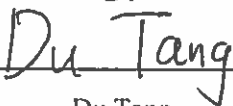SUPRAMOLECULAR ASSEMBLY STATE SWITCHING AND WATER ABSORPTION OF

DEEP-CAVITY CAVITANDS

AN ABSTRACT

SUBMITTED ON THE TWENTY-FOURTH DAY OF JANUARY 2020

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
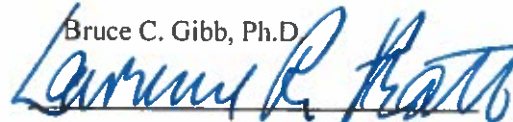
BY

Du Tang

Du Tang

APPROVED: _____

Henry S. Ashbaugh, Ph.D.

Director

Bruce C. Gibb, Ph.D

Lawrence R. Pratt, Ph.D.

Julie N. L. Albert, Ph.D. .

Abstract

Inspired by the study of supramolecular self-assembly behaviors, hydrophobic effect has been identified as one of the prominent driving forces behind these assemblies. To have a better understanding of the molecular scale forces associated with supramolecular host-guest binding processes, and the guest packing motifs within confinement, a family of water-soluble deep-cavity cavitand host molecules were developed. They possess bowl-shaped binding pockets and can form into distinct assembly states with presence of suitably-sized guest molecules in aqueous environment. Besides, their assembly properties can be tuned by changing the guest molecules and the functionalization on their hydrophobic rim. And the hydration states within their hydrophobic pockets, which are also sensitive to different functionalization, can have significant impact on the binding process. We used Molecular Dynamics (MD) simulation to study the thermodynamics of cavitand's host-guest binding process, and the water absorption within the pockets efficiently.

This research begins by a small application of molecular dynamics simulation to predict Second osmotic viral coefficients of small molecules like methane, ethane, and their alcohol counterparts. The second and third parts of this dissertation are studying the different assembly states progression behaviors of different cavitands with minor structural

modification. We used advanced sampling method to evaluate the association free energy of all the different complexes and therefore quantitatively characterize their stabilities. From that, we built up a network reaction model to describe this whole process and used this model to reproduce and predict the distinct assembly states progression trends for cavitands with different structures. The fourth part of this dissertation provides a systematic study of the hydration states within cavitand pockets and its correlation to different functionalities on the rim. A unique self-dewetting behavior were promoted by further methylation on the cavitand portal. Several analytic models were developed to give a deeper insight into this special two-state like behavior. In the last part of this dissertation, we applied a newly developed advanced sampling technique (INDUS) to measure the different responses of water within cavitand pockets with different functionalization. In addition, the different hydrophobicity in different regions of the cavitand can be qualitatively characterized by this method.

SUPRAMOLECULAR ASSEMBLY STATE SWITCHING AND WATER ABSORPTION OF
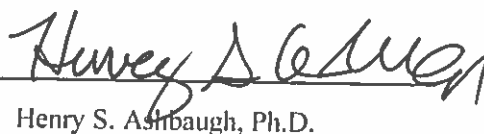
DEEP-CAVITY CAVITANDS

A DISSERTATION

SUBMITTED ON THE TWENTY-FOURTH DAY OF JANUARY 2020

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

_____

Du Tang

APPROVED: _____

Henry S. Ashbaugh, Ph.D.

Director

_____

Bruce C. Gibb, Ph.D.

_____

Lawrence R. Pratt, Ph.D.

_____

Julie N. L. Albert, Ph.D.

ACKNOWLEDGEMENT

during all my time in graduate school. I also thank the Louisiana Optical Network Initiative for their computational resources.

During my time at Tulane, I have the opportunity to interact and build friendship with many brilliant and wonderful people. I would express my sincere appreciation to Dr. James "Wes" Barnett for his guidance in my early stage of research work. He also led me into the wonderful world of scientific programming and computation. I am also grateful for all the fellow students in our group and their help, suggestions, and insights into this work: Yang Wang, Amy Goodson, Dr. Alexander Saltzman, Busayo Alagbe, Dr. Lalit Surampudi, Dr. Bin Meng, Nusrat Muhith, Tobias Dwyer, Hussain Bukannan, Odella Blackmon, and Courtney Delpo. Outside our research group, I also met lots of people who enlighten my life at Tulane. Dr. Liang Tan is the best roommate in the world. He not only introduced me to the world of machine learning and data science, but also gave me a number of advices on Dota and StarCraft. I consider Dr. Ajay Muralidharan as one of the most open-minded thinkers I have ever known, and I enjoyed every minute of academic and non-academic discussions with him. With lots of gratitude, I appreciate the time with Danqia Peng in the world of PokemonGo when I was feeling overwhelmed by science and coding. I also enjoyed so much in the encouraging conversations, road trips, advices for life and career paths, and deep friendship with lots of people in the Department of Chemical and Biomolecular Engineering: Yang Wang, Dr. Mengyuan(Sonya) Zheng, Dr. Yueyun(Maya) Lou, Dr. Lelia Pashazanusi, Dr. Meysam Shahami, Dr. Yueheng Zhang, and Dr. Yang Su. Outside the department, I also met people who made my journey in graduate school more fulfilling. I am so happy to have a lifetime friend like Dr. Shuwen Lin from medical school. She taught me how to actively integrate into life in another country, embrace the

Table of Contents

List of Tables

List of Figures

Chapter1 Introduction

In our nature, particularly in bio-systems, self-assembly behaviors and nano-scale binding are widely utilized to achieve unique and special functions that can hardly be reached in bulk environment. The GroEL-GroES chaperonin complex can bound proteins in its central cavity. And those proteins will undergo conformation rearrangement and result in folding[1]. CCT chaperones can also assist and regulate the proper folding of proteins like tubulin[2]. Similarly, vaults, found in numerous eukaryotic species[3], can form significant associations with nucleus and are localize to them[3,4]. Those vaults proteins are suspected to serve as transporting pipelines between the cytoplasm and the nucleus[4]. DNA can also be encapsulated in viral capsids with presence of high internal pressures. Therefore this tight encapsulation can accumulate enough energy and result in DNA ejection to host molecules[5,6]. Two-partner secretion is also assisted by the recognition and binding between proteins. The two periplasmic POTRA domains of a TpsB transporter FhaC can recognize and associate with their TpsA partners, FHA. The secretion of TpsA will then be mediated by their fast association and dissociation[7]. Protein-ligand binding is also utilized to transport fatty acids across cell membranes. For instance, crystal structure of FadL, one type of outer membrane proteins, was reported to explain its mechanism to bind with long-chain fatty acids (LCFAs) on multiple sites and convey them by spontaneous conformational changes[8]. Similar research was also conducted to study the binding

between rat intestinal fatty acid binding protein (I-FABP) and hydrophobic ligands[9]. Proton conduction can even be achieved by the unique properties of an array of water in hydrophobic cavities of haem-copper oxidases[10]. The conduction pathways can be controlled and switched by redox-state-dependent electric field and therefore determine the proton transfer direction. In a word, this list of the biological systems taking advantage of binding and self-assembly associations goes on.

The aspiration to utilize those self-assembly behaviors and control their functions precisely has stimulated the study and synthesize of those assemblies. And due to the irregular and complicated structures of natural protein molecules' binding sites and cavities, structures with simple shapes and well-defined cavities are drawing attentions. Under this circumstance, the concept of supramolecular chemistry is emerging. In 1960s, Pedersen discovered crown ether, which is the first time scientists were able to synthesize self-assembly host molecules[11]. Crown ether is a cyclic-shape molecule and it can stabilize a sodium ion in the center hole by electrostatic interactions. From then on, the wonderful world of supramolecular chemistry opened its door to us.

Multiple mechanisms and routes, especially non-covalent interactions between molecules, can be applied to form self-assembly containers or capsules. Among those mechanisms, metal coordination is a very strong driving force. Fujita[12] reported a family of three-dimensional assemblies formed by planar *exo*-multidentate organic ligands through metal-coordination. They are "paneled" by square planar metals like Pd^{2+} and Pt^{2+}. Recently, a very large tetravalent Goldberg polyhedra in the same family assembled from 48 palladium ions and 96 bent ligands was also reported[13]. It has square packing, which is different from the hexagonal packing usually found in nature, like graphite.

Alternately, hydrogen bonding can be used to form capsules with interior pockets in non-polar environments. A chiral spherical self-assembly complex was reported by MacGillivray and Atwood[14]. This complex is formed by six calix[4]resorcinarenes and eight water molecules, and those constituents were held together by 60 hydrogen bonds. The water molecules were found to be in a fast exchange with the bulk water[15]. Alcohol molecules like 2-butanol, 2-hexanol, 2-ethylhexanol, and 2-ethylbutanol were found to be either encapsulated within this complex or become part of the structure of it[16]. The fluorescence resonance energy transfer (FRET) was applied to study the dynamics of this self-assembly process[17]. The exchange of resorcinarene monomers can be observed and FRET can occur across the mechanical boundary with a presence of a fluorescent guest inside the capsule. Another cylindrical dimeric resorcinarene capsule in apolar solvents was also reported[18,19]. These nano-sized capsules are held together by a cyclic array of bifurcated hydrogen bonds. The structure of the capsule can even be adjusted depending on the size and configuration of the carbon chains at the bottom of each resorcinarene. Small and mid-sized guest molecules like propionic, pivalic acid, CCl_4, *p*-ethyltoluene, and normal alkanes up to tetradecane can be encapsulated within the capsules[20,21]. Moreover, glycoluril molecules can be inserted and therefore incorporated in the capsule to expand the length of the cavity. As a result, longer alkane chains can accommodate themselves within this extended inner-space[22,23]. Likewise, insertion of propanediurea into this cylindrical capsule can form new capsules with "S" and "banana" shapes[24]. This bent architecture allows the encapsulation of guests with complementary shapes and provides possibility for wider application of concomitant guest molecule recognition.

All these containers are held together by non-covalent driving forces and possessing large inner spaces. In addition, they exhibited a wide variety of applications in molecule recognition and capture. However, containers formed through metal-coordination may also potentially exhibit toxicity. And capsules held by hydrogen bonding are not stable in water since water itself can form hydrogen bonds easily both as a donor and an acceptor. Therefore, it is very unlikely to utilize these mechanisms in biological or environmental applications[25]. Naturally, in order to build more eco-friendly and water-soluble supramolecular self-assembly systems, hydrophobic effect entered people's field of vision. Host molecules can capture guest molecules based on hydrophobic effect with various shapes and applications were discovered and designed.

Cucurbituril is the generic term of a family of supramolecular host molecules with a pumpkin-like shape. They are composed of several glycoluril units, which are connected by methylene bridges[26]. These glycoluril units forms a cyclic multimer with a hydrophobic cavity in the center. In very early studies, it was found that alkylammonium cations can be encapsulated within this hydrophobic cavity[27]. Due to its outstanding binding properties, many designs and applications were developed based on cucurbituril. Unusual dynamic phenomena like kinetic self-sorting behaviors were found within CB[6] and CB[7] in the presence of two-faced guests that contain two distinct binding epitopes[28]. Another important application using cucurbituril is molecular switches[26]. The self-sorting of some certain configurations of the spermine derivatives polyaminated axle can be achieved in the presence of CB[6] and CB[7] driven by different temperatures[29]. A pH-dependent switch was observed in the association between CB[7] and a V-shaped cyanine dye[30]. By

protonation of the aniline group, the CB[7] ring will switch from the cyanine branch of the dye to the aminophenoxy ethyl side.

Another important family of host molecule that utilize hydrophobic effect to form host-guest complex systems is cyclodextrin. They are synthetic products from enzymatic degradation of starch[31]. They exhibited a truncated cone structure with a cyclic chair of glucopyranose units, and formed a hydrophobic cavity in the center[32]. According to different number of glucopyranose units, natural cyclodextrin can be divided into three categories, α-CD, β-CD, and γ-CD. They contain six, seven and eight glucopyranose units respectively[33]. They can capture guest molecules within the central hydrophobic cavities and form inclusion complexes. The guest molecules can be encapsulated include a very wide range of solid, liquid, and gaseous compounds[34]. As a result, cyclodextrin and their derivatives are actively used in the area of drug delivery. For instance, the 9-nitro-camptothecin(9-NC)/hydroxypropyl-β-CD complex exhibited a significantly superior anti-tumor activity compared to that of free 9-NC on tumor-bearing mice[35]. Lonidamine also shows enhanced solubility with a formation of inclusion complexes together with hydroxypropyl and partially methylated cyclodextrins[36]. Some drug delivery carriers were also developed from cyclodextrin-based nanoparticles cyclodextrin grafted polymers[32]. The drug-loading values of camptothecin (CPT) were significantly higher with β-CD nanoparticles compared with those with PLGA and PCL nanoparticles. And the drug release with β-CD nanoparticles was also more controllable[37]. A β-CD functionalized hyperbranched polyglycerol was also reported and can serve as efficient carrier of paclitaxel. They can assemble into dispersive spherical micelles in water, and exhibited not only high loading capacity but also long extended release profile[38].

Though it is a state of prosperity to apply hydrophobic effect as a driving force to trigger self-assembly behaviors in supramolecular chemistry. The fundamentals behind the binding mechanisms of hydrophobic effect remains unclear. The influence brought by host structures, guest conformations, guest packing, and interactions between host and guest is less understood. Therefore, precise and controlled design of hydrophobic self-assembly devices is still very difficult. Traditionally, hydrophobic effect was accepted as a entropy-driven process[39,40]. It was described that water will form a "cage"-like structure around the nonpolar surface. The formation of this "cage"-like structure is because the original hydrogen-bonding network structure of water was break by the introduction of nonpolar molecules. Therefore, with the assembly of nonpolar molecules, the number of solvating waters required is less and leads to an entropic advantage[41]. In other word, the break of the original water hydrogen-bonding network is "minimized". However, many experimental and computational results, especially in the host-guest self-assembly systems with different host cavity shapes, indicate different understandings about the essence of hydrophobic effect. For instance, both by calorimetric study and van't Hoff analysis of HNMR, the inclusion complexation of pyrene into cyclophanes is enthalpically driven[42]. Cucurbituril was also reported to be able to bind adamantane and bicyclo[2.2.2]octane derivatives with a enthalpy-driven process[43]. And molecular dynamics simulation research indicates this is due to the release of high-energy water from its central hydrophobic cavity[44]. With the increase of cavity size, the energy of each individual water molecule encapsulated within the cavity decreases because more hydrogen bonds can be formed with more neighbor water molecules. Conversely, the number of water molecules bounded within the cavity will increase. Therefore, the binding affinity of different sized cucurbiturils can be

determined by these two counteracting effects. And as a result, CB[7] exhibits the largest binding affinity. This theory of the enthalpic driving force from the replacement of high-energy water in cavities by guest molecules was also extended to other hydrophobic hosts like cyclodextrins and calix[4]arene[41]. A theoretical study of an idealized cavity-ligand model also showed that the repulsion of disorganized water from the hydrophobic environment of cavity to bulk water actually determined the thermodynamic of the binding process. And this cavity-ligand binding is also enthalpy dominating, instead of entropy dominating[45].

In a word, though there is no universal theory, water plays a vital role in the whole self-assembly and binding behaviors of supramolecular chemical systems and biosystems. Water may exhibit special behaviors within binding cavities and unique wetting and dewetting phenomenon was also observed and studied. For instance, positionally disordered water molecules were initially found inside the large hydrophobic cavity of human interleukin-1β (hIL-1β) protein by NMR[46]. But further molecular dynamics simulation investigation showed these water molecules in the central cavity of hIL-1β are thermodynamically unstable. The transfer free energy of water molecules from bulk phase to the hydrophobic cavity was shown to be positive[47]. A similar dewetting phenomenon was also reported in the mutant T4 lysozyme cavities[48]. The cavity was found to be empty at ambient pressure, and on average two water molecules were found occupying the cavity at an elevated pressure of 200 MPa. And this transition behavior was also shown as a cooperative transition[49] and was believed to be an important contributor in the pressure-induced destabilization of protein hydrophobic cores[48]. A large and completely dry ligand-

binding cavity with a volume of 315 $Å^3$ was also found and thoroughly studied in bovine

β-lactoglobulin (BLG), which provides the prerequisite of a productive ligand binding[50].

The wetting and dewetting behaviors and mechanisms also largely depended on the shapes and sizes of cavities. In the case of carbon nanotubes (CNTs), water molecules are more stable inside the CNTs than in the bulk[51]. An entropy stabilized vapor-like phase of water was founded in small-sized CNTs, while an enthalpy stabilized solid-like phase of water was found in middle-sized CNTs. In large-sized CNTs, water behaves more like bulk and liquid phase. Likewise, thermodynamics and water cluster structures are found to be highly sensitive to the sizes of graphene-like spherical cavities and that of rigid fullerenes[52]. With the increase of cavity size, the hydration state with more water molecules are more favorable. And the van der Waals interaction between water and cavity walls also influenced the hydration states a lot. In fullerene cages, a linear relationship between maximum encapsulated water molecule numbers and the cavity volumes were also concluded[53]. The thermodynamics of wetting behaviors were also influenced by local or external electric field. In CNTs, a two-state behavior was also observed where there are two local minimums in the water filling state distribution. The stronger electric fields along the tube axis will shift the filling equilibrium towards the filled state[54]. In a generic hydrophobic nanopore, under equilibrium conditions it is impermeable to water, but a strengthened electric field will drive water molecules inside[55].

As described above, the water molecules inside hydrophobic binding cavities and the wetting/dewetting behaviors inside those binding sites are nonnegligible in the self-assembly behaviors. This is a sophisticated process currently without universal theory and can be affected by numerous factors. To address this complexity and study the

hydrophobic-driven self-assembly behavior systematically, a supramolecular deep-cavity host molecule, octa-acid, and its derivatives are developed and studied[56]. They are formed by three rows of aromatic rings that build a bowl-shaped hydrophobic pocket. This pocket is approximately 8 Å in diameter and 8 Å deep. The foot and rim of the molecule are decorated with eight carboxylic-acid groups to enable water solubility. Via hydrophobic effect, they can form distinct complexes with the presence of different guest molecules. Initially, short hydrocarbon molecules like propane and butane can form quaternary complex (2:2) with octa-acid, while ethane only very weakly binds to the hosts. Therefore, it is proposed that deep-cavity cavitand can serve as gaseous-phase hydrocarbons separator[57]. The large difference between cavitand's binding affinities with different structured isomeric long-chain esters was also reported. Cavitands strongly bind with methyl ester and have very weakly binding with ethyl-ester and propyl-ester. This results from a combined contribution from the steric clash of the ester group deeply inside the pole-region of the capsules and $-CH_3\cdots\pi$ interactions between host inner walls and methyl groups of the ester. Therefore, in a competition experiment, most of the strong binders were captured within capsules, but weaker binders primarily remained in bulk medium[58]. The potential usage of deep-cavity cavitand as nano-scale reactors was also intensively studied. One typical instance is the photolysis of dibenzyl ketone derivatives[59,60]. The reaction occurred within capsule will result in unusual products since the radical has to rotate and reorient itself to better accommodate within capsules.

Further research shows this guest effects and host effects both contributed to this binding behavior[61]. For instance, a series of cyclic and acyclic carboxylates can bind with octa-acid and form 1:1 complex. Isothermal Titration Calorimetry (ITC) results showed, for both

cyclic and acyclic carboxylates, there is a linear inversely proportional relationship between the binding free energy and number of carbon atoms in the guest molecule. But in terms of enthalpy and entropy, different trends were observed. For acyclic guests, both entropy and enthalpy terms are more favorable with increase of carbon atom number. In contrast, for cyclic guests, entropy term is less favorable with more carbon atoms, and there is a enthalpy-entropy compensation[62]. This phenomenon was attributed to the integrated effect from increasing host-guest contacts and the restriction motion of larger guests[61].With presence of hydrocarbons, octa-acid can form stable dimeric capsules. In details, small guests from propane through heptane formed 2:2 quaternary complexes. Octane serves as a special guest in that it formed a mixture of 2:2 and 2:1 complexes[63]. Nonane is the shortest guest that can form stable 2:1 complexes with octa-acid. And smooth transitions of guest packing motifs within capsules can be observed as a function of chain length[64]. Short guests like nonane and decane still exhibit fully extended conformations. Middle-sized guests such as tetradecane adopted a helical packing motif enriched in gauche dihedral angles. The dominating packing motif gradually transformed to a hairpin motif in the case of longer guests like heptadecane or octadecane. For guests longer than tricosane, they showed a spinning top packing motif, where the middle part of the guest underwent a hairpin turn and was exposed partially to bulk solvent. Demonstrated by computational work, the packing under confinement to minimize the contact between nonpolar host-guest systems and water is the main driving force of this conformation transition[65]. The polarities of guest molecules can also influence the formation of capsules with octa-acid. In general, the guests with less oxygen atoms in its chain, such as tridecane and 1-dodecanol formed

stable 2:1 complexes with hosts. In contrast, guests with higher polarities can only form unstable capsules relative to NMR timescale or simple 1:1 complex[66].

Modification on deep-cavity cavitand structures also has significant impact on their binding behaviors. Tetra-*endo*-methyl octa-acid (TEMOA) is a derivative of octa-acid. It has four additional methyl groups on *endo* positions of the hydrophobic rim. Therefore, it has a narrower portal but a deeper pocket. It exhibited a very unusual nonmonotonic assembly behavior with hydrocarbons[67]. Initially, it will form monomeric assembly with methane and ethane. With the increase of alkane chain length, dimeric assemble was triggered by heptane. Unlike octa-acid, monomeric assembly emerged again with heptane and octane. For long guests from nonane, just as in the case of octa-acid, stable 2:1 complexes were formed. Another strange unusual behavior related to TEMOA was also reported with hetero-capsule formation[68]. With the mixture of TEMOA and octa-acid, the extent of hetero-capsule formation was different for various guests. In the case of heptane and octane, much less proportion of host molecules formed into hetero-capsules than with presence of pentane, hexane, and long guests from nonane to tetradecane. With long alkane guests, TEMOA also exhibited other assembly pathways. They can assemble into tetrameric capsule encapsulating two guest molecules from heptadecane to eicosane. With even longer guests such as pentacosane and hexacosane, a hexameric 6:3 complex was observed[69]. In a word, small modifications on the deep-cavity cavitand based host-guest self-assembly systems will largely impact and determine the binding behavior and the thermodynamics behind it.

To have a deeper investigation and understanding of the guest packing, host effects, water absorption behavior, and their contributions to the thermodynamics of this self-assembly

process, we utilized Molecular Dynamics (MD) simulation throughout the whole dissertation. Molecular Dynamics simulation can provide us a molecular-level picture with deep insight of how molecules interact with each other and how they adopt their configurations for the self-assembly process. In general, a potential energy function used in MD simulation is illustrated as below[70]:

$$U = \sum_{bonds} k_r(r - r_{eq})^2 + \sum_{angles} k_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{v_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j}[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\varepsilon R_{ij}}]$$

This function describes the sum of potential energy in the whole simulation system. The first three summation terms are bonded interactions. They represent bond interactions, angle interactions and dihedral interactions respectively. The last summation term represents pairwise non-bonded interactions. In details, in the last term, the first two terms are Lennard-Jones potential describing van der Waals interaction, dispersion interaction, and repulsion between two atoms. The third term represents the electrostatic interaction between two atoms. In practice, all these parameters in the above function will be provided in specifically chosen forcefields. Ewald summation is usually applied to treat long-range electrostatic interactions[71].

In a typical MD simulation, we input an initial configuration of the whole system, where all the atoms have their initial positions. From the chosen forcefield, a potential interaction of each atom as a function of atom positions can be obtained. And initial velocities are randomly generated for all atoms. Therefore, the force on each atom can be calculated from its potential interactions. By numerically solving Newton's equations of motion, the movement of each atom can be evaluated. In each step, the position, velocity, and potential energy of each atom is updated. All the information plus temperature and pressure of the

system provided us a dynamic picture of the simulated system[72]. By every a few steps, the data of all the information will be saved for analysis.

The outline of this dissertation is listed here. In chapter 2, we described an application of molecular simulation to predict variables related to molecular-scale hydrophobic interactions. Second osmotic viral coefficients of methane and ethane, along with their alcohol counterparts, were predicted from simulation and compared with experimental results. In chapter 3, a systematic series of simulations were conducted to evaluate the potential of mean force between deep-cavity cavitand hosts and long *n*-alkane guests. A network reaction model was built up to predict the distribution of different assembly complexes of octa-acid and TEMOA, and therefore to explain the unusual nonmonotonic progression of assembly states of TEMOA with increasing guest length. Follow on in chapter 4, a similar prediction of the assembly states progression was made for a newly developed deep-cavity cavitand derivative. This new host molecule was named as tetra-*exo*-methyl octa-acid (TEXMOA) since it has four methyl units pointing upwards on *exo* positions of its hydrophobic rim. This minor structural modification may lead to significantly difference not only in assembly behaviors, but also in its pocket hydration and water absorption. And as described above, water itself is very important in the driving force of supramolecular self-assembly behaviors. Thus, in chapter 5, we presented a detailed study about wetting and dewetting phenomenon in multiple types of deep-cavity cavitand molecules. All the host molecules studied have similar structures with octa-acid, but minor modifications were made on their functional groups on the rim. The probability distributions of observing certain water molecules within the hydrophobic pocket were evaluated, and a thermodynamic model was fitted to give a deeper insight of this

phenomenon. In chapter 6, a newly developed technique named indirect umbrella sampling (INDUS) was applied to deep-cavity cavitand molecules to give a finer illustration of water behaviors inside the pocket and around the whole host molecule.

From Appendix A through Appendix E, supplemental materials like experimental details, additional simulations and extra calculations were presented for chapter 2 through chapter 6. In Appendix F, some related simulation and calculation to the drying of hydrophobic pocket of octa-acid, TEMOA, and TEXMOA were discussed in detail. More interpretations will also be included.

Chapter2
Second Osmotic Virial Coefficients of Short Alkanes and Their Alcohol Counterparts in Water as a Function of Temperature

2.1   Summary

Second osmotic virial coefficient is not only a good implication of inter-solute interactions in solution, but also a measure of hydrophobicity of small molecules. Here, we provide a series of molecular dynamic simulations for small alkanes, methane and ethane, and their alcohol counterparts in water for a wide range of temperatures from 275K to 370K. The second osmotic virial coefficient, B2, for those solutes were evaluated using an augmented form integral. The temperature dependence of B_2 are in good consistent with available experimental results. The aggregation states and hydrophobic interaction reflected by B_2 values are also in agreement with experimental observations. The accuracy of our simulations indicated that the prediction of B2 values from simulations in a closed ensemble and in finite concentration by extrapolation is feasible and can be applied to study the hydrophobic interactions between small solutes.

2.2   Introduction

The thermodynamic driving force of many different aqueous phase assembly processes can be acquired from hydrophobic interactions. However, due to the very low solubility of many non-polar molecules in water, it is very hard to explore the molecular-scale hydrophobic interactions in experiments without any polar groups. While we can add polar groups into the solutions to increase the solution concentrations, much of our understanding of molecular-scale hydrophobic interactions ultimately comes from molecular simulations. Molecular dynamics simulation can help us overcome the solubility limits. And the main method applied to quantitatively study the hydrophobic interactions is the solute-solute radial distribution functions in solution, $g_{ss}(r)$. It cannot be obtained directly from the experiments. But the thermodynamic consequences of solute-solute interactions can be evaluated from experiments by the viral expansion of the osmotic pressure as following.

$$\frac{\Pi}{RTc_s} = 1 + B_2 c_s + order(c_s^2) \tag{2.1}$$

Here Π is the osmotic pressure. RT is the product of the ideal gas constant and absolute temperature. B_2 is the second osmotic virial coefficient, and c_s is the solute concentration. B_2 depends on molecular-level solute pair correlations in aqueous solution. It can be seen as a measure of solution non-ideality arising from pair correlations. From McMillian-Mayer theory[73], a fundamental connection between B_2 and $g_{ss}(r)$ can be expressed as below.

$$B_2 = \frac{1}{2} \int [1 - g_{ss}(r)] dr \tag{2.2}$$

From the relation above, B_2 can be expected to be negative when solute-solute interactions are attractive in solution. Conversely, repulsive interactions will make B_2 more positive. Though it is very hard to get well-converged values for B_2, an augmented form for this integral can be used to evaluate it[74]. And some recent researches, including some previous researches from our group, showed that we are able to get accurate results for simple non-polar gas molecules in water[75–78]. The results for B_2 obtained from that integral in augmented form is in agreement with those determined from the solute chemical potential's concentration dependence extrapolated to infinite dilution[76], despite that simulations were conducted at small but finite concentrations.

We provide a series of molecular simulations and expand this study on B_2 to methane and ethane as well as their alcohol counterparts, methanol and ethanol. The results of alcohol can be directly compared with experiments. Therefore, we can evaluate the abilities of molecular dynamic simulations to predict molecular-scale hydrophobic interactions.

2.3   Methods

All the simulations of methane, ethane, methanol, and ethanol were conducted in isothermal-isobaric ensemble. We used the GROMACS 5.1.3[79] software package to do the simulations. The temperatures and pressures were coupled by the Nosé-Hoover thermostat[80,81] and the Parrinello-Rahman barostat[82], respectively. The temperature was from 275K to 370K in 5 K increments, and the pressure was kept at 1 bar. Water molecules were modeled using TIP4P/2005 potential[83]. And all the solutes were modeled by the TraPPE-United Atom potentials[84,85]. It has demonstrated that those force-fields can have

the best descriptions and illustrations of the equations-of-states of the small alcohol and alkane molecules. The internal constraints for solutes were maintained by the LINCS algorithm[86], while that for water molecules were maintained using the SETTLE[87]. A truncation cutoff of 14 Å for non-bonded Lennard-Jones interactions were also used. The cross interactions between Lennard-Jones sites were calculated by Lorentz-Berthelot combining rules[88]. The long-range electrostatic interactions were calculated using the particle mesh Ewald summation method[71]. In each simulation, 40 solute molecules were solvated in a box of 4000 TIP4P/2005 water molecules. A 100 ns production run was conducted to get trajectory and evaluate equilibrium averages, following a 5 ns of equilibration simulation. The pair-correlations determined from these simulations, though at finite concentration, have been found to be consistent with those results at infinite dilution[76]. The pair-correlations were determined based on the solute center-of-mass. Although the radial distribution function used in eq. (2.2) should be calculated in the grand canonical ensemble, previous simulation results have demonstrated those determined from closed ensembles yield thermodynamically consistent results when the integrals are extrapolated from finite volume results that effectively mimic an open system[76,77,89,90]. Statistical errors were determined using jackknife resampling. The second osmotic virial coefficients of methane, ethane, methanol, and ethanol calculated from eq. (2.2) are reported in appendix.

2.4   Results and Discussion
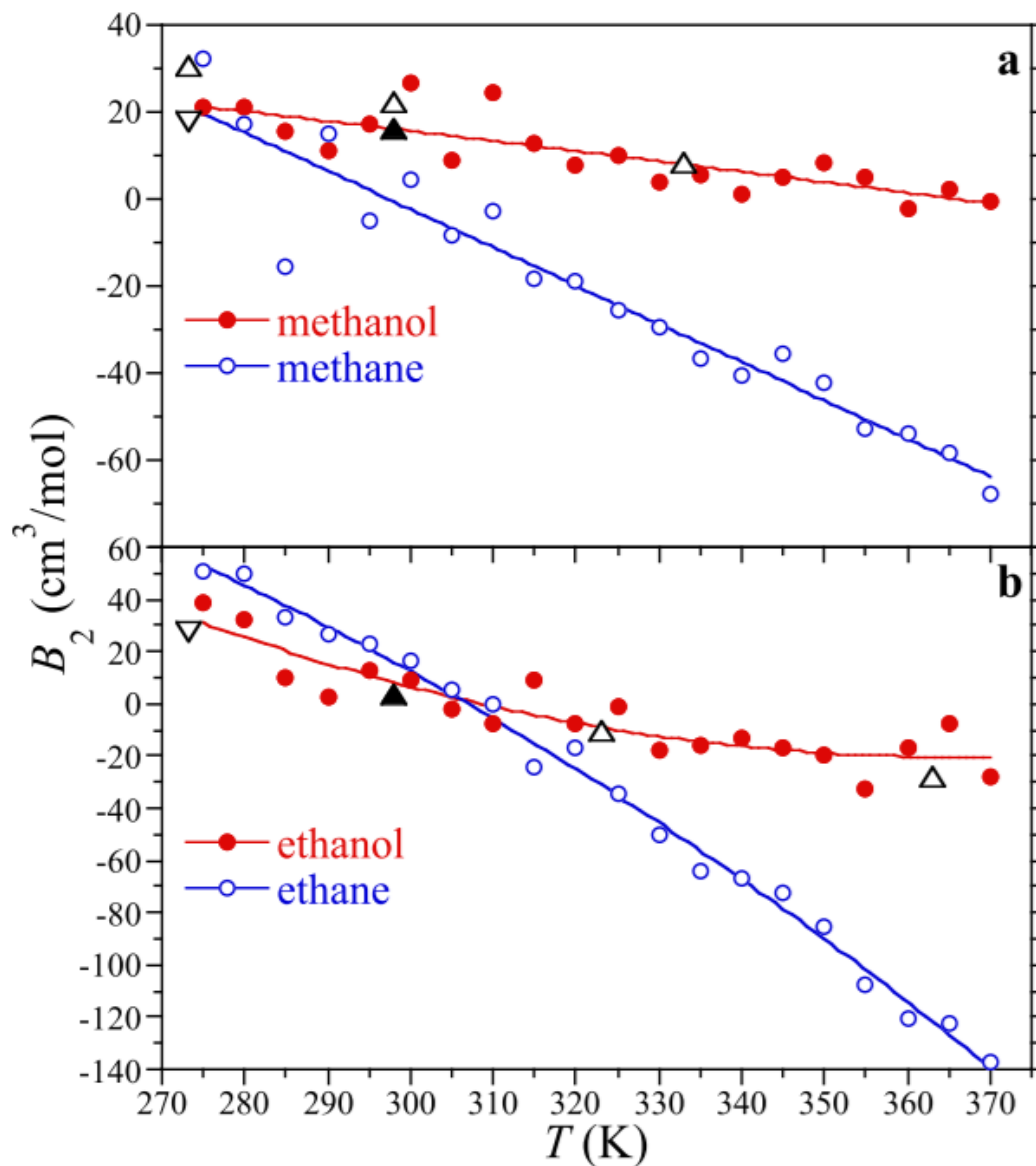
Figure 2.1: Second osmotic virial coefficients of methane and methanol (a), and ethane and ethanol (b) as a function of temperature. The symbols are defined in the figure legends. All the simulation results are compared with experiment results for methanol and ethanol from references. The symbols of experiments results are black and white triangles in the figure. The curves are just serving as a guide for the eye. The error bars are comparable to the figure symbols.

The second osmotic viral coefficients are plotted in figure2.1. Generally, B_2 for methane and ethane decrease from repulsive region (positive values) to attractive region (negative values) with the increase of temperature. That means when temperature increases, the solutes prefer to aggregate together in solution. For methane, it changed sign from positive to negative just above room temperature. For ethane, the B_2 is in a magnitude approximately twice that of methane, indicating ethane is a much larger non-polar solute. It also changed sign at ~305K. The simulation may not fully converge after 100 ns because the scatter in our results is larger than the statistical uncertainty determined in jackknife resampling. But the magnitudes and trends with temperature for both methane and ethane are consistent with each other, which suggests that our simulation are accurate, and our comparisons are valid.

For methanol and ethanol, the B_2 values are generally more positive than that of their alkane counterparts. Due to the polar hydroxyl group, they are not as hydrophobic as methane and ethane. Therefore, the solubility of methanol and ethanol is always much higher. In the case of methanol, B_2 is predicted to be in repulsive region during the whole temperature range. But it will still decrease with temperature. In the case of ethanol, it will ultimately change sign and go to attractive region near room temperature. But in total, their temperature dependence is much weaker than that of alkanes. The experimental results of B_2 for methanol and ethanol are obtained from some references and reported in figure2.1[91–93]. The black triangles are for results from ref.93. The white triangles and white inverted triangle are for results from ref.91 and ref.92, respectively. It can be concluded that the experiments and simulations are in excellent agreement with other in a wide range of

temperatures. Thus, we are confident that our simulations accurately capture the relationship between temperature and the solvent-mediated interactions between solutes.

At the critical point, $RT_{crit}/\varepsilon = 1.326$, Lennard-Jones (LJ) fluid has a second gas-phase virial coefficient of $B_2 = -3.20 \, \sigma^3$. Here, σ and ε are the LJ diameter and well-depth. Indicated by methane and ethane's van der Waals volumes, the van der Waals diameters for them are 3.8 Å and 4.15 Å respectively. So, the critical LJ B_2 values are therefore approximately -105.7 cm^3/mol and -137.7 cm^3/mol for methane and ethane. And the B_2 is required to be significantly lower than its critical value for those solutes to phase separate into a liquid-like drop in water. These critical LJ B_2's is more negative than those reported in figure 2.1 for methane and ethane, let alone methanol and ethanol, for almost the whole range of temperatures simulated. The only exception appears near water's boiling point for ethane.

In addition, at room temperature, the B_2 for both alkanes and ethanol is nearly zero. For methanol, it is positive, and its absolute value is small, which is less than half of its molar volume. It implies that methanol molecules are not strongly attracted to each other at room temperature. Instead, they are dispersed in solution with a near ideal gas-like osmotic pressure. The experiment using Raman hydration shell spectroscopy and polarization-resolved femtosecond infrared measurements also exhibited that short chain alcohol molecules are randomly distributed in water[94].

2.5   Conclusions

We conducted the simulations of short alkanes, methane and ethane, and their alcohol counterparts in water with a wide range of temperatures. The second osmotic viral coefficients were evaluated from simulation and they are in excellent agreement with available experimental results, which means the simulations correctly reproduce the inter-solute interactions in water. We may also conclude that the underlying description of hydrophobic interactions is also accurate by extrapolation. The hydrophobic interactions between small alkanes and alcohols are not strong enough for them to assemble in water in almost the whole range of simulated temperatures. Cooperative interactions between larger non-polar surfaces might be essential for self-assembly as concluded in the analysis of Lum, Chandler, and Weeks[95].

Chapter3
Guest Controlled Nonmonotonic Deep Cavity Cavitand Assembly State Switching

3.1   Summary

Octa-acid (OA) and tetra-*endo*-methyl octa-acid (TEMOA) are water-soluble, deep-cavity cavitand hosts with nanometer-sized hydrophobic pockets that readily bind nonpolar guests, like the n-alkanes. Experimentally the host-guest stoichiometry in these assemblies depends sensitively on the length of encapsulated n-alkanes. OA, for instance, exhibits a progression of 1:1 to 2:2 to 2:1 host-guest complexes (X: Y where X is the number of hosts and Y is the number of guests) with increasing alkane chain length from methane to tetradecane. Despite differing from OA by only the addition of 4 methyl groups, TEMOA exhibits a non-monotonic progression of states from 1:1 to 2:2 to 1:1 to 2:1 with increasing guest length. Understanding the interplay between host chemistry and the equilibrium structures formed can impact the potential applications for these supramolecular assemblies. Here we present a molecular simulation study of n-alkane association with both OA and TEMOA to form the full range of structures observed to uncover the reasons underlying their distinct assemblies. Potentials-of-mean force between hosts and guests, determined via umbrella sampling, are used to characterize their association free energies. The free energies so determined are subsequently used within the context of a reaction network model to rationalize the sequence of assemblies formed. Our modeling results

show that TEMOA's added methyl units constrict the opening of the binding pocket, limiting the conformations available to encapsulated guests and thereby shifting the balance between monomeric and dimeric capsules. And the comparison between the experimental results and our model prediction demonstrated that our model accurately reproduced the experiments. And a very small chemical modification on supramolecular assembly molecule can largely impact their structures and properties.

3.2   Introduction

Compartmentalization can be utilized in many novel ways. For instance, it can be applied to organize matter and broaden the utility of X-ray crystallography[96–97], some novel separation protocols[57,58,98], and control chemical catalysis[99,100]. And self-assembly provides a very efficient way to engender compartmentalization and even precisely control the spatial positions of atoms. Many structurally well-defined supramolecular complexes assembled by various methods are manipulated and formed[61,101]. Of those different methods, utilizing the hydrophobic effect is the least explored, but also may provide a chance to get insight of the relationship between solute shape, solvation, and the hydrophobic interactions[102,103].

Gibb has explored the complexation and assembly of deep-cavity cavitands such as octa-acid (OA)[56,63], and tetra-endo-methyl-octa-acid (TEMOA)[69]. These water-soluble host molecules are formed by three rows of aromatic rings, which form a bowl-shape hydrophobic pocket. It is approximately 8 Å in deep and 8 Å in diameter. Their rim and feet are decorated with eight carboxylic-acid groups to enable water solubility. TEMOA

differs from OA by the presence of four methyl groups, which narrows the portal and deepening the pocket, projecting to the hydrophobic rim. This seemingly minor structural modification dramatically changes the assembly behaviors of the hosts. Let's consider OA and TEMOA complexation with n-alkane guests. Alkanes are very hydrophobic, and thus they can be encapsulated into the hydrophobic pockets of both hosts and form capsules[63,67,69]. In the case of OA, the relationship between guest size and assembly state is very straightforward[67]: methane (C_1) does not bind to the host; ethane (C_2) forms 1:1 complexes (*X:Y*, where *X* is the number of hosts and *Y* is the number of guests); the guests from propane (C_3) to n-octane (C_8) forms 2:2 complexes; and longer guests from nonane (C_9) forms 2:1 complexes. In the case of TEMOA, it exhibits a totally different nonmonotonic trend[67]: C_1 and C_2 form 1:1 complexes; guests from C_3 to C_6 form 2:2 complexes; 1:1 complexes reappear with the presence of guests of C_7 and C_8; and longer guests from C_9 still form 2:1 complexes. Moreover, it has been observed that TEMOA will also form into larger capsules like 4:2 tetrameric and 6:3 hexameric complexes[57]. But OA only forms monomeric and dimeric complexes. Besides, different guests will exhibit different packing motifs inside different complexes.
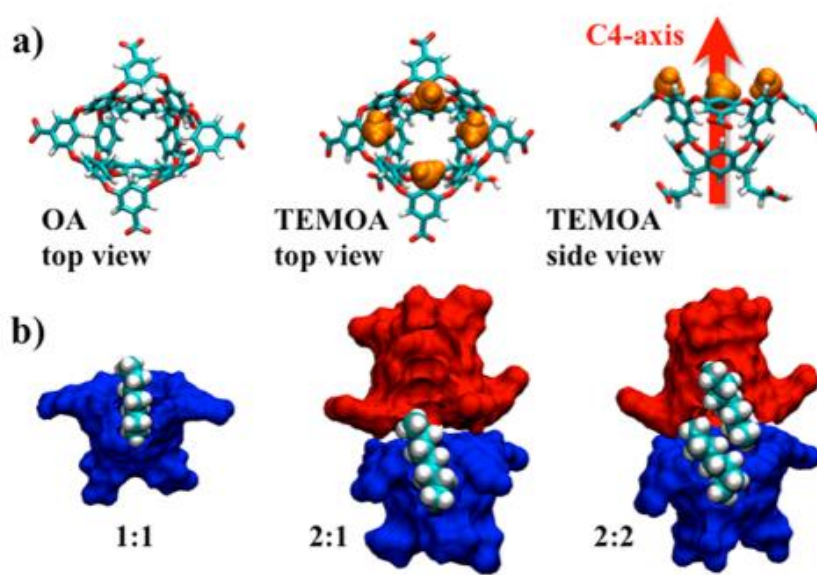
Figure 3.1: Illustrations of the host and host/guest complex structures. (a) Top views of OA and TEMOA hosts and side views of TEMOA. The main body of cavitand is shown as a wire frame view. The four methyl groups on rim of TEMOA are highlighted as orange van der Waals surfaces. The C_4-axis of symmetry is also pictured in the TEMOA side view. (b) 1:1, 2:1, and 2:2 complexes between OA and C_8. The cavitand host is illustrated as red and blue surfaces. The C_8 guests are illustrated by the van der Waals surfaces. The side of the complexes has been peeled back to show the guests within it.

In a recent computational study conducted in our group, Barnett, Gibb, and Ashbaugh[65] used a series of systematic molecular dynamics simulations to relate the encapsulated alkane packing motifs within confinement with the conformational strain inferred from guest packing free energies. They found flexible alkane guests will adopt a succession of conformational motifs, from extended, to helical, to hairpin, to spinning-top structures with an increase of guest length in OA. And the result of this simulation work is very consistent with the result observed in experiments. Simulations are also used to investigate the structural changes of alkanes and polyethylene glycol within carbon nanotubes[104]. These studies showed the results of simulation of supramolecular self-assembly systems are accurate and reliable. And they can be used to guide the design of cavitand/guest assemblies from the bottom-up.

Here we will present a systematic molecular simulation and thermodynamic analysis of the assembly of OA and TEMOA with alkanes to assemble into a range of monomeric and dimeric complexes in water. We consider n-alkane guests from methane (C_1) to hexadecane (C_{16}) encapsulated within OA and TEXMOA to form 1:1, 2:1, and 2:2 complexes. The association free energies of different complexes will be evaluated from the potential of mean force (PMFs) between complex constituents along designated coordinates. Then we develop a reaction network model from these PMFs and calculate the relative populations and distribution of different complexes as a function of guest size to find the distinct trend determined from experiments for OA and TEMOA. Finally, we will study the guest packing inside the complexes to find the thermodynamic source of the different assembly patterns.

3.3   Methods

We used GROMACS 5.1[79] package to perform the molecular dynamics simulations of n-alkanes captured within OA and TEMOA pockets. The alkanes were modeled using the L-OPLS all-atom force field[105], which is very good at reproducing the conformational and thermodynamic properties of long alkanes. The alkane guests simulated are from methane (C_1) to hexadecane (C_{16}). Generalized Amber Force Field (GAFF)[70] was used to model OA and TEMOA. The partial charges were obtained from the AM1-BCC[106] calculations. The network charge of each cavitand was set to be -6e to match the protonation state at pH 7[107]. So, we deprotonated the four benzoic acid groups around the rim of cavitand and two of the four benzoic acid groups on the feet. To neutralize this negative charge during simulation, we include six sodium cations modeled by GAFF[70] per host molecule. TIP4P-

EW[108] potential was used to model water molecules. Nonbonded Lennard-Jones interactions were truncated beyond a separation of 9 Å with a mean-field dispersion correction for longer-range contributions to the energy and pressure. Particle mesh Ewald Summation method[71] with a real space cutoff of 9 Å were applied to calculate the electrostatic interactions. All the simulations were conducted in isothermal-isobaric ensemble at 298.15 K and 1bar. The temperature and pressure were coupled by Nosé-Hoover thermostat[80,81] and Parrinello−Rahman barostat[82] respectively. Bonds involving hydrogens for the hosts and guests were constrained using the LINCS algorithm[86], while water was held rigid using SETTLE[87]. The equations of motion were integrated using a time step of 2 fs.

Potential of mean force quantifies the association free energy between constituents of cavitand supramolecular complexes along a designated reaction coordinate. Here, it is colinear with the host's 4-fold (C_4) rotational axis of symmetry. Thus, they can be used to characterize the stability of the different alkane/cavitand complexes. Three distinct series of PMFs were taken into consideration here: the interaction between a cavitand and a single alkane guest to form a 1:1 complex; the interaction between a 1:1 complex and an empty cavitand host to form a 2:1 complex; and the interaction between two 1:1 complexes to form a 2:2 complex. In the first set of simulations, the PMF was determined between a cavitand and a single alkane (C_1 to C_{16}) from bulk water into the hydrophobic pocket along the C_4 axis. The cavitand and the guest were solvated by 2600 water molecules in a cubic simulation box. Two virtual sites were created on the cavitand molecule. The first one was determined by the average position of the four carbon atoms connecting the four feet of the cavitand to the bottom row of aromatic rings. And the second virtual site was determined

by the average positions of the four carbon atoms at the top of the second row of aromatic rings. To align the cavitand to the designated reaction coordinate, we restrained the virtual site at the bottom to the z-axis of the simulation box with a harmonic force constant of 100000 kJ/ (mol nm^2). There was also a vector connecting the bottom virtual site to the top virtual site. We restrained this vector along z-axis with a harmonic angular constraint of 50000 kJ/ (mol). So, the C$_4$ axis of cavitand was parallel with z-axis. The PMF was evaluated from a series of overlapping windows spanning from bulk water into the hydrophobic pocket using umbrella sampling. The guest center was also restrained to z-axis using another harmonic force constant of 100000 kJ/ (mol nm^2). The center of guest was defined as following: for the guests with odd number of carbon atoms, the middle carbon (carbon number $(n+1)/2$) along the carbon chain was regarded as the center of guest; for the guests with even number of carbon atoms, a virtual site was put between the $n/2$ and $n/2+1$ carbons serving as the guest center. Windows were simulated from 5 Å deep inside the host pocket, which was measured from the center of the top plane defined by the same four carbon atoms used to determine the virtual site at top, to 15 Å out into bulk water. Forty windows lied along the z-axis of simulation box with a harmonic umbrella potential minimum separated in 0.5 Å increments and a force constant of 15000 kJ/ (mol nm^2). The simulation in each window was equilibrated by 1 ns, and then was run by 15 ns to get the trajectory and thermodynamic data. Configurations were saved every 0.2 ps for analysis. The PMF were reconstructed using the weighted histogram analysis method (WHAM)[109].

For the case of 2:1 complexes, we determined the PMF between two cavitand hosts with a guest molecule (C$_1$ to C$_{16}$) sitting inside it. The two cavitands were facing each other and their C$_4$ axis were restrained by the same setup used in the case of 1:1 complexes in mutual

alignment with each other. The guest was placed in one of these two cavitands without any restraint. During the simulation, it will stay inside by hydrophobic interactions. Here, we controlled the distance between the two virtual sites on top of the two cavitand hosts. The overlapping windows were simulated from where this distance is zero to the place where the distance is 13 Å. Twenty-five windows were used along z-axis with a harmonic umbrella potential minimum separated in 0.5 Å increments and a force constant of 15000 kJ/ (mol nm^2).2600-3000 water molecules were used here to achieve enough solvation. All the other setup was the same as what was used in the case of 1:1 complexes. And we also did the PMF between two empty hosts without any guest molecule as a basis.

Lastly, the PMF between two cavitand hosts with two guests was also determined. The setup here was very similar with that in 2:1 complexes. The only main difference is now there is one alkane guest in each of cavitand hosts. The range of guests we studied also varied. In the case of OA, guests considered were from C_1 to C_{11}. Longer guests will have repulsive interactions within cavitand dimer, which was large enough to lead the destabilization of capsule and made GROMACS package keep complaining about it. In the case of TEMOA, this range, which is from C_1 to C_9, was even narrower. Also, no additional restraints were applied to the guests. All the other setup was the same as what was used in the case of 2:1 complexes.
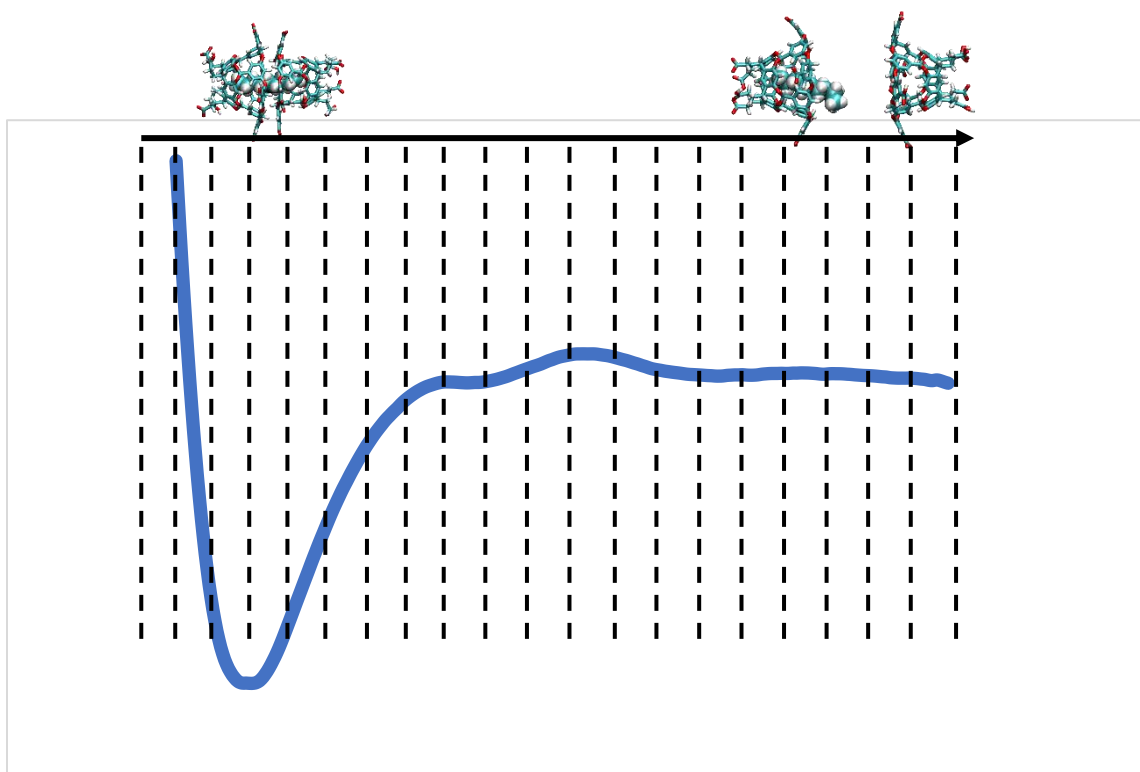
Figure 3.2: Schematic picture for umbrella sampling. Overlapping windows are along C_4-axis of cavitand. And there is a harmonic biasing potential in each window to control the distance between two cavitand hosts.

3.4 Results and Discussion

All the PMFs of different complexes were evaluated. Here, as shown in figure 3.3, the three PMFs between nonane (C_9) and OA are illustrated. The PMFs for the whole range of guests can be found in appendix.
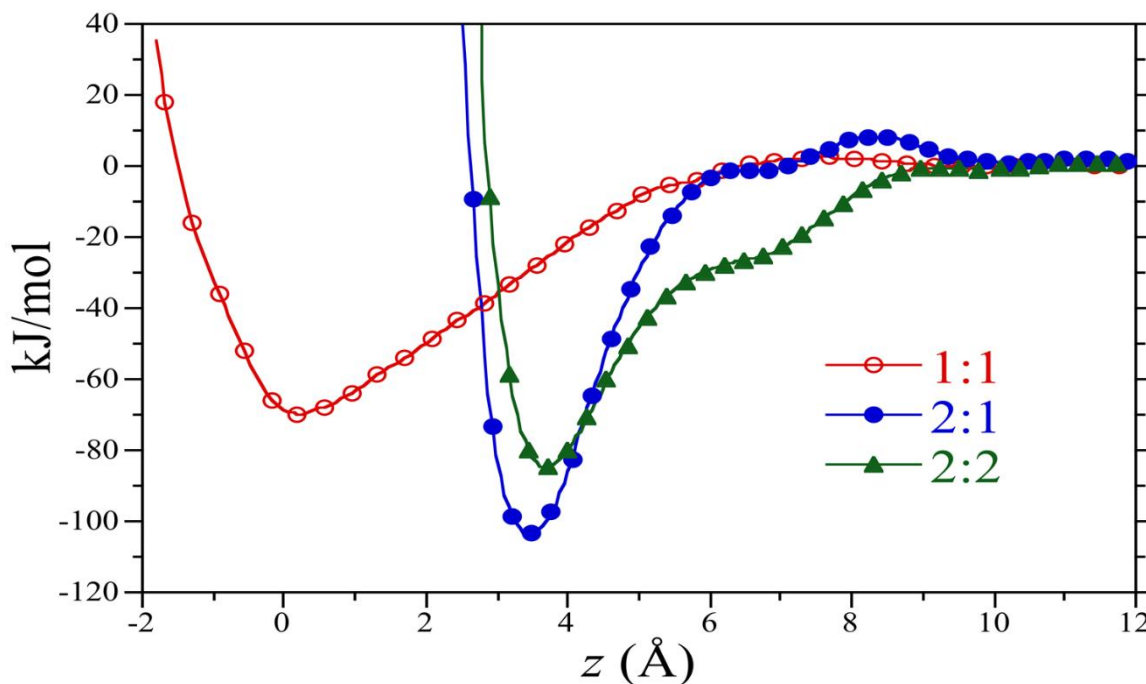
Figure 3.3: Potentials of mean force for association of 1:1, 2:1, and 2:2 complexes between nonane and OA. $Z = 0$ Å means the position of the face of the cavitand, which is on the left-hand side and facing right. Symbols for the 1:1, 2:1, and 2:2 interactions are defined in the figure legend. We omitted the error bars for clarity.

In general, the interaction between C_9 and OA is very attractive. For 1:1 complex, the minimum of the PMF is at $z = 0$ Å, which means alkane guest prefers to stay inside the hydrophobic pocket. This attractive interaction extends out into bulk water for ~5 Å due to the length of C_9 chain. After that, it is too far away for them to have notable interactions so that PMF becomes flat. The attractive well depth is -70 kJ/mol, which is much lower than the unfavorable solvation free energy of C_9 in water (17 kJ/mol). And it can be concluded that 1:1 complex is stable in water.

The attractive well depth of 2:1 complex PMF is much deeper than that of 1:1 PMF, and this more attractive association free energy has many contributions. First, more attractive interactions are gained between the guest and the second cavitand. So, the portion of guest chain that was exposed to bulk water in 1:1 complex is now also buried in hydrophobic environment. Secondly, the aromatic rings forming the rim of each cavitand, which are also hydrophobic groups, are also having favorable interactions with each other. The PMF minimum equilibrium position lies at $z = 3.5$ Å, which is approximately the diameter of one carbon atom.

In the case of 2:2 complex, the attractive well depth lies between that of the 1:1 and 2:1 PMFs. As we can think of, under nanoscale confinement, the inner space of an OA dimer hydrophobic pocket is limited. So, the packing of 2 C_9 chains within it is far more frustrated and unfavorable than packing of only one C_9 chain. Therefore, though more attractive interactions are gained between hosts and that additional alkane chain, The PMF minimum equilibrium position lies ~0.4 Å further to the right than that of 2:1 complex. This is another reflection of the frustration caused by guest packing and the capsule is less stable. Besides, 2:2 complex interaction has longer range, which is up to ~9 Å into bulk water. This is since the ends of the 2 C9 chains will extend out and wiggle in bulk water. This interaction extension comes from the interaction between the two ends.

Furthermore, PMFs' attractive well depth is able to provide useful information about the association of different complexes. So, we use the minima of attractive well depth to quantify the relative stability of the cavitand/C_n complexes. A deeper or more negative well depth minimum indicates a more stable complex. And the formation free energy minima

of the 1:1 ($\omega_{1:1}$), 2:1 ($\omega_{2:1}$), and 2:2 ($\omega_{2:2}$) complexes as a function of guest size are plotted below in figure 3.4.
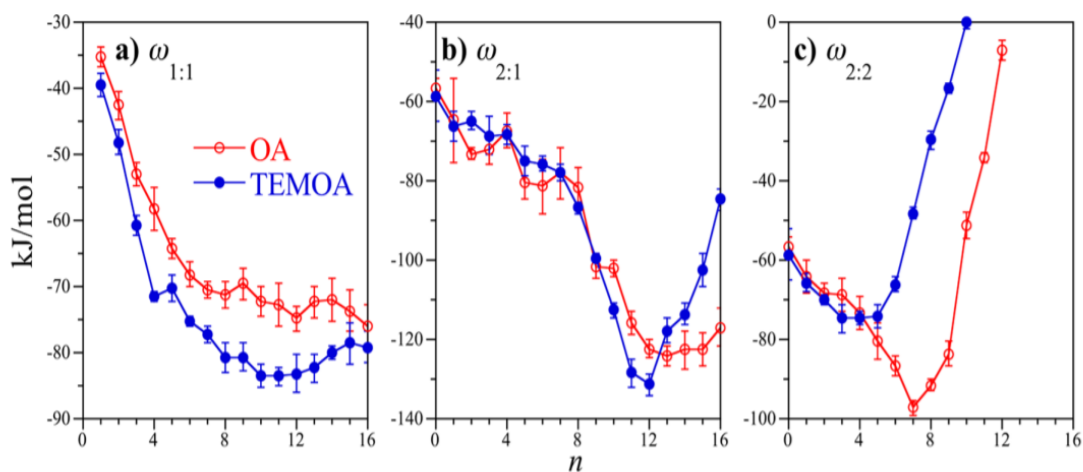
Figure 3.4: Potential of mean force minima determined for cavitand/alkane complex as a function of guest size. Part a, b, and c reports minima for 1:1, 2:1, and 2:2 complexes respectively. Symbols for OA and TEMOA are defined in legend. Error bars show one standard deviation.

In the case of 1:1 complex, both for OA and TEMOA, the PMF minima is initially a sharply decreasing function with guest size up to C_6. Beyond that, it becomes relatively flat. As we know, the depth of the cavitand hydrophobic pocket is approximately the same as the length of C_6. When the guest chain is getting longer and filling up the inner space of the pocket, it is gaining more and more attractive van der Waals interactions with the host. That is why the PMF minima has such a deep decrease firstly. When the guest chain length exceeds the pocket depth and begins to extend out into bulk water, it is not gaining van der Waals interactions anymore so that it begins to have less contribution to the total association free energy. Generally, the 1:1 association free energy of TEMOA is slightly more attractive than that of OA. We can ascribe it to the additional van der Waals interaction between guests and the methyl groups on the rim of TEMOA.

The 2:1 complex PMF minima is more informative. And you may notice that the formation free energy of an empty 2:0 cavitand complex, which is ~-58 kJ/(mol), is already favorable. Beginning from that point, for both OA and TEMOA, the PMF minima slowly decreases as a function of guest size up to C_8. And it is very hard to distinguish the PMF minima of OA and TEMOA within statistical uncertainty up to this point. Starting from C_9, the PMF minima for both cavitands take a big sharp turn and become remarkably attractive. This is consistent with the experimental results[63,67] for both cavitand where 2:1 complexes become stable from C_9. After that, the behavior of OA and TEMOA 2:1 PMF minima starts to be distinct from each other. In the case of OA, it reaches and exhibits a broad minimum for alkanes from C_{12} to C_{15}. And the 2:1 formation free energy for OA and C_{16} has a slight increase. From C_9, the guest carbon chain has spanned and filled up the interior space of OA dimer. Therefore, they acquire enough attractive van der Waals interactions with both hosts to deep the attractive well depth and make the complex very stable. And according to the previous research in our group, the end-to-end length of a OA dimer inner pocket is comparable to the length of a C_{11} chain[65]. And up to C11, the guest is still adopting a fully-extended packing motif where most of the dihedral angles along the carbon chain exhibit a trans morphology. Starting from C_{12}, the alkane chain length exceeds the end-to-end length of OA dimer. The end methyl groups begin to collide with the bottom of the both opposing cavitands' interior pockets. It will force the bounded guest to adopt a more compact packing motif like a helical motif with more gauche dihedrals to accommodate themselves inside OA dimer. And this broad minimum from C_{12} to C_{15} can be ascribed to a competition between the more favorable van der Waals interactions acquired between host and guest and the unfavorable guest conformations the guests adopt. In previous

research, the C17 will adopt a hairpin-like packing motif in OA dimer, which is more unfavorable. That can explain the little uptick for 2:1 C_{16} PMF minima.

In the case of TEMOA, after the dramatic decrease from C_9, the 2:1 PMF minima reached a very narrow minimum at C_{12}. After that, it exhibits a quick increase and becomes very repulsive. As a result of the four endo-methyl groups on rim of TEMOA, it will significantly restrict the packing motifs available of the guests, especially around the equatorial region of the dimer. In other word, TEMOA dimers will only favor fully extended packing motifs. This inability of TEMOA to pack longer alkane guests is reflected in the remarkable increase of PMF minima beyond C_{12}. It will finally lead to the destabilization of TEMOA dimer. Actually, in experiments, larger 4:2 tetrameric TEMOA complexes with guests longer than tetradecane were observed[69]. That is because longer guests need larger inner space to accommodate themselves.

The 2:2 complex PMF minima also initially weakly decreases with the encapsulated guest size. From no guest to C_4, the PMF minima shows no difference between OA and TEMOA. In the case of OA, the association free energy minima reach a minimum at C_7, and afterwards increases up drastically and becomes remarkable repulsive. This result is comparably consistent with the 2:1 PMF result since the length of one C_{14} chain is approximately the same as the total length of two C_7 chains. In difference to OA, the 2:2 TEMOA PMF minima reach a broad minimum very early from C_3 to C_5. After that, it also goes up very quickly for C_6 and longer guests. This guest length is consistent with the guest length where the transition from 2:2 to 1:1 occurred in the experiment[67]. In total, the comparison between the 2:2 complexes formation free energy of OA and TEMOA indicates that the 2:2 complexes are stable in a wider range of guests for OA than for

TEMOA. It can be interpreted that this difference is a reasonable explanation for the distinct assembly patterns of those two cavitands.

In order to better understand the equilibrium between hosts and guests to form into different complexes, we developed a reaction network model to describe the whole self-assemble process. Four reaction functions are used below to illustrate the competitive equilibrium cavitand hosts and alkane guests. G means alkane guests, and distinct complexes are denoted using 1:1, 2:1, and 2:2 respectively. We also consider the formation of empty dimer complexes for completeness. It is denoted by 2:0. And 1:0 is for empty cavitand.

$$1:0 + G \leftrightarrow 1:1 \tag{3.1}$$

$$1:0 + 1:0 \leftrightarrow 2:0 \tag{3.2}$$

$$1:1 + 1:0 \leftrightarrow 2:1 \tag{3.3}$$

$$1:1 + 1:1 \leftrightarrow 2:2 \tag{3.4}$$

From the four reactions above, the corresponding equilibrium reaction constants can be derived as below.

$$\frac{[1:1]}{[1:0][G]} = K_{1:1} \tag{3.5}$$

$$\frac{[2:0]}{[1:0]^2} = K_{2:0} \tag{3.6}$$

$$\frac{[2:1]}{[1:1][1:0]} = K_{2:1} \tag{3.7}$$

$$\frac{[2\!:\!2]}{[1\!:\!1]^2} = K_{2:2} \tag{3.8}$$

All the equilibrium constants can actually be evaluated from the multidimensional integrals over all separations and orientations of all the PMF between cavitand hosts and alkane guests[110–112]. But the PMF evaluated in our simulation, as described in methods session, is only along one designated reaction coordinate. And the orientation of the cavitand was also fixed during the simulation. So, it misses a large piece of information in the total integration domain. However, it can still be related to the Boltzmann weighting of the PMF minima. In a word, the equilibrium constants are proportional to the Boltzmann weighting of the PMF minima as below.

$$K_{1:1} = \alpha \exp(-\frac{\omega_{1:1}}{RT}) \tag{3.9}$$

$$K_{2:x} = \beta \exp(-\frac{\omega_{2:x}}{RT}) \tag{3.10}$$

Here, $K_{1:1}$ is the equilibrium constant for the formation of 1:1 complex. Correspondingly, $K_{2:x}$ is for the formation of dimeric complexes (x = 0, 1, or 2 for distinct complexes). We introduced two pre-factors α, and β for monomeric and dimeric complexes respectively. They are adjustable parameters and can be used to make up all the missing information in the potential of mean force calculation. That missing information includes that from all the noncomplex forming reaction paths and those from other host orientationally degrees of freedom. We also assume the values of α and β are the same for both OA and TEMOA. Besides, β is independent of the number of bounded guests. So, the biasing introduced from these two parameters can be avoided. Here is what we are going to do. Firstly, we will

adjust the values of α and β to reproduce the experimental assembly behavior of OA with alkanes, which displays a very straightforward monotonic transition from monomeric to dimeric complexes with an increase of guest size, and get reasonable values for α and β. Afterwards, by using the same values of α and β, the assembly behavior of TEMOA complexes with alkane guests can be predicted. And we will compare our prediction with the experimental results.

From the mass balance of cavitand hosts in bulk solvent, the total concentration of cavitand is listed as below.

$$[1]_{total} = [1{:}0] + [1{:}1] + 2([2{:}0] + [2{:}1] + [2{:}2]) \tag{3.11}$$

The total concentration of cavitand, $[1]_{total}$ is assumed to 3 mM, which is used in a very typical experiment. The bulk alkane guest concentration in water was determined from the following relation[113,114].

$$[G] = \frac{P_G}{RT}\exp(-\frac{\mu_G^{ex}}{RT}) \tag{3.12}$$

P_G is the partial pressure of alkane guest in gas phase. μ_G^{ex} is the excess chemical potential of the guest in water (i.e. The solvation free energy of alkane guest). RT is the product of temperature and gas constant. For those guests (from C_1 to C_4) whose boiling points are below 25 °C, the guest gas pressure was assumed to be 1 atm. For other guests whose boiling points are above 25 °C, the guest partial pressures were assumed to adopt the vapor pressure of the pure alkane liquids at 25 °C. Thus, the partial pressures at 25 °C can be determined from Wagner correlation fits to experiment[115]. The excess chemical potentials at infinite dilution in water were evaluated from standard free energy perturbation

techniques. We applied this method to L-OPLS alkanes in TIP4P/EW water. All the simulation details, guest partial pressures, excess chemical potential obtained from simulation, and final guest concentrations are included in appendix.

All the concentrations of different complexes can be expressed by an equation of corresponding equilibrium constants and concentrations of guest and empty cavitand. After we substitute them into the mass balance equation of cavitand in water, the whole reaction network model will be in a form of quadratic equation in concentration of empty cavitand.

$$2(K_{2:0} + K_{2:1}K_{1:1}[G] + K_{2:1}K_{1:1}^2[G]^2)[1:0]^2$$

$$+(1 + K_{1:1}[G])[1:0] - [1]_{total} = 0 \tag{3.13}$$

When this equation is solved, we can get the concentration of free guest in solvent. From that, the concentration of all the complexes can be evaluated. Therefore, the distribution of all assembly states for all guests can be determined.

The finally fitted values of α and β were 2×10^{-4} and 8×10^{-11} M^{-1}. The distribution of 1:0, 1:1, 2:1, and 2:2 complexes between cavitand (OA and TEMOA) and alkane guests are reported below in figure 3.5.
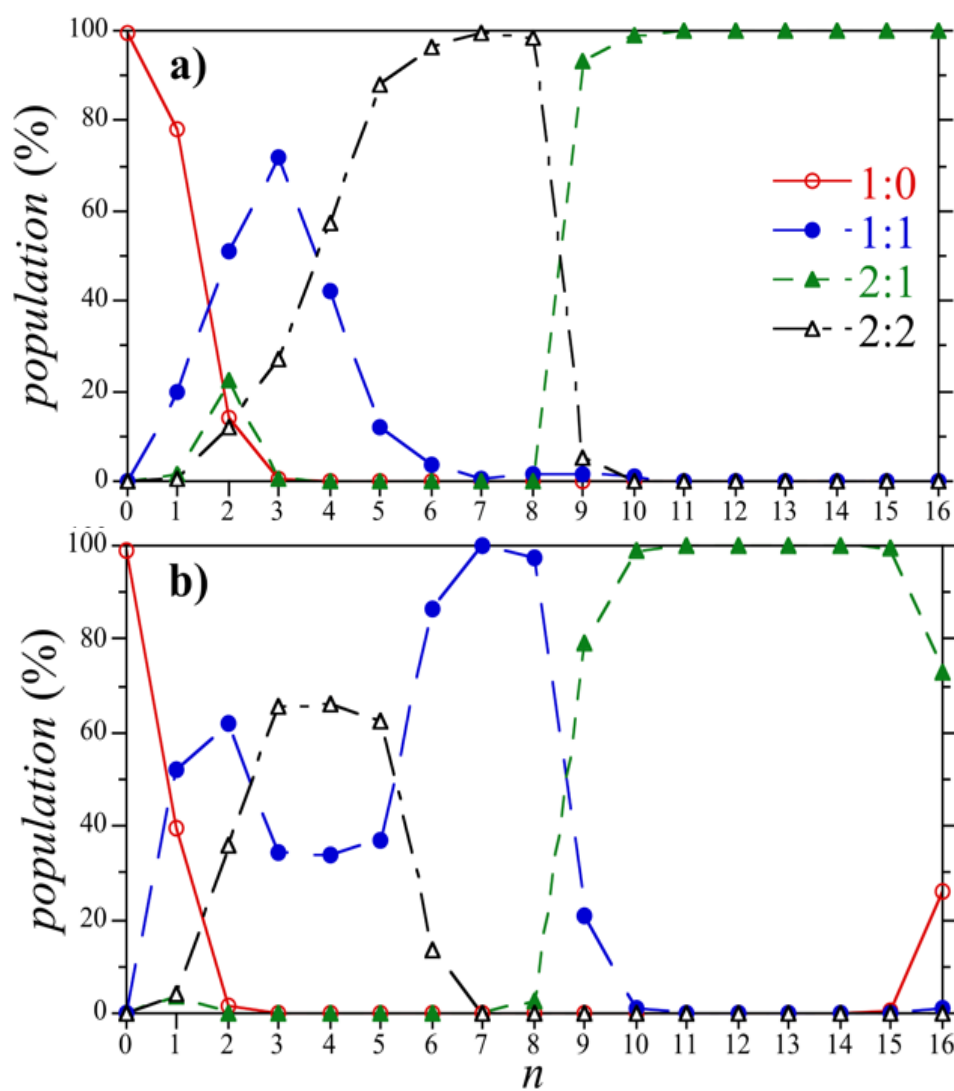
Figure 3.5: (a) Distribution of OA/alkane assembly states between 1:0, 1:1, 2:1, 2:2 complexes as a function of guest size. (b) Predicted distribution of TEMOA/alkane assembly states between 1:0, 1:1, 2:1, 2:2 complexes as a function of guest size. Symbols for the 1:0, 1:1, 2:1, 2:2 complexes are defined by figure legend. The population of 2:0 complexes was below 0.1% so that they are not reported. Error bars are not reported for clarity.

The progression of dominant OA complexes as a function of guest size is: empty single OA for no guest and C_1; 1:1 complexes from C_2 and C_3; 2:2 complexes for guests from C_4 to C_8; and 2:1 complexes for C_9 and longer guests. This result is fundamentally consistent with the experiments[63,67]. The transition between 2:2 complexes and 2:1 complexes also lies between C_8 and C_9. The population of 2:0 complexes is negligible for the whole range of guest size. Therefore, it is not reported in figure 3.5. The only exception is C3 mainly formed 2:2 complex in experiments.

The progression of dominant TEMOA complexes as a function of guest size is: empty single TEMOA for no guest; 1:1 complexes for C_1 and C_2; 2:2 complexes for guests from C_3 to C_5; then 1:1 complexes reappear for guests from C_6 to C_8; and finally, 2:1 complexes are dominating for C_9 and longer guests. Thus, an obvious non-monotonic trend is determined from our simulation. 1:1 complexes are dominating again for medium-sized alkane guests. Our simulation also captured some other notable details in experiments. In experiments, C_1 formed 1:1 complex with TEMOA. But with OA, this binding is much weaker, and it didn't form 1:1 complex. It can be ascribed to the additional attractive van der Waals interactions between C_1 and the four endo-methyl substituents of TEMOA. Our simulation is in agreement with the experiments. 1:1 TEMOA complex is dominant for C_1, while 1:0 empty OA is dominant for C_1. The 1:1 complexes PMF minima also exhibits the more enhanced interaction between C_1 and TEMOA. Another interesting observation is the destabilization of 2:1 TEMOA complex with C_{16}. The population of 2:1 TEMOA complex decreases from nearly 100% to ~71% at C_{16}. In experiments, OA will assemble into 2:1 complexes with a wide range of guests from C_9 to C_{25}[64]. However, for TEMOA, they will prefer larger assembly states like 4:2 tetrameric and 6:3 hexameric assemblies with guests

longer than C_{14}[69]. This destabilization of TEMOA dimer can be attributed to the increasingly repulsive PMF minima for 2:1 complex formation free energy beyond guests longer than C_{12}. This decrease of TEMOA 2:1 complex population is accompanied by the increase of TEMOA 1:0 complex to ~26%. But the TEMOA 1:1 complex only increases to ~1%. The preferential growth of empty TEMOA population results from the diminishing solubility of alkane with increasing length. According to eq. (3.12), the solubility of C_{16} is on the order of 2pM. Therefore, the reaction network model will predict a huge increase of 1:0 complex population. Actually, this increase begins from around C_{12} or C_{13}. But the population of 2:1 complex is too significant (~100%) to make this increase notable.

We can evaluate the average aggregation number of OA and TEMOA, which is the average number of complexed hosts averaged over all assembly states, as a function of guest size directly from the cavitand/alkane assembly states distribution. And it is reported below in figure 3.6.
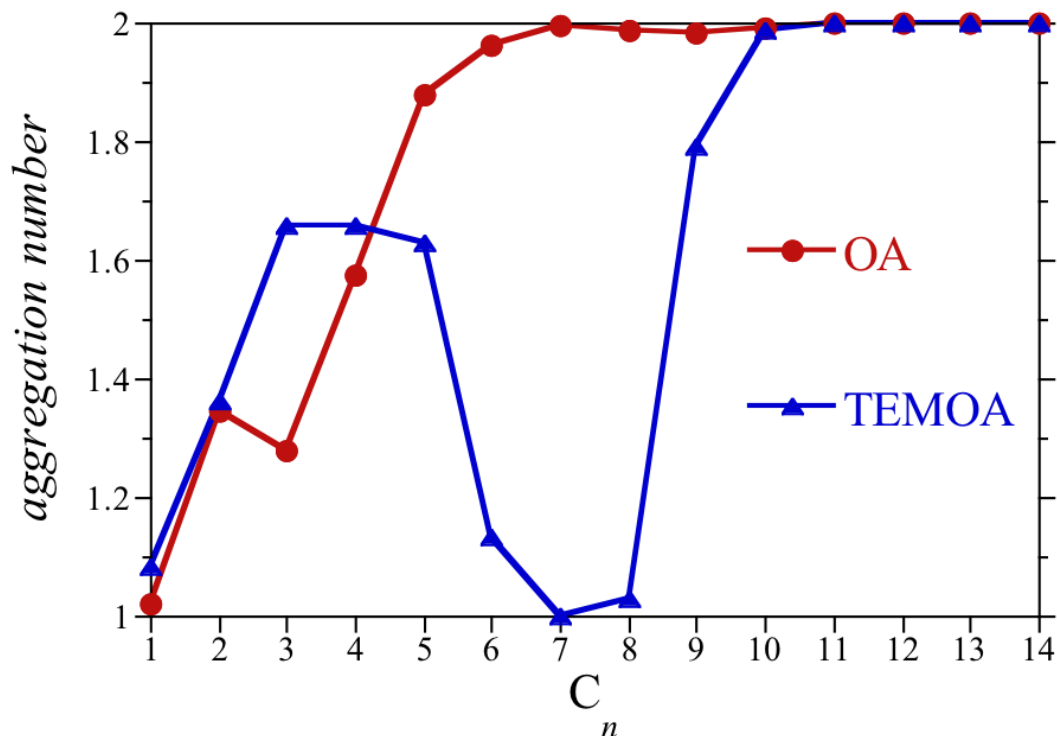
Figure 3.6: Average cavitand aggregation number in cavitand/alkane complexes as a function of the alkane guest size. Results for OA and TEMOA are defined by the figure legend. Error bars are omitted for clarity.

In the case of OA, the average aggregation number gradually increases from one to two over the range of guests from C_1 to C_6. After that, it stays at two and never drop down. So, clearly the average aggregation number increases monotonically. There is a hump at C2, which is potentially an artifact or due to the statistical uncertainty since its size is smaller than the error bars from C_1 to C_3. In contrast to OA, the aggregation number of TEMOA is non-monotonic. It initially increases from 1 to ~1.6 from C_1 to C_5. Then it drops back to 1 for C_6 through C_8. Then TEMOA reassemble into dimers for C9 and longer guests. All these predictions are excellently consistent with the variation of cavitand/alkane assembly behaviors and complex volumes with change of alkane chain length exhibited in experiments.
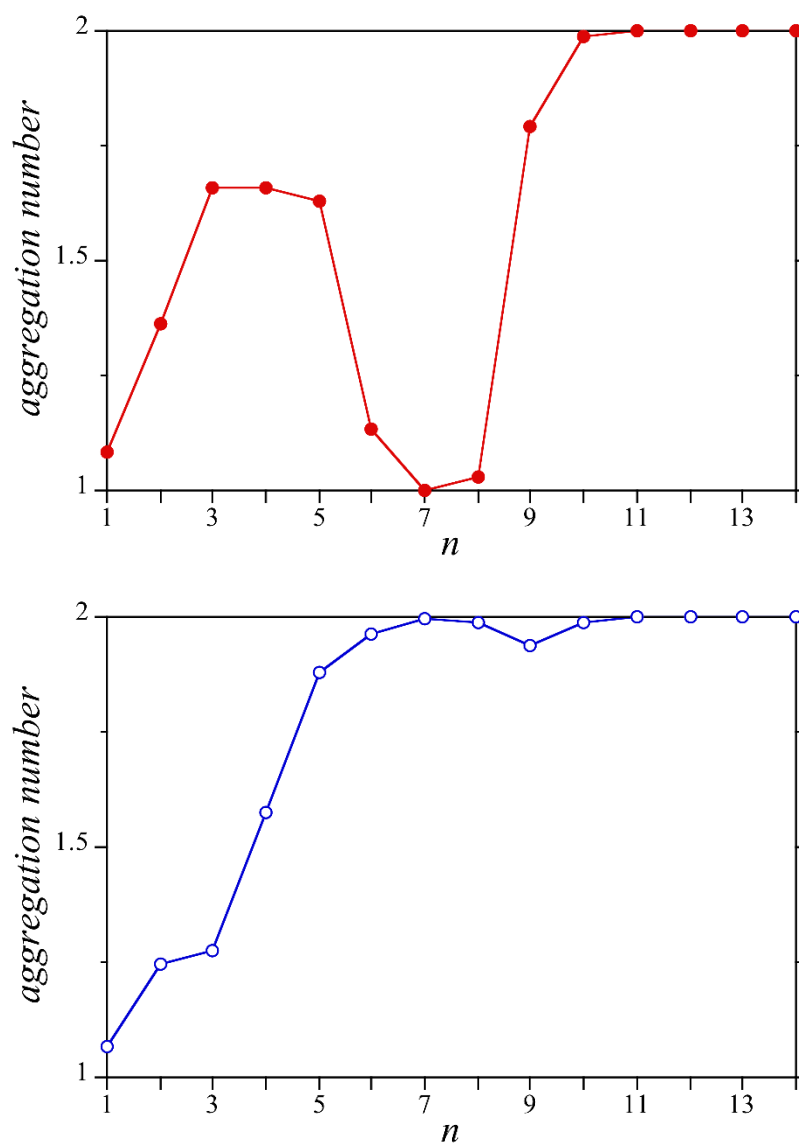
Figure 3.7: Predicted OA average aggregation numbers as a function of guest size using the $\omega_{2:2}$ results of TEMOA are reported on top. Predicted TEMOA average aggregation numbers as a function of guest size using the $\omega_{2:2}$ results of OA are reported at bottom. Error bars are omitted for clarity.

Remember that the 2:2 complexes association free energy of OA and TEMOA exhibited distinct trends. It reached the minimum very early for TEMOA with shorter guests. Therefore, it can be hypothesized that the differences in 2:2 association free energy actually determines the totally different assembly behaviors of OA and TEMOA. We made several calculations to test this hypothesis. In the figure 3.7 shown above, we firstly substituted the

$\omega_{2:2}$ results of TEMOA into the reaction network model of OA. And thereafter OA also exhibits a monotonic assembly behavior trend. Conversely, we also substitute the $\omega_{2:2}$ results of OA into the reaction network model of TEMOA. In this case, the nonmonotonic aggregation behavior is suppressed, where it monotonically transits between monomeric to dimeric complexes. These results demonstrated that this non-monotonic assembly behavior of TEMOA is controlled by the formation of 2:2 complexes.
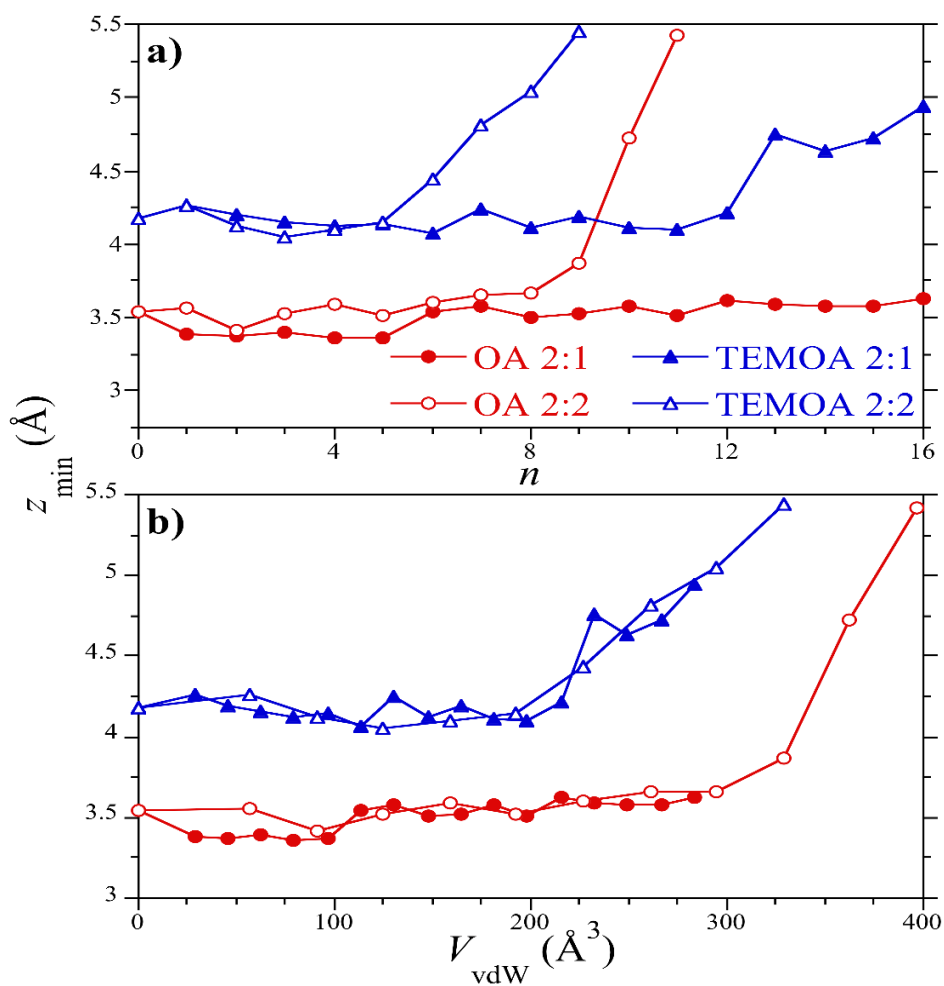
Figure 3.8: (a) Equilibrium separation distance between cavitand hosts of 2:1 and 2:2 complexes at where PMF minimum occurred as a function of guest size. (b) Equilibrium separation distance between cavitand hosts of 2:1 and 2:2 complexes at where PMF minimum occurred as a function of total encapsulated guest van der Waals volume. Symbols are defined in the legend.

In figure 3.8(a), the inter-cavitand separations of 2:1 and 2:2 complexes at where the PMF minima occurred are plotted as a function of guest size. For OA 2:1 complex, the equilibrium separation is ~3.5 Å and independent of guest size up to C_{16}. That means for the whole range of guests, the 2:1 complexes are stable. On the other hand, for TEMOA 2:1 complex, the equilibrium separation is initially flat at ~4.2 Å up to C_{12}. Beyond that, it increases with the guest chain length. In comparison to OA, the initial separation distance is slightly larger because of the protrusion of the endo-methyl substituents of TEMOA above the cavitand rim. For longer guests, the frustration of guest packing will lead to a destabilization of TEMOA dimer so that the separation distance will begin to increase. This is also reflected in the corresponding increase of $\omega_{2:1}$ of TEMOA for guests longer than C_{12}. The reason is guests encapsulated in TEMOA dimer can only adopt predominantly extended conformations. But within OA, a succession of packing motifs, include some more compact packing motifs, can be adopted for longer guests. The limited flexibility of the guest within TEMOA dimer will have very negative contribution to the association free energy and finally make the dimer crack open.

For 2:2 OA complex, the inter-cavitand equilibrium separation is independent of guest chain length, also at ~3.5 Å, up to C8. With longer guests starting from C9, it systematically increases with guest size. This is coincident with the fact that $\omega_{2:1}$ of OA reached the minimum at C_7. For 2:2 TEMOA complex, the inter-cavitand equilibrium separation is also initially flat up to C_5, and then begin to rise. It is also consistent with the narrower stability window displayed in the formation free energy result of this complex.

In figure 3.8 (b), the equilibrium separation distance between cavitand hosts of 2:1 and 2:2 complexes are plotted as a function of total encapsulated guest van der Waals volume. The

results of 2:1 complexes and 2:2 complexes are in great agreement with each other. The 2:1 and 2:2 OA complexes equilibrium separation distances are the same up to a total encapsulated guest volume of ~300 Å^3, which is approximately the volume of one C_{16} chain or two C_8 chains. From that point, you may anticipate that for OA 2:1 complex, the separation distance will begin to increase from C_{18} since it started to increase for OA 2:2 complex from C_9. However, in experiments, OA still formed stable 2:1 complexes with C18. Both experiments[64] and previous simulation[65] in our group found that C_{18} adopted a hairpin-shaped packing motif with OA dimer, which was not observed in guests in 2:2 complexes.

For TEMOA 2:2 complexes, an excellent agreement between the 2:1 and 2:2 complexes' separation distances are exhibited for whole range of guests. They both begin to increase beyond a total encapsulated guest volume of 220 Å^3. This volume lies between the volume of one single C_{12}, which is where the 2:1 TEMOA PMFs minima function reaches the minimum, and the volume of two C_6 chains, which is the guest size where 2:2 TEMOA complexes firstly crack open. Beyond this volume, the equilibrium separations for the 2:1 and 2:2 TEMOA complexes also match with each other quantitatively. So, the alkane guest packing is a determinant of the equilibrium separations. And this equilibrium separation also depends largely on cavitand species. That indicates different cavitands with different internalized spaces dramatically influence guest packing.

Since the only difference between TEMOA and OA are the four methyl groups on the rim of TEMOA. The different guest packing conditions can be ascribed to those rim substitutes. More detailly, the narrowing of TEMOA's portal is significantly limiting the packing motifs available to the encapsulated guests, and therefore leads to different assembly

patterns. Here, we evaluate the lateral displacement of guest carbons at the seam between

cavitand dimers in 2:1 complexes to quantify the width of the cavitand portals. The result

is shown below in figure 3.9.
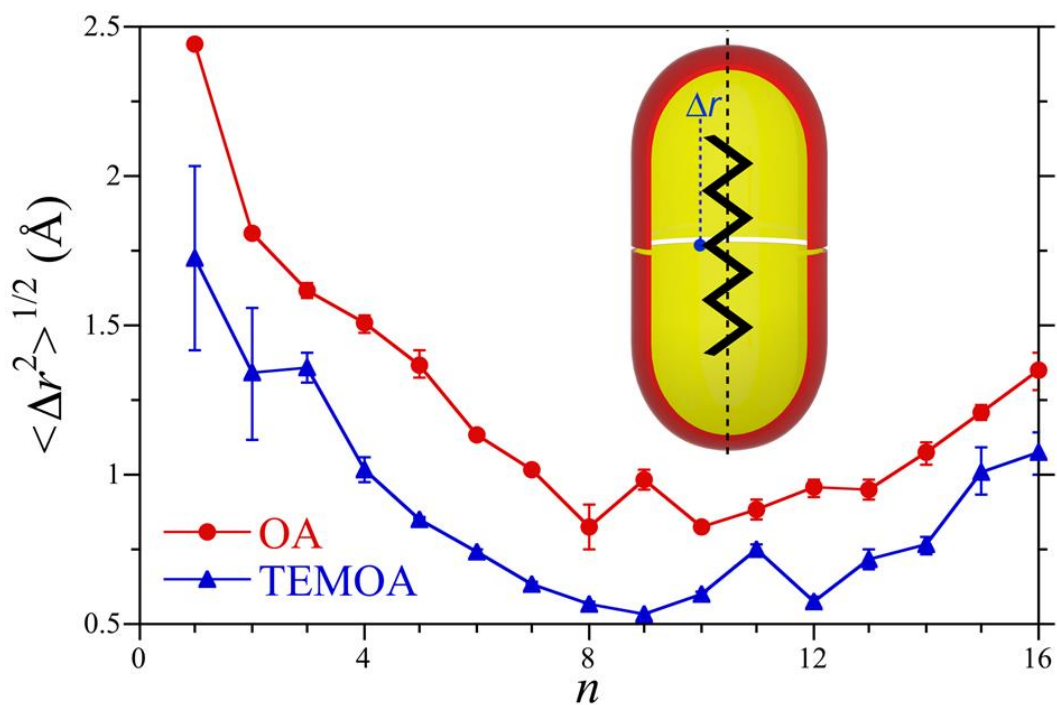
Figure 3.9: Root mean square lateral displacement from the dimer centerline for the guest carbon closet to the seam in 2:1 complexes as a function of guest size. Symbols are defined in the legend. Error bars indicate one standard deviation. The inset picture illustrates the lateral displacement, Δr. The centerline is defined by the line connecting the two virtual sites on the top plane of cavitand, which are the centers of both cavitand mouths.

Firstly, we identify the guest carbon closet to the midpoint of the cavitand dimer. Then the distance of this carbon atom from the centerline, which is connecting the centers of both cavitands' mouths, is measured. Lastly, we plot the root mean square displacements as a function of guest size. For the whole range of guests, this displacement is largely restricted in case of TEMOA compared to OA. For both OA and TEMOA, they have the largest lateral displacement with C_1, and this is because methane is small enough to explore both the dimer seam and the pocket depths. Beginning from C_1, the both displacements initially decrease with the guest size. This is clearly due to that small guests are free to explore the whole seam area, while portions of larger guests are forced to stay in the equatorial region. Interestingly, the displacement of both OA and TEMOA will pass through a minimum and increase again with guest size. For the case of OA, this increase starts at C_{11} because longer guests will adopt a helical packing motif instead of a trans-enriched extended motif. In the case of TEMOA, helical packing motifs were never observed. But due to the restriction around the cavitand portal, the increase of this displacement is actually a result of the destabilization of TEMOA dimer.

3.5 Conclusions

We conducted molecular dynamics simulation to explain the surprising nonmonotonic switching between monomeric and dimeric assembly states of deep-cavity cavitand TEMOA with n-alkane guests of increasing length. The simulations showed that while the association free energies of monomeric 1:1 complexes of OA and TEMOA are qualitatively similar, 2:1 and 2:2 complexes have significant different association free energy behaviors for OA and TEMOA. The four endo-methyl substituents on rim of

TEMOA largely narrow the portal down, and therefore destabilize 2:1 TEMOA complexes with C_{12} and longer guests. And the conformational freedom of encapsulated guests within TEMOA dimer are restricted. This restriction will also have very negative contributions to the formation free energy of 2:2 TEMOA complexes with short and middle-sized guests. From our reaction network model incorporated with all the association free energies, this calculation accurately captured and predicted OA's monotonic and TEMOA's non-monotonic assembly behaviors. Our simulation also indicated TEMOA 2:1 complexes begin to crack open for guests longer than C_{16}, which is consistent with the experiments where TEMOA formed larger multimeric complexes with those long guests.

This work demonstrated one typical example that small chemical modification on constituents of a supramolecular self-assembly system will dramatically affect its structures and properties. And molecular dynamics simulation can play an important role in decoding the impact of different chemistry structures and helping the bottom-up design of self-assembly system. Moreover, interpreting the guest packing is crucial in determining the assembly states of self-assembly system and therefore switching their properties.

Chapter4
The Progression of Assembly States as Function of Guest Sizes for a Newly Developed Deep-cavity Cavitand

4.1   Summary

Tetra-*exo*-methyl octa-acid (TEXMOA) is another newly developed derivative in the family of deep-cavity cavitand host molecules. It has four methyl units pointing upwards on the *exo* positions of hydrophobic rim. As shown in experiments, the addition of four methyl groups on *endo* positions of the rim will change OA's monotonic progression of assembly states as function of alkane chain length into TEMOA's nonmonotonic progression trend. In the last chapter, we presented a systematic computational study to show that this unusual nonmonotonic assembly progression can be attributed to the narrowed portal with four methyl groups. Therefore, the packing of two guest molecules within dimeric capsules became very frustrated and 2:2 assembly states were suppressed in the case of TEMOA. Here, the similar study and modeling was expanded to the assembly states progression of TEXMOA. Our simulation showed that a monotonic assembly states progression similar with that of OA was predicted. To some extent, the methylation of *exo* positions shifted the transition of assembly states towards longer guests. But the trend was not fundamentally changed.

4.2 Introduction

Self-assembly behaviors and the process of forming into complexes are playing a nonnegligible role in many bio-systems such as the chaperonin complex regulating protein folding[2], viral capsids stimulating DNA injections[5,6], and other protein-ligand binding to achieve many basic functions in living organisms[116,117]. To study, and further make use of those unique self-assembly behaviors, various mechanisms, such as metal coordination[12,13] and hydrogen bonding[14,15,22–24], were utilized and developed to generate well-defined supramolecular self-assembly systems. Hydrophobic effect was also identified as an important mechanism affecting the thermodynamics of biomolecular recognition[50,118–120]. And it is more easily to be utilized in biological applications mainly with aqueous environments due to its nontoxicity and its ability to distinguish polar and nonpolar groups.

On the other hand, the design of supramolecular self-assembly system demands refinement. A small modification on the structure or arrangement of this molecular system may lead to significant different properties and behaviors. For instance, cucurbit[n]urils (CB_n) are formed by a group of glycolurils connected by methylene bridges[26,27,121]. They exhibited significantly different binding affinities with different numbers of glycoluril units[122]. Computational work suggested that this is mutually determined by two counteracting effects. With increase of cavity size, the energy of individual water molecule decreases because they can form more hydrogen bonds with neighboring waters. But meanwhile more water molecules can be encapsulated in a larger cavity. As a result from these two factors, CB[7] stands out with the highest binding affinity because it obtains the maximum free energy gain by expelling all the waters out of the cavity[44]. Another example is the family of resorcinarene-based velcrands. In aqueous environment, they are presented in a

vase-like conformation by connecting the nitrogens of their benzimidazole groups with hydrogen-bonding of four water molecules. With this deep concave pocket, the original type of velcrand can bind with a wide range of suitably-sized guests including short tetraalkylammonium, N,N-Dimethylpyrrolidine, and adamantane[123]. Interestingly, a tetracarboxylate analogous form of the original velcrand exhibited a different binding preference. Cyclopentane is preferentially bound over cyclohexane in this analogous form velcrand. While for original type, the reverse is true. And the thermodynamics and kinetics behind the binding mechanism was modified by structural alternation on host molecules[124]. Another different form of velcrand derivative can form into complexes and bind with alkane guests. Short guest from heptane through undecane will get bound within the pocket and form a 1:1 complex[125]. While longer guests like tetradecane and heptadecane can trigger dimerization and formed 2:1 capsules with the hosts[126]. With a simple methylation of the benzamidazolones of this velcrand derivative, the assembly behaviors can be significantly adjusted. Utilizing the wider space provided by the methyl groups, host molecules can form 1:1 complex with longer guests from dodecane to tetradecane with a folded U-shaped motif. The methylene groups in the middle of the carbon chains resided the deepest inside the pockets[127].

To gain more understanding of the sophisticated relationship among supramolecular system structures, their assembly behaviors, and the driving mechanism behind it, a family of water-soluble deep-cavity cavitand host molecules were developed. They are formed by three rows of aromatic rings and there is a bowl-shaped hydrophobic pocket in the center. There are eight carboxylic acid groups on the feet and hydrophobic rim of cavitand to ensure water solubility[56]. The binding behaviors of octa-acid (OA, shown in Figure 4.1a)

and tetra-*endo*-methyl-octa-acid (TEMOA, shown in Figure 4.1b) were studied experimentally[63] and computationally, as described in chapter 3. Significantly different assembly behaviors were observed for these two host molecules, though the only difference between their structures is the four methyl units on *endo* positions of TEMOA[67]. In the case of OA, there is a straightforward monotonic relationship between assembly states and guest lengths. Methane doesn't bind with the hosts, and ethane forms a monomeric 1:1 complex (denoted *X:Y*, where *X* is number of cavitands, and *Y* is number of *n*-alkane guests). With the increase of guest size, OA can form into dimeric 2:2 complexes with guests from propane through octane, where octane actually forms a mixture of 2:2 and 2:1 complexes. From nonane, all the longer guests form 2:1 complexes. In the case of TEMOA, an unusual nonmonotonic progression of assembly states were observed. 1:1 complexes were formed with methane and ethane, while 2:2 complexes were formed with guests from propane to hexane. 1:1 complex reemerged with the presence of heptane and octane. Finally, nonane and longer guests will form 2:1 complexes. Our computational work, as described in chapter 3, demonstrated that the packing of two alkane guests within TEMOA is very frustrated due to the narrowed portal by those four methyl units, and therefore the formation of 2:2 complex actually controls the whole assembly behavior.

Since functionalization can lead to totally distinct assembly behaviors, the location of the functionalization potentially can be another important influencing factor. More importantly, we want to investigate the ability of our previous reaction network model to predict the assembly behavior of different cavitands before experiments. Here we performed a very similar simulation work on another derivative of OA, which is named as tetra-*exo*-methyl-octa-acid (TEXMOA, shown in Figure 4.1c). It also has four methyl groups on the rim, but

they are on *exo* positions instead of *endo* positions. Thus, the methyl groups are extending outwards, and the month is not narrowed. Association free energies of distinct TEXMOA complexes (1:1,2:1 and 2:2) were evaluated based on potential of mean force between different constitutes by molecular dynamics simulation. A reaction network model was built up based on these association free energies to predict the distribution of different assembly states. And this prediction was compared with the assembly state distribution of OA and TEMOA.
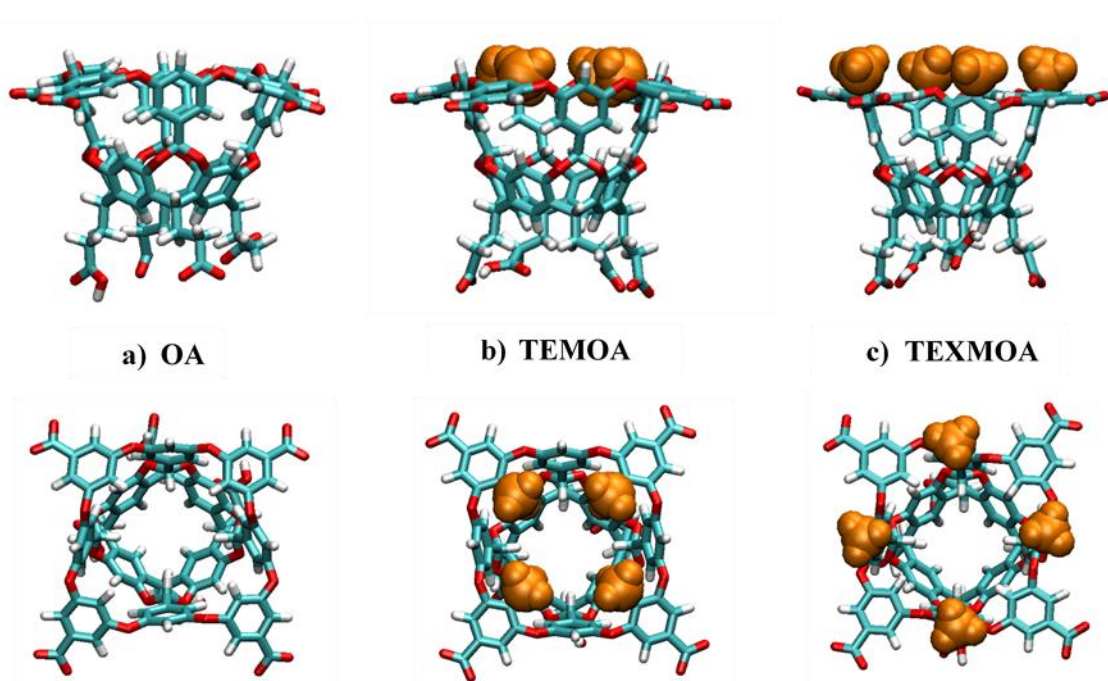
Figure 4.1: Top view and side view of cavitands involved in this study: a) structures of octa-acid (OA) b) structures of tetra-*endo*-methyl-octa-acid (TEMOA) c) structures of tetra-*exo*-methyl-octa-acid.

]

4.3 Methods

The simulation details were exactly the same as that described in chapter 3. All the molecular dynamics simulation of TEXMOA complexes with *n*-alkane guests in water were performed using GROMACS 5.1[79]. The cavitand was simulated using the

Generalized Amber Force Field (GAFF)[70] with partial charges obtained from AM1-BCC calculations[106]. Each cavitand has a network charge of -6 e to match the protonation state at pH 7[107]. The *n*-alkane molecules were modeled using the L-OPLS all-atom force field[105], which accurately captures the thermodynamic and conformational properties of long alkanes. Guests from methane (C_1) to hexadecane (C_{16}) were studied. In the simulation box, six sodium cations modeled by GAFF were included for each cavitand molecule to make the whole system neutral. Water molecules were modeled using TIP4P-EW potential[108]. Electrostatic interactions were treated using the particle mesh Ewald Summation method[71] with a cutoff of 9 Å in real space. Non-bonded Lennard-Jones interactions were truncated beyond a separation of 9 Å with a mean-field dispersion correction for longer range contributions to the energy and pressure. All the simulations were in an isothermal-isobaric ensemble at 298.15 K and 1 bar. The temperature and pressure were coupled by Nosé-Hoover thermostat[80,81] and Parrinello-Rahman barostat[82], respectively. All the bonds involving hydrogens in cavitands and alkane guests were constrained using the LINCS algorithm[86], while water was held rigid using the SETTLE algorithm[87]. The equations of motion were integrated using a time step of 2 fs.

The potentials-of-mean force (PMF) between hosts and guests were used to quantitatively characterize the relative stability of different complexes. PMF is part of the association free energy between constituents of the self-assembly complexes along with a reaction trajectory, in this case it lies along the host's four-fold (C_4) rotational axis of symmetry. In these simulations, the hosts and guests were solvated by 3000-5000 water molecules. Restraint potentials were applied to two dummy atoms along the C_4-axis of each host to perfectly align the cavitand along the z-axis of the simulation box. The first dummy atom

was determined by the central position of the four carbon atoms connecting the four feet of the cavitand to the first row of aromatic rings. The second dummy atom was determined by the average position of the four carbon atoms on the second row of aromatic rings lying closet to the cavitand portal. The "bottom" dummy atom was spatially restrained on the z-axis of simulation box with a harmonic force constant of 100000 kJ/ (mol nm^2). The vector connecting the two dummy atoms was also restrained to align with the z-axis using an angle restraint. The angle between this vector and the z-axis of simulation box was restrained to be zero degree with a harmonic angular constraint of 50000 kJ/(mol).

In general, three distinct PMFs were considered. The first one is the interaction between a single empty cavitand and one alkane guest to form a 1:1 complex. In this first set of simulations, we determined the PMF between one TEXMOA and one alkane (C_1 to C_{16}). A series of overlapping windows from bulk water into the host hydrophobic pocket were used. The guest center was also restrained to the C_4-axis of simulation box with a harmonic force constant of 100000 kJ/ (mol nm^2). The center of the alkane guest was taken as the middle carbon along the backbone for guests with an odd number of carbon atoms. For those with an even number of carbon atoms, the average position of the $n/2$ and $n/2+1$ carbon atoms was set as a dummy atom. This dummy atom was then restrained to the z-axis by the same potential constant. The simulations in overlapping windows were conducted from 5 Å inside the hydrophobic pocket to 12.5 Å out into bulk water. This distance was measured based on the center of the top plane of cavitand, i.e. the position of the second dummy atom. Totally there are thirty overlapping windows. The minimums of harmonic umbrella potentials were separated in 0.5 Å increments, and the force constant is 15000 kJ/ (mol nm^2). Simulation in each window was firstly equilibrated for 1 ns, followed

by a 15 ns production run. Configurations were saved every 0.2 ps for analysis, and in total there are 750 frames in each window. The PMF of the association of 1:1 complex was reconstructed by the weighted histogram analysis method (WHAM)[109].

The second one is the interaction between a 1:1 complex and another empty cavitand to form a 2:1 complex. The PMF was determined between two TEXMOA cavitands. One of the hosts was in a 1:1 complex with a single guest (C_1 to C_{16}) inside it. The two cavitands were facing each other with their C_4-axes restrained in mutual alignment with the z-axis of simulation box. This restraint potential was the same as that described in the setup of 1:1 complex simulation. The alkane guest was free inside one of the cavitands' pocket without any restraint. During the simulation, it will stay inside the pocket by hydrophobic interaction. Overlapping windows were applied from where the distance between the two top dummy atoms of cavitands is zero, to the state where this distance is 15.5 Å. Thirty-one windows in total were simulated here. Again, the minimums of harmonic umbrella potentials were separated in 0.5 Å increments, and the force constant is 15000 kJ/ (mol nm^2). All the other simulation details were same as that in the case of 1:1 complex.

Finally, we also consider the interaction between two 1:1 complexes to form a 2:2 complex. We determined the PMF between two TEXMOA cavitands, where both hosts were in a 1:1 complex with a single alkane guest (C_1 to C_{13}) inside it. As described above, the two cavitands were oriented to face one another. Their C_4-axes were restrained in mutual alignment with the z-axis of simulation box. Once again, no restraints were applied to the guests. The rest of the simulation setup was identical to that of 2:1 complex PMF simulation.

In chapter 3, a reaction network model was developed to predict the distribution of different assembly states with presence of guests with different lengths. We will reemphasize the important entities in this model. Four reactions were involved in the whole equilibrium assembly process between n-alkane guests and all the host/guest complexes.

$$1\!:\!0 + G \;\leftrightarrow\; 1\!:\!1 \tag{4.1}$$

$$1\!:\!0 + 1\!:\!0 \;\leftrightarrow\; 2\!:\!0 \tag{4.2}$$

$$1\!:\!1 + 1\!:\!0 \;\leftrightarrow\; 2\!:\!1 \tag{4.3}$$

$$1\!:\!1 + 1\!:\!1 \;\leftrightarrow\; 2\!:\!2 \tag{4.4}$$

Therefore, the equilibrium constants of these four reactions, $K_{1:1}$, $K_{2:0}$, $K_{2:1}$, and $K_{2:2}$, can be denoted by the corresponding quotients of concentrations of reaction product and the product of the concentrations of all reactants. Meanwhile, those equilibrium constants can be related to the Boltzmann weighting of the minima in the PMFs we simulated, and the functions are listed below.

$$K_{1:1} = \alpha \exp(-\frac{\omega_{1:1}}{RT}) \tag{4.5}$$

$$K_{2:x} = \beta \exp(-\frac{\omega_{2:x}}{RT}) \tag{4.6}$$

Here $x = 0, 1,$ or 2 means number of encapsulated guests. We introduced a monomeric pre-factor α and a dimeric pre-factor β to make up the contributions missed in the PMF simulations here. These contributions include different reaction trajectories where no complex was formed, and different orientations the two hosts could be in[110–112]. In chapter 3, we tuned these two pre-factors to reproduce the assembly progression trend of OA

observed in experiments. From there we determined $\alpha = 2 \times 10\text{-}4$ M-1 and $\beta = 8 \times 10\text{-}11$ M-1. In this calculation we will continue to use these two values to keep consistence.

Here we still assume the total concentration of cavitand hosts is 3 mM in solvent. The host molecules were distributed into all distinct types of assembly complexes so that this total concentration equals the summation of the multiplication product of each assembly state's concentration and the corresponding number of cavitand hosts in it.

$$[1]_{total} = [1:0] + [1:1] + 2([2:0] + [2:1] + [2:2]) \tag{4.7}$$

As in chapter 3, the bulk alkane guest concentration was determined from the relationship below.

$$[G] = \frac{P_G}{RT} \exp(-\frac{\mu_G^{ex}}{RT}) \tag{4.8}$$

Here μ_G^{ex} is the excess chemical potential of the guest in water, RT is the product of temperature and gas constant, and P_G is the gas phase alkane partial pressure. For short alkane guests from methane through butane, whose boiling points are below 298.15 K, the guest pressure was assumed to be 1 atm. In the case of longer guests, the guest partial pressure was assumed to be the vapor pressure of the pure alkane liquid at 298.15 K. The details about how to obtain the alkane partial pressures and alkane guest excess chemical potentials can be found in chapter 3 and Appendix B. And finally, based on mass balance of cavitand host molecules, a quadratic equation of free host concentration can be derived as below.

$$2(K_{2:0} + K_{2:1}K_{1:1}[G] + K_{2:1}K_{1:1}^2[G]^2)[1:0]^2$$

$$+(1 + K_{1:1}[G])[1:0] - [1]_{total} = 0 \tag{4.9}$$

Once the free host concentration was solved, the concentrations of all the other assembly states can be determined.

4.4 Results and Discussion
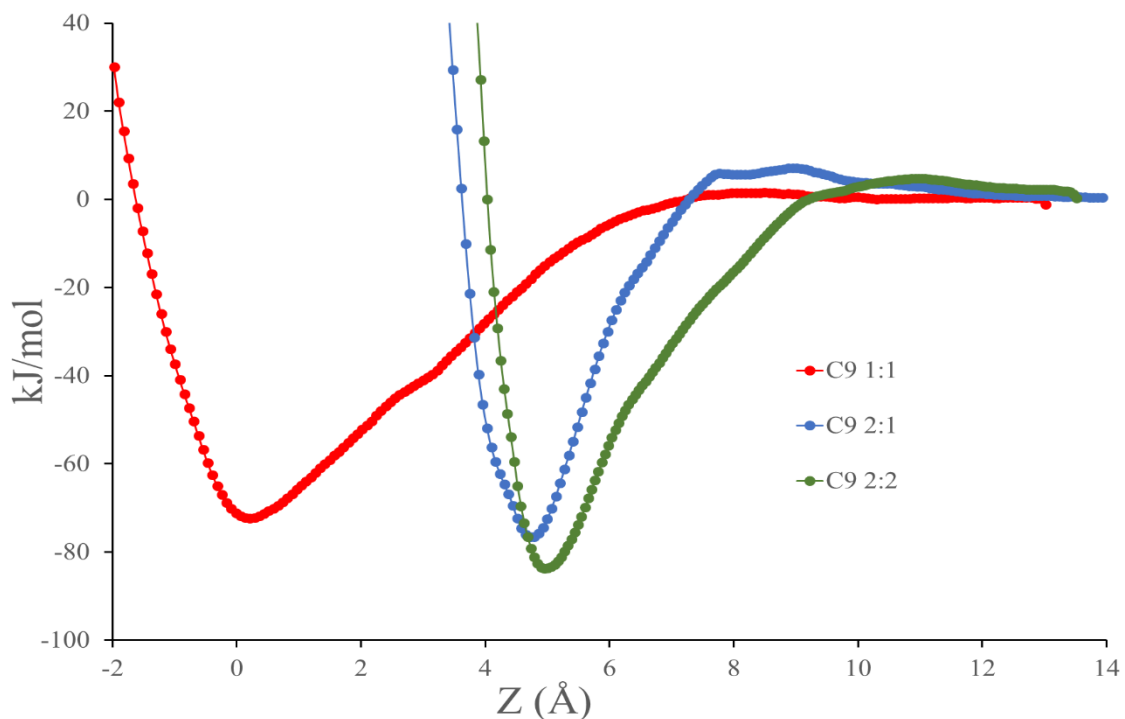
Figure 4.2: Potentials of mean force for association of 1:1, 2:1, and 2:2 complexes between nonane and TEXMOA. Z = 0 Å means the position of the face of the cavitand, which is on the left-hand side and facing right. Symbols for the 1:1, 2:1, and 2:2 interactions are defined in the figure legend. We omitted the error bars for clarity.

In Figure4.2, the potential of mean force between C_9 and TEXMOA from bulk water into the hydrophobic pocket along the C_4-axis to form a 1:1 complex is shown in red curve. Clearly, the interaction between the alkane and TEXMOA is attractive. The minimum of the PMF attractive well sits around the portal of TEXMOA (z ≈0 Å). This attractive interaction extends out into bulk water by ~7 Å. Compared to OA, it extends further into

the bulk, partially due to the additional favorable van der Waals interactions between the alkane and those four methyl groups pointing forward. The attractive well depth (-72 kJ/mol) is also deep enough to overcome the unfavorable solvation free energy of C_9 in water (17 kJ/mol), which indicates a stable 1:1 complexation. This well depth is comparatively the same as that of 1:1 complex of OA and C_9 (-70 kJ/mol).

The attractive well depth of 2:1 TEXMOA/C_9 complex (-77 kJ/mol) is not significantly deeper than that of 1:1 TEXMOA/C_9 complex. But it still indicates more free energy gain by burying the part of C_9 chain that is exposed to bulk solvent while forming monomeric complex. This deeper PMF well depth can also be attributed to the association between the aromatic rings forming the cavitand portal. Those hydrophobic groups prefer to get to a relative hydrophobic environment rather than stay in bulk water. The PMF minima of 2:1 complex lies at ~4.8 Å, which is significantly larger compared to that in 2:1 OA/C_9 complex. TEXMOA has four methyl units at *exo* positions on its rim. Those extending-out functional groups will add some steric hindrance so that the two cavitands cannot associate so tightly as OA did in 2:1 complex. And this looser contact also contributes to the less attractive free energy between the two cavitands.

In the case of 2:2 TEXMOA/C_9 complex, the PMF minima is at ~-84 kJ/mol. This is even deeper than that of 2:1 TEXMOA/C_9 complex. The distance between 2 cavitands' portals when PMF minima occurs is ~4.9 Å, which is only about 0.1 Å further to the right than that for 2:1 TEXMOA/C_9 complex. Though the extent of the 2:2 TEXMOA complex interaction still has a longer range than that of 2:1 TEXMOA complex. This range different is only about 1 Å, which means a much tinier difference than that in the case of OA. All of these reflect that the packing of 2 C_9 chains within a TEXMOA dimeric capsule is not as

frustrated as that within an OA dimer. Note that the contact between two TEXMOA cavitands is looser and the equilibrium distance is larger. There is more available space within TEXMOA dimer to accommodate 2 C_9 chains. The additional C_9 molecule doesn't have much repulsive contribution to the association free energy. In addition, more attractive interactions between the hydrophobic pocket and that additional alkane chain contributes to this deeper well depth. All of these can explain why 2:2 TEXMOA/C_9 complex exhibits a deeper well depth than 2:1 TEXMOA/C_9 complex.

We can quantitatively characterize the strength of binding between different constituents by the association free energy of each distinct assembly complex. And to some extent, the association free energy can be quantified by the minima of potential of mean force. In figure4.3, the PMF minima for the 1:1, 2:1, and 2:2 complexes as a function of number of carbon atoms in guest carbon chain was plotted. In 1:1 complex, for short sized guests from
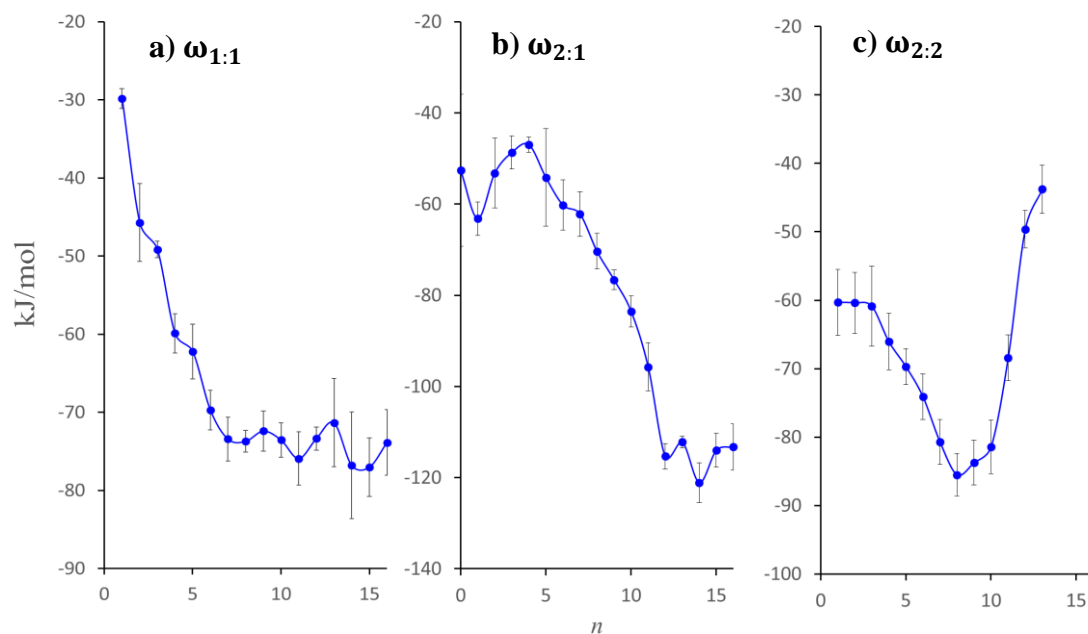
Figure 4.3: Potential of mean force minima determined for TEXMOA/alkane complex as a function of guest size. Part a, b, and c reports minima for 1:1, 2:1, and 2:2 complexes respectively. Error bars show one standard deviation.

methane to heptane, the association between host and guest becomes increasingly attractive with the increase of guest length. More methylene units can be encapsulated and contact with the hydrophobic inner pocket, and therefore more favorable van der Waals interactions were formed. Compared to OA and TEMOA, this decrease of association free energy has a longer range because additional interactions can be formed between longer guest and the four methyl units extending into bulk water. But begin from guests slightly longer than the hydrophobic pocket, this increasingly attractive trend stopped, and it exhibited a flat plateau from octane. Longer guests extend out into bulk water so that no additional favorable interactions can be acquired. The PMF minima didn't show significant changes for all the guests longer than octane.

In the case of 2:1 complex, in general, the association free energy for guests in the whole range is attractive. For short alkanes from methane to pentane, it exhibited an increasing trend. But relatively larger error bars indicate that those association free energy values are still in neighborhood. This association free energy is more positive compared to that of OA and TEMOA because of the larger equilibrium distances and looser contact between the two cavitands. But similarly, from heptane, the association free energy underwent a very drastic decrease. The 2:1 complex is further stabilized by the guests who can extend into the hydrophobic pocket of the second cavitand. They can fill up the inner space and gain benefits from two host molecules at the same time. This association free energy reaches a broad minimum centered at tetradecane. In essence, despite the four methyl groups on *exo* positions, this association free energy profile of TEXMOA 2:1 complex with a drastic decrease followed by a broad minimum is very similar with that of OA. Unlike TEMOA, these methyl groups extend outwards so that the portal of TEXMOA is not narrowed down.

Therefore, longer guests can accommodate themselves within TEXMOA capsules using more compact packing motifs. Similarly, this competition between more advantages from more favorable interactions and the disadvantages from frustrating guest conformations contributed to this broad minimum. This broad minimum emerged with longer guests in case TEXMOA, and it is also due to the larger space in the capsule brought by the longer distances between two cavitands.

The association free energy of 2:2 TEXMOA complex was initially flat from methane to propane. Large error bars indicate relatively larger statistical fluctuation here. Beginning from butane, it exhibited a sharp decrease and became more and more attractive. A broad minimum also emerged from octane to decane. After that, formation of 2:2 complex became strongly unfavorable and the association free energy profile increased very quickly. Compared to TEMOA and OA, this turning point appeared very late. 2:2 complexes with longer guests are much more stable in the case of TEXMOA. As what we discussed above, the larger equilibrium distance between two TEXMOA cavitands provides more interior space for the packing of 2 longer carbon chains within dimeric capsules. And according to chapter 3, this difference in 2:2 complex association free energy profile may lead to totally different assembly patterns.

From the network reaction model detailly described in chapter 3 and in methodology part here, the predicted equilibrium distribution of TEXMOA/C_n complexes as a function of alkane chain length is plotted in figure 4.4. In general, it is a monotonic progression with increasing guest size. From the beginning, methane didn't bind with the host. Ethane also exhibited a weak inclination to bind, and 1:0 complex, which means an empty cavitand
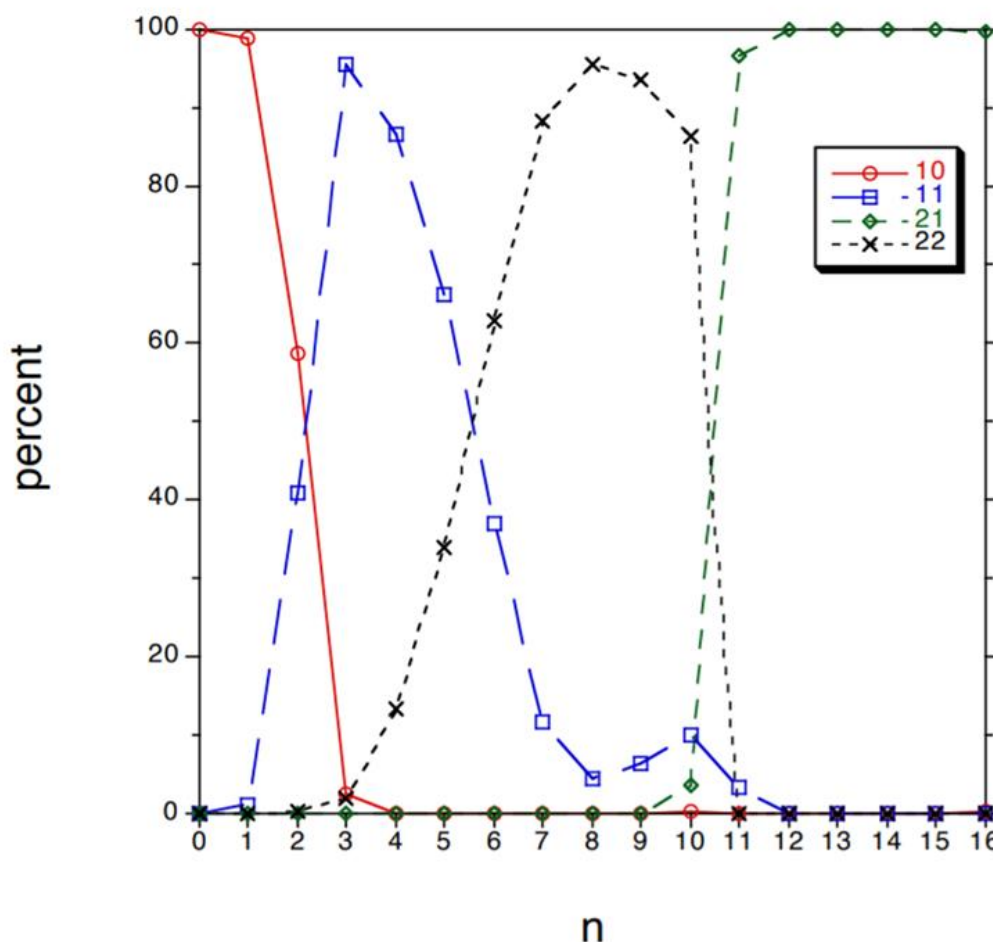
Figure 4.4: Predicted distribution of TEXMOA/alkane assembly states between 1:0, 1:1, 2:1, 2:2 complexes as a function of guest size. Symbols for the 1:0, 1:1, 2:1, 2:2 complexes are defined by figure legend. The population of 2:0 complexes was below 0.1% so that they are not reported. Error bars are not reported for clarity.

monomer, dominated. With the increase of guest size, monomeric 1:1 complex dominated for guests from propane to pentane, dimeric 2:2 complex dominated for guests from hexane

to decane, and finally dimeric 2:1 complex dominated for guests from undecane to hexadecane. A negligible population of 2:0 complex, which means an empty capsule, was also predicted and it was not plotted in figure 4.4. And there is also no empty cavitand dimers observed in experiments. Though TEXMOA exhibited a very similar assembly behavior with that of OA, some significant differences can still be observed. For small sized guests like ethane, the tendency to assemble into 1:1 complex is weaker. For TEXMOA, all the transition between distinct assembly states were shifted and occurred with longer guests. The transition from 1:1 complex to 2:2 complex lies between pentane and hexane, and the transition from 2:2 complex to 2:1 complex occurs between decane and undecane. This difference can be attributed to the different PMF minima. The formation free energy of TEXMOA 2:1 complex with middle-sized guests like pentane is much more unfavorable than that of OA. And conversely, the formation free energies of TEXMOA 2:2 complex, especially with long guests such as decane and undecane, are more attractive and negative compared to OA.

This monotonic assembly trend can be more easily recognized from the average host molecule aggregation number in each TEXMOA assembly complex, which is shown in figure4.5 below. From no guest to ethane, the aggregation number is near one because their inclination to bind with TEXMOA is very weak. The aggregation number increases from one to near two over the range of guests from propane to heptane. During this stage, majority of TEXMOA firstly form into 1:1 complex with propane and butane. Progressively, a mixture of 1:1 and 2:2 complexes were formed with presence of pentane and hexane. With heptane, 2:2 complex dominated. From octane, the aggregation number stays around two, which indicates TEXMOA mainly assemble into dimeric capsules. It is

clearly that methylation on *exo* positions doesn't have fundamental impact on deep-cavity cavitand's assembly behavior.
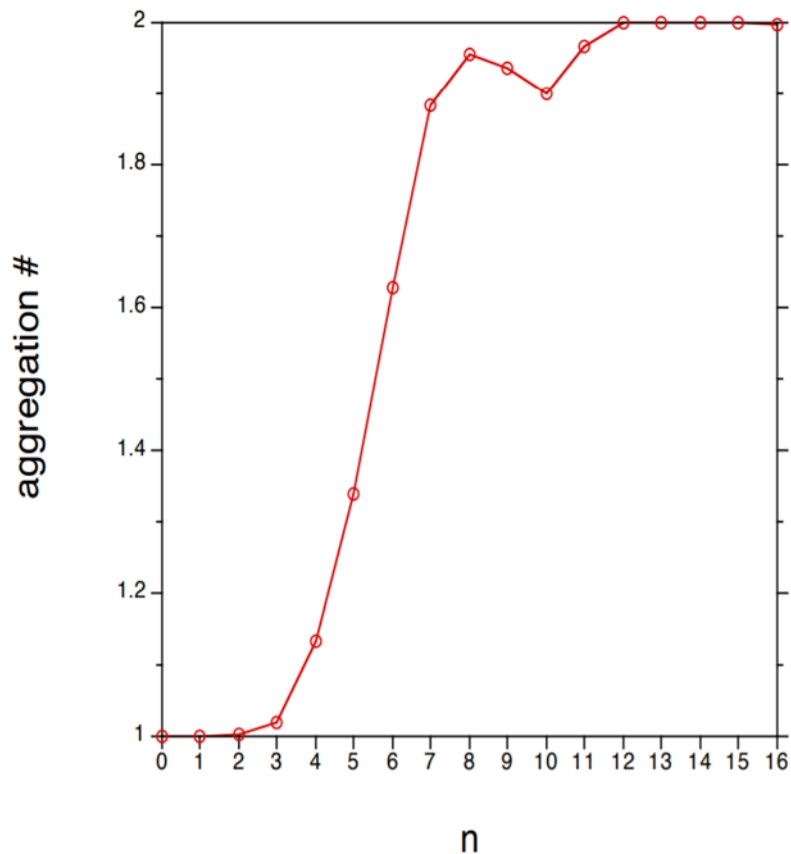
Figure 4.5: Average TEXMOA aggregation number in TEXMOA/alkane complexes as a function of the alkane guest length.

The minor change brought by methyl groups on exo positions can be explained by their structures. To get a deeper understanding of it, the equilibrium separation distances between two cavitands in dimeric capsules are plotted in figure 4.6 as a function of guest sizes. The equilibrium separation of 2:1 TEXMOA complexes is ~4.8 Å. It is independent of guest size in the whole range of alkane guests. Their equilibrium separations are significantly larger than that of 2:1 OA complexes (~3.5 Å). This larger inter-cavitands distance directly reflects the looser contact and less interactions between two hosts. As

mentioned above, it is the main contributor to the less attractive association free energy of 2:1 TEXMOA complexes with short and middle-sized guests, compared to that of 2:1 OA complexes. But for longer guests, larger inner space in a hydrophobic environment is generated to accommodate them within the capsules. Some compact packing motifs can be adopted since the month of TEXMOA is not narrowed down.

In the case of 2:2 TEXMOA complexes, up to heptane, the inter-cavitands equilibrium distance stays at 0.48 Å. This is consistent with the separation distance in 2:1 complex. From octane it starts to slightly increase, but it is still in the neighborhood of 0.48 Å. Beginning from tridecane, 2:2 TEXMOA complexes became instable with a dramatically
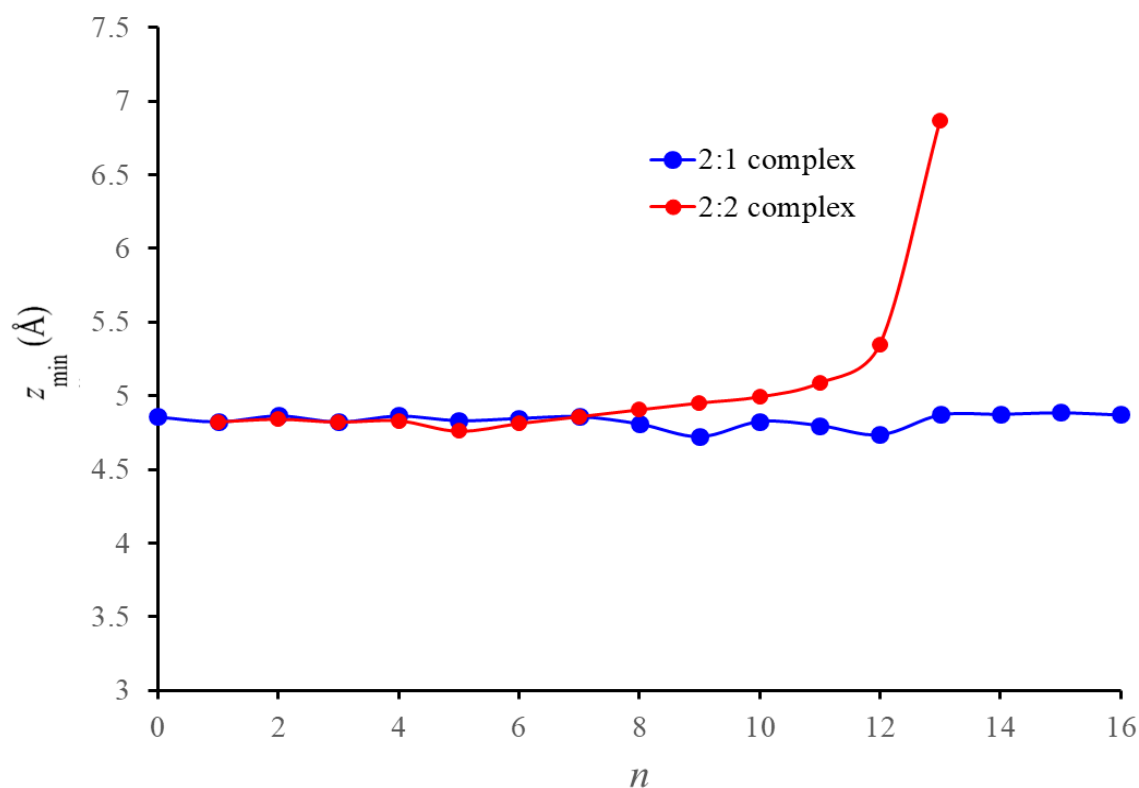
Figure 4.6: Equilibrium separation distance between cavitand hosts of 2:1 and 2:2 TEXMOA complexes at where PMF minimum occurred as a function of guest size.

increasing separation distance. In contrast to OA, likewise, due to the larger inner space resulted from the four methyl groups extending forwards, 2:2 TEXMOA dimer is stable for a wider range of guests including nonane and decane. This explains the shift of the transition between 2:2 and 2:1 complexes towards longer guest regions.

4.5   Conclusion

We conducted and presented a molecular dynamics simulation of the assembly and complexation of a modified deep-cavity cavitand derivative, TEXMOA. A monotonic assembly behavior was predicted from the network reaction model we developed. It progressively transits from 1:1 complex to 2:2 complex, then finally to 2:1 complex with the increase of guest sizes. The methylation on *exo* positions of the hydrophobic rim doesn't introduce fundamental change in assembly profiles since the *exo*-pointing methyl units extend the distances between cavitands and create a more comfortable environment for longer guests. As a result, 2:2 complexes are stabilized for a wider range of alkane guests, including some that will trigger disassociation of 2:2 complex in TEMOA. Furthermore, compared to OA, the transition between distinct assembly states are shifted towards longer guests also because of the difference in association free energies introduced from the stretched capsules. From this work, it can be concluded that even the same functionalization on a different position of supramolecular host molecules will lead to significantly different assembly behaviors. A minor change in supramolecular system design actually affects the whole body.

Chapter5
Pressure Induced Wetting and Dewetting of the Non-polar Pocket of Deep-Cavity Cavitands in Water

5.1   Summary

Hydrophobic interactions drive the binding of non-polar ligands to the oily pockets of proteins and supramolecular species in aqueous solution. As such, the wetting of host pockets is expected to play a critical role in determining the thermodynamics of guest binding. Here we use molecular simulations to examine the impact of pressure on the wetting and dewetting of the non-polar pockets of a series of deep-cavity cavitands in water. The portals to the cavitand pockets are functionalized with both non-polar (methyl) and polar (hydroxyl) units oriented either pointing upwards or inwards towards the pocket. We find wetting of the pocket is favored by the hydroxyl units and dewetting is favored by the methyl units. The distribution of waters in the pocket is found to exhibit a two-state like equilibrium between wet and dry states with a free energy barrier between the two states. Moreover, we demonstrate that the pocket hydration of the cavitands can be collapsed onto a unified absorption isotherm by assuming the effective pressures within each cavitand pocket differ by a shift pressure that depends on the chemical identity and number of functional units placed about the portal. These observations lead to the development of a two-state capillary evaporation model that accurately describes the equilibrium between states and natural gives rise to the effective shift pressures observed from simulation. This

work demonstrates that the hydration of supramolecular host pockets can be tuned following simple design rules that in turn are expected to impact the thermodynamics of guest complexation.

5.2 Introduction

Near water vapor/liquid coexistence, idealized, purely repulsive hydrophobic surfaces in aqueous solution are predicted to be enshrouded within a vapor-like layer as a result of the preferential self-hydration of water over interfacial wetting[95,128–132]. Confining water between non-polar mesoscale surfaces or within non-polar pockets is subsequently predicted to nucleate solvent evaporation, magnifying attractive hydrophobic interactions between non-polar moieties[95,133–137]. This nebulous vapor layer, however, is suppressed by ubiquitous van der Waals interactions that draw liquid water back into contact with the solute surface[138]. Nevertheless, water density fluctuations next to realistic, extended hydrophobic surfaces have been shown to be akin to those at a vapor/liquid interface[139], indicating the neighboring solvent density alone is insufficient for quantifying the hydrophobicity of meso-scale and larger interfaces[140].

The potential relevance of hydrophobic dewetting phenomena is highlighted by experimental and theoretical results indicating that the non-polar cavities of some proteins spontaneously dry in water to impact their function. For instance, crystallographic and molecular simulation studies of the L99A mutant of T4 lysozyme provided some of the first evidence of a drying transition within a protein's non-polar cavity[49]. While dry at atmospheric pressure, water was shown to cooperatively fill lysozyme's internal non-polar cavity with increasing pressure (~1000 bar) providing a mechanism for protein

denaturation. Alternately, the conduction of sodium across pentameric ligand-gated ion channels is regulated by subtle iris-like conformational changes about the transmembrane pore[141]. In the open-state the pore is fully hydrated, but in the closed-state hydrophobic isoleucine residues constrict the aperture and induce drying of a 15 Å long segment of the channel. This blocks free ion passage despite the fact the pore is still large enough to permit solvent egression. Similar vapor-lock gating mechanisms have been proposed for the nicotinic acetylcholine receptor[142] and *Escherichia Coli*'s small mechanosensitive channel[143]. Finally, NMR and molecular dynamics simulation studies have found that the ligand-binding pocket of bovine *β*-lactoglobulin is largely devoid of water at ambient conditions, with only fleeting filling events[50]. Dewetting of this hydrophobic pocket is thereby expected to impact the kinetics and thermodynamics of fatty acid binding to this apoprotein.

Considering the wetting of a hemispherical non-polar pocket in water, Setny *et al.*[144] examined the stability of water in the pocket is dependent on the approach of non-polar ligands. When the ligand was far away, water was found to freely fluctuate between all possible hydration states from zero (dry) to ten (wet) waters within the pocket with practically no free energy differences between states. As the ligand approached the pocket opening the hydration free energy landscape was found to dramatically change to a bimodal distribution between wet and dry states with a barrier between them. Upon ligand binding the dry state becomes dominant. This drying-mediated ligand binding in this model system was shown to be dominated by a favorable enthalpy of association resulting from gaining water-water attractive interaction upon release from the pocket and opposed by an unfavorable entropy of association[45,145]. This observation stands in opposition to traditional

descriptions of hydrophobic interactions, which associate non-polar solute aggregation with a strongly favorable entropy resulting from the release of structured waters from non-polar surfaces[146]. Indeed, enthalpically driven association has been experimentally observed for guests binding to proteins and supramolecular species in water[42,119]. Nau and co-workers[41,147] have interpreted this "non-classical hydrophobic interaction" in terms high energy waters that forfeit hydrogen-bonds within non-polar pockets. They have demonstrated a strong correlation between lost hydrogen-bonds and the binding constants between a range of supramolecular hosts and non-polar guests. This correlation, however, does not account for the equilibrium between wet and dry states observed by Setny *et al*[144]. and theoretically predicted for water under non-polar confinement.

Deep-cavity cavitands are a class of water soluble, bowl-shaped supramolecular host species possessing a non-polar guest binding pocket that is approximately 8 Å deep with a portal opening approximately 8 Å wide (Figure 5.1)[56,61]. These supramolecular hosts resemble the model hydrophobic pockets considered by Setny and co-workers[45,144,145]. Appropriately sized non-polar and amphipathic guests readily bind to cavitand pocket facilitated by hydrophobic interactions. As such, understanding the hydration of the cavitand pocket is expected to help elucidate the role of the solvent in driving guest recognition and binding[107]. Recently, we demonstrated through a combination of both experiment and simulation that the hydration of the non-polar pocket of cavitands depends sensitively on the functionalization of the host portal, and the detailed information can be found in Appendix F. In particular, it was established that while water wets the pocket of the non-functionalized parent cavitand octa-acid (OA, Figures 5.1 and 5.2), the pocket of tetra-endo-methyl octa-acid (TEMOA, Figures 5.1 and 5.2), differing from OA by four

methyl units oriented towards the pocket, is dry during a significant fraction of the time. Dewetting of the pocket was found to be consistent with thermodynamic equilibrium between liquid and vapor states driven by interfacial forces. The drying of TEMOA was manifest in stronger binding of sodium alkyloate guests compared to OA, favored by a dominant enthalpy and opposed by a weaker entropy in agreement with the "non-classical hydrophobic interaction". This work illustrated that hydration of supramolecular hosts is a potential synthetic design target for tailoring guest binding, subject to subtle changes in the environment ringing the binding site.

Seeking to expand our prior work, we report here a molecular dynamics simulation study of the impact of the functionalization of deep-cavity cavitands on the hydration of their non-polar binding pockets. In addition to the parent, non-functionalized, cavitand octa-acid (OA), we consider seven additional cavitands with varying functionalities (Figure 5.1): Mono-*endo*-methyl octa-acid (MEMOA), 1,3-di-*endo*-methyl octa-acid (DEMOA), tri-*endo*-methyl octa-acid (TrEMOA), tetra-*endo*-methyl octa-acid (TEMOA), tetra-*exo*-methyl octa-acid (TEXMOA), tetra-*endo*-hydroxyl octa-acid (TEHOA), and tetra-*exo*-hydroxyl octa-acid (TEXHOA). The portals of the functionalized cavitands are ringed by either non-polar (methyl) or polar (hydroxyl) units that are oriented either at inward pointing *endo* positions or upward pointing *exo* positions relative to the host pocket (Figures 5.1 and 5.2). In addition, the series OA, MEMOA, DEMOA, TrEMOA, and TEMOA examines the impact of increasing the hydrophobicity of the portal. We analyze pocket hydration in terms of the spatial distribution of waters about the entire host, pocket hydration probability distributions, and the mean water absorption within the pocket. The thermodynamics of water absorption are subsequently quantified within the context of a

hydration distribution model. Finally, we construct a two-state capillary evaporation model based on our results to gain insight into the forces governing the extent of pocket hydration and the role of the functional units ringing the portal.

5.3   Method

Molecular dynamics simulation of a range of functionalized cavitands in water were performed using GROMACS 5.1[79]. Simulations were performed in the isothermal-isobaric ensemble at 25 °C with pressures ranging from -1000 bar to 2500 bar. The temperature and pressure were controlled using a Nosé-Hoover thermostat[80,81] and Parrinello-Rahman barostat[82], respectively. Water was modeled using the TIP4P/EW potential[108]. The cavitands and their counterions were modelled using the Generalized Amber Force Field[70]. The charge of each cavitand was set to be -6e to match the protonation state at pH 7[107]. The four benzoic acid groups around the rim of cavitand and two of the four benzoic acid groups on the feet were deprotonated (Figure 5.1). And the partial charges of each atom were obtained by AM1-BCC calculations following geometry optimization[106]. Eight distinct cavitands with varying functional units added about the portal to the hydrophobic pocket were simulated (Figure 5.1): Octa-acid (OA), mono-*endo*-methyl-octa-acid (MEMOA), 1,3-di-*endo*-methyl-octa-acid (DEMOA), and tri-*endo*-methyl-octa-acid (TrEMOA), tetra-*endo*-methyl-octa-acid (TEMOA), tetra-*exo*-methyl-octa-acid (TEXMOA), tetra-*endo*-hydroxyl-octa-acid (TEHOA), and tetra-*exo*-hydroxyl-octa-acid (TEXHOA). These cavitands exhibit varying degrees of methylation (OA, MEMOA, DEMOA, TrEMOA, TEMOA), orientation of methyl units relative to the portal (TEXMOA), and orientations of polar hydroxyl units about the portal (TEHOA and TEXHOA). Non-bonded Lennard-

Jones interactions were truncated beyond a separation of 9 Å with a mean-field dispersion correction for longer-range contributions to the energy and pressure. Cross Lennard-Jones interactions between unlike species were determined using Lorentz-Berthelot combining rules[148]. Particle mesh Ewald Summation with a real space cutoff of 9 Å was applied to calculate electrostatic interactions[71]. Bond-lengths involving hydrogens for the cavitands were held fixed using the LINCS algorithm[86], while water was held rigid using SETTLE[87]. In each simulation, a single cavitand was solvated in 2000 water molecules. Production runs were conducted for 200 ns following an equilibration run of 5 ns. The equations of motion were integrated using a 2-fs time step. Configurations were saved every 1 ps for post-simulation analysis of thermodynamic averages.

Water densities about the cavitands were determined by assuming the hosts effectively symmetric about their C_4-axis (Figure 5.2). This approximation allows us to apply a cylindrical coordinate system to cavitand to bin water densities about the cavitand and characterize pocket hydration.

The number of waters within the cavitand pockets, n, was determined by constructing a hexahedron bounding the pocket and counting the number of waters within the polyhedron. The top face of the hexahedron was determined by constructing a plane fitted through the four diphenyl ether oxygen atoms ringing the pocket portal, while the bottom face was determined by the carbon atoms as the bottom of the pocket. The remaining 4 faces were determined by planes passing through oxygen atoms at the portal and carbon atoms at the bottom following the 4-fold rotational symmetry of the cavitand. Hydration state probability distribution functions were determined from analysis of the number of waters in the pocket over saved configurations.

Finally, we considered the impact of the number of waters with a host pocket on its partial molar volume. The partial molar volume of OA as a function of the pocket hydration state was determined by taking the difference in the average simulation volume of the cavitand with n waters in the pocket less the average volume of a simulation box of water with the same number of waters in it

$$v(n) = \langle V(n) \rangle_{c+w} - \langle V \rangle_{w}, \tag{5.1}$$

Where the angled brackets ($\langle ... \rangle$) indicate simulation averages of the system volume determined with $(c + w)$ and without (w) the cavitand, while the argument n indicates the number of waters in the host pocket.
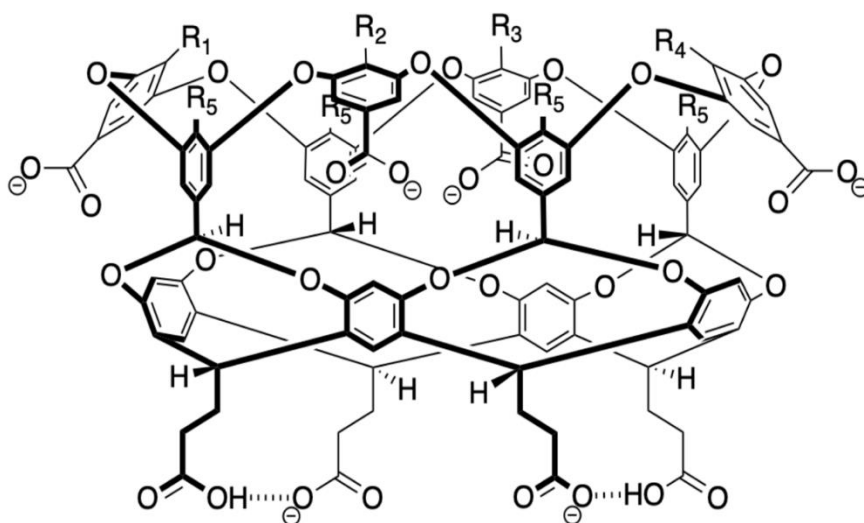
octa-acid (OA): $R_1 - R_5 = H$
mono-*endo*-methyl octa-acid (MEMOA): $R_1 = CH_3$, $R_2 - R_5 = H$
1,3-di-*endo*-methyl octa-acid (DEMOA): $R_1 = R_4 = CH_3$, $R_2 = R_3 = R_5 = H$
tri-*endo*-methyl octa-acid (TrEMOA): $R_1 - R_3 = CH_3$, $R_4 - R_5 = H$
tetra-*endo*-methyl octa-acid (TEMOA): $R_1 - R_4 = CH_3$, $R_5 = H$
tetra-*exo*-methyl octa-acid (TEXMOA): $R_1 - R_4 = H$, $R_5 = CH_3$
tetra-*endo*-hydroxyl octa-acid (TEMOA): $R_1 - R_4 = OH$, $R_5 = H$
tetra-*exo*-hydroxyl octa-acid (TEXHOA): $R_1 - R_4 = H$, $R_5 = OH$

Figure 5.1: Chemical structures of the eight deep-cavity cavitands simulated here.
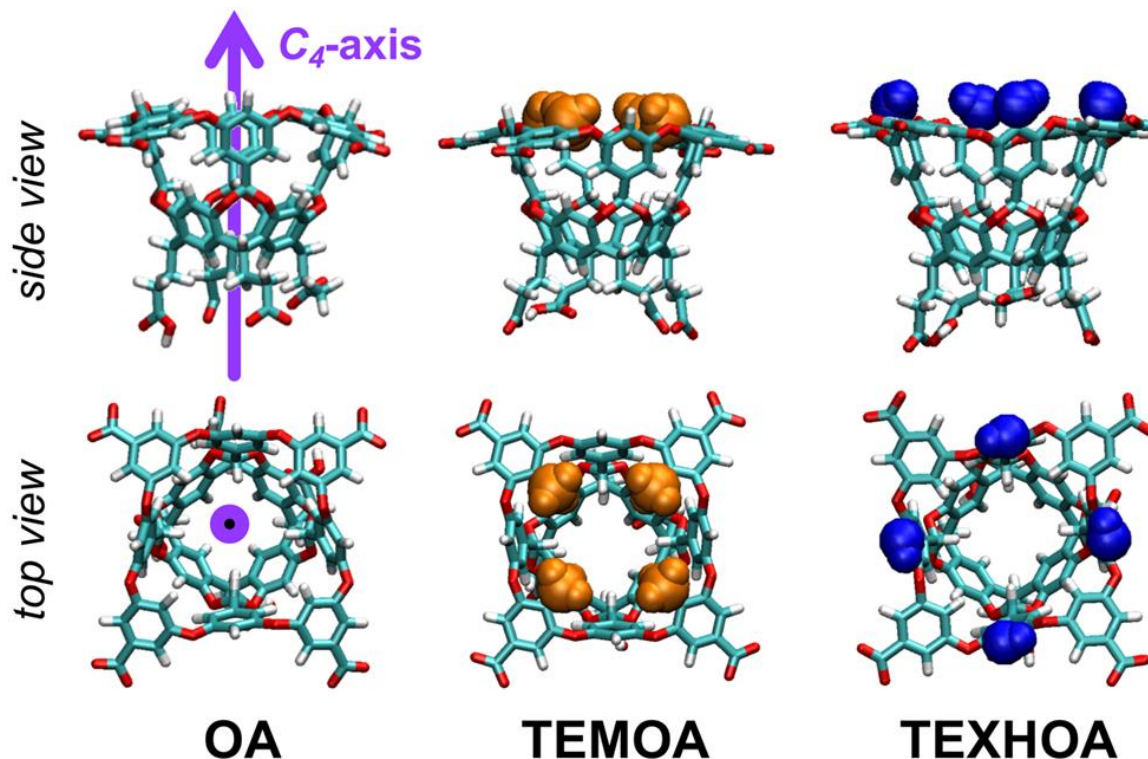
Figure 5.2: Molecular snapshots of the parent cavitand, octa-acid (OA), and an *endo*-functionalized, tetra-*endo*-methyl octa-acid, and *exo*-functionalized cavitand (TEMOA), tetra-*exo*-hydroxyl octa acid (TEHOA), host. The body of the parent cavitand is illustrated in licorice format, while the functional units are illustrated using a van der Waals representation. The cavitands are shown from the side showing the height of the walls of the pocket from the top and looking down into the pocket through the portal. The four-fold C_4-axis of rotational symmetry is denoted by the purple arrow pointing through the side view and up from the pocket of OA.

5.4   Results and Discussion

Cavitand Pocket Hydration. The average water density about OA, TEXMOA, TEMOA, and TEHOA is illustrated in Figure 5.3. The density in these plots is cylindrically averaged about the host C_4-axis of symmetry to clearly visualize the water distribution. The cavitands sit in the U-shaped black outline in these plots, corresponding to the location of

the heavy atoms of the host where the water density is zero. Water readily packs about the outside of the cavitands as indicated by the high-density regions ($\rho > 1.2$ g/cm^3, denoted in blue) from the rim of the pocket to the foot of the cavitand, forming approximately three water shells. In the case of OA (Figure 5.3a), water readily enters the hydrophobic pocket and packs against the inner walls of the host, however there is a region of depleted water density along the centerline of the cavitand within the pocket where the density drops to zero. This depletion is attributable to the van der Waals attractions between the waters and cavitand drawing the solvent to the inner walls of the host. One of the most significant absorption regions for water is the bottom of the pocket, where a large increase in water density is observed. This absorption site is big enough for a single water, which enjoys hydrogen-bonding with the four inner pointing hydrogens at the bottom of the pocket (Figures 5.1 and 5.2).). The distribution of water about TEXMOA (Figure 5.3b) closely follows that of OA, with only a slight depletion of waters near the rim of the pocket due to the presence of the four *exo*-pointing methyl units at the top of the cavitand. More significant perturbations in the water density are observed for TEMOA (Figure 5.3c). While the packing of water on the exterior of TEMOA is comparable to that for OA and TEXMOA, the water density along the inner walls of TEMOA's pocket is reduced. We do observe, however, a slight increase in the water density at the top of the pocket, which we believe can be attributed to favorable attractive interactions between water and the four *endo*-pointing methyls. For TEHOA we find water rewets the cavitand pocket as indicated by the increase in water density along the inner walls this host (Figure 5.3d), comparable to the wetting of OA and TEXMOA. An additional increase in water density is observed at the top of TEHOA's pocket ($z \approx 5$ Å) supported by hydrogen-bonding between the

endo hydroxyl units and water that makes pocket hydration favorable. We conclude from the analysis of water's density that while the distribution of water outside of all the cavitands are effectively the same, the hydration of the hydrophobic pocket depends sensitively on the rim functionalization.
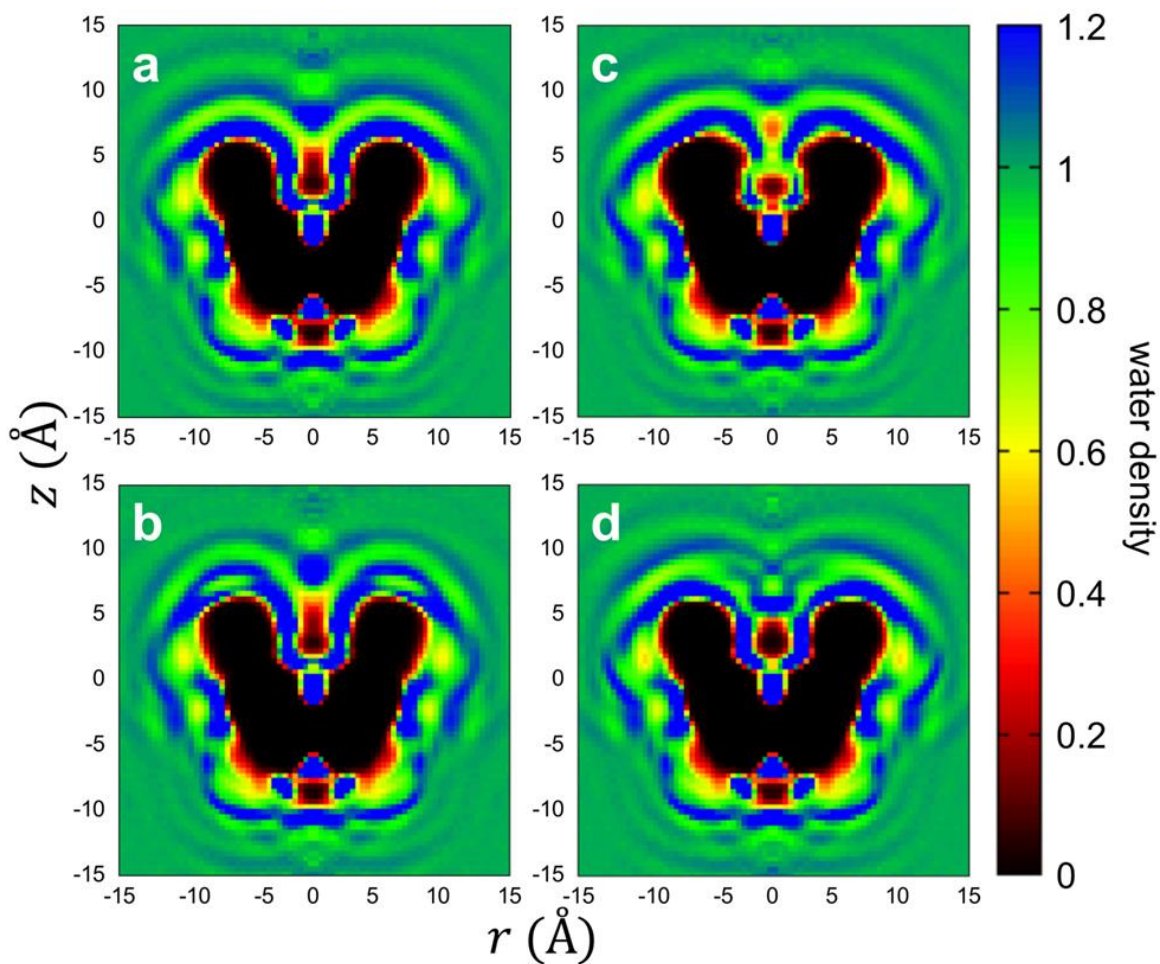
Figure 5.3: Average water density distribution about the hosts OA (a), TEXMOA (b), TEMOA (c), and TEHOA (d). The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figure.

The hydration of the host pocket is dynamic, with the number of waters within the cavitand continuously fluctuating. To capture the fluctuating nature of the pocket hydration, it is more informative to consider the probability distribution of hydration states. Figure 5.4 shows the hydration number probability distribution for water in OA, MEMOA, DEMOA, TrEMOA, and TEMOA at 1 bar and 2500 bar, illustrating the impact of increasing the hydrophobicity of the pocket portal. At atmospheric pressure OA exhibits a unimodal probability distribution (Figure 5.4a), with 4 waters being the most probable hydration state. This distribution is not symmetric as would be expected for a normal distribution but is skewed slightly towards lower occupancy states. As such, while the probability of observing $8 (= 4 + 4)$ waters in OA's pocket is effectively zero, the probability of observing an empty pocket $(0 = 4 - 4)$ is finite. This asymmetry is accentuated as the portal methylation increases. While we never observe more than 8 waters in the pocket, the probability of observing an empty pocket systematically increases from 1% for OA to 36% for TEMOA. Moreover, increasing portal methylation changes the distribution from unimodal for OA and MEMOA to bimodal for DEMOA, TrEMOA, and TEMOA. The second peak of the bimodal distributions occurs for the empty pocket, with a minimum between the two peaks occurring at $n = 1$. The position of the primary peak shifts from 4 waters for OA through TrEMOA to 3 waters for TEMOA. While not a true phase equilibrium between a liquid and vapor, the rise of the second, low-density state with increasing endo-methylation is reminiscent of capillary evaporation driven by decreasing pore wettability (see below). While not shown here, the hydration distribution for water in TEXMOA is practically indistinguishable from that of OA. This suggests that pointing the methyl units away from the pocket limits their influence on pocket hydration.
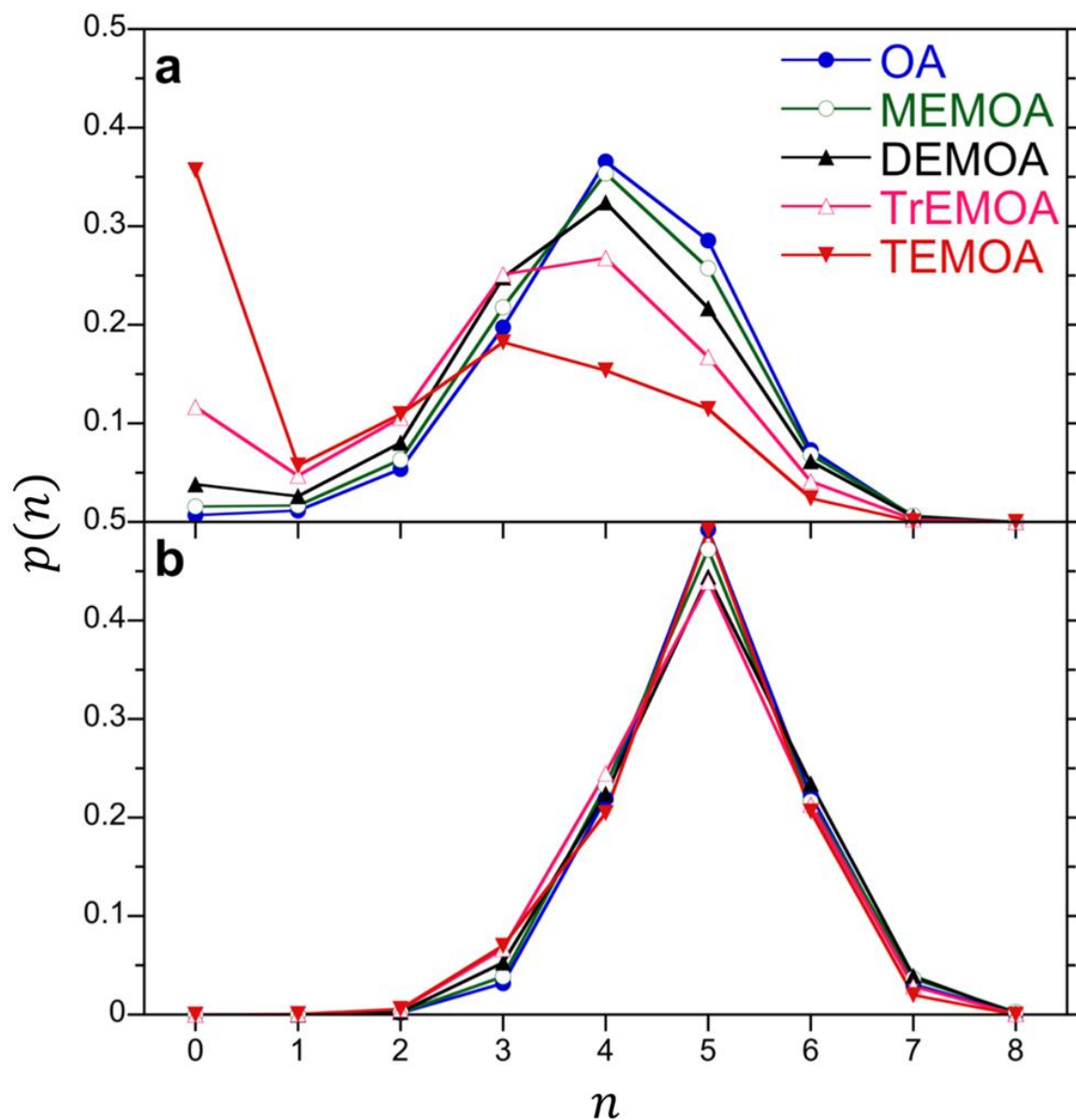
Figure 5.4: Hydration number probability distributions for water in the pockets of OA and *endo*-methyl functionalized hosts (MEMOA, DEMOA, TrEMOA, and TEMOA) at 25° and pressures of 1 bar (a) and 2500 bar (b). The figure symbols are defined in the legend. Error bars are comparable in size or smaller than the symbols.

It may be thought that the increasing methylation may impact the shape of the pocket, thereby playing a role in the hydration state distributions observed at atmospheric pressure. If we increase the pressure to 2500 bar, the dry ($n = 0$) state is suppressed and the distributions are unimodal (Figure 5.4b). The most probable pocket hydration state at elevated pressure is 5 waters as a result of pressure pushing more waters into the pocket. While there are subtle differences between the hydration distributions for all the cavitands, the shapes of the distributions are practically the same. This suggests no significant role for portal functionalization on the shape and volume of the pocket available for water absorption.
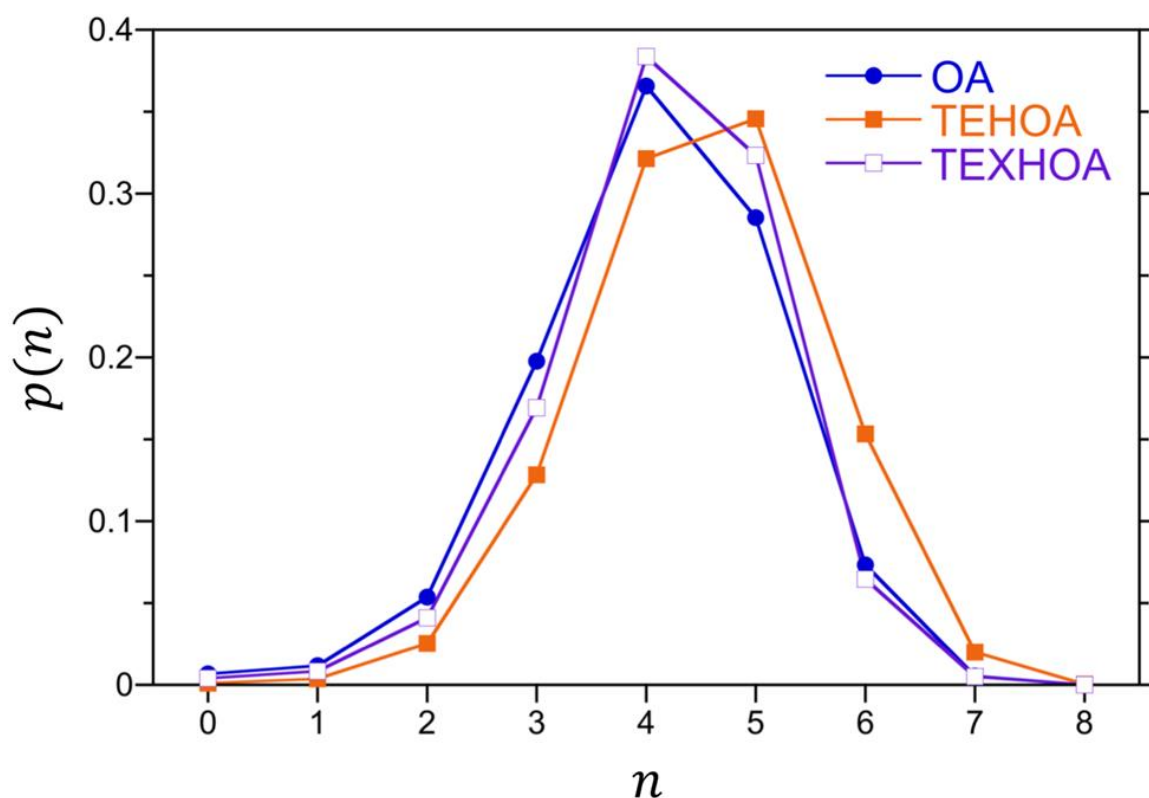
Figure 5.5: Hydration number probability distributions for water in the pockets of OA and the hydroxyl functionalized hosts (TEHOA and TEXHOA) at 25°C and 1 bar. The figure symbols are defined in the legend. Error bars are comparable in size or smaller than the symbols.

In difference to methylation, adding polar hydroxyl units about the portal can drive the host pocket to be more favorably wetted. The impact of functionalizing the portal with hydroxyl units placed at the *exo-* and *endo*-positions at atmospheric pressure is shown in Figure 5. While the methyl units play almost no role in the hydration of TEXMOA, for TEXHOA we find the larger occupancy states are preferentially stabilized compared to the OA. As such, the $n = 4$ and 5 occupancy states are slightly more probable for TEXHOA than OA. For TEHOA, the shift to higher pocket occupancy states is even more profound with $n = 5$ becoming the more probable hydration state, with the distribution becoming more symmetric.

The pocket hydration state can be manipulated by changing the system pressure, as suggested in Figure 5.4. In Figure 5.6 we provide a more detailed account of the pressure dependence of the pocket hydration distribution for OA and TEMOA. While we only observed a unimodal hydration distribution for OA above, Figure 5.6a shows that a second maximum in the distribution can be obtained by dropping the pressure to negative values, placing the water under tension (metastable). Specifically, as the pressure drops the probability of observing OA to be empty increases from 1% at 1 bar to 12% at -750 bar. The first hint of a secondary maximum for the empty cavitand occurs at -500 bar (not shown), where the probability of observing an empty cavitand is more probable than observing a single water in the pocket (i.e., $p(0) > p(1)$). The empty cavitand is even more probable at -750 bar. Similar to the hydration distribution for TEMOA reported above (Figure 5.4a), OA under tension exhibits another maximum at $n = 4$ separated from the empty state by a minimum (free energy barrier) at $n = 1$. While the empty cavitand presumably could further stabilized by dropping the pressure further, the simulations

became too unstable and failed. The pressure induced stabilization of the empty cavitand is even more prominent for TEMOA (Figure 5.6b), which is empty ~90% of the time at -750 bar. At this pressure, the hydration probability distribution for TEMOA is a monotonically decreasing function of the occupation state and a second maximum is not observed. Comparing OA and TEMOA, the hydration distributions for OA correspond roughly to that for TEMOA at a pressure 1000 bar greater. Overall the pressure
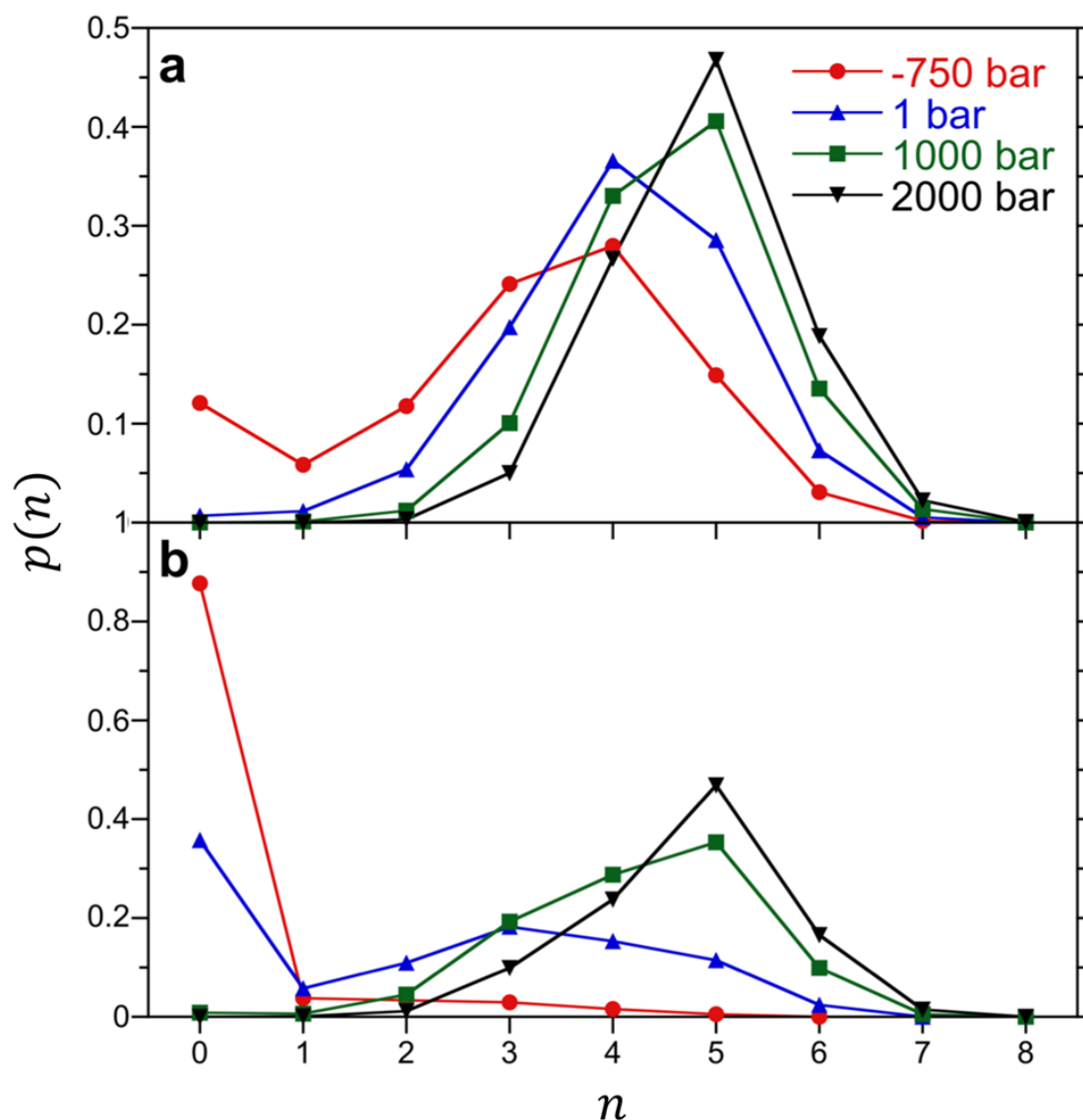
Figure 5.6: Impact of pressure on the hydration number probability distributions for water in the pockets of OA (a) and TEMOA (b) at 25°C. The figure symbols are defined in the legend. Error bars are comparable in size or smaller than the symbols.

dependencies of the hydration distributions for OA and TEMOA are consistent with water exhibiting two-state like equilibria between a dewetted/empty state and wetted/filled state.

Thermodynamics of Pocket Hydration. Rather than the individual hydration state probabilities, an alternate way of examining pocket hydration is to consider the average number of waters in a cavitand pocket as a function of pressure. The average number of waters in the pocket is determined by the sum

$$\langle n \rangle = \sum_{i=1}^{\infty} i p(i). \tag{5.2}$$

We compare $\langle n \rangle$ as a function of pressure for all the simulated cavitands in Figure 5.7. Generally speaking, the mean hydration number is an increasing function of pressure for all the cavitands. The hydrophobic, *endo*-functionalized cavitands typically have lower occupation numbers than OA at constant pressure, with $\langle n \rangle$ systematically decreasing with increasing methylation. Conversely, TEHOA exhibits greater pocket hydration than OA over the entire pressure range. The impact of functionalization of the *exo*-position by either methyl or hydroxyl units, on the other hand, only exerts a minor perturbation on the pocket hydration relative to OA. For the most hydrophobic pockets (e.g., DEMOA through TEMOA), the water binding isotherm appears sigmoidal, indicating water absorption is cooperative. As the pocket hydrophobicity decreases, however, the water binding isotherms systematically shift to lower pressures, obscuring the sigmoidal shape of the binding isotherms for the more hydrophilic cavitands (e.g., TEHOA). Taken together, these binding isotherms suggest water wets the pocket following two-state-like behavior as suggested by the probability distributions described above.
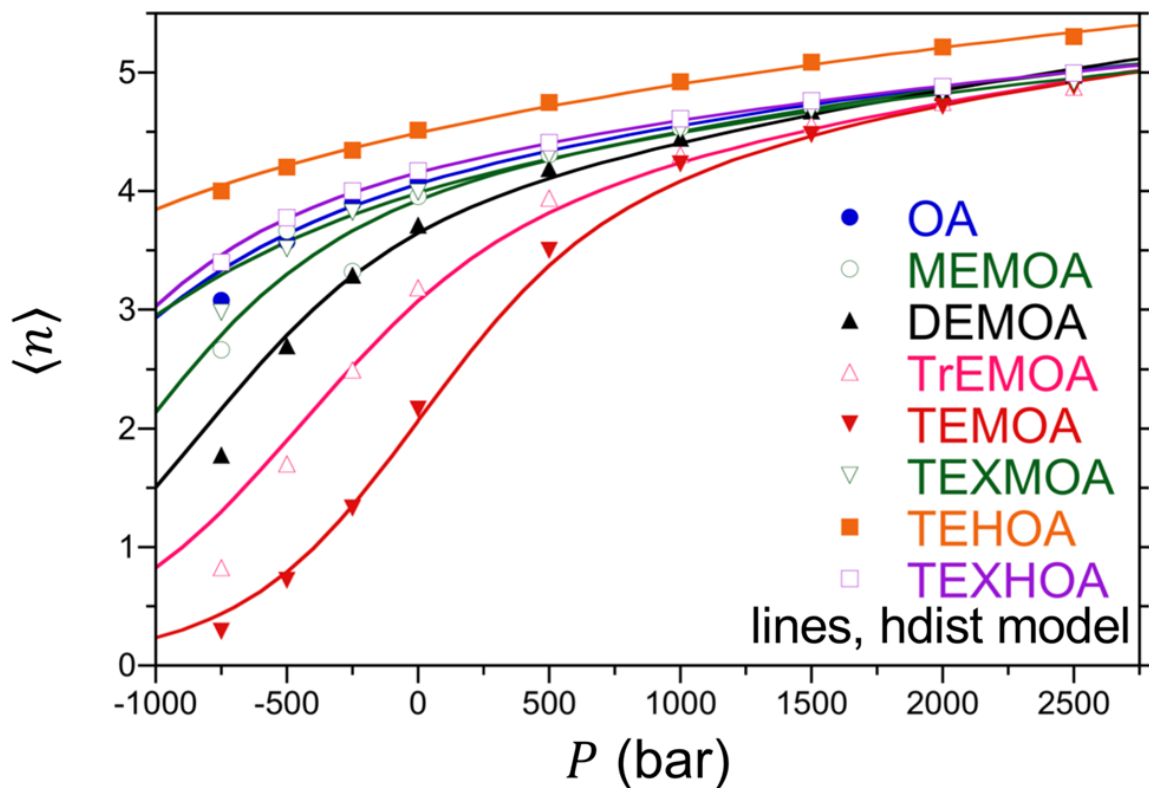
Figure 5.7: Mean pocket hydration numbers for all simulated cavitand hosts as a function of pressure at 25°C as determined from simulation and fits to the hydration distribution model (hdist). The figure symbols are defined in the legend. Error bars are comparable in size or smaller than the symbols.

We can construct an analytical expression for the absorption isotherm of an individual cavitand by considering the pressure dependence of the pocket hydration free energies. Specifically, the relative hydration free energy of a host pocket with n waters in it is

$$\Delta G(n) = -kTln\left(\frac{p(n)}{p(4)}\right), \qquad (5.3)$$

where we have chosen $n = 4$ as the reference state from which we measure the free energy. Assuming that for each hydration state the water inside the cavitand pocket is incompressible, the pressure dependence of the pocket hydration free energies can be expressed as

$$\Delta G_P(n) = \Delta G_1(n) + (P - 1 \, bar) \times \Delta v(n), \qquad (5.4)$$

where $\Delta G_1(n)$ is the pocket hydration free energy at 1 bar pressure, and $\Delta v(n)$ is the relative pocket hydration volume. We can subsequently obtain $\Delta G_1(n)$ and $\Delta v(n)$ by linear regression of the pocket free energies from simulation obtained following eq. (5.3) as a function of pressure. The occupation probabilities can subsequently be recovered from eq. (5.4) following the relationship

$$p(n) = \alpha exp\left(-\frac{\Delta G_P(n)}{kT}\right), \qquad (5.5a)$$

where the normalization constant α is determined as

$$\alpha = \frac{1}{\sum_{n=0}^{\infty} exp\left(-\frac{\Delta G(n)}{kT}\right)}. \qquad (5.5b)$$

Taken together we refer to this as the hydration distribution model (hdist).

The relative hydration free energies of OA for a number of different hydration states (n) are plotted as a function of pressure in Figure 5.8. The free energies are approximately linear functions pressure as described by eq. (5.4), consistent with the assumption that the waters in the pocket are to a first approximation incompressible. Qualitatively similar results were obtained for all the cavitands. In Figures 5.9 and 5.10 we plot $\Delta G_1(n)$ and $\Delta v(n)$, respectively, for OA obtained from least squares fitting of the hydration distribution model. As suggested by the hydration state distributions discussed above, $\Delta G_1(n)$ for OA (Figure 5.9) is approximately parabolic (i.e., Gaussian distribution) near its minimum at n = 4. The free energy for the dry state ($n = 0$), however, is significantly lower than would be anticipated based on a quadratic polynomial. The fitted hydration volumes (Figure 5.10), on the other hand, are a decreasing function of the number of waters within the pocket, consistent with the partial molar volume of the cavitand decreasing as the pocket fills with

water. If the volumes of water within the pocket matched that of water in the bulk, then we would expect $\Delta v(n)$ to be given as

$$\Delta v(n) \approx (4 - n) \times v_w, \tag{5.6}$$

where $v_w = 18.06 \text{ cm}^3/\text{mol}$ corresponds to the molar volume of bulk water. The bulk water approximation semi-quantitatively captures the n of the hydration volume, however, the fitted values of $\Delta v(n)$ deviate from linearity (Figure 5.10). Specifically, $\Delta v(n)$ is concave up, with the volume of the first water that enters the cavity, $\Delta\Delta v(0,1) = \Delta v(0) - \Delta v(1) = 25.0 \text{ cm3/mol}$, considerably greater than that of the bulk, and the volume of the seventh water that enters the cavity, $\Delta\Delta v(6,7) = \Delta v(6) - \Delta v(7) = 6.6 \text{ cm3/mol}$, considerably less than that of the bulk. Thus, water is compressed as we try to squeeze more and more waters into the pocket. We obtain excellent agreement comparing the changes in the relative partial molar volumes of OA as a function of the number of waters in its pocket determined directly from simulation (eq. (5.1)) against that for the hydration distribution model, giving us confidence that the model accurately reflects the physical volume of the cavitand. Interestingly, while $\Delta G_1(n)$ is distinct for each of the cavitands, $\Delta v(n)$ for all the cavitands are approximately equal to one another (Figure 5.10 inset).
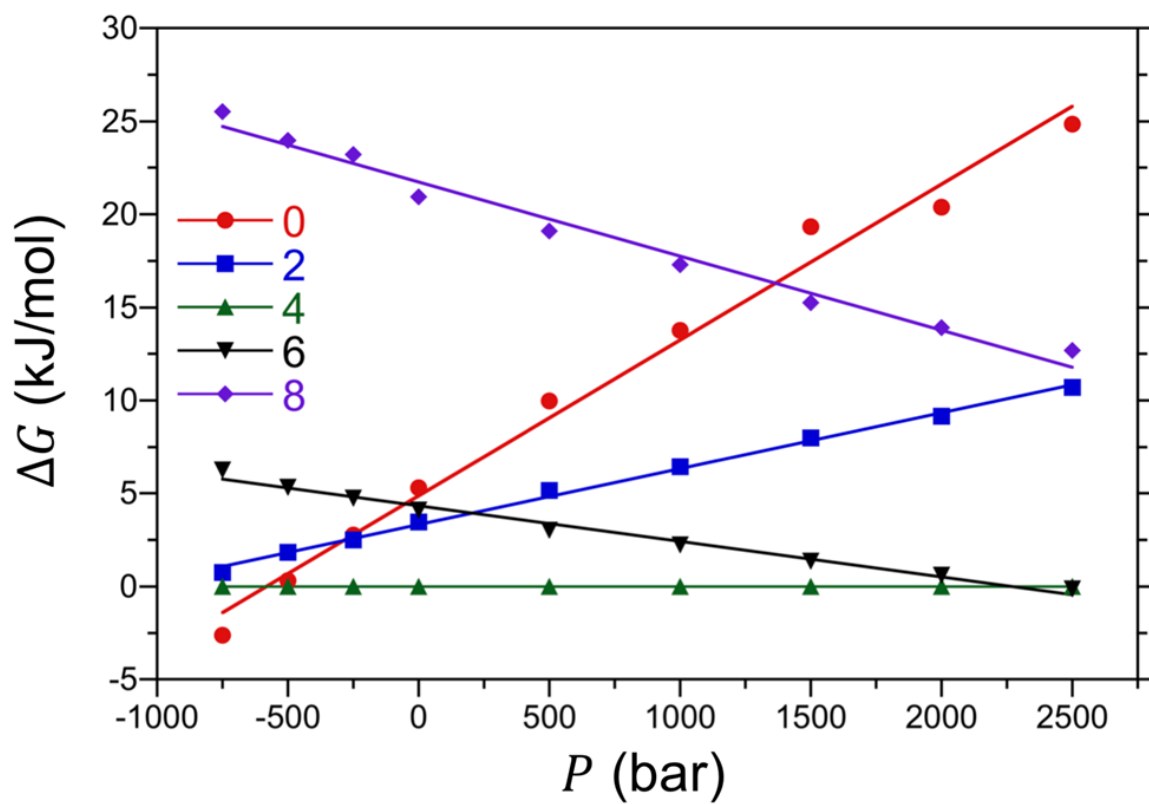
Figure 5.8: Relative pocket hydration free energies for DEMOA for water occupancy states of $n = 0$, 2, 4, 6, and 8 evaluated using eq. (5.3) as a function of pressure. Points indicate simulation results and lines indicate fits to eq. (5.4). The figure symbols are defined in the legend. Error bars are neglected for clarity.
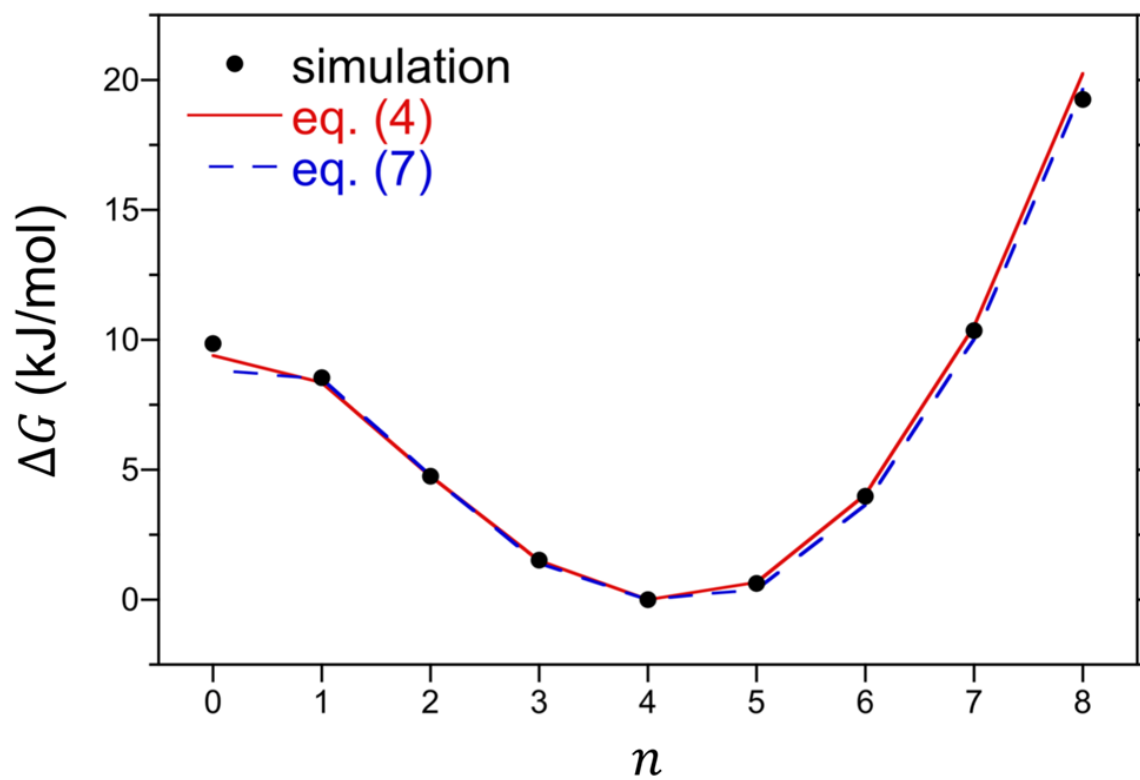
Figure 5.9: Relative pocket hydration free energies for OA at 25°C and 1 bar evaluated from simulation (via eq. (5.3)) and fits to the hydration distribution (hdist, eq. (5.4)) and unified distribution (udist, eq. (5.7)) models. The figure symbols are defined in the legend. Simulation error bars are comparable in size or smaller than the symbols.
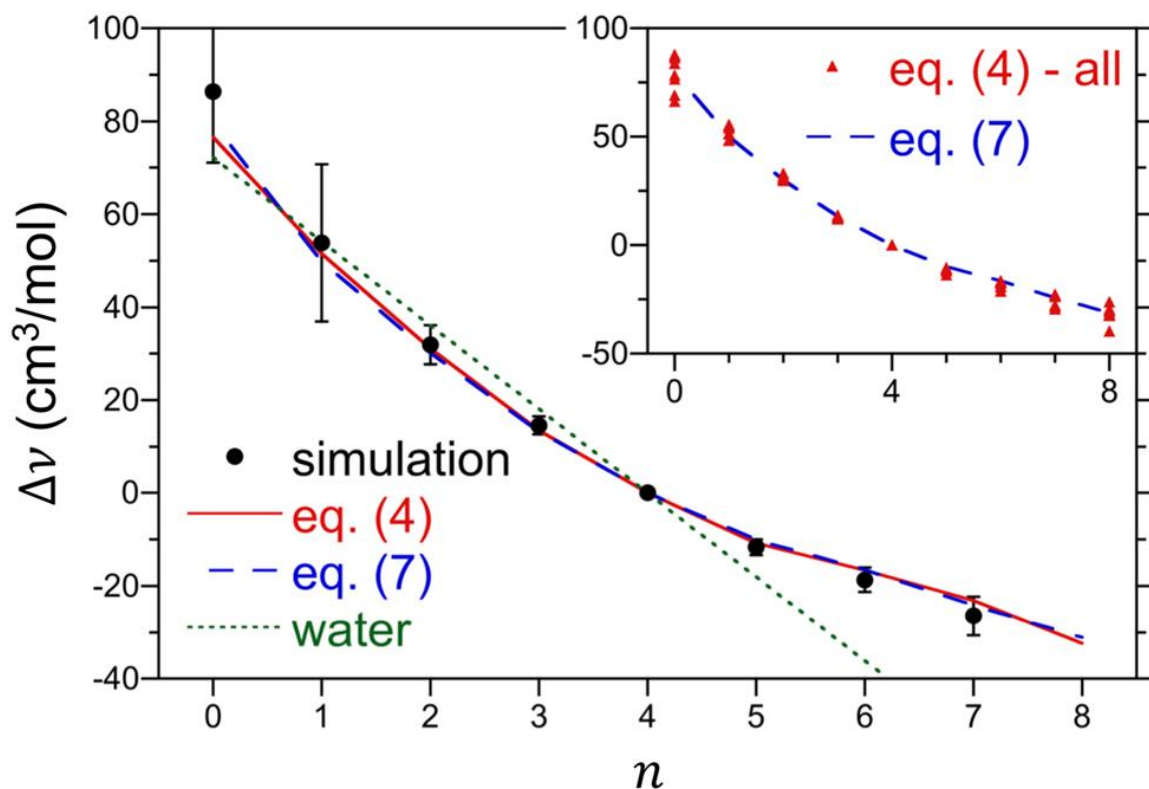
Figure 5.10: Relative pocket hydration volumes for OA at 25 °C evaluated directly from simulation at 1 bar (eq. (5.1)) and fits to the hydration distribution (hdist, eq. (5.4)) and unified distribution (udist, eq. (5.7)) models. We compare these results against the relative volumes one would obtain using the bulk volume of water following eq. (5.6). The inset figure compares the volumes determined from eq. (5.4) for all the simulate hosts against that determined from the unified distribution model (eq. (5.7)). The figure symbols are defined in the legend. Simulation error bars in the main figure indicate one standard deviation.

We compare the mean pocket hydration numbers obtained from the hydration distribution model against those determined from simulation in Figure 5.7. Overall, the model does an excellent job describing the water absorption isotherms to the cavitand pockets, validating the assumptions underlying the model. Moreover, the model allows us to evaluate the hydration numbers at pressures not considered in our simulations and more clearly observe the sigmoidal shapes of the binding isotherms.

The differences between the binding isotherms for all the hosts reported in Figure 5.7 to a first approximation appear to be horizontally shifted relative to one another. In addition,

the relative cavitand volumes obtained from fitting eq. (5.4) to the simulation results are quantitatively similar to one another (Figure 5.10 inset). We therefore propose that the water absorption isotherms for all the hosts can be fitted to a unified absorption isotherm following the functional form for the free energy

$$\Delta G_P(n) = \Delta G_1^*(n) + (P + P_{shift} - 1\ bar)\Delta v^*(n), \tag{5.7}$$

where $\Delta G_1^*(n)$ and $\Delta v^*(n)$ correspond to the universal pocket hydration free energy and hydration volume, and P_{shift} is a host dependent pressure shift measured relative to OA (i.e., $P_{shift} = 0$ for OA). We refer to this as the unified distribution model (udist). The parameters in eq. (5.7) have subsequently been obtained by performing a non-linear least squares fit to the pocket hydration free energies of all the simulated cavitands.

The unified distribution model is compared against the mean hydration numbers for all the cavitands in Figure 5.11 collapsed onto a unified absorption isotherm using the fitted shift pressures. Overall the unified absorption isotherm achieves an excellent quantitative description of water absorption over the entire pressure range. The fitted results for $\Delta G_1^*(n)$, $\Delta v^*(n)$, and P_{shift} are reported in Figures 5.9, 5.10, and 5.12, respectively. We observe excellent agreement between the universal results for $\Delta G_1^*(n)$ and $\Delta v^*(n)$ with $\Delta G_1(n)$ and $\Delta v(n)$ for OA (Figures 5.9 and 5.10), establishing OA as an excellent reference host to describe water absorption. P_{shift}, on the other hand, shows distinct trends with the degree and polarity of the portal functionalization. The largest effect is observed for *endo*-methyl functionalization, with P_{shift} monotonically decreasing from OA to TEMOA with each added methyl unit, corresponding to increasing destabilization wetted state by methyl functionalization. In the case of *endo*-hydroxyl functionalization, TEHOA exhibits a positive value for P_{shift}, corresponding to stabilization of the wetted state within the

pocket. The magnitude of P_{shift} for TEHOA is only about 60% that for TEMOA, however, indicating that *endo*-methyl functionalization plays a greater role on pocket hydration than *endo*-hydroxyl functionalization. As for the *exo*-methyl units, TEXMOA exhibits a negative value of P_{shift}, although its magnitude is less than that for MEMOA despite the fact TEXMOA has four times as many methyl units. Similarly, TEXHOA exhibits a positive value for P_{shift} that is lower in magnitude than that for TEHOA. We conclude then that while *exo*-functionalization plays a role in stabilizing/destabilization water within pocket, the magnitude of this effect is not as significant as that observed for *endo*-functionalization.

Figure 5.13 illustrates the pocket hydration free energies of the *endo*-functionalized hosts TEHOA, OA, MEMOA, DEMOA, TrEMOA, and TEMOA in order of increasing hydrophobicity at 1 bar from simulation and fitted to the unified distribution model. Overall the unified distribution model provides an excellent quantitative descript of the pocket hydration free energies for these cavitands, supporting the proposition that the impact of portal functionalization can be treated as a shift in the effective pressure. Importantly, increasing the hydrophobicity of the pocket portal leads to systematic stabilization of the dry state as indicated by the increasing depth of $\Delta G_1(0)$. Even for TEHOA, however, the free energy exhibits a shoulder at $n = 0$, suggesting a latent inclination for TEHOA's pocket to dry. This inclination subsequently becomes dominant for TEMOA. For the hosts MEMOA through TEMOA, the model predicts the dry state is separated from the wet state by a free energy barrier (i.e., maximum) at $n = 1$ This supports the proposition that the water absorption within the cavitand pockets is akin to a two-state like vapor/liquid transition, albeit on a microscopic scale. Presumably if we could make the pocket more

hydrophobic, we would expect to observe only a single free energy minimum for the dry state. The lack of additional endo sites and complication of adding conformational degree of freedom for larger functional units, however, suggests that this may not be easily to achievable.
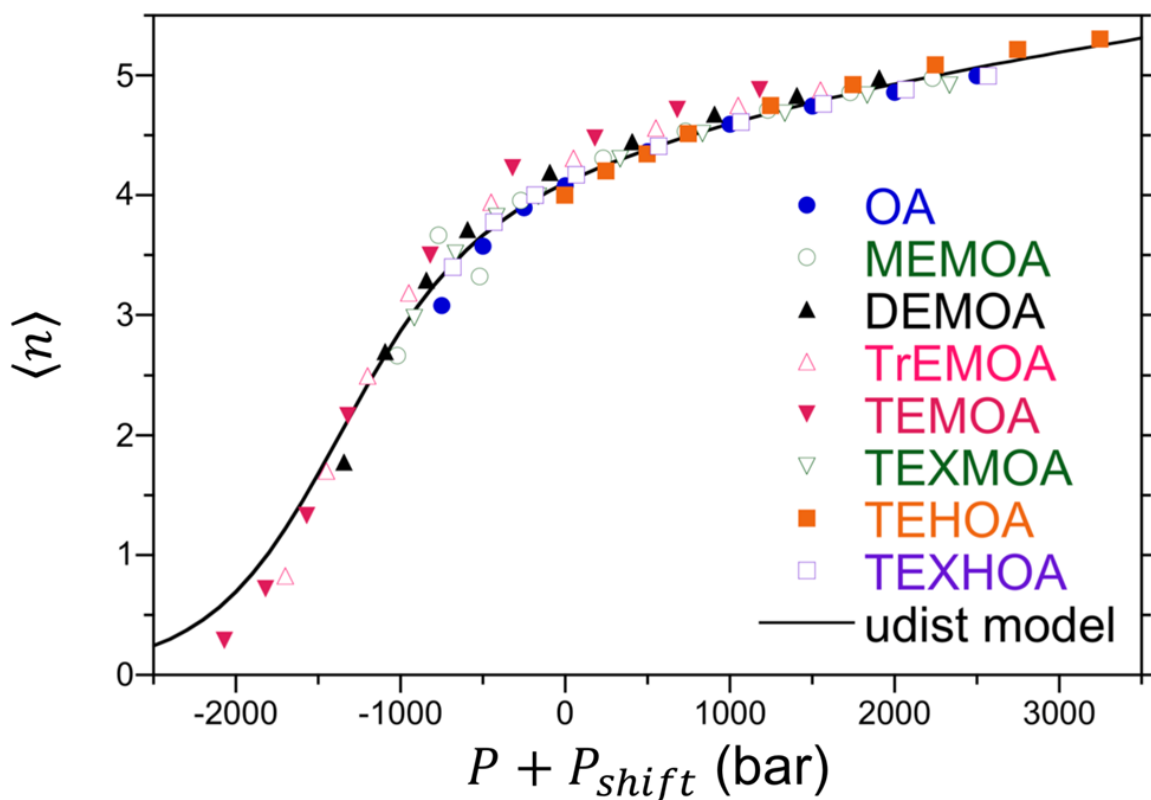
Figure 5.11: Unified absorption isotherm for all the simulated cavitand hosts as a function of effective pocket pressure ($P + P_{shift}$) at 25°C. Points indicate simulation results and lines indicate fits to the unified distribution model (eq. (5.7)). The figure symbols are defined in the legend. Error bars are comparable in size or smaller than the symbols.
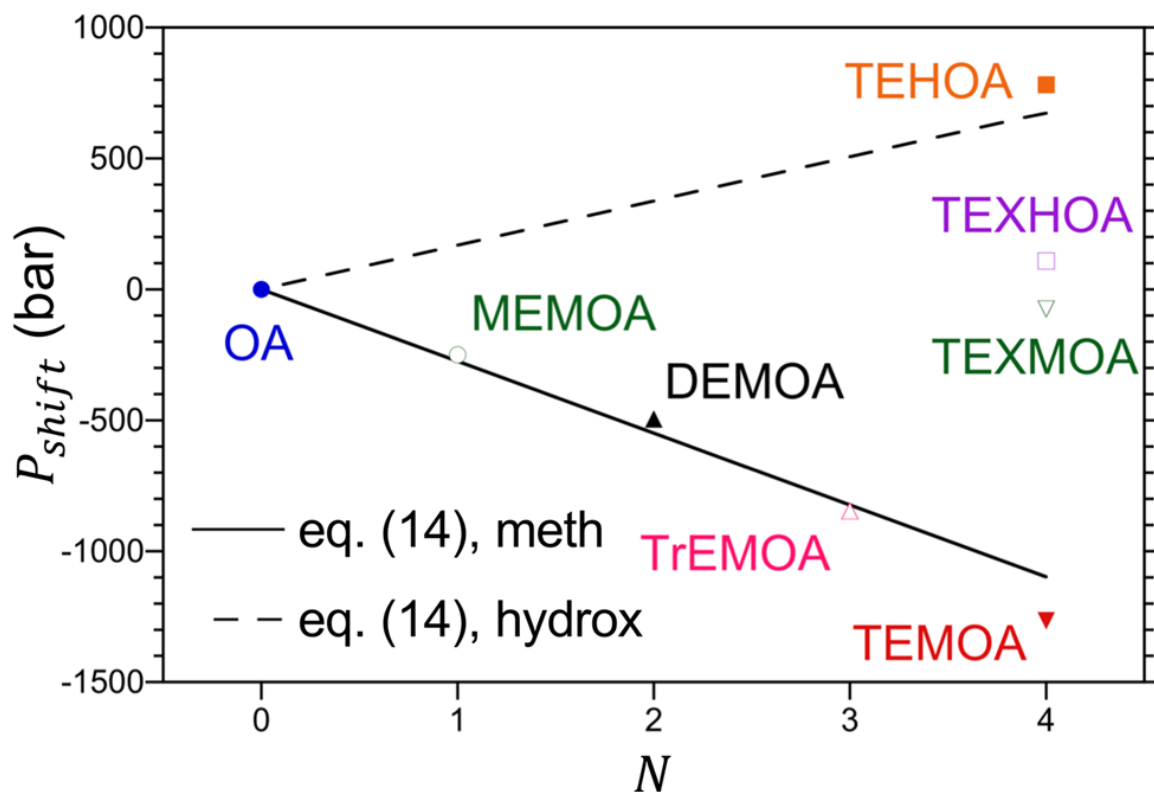
Figure 5.12: Shift pressures for the unified distribution model (eq. (5.7)) and capillary evaporation models for all the hosts as a function of the number of functional units N. The shift pressures are measured relative to the parent cavitand OA. The points indicate shift pressures for the unified distribution model (eq. (5.7)) obtained from fits to all the simulated hosts. The specific hosts are identified next to the points. The lines indicated shift pressures obtained for the capillary evaporation model (eq. (5.14)) for *endo*-methyl and *endo*-hydroxyl functionalization. The capillary evaporation model assumes the shift pressures of the *exo*-functionalized hosts are zero.
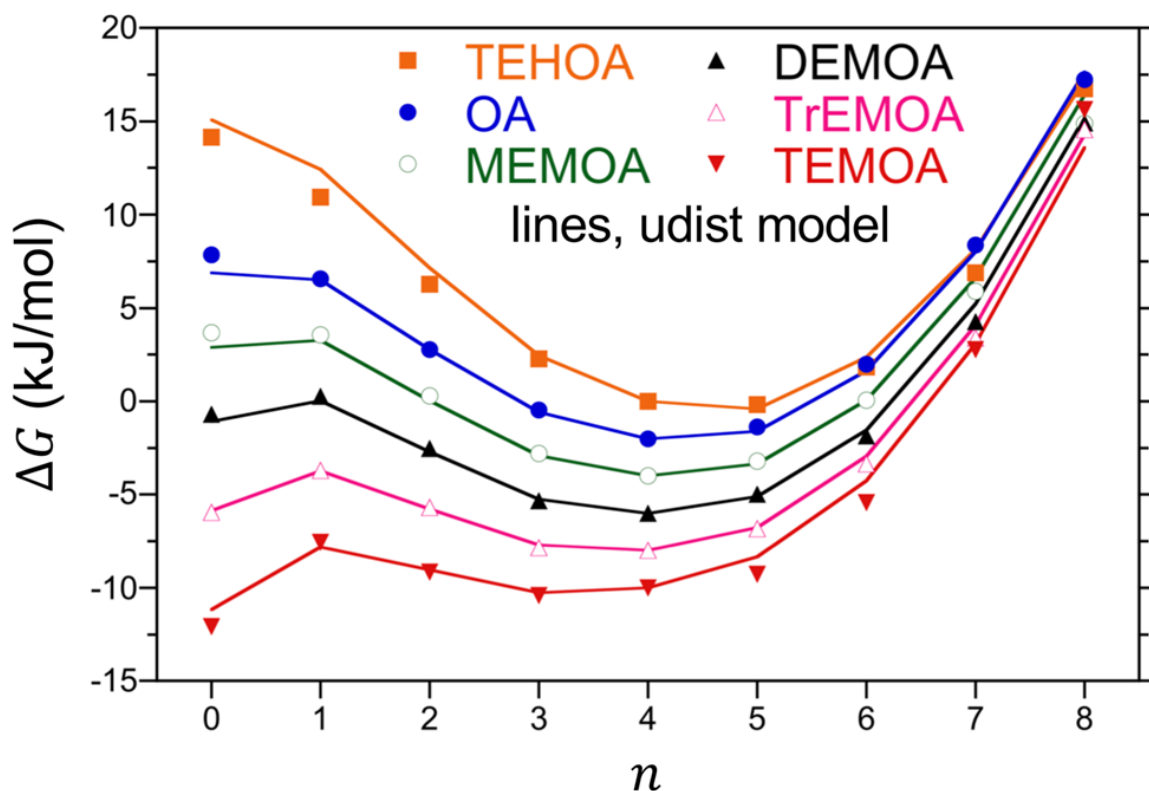
Figure 5.13: Impact of increasing portal hydrophobicity on relative pocket hydration free energies for the series of *endo*-functionalized hosts TEHOA, OA, MEMOA, DEMOA, TrEMOA, and TEMOA at 25°C and 1 bar. Points indicate simulation results (eq. (5.3)) and lines indicate fits to the unified distribution model (eq. (5.7)). The figure symbols are defined in the legend. Simulation error bars are comparable in size or smaller than the symbols. Results are successively shifted downward by 2 kJ/mol from TEHOA to TEMOA for clarity.

Capillary Evaporation Model of Cavitand Hydration. Here we describe a capillary evaporation model in an effort to rationalize the impact of portal functionalization on water absorption into cavitand pockets. This model is based on the idea that the system thermodynamics can be described using interfacial free energies associated with the cavitand surfaces contacting water or vacuum (water absence), which is clearly an approximation when applied to the molecular-level. Nevertheless, this model reproduces many of the salient elements of water absorption described above. Our physical picture of the cavitand in water underlying our capillary evaporation model is illustrated in Figure

5.14. Here, the parent cavitand, OA, is assumed to be a bowl-shaped pocket (indicated in gray) with an inner surface area of A_{pock}. The portal to the binding pocket (indicated in green) has a cross sectional area of A_{port} in the absence of any added endo-functional units (methyl or hydroxyl). Finally, each endo-functional unit (indicated in red) ringing the pocket blocks an area of A_{func} from the portal. The volume of the cavitand, v_{cav}, is defined by the region bonded by the pocket and portal. We neglect the impact of *exo*-functional units in this model, which were shown above to have a minimal impact on water absorption relative to OA.
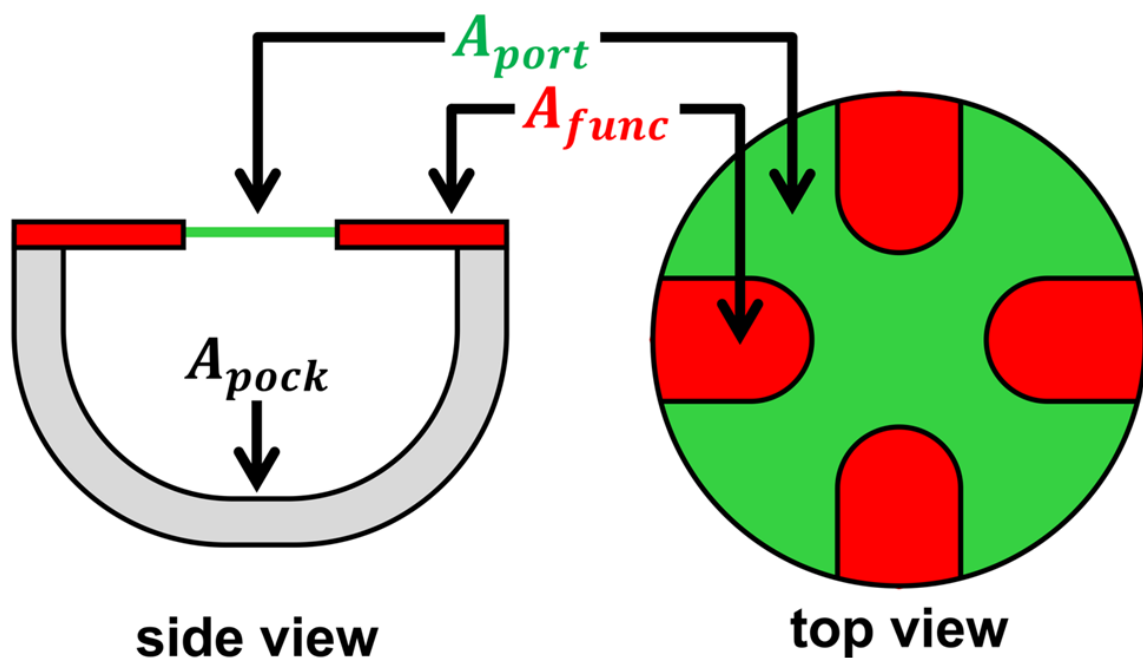
Figure 5.14: Schematic illustration of the *endo*-functionalized hosts used to develop the capillary evaporation model. The host pocket, portal, and functional units are identified in gray, green, and red, respectively, along with their corresponding areas, A_{pock}, A_{port}, and A_{func}. The cavitand volume, v_{cav}, is bounded by the pocket and portal surfaces.

The cavitand pocket is open to bulk water, permitting free exchange through the portal. In this case, water absorption equilibrium is described by the grand potential of the open system associated with the cavitand volume[133,134,149,150]. The grand potential of the dry pocket with N endo-functional units is

$$\Omega_{dry} = \gamma_{cv}A_{pock} + \gamma_{lv}(A_{port} - NA_{func}) + \gamma_{fv}NA_{func}, \tag{5.8}$$

where γ_{cv}, γ_{lv}, and γ_{fv} are the cavitand/vacuum, liquid water/vacuum, and functional unit/vacuum interfacial free energies, and N ($= 0$ to 4) is the number of endo functional units. The dry state here is assumed to be a vacuum rather than a vapor, as would normally be assumed for larger volumes. The pressure inside the dry pocket is zero (i.e., vacuum) and therefore does not contribute to eq. (5.8). The grand potential of the wet pocket is

$$\Omega_{wet} = \gamma_{cl}A_{pock} + \gamma_{fl}NA_{func} - Pv_{cav}, \tag{5.9}$$

where γ_{cl} and γ_{fl} are the cavitand/liquid water and functional unit/liquid water interfacial free energies, respectively, and P is the bulk pressure.

Given that the probability of observing the dry or wet state is proportional to the Boltzmann weighting of their respective grand potentials (i.e., $p_{dry/wet} \propto \exp(-\Omega_{dry/wet}/kT)$), the probability of observing the wet state is

$$p_{wet} = \frac{\exp(-\Omega_{wet}/kT)}{\exp(-\Omega_{wet}/kT) + \exp(-\Omega_{dry}/kT)}$$

$$= \frac{1}{1+\exp(\Delta\Omega/kT)}, \tag{5.10}$$

where

$$\Delta\Omega = \Omega_{wet} - \Omega_{dry}$$

$$= (\gamma_{fl} + \gamma_{lv} - \gamma_{fv})NA_{func} + [(\gamma_{cl} - \gamma_{cv})A_{pock} - \gamma_{lv}A_{port}] - Pv_{cav}. \tag{5.11}$$

Thus, based on this model water is expected to follow a Fermi function, characteristic of two-state thermodynamic equilibrium. Considering that v_{cav} is positive, eq. (5.10) predicts the cavitand is dry for low pressures (i.e., $p_{wet} = 0$ for low or even negative P), and wetting occurring with increasing pressure (i.e., $p_{wet} = 1$ for large P). This agree with our simplest expectations from above. The transition point between dry and wet states ($\Delta\Omega = 0$) occurs at the pressure

$$P_{trans} = \frac{(\gamma_{fl}+\gamma_{lv}-\gamma_{fv})NA_{func}+[(\gamma_{cl}-\gamma_{cv})A_{pock}-\gamma_{lv}A_{port}]}{v_{cav}}, \tag{5.12}$$

which may occur at positive or negative pressures depending on the degree of rim functionalization, and values of the associated interfacial free energies.

Assuming that v_{cav} is the same across all the hosts examined, the capillary evaporation model provides a direct rationalization of the collapse of the water absorption isotherms onto a universal curve and the shift pressures. Assuming OA ($N = 0$) as a reference host, the shift pressure can be determined by equating eq. (5.11) for a functionalized host with that of OA at a pressure of $P + P_{shift}$

$$\left(\gamma_{fl} + \gamma_{lv} - \gamma_{fv}\right)NA_{func} + \left[(\gamma_{cl} - \gamma_{cv})A_{pock} - \gamma_{lv}A_{port}\right] - Pv_{cav}$$
$$= \left[(\gamma_{cl} - \gamma_{cv})A_{pock} - \gamma_{lv}A_{port}\right] - (P + P_{shift})v_{cav}, \tag{5.13}$$

which yields

$$P_{shift} = -\frac{(\gamma_{fl}+\gamma_{lv}-\gamma_{fv})A_{func}}{v_{cav}}N. \tag{5.14}$$

Based on this model we expect the shift pressure to be a linear function of the degree of *endo*-functionalization, qualitatively similar to that reported in for the sequence OA to TEMOA in Figure 5.12. The negative shift pressures observed for methyl functionalization

can be explained within the context of eq. (5.14). Specifically, we expect for liquid water that γ_{fl} is positive, that water unfavorably wets hydrophobic functional units, and that γ_{lv} is greater than γ_{fv} given the lower surface tension of water compared to that of liquid alkanes. As a result, P_{shift} is anticipated to be negative for methyl unit functionalization, as observed in Figure 5.12. For the hydroxyl units to exhibit a positive shift pressure, we must have the combination of terms $\gamma_{fv} - \gamma_{fl}$ to be greater than γ_{fl}. This could result from a combination of wetting of the hydroxyl unit by water to be very favorable ($\gamma_{fl} < 0$) and exposure of the unit to vacuum being exceptionally unfavorable ($\gamma_{fv} > \gamma_{lv}$). A definitive rationalization for the positive shift pressures of the hydroxyl units, however, is not immediately apparent based on macroscopic considerations.

A shortcoming of applying the capillary evaporation model as described above is that eq. (5.10) only describes the probability of observing a dry or wet state, while the absorption isotherms reported above indicate continued filling of the cavitand with increasing pressure above the wetting transition. This suggests the water within the pocket is compressible. We subsequently model the mean wet state pocket water occupancy as

$$n_{wet} = n_0 \exp[\kappa_0 (P + P_{shift})], \qquad (5.15)$$

where n_0 is the mean wet state occupancy at zero pressure and κ_0 is the compressibility of water within the host pocket. While we assumed zero compressibility of the individual water occupancy states in eqs. (5.4) and (5.7) above, the hydration distribution models described above do capture the compression of waters within the pocket (e.g., Figures 5.7 and 5.15). This is a result of the fact that equilibrium tips towards states with lower volumes in response to increasing pressure (e.g., Figure 5.10), compressing waters within the pocket as described by eq. (5.15). The effective pressure of water inside the wetted pocket in this

expression is taken as $P + P_{shift}$. This is necessary for the model to achieve a universal collapse. In this case, the shift pressure can be interpreted as being analogous to the Laplace pressure, the pressure differential across a curved interface.

Given that the application of macroscopic interfacial free energies down to molecular-scale phenomena is questionable and that the individual interfacial free energies are not independent of one another in eq. (5.11), we lump them together as

$$\alpha = \left(\gamma_{fl} + \gamma_{lv} - \gamma_{fv}\right)A_{func} \tag{5.16a}$$

and

$$\beta = \left(\gamma_{cl} - \gamma_{cv}\right)A_{pock} - \gamma_{lv}A_{port}. \tag{5.16b}$$

The mean pocket occupancy for the capillary evaporation model is finally determined as

$$\langle n \rangle = p_{wet}n_{wet} + (1 - p_{wet})n_{dry} = p_{wet}n_{wet}$$

$$= \frac{n_0\exp[\kappa_0(P+P_{shift})]}{1+\exp\{[\beta-\left(P+P_{shift}\right)v_{cav}]/kT\}}, \tag{5.17}$$

where the mean dry state occupancy, n_{dry}, is zero. The shift pressure in here is determined as $P_{shift} = -\alpha N/v_{cav}$, where α adopts different values for methyl and hydroxyl functional units. In this expression, $\alpha, \beta, n_0, \kappa_0$, and v_{cav} are treated as fitting parameters. To connect the volume of the cavitand pocket to the molecular scale packing of water within its confines, we assume

$$v_{cav} = n_0/\rho_w, \tag{5.18}$$

where ρ_w here is the bulk number density of water $(1/\rho_w = 30 \text{ Å}^3)$ at atmospheric pressure.

In Figure 5.15 we compare the least squares fit of the capillary evaporation model (fit parameters reported in Table 5.1) as a function of $P + P_{shift}$ against the simulation results for water absorption in the endo-functionalized hosts: OA, MEMOA, DEMOA, TrEMOA, TEMOA, and TEHOA. Overall, the model provides a near quantitative description of water absorption as a function of pressure, supporting the assumption water absorption within hosts is well described as a two-state equilibrium. Moreover, we find semi-quantitative agreement between P_{shift}'s obtained from the capillary evaporation model against those obtained by fitting eq. (5.7) (Figure 5.12). Consideration of specific parameters reported in Table 5.1 further supports the reasonableness of the proposed capillary evaporation model. Given the fact that water loses hydrogen-bonding partners when it enters the host pocket, it is sensible to expect adsorbed waters to be more compressible. This is born out comparing the fitted compressibility of water in the pocket against that for bulk TIP4P/Ew water, 4.75×10^{-5} bar^{-1}, that is the fitted compressibility is 50% greater than the bulk compressibility. For a wetted OA pocket at atmospheric pressure with $n_0 = 4.269$ waters, the universal binding model described in the previous section anticipates a pocket volume (e.g., Figure 5.10) of $\Delta\Delta v^*(0, n_0) = \Delta v^*(0) - \Delta v^*(4.269) = 82.2$ cm3/mol $= 136$ Å^3 ($\Delta v^*(4.269)$ was determined by interpolation between $\Delta v^*(4)$ and $\Delta v^*(5)$). This volume is in good agreement with that obtained by fitting v_{cav}, differing only by ~7%. It may be thought, however, that the assumption that the cavitand volume is described by eq. (5.18) is erroneous. If we relax this assumption and make v_{cav} itself a fitting parameter we obtain a best fit value of 128.9 Å^3, differing by less than 1%. Taken together these observations support our proposition that drying of supramolecular host pockets is driven by capillary evaporation.
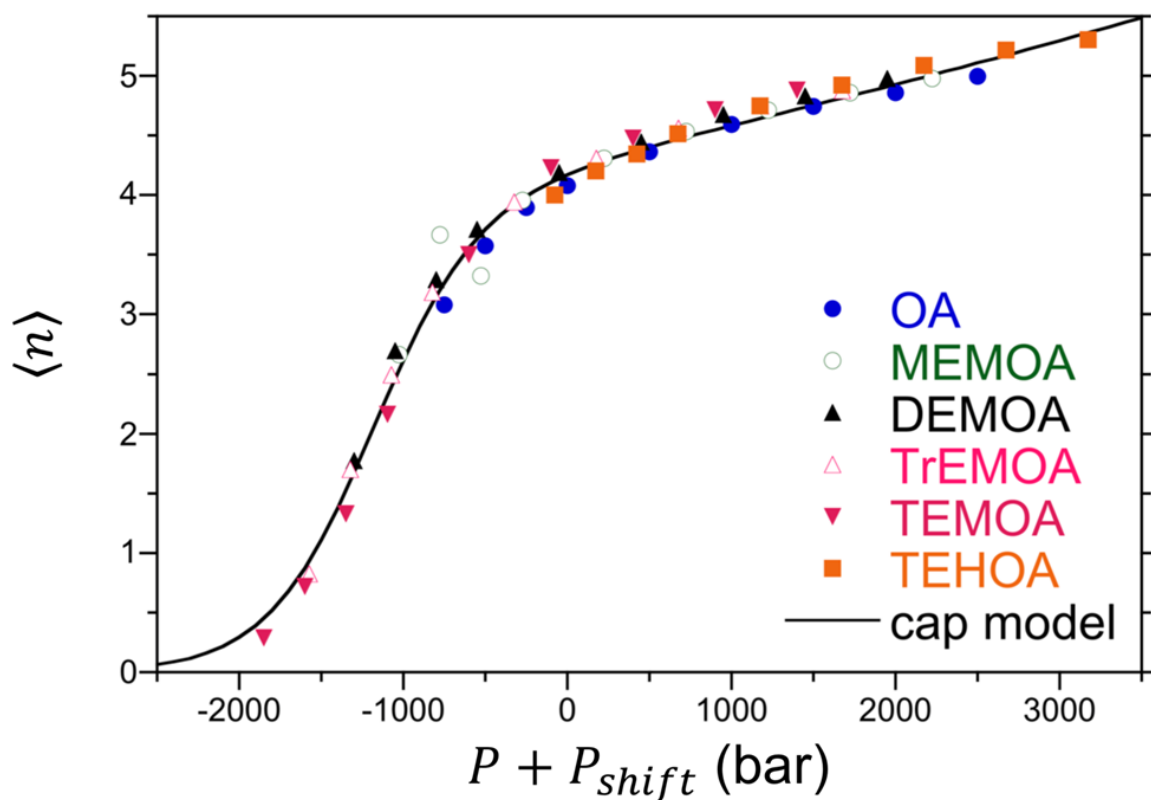
Figure 5.15: Comparison between the mean pocket hydration numbers as a function of the effective pocket pressure $(P + P_{shift})$ for the endo-functionalized hosts determined from simulation and the capillary evaporation model (eq. (5.17)) at 25°C. The figure symbols are defined in the legend. Error bars are comparable in size or smaller than the symbols.

Table 5.1: Capillary evaporation model parameters obtained by least squares fitting of simulation results for OA, MEMOA, DEMOA, TriEMOA, TEMOA, and TEHOA.

parameter	value
n_0	4.269
κ_0	7.185×10^{-5} bar^{-1}
$v_{cav} = n_0/\rho_w$	128.1 Å^3
$-\alpha_{meth}/v_{cav}$	-274.4 bar
$-\alpha_{hydrox}/v_{cav}$	168.2 bar
β	-9.341 kJ/mol

5.5   Conclusion

In conclusion, we have reported a molecular simulation and theoretical analysis of the absorption of water into the hydrophobic pockets of deep-cavity cavitands in solution over a broad range of pressures. The portals of these cavitands were functionalized with hydrophobic (methyl) and hydrophilic (hydroxyl) units, that adopted either inwards (*endo*) or upwards (*exo*) orientations with respect to the pocket. Our simulations found that wetting of the pocket, as captured by the mean water occupancy, is directly controlled by hydrophobicity of the portal, with more hydrophobic units tilting equilibrium towards lower pocket occupancies and hydrophilic units tilting toward higher pocket occupancies. Directing the functional units in an *exo* (upwards) orientation from the pocket significantly reduces their impact on pocket occupancy, indicating the effect on pocket hydration is local.

From a thermodynamic perspective, water absorption within host pockets appears to be governed by equilibrium between dry (empty) and wet (filled) states, separated by a free energy barrier associated with the first water entering the host pocket. Filling of the pocket is driven by lowering of the host volume with increasing pressure, in accordance with Le Chatlier's principle. Considering the pocket occupancy distributions, our analysis demonstrated that water absorption across all hosts could be collapsed onto a unified absorption isotherm, in which the absorption isotherms of chemically distinct cavitands could be mapped onto one another via a shift pressure. The shift pressure is positive for hydrophilic functional units and progressively negative for hydrophobic functional units. Finally, we demonstrated that the absorption equilibrium could be described by a two-state capillary evaporation model, which describes equilibrium as resulting from a balance of interfacial free energies for wetting/drying of the internal surfaces of the cavitand and the bulk system pressure. This model captures many of the salient features observed from our simulations, including the collapse of the absorption isotherms for a range of cavitands onto a unified isotherm and the signs of the shift pressures for the hydrophobic and hydrophilic functional units.

This work demonstrates that the wetting behavior of a supramolecular host pocket can be directly tuned by functionalizing the portal to the binding site. In particular, the pocket can be dewetted using hydrophobic functional units or more strongly wetted using hydrophilic functional units. This, in turn, is sure to impact the binding of guests by the host since evacuation of the pocket is a necessary step for the guest to settle into the binding site.

Chapter6
Characterization of wetting and dewetting behaviors within deep-cavity cavitand pockets using Indirect Umbrella Sampling (INDUS)

6.1 Summary

The wetting and dewetting behaviors are vital factors controlling the thermodynamics and driving forces of the host-guest binding and assembly behaviors led by hydrophobic effect. These behaviors can be affected by the hydrophobicity distribution on the surface and within the binding sites of host molecules. Here we use molecular dynamics simulation and advanced sampling techniques (INDUS) to study the wetting behaviors and hydrophobicity distribution within pockets of different deep-cavity cavitand host molecules. From the cylindrical water density distributions, we find waters in hydrophobic pockets responded differently to gradually increasing biasing potentials with different functionalization on the rim of cavitand. These different responses were further verified by the profile of water numbers in the whole observation volume as a function of biasing potential strengths. In addition, we demonstrated that individual locations within the pockets exhibit differences in hydrophobicity. This difference was reflected in different levels of ease to displace water molecules. Moreover, this ease of expelling waters corresponds to the hydrogen bonding distribution patterns of individual water within pockets. This work demonstrated that the hydration of supramolecular host pockets is

complicated even with minor structural modification. And INDUS is efficient in serving as probes for these sophisticated behaviors.

6.2   Introduction

Water plays a nonnegligible role in the molecular binding and surface behaviors, especially in the form of hydrophobic effect. This significance is being amplified for biomolecules such as protein. Hydrophobicity is a dominating driving force for protein folding[151]. The hydrophobic side chains of protein can be guided by the polar water environment and form packed hydrophobic core[152]. The long-range water-mediated force and hydration structure were demonstrated to determine the protein aggregation propensity. And it leads to different solubilities for positively and negatively charged protein residues[153]. Not only restricted to be near surface, water molecules can enter the confined regions and cavities of proteins. The release of trapped water was shown to be a vital step in the formation of amyloid fibrils, which is associated with a variety of neurodegenerative diseases[154]. The solvation of protein cavities will induce entropic and enthalpic penalties and therefore stabilize the protein-ligand complexes[155]. A recent study of hydrogen bonding structures revealed hydration within the pocket of bovine trypsin with a lack of hydrogen bonds. The ligand binding was also fueled by this phenomenon[156]. As mentioned in previous chapters, self-dewetting also occurred within protein pockets. Under ambient pressure, both the pocket of the mutant L99A of lysozyme and the binding cavity of bovine β-lactoglobulin were found to be mostly dry[49,50]. It will place significant influence on the protein-ligand binding mechanisms.

Therefore, quantifying the local hydrophobicity near molecular surface and inside binding cavities can provide us a detailed picture of molecular scale interactions and how waters respond to a specific environment. Observed in both experiments and computational studies, enhancement of water density fluctuations and cavity formation were reported near hydrophobic surfaces[95,132,157,158]. Based on the quantitative correlation between the water contact angle on a particular surface and the cavity formation free energy near it[159], Patel developed a method to evaluate the hydration free energies, μ_v^{ex}, of cavities of arbitrary shapes and sizes[160]. By applying an unfavorable biasing potential proportional to the number of water molecules in the observation volume v, ϕN_v, the waters can be displaced from the probe volume. The free energy of removing all the waters from an observation volume, v, just equals to the free energy needed to create a cavity with the same shape and size. Therefore, the hydrophobicity in a certain location near a hydrophobic surface can be characterized by the quantitative relationship relating the hydration free energy, μ_v^{ex}, to the average number of water molecules in the volume, $\langle N_v \rangle_\phi$, as a function of ϕ. The brief derivation is from Patel's work[160] and described here.

When an unfavorable potential coupled linearly to number of waters in v is applied, the Hamiltonian of the whole system became:

$$H_\phi = H_0 + \phi N_v \tag{6.1}$$

Here H_0 is the Hamiltonian for a normal system without any extra biasing potentials. If the partition function associating with H_ϕ is Q_ϕ. Then Q_ϕ will be:

$$Q_\phi = \sum_{N_v} \exp(-\beta(H_0 + \phi N_v)) \tag{6.2}$$

Based on this, the first derivative of $\ln Q_\phi$ with respect to $-\beta\phi$ is the mean of N_v.

Likewise, the second derivative is the variance of N_v.

$$\frac{\partial \ln Q_\phi}{\partial(-\beta\phi)} = \frac{1}{Q_\phi}\sum_{N_v}\exp(-\beta H_0 - \beta\ \phi N_v) * N_v = \langle N_v\rangle_\phi \tag{6.3}$$

$$\frac{\partial^2 \ln Q_\phi}{\partial(-\beta\phi)^2} = \frac{\partial\langle N_v\rangle_\phi}{\partial(-\beta\phi)} = \langle N_v^2\rangle_\phi - \langle N_v\rangle_\phi^{\ 2} \tag{6.4}$$

If integrating equation 6.3, you can have:

$$\int_0^\phi \ln Q_{\phi'} = -\int_0^\phi \langle N_v\rangle_{\phi'} * \beta\ d\phi' \tag{6.5}$$

$$\frac{Q_\phi}{Q_0} = \exp(-\beta\int_0^\phi\langle N_v\rangle_{\phi'}\ d\phi') \tag{6.6}$$

Besides, the expression on the left, Q_ϕ/Q_0, is related to the probability, $P_v(N)$, of

observing N water molecules in observation volume, v, without biasing potential, ϕ.

$$\frac{Q_\phi}{Q_0} = \frac{\sum_{N_v}\exp(-\beta(H_0+\phi N_v))}{\sum_{N_v}\exp(-\beta H_0)} = \langle\exp(-\beta\phi N_v)\rangle_0 = \sum_N P_v(N) * \exp(-\beta\phi N) \tag{6.7}$$

By defining $z = \exp(-\beta\phi)$, equation 6.7 can be simplified as:

$$\frac{Q_\phi}{Q_0} = \sum_N P_v(N) * z^N = P_v(0) * z^0 + P_v(1) * z^1 + \cdots + P_v(N) * z^N \tag{6.8}$$

If $\phi \to \infty, z \to 0$:

$$P_v(0) = \frac{Q_\phi}{Q_0}\Big|_{z\to0}\ ,\ P_v(N) = \frac{1}{N!}\frac{\partial^N}{\partial z^N}\left(\frac{Q_\phi}{Q_0}\right)\Big|_{z\to0} \tag{6.9}$$

If we consider equation 6.6 and 6.9, together with $\beta\mu_v^{ex} = -\ln P_v(0)$[161]:

$$\exp\left(-\beta\int_0^\infty\langle N_v\rangle_\phi\ d\phi\right) = P_v(0)\ ,\ \int_0^\infty\langle N_v\rangle_\phi\ d\phi = \mu_v^{ex} \tag{6.10}$$

The final expression in equation 6.10 is the central result in this derivation. It is also the most important equation applied in this Indirect Umbrella Sampling method (INDUS). Basically, during this process, biasing potentials with different ϕ values were applied to the observation volume. When ϕ is larger, the number of waters within this volume is less. By applying a range of ϕ values until ϕ is very large or there is no water in the observation volume, we can simulate the process of creating a cavity in solvent. And the free energy, μ_v^{ex}, required to create this cavity can be evaluated by doing this integral over the whole range of ϕ values from zero. In addition, with any value of ϕ, we can characterize how water molecules respond to biasing potentials with different strengths.

In this chapter, we will apply INDUS method to two hosts in the family of deep-cavity cavitand molecules, octa-acid (OA) and tetra-endo-methyl octa-acid (TEMOA), which have well-defined binding sites to mimic protein cavities. Their structures are illustrated in figure 6.1 and their differences are highlighted in figure 6.2. As introduced in detail in previous chapters, they are both supramolecular host molecules possessing hydrophobic pockets in the center and formed by three rows of aromatic pockets[56,67]. This pocket is roughly 8-9 Å in diameter and 8 Å in deep. Eight carboxylic groups are decorating the rim and the feet to assure their water solubilities. Distinct assembly complexes can be formed by encapsulating different sized guests within the non-polar environment of the pockets[56,57,63,64,66,68,69]. As described in Chapter 5 and in Appendix F, water absorption inside the pockets can be altered by the methylation on *endo*-positions of the rim. And it places significant influence on the host-guest binding behaviors. Here, the local hydrophobicity around the cavitands and inside the binding pockets can be investigated by creating cavities centered on heavy atoms of cavitand. The water density and hydrogen

bonding distributions were evaluated to illustrate this water-expelling process and measure the ease to displace water molecules in certain locations. Moreover, the susceptibility of hydration numbers in the whole observation volume of both cavitands were also calculated. And we showed their profiles matched the universal shape of that of many proteins[162].

1: R = H (OA)
2: R = Me (TEMOA)
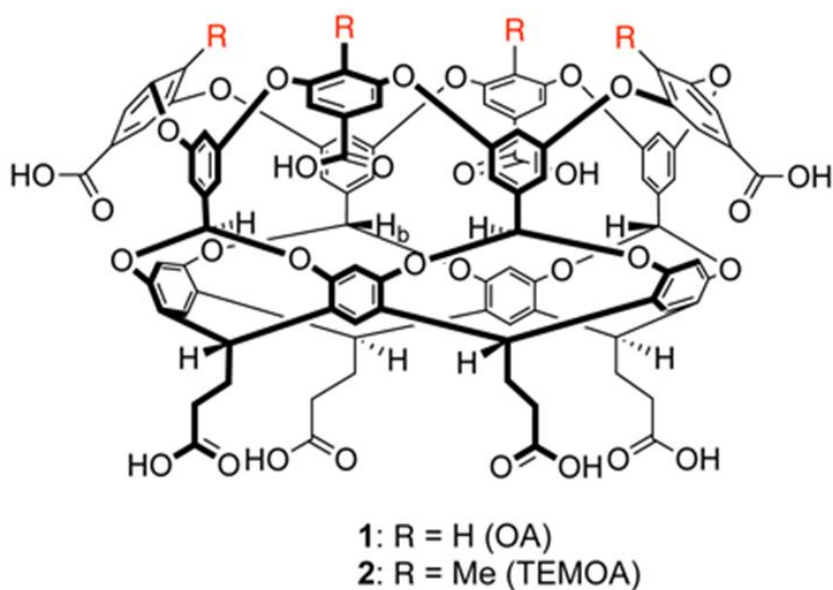
Figure 6.1: structures of octa-acid (OA), tetra-endo-methyl octa-acid (TEMOA).
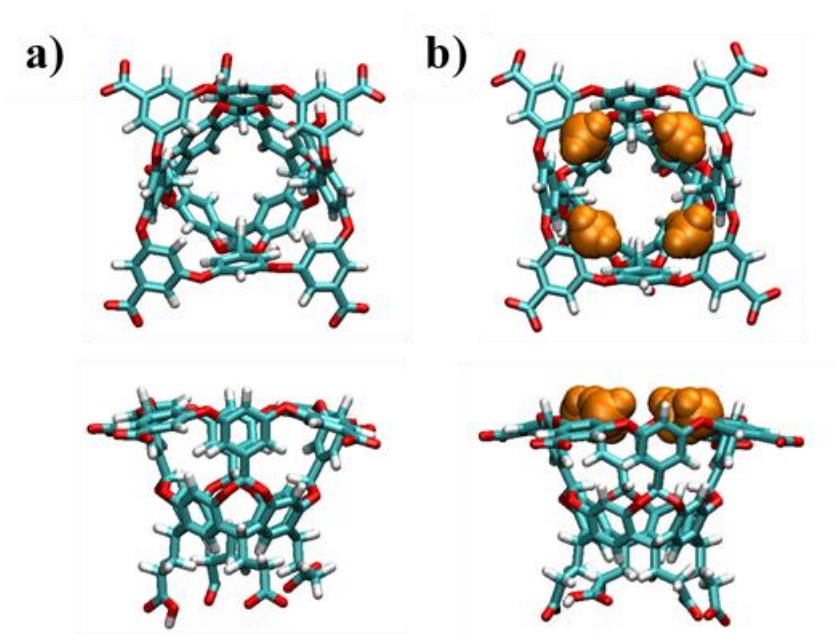
Figure 6.2: Molecular snapshots of octa-acid (OA) (a), and, tetra-endo-methyl octa-acid (TEMOA) (b). The body of the parent cavitand is illustrated in licorice format, while the methyl units are highlighted using a van der Waals representation. Top views and side views are both shown.

6.3 Method

All the molecular dynamics simulation with Indirect Umbrella Sampling method (INDUS)[163] were conducted by a properly modified package of GROMACS 2016[79]. For each simulation, one single cavitand was solvated in 2994 water molecules. The cavitand molecule was modeled using the Generalized Amber Force Field (GAFF)[70] with partial charges obtained from the AM1-BCC calculations[106]. To be consistent with the protonation state at pH 7[107], the total net charge of a single cavitand was set to be -6e. The four benzoic acid groups ringing the rim of the cavitand and two of the four benzoic acid groups on the feet were deprotonated to get the net charge state. For each single cavitand, six sodium cations were put into the system to make the whole simulation box neutral. Those sodium

cations were also modeled using GAFF[70]. TIP4P-EW potential[108] was used to model the water molecules. Nonbonded Lennard-Jones interactions were truncated using a cutoff of 9 Å. The short-range electrostatic interactions were also truncated with a real space cutoff of 9 Å. While the long-range electrostatic interactions were treated by the particle mesh Ewald Summation method[71]. All the simulations were performed in the isothermal-isobaric ensemble at 298.15 K and 1 bar. The temperatures were coupled using the Nosé-Hoover thermostat[80,81], while the pressures were controlled using Parrinello-Rahman barostat[82]. The equations of motion were integrated with a time step of 2 fs. The oxygen-hydrogen bonds in water molecules were held rigid using SETTLE[87], while all the other bonds involving hydrogens were constrained using the LINCS algorithm[86].

To apply the unfavorable biasing potentials, all the atoms of cavitand molecule was held fixed by a harmonic restraint with a spring constant of 1000 kJ/(mol nm^2) in three dimensions. Therefore, the cavitand was restrained in the center of the simulation box. Centered on each heavy atom (carbon and oxygen) of cavitand, a biasing potential was applied to a spherical observation volume and bias the number of coarse-grained water molecules[160,163,164]. To use a spherical observation volume, the inner shell radius, RLOWs, and the outer shell radius, RHIGHs, were set to be -0.5 nm and 0.6 nm, respectively. Therefore, the effective cavity radii for both oxygen and nitrogen atoms were 0.6 nm. With the presence of this biasing potential, the Hamiltonian regulating the system is: $\mathcal{H}_\phi = \mathcal{H}_0 + U_\phi(\widetilde{N_v})$, where $\mathcal{H}_0$ is the Hamiltonian of the unbiased system, $\widetilde{N_v}$ is the coarse-grained water number, and U_ϕ is the biasing potential parameterized by ϕ, which regulated the strength of the biasing potential[164]. The Gaussian coarse-graining function with a standard deviation of $\sigma = 0.01$ nm and a truncation cutoff of $r_c = 0.02$ nm was used.

Specifically, the biasing potential is in a combination form of a harmonic part and a linear part[157,160]:

$$U_\phi\big(\widetilde{N_v}\big) = \frac{\kappa}{2}(\widetilde{N_v} - N^*)^2 + \phi\widetilde{N_v} \qquad (6.11)$$

Here since each observation volume is small, we set $\kappa = 0$. It became a simple linear type potential. The parameter ϕ denotes the strength of this potential. A series of windows with increasing ϕ values were simulated. The ϕ values ranges from 0 to 10 kJ/mol with a 1 kJ/mol increment. In the case of OA, some inner states with ϕ from 2.1 kJ/mol to 3.9 kJ/mol with an increment of 0.1 kJ/mol were also simulated. While for TEMOA, biasing simulations with ϕ from 0.1 kJ/mol to 1.9 kJ/mol with the same finer increment of 0.1 kJ/mol were conducted. All the biasing simulations were equilibrated for 1 ns followed by 10 ns production run. The number of waters, N_v, and coarse-grained water numbers, $\widetilde{N_v}$, observed in volume v as a function of simulation time were recorded for each window. The susceptibility $\chi_v = -\frac{\partial <N_v>_\phi}{\partial(\beta\phi)} = < \delta N_v^2 >_\phi$ can also be obtained from those data. The probability of observing N water molecules in volume v, $P_v(N)$, can be further evaluated by WHAM algorithm[109].

The wetting behavior around the cavitand and inside the hydrophobic pocket under the impact of the biasing potentials were characterized via several different techniques. First of all, the water density about the cavitand was evaluated in a cylindrical symmetric coordinate by assuming their rotational symmetry about their C_4-axes. The water density can be accounted within bins of this coordinate and averaged over the azimuthal angle q. In a similar way, the number of hydrogen bonds of each water molecule formed with other water molecules can also be determined and we applied this cylindrical coordinate system

to cavitand to bin hydrogen bonding numbers about the cavitand. Specifically, for each water molecule, the number of hydrogen bonds it formed as an acceptor or a donor with other water molecules can be determined separately. We used a geometrical criterion to determine if a hydrogen bond is formed. When the donor-acceptor distance is smaller than or equal to 0.3 nm, and the hydrogen-donor-acceptor angle is smaller than or equal to 30°, a hydrogen bond is believed to exist. And if this donor or acceptor belongs to a certain water molecule, the number of its hydrogen bonds as a donor or acceptor will correspondingly add one.

In the same cylindrical symmetric coordinate, we can also measure the ease of expelling the water from the observation volume with the biasing potential in different regions. In a certain window, if the water density in a bin was firstly observed to drop below 5% of its initial level without any biasing potential, the bin was defined to be "empty" in this window. Therefore, each bin was labeled as the window number where it was firstly observed to be "empty", which indicated the strength of the biasing potential needed to expel water from that region.

6.4   Result and Discussion

In figure 6.3, the average water density around and inside OA in the presence of increasing biasing potentials on each heavy atom (oxygen and carbon) of cavitand are illustrated. The number in the window label indicates the strength of the linear biasing potential (i.e., the ϕ value). Window 1 means $\phi = 1$ kJ/mol, and Window 5 means $\phi = 5$ kJ/mol, etc. The water densities were cylindrically averaged about the cavitand's C_4-axis of symmetry and

the azimuthal angle θ. In window 0, which means no extra repulsive biasing potential was applied, it is a normal simulation of a single OA solvated in water, except that the cavitand was manually restrained in the center of the simulation box. The big black U-shape region indicates where the cavitand was placed. It is the location where all the heavy atoms of cavitand occupied and therefore the water density is zero. It can be clearly observed that approximately three water solvation shells were formed around the cavitand indicated by the high-density blue regions ($\rho > 1.2$ g/cm^3). The most prominent one is that nearest blue layer wrapping the cavitand from the hydrophobic rim to the feet of cavitand at bottom. It is also not surprising to see water molecules, as polar solvents, prefer to stay near the deprotonated carboxylic acid groups on the rim and feet. Near the junction area of the first and the second layers of aromatic rings, where there are eight oxygen atoms serving as the jointing points, waters are also observed to sticking there. Water enters the inner pocket and prefers to adhere to the inside walls of the cavitand. In contrast, there is a depletion of water in the upper middle part of the pocket denoted by black color. In the previous chapter, this was attributed to the favorable van der Waals interactions between waters and cavitand inner walls attracting water molecules to stay near the walls.
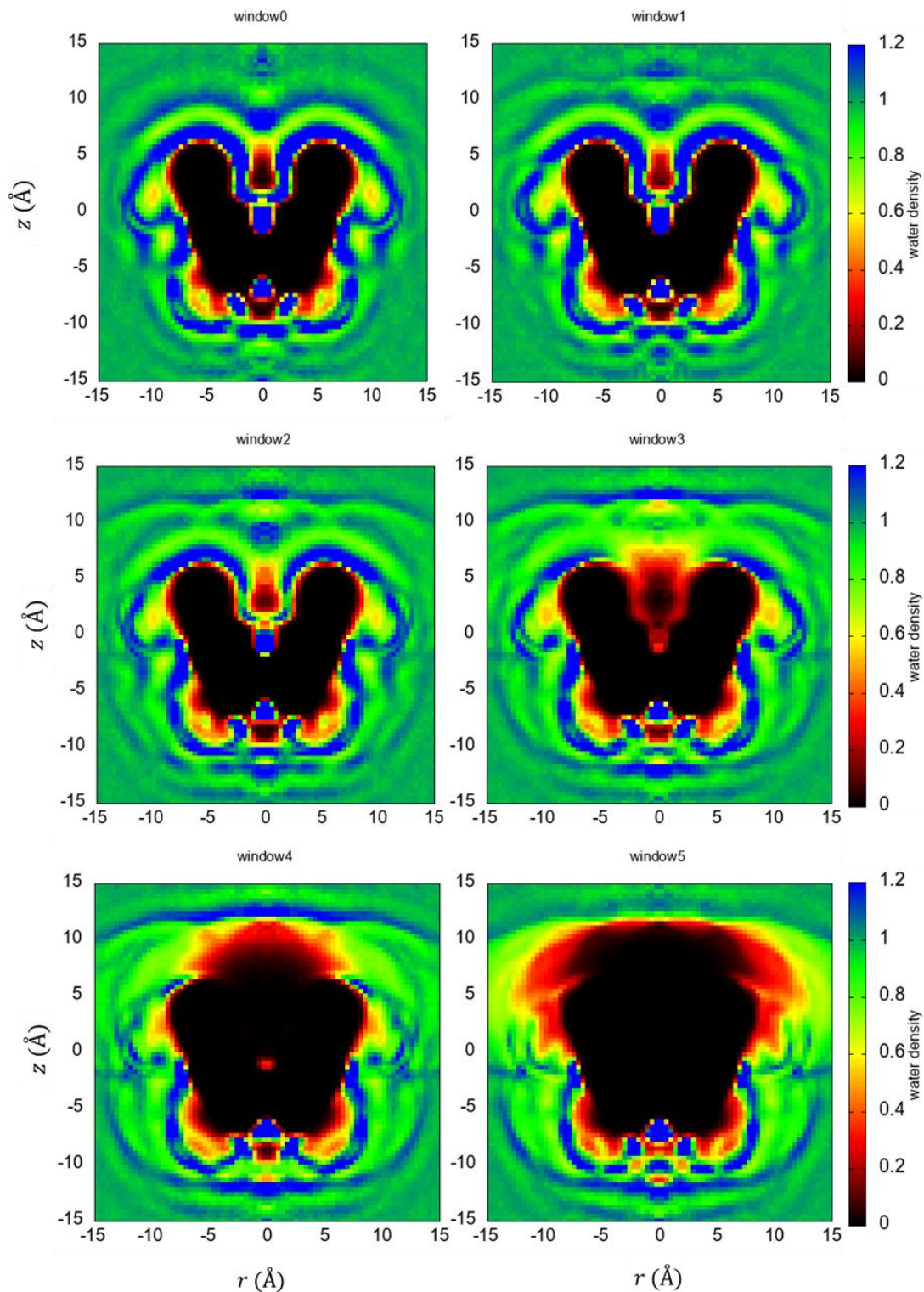
Figure 6.3: Average water density distributions about OA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.

At the bottom of the hydrophobic pocket, there is a region showing a significant high density of water. This region can accommodate one water molecule, and it can easily form hydrogen bonds with the four hydrogens on the waist of cavitand pointing inwardly (Figure 6.1 and Figure 6.2). In window 1, the water density distribution doesn't show significantly different picture compared to that of window 0. This figures the biasing potentials of $\phi = 1$ kJ/mol were not strong enough to expel water out of the pocket. When ϕ was increased to 2 kJ/mol, a clear depletion of water near the inner side walls can be observed. Those blue regions were partitioned and became significantly narrower. When ϕ reached 3 kJ/mol, the turning point appeared. Almost all the water molecules inside the pocket, no matter which portion they originally stayed, were removed. A large region that is void of water and in black color emerged in the center of the pocket surrounded by wide red regions. Occasionally, there were water molecules remain inside, but the water density is already very thin. Moreover, the first and second solvation shells were also affected and became blurred. In window 4 and window 5, under the impact of very strong biasing potentials, the black void region spread like a volcanic eruption. No water molecules can be found inside the pocket or near the top of cavitand. The first layer of solvation shell disappeared, and the second layer was pushed far away from the cavitand and almost coincided with the third solvation shell. In contrast, the water near bottom of the cavitand were not largely affected. That might be because there is a denser distribution of heavy atoms on upper part of the cavitand. Thus, with the same ϕ value, the area near the upper part underwent stronger biasing potentials synthetically. But it can also be attributed to the strong hydrophobicity on rim and the second layer of aromatic rings.

By more windows with biasing potentials with ϕ from 2.0 kJ/mol to 3.9 kJ/mol in 0.1 kJ/mol increments, the water activity profile under the impact of external repulsive potentials can be more carefully observed. As shown in figure 6.4, some representative windows were selected, and it is also easier for us to compare the abilities of different locations inside the pocket to accommodate and adsorb water molecules. Initially, the water molecules near the inner walls left the pocket very early. Though they can still form favorable van der Waals interactions with cavitand atoms and other water molecules in neighborhood, once their neighbors were expelled out of the pocket, this interaction will get broken. Especially for the water molecule near the cavitand portal who may serve as the connector between the encapsulated water molecules and the external water network, they stayed near the hydrophobic rim, and with additional biasing potentials they can easily get out into the bulk. As a result, this connection will be cut off in a relative early stage. In contrast, you can still find a significant water density at the bottom of cavitand when $\phi =$ 2.6 or 2.7 kJ/mol. The water density was even not zero when ϕ was as high as 3.1 kJ/mol. As mentioned before, the stabilization of water molecules at the bottom is mainly ascribed to the strong hydrogen-bonding with the inner pointing four hydrogens. The result here also indicates that the water molecule at the bottom is isolated. Its stabilization doesn't quite depend on the interactions with other water molecules inside.
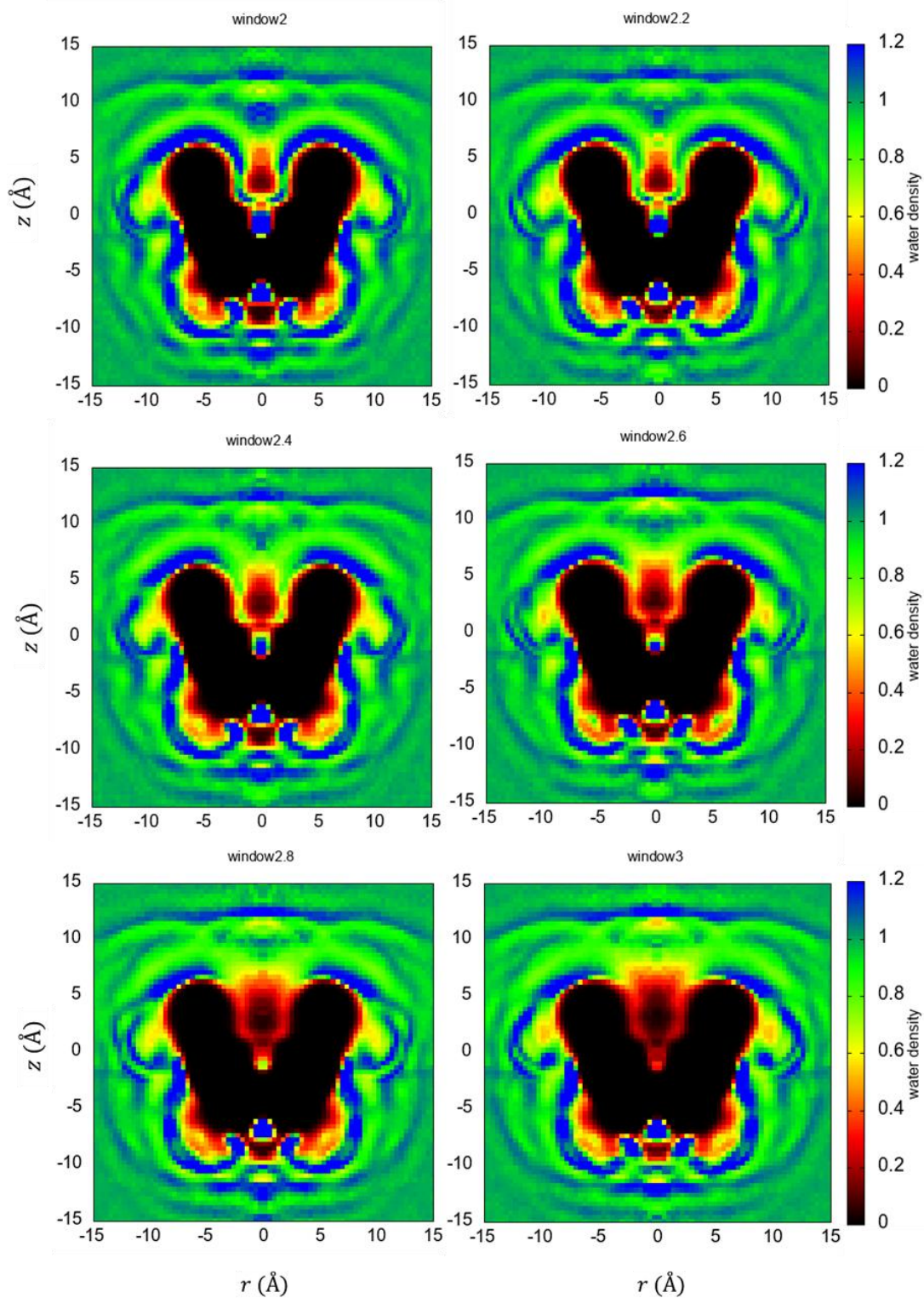
Figure 6.4: Average water density distributions about OA in presence of biasing potentials on each heavy atom of the cavitand with increasing strength in finer increments. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.

As observed in previous chapters and in figure 6.5, when $\phi = 0$, (i.e., no extra biasing potential), the water density inside the inner pocket of TEMOA is significantly lower. The waters encapsulated are more isolated, and there is an interface across the cavitand portal splitting the dry and wet regions. There is still a region of depleted water density along the centerline of the pocket. But relatively, this region is deeper inside the pocket in the case of TEMOA, and correspondingly, the water density in the upper pocket is higher due to more van del Waals interactions with the four additional methyl groups on the rim. The distribution of water density outside TEMOA pocket is very similar with that in the case of OA.

Once the biasing potential was in effect, the inner pocket of TEMOA was evacuated with even very small potential strength, $\phi = 1$ kJ/mol. It is nature to observe this rapid evacuation when combined with TEMOA's tendency to dry itself under ambient pressure and having less hydration number within the pocket. The more stable waters at bottom of the pocket were also removed when ϕ reached 2 kJ/mol. With the four additional methyl units on *endo* positions of the rim, the evacuation process of the waters in inner pocket is actually still very similar with that in OA, but with a few windows in advance. However, the water molecules outside and near the top of TEMOA were not easier to get expelled from the observation volume. When ϕ was increased from 2 kJ/mol to 3 kJ/mol, only the water density right above the pocket had a significant decrease. The speed of this evacuation had noticeably slowed down. When ϕ was further elevated to 5 kJ/mol, the water density distribution was almost identical to that of OA with the same biasing potential strength. Thus, the methylation of *endo* positions on the rim of cavitand significantly influenced the wetting behaviors inside the pocket and near portal, but the

solvation shell around the cavitand was not affected. Likewise, even with a large biasing potential, the water shells near the bottom of TEMOA were not largely impacted. The firm interactions with carboxylic acid groups can help stabilize those waters.

After investigating more windows, as some of them are shown in figure 6.6, with ϕ values between 0 kJ/mol and 2.0 kJ/mol and finer increments, we found that the waters staying near the inner-walls were removed when ϕ = 0.4 kJ/mol. In window 0.4, The water depletion region occupied the most portion of the inner space, indicated by a huge black and red area along the centerline. This water depletion region quickly spread to the whole pocket with further increase of ϕ. As what we observed in OA, the waters staying at bottom have much stronger resistance to the biasing potentials. Though the blue region rich in water was also shrinking with the increase of biasing potential strengths, it can still be detected when ϕ = 1.2 kJ/mol. In fact, even when ϕ = 1.9 kJ/mol (shown in Appendix D), a small blue region can still be observed in the same location at bottom. Through all the windows, some back and forth phenomenon, i.e., water density unexpectedly increased with the strengthening of biasing potential compared to previous windows, can be noticed in small local domain. This can be partially ascribed to the small increment of ϕ, which is only 0.1 kJ/mol. Not all the water densities in every single location have the same sensitivity to a small perturbation of the biasing potential, but the general trend followed as what we expected.
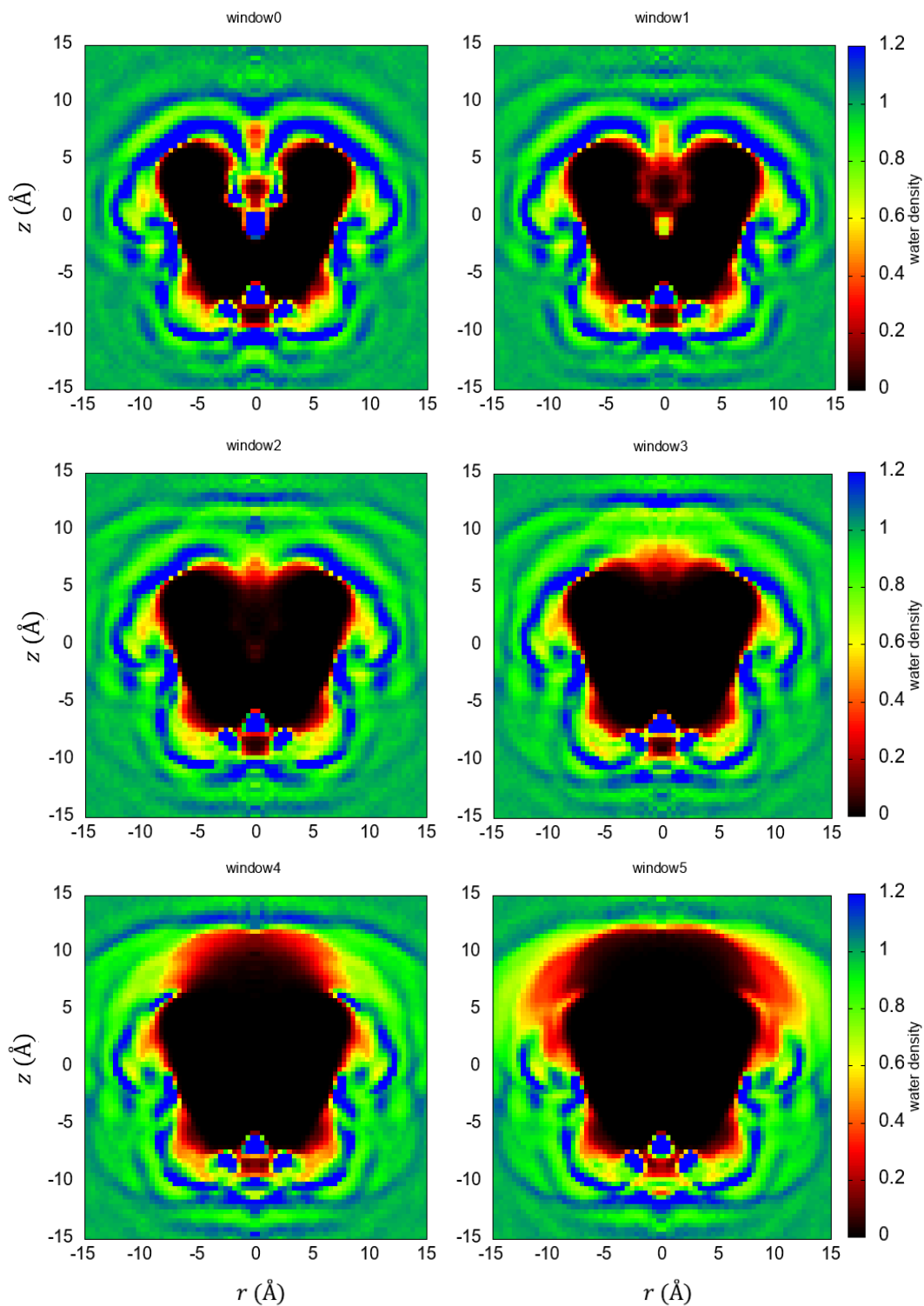
Figure 6.5: Average water density distributions about TEMOA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
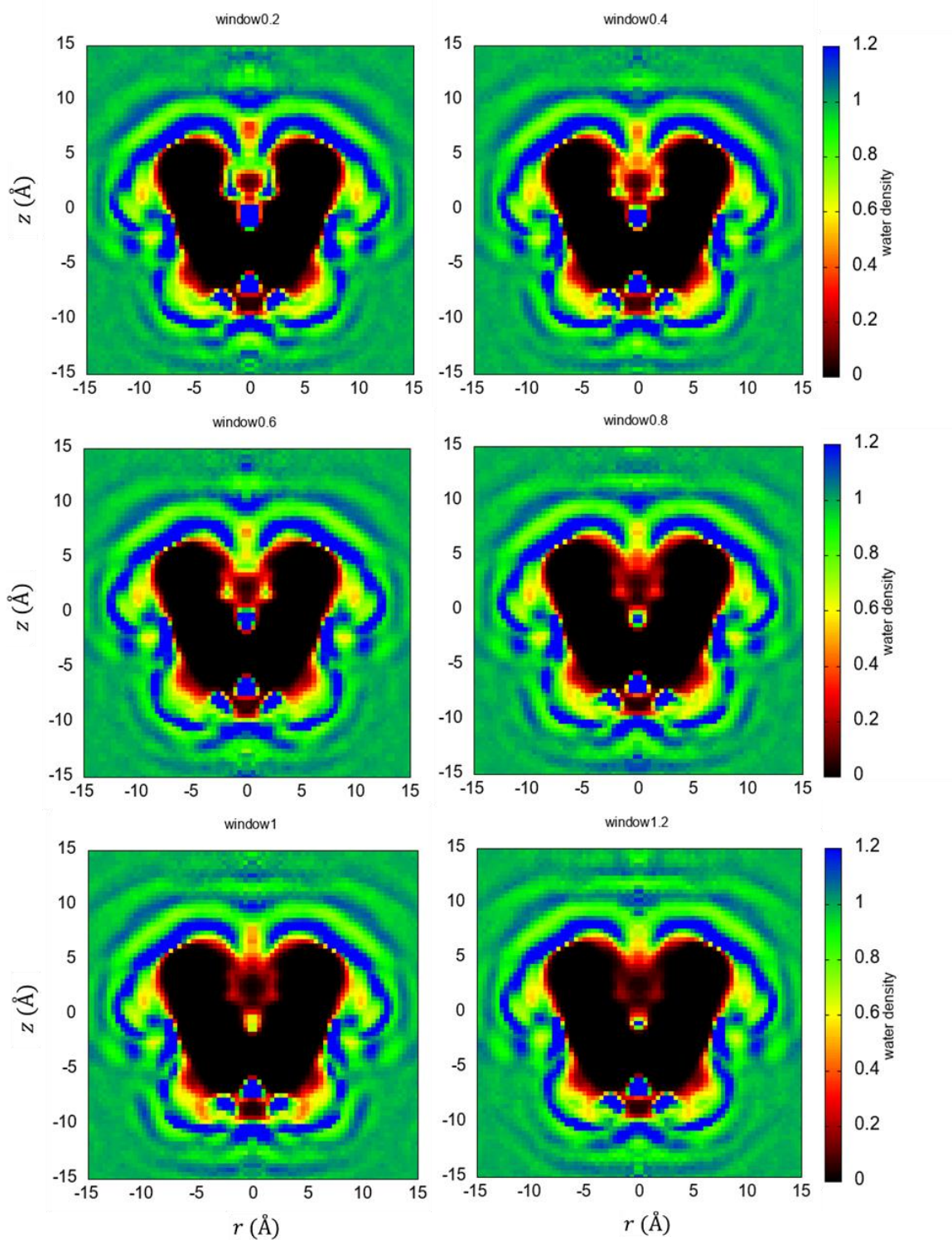
Figure 6.6: Average water density distributions about TEMOA in presence of biasing potentials on each heavy atom of the cavitand with increasing strength in finer increments. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.

In general, the ease of expelling water from a volume of observation can be used to characterize the hydrophobicity of a certain location. Here, we evaluated the sequence of evacuation occurrence in different regions inside and outside the hydrophobic pocket. Basically, in each window, if the local water density in a certain bin decreased by more than 95% compared with the last window, we assumed that this bin was evacuated in this window and will be labeled by the window number, which is also the value of ϕ, the value of the parameter regulating the strength of biasing potential applied in this window. For a more hydrophobic region, it is easier to expel water from this region so the evacuation will happen earlier. By this method, the ease of evacuation in different locations was mapped in figure 6.7 and was indicated by different colors. The white region indicating the locations of all the heavy atoms of cavitand hosts where water never occupied. The black region outside the cavitand denotes, which is labeled as -1, the locations where water density never decreased by 95% compared to previous window. Those locations are far away from the hydrophobic pocket and the portal of cavitand. This means though large biasing potentials, even with ϕ of 10 kJ/mol, can affect the first and second solvation shells, they were not able to dry the region far away from the cavitand thoroughly. The color map inside the hydrophobic pocket is also consistent with what we covered previously in the cylindrical water density distributions. Initially, in window 1 and 2, waters were only expelled in local positions that are nearest to the inner walls. When ϕ reached to 3 kJ/mol, the expelling has spread to a large portion of side and lower pocket. With the further strengthening of biasing potentials, the evacuation also occurred in the middle, upper parts of the pocket, and in the regions near the portal. When ϕ is larger than 5 kJ/mol, gradually,

the water density above the cavitand rim, and near the side and feet of the cavitand, exhibited a large extent of decrease.
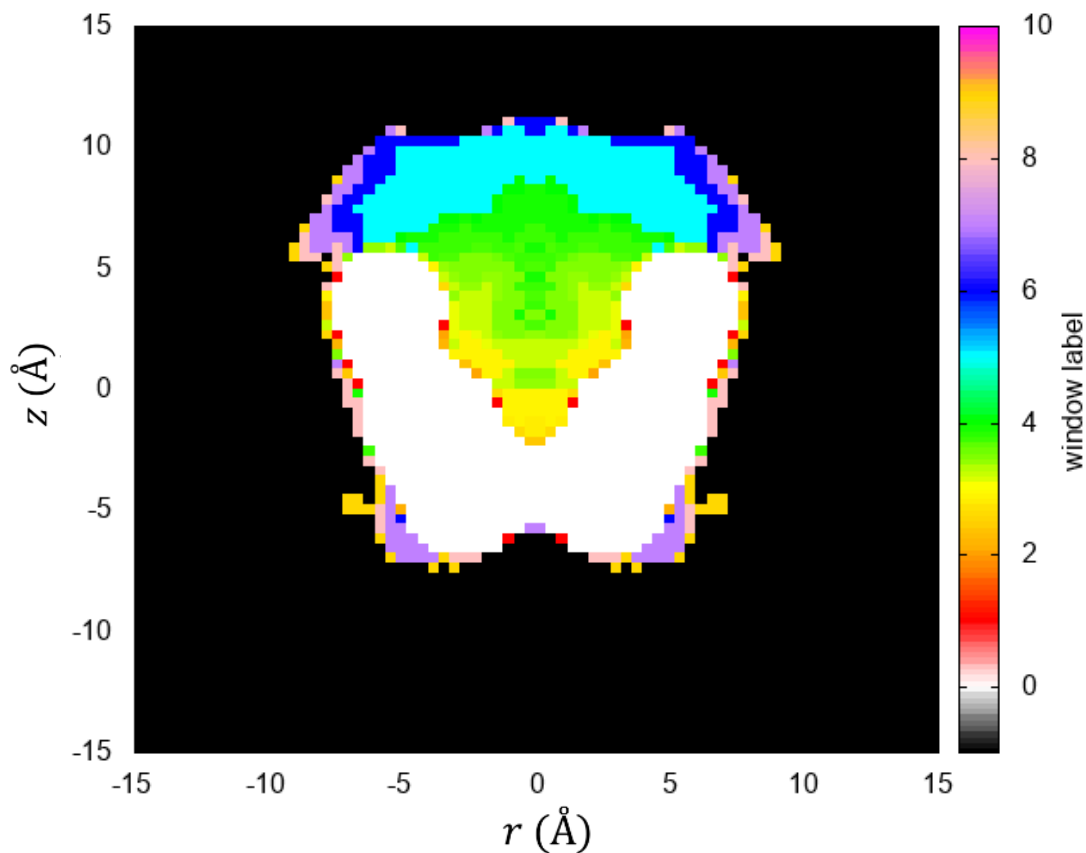
Figure 6.7: The map of evacuation sequence inside and around OA in cylindrical symmetric coordinate. All the regions are labeled as the window number where the local water density dropped by more than 95% compared to the previous window. This label follows the color key on the righthand side of the figure.

In the case of TEMOA, as shown in figure 6.8, the behavior of water outside the pocket, especially above the portal, is nearly identical to that of OA. The waters around the feet of TEMOA were more unlikely to be removed even with a strong biasing potential, where ϕ reached 10 kJ/mol. TEMOA's inner pocket exhibited a totally different behavior. The water molecules near the side walls were instantly removed in the first window that the

biasing potentials took effect. When ϕ reached 2 kJ/mol, the more stable waters at bottom of the pocket and the waters near the centerline were also expelled out of the observation volume. The water density of the whole interior space, including the water depletion region in the upper part of the pocket, underwent a significant decrease before ϕ was elevated to 4 kJ/mol. The wetting behaviors and its reaction to the external potentials inside the hydrophobic pocket are sensitive to the methylation of the endo positions on the rim. And it is also consistent with the observation of self-dewetting behavior of TEMOA discussed in detail in chapter 4.
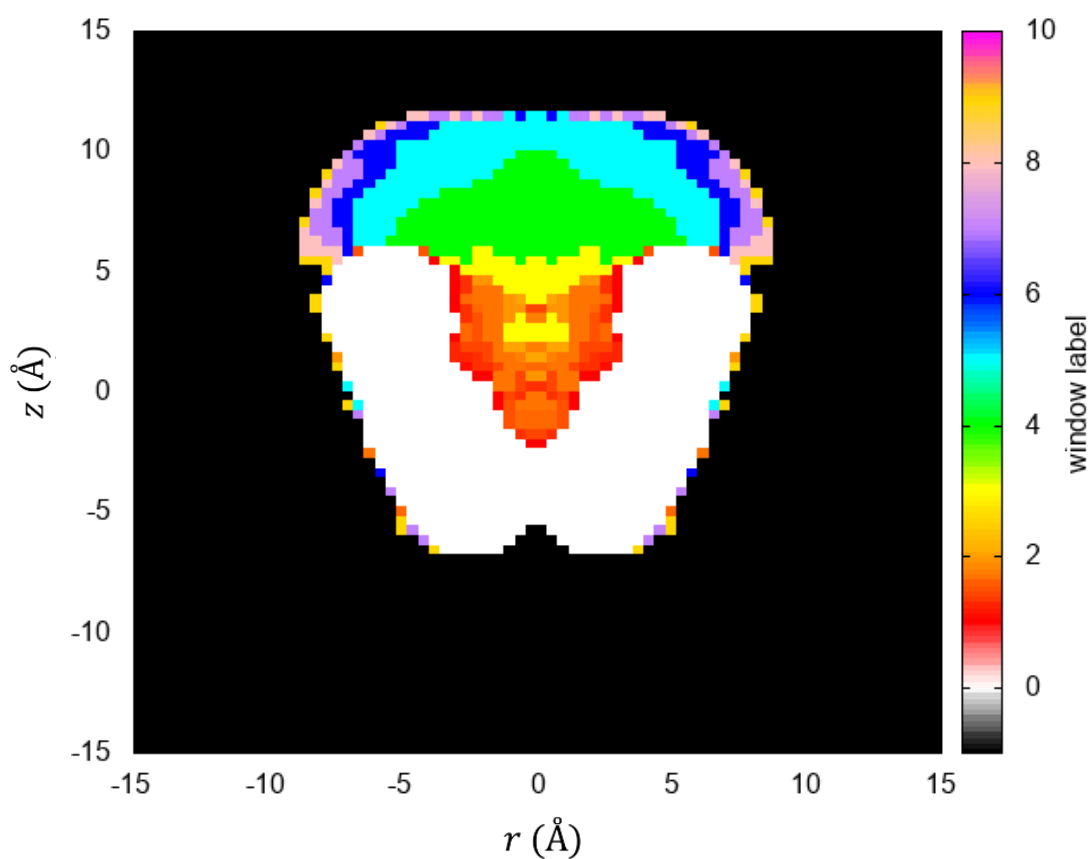
Figure 6.8: The map of evacuation sequence inside and around TEMOA in cylindrical symmetric coordinate. All the regions are labeled as the window number where the local water density dropped by more than 95% compared to the previous window. This label follows the color key on the righthand side of the figure.

From the distribution of hydrogen bond number of each water molecule inside the pocket and around the cavitand, a familiar pattern can be recognized. In the case of OA, which is shown in figure 6.9, the regions in red means the average number of hydrogen bonds of a water molecule forms in this bin is larger than or equal to three, which is the case in the whole bulk area around the cavitand. Very interestingly, in the water-depleted area, which is in the upper middle part of the pocket, the average hydrogen bond number is also near the level of the bulk water. The tendency to form hydrogen bonds exhibits a weakening trend when it is approaching the cavitand walls. In the area near the wall, where the water prefers to stay, the average hydrogen bond number drops to around 1.5, which is shown by green color. The water density near the upper pocket, where waters can form more hydrogen bonds, is much lower than that near the cavitand wall where less hydrogen bonds can be formed. It indicates that water sacrificed the hydrogen bonding at the portal of cavitand and formed attractive van der Waals interactions with the cavitand walls. The hydrogen bonding distribution inside the pocket corresponds to the pattern shown in the evacuation regions in figure 6.7. In general, water molecules can be expelled more easily in the regions where less hydrogen bonds can be formed. It can be noticed that the green region near the wall in figure 6.9 fits with the yellow region in figure 6,7, where waters were expelled when $\phi = 3$kJ/mol. The red region in the upper pocket near the portal also has the same shape as the green region where waters were evacuated in window 4. In this calculation, the hydrogen bonds between waters in pocket and the atoms of cavitand were not accounted. For the water molecules staying at the bottom of the pocket, if we include their hydrogen bonding with the four inner pointing hydrogens at the bottom of the pocket,

they can form more hydrogen bonds than their neighbors near the side wall. Therefore, it needs stronger biasing potential to remove the waters at bottom.
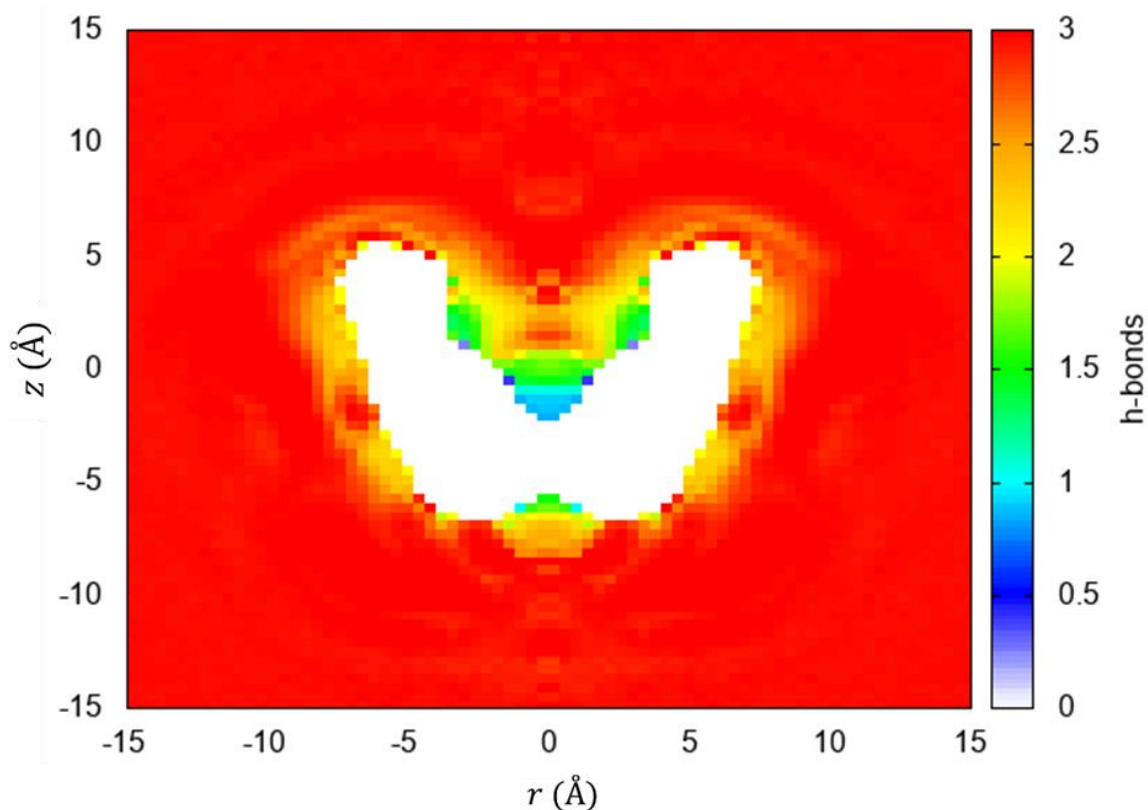
Figure 6.9: The distribution of average water-water hydrogen bonds of each water molecule inside and around OA in cylindrical symmetric coordinate. All the regions are labeled as the number of hydrogen bonds. This label follows the color key on the righthand side of the figure.

In the case of TEMOA, as shown in figure 6.10, generally less hydrogen bonds can be formed within the pocket due to a drier environment with much less hydration numbers. But it can still be observed that van der Waals interactions with the cavitand attracted waters to get inside the pocket in expense of hydrogen bonding. In the two segregated water-depleted regions near the portal and in the middle of the pocket, respectively, 2-2.5 hydrogen bonds can be formed for each water in average. In contrast, the water-rich regions

shown by green color can only form approximately 1.5 hydrogen bonds with neighbor waters. This green region roughly convers two areas. One is in the middle pocket separating the two water-depleted regions, along with the areas near the side wall. Another is the whole lower part of the pocket. These two areas can both be evacuated very easily according to figure 6.8 with a value of ϕ less than or equal to 2 kJ/mol. By comparison, the yellow regions with 2-2.5 hydrogen bonds cannot be evacuated until ϕ reached 3kJ/mol. Thus, in both OA and TEMOA, water molecules can be expelled with weaker biasing potentials in the regions where less hydrogen bonds can be formed with neighbor waters.
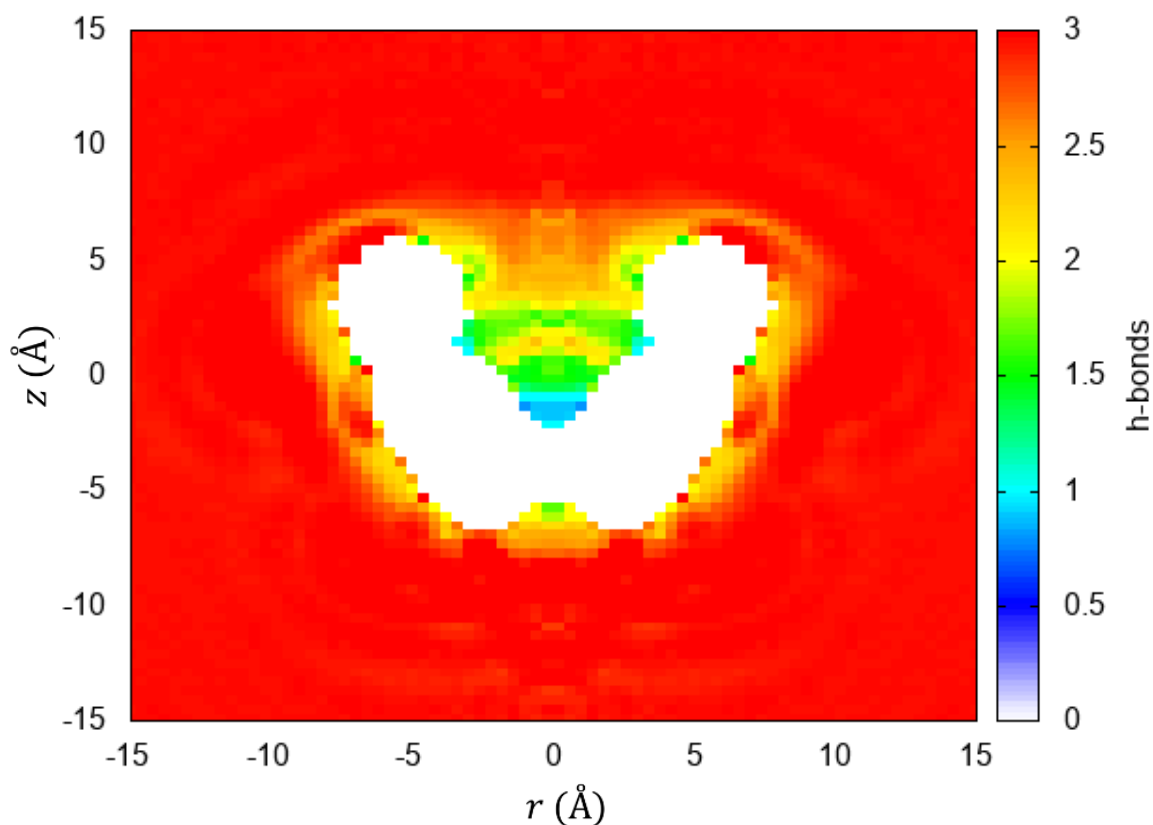
Figure 6.10: The distribution of average water-water hydrogen bonds of each water molecule inside and around TEMOA in cylindrical symmetric coordinate. All the regions are labeled as the number of hydrogen bonds. This label follows the color key on the righthand side of the figure.

Figure 6.11 (a) and 6.12 (a) shows how water molecules inside the pocket and in the whole solvation shell respond to the unfavorable biasing potential, $\phi \widetilde{N_v}$, in OA and TEMOA . $\langle N_v \rangle_\phi$ is defined as the average number of water molecules in the whole observation volume with a certain ϕ value. As described in the methodology part, this observation volume is defined as the union of all the spherical volumes centered on the initial positions of all the heavy atoms (carbon and oxygen) of cavitand. For both OA and TEMOA, the radius of each spherical sub-volume was chosen to be 0.6 nm. This whole observation volume contains approximately 260 water molecules in total. By applying biasing potentials with different strengths, the perturbation of water molecules within the affected volumes can be observed as a function of $\beta \phi$, where $\beta = 1/k_B T$. Here k_B is the Boltzmann constant and T is the temperature. In both cases of OA and TEMOA, the decrease of $\langle N_v \rangle_\phi$ with increasing ϕ is in a sigmoid shape, which is similar with that near a hydrophobic CH_3-terminated SAM (self-assembled monolayer) surface[162]. The water molecules were not expelled gradually. Instead, initially the displacement of water molecules is slow when the biasing potential is weak. When the strength of biasing potential accumulated to a relatively large level, a large proportion of water will be displaced collectively. This phenomenon can be observed more obviously in figure 6.11 (b) and 6.12 (b), where the susceptibility of total water number in the observation volume to the biasing potential strength, $\chi_v = -\frac{\partial <N_v>_\phi}{\partial (\beta \phi)}$, was plotted as a function of $\beta \phi$. It shows a prominent peak around $\phi = 2k_B T$ where the biasing potential is strong enough to trigger this collective dewetting in the solvation shells of cavitand. Very interestingly, Amish et
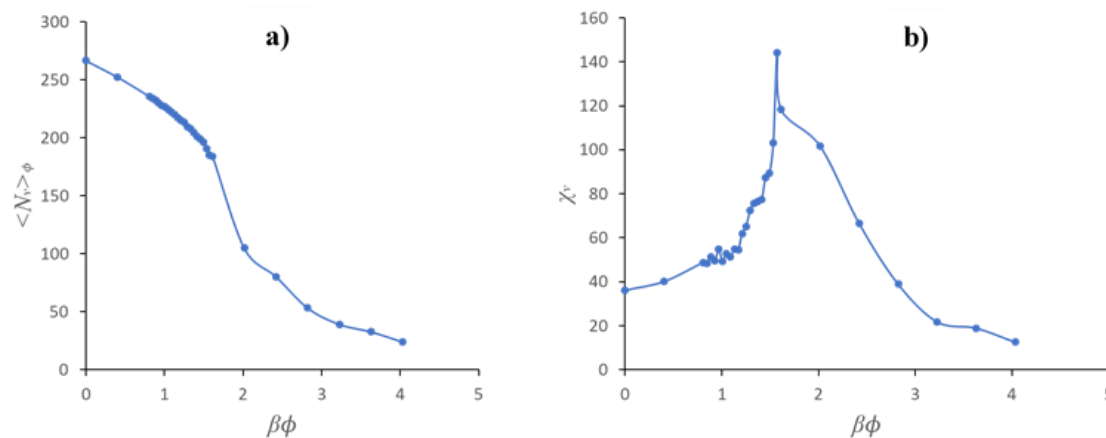
Figure 6.11: (a) Average number of water molecules in the observation volume inside and around OA affected by biasing potentials, $\langle N_v \rangle_\phi$, as a function of $\beta\phi$. (b) Corresponding susceptibility, χ_v, in the same observation volume as a function of $\beta\phi$.
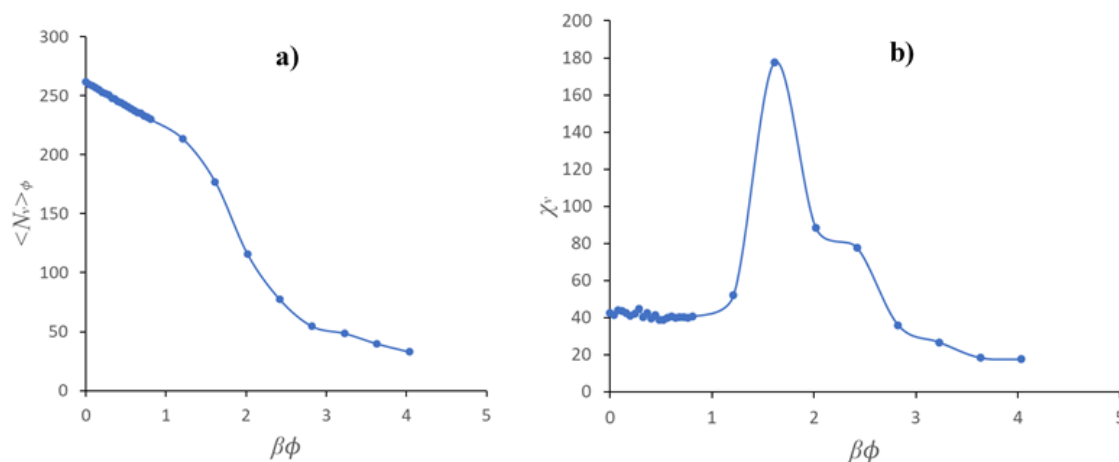
Figure 6.12: (a) Average number of water molecules in the observation volume inside and around TEMOA affected by biasing potentials, $\langle N_v \rangle_\phi$, as a function of $\beta\phi$. (b) Corresponding susceptibility, χ_v, in the same observation volume as a function of $\beta\phi$.

al[162] and coworkers stated that this peak in χ_v around $\phi = 2k_BT$ is an universal signature of protein surfaces. No matter the sizes, the biological functions, and the distribution of hydrophobic and hydrophilic regions on the surface, this collective dewetting phenomenon is a general property for proteins. Our results here on the surfaces of cavitands, which has well defined hydrophobic pockets and is studied as a mimic of protein binding sites, agree with their conclusions. And this general susceptibility was attributed to the 0.5 ratio of

hydrophobic heavy atoms on all the different protein surfaces. These atoms make the whole protein surface susceptible to those unfavorable biasing potentials[162].

Though, on the macro level, waters in the solvation exhibited similar susceptibility in the solvation shells of both cavitands. Slight differences can be noticed and indicated different hydration properties within the hydrophobic pockets. In the case of OA, the decrease of $\langle N_v \rangle_\phi$ is concave down when $\beta\phi$ is small. In contrast, the profile of $\langle N_v \rangle_\phi$ in the case of TEMOA shows a slight concave-up shape in the early stage. This early stage corresponds to the evacuation of the inner pockets and the areas above the portals of OA ($\phi = 2.0\text{-}4.0$ kJ/mol) and TEMOA ($\phi = 0\text{-}2.0$ kJ/mol). Here, $\langle N_v \rangle_\phi$ and χ_v were plotted as a function of $\beta\phi$ with a smaller increment to have a finer illustration of the drying process. For OA, the concave-down shape of $\langle N_v \rangle_\phi$ profile indicates the evacuation was accelerated with strengthened biasing potentials. The drying of the pocket requires the accumulation of biasing potential strengths to a certain degree. And after the drying of the inner pocket, the drastic displacement of water molecules in solvation shells immediately started. For TEMOA, the concave-up shape of $\langle N_v \rangle_\phi$ profile tells a different story. Once the biasing potentials took effect, the evacuation of the inner pocket occurred in the very beginning of the whole process. However, after that the displacement of waters in the rest of observation shells didn't start directly. It requires additional accumulation of biasing potential strengths. All of these are consistent with what we found in the cylindrical water density distributions in each window. And this difference is also reflected in the plot of χ_v. In the case of OA, from the beginning, it is an increasing function of $\beta\phi$, and the slope gets steeper until the prominent peak appears. For TEMOA, in general χ_v is flat as a function of $\beta\phi$ in the early stage. But it also has fluctuations and a gentle descent trend can be noticed. This

inconspicuous local minimum appears around where $\phi = 1k_BT$. This difference is not very obvious due to the small capacity of inner pockets for water molecules compared with the bulk. But it verifies our observations before.

6.5 Conclusion

In conclusion, we utilized an advanced sampling method (INDUS) to apply biasing potentials with controllable strengths centered in every heavy atom of deep-cavity cavitands. From that, the hydrophobicity and the ease of displacing water molecules in different locations can be quantified and compared. The cavitands studied here are OA (octa-acid) and TEMOA (tetra-*endo*-methyl-octa-acid). TEMOA differs from OA by having four methyl positions on its hydrophobic rim that adopted inwards (*endo*) orientations with respect to the pocket. Our simulations found that water molecules in the inner pockets responded differently to the unfavorable biasing potentials. Within OA, the inner water molecules cannot be evacuated until the strengths of biasing potentials were accumulated to a certain extent. In contrast, it is very easy to displace the water molecules within the hydrophobic pocket of TEMOA. The drying occurred immediately when the biasing potential took effect. The bulk waters in the solvation shells of both cavitands exhibited very similar behaviors under the affection of biasing potentials. This observation corresponds to the findings in previous chapters that TEMOA has a strong tendency to dewet its inner pockets by itself. And here their different reactions to external unfavorable potentials were also verified by a variety of distinct computational characterizations.

By studying the water density distribution within the pockets under effects from potentials with varying strengths, water molecules resided in different regions of the pocket showed diverse responses. Basically, the waters near the side walls of cavitands are the first to be removed from pocket. The waters staying at the bottom of the pocket are more durable to the increasing potentials. The waters near the water-depleted areas of the upper pocket are the last ones to be displaced.

Finally, we also found there is connection between the hydrogen bonding distribution for each individual water within the pocket and the ease to displace it. In general, in a certain location, water molecules can be expelled out of the pocket easily if it is harder for them to form hydrogen bonding with their neighbor waters. In addition, we found evidence showing water molecules were attracted into the pocket by interactions with atoms on cavitands and sacrificed their hydrogen bonding with other water molecules.

Chapter7 Conclusion

In Chapter 2, we reported a precise analysis of the second osmotic viral coefficients for small hydrophobic moieties such methane and ethane, and also for their alcohol counterparts, methanol and ethanol. Over a wide range of temperatures, our simulations quantitatively agree with available experimental results. It demonstrated that molecular simulation accurately captures the molecular-scale hydrophobic interactions, and the second osmotic virial coefficient is a good way to measure and quantify it.

In Chapter 3, we turned to deep-cavity cavitand systems. By advanced sampling methods (Umbrella Sampling), the association free energy of distinct assembly states of two cavitands (OA and TEMOA) with different sized *n*-alkane guests can be characterized. There are three different assembly states, 1:1, 2:1, and 2:2 complexes. By a network reaction model, the progression and the probability distribution of assembly states as a function of guest size can be predicted. This model accurately reproduces the distinct progression trends observed experimentally for OA and TEMOA. In addition, we showed that the methyl units on *endo*-positions of TEMOA significantly narrowed its opening portals and make the packing of two *n*-alkane guests very frustrated. Therefore, the balance between monotonic and dimeric complexes are disrupted and a non-monotonic assembly-state progression trend emerges. This work is extended in Chapter 4, and we applied the same model and technique to a newly developed cavitand host molecule, TEXMOA. It can

be predicted that, by shifting the methyl units from *endo*-positions to *exo*-positions, the non-monotonic progression trend was suppressed. All the transition between states occurred with longer guests compared to OA. All of these provide a guidance on the design of supramolecular hos/guest systems can be based on very small structural modifications on hosts and precise choose of guests.

The hydration behaviors were studied by simulation and models in Chapter 5. By a systematic study of a series of cavitand host molecules with different extents of functionalization (methyl or hydroxyl) on either *exo-* or *endo*-positions of rim, we found the hydration condition within pocket can be controlled by functionalization. In general, wetting state is favored by hydroxylation on endo-positions, while dry state is favored by methylation correspondingly. From probability distribution of hydration numbers, a two-state behavior can be observed and there is a free energy barrier at $n=1$ between the dry and wet states. The hydration condition is largely dependent on system pressures, and we also showed the absorption isotherm of different cavitands can be shifted to a unified shape by assuming the effective pressures for each cavitand are differed by a shift pressure. Finally, a capillary evaporation model was developed to explain the equilibrium between the dry and wet states within pockets. And this equilibrium is governed by the balance of interfacial free energies for drying/wetting the interior spaces. This result provides us a potential new design route to shift the wetting states within molecular binding states. The first observation of the self-dewetting phenomenon for TEMOA and its impact on host-guest binding can be found in Appendix F.

Lastly, in Chapter 6, we applied a new advanced sampling approach (INDUS) to qualitatively characterize the hydrophobicity on the whole surface and pocket of cavitand.

In the presence of unfavorable biasing potentials, water molecules exhibited totally different behaviors. In general, waters inside TEMOA are more easily to be displaced compared to that within OA. We also observed different eases of expelling water from individual locations and it is also connected with the opportunities to form hydrogen-bonding. The indication of water entering the pocket for favorable van der Waals interactions in expense of hydrogen bonds with neighbor water molecules were also noticed. Finally, on surface of cavitand, we also observed a collective dewetting phenomenon, which is regarded as a general property for proteins.

For the future work, lots of things can be done and investigated by molecular simulation for the family of cavitands. The binding with anion guests is one of them since ion binding is a sophisticated process. This binding can occur not only within pockets but also may happen in other sites like the feet or rim. Its impact on binding and reactions happening within cavitand pocket is another huge and important topic. On the assembly states side, the formation of hetero-dimers (two different types of cavitands in one dimer) can be studied by umbrella sampling, and even with wider range of guests (not limited to long n-alkane guests). The work about INDUS reported in the Chapter 6 is a preliminary exploratory study. Deeper insights can be provided for the hydration both within pocket and on surface if a detailed quantitative hydrophobicity map can be prepared for cavitand based on this technique. In a word, the family of cavitands and its binding or hydration has lots of potential work in the aspect of molecular simulation.

Appendix A
Supporting information for "Second Osmotic Virial Coefficients of Short Alkanes and Their Alcohol Counterparts in Water as a Function of Temperature"

The second osmotic viral coefficients of methane, ethane, methanol, and ethanol were determined from calculation using equation 2.2 and are reported in Table A1. The simulation results for methane were previously reported in former research of our group[76].

Table A 1: Second osmotic virial coefficients of methanol, methane, ethanol, and ethane in water determined from molecular simulation as a function of temperature. The numbers in parentheses indicate a statistical error of one standard deviation.

T(K)	B_2 (cm^3/mol)			
	methanol	methane	ethanol	ethane
275	21.3 (1.5)	32.2 (1.6)	38.9 (2.5)	50.9 (2.4)
280	21.2 (2.6)	17.5 (2.2)	32.1 (2.0)	49.7 (2.6)
285	15.8 (1.8)	-15.4 (2.2)	10.2 (2.1)	33.1 (3.0)
290	11.4 (2.0)	15.2 (2.3)	2.2 (2.4)	26.3 (1.6)
295	17.4 (2.2)	-5.2 (2.2)	12.9 (2.2)	22.6 (1.7)
300	26.7 (1.5)	4.5 (2.1)	9.5 (1.6)	16.5 (2.6)
305	9.1 (1.4)	-8.1 (2.3)	-2.4 (1.8)	5.8 (2.4)
310	24.2 (1.9)	-3.0 (2.3)	-7.2 (2.5)	-0.4 (1.8)
315	13.0 (2.3)	-18.4 (2.3)	8.9 (2.4)	-24.6 (2.6)
320	7.8 (1.9)	-18.7 (2.3)	-7.1 (2.5)	-16.4 (1.9)
325	10.2 (1.9)	-25.4 (2.4)	-0.7 (2.1)	-34.4 (2.0)
330	3.7 (1.2)	-29.3 (2.3)	-17.3 (2.7)	-49.7 (2.3)
335	5.6 (1.1)	-36.6 (2.3)	-15.9 (1.8)	-63.7 (1.6)
340	1.4 (1.9)	-40.3 (2.4)	-13.0 (2.7)	-66.6 (2.5)
345	5.3 (1.1)	-35.3 (2.2)	-16.8 (2.8)	-72.2 (2.4)
350	8.3 (1.7)	-42.4 (2.2)	-19.6 (2.0)	-85.1 (1.9)
355	5.1 (1.5)	-53.0 (2.3)	-32.1 (2.7)	-108.0 (2.8)
360	-2.4 (1.9)	-53.6 (2.5)	-16.9 (2.2)	-120.9 (2.8)
365	2.2 (2.2)	-58.1 (2.5)	-7.8 (2.5)	-122.0 (1.7)
370	-0.5 (1.5)	-68.0 (2.5)	-28.0 (1.7)	-137.2 (2.1)

Appendix B
Supporting information for "Guest Controlled Nonmonotonic Deep Cavity Cavitand Assembly State Switching"

The potential of mean force for association of 1:1, 2:1, and 2:2 complexes of OA and nonane shown in figure 3.3 are drawn from a systematic analysis of the PMFs of all three complexes between cavitands and alkane guests. The cavitands are OA and TEMOA. The guests are all hydrocarbon molecules from methane to hexadecane. The PMF minima and the corresponding separation distances between two cavitands at where the PMF minima occurs shown in figure 3.4 and 3.8 are also directly from this study. Here, all the PMFs between one single cavitand (OA or TEMOA) and one n-alkane (C_1 to C_{16}) along the C_4-axis to form 1:1 complexes are presented in Figure B1. Likely, the PMFs between two cavitands (OA or TEMOA) with one n-alkane guest (C_1 to C_{16}) sitting inside one of the cavitands along the collinear C_4-axes to form 2:1 complexes are presented in Figure B2. Figure B3 illustrates the PMFs between two cavitands (OA or TEMOA) with one n-alkane guest (C1 to C11 for OA, and C1 to C9 for TEMOA) sitting in the non-polar pocket of each of the cavitands along the collinear C_4-axes to form 2:2 complexes. Figure B4 exhibits the PMFs between two empty cavitands (OA or TEMOA) along the collinear C_4-axes to form 2:0 complexes.
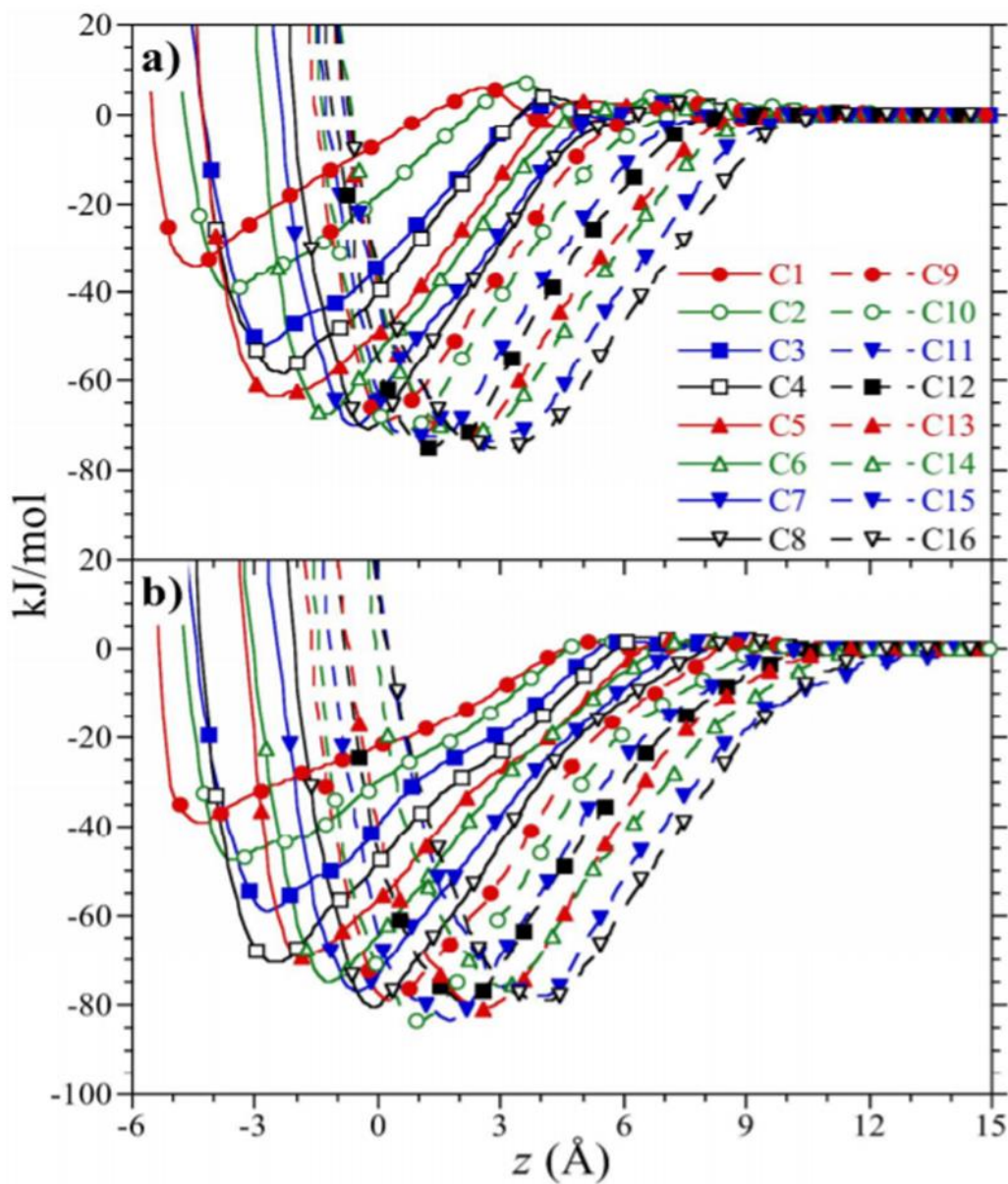
FigureB 1: Potential of mean force between a single cavitand and one single n-alkane guest to form a 1:1 complex. The results of OA and TEMOA are shown in a and b, respectively. The legend in the upper figure is illustrating the legend for different guests. This legend also applies to all the PMF figures shown below. We also didn't show the error bars for clarity.
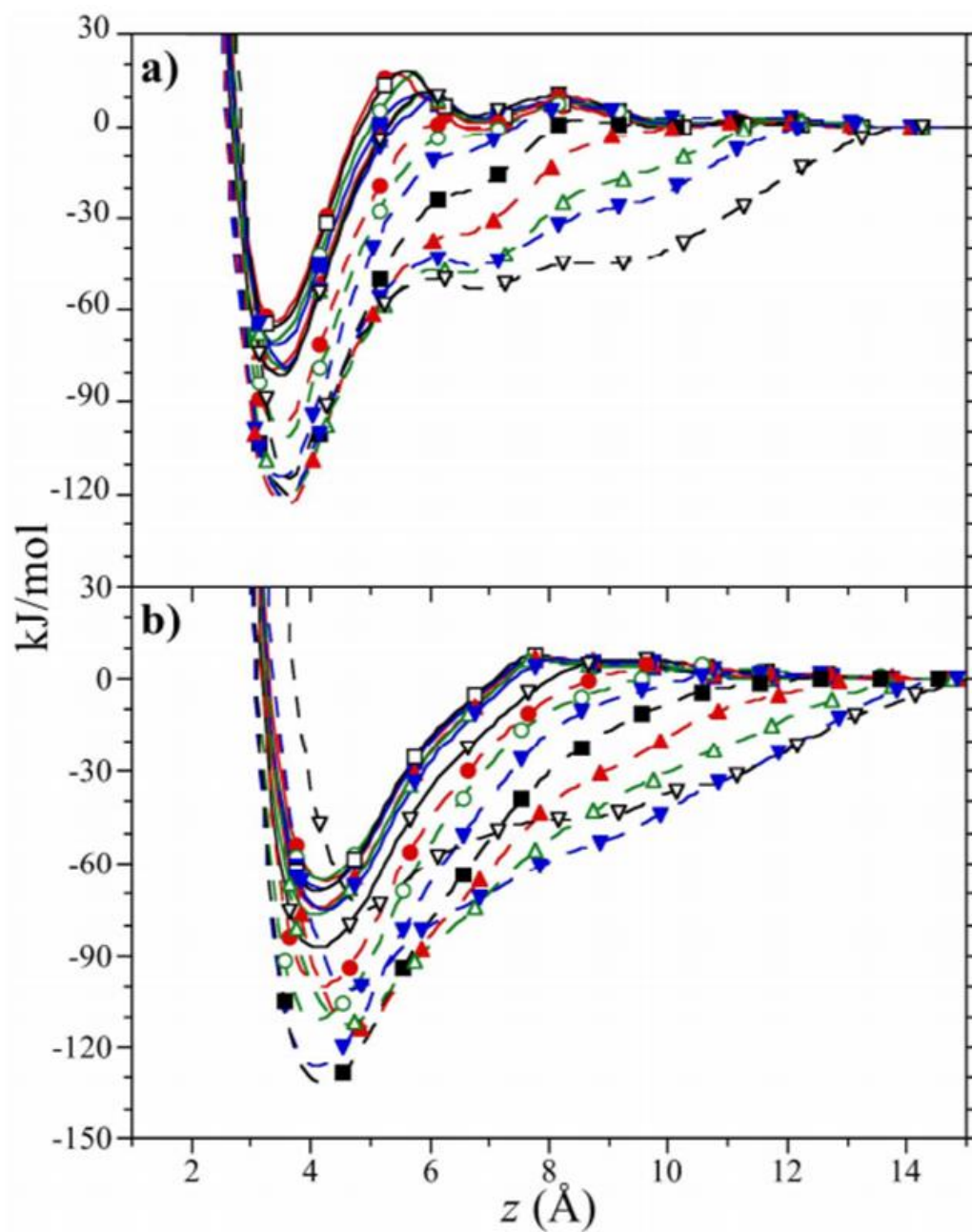
FigureB 2: Potential of mean force between two cavitands with one single n-alkane guest sitting in the hydrophobic pocket of one of the cavitands to form a 2:1 complex. The results of OA and TEMOA are shown in a and b, respectively. We also didn't show the error bars for clarity.
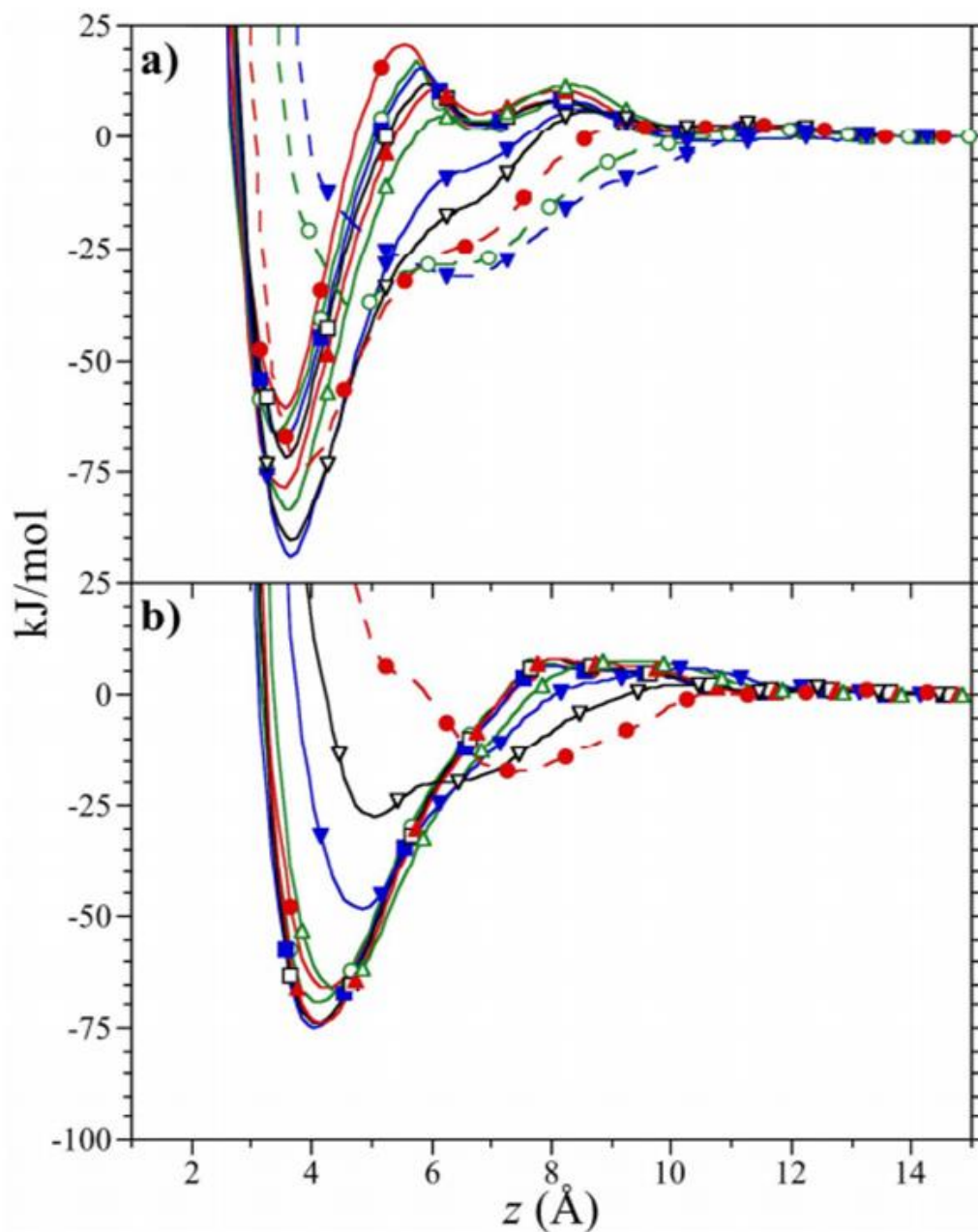
FigureB 3: Potential of mean force between two cavitands with one single n-alkane guest sitting in the hydrophobic pocket of each of the cavitands to form a 2:1 complex. The results of OA and TEMOA are shown in a and b, respectively. The alkane guests considered for OA are from C_1 to C_{11}. And guests from C_1 and C_9 are considered for TEMOA. Longer guests will lead to unstable complexes in simulation. We also didn't show the error bars for clarity.
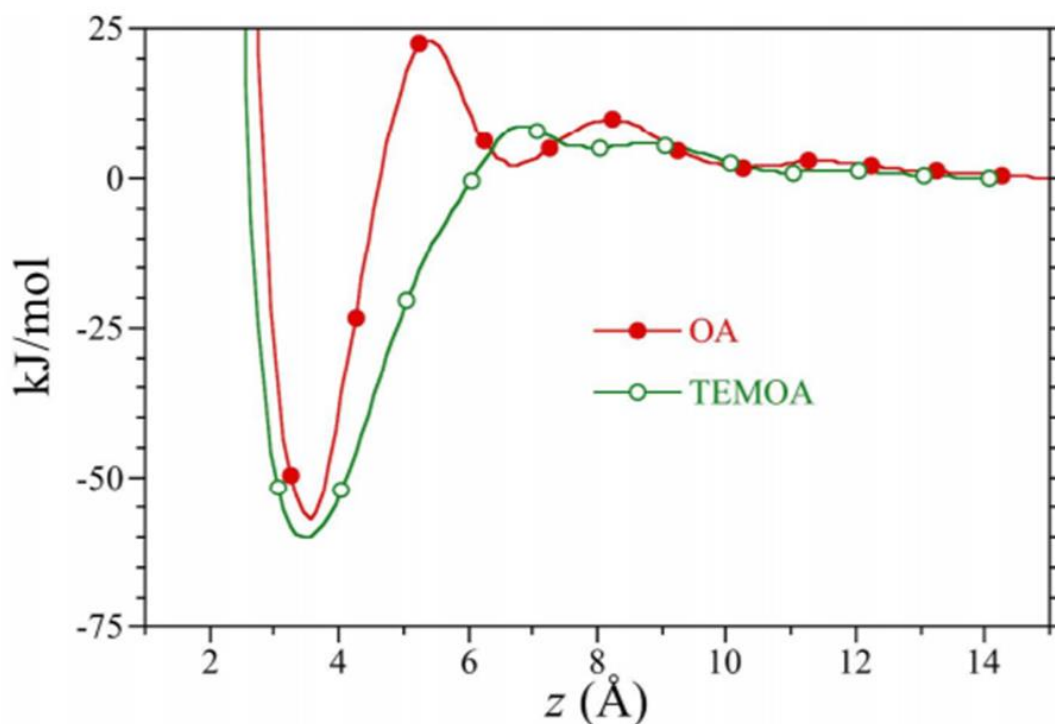
FigureB 4: Potential of mean force between two empty cavitands to form an empty dimer. The legend is identifying the different colors and symbols for OA and TEMOA. The error bars are also neglected for clarity.

In chapter 3, in order to evaluate the molar concentration of cavitand in aqueous solution at 298.15 K and 1 bar, the excess chemical potential of each n-alkane guest need to be calculated. The excess chemical potential, which is also the solvation free energy of the n-alkane guests can be evaluated by the Bennett acceptance ratio method[165,166]. The molecular simulations to calculate the solvation free energy were conducted for one single n-alkane guest in 500-1000 TIP4P-EW water molecules. The number of waters depends on the length of n-alkane guest. All the other simulation setups (forcefields, nonbonded Lennard-Jones interactions, electrostatic interactions, dynamic integration scheme,

thermostats, barostat, etc.) are exactly the same as that described in chapter 3. During the simulation, the interactions between water and n-alkane guest were gradually turned off. A coupling parameter, λ, was introduced to describe the ratio remnant interaction to the full interaction. There are two main phases in the whole transformation process. Firstly, the partial charges on all the carbon and hydrogen atoms of n-alkane guests were turned off in 0.25λ increments from 1 to 0. Here 1 means full interactions and 0 means all interactions are turned off. After that, the Lennard-Jones interactions between water and n-alkane were turned off in 0.1λ increments from 1 to 0. Therefore, there are totally 16 states in the whole process. The simulation in each state was run by 5 ns for data production following a 1 ns equilibration run.

Table B1 below shows the gas phase partial pressure (the determination of guest partial pressure has been described in chapter 3), excess chemical potential, and molar solubility of all the n-alkane guests. Though the guest concentrations are very small in water, as described in chapter 3, their associations with cavitands are large enough to attract guests into pockets of cavitand to form complexes.

TableB 1: n-alkane gas phase pressures (PG), excess chemical potentials (μ_G^{ex}), and molar guest solubility ([G]) in water at 298.15 K. Standard errors in the last two digits for the excess chemical potential and guest concentration are reported in parentheses.

guest	P_G (bar)	μ^{ex}_G (kJ/mol)	[G] (M)
C_1	1.013	9.687(42)	$8.21(14)\times10^{-4}$
C_2	1.013	10.272(67)	$6.48(18)\times10^{-4}$
C_3	1.013	11.057(88)	$4.72(17)\times10^{-4}$
C_4	1.013	12.61(11)	$2.524(12)\times10^{-4}$
C_5	6.834×10^{-1}	13.82(12)	$1.045(51)\times10^{-4}$
C_6	2.018×10^{-1}	14.92(14)	$1.98(11)\times10^{-5}$
C_7	6.105×10^{-2}	15.68(17)	$4.41(30)\times10^{-6}$
C_8	1.875×10^{-2}	16.59(18)	$9.37(68)\times10^{-7}$
C_9	5.837×10^{-3}	16.94(20)	$2.53(20)\times10^{-7}$
C_{10}	1.834×10^{-3}	18.66(21)	$3.98(34)\times10^{-8}$
C_{11}	5.803×10^{-4}	19.64(23)	$8.49(77)\times10^{-9}$
C_{12}	1.856×10^{-4}	19.86(26)	$2.49(26)\times10^{-9}$
C_{13}	5.934×10^{-5}	21.75(28)	$3.71(41)\times10^{-10}$
C_{14}	1.941×10^{-5}	22.40(30)	$9.3(11)\times10^{-11}$
C_{15}	6.295×10^{-6}	22.85(31)	$2.52(31)\times10^{-11}$
C_{16}	2.056×10^{-6}	24.02(34)	$5.13(71)\times10^{-12}$

Appendix C
Supporting information for "The Progression of Assembly States as Function of Guest Sizes for a Newly Developed Deep-cavity Cavitand"

The potential of mean force for association of 1:1, 2:1, and 2:2 complexes of TEXMOA and nonane shown in figure 4.2 are chosen from a systematic analysis of the PMFs of all three complexes between TEXMOA and alkane guests. The PMF minima and equilibrium inter-cavitand separations reported in figure 4.3 and 4.6 were also taken from the PMFs listed here. Figure C1 reports the PMFs between an empty TEXMOA with an n-alkane guest along cavitand's C_4-axis to form 1:1 complexes. The guests studied here ranges from C_1 to C_{16}. Similarly, Figure C2 reports the PMFs between an empty TEXMOA and a 1:1 complex, which has a single guest (also from C_1 to C_{16}) residing inside the inner pocket freely, along the C_4-axis to form 2:1 complexes. Figure C3 reports the PMFs between two 1:1 TEXMOA complexes with a single guest in each pocket (from C_1 to C_{13}) along the C_4-axis to form 2:2 complexes. The PMF between two empty TEXMOA hosts along the C_4-axis to form 2:0 complexes are reported in Figure C4.

FigureC 1: Potential of mean force between a single n-alkane guest and a TEXMOA to form a 1:1 complex in water. The figure symbols are identified in the legend. Error bars are neglected for clarity.
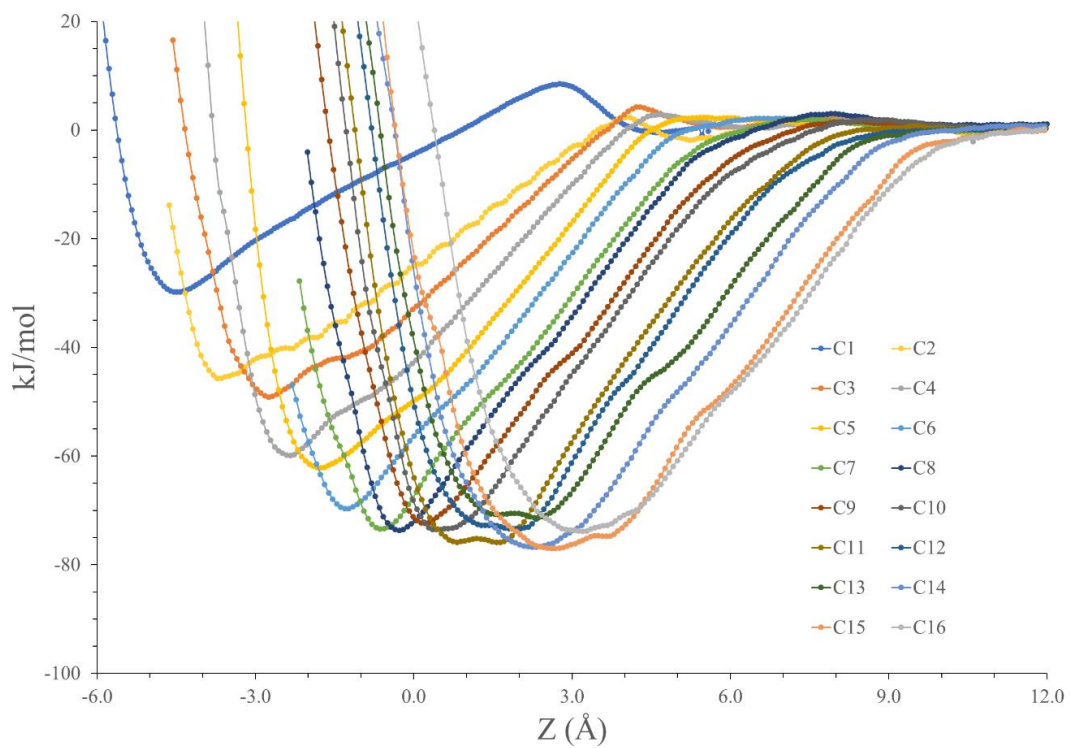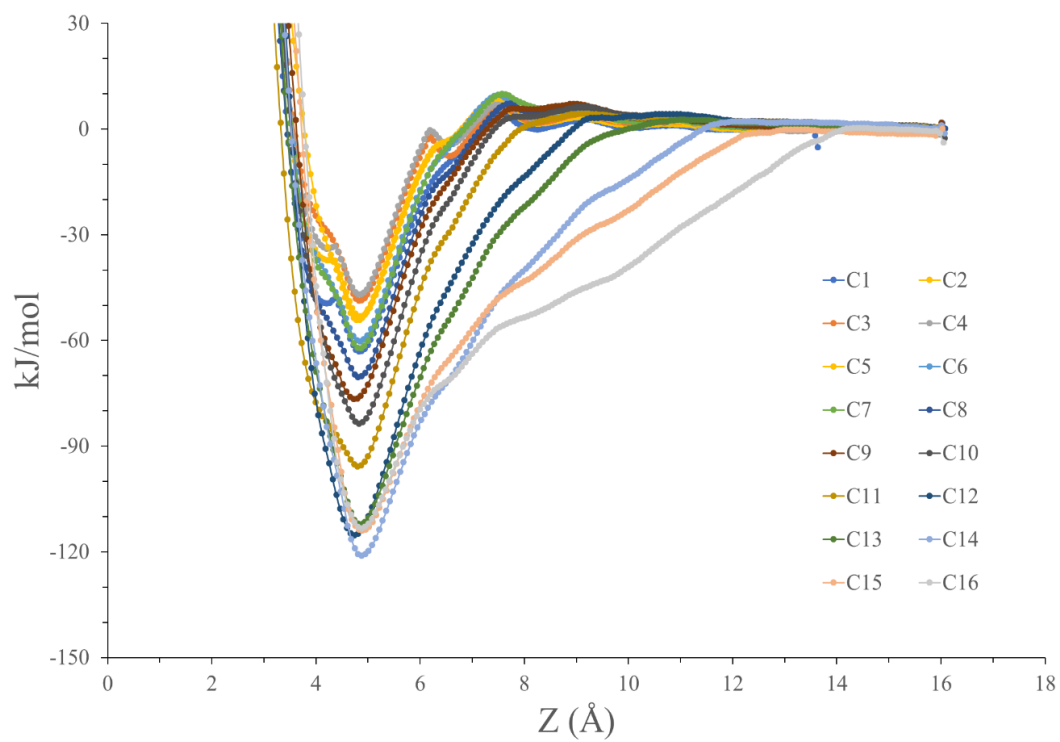
FigureC 2: Potential of mean force between a single n-alkane guest and a TEXMOA to form a 2:1 complex in water. The figure symbols are identified in the legend. Error bars are neglected for clarity.
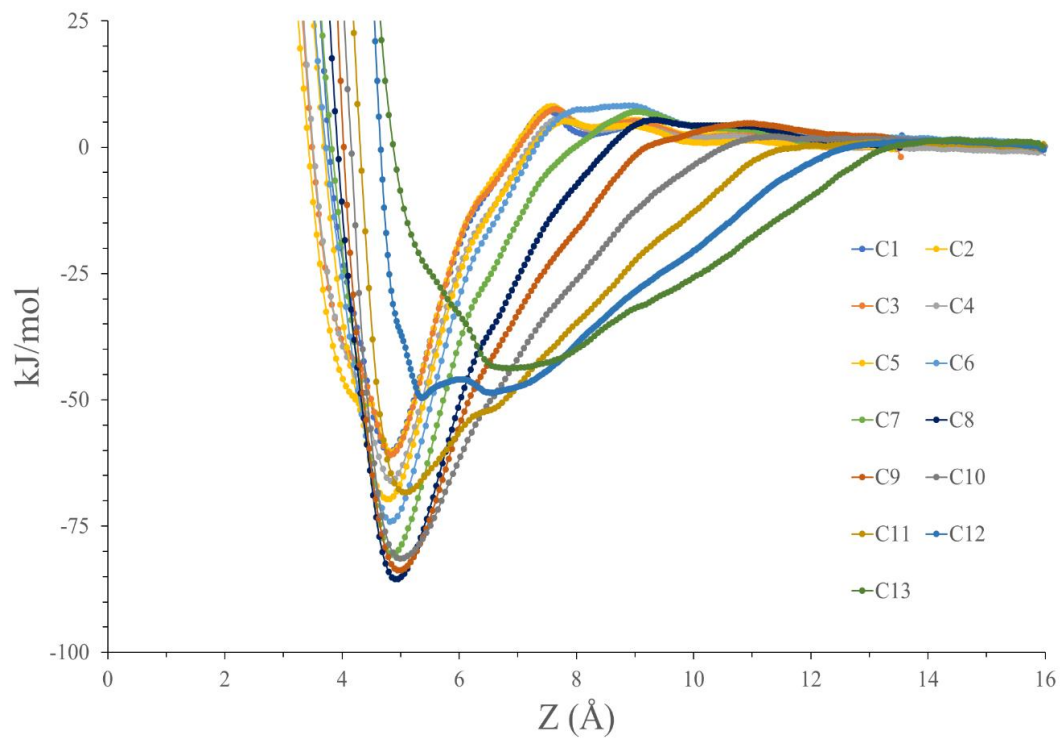
FigureC 3: Potential of mean force between a single n-alkane guest and a TEXMOA to form a 2:2 complex in water. The figure symbols are identified in the legend. Error bars are neglected for clarity.
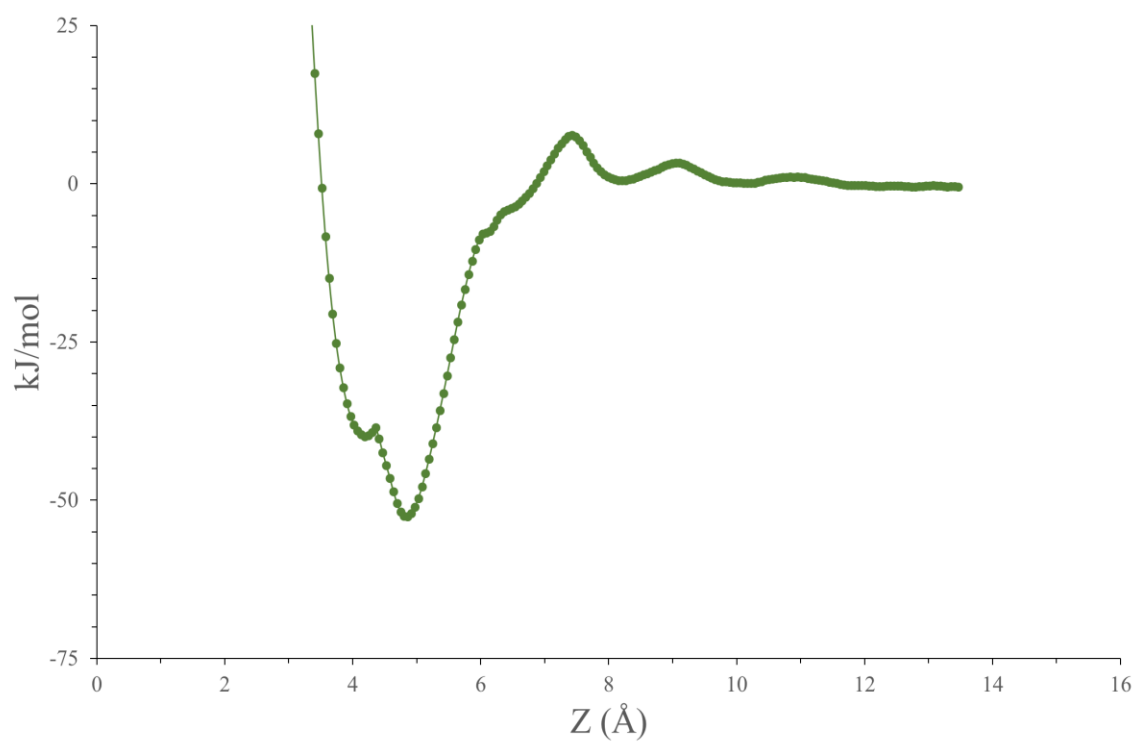
FigureC 4: Potential of mean force between two empty TEXMOA hosts to form a 2:0 complex in water. Error bars are neglected for clarity.

Appendix D

Supporting information for "Water Absorption and Dewetting Phenomenon in Deep Cavity Cavitand and The Effect of Modification of Cavitand Structure"

Figure D1-D8 present all the hydration distributions in the non-polar pockets.
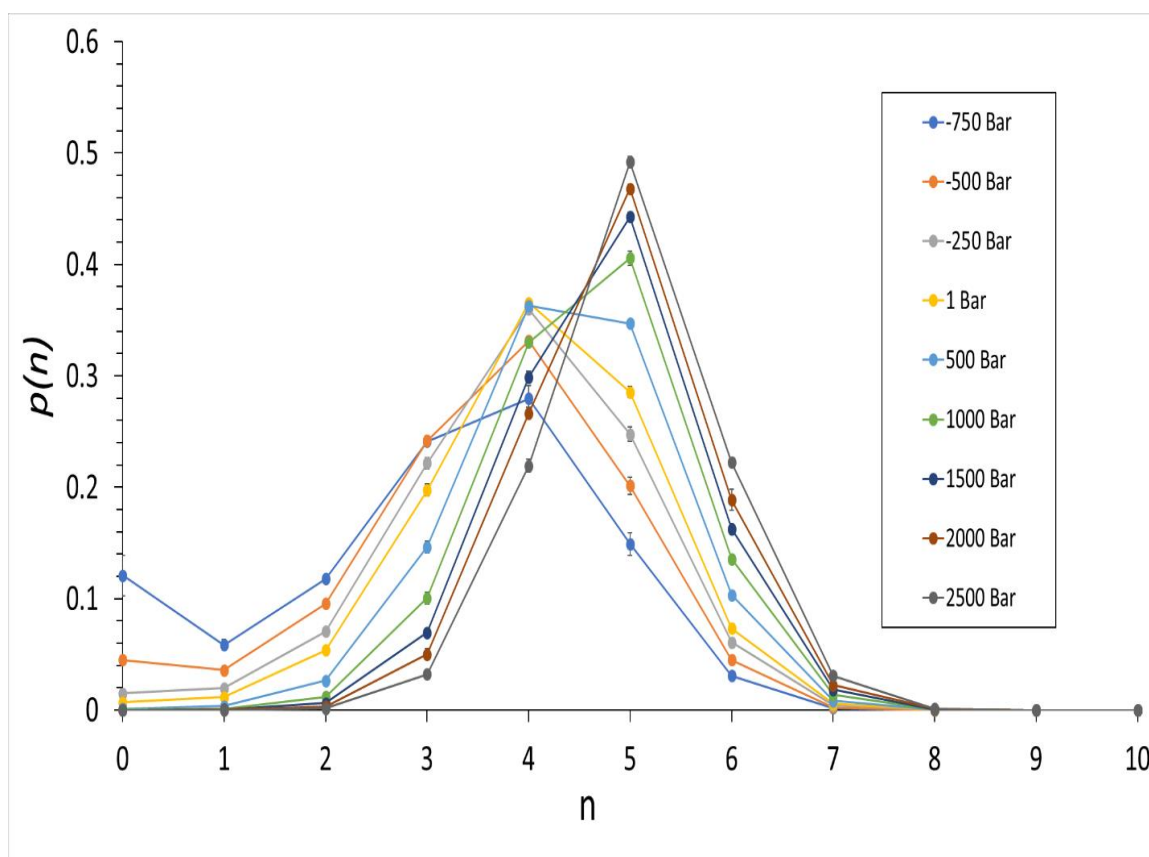
FigureD 1: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of OA as a function of the number of waters, n, determined from simulations at 298.15 K. The pressure ranges from -750 bar to 2500 bar. The colors indicating different pressures are identified in the figure legend. The error bars in the figure indicate one standard deviation.
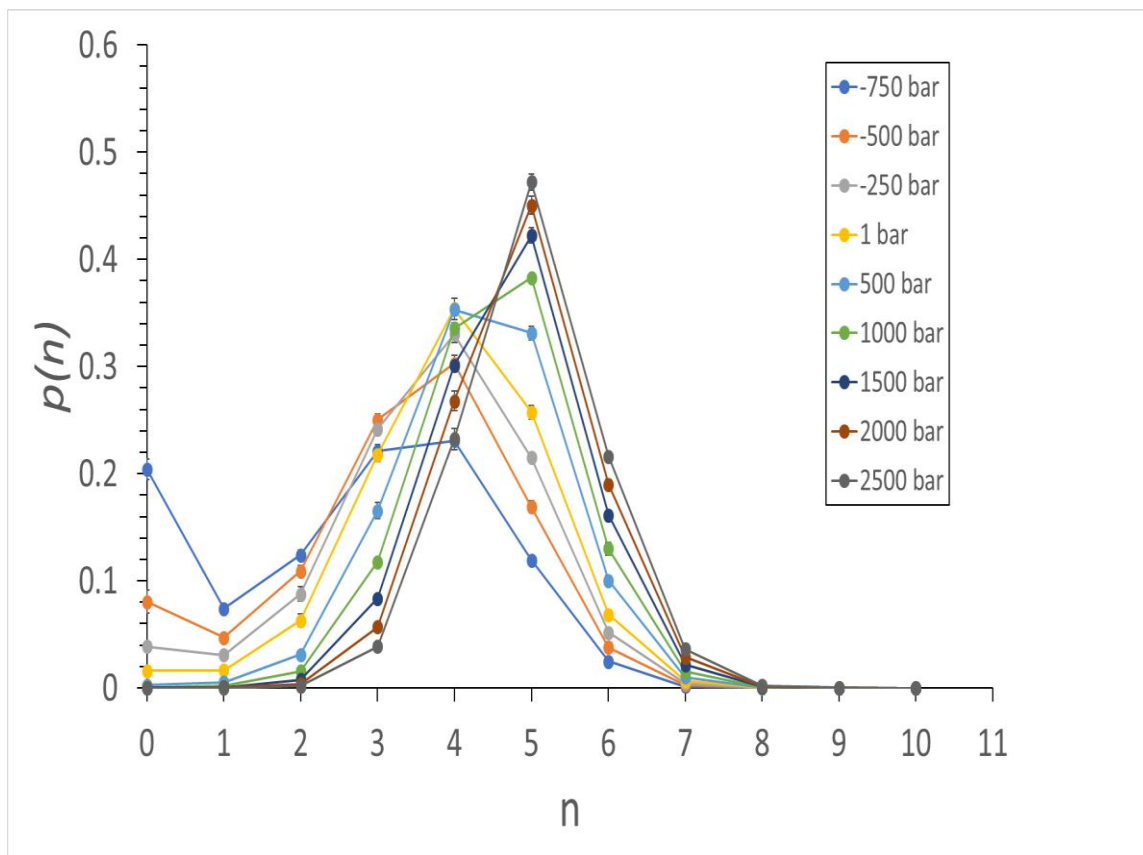
FigureD 2: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of MEMOA as a function of the number of waters, n, determined from simulations at 298.15 K. The pressure ranges from -750 bar to 2500 bar. The colors indicating different pressures are identified in the figure legend. The error bars in the figure indicate one standard deviation.
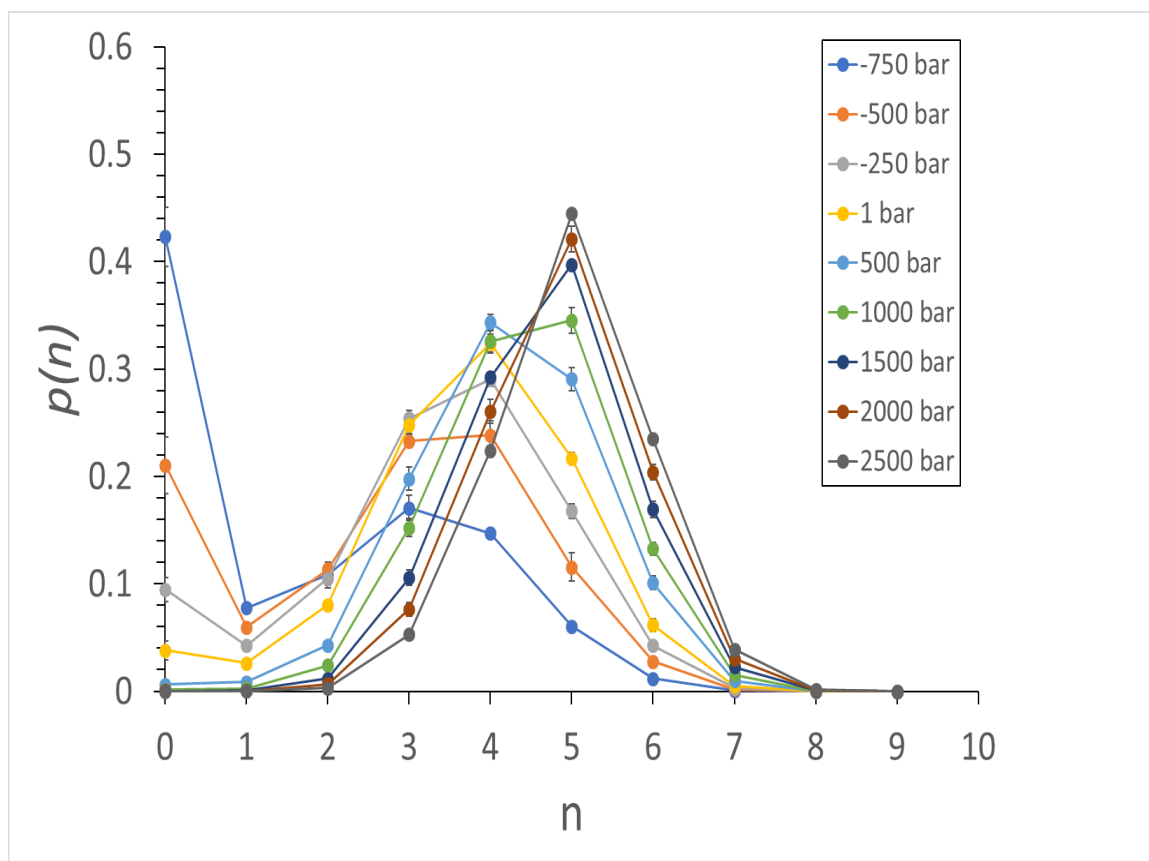
FigureD 3: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of DEMOA as a function of the number of waters, n, determined from simulations at 298.15 K. The pressure ranges from -750 bar to 2500 bar. The colors indicating different pressures are identified in the figure legend. The error bars in the figure indicate one standard deviation.

FigureD 4: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TriEMOA as a function of the number of waters, n, determined from simulations at 298.15 K. The pressure ranges from -750 bar to 2500 bar. The colors indicating different pressures are identified in the figure legend. The error bars in the figure indicate one standard deviation.

FigureD 5: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TEMOA as a function of the number of waters, n, determined from simulations at 298.15 K. The pressure ranges from -750 bar to 2500 bar. The colors indicating different pressures are identified in the figure legend. The error bars in the figure indicate one standard deviation.

FigureD 6: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TEXMOA as a function of the number of waters, n, determined from simulations at 298.15 K. The pressure ranges from -750 bar to 2500 bar. The colors indicating different pressures are identified in the figure legend. The error bars in the figure indicate one standard deviation.

FigureD 7: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TEHOA as a function of the number of waters, n, determined from simulations at 298.15 K. The pressure ranges from -750 bar to 2500 bar. The colors indicating different pressures are identified in the figure legend. The error bars in the figure indicate one standard deviation.

FigureD 8: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TEXHOA as a function of the number of waters, n, determined from simulations at 298.15 K. The pressure ranges from -750 bar to 2500 bar. The colors indicating different pressures are identified in the figure legend. The error bars in the figure indicate one standard deviation.

The linear models relating the change of Gibbs free energy from certain hydration state to the reference state to pressures for all the cavitands are shown here in figure D9-D16.
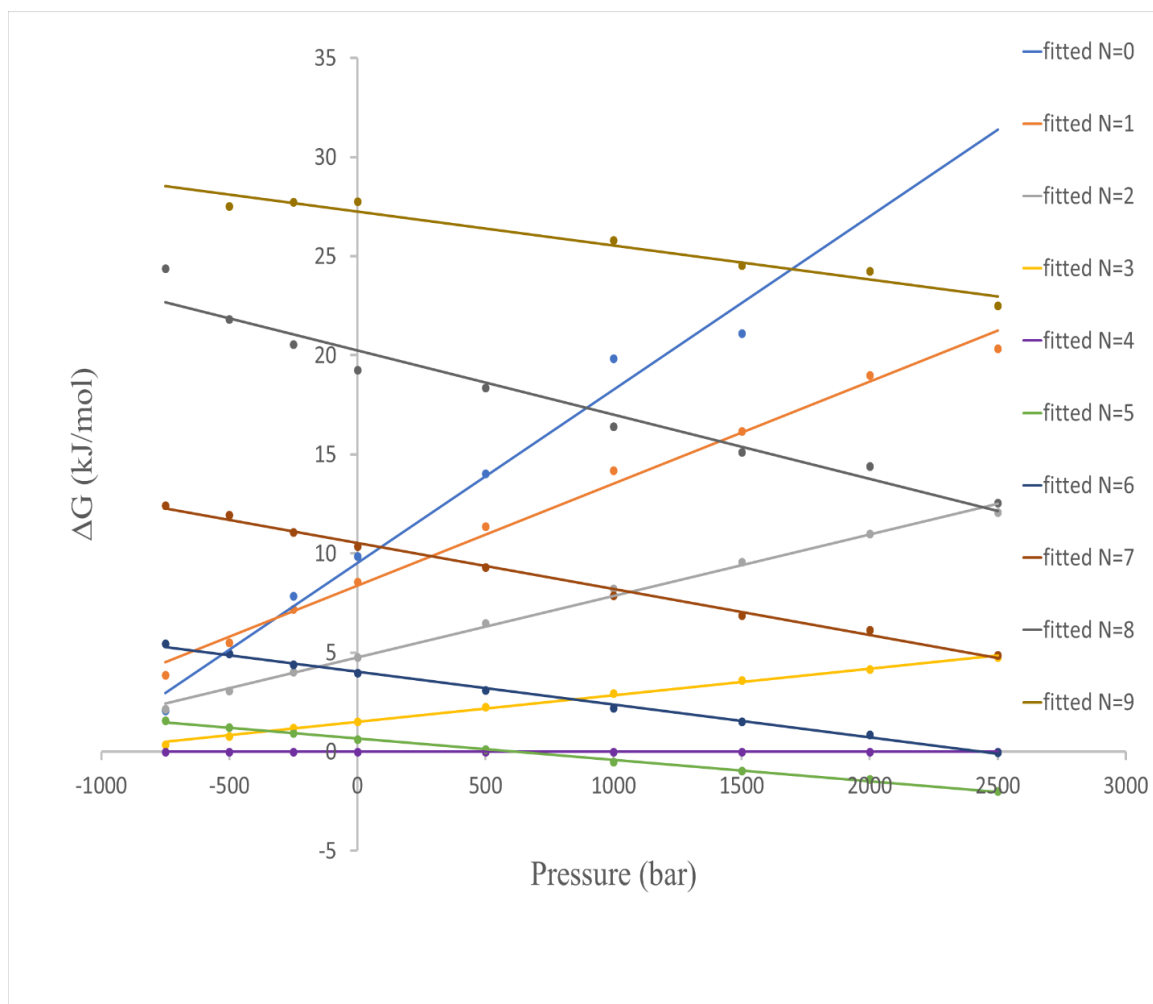
FigureD 9: Change of free energy from certain states and the reference state with different pressures (data points), and the linear trendline fitted to the data (line). The cavitand here is OA, and temperature is 298.15 K. The colors indicating different hydration numbers are identified in the figure legend.
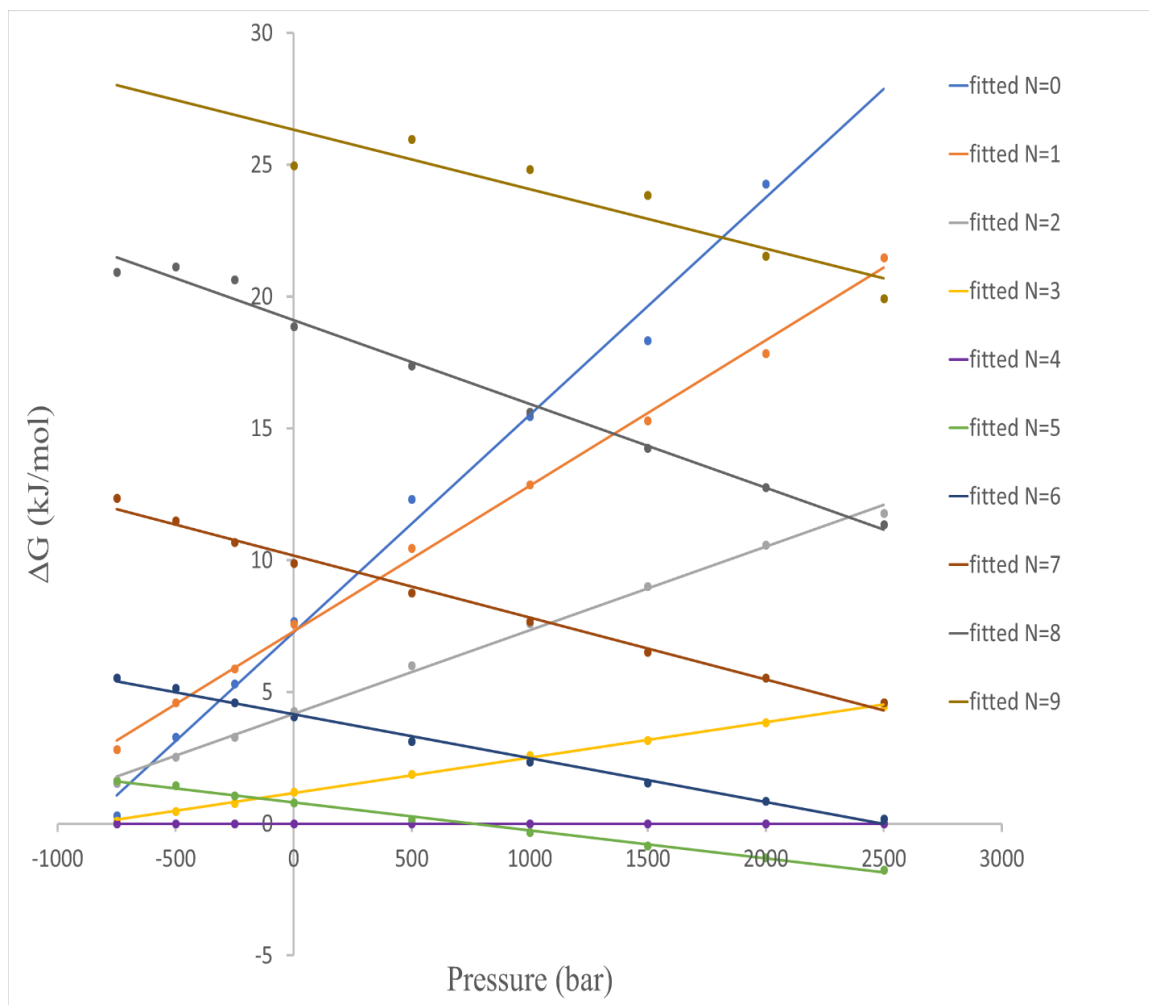
FigureD 10: Change of free energy from certain states and the reference state with different pressures (data points), and the linear trendline fitted to the data (line). The cavitand here is MEMOA, and temperature is 298.15 K. The colors indicating different hydration numbers are identified in the figure legend.
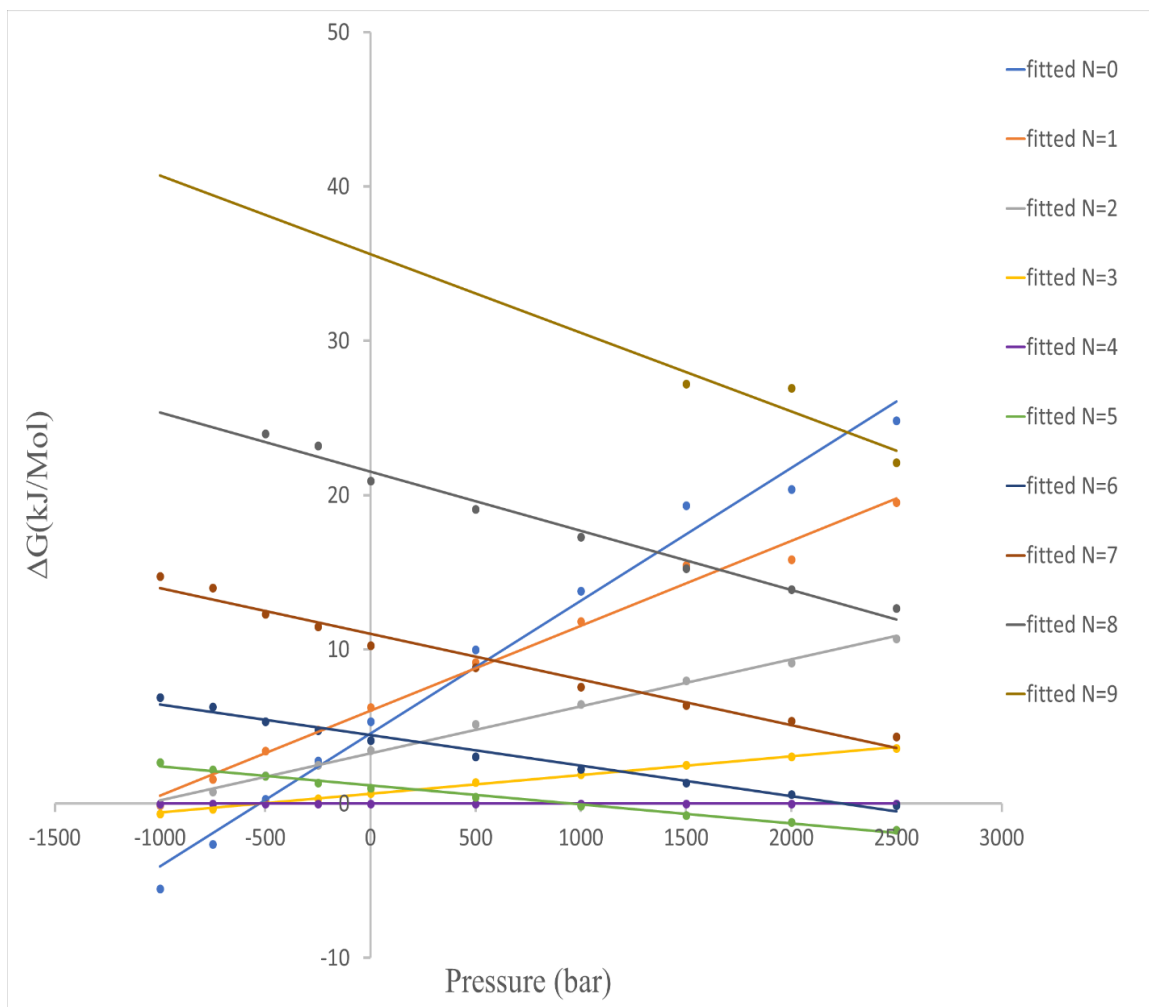
FigureD 11: Change of free energy from certain states and the reference state with different pressures (data points), and the linear trendline fitted to the data (line). The cavitand here is DEMOA, and temperature is 298.15 K. The colors indicating different hydration numbers are identified in the figure legend.
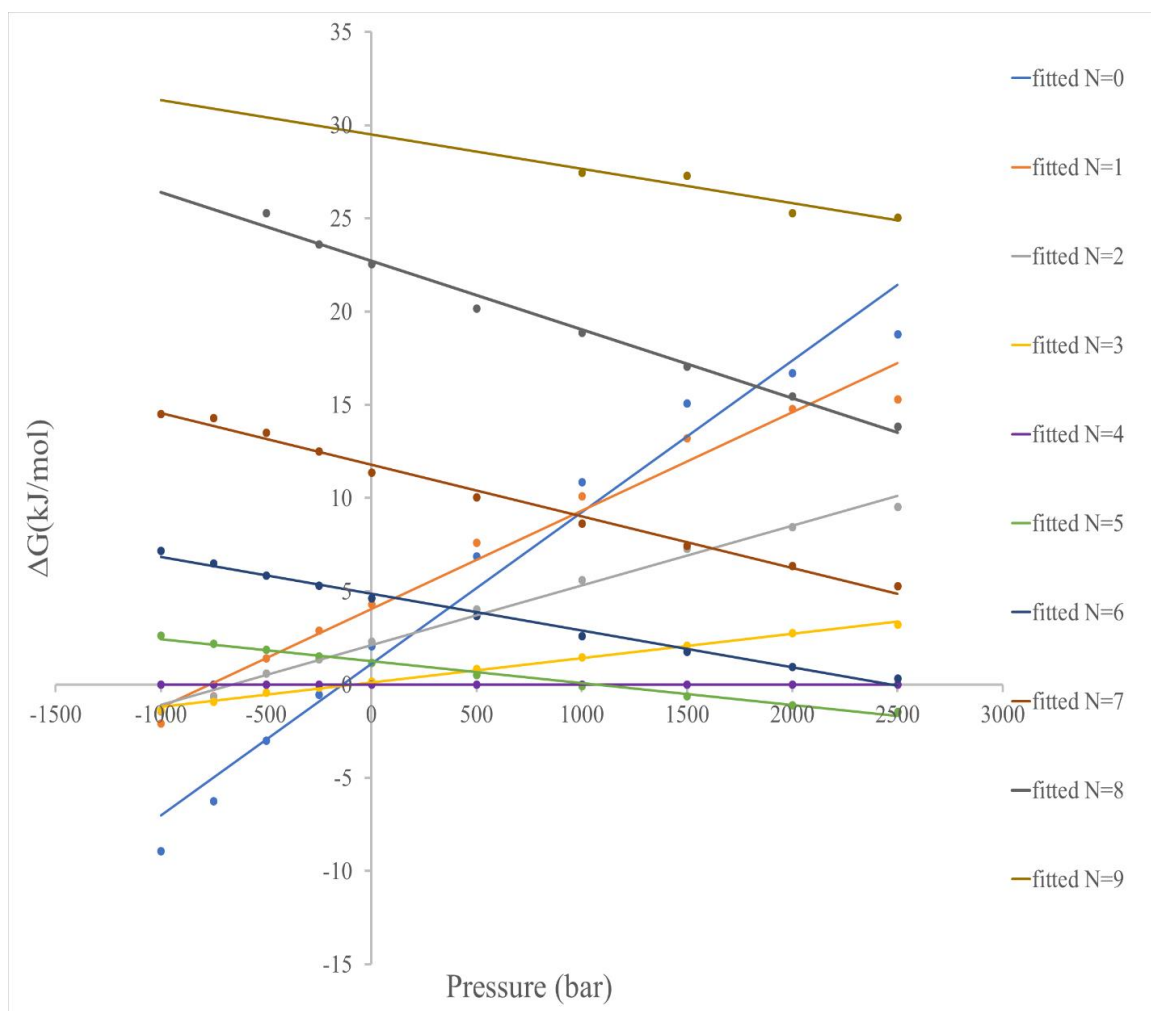
FigureD 12: Change of free energy from certain states and the reference state with different pressures (data points), and the linear trendline fitted to the data (line). The cavitand here is TriEMOA, and temperature is 298.15 K. The colors indicating different hydration numbers are identified in the figure legend.
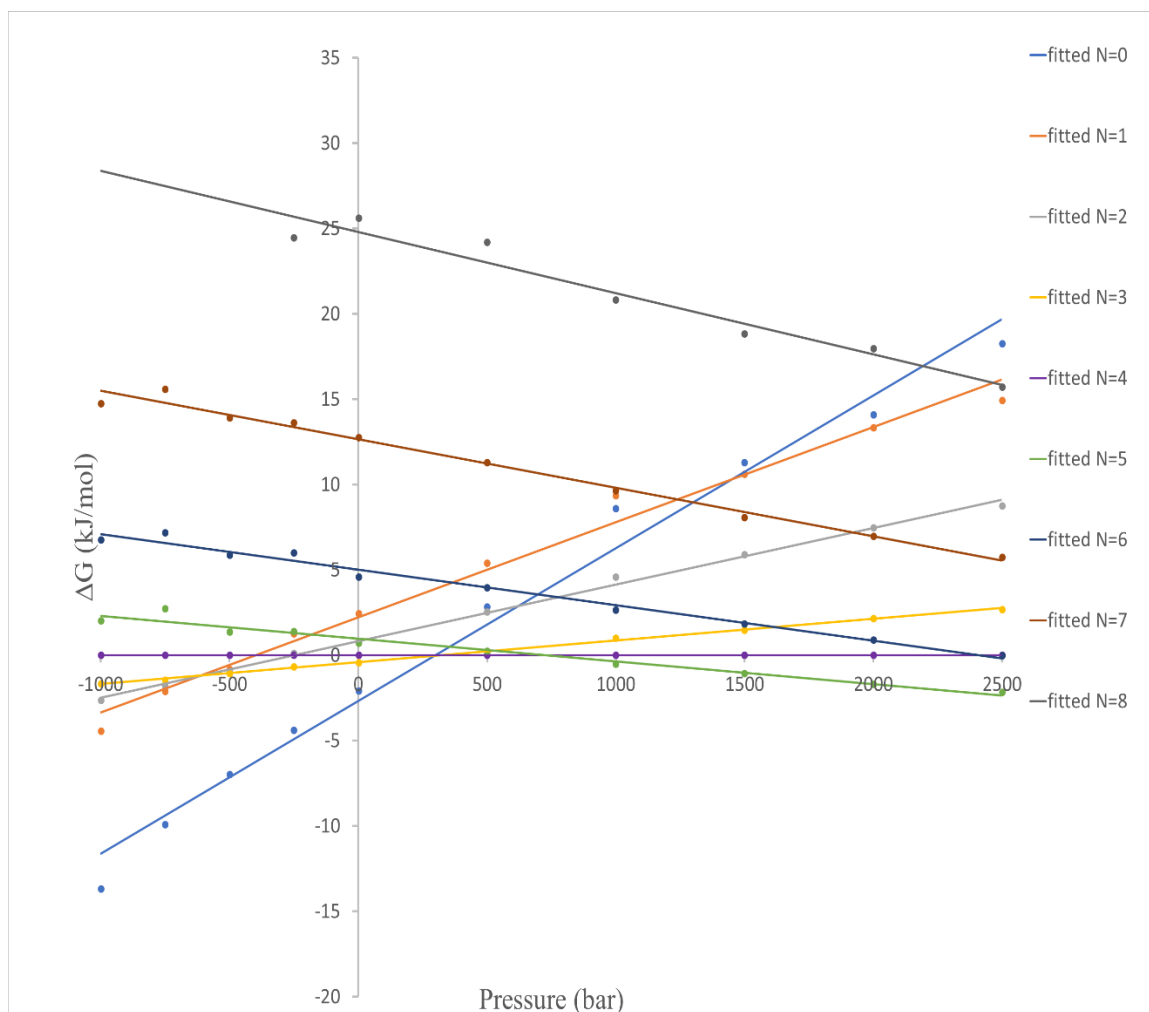
FigureD 13: Change of free energy from certain states and the reference state with different pressures (data points), and the linear trendline fitted to the data (line). The cavitand here is TEMOA, and temperature is 298.15 K. The colors indicating different hydration numbers are identified in the figure legend.
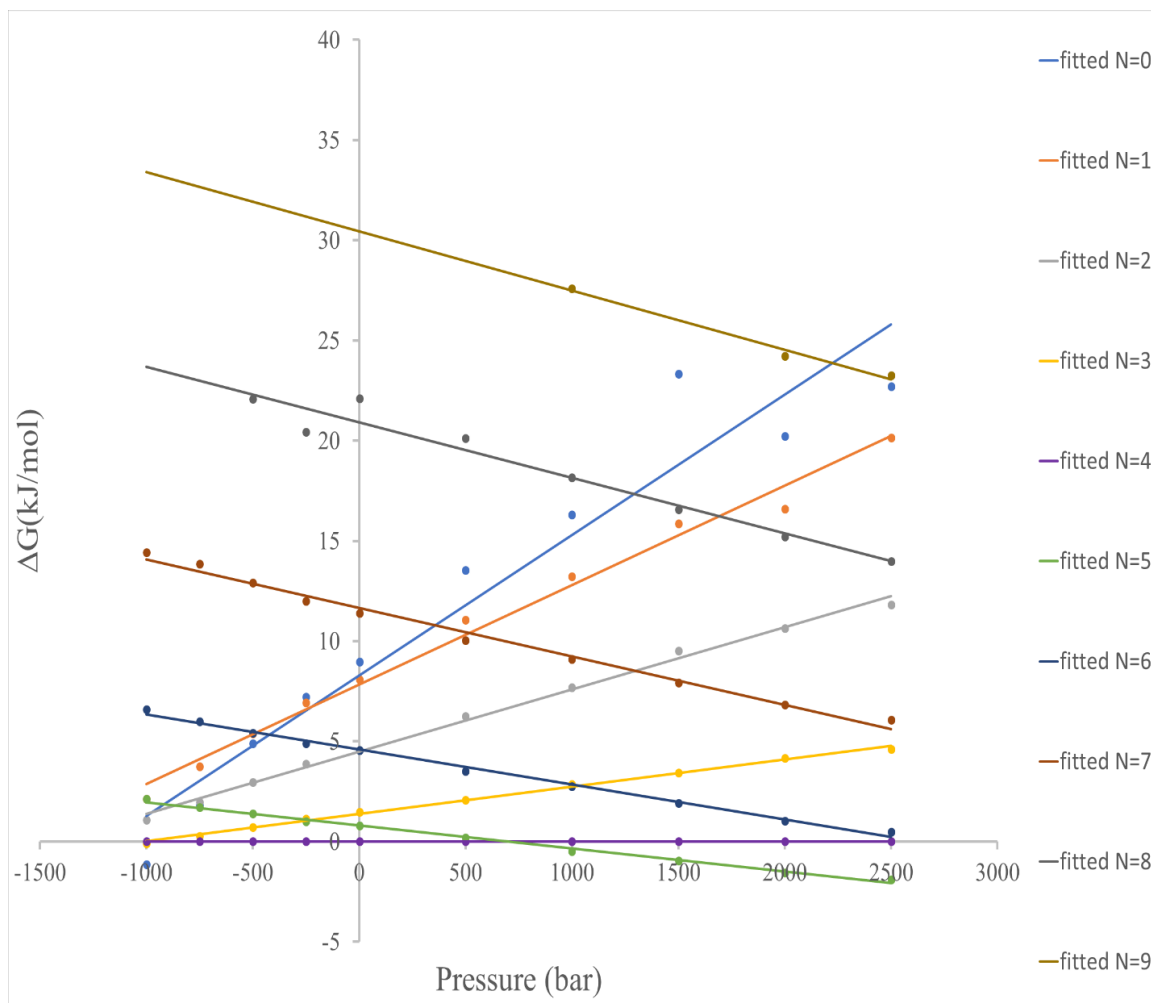
FigureD 14: Change of free energy from certain states and the reference state with different pressures (data points), and the linear trendline fitted to the data (line). The cavitand here is TEXMOA, and temperature is 298.15 K. The colors indicating different hydration numbers are identified in the figure legend.
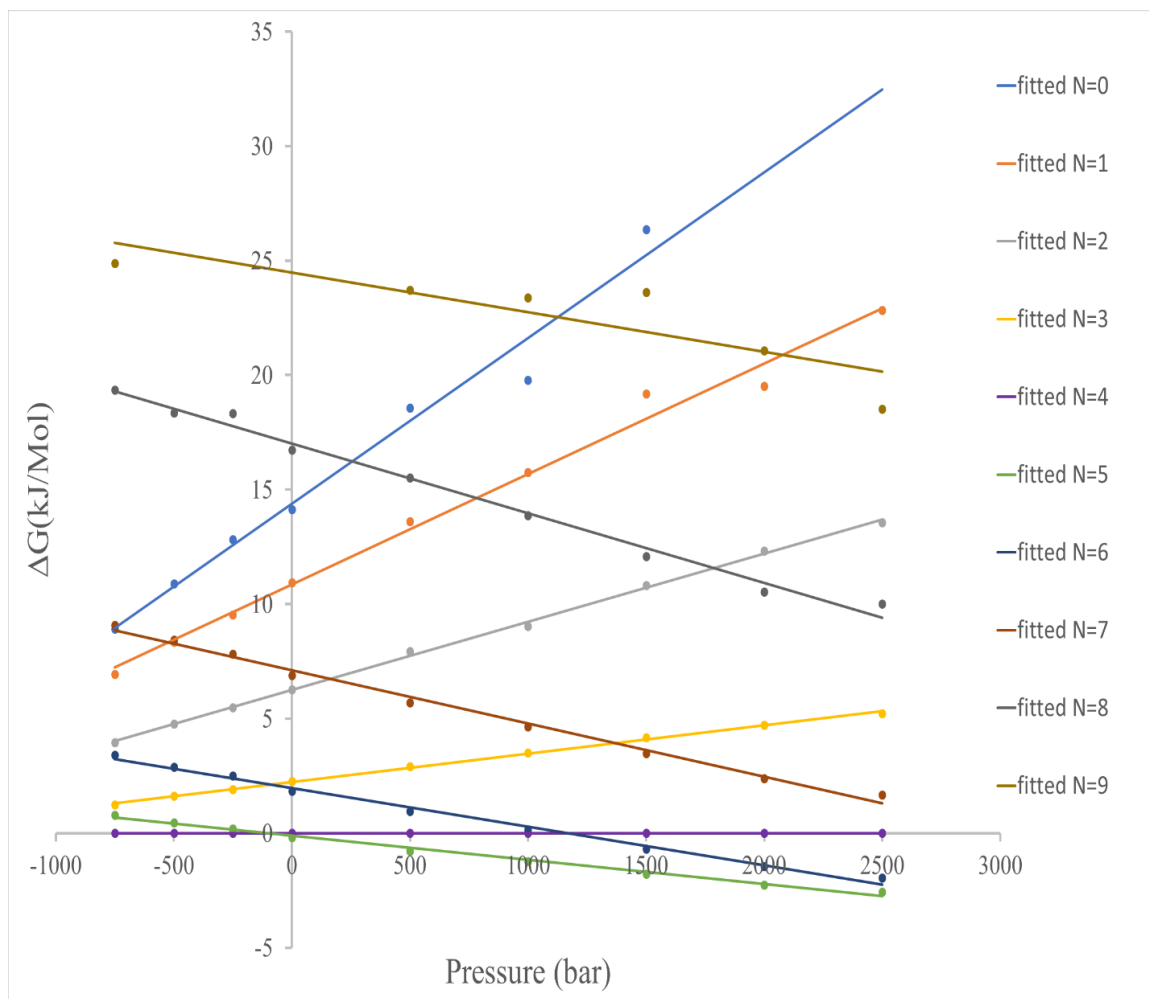
FigureD 15: Change of free energy from certain states and the reference state with different pressures (data points), and the linear trendline fitted to the data (line). The cavitand here is TEHOA, and temperature is 298.15 K. The colors indicating different hydration numbers are identified in the figure legend.
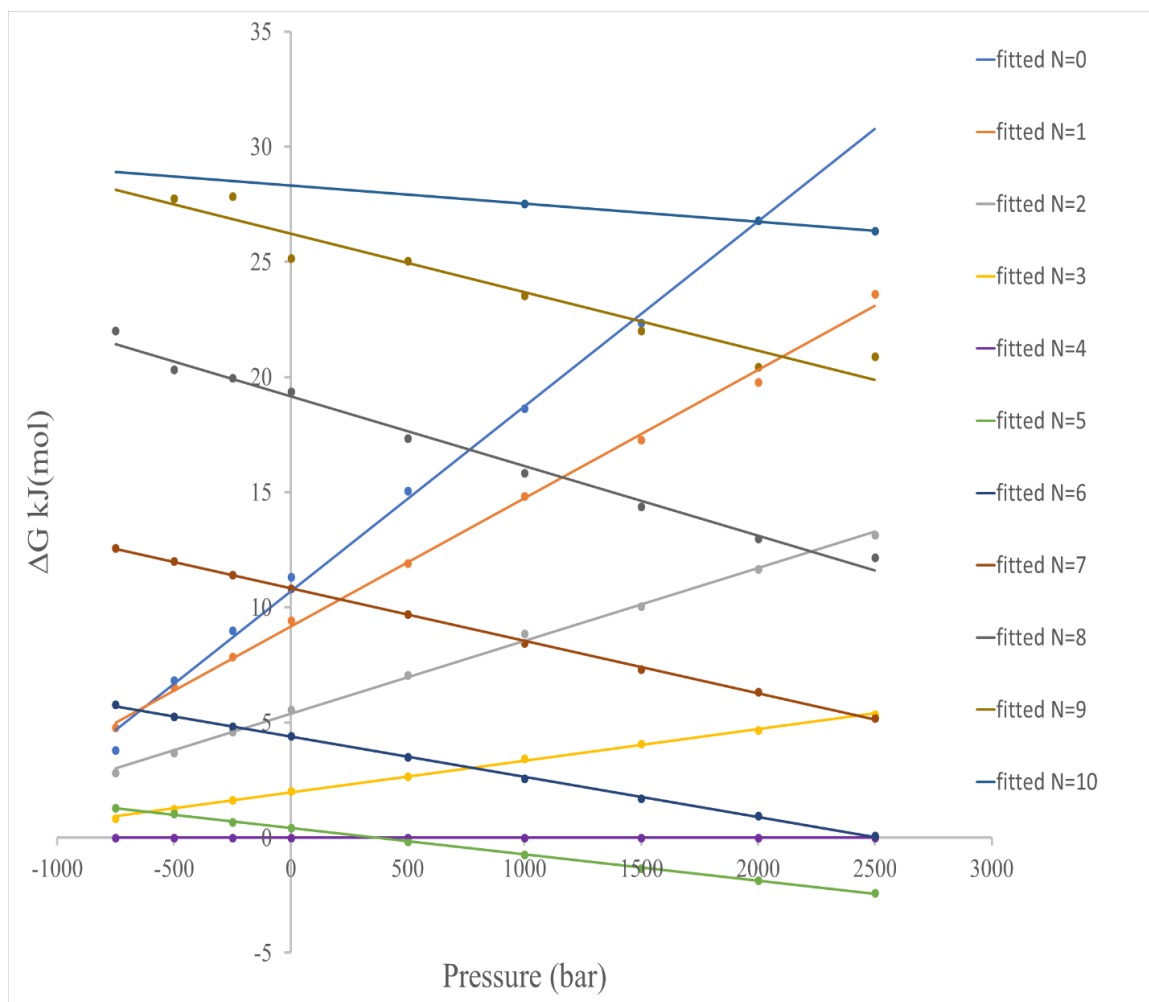
FigureD 16: Change of free energy from certain states and the reference state with different pressures (data points), and the linear trendline fitted to the data (line). The cavitand here is TEXHOA, and temperature is 298.15 K. The colors indicating different hydration numbers are identified in the figure legend.

The figures D17-D24 present the change of volume from a certain hydration state to the reference state (n=4) for all the cavitands. This volume difference is directly got from the subtraction of two simulation box volumes.
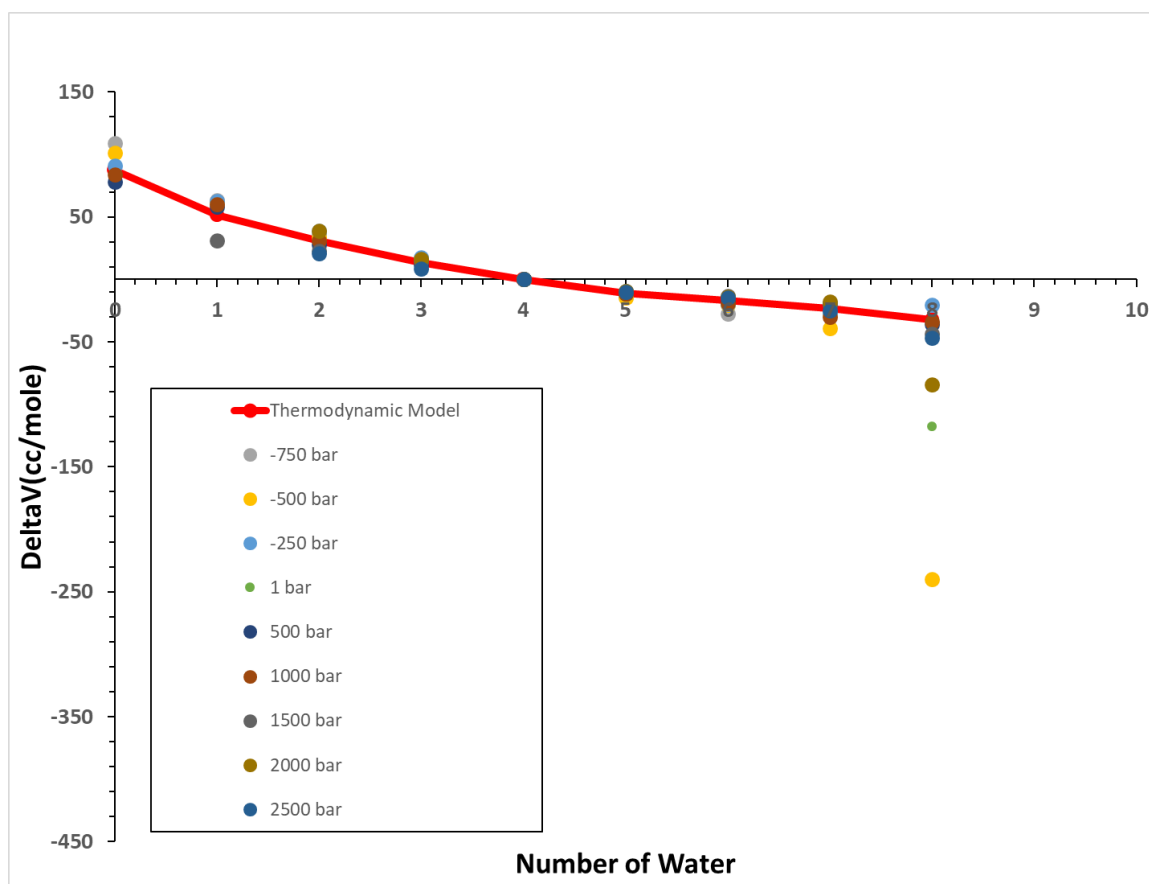
FigureD 17: The change of system volume with OA in water from all hydration states to reference state as a function of number of water molecules in pocket. The data points are directly obtained from simulation data in cubic centimeters per mole of cavitand measured in Gromacs. The colors indicating different pressures are identified in the figure legend. The thick red curve is the change of volume from the slopes of fitted model predictions.
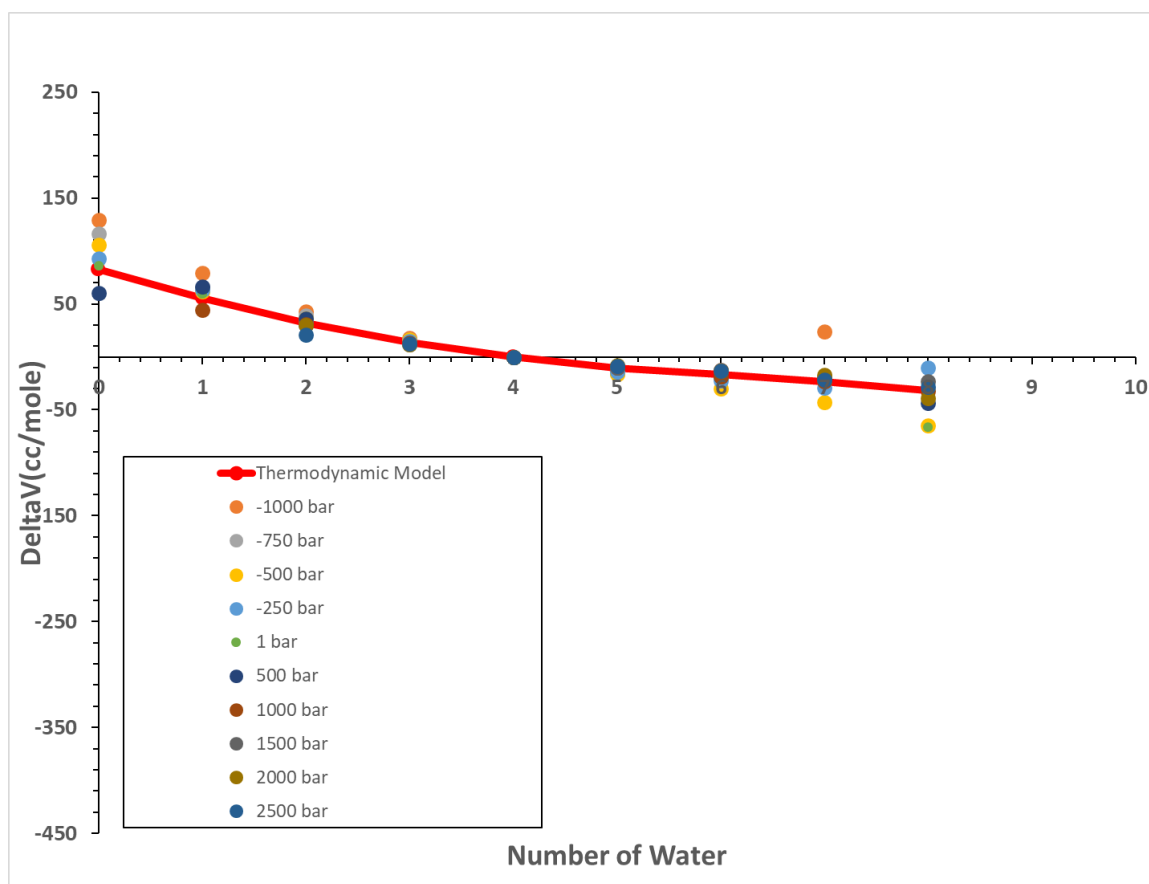
FigureD 18: The change of system volume with MEMOA in water from all hydration states to reference state as a function of number of water molecules in pocket. The data points are directly obtained from simulation data in cubic centimeters per mole of cavitand measured in Gromacs. The colors indicating different pressures are identified in the figure legend. The thick red curve is the change of volume from the slopes of fitted model predictions.
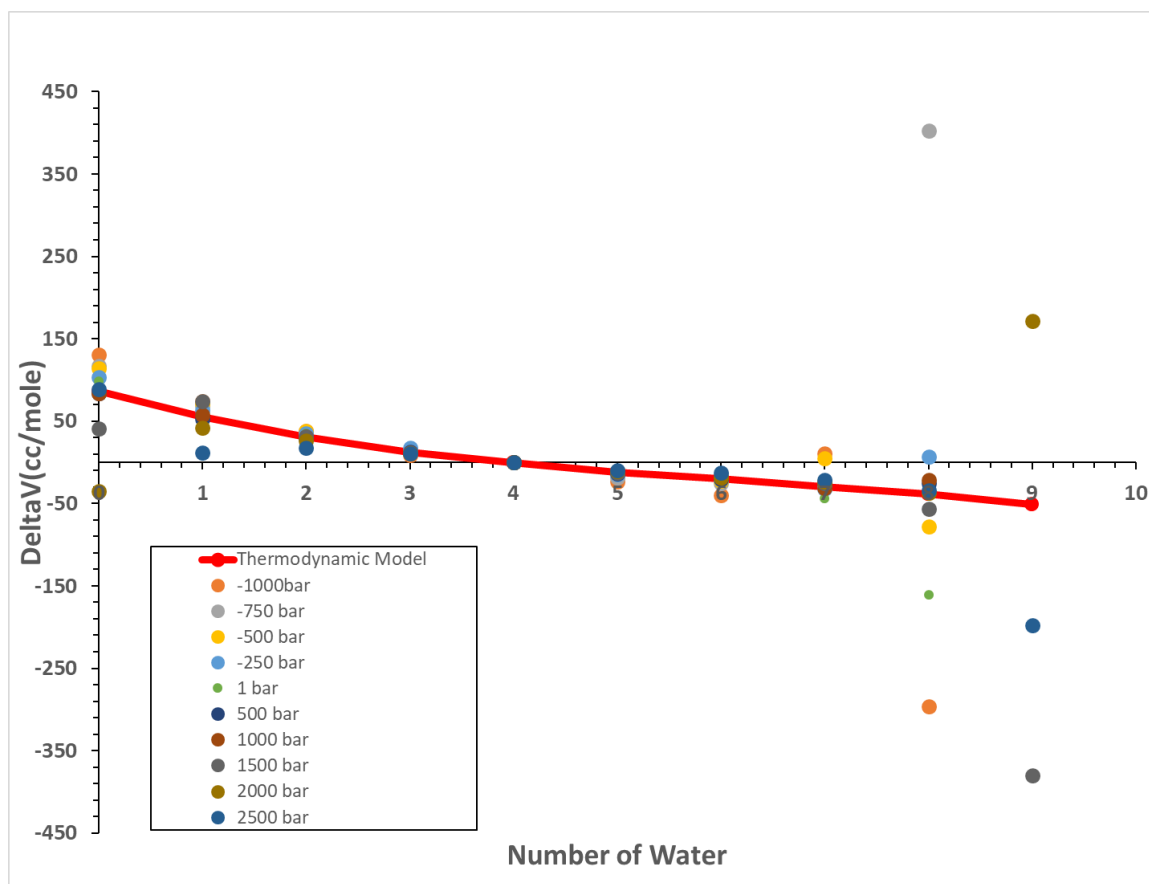
FigureD 19: The change of system volume with DEMOA in water from all hydration states to reference state as a function of number of water molecules in pocket. The data points are directly obtained from simulation data in cubic centimeters per mole of cavitand measured in Gromacs. The colors indicating different pressures are identified in the figure legend. The thick red curve is the change of volume from the slopes of fitted model predictions.
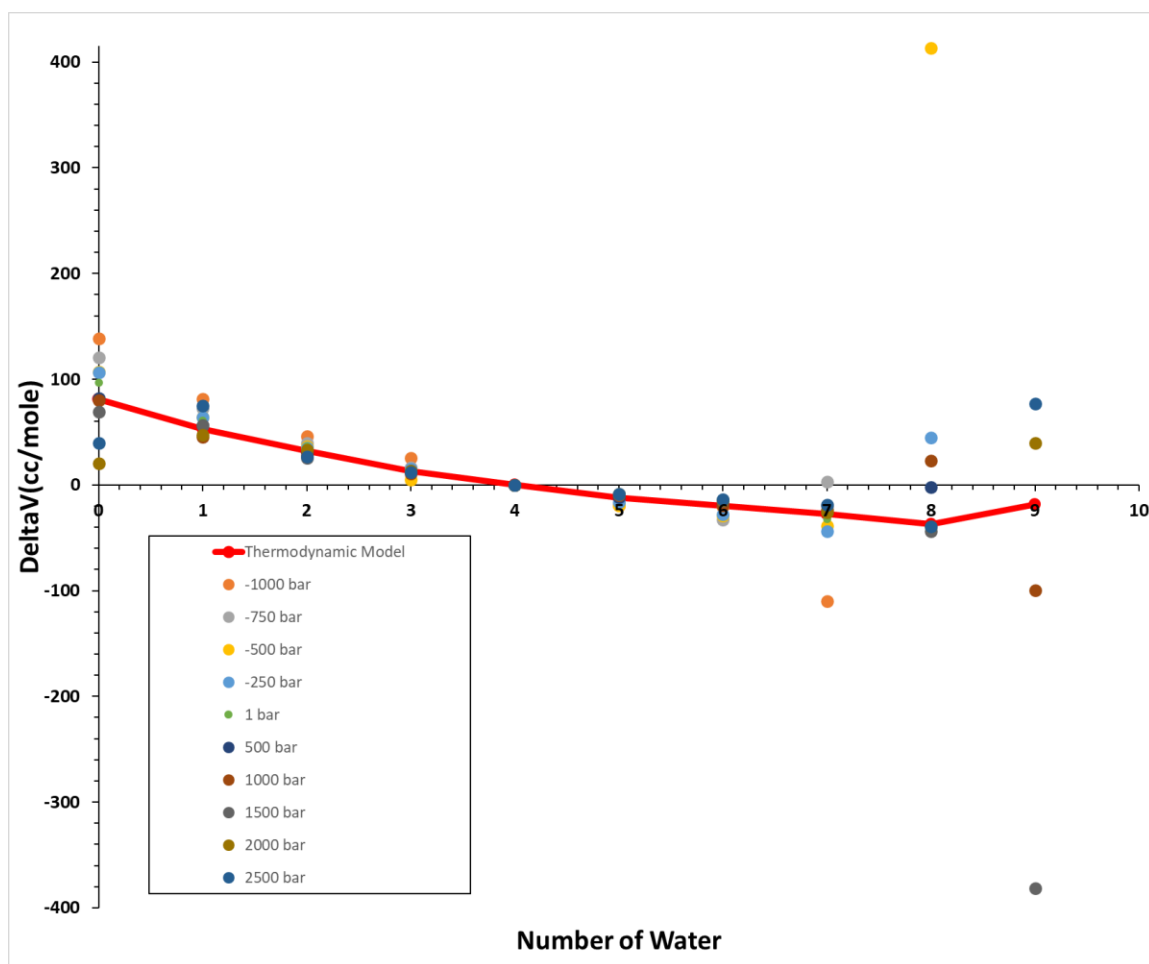
FigureD 20: The change of system volume with TriEMOA in water from all hydration states to reference state as a function of number of water molecules in pocket. The data points are directly obtained from simulation data in cubic centimeters per mole of cavitand measured in Gromacs. The colors indicating different pressures are identified in the figure legend. The thick red curve is the change of volume from the slopes of fitted model predictions.
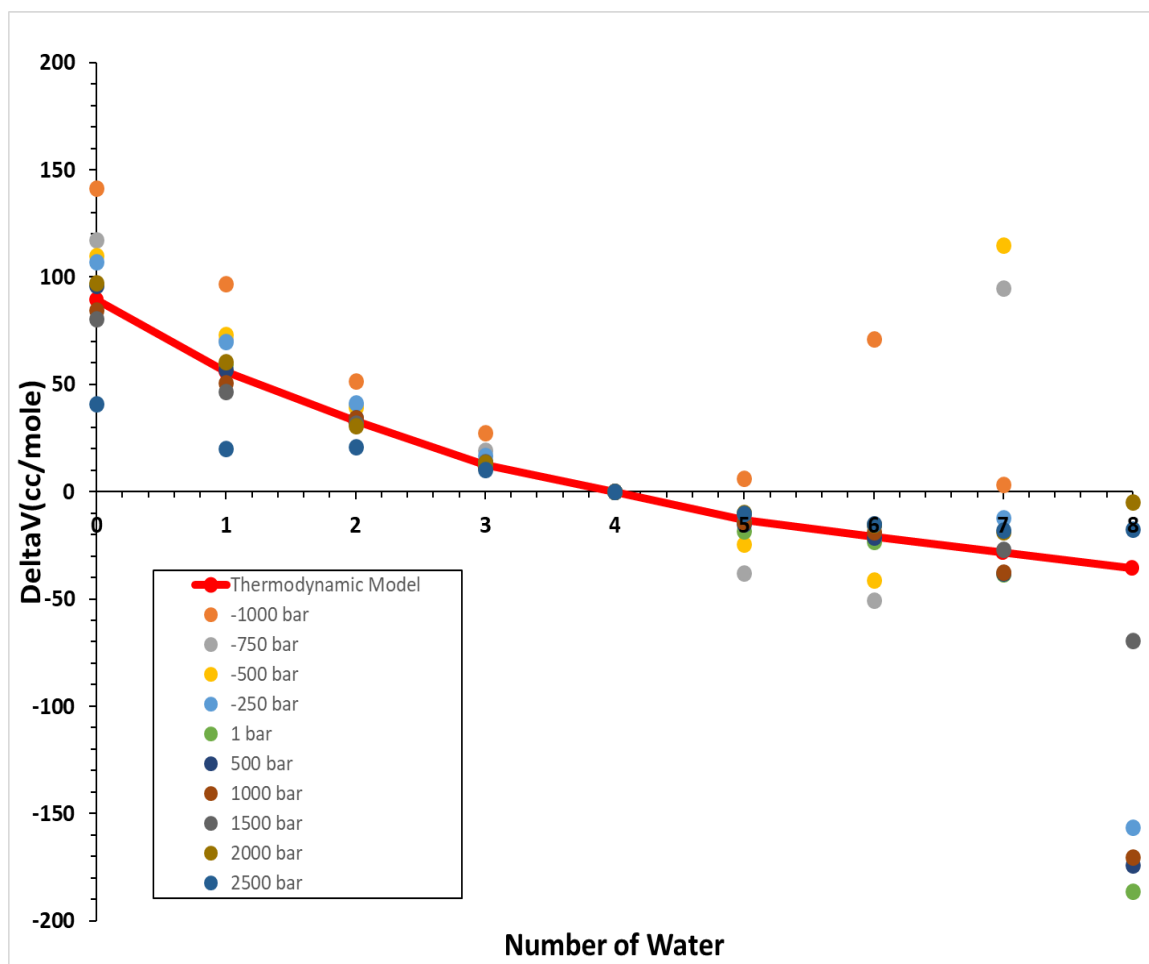
FigureD 21: The change of system volume with TEMOA in water from all hydration states to reference state as a function of number of water molecules in pocket. The data points are directly obtained from simulation data in cubic centimeters per mole of cavitand measured in Gromacs. The colors indicating different pressures are identified in the figure legend. The thick red curve is the change of volume from the slopes of fitted model predictions.
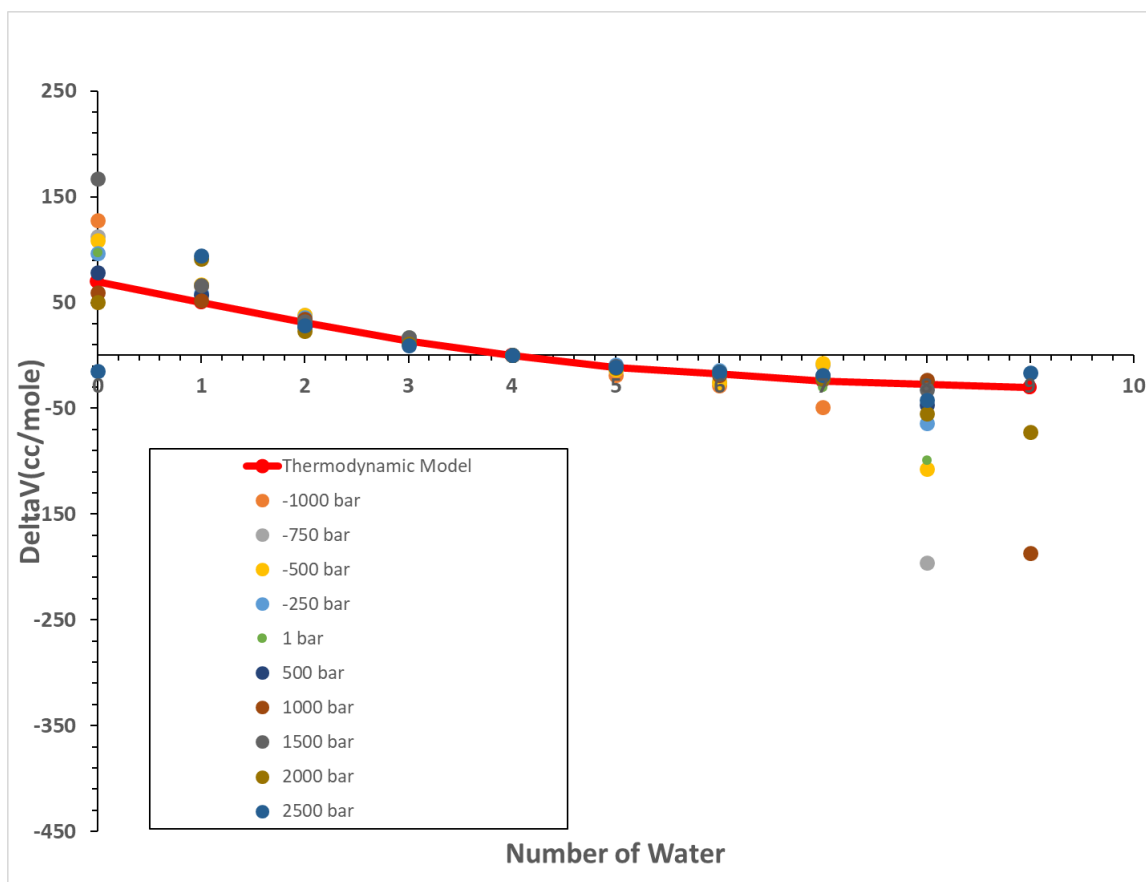
FigureD 22: The change of system volume with TEXMOA in water from all hydration states to reference state as a function of number of water molecules in pocket. The data points are directly obtained from simulation data in cubic centimeters per mole of cavitand measured in Gromacs. The colors indicating different pressures are identified in the figure legend. The thick red curve is the change of volume from the slopes of fitted model predictions.
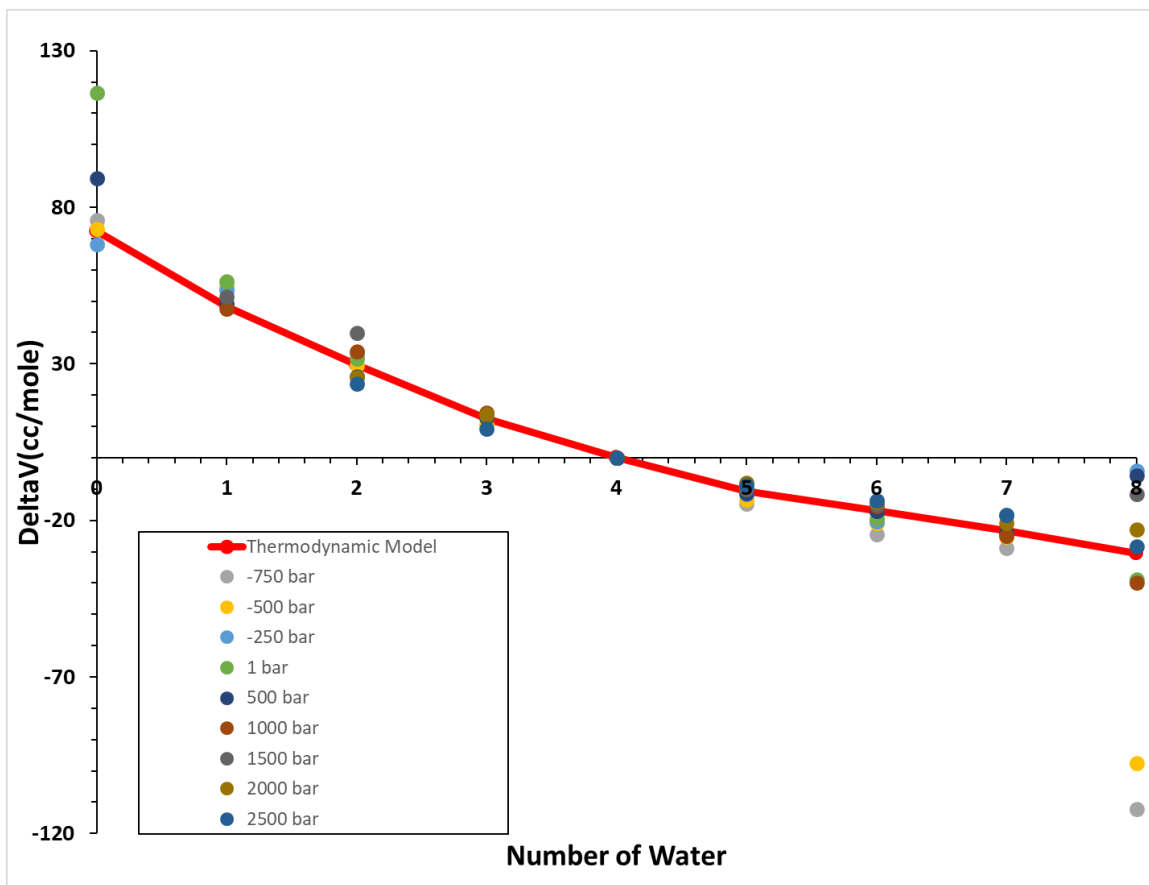
FigureD 23: The change of system volume with TEHOA in water from all hydration states to reference state as a function of number of water molecules in pocket. The data points are directly obtained from simulation data in cubic centimeters per mole of cavitand measured in Gromacs. The colors indicating different pressures are identified in the figure legend. The thick red curve is the change of volume from the slopes of fitted model predictions.
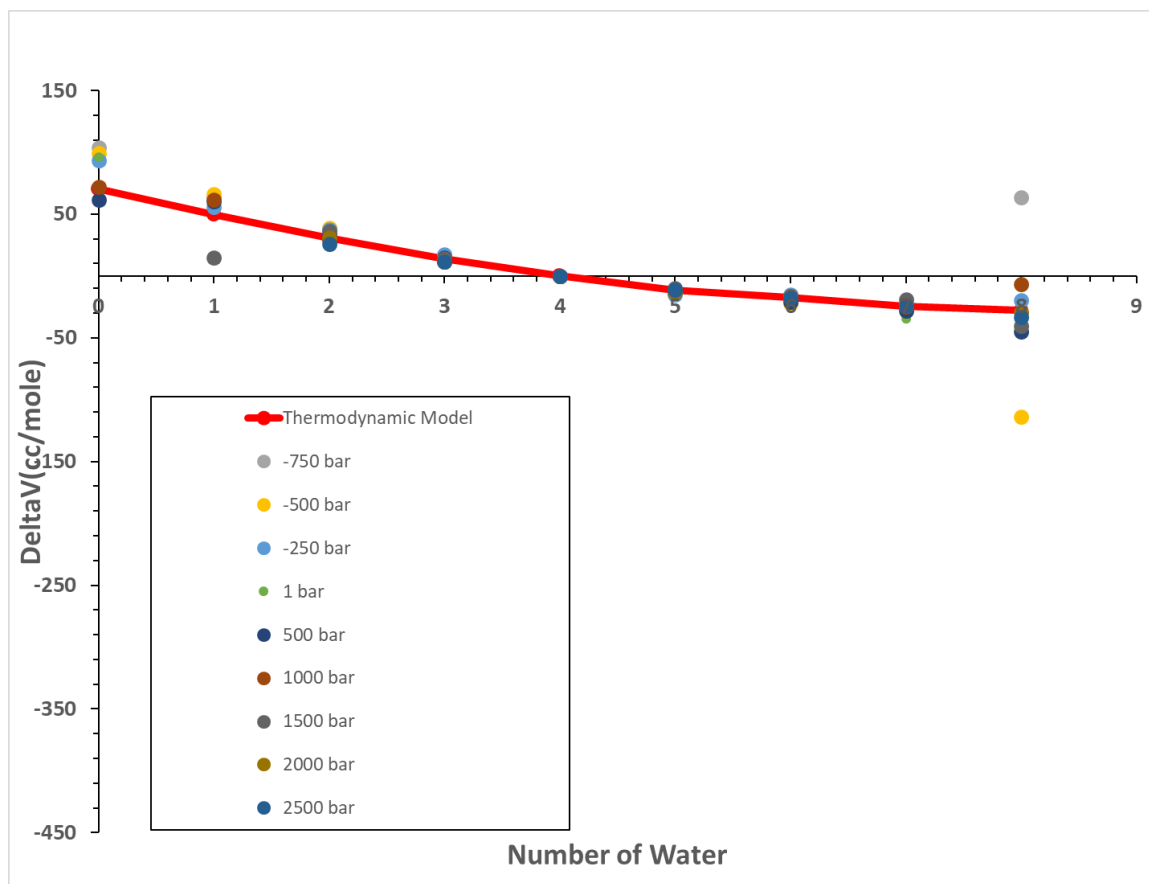
FigureD 24: The change of system volume with TEXHOA in water from all hydration states to reference state as a function of number of water molecules in pocket. The data points are directly obtained from simulation data in cubic centimeters per mole of cavitand measured in Gromacs. The colors indicating different pressures are identified in the figure legend. The thick red curve is the change of volume from the slopes of fitted model predictions.

The comparisons between the fitted probability distributions of hydration numbers (the curve) and the same distributions directly from simulation (the data points) of all the cavitands are presented here in figure D25-D32.
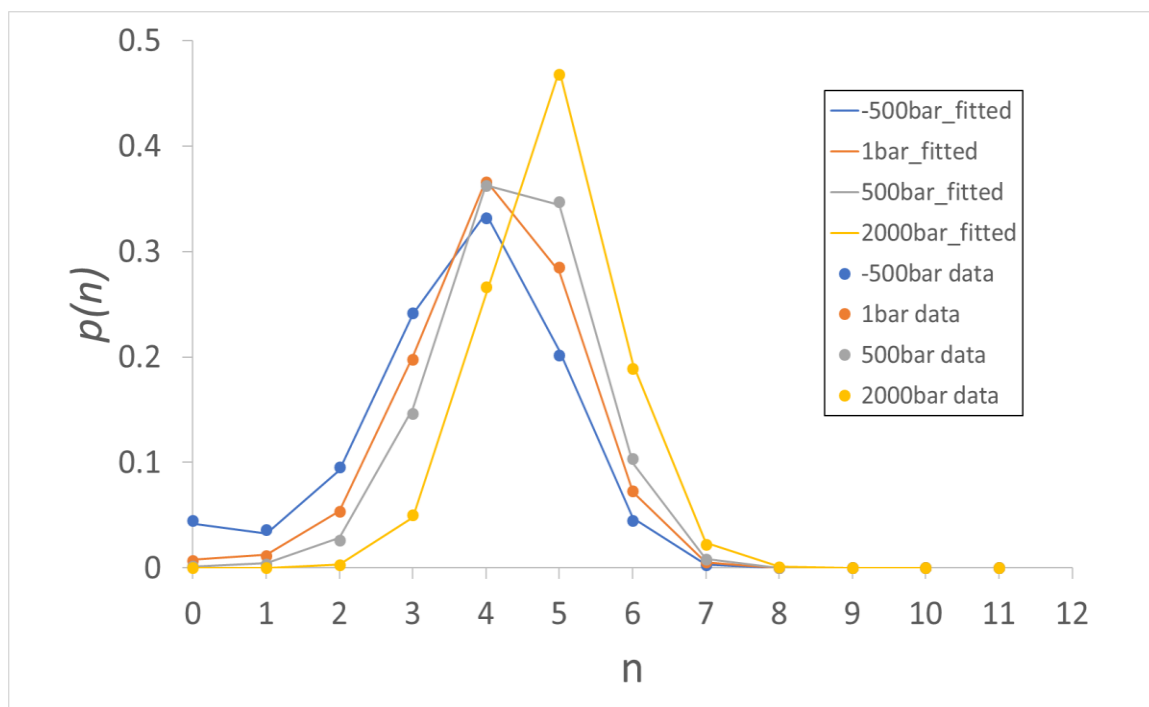
FigureD 25: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of OA as a function of hydration number, n. The data directly got from simulation (the data points) and the distribution evaluated from the fitted thermodynamic model (the curve) are both illustrated and compared with each other in the figure. The pressures are -500 bar, 1 bar, 500 bar, and 2000 bar. The colors indicating different pressures are identified in the figure legend.
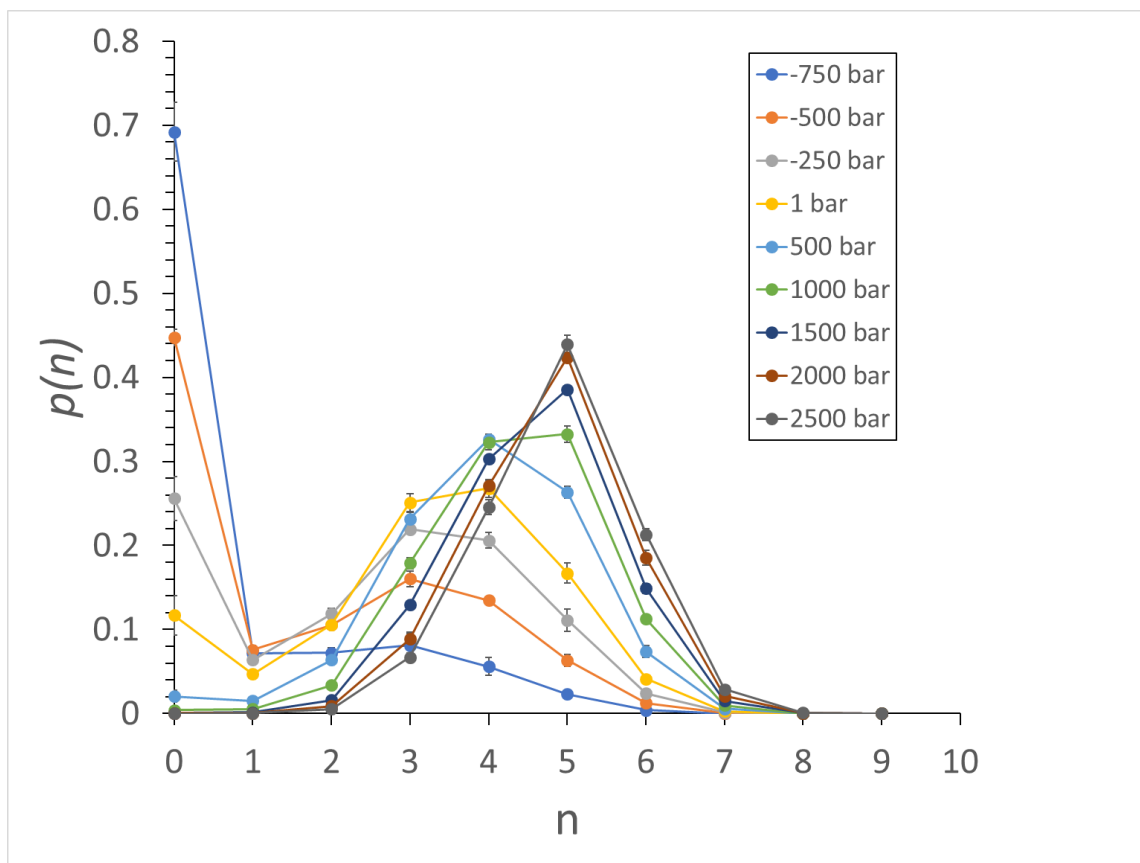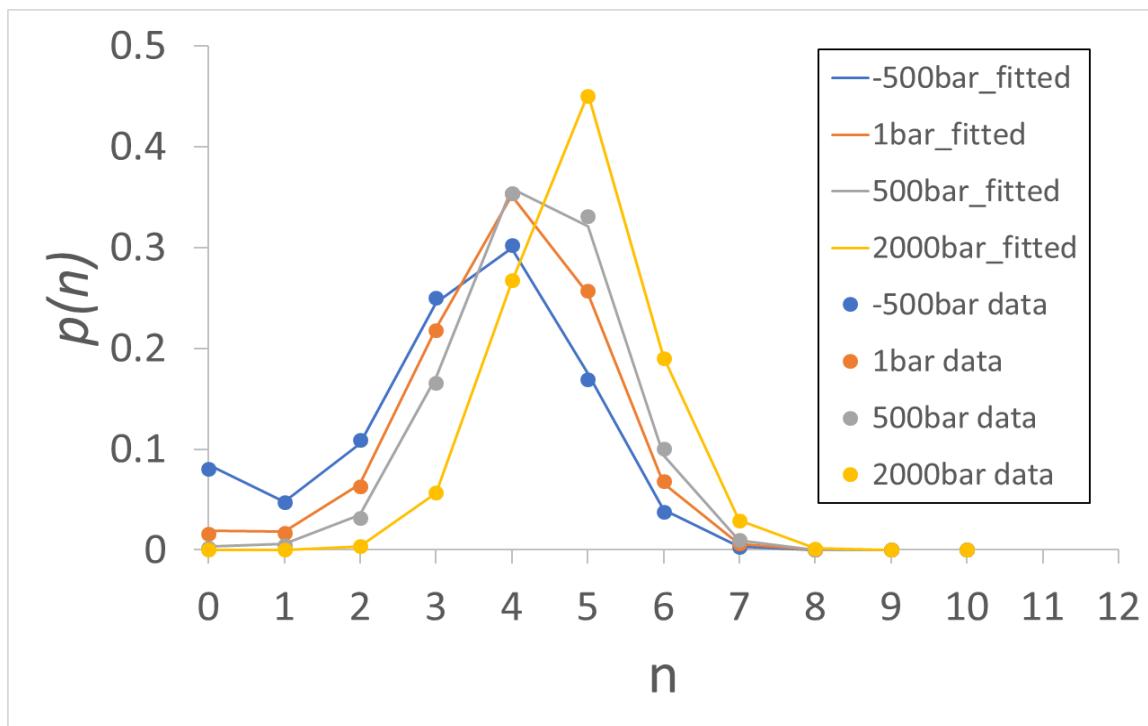
FigureD 26: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of MEMOA as a function of hydration number, n. The data directly got from simulation (the data points) and the distribution evaluated from the fitted thermodynamic model (the curve) are both illustrated and compared with each other in the figure. The pressures are -500 bar, 1 bar, 500 bar, and 2000 bar. The colors indicating different pressures are identified in the figure legend.
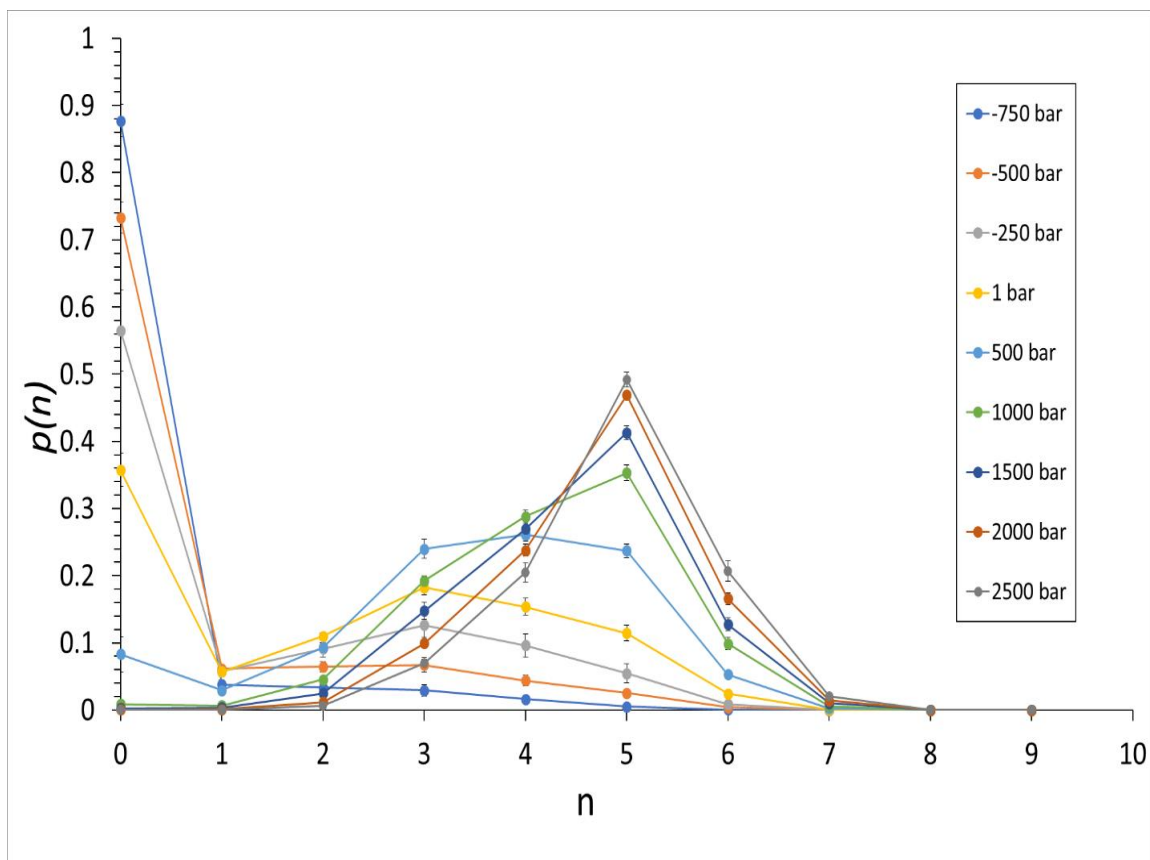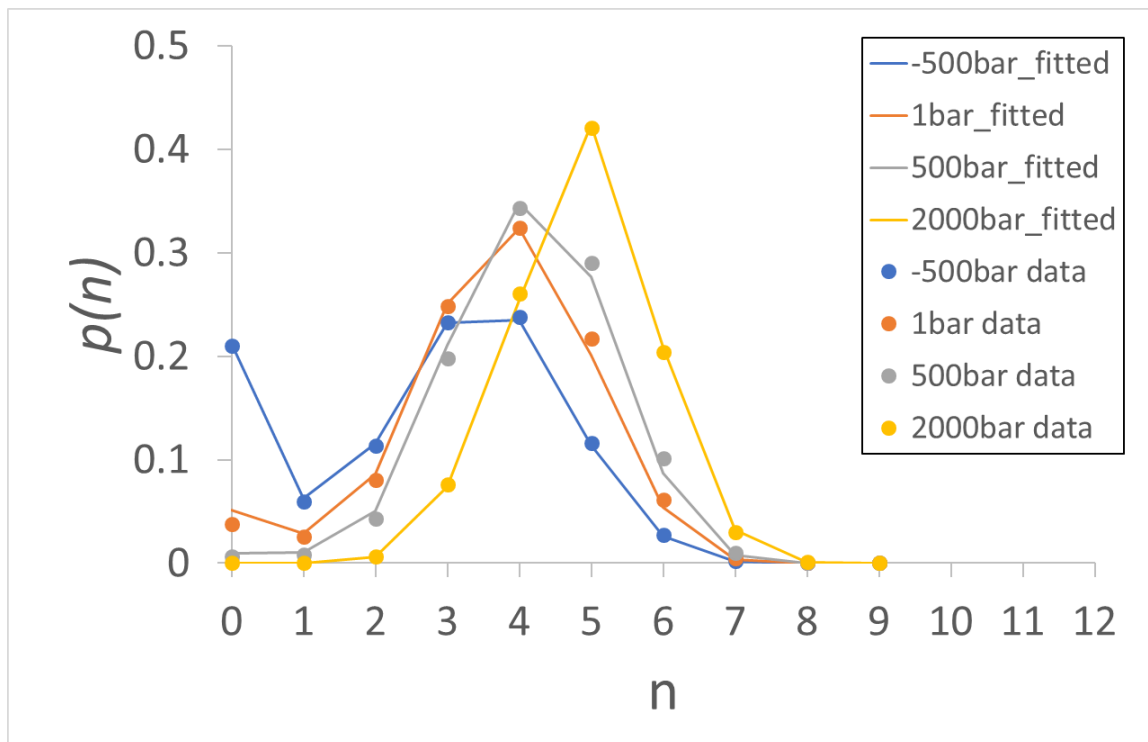
FigureD 27: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of DEMOA as a function of hydration number, n. The data directly got from simulation (the data points) and the distribution evaluated from the fitted thermodynamic model (the curve) are both illustrated and compared with each other in the figure. The pressures are -500 bar, 1 bar, 500 bar, and 2000 bar. The colors indicating different pressures are identified in the figure legend.
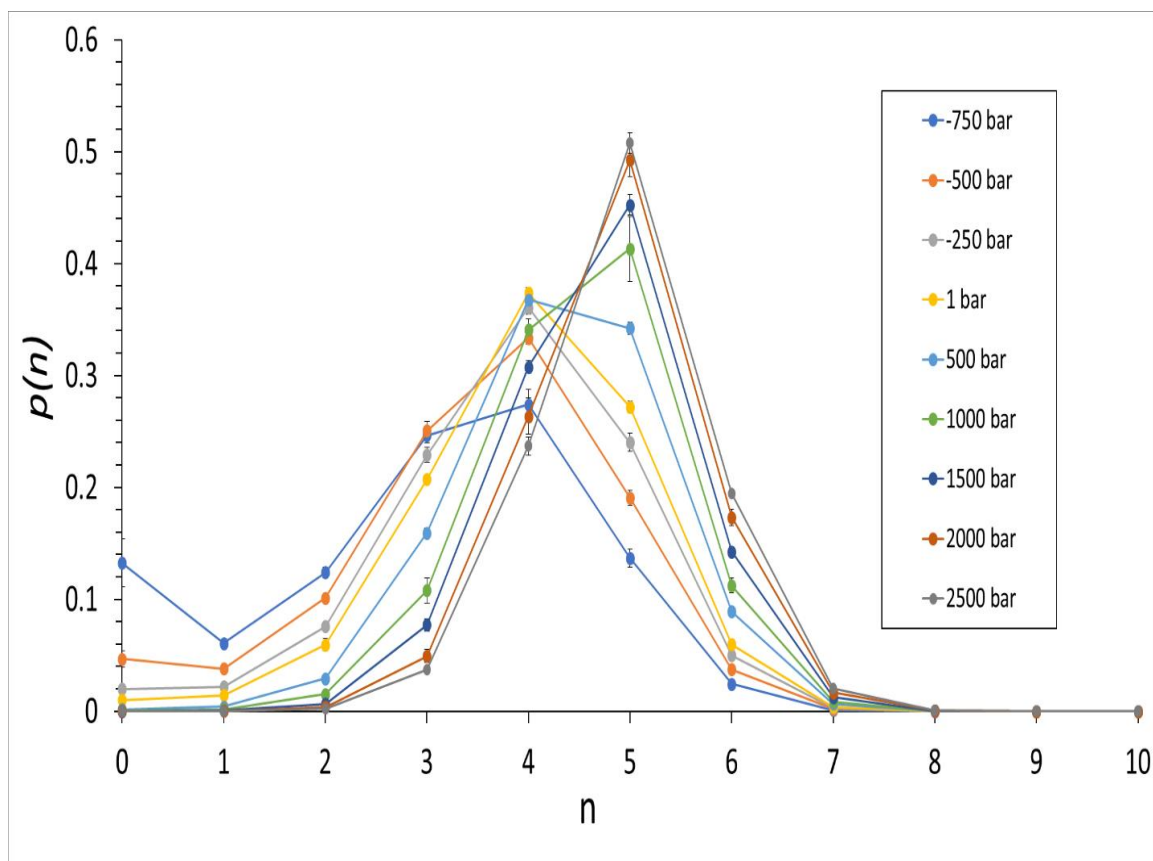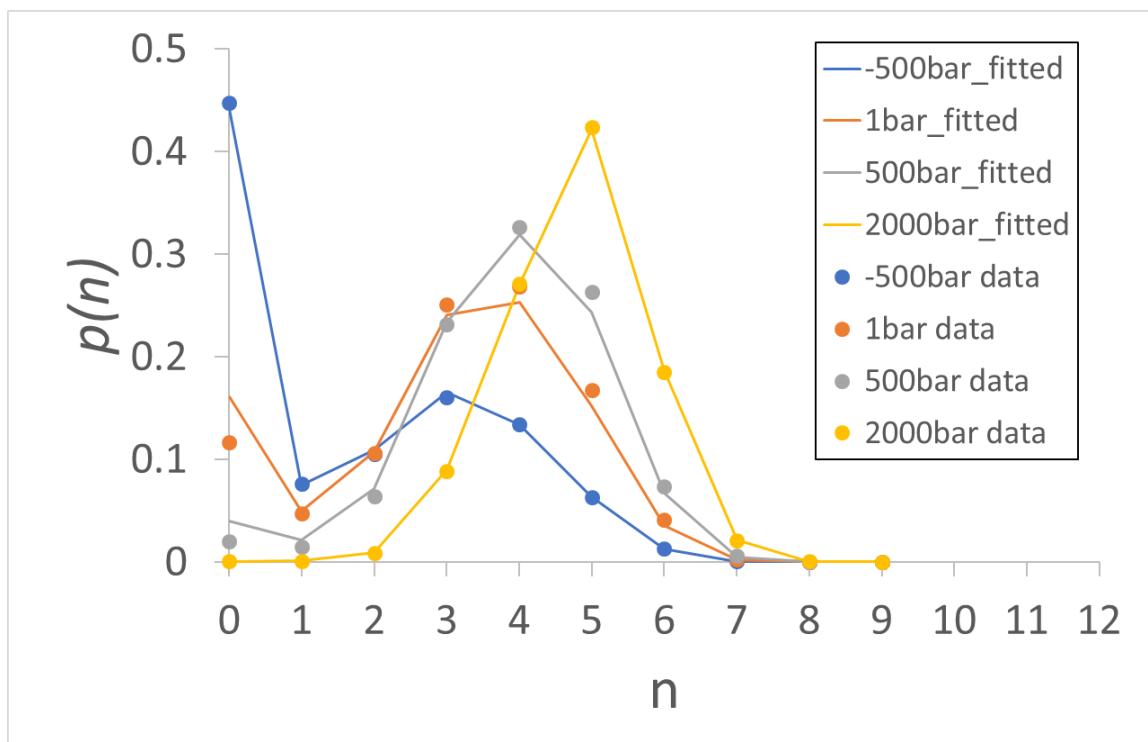
FigureD 28: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TriEMOA as a function of hydration number, n. The data directly got from simulation (the data points) and the distribution evaluated from the fitted thermodynamic model (the curve) are both illustrated and compared with each other in the figure. The pressures are -500 bar, 1 bar, 500 bar, and 2000 bar. The colors indicating different pressures are identified in the figure legend.
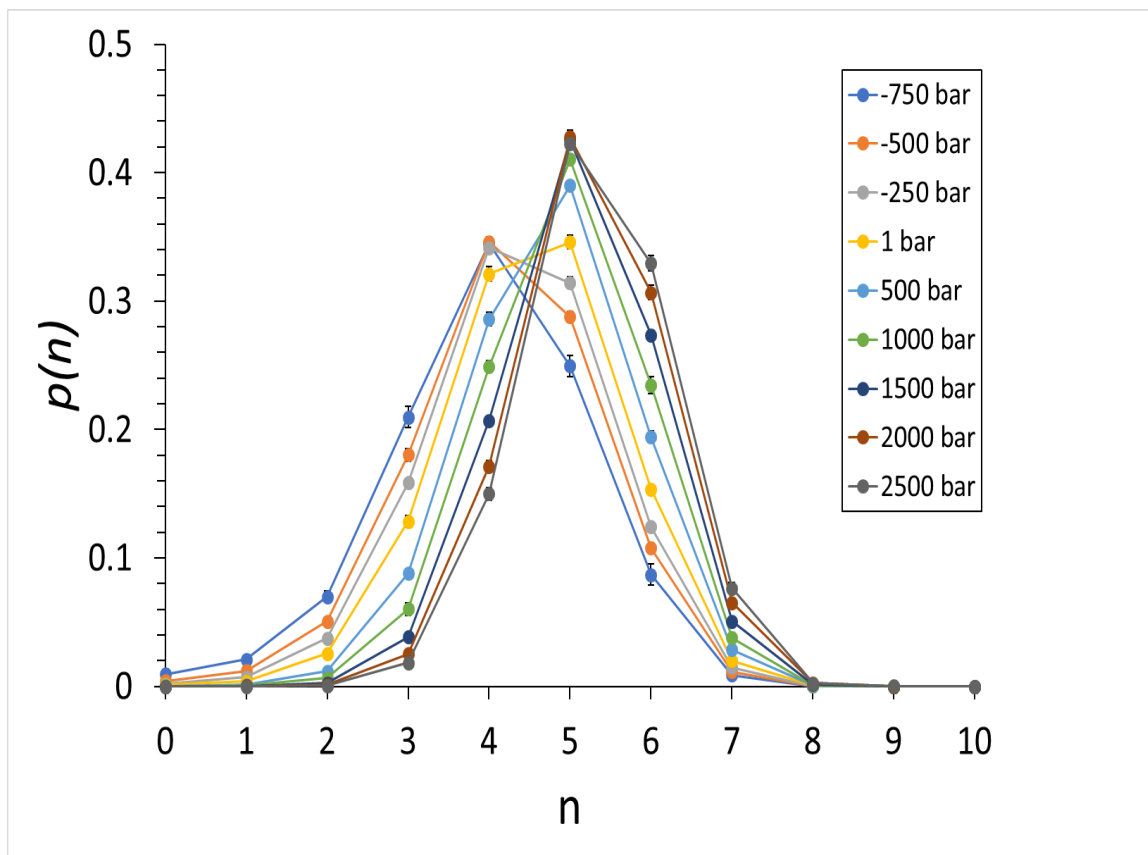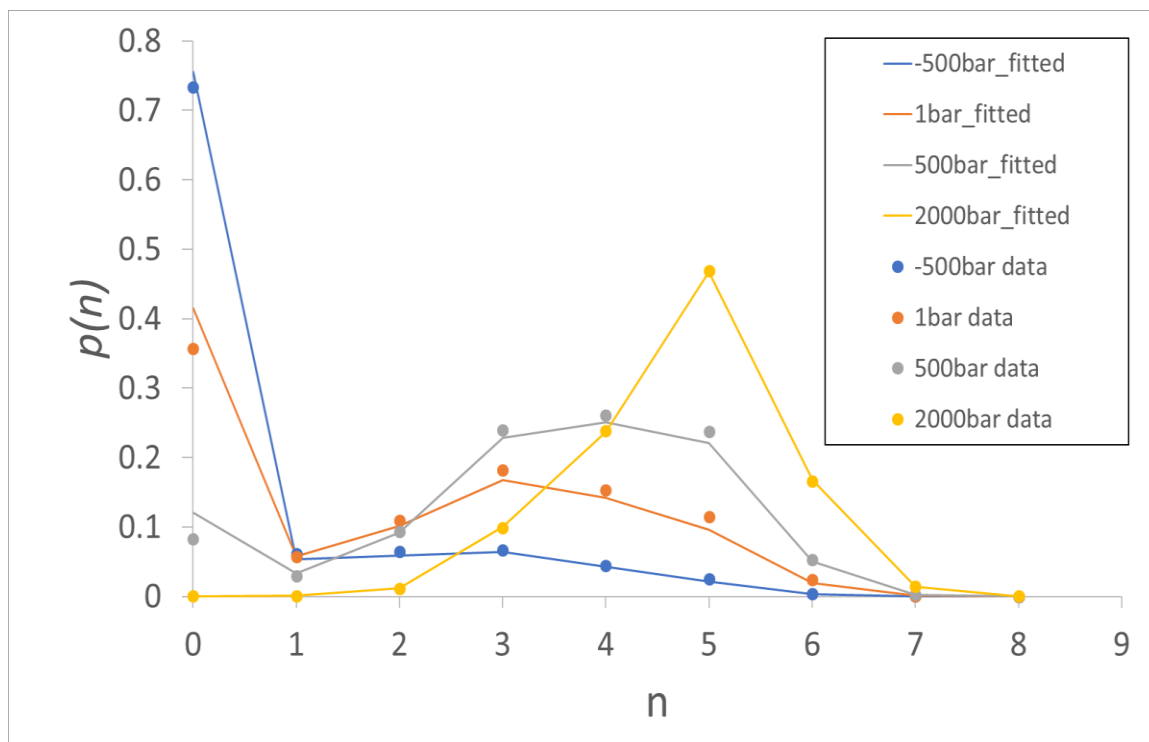
FigureD 29: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TEMOA as a function of hydration number, n. The data directly got from simulation (the data points) and the distribution evaluated from the fitted thermodynamic model (the curve) are both illustrated and compared with each other in the figure. The pressures are -500 bar, 1 bar, 500 bar, and 2000 bar. The colors indicating different pressures are identified in the figure legend.
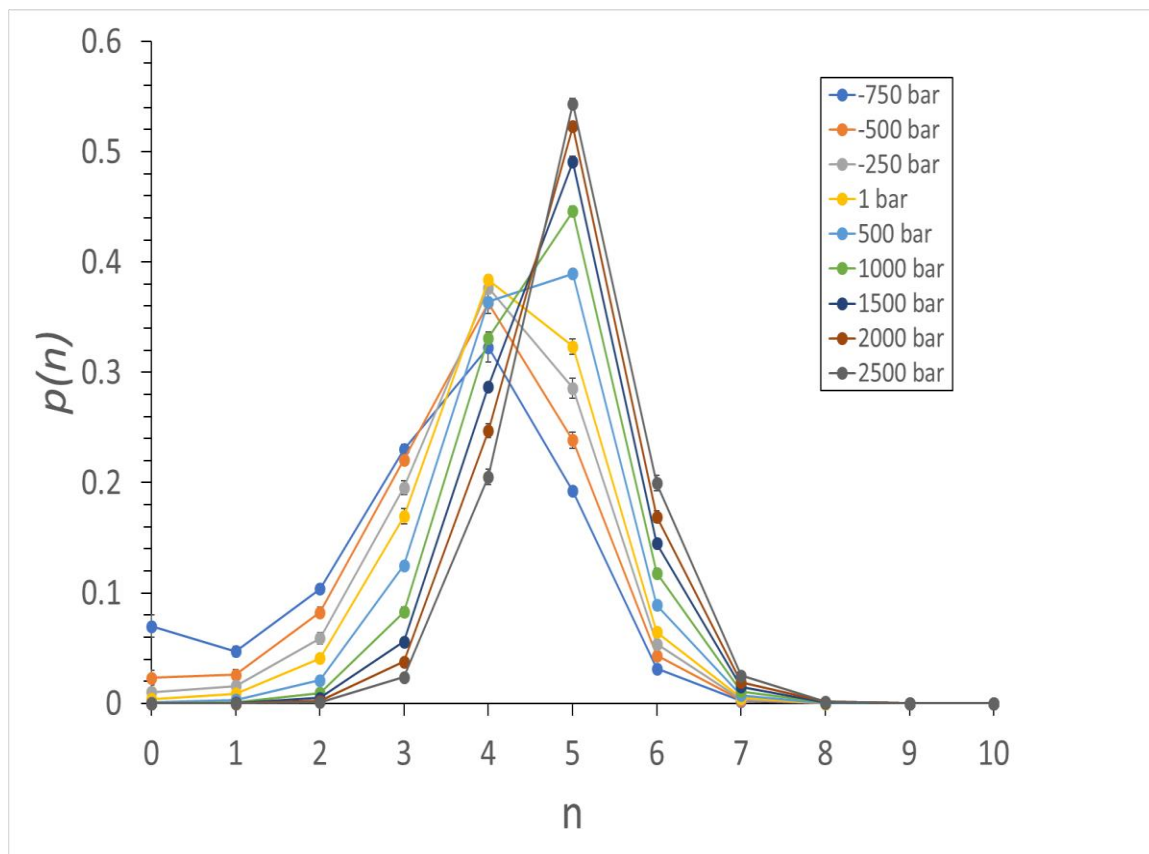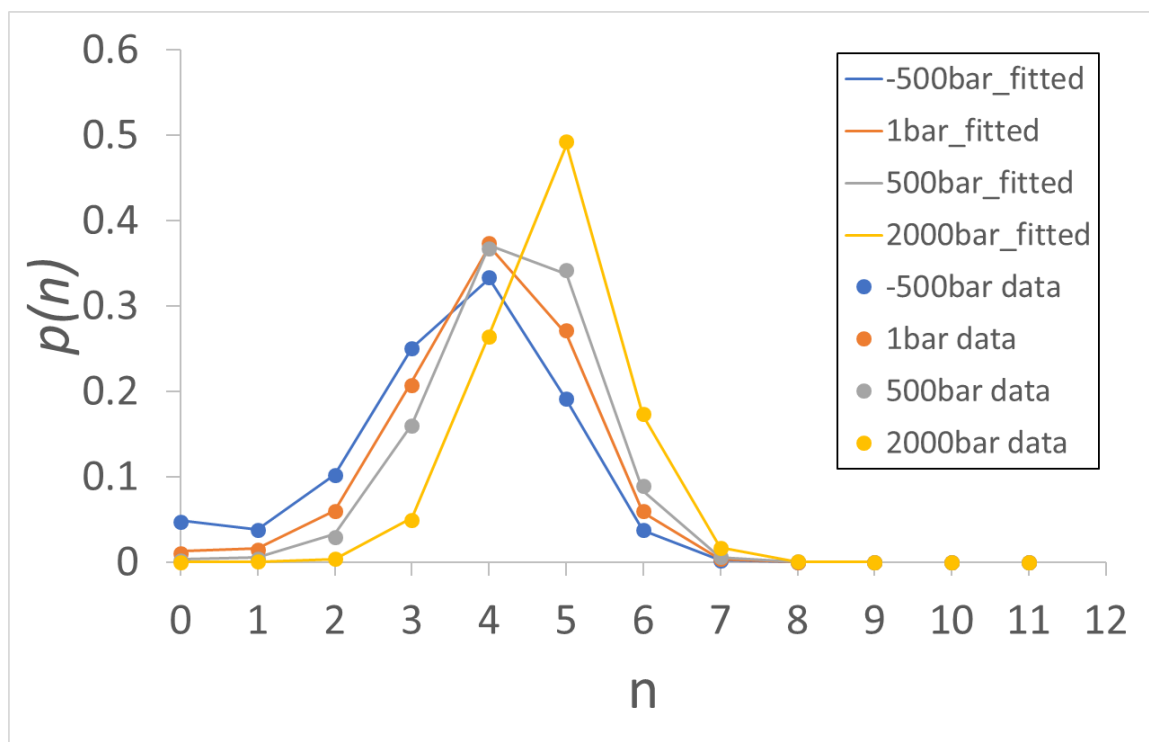
FigureD 30: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TEXMOA as a function of hydration number, n. The data directly got from simulation (the data points) and the distribution evaluated from the fitted thermodynamic model (the curve) are both illustrated and compared with each other in the figure. The pressures are -500 bar, 1 bar, 500 bar, and 2000 bar. The colors indicating different pressures are identified in the figure legend.
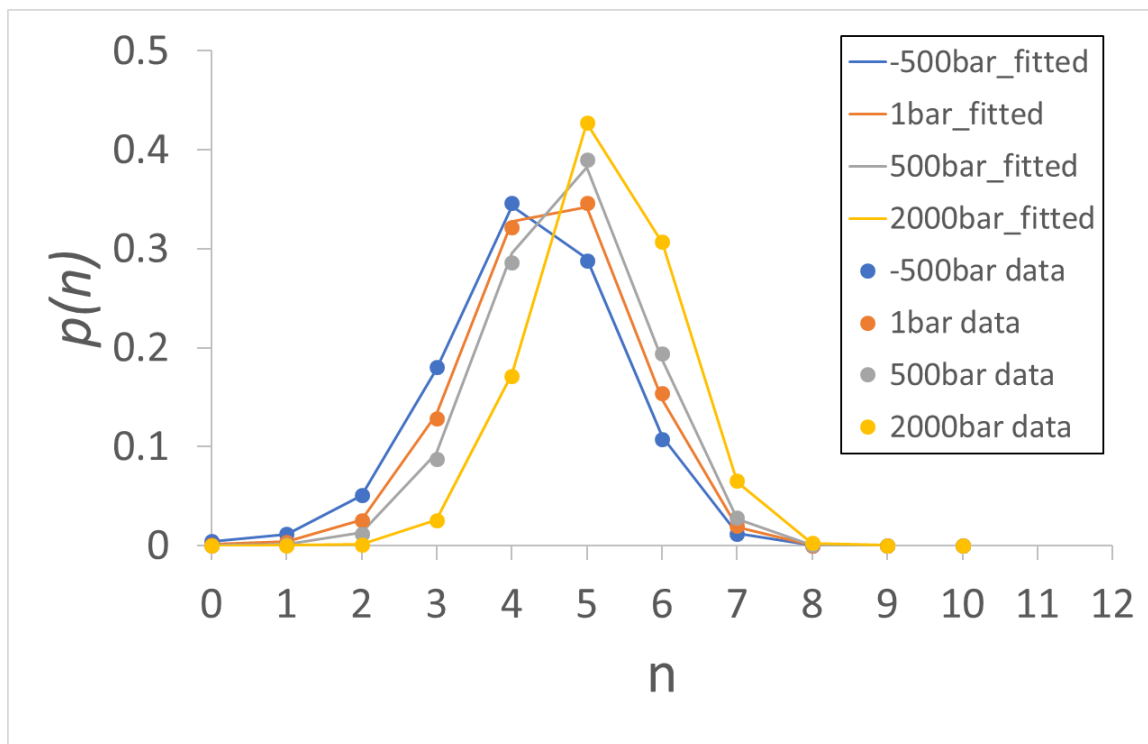
FigureD 31: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TEHOA as a function of hydration number, n. The data directly got from simulation (the data points) and the distribution evaluated from the fitted thermodynamic model (the curve) are both illustrated and compared with each other in the figure. The pressures are -500 bar, 1 bar, 500 bar, and 2000 bar. The colors indicating different pressures are identified in the figure legend.
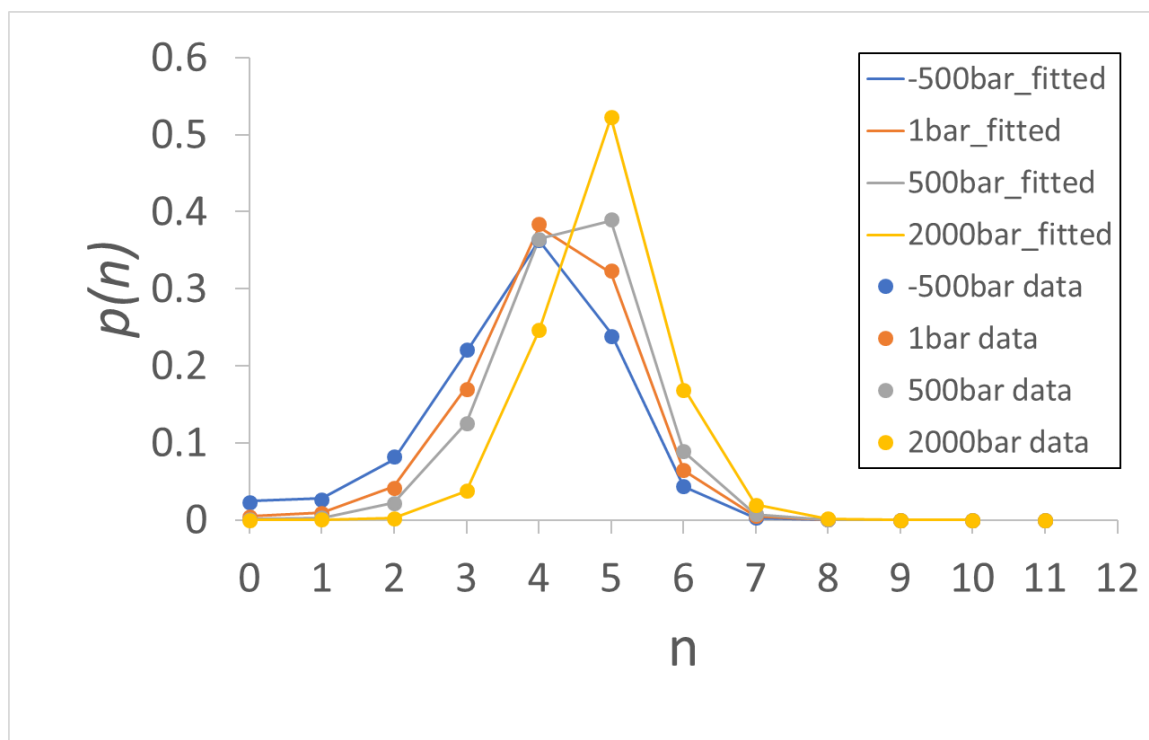
FigureD 32: The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of TEXHOA as a function of hydration number, n. The data directly got from simulation (the data points) and the distribution evaluated from the fitted thermodynamic model (the curve) are both illustrated and compared with each other in the figure. The pressures are -500 bar, 1 bar, 500 bar, and 2000 bar. The colors indicating different pressures are identified in the figure legend.

Appendix E
Supporting information for "Characterization of wetting and dewetting behaviors within deep-cavity cavitand pockets using Indirect Umbrella Sampling (INDUS)"

The water density distributions inside and around cavitand host molecules in cylindrical symmetrical coordinates shown in figure 6.3-6.6 are drawn from a systematical analysis of the responses of waters to unfavorable biasing potentials with gradually increasing strengths. These biasing potentials were centered on each heavy atom of the cavitands. Here figure E1 and E2 illustrate the water density distributions inside and around OA in the same cylindrical coordinate. The strength regulating parameter ϕ ranges from 0 to 10 kJ/mol with a 1 kJ/mol increment. Similarly, figure E3 and E4 show the water density distribution in presence of TEMOA with the same strengths of biasing potentials. In addition, more windows with a finer changing process of ϕ values were simulated to investigate the impact of unfavorable potentials on the hydration of inner pockets. Figure E5 through E7 exhibit the water density distributions inside and around OA with ϕ values from 2.1 kJ/mol to 3.9 kJ/mol with an increment of 0.1 kJ/mol. Correspondingly, figure E8 through E10 report the water density distributions in case of TEMOA with ϕ values from 0.1 kJ/mol to 1.9 kJ/mol.
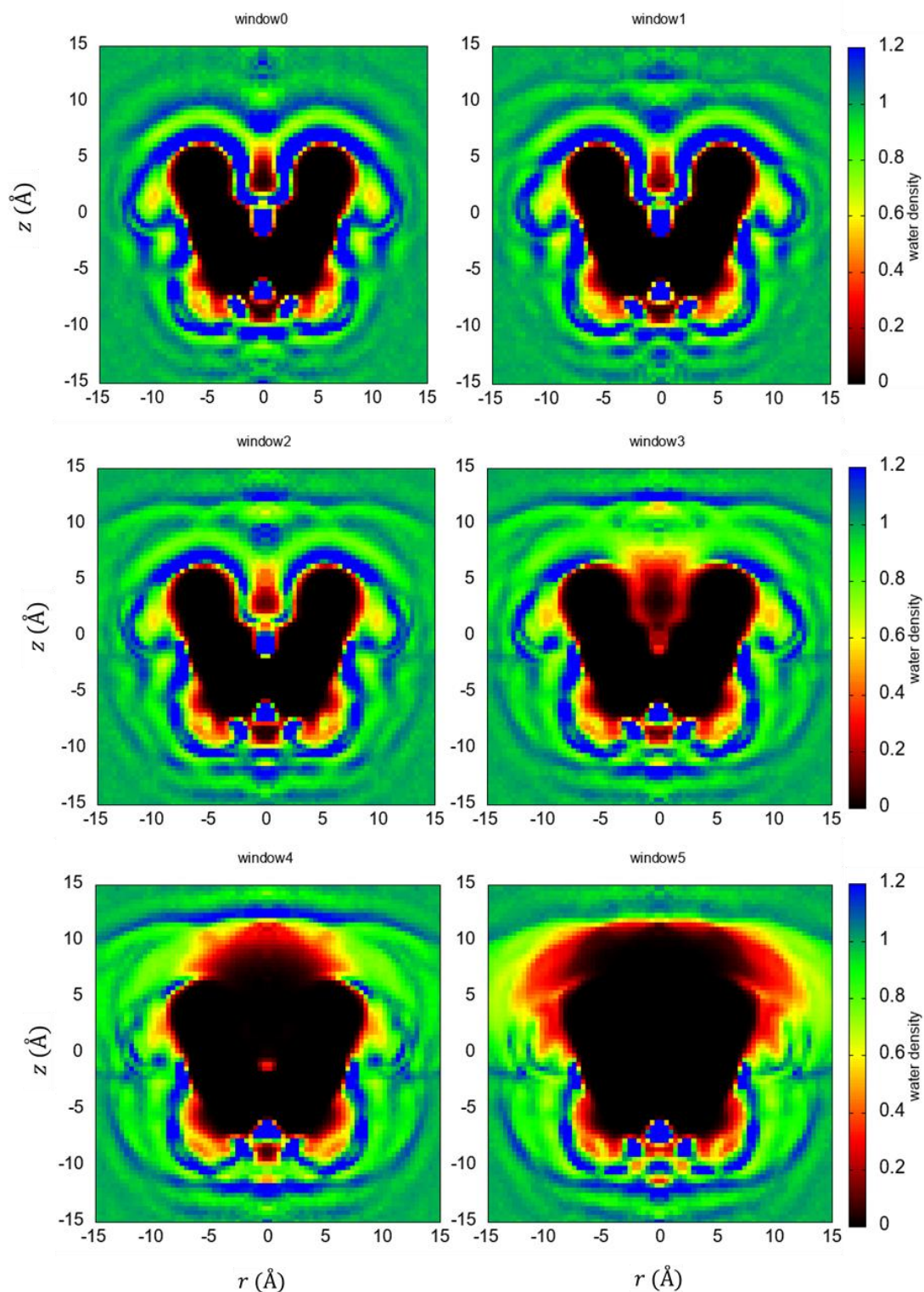
FigureE 1 Average water density distributions about OA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 0-5 kJ/mol with an increment of 1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
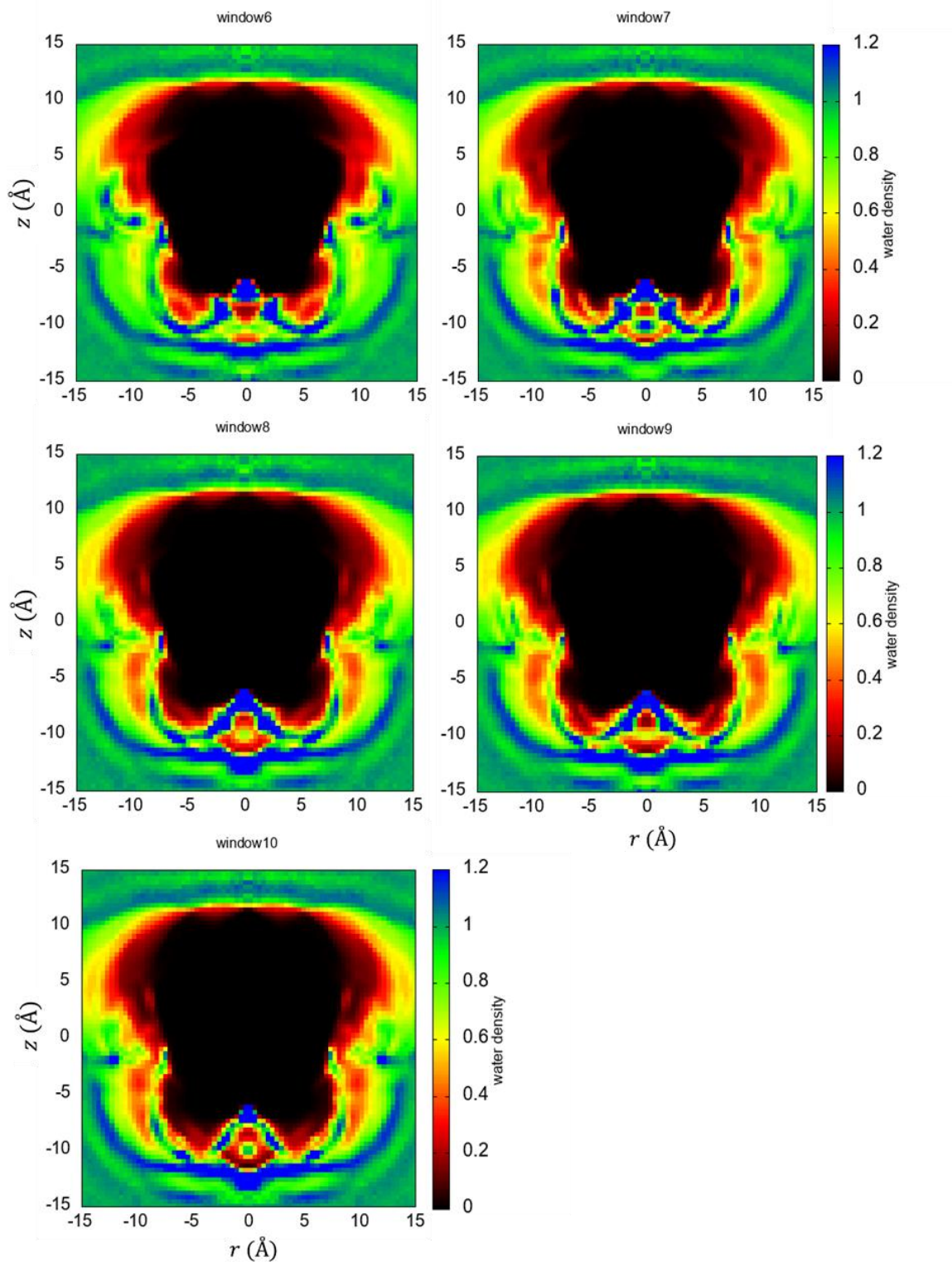
FigureE 2: Average water density distributions about OA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 6-10 kJ/mol with an increment of 1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
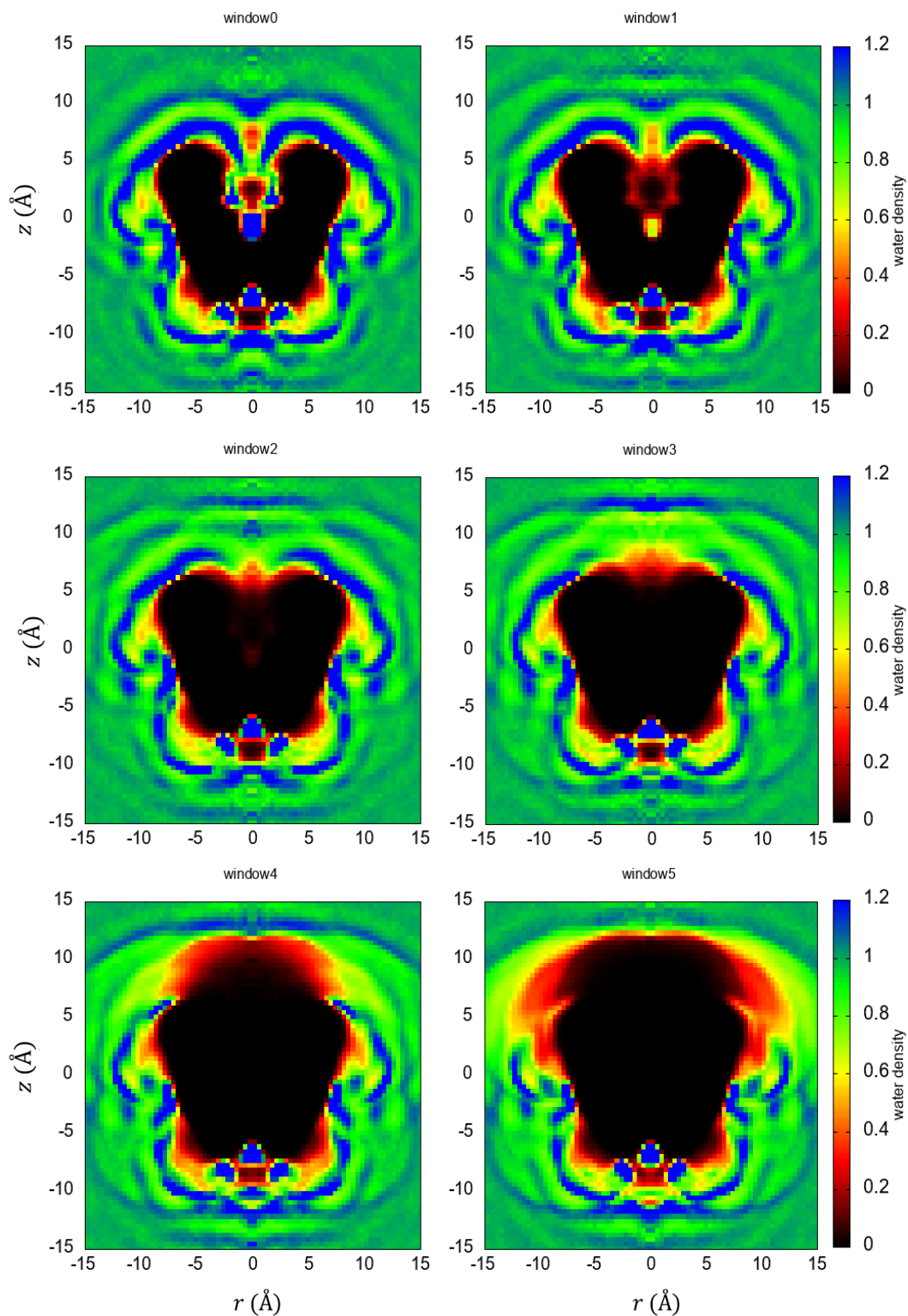
FigureE 3: Average water density distributions about TEMOA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 0-5 kJ/mol with an increment of 1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
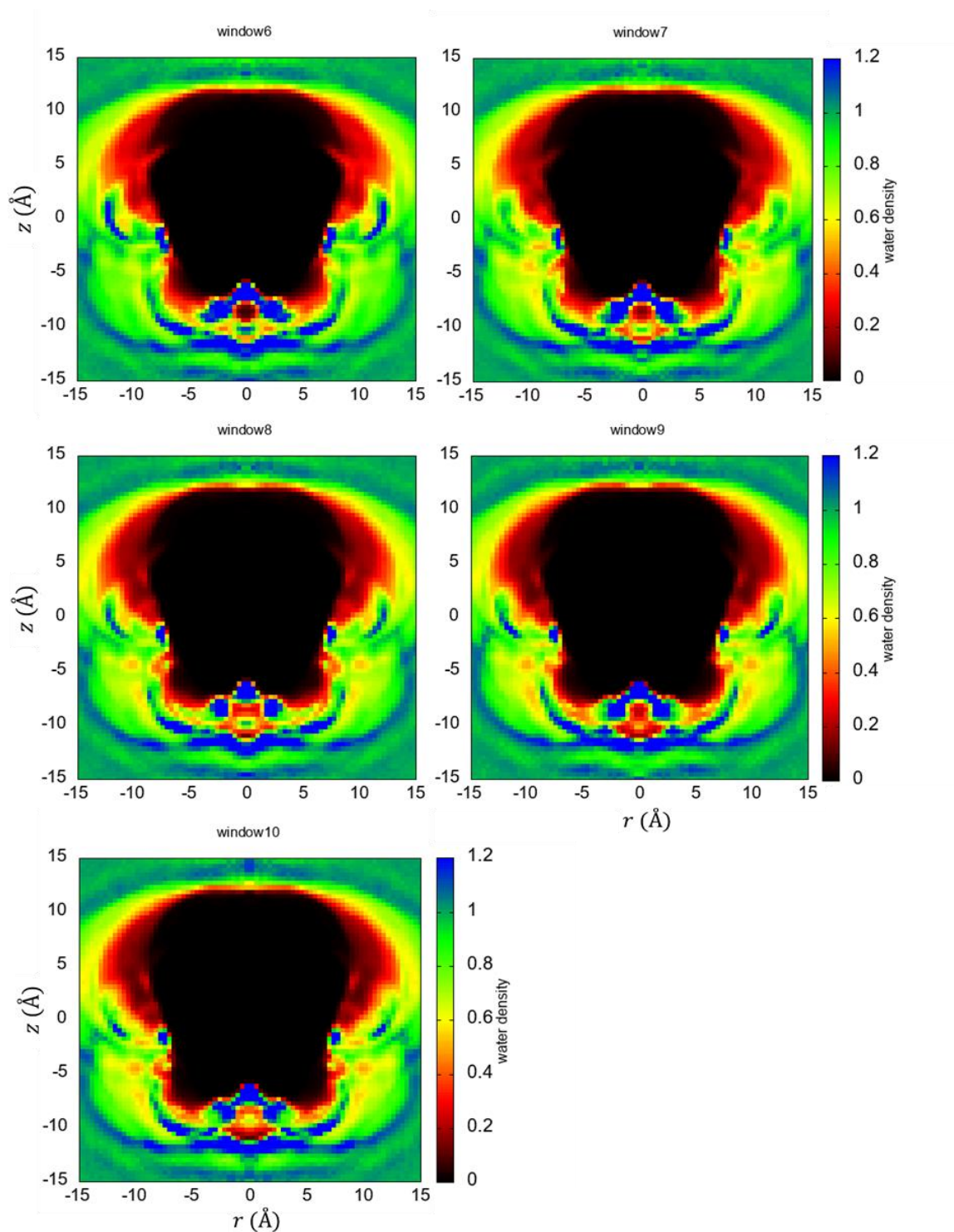
FigureE 4: Average water density distributions about TEMOA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 6-10 kJ/mol with an increment of 1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
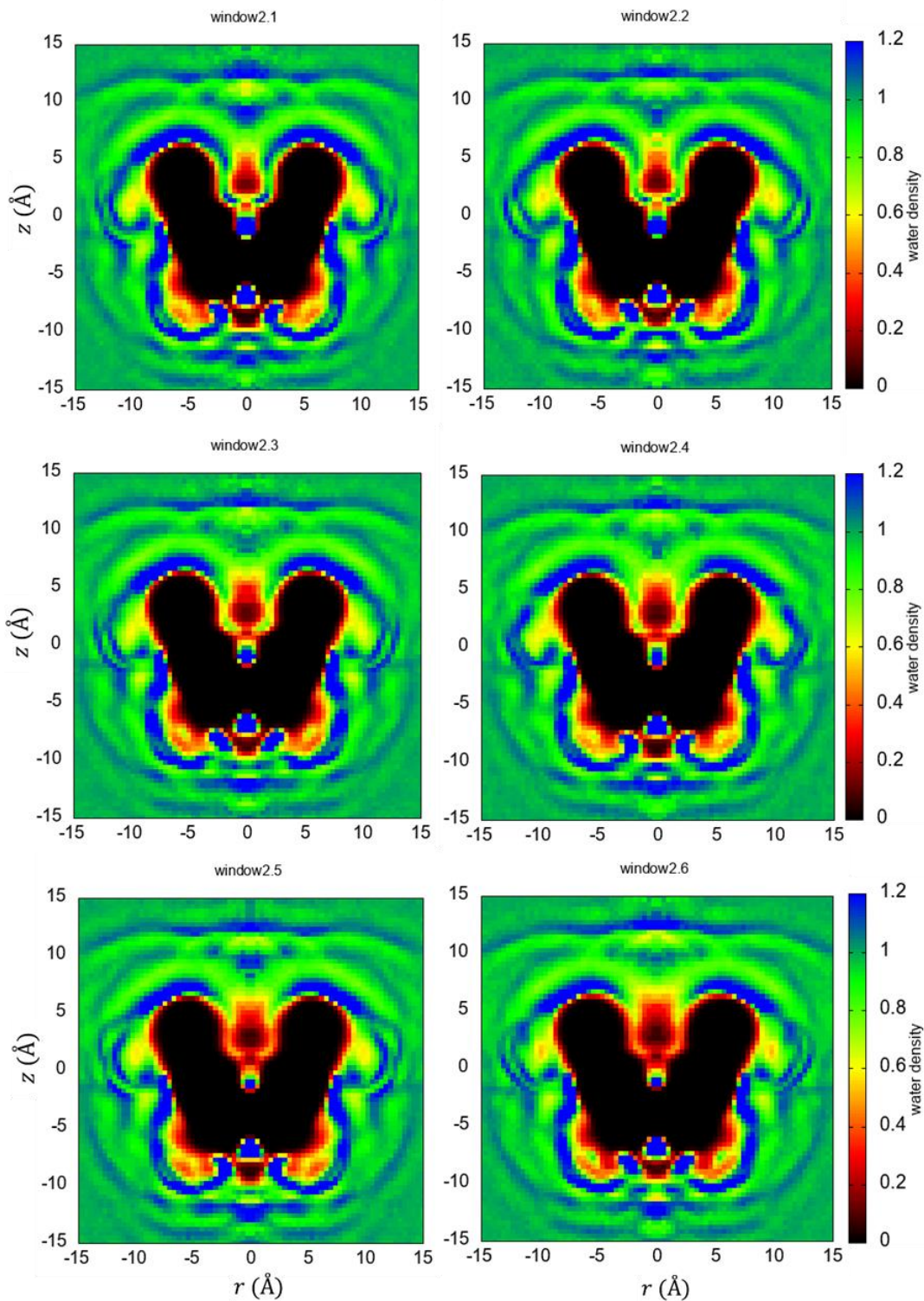
FigureE 5: Average water density distributions about OA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 2.1-2.6 kJ/mol with an increment of 0.1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
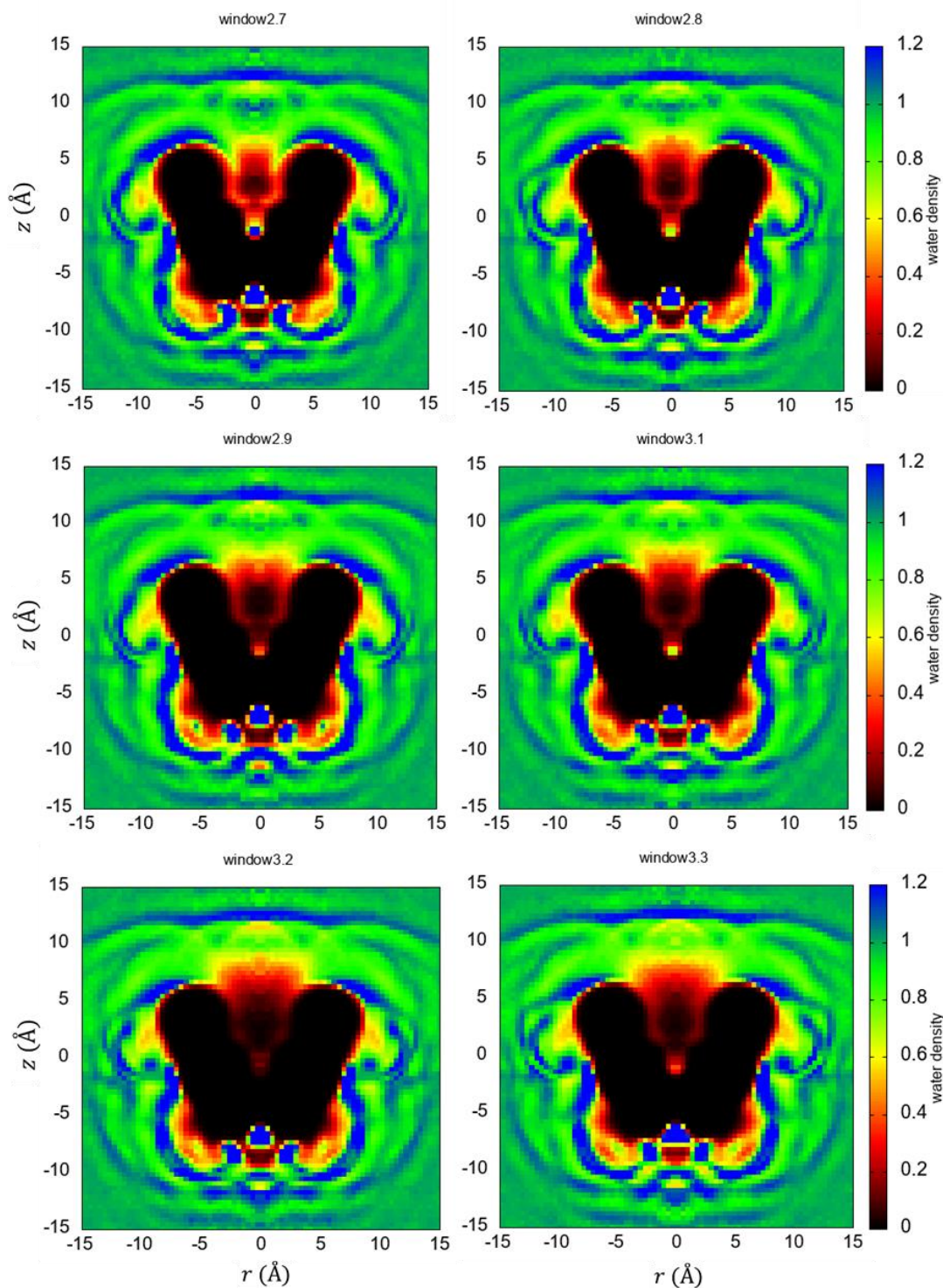
FigureE 6: Average water density distributions about OA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 2.7-2.9 kJ/mol and 3.1-3.3 kJ/mol with an increment of 0.1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
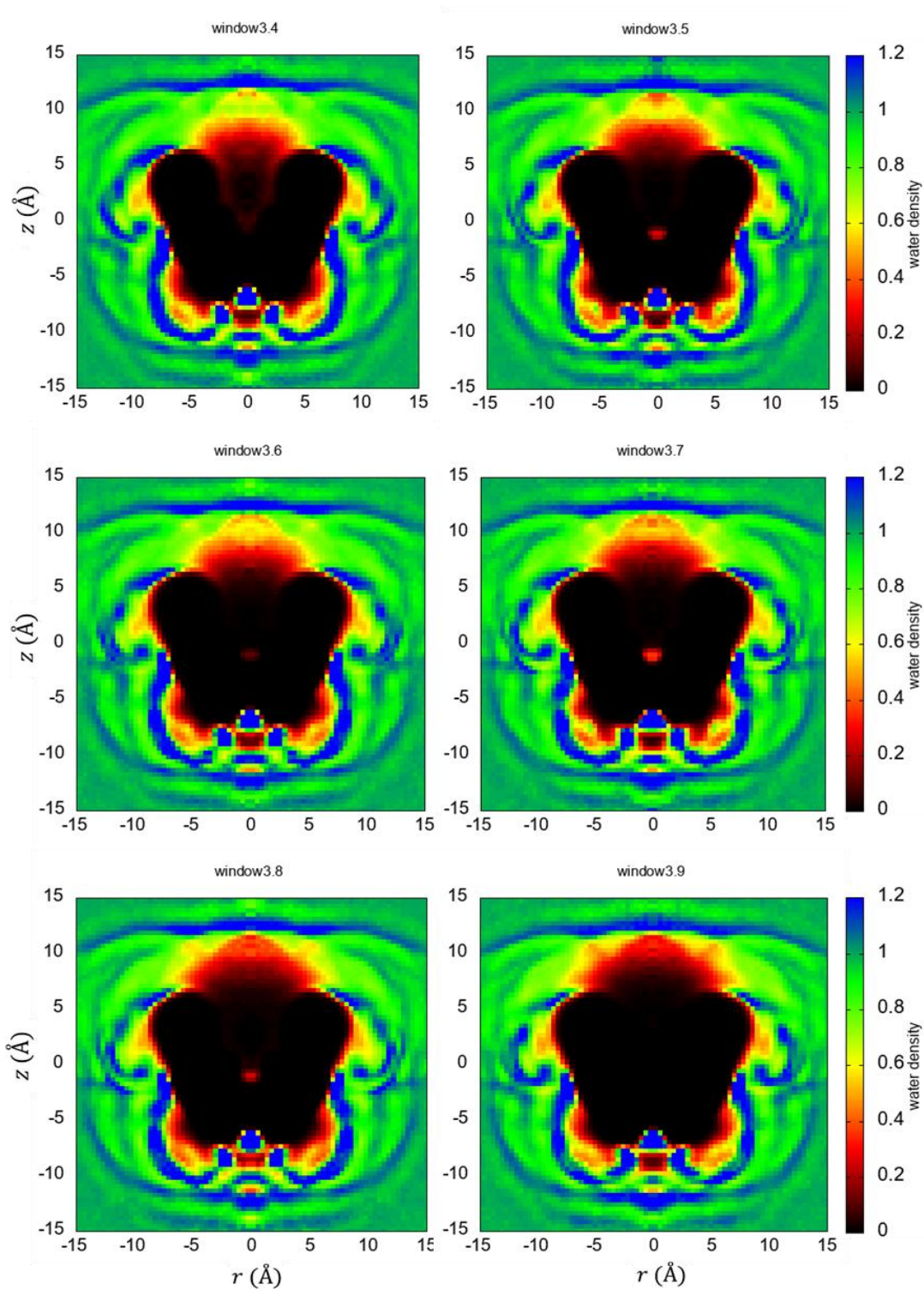
FigureE 7: Average water density distributions about OA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 3.4-3.9 kJ/mol with an increment of 0.1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
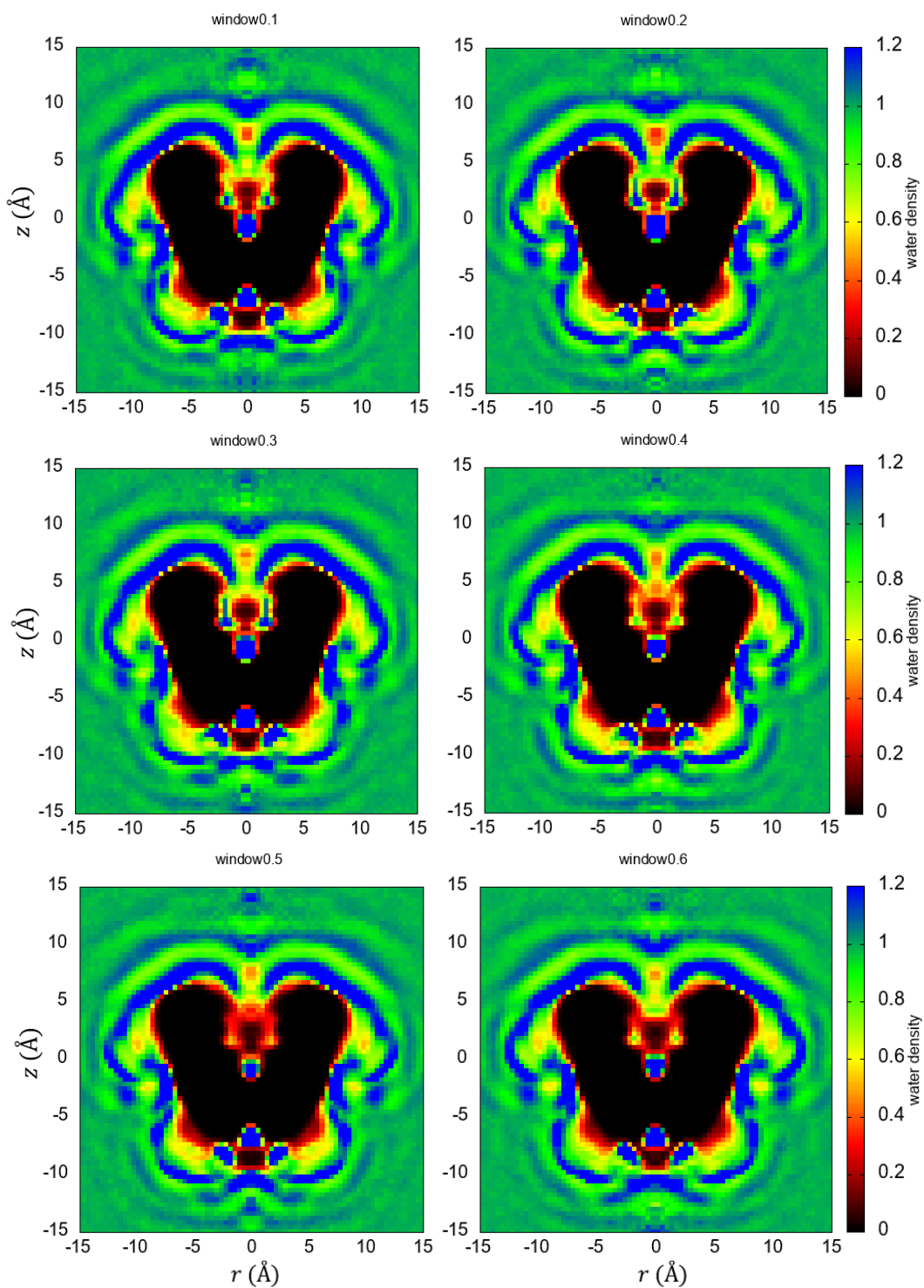
FigureE 8: Average water density distributions about TEMOA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 0.1-0.6 kJ/mol with an increment of 0.1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
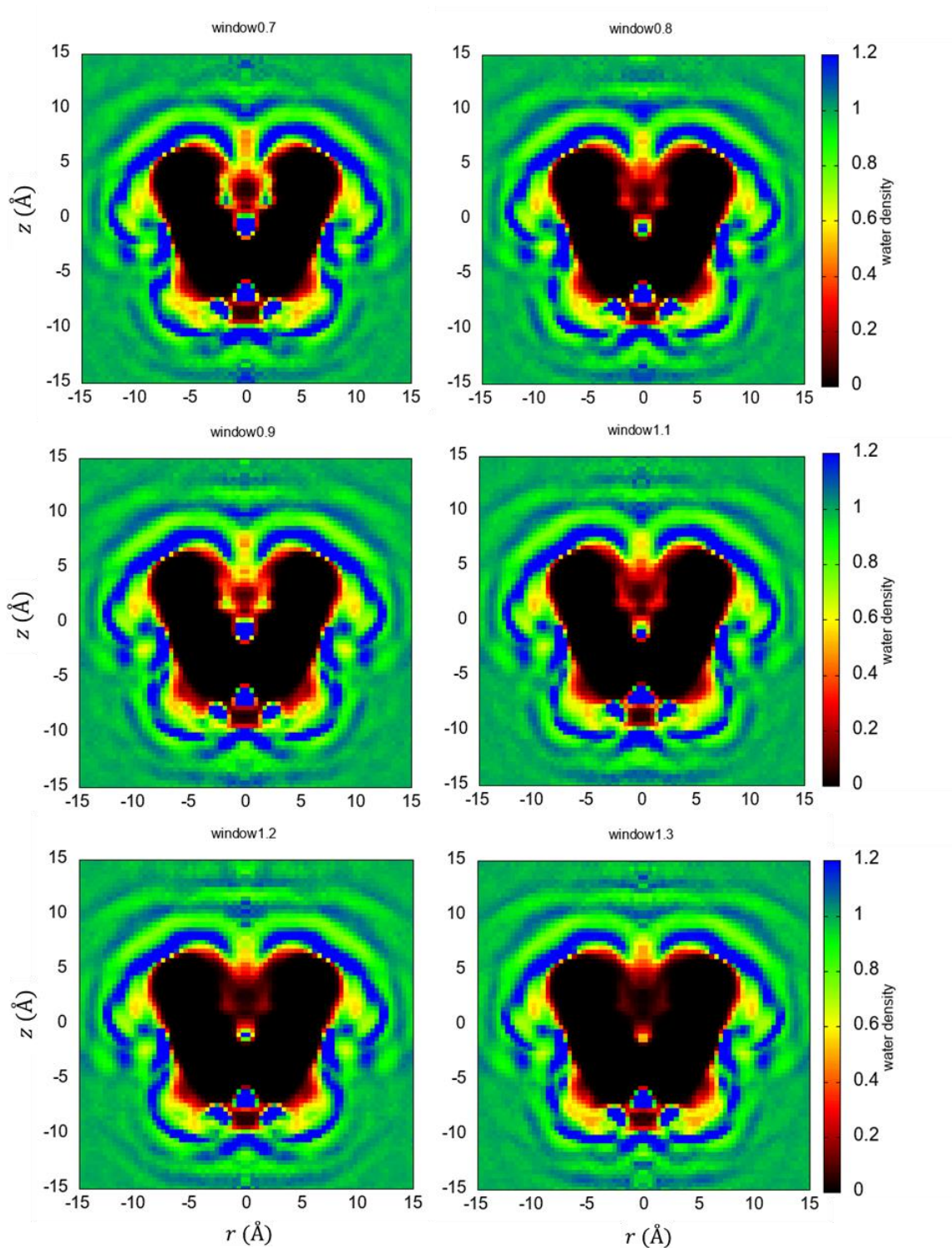
FigureE 9: Average water density distributions about TEMOA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 0.7-0.9 kJ/mol and 1.1-1.3 kJ/mol with an increment of 0.1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.
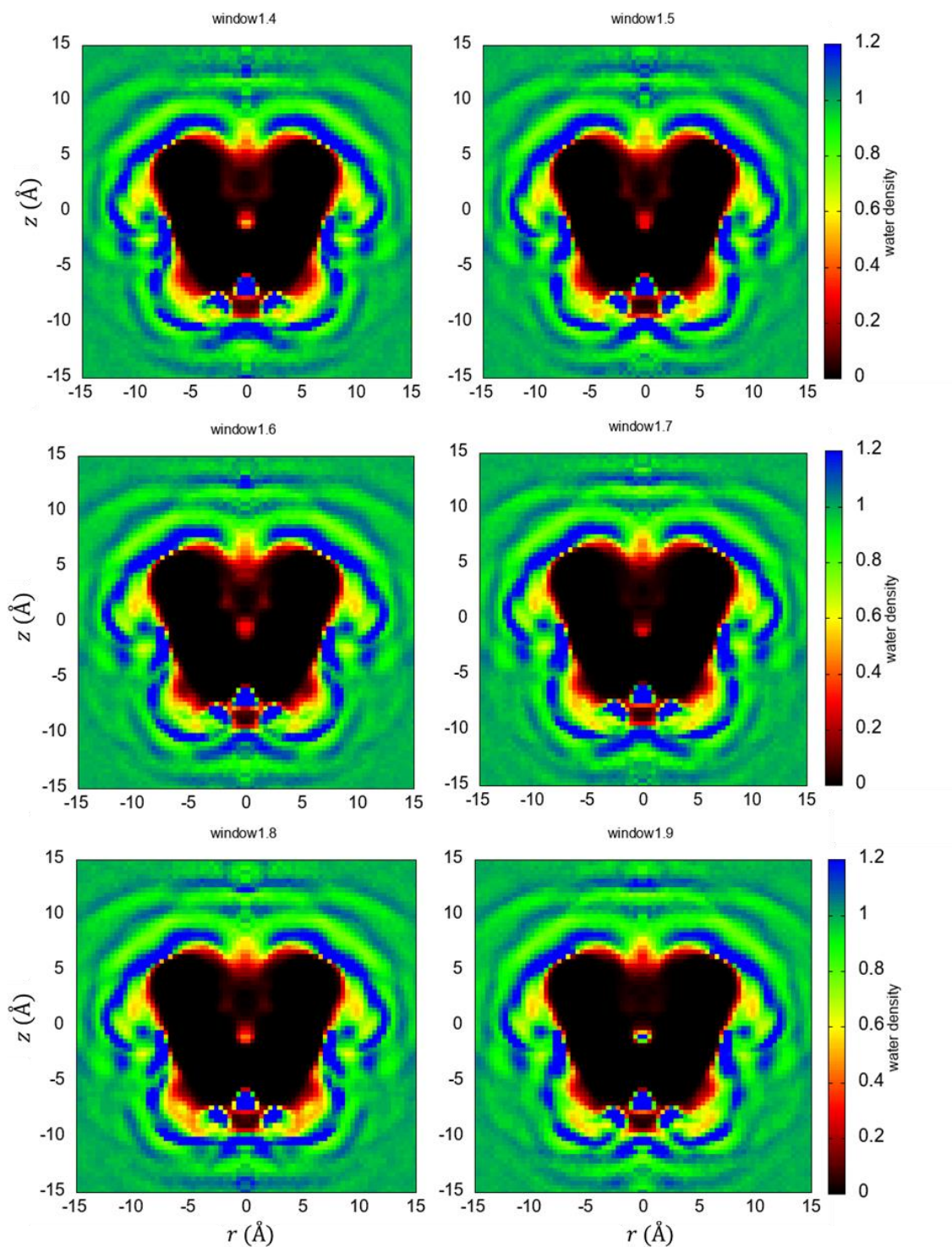
FigureE 10: Average water density distributions about TEMOA in presence of biasing potentials on each heavy atom of the cavitand with increasing strengths. The value of ϕ is from 1.4-1.9 kJ/mol with an increment of 0.1 kJ/mol. The water density is cylindrically averaged about the C_4-axis of symmetry for each of the hosts, with r corresponding to the radial distance from the C_4-axis and z indicating the vertical rise relative the cavitand center-of-mass. The densities follow the color key on the righthand side of the figures.

Appendix F

Customizable Drying of Hydrophobic Pockets and its Influence on Guest Binding

F.1 Introduction

As the adage goes "oil and water do not mix", water molecules are predicted to dewet large-scale non-polar solutes[128,131,132]. Unique phenomenon and attractive interactions still exist between oil and water. For instance, water was found to be able to transmit through the carbon nanotube and exhibited a dry-filled two-state transition[167]. Within the inner cavity and space of supramolecular host molecules like cyclodextrin and cucurbiturils, water molecules can stay and be even shielded inside. These has been called "high-energy" waters and can lead to significant impact on the host/guest binding processes[41]. Demonstrated both by experiments[119] and simulations[45], the displacement and evacuation of these water molecules from the cavities and pockets can make the binding process dominated by enthalpy components. It is in contrast to the classical mechanism behind hydrophobic interaction stating the main driving force is the entropically favored release of more structured waters near hydrophobic environment[146]. Particularly, in biological area this dewetting phenomenon is more prominently detected showing some non-polar cavities

of proteins are dry when dissolved in water. Cooperative water filling was observed in a hydrophobic cavity of lysozyme by high-pressure crystallography and molecular simulation[49]. This ~160 $Å^3$ containing mutant L99A was found to be empty at ambient pressure and can accommodate up to 4 water molecules under elevated pressures. While the shape of the protein was not changed during the whole process. A similar phenomenon was shown in the larger nonpolar binding cavity (~315 $Å^3$) of bovine $β$-lactoglobulin[50]. By MRD, ^{13}C NMR, and molecular simulation, it can be concluded that this large pocket is completely dry. Thus, it could potentially bind with non-polar guests and ligands very efficiently.

To mimic the hydrophobic environment of protein cavities and to investigate the mechanism behind the binding process, three deep-cavity cavitand host molecules were developed. In figure F1, the structures of Octa-acid (OA, host **1**), tetra-*endo*-methyl octa-acid (TEMOA, host **2**), and tetra-*exo*-methyl octa-acid (TEXMOA, host **3)** are illustrated. All of them are formed by three rows of aromatic rings and has a bowl-shaped hydrophobic pocket. The feet and rim of cavitand are decorated by eight carboxylic acid groups to enable water solubility. The hydrophobic pocket is approximately 8 Å in diameter and 8-9 Å deep. TEMOA differs from OA by having four methyl units on the rim and they are pointing inwards. The synthesis and host-guest binding of OA and TEMOA were reported previously[56,67,69,168]. In general, with four methyl units on *endo*-positions, the portal is narrowed, and the packing of guests is restricted. The binding kinetics, guest affinities, and assembly states will be affected[67,169]. TEXMOA is a newly developed derivative of OA, and its only structural variation is moving the four methyl units from the *endo*-positions to the *exo*-positions of rim. Those methyl units are pointing upwards, but essentially the

hydrophobic pockets are identical to that of OA. Both OA and TEMOA are able to form 2:1 dimeric capsules with suitably sized hydrophobic guests[61]. TEMOA can even form into tetrameric and hexameric complexes[69]. Besides, they can also form into 1:1 complexes with anions like perchlorate and amphiphiles such as short chain fatty acids[169–171].

Here we studied the hydration inside the pockets of all three cavitands using molecular dynamics simulation. The influence of temperature and pressure on the hydration conditions were also investigated. The partial molar volume of each cavitand was also evaluated from simulation to figure out the differences of the pocket shapes. As a result, we reveal the evidence of wetting/dewetting within the hydrophobic pockets. And this hydration is dominantly controlled by the positions of the methyl groups relative to the portals. In an effort to study its impact on host/guest binding behaviors, the drying free energy of the pockets were evaluated from simulation results. The association free energy between a cavitand and a single guest molecule to form a 1:1 complex was calculated by umbrella sampling. These free energy terms were also break down to enthalpic and entropic components to have a better understanding of the connection between hydration and binding. Several experimental results using densimetry and calorimetry were conducted by our collaborators' groups to have a comparison.
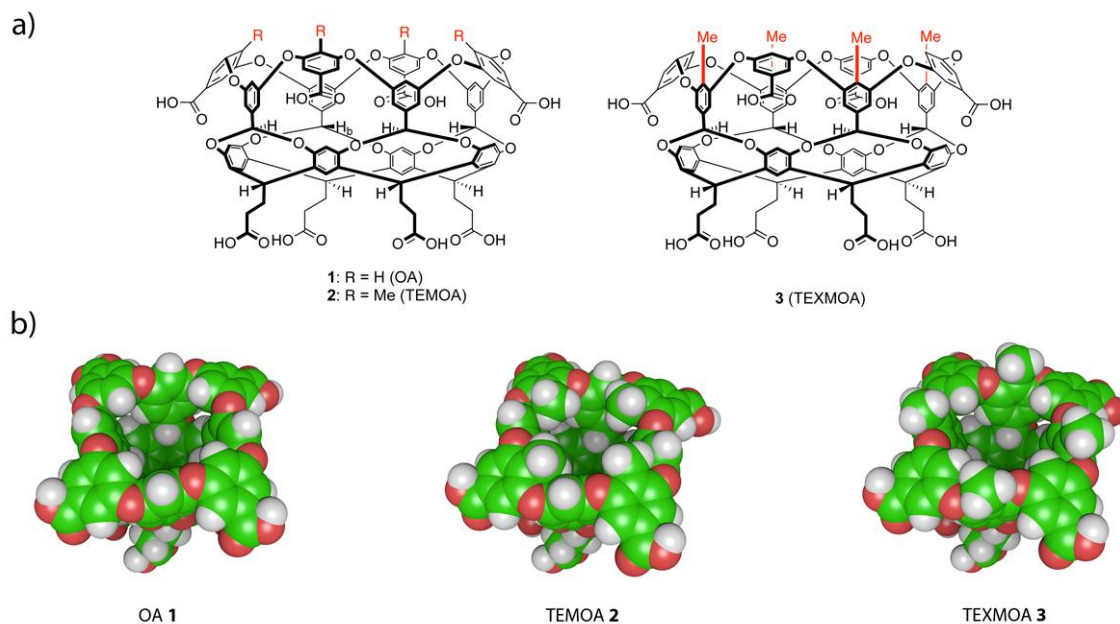
FigureF 1: a) structures of octa-acid (OA, host **1**), tetra-*endo*-methyl octa-acid (TEMOA, host **2**), and tetra-*exo*-methyl octa-acid (TEXMOA, host **3**); **b)** Space-filling representations of three hosts and highlighting the structural effects of methylation on the *endo*-positions (host **2**) and *exo*-positions (host **3**).

F.2 Methods

We conducted molecular simulations of cavitands in water using GROMACS 5.1[79] to study their inner hydration numbers and partial molar volumes. All the simulations were performed in isothermal-isobarric ensemble with a systematic combination of different temperatures and pressures. Simulations were conducted for all three cavitands at 1 bar and a range of temperatures from 278.15 K to 368.15 K with an increment of 10 K. Besides, to investigate the impact of pressure, simulations were also conducted for them at 298.15 K and a range of pressures from 1 bar to 2500 bar with a 500-bar increment. The temperature and pressure were controlled by Nosé-Hoover thermostat[80,81] and Parrinello-Rahman barostat[82], respectively. The cavitands were modeled using the Generalized Amber Force

Field (GAFF)[70]. To match the deprotonation state at pH 7[107], the network charge of cavitand was set to be -6e. The four benzoic acid groups around the rim of cavitand and two of the four benzoic acid groups on the feet were deprotonated. And to make the whole system neutral, six sodium counterions, which are also modeled by GAFF, were introduced for each cavitand. The partial charges of each atom were initially obtained by AM1-BCC calculations following geometry optimization[106]. As described in the discussion part below, to take the impact of electronic polarization on cavitand volumes into consideration and have a comparison with experimental data, we rescale the partial charges, including the charges of sodium counterions, by multiplying 75%. Waters were modeled using the TIP4P/EW potential[108]. A 9 Å cut-off was used to truncate the non-bonded Lennard-Jones interactions with a mean-field dispersion correction for longer-range contributions to the energy and pressure. Long range electrostatic interactions were treated using particle-mesh Ewald summation[71] with a real space cutoff of 9 Å. Bonds involving hydrogens in the hosts were constrained using the LINCS algorithm[86]. The bonds in water molecules were held rigid using SETTLE[87]. The equations of motion were integrated using a 2-fs time step. All the simulations were equilibrated by 5 ns followed by a 100 ns production run. Configurations were saved every 1 ps for analysis afterwards.

We assumed a cavitand as a hexahedron constructed from six planes formed by atoms on the cavitand, and therefore to determine the number of waters within a cavitand. The top face of the cavitand pocket was determined by constructing a plane through the eight diphenyl ether oxygen atoms on the rim of the cavitand. While the bottom face of the cavitand was constructed by fitting the plane based on the four carbon atoms connecting the cavitand feet and the last row of aromatic rings at bottom. The four planes on the side

were determined by three sites. Two of them are the diphenyl ether oxygens connecting a single benzoic acid moiety to the second layer of aromatic rings. Another site is the average position of the two correspondingly nearest carbon atoms that define the bottom plane. For a specific water molecule, if its oxygen atom was found to be within this constructed hexahedron, it was considered to be within the inner pocket of cavitand. The average positions of the atoms defining the top and bottom plane are illustrated by two dummy sites highlighted in figure F2. The line passing through these two dummy sites is defined as the four-fold rotational axis of symmetry (C_4-axis) for the cavitand.

In addition, the potential of mean force (PMF) between the assembly constituents of 1:1 guest/cavitand complex was also evaluated by umbrella sampling. Here we focused on the interaction between a single hexanoate and a single cavitand to form a 1:1 complex. All three cavitand hosts, OA, TEMOA, and TEXMOA were considered. The hexanoate guest was modeled by the OPLS all-atom force field[172]. Seven sodium cations, modeled with GAFF[70], were put in the simulation box to compensate for the total -7e net charges, six of them are from the cavitand and one is from the carboxylate ion of hexanoate. The PMF were evaluated from bulk water into the cavitand pocket along the host's C_4-axis. The cavitand and the guest were solvated by 2669-2674 water molecules modeled using TIP4P-EW potential[108]. As described in previous chapters, two dummy atoms were created for a cavitand. The first one was determined according to the average position of the four carbon atoms connecting the feet to the bottom row of aromatic rings. The second one was determined by the average positions of the four carbon atoms on the second row of aromatic rings and are closest to the open month. The vector connecting these two dummy atoms were restrained to be parallel with the z-axis of the simulation box with an angular restraint

of 50000 kJ/(mol). Another harmonic position restraint of 100000 kJ/(mol nm^2) was applied to keep the bottom dummy atom on z-axis. Therefore, the C$_4$-axis of cavitand always coincided with the z-axis. Forty overlapping windows from -7 Å deep-inside the cavitand pocket to 13 Å out into the bulk solvent. This distance was measured from the top dummy atom, which represents the center of the portal of cavitand defined based on the positions of the four carbon atoms on the second row of aromatic rings and are closest to the open month. The center of the hexanoate was also restrained on z-axis by a potential with a harmonic force constant of 100000 kJ/(mol nm^2). This guest center was defined as the position of a dummy atom, which was determined as the average positions of the third and fourth carbon atoms along the chain backbone. In this series of overlapping windows, the harmonic umbrella potential minimums were separated in 0.5 Å increments with a force constant of 15000 kJ/(mol nm^2). In general, all the simulations were conducted in the isothermal-isobaric ensemble. Temperature and pressure were coupled using Nosé-Hoover thermostat[80,81] and Parrinello-Rahman barostat[82]. Pressure was at 1 bar. Replica exchange calculations[173] were performed in each window with 10 replicas spanning the 283.15-373.15 K temperature range with an increment of 10 K. The system in each window was equilibrated for 1 ns followed by a production run of 15 ns. The PMF for the formation of 1:1 cavitand/hexanoate complex was reconstructed using the weighted histogram analysis

method (WHAM)[109]. The other simulation details were set up the same as other simulations described above.
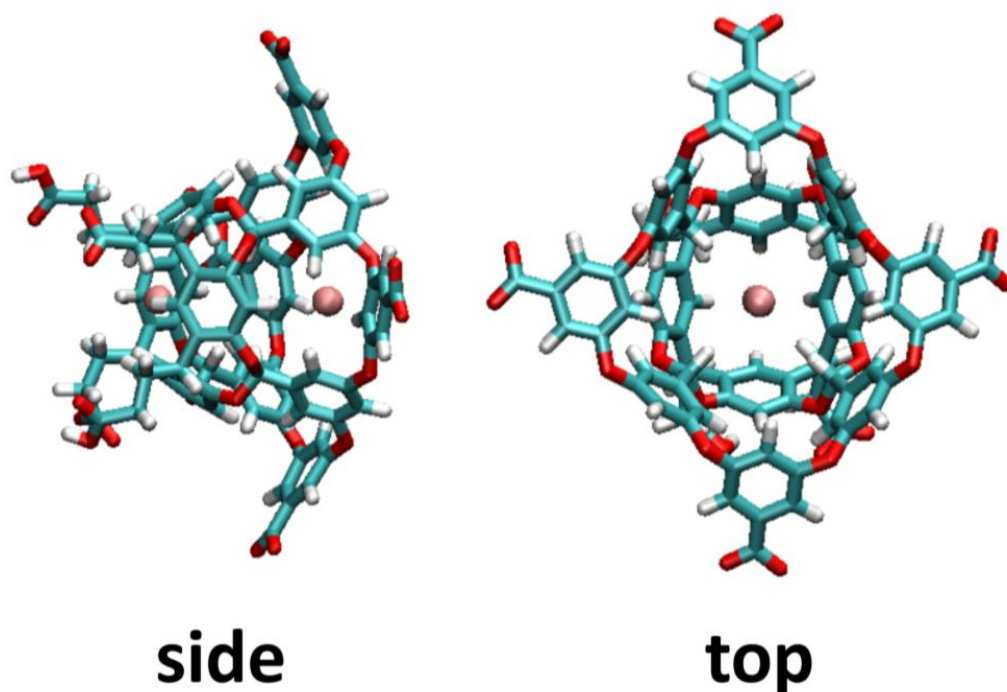
FigureF 2: Structure used for the host octa-acid (OA). The top and bottom planes that define the cavitand pocket are shown in the side and top views. The pink "dummy" site at the top rim of the pocket represents the average position of the eight (diphenyl ether) oxygens that connecting the benzoic acid groups to the second layer of aromatic rings. The pink "dummy" site at bottom is determined by the average position of the four carbon atoms connecting the four feet to the first layer of aromatic rings. The four planes on the side are determined by the four-fold rotational symmetry of the cavitand (top view). Each of them is formed by the two oxygens connecting each benzoic acid moiety to the rim of the cavitand, and the average position of the two closest carbon atoms forming the bottom plane. The C_4-axis is defined as the line passing through the two dummy sites.

Finally, we also investigated how hydrogen bonding of waters inside the cavitand pocket was affected by the different hydration conditions. Specifically, for each water molecule within the pocket, we can determine the number of hydrogen bonds it formed with neighbor water molecules as an acceptor or a donor. When the distance between one donor and one acceptor (two oxygens) is smaller than or equal to 0.35 nm, and at the same time the hydrogen-donor-acceptor angle is smaller than or equal to 30°, a hydrogen bond is

considered to be exist. And when this donor or acceptor belongs to a water molecule inside the pocket, we add one to its number of hydrogen bonds as a donor or acceptor correspondingly. In the end, the average total number of hydrogen bonds for individual water within pocket can be obtained for each simulation frame. And according to different hydration numbers within pocket of each frame, we can account the mean number of hydrogen bonds per water molecule in the hydrophobic pocket as a function of distinct hydration numbers.

F.3   Results and discussion

We firstly investigated the partial molar volumes of each cavitand from simulation. It is defined as the difference between the average total simulation box volume of the system with a single cavitand solvated in water with its counter ions and that of a system with pure water. The number of waters in both systems are the same. The detailed mathematical expression is described as below:

$$\bar{v} = \langle V \rangle_{host+water} - \lambda \langle V \rangle_{water} \qquad \text{(F.1)}$$

Here it is approximating the partial molar derivative as a finite difference between two systems. The pressure, temperature, and number of water molecules are all held fixed. $\langle V \rangle_{host+water}$ and $\langle V \rangle_{water}$ refer to the average volumes of simulation systems with and without the solvated cavitand, respectively. In order to make sure the number of waters are the same in these two terms, a factor $\lambda = N_{w,host+water}/N_{w,water}$ is introduced to rescale the volume of the pure water simulations and treat for different water numbers in

simulations with and without cavitand. The comparison between the partial molar volume from simulations and experiments are listed in table F1.

TableF 1: Partial molar volumes of single cavitand, $\bar{v}_i$, and partial molar volume differences, $\Delta\bar{v}_{ij} = \bar{v}_i - \bar{v}_j$, from simulations using scaled charge and experiments at 298.15 K and ambient pressure. Experimental results were obtained from the average of two independently prepared samples. The reported errors in experimental results are the maximum deviation from the average. While the reported errors in simulation results represent for one standard deviation.

	$\bar{v}_1$ (cm^3/mol)	$\bar{v}_2$ (cm^3/mol)	$\bar{v}_3$ (cm^3/mol)	$\Delta\bar{v}_{31}$ (cm^3/mol)	$\Delta\bar{v}_{21}$ (cm^3/mol)	$\Delta\bar{v}_{23}$ (cm^3/mol)
experiment	1083 ± 3	1245 ± 12	1157 ± 6	73 ± 7	162 ± 12	89 ± 13
simulation	1108 ± 5	1256 ± 5	1189 ± 5	81 ± 3	148 ± 3	67 ± 2

Table F1 shows the partial molar volumes, $\bar{v}_i$, for three distinct types of cavitands (i=1, 2, or 3), and volume differences, $\Delta\bar{v}_{ij} = \bar{v}_i - \bar{v}_j$, for hosts i and j (i, j = 1, 2, or 3, and $i \neq j$) both from simulations and experiments. The experiments were conducted in our collaborator, Dr. Dor Ben-Amotz's group at Purdue University. In general, the cavitand solution densities can be measured by an Anton Parr DSA 5000 density and sound velocity meter. The inverse densities of aqueous solutions of all three cavitands with multiple cavitand weight fractions (between ~0.2% and 1%) can be determined. A simple linear relationship between inverse densities and cavitand weight fractions can be fitted to the experimental data. The resulted slope and interception can thus be used to determine the cavitand partial molar volumes at infinite dilution using the equation F.2.

$$\bar{v}_1^\infty = M(\frac{1}{\rho_w^0} + S_w^0)$$

(F.2)

Here $1/\rho_w^0$ (cc/g) and S_w^0 (cc/g) are the intercept and slope of the linear relationship in experimental data, $1/\rho$ vs. w, respectively. $1/\rho_w^0$ is the inverse density of the pure solvent, which is the aqueous sodium hydroxide solution with pH $= 12.7 \pm 0.2$. M is the cavitand molar mass. $\bar{v}_1^\infty$ is the partial molar volume at infinite dilution. The solid cavitands are protonated acids. Thus, this observed volumes have two contributions. One is the partial molar volume of the ionized cavitand. The other is the volume change during the deprotonation of the cavitand and the conversion of OH^- to H_2O. Therefore, we need to assume all the cavitands have the same extent of ionization.

Very interestingly, it can be found that there is a strong dependence of the partial molar volume differences on the positions of host methylation about the portal. The partial molar volume difference, $\Delta\bar{v}_{31}$, between TEXMOA and OA (host **3** and host **1**) is 73 ± 7 cm^3/mol from experiment. As a comparison, the reported volume difference between alanine and glycine amino acid side chains is 18.3 cm^3/mol[174]. As we know, the structural difference between alanine and glycine amino acid side chains is one additional methylation. And this volume difference is just a quarter of $\Delta\bar{v}_{31}$ ($\approx 4 \times 18.3$ cm^3/mol $= 73.2$ cm^3/mol). It indicates that the only contribution to the volume difference between TEXMOA and OA is the methylation on the rim. However, the partial molar volume difference, $\Delta\bar{v}_{21}$, between TEMOA and OA (host **2** and host **1**) is 162 ± 12 cm^3/mol, which is even larger than twice of the volume increment brought by the additional four methyl groups on rim. Our simulation results are remarkably consistent with the experiments. The largest difference is at most 2%. In simulation, the partial molar volume difference, $\Delta\bar{v}_{31}$, between simulated TEXMOA and OA is 81 ± 3 cm^3/mol, which is very closed to the volume difference resulting from four additional methylation. In contrast, the volume difference, $\Delta\bar{v}_{21}$,

between TEMOA and OA is 148 ± 3 cm^3/mol. This is also significantly larger than that of four methyl groups and in the neighborhood of the experimental difference. Thus, simulations also captured the abnormal volume difference between TEMOA and OA. The combination of observations from experiments and simulations is indicating that there are other hidden contributions to this additional volume difference.

Previously, we also compared the cavitand partial molar volumes determined from simulations using the full -6e charge with that from experiments. As shown in table F2, it turns out that those volumes with full charge are significantly smaller than experimentally observed ones. This inconsistence can be potentially ascribed to the neglect of electronic polarization in our simulations. A rescaling of the charges for ionic species by multiplying 75% was recommended by Jungwirth and coworkers to address this issue on aqueous electrolyte solutions[175,176] and aqueous electrolytes near non-polar interfaces[177]. The polarization of electronic degrees of freedom as a continuum dielectric with an effective electronic dielectric constant of $\varepsilon_e = 1.78$[178] can be approximated by this scaling factor. And the consistence between experiments and simulations were dramatically improved after applying this scaling factor. According to Kirkwood-Buff theory[179], at infinite dilution, the partial molar volume of a solute can be evaluated by an integral over solute/solvent correlations. Thus, this scaling factor is also applied for cavitand in all the simulations involved in this chapter.

TableF 2: Partial molar volumes of single cavitand, $\bar{v}_i$, and partial molar volume differences, $\Delta\bar{v}_{ij} = \bar{v}_i - \bar{v}_j$, from simulations using their full charge and experiments at 298.15 K and ambient pressure. Experimental results were obtained from the average of two independently prepared samples. The reported errors in experimental results are the maximum deviation from the average. While the reported errors in simulation results represent for one standard deviation.

	$\bar{v}_1$ (cm^3/mol)	$\bar{v}_2$ (cm^3/mol)	$\bar{v}_3$ (cm^3/mol)	$\Delta\bar{v}_{31}$ (cm^3/mol)	$\Delta\bar{v}_{21}$ (cm^3/mol)	$\Delta\bar{v}_{23}$ (cm^3/mol)
experiment						
	1083 ± 3	1245 ± 12	1157 ± 6	73 ± 7	162 ± 12	89 ± 13
simulation						
	953 ± 5	1074 ± 6	1026 ± 7	73 ± 3	121 ± 3	48 ± 4
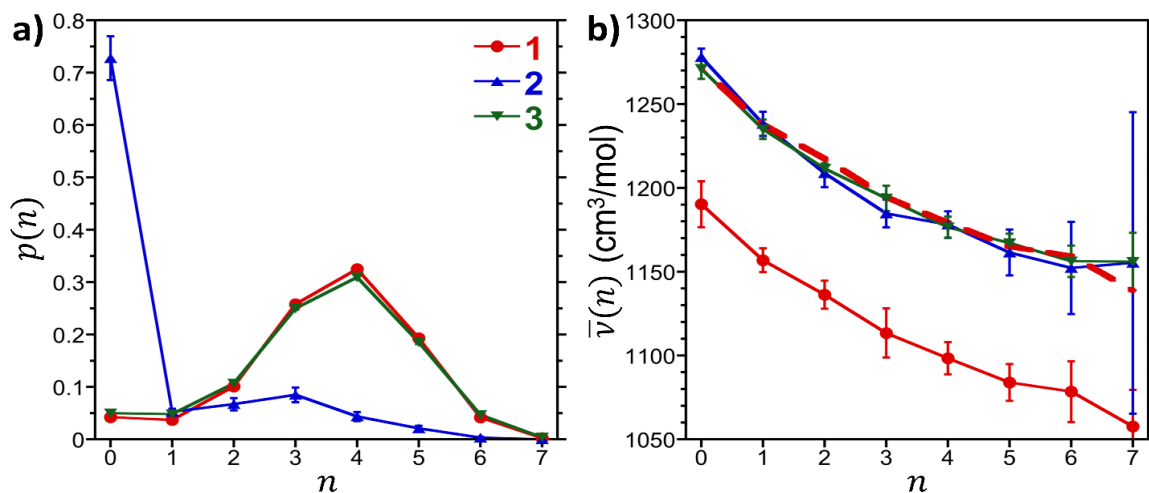
FigureF 3: a) The probability distribution, p(n), of number of water molecules within the hydrophobic pocket of cavitand as a function of hydration number, n. They are evaluated from simulations at 298.15 K and 1 bar using scaled charge. The symbols for hosts **1-3** are identified in the figure legend. **b)** Partial molar volumes of cavitand, $\bar{v}(n)$, as a function of hydration numbers within hydrophobic pocket determined from the same simulations. The symbols for hosts **1,2**, and **3** are defined in figure F3a. The thick, dashed, red line corresponds to the results for **1** shifted up by $\Delta = 81$ cm³/mol. The error bars in both figures indicate one standard deviation.

Figure F3a is showing the probability distribution of number of water molecules within cavitand's hydrophobic pocket. In the case OA (host **1**), it exhibits a nearly symmetric and unimodal shape. The most prominent peak is centered at n=4 waters inside the pocket. But actually, it has a slightly skewness towards lower occupancy states. With increasing hydration numbers, the probability quickly falls to 0 when $n \geq 8$. When number of waters

decreases, a second, weak peak can be observed, and the distribution profile extends a long tail to the left. There is an approximately 5% probability $(= p_1(0))$ to observe a completely dry pocket. In the case of TEXMOA (host **3**), the probability distribution of hydration number within pocket is found to be nearly identical to that of OA. The profile of pocket hydration distribution is not significantly affected by the four methyl units on *exo*-positions of rim. In contrast, the hydration number distribution of TEMOA (host **2**) is impacted dramatically by its four methyl units on *endo*-positions. It transforms from the nearly gaussian shape of OA and TEXMOA into a very apparent bimodal shape. Two local maxima of probability can be observed at $n=0$ and $n=3$. And the pocket of TEMOA is empty during ~73% $(= p_2(0))$ of the overall simulation time. Therefore, it can be concluded that whether the methylation occurs on *exo*-positions (pointing upwards) or *endo*-positions (pointing inwards) will significantly change the hydration number distribution within the inner-pockets of cavitand.

TableF 3: Average number of water molecules within cavitands' pockets, $\langle n \rangle = \sum i * p(i)$, determined from simulations at 298.15 K and 1 bar with scaled charge. Reported errors indicate one standard deviation.

host	$\langle n \rangle$
1	3.55 ± 0.04
2	0.74 ± 0.06
3	3.48 ± 0.03

This difference in pocket hydration can be promptly reflected from the average number of water molecules within the pockets of these three distinct cavitands, as reported in table F3. In average, both OA and TEXMOA can capture around 3.5 waters, while the mean hydration number is only 0.74 for TEMOA.

To better explain these distinct hydration behaviors, we need to probe the impact of different methylation positions on the cavitand pocket shapes. Hence, we examined the pocket hydration of all three cavitands under elevated pressure. Under very high pressures, water molecules can be pushed into the pocket, and the dry states stabilized by capillary evaporation will be overcome. The probability distributions of hydration numbers within hydrophobic pockets for all three types of cavitands under a pressure as high as 2500 bar were reported in figure F4. All the hydration number probability distributions exhibit symmetric and Gaussian-like shapes. The centers of the distributions now are shifted to $n=5$ since more waters are pushed into the pockets. All the dry states ($n=0$) are suppressed, even for the probability distribution of TEMOA, which has a ~73% probability to observe dry states under ambient pressure. Despite some subtle differences, the three hydration distribution profiles are almost identical to each other.
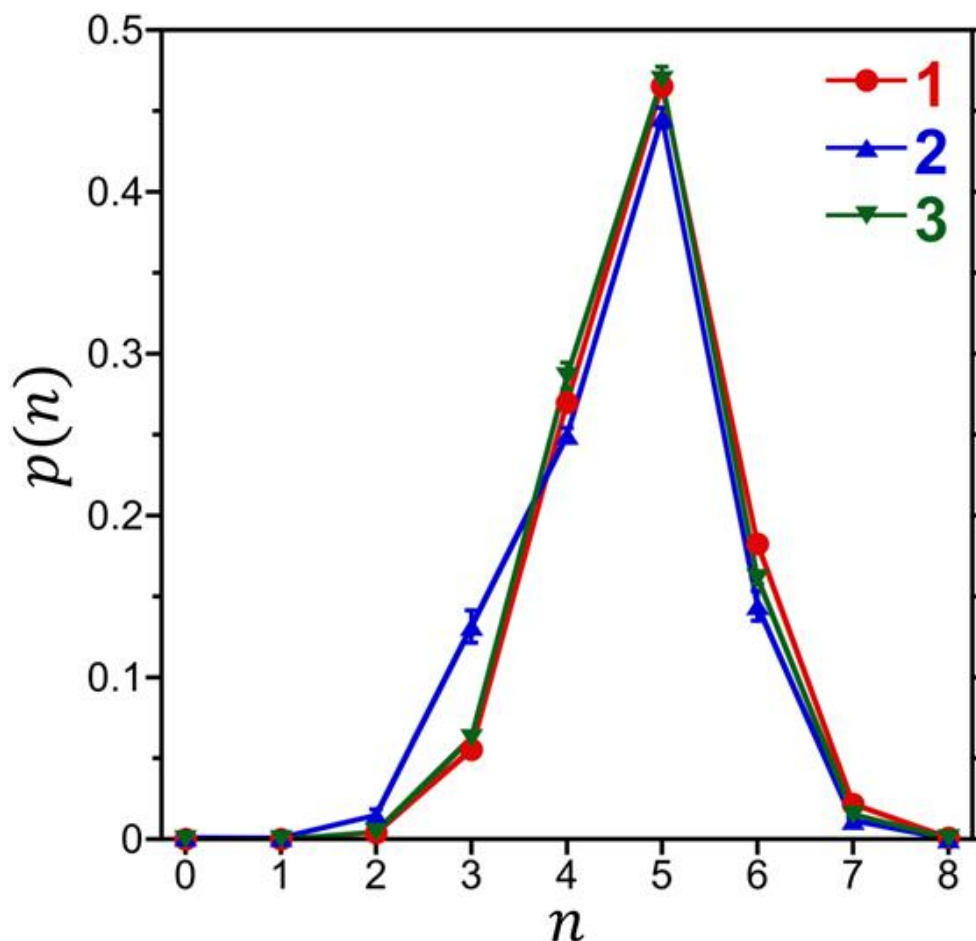
FigureF 4: The probability distributions, p(n), of number of water molecules within the hydrophobic pocket of cavitand as a function of hydration number, n. They are evaluated from simulations using scaled charges at 298.15 K and 2500 bar. The symbols for hosts **1-3** are identified in the figure legend. Error bars indicate one standard deviation.

These similar hydration behaviors demonstrated that the methyl units on *endo*-positions of the rim doesn't significantly change the shapes of pockets. They are able to accommodate water molecules of similar quantity once the dry states are overcome by sufficiently high pressure. It can be further verified by the average hydration numbers within all three cavitands under pressure of 2500 bar. As shown in table F4, they are all around 4.7 and in excellent agreement with each other. Moreover, if we turn back to partial molar volumes. Under 2500 bar, the partial molar volume difference (reported in table F5) between TEMOA and OA (host **2** and **1**) is ~72 cm^3/mol. This is not only very closed to the partial

molar volume difference between TEXMOA and OA (host **3** and **1**) under same pressure, but also almost identical to that of four additional methyl groups. All of these results lead to one conclusion that the positions of the additional methyl units on the rim don't affect the shape or capacity of hydrophobic pockets. The additional partial molar difference between TEMOA and OA under ambient pressure is from different hydration conditions. There is a tendency for TEMOA to self -dewet its pocket and easily to exclude water molecules. Once waters are pushed into the pocket and it has the same hydration statistics as that of OA and TEXMOA, the volume increment from the four added methyl groups remains the only source of partial molar volume difference.

TableF 4: Partial molar volumes of single cavitand, $\bar{v}_i$, and partial molar volume differences, $\Delta \bar{v}_{ij} = \bar{v}_i - \bar{v}_j$, from simulations using scaled charge and experiments at 298.15 K and 2500 bar. The reported errors in simulation results represent for one standard deviation.

	$\bar{v}_1$ (cm^3/mol)	$\bar{v}_2$ (cm^3/mol)	$\bar{v}_3$ (cm^3/mol)	$\Delta \bar{v}_{31}$ (cm^3/mol)	$\Delta \bar{v}_{21}$ (cm^3/mol)	$\Delta \bar{v}_{23}$ (cm^3/mol)
simulation						
	1063 ± 4	1135 ± 8	1036 ± 6	73 ± 7	72 ± 9	-1± 10

TableF 5: Average number of water molecules within cavitands' pockets, $\langle n \rangle = \sum i * p(i)$, determined from simulations at 298.15 K and 2500 bar with scaled charge. Reported errors indicate one standard deviation.

host	$\langle n \rangle$
1	4.84 ± 0.03
2	4.60 ± 0.07
3	4.77 ± 0.06
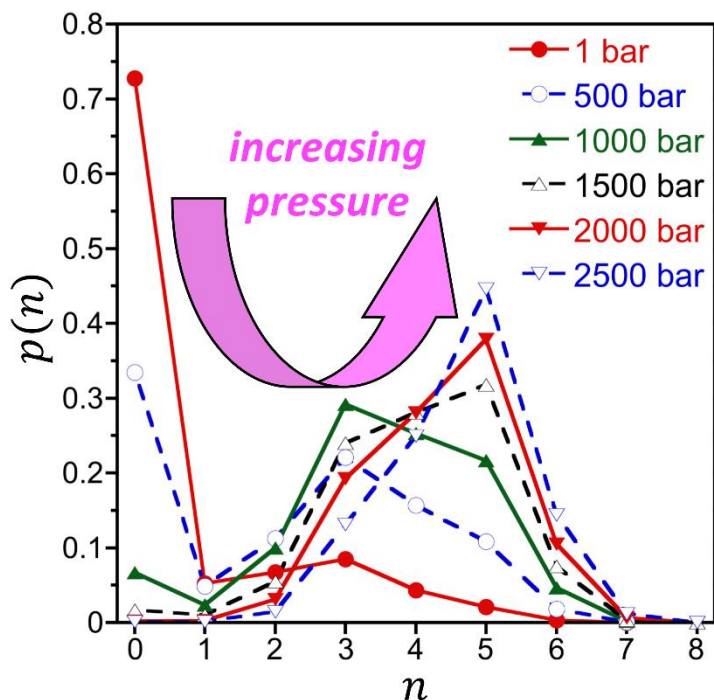
FigureF 5: The probability distributions, $p(n)$, of number of water molecules within the hydrophobic pocket of cavitand as a function of hydration number, n, and pressure. They are evaluated from simulations using scaled charges at 298.15 K. The pressure ranges from 1 bar to 2500 bar. The large inset arrow reflects the impact of pressure on $p(n)$. The symbols for the different pressure are identified in the figure legend. Error bars are neglected for clarity.

The probability distribution of the dry states and the wet sates can actually be controlled by pressures. And this has been detailly covered in the discussion part of chapter 5. Figure F5 illustrates the probability distributions of hydration numbers inside TEMOA (host **2**) as a function of pressure. In general, when pressure is lower, the bimodal shape of the distribution is more prominent. The probability to observe dry pocket drops from ~73% to ~34% after the pressure is elevated from 1bar to 500 bar. With the increase of pressure, the nature of the second local maximum is suppressed. When pressure reaches 1500 bar, basically the probability distribution is transformed into a nearly unimodal shape. In the bimodal shape of the hydration distribution profile, the two local maxima at $n=0$ and at $n=4$ or 5 are separated by a minimum (free energy barrier) at $n=1$. It indicates that the

probability of observing a completely dry pocket is higher than that of a pocket with one single water inside. The filling of the cavitand pocket is a cooperative effect similar with the reported drying phenomena of buried protein cavities.

The driving force of TEMOA's self-dewetting phenomenon is discussed in chapter 5. Generally speaking, it can be simplified using capillary evaporation model and the system thermodynamics can be described by interfacial free energies associated with the surfaces of cavitands contacting other phases (water or vacuum). From a macroscopic perspective, the free energy of a wetted state is proportional to the surface area of the pocket (with a proportionality constant equals to the specific solid/water interfacial tension). Likewise, the free energy of a dry state is proportional to the area of the air/water interface across the portal accompanied by a corresponding air/water interfacial tension[132,134]. A certain state with lower free energy will be favored by the system. Therefore, the dry state will be favored if the pocket has a larger pocket surface area, while the wet state will be favored with a larger tension in air/water interface at portal. In the comparison between TEMOA and TEXMOA, when the methyl groups are displaced from *exo*-positions to *endo*-positions, the air/water interface area at portal is narrowed since the opening of pocket is smaller. Meanwhile, the capacity of pocket is not altered but the pocket surface area is partially increased. All of these mutually contributed to the system free energy and make the dry state preferred over the wet state. This macroscopic picture of surface tensions and free energy is questionable, especially when it is applied to cavitands with nano-scale pockets. However, our simulation results reach a qualitatively agreement with this capillary evaporation proposition. And a detailed capillary evaporation model is developed in chapter 5.

Figure F3b reports the average partial molar volume of all three cavitands as a function of hydration number within their pockets. Naturally, the partial molar volume is smaller when more waters are occupying the pocket. Thus, it is a decreasing function of the pocket hydration number for all the hosts. When the hydration condition is the same, the partial molar volumes of TEXMOA and TEMOA are basically the same. This further demonstrated that the partial molar volume difference between TEXMOA and TEMOA under ambient pressure doesn't originate from intrinsic difference in the packing of the waters inside or inner structures of their pockets. Additionally, if we shift the partial molar volume profile of OA upwards by ~81 cm^3/mol, which is the reported partial molar volume difference between TEXMOA and OA from simulation, an excellent consistence between the shifted curve and the profiles of TEMOA and TEXMOA can be observed. Apparently, accommodating same number of waters inside, the primary difference in partial molar volumes of OA with TEMOA and TEXMOA only comes from the physical volume of the four additional methyl units. Combining experiments and computational work, our results verify that the methylation on *endo*-positions of rim can increase the preference of cavitand inner pocket to completely dry states over relatively wet states. And it leads to the anomalously large partial molar volumes of TEMOA over the other two cavitands.

A practical question regarding this cavity self-dewetting can be asked is: how does it affect the host-guest binding? Here a combination of simulations and experiments are used to investigate this impact. In the aspect of experiments, Isothermal Titration Calorimetric (ITC) experiments were performed in our collaborator, Dr. Bruce Gibb's group, to evaluate the thermodynamics of the binding processes of selected n-alkyl carboxylates guests within hydrated OA and poorly hydrated TEMOA. ITC can directly report the enthalpy of

complexation (ΔH), and then yields the Gibbs free energy (ΔG), entropy contribution ($-T\Delta S$), and heat capacity (ΔC_p). As shown in table F6, not surprisingly, for both OA and TEMOA, larger-sized guests have higher binding affinities. The binding is always stronger (more negative ΔG) with TEMOA for all the guests compared to OA. Interestingly, this binding process with TEMOA is more exothermic (relatively more negative (ΔH) and more penalized in entropy (more positive $-T\Delta S$). This category of binding phenomena enthalpically dominated and more penalized entropically is described as "non-classical hydrophobic effect"[180]. In addition, during the binding process, there is a heat capacity loss in the whole system for both hosts upon host and guest desolvation. But in general, with whole range of guests, this heat capacity loss is more prominent in the case of TEMOA (very negative values in $\Delta\Delta C_p$).

TableF 6: Thermodynamics for the complexation of fatty acid to OA (host 1) and TEMOA (host 2) at 298.15 K and their corresponding differences.

Host 1

Guest	K_a (M^{-1})	ΔG (kcal/mol)	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔC_p(cal/mol K)
hexanoate	6.05×10^3	-5.16	-5.74	0.59	-68.3
heptanoate	3.80×10^4	-6.24	-6.49	0.24	-89.0
octanoate	1.37×10^5	-7.00	-6.12	-0.89	-115.6
nonanoate	3.37×10^5	-7.54	-6.46	-1.08	-133.3
decanoate	6.33×10^5	-7.91	-6.48	-1.43	-154.5

Host 2

Guest	K_a (M^{-1})	ΔG (kcal/mol)	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔC_p(cal/mol K)
hexanoate	2.70×10^4	-6.04	-7.53	1.49	-95.1
heptanoate	2.35×10^5	-7.32	-8.54	1.21	-119.3
octanoate	1.49×10^6	-8.42	-8.69	0.27	-143.3
nonanoate	3.51×10^6	-8.92	-9.67	0.75	-164.9
decanoate	4.40×10^6	-9.05	-9.08	0.03	-182.1

Difference(2-1)

Guest	K_a (M^{-1})	$\Delta\Delta G$ (kcal/mol)	$\Delta\Delta H$ (kcal/mol)	$-T\Delta\Delta S$ (kcal/mol)	$\Delta\Delta C_p$(cal/mol K)
hexanoate	--	-0.88	-1.79	0.90	-26.8
heptanoate	--	-1.08	-2.05	0.97	-30.3
octanoate	--	-1.42	-2.57	1.16	-27.7
nonanoate	--	-1.38	-3.21	1.83	-31.6
decanoate	--	-1.14	-2.60	1.46	-27.6

To figure out the reasons for those difference in binding thermodynamics between OA and TEMOA, the drying free energies of all three cavitands were evaluated based on the hydration probability distributions from simulations in a range of temperatures. In figure F6, the free energies of water evacuation process from the hosts pockets, defined as $\Delta G_{dry} = -kTlnp(0)$, is plotted as a function of temperature. Roughly speaking, OA and TEXMOA exhibit comparably the same trends in the drying free energy. They are both linearly decreasing functions with the increase of temperature. On the other hand, the drying free energy of TEMOA is quite flat. And from a

macroscopic perspective, the drying of TEMOA's pocket is much less unfavorable (less positive ΔG_{dry}), while this evacuation process in OA and TEXMOA is more penalized. The drying enthalpies and entropies can be extracted from the total drying free energies by a linear fitting (*i.e.*, $\Delta G_{dry} = \Delta H_{dry} - T\Delta S_{dry}$, where ΔH_{dry} and ΔS_{dry} are both independent of temperature). ΔH_{dry} is the interception and $-\Delta S_{dry}$ is the slope. Those thermodynamic data are reported in table F7.

Detailly, For OA and TEXMOA, this drying is more favorable in entropy (more negative $-T\Delta S_{dry}$) and unfavorable enthalpically (positive ΔH_{dry}). And quantitatively, the entropic and enthalpic contributions to their drying free energies are both nearly the same. In the case of TEMOA, the drying enthalpy and entropy are both smaller by an order of magnitude. Considering the fact that there are fewer water molecules within the pocket, the less contributions from enthalpy and entropy can be explained reasonably. It is interesting to observe that for all three cavitands, the drying process is enthalpically penalized. Despite the loss of hydrogen bonds when interfacial water molecules enter the pocket (*vide infra*), favorable van der Waals interactions can be built between waters and atoms on cavitand. The combined influence of these two factors leads to this enthalpically penalized drying.
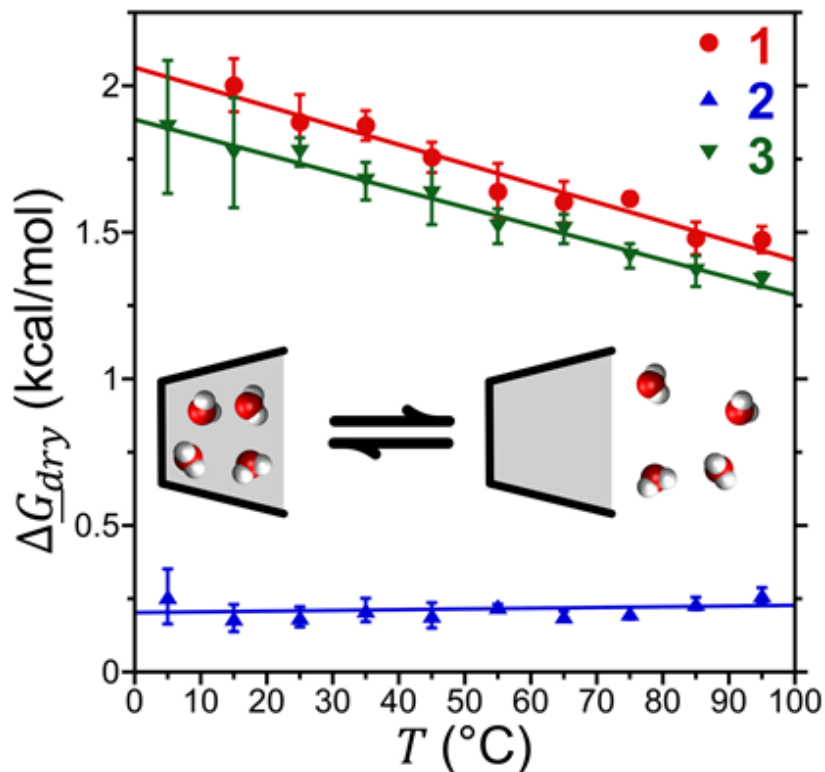
FigureF 6: Drying free energy ($\Delta G_{dry} = -kT\ln p(0)$) of all three cavitands as a function of temperatures determined from simulations using scaled charges. The symbols for hosts 1-3 are identified in the figure legend. The linear fits of the drying free energies to the expression $\Delta G_{dry} = \Delta H_{dry} - T\Delta S_{dry}$ are shown by the lines. ΔH_{dry} and ΔS_{dry} are both assumed to be independent of temperature. The error bars indicate one standard deviation.

TableF 7: Cavitand drying free energies, enthalpies, and entropies at 298.15 K as determined from linear fits of simulation drying free energies (Figure F6). Thermodynamic properties reported in groups of kcal/mol.

	1	2	3	2-1	3-1
ΔG_{dry}	1.90 ± 0.02	0.21 ± 0.02	1.74 ± 0.01	-1.69 ± 0.03	-0.16 ± 0.03
ΔH_{dry}	3.86 ± 0.20	0.13 ± 0.18	3.53 ± 0.09	-3.73 ± 0.27	-0.33 ± 0.26
$-T\Delta S_{dry}$	-1.96 ± 0.18	0.08 ± 0.16	-1.79 ± 0.08	2.04 ± 0.24	0.17 ± 0.23
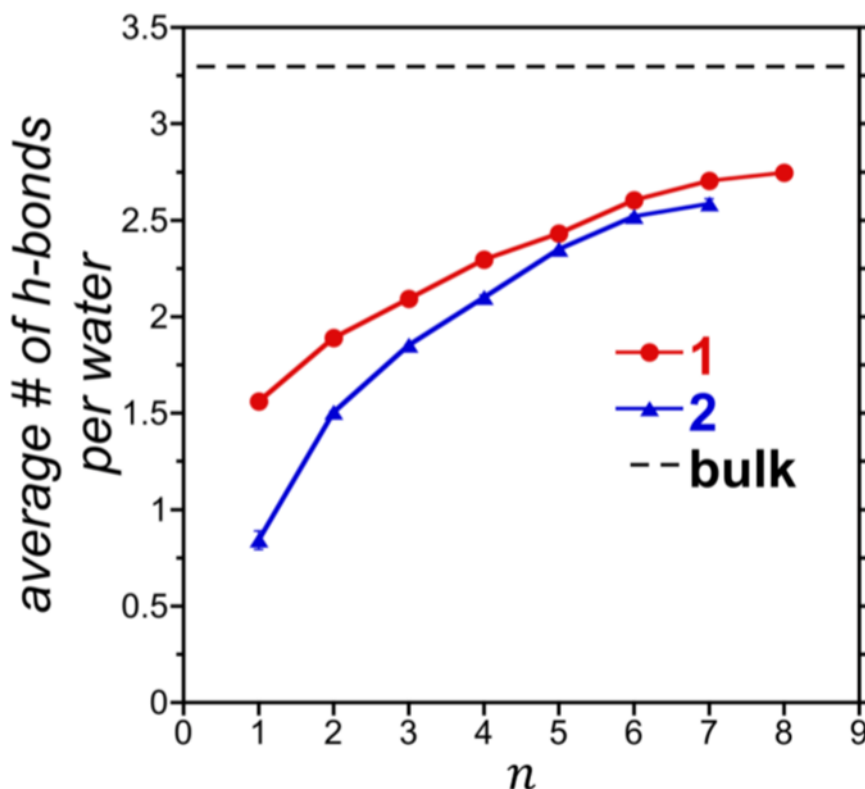
FigureF 7: Average number of water-water hydrogen bonds for individual water molecule within pockets of OA (host 1) and TEMOA (host 2) as a function of hydration number. They are evaluated in simulations at 298.15 K and 1 bar using scaled charges. The corresponding average hydrogen number in bulk water is represented by the horizontal dash line.

In figure F7, it can be confirmed that the encapsulation of water molecules within cavitand

pocket will reduce their opportunities for hydrogen bonding. It reports the average number

of water-water hydrogen bonds formed for each individual water within pocket as a

function of hydration number. The calculation is also done for bulk water. In general, the

number of hydrogen-bonds for each water in pocket is significantly lower compared to

bulk. When the hydration number is smaller or in a relative drier state, the opportunity to

form hydrogen bonds is further reduced since there are fewer neighbor waters. Besides,

within TEMOA (host **2**), the number of hydrogen-bonds is lower for each water compared

to that in OA (host **1**) in whole range of hydration conditions. Though at high hydration numbers, they appear to converge to one another. It further corresponds to that the drying process of TEMOA is less enthalpically penalized and waters are in states with higher energy. The drying of OA is very unfavorable enthalpically, and this might be due to the van der Waals interactions or even additional hydrogen bonding with atoms in cavitand. Water is sacrificing its hydrogen bonding with neighbor waters during the process of absorption to pocket of OA.

If we compare the data presented in table F6 and F7, an excellent agreement can be observed both in sign and magnitudes of these thermodynamic contributions. For OA, the binding process is less enthalpically favored (relatively more positive ΔH) and less entropically penalized (more negative $-T\Delta S$). It is due to its higher water occupancy since the drying process itself is also very unfavorable in enthalpy (positive ΔH) and favorable in entropy (negative $-T\Delta S$). The release of water contributes to the whole complexation process and more dominantly it places extra penalty on the part of enthalpy. Therefore, the total binding is weaker. In case of TEMOA, due to the absence of water within pocket, the overall guest binding process is stronger. When the guest is approaching the pocket of TEMOA, it doesn't need to compete with water molecules and can skip the water expelling step. Thus, it can avoid that large enthalpic penalty.

This impact of solvent in pocket on host-guest binding can also be inferred from the potential-of-mean force (PMF), which can quantitatively describe the change of free energy between two interacting groups along a certain pre-defined reaction coordinate. Figure F8 is showing the PMF between a hexanoate guest and a single empty cavitand (OA and TEMOA) to form a 1:1 complex along the C_4-axes of symmetry. The distance z is

measured from the center of the guest to the center of the cavitand portal. Hence, the distance is negative ($z < 0$) when the center of the guest is inside the pocket, while a positive distance ($z > 0$) means the center of the guest is outside.
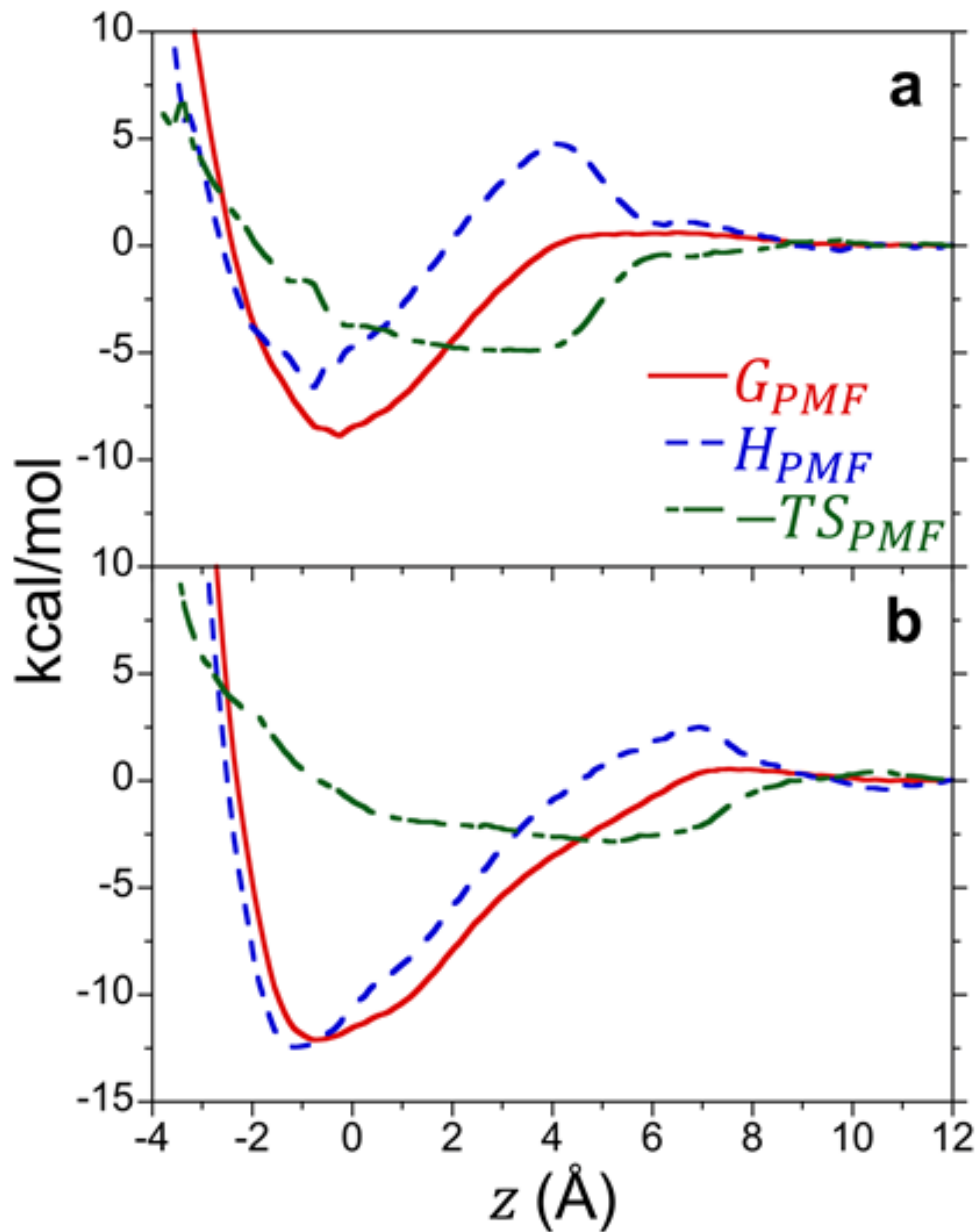
FigureF 8: Potential of mean force between a hexanoate and a single cavitand (OA in **a** and TEMOA in **b**) to form a 1:1 complex in water at 298.15 K and 1 bar. It is measured along the host C4-axis of symmetry as a function of distance z, which is the distance between the hexanoate center-of-mass and the center of cavitand portal. The PMF (G_{PMF}), was also broken down into its enthalpic (H_{PMF}) and entropic ($-TS_{PMF}$)components assuming the free energy is linear with temperature. The lines are identified in the legend of **a**. Error bars have been neglected for clarity.

In both cases, a deep attractive well can be observed where the center of the hexanoate guest is bound inside the pockets ($z < 0$). Thus, overall, the host-guest binding process is favorable and the 1:1 complex is stable in water. Within TEMOA, PMF exhibits a deeper well depth indicating the binding is stronger. It is consistent with what is reported (more negative ΔG for TEMOA) from experiments in Table F6. Considering the PMF minima when they are break down into their enthalpic and entropic components, in both cases they are dominantly favored by the enthalpy and only weakly favored or even opposed by the entropy. Comparing the two PMFs, at minimum, it is more enthalpically favored in case of TEMOA over OA. Meanwhile, it is less strongly favored by the entropy for TEMOA. It can be further demonstrated, combining all the thermodynamic data in binding and pocket drying (Table F6 and F7), and together with the PMF, the presence of less water molecules in TEMOA leads to a significant impact on binding. It doesn't need to account for the additional enthalpic penalty to expel the water and as a result the binding is stronger.

This guest approaching process can be visualized by the plot of average hydration number within pocket as a function of the distance between the guest center of mass and the center of the pocket portal. This can be evaluated from the PMF and illustrated in figure F9. In case of OA, when guest was outside, the average hydration number is stable at n=4 with very minor fluctuation. When one end of the guest entered, around $z = {\sim}4\text{Å}$, the waters were expelled out and the average hydration number drastically decreased to zero. Correspondingly, the PMF started to decrease from zero and became favorable from the same place. On the other hand, the average hydration number within TEMOA exhibited large variance when guest was outside. Not surprisingly, the average water numbers are significantly lower than that of OA even the guest was still far away. Due to the self-

dewetting behavior, waters can be displaced from the pocket very easily so that it started to decrease rapidly at where the guest center-portal center distance was still ~ 7Å. That can partially explain, for TEMOA, the extent of the PMF's attractive well has a longer range. It began to decrease also at $z = \sim 7$Å. For both cavitands, the enthalpic component (H_{PMF}) has an unfavorable hump at where waters started to be evacuated. And this hump is more prominent in OA compared to that of TEMOA. Relevantly, the entropic component ($-TS_{PMF}$) became relatively more favorable at the same node.
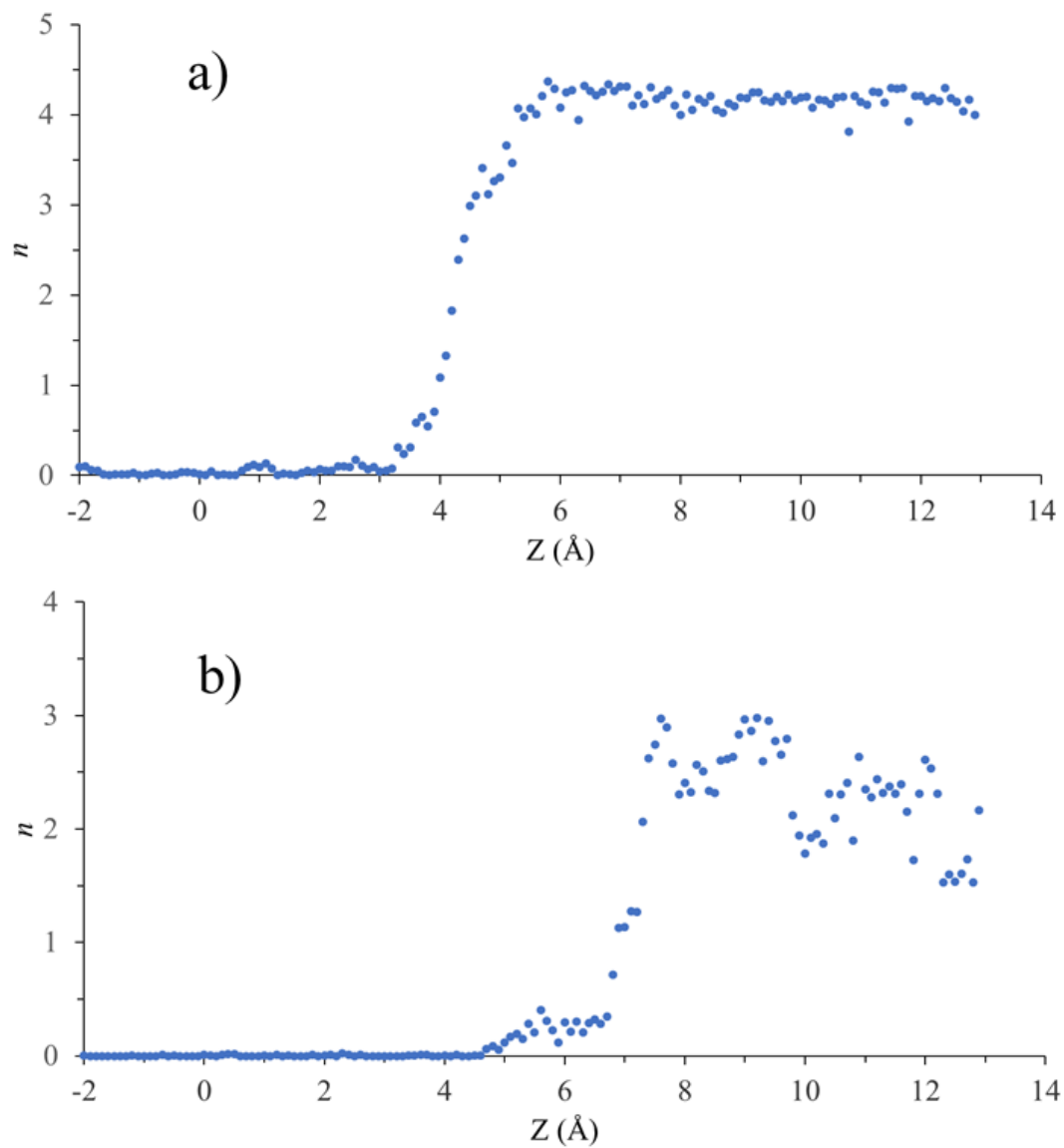
FigureF 9: Average hydration number within OA(a) and TEMOA (b) as a function of the distance z. This distance is measured between the center-of-mass of hexanoate guest and the center of the cavitand portal.

F.4   Conclusion

In conclusion, our findings both in experiments and simulations are the first self-dewetting behaviors observed within a synthetic host in aqueous solution. When there is no methylation on the rim (OA) or the methyl units are at *exo*-positions (TEXMOA), the pocket is more wet, and the average hydration number is high. When the methylation occurs at those pointing-inward *endo*-positions (TEMOA), the evacuation of water is triggered in a way similar with capillary evaporation. It has a significant impact on guest binding. The dry pocket of TEMOA will make the binding in total more favorable. It is less penalized in enthalpy and more unfavorable in entropy. When guest is approaching the binding pocket, there are fewer waters to compete with it. Therefore, it needs to pay less enthalpically to expel hydration water out of the pocket. We also proposed a macroscopic perspective regarding this phenomenon, where there is a balance between the combined effect from attractive van der Waals interactions and limited hydrogen-bonding opportunities to form a cavitand/water interface, against the formation of air/water interface at the portal. These results potentially provide us a new approach to design host-guest binding systems. Specifically, structural modification can be conducted on host pockets or binding sites to change its hydration propensity and wetting tendency, therefore to shift the balance towards dry or wet states and affect the total binding behaviors.

References

(1)   Mayhew, M.; da Silva, A. C. R.; Martin, J.; Erdjument-Bromage, H.; Tempst, P.; Hartl, F. U. Protein Folding in the Central Cavity of the GroEL–GroES Chaperonin Complex. *Nature* **1996**, *379* (6564), 420–426.

(2)   Muñoz, I. G.; Yébenes, H.; Zhou, M.; Mesa, P.; Serna, M.; Park, A. Y.; Bragado-Nilsson, E.; Beloso, A.; de Cárcer, G.; Malumbres, M.; et al. Crystal Structure of the Open Conformation of the Mammalian Chaperonin CCT in Complex with Tubulin. *Nat. Struct. Mol. Biol.* **2011**, *18* (1), 14–19.

(3)   Tanaka, H.; Kato, K.; Yamashita, E.; Sumizawa, T.; Zhou, Y.; Yao, M.; Iwasaki, K.; Yoshimura, M.; Tsukihara, T. The Structure of Rat Liver Vault at 3.5 Angstrom Resolution. *Science (80-.).* **2009**, *323* (5912), 384 LP – 388.

(4)   Chugani, D. C.; Rome, L. H.; Kedersha, N. L. Evidence That Vault Ribonucleoprotein Particles Localize to the Nuclear Pore Complex. *J. Cell Sci.* **1993**, *106* (1), 23 LP – 29.

(5)   Purohit, P. K.; Kondev, J.; Phillips, R. Mechanics of DNA Packaging in Viruses. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (6), 3173–3178.

(6)   Purohit, P. K.; Inamdar, M. M.; Grayson, P. D.; Squires, T. M.; Kondev, J.; Phillips, R. Forces during Bacteriophage DNA Packaging and Ejection. *Biophys. J.*

2005, *88* (2), 851–866.

(7)   Delattre, A.-S.; Saint, N.; Clantin, B.; Willery, E.; Lippens, G.; Locht, C.; Villeret, V.; Jacob-Dubuisson, F. Substrate Recognition by the POTRA Domains of TpsB Transporter FhaC. *Mol. Microbiol.* **2011**, *81* (1), 99–112.

(8)   van den Berg, B.; Black, P. N.; Clemons, W. M.; Rapoport, T. A. Crystal Structure of the Long-Chain Fatty Acid Transporter FadL. *Science (80-.).* **2004**, *304* (5676), 1506 LP – 1509.

(9)   Sacchettini, J. C.; Gordon, J. I.; Banaszak, L. J. Crystal Structure of Rat Intestinal Fatty-Acid-Binding Protein: Refinement and Analysis of the Escherichia Coli-Derived Protein with Bound Palmitate. *J. Mol. Biol.* **1989**, *208* (2), 327–339.

(10)   Wikström, M.; Verkhovsky, M. I.; Hummer, G. Water-Gated Mechanism of Proton Translocation by Cytochrome c Oxidase. *Biochim. Biophys. Acta - Bioenerg.* **2003**, *1604* (2), 61–65.

(11)   Pedersen, C. J. The Discovery of Crown Ethers (Noble Lecture). *Angew. Chemie Int. Ed. English* **1988**, *27* (8), 1021–1027.

(12)   Fujita, M.; Umemoto, K.; Yoshizawa, M.; Fujita, N.; Kusukawa, T.; Biradha, K. Molecular Paneling Coordination. *Chem. Commun.* **2001**, No. 6, 509–518.

(13)   Fujita, D.; Ueda, Y.; Sato, S.; Mizuno, N.; Kumasaka, T.; Fujita, M. Self-Assembly of Tetravalent Goldberg Polyhedra from 144 Small Components. *Nature* **2016**, *540*, 563.

(14)   MacGillivray, L. R.; Atwood, J. L. A Chiral Spherical Molecular Assembly Held

Together by 60 Hydrogen Bonds. *Nature* **1997**, *389* (6650), 469–472.

(15)   Avram, L.; Cohen, Y. Spontaneous Formation of Hexameric Resorcinarene Capsule in Chloroform Solution as Detected by Diffusion NMR. *J. Am. Chem. Soc.* **2002**, *124* (51), 15148–15149.

(16)   Slovak, S.; Avram, L.; Cohen, Y. Encapsulated or Not Encapsulated? Mapping Alcohol Sites in Hexameric Capsules of Resorcin[4]Arenes in Solution by Diffusion NMR Spectroscopy. *Angew. Chemie Int. Ed.* **2010**, *49* (2), 428–431.

(17)   Barrett, E. S.; Dale, T. J.; Rebek, J. Assembly and Exchange of Resorcinarene Capsules Monitored by Fluorescence Resonance Energy Transfer. *J. Am. Chem. Soc.* **2007**, *129* (13), 3818–3819.

(18)   Ebbing, M. H. K.; Villa, M.-J.; Valpuesta, J.-M.; Prados, P.; de Mendoza, J. Resorcinarenes with 2-Benzimidazolone Bridges: Self-Aggregation, Self-Assembled Dimeric Capsules, and Guest Encapsulation. *Proc. Natl. Acad. Sci.* **2002**, *99* (8), 4962 LP – 4966.

(19)   Barrett, E. S.; Dale, T. J.; Rebek, J. Self-Assembly Dynamics of a Cylindrical Capsule Monitored by Fluorescence Resonance Energy Transfer. *J. Am. Chem. Soc.* **2007**, *129* (28), 8818–8824.

(20)   Scarso, A.; Shivanyuk, A.; Rebek, J. Individual Solvent/Solute Interactions through Social Isomerism. *J. Am. Chem. Soc.* **2003**, *125* (46), 13981–13983.

(21)   Scarso, A.; Trembleau, L.; Rebek Jr., J. Encapsulation Induces Helical Folding of Alkanes. *Angew. Chemie Int. Ed.* **2003**, *42* (44), 5499–5502.

(22)   Ajami, D.; Rebek Jr., J. Longer Guests Drive the Reversible Assembly of Hyperextended Capsules. *Angew. Chemie Int. Ed.* **2007**, *46* (48), 9283–9286.

(23)   Ajami, D.; Rebek, J. Expanded Capsules with Reversibly Added Spacers. *J. Am. Chem. Soc.* **2006**, *128* (16), 5314–5315.

(24)   Tiefenbacher, K.; Ajami, D.; Rebek Jr., J. Self-Assembled Capsules of Unprecedented Shapes. *Angew. Chemie Int. Ed.* **2011**, *50* (50), 12003–12007.

(25)   Barnett, J. W.; Tang, D.; Gibb, B. C.; Ashbaugh, H. S. Alkane Guest Packing Drives Switching between Multimeric Deep-Cavity Cavitand Assembly States. *Chem. Commun.* **2018**, *54* (21), 2639–2642.

(26)   Masson, E.; Ling, X.; Joseph, R.; Kyeremeh-Mensah, L.; Lu, X. Cucurbituril Chemistry: A Tale of Supramolecular Success. *RSC Adv.* **2012**, *2* (4), 1213–1247.

(27)   Freeman, W. A.; Mock, W. L.; Shih, N. Y. Cucurbituril. *J. Am. Chem. Soc.* **1981**, *103* (24), 7367–7368.

(28)   Mukhopadhyay, P.; Zavalij, P. Y.; Isaacs, L. High Fidelity Kinetic Self-Sorting in Multi-Component Systems Based on Guests with Multiple Binding Epitopes. *J. Am. Chem. Soc.* **2006**, *128* (43), 14093–14102.

(29)   Masson, E.; Lu, X.; Ling, X.; Patchell, D. L. Kinetic vs Thermodynamic Self-Sorting of Cucurbit[6]Uril, Cucurbit[7]Uril, and a Spermine Derivative. *Org. Lett.* **2009**, *11* (17), 3798–3801.

(30)   Zhang, H.; Wang, Q.; Liu, M.; Ma, X.; Tian, H. Switchable V-Type [2]Pseudorotaxanes. *Org. Lett.* **2009**, *11* (15), 3234–3237.

(31)   Crini, G. Review: A History of Cyclodextrins. *Chem. Rev.* **2014**, *114* (21), 10940–10975.

(32)   Gidwani, B.; Vyas, A. A Comprehensive Review on Cyclodextrin-Based Carriers for Delivery of Chemotherapeutic Cytotoxic Anticancer Drugs. *Biomed Res. Int.* **2015**, *2015*, 198268.

(33)   Vyas, A.; Saraf, S.; Saraf, S. Cyclodextrin Based Novel Drug Delivery Systems. *J. Incl. Phenom. Macrocycl. Chem.* **2008**, *62* (1), 23–42.

(34)   Radhouan, M.; Abderrahim, R. Applications of Cyclodextrins: Formation of Inclusion Complexes and Their Characterization. *Int. J. Adv. Res.* **2015**, *3*, 1030.

(35)   Jiang, Y.; Jiang, X.; Law, K.; Chen, Y.; Gu, J.; Zhang, W.; Xin, H.; Sha, X.; Fang, X. Enhanced Anti-Tumor Effect of 9-Nitro-Camptothecin Complexed by Hydroxypropyl-β-Cyclodextrin and Safety Evaluation. *Int. J. Pharm.* **2011**, *415* (1), 252–258.

(36)   Lahiani-Skiba, M.; Bounoure, F.; Fessi, H.; Skiba, M. Effect of Cyclodextrins on Lonidamine Release and In-Vitro Cytotoxicity. *J. Incl. Phenom. Macrocycl. Chem.* **2011**, *69* (3), 481–485.

(37)   Çirpanli, Y.; Bilensoy, E.; Lale Doğan, A.; Çaliş, S. Comparative Evaluation of Polymeric and Amphiphilic Cyclodextrin Nanoparticles for Effective Camptothecin Delivery. *Eur. J. Pharm. Biopharm.* **2009**, *73* (1), 82–89.

(38)   Zhang, X.; Zhang, X.; Wu, Z.; Gao, X.; Cheng, C.; Wang, Z.; Li, C. A Hydrotropic β-Cyclodextrin Grafted Hyperbranched Polyglycerol Co-Polymer for

Hydrophobic Drug Delivery. *Acta Biomater.* **2011**, *7* (2), 585–592.

(39)   Tanford, C. How Protein Chemists Learned about the Hydrophobic Factor. *Protein Sci.* **1997**, *6* (6), 1358–1366.

(40)   Kauzmann, W. Some Factors in the Interpretation of Protein Denaturation11The Preparation of This Article Has Been Assisted by a Grant from the National Science Foundation.; Anfinsen, C. B., Anson, M. L., Bailey, K., Edsall, J. T. B. T.-A. in P. C., Eds.; Academic Press, 1959; Vol. 14, pp 1–63.

(41)   Biedermann, F.; Nau, W. M.; Schneider, H.-J. The Hydrophobic Effect Revisited—Studies with Supramolecular Complexes Imply High-Energy Water as a Noncovalent Driving Force. *Angew. Chemie Int. Ed.* **2014**, *53* (42), 11158–11171.

(42)   Smithrud, D. B.; Wyman, T. B.; Diederich, F. Enthalpically Driven Cyclophane-Arene Inclusion Complexation: Solvent-Dependent Calorimetric Studies. *J. Am. Chem. Soc.* **1991**, *113* (14), 5420–5426.

(43)   Moghaddam, S.; Yang, C.; Rekharsky, M.; Ko, Y. H.; Kim, K.; Inoue, Y.; Gilson, M. K. New Ultrahigh Affinity Host−Guest Complexes of Cucurbit[7]Uril with Bicyclo[2.2.2]Octane and Adamantane Guests: Thermodynamic Analysis and Evaluation of M2 Affinity Calculations. *J. Am. Chem. Soc.* **2011**, *133* (10), 3570–3581.

(44)   Biedermann, F.; Uzunova, V. D.; Scherman, O. A.; Nau, W. M.; De Simone, A. Release of High-Energy Water as an Essential Driving Force for the High-Affinity Binding of Cucurbit[n]Urils. *J. Am. Chem. Soc.* **2012**, *134* (37), 15318–15323.

(45)   Setny, P.; Baron, R.; McCammon, J. A. How Can Hydrophobic Association Be Enthalpy Driven? *J. Chem. Theory Comput.* **2010**, *6* (9), 2866–2871.

(46)   Ernst, J. A.; Clubb, R. T.; Zhou, H.-X.; Gronenborn, A. M.; Clore, G. M. Demonstration of Positionally Disordered Water Within a Protein Hydrophobic Cavity by NMR. *Science (80-.).* **1995**, *267* (5205), 1813–1817.

(47)   Yin, H.; Feng, G.; Clore, G. M.; Hummer, G.; Rasaiah, J. C. Water in the Polar and Nonpolar Cavities of the Protein Interleukin-1β. *J. Phys. Chem. B* **2010**, *114* (49), 16290–16297.

(48)   Collins, M. D.; Quillin, M. L.; Hummer, G.; Matthews, B. W.; Gruner, S. M. Structural Rigidity of a Large Cavity-Containing Protein Revealed by High-Pressure Crystallography. *J. Mol. Biol.* **2007**, *367* (3), 752–763.

(49)   Collins, M. D.; Hummer, G.; Quillin, M. L.; Matthews, B. W.; Gruner, S. M. Cooperative Water Filling of a Nonpolar Protein Cavity Observed by High-Pressure Crystallography and Simulation. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (46), 16668–16671.

(50)   Qvist, J.; Davidovic, M.; Hamelberg, D.; Halle, B. A Dry Ligand-Binding Cavity in a Solvated Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (17), 6296–6301.

(51)   Pascal, T. A.; Goddard, W. A.; Jung, Y. Entropy and the Driving Force for the Filling of Carbon Nanotubes with Water. *Proc. Natl. Acad. Sci.* **2011**, *108* (29), 11794 LP – 11798.

(52)   Vaitheeswaran, S.; Yin, H.; Rasaiah, J. C.; Hummer, G. Water Clusters in

Nonpolar Cavities. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (49), 17002–17005.

(53)   Wang, L.; Zhao, J.; Fang, H. Water Clusters Confined in Nonpolar Cavities by Ab

Initio Calculations. *J. Phys. Chem. C* **2008**, *112* (31), 11779–11785.

(54)   Vaitheeswaran, S.; Rasaiah, J. C.; Hummer, G. Electric Field and Temperature

Effects on Water in the Narrow Nonpolar Pores of Carbon Nanotubes. *J. Chem.*

Phys. **2004**, *121* (16), 7955–7965.

(55)   Dzubiella, J.; Allen, R. J.; Hansen, J.-P. Electric Field-Controlled Water

Permeation Coupled to Ion Transport through a Nanopore. *J. Chem. Phys.* **2004**,

120 (11), 5001–5004.

(56)   Gibb, C. L. D.; Gibb, B. C. Well-Defined, Organic Nanoenvironments in Water:

The Hydrophobic Effect Drives a Capsular Assembly. *J. Am. Chem. Soc.* **2004**,

126 (37), 11408–11409.

(57)   Gibb, C. L. D.; Gibb, B. C. Templated Assembly of Water-Soluble Nano-

Capsules: Inter-Phase Sequestration, Storage, and Separation of Hydrocarbon

Gases. *J. Am. Chem. Soc.* **2006**, *128* (51), 16498–16499.

(58)   Liu, S.; Gan, H.; Hermann, A. T.; Rick, S. W.; Gibb, B. C. Kinetic Resolution of

Constitutional Isomers Controlled by Selective Protection inside a Supramolecular

Nanocapsule. *Nat. Chem.* **2010**, *2*, 847.

(59)   Sundaresan, A. K.; Ramamurthy, V. Making a Difference on Excited-State

Chemistry by Controlling Free Space within a Nanocapsule: Photochemistry of 1-

(4-Alkylphenyl)-3-Phenylpropan-2-Ones. *Org. Lett.* **2007**, *9* (18), 3575–3578.

(60)   Sundaresan, A. K.; Ramamurthy, V. Consequences of Controlling Free Space within a Reaction Cavity with a Remote Alkyl Group: Photochemistry of Para-Alkyl Dibenzyl Ketones within an Organic Capsule in Water. *Photochem. Photobiol. Sci.* **2008**, *7* (12), 1555–1564.

(61)   Jordan, J. H.; Gibb, B. C. Molecular Containers Assembled through the Hydrophobic Effect. *Chem. Soc. Rev.* **2015**, *44* (2), 547–585.

(62)   Sun, H.; Gibb, C. L. D.; Gibb, B. C. Calorimetric Analysis of the 1:1 Complexes Formed between a Water-Soluble Deep-Cavity Cavitand, and Cyclic and Acyclic Carboxylic Acids. *Supramol. Chem.* **2008**, *20* (1–2), 141–147.

(63)   Gibb, C. L. D.; Gibb, B. C. Straight-Chain Alkanes Template the Assembly of Water-Soluble Nano-Capsules. *Chem. Commun.* **2007**, No. 16, 1635–1637.

(64)   Liu, S.; Russell, D. H.; Zinnel, N. F.; Gibb, B. C. Guest Packing Motifs within a Supramolecular Nanocapsule and a Covalent Analogue. *J. Am. Chem. Soc.* **2013**, *135* (11), 4314–4324.

(65)   Barnett, J. W.; Gibb, B. C.; Ashbaugh, H. S. Succession of Alkane Conformational Motifs Bound within Hydrophobic Supramolecular Capsular Assemblies. *J. Phys. Chem. B* **2016**, *120* (39), 10394–10402.

(66)   Gibb, C. L. D.; Gibb, B. C. Guests of Differing Polarities Provide Insight into Structural Requirements for Templates of Water-Soluble Nano-Capsules. *Tetrahedron* **2009**, *65* (35), 7240–7248.

(67)   Gan, H.; Benjamin, C. J.; Gibb, B. C. Nonmonotonic Assembly of a Deep-Cavity

Cavitand. *J. Am. Chem. Soc.* **2011**, *133* (13), 4770–4773.

(68) Gan, H.; Gibb, B. C. Guest-Controlled Self-Sorting in Assemblies Driven by the Hydrophobic Effect. *Chem. Commun.* **2012**, *48* (11), 1656–1658.

(69) Gan, H.; Gibb, B. C. Guest-Mediated Switching of the Assembly State of a Water-Soluble Deep-Cavity Cavitand. *Chem. Commun.* **2013**, *49* (14), 1395–1397.

(70) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.

(71) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.

(72) Frenkel, D.; Smit, B. Chapter 4 - Molecular Dynamics Simulations; Frenkel, D., Smit, B. B. T.-U. M. S. (Second E., Eds.; Academic Press: San Diego, 2002; pp 63–107.

(73) McMillan, W. G.; Mayer, J. E. The Statistical Thermodynamics of Multicomponent Systems. *J. Chem. Phys.* **1945**, *13* (7), 276–305.

(74) Schnell, S. K.; Englebienne, P.; Simon, J.-M.; Krüger, P.; Balaji, S. P.; Kjelstrup, S.; Bedeaux, D.; Bardow, A.; Vlugt, T. J. H. How to Apply the Kirkwood–Buff Theory to Individual Species in Salt Solutions. *Chem. Phys. Lett.* **2013**, *582*, 154–157.

(75) Koga, K. Osmotic Second Virial Coefficient of Methane in Water. *J. Phys. Chem. B* **2013**, *117* (41), 12619–12624.

(76) Ashbaugh, H. S.; Weiss, K.; Williams, S. M.; Meng, B.; Surampudi, L. N. Temperature and Pressure Dependence of Methane Correlations and Osmotic Second Virial Coefficients in Water. *J. Phys. Chem. B* **2015**, *119* (20), 6280–6294.

(77) Chaudhari, M. I.; Sabo, D.; Pratt, L. R.; Rempe, S. B. Hydration of Kr(Aq) in Dilute and Concentrated Solutions. *J. Phys. Chem. B* **2015**, *119* (29), 9098–9102.

(78) Chaudhari, M. I.; Holleran, S. A.; Ashbaugh, H. S.; Pratt, L. R. Molecular-Scale Hydrophobic Interactions between Hard-Sphere Reference Solutes Are Attractive and Endothermic. *Proc. Natl. Acad. Sci.* **2013**, *110* (51), 20557 LP – 20562.

(79) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

(80) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81* (1), 511–519.

(81) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31* (3), 1695–1697.

(82) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.

(83) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123* (23), 234505.

(84) Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *J. Phys. Chem. B* **1998**, *102* (14), 2569–

2577.

(85)   Chen, B.; Potoff, J. J.; Siepmann, J. I. Monte Carlo Calculations for Alcohols and Their Mixtures with Alkanes. Transferable Potentials for Phase Equilibria. 5. United-Atom Description of Primary, Secondary, and Tertiary Alcohols. *J. Phys. Chem. B* **2001**, *105* (15), 3093–3104.

(86)   Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.

(87)   Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13* (8), 952–962.

(88)   Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 1987.

(89)   Chitra, R.; Smith, P. E. Molecular Association in Solution: A Kirkwood−Buff Analysis of Sodium Chloride, Ammonium Sulfate, Guanidinium Chloride, Urea, and 2,2,2-Trifluoroethanol in Water. *J. Phys. Chem. B* **2002**, *106* (6), 1491–1500.

(90)   Schnell, S. K.; Liu, X.; Simon, J.-M.; Bardow, A.; Bedeaux, D.; Vlugt, T. J. H.; Kjelstrup, S. Calculating Thermodynamic Properties from Fluctuations at Small Scales. *J. Phys. Chem. B* **2011**, *115* (37), 10911–10918.

(91)   Lepori, L.; Matteoli, E.; Hamad, E. Z.; Mansoori, G. A. *Fluctuation Theory of Mixtures*; Matteoli, E., Mansoori, G. A., Eds.; Taylor & Francis, New York, 1990,

1990.

(92)   Ben-Naim, A. Inversion of the Kirkwood–Buff Theory of Solutions: Application to the Water–Ethanol System. *J. Chem. Phys.* **1977**, *67* (11), 4884–4890.

(93)   Blokzijl, W.; Engberts, J. B. F. N. Hydrophobic Effects. Opinions and Facts. *Angew. Chemie Int. Ed. English* **1993**, *32* (11), 1545–1579.

(94)   Rankin, B. M.; Ben-Amotz, D.; van der Post, S. T.; Bakker, H. J. Contacts Between Alcohols in Water Are Random Rather than Hydrophobic. *J. Phys. Chem. Lett.* **2015**, *6* (4), 688–692.

(95)   Lum, K.; Chandler, D.; Weeks, J. D. Hydrophobicity at Small and Large Length Scales. *J. Phys. Chem. B* **1999**, *103* (22), 4570–4577.

(96)   Inokuma, Y.; Arai, T.; Fujita, M. Networked Molecular Cages as Crystalline Sponges for Fullerenes and Other Guests. *Nat. Chem.* **2010**, *2*, 780.

(97)   Inokuma, Y.; Yoshioka, S.; Ariyoshi, J.; Arai, T.; Hitora, Y.; Takada, K.; Matsunaga, S.; Rissanen, K.; Fujita, M. X-Ray Analysis on the Nanogram to Microgram Scale Using Porous Complexes. *Nature* **2013**, *495*, 461.

(98)   Browne, C.; Ramsay, W. J.; Ronson, T. K.; Medley-Hallam, J.; Nitschke, J. R. Carbon Dioxide Fixation and Sulfate Sequestration by a Supramolecular Trigonal Bipyramid. *Angew. Chemie* **2015**, *127* (38), 11274–11279.

(99)   Hastings, C. J.; Backlund, M. P.; Bergman, R. G.; Raymond, K. N. Enzyme-like Control of Carbocation Deprotonation Regioselectivity in Supramolecular Catalysis of the Nazarov Cyclization. *Angew. Chemie* **2011**, *123* (45), 10758–

10761.

(100) Brown, C. J.; Toste, F. D.; Bergman, R. G.; Raymond, K. N. Supramolecular Catalysis in Metal–Ligand Cluster Hosts. *Chem. Rev.* **2015**, *115* (9), 3012–3035.

(101) Ajami, D.; Rebek, J. More Chemistry in Small Spaces. *Acc. Chem. Res.* **2013**, *46* (4), 990–999.

(102) Hillyer, M. B.; Gibb, B. C. Molecular Shape and the Hydrophobic Effect. *Annu. Rev. Phys. Chem.* **2016**, *67* (1), 307–329.

(103) Ben-Amotz, D. Water-Mediated Hydrophobic Interactions. *Annu. Rev. Phys. Chem.* **2016**, *67* (1), 617–638.

(104) Wanjari, P. P.; Sangwai, A. V; Ashbaugh, H. S. Confinement Induced Conformational Changes in N-Alkanes Sequestered within a Narrow Carbon Nanotube. *Phys. Chem. Chem. Phys.* **2012**, *14* (8), 2702–2709.

(105) Siu, S. W. I.; Pluhackova, K.; Böckmann, R. A. Optimization of the OPLS-AA Force Field for Long Hydrocarbons. *J. Chem. Theory Comput.* **2012**, *8* (4), 1459–1470.

(106) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23* (16), 1623–1641.

(107) Ewell, J.; Gibb, B. C.; Rick, S. W. Water Inside a Hydrophobic Cavitand Molecule. *J. Phys. Chem. B* **2008**, *112* (33), 10272–10279.

(108) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.;

Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120* (20), 9665–9678.

(109) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011–1021.

(110) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. *Biophys. J.* **1997**, *72* (3), 1047–1069.

(111) Woo, H.-J.; Roux, B. Calculation of Absolute Protein–Ligand Binding Free Energy from Computer Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6825 LP – 6830.

(112) Luo, H.; Sharp, K. On the Calculation of Absolute Macromolecular Binding Free Energies. *Proc. Natl. Acad. Sci.* **2002**, *99* (16), 10399 LP – 10404.

(113) Widom, B. Potential-Distribution Theory and the Statistical Mechanics of Fluids. *J. Phys. Chem.* **1982**, *86* (6), 869–872.

(114) POLLACK, G. L. Why Gases Dissolve in Liquids. *Science (80-.).* **1991**, *251* (4999), 1323 LP – 1330.

(115) Vapour Pressures up to Their Critical Temperatures of Normal Alkanes and 1-Alkanols . *Pure and Applied Chemistry* . 1989, p 1395.

(116) Du, X.; Li, Y.; Xia, Y.-L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int.*

J. Mol. Sci. **2016**, *17* (2), 144.

(117) Janin, J. Protein-Protein Recognition. *Prog. Biophys. Mol. Biol.* **1995**, *64* (2), 145–166.

(118) Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008**, *108* (1), 74–108.

(119) Snyder, P. W.; Mecinovic, J.; Moustakas, D. T.; Thomas, S. W.; Harder, M.; Mack, E. T.; Lockett, M. R.; Héroux, A.; Sherman, W.; Whitesides, G. M. Mechanism of the Hydrophobic Effect in the Biomolecular Recognition of Arylsulfonamides by Carbonic Anhydrase. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (44), 17889–17894.

(120) Myslinski, J. M.; DeLorbe, J. E.; Clements, J. H.; Martin, S. F. Protein-Ligand Interactions: Thermodynamic Effects Associated with Increasing Nonpolar Surface Area. *J. Am. Chem. Soc.* **2011**, *133* (46), 18518–18521.

(121) Lagona, J.; Mukhopadhyay, P.; Chakrabarti, S.; Isaacs, L. The Cucurbit[n]Uril Family. *Angew. Chemie Int. Ed.* **2005**, *44* (31), 4844–4870.

(122) Jayaraj, N.; Zhao, Y.; Parthasarathy, A.; Porel, M.; Liu, R. S. H.; Ramamurthy, V. Nature of Supramolecular Complexes Controlled by the Structure of the Guest Molecules: Formation of Octa Acid Based Capsuleplex and Cavitandplex. *Langmuir* **2009**, *25* (18), 10575–10586.

(123) Biros, S. M.; Ullrich, E. C.; Hof, F.; Trembleau, L.; Rebek, J. Kinetically Stable Complexes in Water: The Role of Hydration and Hydrophobicity. *J. Am. Chem.*

Soc. **2004**, *126* (9), 2870–2876.

(124) Hooley, R. J.; Van Anda, H. J.; Rebek, J. Cavitands with Revolving Doors Regulate Binding Selectivities and Rates in Water. *J. Am. Chem. Soc.* **2006**, *128* (12), 3894–3895.

(125) Zhang, K.-D.; Ajami, D.; Gavette, J. V; Rebek, J. Complexation of Alkyl Groups and Ghrelin in a Deep, Water-Soluble Cavitand. *Chem. Commun.* **2014**, *50* (38), 4895–4897.

(126) Gavette, J. V; Zhang, K.-D.; Ajami, D.; Rebek, J. Folded Alkyl Chains in Water-Soluble Capsules and Cavitands. *Org. Biomol. Chem.* **2014**, *12* (34), 6561–6563.

(127) Zhang, K.-D.; Ajami, D.; Gavette, J. V; Rebek, J. Alkyl Groups Fold to Fit within a Water-Soluble Cavitand. *J. Am. Chem. Soc.* **2014**, *136* (14), 5264–5266.

(128) Stillinger, F. H. Structure in Aqueous Solutions of Nonpolar Solutes from the Standpoint of Scaled-Particle Theory. *J. Solution Chem.* **1973**, *2* (2), 141–158.

(129) Weeks, J. D.; Katsov, K.; Vollmayr, K. Roles of Repulsive and Attractive Forces in Determining the Structure of Nonuniform Liquids: Generalized Mean Field Theory. *Phys. Rev. Lett.* **1998**, *81* (20), 4400–4403.

(130) Hummer, G.; Garde, S. Cavity Expulsion and Weak Dewetting of Hydrophobic Solutes in Water. *Phys. Rev. Lett.* **1998**, *80* (19), 4193–4196.

(131) Ashbaugh, H. S.; Pratt, L. R. Colloquium: Scaled Particle Theory and the Length Scales of Hydrophobicity. *Rev. Mod. Phys.* **2006**, *78* (1), 159–178.

(132) Chandler, D. Interfaces and the Driving Force of Hydrophobic Assembly. *Nature*

2005, *437* (7059), 640–647.

(133) Cerdeiriña, C. A.; Debenedetti, P. G.; Rossky, P. J.; Giovambattista, N. Evaporation Length Scales of Confined Water and Some Common Organic Liquids. *J. Phys. Chem. Lett.* **2011**, *2* (9), 1000–1003.

(134) Ashbaugh, H. S. Solvent Cavitation under Solvophobic Confinement. *J. Chem. Phys.* **2013**, *139* (6), 64702.

(135) Altabet, Y. E.; Haji-Akbari, A.; Debenedetti, P. G. Effect of Material Flexibility on the Thermodynamics and Kinetics of Hydrophobically Induced Evaporation of Water. *Proc. Natl. Acad. Sci.* **2017**, *114* (13), E2548 LP-E2555.

(136) Remsing, R. C.; Xi, E.; Vembanur, S.; Sharma, S.; Debenedetti, P. G.; Garde, S.; Patel, A. J. Pathways to Dewetting in Hydrophobic Confinement. *Proc. Natl. Acad. Sci.* **2015**, *112* (27), 8181 LP – 8186.

(137) Truskett, T. M.; Debenedetti, P. G.; Torquato, S. Thermodynamic Implications of Confinement for a Waterlike Fluid. *J. Chem. Phys.* **2001**, *114* (5), 2401–2418.

(138) Huang, D. M.; Chandler, D. The Hydrophobic Effect and the Influence of Solute−Solvent Attractions. *J. Phys. Chem. B* **2002**, *106* (8), 2047–2053.

(139) Patel, A. J.; Varilly, P.; Jamadagni, S. N.; Hagan, M. F.; Chandler, D.; Garde, S. Sitting at the Edge: How Biomolecules Use Hydrophobicity to Tune Their Interactions and Function. *J. Phys. Chem. B* **2012**, *116* (8), 2498–2503.

(140) Godawat, R.; Jamadagni, S. N.; Garde, S. Characterizing Hydrophobicity of Interfaces by Using Cavity Formation, Solute Binding, and Water Correlations.

Proc. Natl. Acad. Sci. **2009**, *106* (36), 15119 LP – 15124.

(141) Zhu, F.; Hummer, G. Pore Opening and Closing of a Pentameric Ligand-Gated Ion Channel. *Proc. Natl. Acad. Sci.* **2010**, *107* (46), 19814 LP – 19819.

(142) Beckstein, O.; Sansom, M. S. P. Ahydrophobic Gate in an Ion Channel: The Closed State of the Nicotinic Acetylcholine Receptor. *Phys. Biol.* **2006**, *3* (2), 147–159.

(143) Anishkin, A.; Sukharev, S. Water Dynamics and Dewetting Transitions in the Small Mechanosensitive Channel MscS. *Biophys. J.* **2004**, *86* (5), 2883–2895.

(144) Setny, P.; Wang, Z.; Cheng, L.-T.; Li, B.; McCammon, J. A.; Dzubiella, J. Dewetting-Controlled Binding of Ligands to Hydrophobic Pockets. *Phys. Rev. Lett.* **2009**, *103* (18), 187801.

(145) Baron, R.; Setny, P.; McCammon, J. A. Water in Cavity-Ligand Recognition.(Report). *J. Am. Chem. Soc.* **2010**, *132* (34), 12091–12097.

(146) Hummer, G. Molecular Binding: Under Water's Influence. *Nat. Chem.* **2010**, *2* (11), 906–907.

(147) Biedermann, F.; Vendruscolo, M.; Scherman, O. A.; De Simone, A.; Nau, W. M. Cucurbit[8]Uril and Blue-Box: High-Energy Water Release Overwhelms Electrostatic Interactions. *J. Am. Chem. Soc.* **2013**, *135* (39), 14879–14888.

(148) Frenkel, D.; Smit, B. *Understanding Molecular Simulation : From Algorithms to Applications. 2nd Ed*; 1996; Vol. 50.

(149) Evans, R. Fluids Adsorbed in Narrow Pores: Phase Equilibria and Structure. *J.*

Phys. Condens. Matter **1990**, *2* (46), 8989–9007.

(150) Huang, X.; Margulis, C. J.; Berne, B. J. Dewetting-Induced Collapse of Hydrophobic Particles. *Proc. Natl. Acad. Sci.* **2003**, *100* (21), 11953 LP – 11958.

(151) Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **1990**, *29* (31), 7133–7155.

(152) Levy, Y.; Onuchic, J. N. WATER MEDIATION IN PROTEIN FOLDING AND MOLECULAR RECOGNITION. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35* (1), 389–415.

(153) Chong, S.-H.; Ham, S. Interaction with the Surrounding Water Plays a Key Role in Determining the Aggregation Propensity of Proteins. *Angew. Chemie Int. Ed.* **2014**, *53* (15), 3961–3964.

(154) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* **2012**, *45* (1), 83–92.

(155) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for Molecular Recognition Exploiting Hydrophobic Enclosure in Protein–Ligand Binding. *Proc. Natl. Acad. Sci.* **2007**, *104* (3), 808 LP – 813.

(156) Schiebel, J.; Gaspari, R.; Wulsdorf, T.; Ngo, K.; Sohn, C.; Schrader, T. E.; Cavalli, A.; Ostermann, A.; Heine, A.; Klebe, G. Intriguing Role of Water in Protein-Ligand Binding Studied by Neutron Crystallography on Trypsin Complexes. *Nat. Commun.* **2018**, *9* (1), 3559.

(157) Patel, A. J.; Varilly, P.; Chandler, D. Fluctuations of Water near Extended

Hydrophobic and Hydrophilic Surfaces. *J. Phys. Chem. B* **2010**, *114* (4), 1632–1637.

(158) Varilly, P.; Patel, A. J.; Chandler, D. An Improved Coarse-Grained Model of Solvation and the Hydrophobic Effect. *J. Chem. Phys.* **2011**, *134* (7), 74109.

(159) Patel, A. J.; Varilly, P.; Jamadagni, S. N.; Acharya, H.; Garde, S.; Chandler, D. Extended Surfaces Modulate Hydrophobic Interactions of Neighboring Solutes. *Proc. Natl. Acad. Sci.* **2011**, *108* (43), 17678 LP – 17683.

(160) Patel, A. J.; Garde, S. Efficient Method To Characterize the Context-Dependent Hydrophobicity of Proteins. *J. Phys. Chem. B* **2014**, *118* (6), 1564–1573.

(161) Widom, B. Some Topics in the Theory of Fluids. *J. Chem. Phys.* **1963**, *39*, 2808–2812.

(162) Rego, N. B.; Xi, E.; Patel, A. J. Protein Hydration Waters Are Susceptible to Unfavorable Perturbations. *J. Am. Chem. Soc.* **2019**, *141* (5), 2080–2086.

(163) Patel, A. J.; Varilly, P.; Chandler, D.; Garde, S. Quantifying Density Fluctuations in Volumes of All Shapes and Sizes Using Indirect Umbrella Sampling. *J. Stat. Phys.* **2011**, *145* (2), 265–275.

(164) Xi, E.; Remsing, R. C.; Patel, A. J. Sparse Sampling of Water Density Fluctuations in Interfacial Environments. *J. Chem. Theory Comput.* **2016**, *12* (2), 706–713.

(165) Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* **1976**, *22* (2), 245–268.

(166) Klimovich, P. V; Shirts, M. R.; Mobley, D. L. Guidelines for the Analysis of Free

Energy Calculations. *J. Comput. Aided. Mol. Des.* **2015**, *29* (5), 397–411.

(167) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. Water Conduction through the Hydrophobic Channel of a Carbon Nanotube. *Nature* **2001**, *414* (6860), 188.

(168) Liu, S.; Whisenhunt-Ioup, S. E.; Gibb, C. L. D.; Gibb, B. C. An Improved Synthesis of "octa-Acid" Deep-Cavity Cavitand. *Supramol. Chem.* **2011**, *23* (6), 480–485.

(169) Sullivan, M. R.; Sokkalingam, P.; Nguyen, T.; Donahue, J. P.; Gibb, B. C. Binding of Carboxylate and Trimethylammonium Salts to Octa-Acid and TEMOA Deep-Cavity Cavitands. *J. Comput. Aided. Mol. Des.* **2017**, *31* (1), 21–28.

(170) Gibb, C. L. D.; Gibb, B. C. Binding of Cyclic Carboxylates to Octa-Acid Deep-Cavity Cavitand. *J. Comput. Aided. Mol. Des.* **2014**, *28* (4), 319–325.

(171) Gibb, C. L. D.; Gibb, B. C. Anion Binding to Hydrophobic Concavity Is Central to the Salting-in Effects of Hofmeister Chaotropes. *J. Am. Chem. Soc.* **2011**, *133* (19), 7344–7347.

(172) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225–11236.

(173) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(174) Hedwig, G. R.; Hinz, H.-J. Group Additivity Schemes for the Calculation of the Partial Molar Heat Capacities and Volumes of Unfolded Proteins in Aqueous

Solution. *Biophys. Chem.* **2002**, *100* (1), 239–260.

(175) Kohagen, M.; Mason, P. E.; Jungwirth, P. Accurate Description of Calcium

Solvation in Concentrated Aqueous Solutions. *J. Phys. Chem. B* **2014**, *118* (28),

7902–7909.

(176) Mason, P. E.; Wernersson, E.; Jungwirth, P. Accurate Description of Aqueous

Carbonate Ions: An Effective Polarization Model Verified by Neutron Scattering.

J. Phys. Chem. B **2012**, *116* (28), 8145–8153.

(177) Vazdar, M.; Pluhařová, E.; Mason, P. E.; Vácha, R.; Jungwirth, P. Ions at

Hydrophobic Aqueous Interfaces: Molecular Dynamics with Effective

Polarization. *J. Phys. Chem. Lett.* **2012**, *3* (15), 2087–2091.

(178) Leontyev, I.; Stuchebrukhov, A. Accounting for Electronic Polarization in Non-

Polarizable Force Fields. *Phys. Chem. Chem. Phys.* **2011**, *13* (7), 2613–2626.

(179) Kirkwood, J. G.; Buff, F. P. The Statistical Mechanical Theory of Solutions. I. *J.

Chem. Phys.* **1951**, *19* (6), 774–777.

(180) Cremer, P. S.; Flood, A. H.; Gibb, B. C.; Mobley, D. L. Collaborative Routes to

Clarifying the Murky Waters of Aqueous Supramolecular Chemistry. *Nat. Chem.*

2017, *10*, 8.

Biography

In 1990, Du Tang was born in a beautiful city, TaiCang, which is affiliated to SuZhou in eastern China. He received his B.S. in Materials Engineering from East China University of Science & Technology in 2012. After that, he moved to United States and attended Graduate School at University of Missouri-Columbia. During that time, he focused on the research in synthesis of novel nanomaterials using microwave irradiation. Du received his M.S. in Chemical Engineering from University of Missouri-Columbia in the May of 2014. He joined the Ph.D. program in the Department of Chemical & Biomolecular Engineering in August of the same year, where he focuses on studying the supramolecular self-assembly behaviors using computational simulation methods. After receiving his Ph.D. he plans to pursue industrial positions related to computation, data science, and scientific programming.