How a Hemicarcerand Incarcerates Guests at Room Temperature
Decoded with Modular Simulations

Katherine G. McFerrin\textsuperscript{1,2} and Yuan-Ping Pang\textsuperscript{1\textdegree}

\textsuperscript{1}Computer-Aided Molecular Design Laboratory, Mayo Clinic, Rochester, MN, USA

\textsuperscript{2}Undergraduate intern from Carleton College, Northfield, MN, USA

\textsuperscript{\textdegree}E-mail: pang@mayo.edu
Abstract

Hemicarcerands are host molecules created to study constrictive binding with guest molecules for insights into the rules of molecular complexation. However, the molecular dynamics simulations that facilitate such studies have been limited because three-dimensional models of hemicarcerands are tedious to build and their atomic charges are complicated to derive. There have been no molecular dynamics simulations of the reported water-soluble hemicarcerand (Octacid$_4$) that explain how it uniquely encapsulates its guests at 298 K and keeps them encapsulated at 298 K in NMR experiments. Herein we report a modular approach to hemicarcerand simulations that simplifies the model building and charge derivation in a manner reminiscent of the approach to protein simulations with truncated amino acids as building blocks. We also report that apo Octacid$_4$ in water adopts two clusters of conformations, one of which has an equatorial portal open thus allowing guests to enter the cavity of Octacid$_4$, in microsecond molecular dynamics simulations performed using the modular approach at 298 K. Under the same simulation conditions, the guest-bound Octacid$_4$ adopts one cluster of conformations with all equatorial portals closed thus keeping the guests incarcerated. These results explain the unique constrictive binding of Octacid$_4$ and suggest that the guest-induced host conformational change that impedes decomplexation is a previously unrecognized conformational characteristic that promotes strong molecular complexation.
Introduction

Hemicarcerands comprise two identical bowl-shaped resorcinarene fragments tethered with four linkers (Fig. 1). These molecules, containing four equatorial portals in the linker region and two axial portals in the resorcinarene region, are developed as hosts to encapsulate small-molecule guests in a typical way that the guests enter and exit the host cavity at high temperatures and remain in the cavity at low temperatures\textsuperscript{1,2}. The binding characteristics and solubility of hemicarcerands are governed mostly by the linker structures. Octaacid\textsubscript{4} (Octacid\textsubscript{4}; tethered with a 4,6-dimethylisophthalic acid fragment as shown in Fig. 1) is reportedly a unique hemicarcerand that is water soluble and can form host-guest complexes (hemicarceplexes) with small molecules in water without the need to raise temperature\textsuperscript{3}. Therefore, the Octacid\textsubscript{4}-like hemicarceplexes are useful model systems for the study of constrictive binding, which is a type of molecular complexation that is affected by the activation energy required for a guest to enter the host cavity through a size-restricting portal\textsuperscript{4,6}, to obtain insights into the rules of molecular complexation.

To date, only a few molecular dynamics (MD) simulations of hemicarcerands or hemicarceplexes have been reported for the facilitation of constructive binding studies\textsuperscript{7-10}. This scarcity is partly because hemicarcerands have more than 200 atoms, making their three-dimensional models tedious to build. The scarcity is also due to the complexity of hemicarcerands, all of which have multiple conformations\textsuperscript{8,11}. As shown in Fig. 2, variations of the torsions along the four linkers of Octacid\textsubscript{4} can result in many distinct conformations, which complicates the derivation of the conformation-dependent atomic charges of the hemicarcerand. These technical complexities, as detailed below, may explain why there have hitherto been no aqueous MD simulations of Octacid\textsubscript{4} to understand how it uniquely encapsulates various small-molecule guests at 298 K and keeps them encapsulated at the same temperature such that the bound guests can be differentiated in their NMR spectra from those in the bulk phase\textsuperscript{3}.

Herein we report a modular method that simplifies the model building and charge derivation for MD simulations of Octacid\textsubscript{4} and its hemicarceplexes. We also report the characterization of the host
cavity and motions of the encapsulated guests that explains the unique constrictive binding of Octacid4 and offer new mechanistic insight into molecular complexation.

Results

The modular approach. In addition to the general technical complexities in hemicarcerand simulations noted above, the charge derivation for water-soluble hemicarcerands such as Octacid4 is particularly complicated. Because of the need to balance the atomic charges of the water-soluble hemicarcerand with the charges of the aqueous solvent and the charges of the small-molecule guest when using an AMBER forcefield such as FF12MClm12, the hemicarcerand charges need to be derived from ab initio calculations using the HF/6-31G* basis set that uniformly overestimates the polarity of the molecule. This is because (1) aqueous solvent models (such as the widely used TIP3P empirical water model) include polarization effects due to the empirical calibration to reproduce the density and enthalpy of liquid vaporization13, and (2) the small-molecule guest bears the restrained electrostatic potential (RESP) charges that are derived from ab initio calculations using the HF/6-31G* basis set4446. Further, to obtain the atomic charges of a water-soluble hemicarcerand without any bias toward one particular conformation of the molecule, the hemicarcerand charges need to be obtained from ab initio calculations at the HF/6-31G*//HF/6-31G* level with multiple conformations of the molecule followed by using the Lagrange multiplier to force identical charges on equivalent atoms in these conformations46. These ab initio calculations and the Lagrange multiplier are computation-demanding and labor-intensive, respectively. For example, the ab initio calculation of one Octacid4 conformation at the HF/6-31G*//HF/6-31G* level took ~168 CPU hours using the computers at the University of Illinois Urbana-Champaign National Center for Supercomputing Applications.

During the course of our Octacid4 simulations, we recognized that the technical complexities described above differ little from those of MD simulations of proteins with large numbers of atoms
and conformations. One known approach to circumventing the complexities of protein MD simulations is to perform the simulations using the modular concept underlying the second-generation AMBER forcefield\textsuperscript{7}. This forcefield builds a protein model with truncated amino acids as building blocks and derives the atomic charges of each building block from its multiple conformations using the Lagrange multiplier to force identical charges on equivalent atoms in these conformations\textsuperscript{6}. Given the effectiveness of the modular concept as demonstrated by the reported autonomous miniprotein folding MD simulations that achieved agreements between simulated and experimental folding times within factors of 0.6--1.4\textsuperscript{8}, we used to the modular concept to simplify the model building and charge derivation for hemicarcerands and devised a modular method for hemicarcerand simulations. We describe below two key attributes of our method for performing the MD simulations of Octacid\textsubscript{4}.

**The building block of Octacid\textsubscript{4}**. Because Octacid\textsubscript{4}, similar to most hemicarcerands, has C\textsubscript{4} symmetry, we divided it into four identical building blocks (termed HC\textsubscript{1}; Fig. 3). In this building block, C\textsubscript{1} and C\textsubscript{20} are designated as the head and tail atoms of the residue, respectively, similar to the designation of head and tail atoms in a truncated amino acid residue for protein simulations. Topology-wise, we built Octacid\textsubscript{4} as shown in Fig. 3 with a sequence of HC\textsubscript{1}-HC\textsubscript{1}-HC\textsubscript{1}-HC\textsubscript{1}.

First, we assembled the four residues into a linear molecule. This assembly followed the same scheme as that for assembling protein residues—namely, constructing a covalent bond between the tail atom of the preceding residue and the head atom of the following residue. Second, we converted the linear molecule to a cyclic molecule by specifying a cross-link (viz., constructing a covalent bond) between the head atom of the first HC\textsubscript{1} and the tail atom of the last HC\textsubscript{1}. Third, we converted the cyclic molecule to Octacid\textsubscript{4} by specifying (i) four cross-links between C\textsubscript{4} of the preceding HC\textsubscript{1} and C\textsubscript{18\textsubscript{a}} of the following HC\textsubscript{1}, (ii) four cross-links between O\textsubscript{5} of the preceding HC\textsubscript{1} and C\textsubscript{17\textsubscript{a}} of the following HC\textsubscript{1}, and (iii) four cross-links between O\textsubscript{16\textsubscript{a}} of the preceding HC\textsubscript{1} and C\textsubscript{6\textsubscript{b}} of the following HC\textsubscript{1}. 

- 5 -
The Octacid$_4$ sequence and all cross-links can be specified all together, thus making the hemicarcerand model building conceptually simple, algorithmic, and generalizable. Notably the HC$_1$ building block can, in the same manner, be further divided into three sub-building blocks for hemicarcerands without C$_4$ symmetry.

**Atomic charges of HC$_1$.** Given the HC$_1$ residue defined above, we obtained the atomic charges of Octacid$_4$ via the RESP charge derivation for HC$_1$ with the following specific conditions, in addition to the general conditions that are the same as those for protein RESP charges (such as the two-stage fitting for the methyl and methylene groups and forcing identical charges on equivalent intramolecular atoms)\(^{16,17}\).

First, we converted HC$_1$ to HCD$_1$ by attaching blocking groups to the junction atoms of HC$_1$ (Fig. 3) as these groups are needed to mimic the polar and aromatic groups abutting the junction atoms in Octacid$_4$. We then performed the ab initio calculation of HCD$_1$ to derive the HCD$_1$ charges by using the Lagrange multiplier to force (1) HCD$_1$ to have a net charge of $-2$ and (2) the total charge of all blocking groups to be zero, similar to the protein charge derivation from ab initio calculations of the acetyl- and N-methyl-blocked amino acids whose blocking group are used to mimic adjacent residues of the central amino acid\(^6\). Second, to balance the hemicarcerand charges with those of the solvent and the guest, we obtained the HCD$_1$ RESP charges from ab initio calculations of HCD$_1$ at the HF/6-31G*//HF/6-31G* level. Third, to avoid any bias toward one particular conformation of Octacid$_4$, we obtained the HCD$_1$ RESP charges from ab initio calculations using (1) two HCD$_1$ conformations taken from the representative conformations of Octacid$_4$ in vacuo with all equatorial portals open (abbreviated as the open conformation; Fig. 2A) and all equatorial portals closed (abbreviated as the closed conformation; Fig. 2B) and (2) the Lagrange multiplier to force identical charges on equivalent atoms in the two conformations akin to the charge derivation method for proteins that uses the alpha helical and beta strand side-chain conformations\(^6\). Last, we extracted the HC$_1$ charges from the HCD$_1$ charges.
In contrast to the computation-demanding and labor-intensive charge derivation for the intact Octacid\textsubscript{4} molecule, the HCl charge derivation can be readily performed with HCD\textsubscript{1}. For example, the \textit{ab initio} calculation of each of the two HCD\textsubscript{1} conformations at the HF/6-31G*//HF/6-31G* level took 31–40 CPU hours, considerably less than the \textasciitilde168 CPU hours for Octacid\textsubscript{4} noted above.

**Characterization of the Octacid\textsubscript{4} cavity.** Using the modular method described above, we performed 16 sets of 20 316-ns distinct, independent, unrestricted, unbiased, and classical isobaric–isothermal MD simulations of Octacid\textsubscript{4} and its hemicarceplexes with seven small-molecule guests at 298 K and 340 K. These simulations were performed with statistical relevance to investigate the change in the cavity volume of \textit{apo} and guest-bound Octacid\textsubscript{4} in water. A set of 20 simulations for each guest had an aggregated simulation time of 6.32 \textmu s. The seven guests were dimethyl sulfoxide (DMSO), ethyl acetate (EtOAc), dimethyl acetamide (DMA), 1,4-dioxane, diethylammonium (DEA), \textit{p}-xylene, and naphthalene. These molecules were selected according to their calculated maximal dimensions and molar volumes (Table 1) from the 14 reported guests that formed hemicarceplexes with Octacid\textsubscript{4}.

As apparent in Fig. 4, the Octacid\textsubscript{4} cavity is confined by a set of atoms shown with the sphere model. Therefore, we used the variation in the radius of gyration (Rg) of these atoms to estimate the change in cavity volume. Overall, the standard errors for the average Rgs of the cavity for the 16 sets of MD simulations were either 0.01 \textAA or 0.02 \textAA (Table 1), demonstrating the convergence for each simulation of the 16 sets and the statistical rigor of these simulations.

For the simulations of the \textit{apo} Octacid\textsubscript{4} in water at 298 K, we found that the Rg of the cavity remained at \textasciitilde5.4 \textAA and periodically spiked to \textasciitilde5.7 \textAA (Fig. 5A and Fig. S1), while the cavity Rgs of the open and closed conformations of Octacid\textsubscript{4} in vacuo were 5.86 \textAA and 5.37 \textAA, respectively (Table 1). The average and standard error of the cavity Rg for the 20 simulations at 298 K were 5.44 \textAA and 0.02 \textAA, respectively (Table 1). For the simulations at 340 K, the Rg also remained at \textasciitilde5.4 \textAA and spiked to \textasciitilde5.7 \textAA, but the frequency of the spikes was much higher at 340 K than that at 298 K (Fig. 5B and Fig. S1), and the average and standard error of the Rg at 340 K were 5.45 \textAA and 0.02 \textAA, respectively (Table 1).
These time series of Rg and their average values indicate that (1) *apo* Octacid\(_4\) in water adopted two clusters of conformations, one similar to the closed conformation of *apo* Octacid\(_4\) in vacuo with a small cavity and the other similar to the open conformation of *apo* Octacid\(_4\) in vacuo with a large cavity, (2) the large-cavity conformations were much less populated in water than the small-cavity conformations, (3) the population of the large-cavity conformations in water increased as temperature increased, and (4) *apo* Octacid\(_4\) in water underwent cavity expansion and contraction at both temperatures.

For the simulations of the DMSO, EtOAc, and DMA hemicarceplexes, we found a cavity Rg of \(\sim 5.4\ \text{Å}\) with a few or no spikes of \(>5.7\ \text{Å}\) at 298 K and 340 K (Fig. 5 and Fig. S1); the respective average and standard error of the cavity Rg of each set of the 20 MD simulations at 298 K were 5.41 Å and 0.02 Å for DMSO, 5.40 Å and 0.01 Å for EtOAc, and 5.40 Å and 0.01 Å for DMA (Table 1). The corresponding values at 340 K were 5.41 Å and 0.02 Å for DMSO, 5.41 Å and 0.01 Å for EtOAc, and 5.41 Å and 0.01 Å for DMA (Table 1). For the simulations of the 1,4-dioxane hemicarceplex, we found a cavity Rg of \(\sim 5.35\ \text{Å}\) with no spikes at either temperature (Fig. 5 and Fig. S1); the respective average and standard error of the cavity Rg for the simulations at 298 K were 5.37 Å and 0.01 Å, and the corresponding values at 340 K were 5.38 Å and 0.01 Å (Table 1). These time series of Rg and their average values indicate that (1) the hemicarceplexes involving relatively compact guests in water all adopted one cluster of conformations that is similar to the closed conformation of *apo* Octacid\(_4\) in vacuo except for DMSO•Octacid\(_4\) (which adopted two clusters conformations that resemble the open and closed conformations of *apo* Octacid\(_4\) in vacuo, but the population of the open conformation of DMSO•Octacid\(_4\) was much lower than that of *apo* Octacid\(_4\) in water), and (2) these hemicarceplexes mostly kept their cavities contracted with all equatorial portals closed.

For the simulations of the DEA, \(p\)-xylene, and naphthalene hemicarceplexes, we found a cavity Rg of \(\sim 5.6\ \text{Å}\) with a few or no spikes of \(>5.7\ \text{Å}\) at 298 K and 340 K (Fig. 5 and Fig. S1); the respective average and standard error of the cavity Rg of the simulations at 298 K were 5.64 Å and 0.02 Å for DEA, 5.56 Å and 0.02 Å for \(p\)-xylene, and 5.65 Å and 0.02 Å for naphthalene. The corresponding
values at 340 K were 5.65 Å and 0.02 Å for DEA, 5.59 Å and 0.02 Å for p-xylene, and 5.66 Å and 0.02 Å for naphthalene (Table 1). These time series of Rg and their average values indicate that (1) the hemicarceplexes involving relatively bulky guests in water all adopted one cluster of conformations with their cavities larger than that of the closed conformation of apo Octacid4 in vacuo and smaller than that of the open conformation of apo Octacid4 in vacuo, and (2) these hemicarceplexes mostly kept their cavities contracted with all equatorial portals closed.

**Characterization of guest motion in Octacid4.** We also performed conformational cluster analyses on the simulations at 340 K for apo Octacid4 and the seven hemicarceplexes. All hemicarceplex simulations were performed using an initial conformation in which the host adopted the open conformation of apo Octacid4 in vacuo (Fig. 2A) and the guest was manually docked into the host cavity with its maximal dimension perpendicular to the axial axis (viz., the axis passing the two axial portals). Interestingly, as detailed below, in the aqueous MD simulations, all seven guests adopted a new orientation with their maximal dimensions parallel to the axial axis (Fig. 6). Similar results were obtained (Fig. S2) when the analyses were performed using the simulations at 298 K.

For apo Octacid4 in water, the most populated conformation (occurrence: 97.2%) in the MD simulations was nearly identical to the closed conformation in vacuo (carbon root mean square deviation: 0.34 Å; Figs. 2A and 2C), whereas the second most populated conformation (occurrence: 1.9%) was somewhat different from the open conformation in vacuo (carbon root mean square deviation: 1.49 Å) in that the solution conformation had one equatorial portal open and the opposite portal closed (Figs. 2B and 2D).

For the DMSO hemicarceplex (Fig. 6A), the most populated conformation had the guest oxygen atom pointing to the equatorial portals and the two guest methyl groups pointing to the axial portals; in the average conformation of the largest conformation cluster, the guest was shrunk to a ball with the oxygen atom on one side and the two overlapping carbon atoms on the other, indicating that the guest had rotated around multiple axes.
For the EtOAc hemicarceplex (Fig. 6B), the most populated guest conformation (population: 94.7%) was not fully extended (with C1-C2-O1-C3 and C2-O1-C3-C4 being $-114^\circ$ and $81^\circ$, respectively), and the population of the fully extended guest conformation with corresponding torsions of $-180^\circ$ and $180^\circ$ was 0.1%; the most populated guest conformation had its two oxygen atoms pointing to the equatorial portals and its two methyl groups pointing to the axial portals; in the average conformation of the largest conformation cluster, the guest was shrunk to a short rod with the oxygen atoms in the middle and the two carbon atoms on both ends of the rod, indicating that the guest had rotated frequently around the axial axis and less frequently around the equatorial axis (the axis passing the two opposing equatorial portals).

For the DMA hemicarceplex (Fig. 6C), the most populated conformation had the guest oxygen atom pointing to the equatorial portals and the two trans methyl groups of the guest pointing to the axial portals; in the average conformation of the largest conformation cluster, the guest was shrunk to a short rod with (1) the oxygen atom on one side of the middle region, (2) the carbon atom trans to the oxygen atom on the other side of the middle region, and (3) the two trans carbon atoms on both ends of the rod, indicating that the guest had rotated frequently around the axial axis and less frequently around the equatorial axis.

The 1,4-dioxane hemicarceplex (Fig. 6D) had one conformation cluster. The most populated guest conformation adopted the energetically stable chair conformation with the two guest oxygen atoms pointing to the equatorial portals and the two methylene groups on one side of the guest pointing to an axial portal while those on the other side pointing to the opposite axial portal. In the average conformation of the cluster, the guest was shrunk to a ball with two oxygen atoms on opposite sides, indicating that the guest had rotated around multiple axes.

The DEA, $p$-xylene, and naphthalene hemicarceplexes (Figs. 6E–6G) had one conformation cluster each. The conformations of these clusters had (1) the DEA nitrogen atom pointing to the equatorial portals and the two DEA methyl groups pointing to the axial portals, (2) the two $p$-xylene methyl groups pointing to the axial portals, and (3) the naphthalene $\beta$-carbon atoms pointing to the
axial portals. In the average conformation of each cluster, the guest was shrunk to a long rod whose length was the same as that of the guest, indicating that these guests all had rotated exclusively around the axial axis.

**Discussion**

**Support for the modular approach.** Our characterization studies reveal that (1) apo Octacid4 adopted two clusters of conformations in aqueous MD simulations at 298 K, (2) one cluster had the conformations (population: 97.2%) nearly identical to the one in vacuo with all equatorial portals closed, and (3) the other had the conformations (population: 1.9%) somewhat different from the open conformation of apo Octacid4 in vacuo in that the solution conformation has one equatorial portal open and the opposite portal closed. The population ratio (97.2/1.9) of the two clusters according to the conformational cluster analysis is consistent with the frequency for the conformational exchanges depicted in Fig. 5, which was derived from the Rg-based analysis. In addition, all solution conformations derived from the MD simulations have one resorcinarene fragment rotated by \( \sim 22^\circ \) with respect to the other, which is consistent with reported crystal structures of different hemicarceplexes (involving the same or different guests) all of which have one resorcinarene fragment rotated by \( 13-21^\circ \) with respect to the other\(^{14,19,20}\). Moreover, the less populated aqueous Octacid4 conformations with one equatorial portal open and the opposite portal closed are also consistent with the reported V-shaped hemicarcerand conformation proposed in the sliding-door mechanism for hemicarceplex formation\(^{7,21}\). These internal and external consistencies indicate the plausibility of the aqueous conformations of Octacid4 obtained a priori from our modular-approach–based MD simulations and lend credence to our modular approach to hemicarcerand simulations for probing solution conformations relevant to the following mechanistic insights.
**How Octacid₄ incarcerates its guests.** Our characterization studies also show that *apo* Octacid₄ in water mostly adopted a cluster of conformations with small cavities whose equatorial portals are all closed and that, even at 298 K, this host also periodically and transiently adopted another cluster of conformations that have large cavities with at least one equatorial portal open. These periodic and transient conformations with an open equatorial portal explain how Octacid₄ reportedly encapsulated guests in only a few minutes at 298 K in the NMR experiments, but do not explain how Octacid₄ reportedly kept its guests encapsulated at the same temperature. Perhaps the guests were incarcerated by strong intrinsic binding (viz., the complexation governed by strong nonbonded intermolecular interactions). However, the involvement of strong intrinsic binding is debatable given the weak interaction of the cavity with DMSO or 1,4-dioxane as indicated by their free spins as described above (Table 1 and Figs. 6A and 6D). An alternative explanation is therefore needed.

Unexpectedly, our characterization studies show that upon complexation, Octacid₄ adopted only one cluster of conformations. These conformations have, depending on the size of the guest, a contracted or expanded cavity relative to that of *apo* Octacid₄ in water. The conformations with contracted cavities have the equatorial portals all closed because the guests are relatively compact and subsequently the weak nonbonded intermolecular interactions of the host cavity with a small guest shift the host conformations with an open equatorial portal to the conformations with all equatorial portals closed. The conformations with slightly expanded cavities also have the equatorial portals all closed because the guests are relatively bulky and these guests jam the cavity and consequently prevent the host from adopting the V-shaped sliding-door conformations shown in Figs. 2D and 2E to open any of the equatorial portals. Therefore, regardless of the size of the guest, the guest-bound Octacid₄ in water adopts one cluster of conformations with closed equatorial portals. These conformations satisfactorily explain how Octacid₄ can incarcerate a range of guests at 298 K, which is the temperature at which the guests enter the cavity.
Implication. Collectively, our studies suggest that Octacid$_4$ can open one of its equatorial portals without elevating temperature before complexation and close all of its equatorial portals without lowering temperature after complexation. These interesting and unexpected capabilities enable the unique constrictive binding of Octacid$_4$ with a range of small-molecule guests and suggest further that the guest-induced host conformational change that impedes decomplexation is a previously unrecognized conformational characteristic that is conducive to strong molecular complexation. This characteristic could broaden the theoretical dissection of the experimentally observed complexation affinity and the design of new complex systems for materials technology, data storage and processing, catalysis, drug design and delivery, and medicine by accounting for not only intrinsic binding (which is limited because, confined by the cost of chemical synthesis, only a small number of functional groups can be introduced to improve nonbonded intermolecular interactions) but also constrictive binding (which requires one or a few functional groups to trigger the formation of a host conformation that hinders decomplexation as demonstrated by the Octacid$_4$ hemicarceplexes).

Methods

Molecular dynamics simulations. Hemicarcerand Octacid$_4$ or its hemicarceplex neutralized with sodium ions was solvated with the TIP3P water\textsuperscript{13} (“solvatebox molecule TIP3BOX 8.2”) using tLEaP of the AmberTools 16 package (University of California, San Francisco) and then energy-minimized for 100 cycles of steepest-descent minimization followed by 900 cycles of conjugate-gradient minimization to remove close van der Waals contacts using SANDER of the AMBER 11 package (University of California, San Francisco), forcefield FF12MCM\textsuperscript{12}, and a cutoff of 8.0 Å for nonbonded interactions. The resulting system was heated from 5 K to 298 or 340 K at a rate of 10 K/ps under constant temperature and constant volume, and then equilibrated for $10^6$ timesteps under constant temperature of 298 K or 340 K and constant pressure of 1 atm employing isotropic molecule-based scaling. Finally, a set of 20 distinct, independent, unrestricted, unbiased, isobaric–isothermal, and 316-
ns MD simulations of the equilibrated system was performed for Octacid$_4$ or its hemicarceplex using PMEMD of the AMBER 16 package (University of California, San Francisco), forcefield FF12MClm$^{12}$, and a periodic boundary condition at 298 K or 340 K and 1 atm. The 20 unique seed numbers for initial velocities of the 20 simulations were taken from Ref. $^{22}$. All simulations used (i) a dielectric constant of 1.0, (ii) the Berendsen coupling algorithm$^{23}$, (iii) the particle mesh Ewald method to calculate electrostatic interactions of two atoms at a separation of $>8$ Å$^{24}$, (iv) $\Delta t = 1.00$ fs of the standard-mass time$^{22,25}$, (v) the SHAKE-bond-length constraint applied to all bonds involving hydrogen, (vi) a protocol to save the image closest to the middle of the “primary box” to the restart and trajectory files, (vii) a formatted restart file, (viii) the revised alkali ion parameters$^{26}$, (ix) a cutoff of 8.0 Å for nonbonded interactions, (x) a uniform 10-fold reduction in the atomic masses of the entire simulation system (both solute and solvent)$^{22,25}$, and (xi) default values of all other inputs of PMEMD.

The *apo* Octacid$_4$ conformation with all equatorial portals open (Fig. 2B) was used as the initial conformation for the *apo* Octacid$_4$ MD simulations. This conformation was obtained from energy minimization for 200 cycles of steepest-descent minimization followed by 2800 cycles of conjugate-gradient minimization using SANDER, forcefield FF12MClm$^{12}$, and a cutoff of 30.0 Å for nonbonded interactions. For each hemicarceplex, the guest was manually docked into the cavity of the *apo* Octacid$_4$ conformation with all equatorial portals open (Fig. 2B) in such a way that the maximal dimension of the guest was perpendicular to the axial axis (viz., the axis passing the two axial portals); the resulting hemicarceplex was then energy minimized using the same minimization protocol for obtaining the *apo* Octacid$_4$ conformation; the energy-minimized hemicarceplex was finally used as the initial conformation for the hemicarceplex MD simulations.

The forcefield parameters for the Octacid$_4$ building block and for all small-molecule guests except dimethyl sulfoxide were obtained from ab initio calculations of the molecules at the HF/6-31G$^*$/HF/6-31G$^*$ level using (i) a published procedure for the charge derivation$^{16}$ and (2) the arithmetic average from multiple conformations for the bond, angle, and torsion parameters. The
forcefield parameters for dimethyl sulfoxide were obtained from Ref. 27. The forcefield parameters of FF12MClm are available in the Supporting Information of Ref. 25.

All simulations were performed using computers at the University of Minnesota Supercomputing Institute and the Mayo Clinic high performance computing facility at the University of Illinois Urbana-Champaign National Center for Supercomputing Applications.

**ab Initio calculations.** All ab initio calculations were performed using Gaussian 98 (Revision A.7; Gaussian, Inc. Wallingford, CT), except those for the intact Octacid4 at the HF/6-31G°//HF/6-31G° level were done using Gaussian 16 (Revision C.01; Gaussian, Inc. Wallingford, CT).

**Conformational cluster analysis.** The conformational cluster analyses were performed using CPPTRAJ of the AmberTools 16 package (University of California, San Francisco) with the average-linkage algorithm, epsilon of 1.0 Å, and root mean square coordinate deviation on all carbon atoms of Octacid4 or its hemicarceplex. Centering the coordinates of the hemicarceplex and imaging the coordinates to the primary unit cell were performed prior to the cluster analyses.

**Root mean square deviation and radius of gyration.** Carbon root mean square deviations were calculated manually using ProFit V2.6 (http://www.bioinf.org.uk/software/profit/). Radius of gyration was calculated using CPPTRAJ.

**Data availability**

All data generated for this study are included in this Article and its Supplementary Information.

**References**
1. Tanner, M. E., Knobler, C. B. & Cram, D. J. Hemicarcerands permit entrance to and egress from their inside phases with high structural recognition and activation free-energies. *J. Am. Chem. Soc.* **112**, 1659–1660 (1990).

2. Warmuth, R. & Yoon, J. Recent highlights in hemicarcerand chemistry. *Acc. Chem. Res.* **34**, 95–105 (2001).

3. Yoon, J. Y. & Cram, D. J. The first water-soluble hemicarceplexes. *Chem. Commun.*, 497–498 (1997).

4. Cram, D. J., Tanner, M. E. & Knobler, C. B. Host-guest complexation. Guest release and capture by hemicarcerands introduces the phenomenon of constrictive binding. *J. Am. Chem. Soc.* **113**, 7717–7727 (1991).

5. Quan, M. L. C., Knobler, C. B. & Cram, D. J. Constrictive binding by an octalactone hemicarcerand. *J. Chem. Soc., Chem. Commun.*, 660–662 (1991).

6. Quan, M. L. C. & Cram, D. J. Constrictive binding of large guests by a hemicarcerand containing 4 portals. *J. Am. Chem. Soc.* **113**, 2754–2755 (1991).

7. Nakamura, K. & Houk, K. N. Mechanism and thermodynamics of guest escape from a carcerand. *J. Am. Chem. Soc.* **117**, 1853–1854 (1995).

8. Sheu, C. M. & Houk, K. N. Molecular mechanics and statistical thermodynamics studies of complexes of a flexible hemicarcerand with neutral guests. *J. Am. Chem. Soc.* **118**, 8056–8070 (1996).

9. Vazdar, M., Mitchell, A. S., Warrener, R. N. & Margetic, D. High pressure promoted exchange of guests from hemicarceplexes. *Tetrahedron* **71**, 550–553 (2015).

10. Aulin, Y. V., Liu, M. D. & Piotrowsiak, P. Ultrafast vibrational cooling inside of a molecular container. *J. Phys. Chem. Lett.* **10**, 2434–2438 (2019).

11. Liddell, M. J., Margetic, D., Mitchell, A. S. & Warrener, R. N. An AM1 semiempirical study of host-guest complexation in hemicarcerand complexes. *J. Comput. Chem.* **25**, 542–557 (2004).
12. Pang, Y.-P. FF12MC: a revised AMBER forcefield and new protein simulation protocol. *Proteins* **84**, 1490–1516 (2016).

13. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

14. Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges - the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).

15. Cornell, W. D., Cieplak, P., Bayly, C. I. & Kollman, P. A. Application of RESP charges to calculate conformational energies, hydrogen-bond energies, and free-energies of solvation. *J. Am. Chem. Soc.* **115**, 9620–9631 (1993).

16. Cieplak, P., Cornell, W. D., Bayly, C. & Kollman, P. A. Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem.* **16**, 1357–1377 (1995).

17. Cornell, W. D. *et al.* A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).

18. Pang, Y.-P. How fast fast-folding proteins fold in silico. *Biochem. Biophys. Res. Commun.* **492**, 135–139 (2017).

19. Cram, D. J., Blanda, M. T., Paek, K. & Knobler, C. B. Host-guest complexation.63. Constrictive and intrinsic binding in a hemicarcerand containing 4 portals. *J. Am. Chem. Soc.* **114**, 7765–7773 (1992).

20. Robbins, T. A., Knobler, C. B., Bellew, D. R. & Cram, D. J. Host-guest complexion.67. A Highly adaptive and strongly binding hemicarcerand. *J. Am. Chem. Soc.* **116**, 111–122 (1994).

21. Houk, K. N., Nakamura, K., Sheu, C. M. & Keating, A. E. Gating as a control element in constrictive binding and guest release by hemicarcerands. *Science* **273**, 627–629 (1996).

22. Pang, Y.-P. Use of 1–4 interaction scaling factors to control the conformational equilibrium between α-helix and β-strand. *Biochem. Biophys. Res. Commun.* **457**, 183–186 (2015).
23. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Di Nola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
24. Darden, T. A., York, D. M. & Pedersen, L. G. Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
25. Pang, Y.-P. Low-mass molecular dynamics simulation for configurational sampling enhancement: More evidence and theoretical explanation. *Biochem. Biophys. Rep.* **4**, 126–133 (2015).
26. Joung, I. S. & Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).
27. Fox, T. & Kollman, P. A. Application of the RESP methodology in the parametrization of organic solvents. *J. Phys. Chem. B* **102**, 8070–8079 (1998).
28. Shao, J., Tanner, S. W., Thompson, N. & Cheatham III, T. E. Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.* **3**, 2312–2334 (2007).

**Acknowledgements**

This work was supported by the US Army Research Office (W911NF-16-1-0264) and the Mayo Foundation for Medical Education and Research. Responsibility for the information and views in this study lies entirely with the authors. The authors acknowledge the computing resources provided by the University of Minnesota Supercomputing Institute and the Mayo Clinic high performance computing facility at the University of Illinois Urbana-Champaign National Center for Supercomputing Applications.

**Author contributions**
K.G.M. performed the literature search; prepared the z-matrices of the HCl and HCD$_1$ residues; performed energy minimization of HCl, HCD$_1$, and Octacid$_4$; analyzed the MD simulation result of $p$-xylene•Octacid$_4$; contributed to revisions of the manuscript. Y.-P.P. conceived the modular method; designed the HCl and HCD$_1$ residues and all protocols for MD simulation and analysis; performed the remaining computational study; analyzed the data; wrote the manuscript.

**Competing interests**

The authors declare no competing interests.

**Supplementary information**

Fig. S1. Time series of radius of gyration of the Octacid$_4$ cavity for all 20 distinct and independent MD simulations at 298 K or 340 K.

Fig. S2. The most populated conformations of seven Octacid$_4$ hemicarceplexes in the aqueous MD simulations at 298 K. The representative and average conformations in the largest conformation cluster of 20 MD simulations for each hemicarceplex are shown in the left and right panels, respectively. The sulfur, oxygen, nitrogen, and carbon atoms are in yellow, red, blue, and green, respectively. Counter ions and water molecules of aqueous Octacid$_4$ are not displayed for clarity. (A) DMSO•Octacid$_4$. (B) EtOAc•Octacid$_4$. (C) DMA•Octacid$_4$. (D) 1,4-Dioxane•Octacid$_4$. (E) DEA•Octacid$_4$. (F) $p$-Xylene•Octacid$_4$. (G) Naphthalene•Octacid$_4$.

**Figure legends**

Fig. 1. Chemical structures of a generic hemicarcerand and Octacid$_4$.

Fig. 2. Distinct conformations of Octacid$_4$. Distances shown by thin dashed lines to indicate the
open or closed equatorial portal are in Angstroms. Oxygen atoms are in red. Carbon atoms are in lime, green cyan, limon, cyan, marine, or yellow orange. Counter ions and water molecules of aqueous Octacid$_4$ are not displayed for clarity. (A) Energy-minimization–derived Octacid$_4$ in vacuo with all equatorial portals closed. (B) Energy-minimization–derived Octacid$_4$ in vacuo with all equatorial portals open. (C) The most populated Octacid$_4$ conformation with all equatorial portals closed in the aqueous MD simulations. (D) The second most populated Octacid$_4$ conformation with one equatorial portal open in the aqueous MD simulations. (E) An Octacid$_4$ conformation with one equatorial portal open in the aqueous MD simulations. (F) Another Octacid$_4$ conformation with equatorial portals open in the aqueous MD simulations.

Fig. 3. Chemical structures of HC$_1$ and HCD$_1$ and the assembly of HC$_1$ into Octacid$_4$. Thick lines indicate the bonds between main-chain atoms. Thin lines indicate the bonds between side-chain atoms or between a main-chain atom and a side-chain atom. Thin dashed lines and colored lines indicate the intra- and inter-residue bonds, respectively.

Fig. 4. Labels of the atoms in the HC$_1$ residue that confine the Octacid$_4$ cavity.

Fig. 5. Time series of radius of gyration of the Octacid$_4$ cavity for the first of 20 distinct and independent MD simulations. (A) Simulations at 298 K. (B) Simulations at 340 K.

Fig. 6. The most populated conformations of seven Octacid$_4$ hemicarceplexes in the aqueous MD simulations at 340 K. The representative and average conformations in the largest conformation cluster of 20 MD simulations for each hemicarceplex are shown in the left and right panels, respectively. The sulfur, oxygen, nitrogen, and carbon atoms are in yellow, red, blue, and green, respectively. Counter ions and water molecules of aqueous Octacid$_4$ are not displayed for clarity. (A) DMSO•Octacid$_4$. (B) EtOAc•Octacid$_4$. (C) DMA•Octacid$_4$. (D) 1,4-Dioxane•Octacid$_4$. (E) DEA•Octacid$_4$. (F) $p$-Xylene•Octacid$_4$. (G) Naphthalene•Octacid$_4$. 
TABLE 1. Static and dynamic properties of Octacid4 and its complexes.

| Structure                      | MDG (Å) | GMV (cm³/mol) | Temp (K) | Avg Rg (Å) Mean ± SE | SAGS         | MoGIO          |
|--------------------------------|---------|----------------|----------|-----------------------|--------------|-----------------|
| Open Apo Octacid4 (vacuo)     | —       | —              | —        | 5.86                  | —            | —               |
| Closed Apo Octacid4 (vacuo)   | —       | —              | —        | 5.37                  | —            | —               |
| Apo Octacid4 (water)          | —       | —              | 298      | 5.44 ± 0.02           | —            | —               |
|                               |         |                | 340      | 5.45 ± 0.02           | Ball         | Free spin       |
| DMSO-Octacid4 (water)         | 4.5     | 49             | 298      | 5.41 ± 0.02           | Ball         | Free spin       |
|                               |         |                | 340      | 5.41 ± 0.02           | Ball         | Free spin       |
| 1,4-Dioxane-Octacid4 (water)  | 4.8     | 74             | 298      | 5.37 ± 0.01           | Ball         | Free spin       |
|                               |         |                | 340      | 5.38 ± 0.01           | Ball         | Free spin       |
| DMA-Octacid4 (water)          | 5.3     | 73             | 298      | 5.40 ± 0.01           | Short rod    | Partially free spin |
|                               |         |                | 340      | 5.41 ± 0.01           | Short rod    | Partially free spin |
| EtOAc-Octacid4 (water)        | 6.5     | 71             | 298      | 5.40 ± 0.01           | Short rod    | Partially free spin |
|                               |         |                | 340      | 5.41 ± 0.01           | Short rod    | Partially free spin |
| DEA-Octacid4 (water)          | 6.8     | 72             | 298      | 5.64 ± 0.02           | Long rod     | Axial spin      |
|                               |         |                | 340      | 5.65 ± 0.02           | Long rod     | Axial spin      |
| p-Xylene-Octacid4 (water)     | 6.9     | 82             | 298      | 5.56 ± 0.02           | Long rod     | Axial spin      |
|                               |         |                | 340      | 5.59 ± 0.02           | Long rod     | Axial spin      |
| Naphthalene-Octacid4 (water)  | 7.1     | 115            | 298      | 5.65 ± 0.02           | Long rod     | Axial spin      |
|                               |         |                | 340      | 5.66 ± 0.02           | Long rod     | Axial spin      |

MDG: The maximal dimension of the guest calculated using Gaussian 98. GMV: Guest molar volume calculated using Gaussian 98. Temp: Temperature. Avg Rg: Average radius of gyration from 20 independent and distinct simulations. SAGS: The shape of the average guest structure. MoGIO: The motion of the guest incarcerated in Octacid4. Axial spin: rotation around the axial axis. Free spin: rotations evenly around multiple axes. Partially free spin: rotations around one axis more frequently than around the orthogonal axis.
Fig. 1

Octacid4: $X = \text{[structure]}$; $Y = \text{CH}_2$; $Z = \text{CH}_3$
Fig. 4
Fig. 5a

Octacid4 at 298 K

Dimethylsulfoxide@Octacid4 at 298 K

Ethylacetate@Octacid4 at 298 K

Dimethylacetamide@Octacid4 at 298 K

1,4-Dioxane@Octacid4 at 298 K

Diethylammonium@Octacid4 at 298 K

p-Xylene@Octacid4 at 298 K

Naphthalene@Octacid4 at 298 K
Fig. 5b

Octacid4 at 340 K

Dimethylsulfoxide@Octacid4 at 340 K

Ethylacetate@Octacid4 at 340 K

Dimethylacetamide@Octacid4 at 340 K

1,4-Dioxane@Octacid4 at 340 K

Diethylammonium@Octacid4 at 340 K

p-Xylene@Octacid4 at 340 K

Naphthalene@Octacid4 at 340 K
