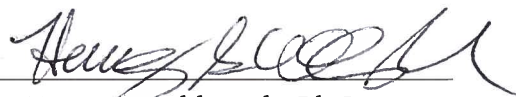SUPRAMOLECULAR ASSEMBLIES OF DEEP-CAVITY CAVITANDS STABILIZED BY THE
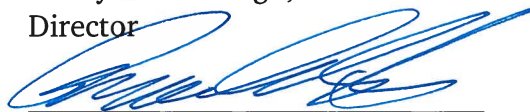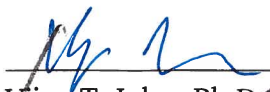
HYDROPHOBIC EFFECT

AN ABSTRACT

SUBMITTED ON THE TWENTY-THIRD DAY OF MARCH 2017
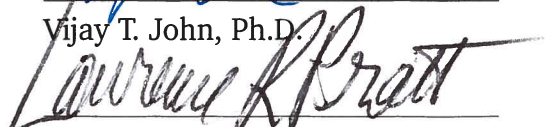
TO THE DEPARTMENT OF CHEMICAL AND BIOMOLECULAR ENGINEERING

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF SCIENCE AND ENGINEERING

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY

BY

_____
James W. Barnett

APPROVED: _____
Henry S. Ashbaugh, Ph.D.
Director
_____
Bruce C. Gibb, Ph.D.
_____
Vijay T. John, Ph.D.
_____
Lawrence R. Pratt, Ph.D.

ABSTRACT

Since the mid-20th century supramolecular chemistry has become a thriving field in synthetic chemistry. Supramolecular assemblies are assemblies of molecules formed and stabilized by non-covalent interactions. Deep-cavity cavitands, bowl-shaped molecules, are one such class of molecules that form assemblies using the hydrophobic effect when in the presence of suitable hydrophobic guests in an aqueous environment. Computer simulations allow us to study these assemblies at the molecular level and provide valuable insight into both the thermodynamics of assembly as well as provide information relevant to the design of the next generation of deep-cavity cavitands.

This research begins by investigating dimeric capsules of a deep-cavity cavitand known as Octa-acid (OA). We use Molecular Dynamics to study a homologous series of *n*-alkane guests in order to learn some of the "rules" of guest packing. Additionally we use a machine learning technique to harvest a dominant conformation from each simulation and compare computed chemical shifts of that structure with experimental chemical shifts.

The second part of this dissertation looks into multimeric systems formed by one of OA's derivatives known as Tetra-endomethyl Octa-acid (TEMOA). The entrance to the binding pocket of TEMOA is narrower than OA due to four methyls being added to its rim. TEMOA forms not only dimers, but also tetramers and hexamers, depending

on the guest size. We use free energy techniques to show that guest packing primarily drives the transitions between each assembly state. Additionally we obtain the interior volumes of each multimer and demonstrate that they now approach that of structures formed by other means. We give insight into why TEMOA forms multimeric systems and OA does not.

The last section of this dissertation compares the interior hydration characteristics of OA and TEMOA. We show that the small structural changes from OA to TEMOA promote a large change in wetting/dewetting behavior inside the binding pocket. Normally OA is full of water in its interior, but TEMOA exhibits a two-phase behavior. Here we also demonstrate a simple bridge between simulation and experiment to validate our findings by using partial molar volume calculations.

SUPRAMOLECULAR ASSEMBLIES OF DEEP-CAVITY CAVITANDS STABILIZED BY THE

HYDROPHOBIC EFFECT

A DISSERTATION

SUBMITTED ON THE TWENTY-THIRD DAY OF MARCH 2017

TO THE DEPARTMENT OF CHEMICAL AND BIOMOLECULAR ENGINEERING

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF SCIENCE AND ENGINEERING

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY
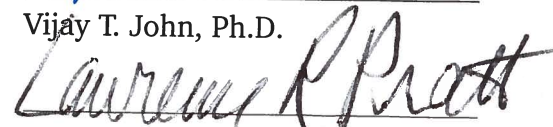
BY

_____
James W. Barnett

APPROVED: _____
Henry S. Ashbaugh, Ph.D.
Director

_____
Bruce C. Gibb, Ph.D.

_____
Vijay T. John, Ph.D.

_____
Lawrence R. Pratt, Ph.D.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

For over a century the covalent bond dominated the field of synthetic chemistry beginning when, in 1826, Friedrich Wöhler[1] synthesized urea. Since that time chemists and engineers have sought new reactions to covalently link and unlink reactants together into products.[2] Nature, however, is well-known to routinely utilize not only covalent interactions but also non-covalent interactions, with Hermann Emil Fischer as early as 1894 suggesting that enzyme-substrate interactions take the form a of a "lock" and "'key".[3]

It was not until the mid-20th century that chemists were able to utilize such non-bonded interactions to form larger assemblies. At that time Pederson, Cram, and Lehn[1] formed cyclic molecules that assembled with an ion in their centers (Figure 1.1). These cyclical molecules with ions non-covalently bound to their interiors gave birth to what is now known as "supramolecular chemistry", and in 1987 Pedersen, Cram, and Lehn were awarded the Nobel prize because they had "elucidated the factors that determine the ability of the molecules to recognize each other and fit into one another like a key fits a lock".[4]
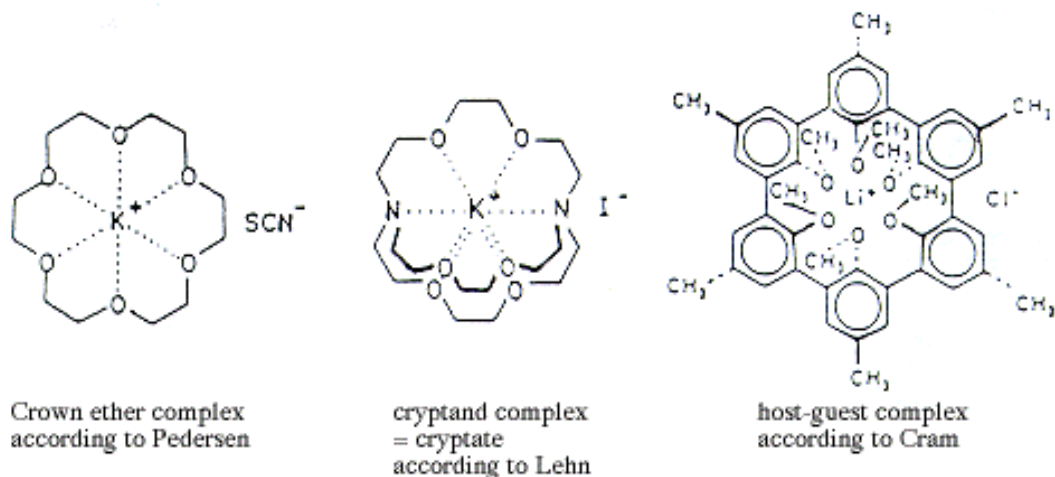
Figure 1.1: Guest-host complexes created by Pedersen, Lehn, and Cram, shown to selectively bind with a positive metal ion.[4]

This field of supramolecular chemistry uses weaker, non-covalent forces to form larger assemblies, and such assemblies are often called "supramolecules". In addition to the driving force of metal coordination used by Pedersen, several other driving forces can be involved in the formation of supramolecular assemblies. These include, but are not limited to, hydrogen bonding, π-π stacking, and the hydrophobic effect.[5,6] Such assembly processes are often called "self-assembly", and such assemblies are sometimes known as "guest-host" assemblies when a larger molecule denoted as the "host" encloses or surrounds a smaller molecule called the "guest" (cf. Cram's assembly in Figure 1.1).

Eventually Cram would popularize the term "cavitand" to describe a specific class of such supramolecular hosts (Figure 1.2). Cavitands are synthetic organic molecules containing cavities larger than some other smaller atoms or molecules. Unlike crown ethers, which only become rigid upon complexation, cavitands are rigid structures after synthesis.[5]
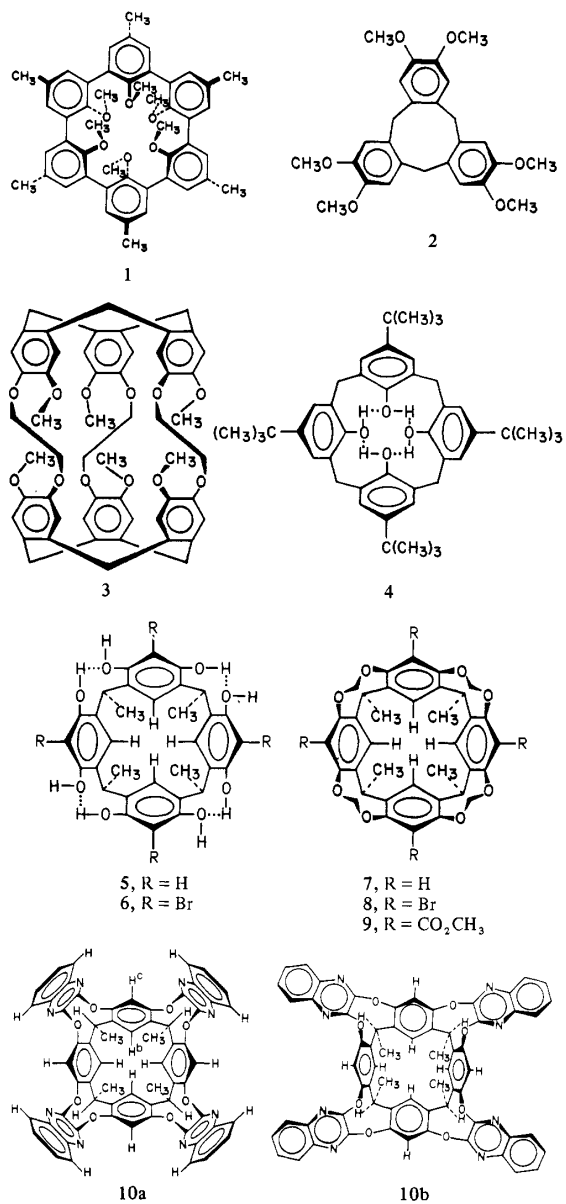
Figure 1.2: Several molecules called "cavitands" by Cram. Reprinted (adapted) with permission from Ref. 6. Copyright 1982 American Chemical Society.

This disseration uses computer simulations to investigate the assemblies of two cavitands which are part of a family known as "deep-cavity cavitands." Deep-cavity cavitands are bowl-shaped and are built upon a resorcin[4]arene base. These molecules utilize the hydrophobic effect—the low solubility of nonpolar molecules in water and their tendency to associate together in such aqueous environments—in order to self-assemble into dimeric, or even larger, assemblies when in the presence of suitable guests, depending on the chemistry of the cavitand. Unlike previous cavitands, deep-cavity cavitands have larger, bowl-shaped hydrophobic pockets.[7] One of the simplest of the deep-cavity cavitands is Octa-acid (OA) (see Figure 2.1), and Chapter 2 introduces the structural details of OA along with some of its realized and potential applications. We investigate both OA and one of it's derivatives known as Tetra-endomethyl Octa-acid (TEMOA, introduced in Chapter 3) throughout this dissertation using computer simulations.

Computer simulations provide an opportunity to investigate the molecular-level details of deep-cavity cavitand systems, both supplementing experimental results and guiding the design of new experiments. Molecular Dynamics (MD) is one specific method that can be used to simulate these systems.[8] MD integrates Newton's equations of motion over discreet time steps using a potential energy function of the following form:[9]

$$U = \sum_{bonds} K_b(b - b_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + cos(\eta\phi - \gamma)]$$
$$+ \sum_{i<j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon r_{ij}} \right)$$

Here the first three summations represent bonded interactions, and the last summation represents non-bonded interactions, including terms for both van der Waals interactions and electrostatic interactions. The constants in the above formula are determined by the specific force field used. In addition, long-range van der Waals interactions are typically truncated after a specified distance, and long-range electrostatic interactions are usually treated with variations of an Ewald summation.[10] The force field and simulation parameters used in this dissertation are described in each chapter.

After a force field is chosen an initial starting configuration of the molecular system of interest is created. As a Molecular Dynamics simulation is running data is saved for further analysis. This data can include the positions and velocities of the particles, as well as quantities such as the average temperature and pressure which correspond to their thermodynamic quantities.[11] The data is not usually saved at every step since correlations need to be removed.

In addition to regular MD simulations, we utilize free energy simulations[12] which can be used to determine the change of free energy in a system when some variable is changed. Often this is used in the case of coupling or decoupling a molecule with the surrounding system, for example, coupling a solute with a system of solvent, thereby determining the free energy of solvation of the solute. Simulations must be run at several discrete states of the variable of interest, often denoted as λ. After running simulations at all states, completing a thermodynamic cycle, techniques such as thermodynamic integration and the Bennett Acceptance Ratio[13] are used to compare the differences in free energy between two states. Then, summing all of the differences between each successive pair of states, the total free energy change between the beginning and ending

state can be estimated. Not only can the reaction coordinate in a free energy simulation be the interaction of a molecule with the rest of the system, it can also be a distance between two or more molecules. In that case a umbrella sampling and the weighted histogram analysis method (WHAM)[14] can be used to determine what is known as the potential of mean force. Free energy simulations are used in several different portions of this dissertation to describe deep-cavity cavitand systems.
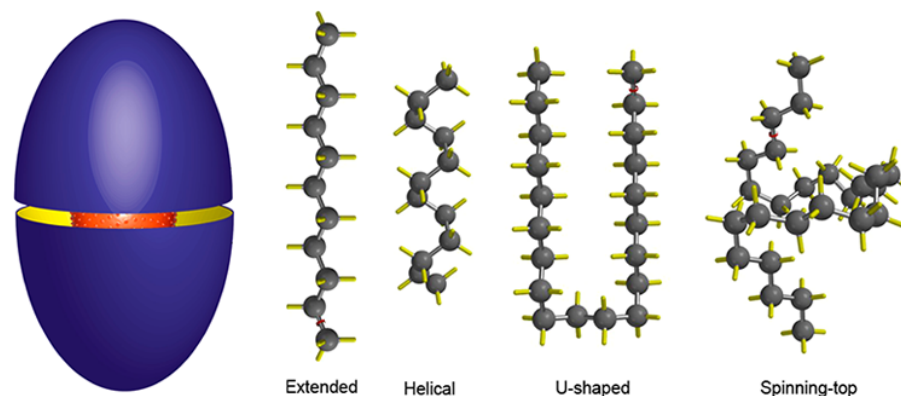
Figure 1.3: Schematic representation of an OA dimer with guest inside as well as four packing motifs inferred from NMR experiment. Reprinted with permission from Ref. 15. Copyright 2013 American Chemical Society.

1.1 Outline

Due to the relative simplicity of Octa-acid and its propensity to form dimeric capsules around hydrophobic guests,[15] we are able to effectively utilize simulations to explore the effects of confinement. Confinement is utilized throughout Nature in systems such as chaperonins,[16] capsids,[17] and vaults,[18] and OA allows us to study confinement, but at a much smaller scale. In Chapter 2 we discuss how we use computer simulations to investigate OA dimers with a homologous series of n-alkane guests confined in their

interiors, investigating how the hosts affect the guests. We use principal components analysis to reduce the dimensionality of our simulations and find a representative structure of each n-alkane guest simulated. The chemical shifts of each representative structure are calculated and show good agreement with previous experimental observations of length-dependent guest-packing motifs (Figure 1.3). Time averages are also shown to correspond with these representative structures, showing a length-dependent trend in their packing motifs.

One of the challenges of using the hydrophobic effect in self-assembly is creating assemblies comparable in size to other driving forces. In Chapter 3 we look at one of OA's derivates, Tetra-endoymethyl Octa-acid (TEMOA). TEMOA is chemically identical to OA except that four hydrogens on the rim have been changed to methyls. Unlike OA, however, TEMOA forms larger multimeric systems.[19] A schematic representation of TEMOA multermic assemblies is shown in Figure 1.4. We investigate the stability of TEMOA-alkane systems using a free energy technique. We show that the transition from one assembly state to the next, larger assembly state is driven by the guests' packing. Additionally the volumes of these multimeric systems begin to approach assemblies formed by other driving forces, such as hydrogen bonding.

In Chapter 4, we continue to look into the differences of OA and TEMOA. We investigate the water inside the binding pocket and discuss how water molecules interact with the interior of these two deep-cavity cavitands. We show that unlike OA, TEMOA is sometimes dry within its interior. Additionally we investigate pressure and temperature effects on this dewetting phenomenon. We also show that TEMOA has a larger partial molar volume than one would expect based on structure alone, due to its tendency to
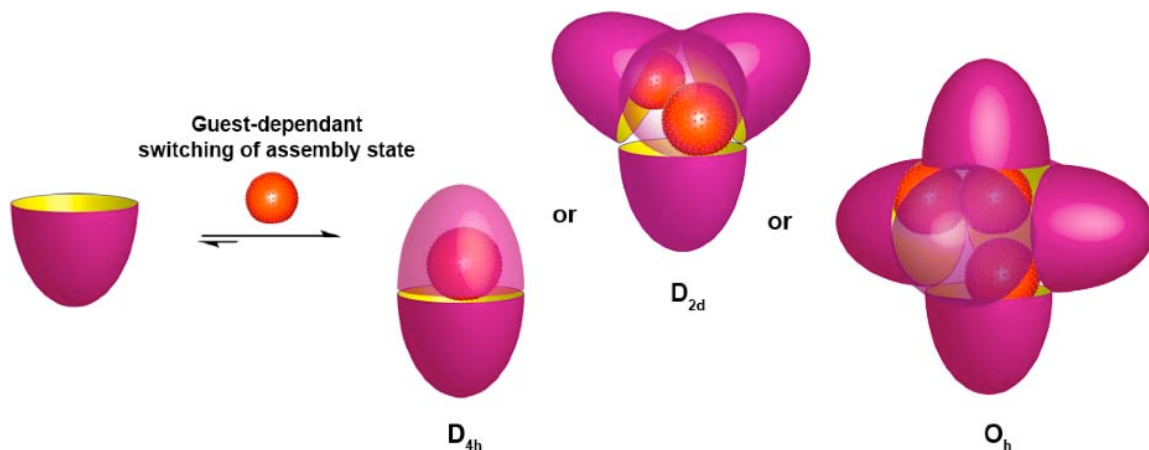
Figure 1.4: Picture representation of various assemblies that can form with TEMOA and a hydrophobic guest.

evacuate its interior waters This value is a possible bridge between our simulations and experiment.

The appendices detail work not related to deep-cavity cavitands, but nonetheless are relevant to supramolecular systems. Appendix A details another supramolcular host which utilizes both the hydrophobic effect and hydrogen bonding. This host is known as a macrocyclic tetralactam, which is made up of two anthracene walls and contains four amide groups points into its interior. Such a molecule that contains both hydrophobic and hydrophilic portions is termed "amphiphilic" due to its dual nature. We specifically investigate how waters interact with the interior region as well as how squaraine dyes bind to the host.

In Appendix B we revisit experimental estimates of interactions between two methanes. In the 1970's Ben-Naim developed experimental estimates for two methanes in the overlap region.[20] Since then, others have shown that such estimates do not line up with results from computers simulations.[93b] Here we show that by making a simple free energy perturbation calculation we can reconcile these estimates with simulation results.

Chapter 2

SUCCESSION OF ALKANE CONFORMATIONAL MOTIFS BOUND WITHIN HYDROPHOBIC SUPRAMOLECULAR CAPSULAR ASSEMBLIES

2.1   Summary

n-Alkane encapsulation experiments within dimeric Octa-acid cavitand capsules in water reveal a succession of packing motifs from extended, to helical, to hairpin, to spinning top structures with increasing chain length. Here, we report a molecular simulation study of alkane conformational preferences within these host-guest assemblies to uncover the factors stabilizing distinct conformers. The simulated alkane conformers follow the trends inferred from ^{1}H NMR experiments, while guest proton chemical shifts evaluated from Gauge Invariant Atomic Orbital calculations provide further evidence our simulations capture guest packing within these assemblies. Analysis of chain length and dihedral distributions indicates that packing under confinement to minimize nonpolar guest and host interior contact with water largely drives the transitions. Mean intramolecular distance maps and transfer free energy differences suggest the extended and helical motifs are members of a larger family of linear guest structures, for which the guest gauche population increases with increasing chain length to accommodate the chains within the complex. Breaks observed between the helical/hairpin and

hairpin/spinning top motifs, on the other hand, indicate the hairpin and spinning top conformations are distinct from the linear family. Our results represent the first bridging of empirical and simulation data for flexible guests encapsulated within confined nanospaces, and constitute an effective strategy by which guest packing motifs within artificial or natural compartments can be rationalized and/or predicted *a priori*.

2.2   Introduction

Living systems regularly employ nanoscale confinement to enable processes from catalysis and transport to storage and protection. Chaperonins like GroEL/GroES[16] and CCT[21] internalize misfolded proteins within their barrel-shaped interiors and catalyze refolding into the required tertiary structure. By similar means, eukaryotic vault organelles are suspected in affecting mRNA transport between the nucleus and cytoplasm,[18] while nucleic acids are protected from the extracellular environment within viral capsids and undergo significant conformational changes upon encapsulation that raises the internal pressure to aid viral gene injection into host cells.[17] More pertinent to the results described here, fatty acid binding proteins play key roles as transporters of this important class of molecules, and there is still much ambiguity about how fatty acid guest packing affects its binding and release thermodynamics[22] from proteins such as FadL.[23] The utility of these types of biotic capsules has stimulated investigation of abiotic, wholly synthetic systems that self-assemble into supramolecular containers. Applications for such entities are wide ranging, including: bioimaging, photodynamic therapy, drug and gene delivery.[24] Moreover, many artificial enzyme mimics are host-guest systems that employ reactant confinement to affect catalysis.[25]
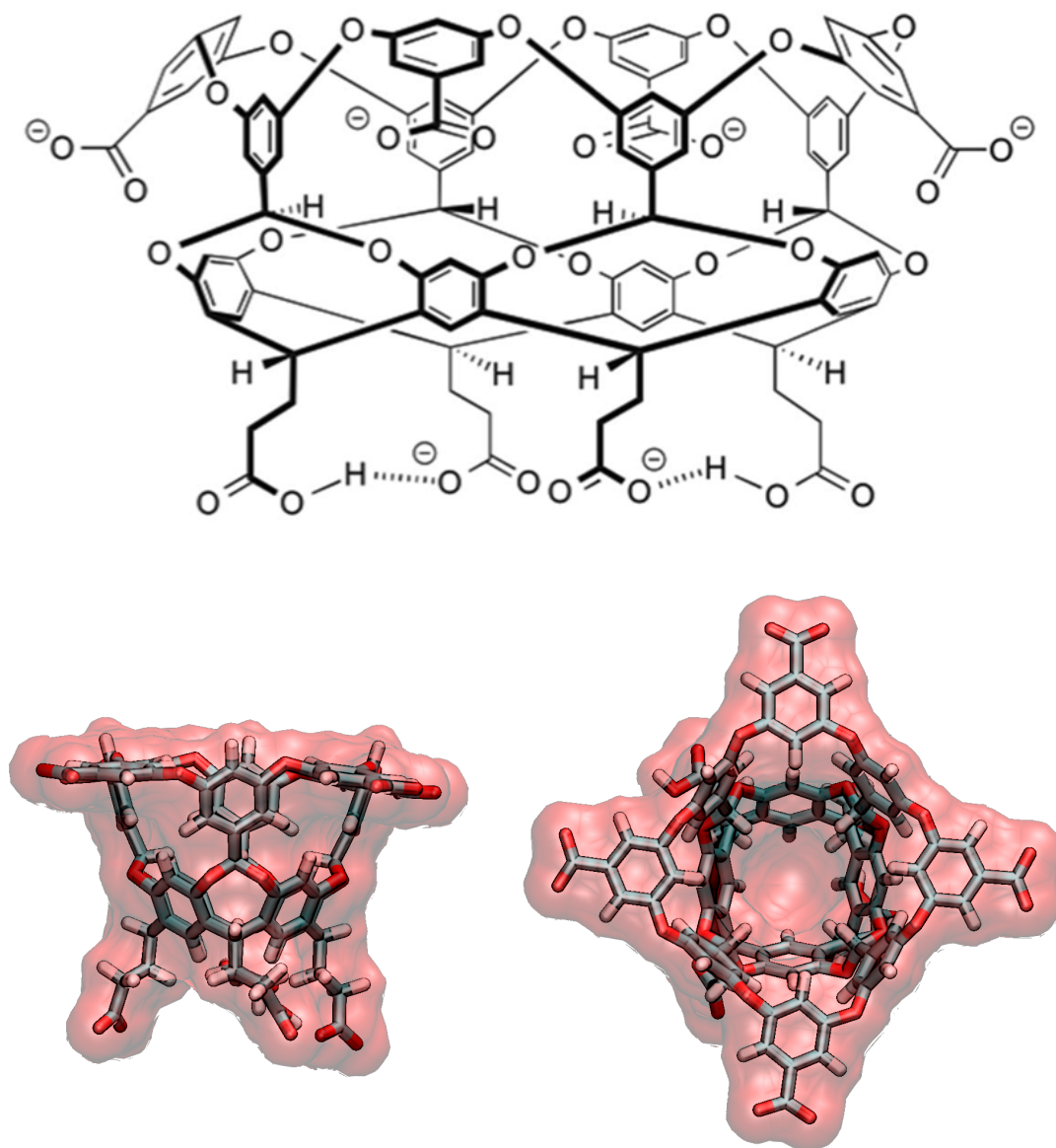
Figure 2.1: Structure of the Octa-acid deep-cavity cavitand host. The host-guest assemblies studied here are comprised of two host cavitands forming a dimer with an alkane guest encapsulated within the dimer's interior. *Top:* Chemical structure of an individual OA host. The host possesses three rows of aromatic rings that build up the concave binding pocket. The mouth of the cavitand at the top of this structure is rimmed with four carboxylic acid coating groups that are presumed to be fully deprotonated at pH 7. The foot of the cavitand possess four carboxylic acid coating groups with only two presumed to be deprotonated at pH 7. *Bottom:* Side and top views of an empty Octa-acid cavitand. The cavitand is depicted as a frame structure encased within a transparent solvent-excluded volume.

Octa-acid deep-cavity cavitand (OA) is a host molecule possessing three rows of aromatic rings that build a concave hydrophobic pocket approximately 7 Å in diameter and 7 Å deep (Figure 2.1). Conversely, the exterior rim and foot of the bowl are decorated with eight carboxylic acid coating groups that enable dissolution in water. Via the hydrophobic effect, OA readily assembles into dimeric capsules in the presence of suitably sized guests, forming 2:1 or 2:2 host-guest complexes[24c] that have been used as selective, yoctoliter-scale reactors,[26] and in the physical separation of hydrocarbon gases[27] and chemical resolution of constitutional isomers.[28] In difference to complex biological containers, the simplicity of OA provides an excellent test bed for studying both the forces associated with guest packing as well as the guest conformational motifs under molecular confinement that are not typically observed for unbound species.[29] Thus, phenyl-substituted hydrocarbons with short alkyl chains exhibit extended conformations in OA dimers, while longer chains packed with concavities or nano-tubes tend to preferentially adopt helical conformations and/or become folded.[30]

In a recent study of the effect of n-alkane packing within OA dimeric, 2:1 host-guest complexes in water, Gibb and coworkers observed a succession of conformational motifs with increasing guest chain length as inferred by one- and two-dimensional NMR experiments.[15] Beginning with n-nonane,[7a] the shortest guest that stabilizes a 2:1 complex, the alkane exhibited an extended motif enriched in trans dihedral conformations and the guest spanning the interstice from one cavitand to its partner. This extended motif compresses into a helical gauche-enriched motif with increasing chain length for guests such as n-tetradecane. Around n-heptadecane and n-octadecane, the guests adopt a turn and transition from a helical to a hairpin motif, with the two guest ends in a single

cavitand and a hairpin turn in the other. Finally, guests longer than n-tricosane undergo a final transition to a motif resembling a spinning top, with the two ends re-segregated to the opposing cavitand bottoms and the hairpin turn partially exposed to water.

Compared to larger, more complex biological vessels, OA assemblies are readily amenable to simulation to validate the experimental interpretation of guest packing and probe the molecular forces that shape their conformations under confinement. Here we present a systematic simulation study of dimeric OA host complexes with n-alkane guests (denoted C_n where n is the number of carbons) in water, with guests ranging in size from n-nonane (C_9) to n-pentacosane (C_{25}). We characterize the dominant guest structures by performing dihedral principal component analysis (DPCA) and predict changes in the proton chemical shifts of encapsulated guests using Gauge Invariant Atomic Orbital (GIAO) calculations. Conformationally dependent averages like the guest end-to-end lengths and vacuum-to-host complex transfer free energies are analyzed to characterize the guest conformational motifs, host-imposed packing constraints, and the thermodynamic forces stabilizing the observed structures.

2.3   Methods

Molecular Dynamics simulations of 2:1 host-guest systems containing an n-alkane guest encapsulated within an Octa-acid (OA) cavitand dimer were performed using GROMACS 4.6.[31] The homologous series of linear alkane guests (denoted C_n, where n is the number of carbons in the chain) were simulated from nonane (C_9) to pentacosane (C_{25}). The alkanes were modeled using the L-OPLS all-atom force-field,[32] which accurately captures the liquid state properties and conformational preferences of long alka-

nes. The OA cavitands were modeled using the Generalized Amber Force Field (GAFF)[9] with partial charges taken from the AM1-BCC calculations reported in Ref. 33. The OA structure in protein database format and interaction potential details are reported in Appendix C. Following Ewell, Gibb, and Rick,[34] the net charge on each OA was assumed to be -6e to reproduce the expected protonation state at pH 7. The four acidic coating groups ringing the hydrophobic pocket at the top of OA and two acids diagonal to one another at the foot of OA were deprotonated (Figure 2.1a). The carboxylic acid charges were neutralized by a total of 12 sodium cations (6 per OA) modeled using GAFF.[9] Water was modeled using the TIP4P-Ew potential.[35]

Each cavitand-alkane assembly was initially simulated in vacuum to equilibrate the encapsulated chain conformations before solvation of the complex in water. No periodic boundary conditions or potential cut-offs were employed in the vacuum simulations. Replica Exchange Molecular Dynamics[36] was used to extensively sample the conformational landscape of the confined alkanes and equilibrate the initial chain conformations used in our production runs carried out for the complexes in water. Four replicas were simulated at 280, 315, 353, and 394 K, chosen to achieve an exchange rate of âĽij5%.[37] The temperature was controlled using the Nosé-Hoover thermostat.[38] Simulations were performed for 5 μs, integrating the equations of motion using the leapfrog algorithm with a time step of 2 fs. Exchange attempts were carried out every 2 ps. To ensure the complexes did not disassemble during the vacuum simulations, harmonic restraints were applied between eight pairs of carbons ringing the mouth of the OA dimer. The harmonic bond length and spring constant were 4 Å and 15 kJ/(mol Å^2), respectively.

The conformational analysis reported below was performed for complexes simulated in water at 333 K (60 °C) and 1 bar pressure. This temperature conforms to the temperature at which the alkane/cavitands were prepared experimentally to melt the longer chain alkanes and get them into the capsules. Initial host-guest configurations for the simulations in water below were harvested from the final frame of the simulation replica at 315 K and hydrated with between 2563 and 5932 waters. The harmonic restraints imposed in the vacuum simulations above were released in these simulations to allow the conformations of the longest chains to further relax. This is most significant for the chains adopting the spinning top motif in which the guest partially exposes itself to the solvent. The temperature and pressure in solvent were controlled using the Nosé-Hoover thermostat[38] and Parrinello-Rahman barostat,[39] respectively. Short-range van der Waals interactions were cut off beyond 9 Å with a mean-field dispersion correction for longer-range forces. Long-range electrostatic interactions were evaluated using particle-mesh Ewald summation with a real-space cutoff of 9 Å.[10d] Following at least 1 ns of equilibration, each system was simulated for 30 ns and 30,000 frames were saved for analysis of conformational averages.

To obtain a representative structure for the dominant conformational motif of each encapsulated alkane, we performed dihedral principal component analysis (DPCA) on the set of dihedral angles down the guest carbon backbones extracted from simulation.[40] This dimensional reduction technique has been successfully applied to study not only the structures of proteins but of n-alkanes in bulk aqueous solution.[41] DPCA was used here to determine the two dominant eigenvectors of the dihedral covariance matrix that exhibit the greatest variance during our simulations. Probability maps of

the coefficients—the principal components—of those two eigenvectors were evaluated from the transformation of the original dihedral coordinates from simulation into the eigenvetor coordinate system. The simulation configuration whose principal components lie closest to the most probable value was subsequently judged to represent the dominant chain conformation. We use that simulation configuration to assess the preferred conformational motif of the guest and to evaluate the bound guest ^{1}H NMR chemical shifts.

^{1}H NMR chemical shifts are sensitive to the local environment the analyte hydrogens are subjected. Subsequently, alkane hydrogen depths and positions within cavitand interiors are empirically correlated with perturbations in ^{1}H NMR chemical shifts upon transfer from solution into the complex. These changes in the experimental ^{1}H NMR chemical shifts have been used to interpret the conformation of guests within supramolecular complexes. Chemical shifts can be computed from isotropic shielding constants determined from ab initio calculations to validate the conformations inferred from experiment against molecular scale structures evaluated from simulations. Specifically, we performed Gauge Invariant Atomic Orbital calculations (GIAO)[42] on cavitand-alkane complexes using Gaussian03[43] at the B3LYP/6-31G(d,p) level to determine isotropic shielding constants. The chemical shifts (δ) of bound and free guest protons bonded to alkane carbon i are calculated as the difference between the isotropic shielding constant (σ) for a reference compound (*e.g.*, tetramethylsilane) and the computed shielding constant of the specified proton

$$\delta_i^{bound} = \sigma^{ref} - \sigma_i^{bound}$$

and

$$\delta_i^{free} = \sigma^{ref} - \sigma_i^{free}$$

The change in chemical shift between a free to bound guest proton ($\Delta\delta_i$) subsequently reduces to

$$\Delta\delta_i = \delta_i^{bound} - \delta_i^{free} = \sigma_i^{free} - \sigma_i^{bound}$$

To evaluate the isotropic shielding constants of the protons on the free guest we performed geometry optimization for each alkane in the all trans conformation. GIAO calculations were performed on the optimized unbound guest in a solvent reaction field using the polarizable continuum model. Isotropic shielding constants obtained from this calculation were subsequently averaged based on the internal symmetry of the alkane. To evaluate the shielding constants of the bound guest protons we used the dominant encapsulated alkane conformation as determined by DPCA above. No minimization of the host-guest structure was performed in order to preserve the alkane position within the complex determined from simulation. While the experimentally measured chemical shifts are an ensemble average over guest conformations, the capacity to perform *ab initio* calculations over a large number of simulation configurations—30,000 for each alkane—is beyond current computational capabilities. We, therefore, limit our analysis to the dominant guest conformation. GIAO calculations on structures obtained from molecular simulations have been previously reported to examine conformational contributions to the chemical shifts of a range of systems.[44]

To characterize the thermodynamic contributions to the guest conformational preferences we evaluated the vacuum-to-capsule transfer free energy differences to grow

the alkanes from n-1 to n carbons. Free energy differences were evaluated for all the guests simulated above from C_9 to C_{25}. These simulations were carried out using GRO-MACS 5.0 at 333 K and 1 atm pressure in 2666 to 4968 waters. van der Waals interactions were smoothly cutoff from 10 to 12 Å, while the real space contribution to electrostatic interactions were cutoff at 12 Å. Bennett's acceptance ratio was used to evaluate free energy differences.[45] Alchemical transformations in vacuo and within the capsule interior were conducted over thirty-six different states using a λ-coupling parameter approach. Each state was simulated for 5 ns following equilibration for at least 1 ns. The transformations were conducted over four phases to evaluate separate electrostatic and van der Waals contributions to the free energy. In the first phase the charges on a terminal methyl of the n-1 carbon alkane were turned off in 0.25-λ increments from 1 to 0, were 1 indicates full interactions and 0 indicates no interactions. In the second phase the new terminal methyl group van der Waals interactions of the n-alkane were turned on in 0.05-λ increments. In the third phase the electrostatic interactions of the methylene adjacent to the new terminal methyl unit were turned on in 0.25-λ increments, while in the fourth phase the electrostatic interactions of the methyl unit were turned on in 0.25-λ increments. Reverse calculations incrementally deleting the chain were not performed.

2.4 Results and discussion

2.4.1 Dominant encapsulated alkane conformers

Simulation snapshots of the dominant conformational motifs of C_{11}, C_{14}, C_{18}, and C_{25} encapsulated within the deep-cavity cavitand dimer complex as ranked by DPCA show

distinct differences with increasing chain length (Figure 2.2). C_{11}, for instance, adopts the minimal intramolecular energy extended motif with its two ends stretched from the bottom of one cavitand into the other. While C_{14} similarly spans the two cavitands, the chain appears to adopt twisted gauche dihedral conformations with an overall helical structure. Correlations between gauche conformations ultimately give rise to a turn near the middle of the chain so that C_{18} adopts a hairpin motif with the two ends directed into one cavitand and the turn directed into its partner. Eventually the dimer's internal volume cannot fully accommodate longer guests like C_{25} and the complex herniates to expose a hairpin turn to water while, at the same, time the two ends of the alkane re-segregate to the opposing ends of the capsule, and the guest adopts a structure reminiscent of a spinning top. The dominant conformers from C_9 to C_{25} as determined from DPCA are illustrated in Figure 2.3. These guests follow the succession of conformational motifs discussed above: C_9 to C_{11} exhibit extended conformations, C_{12} to C_{16} exhibit helical conformations, C_{17} to C_{19} exhibit hairpin conformations, and C_{20} to C_{25} exhibit spinning top conformations. These conformations follow the experimental progression, although the spinning top motif precedes experiment at C_{20} compared to C_{23}.

Experimental guest conformations are characterized by perturbations in the proton chemical shifts ($\Delta\delta$), which are themselves sensitive to the average depth a proton is located within a cavitand. Predicted ^{1}H NMR $\Delta\delta$ values for the dominant encapsulated guest conformations evaluated from GIAO calculations are plotted against the experimental results reported in Refs. 24c and 15 in Figure 2.4a. The calculated and experimental $\Delta\delta$'s are strongly correlated over a span of values from -4 to 0 with a

| C$_{11}$ (extended) | C$_{14}$ (helical) | C$_{18}$ (hairpin) | C$_{25}$ (top) |
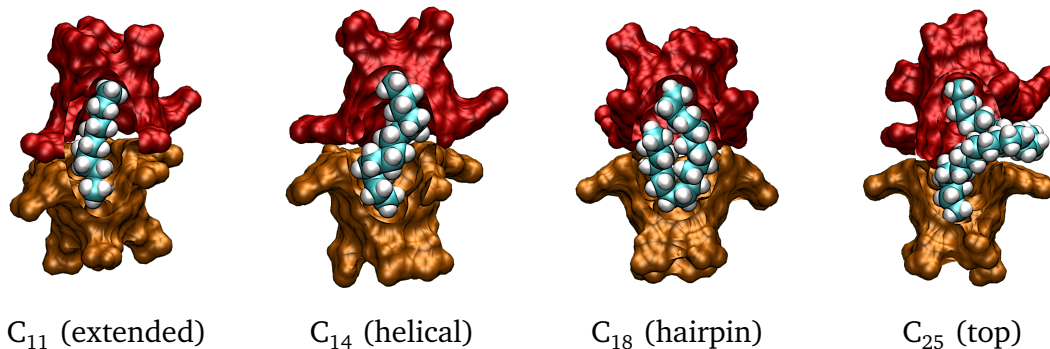
Figure 2.2: Representative simulation snapshots obtained from DPCA ranking of the *n*-alkanes undecane (C$_{11}$), tetradecane (C$_{14}$), octadecane (C$_{18}$), and pentacosane (C$_{25}$) encapsulated within OA cavitand dimers in water at 333 K and 1 bar pressure. These alkanes exhibit the conformational motifs extended (C$_{11}$), helical (C$_{14}$), hairpin (C$_{18}$), and spinning top (C$_{25}$). The two OA cavitands are drawn as red and orange solvent accessible surfaces, while the alkanes are illustrated using the van der Waals representation. A dividing plane has been applied to the cavitands to more clearly shown the guest inside the assembly. The waters hydrating the complexes have been omitted for clarity.

root mean square difference of 0.59. Linear regression between the simulated and experimental $\Delta\delta$ values yields a slope of 1.09, in good agreement with the parity slope of 1 and giving confidence that the simulations are representative of experiment. A potential source of scatter in this comparison arises from the fact that we have only considered one dominant conformation for each alkane while truly the guests dynamically fluctuate. Moreover, the protons near the turn of the spinning top motif are partially exposed to water while the solvent was excluded from our *ab initio* calculations. Given the computational expense for performing GIAO calculations on multiple chain conformations within the OA dimer, we are constrained to considering only a limited number of chain conformations. If the solvent exposed protons—indicated by the open circles in Figure 2.4a—are excluded, however, we find improved agreement with a root
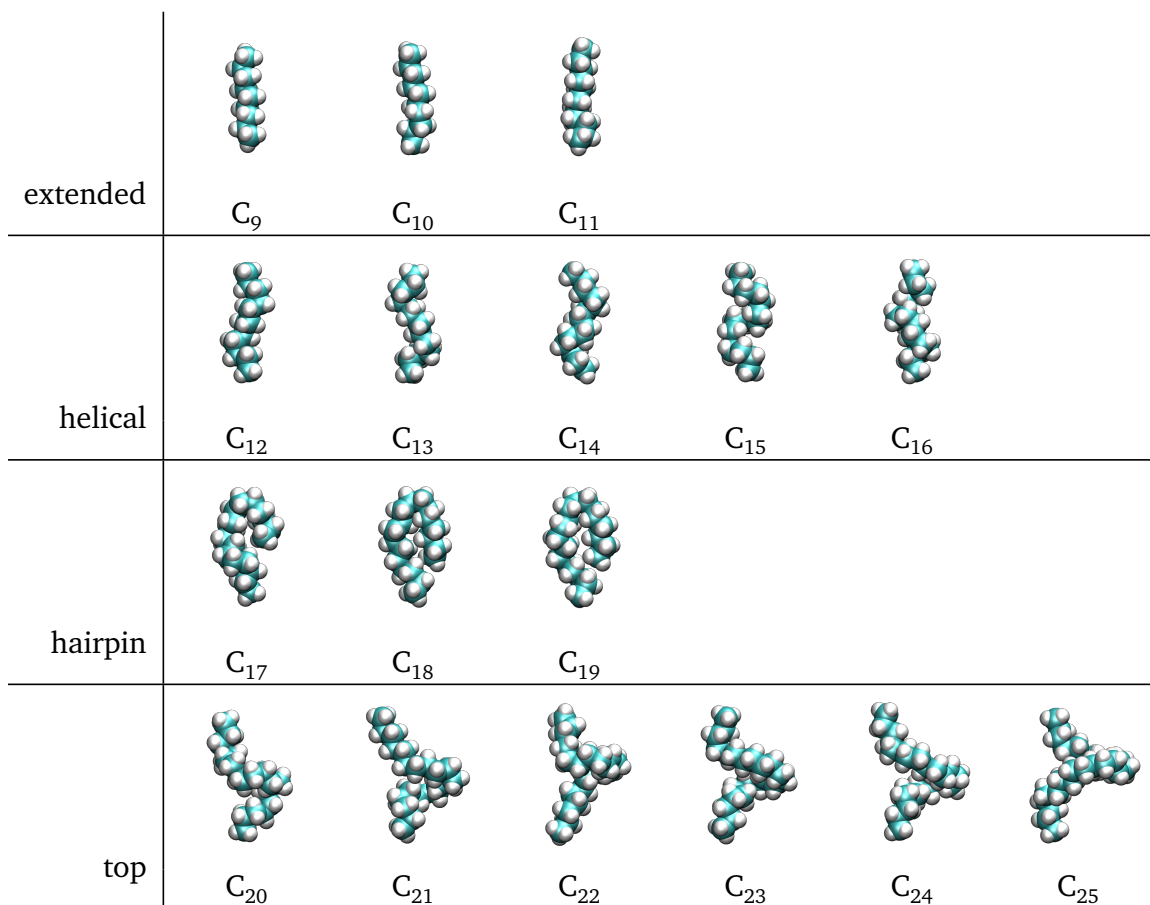
Figure 2.3: Snapshots obtained from DPCA ranking of guest conformational motifs within OA complexes for *n*-alkanes from nonane to pentacosane. The extended motif is observed for alkanes from C_9 to C_{11}. The helical motif is observed from C_{12} to C_{16}. The hairpin motif is observed from C_{17} to C_{19}. The spinning top motif is observed from C_{20} to C_{25}. These representative snapshots were harvested from simulations of the guest encapsulated OA cavitand dimers in water at 333 K and 1 bar pressure. The cavitands and water have been omitted to clearly show the alkane conformation.
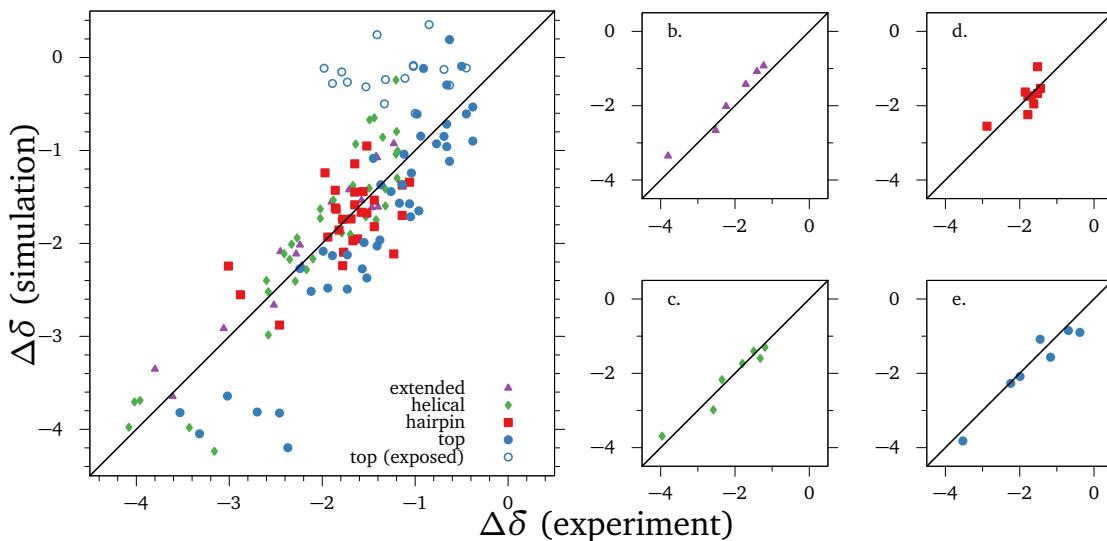
Figure 2.4: Comparison between the changes in the ^{1}H NMR chemical shifts ($\Delta\delta$'s) upon encapsulation reported from experiment[15] and predicted from GIAO calculations performed on the simulation guest structures reported in Figure 2.3. *a.* Overall comparison for alkanes simulated. *b.* Comparison of results for the dominant extended motif of C_{11}. *c.* Comparison of results for the dominant helical motif of C_{14}. *d.* Comparison of results for the dominant hairpin motif of C_{18}. *e.* Comparison of results for the dominant spinning top motif of C_{25}.

mean square difference between the simulated and experimental $\Delta\delta$'s is 0.46 and a regression slope of 1.02.

While Figure 2.4a identifies results for guests exhibiting the extended, helical, hairpin, and spinning top motifs, the scatter in the plot makes it difficult to distinguish between them. We have subsequently broken out the results for C_{11}, C_{14}, C_{18}, and C_{25} in Figures 2.4b–e , which individually adopt one of the four identified encapsulated guest motifs as shown in Figure 2.2. $\Delta\delta$ parity plots for each alkane are provided in Appendix C. The root mean square differences between the simulated and experimental $\Delta\delta$ values for C_{11}, C_{14}, C_{18}, and C_{25} are 0.23, 0.23, 0.30, and 0.31, respectively, and are generally better than for the overall comparison. This suggests that while there is

some discrepancy between the experimental and simulated conformations at the borders between motifs (*e.g.*, as noted for the hairpin and spinning top transition above) potentially due to sampling of multiple motifs near the transition, we find that guests that lie firmly in the range of guests exhibiting a specific motif affect a better comparison with experiment (*e.g.*, C_{14} lies in the middle of the range of guests exhibiting the helical motif). We conclude then that the simulations successfully track the succession of guest conformers inferred from ^{1}H NMR experiments.

2.4.2   Characterization of encapsulated guest conformational averages

To this point we have focused only on single representative conformations for each alkane, while in reality each guest dynamically samples a range of conformations. To establish that the conformational ensemble of each alkane tracks the motifs identified above, we consider how guest conformational averages vary with chain length. For instance, the root mean square end-to-end length of the guests in the extended regime from C_9 to C_{11} grows by approximately 1.12 Å with each added carbon unit (Figure 2.5), comparable to the incremental change in the length of an alkane in an all trans conformation of 1.26 Å. The average population of trans dihedral conformations over this range of guests is 65% (Figure 2.6), which is greater than that of an individual C_{17} chain in bulk water at the same temperature and pressure (55%) and is consistent with the assertion that confinement prefers elongated structures in this regime. At C_{11} the end-to-end length of the guest is 11.2 Å, which corresponds to the internal major axis length of an empty dimer available to the alkane when the van der Waals radius of the methyl units (∼1.9 Å) are taken into account. Alkanes between C_{12} and C_{16} subsequently adopt

a helical motif to fit within the dimer assembly and the population of trans dihedrals significantly drops well below the bulk value with increasing chain length (Figure 2.6). The mean end-to-end length of the guests only increases by 0.15 Å per carbon in this regime (long-dashed line in Figure 2.5), consistent with the chains adopting a compressed structure. Tilting of the guest and off-axis slipping of the dimerized cavitands relative to one another easily accommodate the slight end-to-end growth with chain length. Gauche compression eventually overwhelms the guest structure, however, and starting with C_{17} the alkanes enter the hairpin regime. The end-to-end length drops precipitously for C_{17}, C_{18}, and C_{19} compared to the helical conformers indicative of the two methyl ends oriented in the same direction (Figure 2.5) while the trans dihedral population partially recovers to relieve the gauche strain (Figure 2.6). For alkanes longer than C_{19}, the end-to-end distance jumps back up to the helical guest trend as a result of redirection of the methyl ends into the opposing cavitands in the spinning top conformers (Figure 2.5). The trans dihedral population for the spinning top alkanes is comparable to that for the hairpin (Figure 2.6); however, as a result of retention of the turn that is now sandwiched between the two cavitands and partially extruded into the solvent (Figure 2.2).

It is interesting to note that the ratio of the end-to-end length of an all trans alkane to its perfect helical gauche conformation is 1.42, which corresponds to the ratio of the number of carbons in the guest at the helical/hairpin transition to that at the extended/helical transition ($16/11 = 1.45$). This observation is consistent with the hypothesis that guest packing under confinement drives the conformational changes.
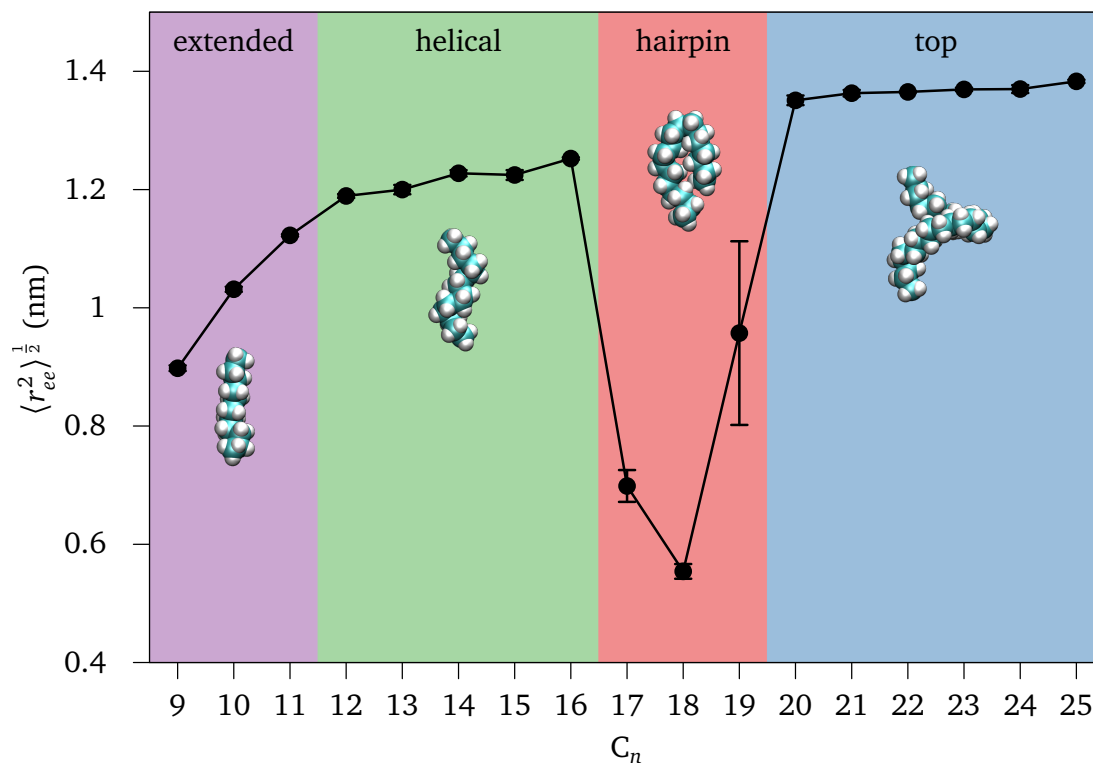
Figure 2.5: Average end-to-end length of the encapsulated alkanes as a function of their chain length. The averages were gathered from simulations of the OA dimer complex in water at 333 K and 1 bar pressure. The simulation results are reported as the points connected by the solid line. Error bars indicate one standard deviation in the simulation results, though in many cases the error bars are smaller than the figure symbols. The purple, green, red, and blue shading indicates the dominant conformational motif observed from extended, to helical, to hairpin, to spinning top, respectively.
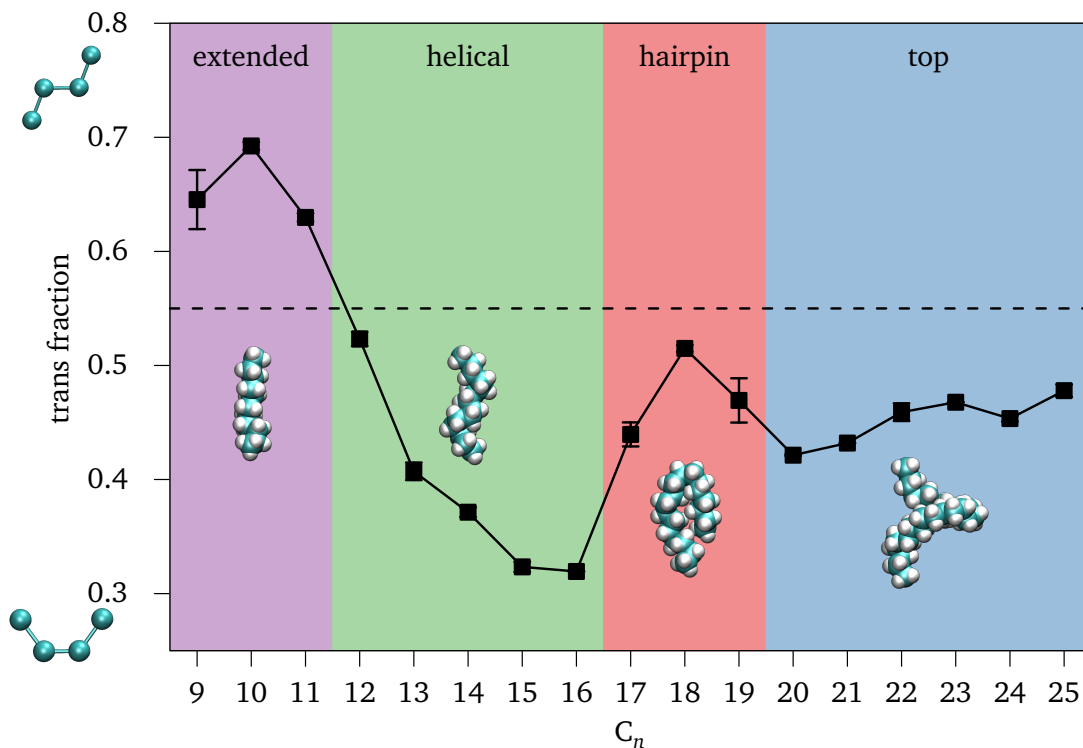
Figure 2.6: The population of trans dihedral conformations of the encapsulated alkanes as a function of their chain length. The averages were gathered from simulations of the OA dimer complex in water at 333 K and 1 bar pressure. The simulation results are reported as the points connected by the solid line. Error bars indicate one standard deviation in the simulation results, though in many cases the error bars are smaller than the figure symbols. The horizontal dashed line indicates the average fraction of trans dihedrals observed from simulations of C_{17} in bulk water outside the dimer capsule at 333 K and 1 bar pressure. The purple, green, red, and blue shading indicates the dominant conformational motif observed from extended, to helical, to hairpin, to spinning top, respectively.

Beyond the end-to-end length, additional molecular detail on the alkane conformation can be obtained from plots of the mean distance between carbons i and j down the guest backbone $\langle r_{ij} \rangle$. Plotted as a function of i and j these mean distance maps are symmetric down the $i = j$ diagonal with $\langle r_{ij} \rangle = 0$, while the off-diagonal elements report details of the chain conformation. For the n-alkanes we also expect the distance map to be bisymmetric about the $j = n + 1 - i$ cross-diagonal (n is the number of carbons in the alkane C_n) due to the indistinguishablility of counting down the alkane from either end. This symmetry, however, can be suppressed during finite simulations due to kinetic barriers under confinement. Mean distance maps for C_{11}, C_{14}, C_{18}, and C_{25} plotted in Figure 2.7 show systematic trends consistent with the progression of alkane conformational motifs. Maps for all guests provided in Appendix C follow the trends discussed here. For example, the mean distance between carbons for the extended guest C_{11} (Figure 2.7a) are essentially linear with the relative difference in position of the two carbons on the chain (*i.e.*, $\langle r_{ij} \rangle \propto |i - j|$), indicative of an extended conformation. As a result the distance map shows a banded structure parallel to the $i = j$ diagonal. The distance map for the helical alkane C_{14} (Figure 2.7b) is comparable to that for C_{11} since the helix similarly propagates in one-dimension with added carbons, although minor deviations in the $|i - j|$ proportionality indicate conformational perturbations. A more significant change is observed for the hairpin guest C_{18} (Figure 2.7c), for which off-diagonal wings sprout off the banded structure of the extended and helical motifs as a result of the hairpin turn bringing more distantly connected carbons closer to one another. The turn appears to be centered about carbon 7 of C_{18}, as determined by the intersection of the wing with the diagonal, and indicates the guest is best thought of

as a low symmetry J-like structure as seen for the dominant hairpin conformations illustrated in Figure 2.3. Based on the turn position, the longest leg of the J consists of 11 (= 18 - 7) carbons, in agreement with the length of the longest extended chain C_{11} that fits within the cavitand dimer. The mean distances plotted in Figure 2.7c averaged over the entire simulation trajectory are not symmetric about the cross-diagonal, however, due to barriers for cooperative translation of the turn along the chain backbone from between carbons 7 and 8 up to carbons 10 and 11. The off-diagonal wings remain for the distance map of the spinning top guest C_{25} (Figure 2.7d). These wings are not as broad as those of the hairpin although they do exhibit cross-diagonal symmetry. In difference to C_{18} the most distantly connected carbons 1 and 25 are also the most distant carbons in the chain as a result of the re-segregation of the methyl ends into the opposing cavitand bowls. The wings for C_{25} display low separation islands along the cross-diagonal centered about carbons 6 and 19. These islands can be attributed to the turn extruded towards the bulk solution pinching back around before the two ends (ranging from carbons 1 to 6 and from 19 to 25) partition into the two cavitand bowls. The lengths of these two end stretches (between 5 and 6 carbons) are approximately half that of the longest extended chain, C_{11}, consistent with the two halves of the guest anchored into either cavitand being in predominantly trans conformations. The bisymmetry of the spinning top motif indicates the barriers for translation of the turn observed for the hairpin are not present in that of the turn of the spinning top, which is partially exposed to the solvent and not fully confined by the capsule.
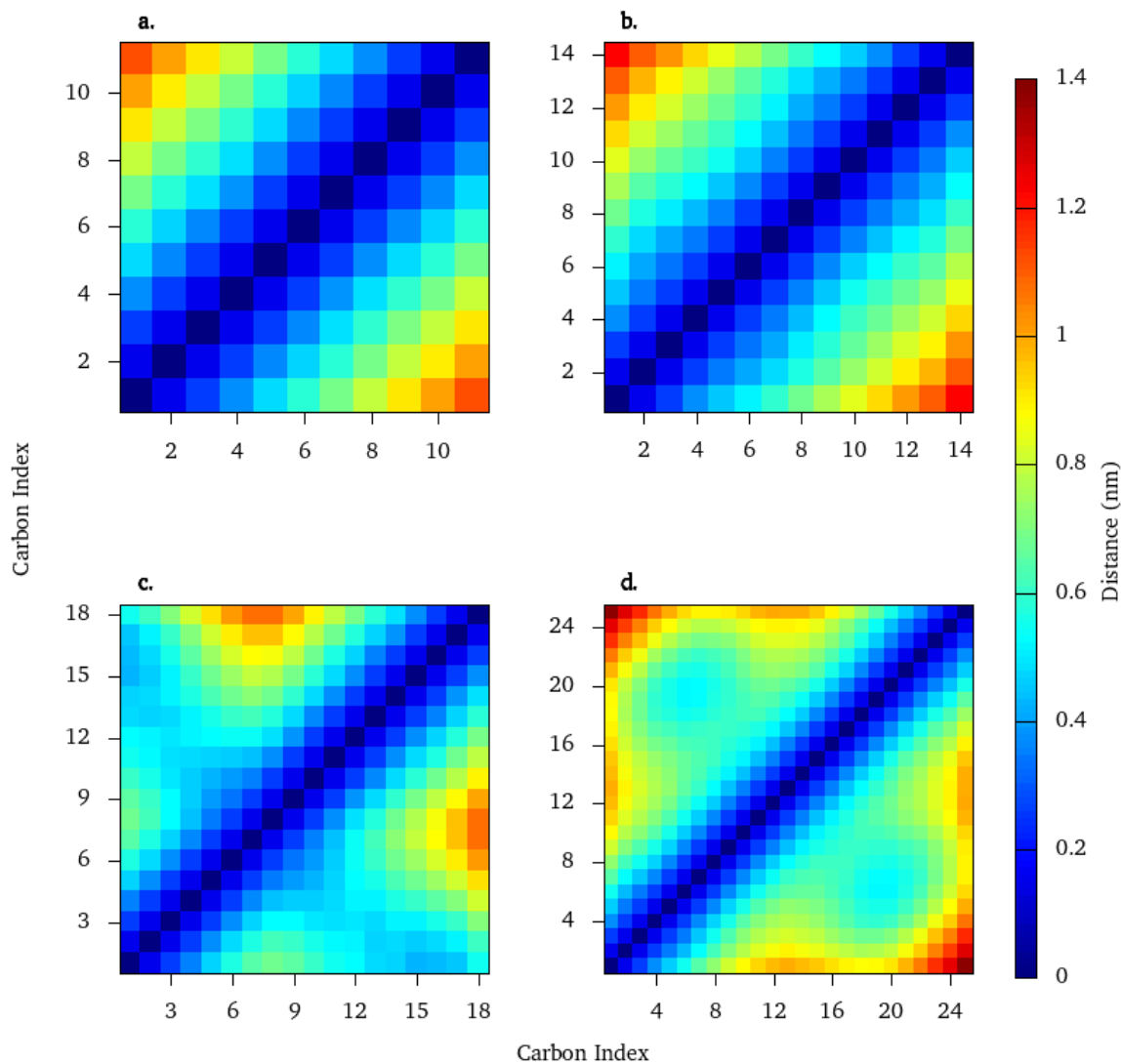
Figure 2.7: Mean distance maps reporting the average distance between two specified carbons (denoted by their carbon index) of alkane guests within the OA dimer complex. Mean distances are keyed to the color scale given on the right-hand side of the figures. The figures a, b, c, and d report results for C_{11}, C_{14}, C_{18}, and C_{25}, respectively. The averages were gathered from simulations of the OA dimer complex in water at 333 K and 1 bar pressure.

2.4.3 Guest vacuum-to-complex transfer free energies

Minimization of the guest free energy ultimately determines its conformation. The free energy increments for transforming C_{n-1} into C_n show four regimes that follow the conformational transitions discussed above (Figure 2.8). The free energies of transforming C_9 to C_{10}, C_{10} to C_{11}, and C_{11} to C_{12} are favorable (negative) and approximately independent of the chain length. While guests C_9 to C_{11} exhibit an extended structure, C_{12} lies on the border between the extended and helical motifs and correspondingly the free energy increments from C_{12} to C_{16}, systematically increase for alchemical changes associated with the helical motif. This indicates growing strain due to the increasing unfavorable gauche dihedral population reported in Figure 2.6. If the free energy trend observed for the extended chains could be extrapolated into the helical regime we would expect the chains to maintain the extended conformation from C_9 to C_{16} since the extended free energy increment is lower than those for the helical chains. Rather we believe the transition between the extended and helical motifs is not a true thermodynamic transition, but reflects strained packing of the effectively one-dimensional guest within the dimer as it collides with both sides of the container. The extended and helical structures can better be thought of as members of a unified âĂŸlinearâĂŹ motif whose free energy increments follow the solid line indicated in Figure 2.8. This interpretation is consistent with the gradual growth of the mean end-to-end distance with increasing chain length over the extended and helical regimes (Figure 2.5) and the similarity of their distance maps (Figures 6a and b).
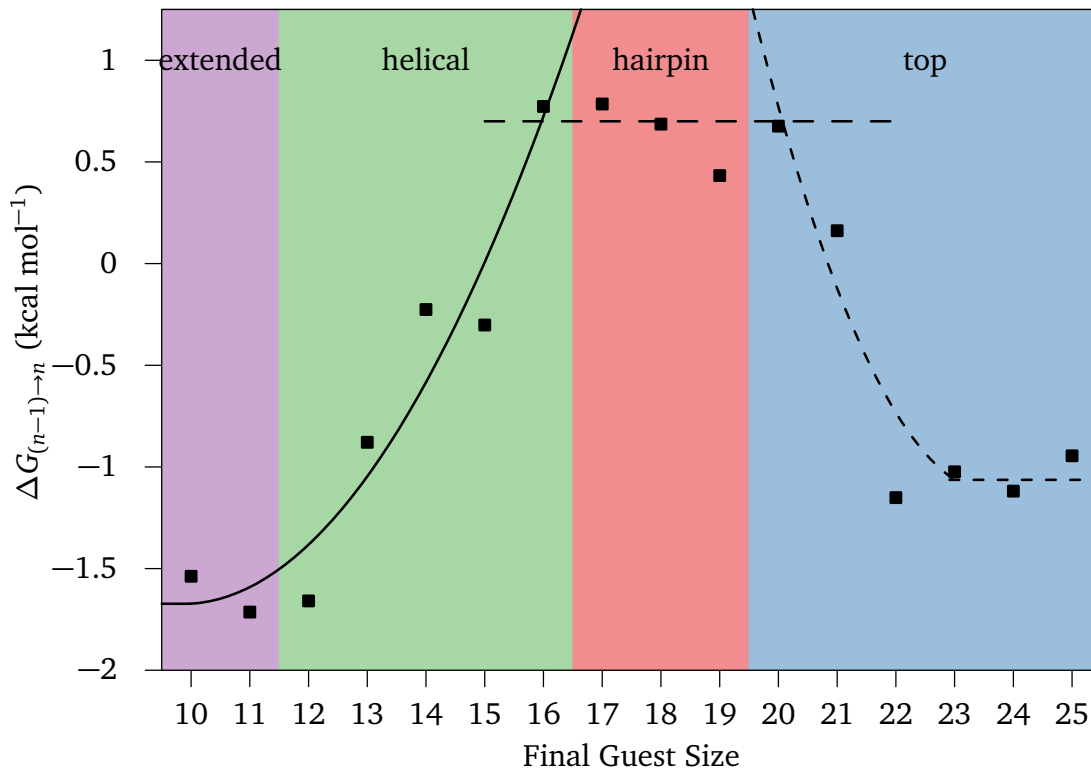
Figure 2.8: Incremental excess free energy change for transferring an alkane of length n-1 to n from vacuum to the cavitand dimer interior. The free energies were gathered from simulations of OA dimer complexes in water at 333 K and 1 bar pressure. The simulation results are reported as the points. The error bars in the calculated free energies are comparable in size to the symbols. The shading indicates the dominant conformational motifs observed. The lines drawn in this figure serve only as guides to the eye.

In contrast, breaks in the free energy consistent with true conformational transitions are observed for chain lengths beyond the helical regime. For instance, the free energy increments associated with the hairpin motif (C_{17} to C_{19}) are effectively independent of chain length, being positive and unfavorable (Figure 2.8). The transition between the linear and hairpin regimes is approximately determined by the intersection of the solid and long-dashed curves in Figure 2.8, where jumping from the extrapolated extend/helical curve down to the hairpin curve minimizes the incremental free energy for chains longer than C_{16}. A second break in the incremental free energy is observed in the neighborhood of C_{20}, indicative of the transition between the hairpin and spinning top regimes. In the spinning top regime the incremental free energy (short-dashed curve in Figure 2.8) drops from the hairpin plateau attaining a new negative free energy plateau for the longest chains simulated. We note that the breaks we observe between the extend/helical, hairpin, and spinning top regimes are not sharp as we would expect for a first order transition and suggests coexistence between motifs for the alkane chains near the identified boundaries.

2.5 Conclusion

Our molecular simulations of alkane guest conformational equilibrium confined within dimeric OA host assemblies support the conclusion that the interplay between the end-to-end length of the fully extended chain and the internal length of the capsule along its major axis is a significant determinant of the dominant guest conformation. Specifically, alkanes shorter than the length necessary to span the capsuleâĂŹs major axis (C_{11} and shorter) predominantly adopt an extended motif enriched in trans dihedral confor-

mations compared to the free guest. Guests between C_{12} and C_{16} assume an increasing population of gauche dihedrals as the guest is compressed into a helical motif. The transition between these two motifs is continuous, however, as observed in the changes in the end-to-end length over this range of guests and the similarities between their mean carbon-carbon distance maps, indicating they may be thought of as members of a larger linear motif family which propagates along one-dimension with increasing chain length. Nevertheless, the helical motif is more strained than its extended brethren as indicated by their increasing free energies with chain length, resulting from the increase in the fraction of gauche dihedrals and collisions with the interior walls of the capsule. Between C_{16} and C_{17}, however, the guest undergoes a more dramatic conformational change from the helical to hairpin motifs where the guest adopts a turn that directs the two terminal methyls into one cavitand while the turn sits in the other host bowl. Following the initiation of the helical motif after C_{11}, we believe this transition is driven by the coincidence of the length of the capsuleâĂŹs major axis with length of a per-fectly helical (all gauche) C_{16} guest. C_{17} subsequently adopts the hairpin conformation in an effort to maintain the capsuleâĂŹs integrity. The onset of the hairpin motif for guests C_{17} to C_{19} is accompanied with discontinuous changes in the guest end-to-end lengths and mean carbon-carbon distance maps, as well as a break in the incremental transfer free energy that indicates the hairpin motif is the minimal free energy motif for these guests. Guest packing, even in the folded hairpin motif within the capsule, with increasing chain length eventually becomes untenable, however, and the host capsule herniates and partially exposes the guest to water in the spinning top motif observed for chains longer than C_{19}. In difference to the hairpin, the terminal methyl units of the

guest in the spinning top motif are directed into the two opposing cavitand halves while the turn is sandwiched between the cavitands and towards the solvent. Like the hairpin, the spinning top motif is accompanied with discontinuous changes in its structural measures and a break in the incremental free energy indicating a new stable state.

The computational strategy utilized here confirms the experimental inferred structures of encapsulated alkanes, and provides a clearer picture of how packing under confinement stabilizes specific guest conformational motifs.

Chapter 3

ALKANE GUEST LENGTH DRIVEN ASSEMBLY TRANSITIONS BETWEEN
MULTIMERIC DEEP-CAVITY CAVITAND HOST COMPLEXES

3.1   Summary

Experiments of Tetra-endomethyl Octa-acid (TEMOA) deep-cavity cavitands demonstrate that, unlike Octa-acid, TEMOA is able to form multimeric assemblies with n-alkane guests. We report here the results of molecular simulations of TEMOA and a series of n-alkane guests. Free energies of transferring guests into dimeric capsules are comparable to the results from potentials-of-mean force. We use this transfer process not only in dimeric systems, but also in tetrameric and hexameric assemblies, demonstrating that packing drives the relative stability of one aggregation state over. We make this connection between packing and assembly stability by calculating packing fractions of each assembly state, and at the same time, demonstrating the size of the assemblies.

3.2   Introduction

Nature routinely utilizes biological confinement within channels and containers to realize an array of outcomes unachievable in the bulk medium. Water constrained to single-file hydrophobic channels in proteins like bacteriorhodopsin and cytochrome c oxidase

aids proton translocation across cellular membranes along hydrogen-bonded proton "wires".[46] Chaperonins like GroEL/GroES mediate protein folding through the encapsulation of unfolded and misfolded proteins under close confinement to catalyze refolding before release.[16] Barrel-shaped, eukaryotic vault protein complexes are thought to aid intracellular mRNA trafficking.[18] Capsids, on the other hand, hold genetic material under pressure to protect viral genomes from the extracellular environment before injection into suitable hosts.[17]

In contrast to chaperonins, vaults, and capsids whose interior volumes can span several hundred thousand cubic angstroms or more, the largest synthetic molecular assemblies achieved to date are only on the order of thousands of cubic angstroms in volume. A variety of forces stabilize man-made self-assembled containers. Hydrogen-bonding, for example stabilizes banana- and S-shaped cavitand complexes with internal volumes of up to 800 Å^3 in mesitylene,[47] while similarly stabilized resorcinarene hexamer complexes have been crystalized from aromatic solvents with internal volumes up to 1,375 Å^3.[48] Giant rhombicuboctahedra $M_{24}L_{48}$ metal coordination complexes built from palladium ions and pyridines, on the other hand, have been created with internal volumes as large as 23,000 Å^3.[49] Although hydrogen-bonding and metal coordination can serve as the foundations for building synthetic assemblies with impressive internal volumes, these complexes are not as suitable for biological or environmental applications due to their instability in aqueous environments.

Assemblies stabilized by hydrophobic interactions open up the potential to construct environmentally friendly, biocompatible containers. For example, amphiphilic cog-like molecules have been shown to form hexameric assemblies with inner volumes of 500

$Å^3$,[50] while bowl-shaped deep-cavity Octa-acid cavitands (OA) complexed with hydrophobic guests in water form dimeric capsules with volumes of 650 $Å^3$.[7] Even though these examples are biocompatible, they are smaller than the hydrogen-bond and metal coordination complexes discussed above, limiting their potential use in applications like bioimaging, photodynamic therapy, and drug delivery.[24] More recently, it has been reported that larger tetrameric and hexameric deep-cavity cavitand hydrophobically-driven assemblies depending on the length of the encapsulated alkane guests with internal volumes on the order of 3000 $Å^3$.[19] The deep-cavity cavitand Tetra-endomethyl Octa-acid (TEMOA) used in that study is based on OA where the four hydrogens on the rim of the hydrophobic pocket pointing towards the mouth (*endo*) have been replaced with methyl units (Figure 3.1). Like OA, TEMOA is rendered water soluble as a result of the eight carboxylic acid coating groups decorating the rim and foot of the bowl.

In this paper we investigate the driving forces behind the self-assembly of TEMOA-alkane complexes using molecular simulations. We hypothesize that packing of the alkane guests within the TEMOA complexes provides the primary thermodynamic impetus driving the transitions between the dimeric, tetrameric, and hexameric assemblies. By utilizing free energy simulations of transferring guests ranging in size from C_1 through C_{32} from the gas phase into pre-formed assemblies in water, we demonstrate that the transitions between assembly states are primarily caused by guest-packing. Guest packing coefficients are calculated and shown to indicate the stability of transition points between assembly states.
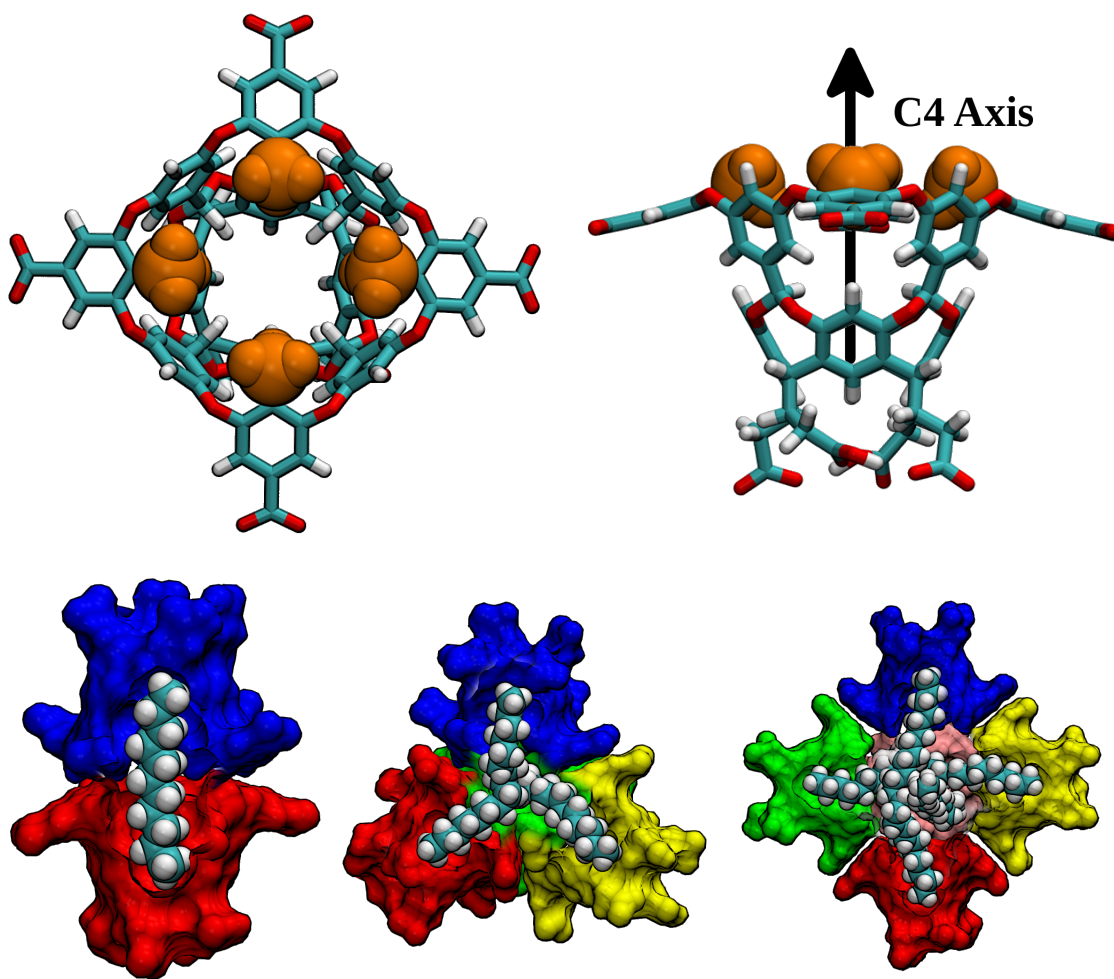
Figure 3.1: TEMOA structures. *Top:* Host monomer in licorice representation with endo methyls highlighted in VDW representation. *Bottom, from left to right:* Dimer, tetramer, and hexamer host-guest complexes. The exterior of one side of the hosts has been removed to allow the reader to see the interior of the complexes.

3.3 Methods

Molecular Dynamics simulations of n-alkane guests encapsulated inside TEMOA dimers, tetramers, hexamers, and octamers in water were performed using GROMACS 5.[31] Water was modeled using the TIP4P/EW force-field.[35] TEMOA Lennard-Jones interactions were modeled using the General Amber Force Field (GAFF)[51] with partial charges calculated using ACPYPE[52] following the AM1-BCC fitting procedure. The net charge on each host at neutral pH was set to -6e, with six GAFF sodium counter cations added for each host. Alkane guests were modeled using the L-OPLS force-field.[32] Lennard-Jones cross interactions were evaluated using Lorentz-Berthelot combining rules.[53] Bonds involving hydrogens for the hosts and guests were constrained using the LINCS algorithm,[54] while bonds within water molecules were held rigid using SETTLE.[55] Host-guest complexes were hydrated by 3,500 to 12,500 waters depending on the size of the assembly (dimer, tetramer, hexamer, or octamer). The number of alkane guests within the TEMOA complex was one, two, three, or four for the dimer, tetramer, hexamer, and octamer complexes, respectively, to maintain a two-to-one host-guest ratio as determined experimentally.[19] Simulations were performed in the isothermal-isobaric ensemble at 25 °C at 1 bar, with the temperature and pressure maintained using the Nosé-Hoover thermostat[38] and Parinello-Rahman barostat.[39] Lennard-Jones interactions were cut-off of 9 Å with a mean-field long-range correction. Long-range electrostatic interactions were evaluated using particle-mesh Ewald summation[10d] with a real-space cut-off beyond 9 Å. The equations-of-motion were integrated with a time step of 2 fs.

Potentials-of-mean force for drawing two TEMOA's together with a single guest alkane in the pocket of one of the host cavitands to form a dimer complex were evaluated using umbrella sampling. Anywhere from 29 to 41 windows separated by 0.5 Å increments with a spring constant of 150 kJ mol^{-1} Å^{-2} were used depending on the length of the guest. Each window was simulated for 5 ns after at least 1 ns of equilibration. In the case of the alkane interacting with a TEMOA or two TEMOA's interacting with one another, the host and guest or both hosts were held fixed along the primary C4 axis of symmetry centered about the host's hydrophobic pocket. The free energy profile was reconstructed from the umbrella windows using the weighted histogram analysis method (WHAM) as implemented in GROMACS' analysis tools.[14]

In order to determine the impact of alkane packing on the stability of the multimeric TEMOA complexes, the free energy of transferring the guests from vacuum into the complex interior was computed. During these simulations the TEMOA complexes were held together using harmonic restraints (spring constant 150 kJ mol^{-1} Å^{-2} and bond length 3.4 Å) between specified carbons on the host rims to hold the complexes together for guest lengths shorter than that necessary to stabilize the assembly. Incremental free energy differences between guests with n to $n+1$ carbon units were evaluated from no included guest to C_{18} for the dimer, C_{26} for the tetramer, and C_{32} for the hexamer and octamer. Free energies were evaluated using a coupling-parameter technique with 21 different states, using the Multi-state Bennett Acceptance Ratio to calculate free energy differences.[56] Each intermediate state was simulated for 5 ns following 1 ns for equilibration.

To evaluate guest packing fractions, the internal volumes of the empty complexes and the van der Waals volumes of the guests was determined by Monte Carlo integration. The internal volume was averaged over 25,000 simulation snapshots, where 100,000 random points were inserted into a box bounding the empty complex, and the volume was determined as the bounding box volume multiplied by the fraction of points inside the internal cavity bound by the van der Waals volume of the cavitands. The van der Waals volume of the n-alkanes was determined using Bondi's values for methyl and methylene groups.[57]

3.4   Results and discussion

Perhaps the simplest capsular assembly process is the formation of a 2:1 TEMOA/guest complex. To quantify the propensity for this dimeric capsule to form, we have evaluated the potential-of-mean force (PMF) between two TEMOA's with their open pockets facing one another and mutually constrained along their C4 axes. In the case in which no guest has been added to the cavitands, we find that the PMF between TEMOA's is attractive (Figure 3.2a). Specifically, the PMF displays a deep attractive well when the two cavitand faces come into contact at a separation of r $\sim$ 4 Å to form a 2:0 complex. Empty dimers are not observed experimentally, however, indicating that the free energy for complex formation along this trajectory is incomplete, missing potentially repulsive contributions like the entropy penalty associated with bringing the two cavitands into mutual alignment along their C4 axes. Nevertheless, the depth of the attractions strongly suggests that hydrophobic interactions between the hydrophobic

cavitand faces is quite strong, and that if the attraction can be made larger, perhaps the complex can be stabilized.

Experimentally, hydrophobic guests which have the length of C_9 and longer cause TEMOA to create and stabilize 2:1 host-guest complexes. We subsequently surmise that by adding guests of increasing length can increase attractions within the assembly. When C_5 is bound to one of the interacting TEMOA's in a 1:1 complex the depth of the attraction at contact increases slightly (Figure 3.2a), indicating that encapsulated guests can play a role in stabilizing the complex. The depth of the attraction is even greater in the case of C_{10}, nearly doubling the depth of the attractive well relative to the empty dimer. Moreover, in the case of C_5 the range of the attraction is comparable to that of the empty complex, although the range increases significantly for C_{10}. In Chapter 2's study of alkanes bound within an OA dimer, we found that the depth of a single cavitand was between 5 to 6 alkane carbon units long. We can then interpret the observed differences between C_5 and C_{10} as resulting from C_{10} bridging between the two TEMOA's, while removing a significant fraction of exposed C_{10} area from water as the second TEMOA caps the guest extended from its partner's pocket into solution. C_5, on the other hand, only plays a weak perturbative role largely sitting only within a single cavitand over the association trajectory. The change in the well-depth from C_5 to C_{10} is considerably greater than that from no guest to C_5, consistent with the experimental observation that guests C_9 and longer are required to stabilize 2:1 complexes.

By increasing the guest length further to C_{15} (Figure 3.2a), we find that while the range of the attraction grows further, the depth of the attractive well is less than that for C_{10}, and the position of the contact minimum is shifted out slightly to larger sep-
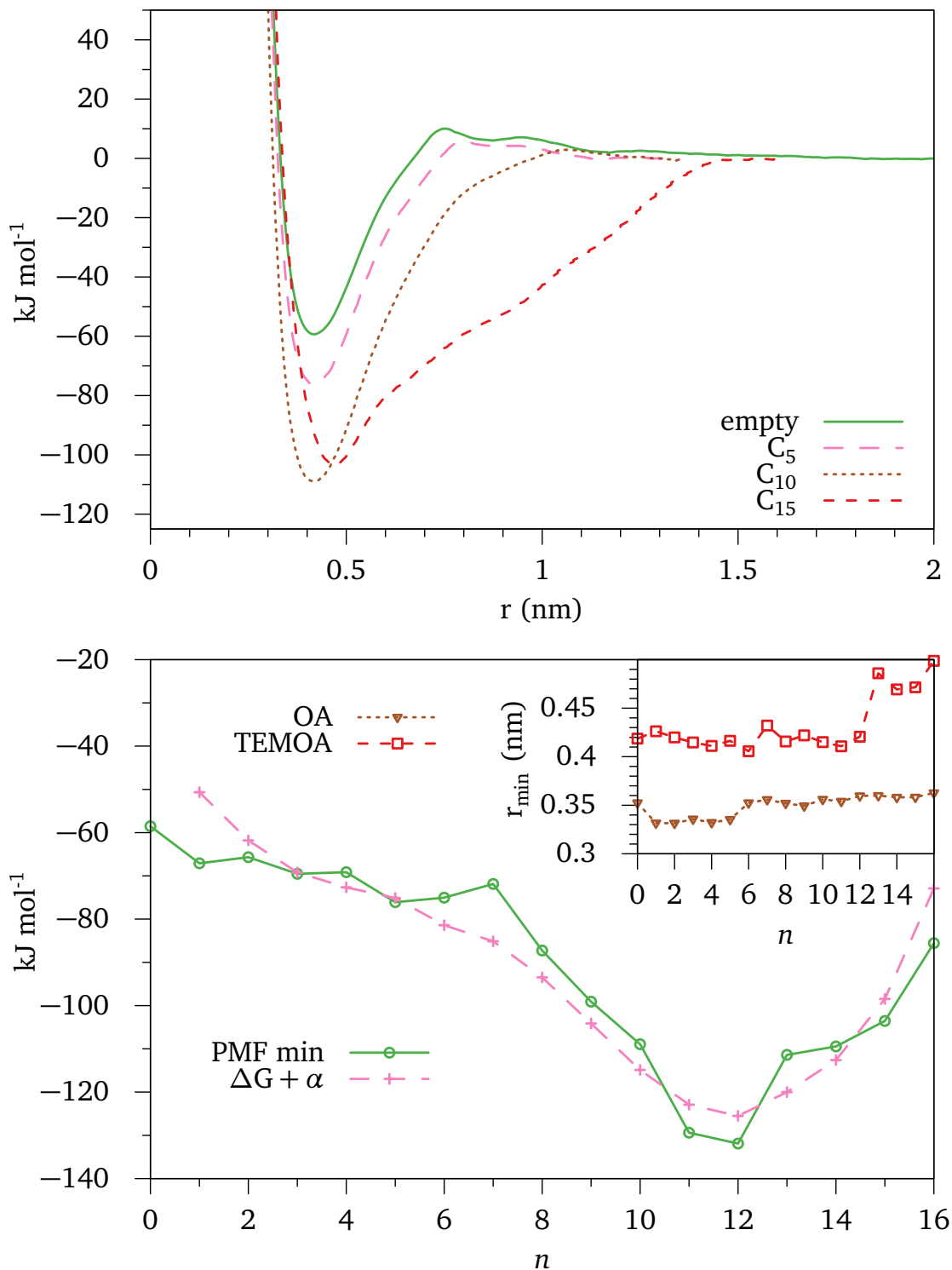
Figure 3.2: *a.* TEMOA-TEMOA potentials of mean force (PMF) with one guest added to the hydrophobic pocket of the host. *b.* Minimum of the PMF's as a function of chain length compared to the vacuum-to-dimer transfer free energies obtained in Figure 3.3. The transfer free energy has been shifted down by 5.05 kJ/mol to facilitate comparison. The inset shows the distance between two hosts at the PMF minimum as a function of chain length.
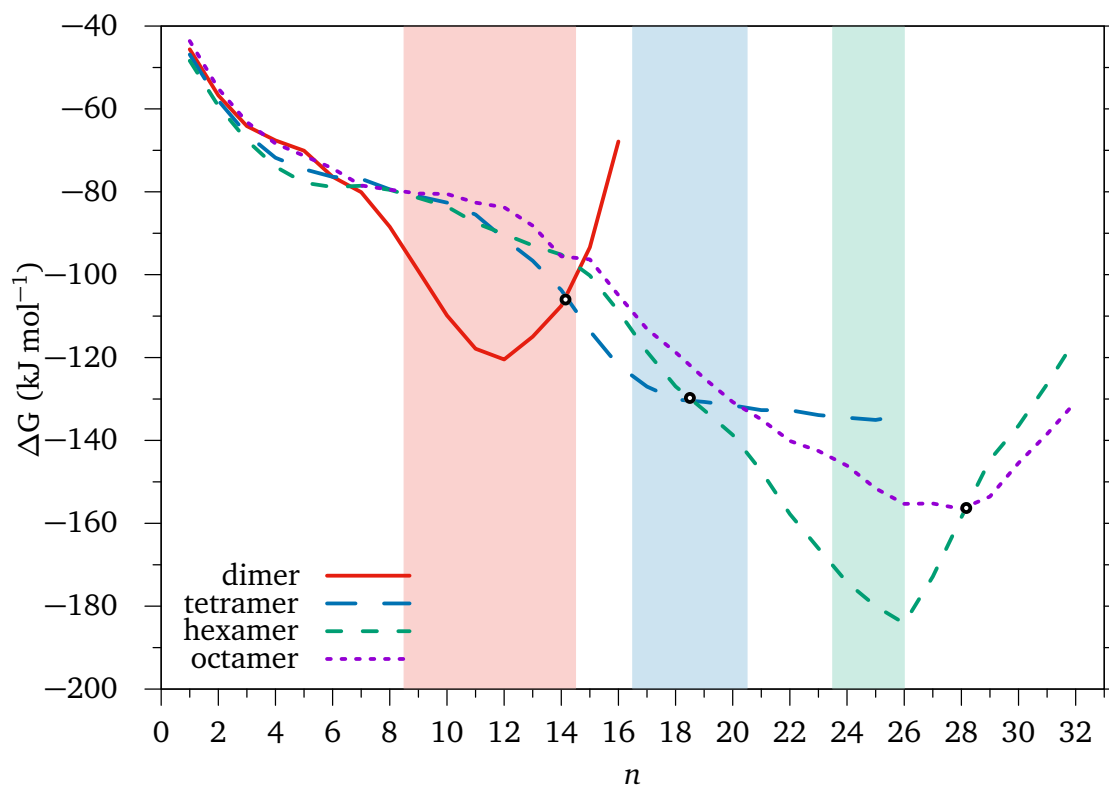
Figure 3.3: Vacuum-to-complex transfer free energies as a function of the length of the guest. Circles indicate crossing points of different assembly states, and shaded regions indicate stable complexes in experiment.

arations. The interior span of an empty OA complex was previously established to be approximately 11 carbons in length—twice the depth of a single cavitand of 5 to 6 carbons. We surmise then that the decrease in the attraction results from the host complex forcing the alkane guest to adopt strained conformations while concurrently forcing the cavitands apart.

If we consider the depth of the 2:1 PMF attractive well as indicative of the complex stability, we can readily compare the 2:1 dimerization guest preferences by plotting the attractive well-depth as a function of the guest length (Figure 3.2b). We find that the depth of the attraction is only weakly decreasing for guests smaller than C_8. The depth of the attraction begins to increase more significantly beyond C_7 and reaches a minimum for C_{12}. These observations correlate with the experimental observation that TEMOA dimers are stable for guests ranging in length between C_9 and C_{14}. Indeed, we observe that C_8 and C_{16} have a comparable attractive strength, while those with lengths in between have a deeper attraction. This approximately identifies the penalty required to assemble the TEMOA dimer as comparable to the free energy difference between the empty complex and that including C_8 or C_{16}. The position of the PMF minimum as a function of guest length is reported in the inset to Figure 3.2b. Here we find that the minimum is approximately constant up to guests as long as C_{12}. Beyond C_{12}, the minimum shifts out to greater separations, indicating that longer guests begin to push the hosts apart and to expose the hydrophobic surface of the guest. This increase in the separation is consistent with the idea that the internal confinement is straining the guest to force the capsule open. In contrast, if we evaluate the position of the PMF minimum for forming a 2:1 OA/alkane complex we find that the minimum is practically

independent of the guest length even up to C_{16} (Figure 3.2b, inset). Previously, in Chapter 2, we found that in OA complexes, the guests are permitted to explore a wide range of conformations to fit within the capsule. Our results here suggest that the inward pointing methyl units limit the conformations available to the guests in TEMOA, making them more rigid at the cavitand mouth and thereby help push the hosts apart. We note that the minimum position in TEMOA is slightly shifted to greater separations than OA by 0.7 Å as a result of the protrusion of the methyl units above the rim of the host's pocket.

Although the PMF between two TEMOA's does provide insight into the stability of dimer capsules, the formation of multimeric capsules (*e.g.*, tetramers and hexamers) would require a more complex, multi-dimensional reaction trajectory. To circumvent this difficulty, we hypothesize that the packing of guest chains into the complex interior is the dominant contribution stabilizing distinct multimers. We quantify this contribution by the free energy of transferring guests from vacuum into a pre-formed assembly in water with the appropriate host-guest stoichiometry. The transfer free energy of an alkane into the interior of a TEMOA dimer is reported as a function of the length of the alkane in Figure 3.3. Qualitatively this transfer free energy resembles the dimerization PMF minima, with the most favorable free energies lying in the range of guests from C_9 to C_{15}.

The differences between these two free energy profiles can be illuminated by considering the thermodynamic cycle for assembling the dimeric complex (Figure 3.4). In this cycle we consider two potential paths for assembling the dimer starting from disassociated cavitands with the guest initially in the gas phase (A to E in Figure 3.4).

Along the first path the guest is first transferred from the gas phase into solution (A to B), the guest is subsequently pulled into one of the cavitands (B to C), and finally an empty cavitand and a cavitand with a bound guest are brought together to form the 2:1 host-guest capsule (C to E). Along the second path the two empty cavitands are dimerized in solution to make an empty capsule (A to D), and subsequently the guest is transferred from the gas phase into the empty capsule (D to E). The dimerization free energies reported in Figure 3.2 correspond to the free energy difference from C to E, while the transfer free energies reported in Figure 3.3 correspond to the difference from D to E. A semi-quantitative comparison between the transfer free energy and the 2:1 PMF minima can be achieved by shifting the dimer transfer free energies by -5.05 kJ mol^{-1} (Figure 3.2b), indicating that the variation of the PMF minima is largely influenced by the transfer free energy despite the differences in the paths to assemble these free energies contribute (Figure 3.4). Moreover, in trying to determine how the guest length impacts the relative stability of the multimeric complexes, the second path may be more appropriate since all the chain length-dependent effects are captured in the transfer process (D to E), while additional chain length-dependent contributions beyond the PMF between cavitands enter into the complex stability from steps A to B and B to C along the first path. We, therefore, focus below on the free energy of transfer as a determinant for the switching between multimeric assembly states

Following our contention that the guest transfer into an empty complex can be used to distinguish the relative stability of multimeric assemblies, we compare the free energies of transferring alkanes into the dimeric, tetrameric, and hexameric complexes in Figure 3.3. We note that while the number of alkanes transferred into each capsule
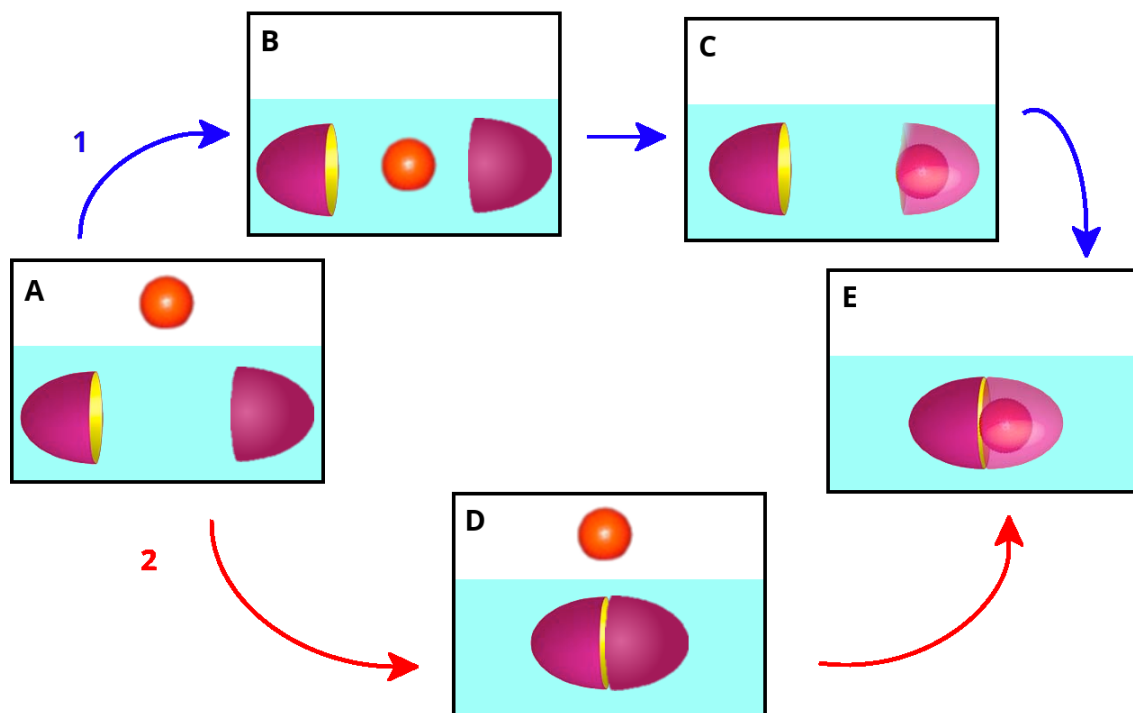
Figure 3.4: Schematic of thermodynamic paths from a state with one guest in the gas phase and two TEMOA's in solvent (A) to a state where the guest is in the interior of a dimeric TEMOA complex (E). Path 1 follows the process of first transferring the guest into water (A to B), then allowing the guest to bind to one TEMOA (B to C), and finally having the other TEMOA complex with that assembly (C to E, cf. Figure 3.2a). Path 2 follows the process of first assembling the two empty TEMOA's (A to D, cf. Figure 3.2a), and then transferring the guest from the gas phase to the interior of the assembly (D to E, cf. Figure 3.4).

corresponds to a 2-to-1 stoichiometric ratio, the free energies are reported on a per alkane basis to affect an appropriate comparison. Although the transfer free energies into each capsule are similar for chains shorter then C_7, they diverge beyond this guest length. The packing free energy within the dimers drops below those of the larger capsules over the range of lengths from C_8 to C_{14}, suggesting packing favors dimers for the shorter alkanes studied. Between C_{14} and C_{15}, the transfer free energy of alkanes into the 4:2 tetramer becomes the lowest. The transfer free energy into the tetramers is the minimum between C_{15} and C_{18}. Subsequently, between the C_{18} and C_{19} the transfer free energy is lowest for the 6:3 hexamer.

The succession of stable assembly states described above follows that observed experimentally. Comparing the crossing points against the experimental stability ranges, however, we find that our simulations slightly under predict observation. For example, experimentally dimers are observed for guests ranging in size from C_9 to C_{14}, while tetramers are observed between C_{17} and C_{20}. No specific TEMOA aggregation numbers were experimentally observed for C_{15} and C_{16}, presumably because both assembly states coexist for these guest lengths. In this case, we might expect the transfer free energy crossing between dimers and tetramers to fall between C_{15} and C_{16}, while here our simulations find it lies between C_{14} and C_{15}. Similarly, tetramers are not distinctly resolved experimentally beyond C_{20}, while hexamers are observed for C_{24} to C_{26}. Guests longer that C_{26} were not examined experimentally, so the upper stability bound for the hexamers is not known yet. From our simulations, we predict the transfer free energy between tetramers and hexamers crosses between C_{18} and C_{19}, rather than near C_{22} as suggested by experiment. One of the main reasons we may anticipate that our

simulations transition points occur at guests shorter than the experimental transition points is the fact that we held the assemblies fixed as we grew the alkanes from nothing to their full lengths by adding successive CH2 units when evaluating the transfer free energies. This was done in order to stabilize the capsules during the incremental transfer process, which are expected to fall apart for the shortest alkanes considered. This rigidifies the complexes by constraining the TEMOA's and minimizes shape fluctuations, thereby shifting the crossing points to shorter alkane lengths. In addition, we have only evaluated a portion of the total assembly free energy, neglecting contributions like the empty complex assembly free energies (*e.g.*, A to D in Figure 3.4), which could potentially push the transfer free energy crossings to different lengths. Nevertheless, the agreement between simulation and experiment for the order of assembly states and the minimal perturbation in the crossing points strongly suggests that the transfer free energy is the dominant contribution to the stability of the TEMOA assemblies observed.

Following our attribution of alkane packing within the internal complex spaces as being a driver for assembly state transitions, we wish to examine the correlation between the transitions and the packing fractions of guests within the complexes. In Table 3.1 we report the internalized volume for the dimeric, tetrameric, and hexameric complexes. It can be seen that increasing the TEMOA aggregation number results in non-additive increases in the internalized volume. This results from the fact that the complexes internalize an increasing amount of volume within the central space between the hosts as the aggregation number increases (Figure 3.5). Subtracting twice the internalized volume of a dimer from that of the tetramer, we find the tetramer holds 748 Å^3 more space (Table 3.1). Similarly, subtracting three times the internalized dimer vol-

complex	$V_{internal}$ (Å³)	$V_{central}$ (Å³)
dimer	587	
tetramer	1922	748
hexamer	4020	2259
octamer	3353	1619

Table 3.1: Internal and central volumes of each TEMOA complexes based on Monte Carlo integration. The central volume of the tetramer is determined by subtracting the two times the volume of the dimer, and the central volume of the empty hexamer is determined by subtracting three times the volume of the dimer. Figure 3.5 is a graphical representation of these. Due to the flexibility of the assembly, the octamer volume calculations below expected values.

ume from that of the hexamer, we find the tetramer holds 2259 Å³ more space. These

increasing volumes allow the tetramer to accommodate longer guests than the dimer

and the hexamer to accommodate longer guests that the tetramer.

The guest packing fraction within the complex is determined by the ratio of the van

der Waals volume (V_{vdW}) of the N bound guests divided by the complex internal volume

$$\eta = \frac{N V_{vdW}}{V_{internal}}.$$ (3.1)

We plot the packing fraction against the guest chain lengths for the dimers, tetramers,

and hexamers in Figure 3.6. Clearly the larger complexes have more free space for

the guests to pack in, as indicated by the lower packing fractions at fixed guest lengths

with increasing cavitand aggregation numbers. If we consider the guest lengths for

which each complex is first stable, identified by the crossing points in Figure 3.3 for the

tetramer and hexamer, and assumed to be 8.5 for the dimer as determined experimen-

tally and confirmed from our simulations, we find that this critical packing fraction to
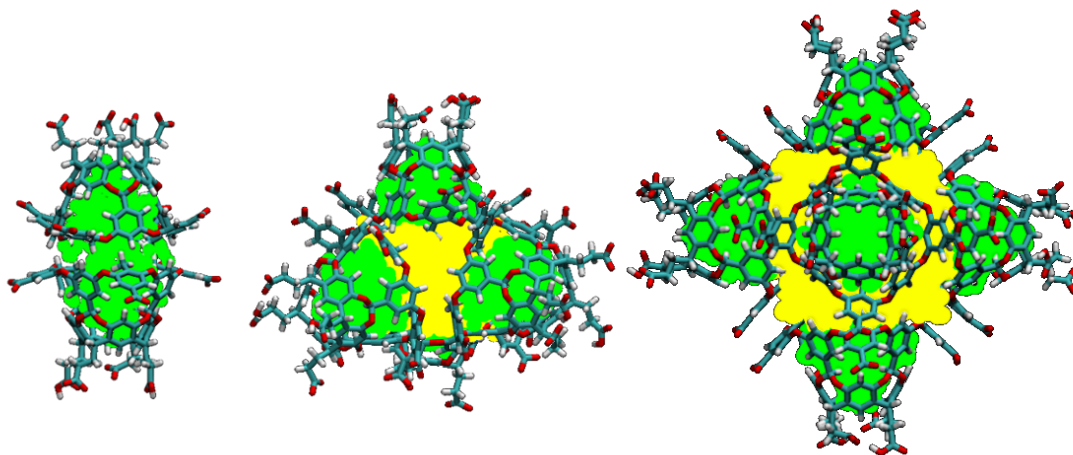
Figure 3.5: Graphical representation of the internal volumes for the dimer, tetramer, and hexamer complexes. Green indicates volumes within individual hosts, and yellow indicates the remaining volume, corresponding with V_{central} in Table 3.1.

stabilize a complex is approximately $\eta = 0.257$ and is nearly independent of the complex aggregation number. If we attempt to determine the maximum packing fraction for which a complex is stable, however, we do not observe a similarly good correlation. For example, at the point the tetramer becomes stable, the alkane packing fraction within the dimer is approximately $\eta = 0.44$. For the packing fraction within the tetramer to reach this level, the alkane chain length must reach C_{25}, which is beyond the point at which the hexamer becomes stable. Based on this correlation, we conclude that the stability of a complex necessitates a minimal packing fraction.

To this point we have examined the range of experimentally observed assemblies. The upper stability of the 6:3 TEMOA complex as a function of the guest length, however, has not yet been established. Given the semi-quantitative agreement between our calculations of the guest transfer free energy crossings between assemblies of differing size and the experimental transitions between complex polymorphs, we propose that
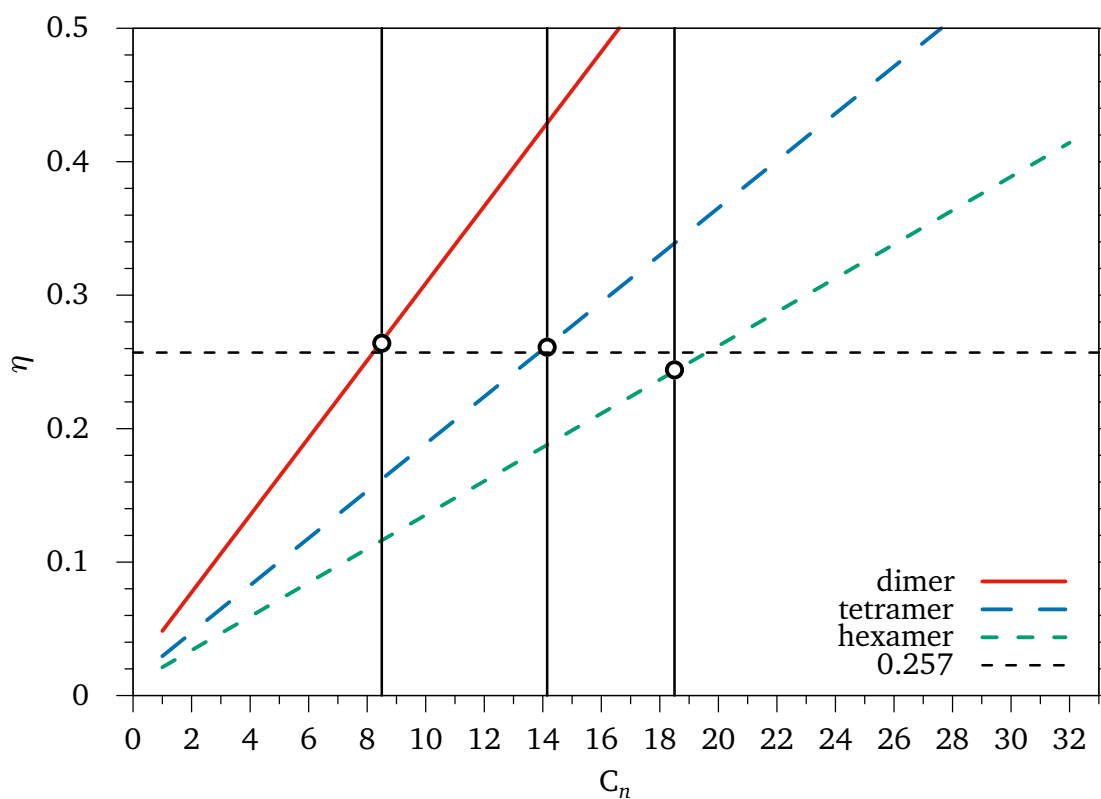
Figure 3.6: Packing fractions as a function of chain length. Circles indicate assembly-state transition points from Figure 3.4 for the tetramer and hexamer. The dimer transition point is assumed to be 8.5 based on previous experimental observations.

our simulations can provide insight into the stability bounds for the TEMOA hexamer. Following the progression of complex structures observed, it is natural to suppose that the next most stable complex with increasing alkane length would be an 8:4 octameric assembly. To this end, we have performed guest transfer free energy calculations into an empty octamer in water. This octamer is assumed to adopt the shape of a Platonic octahedron, with each host's binding pocket entrance facing inwards into one of the eight faces of the polyhedron. As for the smaller assemblies, we stabilized the complex structure using harmonic restraints between hosts that share an edge between the octahedron faces.

The free energies evaluated from the guest transfer simulations into the octamer are presented in Figure 3.3. Over the range of experimentally investigated guest lengths up to C_{26}, the free energy of transfer into the octamer is greater than of that into smaller assemblies, consistent with the octameric complexes not being observed. Based on our calculations, however, we predict that between C_{28} and C_{29} the free energy of transfer into the octamer becomes lower than of that into the hexamer, beyond which we may anticipate the octamer is stable. Following the tendency of our computed transition points to occur at guest lengths slightly shorter than that observed experimentally, we anticipate that the hexamer-to-octamer transition may experimentally occur closer to guests of the size of C_{32}. Our calculation suggests the experimental procedure would need to be altered, since, when the TEMOA/alkane complexes are prepared experimentally, they are held at a temperature of 60 °C to ensure the hydrocarbon phase in contact with the aqueous TEMOA solution is melted. Specifically, the melting point of C_{32} is 70 °C, requiring a higher preparation temperature. Subsequently, if longer guests

are needed to stabilize even larger complexes, even higher temperatures would be required and would form an upper bound—bounded by the boiling point of water—to the largest assemblies that could be prepared.

If we consider the proposed correlation between guest packing and the stability of the octameric complex, we are immediately confronted with a difficulty. Specifically the internalized volume of the empty octamer of 3353 Å^3 falls well below that of the hexamer (Table 3.1). In difference to the dimer, tetramer, and hexamer, the octamer is not rigid when stabilized by harmonic edge restraints, such that the octamer's volume swells with guest length. Moreover, the precise determination of the octamer volume is complicated by the gaps between hosts at the octahedron vertices, which are large enough to permit partial guest protrusion, in difference to the smaller complexes. We hypothesize that this results in the minimum guest transfer free energy within the octamer near C_{28} being greater than the hexamer minimum near C_{26} (Figure 3.3). These difficulties do not permit a comparison of the stability of the octameric complex on the packing fraction correlation (Figure 3.6), which we surmise may best apply to rigid complexes.

3.5   Conclusion

Using computer simulations we have demonstrated that guest packing provides a controlling thermodynamic determinant for selecting the dominant TEMOA/alkane complex aggregation state. Specifically, we found that the free energies of transferring alkane guests into pre-formed multimeric TEMOA complexes in water with a 2-to-1 host-guest stoichiometry exhibits a sequence of free energy minima with increasing

guest length indicating that dimeric complexes are favored for guests from C_9 to C_{14}, tetrameric complexes are favored from C_{15} to C_{18}, hexameric complexes are favored from C_{19} to C_{28}, and octameric complexes are favored for guests C_{29} and longer. The succession of dimer to tetramer to hexamer complex stabilities follows that experimental observation, although our simulations predict the transitions between multimers occur at slightly shorter guest lengths than experiment. We believe the discrepancies between simulation and experiment potentially reflect the effects of necessary imposed constraints on the host geometry that rigidify the complex structures to be able to evaluate transfer free energies, as well as the neglect of contributions like inter-host assembly free energies of empty multimeric complexes. Nevertheless, our results strongly indicate that guest packing within complex interiors largely determines assembly stability. Further, our simulations predict that octameric complexes may be experimentally achievable with even longer guests; however, the temperatures required to melt the alkane guests may make their preparation impractical. Complex stability as a function of guest length from dimers to hexamers is well correlated with a critical guest packing fraction of $\eta \approx 0.255$, consistent with our hypothesis that guest packing is a controlling determinant for the assembly aggregation state. This empirical correlation fails, however, in the case of the octamer, which we believe reflects the lack of complex rigidity compared to the dimeric, tetrameric, and hexameric assemblies. In conclusion, our simulations have illustrated the important role of guest packing in determining the ultimate stability host-guest assemblies, and points to the potential for using simulations as a tool for the design of new supramolecular complexes.

Chapter 4

HYDRATION CHARACTERISTICS OF THE INTERIOR OF DEEP-CAVITY CAVITANDS OCTA-ACID AND TEMOA

4.1 Summary

Concave surfaces have been demonstrated both experimentally and computationally to bring about enthalpy-driven evaporation within confined areas. Here we report on molecular simulations used to study deep-cavity cavitands Octa-acid (OA) and Tetra-endoymethyl Octa-acid (TEMOA), which only differ in the addition of methyl groups to the *endo* positions of the binding pocket entrance's rim. These methyls effectively narrow the entrance to the binding pocket of the deep-cavity cavitand. We demonstrate that waters within TEMOA exhibit a two-phase behavior not seen within OA. This behavior is also not seen in Tetra-exomethyl Octa-acid (TExMOA), identical to TEMOA except with the methyl groups pointing up in the *exo* positions instead of the pointing inward, toward the C4 axis. Pressure and temperature are also used to tune the wetting behavior of TEMOA. These results are shown to be connected to the partial molar volume of each deep-cavity cavitand, with TEMOA being "larger" than expected based on structure alone, due to its accessing the empty state much more frequently than

OA. Lastly we demonstrate how TEMOA's ability to dewet its interior affects its binding strength.

4.2 Introduction

The classical hydrophobic effect—the low solubility of hydrophobes in water due to entropic penalties—drives many processes. Protein folding and the self-assembly of micelles are among the many examples that utilize water-mediated interactions driven by the hydrophobic effect.[58] In contrast, many processes are thought to be driven by the "non-classical" hydrophobic effect—the low solubility of water in a non-polar medium or region. This effect often occurs when water is in a confined space within a hydrophobic region.[59] The non-classical hydrophobic effect is especially prevalent in biological systems which contain nanoconfined areas. Perhaps the most prevalent example is water within the cell. Water's space is limited within the cell due to macromolecules which crowd the cell's interior and are separated from each other on the order of a nanometer.[60]

The "non-classical" hydrophobic effect drives these interactions largely in part because nanoscopic hydrophobic confinement brings about changes to the thermodynamic properties of water, especially in regard to phase transitions.[61] Evaporation of water from within a confined space can occur at thermodynamic conditions where bulk water would be in the liquid phase.[62] For example, concave surfaces have been shown to bring about enthalpy-driven dewetting when a spherical hydrophobe is brought near to it in computer simulations.[63] Likewise, ligands often displace water in protein binding sites before binding. In rare cases, spontaenous dewetting can also occur without a

hydrophobe in proximity. For example, β-lactoglobulin's large binding cavity has been found to be completely dry.[64] Additionally, simulations using test particle insertion have shown that at standard conditions water exhibits a two-state behavior when it is within nonpolar graphene-like spherical cavities, as well as within fullerenes, when the diameter of such cavities is larger than 1 nm; when such cavities are smaller than 1 nm, their interiors have been shown to dry at the same conditions.[65] Isoreticular metal-organic frameworks have also been shown to be able to change the critical point of non-water pure substances such as benzene and xylene, showing that the change of thermodynamic properties and phase behavior comes from the confinement of the substance.[66]

The thermodynamics of evaporation of confined water can be controlled by changing the excess chemical potential of water on the inside and/or outside of the confined area.[67] For example, pressure and temperature can be used to tune the evaporation of water between two hydrophobic plates. Simulations have shown that under standard conditions when water is confined between two hydrophobic plates closer than some critical drying distance, capillary evaporation will occur. At higher pressures[68] and lower temperatures[69] this critical drying distance becomes shorter, indicating that the liquid phase of the interior water has become more stable at those thermodynamic state points. Carbon nanotubes are another example of such tunability, but in regard to a solid/liquid transition.[70]

Simple hydrophobic concave synthetic molecules allow for the experimental and computational study of the effect of hydrophobic confinement on water. One such concave synthetic receptor is Octa-acid (OA), a deep-cavity cavitand, which is a bowl-shaped molecule. OA is made up of three rows of benzene rings and has a depth of

0.7 nm. Hydrophobic guests bind to OA when in the presence of water, and can form assemblies with 1:1, 2:2, and 2:1 host-to-guest ratios.[24c] OA has been used as a reaction vessel[26] and a hydrocarbon gas separator.[27]
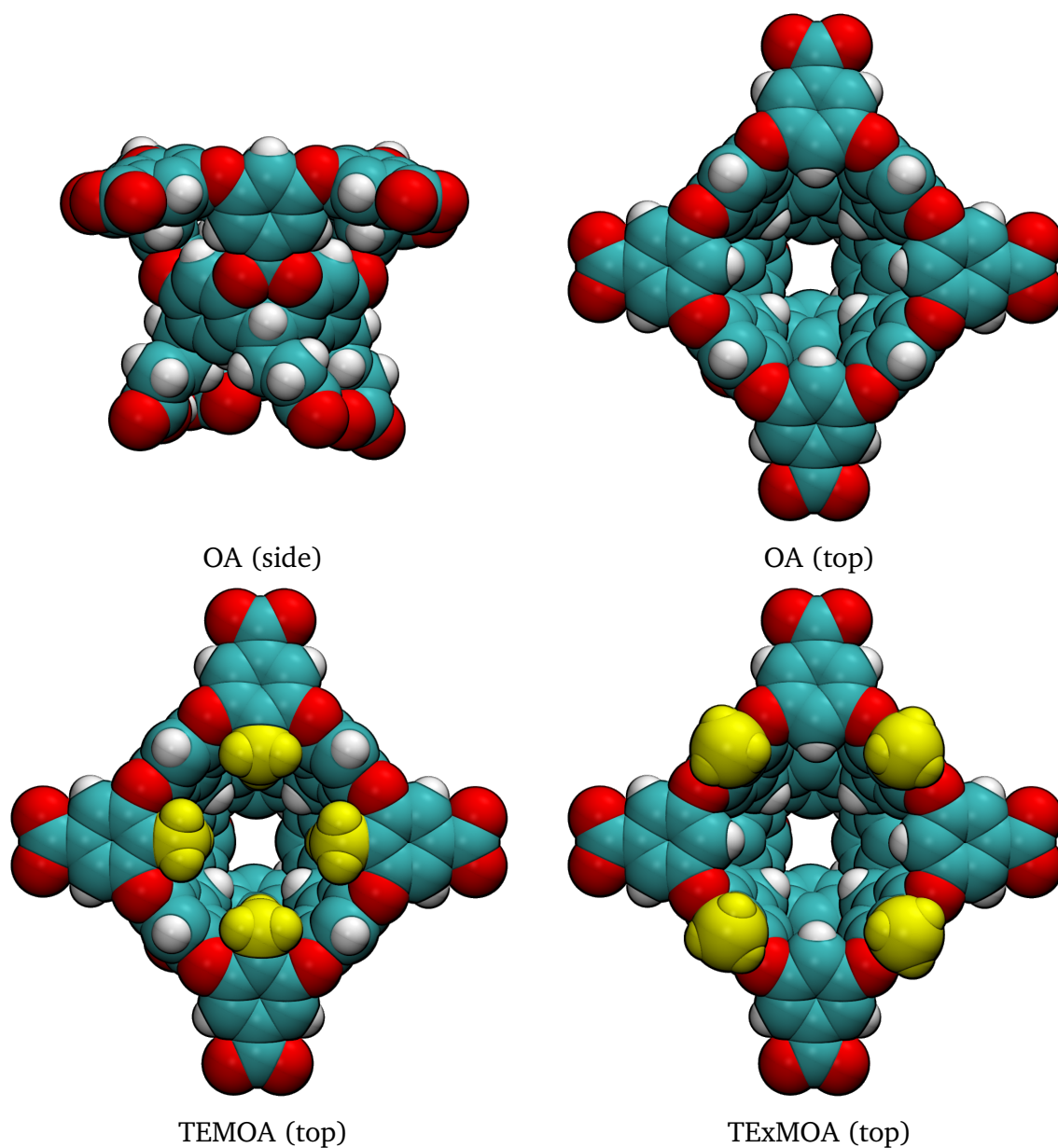
Figure 4.1: Side view of OA (top left), top view of OA (top right), top view of TEMOA with *endo* methyls highlighted (bottom left), top view of TExMOA with *exo* methyls highlighted (bottom right).

Previous simulations have shown that OA evacuates its interior waters when a hydrophobe is nearby; this indicates that the confined water within OA is near a vapor-liquid transition[34] consistent with other studies of concave binding sites.[63] This suggests that a modification to the entrance to OA's binding pocket might cause spontaneous dewetting inside. Tetra-endomethyl Octa-acid (TEMOA) is identical to OA except that methyls are in the *endo* positions instead of hydrogens (indicated in yellow in Figure 4.1). Unlike OA, TEMOA forms not only dimers, but also multimers.[19] This modification of adding methyls to the rim of the deep-cavity cavitand narrows the entrance to the cavitand. Another derivative of OA is Tetra-exomethyl Octa-acid (TExMOA), which is identical to OA except that methyls are in the *exo* positions on the rim (also indicated in yellow in Figure 4.1). In this case the entrance to the interior does not narrow.

These deep-cavity cavitands give us an opportunity to study the effect of confinement on water using real-world synthetic molecules. In this study we show that the modification of OA to TEMOA leads to spontaneous evaporation of TEMOA's interior waters at standard conditions, and that TExMOA does not exhibit such behavior. We use simulations to show that by changing the pressure and temperature, we can tune the evaporation of TEMOA's interior. We also show that the partial molar volume of TEMOA is larger than can be explained by just its additional methyl groups alone, and that this can serve as a bridge to experimental observations. Unlike previous studies on the evaporation of water within a hydrophobic cavity, we directly observe such evaporation computationally in a real synthetic system.

4.3 Methods

We performed Molecular Dynamics simulations of OA, TEMOA, and TExMOA in water; of pure water; of each cavitand with an *n*-hexane guest in water; and of a *n*-hexane solute in water. ACPYPE[52] was used to calculate partial charges for each cavitand by using an AM1-BCC fitting procedure, and each cavitand was modeled using the General Amber Force Field (GAFF).[51] The net charge on each cavitand was assumed to be -6*e*, following previous work on OA.[34] In order to make each system charge-neutral, six GAFF sodium ions per cavitand host were added.[51] *n*-hexane was modeled using L-OPLS,[32] and TIP4PEW[35] was used for all water molecules. All bonds involving hydrogens were turned into constraints using LINCS,[54] and SETTLE[71] was used to keep bonds within water molecules rigid. Initial solute configurations were hydrated with between 2,135 and 2,994 waters.

These systems were simulated in the isothermal-isobaric ensemble at 298.15 K and 1 bar using the Nosé-Hoover thermostat[38] and the Parrinello-Rahman[39] barostat, respectively. Additionally, simulations of these systems were performed at 500 bar, 1000 bar, 1500 bar, and 2000 bar at standard temperature. Temperature was also varied at 278.15 K, 298.15 K, 343.15 K, and 368.15 K at standard pressure. Short-range van der Waals interactions were cut off beyond 9 Å where a long-range dispersion correction was used for energy and pressure. Real-space electrostatic interactions were also cut off at 9 Å and Particle-mesh Ewald summation[10d] was used to evaluate long-range electrostatic interactions beyond the cutoff. Each system was simulated for 100 ns after at

least 1 ns of equilibration. Trajectories of each system were saved every 4 ps, resulting in 25,000 frames being saved for each thermodynamic state.

For systems containing a single deep-cavity cavitand in water, the distribution of waters inside the cavitand was calculated. In order to determine whether or a not a water molecule was inside the cavitand, several atoms on each side of the cavitand were selected and a plane was fit using singular value decomposition. Six planes in total were used for each deep-cavity cavitand, with the plane at the entrance to the binding pocket being defined by the oxygens on the rim of the cavitand, following previous work.[34] A block bootstrap method was used to calculate the uncertainties in the average number of waters inside each cavitand, as well as within each probability distribution. Each uncertainty analysis used 200 bootstrap iterations and 5 blocks.

In order to determine the partial molar volume of each solute, the average volume of each simulation containing a cavitand was computed, as well as the average volume of a simulation of pure water. The partial molar volume was then calculated as:

$$\bar{V}_{solute(s)} = \langle V \rangle_{solute(s)+wat.} - \langle V \rangle_{wat.} \tag{4.1}$$

where $\langle V \rangle_{wat.}$ was scaled to match the same number of waters in the simulations with a cavitand. Additionally the volume of association between each cavitand and a hexane was calculated as follows:

$$\bar{V}_{assoc.} = \bar{V}_{cav.+hex.} - (\bar{V}_{cav.} + \bar{V}_{hex.}) \tag{4.2}$$

Furthermore, the effect of pressure on the Gibbs free energy change during association was computed according to the following relation:

$$\bar{V}_{assoc.} = \frac{\delta \Delta \bar{G}_{assoc.}}{\delta p} \qquad (4.3)$$

Thus, by proper integration of the volume of association, the change of association Gibbs free energy can be obtained. In order to verify this prediction, the potential of mean force of one cavitand (OA or TEMOA) and one n-hexane guest was evaluated by umbrella sampling. Between 30 and 40 overlapped windows at 0.5 Å-increments and a spring constant of 150 kJ mol^{-1} Å^2 were used. The center of the top of the cavitand and the center of the hexane guest were used as reference points. The center of the hexane is defined as the center of the bond connecting its two center-most carbon atoms. The center of the hexane was fixed along the z-axis of simulation box, and the C4 axis of the cavitand was also restrained along to the z-axis. The pulling process was performed from the deep-cavity cavitand to bulk solvent along z-axis. WHAM[14] was used to unbias the windows.

4.4 Results and discussion

In order to understand the hydration characteristics of the binding pockets of the deep-cavity cavitands, we directly measure the number of waters residing inside the pocket throughout each simulation. The distributions of interior waters for each cavitand show that there are distinct differences between OA and TEMOA, as well as between TExMOA and TEMOA (Figure 4.2). At 1 bar, OA and TExMOA both show a Gaussian-like distri-

bution, with TExMOA's curve slightly shifted to the left of OA's. The peak of both curves occurs at 4 waters, matching with a probability of approximately 38%. At 1 bar there are on average 4.072 ± 0.029 within OA and 3.970 ± 0.026 within TExMOA. On average there are 0.101 ± 0.055 less waters inside TExMOA than OA. We attribute this very small effect to the addition of methyls groups in the *exo* positions on the rim.

In contrast, TEMOA has a bimodal distribution at standard conditions. Unlike both OA and TExMOA, TEMOA is empty 35% of the time. On average, there are 2.165 ± 0.133 waters inside of TEMOA, which is 1.907 ± 0.162 fewer than those within OA. The addition of methyls to the *endo* positions has a much more dramatic effect on the hydration characteristics of the deep-cavity cavitand interior than adding them to the *exo* positions as in TExMOA. The main structural difference between TEMOA and TExMOA is that the methyls in TEMOA are pointing inward toward the C4 axis, narrowing the diameter of the entrance to the binding pocket. Perhaps this can be thought of in terms of previous simulations of hydrophobic plates being brought near each other. As the plates are brought closer together past a critical drying distance, capillary evaporation occurs.[68] Here we have effectively shortened the "critical drying distance" of the entrance of the deep-cavity cavitand by narrowing the entrance.

This drying effect can be tuned by manipulating the thermodynamic conditions of the simulation. By increasing the pressure to 2000 bar, the water occupancy state of TEMOA can be changed (Figure 4.2, bottom). As one would expect, there are slightly more waters inside both of OA and TExMOA when compared to TEMOA. However, the bimodal distribution seen at 1 bar for TEMOA has vanished, and its distribution is now nearly identical to the other two deep-cavity cavitands'. At this pressure there are 4.821
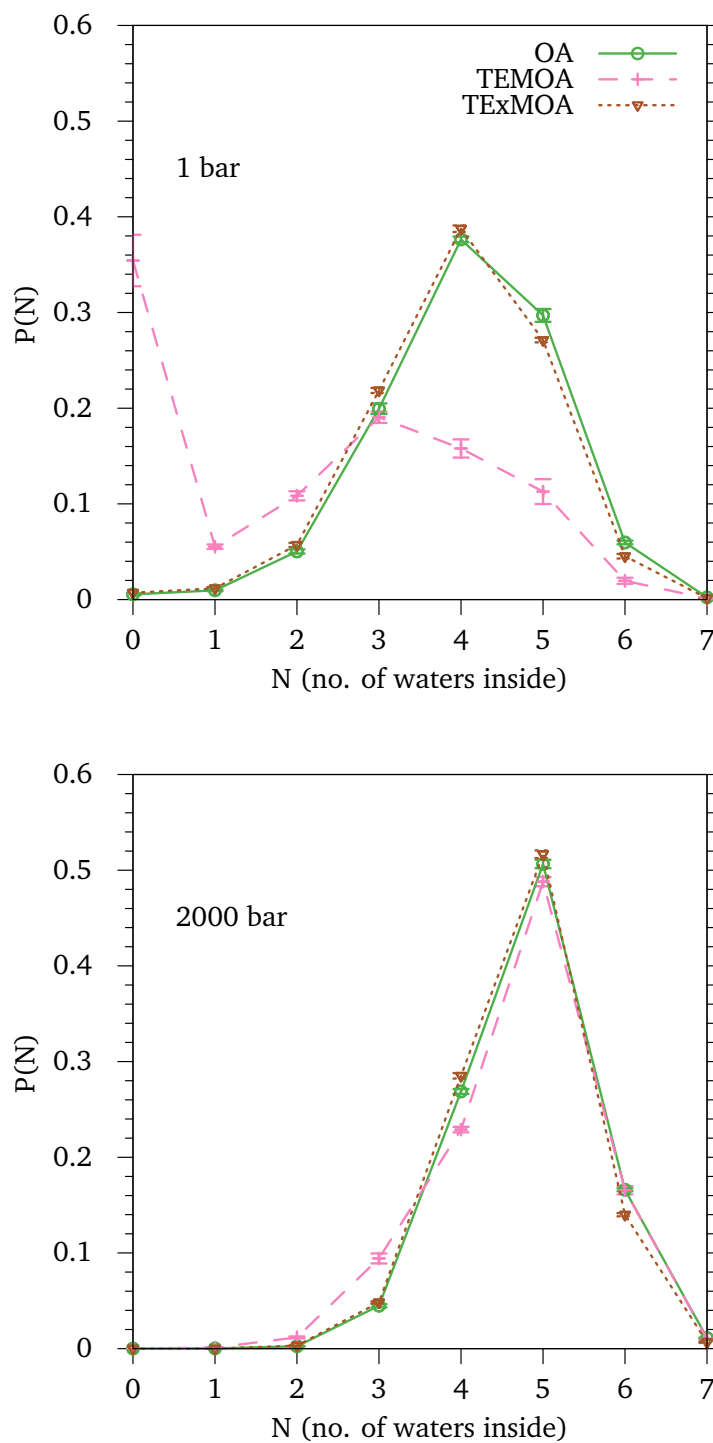
Figure 4.2: The probability distribution of the number of waters inside each cavitand for 1 bar (top) and 2000 bar (bottom).

± 0.025, 4.729 ± 0.029, and 4.762 ± 0.024 waters on average inside of OA, TEMOA, and TExMOA, respectively. The elimination of dewetting for TEMOA at high pressures is consistent with the fact that the properties of water near hydrophobic surfaces become nearly the same as that of water near hydrophilic surfaces at 2 kbars.[68]

Temperature can also control the water occupancy state of TEMOA. At 278.15 K, the tendency for TEMOA to dewet its interior decreases (Figure 4.3). TEMOA is empty of water 23% of the time with an average of 2.945 ± 0.164 waters inside at 278.15 K. OA and TExMOA also saw a similar increase in the number of waters inside their interiors with averages of 4.243 ± 0.026 and 4.206 ± 0.024, respectively. The supporting information in Appendix D details how higher temperatures cause TEMOA to dewet at higher percentages. These findings are consistent with the tendency for confined water to become more stable in the liquid phase relative to the vapor phase upon cooling.[69] Electric fields also have an effect on the interior waters; the interested reader is directed to Appendix D.

In order to bridge our computational observations of TEMOA being dry a large percentage of the time with experiment, we first establish the connection between water occupancy state and partial molar volume differences. The change in partial molar volume ($\Delta\bar{V}$) between TEMOA and OA is correlated with the difference in the number of interior waters ($\Delta\langle N \rangle$) between these systems at the various pressures sampled (Figure 4.4). By increasing the pressure of the system we are able to push water molecules inside of TEMOA and bring its partial molar volume closer to that of TExMOA and OA. In contrast, $\Delta\bar{V}$ only slightly changes with pressure for TExMOA, and $\Delta\langle N \rangle$ remains nearly constant for the same system. $\Delta\bar{V}$ for TEMOA shows a much more drastic change as
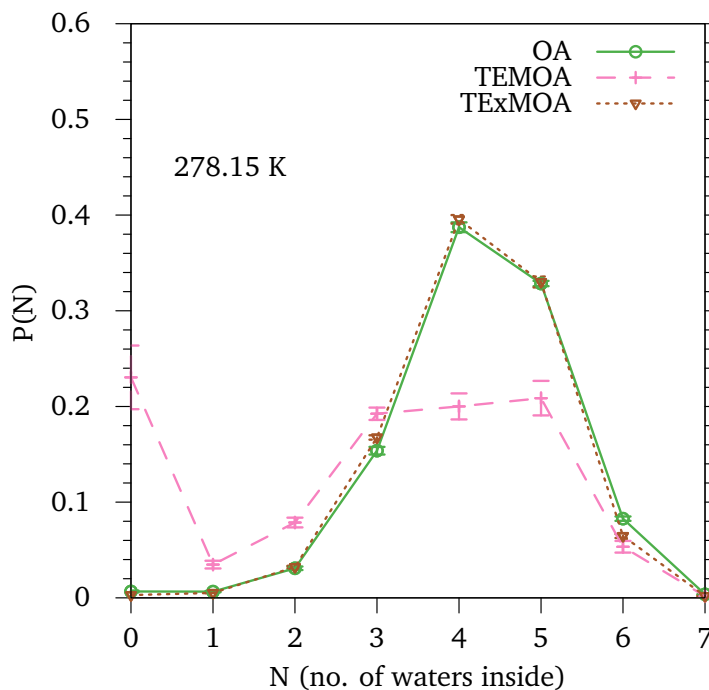
Figure 4.3: The probability distribution of the number of waters inside each cavitand for 1 bar and 278.15 K.

pressure increases. As the number of waters inside of TEMOA becomes closer to that of OA (*i.e.,* approaches 0), the partial molar volume difference between TEMOA and OA also decreases, but it does not approach 0. A linear fit was performed on the data for TEMOA with an intercept of 66.99 cm^3 mol^{-1}, which can be associated with the addition of four methyl groups when compared to OA. The remaining contribution of 53.7 cm^3 mol^{-1} can then be attributed to the fact that TEMOA is empty nearly 35% of the time unlike OA. These findings suggest that the partial molar volume differences between TEMOA and OA can be used as a bridge between experiment and computer simulation results.

The differences in partial molar volumes between OA and TEMOA (Figure 4.5a), and OA and TExMOA (Figure 4.5b) at standard conditions can be compared with four
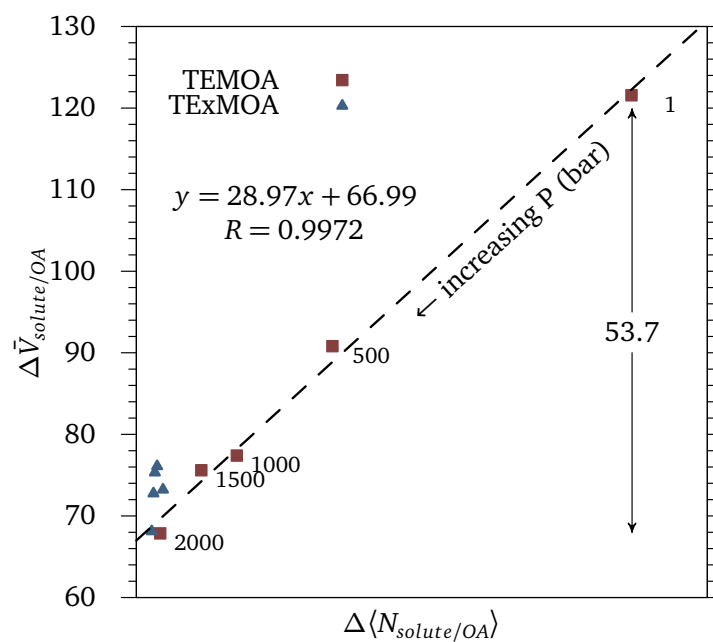
Figure 4.4: Difference in the number of interior waters between OA and TEMOA, and OA and TExMOA versus the difference in the partial molar volumes for the same systems.

Solute	$\bar{V}$ (cm^3 mol^{-1})	$\Delta\bar{V}$ (cm^3 mol^{-1})
OA	953.0	
TEMOA	1074.6	121.6
TExMOA	1026.3	73.3
alanine[72]	42.54	
glycine[72]	60.19	70.6

Table 4.1: Partial molar volumes ($\bar{V}$) of each cavitand and of two amino acids as a comparison. Since both TEMOA and TExMOA have four additional methyl groups, we compare the change in partial molar volume with these two cavitands and OA with 4 times the change in partial molar volume of glycine and alanine.

times the difference in partial molar volumes between glycine and alanine—identical amino acids with the exception of an additional methyl for glycine (Figure 4.5c).[72] The differences between OA and TEXMOA as well as four times the difference between alanine and glycine are in line with the expected value of 66.99 cm^3 mol^{-1} from Figure 4.4, indicated by the dashed line. Additionally one would expect $\Delta\bar{V}$ of OA/TEMOA and OA/TExMOA to be approximately four times that of $\Delta\bar{V}$ for alanine/glycine, since the only structural change between OA and both of its deep-cavity cavitand derivatives is the substitution of four hydrogens into four methyl groups. This is observed for $\Delta \bar{V}$ of TExMOA and OA; the difference between TEMOA and OA, however, is 51.0 cm^3 mol^{-1} larger than $4\Delta\bar{V}$ of glycine and alanine. We attribute this to TEMOA's accessing the empty-water state much more often than both OA and TExMOA.

TEMOA's partial molar volume dependency on accessing the empty state frequently can be shown by observing the partial molar volume of each deep-cavity cavitand as a function of water occupancy state (Figure 4.6, top). This value is statistically the same at each water occupancy state for both TEMOA and TExMOA. OA's partial molar volume is parallel to both TEMOA's and TExMOA's as a function of water occupancy
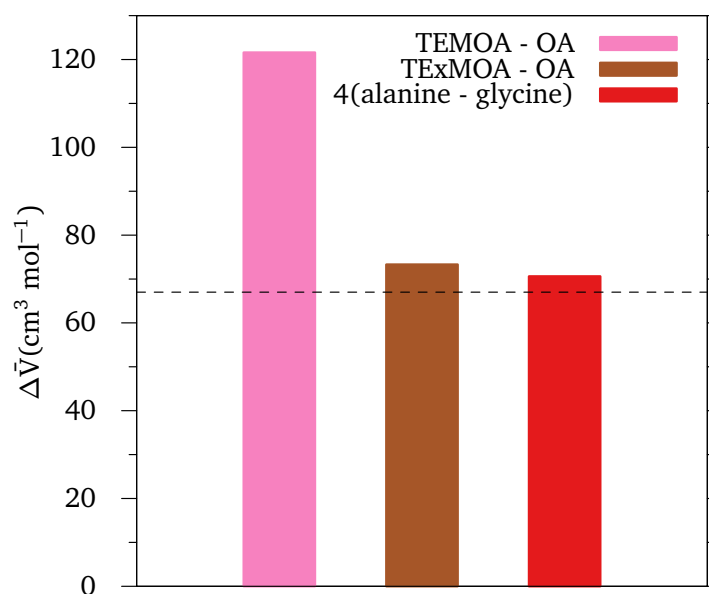
Figure 4.5: The difference in partial molar volume between OA and TEMOA (left); the difference in partial molar volume between OA and TExMOA (middle); four times the difference in partial molar volume between alanine and glycine, data from Ref. 72 (right). The dashed line indicates the value of the intercept of Figure 4.4. See Table 4.1 for data.

state, indicating that the difference in partial molar volume at each state between OA
and the other two deep-cavity cavitands is approximately constant. This is indicated by
the difference in the partial molar volume of OA and the other two deep-cavity cavitands
as a function of the number of waters inside their interiors (Figure 4.6, bottom). Each
of these differences is approximately constant at ~ 70 cm^3 mol^{-1}. This is expected
since both TEMOA and TExMOA are structurally similar enough that their partial molar
volumes are the same when at the same water occupancy state. The intercept from the
fit in Figure 4.4 (indicated by dashed line in Figure 4.6, bottom) is slightly lower than
the observed difference here due to the influence of high pressures in those previous
simulations, and is therefore consistent with the observations here.

Having established that the partial molar volume differences between OA and TEMOA
are indicative of TEMOA's tendency to evaporate its interior, we wish to examine this
attribute effect on host/guest association strength. In addition to spontaneously evac-
uating its interior of water TEMOA shows distinct volume of association behavior with
hexane—unlike OA or TExMOA (Figure 4.7). At standard conditions, both OA and TEx-
MOA have positive volumes of association with hexane. TEMOA, in contrast, has a neg-
ative volume of association with hexane which can be attributed to TEMOA's tendency
to be periodically dry throughout the simulation, unlike OA or TExMOA. From the in-
tegration of volume of association between one cavitand and one hexane from 1 bar to
1000 bar, the effect of pressure on the Gibbs free energy change during association was
obtained (Figure 4.8). For OA, while increasing pressure from 1 bar to 1000 bar, the as-
sociation Gibbs free energy increases by 1.12 kJ mol^{-1}. In the case of TEMOA, however,
a reverse tendency was predicted, where the association Gibbs free energy changed by
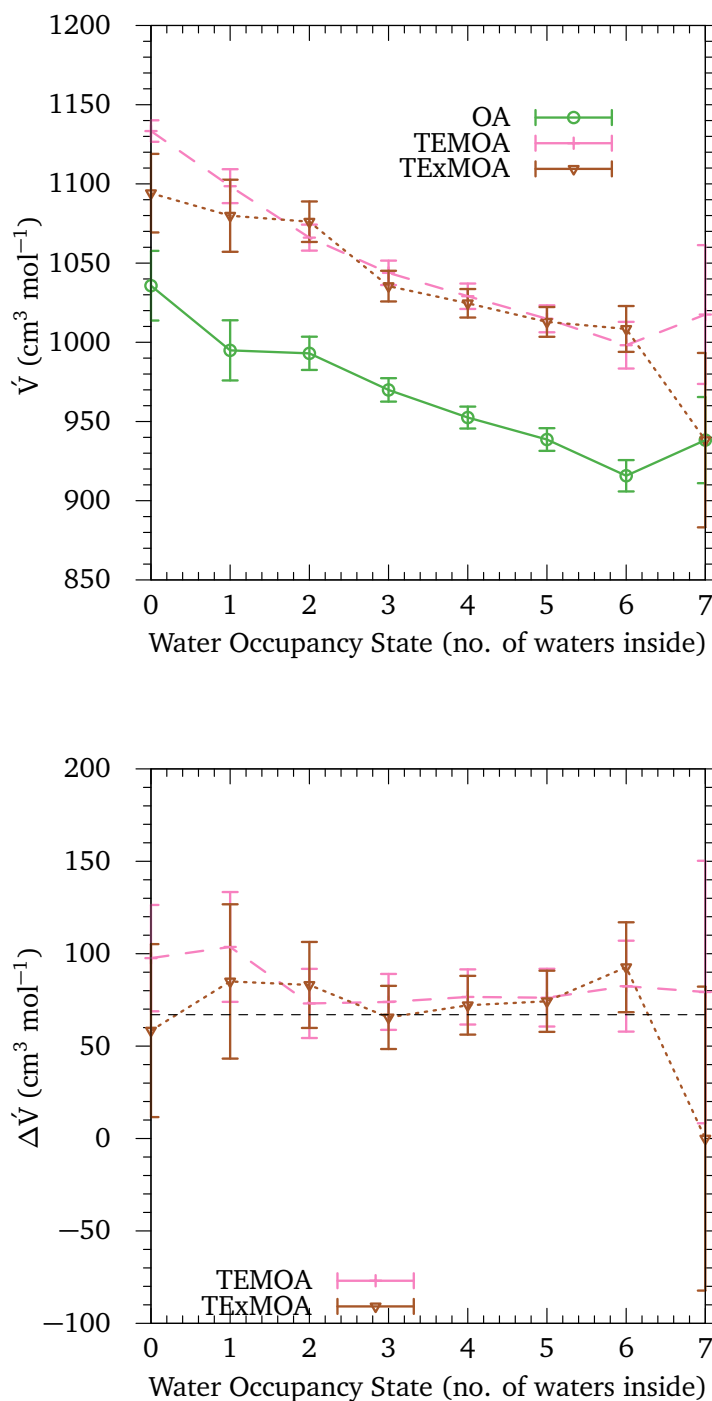
Figure 4.6: *Top:* The partial molar volume of each deep-cavity cavitand as a function of the number of waters within each deep-cavity cavtiand. *Bottom:* The difference in the partial molar volume of TEMOA and OA along with the same difference between TExMOA and OA as a function of the number of waters within each deep-cavity cavitand. The dashed line indicates the intercept from Figure 4.4.

-2.10 kJ mol^{-1}. The change in the minimum of the potential of mean force (PMF) of one cavitand and one *n*-hexane guest from 1 bar to 1000 bar corresponds with our predictions based on the volume of association. For OA, the minimum of PMF depth, which indicates both the preferred position of the cavitand and of guest, increased by 1.23 kJ mol^{-1} with the increase of pressure. This means increasing pressure makes the binding process less favorable. Additionally, this increase is consistent with the 1.12 kJ mol^{-1} increase from our prediction based on the volume of association above. For TEMOA, where a reverse result was observed, the PMF depth minimum was changed by -2.65 kJ mol^{-1} showing that increased pressure is making the binding process more favorable. Therefore, both in tendency and magnitude, the change of the PMF minimum for OA and TEMOA with pressure are consistent with the prediction from the change of volume of association. The slight difference between the values can be attributed to the fact that the PMF minimum is not the exact binding free energy; however, it is an important predictor of the binding free energy and the whole binding thermodynamic.

4.5 Conclusion

In this study we have shown that a small structural change in Octa-acid—adding methyl groups to the *endo* positions, forming TEMOA—results in large changes in the hydration characteristics of the binding pocket. Unlike OA, TEMOA is empty of water 35% of the time. A deep-cavity cavitand with methyls in the *exo* positions, TExMOA, does not have this characteristic. We also have shown that this dewetting phenomenon inside of TEMOA can be controlled with thermodynamic variables such as temperature and pressure, falling in line with previous studies on capillary evaporation.
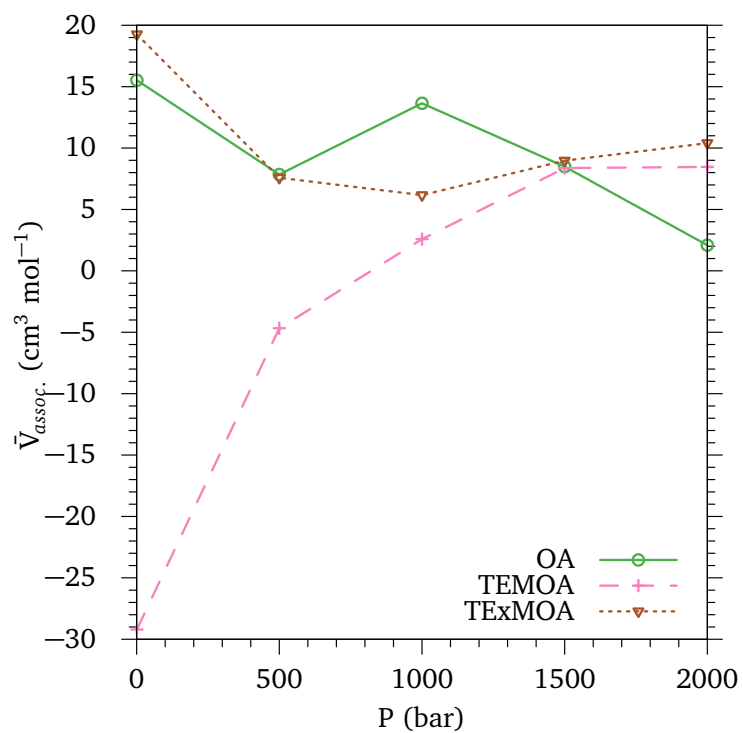
Figure 4.7: Volume of association between each cavitand and a hexane guest as a function of pressure.
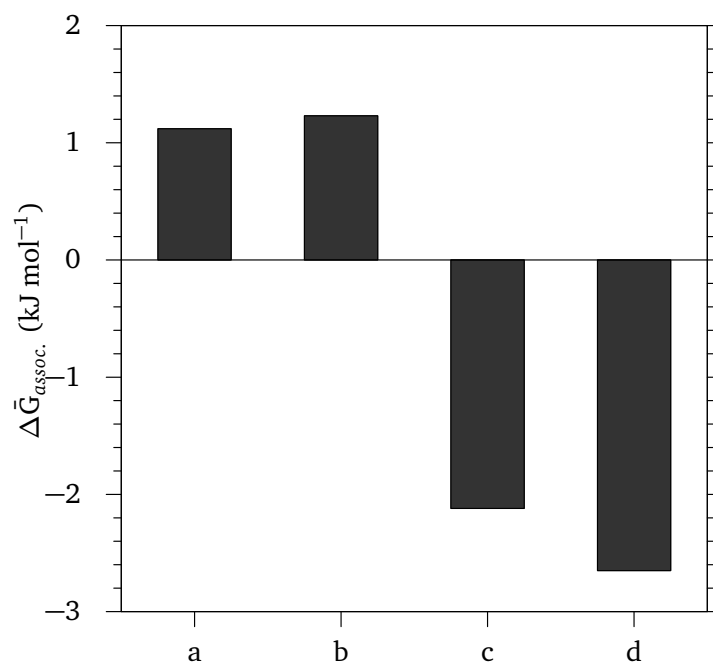
Figure 4.8: *a.* Free energy of association of OA and hexane calculated from integrating the volume of association between 1 bar and 1000 bar (Figure 4.7). *b.* Change in the PMF minimum for OA /hexane association from 1 bar to 1000 bar. *c.* Free energy of association of TEMOA and hexane calculated from integrating the volume of association between 1 bar and 1000 bar (Figure 4.7). *d.* Change in the PMF minimum for TEMOA/hexane association from 1 bar to 1000 bar. Full PMF's are shown in Figure D.8.

Most importantly we have introduced a simple bridge to experiment—the partial molar volume differences between TEMOA and OA. We have shown that TEMOA is "too large" based on structure alone and that this is attributed to TEMOA's interior accessing the empty state frequently. Additionally we have shown that TEMOA associates with hydrophobic guests more strongly than OA due to its tendency to be dry in its interior. We are currently working with our experimental collaborators to validate our findings.

Chapter 5

CONCLUSION

In Chapter 2 we used computer simulations to investigate the molecular-level details of a series of homologous guests inside of an Octa-acid dimer. We showed there were four, length-dependent guest packing motifs—extended, helical, hairpin, and spinning top. Additionally we showed that guests pack according to the "rule of 5 or 6"; that is, five or six carbons fill each side of the dimer complex fully extended in each of these motifs. Additionally the usefulness of principal component analysis when choosing dominant structures from our simulations was demonstrated. The structures harvested from these calculated had computed chemical shifts that were highly correlated with those of experiment.

In Chapter 3 we investigated multimeric TEMOA systems, showing that guest packing drives the transitions between various assembly states. Volumes of these assemblies were shown the be larger than those formed by the hydrophobic effect. Additionally packing coefficients were shown to describe the transitions between each assembly state.

The hydration characteristics of the interiors of three deep-cavity cavitands were studied in Chapter 4. There we showed that the simple addition of methyl groups to

the *endo* positions of OA—thus forming TEMOA—resulted in a two-phase behavior of the cavitand's interior waters. This was not true for methyls in the *exo* positions of the cavitand (TExMOA). We also showed that TEMOA's partial molar volume is larger than one would expect based just on structure alone. This characteristic can be attributed to TEMOA accessing the empty state much more often than either OA or TExMOA, and it indicates that partial molar volumes can serve as a simple bridge between simulation and experiment. It was also demonstrated that TEMOA showed a stronger binding association with a hydrophobic molecule than OA.

5.1 Future work

Several areas remain to be investigated for Octa-acid and its derivatives. In this dissertation we have simulated and studied 2:1 deep-cavity cavitand host-guest complexes (Chapters 2 and 3) as well as 1:0 host-guest complexes (*i.e.*, empty cavitands, Chapter 4). However, we have neglected shorter guests that can form 2:2 and 1:1 complexes. Experimental studies have shown nonmonotonic assembly profiles for TEMOA assemblies with *n*-alkane guests at shorter lengths, unlike OA.[73] Simulations can provide further insight into this unusual property and how it may be connected to the differences in the hydration characteristics of the interiors of these deep-cavity cavitands that we discussed in Chapter 4.

Additionally ion effects on deep-cavity cavitand assemblies have not been systematically studied with computers simulations. Experiments have shown that salts affect the binding of anion guests with Octa-acid.[74] Simulations should be used in order bet-

ter understand the effects of electrostatics not only on the binding properties of these

deep-cavity cavitands, but also on how assembly states may be affected.

Appendix A

ASSOCIATION OF AMPHIPHILIC MACROCYCLIC TETRALACTAM HOSTS WITH SQUARAINE DYE

A.1   Introduction

Amphiphilic molecules are ubiquitous within biological systems and in nature, utilizing their dual nature in many processes and structures. Phospholipids, for example, which contain a hydrophobic tail and hydrophilic head, are the main building blocks of biological membranes. Another notable example is proteins, which contain both hydrophobic and polar side chains that aid in the self-assembly process of folding. More pertinent to the results here is that binding processes in biological systems often utilize amphiphilic receptors. The classic biological example of an extremely strong binding processes using both hydrophobic and polar groups is biotin binding with the streptavidin protein $(\log K_a \approx 13 - 15)$.[75]

In contrast to naturally-occurring amphiphilic receptors such as these, most synthetic receptors are uncharged and primarily utilize the hydrophobic effect for binding. Perhaps the most thoroughly studied among hydrophobically-driven synthetic receptors are cyclodextrins[76] and cucurbiturils,[77] ring-shaped molecules containing a hydrophobic interior. In addition to these cyclical molecules, both experimental and computa-

tional studies have been performed on bowl-shaped concave receptors—such as Octa-acid[34] and Rebek's capsular host[30]—which utilize the hydrophobic effect in order to bind non-polar and negatively charged guests.

Although rare, interest in designing and synthesizing synthetic amphiphilic cavities—often designated as cooperatively enhanced receptors (CER)—has increased in recent years.[78] Recently Smith and coworkers created a synthetic organic-soluble amphiphilic cavity known as a macrocyclic tetralactam host.[79] Such tetralactam hosts are made up of two anthracene walls with four amide residues pointing into the interior of the binding cavity which is located in the center of host. They showed that the host can encapsulate a squaraine dye with extremely high affinity ($\log K_a = 5.2$) in chloroform and that such complexation brings about emergent properties such as "improved chemical stabilities, red-shifted absorption/emission maxima, and different cell localization propensities."[79]

More recently Smith developed a water-soluble analogue of the tetralactam host which differs from the organic-soluble host only in its peripheral appendages.[80] The water-soluble version was shown to be able to encapsulate deep-red fluorescent squaraine dyes flanked with PEG chains. The complexation had association constants about 1000 times higher than those of its organic-soluble analogue. Furthermore they showed that by adding different *n*-alkyl substituents to the ends of the dye they could maintain the high association constant while changing the threading kinetics. These substituent stopper groups are essentially "speed bumps" that control threading rates through the macrocycle.[81]

In this study we use simulations to investigate the water-soluble tetralactam host's binding with squaraine dye guests. Water molecules within hydrophobic containers
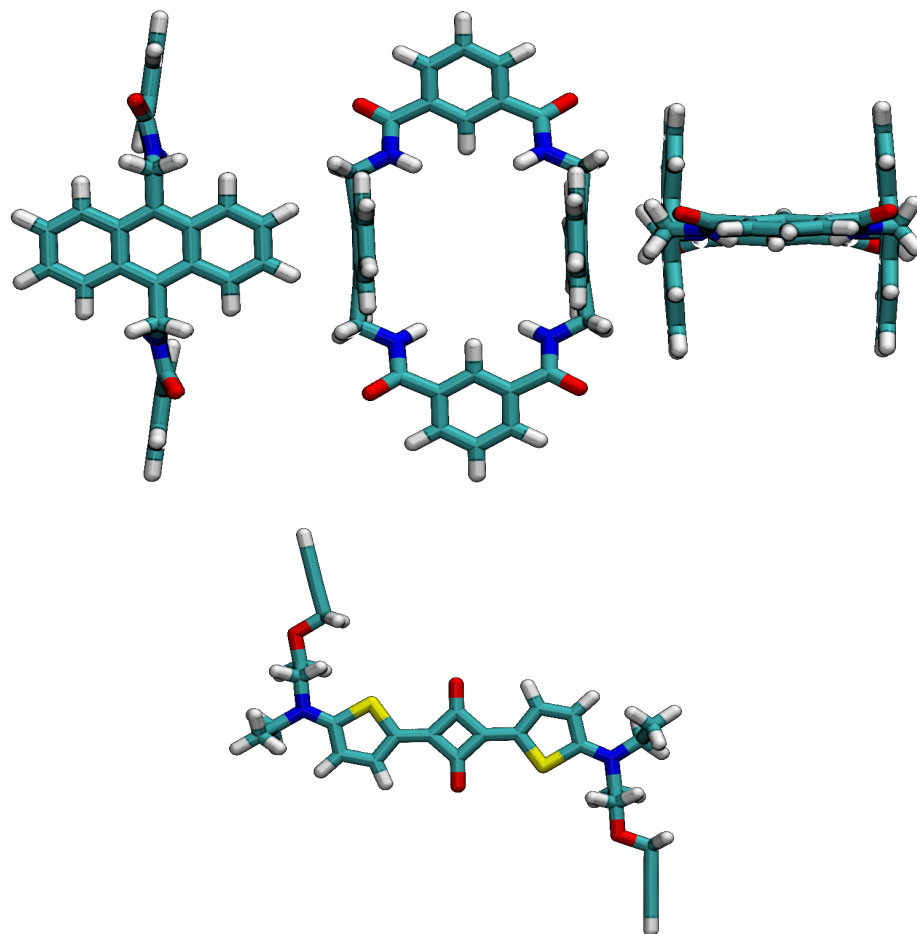
Figure A.1: *Top:* Three different views of the tetralactam. *Bottom:* Squaraine dye.

such as cyclodextrins are often energetically frustrated. Is this the case with waters within an amphiphilic molecule such as the tetralactam? First we investigate the behavior of water within a free tetralactam host in order to answer this question and also to determine water's role in stabilizing a specific conformation of the host. From there we use potential-of-mean-force calculations to describe the binding between the tetralactam host and various squaraine dye guests with stopper groups to investigate their relationship to the kinetic "speed bumps" in the aforementioned study.[81] Lastly, we break down the various contributions to the binding process. Will a hydrophobic interior inside the tetralactam bring about favorable binding with a hydrophobic squaraine? Or, do the electrostatic interactions provide a crucial impetus for the binding process?

A.2   Methods

We simulated various versions of the water-soluble tetralactam host and squaraine dye guest using GROMACS 5.1.[31] Simulations of a free tetralactam host in water were performed with four different variations to its structure: unmodified, with amide hydrogen charges turned off, with two methoxy groups added to the each end of each anthracene, and with both methoxy's added and amide hydrogens turned off.

We used ACPYPE[52] to calculate partial charges for both the host and guest by using an AM1-BCC fitting procedure, and hosts and guests were modeled using the General Amber Force Field (GAFF).[51] The TIP4PEW[35] water model was utilized for all water molecules, with all bonds within each water molecule kept rigid using SETTLE[71] and all solute bonds involving hydrogens transformed into constraints using LINCS.[54] We hydrated initial configurations with between 1,100 and 1,582 waters for regular simu-

lations and between 5,317 an 5,327 waters for potential-of-mean-force (PMF) simulations.

These systems were simulated in the isothermal-isobaric ensemble at 25 °C and 1 bar using the Nosé-Hoover thermostat[38] and the Parrinello-Rahman[39] barostat, respectively. Short-range van der Waals interactions were cut off beyond 9 Å where a long-range dispersion correction was used. Real-space electrostatic interactions were also cut off at 9 Å, using Particle-mesh Ewald summation[10d] to evaluate long-range electrostatic interactions beyond the cutoff. Each of these systems, as well as bulk water, was simulated for 100 ns after at least 1 ns of equilibration, using a 1 fs time step.

The indirect umbrella sampling method (INDUS)[82] was used to insert benzene-sized cavities around the tetralactam. The probability of inserting a cavity with no water molecules inside of it ($P_V(0)$) was used to determine the free energy of transferring each cavity from bulk water to positions on the outside of the anthracenes, as well as inside, with each benzene-shaped cavity aligned parallel to the aromatic ring nearest it on the tetralactam.

Additionally, we performed PMF simulations on a variety of tetralactam-squaraine systems with the unmodified host, where the "stopper" groups of the squaraine were changed (Figure A.2). PMF simulations were also performed with the ethyl variation of the squaraine dye with all charges of the squaraine turned off with the unmodified host, as well as PMF simulations with the ethyl variation of the squaraine with the host's amide hydrogens turned off. PMF simulations of dimethoxy-ethane with the tetralactam host were additionally performed. Lastly PMF simulations of the ethyl variation of the squaraine dye with the unmodified host in the gas phase, both with the guest with
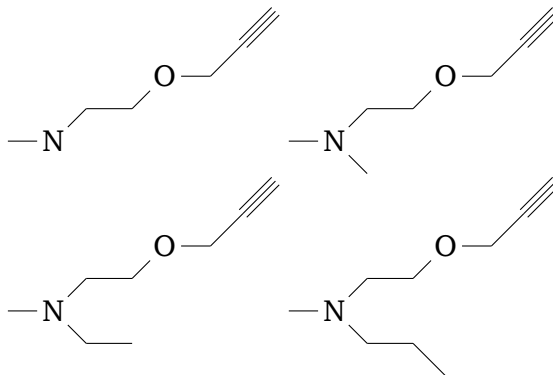
Figure A.2: Stopper groups of the squaraine dye used in PMF's. From left to right: hydrogen, methyl, ethyl, propyl

normal charges and with the guest without charges, were performed. All PMF simulations had windows spaced at 0.005-nm increments on both sides of the host using a spring constant of 2,000 kJ mol^{-1} nm^{-1}, for a total of 800 windows for each system. We kept both the entrance to the host and the longest dimension of the squaraine guest aligned to the z-axis of the box in PMF simulations. Each window was simulated for 30 ns after at least 2 ns of equilibration.

A.3   Results and discussion

A.3.1   Free tetralactam in water

Our simulations give insight into how how a free tetralactam host behaves in water and how waters within the tetralactam behave. Throughout our simulations, the tetralactam host switches between two conformational motifs—one where the two anthracene groups are nearly aligned, and one where these two groups are not aligned. The cosine of the angle between the two anthracene groups describe these conformational changes, with $\cos\theta = 1.0$ being completely parallel (Figure A.3). For the unmodified tetralac-

tam there is a bimodal distribution separated by a minimum near 0.8. We consider the anthracenes aligned for values larger than this. For values less than this the anthracenes form a funnel-like shape with one entrance much larger than the other, with a peak centered near 0.6. Using 0.8 as the cutoff between the two types of observed conformations, we calculated that 72.8% of the time the anthracenes were aligned for the unmodified host, and on average there were 4.481 ± 0.040 waters within the host. A water was considered to be within the macrocycle if it was in between the two ends of the anthracene entrances and was not more than 0.3 nm above or below the lines bisecting the centers of the anthracenes. When the two anthracenes were aligned, 4.910 ± 0.006 waters were within the interior, and when $\cos(\theta)$ was less than 0.8, 3.336 ± 0.010 were present. In other words, as the free tetralactam hosts goes from an "aligned" configuration to a "funnel" configuration, about 1.5 waters are expelled from inside of the tetralactam.

The water within the unmodified host also has a higher potential energy than bulk water (cf. Appendix E). Bulk water has a potential energy of -92.970 kJ mol^{-1}, but water within the tetralactam has a potential energy of -75.209 kJ mol^{-1}. When we consider only the case when the two anthracenes are aligned, the energy is slightly higher (-76.573 kJ mol^{-1}); however, the potential energy of water within the tetralactam when the anthracenes are not aligned is even greater (-71.568 kJ mol^{-1}), indicating that water is somewhat more perturbed when the tetralactam is in the "funnel" shape.

By turning off the interior amide hydrogens of the tetralactam, we can investigate their influence on the hydration of the interior of the host. For the tetralactam with amide hydrogen charges turned off the peak in Figure A.3 near 0.6 increases and the
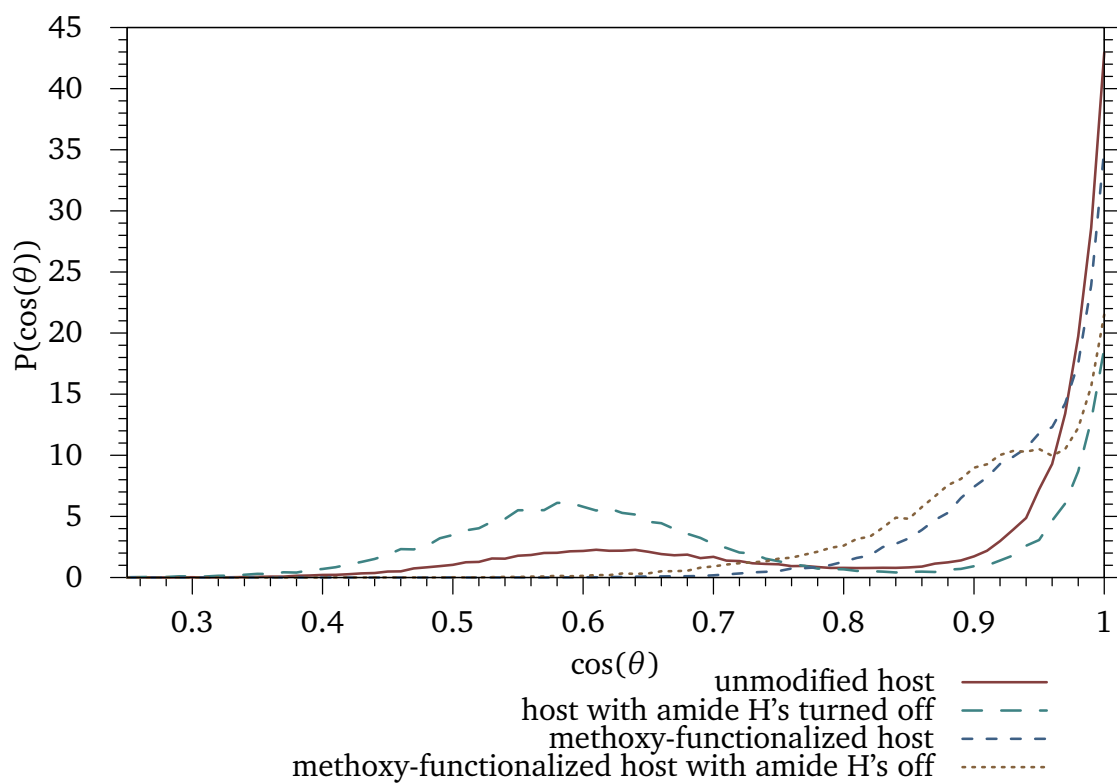
Figure A.3: Distribution of the cosine of the angle between the anthracenes.

value near 1.0 decreases drastically. This indicates that the anthracenes are in alignment much less often with the interior amide hydrogen charges turned off, with the alignment conformation only occurring 33.2% of the time. On average 3.249 ± 0.078 waters were within the macrocycle, which is approximately 1.2 waters less than inside an unmodified host. Not surprisingly, waters within the host with amide hydrogens turned off have a potential energy of about 5 kJ mol^{-1} more than those within an unmodified host, indicating that the waters are more frustrated without being able to hydrogen bond to an amide inside of the host (Table E.3).

We used 9 benzene-shaped empty cavities to effectively "map" the hydrophobicity of both the regular tectralactam and the tetralactam with its interior amide hydrogens turned off. Figure A.4 shows the free energy of transferring such probes from bulk water to various locations in and around the host, averaging values that are in symmetric positions. It is more energetically favorable to transfer such cavities to the exterior and interior edge of the non-modified tetralactam when compared to the interior center. This is attributed to the fact that the four amide hydrogens are pointed to the interior of the tetralactam host. This is made even more evident when the amide hydrogen charges are turned off. In that case the transfer of a benzene-sized empty cavity to the interior center region of the tetralactam is over 8 kT more favorable than the same such transfer to the normal tetralactam.

Both the fewer waters within the host with amide hydrogens turned off and the higher energy of those waters are from the loss of the ability of waters to hydrogen bind to the amide hydrogens. Additionally, the tetralactam's anthracenes' lack of alignment of the anthracenes when the amide hydrogens are turned off is due to the ab-
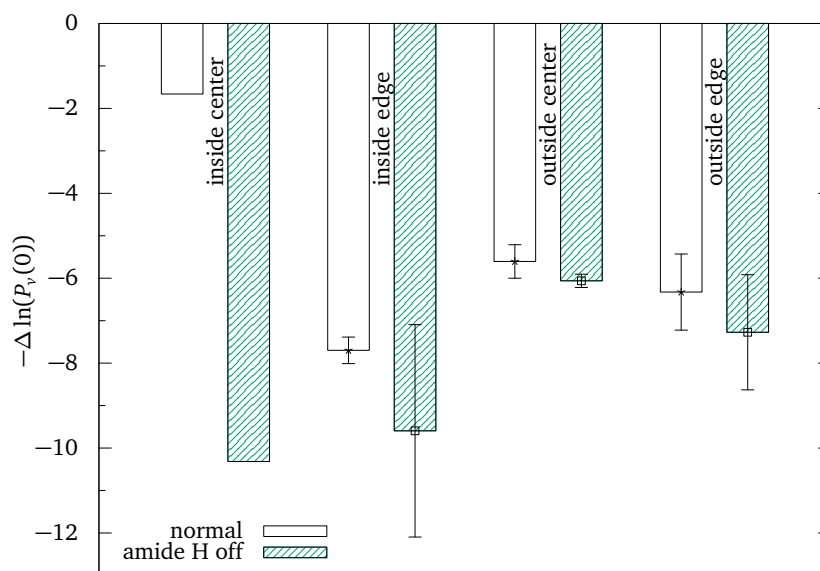
Figure A.4: The free energy (kT) of transferring a benzene-sized probe from bulk water to different locations in and around the tetralactam. Values for symmetric positions were averaged. Error bars indicate the standard deviation in symmetric quantities. The central interior location has no symmetric position. White bars indicate the free energy for the normal tetralactam, and green, shaded bars indicate that for the tetralactam with amide hydrogens turned off.

sence of water molecules inside its interior. To demonstrate this we added methoxy groups to the ends of the anthracenes of the tetralactam. By adding these methoxy groups we force the macrocycle to stay aligned 96.2% of the time throughout the simulation. When the charges of the amide hydrogens are turned off with the addition of methoxys, the alignment percentage drops to 88.8%. However, even though the anthracenes are not aligned as often, there is a decrease in the average number of interior waters from 4.550 ± 0.014 in a methoxy-functionalized host to 3.618 ± 0.024 in a methoxy-functionalized host with amide hydrogens turned off. This means that there are fewer waters in the methoxy-functionalized host with the amide hydrogen charges turned off than there are within the methoxy-functionalized host with the hydrogens turned on. Additionally, waters within the methoxy-functionalized host have a slightly higher potential energy than an unmodified host with a value of -74.053 kJ mol^{-1}, and waters within the methoxy-functionalized host with amide hydrogens turned off are the most perturbed, having a potential energy of -68.042 kJ mol^{-1}.

The average number of waters within the methoxy-functionalized host and amide hydrogen's turned off are not as low as those within a host with just the methoxy modification; however, the distribution of waters shows that there is a small percentage of the time that the interior is completely dry in the methoxy case with amide hydrogens off as demonstrated in Figure A.5. This means that when the amide hydrogen charges are turned off the water molecules are able to leave the interior even with the anthracene groups forcibly aligned with the methoxy groups. This shows that the lack of anthracene alignment that periodically occurs in the unmodified host—and which is exacerbated when the tetralactam's amide hydrogens are turned off—is due to water
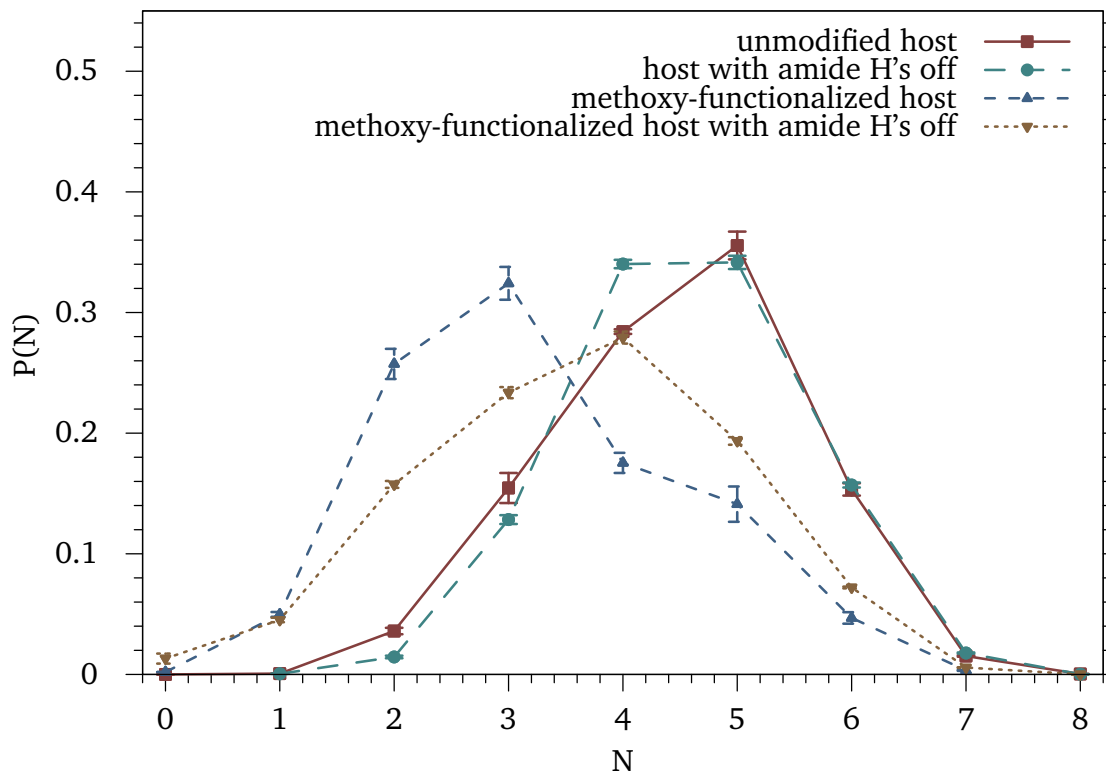
Figure A.5: Distribution of waters inside each version of the tetralactam host.

leaving the interior and not vice versa. In other words, when we remove the hydrogen-bonding ability of the interior of the tetralactam host and create a purely hydrophobic pocket, we create a situation in which the anthracenes are no longer stabilized by the interior water.

A.3.2 Potentials-of-mean force between tetralactam and squaraine dyes

Now that we have examined a free macrocyclic tetralactam host and how interior waters stabilize the structure in an aligned conformation, we investigate the binding between a tetralactam and various squaraine dyes. Figure A.6a shows the potentials of mean force for the unmodified tetralactam and the squaraine dye with different stopper

groups. When no stopper group is present, there is no barrier to entering the interior of the tetralactam. However, by adding the methyl stopper groups a barrier near 0.75 nm emerges. The barrier for the squaraine with the ethyl stopper group is larger than the one corresponding for the methyl stopper group, and the propyl stopper groups' barrier remains approximately the same size as the one for the ethyl stopper group. These barriers demonstrate that adding such stopper groups causes the squaraine to have difficulty in entering the tetralactam host, corresponding with the "speed bumps" observed in experiment.[81] However, the minimum in the PMF is still on the order of -50 to -60 kJ mol^{-1}, in agreement with the idea that the n-alkyl groups affect the binding kinetics and not necessarily the binding strength. In previous experimental studies polyethylene glycol (PEG) chains were functionalized to the ends of squaraine dyes; these dyes were shown to thread through the tetralactam host.[80] In this study we approximate the binding effects of PEG chains by studying the PMF between a dimethoxyethane guest and tetralactam host. Unlike the squaraine dyes with stopper groups, dimethoxyethane has no free energy barrier when transferred from bulk water to the interior of the host, indicating that the PEG chains easily thread through the tetralactam host while the stopper groups cause the "speed bumps".

In addition to the free energy barriers located near 7.5 Å, we also see a smaller barrier near the center of the tetralactam host, while the minimum in the free energy profile is near 0.15 nm. The amide hydrogens within the tetralactam do not point to the exact center of the host but instead point slightly outward toward the oxygens on the squaraine dye, causing this non-centered minimum. Figure A.6b shows that the barrier near the center is increased when changes are made to the host and guest.
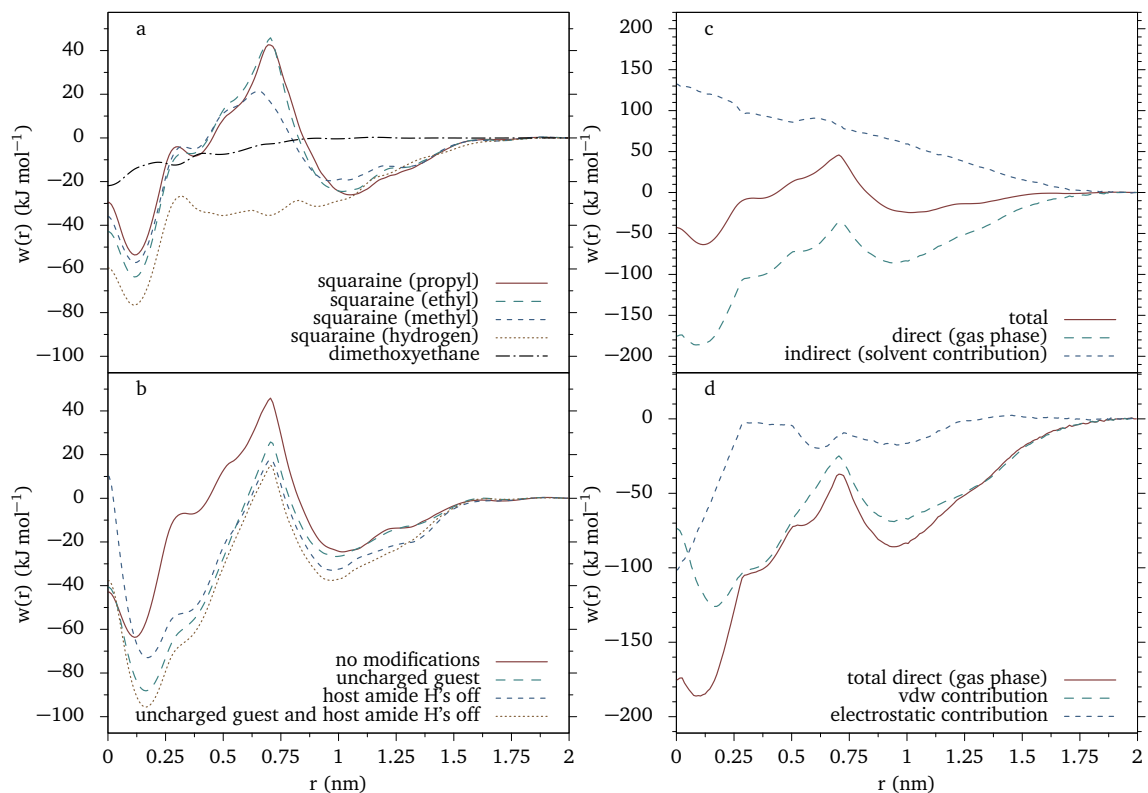
Figure A.6: *a.* Potential of mean force between lactam host and squaraine guests with various stopper groups. Additionally the PMF between the lactam host and dimethoxyethane is shown. *b.* PMF between lactam host and squaraine-ethyl guest with various modifications. *c.* Breakdown of PMF between lactam host and squaraine-ethyl guest into direct and indirect contributions. *d.* Breakdown of direct contribution of c. into van der Waals and electrostatic contributions.

When all charges are removed from the ethyl version of the squaraine guest, this central barrier increases by a factor of 2.5 which indicates that there is hindrance due to van der Waals interactions that increases due to the loss of hydrogen bonding between the host and guest. When the host's amide hydrogens are turned off and the squaraine charges are normal, the barrier increases by a factor of 4. In this case there is still a loss of hydrogen bonding between the host and guest, but the squaraine is still able to hydrogen bond to water molecules. When both the lactam amide hydrogen charges and the squaraine charges are turned off, the barrier has a similar value to the system with just an uncharged guest.

In addition to the need to overcome steric forces in the tetralactam-squaraine binding, the binding mechanism also competes with the hydrogen-bonding of waters to the oxygens on the squaraine dye. Figure A.7 shows the number of water hydrogen bonds formed with the ethyl version of the squaraine dye as a function of distance along the z-axis from the center of the tetralactam host. The unmodified tetralactam strips the waters from the squaraine guest as the guest enters the macrocyclic structure. The tetralactam with amide hydrogen's turned off does not have the same effect. Although hydrogen bonds with water are reduced as the squaraine guest enters the structure, they do not ultimately go to zero like they do with the unmodified host. This indicates that the water hydrogen bonds are a key factor in the free energy barriers seen in Figure A.6b and that the electrostatic contribution from the amide hydrogens are essential to stripping off the waters in order for binding to be achieved.

Thus far we have shown that the amphiphilic nature of the tetralactam host is essential not only for its stability in water, but also for binding with the squaraine dye.
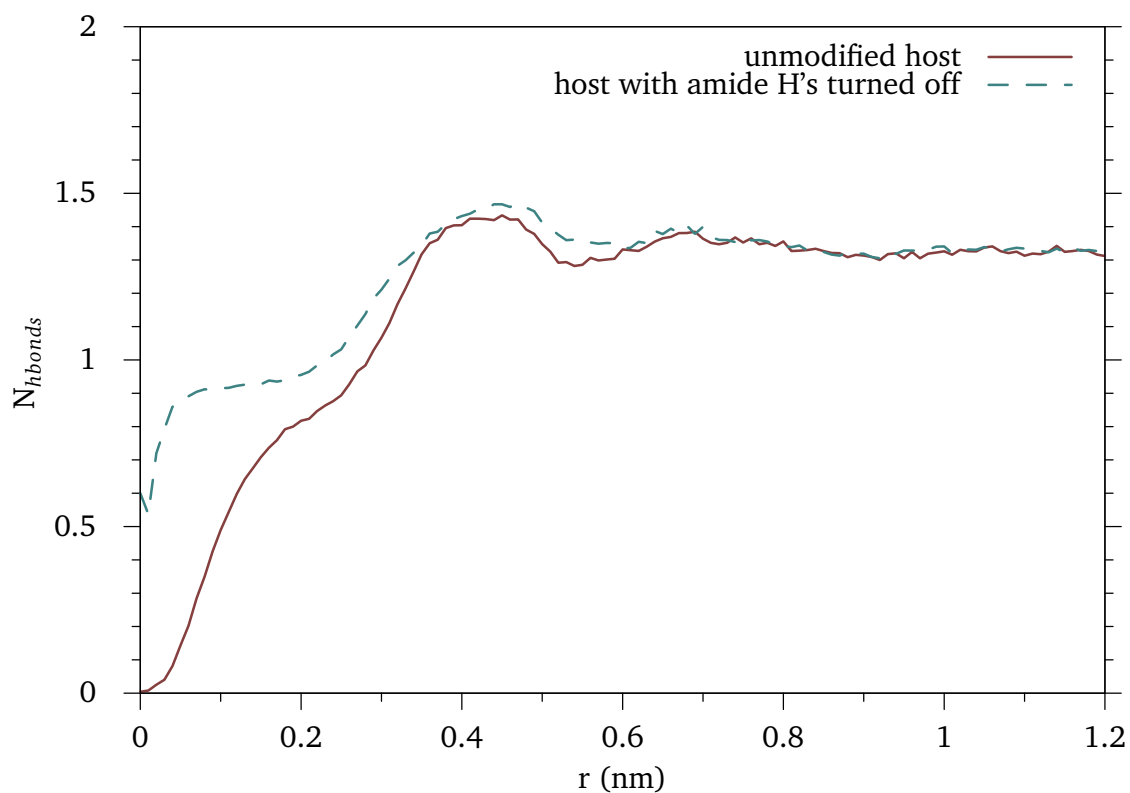
Figure A.7: The number of water hydrogen bonds with a squaraine (ethyl) oxygen as a function of distance along the z-axis from the center of the tetralactam host. As the squaraine dye enters the host the waters are stripped from hydrogen-bonding with it since the squaraine prefers to bind with the lactam amide groups. When the lactam amide hydrogen charges are turned off, more waters tend to stay hydrogen-bonded to the squaraine dye, making it more difficult for the squaraine to bind with the host (see Figure A.6).

To investigate this further, we show the direct and indirect contributions to the potential of mean force between the tetralactam and the ethyl version of the squaraine dye in Figure A.6c. As expected, the indirect, or solvent contribution, is the driving factor between the association of the two molecules. When that indirect contribution is broken down further between electrostatic and van der Waals contributions (Figure A.6d), we can see that up until about 0.25 nm the electrostatics contribute very little to the association. From there the electrostatics contribute favorably for association, all the way until the center of the macrocycle. On the other hand, the van der Waals show the characteristic barrier seen in Figure A.6a at about 0.75 nm after which the van der Waals interaction contributes favorably until the minimum near 0.15 nm. From there the van der Waals contribution increases unfavorably until the center of the macrocycle. We conclude from this that, although van der Waals interactions primarily drive the initial association past the "speed bump", there is also a steric hindrance when the guest reaches the center of the macrocycle that must be overcome by the electrostatic contributions. When coupled with Figure A.6b, this indicates that the main contributer to overcoming steric forces at the center are amide hydrogen charges interacting with the guest.

A.4 Conclusion

In summary, the results of this study have shown the amphiphilic nature of the water-soluble macrocyclic tetralactam host. A free tetralactam host is stabilized in its parallel conformation by interior waters, and the interior waters themselves are perturbed, exhibiting the "non-classical" hydrophobic effect. When a squaraine dye is brought nearby

such that the two associate, the host strips the dye of any waters that are hydrogen-bonded to it, and the electrostatic interactions of amide tetralactam residues play a key role in binding with the squaraine dye. Additionally the "stopper" groups that were shown to change association kinetics in previous experimental studies create free energy barriers for the entry of the squaraine guest into the macrocycle host. This look into the details of the water-soluble tetralactam host and its association with squaraine dyes allows for greater understanding of the molecular mechanisms between these two molecules.

Appendix B

RESOLVING SOLVOPHOBIC INTERACTIONS INFERRED FROM EXPERIMENTAL SOLVATION FREE ENERGIES AND EVALUATED FROM MOLECULAR SIMULATIONS

B.1 Summary

Ben-Naim estimated the solvent-mediated interaction between methanes based on experimental solvation free energy differences between chemically similar hydrocarbons. Interactions were predicted to be strongest in water, dominated by characteristic entropic gains. We use molecular simulations in combination with an empirical interpolation procedure that bridges interactions from outside methane's excluded volume down to overlap to test Ben-Naim's estimates. While Ben-Naim's approach captures many distinctive trends, the alchemical differences between methane and a methyl unit play a non-trivial role on the predicted association strength and the sign of enthalpic and entropic components of the interaction free energy in water and ethanol.

B.2 Introduction

The substantive attraction between non-polar moieties in water inferred from micellization and protein folding experiments is regularly invoked as a principal driving force for aqueous phase assembly.[58] Experimental examination of hydrophobic interactions

in isolation from other water-mediated forces is hampered by the sparing solubility of purely non-polar species, however, necessitating the inclusion of polar and charged functional groups to increase solute concentrations and raise the observable response to measurable levels. Resultantly, much of our understanding of molecular-level hydrophobic interactions is inferred from experimental studies of the hydration of isolated non-polar species and the assembly of amphipathic species, as well as theoretical and simulation studies of simple non-polar species in water. Computationally, the interaction between methanes, quantified by their potential-of-mean force (PMF) in water, has served as a foundation for understanding hydrophobic interactions.[83] The minimal solubility of even simple solutes like methane makes direct experimental resolution of the PMF by neutron scattering experiments intractable, as noted above. Recent simulation studies have demonstrated that the second-osmotic virial coefficient for simple non-polar solutes in water, which depends on the integrated PMF, can be reliable evaluated, opening up new avenues for experimental validation.[84] In addition, new spectroscopic probes of the hydration shells of hydrophobic moieties and their aggregates have been developed that are providing microscopic structural insights that can potentially bridge the experimental and computational divide.[85]

In a series of papers from the 1970s,[20] Ben-Naim proposed an experimental estimate of molecular-scale hydrophobic interactions based on the hydration thermodynamics of small hydrocarbon species. His insight was that solvent-mediated contributions to the PMF are related to solvation free energies differences. The solvent-mediated interaction between methanes was subsequently estimated as the solvation free energy difference between ethane and two methanes at infinite dilution. While this estimate

only provides the solvent-mediated interaction at a separation equal to the carbon-carbon bond length of ethane ($\sim$1.54 Å), which is physically prohibited to non-bonded groups by excluded volume interactions, Ben-Naim's work underscored important differences between interactions in water and a number of organic solvents. Specifically, this interaction was strongest in water and grew with increasing temperature, while the other solvents examined exhibited comparatively weaker, temperature insensitive interactions.[20] Broken down into its thermodynamic components, the methane interaction in water was found to be entropically favored and enthalpically disfavored at ambient temperatures, while in organic solvents the interaction was dominated by a favorable association enthalpy with minor entropic perturbations. Similar thermodynamic conclusions have been drawn when this procedure has been applied to a wider range of solutes, including the interaction between the methyl units in xylene as evaluated by the solvation free energy difference between p- and o-xylene.[20d] Smith and Haymet[83c] concluded on the basis of simulations of the hydrated methane PMF and its temperature derivatives that Ben-Naim's estimates are at least qualitatively correct, although their comparison relied on visual extrapolation of simulation results that are available for methane separations greater than $\sim$3.5 Å down to the ethane carbon-carbon bond length. Wu and Prausnitz[86] subsequently utilized Ben-Naim's procedure to derive an effective molecular-scale pair-wise additive hydrophobic interaction potential consistent with the hydration free energies of a range of saturated hydrocarbons. Ben-Amotz[93b] has recently pointed out that if Ben-Naim's experimental estimates are quantitatively correct, molecular simulations appear to over predict the magnitude of the hydrophobic attraction and that this discrepancy needs resolution.

In a recent molecular simulation study of methane interactions in water over a broad range of temperatures and pressures, Ashbaugh and coworkers[83e] developed an interpolative Ansatz for bridging solvent-mediated interactions evaluated from radial distribution functions passively observed during simulations and the value at direct overlap obtained from particle insertion free energies. Specifically, they proposed the solvent-mediated contribution to the PMF between solutes is described over all separations from overlap out to infinity by the equation

$$\omega_{mm}^{j}(r) = f(r)\omega_{mm}^{j}(r)|_{cubic} + [1 - f(r)]\omega_{mm}^{j}(r)|_{sim}. \tag{B.1}$$

The function $\omega_{mm}^{j}(r)|_{sim}$ is the solvent-mediated contribution to the methane-methane PMF determined from molecular simulations in solvent j following the expression $\omega_{mm}^{j}(r)|_{sim} = -kT \ln g_{mm}^{j}(r) - \phi_{mm}(r)$ (Here kT is the product of Boltzmann's constant and the temperature, $g_{mm}^{j}(r)$ is the methane-methane radial distribution function in j, and $\phi_{mm}(r)$ is the gas phase pair interaction between methanes). The solvent-mediated PMF evaluated from radial distribution functions is accurate only for separations greater than ~ 3.5 Å as a result of poor/no sampling in the excluded volume regime. The function $\omega_{mm}^{j}(r)|_{cubic}$ is a fitted cubic function that bridges from the methane overlap value determined from simulation out to separations of ~ 5 Å (evaluated as $\omega_{mm}^{j}(r)|_{cubic} = \omega_{mm}^{j}(0) + ar + br^2 + cr^3$, where $\omega_{mm}^{j}(0)$ is the value of the PMF at direct methane overlap, and a, b, and c are fitted variables smoothly interpolating between overlap and the solvent contribution extracted from the radial distribution function). The function $f(r)$ is a cubic switching function that changes values from 1 to 0 over the range R_{lower}

(=3.5 Å) to R_{upper} (=5 Å) to smoothly transition between $\omega_{mm}^{j}(r)|_{cubic}$ and $\omega_{mm}^{j}(r)|_{sim}$. Additional information on the development and fitting of Equation B.1 can be found in Ref. 83e. An advantage of using Equation B.1 to obtain the solvent contribution to the PMF over more traditional approaches, like free energy perturbation (*e.g.,* Ref. 83c), is that the PMF over a wide range of separations can be determined from the interpolation formula using results from a single simulation rather than from multiple windows.

The capacity of Equation B.1 to describe methane interactions over all separations opens up the potential to quantitatively compare Ben-Naim's experimental estimates for the solvent contribution to the PMF in the unobserved overlap regime against molecular simulations. In this paper we compare Ben-Naim's predictions for methane interaction free energies, enthalpies, and entropies in water and ethanol against interpolated molecular simulation results obtained from Equation B.1 to assess the fidelity of simulation models at capturing experiment and vice versa.

B.3   Simulation methodology

Simulations were performed using the GROMACS 5.0 simulation package.[31c] Simulations were conducted in the isothermal-isobaric ensemble with the temperature and pressure controlled using the Nosé-Hoover thermostat[38] and Parrinello-Rahman barostat,[39] respectively. Simulations were conducted at temperatures from 5 °C to 65 °C in 5 °C increments at 1 bar pressure. We examined solutions of methane, ethane, p- and o-xylene in water and ethanol. Water was modeled using the TIP4P/2005 potential.[87] Ethanol and the organic solutes were modeled using the TraPPE-United Atom potential.[88] Cross interactions were evaluated using standard Lorentz-Bethelot mixing rules.

The methane and methyl group interactions with water's oxygen utilized methane-oxygen and methyl-oxygen Lennard-Jones interaction diameters of 3.436 Å and 3.446 Å and well-depths of 1.023 kJ/mol and 0.8115 kJ/mol, respectively, while interactions for xylene's aromatic carbons with water were obtained using Lorentz-Berthelot combining rules.[53] These methane and methyl group interactions were optimized to reproduce the experimental hydration free energies of methane and ethane. We note that while the methane parameters were previously developed for the Hydrophobic Hydration potential,[89] the methyl units were re-optimized here to yield improved agreement with experiment. Cross interactions between the hydrocarbons and ethanol were evaluated using Lorentz-Berthelot combining rules. Short range van der Waals interactions were truncated at 10 Å with mean-field energy and pressure dispersion corrections applied beyond the cut-off. Electrostatic interactions were evaluated using particle-mesh Ewald summation.[10c] Internal constraints for water were held fixed using the SETTLE[71] algorithm. The LINCS algorithm[54] was used to constrain solute bonds in ethanol and ethane, while SHAKE,[90] which is computationally slower than LINCS, was required for the xylene simulations in order to constrain both angles and bonds in the xylene. While the TraPPE-UA potential specifies rigid improper dihedral angles to maintain the planarity of xylene's aromatic ring, these could not be readily implemented in GROMACS. Subsequently, OPLS-UA aromatic dihedral and improper dihedral angles were used for xylene. Equations of motion were evaluated using a 2 fs time step.

In the first set of simulations performed, 10 methanes were placed in simulation boxes of either 1000 waters or 307 ethanols. Simulations were conducted for 100 ns following at least 10 ns of equilibration. Configurations were saved every picosecond for

evaluation of methane-methane radial distribution functions and methane overlap free energies. Methane overlap free energies required for Equation B.1, *i.e.*, $\omega_{mm}^{j}(0)$, were determined as the excess hydration free energy difference, in Ben-Naim's standard state, of two overlapping methane less that of two methanes in bulk solution. We evaluated these free energies using Widom's test particle insertion method.[91] In practice the free energy difference is determined as the free energy for inserting a test methane directly on top of a methane in solution with those two methanes not interacting with one another less the free energy of randomly inserting a single methane in bulk solution. In the case of direct insertion on top of a pre-existing methane to evaluate the overlap free energy, insertions were performed over each of the 10 methanes in each saved configuration. For evaluation of the free energy of an individual methane in solution, 10,000 random insertions were attempted over each saved configuration.

In the second set of simulations performed, the excess hydration free energies and free energy differences were evaluated for methane, ethane, and p- and o-xylene at infinite dilution in water and ethanol. In these simulations we considered one solute in a box of 280 to 400 waters (depending on the solute size) or 95 to 151 ethanols. Solvation free energies were evaluated via the multistate Bennett acceptance ratio (MBAR) method, using the python alchemical analysis script frontend for pymbar for the computation and uncertainty analysis.[56] Interactions between initial and final states in these free energy calculations were transformed using a coupling parameter λ from 0 (initial state) to 1 (final state) in increments of 0.1 for a total of eleven states to evaluate the free energy. Averages for MBAR analysis were conducted over simulations of 10 ns at each value of λ following at least 1 ns for equilibration. For methane and ethane

the full solvation free energy relative to the vacuum in Ben-Naim's standard state was determined for each solute, while the free energy difference between p- and o-xylene was determined via alchemical shifting of one of the methyl units from the para to ortho position on the aromatic ring relative to a fixed methyl unit. To determine the impact of the interaction differences between methane and the methyl units of ethane, we evaluated free energies for transforming ethane's methyl units into methane units using Widom's insertion technique.[91] We refer to this ethane comprised of two bonded methanes as m/ethane below.

The free energies and PMFs evaluated from our simulations described above were broken down into their entropic and enthalpic components by fitting their temperature dependencies to the function

$$g = \alpha + \beta(T - T_0) + \gamma T \ln \frac{T}{T_0}, \tag{B.2}$$

which assumes a constant heat capacity. In this expression g is either the solute's excess solvation free energy or the PMF, $\omega_{mm}^{j}(r)$, while T_0 is a reference temperature taken here to be 298.15 K (25 °C). The parameters α, β, and γ are fitting constants in the case of solvation free energies, or separation-dependent fitting functions in the case of the PMF. Entropies and enthalpies were evaluated from the temperature derivatives of Equation B.2, i.e., $s = -\delta g/\delta T|_P$ and $h = \delta(g/T)/\delta(1/T)|_P$.

B.4 Results and discussion

The solvent-mediated interaction between two methanes in water at 25 °C is reported in Figure B.1. The portion of the PMF determined directly from the radial distribution function evaluated from molecular simulation, $\omega_{mm}^{wat}(r)|_{sim}$, extend from bulk solution down to a separation of 3.15 Å . Barring the sharp down turn in $\omega_{mm}^{wat}(r)|_{sim}$ below r < 3.25 Å , which results from poor sampling in this region, the PMF evaluated from simulation at small separations points down towards the methane overlap limit, $\omega_{mm}^{wat}(0)$, determined from methane insertion calculations. Equation B.1 smoothly interpolates the divide between complete overlap and the simulation accessible regime, providing a reasonable description of the water-mediated methane interaction over all separations (Figure B.1). Based on the quality of the description shown here and our previous investigations,[83e] we have confidence that Equation B.1 provides an accurate description of the solvent-mediated between methanes across all separations.

The temperature dependence of the solvent-mediated methane interaction in water and ethanol was subsequently determined by fitting the results of Equation B.1 determined over the temperature range 5–65 °C to Equation B.2. Taking appropriate temperature derivatives of the PMF, the free energy, enthalpy, and entropy of methane association in water and ethanol at 25 °C are reported in Figure B.2. While the methane PMFs in both water and ethanol are attractive down into the excluded volume regime (Figure B.2a and d), the attraction is strongest in water, consistent with the idea that water promotes assembly. While the enthalpy of association is attractive in both solvents (Figure B.2b and e), the enthalpic attraction is stronger in alcohol than it is in
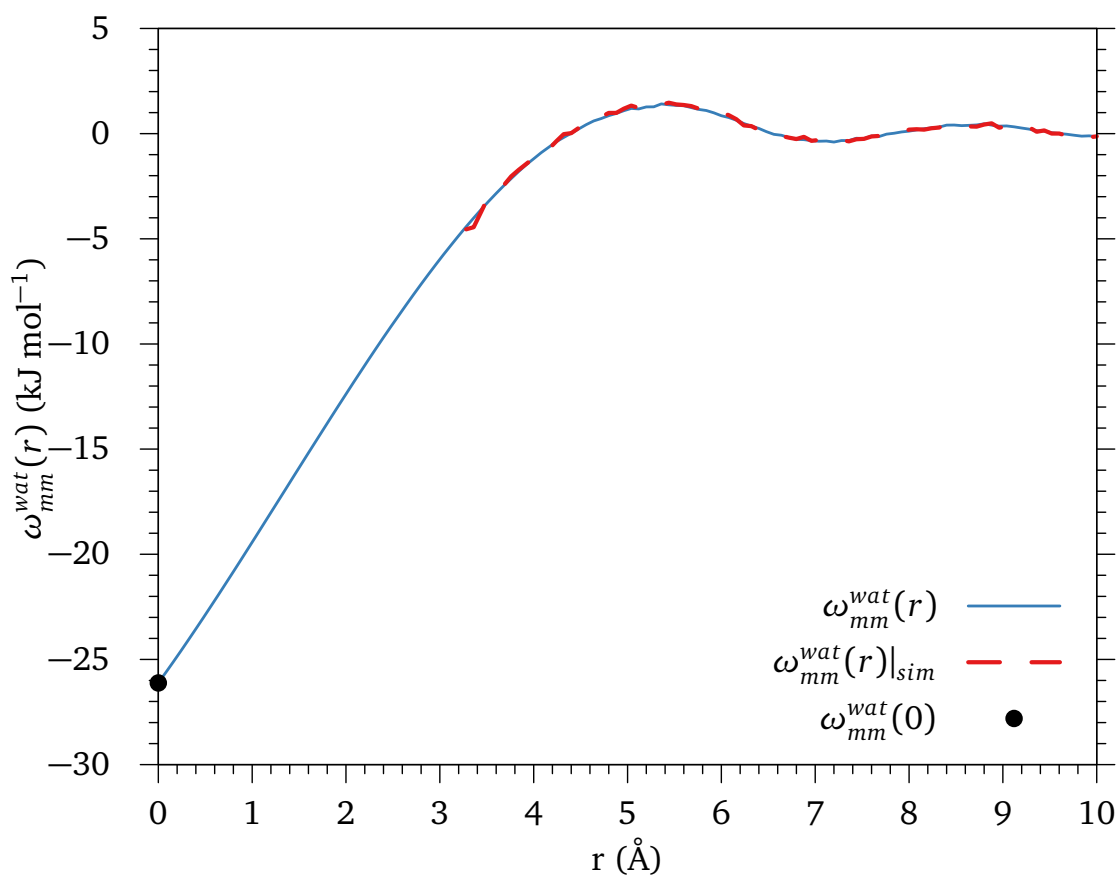
Figure B.1: Water-mediated contribution to the methane-methane potential-of-mean force at a temperature of 25 °C. Results are shown for the potential-of-mean force obtained directly from the simulation radial distribution function, $\omega_{mm}^{wat}(r)|_{sim}$, the free energy at methane overlap $\omega_{mm}^{wat}(0)$, and for the fit of Equation B.1 bridging the simulation and overlap free energies, $\omega_{mm}^{wat}(r)$. The symbols are defined in the figure legend.

water. In water the methane interaction entropy is attractive down into the excluded volume region (Figure B.2c), while in ethanol the interaction entropy is repulsive (Figure B.2f). This interaction entropy disparity between water and organic solvents has been taken as a hydrophobic interaction signature.

Ben Naim suggested that the solvent-mediated interaction between methanes at a fixed separation of 1.54 Å (the bond length of ethane) could be estimated from the difference between the excess chemical potential of ethane in $j =$ water or ethanol less twice that of methane

$$\omega_{mm}^{j}(1.54 \text{ Å}) \approx \delta\mu_{ethane}^{ex,j} = \mu_{ethane}^{ex,j} - 2\mu_{methane}^{ex,j}.$$ (B.3)

This approximation effectively assumes that ethane's two methyl units are equivalent to two methanes. This relationship can be made exact in our simulations by replacing the free energy of ethane for that of m/ethane (that is an ethane with the methyl units alchemically transformed into methanes) in Equation B.3 (*i.e.*, $\omega_{mm}^{j}(1.54 \text{ Å}) = \delta\mu_{m/ethane}^{ex,j} = \mu_{m/ethane}^{ex,j} - 2\mu_{methane}^{ex,j}$). The excess solvation thermodynamic properties for methane, ethane, and m/ethane in water and ethanol at 25 °C needed to estimate the methane interaction thermodynamics at r $= 1.54$ Å following Ben-Naim's prescription are reported in Table B.1. The simulation results for methane and ethane are in excellent agreement with the available experimental numbers,[92] giving us confidence in the fidelity of the computational models used for describing solvation.

The interaction thermodynamics between methanes evaluated from Equation B.3 are reported in Table B.2 and compared against the solvent-mediated PMFs in Fig-
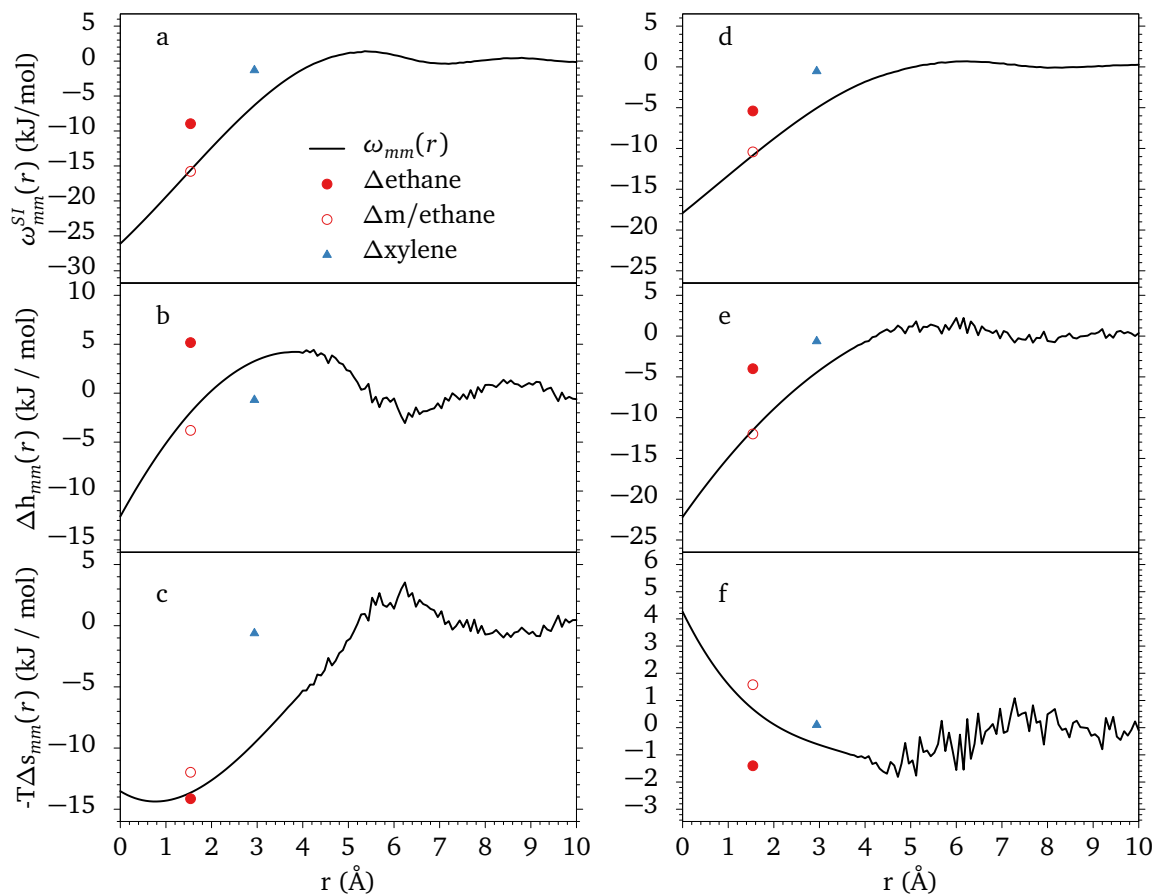
Figure B.2: Free energy (a and d), enthalpy (b and e), and entropy (c and f) of methane association in water (a-c) and ethanol (d-f) at 25 °C and 1 bar as a function of separation from direct overlap (r = 0) to a separation of 12 Å. The figure symbols are defined in the legend on the right hand side of the figure. The entropies and enthalpies were determined by fitting the PMFs and solvation free energies to Equation B.2 and taking the appropriate temperature derivatives.

ure B.2. While we find that the interaction free energies predicted from ethane's solvation free energy, $\Delta\mu_{ethane}^{ex}$, are attractive in both water and ethanol, the predicted interaction free energies are more positive than the solvent-mediated PMF at r = 1.54 Å (Figure B.2a and d). The free energies determined from m/ethane's solvation free energy, $\delta\mu_{ethane}^{ex,j}$, are in quantitative agreement with the interpolated PMF obtained from Equation B.1 in the excluded volume regime (Figure B.2a and d), however, indicating that the alchemical differences between methane and a methyl group is sufficient to account for Ben-Naim's prediction of a lower solvophobic interaction free energy.

Similar to the interaction free energies, the interaction enthalpies in either solvent predicted by Equation B.3 from ethane solvation data are more positive and repulsive than those obtained from the solvent-mediated PMF (Figure B.2b and e). The magnitude of the enthalpy difference is between 8 to 9 kJ/mol at room temperature (Table B.2). This difference is significant enough in water such that the predicted methane interaction enthalpy in water is positive, while we observe a negative interaction enthalpy from the PMF (Figure B.2b) and our results from m/ethane hydration (Table B.2).

In difference to the enthalpy, the interaction entropies predicted from Ben-Naim's approximation are more attractive than that obtained from the solvent-mediated PMF (Figure B.2c and f). The magnitude of this difference accounts for a 2-3 kJ/mol lower entropic contribution predicted by Equation B.3 in both water and ethanol (Table B.2). While smaller in magnitude than the enthalpy, this difference is significant enough in the case of ethanol-mediated interactions that the entropy predicted from Ben-Naim's

approximation is attractive while that obtained from the PMF is repulsive (Figure B.2f), in agreement with the results obtained from m/ethane solvation (Table B.2).

While their diameters are practically the same, the interaction well-depth of methane is deeper than that of an ethane methyl unit in the simulation models used. This observation suggests the reason Ben-Naim's approximation predicts a more repulsive enthalpy is the result of attractive interaction differences of a methyl unit and methane with the solvent. Assuming minimal changes in the solvent structure and volume for alchemically changing ethane's methyl units into methanes, the excess solvation free energy between m/ethane and ethane is approximately given as

$$\mu_{m/ethane}^{ex,j} \approx \mu_{ethane}^{ex,j} + \langle \delta\phi_{e2m} \rangle_{ethane}, \tag{B.4}$$

where the perturbation contribution $\langle \delta\phi_{e2m} \rangle_{ethane}$ is the potential energy difference between treating ethane's methyl units as united-atom methanes versus united-atom methyl units averaged over configurations generated from our ethane simulations. The solvation free energies, enthalpies, and entropies at 25 °C determined from Equation B.4 are in excellent quantitative agreement with the exact values for the m/ethane reported in Table B.1. Resultantly the perturbation corrected interaction thermodynamic predictions are in harmony with those determined from the solvent-mediated PMF (Table B.2). We can conclude that the difference between Ben-Naim's predicted interaction thermodynamics and the solvent-mediated PMF within the excluded volume regime results from interaction differences between a methane and methyl unit with the solvent, and not from changes in the solvent structure around methanes or methyls held together

at a 1.54 Å separation. While Equation B.4 only relies on energetic differences between methyl units and methane, the temperature dependence of this interaction energy difference is sufficient to account for the entropic differences between that predicted from Ben-Naim's estimate and the simulation PMF.

It has been suggested that the solvation free energy difference between o- and p-xylene provides an approximation to the solvent-mediated interaction between methanes at a separation of 2.94 Å (the distance between the methyl units on o-xylene), not as deep into the excluded volume regime as determined from ethane.[93] The solvation thermodynamic differences between the xylene isomers, $\Delta\theta^{ex,j}_{xylene} = \theta^{ex,j}_{o-xylene} - \theta^{ex,j}_{p-xylene}$ (where $\theta = \mu$, h,or s), evaluated from our alchemical change simulations are reported in Table B.2. While we were unable to find experimental results for xylene solvation in ethanol or the temperature derivatives of the free energy in water, the experimental and simulation solvation free energy differences in water are in excellent agreement. The xylene difference free energies in water and ethanol determined from simulation indicate that the methyl group interaction is favorable and strongest in water. Moreover, the simulations indicate that the xylene methyl unit interaction is entropically favorable in water and unfavorable in ethanol. When we compare the methyl unit interaction thermodynamics predicted from the p-/o-xylene difference calculation against the solvent-mediated methane-methane PMFs, however, the comparison is weak (Figure B.2). Specifically, compared to the methane-methane interaction thermodynamics the results estimated from xylene solvation are negligible. It may be noted that the difference between p- and o-xylene more accurately represents the free energy difference of bringing a methyl unit from a separation of 5.86 Å (the methyl unit separation in

p-xylene) to 2.94 Å. Examining the methane-methane PMFs, however, we find that the PMF values at 5.86 Å are significantly closer to zero than at 2.94 Å, suggesting the difference between these two separations is comparable to the free energy of bringing the two carbons together from infinity. We have also performed free energy perturbation calculations to transform xylene's methyl units into methanes; however, this calculation barely changed the results obtained with the methyl group interactions.

We attribute the lower magnitude xylene estimates of the interaction free energy to the incomplete solvation of the methyl units, whose solvation is blocked by xylene's aromatic ring barring the formation of a complete solvation shell about the methyl units for either isomer. Indeed, if we compare the change in the solvent accessible surface area[94] between xylene isomers versus the change in area between bringing two methanes to the methyl unit separation in ortho-xylene we find significant differences. The area difference between ortho- and para-xylene is $\sim$8 Å^2, while that for bringing methanes together is $\sim$76 Å^2 (The van der Waals radii of water, CH3/4, and the aromatic carbons were assumed to be 1.4 Å , 1.9 Å, and 1.7 Å, respectively in this calculation). The difference in solvent exposure for xylene is 10% that of two methanes, suggesting the solvent contribution to the PMF estimated from xylene will be negligible compared to that of two fully hydrated methanes.

B.5   Conclusion

In conclusion, using a previously developed empirical bridge describing the solvent-mediated interaction between methanes outside their mutually excluded volume down to complete overlap, we have constructed a description for the solvent-mediated PMF

over all separations using information garnered simply from molecular simulations of methanes in solution. Comparing the interaction thermodynamics determined from this empirical bridge against earlier experimental estimates of the interaction between methanes at separations of 1.54 Å and 2.94 Å, obtained from the solvation free energy difference between ethane and methane as well as p- and o-xylene, we find distinct differences between these two approaches. Specifically, while Ben-Naim's experimental estimate correctly predicts stronger interactions in water versus ethanol, it under predicts the magnitude of the solvent-mediated attraction in both solvents. Moreover, Ben-Naim's estimate incorrectly predicts methane interactions at 1.54 Å (the carbon-carbon bond length of ethane) are enthalpically unfavorable in water and entropically favorable in ethanol. The difference between Ben-Naim's estimates based on the differences between ethane and methane and the present simulation results were shown to be a result of alchemical distinctions between the methyl units of ethane and methane. When these are taken into account following a simple linear perturbation theory approach to account for the direct interaction difference between ethane's methyl units and methane with the solvent, Ben-Naim's estimate can be reconciled with methane's potential-of-mean force. While interaction estimates made based on the solvation differences between xylene isomers predict stronger interactions in water compared to ethanol, the magnitude of the estimated interactions are weaker than those observed for methane's in solution largely due to the incomplete solvation of the methyl units blocked by xylene's aromatic ring.

We believe these are the first results bringing these two distinct methods for evaluating methane hydrophobic interactions into harmony with one another. Once alchem-

ical differences are taken into account, the agreement between Ben-Naim's estimates and our simulations gives us confidence that simulations using models optimized to reproduce solvation thermodynamic provide an accurate representation of the solvent-mediated interactions.

	Water			Ethanol		
	$\mu_i^{ex,wat}$ (kJ/mol)	$h_i^{ex,wat}$ (kJ/mol)	$-Ts_i^{ex,wat}$ (kJ/mol)	$\mu_i^{ex,alc}$ (kJ/mol)	$h_i^{ex,alc}$ (kJ/mol)	$-Ts_i^{ex,alc}$ (kJ/mol)
Simulation						
Methane	8.30 (0.02)	-11.22 (0.44)	19.52 (0.44)	1.51 (0.01)	-2.15 (0.24)	3.66 (0.24)
Ethane	7.64 (0.02)	-17.27 (0.52)	24.91 (0.52)	-2.38 (0.02)	-8.30 (0.29)	5.29 (0.29)
M/ethane	0.83 (0.03)	-26.24 (0.43)	27.06 (0.42)	-7.40 (0.01)	-16.30 (0.31)	8.90 (0.31)
Ethane + perturbation	0.93 (0.05)	-26.18 (0.55)	27.12 (0.58)	-7.82 (0.03)	-16.71 (0.44)	8.88 (0.43)
Experiment						
Methane	8.39	-11.50	19.89	1.58	-2.81	4.39
Ethane	7.67	-17.47	25.14	-2.50	-9.44	6.94

Table B.1: Excess solvation thermodynamic properties for hydrocarbons in water and ethanol at 25 °C and 1 bar from simulation and experiment. Simulation error bars reported in parentheses correspond to one standard deviation. Experimental results were obtained from Refs. 92.

	Water			Ethanol		
	$\delta\mu_i^{ex,wat}$ (kJ/mol)	$\delta h_i^{ex,wat}$ (kJ/mol)	$-T\delta s_i^{ex,wat}$ (kJ/mol)	$\delta\mu_i^{ex,alc}$ (kJ/mol)	$\delta h_i^{ex,alc}$ (kJ/mol)	$-T\delta s_i^{ex,alc}$ (kJ/mol)
Simulation						
Methane	-8.96 (0.05)	5.17 (0.77)	-14.14 (0.79)	-5.40 (0.04)	-4.00 (0.39)	-1.40 (0.39)
M/ethane	-15.78 (0.06)	-3.80 (0.97)	-11.98 (0.96)	-10.42 (0.03)	-12.00 (0.36)	1.58 (0.35)
Ethane + perturbation	-15.67 (0.07)	-3.74 (0.79)	-11.93 (0.75)	-10.84 (0.03)	-12.41 (0.51)	1.56 (0.50)
Xylene (ortho-para)	-1.31 (0.02)	-0.69 (0.15)	-0.62 (0.15)	-0.53 (0.01)	-0.63 (0.07)	0.10 (0.07)
Experiment						
Ethane	9.10	5.53	-14.63	-5.65	-3.81	-1.84
Xylene (ortho-para)	-1.27	N/A	N/A	N/A	N/A	N/A

Table B.2: Estimates for solvent-mediated methane-methane interaction thermodynamics determined from excess solvation properties at 25 °C and 1 bar following Ben-Naim's difference approach using both simulation and experimental data. Simulation error bars reported in parentheses correspond to one standard deviation. Experimental results were obtained from Refs. 92.

Appendix C

SUPPORTING INFORMATION FOR "SUCCESSION OF ALKANE CONFORMATIONAL MOTIFS BOUND WITHIN HYDROPHOBIC SUPRAMOLECULAR CAPSULAR ASSEMBLIES"

Here we detail the relevant structural and interaction information used to model the Octa-acid host. In addition, we provide parity plots comparing the predicted versus experimental changes in the encapsulated guest proton chemical shift and mean distance plots for all the alkanes simulated. Figure C.1 shows a snapshot of single Octa-acid host rendered using the coordinates reported in Table C.1 in standard Protein Data Bank format. Table C.2 lists the partial charges and non-bonded Generalized Amber Force Field[51] (GAFF) atom types used to model the host electrostatic and Lennard-Jones interactions. Partial charges reported in this table were obtained from AM1-BCC charge fitting following geometry optimization as described in Ref. 33. Figure C.2 compares the experimental $\Delta\delta$'s for each alkane and those evaluated from GIAO calculations performed on the dominant encapsulated guest conformation. Figure C.3 shows the mean-distance plots for each encapsulated alkane.
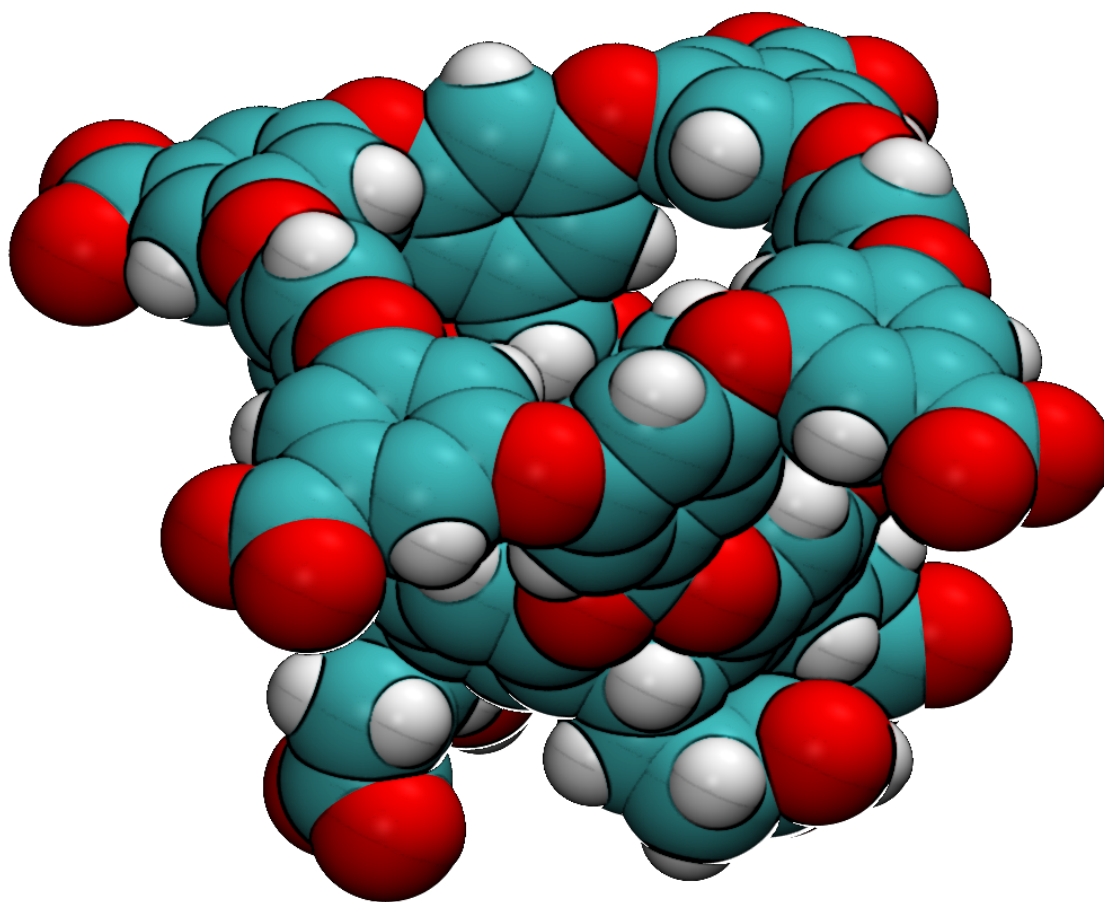
Figure C.1: Snapshot of a single Octa-acid cavitand from the simulations performed here. Protein data bank structure reported in Table C.1. Partial charges and GAFF atom types reported in Table C.2.
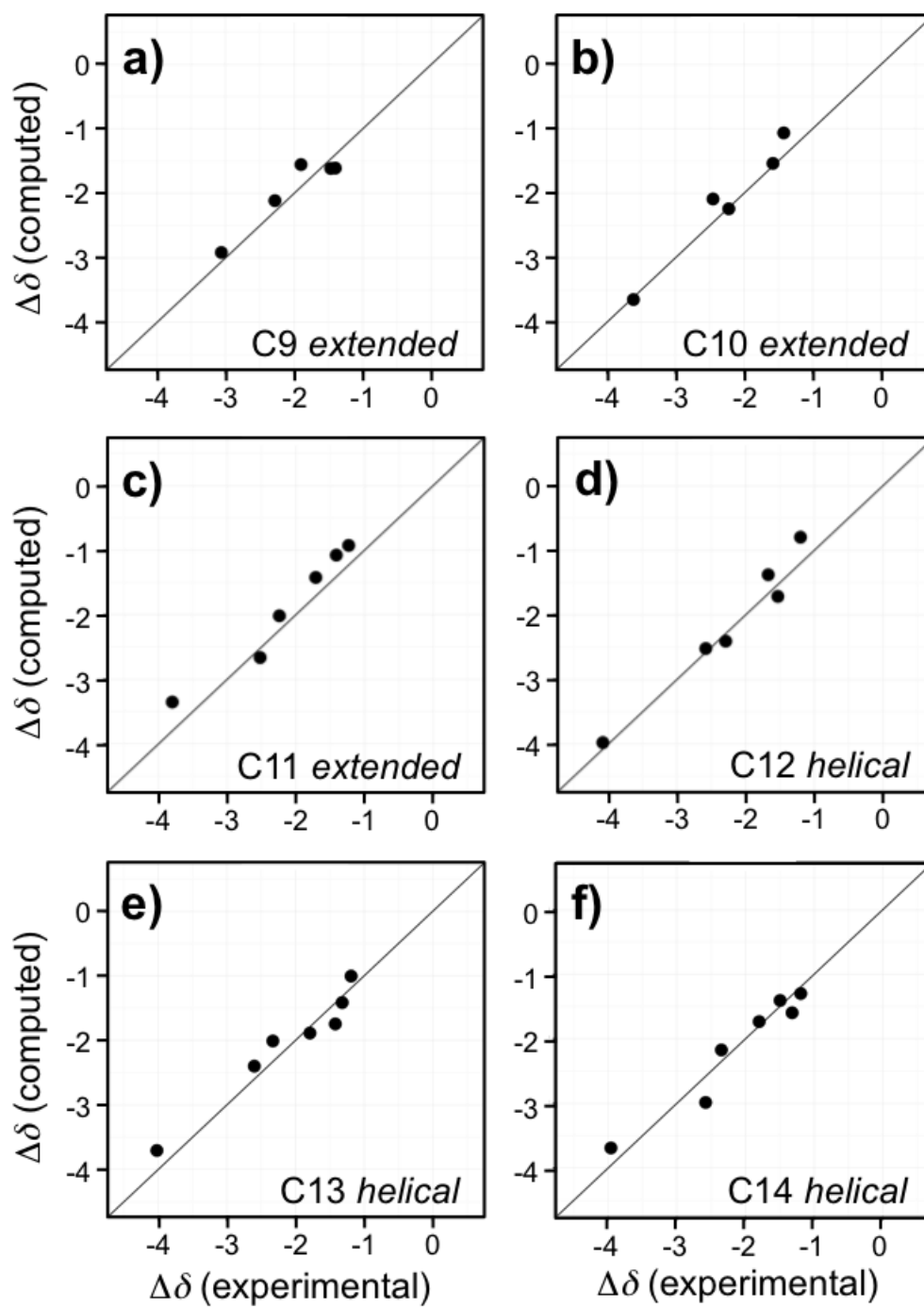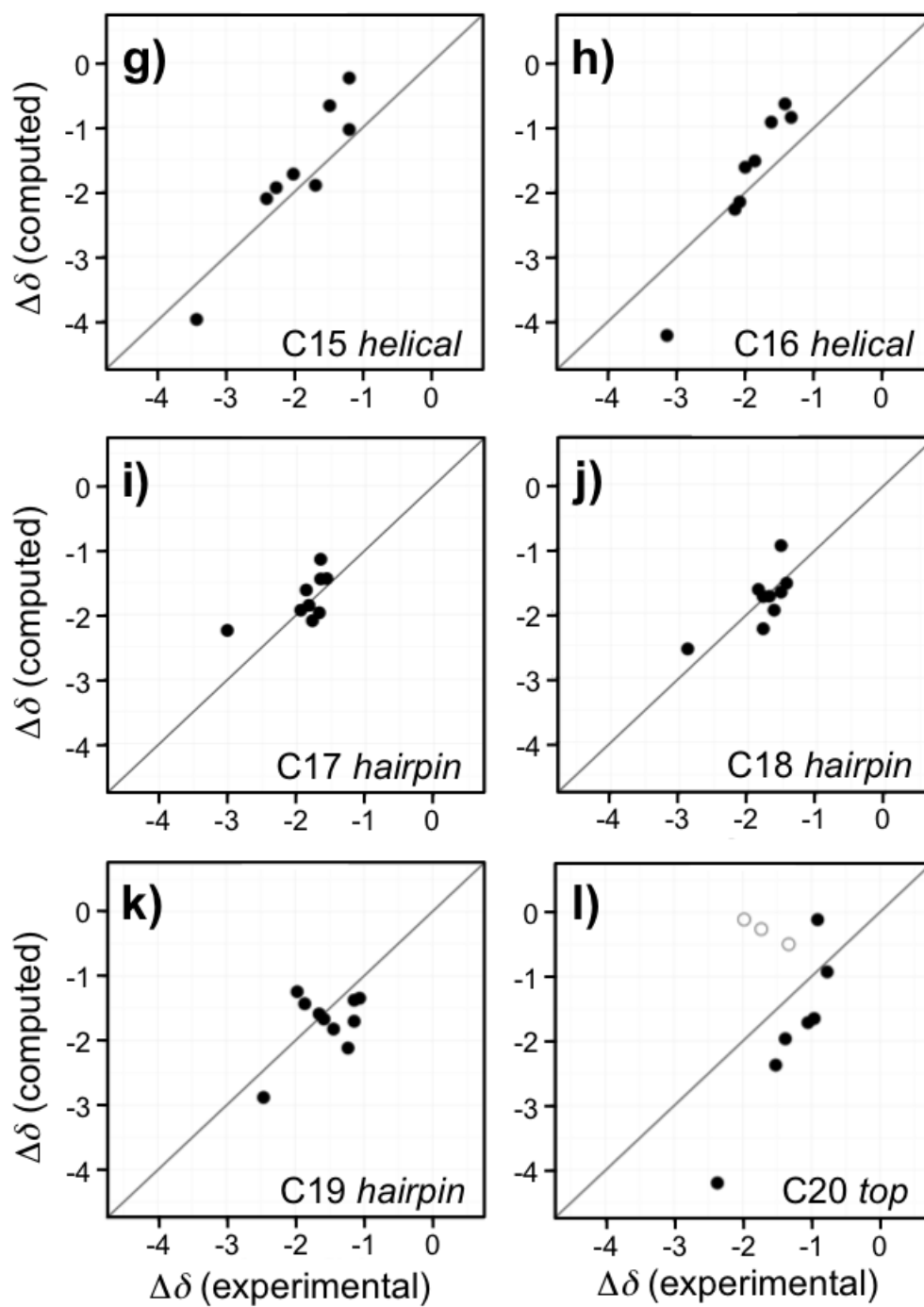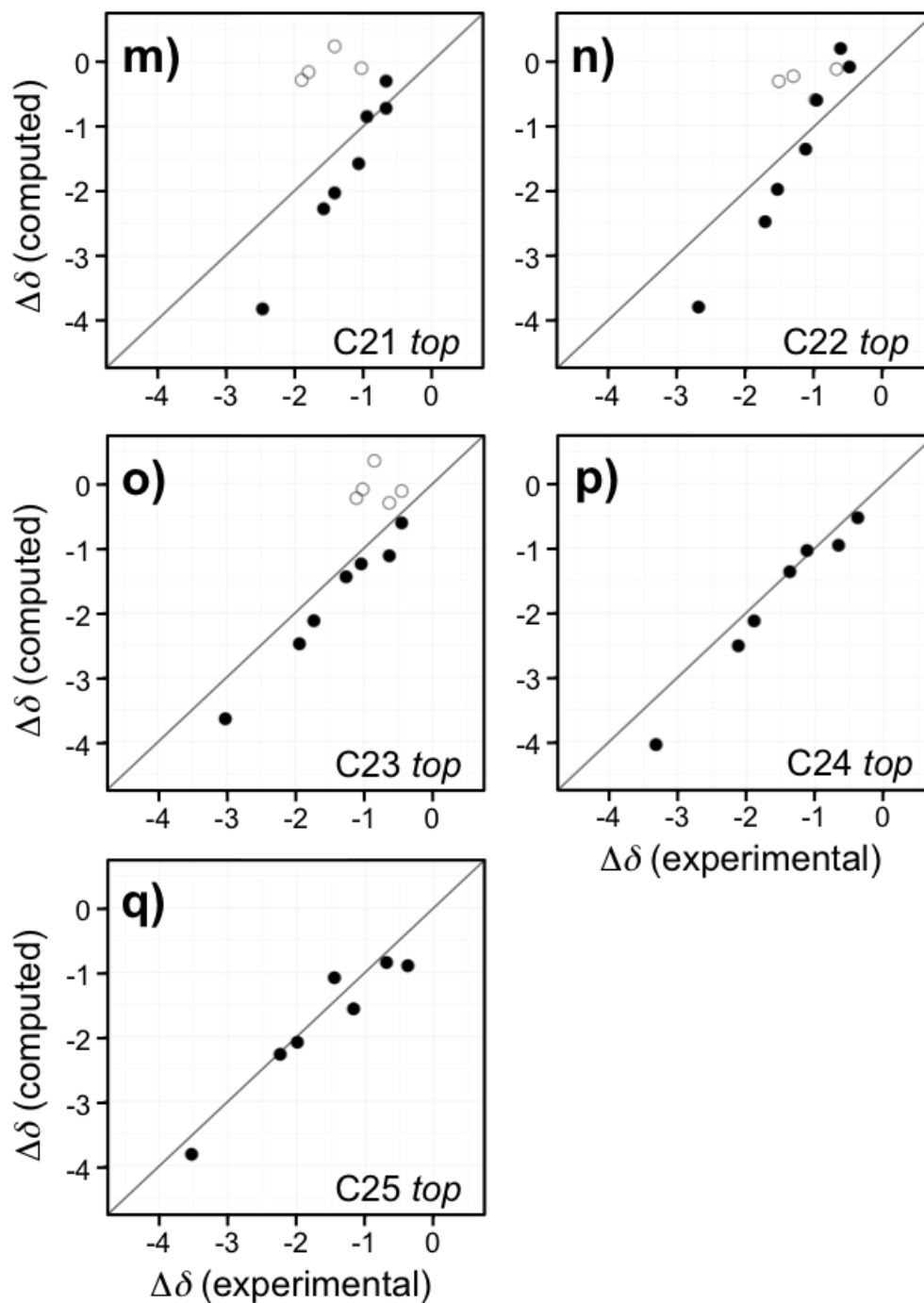
Figure C.2: Comparison between the experimental and GIAO calculated changes in the ^{1}H NMR chemical shifts ($\Delta\delta$'s) for all the simulated alkanes. Figures a through q show results and motif assignments for alkanes C_9 through C_{25}. The filled points indicate protons contained within the OA dimer in the simulation, while the open points indicate protons sandwich between the two cavitands and partially exposed to the solvent.

a) C9 *extended* b) C10 *extended* c) C11 *extended* d) C12 *helical* e) C13 *helical* f) C14 *helical*
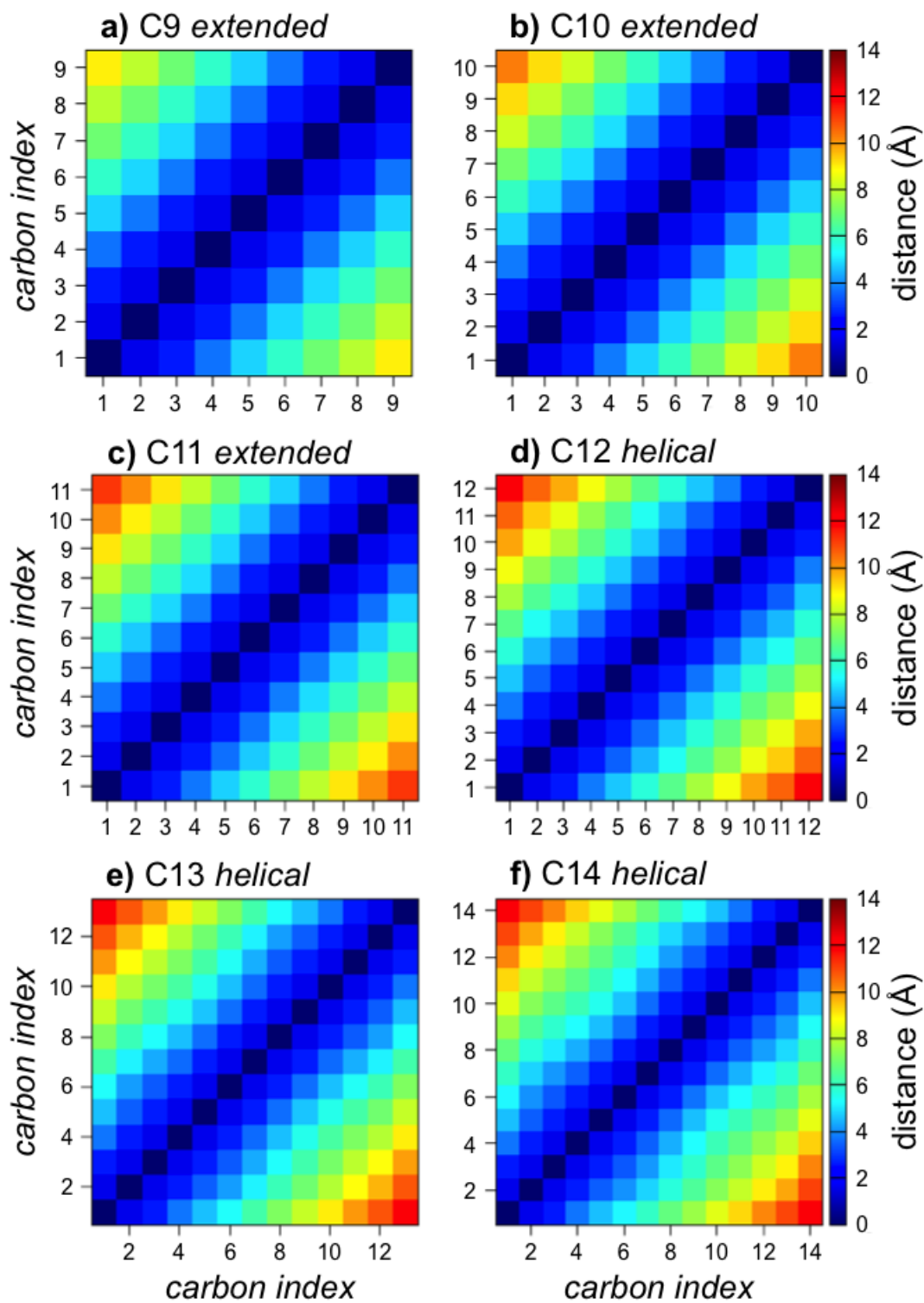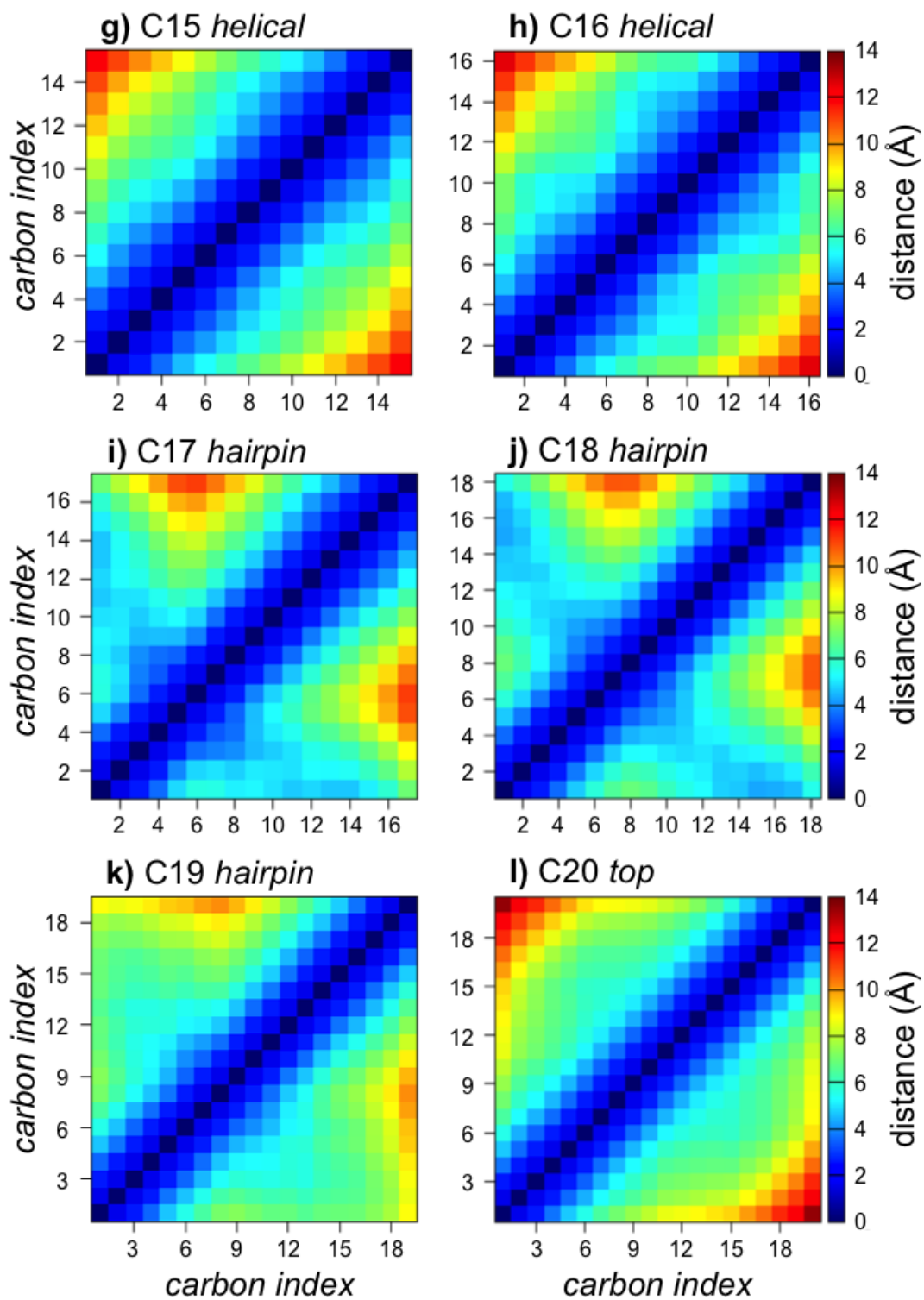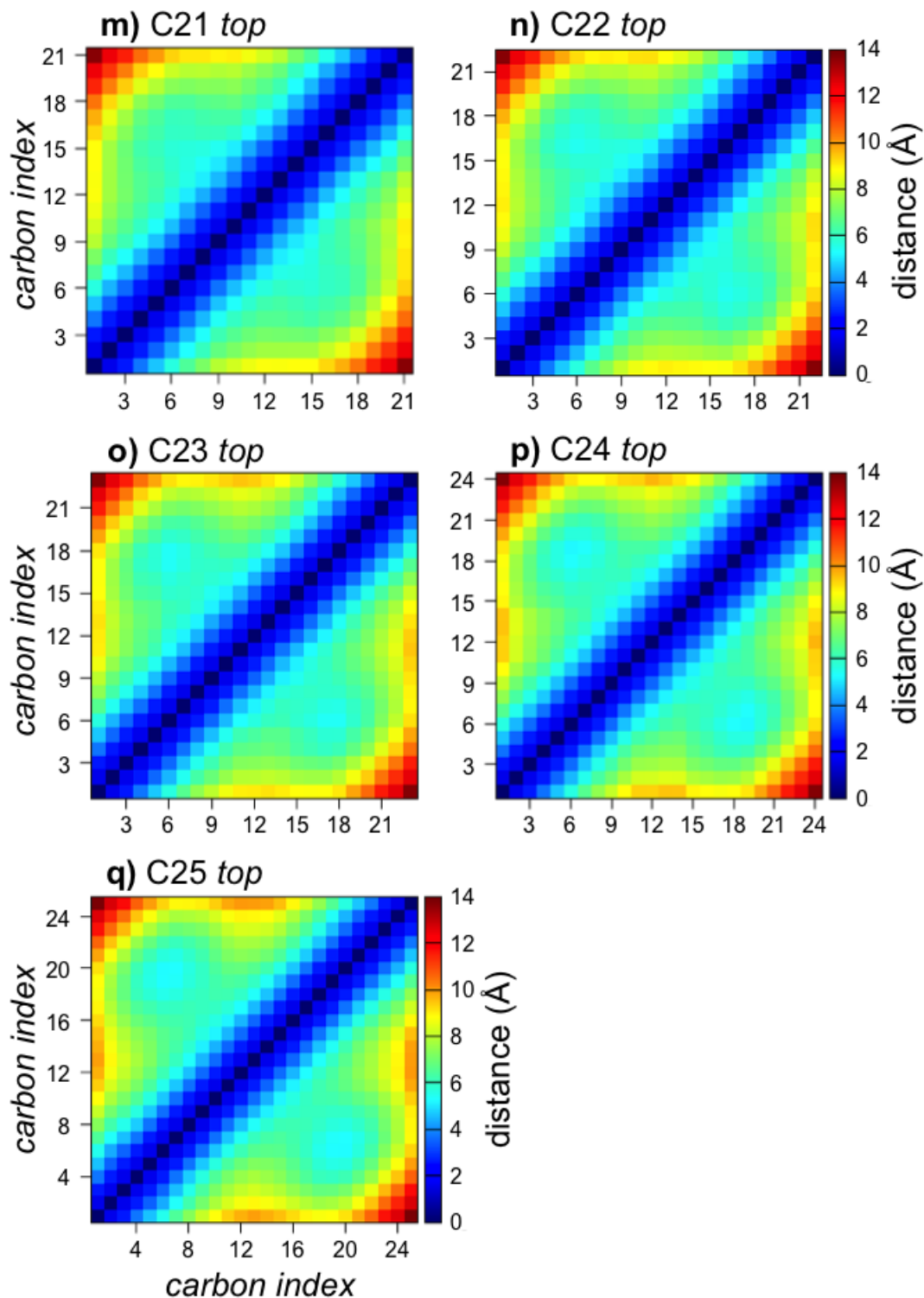
Figure C.3: Encapsulated alkane mean distance maps for all the simulated alkanes. Figures *a* through *q* show results and motif assignments for alkanes C_9 through C_{25}. Distances are keyed to the color key on the right-hand side of the figures.

```
TITLE     OCTA-ACID
AUTHOR    J. W. BARNETT
CRYST1   36.221    35.619    32.285   90.00   90.00   90.00 P 1              1
MODEL         1
ATOM      1  O29 OCT     1      -24.645  20.951  28.114  1.00  0.00           O
ATOM      2  C93 OCT     1      -25.241  20.054  27.507  1.00  0.00           C
ATOM      3  O30 OCT     1      -25.752  19.180  28.211  1.00  0.00           O
ATOM      4  C30 OCT     1      -25.263  20.031  26.008  1.00  0.00           C
ATOM      5  H49 OCT     1      -26.285  20.174  25.655  1.00  0.00           H
ATOM      6  H50 OCT     1      -24.601  20.790  25.666  1.00  0.00           H
ATOM      7  C29 OCT     1      -24.785  18.608  25.587  1.00  0.00           C
ATOM      8  H13 OCT     1      -25.513  17.912  26.110  1.00  0.00           H
ATOM      9  H14 OCT     1      -23.764  18.453  26.013  1.00  0.00           H
ATOM     10  C8  OCT     1      -24.707  18.542  24.065  1.00  0.00           C
ATOM     11  H4  OCT     1      -25.610  18.771  23.672  1.00  0.00           H
ATOM     12  C5  OCT     1      -23.741  19.425  23.413  1.00  0.00           C
ATOM     13  C2  OCT     1      -22.356  19.493  23.887  1.00  0.00           C
ATOM     14  H2  OCT     1      -22.103  18.860  24.797  1.00  0.00           H
ATOM     15  C6  OCT     1      -21.453  20.449  23.424  1.00  0.00           C
ATOM     16  C7  OCT     1      -19.993  20.553  23.929  1.00  0.00           C
ATOM     17  H3  OCT     1      -19.744  21.592  23.673  1.00  0.00           H
ATOM     18  C9  OCT     1      -19.126  19.596  23.035  1.00  0.00           C
ATOM     19  C12 OCT     1      -18.898  18.264  23.513  1.00  0.00           C
ATOM     20  H5  OCT     1      -19.236  17.983  24.497  1.00  0.00           H
ATOM     21  C31 OCT     1      -19.888  20.567  25.471  1.00  0.00           C
ATOM     22  H15 OCT     1      -20.741  21.223  25.758  1.00  0.00           H
ATOM     23  C32 OCT     1      -18.592  21.150  26.073  1.00  0.00           C
ATOM     24  H55 OCT     1      -18.715  21.353  27.151  1.00  0.00           H
ATOM     25  H56 OCT     1      -18.294  22.088  25.541  1.00  0.00           H
ATOM     26  C96 OCT     1      -17.480  20.130  25.841  1.00  0.00           C
ATOM     27  O31 OCT     1      -17.328  19.125  26.565  1.00  0.00           O
ATOM     28  O32 OCT     1      -16.451  20.683  25.308  1.00  0.00           O
ATOM     29  H58 OCT     1      -15.650  20.194  25.644  1.00  0.00           H
ATOM     30  H16 OCT     1      -20.135  19.571  25.914  1.00  0.00           H
ATOM     31  C3  OCT     1      -21.893  21.287  22.358  1.00  0.00           C
ATOM     32  O4  OCT     1      -21.103  22.059  21.580  1.00  0.00           O
ATOM     33  C1  OCT     1      -23.239  21.261  21.977  1.00  0.00           C
ATOM     34  H1  OCT     1      -23.614  21.943  21.254  1.00  0.00           H
ATOM     35  C4  OCT     1      -24.103  20.279  22.328  1.00  0.00           C
ATOM     36  O7  OCT     1      -25.243  20.108  21.532  1.00  0.00           O
ATOM     37  C39 OCT     1      -25.271  18.944  20.723  1.00  0.00           C
ATOM     38  H23 OCT     1      -24.167  18.570  20.599  1.00  0.00           H
ATOM     39  C41 OCT     1      -25.991  19.212  19.394  1.00  0.00           C
ATOM     40  C43 OCT     1      -26.730  18.229  18.767  1.00  0.00           C
ATOM     41  H25 OCT     1      -27.102  17.422  19.332  1.00  0.00           H
ATOM     42  C44 OCT     1      -25.666  20.372  18.686  1.00  0.00           C
ATOM     43  H26 OCT     1      -25.032  21.184  19.118  1.00  0.00           H
ATOM     44  O8  OCT     1      -25.962  18.003  21.560  1.00  0.00           O
ATOM     45  C18 OCT     1      -25.199  17.001  22.162  1.00  0.00           C
ATOM     46  C19 OCT     1      -25.145  15.854  21.426  1.00  0.00           C
ATOM     47  H8  OCT     1      -25.644  15.644  20.500  1.00  0.00           H
ATOM     48  C15 OCT     1      -24.622  17.173  23.398  1.00  0.00           C
ATOM     49  C17 OCT     1      -23.729  16.216  23.824  1.00  0.00           C
ATOM     50  H7  OCT     1      -23.055  16.408  24.606  1.00  0.00           H
```

```
ATOM     51  C20 OCT     1     -23.697  14.948  23.162  1.00  0.00           C
ATOM     52  C21 OCT     1     -22.682  13.878  23.569  1.00  0.00           C
ATOM     53  H9  OCT     1     -23.012  12.921  23.198  1.00  0.00           H
ATOM     54  C35 OCT     1     -22.411  13.676  25.120  1.00  0.00           C
ATOM     55  H19 OCT     1     -21.651  12.886  25.227  1.00  0.00           H
ATOM     56  C36 OCT     1     -23.690  13.265  25.861  1.00  0.00           C
ATOM     57  H51 OCT     1     -23.945  12.244  25.490  1.00  0.00           H
ATOM     58  H52 OCT     1     -24.516  13.952  25.699  1.00  0.00           H
ATOM     59  C94 OCT     1     -23.309  13.205  27.298  1.00  0.00           C
ATOM     60  O27 OCT     1     -22.898  12.187  27.860  1.00  0.00           O
ATOM     61  O28 OCT     1     -23.903  14.053  28.048  1.00  0.00           O
ATOM     62  H57 OCT     1     -23.636  13.674  28.932  1.00  0.00           H
ATOM     63  H20 OCT     1     -21.987  14.642  25.452  1.00  0.00           H
ATOM     64  C16 OCT     1     -24.434  14.822  22.005  1.00  0.00           C
ATOM     65  O2  OCT     1     -24.374  13.739  21.211  1.00  0.00           O
ATOM     66  C40 OCT     1     -23.262  13.638  20.263  1.00  0.00           C
ATOM     67  H24 OCT     1     -22.943  14.701  20.200  1.00  0.00           H
ATOM     68  C47 OCT     1     -23.777  13.281  18.869  1.00  0.00           C
ATOM     69  C49 OCT     1     -25.112  13.258  18.467  1.00  0.00           C
ATOM     70  H28 OCT     1     -25.844  13.527  19.135  1.00  0.00           H
ATOM     71  C50 OCT     1     -22.756  12.906  17.955  1.00  0.00           C
ATOM     72  H29 OCT     1     -21.715  12.947  18.228  1.00  0.00           H
ATOM     73  O1  OCT     1     -22.182  12.815  20.861  1.00  0.00           O
ATOM     74  C24 OCT     1     -21.162  13.511  21.544  1.00  0.00           C
ATOM     75  C28 OCT     1     -20.092  13.758  20.756  1.00  0.00           C
ATOM     76  C25 OCT     1     -19.012  14.497  21.329  1.00  0.00           C
ATOM     77  O3  OCT     1     -17.955  14.777  20.528  1.00  0.00           O
ATOM     78  C38 OCT     1     -17.837  16.069  19.885  1.00  0.00           C
ATOM     79  O6  OCT     1     -17.003  17.006  20.619  1.00  0.00           O
ATOM     80  H22 OCT     1     -18.792  16.587  19.831  1.00  0.00           H
ATOM     81  C53 OCT     1     -17.328  15.927  18.435  1.00  0.00           C
ATOM     82  C55 OCT     1     -17.434  14.754  17.795  1.00  0.00           C
ATOM     83  H31 OCT     1     -17.751  13.895  18.321  1.00  0.00           H
ATOM     84  C56 OCT     1     -16.679  16.975  17.772  1.00  0.00           C
ATOM     85  H32 OCT     1     -16.584  17.899  18.247  1.00  0.00           H
ATOM     86  H12 OCT     1     -20.139  13.402  19.726  1.00  0.00           H
ATOM     87  C27 OCT     1     -21.349  14.131  22.788  1.00  0.00           C
ATOM     88  C26 OCT     1     -20.277  14.886  23.300  1.00  0.00           C
ATOM     89  H11 OCT     1     -20.405  15.442  24.253  1.00  0.00           H
ATOM     90  C23 OCT     1     -19.116  15.108  22.617  1.00  0.00           C
ATOM     91  C22 OCT     1     -17.907  15.943  23.089  1.00  0.00           C
ATOM     92  H10 OCT     1     -17.065  15.534  22.567  1.00  0.00           H
ATOM     93  C33 OCT     1     -17.555  15.749  24.562  1.00  0.00           C
ATOM     94  H17 OCT     1     -16.794  16.466  24.962  1.00  0.00           H
ATOM     95  C34 OCT     1     -16.962  14.358  24.754  1.00  0.00           C
ATOM     96  H53 OCT     1     -16.235  14.102  23.923  1.00  0.00           H
ATOM     97  H54 OCT     1     -17.762  13.591  24.705  1.00  0.00           H
ATOM     98  C95 OCT     1     -16.105  14.313  26.009  1.00  0.00           C
ATOM     99  O25 OCT     1     -14.948  14.007  25.883  1.00  0.00           O
ATOM    100  O26 OCT     1     -16.703  14.093  27.038  1.00  0.00           O
```

```
ATOM    101  H18 OCT    1    -18.383  15.946  25.185  1.00  0.00           H
ATOM    102  C13 OCT    1    -18.092  17.435  22.733  1.00  0.00           C
ATOM    103  C10 OCT    1    -17.632  17.889  21.490  1.00  0.00           C
ATOM    104  C14 OCT    1    -17.789  19.232  21.106  1.00  0.00           C
ATOM    105  H6  OCT    1    -17.376  19.680  20.201  1.00  0.00           H
ATOM    106  C11 OCT    1    -18.528  20.081  21.859  1.00  0.00           C
ATOM    107  O5  OCT    1    -18.949  21.338  21.334  1.00  0.00           O
ATOM    108  C37 OCT    1    -20.239  21.445  20.643  1.00  0.00           C
ATOM    109  H21 OCT    1    -20.700  20.449  20.407  1.00  0.00           H
ATOM    110  C59 OCT    1    -20.156  22.157  19.304  1.00  0.00           C
ATOM    111  C62 OCT    1    -19.247  21.734  18.324  1.00  0.00           C
ATOM    112  H35 OCT    1    -18.606  20.872  18.579  1.00  0.00           H
ATOM    113  C61 OCT    1    -21.012  23.176  19.040  1.00  0.00           C
ATOM    114  H34 OCT    1    -21.644  23.500  19.887  1.00  0.00           H
ATOM    115  C64 OCT    1    -20.986  23.782  17.818  1.00  0.00           C
ATOM    116  C60 OCT    1    -20.082  23.380  16.826  1.00  0.00           C
ATOM    117  H36 OCT    1    -20.000  23.853  15.844  1.00  0.00           H
ATOM    118  C63 OCT    1    -19.270  22.288  17.070  1.00  0.00           C
ATOM    119  O15 OCT    1    -18.620  21.725  15.954  1.00  0.00           O
ATOM    120  C77 OCT    1    -17.385  21.102  16.166  1.00  0.00           C
ATOM    121  C74 OCT    1    -17.282  19.720  15.943  1.00  0.00           C
ATOM    122  H45 OCT    1    -18.174  19.141  15.833  1.00  0.00           H
ATOM    123  C73 OCT    1    -16.246  21.907  16.486  1.00  0.00           C
ATOM    124  H40 OCT    1    -16.361  22.976  16.701  1.00  0.00           H
ATOM    125  C76 OCT    1    -14.978  21.371  16.637  1.00  0.00           C
ATOM    126  C78 OCT    1    -13.844  22.238  17.102  1.00  0.00           C
ATOM    127  O19 OCT    1    -14.005  23.432  17.106  1.00  0.00           O
ATOM    128  O20 OCT    1    -12.726  21.796  17.226  1.00  0.00           O
ATOM    129  C75 OCT    1    -14.889  20.012  16.393  1.00  0.00           C
ATOM    130  H39 OCT    1    -13.923  19.578  16.387  1.00  0.00           H
ATOM    131  C72 OCT    1    -15.976  19.211  16.058  1.00  0.00           C
ATOM    132  O14 OCT    1    -15.684  17.905  15.802  1.00  0.00           O
ATOM    133  C57 OCT    1    -16.215  16.794  16.492  1.00  0.00           C
ATOM    134  C54 OCT    1    -16.555  15.685  15.804  1.00  0.00           C
ATOM    135  H33 OCT    1    -16.278  15.503  14.751  1.00  0.00           H
ATOM    136  C58 OCT    1    -17.079  14.632  16.477  1.00  0.00           C
ATOM    137  O13 OCT    1    -17.373  13.552  15.712  1.00  0.00           O
ATOM    138  C70 OCT    1    -18.353  12.662  15.933  1.00  0.00           C
ATOM    139  C67 OCT    1    -19.719  12.995  15.689  1.00  0.00           C
ATOM    140  H48 OCT    1    -19.955  13.966  15.339  1.00  0.00           H
ATOM    141  C66 OCT    1    -18.147  11.365  16.378  1.00  0.00           C
ATOM    142  H38 OCT    1    -17.128  11.002  16.563  1.00  0.00           H
ATOM    143  C69 OCT    1    -19.189  10.421  16.529  1.00  0.00           C
ATOM    144  C71 OCT    1    -18.914   8.934  16.796  1.00  0.00           C
ATOM    145  O21 OCT    1    -17.762   8.547  16.851  1.00  0.00           O
ATOM    146  O22 OCT    1    -19.758   8.075  16.831  1.00  0.00           O
ATOM    147  C68 OCT    1    -20.503  10.789  16.267  1.00  0.00           C
ATOM    148  H37 OCT    1    -21.384  10.121  16.487  1.00  0.00           H
ATOM    149  C65 OCT    1    -20.804  12.077  15.860  1.00  0.00           C
ATOM    150  O11 OCT    1    -22.181  12.333  15.656  1.00  0.00           O
```

```
ATOM    151  C51 OCT     1     -23.081  12.681  16.643  1.00  0.00           C
ATOM    152  C48 OCT     1     -24.428  12.829  16.225  1.00  0.00           C
ATOM    153  H30 OCT     1     -24.662  12.607  15.238  1.00  0.00           H
ATOM    154  C52 OCT     1     -25.432  13.123  17.125  1.00  0.00           C
ATOM    155  O10 OCT     1     -26.667  13.097  16.675  1.00  0.00           O
ATOM    156  C91 OCT     1     -27.605  13.971  17.078  1.00  0.00           C
ATOM    157  C88 OCT     1     -27.340  15.335  16.849  1.00  0.00           C
ATOM    158  H47 OCT     1     -26.495  15.704  16.258  1.00  0.00           H
ATOM    159  C87 OCT     1     -28.818  13.569  17.621  1.00  0.00           C
ATOM    160  H44 OCT     1     -29.234  12.552  17.748  1.00  0.00           H
ATOM    161  C90 OCT     1     -29.773  14.530  18.007  1.00  0.00           C
ATOM    162  C92 OCT     1     -31.183  14.216  18.457  1.00  0.00           C
ATOM    163  O23 OCT     1     -31.948  15.092  18.775  1.00  0.00           O
ATOM    164  O24 OCT     1     -31.638  13.071  18.388  1.00  0.00           O
ATOM    165  C89 OCT     1     -29.444  15.889  17.814  1.00  0.00           C
ATOM    166  H43 OCT     1     -30.090  16.681  18.080  1.00  0.00           H
ATOM    167  C86 OCT     1     -28.260  16.280  17.193  1.00  0.00           C
ATOM    168  O16 OCT     1     -28.124  17.577  16.880  1.00  0.00           O
ATOM    169  C46 OCT     1     -27.218  18.424  17.487  1.00  0.00           C
ATOM    170  C42 OCT     1     -26.839  19.557  16.777  1.00  0.00           C
ATOM    171  H27 OCT     1     -27.202  19.773  15.787  1.00  0.00           H
ATOM    172  C45 OCT     1     -26.126  20.566  17.374  1.00  0.00           C
ATOM    173  O9  OCT     1     -25.783  21.719  16.655  1.00  0.00           O
ATOM    174  C83 OCT     1     -25.303  22.930  17.230  1.00  0.00           C
ATOM    175  C82 OCT     1     -23.926  23.115  17.270  1.00  0.00           C
ATOM    176  H46 OCT     1     -23.268  22.413  16.913  1.00  0.00           H
ATOM    177  C80 OCT     1     -26.109  23.940  17.770  1.00  0.00           C
ATOM    178  H42 OCT     1     -27.184  23.762  17.836  1.00  0.00           H
ATOM    179  C84 OCT     1     -25.592  25.155  18.211  1.00  0.00           C
ATOM    180  C85 OCT     1     -26.500  26.201  18.806  1.00  0.00           C
ATOM    181  O17 OCT     1     -26.136  27.273  19.163  1.00  0.00           O
ATOM    182  O18 OCT     1     -27.659  25.919  19.091  1.00  0.00           O
ATOM    183  C81 OCT     1     -24.210  25.325  18.060  1.00  0.00           C
ATOM    184  H41 OCT     1     -23.826  26.287  18.436  1.00  0.00           H
ATOM    185  C79 OCT     1     -23.373  24.328  17.617  1.00  0.00           C
ATOM    186  O12 OCT     1     -22.033  24.663  17.503  1.00  0.00           O
TER
ENDMDL
```

Table C.1: Protein data bank structure of a single Octa-acid host illustrated in Figure C.1.

atom	partial charge (e)	GAFF atom type
1 O29	-0.880201	o
2 C93	0.916500	c
3 O30	-0.855801	o
4 C30	-0.189000	c3
5 H49	0.032600	hc
6 H50	0.031500	hc
7 C29	-0.062500	c3
8 H13	0.069100	hc
9 H14	0.041700	hc
10 C8	0.014100	c3
11 H4	0.093800	hc
12 C5	-0.056500	ca
13 C2	-0.099900	ca
14 H2	0.156100	ha
15 C6	-0.083000	ca
16 C7	0.026700	c3
17 H3	0.086300	hc
18 C9	-0.093200	ca
19 C12	-0.095400	ca
20 H5	0.172900	ha
21 C31	-0.069700	c3
22 H15	0.054000	hc
23 C32	-0.134300	c3
24 H55	0.057900	hc
25 H56	0.084400	hc
26 C96	0.639801	c
27 O31	-0.596801	o
28 O32	-0.614601	oh
29 H58	0.514101	ho
30 H16	0.072800	hc
31 C3	0.079100	ca
32 O4	-0.334900	os
33 C1	-0.153800	ca
34 H1	0.157800	ha
35 C4	0.079300	ca
36 O7	-0.336500	os
37 C39	0.370000	c3
38 H23	0.067900	h2
39 C41	-0.046400	ca
40 C43	-0.176400	ca
41 H25	0.159400	ha
42 C44	-0.165000	ca
43 H26	0.164800	ha
44 O8	-0.330000	os
45 C18	0.084400	ca
46 C19	-0.155800	ca
47 H8	0.157700	ha
48 C15	-0.072000	ca
49 C17	-0.095600	ca
50 H7	0.172900	ha

```
 51 C20   -0.093300   ca
 52 C21    0.026800   c3
 53 H9     0.086300   hc
 54 C35   -0.069800   c3
 55 H19    0.054000   hc
 56 C36   -0.134300   c3
 57 H51    0.057900   hc
 58 H52    0.084500   hc
 59 C94    0.639801   c
 60 O27   -0.596701   o
 61 O28   -0.614601   oh
 62 H57    0.514001   ho
 63 H20    0.072800   hc
 64 C16    0.078000   ca
 65 O2    -0.343400   os
 66 C40    0.370500   c3
 67 H24    0.071100   h2
 68 C47   -0.046200   ca
 69 C49   -0.169800   ca
 70 H28    0.163900   ha
 71 C50   -0.179600   ca
 72 H29    0.159700   ha
 73 O1    -0.334800   os
 74 C24    0.079200   ca
 75 C28   -0.153900   ca
 76 C25    0.079300   ca
 77 O3    -0.336600   os
 78 C38    0.370000   c3
 79 O6    -0.329900   os
 80 H22    0.068000   h2
 81 C53   -0.046900   ca
 82 C55   -0.164100   ca
 83 H31    0.164900   ha
 84 C56   -0.175800   ca
 85 H32    0.159300   ha
 86 H12    0.157800   ha
 87 C27   -0.082900   ca
 88 C26   -0.099900   ca
 89 H11    0.156100   ha
 90 C23   -0.056600   ca
 91 C22    0.014200   c3
 92 H10    0.093700   hc
 93 C33   -0.062600   c3
 94 H17    0.069100   hc
 95 C34   -0.189000   c3
 96 H53    0.032600   hc
 97 H54    0.031500   hc
 98 C95    0.916500   c
 99 O25   -0.880301   o
100 O26   -0.855801   o
```

```
101 H18     0.041800     hc
102 C13    -0.072000     ca
103 C10     0.084300     ca
104 C14    -0.155800     ca
105 H6      0.157600     ha
106 C11     0.078000     ca
107 O5     -0.343400     os
108 C37     0.370500     c3
109 H21     0.071200     h2
110 C59    -0.046000     ca
111 C62    -0.170100     ca
112 H35     0.163900     ha
113 C61    -0.179900     ca
114 H34     0.159900     ha
115 C64     0.145600     ca
116 C60    -0.174300     ca
117 H36     0.154300     ha
118 C63     0.127300     ca
119 O15    -0.263400     os
120 C77     0.083200     ca
121 C74    -0.191200     ca
122 H45     0.146200     ha
123 C73    -0.115500     ca
124 H40     0.165300     ha
125 C76    -0.108500     ca
126 C78     0.910501     c
127 O19    -0.841001     o
128 O20    -0.836001     o
129 C75    -0.115300     ca
130 H39     0.165800     ha
131 C72     0.070200     ca
132 O14    -0.265300     os
133 C57     0.142100     ca
134 C54    -0.169800     ca
135 H33     0.154000     ha
136 C58     0.122700     ca
137 O13    -0.262800     os
138 C70     0.084700     ca
139 C67    -0.193200     ca
140 H48     0.146500     ha
141 C66    -0.114800     ca
142 H38     0.165900     ha
143 C69    -0.108800     ca
144 C71     0.910401     c
145 O21    -0.839801     o
146 O22    -0.837901     o
147 C68    -0.115900     ca
148 H37     0.165100     ha
149 C65     0.069700     ca
150 O11    -0.264100     os
```

```
151 C51     0.145500      ca
152 C48    -0.174100      ca
153 H30     0.154400      ha
154 C52     0.127200      ca
155 O10    -0.263500      os
156 C91     0.083100      ca
157 C88    -0.191100      ca
158 H47     0.146200      ha
159 C87    -0.115300      ca
160 H44     0.165300      ha
161 C90    -0.108600      ca
162 C92     0.910501      c
163 O23    -0.836101      o
164 O24    -0.841001      o
165 C89    -0.115200      ca
166 H43     0.165800      ha
167 C86     0.070200      ca
168 O16    -0.265300      os
169 C46     0.142500      ca
170 C42    -0.170500      ca
171 H27     0.154000      ha
172 C45     0.123400      ca
173 O9     -0.262800      os
174 C83     0.084100      ca
175 C82    -0.192500      ca
176 H46     0.146500      ha
177 C80    -0.114300      ca
178 H42     0.166000      ha
179 C84    -0.109000      ca
180 C85     0.910401      c
181 O17    -0.837901      o
182 O18    -0.839701      o
183 C81    -0.115500      ca
184 H41     0.165100      ha
185 C79     0.069300      ca
186 O12    -0.264000      os
```

Table C.2: Partial charges and GAFF atom type for evaluation of electrostatic, Lennard-Jones, and intramolecular interactions. Intramolecular bond, bond-angle, and torsional interactions were taken from the GAFF potential. Atom numbers listed in column 1 follow Table C.1.

Appendix D

SUPPORTING INFORMATION FOR "HYDRATION CHARACTERISTICS OF THE INTERIOR OF DEEP-CAVITY CAVITANDS OCTA-ACID AND TEMOA"

Here we detail water occupancy probabilities for each deep-cavity cavitand at various pressures (Figure D.1), temperatures (Figure D.2), and electric field strengths (Figure D.4). Additionally we show several characteristics of water as a function of depth inside the binding pocket (Figure D.3). We also detail the free energy of inserting alkane-sized empty cavities inside of OA and TEMOA (Figure D.5) as well as a breakdown of the distribution of waters for each system (Figures D.6 and D.7). Lastly we show the entire PMF's used for Figure 4.8 in Figure D.8.
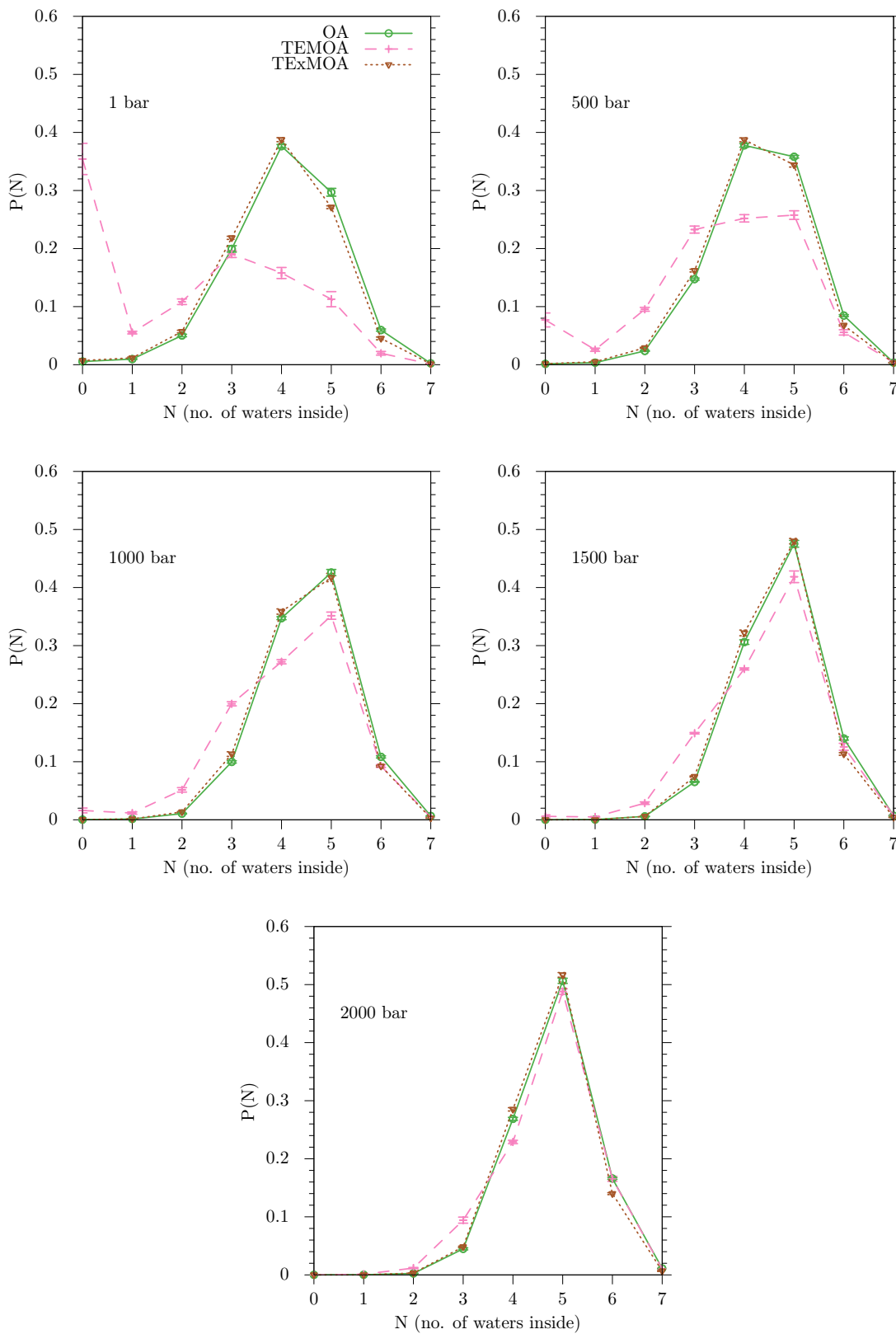
Figure D.1: Distributions of number of waters inside deep-cavity cavitands with pressure varied.
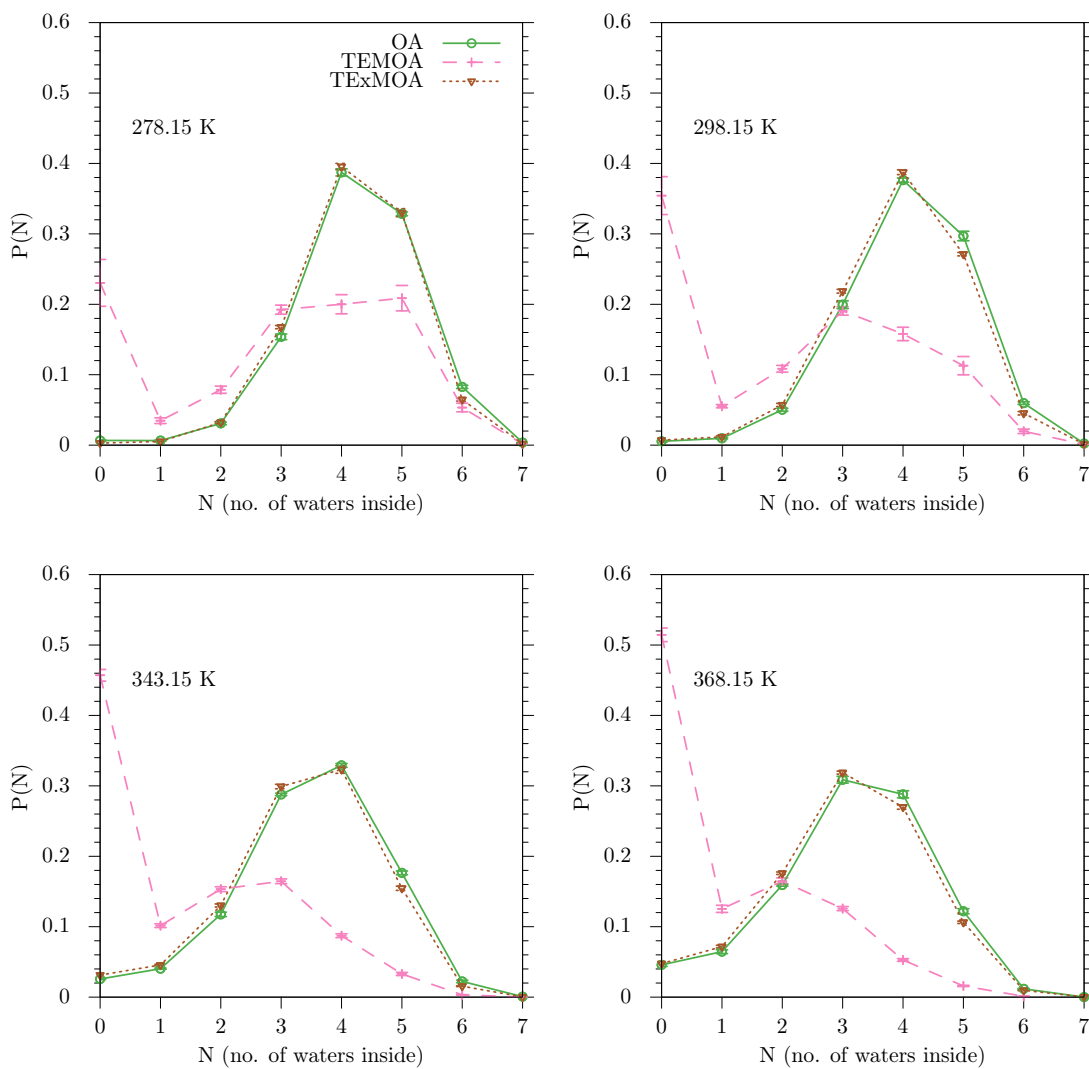
Figure D.2: Distributions of number of waters inside deep-cavity cavitands with temperature varied.
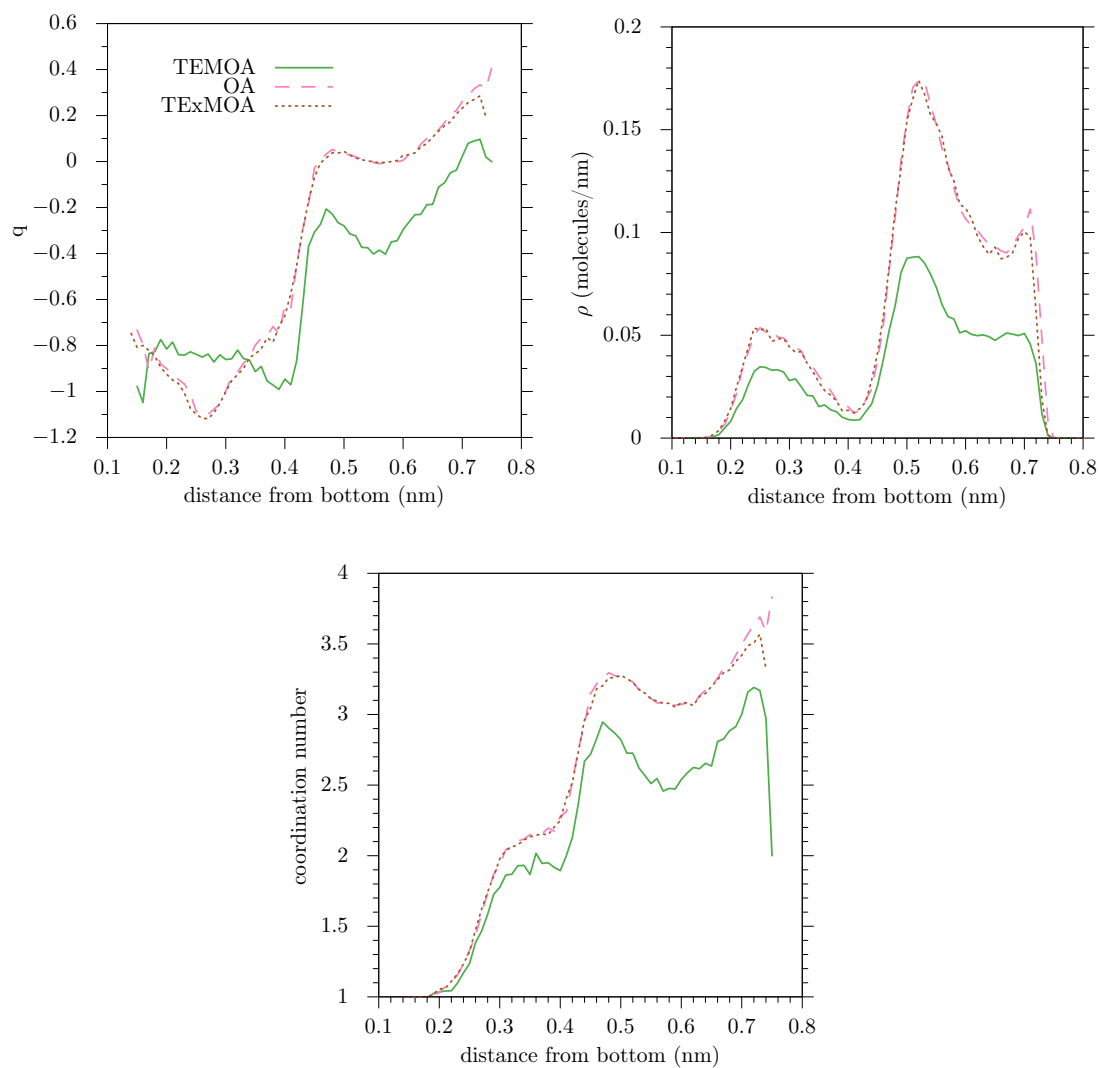
Figure D.3: *Top, left:* Orientational order parameter of water as a function of distance to the bottom of each deep-cavity cavitand. *Top, right:* Density of water as a func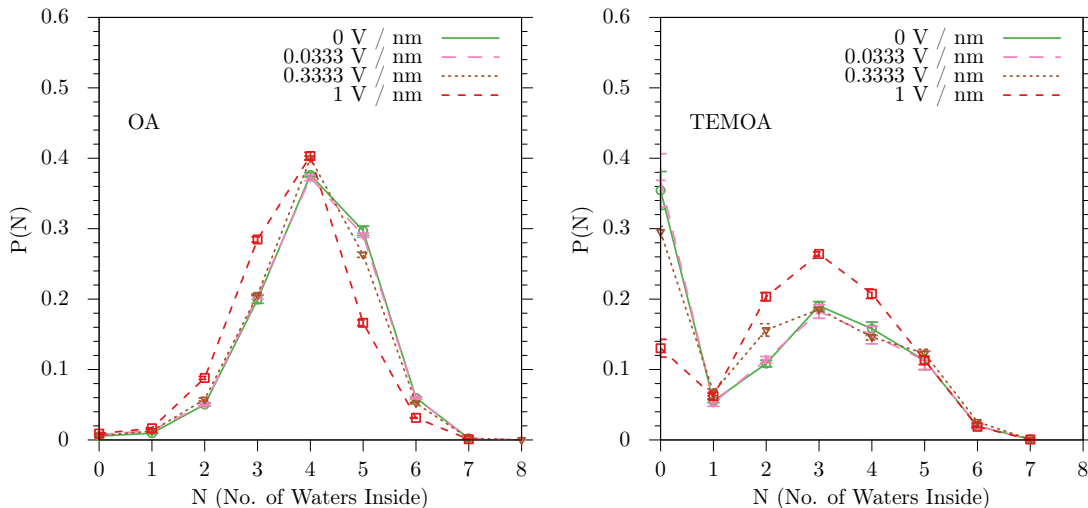tion of distance to the bottom of each deep-cavity cavitand. *Bottom:* Coordination number of water as a function of distance to the bottom of each deep-cavity cavitand.

Figure D.4: Distributions of number of waters inside OA (*left*) and TEMOA (*right*) with electric field applied at various strengths.
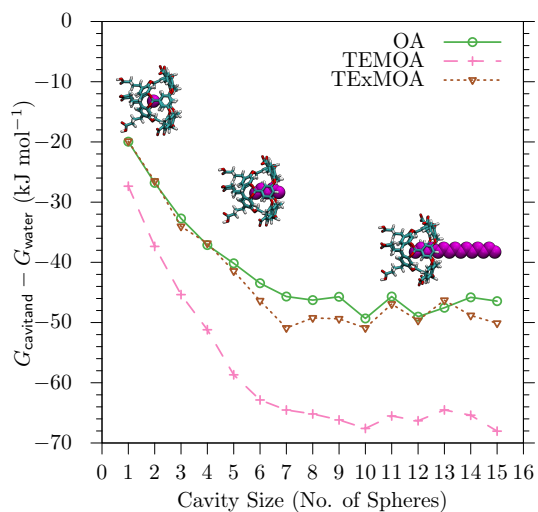
Figure D.5: The free energy of transferring an empty cavity from bulk water into the bottom of a cavitand. Each empty cavity was composed of a series of overlapping spheres a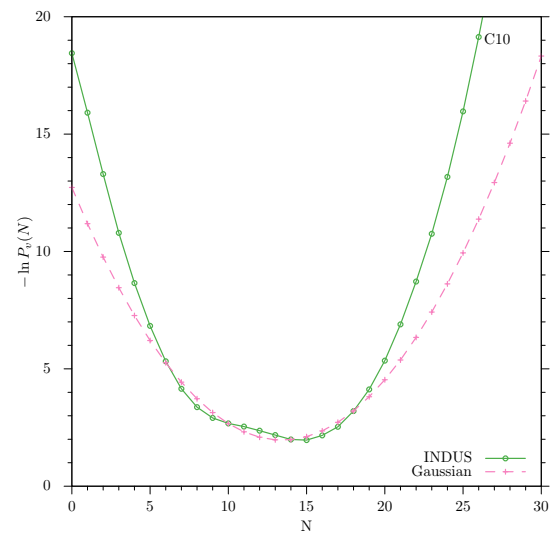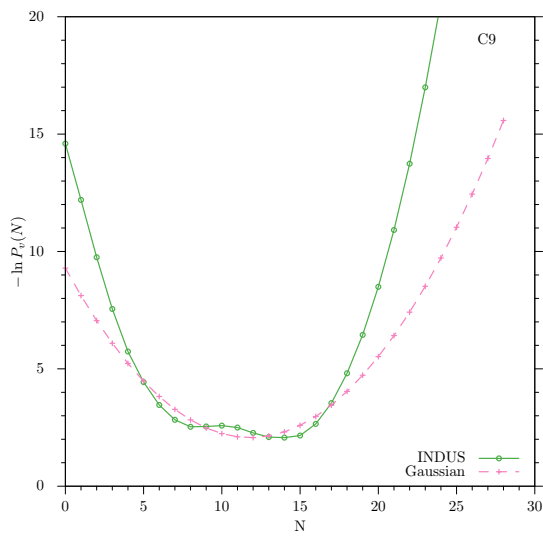pproximating an all-trans alkane. The indirect umbrella sampling technique (INDUS)[82] was used to determine $P_V(0)$, with the y-axis indicating $G = -\ln P_V(0)$. Figures D.6 and D.7 detail each $P_v(N)$ distribution.
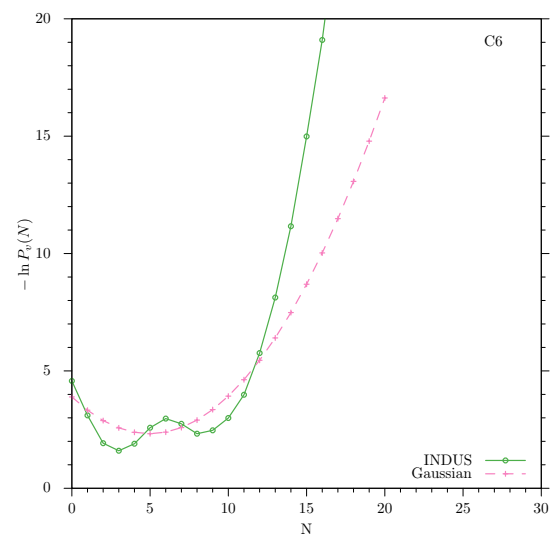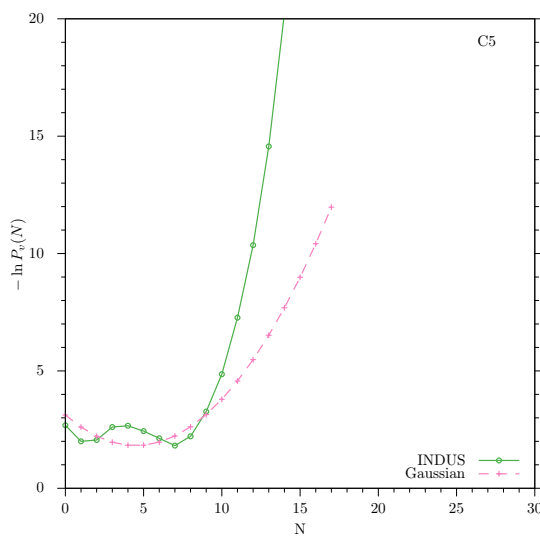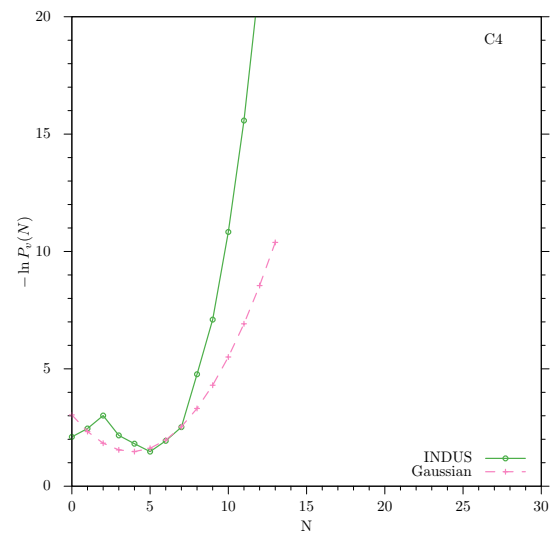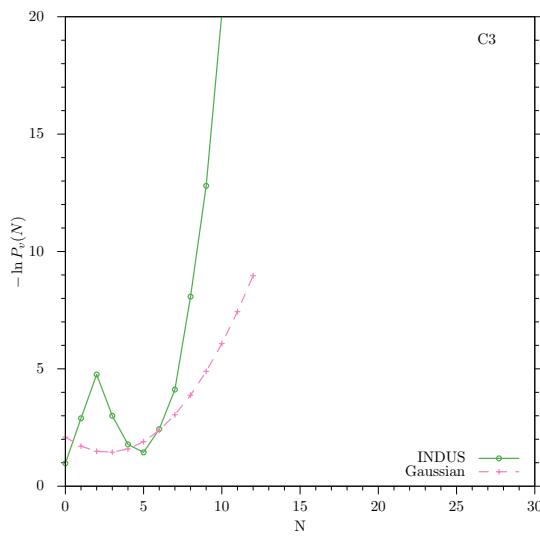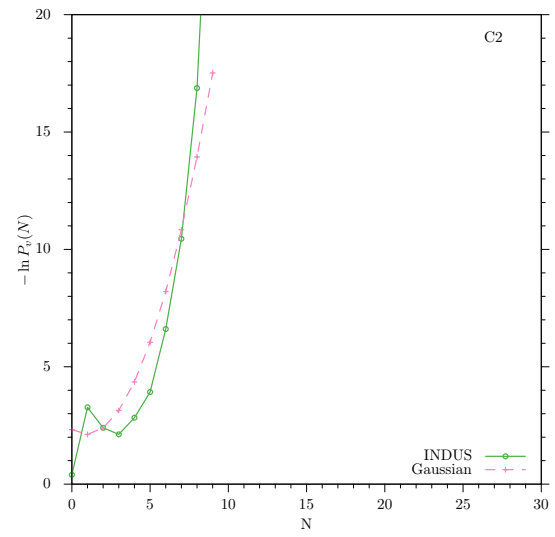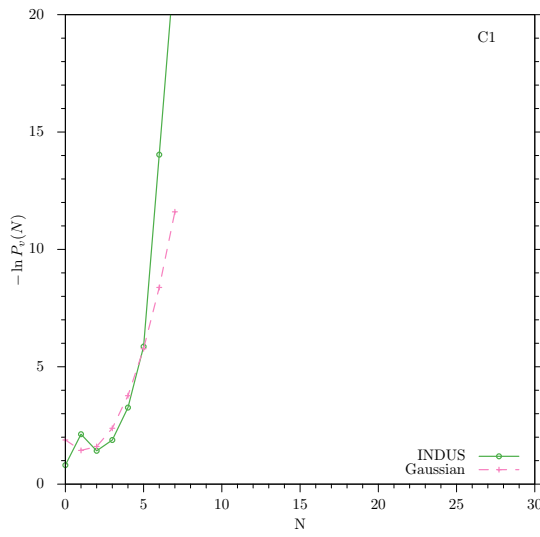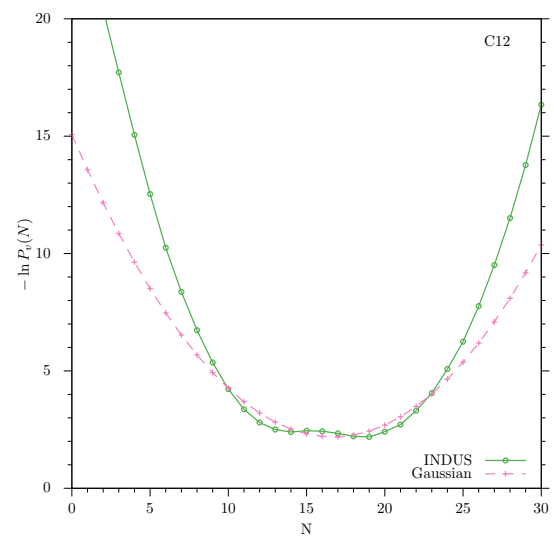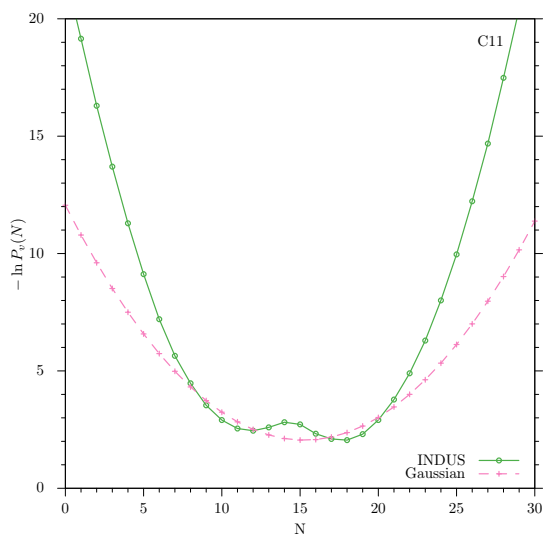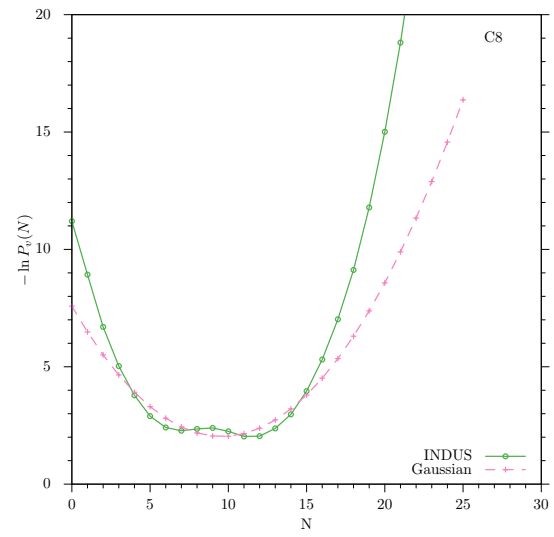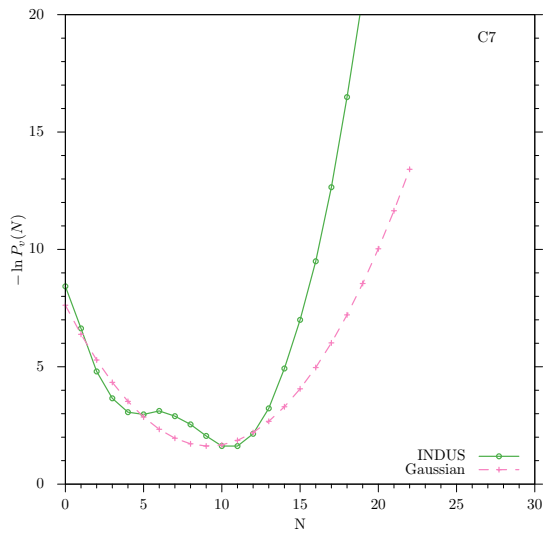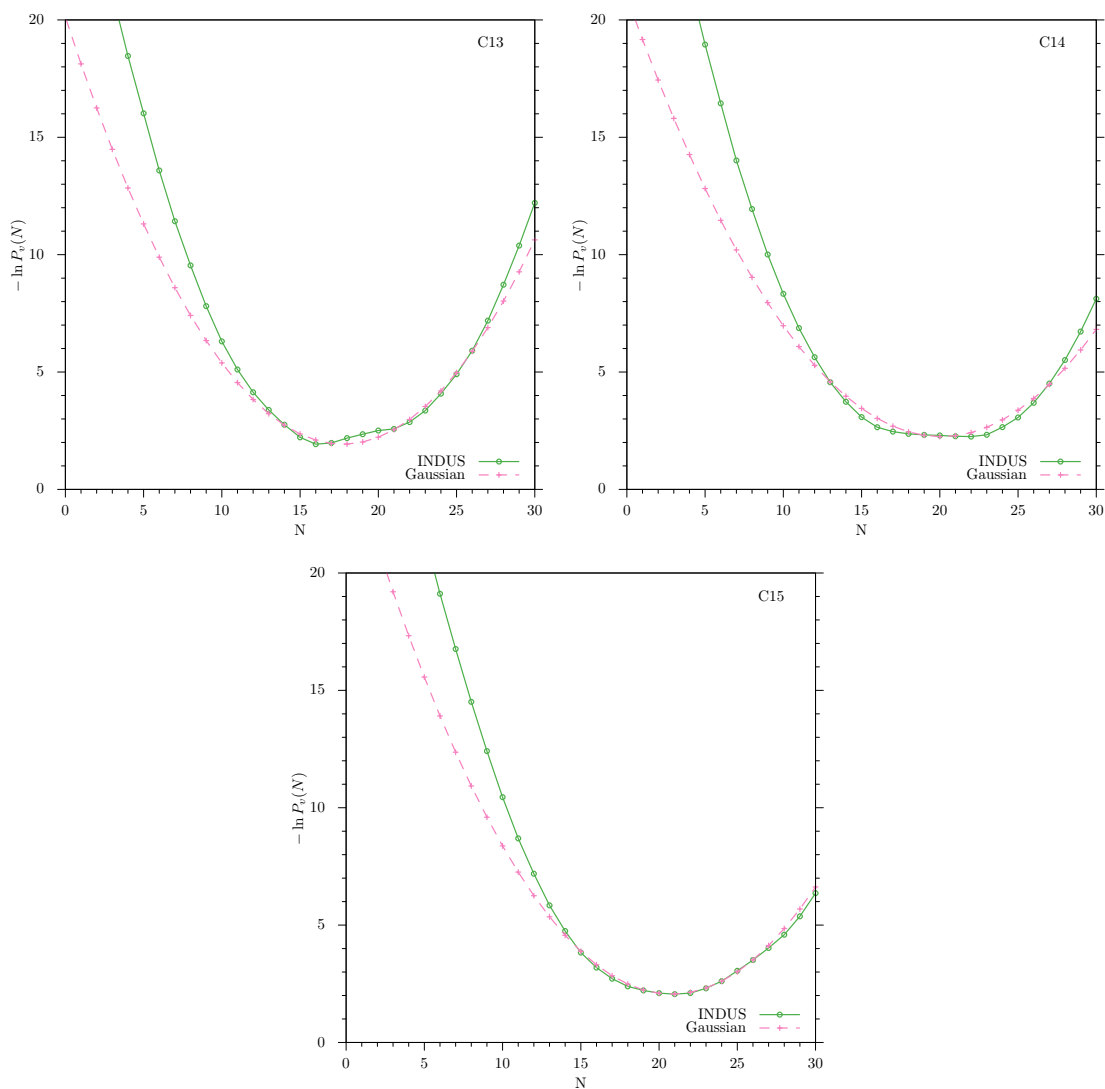
Figure D.6: Probability of observing N waters inside of group of spheres forming an all-trans alkane starting at the bottom of TEMOA using the INDUS technique and using a Gaussian distribution.
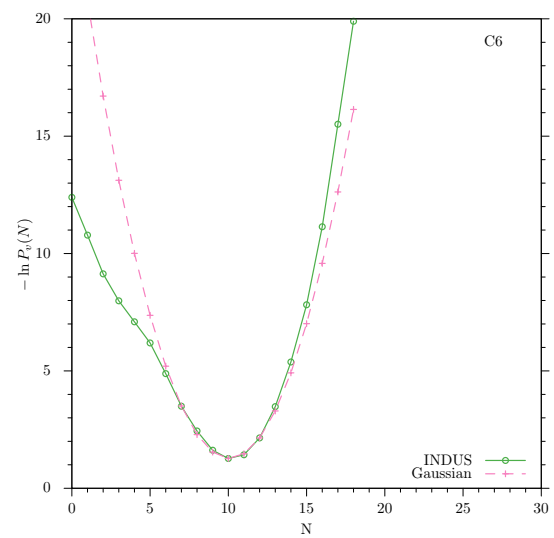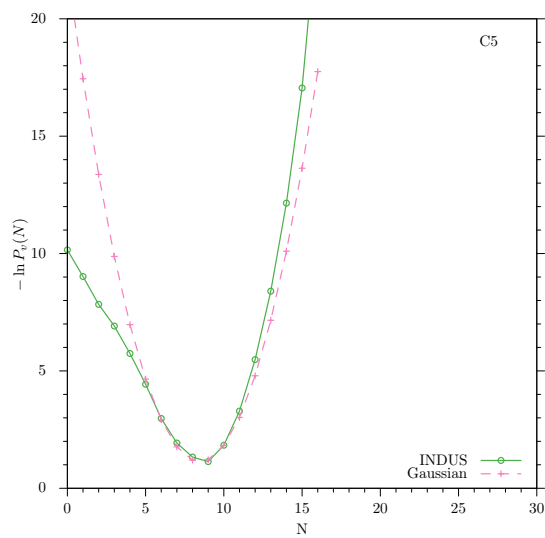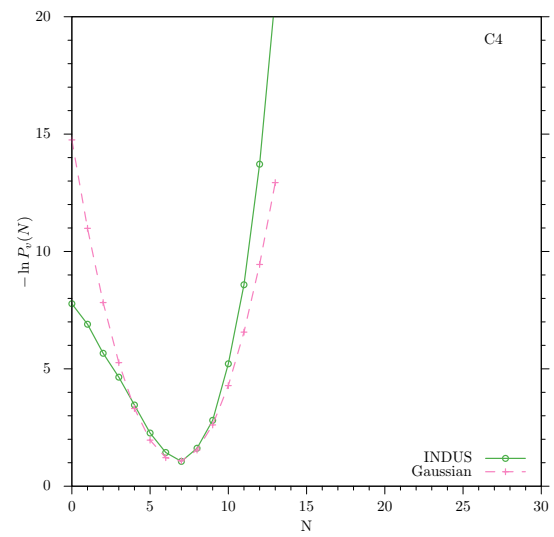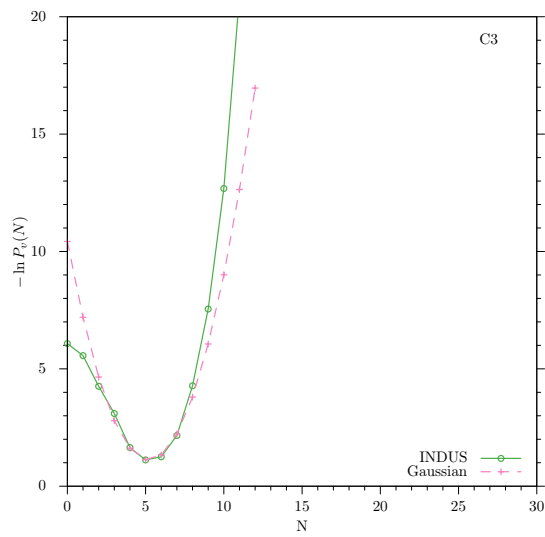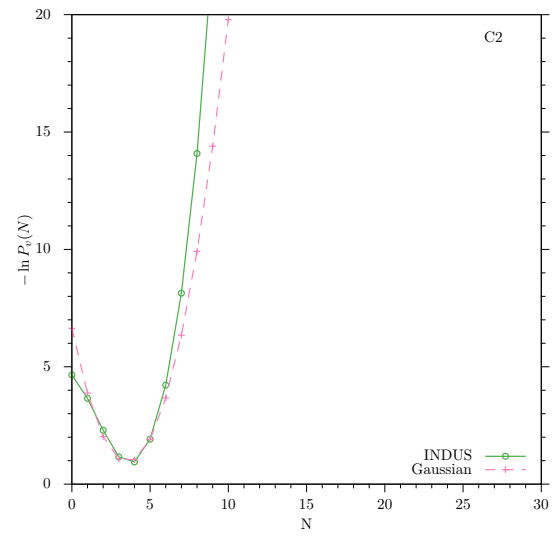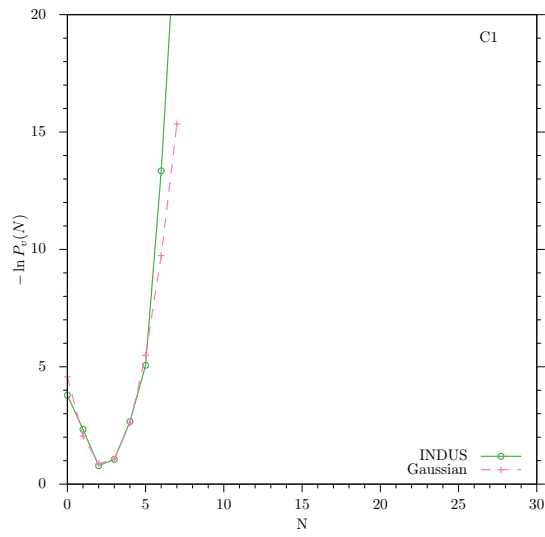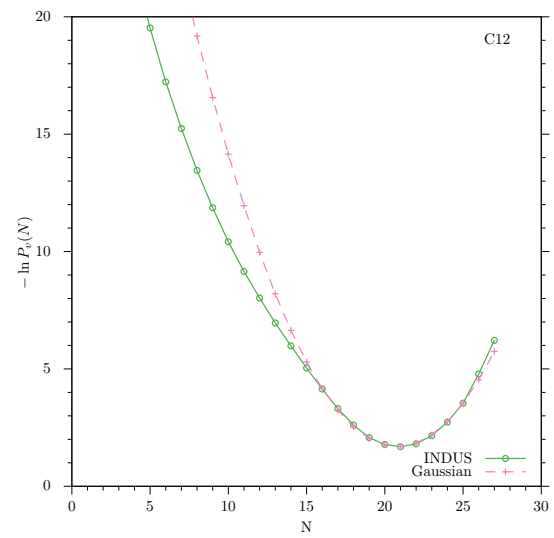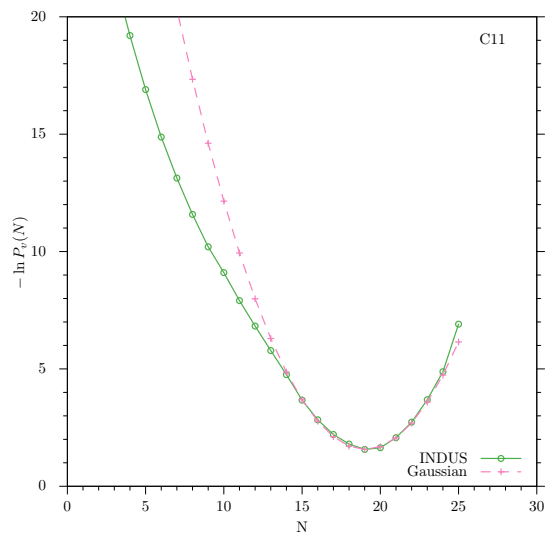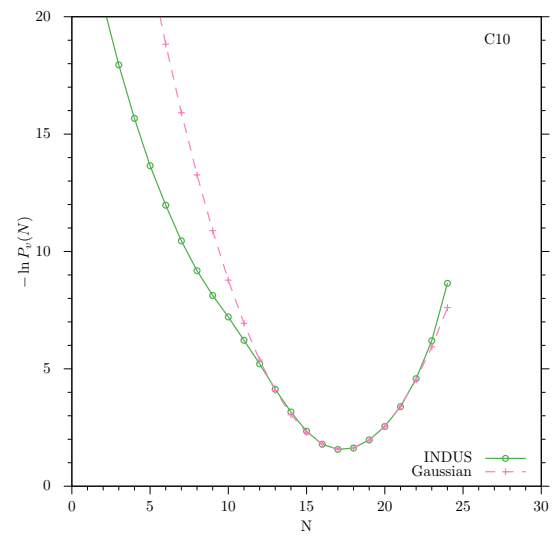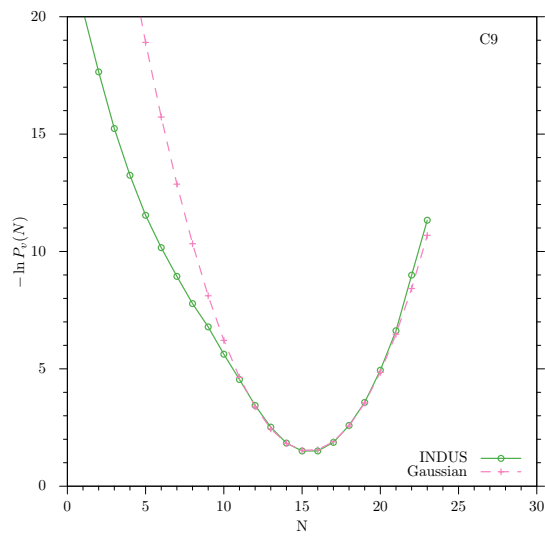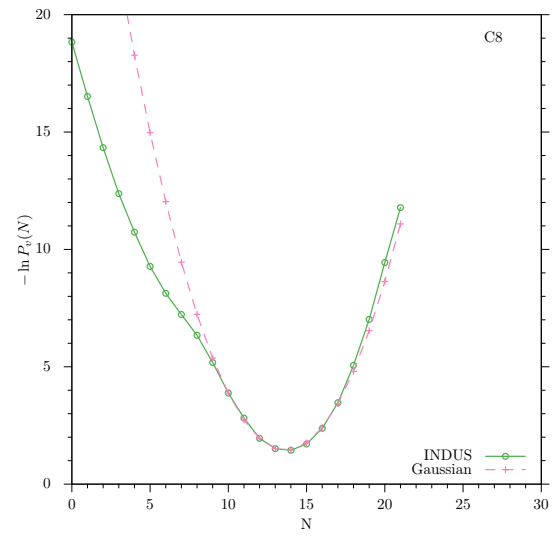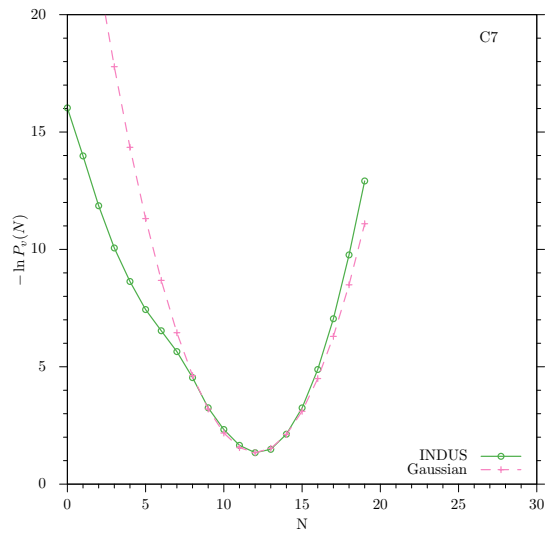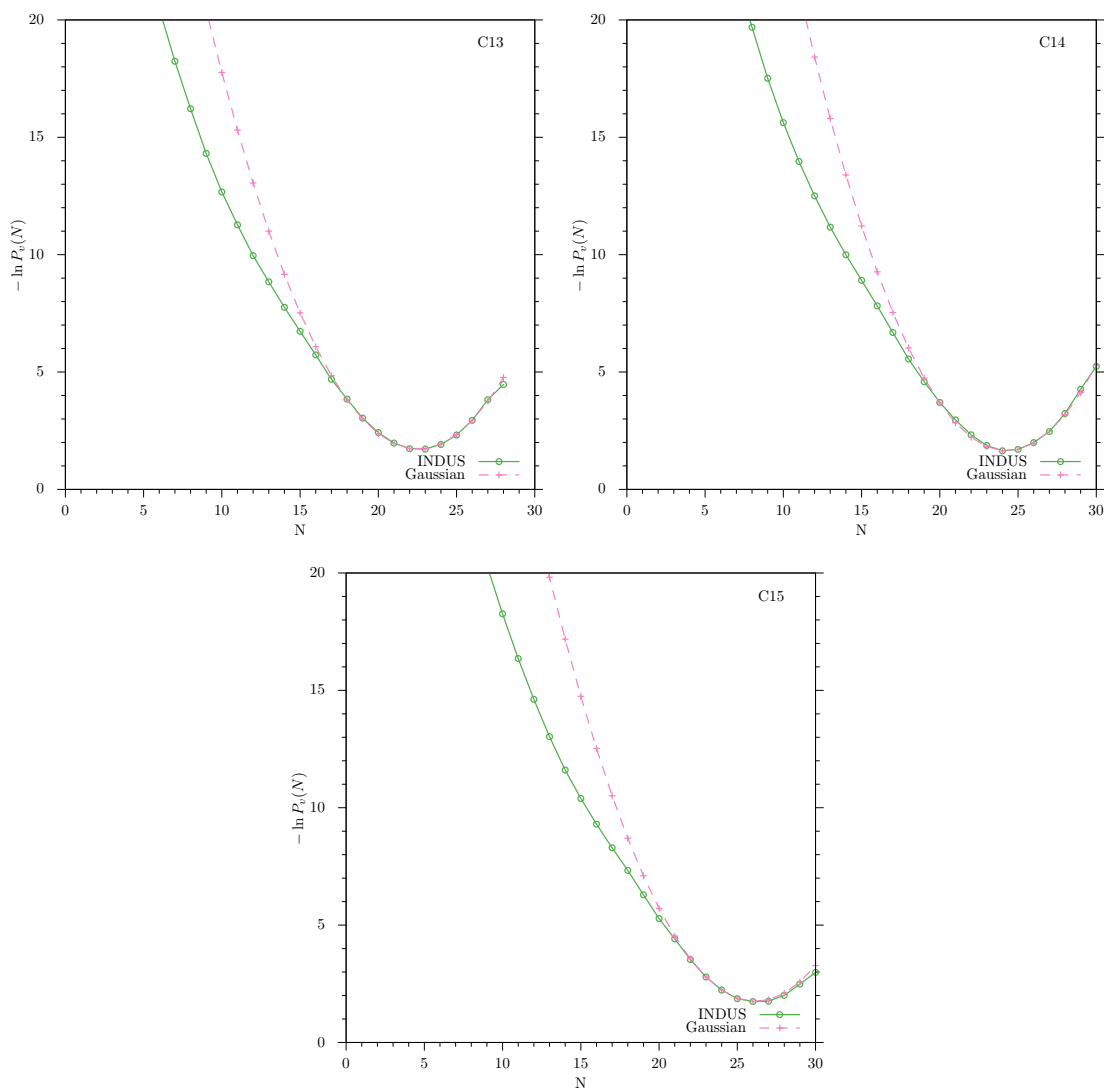
Figure D.7: Probability of observing N waters inside of group of spheres forming an all-trans alkane starting at the bottom of OA using the INDUS technique and using a Gaussian distribution.
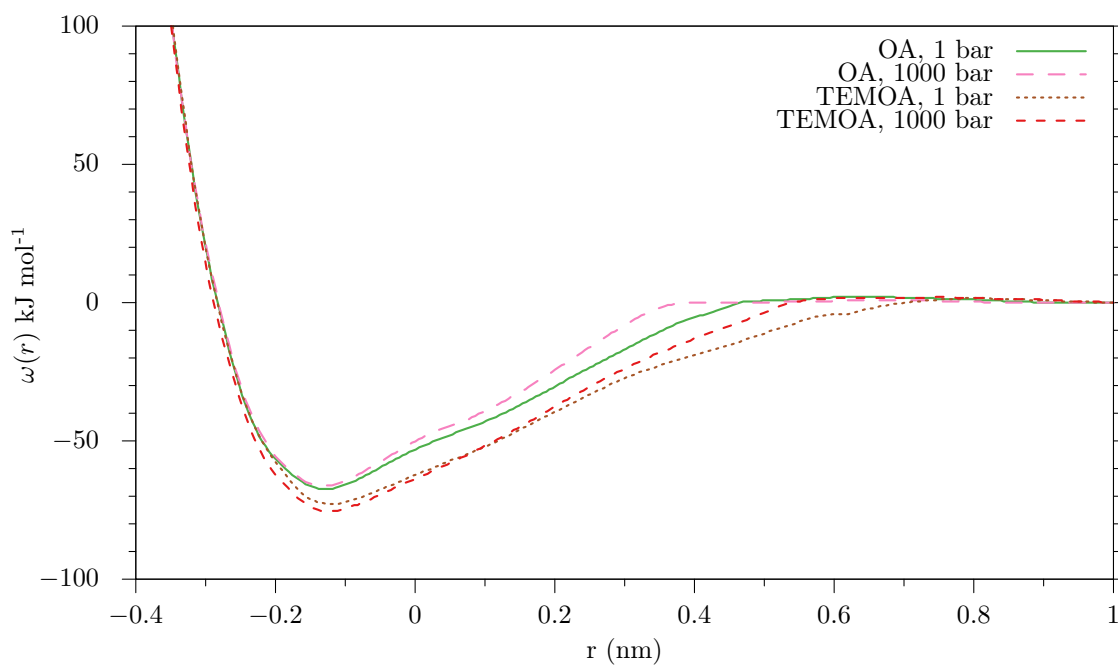
Figure D.8: Potentials-of-mean force between OA and hexane at 1 bar and 1000 bar; and between TEMOA and hexane at 1 bar and 1000 bar.

Appendix E

**SUPPORTING INFORMATION FOR "ASSOCIATION OF AMPHIPHILIC
MACORCYCLIC TETRALACTAM HOSTS WITH SQUARAINE DYE"**

Here we give the detailed results for each free tetralactam system studied. Table E.1
gives a breakdown of number of waters, hydrogen bonds, and probability of the lactam
being in the two configurations for each type of lactam simulated. The following tables
give the potential energy of waters in bulk and inside a regular lactam (Table E.2),
inside a lactam with the amide hydrogens turned off (Table E.3), inside a lactam with
methoxys added (Table E.4), and inside a lactam with both methoxys added and amide
hydrogens turned off (Table E.5).

	Bulk	Inside Normal	Inside Amide H's Off	Inside with Methoxys	Inside with Methoxy's & Amide H's Off
Prob. Aligned		0.728	0.332	0.962	0.888
Prob. Funnel		0.272	0.668	0.038	0.112
N Waters Inside		4.481 ± 0.040	3.249 ± 0.078	4.550 ± 0.014	3.618 ± 0.024
N Waters (aligned)		4.910 ± 0.006	4.612 ± 0.022	4.588 ± 0.014	3.754 ± 0.023
N Waters (funnel)		3.336 ± 0.010	2.570 ± 0.002	3.064 ± 0.029	2.540 ± 0.018
H-Bond Count	3.018 ± 0.003	2.078 ± 0.008	1.989 ± 0.010	2.023 ± 0.004	1.978 ± 0.003
Waters within 0.34 nm	4.668 ± 0.004	2.988 ± 0.007	2.855 ± 0.013	2.809 ± 0.003	2.694 ± 0.003

Table E.1: Comparison of number of waters, hydrogen bonds, and probability of either being "aligned" or "funnel" shaped.

	Bulk	Inside	Inside "Aligned"	Inside "Funnel"
water-water (Coul)	-110.918 ± 0.001	-69.551 ± 0.294	-72.805 ± 0.053	-60.862 ± 0.165
water-water (LJ)	18.285 ± 0.000	15.060 ± 0.056	15.814 ± 0.010	13.046 ± 0.042
water-water total	-92.970 ± 0.000	-54.492 ± 0.251	-56.992 ± 0.052	-47.816 ± 0.141
water-host (Coul)		-13.724 ± 0.153	-12.429 ± 0.045	-17.180 ± 0.061
water-host (LJ)		-6.994 ± 0.016	-7.152 ± 0.014	-6.571 ± 0.023
water-host total		-20.718 ± 0.136	-19.581 ± 0.045	-23.752 ± 0.077
total	-92.970 ± 0.000	-75.209 ± 0.112	-76.573 ± 0.055	-71.568 ± 0.080

Table E.2: Energy of waters in host compared with bulk water (kJ/mol)

	Bulk	Inside	Inside "Aligned"	Inside "Funnel"
water-water (Coul)	-110.918 ± 0.001	-69.248 ± 0.339	-75.434 ± 0.131	-66.161 ± 0.0836
water-water (LJ)	18.285 ± 0.000	14.101 ± 0.101	15.731 ± 0.057	13.287 ± 0.0309
water-water total	-92.970 ± 0.000	-55.147 ± 0.250	-59.703 ± 0.075	-52.874 ± 0.0565
water-host (Coul)		-7.426 ± 0.080	-6.141 ± 0.020	-8.067 ± 0.0167
water-host (LJ)		-8.313 ± 0.011	-8.116 ± 0.009	-8.411 ± 0.0089
water-host total		-15.739 ± 0.078	-14.257 ± 0.022	-16.478 ± 0.0145
total	-92.970 ± 0.000	-70.886 ± 0.163	-73.960 ± 0.065	-69.352 ± 0.0667

Table E.3: Energy of waters in host when amide H's turned off compared with bulk (kJ/mol)

	Bulk	Inside	Inside "Aligned"	Inside "Funnel"
water-water (Coul)	-110.918 ± 0.001	-68.169 ± 0.120	-68.482 ± 0.125	-60.239 ± 0.384
water-water (LJ)	18.285 ± 0.000	15.194 ± 0.040	15.273 ± 0.040	13.187 ± 0.204
water-water total	-92.970 ± 0.000	-52.975 ± 0.087	-53.208 ± 0.093	-47.052 ± 0.183
water-host (Coul)		-12.605 ± 0.046	-12.482 ± 0.035	-15.720 ± 0.111
water-host (LJ)		-8.473 ± 0.011	-8.513 ± 0.009	-7.469 ± 0.052
water-host total		-21.078 ± 0.039	-20.995 ± 0.036	-23.189 ± 0.101
total	-92.970 ± 0.000	-74.053 ± 0.064	-74.203 ± 0.080	-70.241 ± 0.127

Table E.4: Energy of waters in host when methoxys added compared with bulk (kJ/mol)

	Bulk	Inside	Inside "Aligned"	Inside "Funnel"
water-water (Coul)	-110.918 ± 0.001	-67.999 ± 0.0771	-68.438 ± 0.106	-64.490 ± 0.156
water-water (LJ)	18.285 ± 0.000	14.209 ± 0.0303	14.317 ± 0.033	13.347 ± 0.048
water-water total	-92.970 ± 0.000	-53.790 ± 0.0537	-54.122 ± 0.081	-51.143 ± 0.145
water-host (Coul)		-4.625 ± 0.0121	-4.5820 ± 0.013	-4.966 ± 0.051
water-host (LJ)		-9.627 ± 0.0096	-9.684 ± 0.009	-9.170 ± 0.032
water-host total		-14.252 ± 0.0131	-14.267 ± 0.013	-14.136 ± 0.077
total	-92.970 ± 0.000	-68.042 ± 0.0543	-68.388 ± 0.075	-65.279 ± 0.070

Table E.5: Energy of waters in host when methoxys added with amide H's off compared with bulk (kJ/mol)

LIST OF REFERENCES

[1] Elemans, J. A. A. W.; Rowan, A. E.; Nolte, R. J. M. *J. Mater. Chem.* **2003**, *13*, 2661–2670.

[2] Alper, J. *Science* **2002**, *295*, 2396.

[3] Kunz, H. *Angewandte Chemie* **2002**, *41*, 4439–4451.

[4] The Royal Swedish Academy of Sciences, Press Release: The 1987 Nobel Prize in Chemistry. `http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1987/press.html`, (online, Accessed 03 April 2014).

[5] Cram, D. J. *Science* **1983**, *219*, 1177–1183.

[6] Moran, J. R.; Karbach, S.; Cram, D. J. *J. Am. Chem. Soc.* **1982**, *104*, 5826–5828.

[7] (a) Gibb, C. L. D.; Gibb, B. C. *Chem. Comm.* **2007**, 1635–1637; (b) Gibb, C. L. D.; Gibb, B. C. *J. Am. Chem. Soc.* **2004**, *126*, 11408–11409.

[8] (a) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Elsevier LTD, Oxford, 2001; Chapter 4; (b) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, Oxford, 1987; Chapter 3.

[9] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[10] (a) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Elsevier LTD, Oxford, 2001; Chapter 12; (b) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, Oxford, 1987; Chapter 5; (c) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092; (d) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8592.

[11] McQuarrie, D. *Statistical mechanics*; Viva Books: New Delhi, 2003.

[12] (a) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Elsevier LTD, Oxford, 2001; Chapter 7; (b) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, Oxford, 1987; Chapter 7.

[13] Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.

[14] Hubb, J. S.; l. De Groot, B.; Spoel, D. V. D. *J. Chem. Theo. Comp.* **2010**, *6*, 3713–3720.

[15] Liu, S.; Russell, D. H.; Zinnel, N. F.; Gibb, B. C. *J. Am. Chem. Soc.* **2013**, *135*, 4314–4324.

[16] Mayhew, M.; da Silva, A. C. R.; Martin, J.; Erdjument-Bromage, H.; Tempst, P.; Hartl, F. U. *Nature* **1996**, *379*, 420–426.

[17] (a) Purohit, P. K.; Kondev, J.; Phillips, R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *100*, 3173–3178; (b) Purohit, P. K.; Inamdar, M. M.; Grayson, P. D.; Squires, T. M.; Kondev, J.; Phillips, R. *Biophysical Journal* **2005**, *88*, 851–866; (c) Grayson, P.; Evilevitch, A.; Inamdar, M. M.; Purohit, P. K.; Gelbart, W. M.; Knobler, C. M.; Phillips, R. *Virology* **2006**, *348*, 430–436.

[18] (a) Chugani, D. C.; Rome, L. H.; Kerdesha, N. L. *J. Cell. Sci.* **1993**, *106*, 23–29; (b) Tanaka, H.; Kato, K.; Yamashita, E.; Sumizawa, T.; Zhou, Y.; Yao, M.; Iwasaki, K.; Yoshimura, M.; Tsukihara, T. *Science* **2009**, *323*, 384–388.

[19] Gan, H.; Gibb, B. C. *Chem. Commun.* **2013**, *49*, 1395–1397.

[20] (a) Ben-Naim, A. *J. Chem. Phys.* **1971**, *54*, 1387–1404; (b) A. Ben-Naim, M. Y., J. Wilf *J. Phys. Chem.* **1972**, *77*, 95–102; (c) Yaacobi, M.; Ben-Naim, A. *J. Phys. Chem.* **1974**, *78*, 175–178; (d) Ben-Naim, A.; Wilf, J. *J. Chem. Phys.* **1979**, *70*, 771–777; (e) Ben-Naim, A. *Hydrophobic interactions*; Plenum Press, New York, 1980.

[21] Munoz, I. G.; Yebenes, H.; Zhou, M.; Mesa, P.; Serna, M.; Park, A. Y.; Bragado-Nilsson, E.; Beloso, A.; de Carcer, G.; Malumbres, M. *Nat. Struct. Mol. Biol.* **2011**, *18*, 14–19.

[22] Matsuoka, S.; Sugiyama, S.; Matsuoka, D.; Hirose, M.; Lethu, S.; Ano, H.; Hara, T.; Ichihara, O.; Kimura, S. R.; Murakami, S.; Ishida, H.; Mizohata, E.; Inoue, T.; Murata, M. *Angew. Chem. Int. Ed.* **2015**, *54*, 1508–1511.

[23] Van Den Berg, B.; Black, P.; Clemons, W.; Rapoport, T. *Science* **2004**, *304*, 1506–1509.

[24] (a) Cragg, P. J. *Supramolecular Chemistry: from Biological Inspiration to Applications*; Springer, Dordrecht, 2010; (b) Ma, X.; Zhao, Y. *Chem. Rev.* **2015**, *115*, 7794–7839; (c) Jordan, J. H.; Gibb, B. C. *Chem. Soc. Rev.* **2015**, *44*, 547–585.

[25] Raynal, M.; Ballester, P.; Vidal-ferran, A.; Van Leeuwen, P. W. N. M. *Chem. Soc. Rev.* **2014**, *43*, 1734–1787.

[26] (a) Kaanumalle, L.; Gibb, C.; Gibb, B.; Ramamurthy, V. *J. Am. Chem. Soc.* **2004**, *126*, 14366–14367; (b) Kaanumalle, L. S.; Gibb, C. L. D.; Gibb, B. C.; Ramamurthy, V. *J. Am. Chem. Soc.* **2005**, *127*, 3674–3675; (c) Kaanumalle, L. S.; Gibb, C. L. D.; Gibb, B. C.; Ramamurthy, V. *Org. Biomol. Chem.* **2007**, *5*, 236–238;

(d) Natarajan, A.; Kaanumalle, L. S.; Jockusch, S.; Gibb, C. L. D.; Gibb, B. C.; Turro, N. J.; Ramamurthy, V. *J. Am. Chem. Soc.* **2007**, *129*, 4132–+; (e) Gibb, C. L. D.; Sundaresan, A. K.; Ramamurthy, V.; Gibb, B. C. *J. Am. Chem. Soc.* **2008**, *130*, 4069–4080; (f) Sundaresan, A. K.; Gibb, C. L. D.; Gibb, B. C.; Ramamurthy, V. *Tetrahedron* **2009**, *65*, 7277–7288.

[27] Gibb, C. L. D.; Gibb, B. C. *J. Am. Chem. Soc.* **2006**, *128*, 16498–16499.

[28] Liu, S.; Gan, H.; Hermann, A. T.; Rick, S. W.; Gibb, B. C. *Nat. Chem.* **2010**, *2*, 847–852.

[29] Rebek, J., Jr. *Chem. Comm.* **2007**, 2777–2789.

[30] (a) Trembleau, L.; Rebek, J. *Science* **2003**, *301*, 1219–1220; (b) Choudhury, R.; Barman, A.; Prabhakar, R.; Ramamurthy, V. *J. Phys. Chem. B* **2013**, *117*, 398–407; (c) Wanjari, P. P.; Sangwai, A. V.; Ashbaugh, H. S. *Phys. Chem. Chem. Phys.* **2012**, *14*, 2702–2709.

[31] (a) Lindahl, E.; Hess, B.; Van Der Spoel, D. *Molecular Modeling Annual* **2001**, *7*, 306–317; (b) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comp. Chem.* **2005**, *26*, 1701–1719; (c) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447; (d) Berendsen, H.; Spoel, D. V. D.; Drunen, R. V. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

[32] Siu, S. W. I.; Pluhackova, K.; Böckmann, R. A. *J. Chem. Theory Comput.* **2012**, *8*, 1459–1470.

[33] Wanjari, P. P.; Gibb, B. C.; Ashbaugh, H. S. *J. Chem. Phys.* **2013**, *139*.

[34] Ewell, J.; Gibb, B. C.; Rick, S. W. *J. Phys. Chem. B* **2008**, *112*, 10272–10279.

[35] Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

[36] Sugita, Y.; Okamoto, Y. *Chemical Physics Letters* **1999**, *314*, 141–151.

[37] Patriksson, A.; van der Spoel, D. *Physical Chemistry Chemical Physics* **2008**, *10*, 2073.

[38] (a) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511–519; (b) Hoover, W. *Phys. Rev. A* **1985**, *31*, 1695–1697.

[39] Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

[40] (a) Garcia, A. *Phys. Rev. Lett.* **1992**, *68*, 2696–2699; (b) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins - Structure Function and Genetics* **1993**, *17*, 412–425; (c) Mu, Y.; Nguyen, P.; Stock, G. *Proteins-structure Function and Bioinformatics* **2005**, *58*, 45–52; (d) Altis, A.; Nguyen, P. H.; Hegger, R.; Stock, G. *J.*

Chem. Phys. **2007**, *126*; (e) Altis, A.; Otten, M.; Nguyen, P. H.; Hegger, R.; Stock, G. *J. Chem. Phys.* **2008**, *128*.

[41] Ferguson, A. L.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *J. Phys. Chem. B* **2009**, *113*, 6405–6414.

[42] Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789–807.

[43] Frisch, M. J. et al. Gaussian 03, Revision D.02. Gaussian, Inc., Wallingford, Ct, 2004.

[44] (a) Eriksen, J. J.; Olsen, J. M. H.; Aidas, K.; Ågren, H.; Mikkelsen, K. V.; Kongsted, J. *Journal of Computational Chemistry* **2011**, *32*, 2853–2864; (b) Casanovas, J.; Namba, A. M.; León, S.; Aquino, G. L. B.; da Silva, G. V. J.; Alemán, C. *The Journal of Organic Chemistry* **2001**, *66*, 3775–3782; (c) Pérez, M.; Peakman, T. M.; Alex, A.; Higginson, P. D.; Mitchell, J. C.; Snowden, M. J.; Morao, I. *The Journal of Organic Chemistry* **2006**, *71*, 3103–3110.

[45] (a) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268; (b) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. *J. Comput. Aided. Mol. Des.* **2015**, *29*, 397–411.

[46] (a) WikstrÃűm, M. *Current Opinion in Structural Biology* **1998**, *8*, 480–488; (b) WikstrÃűm, M.; Verkhovsky, M. I.; Hummer, G. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2003**, *1604*, 61–65.

[47] Tiefenbacher, K.; Ajami, D.; Rebek, J., Jr. *Angew. Chem. Int. Ed.* **2011**, *50*, 12003–12007.

[48] Macgillivray, L. R.; Atwood, J. L. *Nature* **1997**, *389*, 469–472.

[49] Bunzen, J.; Iwasa, J.; Bonakdarzadeh, P.; Numata, E.; Rissanen, K.; Sato, S.; Fujita., M. *Angew. Chem. Int. Ed.* **2012**, *51*, 3161–3163.

[50] Hiraoka, S.; Nakamura, T.; Shiro, M.; Shionoya, M. *J. Am. Chem. Soc.* **2010**, *132*, 13223–13225.

[51] (a) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174; (b) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*.

[52] Sousa Da Silva, A. W.; Vranken, W. F. *BMC Research Notes* **2012**, *5*, 367.

[53] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, Oxford, 1987.

[54] Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. , *J. Comp. Chem.* **1997**, *18*, 1463–1472.

[55] S. Miyamoto and P. A. Kollman, *J. Comp. Chem.* **1992**, *13*, 952–962.

[56] (a) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*; (b) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 397–411.

[57] Bondi, A. *The Journal of Physical Chemistry* **1964**, *68*, 441–451.

[58] (a) Tanford, C. *The Hydrophobic Effect*, 2nd ed.; John Wiley and Sons, 1980; (b) Blokzijl, W.; Engberts, J. *Angewandte Chemie-international Edition* **1993**, *32*, 1545–1579.

[59] Biedermann, F.; Nau, W. M.; Schneider, H.-J. *Angew. Chem. Int. Ed.* **2014**, *53*, 11158–11171.

[60] Ball, P. *Chem. Rev.* **2008**, *108*, 74–108.

[61] Giovambattista, N.; Rossky, P. J.; Debenedetti, P. G. *Annu. Rev. Phys. Chem.* **2012**, *63*, 179–200.

[62] Hummer, G.; Garde, S. *Phys. Rev. Lett.* **1998**, *80*, 4193–4196.

[63] Setney, P.; Baron, R.; Mccammon, J. A. *J. Chem. Theo. Comp.* **2010**, *6*, 2866–2871.

[64] Qvist, J.; Davidovic, M.; Hamelberg, D.; Halle, B. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6296–6301.

[65] Vaitheeswaran, S.; Yin, H.; Rasaiah, J.; Hummer, G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17002–17005.

[66] Braun, E.; Chen, J. J.; Schnell, S. K.; Lin, L.-C.; Reimer, J. A.; Smit, B. *Angew. Chem. Int. Ed.* **2015**, *54*, 14349–14352.

[67] Rasaiah, J. C.; Garde, S.; Hummer, G. *Annu. Rev. Phys. Chem.* **2008**, *59*, 713–740.

[68] Giovambattista, N.; Rossky, P.; Debenedetti, P. *Phys. Rev. E* **2006**, *73*.

[69] Giovambattista, N.; Rossky, P. J.; Debenedetti, P. G. *J. Phys. Chem. B* **2009**, *113*, 13723–13734.

[70] Strano, K. V. A. S. S. L. W. D. D. K. M. S. *Nat. Nanotechnol.* **2016**,

[71] Miyamoto, S.; Kollman, P. *J. Comput. Chem.* **1992**, *13*, 952–962.

[72] Li, S.; Sang, W.; Lin, R. *The Journal of Chemical Thermodynamics* **2002**, *34*, 1761–1768.

[73] Gan, H.; Benjamin, C. J.; Gibb, B. C. *J. Am. Chem. Soc.* **2011**, *133*, 4770–4773.

[74] (a) Gibb, C. L. D.; Oertling, E. E.; Velaga, S.; Gibb, B. C. *The Journal of Physical Chemistry B* **2015**, *119*, 5624–5638; (b) Gibb, C. L. D.; Gibb, B. C. *Journal of the American Chemical Society* **2011**, *133*, 7344–7347; (c) Carnegie, R. S.; Gibb, C. L. D.; Gibb, B. C. *Angewandte Chemie* **2014**, *126*, 11682–11684.

[75] Weber, P.; Ohlendorf, D.; Wendoloski, J.; Salemme, F. *Science* **1989**, *243*, 85–88.

[76] Szejtli, J. *Chemical Reviews* **1998**, *98*, 1743–1753.

[77] Kim, K.; Selvapalam, N.; Ko, Y. H.; Park, K. M.; Kim, D.; Kim, J. *Chemical Society Reviews* **2007**, *36*, 267–279.

[78] (a) Gunasekara, R. W.; Zhao, Y. *Chem. Commun.* **2016**, *52*, 4345–4348; (b) Rodriguez-Docampo, Z.; Pascu, S. I.; Kubik, S.; Otto, S. *J. Am. Chem. Soc.* **2006**, *128*, 11206–11210; (c) Carrillo, R.; Morales, E. Q.; Martín, V. S.; Martín, T. *The Journal of Organic Chemistry* **2013**, *78*, 7785–7795; (d) Gunasekara, R. W.; Zhao, Y. *J. Am. Chem. Soc.* **2015**, *137*, 843–849; (e) Zhong, Z.; Li, X.; Zhao, Y. *J. Am. Chem. Soc.* **2011**, *133*, 8862–8865; (f) Carrillo, R.; Feher-voelger, A.; Martín, T. *Angew. Chem. Int. Ed.* **2011**, *50*, 10616–10620; (g) Carrillo, R.; Morales, E. Q.; Martín, V. S.; Martín, T. *Chemistry – A European Journal* **2013**, *19*, 7042–7048.

[79] Gassensmith, J. J.; Arunkumar, E.; Barr, L.; Baumes, J. M.; Divittorio, K. M.; Johnson, J. R.; Noll, B. C.; Smith, B. D. *J. Am. Chem. Soc.* **2007**, *129*, 15054–15059.

[80] Peck, E. M.; Liu, W.; Spence, G. T.; Shaw, S. K.; Davis, A. P.; Destecroix, H.; Smith, B. D. *J. Am. Chem. Soc.* **2015**, *137*, 8668–8671.

[81] Liu, W.; Peck, E. M.; Hendzel, K. D.; Smith, B. D. *Org. Lett.* **2015**, *17*, 5268–5271.

[82] (a) Patel, A. J.; Varilly, P.; Chandler, D. *J. Phys. Chem. B.* **2010**, *114*, 1632–1637; (b) Patel, N.; Dubins, D. N.; Pomes, R.; Chalikian, T. V. *J. Phys. Chem. B* **2011**, *115*, 4856–4862.

[83] (a) Pratt, L.; Chandler, D. *J. Chem. Phys.* **1977**, *67*, 3683–3704; (b) Pangali, C.; Rao, M.; Berne, B. *J. Chem. Phys.* **1979**, *71*, 2975–2981; (c) Smith, D.; Haymet, A. *J. Chem. Phys.* **1993**, *98*, 6445–6454; (d) Hummer, G.; Garde, S.; Garcia, A.; Pohorille, a.; Pratt, L. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8951–8955; (e) Ashbaugh, H. S.; Weiss, K.; Williams, S. M.; Meng, B.; Surampudi, L. N. *J. Phys. Chem. B* **2015**, *119*, 6280–6294.

[84] (a) Ashbaugh, H. S.; Weiss, K.; Williams, S. M.; Meng, B.; Surampudi, L. N. *J. Phys. Chem. B* **2015**, *119*, 6280–6294; (b) Koga, K. *J. Phys. Chem. B* **2013**, *117*, 12619–12624; (c) Chaudhari, M. I.; Holleran, S. A.; Ashbaugh, H. S.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 20557–20562; (d) Chaudhari, M. I.; Sabo, D.; Pratt, L. R.; Rempe, S. B. *J. Phys. Chem. B* **2015**, *119*, 9098–9102.

[85] (a) Wilcox, D. S.; Rankin, B. M.; Ben-Amotz, D. *Faraday Discuss.* **2013**, *167*, 177–190; (b) Rankin, B. M.; Ben-Amotz, D.; Van Der Post, S. T.; Bakker, H. J. *J. Phys. Chem. Lett.* **2015**, *6*, 688–692.

[86] Wu, J.; Prausnitz, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9512–9515.

[87] Abascal, J.; Vega, C. *J. Chem. Phys.* **2005**, *123*.

[88] (a) Martin, M.; Siepmann, J. *J. Phys. Chem. B* **1998**, *102*, 2569–2577; (b) Wick, C.; Martin, M.; Siepmann, J. *J. Phys. Chem. B* **2000**, *104*, 8008–8016.

[89] (a) Ashbaugh, H. S.; Truskett, T. M. *J. Chem. Phys.* **2011**, *134*; (b) Ashbaugh, H. S.; Collett, N. J.; Hatch, H. W.; Staton, J. A. *J. Chem. Phys.* **2010**, *132*.

[90] Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327–341.

[91] (a) Widom, B. *J. Chem. Phys.* **1963**, *39*, 2808–2812; (b) Widom, B. *J. Phys. Chem.* **1982**, *86*, 869–872.

[92] (a) Yaacobi, M.; Ben-Naim, A. *J. Phys. Chem.* **1974**, *78*, 175–178; (b) Plyasunova, N.; Plyasunov, A.; Shock, E. *Int. J. Thermophys.* **2004**, *25*, 351–360, 15th Symposium on Thermophysical Properties, Univ Colorado, BOULDER, CO, JUN 22-27, 2003; (c) Shock, E. L. ORganic Compounds HYDration Properties Database. `http://orchyd.asu.edu` (online, accessed August 15, 2016).

[93] (a) Ben-Naim, A.; Wilf, J. *J. Chem. Phys.* **1979**, *70*, 771–777; (b) Ben-Amotz, D. *J. Phys. Chem. Lett.* **2015**, *6*, 1696–1701.

[94] B. Lee, F. R. *J. Mol. Biol.* **1971**, *55*, 379–400.

BIOGRAPHY

James "Wes" Barnett received his B.S. in Mechanical Engineering from Mississippi State University in 2007. After this, he attended New Orleans Baptist Theological Seminary where he received a Master of Divinity in 2011. During that time he also worked for an engineering firm where he tested construction materials used in the New Orleans levee system. Wes joined the Ph.D. program of Tulane's Department of Chemical and Biomolecular Engineering in 2012 and shortly thereafter joined Hank Ashbaugh's research lab. There he received a four-year fellowship from the Louisiana Board of Regents which supported his work on studying deep-cavity cavitands using computer simulations. After receiving his Ph.D. he plans to pursue postdoctoral research.