supramolecular metallocage. Importantly, for the sake of future modeling experiments on such kinds of systems, here, the best results were obtained when deriving the force field parameters using QM calculation, accounting for solvent effects implicitly. The binding Gibbs energies were calculated to be in very good agreement with the experiment.

The binding energies of different guests to the metallocage were found to correlate very well with the volume of the guest (V_{guest}) and the V_{guest}/V_{cavity} ratio. Interestingly, these parameters are available with low computational efforts. These correlations showed an outlier among the guests considered. It is the one that does not fit into the cavity, lying partially outside, despite showing good packing predictions. Calculated packing coefficients for the guests are in the range of 0.55−0.63, close to Rebek's proposal of 0.55, provided that the solvent molecules inside the cavity are included in their evaluation, as revealed by the MD simulations. This work also sustains the validity of Rebek's rule as a general trend to predict favorable encapsulation and also that atomistic modeling such as the one performed here can provide with information on when this rule reaches its limitations (mainly related to the shape and size host−guest complementarity).

The simulations allowed obtaining the atomic description for the encapsulation process. The encapsulation of the guest into the metallocage takes place in two steps. First, an outer surface ion-pair intermediate forms: the guest faces one of the tetrahedron's sides of the host, which turns concave due to the conformation adopted by the naphthalene-base ligands. Then, entering of the guest into the metallocage is regulated by the rotational movements of three naphthalene ligands of the metallocage; rotation of a naphthalene ligand creates two separate cavities. After the guest passes through the transition state, the metallocage closes the aperture and interacts with the guest by its interior surface. In the guest-encapsulated state, water molecules can be encapsulated along with the guest depending on the guest's size, shape, and interactions.

Overall, properly derived force field parameters in combination with statistical mechanics methods allow computing host−guest binding energies for SOCs along with providing an atomic description of the encapsulation process. This offers an interesting option for future development in MOC design.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00348.

Data for analyzing the MD trajectories; plots with correlations of Gibbs binding energies; figures of the encapsulation process; derived force field parameters; and Cartesian coordinates for the optimized structures (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Jean-Didier Maréchal − *Departament de Química and Centro de Innovación en Química Avanzada (ORFEO-CINQA), Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Catalonia, Spain;* orcid.org/0000-0002-8344-9043; Email: jeandidier.marechal@uab.cat

Gregori Ujaque − *Departament de Química and Centro de Innovación en Química Avanzada (ORFEO-CINQA), Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Catalonia, Spain;* orcid.org/0000-0001-5896-9998; Email: gregori.ujaque@uab.cat

Authors

Gantulga Norjmaa − *Departament de Química and Centro de Innovación en Química Avanzada (ORFEO-CINQA), Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Catalonia, Spain*

Pietro Vidossich − *Laboratory of Molecular Modeling and Drug Discovery, Istituto Italiano di Tecnologia, 16163 Genova, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c00348

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

Metal Center Parameter Builder;[36] AMBER 16 Package;[56] (https://ambermd.org) Gaussian09 Program, Revision D.01;[48] (https://gaussian.com); the WHAM Method[64] (http://membrane.urmc.rochester.edu/wordpress/?page_id=126)

■ ACKNOWLEDGMENTS

The authors acknowledge the financial support of the Spanish MINECO-FEDER (grants CTQ-2017-87889-P, PID2020-116861GB-I00, and RED2018-102387-T). UAB is also acknowledged by a PIF grant to G.N., and Generalitat de Catalunya for grant 2017SGR1323 is acknowledged.

■ ABBREVIATIONS

DFT, density functional theory; QM, quantum mechanics; B3LYP-D3, dispersion-corrected Becke three-parameter Lee−Yang−Parr exchange−correlation functional; SMD, solvent model density; PCM, polarizable continuum model; quasi-RRHO, quasi-rigid-rotor-harmonic-oscillator approximation; MD, molecular dynamics; MCPB.py, Python-based metal center parameter builder program; RESP, restrained electrostatic potential method; WHAM, weighted histogram analysis method

■ REFERENCES

(1) Steed, J. W.; Atwood, J. L. *Supramolecular Chemistry*; Wiley: Chichester, U.K., 2009.

(2) 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.

(3) Cook, T. R.; Stang, P. J. Recent Developments in the Preparation and Chemistry of Metallacycles and Metallacages via Coordination. *Chem. Rev.* **2015**, *115*, 7001−7045.

(4) Ballester, P.; Fujita, M.; Rebek, J. Molecular Containers. *Chem. Soc. Rev.* **2015**, *44*, 392−393.

(5) 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.

(6) 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*, 734−777.
(7) Leenders, S. H. A. M.; Gramage-Doria, R.; De Bruin, B.; Reek, J. N. H. Transition Metal Catalysis in Confined Spaces. *Chem. Soc. Rev.* **2015**, *44*, 433−448.
(8) Zhang, D.; Ronson, T. K.; Nitschke, J. R. Functional Capsules via Subcomponent Self-Assembly. *Acc. Chem. Res.* **2018**, *51*, 2423−2436.
(9) Voloshin, Y.; Belaya, I.; Krämer, R. *The Encapsulation Phenomenon: Synthesis, Reactivity and Applications of Caged Ions and Molecules*; Springer International Publishing, 2016.
(10) García-Simón, C.; Garcia-Borràs, M.; Gómez, L.; Parella, T.; Osuna, S.; Juanhuix, J.; Imaz, I.; Maspoch, D.; Costas, M.; Ribas, X. Sponge-like Molecular Cage for Purification of Fullerenes. *Nat. Commun.* **2014**, *5*, 5557.
(11) Li, J.-R.; Zhou, H.-C. Bridging-ligand-substitution strategy for the preparation of metal−organic polyhedra. *Nat. Chem.* **2010**, *2*, 893−898.
(12) Park, J.; Perry, Z.; Chen, Y.-P.; Bae, J.; Zhou, H.-C. Chromium(II) Metal−Organic Polyhedra as Highly Porous Materials. *ACS Appl. Mater. Interfaces* **2017**, *9*, 28064−28068.
(13) Zhu, C.-Y.; Pan, M.; Su, C.-Y. Metal-Organic Cages for Biomedical Applications. *Isr. J. Chem.* **2019**, *59*, 209−219.
(14) Norjmaa, G.; Maréchal, J.-D.; Ujaque, G. Microsolvation and Encapsulation Effects on Supramolecular Catalysis: C-C Reductive Elimination inside $[Ga_4L_6]^{12-}$ Metallocage. *J. Am. Chem. Soc.* **2019**, *141*, 13114−13123.
(15) Norjmaa, G.; Maréchal, J.-D.; Ujaque, G. Reaction Rate Inside the Cavity of $[Ga_4L_6]^{12-}$ Supramolecular Metallocage Is Regulated by the Encapsulated Solvent. *Chem.—Eur. J.* **2020**, *26*, 6988−6992.
(16) Welborn, V. V.; Pestana, L. R.; Head-Gordon, T. Computational Optimization of Electric Fields for Better Catalysis Design. *Nat. Catal.* **2018**, *1*, 649−655.
(17) 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.
(18) Welborn, V. 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.
(19) 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.
(20) Sciortino, G.; Norjmaa, G.; Maréchal, J. D.; Ujaque, G. "Catalysis by Metal Organic Cages: A Computational Perspective". *Supramolecular Catalysis—New Directions and Developments*, ISBN 978-3-527-34902-9; van Leeuwen, W.N.M., Raynal, M., Eds; Wiley-VCH: Weinheim, 2022.
(21) Davis, A. V.; Raymond, K. N. The Big Squeeze: Guest Exchange in an M_4L_6 Supramolecular Host. *J. Am. Chem. Soc.* **2005**, *127*, 7912−7919.
(22) Davis, A. V.; Fiedler, D.; Seeber, G.; Zahl, A.; van Eldik, R.; Raymond, K. N. Guest Exchange Dynamics in an M_4L_6 Tetrahedral Host. *J. Am. Chem. Soc.* **2006**, *128*, 1324−1333.
(23) Mugridge, J. S.; Bergman, R. G.; Raymond, K. N. Does Size Really Matter? The Steric Isotope Effect in a Supramolecular Host-Guest Exchange Reaction. *Angew. Chem., Int. Ed.* **2010**, *49*, 3635−3637.
(24) 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*, 3546−3557.
(25) Limongelli, V. Ligand binding free energy and kinetics calculation in 2020. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2020**, *10*, No. e1455.
(26) Henriksen, N. M.; Fenley, A. T.; Gilson, M. K. Computational Calorimetry: High-Precision Calculation of Host-Guest Binding Thermodynamics. *J. Chem. Theory Comput.* **2015**, *11*, 4377−4394.
(27) Velez-Vega, C.; Gilson, M. K. Overcoming Dissipation in the Calculation of Standard Binding Free Energies by Ligand Extraction. *J. Comput. Chem.* **2013**, *34*, 2360−2371.
(28) Fenley, A. T.; Henriksen, N. M.; Muddana, H. S.; Gilson, M. K. Bridging Calorimetry and Simulation through Precise Calculations of Cucurbituril-Guest Binding Enthalpies. *J. Chem. Theory Comput.* **2014**, *10*, 4069−4078.
(29) Wang, J.; Deng, Y.; Roux, B. Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. *Biophys. J.* **2006**, *91*, 2798−2814.
(30) 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*, 4351−4362.
(31) Hu, L.; Ryde, U. Comparison of Methods to Obtain Force-Field Parameters for Metal Sites. *J. Chem. Theory Comput.* **2011**, *7*, 2452−2463.
(32) Vanduyfhuys, L.; Vandenbrande, S.; Verstraelen, T.; Schmid, R.; Waroquier, M.; Speybroeck, V. 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*, 1015−1027.
(33) Fracchia, F.; Del Frate, G.; Mancini, G.; Rocchia, W.; Barone, V. Force Field Parametrization of Metal Ions from Statistical Learning Techniques. *J. Chem. Theory Comput.* **2018**, *14*, 255−273.
(34) García-Simón, C.; Colomban, C.; Çetin, Y. A.; Gimeno, A.; Pujals, M.; Ubasart, E.; Fuertes-Espinosa, C.; Asad, K.; Chronakis, N.; Costas, M.; Jiménez-Barbero, J.; Feixas, F.; Ribas, X. Complete Dynamic Reconstruction of C_{60}, C_{70}, and $(C_{59}N)_2$ Encapsulation into an Adaptable Supramolecular Nanocapsule. *J. Am. Chem. Soc.* **2020**, *142*, 16051−16063.
(35) 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*, 9792−9802.
(36) Li, P.; Merz, K. M. MCPB.py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, *56*, 599−604.
(37) Seminario, J. M. Calculation of Intramolecular Force Fields from Second-Derivative Tensors. *Int. J. Quantum Chem.* **1996**, *60*, 1271−1277.
(38) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(39) 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*, 1157−1174.
(40) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(41) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(42) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(43) Bayly, C. I.; Cieplak, P.; Cornell, W.; 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*, 10269−10280.
(44) 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*, 247−260.

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381

(45) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(46) 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 Mater. Phys.* **1988**, *37*, 785−789.
(47) 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*, 154104−154119.
(48) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford CT, 2013.
(49) Dunning, T. H.; Hay, P. J. *Modern Theoretical Chemistry*; Schaefer, H. F., III, Eds.; Plenum: New York, 1977; Vol. 3, pp 1−28.
(50) Höllwarth, A.; Böhme, M.; Dapprich, S.; Ehlers, A. W.; Gobbi, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. A Set of d-Polarization Functions for Pseudo-Potential Basis Sets of the Main Group Elements Al-Bi and f-Type Polarization Functions for Zn, Cd, Hg. *Chem. Phys. Lett.* **1993**, *208*, 237−240.
(51) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. 6-31G* Basis Set for Third-Row Atoms. *J. Comput. Chem.* **2001**, *22*, 976−984.
(52) Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chem.—Eur. J.* **2012**, *18*, 9955−9964.
(53) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A., III Calculation of Solvation Free Energies of Charged Solutes Using Mixed Cluster/Continuum Models. *J. Phys. Chem. B* **2008**, *112*, 9709−9719.
(54) Kozlikova, B.; Sebestova, E.; Sustr, V.; Brezovsky, J.; Strnad, O.; Daniel, L.; Bednar, D.; Pavelka, A.; Manak, M.; Bezdeka, M.; Benes, P.; Kotry, M.; Gora, A.; Damborsky, J.; Sochor, J. CAVER Analyst 1.0: Graphic Tool for Interactive Visualization and Analysis of Tunnels and Channels in Protein Structures. *Bioinformatics* **2014**, *30*, 2684−2685.
(55) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera?A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(56) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman, P. A. *AMBER 2016*; University of California: San Francisco, 2016.
(57) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954−9960.
(58) Faller, R.; de Pablo, J. J. Constant pressure hybrid Molecular Dynamics−Monte Carlo simulations. *J. Chem. Phys.* **2002**, *116*, 55−59.
(59) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(60) Yin, J.; Henriksen, N. M.; Slochower, D. R.; Gilson, M. K. The SAMPL5 Host−Guest Challenge: Computing Binding Free Energies and Enthalpies from Explicit Solvent Simulations by the Attach-Pull-Release (APR) Method. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 133−145.
(61) Lee, T.-S.; Allen, B. K.; Giese, T. J.; Guo, Z.; Li, P.; Lin, C.; McGee, T. D.; Pearlman, D. A.; Radak, B. K.; Tao, Y.; Tsai, H.-C.; Xu, H.; Sherman, W.; York, D. M. Alchemical Binding Free Energy Calculations in AMBER20: Advances and Best Practices for Drug Discovery. *J. Chem. Inf. Model.* **2020**, *60*, 5595−5623.
(62) Chodera, J. D. A Simple Method for Automated Equilibration Detection in Molecular Simulations. *J. Chem. Theory Comput.* **2016**, *12*, 1799−1805.
(63) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.
(64) Grossfield, A. *"WHAM: The Weighted Histogram Analysis Method"*, version, 2.0.9. http://membrane.urmc.rochester.edu/wordpress/?page_id=126.
(65) Sgarlata, C.; Mugridge, J. S.; Pluth, M. D.; Zito, V.; Arena, G.; Raymond, K. N. Different and Often Opposing Forces Drive the Encapsulation and Multiple Exterior Binding of Charged Guests to a M4L6 Supramolecular Vessel in Water. *Chem.—Eur. J.* **2017**, *23*, 16813−16818.
(66) Daver, H.; Algarra, A. G.; Rebek, J.; Harvey, J. N.; Himo, F. Mixed Explicit-Implicit Solvation Approach for Modeling of Alkane Complexation in Water-Soluble Self-Assembled Capsules. *J. Am. Chem. Soc.* **2018**, *140*, 12527−12537.
(67) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. *Biophys. J.* **1997**, *72*, 1047−1069.
(68) Procacci, P.; Chelli, R. Statistical Mechanics of Ligand-Receptor Noncovalent Association, Revisited: Binding Site and Standard State Volumes in Modern Alchemical Theories. *J. Chem. Theory Comput.* **2017**, *13*, 1924−1933.
(69) Giovannelli, E.; Cioni, M.; Procacci, P.; Cardini, G.; Pagliai, M.; Volkov, V.; Chelli, R. Binding Free Energies of Host−Guest Systems by Nonequilibrium Alchemical Simulations with Constrained Dynamics: Illustrative Calculations and Numerical Validation. *J. Chem. Theory Comput.* **2017**, *13*, 5887−5899.
(70) Giovannelli, E.; Procacci, P.; Cardini, G.; Pagliai, M.; Volkov, V.; Chelli, R. Binding Free Energies of Host-Guest Systems by Nonequilibrium Alchemical Simulations with Constrained Dynamics: Theoretical Framework. *J. Chem. Theory Comput.* **2017**, *13*, 5874−5886.
(71) Biros, S. M.; Bergman, R. G.; Raymond, K. N. The Hydrophobic Effect Drives the Recognition of Hydrocarbons by an Anionic Metal-Ligand Cluster. *J. Am. Chem. Soc.* **2007**, *129*, 12094−12095.
(72) Pluth, M. D.; Raymond, K. N. Reversible Guest Exchange Mechanisms in Supramolecular Host-Guest Assemblies. *Chem. Soc. Rev.* **2007**, *36*, 161−171.
(73) 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.
(74) Mecozzi, S.; Rebek, J. The 55% Solution: A Formula for Molecular Recognition in the Liquid State. *Chem.—Eur. J.* **1998**, *4*, 1016−1022.
(75) Sgarlata, C.; Raymond, K. N. Untangling the Diverse Interior and Multiple Exterior Guest Interactions of a Supramolecular Host by the Simultaneous Analysis of Complementary Observables. *Anal. Chem.* **2016**, *88*, 6923−6929.

https://doi.org/10.1021/acs.jcim.1c00348
J. Chem. Inf. Model. 2021, 61, 4370−4381