2017

Molecular Dynamics Simulation Study Of Water-Soluble Cryptophane Binding A Variety Guest Objects

Wenhao Liu
University of Pennsylvania, mdsimulate@gmail.com

Molecular Dynamics Simulation Study Of Water-Soluble Cryptophane Binding A Variety Guest Objects

Abstract
Given their signicance to chemical sensing and molecular recognition, host-guest interactions are of broad interest in molecular science. Cryptophanes, which are a kind of spherical cage molecules, are one of the most applicable host molecules. Herein computer simulation is used to address the binding of cryptophanes to monatomic guest species in aqueous environments. Water soluble cryptophanes have high binding anity to Xe-129, which is a powerful contrast agent for magnetic resonance imaging and can be used as an NMR-based biosensor. Molecular details of Xe-cryptophane binding are hard to obtain from experiment, however. Extensive molecular dynamics simulation are carried out to understand the underlying molecular features of the Xe-cryptophane system. Free energy perturbation and adaptive biasing force methods are used to obtain the energetic aspects of cryptophane-Xe interactions: binding anities of individual cryptophanes to Xe and potentials of mean force for Xe to entry the cavity. Additionally, the structure adjustments cryptophane undergoes as it accommodates Xe are investigated in depth. From the simulations, the hydrophobicity of cryptophane interior and the kinetics of water enter the cavity are found to be correlated with the Xe binding anity. These ndings shed light on the design of Xe-binding cryptophanes. A related cryptophane is considered for the selective binding of alkali metal ions. The cryptophane has particularly high binding anities with the large cesium cation.

The Cs-137 isotope is a common, problematic radioactive nuclear ssion byproduct. The

simulations provide insight on the relative anities of alkali cations and the roles of water

molecules that enter the cavity along with ions. The simulations reveal that the high bind-

ing anity of cryptophane to Cs+ may result from the comparatively low solvation energy of Cs+ and its geometric t to the cryptophane cavity.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Chemistry

First Advisor
Jeffery G. Saven

Keywords
Cesium ion, Cryptophanes, Free energy calculation, molecular dynamics simulation, Xenon

MOLECULAR DYNAMICS SIMULATION STUDY OF WATER-SOLUBLE CRYPTOPHANE BINDING A VARIETY GUEST OBJECTS

Wenhao Liu

A DISSERTATION

in

Chemistry

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

Jeffery G. Saven, Professor of Chemistry

Graduate Group Chairperson

Gary A. Molander, Hirschmann-Makineni Professor of Chemistry

Dissertation Committee

Andrew M. Rappe, (Chair), Professor of Chemistry

Feng Gai, Edmund J. and Louise W. Kahn Endowed Term Professor of Chemistry

Amish Patel, Reliance Industries Term Assistant Professor of Chemical and Biomolecular Engineering

ACKNOWLEDGMENT

First of all, I would like to thank my advisor, Dr. Saven. For the past five years, I learned so much from him and benefited a lot from his guidance. His sharp scientific inutition and broad knowledge base always provide me brilliant ideas and shed light to the solutions when I encounter difficulty along the study and research. And his attitude of understanding scientific fact as deep and clear as we can and pay attention to the details really set a great example for me to follow and study. I am pretty sure some of these great characters already take effect on me imperceptibly and they are leading me to be a more excellent people and great career. And during the reserach and study along these years, I always feel satisfied when I obatin some data or find something new from simulations although many of them are subtle. Thank you, Jeff.

Secondly, I would like to thank collaborators in the group and outside the group. Dr. Lu Gao, who is a senior schoolmate to me both undergraduate school and graduate school here. Her expertise in the simulation and programing really helpped me a lot. Professor Ivan J. Dmochowski, who is an oustanding bio-chemist provide me ideas and share the brilliant thoughts from the prospect of the experimentalist.

Thirdly, I want to thank my committee professors, Professor Andrew M. Rappe, Professor Feng Gai and Professor Amish Patel. I want to thank their effort on keep me on track of my projects and their helpful comments to the research at each committe meeting along these years.

Then I want to thank all my labmates and friends here, Dr. Christopher J Lanci, Dr. Chris MacDermaid, Dr. Dr. Christopher Von Bargen, Dr. Lu Gao, Dr. Huixi Violet Zhang, Matthew Eibling, Jose Villegas, Jacqueline Blum, Rui Guo, Dr. Yi Fang, Dr. Wenjun Ouyang, Dr. Yue Zhang, Dr. Mengnan Zhang, Dr. Xing Chen, Dr. Chen Li, Dr. Xinle Liu, Dr. Fan Zheng, Dr. Zheng Shi, Jing Yang, Wenjie Dou, Bo Li, Qi Liu, Chun Liu, Rico Zou, thank for your accompany for the past five years.

Finally, I want to thank my Father and Mother. Without your raise, teach, guide and support, I would not able to this step. And I also want to thank my girl friend and my lovely princess, Zi Yuan, who accompany me, trust me and hold me along the ime in Upenn. Thank you!

ABSTRACT

MOLECULAR DYNAMICS SIMULATION STUDY OF WATER-SOLUBLE CRYPTOPHANE BINDING A VARIETY GUEST OBJECTS

Wenhao Liu

Jeffery G. Saven

Given their significance to chemical sensing and molecular recognition, host-guest interactions are of broad interest in molecular science. Cryptophanes, which are a kind of spherical cage molecules, are one of the most applicable host molecules. Herein computer simulation is used to address the binding of cryptophanes to monatomic guest species in aqueous environments. Water soluble cryptophanes have high binding affinity to Xe-129, which is a powerful contrast agent for magnetic resonance imaging and can be used as an NMR-based biosensor. Molecular details of Xe-cryptophane binding are hard to obtain from experiment, however. Extensive molecular dynamics simulation are carried out to understand the underlying molecular features of the Xe-cryptophane system. Free energy perturbation and adaptive biasing force methods are used to obtain the energetic aspects of cryptophane-Xe interactions: binding affinities of individual cryptophanes to Xe and potentials of mean force for Xe to entry the cavity. Additionally, the structure adjustments cryptophane undergoes as it accommodates Xe are investigated in depth. From the simulations, the hydrophobicity of cryptophane interior and the kinetics of water enter the cavity are found to be correlated with the Xe binding affinity. These findings shed light on the design of Xe-binding cryptophanes. A related cryptophane is considered for the selective binding of alkali metal ions. The cryptophane has particularly high binding affinities with the large cesium cation. The Cs-137 isotope is a common, problematic radioactive nuclear fission byproduct. The simulations provide insight on the relative affinities of alkali cations and the roles of water molecules that enter the cavity along with ions. The simulations reveal that the high binding affinity of cryptophane to Cs^+ may result from the comparatively low solvation energy

of Cs^+ and its geometric fit to the cryptophane cavity.

TABLE OF CONTENTS

ACKNOWLEDGMENT ii

ABSTRACT iv

LIST OF TABLES ix

LIST OF ILLUSTRATIONS xiv

CHAPTER 1 : Introduction 1

CHAPTER 2 : Methods 5

2.1 Molecular dynamics simulation 5

2.2 Free energy perturbation method 6

2.3 Adaptive biasing force 9

2.4 Water map analysis 10

CHAPTER 3 : Xe Affinities of Water-Soluble Cryptophanes and the Role of Confined Water 12

3.1 Abstract 12

3.2 Introduction 13

3.3 Methods 16

3.3.1 Alchemical free energy perturbation. 16

3.3.2 MD simulations. 20

3.3.3 Alchemical Free Energy Perturbation 21

3.4 Result 22

3.5 Discussion 36

3.6 Conclusion 39

CHAPTER 4 : Free Energy of Xe Entry into a Cryptophane Interior 41
4.1 Abstract 41
4.2 Introduction 41
4.3 Methods 44
4.3.1 Molecular models. 44
4.3.2 Molecular dynamics simulation. 44
4.3.3 Adaptive biasing force simulation 45
4.3.4 Structure analysis. 46
4.4 Results and Discussion 47
4.4.1 Potential of Mean Force results. 47
4.4.2 Binding affinity calculation. 49
4.4.3 Trajectory analysis. 50
4.4.4 Dihedrals in linkers. 52
4.5 Conclusions 61

CHAPTER 5 : Molecular basis of Alkali cation affinity for water soluble cryptophane 63
5.1 Abstract 63
5.2 Introduction 63
5.3 Methods 65
5.3.1 Molecular models. 65
5.3.2 Molecular dynamics simulation. 66
5.3.3 Free energy perturbation method 66
5.4 Results and discussion 70
5.4.1 Relative binding affinity calculation results. 70
5.4.2 Host-solvent interaction 72
5.4.3 Guest-solvent interaction 73
5.4.4 Ion-solvent interaction in crytophane environment 73
5.4.5 Ion-solvent interaction in water environment 75
5.4.6 Water map analysis 76

5.5 Conclusions 78

CHAPTER 6 : Design high Xe binding affinity cryptophane 80

6.1 abstract 80

6.2 Introduction 80

6.3 Methods 81

6.3.1 Molecular models. 81

6.3.2 Molecular dynamics simulation. 83

6.3.3 Adaptive biasing force simulation 84

6.4 Results and discussions 85

6.4.1 Potential of Mean Force results. 85

6.4.2 Binding affinity calculation. 89

6.5 Conclusions 89

BIBLIOGRAPHY 90

LIST OF TABLES

TABLE 3.1 : The calculated and experimental binding free energies of Xe and water-soluble cryptophanes. [a] Previously reported values 22

TABLE 4.1 : The experimental and calculated binding affinities between Xe and cryptophanes, calculated PMF minima (ΔW_{min}) and calculated PMF maxima (ΔW_{max}). 49

TABLE 5.1 : Binding energy result table 70

TABLE 5.2 : Dimension of cryptophane-1 in conditions of different guest ions: the distance between top aromatic cap and bottom aromatic cap. . . . 71

TABLE 5.3 : Solvation results of ions in cryptophane cavity and in bulk water . 74

TABLE 6.1 : Calculated binding affinities of derivitives of cryptophane TAAC and TTPC to Xe, calculated PMF minima (ΔW_{min}) and calculated PMF maxima (ΔW_{max}). 85

LIST OF ILLUSTRATIONS

FIGURE 2.1 : Strategy of scaling of λ during free energy perturbation process . 8

FIGURE 3.1 : A) The structures of six water-soluble cryptophanes. B) Renderings of cryptophane molecules with Xe bound. 14

FIGURE 3.2 : Thermodynamic cycle of double decoupling method. All simulations are performed in an aqueous environment. The state "0 " indicates the complete decoupling of Xe atom from its environment when potential energy terms involving the Xe atom vanish. . . . 17

FIGURE 3.3 : Experimental6, 20, 21, 24 and calculated free energies of Xe binding. The solid line is a linear least squares fit. The dotted line corresponds to equality of the experimental and calculated values of ΔG. 23

FIGURE 3.4 : Configurations with water molecules inside cryptophanes m3n3, m2n2 and TAAC. In the lower row, each configuration is rotated 90. 25

FIGURE 3.5 : Distribution of water molecules in cryptophane interior. The distance r is measured between the cryptophane center of mass and a water oxygen nucleus. A) Radial distribution function g(r) for water when Xe is present in the interior (= 0). B) Radial distribution function g(r) for water when Xe is absent from the interior (= 1). C) Average number of water molecules N(r) within a distance r from the cryptophane center of mass in the absence of Xe. Curves are obtained from integrating the corresponding g(r). 26

FIGURE 3.6 : A) Water radial distribution function g(r) about Xe in water, where r is Xe-O distance. B) Sampled configuration depicting first solvation shell (waters within 6Å) of Xe in water. 27

FIGURE 3.7 : Water map clustering for hexa-acid cryptophane m3n3. A) Side view and B) top view differ by a rotation of 90. Five position-based water clusters are indicated by wireframe spheres with radius of 1 and are colored in order of decreasing population: orange, yellow, tan, green and purple. Within each cluster, coordinates of six representative water molecules (red) are rendered. 28

FIGURE 3.8 : Na^+ counter ions (yellow) near the m3n3 cryptophane molecule in the absence of Xe ($\lambda = 1$). A) Top view. B) Side view (90o rotation of configuration in A). C) Schematic illustration of counter ion coordination by the side chains. 29

FIGURE 3.9 : The radial distribution g(r) of counter ions relative to the center of mass of the cryptophane molecule. A) Xe atom is absent ($\lambda = 1$). B) Xe atom is present within the cryptophane ($\lambda = 0$). 30

FIGURE 3.10 :Polar contours for distribution of counter ions. The left and right columns are the polar distributions of counter ions when $\lambda = 0$ and $\lambda = 1$, respectively. The origin is set at the center of mass of the 36 aromatic carbon atoms of the cryptophane; the x-y plane is perpendicular to the vector connecting the two centers of mass of the top and bottom caps, and the x axis is the projection of the vector connecting the origin and the average position of two phenol oxygen atoms on the same linker. On the contour plot, the radial coordinate r is the distance between the origin and a counter ion; θ is the angle between the x axis and the projection of r on the x-y plane. Color scale bar is shown above for g(r, θ. 31

FIGURE 3.11 :Radial distribution functions of atoms associated with simulations of the cryptophanes TTEC, TAAC, TTPC, and m2n2 in the absence of Xe. In each case, r is the distance of the indicated atom type to the center of mass of the cryptophane. Blue (backbone): the aromatic carbon atoms of the cryptophane. Green (ionizable groups in the side chains): for TTEC, the ammonium nitrogen atom N; for TAAC, TTPC and m2n2, the central carbon of the carboxylate group. Red (counter ions): Cl- for TTEC and Na+ for TAAC, TTPC, and m2n2. Cyan (water): the oxygen atoms of water molecules. The blue, green and red distribution functions have been rescaled by dividing by 100, 25 and 6, respectively. . . 32

FIGURE 3.12 :Triazole side chains and solvation of TTEC and TTPC. A) Structures of the triazole side chains of TTEC and TTPC. The triazole N atoms are marked blue, and their different terminal groups are marked red. B) Two configurations of triazole side chains in TTEC. C) Radial distribution function of triazole N1 atoms relative to the mass-weighted cryptopane centers in TTEC (blue) and TTPC (black). D) Radial distribution functions of ammonium N atoms (TTEC, blue) and carboxylate C atoms (TTPC, black) relative to the triazole N2 atoms. 35

FIGURE 3.13 :Correlation between the average number of interior water molecules and experimental free energies of Xe binding.6, 20, 21, 24 The red solid line is a linear least squares fit. The radius cutoff for identifying interior waters was set at 4Å relative to the center of mass of the 6 aromatic rings of cryptophane molecules. 37

FIGURE 4.1 : Structure representations of four cryptophanes: TTEC, TTPC, TAAC and m2n2. 42

FIGURE 4.2 : Representation of analyzed dihedrals. 46

FIGURE 4.3 : Calculated free energy profile. 48
FIGURE 4.4 : Concentration of Xe with respect to the reaction coordinate ξ . . 48
FIGURE 4.5 : Experimental and calculated binding affinities. 50
FIGURE 4.6 : Top-bot distance of all four cryptophanes with Xe at different positions. 51
FIGURE 4.7 : Distributions of dihedrals in TAACs linkers from equilibrium simulation and ABF simulation 53
FIGURE 4.8 : Representative structures of TAAC with Xe bounded inside cavity. 53
FIGURE 4.9 : Representative structures of TAAC with Xe at the pore of cavity. 54
FIGURE 4.10 :Distributions of dihedrals in TAACs sidechains from equilibrium simulation and ABF simulation. 55
FIGURE 4.11 :Distributions of dihedrals in TTPCs linkers from equilibrium simulation and ABF simulation. 56
FIGURE 4.12 :Distributions of dihedrals in TTPCs sidechains from equilibrium simulation and ABF simulation. 57
FIGURE 4.13 :Distributions of dihedrals in TTECs linkers from equilibrium simulation and ABF simulation. 58
FIGURE 4.14 :Distributions of dihedrals in TTECs sidechains from equilibrium simulation and ABF simulation. 59
FIGURE 4.15 :Distributions of dihedrals in m2n2s linkers from equilibrium simulation and ABF simulation. 59
FIGURE 4.16 :Distributions of dihedrals in m2n2s sidechains from equilibrium simulation and ABF simulation. 60

FIGURE 5.1 : Strucure representation of cryptophane-1 65
FIGURE 5.2 : Free energy perturbation calculation strategy 68
FIGURE 5.3 : Scaling of λ during free energy perturbation process 68
FIGURE 5.4 : Radial distribution of water molecules around center mass of cryptophane 71

FIGURE 5.5 : Representative structures of ion-water complex inside cryptophane cavity 72
FIGURE 5.6 : Radial distribution of water molecules around ions in conidtion of cryptophane encapsulate ion 74
FIGURE 5.7 : Radial distribution of water molecules around ions in bulk water . 75
FIGURE 5.8 : Solvation of Cs^+ in bulk water 76
FIGURE 5.9 : Water map clustering for vacant cryptophane. Left panel is the horizontal view and right panel is the perpendicular view. Red wire frame is the most populated cluster, the blue one is the second most populated and the yellow one is the thrid most populated frame. In each cluster, five water molecules with most populated orientation are displayed. 77
FIGURE 5.10 :Water map clustering for cryptophane with Cs occupied inside . . 78

FIGURE 6.1 : Renderings of TAAC and TAAC derivitives with methyl get substituted by ethyl, propyl and isopropyl 82
FIGURE 6.2 : Renderings of TTPC and TTPC with 10 more carbons longer in triazole propionic acid sidechain 82
FIGURE 6.3 : Free energy profile of Xe along the coordinate of entering cavity center (ξ) 86
FIGURE 6.4 : Dihedral distribution of the rotation of alkyl group for TAAC derivitives 86
FIGURE 6.5 : Potential of Mean Force of of TAAC derivitives with smaller volume 88

CHAPTER 1 : Introduction

Since the discovery of the first host-guest systems, cyclic polyethers (crown ethers) that bind with metal ions[1, 2], a wide variety of hosts and guests molecules have been discovered and synthesized. Applications of such systems have included directed self assembly on surfaces of nanomaterials[3–6], delivery of target drugs to biological environments of intereset[7–12] and chemsensors to detect specialized guest molecules[13–16]. In host-guest interactions, the host molecule is normally the more complicated molecule, while the guest molecule is usually the simpler species. Among host molecules, cryptophanes, which are a kind of synthetic spherical cage-like molecule, are of great application. They possess adjustable sidechains and interior volumes to accommodate guest particles. Cryptophanes can bind methane[17, 18], chlorofluorocarbons[18], halogenated alkane derivatives[19–21], rare gases[18, 22] and metal ions[23–25].

Among these guests, Xe-129 is one of the most influential, due to its abilitye to be hyperpolarized with regard to its spin states. Xe-129 can act as a powerful magnetic resonance imaging agent and biosensor, since it has high polarizability and the chemical shift is sensitive to the local molecular environment. The interaction of cryptophane and Xe are necessarily noncovalent and dominated by van der Waals interactions. Cryptophanes can be modified so as to be water soluble and able to bind with other molecules, while maintaining high Xe binding affinity. Cryptophane derivatives are broadly used to carry the Xe and localize it within bio-systems when used in imaging based on Xe magnetic reasonance [26–32].

Although a variety of cryptophane molecules have been synthesized, and their binding affinities to Xe and kinetics related with Xe have been measured through isothermal titration experiments[33–36] and hyperpolzation experiment [33–37], molecular pictures are generally hard to obtain from experimentally. Given the experiment conditions are generally atomsphere and room temperature, and cryptophanes have high binding affinity to Xe and

of broadly use are water soluble. Molecular dynamics simulation of solution with cryptophane and Xe as solute is the ideal method to probe the detailed atomistic picture of this host-guest interaction.

In addition to studying Xe as the guest for cryptophane molecules, alkali metal ions are also of great interest. Experimentalists have discovered cryptophanes with binding affinities for large alkali ions, especially for Cesium$^+$[23–25]. Cs-137 is one of the main byproducts of nuclear fission reactions, and the recent Fukushima Daiichi Nuclear Power Plant accident on March 11, 2011 large amounts of ground water becoming contaminated with Cs-137, affecting hundreds of thousands of people [38, 39]. Cryptophanes with high binding affinity and selectively to cesium would be of great interest. Molecular dynamics simulation are used herein to explore the molecular properties of cryptophanes that bind metal ions and provide insight on features associated with high binding affinity to cesium.

In chapter 2, the theoretical background and implementation details of two simulation methods are reviewed. One is free energy perturbation (FEP) method[40, 41], which is used to estimate the free energy change of a process by dividing the whole process into a series of sub-processes. The free energy change in each sub-process can be calculated by averaging the exponential factor of a difference in Hamiltonians, and theses changes can be iterated begining with an initial state. The other approach is the adaptive biasing force (ABF) method[42–45], which applies an external force to cancel the net force along a coordinate of interest. The system then diffuses along the coordinate and visits states that are less probable due to their high intrinsic free energies. The ABF simulation trajetory would provide us sufficient samples to study the binding and conformational fluctuations, including induced fit, involved in host-guest interactions.[46–48]

In chapter 3, the detailed methodology and findings related to the Xe binding affinities of a series of cryptophanes are described. A thermodynamic cycle which employs the double annihilation method is implemented so as to estimate free energy changes associated with cryptophane binding to Xe. Free energy changes associated with introducing Xe in bulk

water and within the cryptophane interior are considered. In each case, the Xe atom disappears during the course of the simulation and becomes decoupled from the system. The binding affinity results come from simulation are compared with experiment. We discovered that the reaction of cryptophane binding Xe can be understood as displacement of water by Xe within crytophane interior. Additionally, the binding affinity between different cryptophanes are found to be correlated with the water occupancy and the residence times of water molecule inside cavity. The water occupancy and residence time can potentially guide the design of high affinity Xe binding cryptophanes.

In chapter 4, the detailed methodology and findings related to the energetic and structural adjustments along Xe entry and exit of the cryptophane are described. We investigated the potentials of mean force of Xe exit from cryptophanes that process different sidechains. The simulations reveal the energy barrier that the aqueous Xe-cryptophane system needs to overcome for entry and free energies associated with translocation into the interior. From the prospect of conformation change of cryptophanes, we investigated the changes of the interior volume of cryptophane associated with Xe exit from the cavity. We examine conformational properties associated with linking groups and how these are twisted as Xe exits. Sidechains are found to flip so as to encapsulate Xe. The major conformational adjustment for cryptophane to accommodate Xe is the flip of cryptophane's hydrophobic sidechains. The simulations shed light on designing cryptophanes that have larger "door keeper" group, which could yield higher Xe binding affinity.

In chapter 5, the detailed methodology and findings related to the relative binding affinities of cryptophane to alkali metal ions and the molecular picture of this cryptophane binding are described. Due to the nature that electrostatic interaction is dominate when compared with van der Waals interaction, free energy perturbation is used to calculate the binding affinity of ions. The electrostatic interaction of vanishing ions and appearing ions are scaled separately, while the van der Waals interactions of both ions are scaled at the same time. A scaled charge set which considers the static dielectric constant of water is used

in the calculations associated with mutating one ion to another inside the cryptophane. The calculated results are consistent with experimental measurements. The reason the cryptophane has selective binding affinity to Cs^+ is due to the low aqueous solvation free energy of the cation and the complementary size of the cryptophane cavity. The roles of water molecules and their distributions are investigated in this chapter as well.

In chapter 6, the detailed strategy and findings related to the design of better Xe binding affinity cryptophanes are described. Making use of the knowledge obtained from previous simulation studies, we propose two approaches to modifying the cryptophane structure. One is to increase the hydrophobicity of the sidechains, and the other is to change (decrease) the interior volume of the cavity. In this work, a widely studied cryptophane, TAAC, is used as the template to test these strategies. In the strategy of increasing hydrophobicity of sidechains, we increase the size of the alkyl sidechain groups and the length of the triazole linker to an ionizable solubilizing group. Together with the potential of mean force calculation of Xe entering cavity, the structural features of these cryptophanes are also studied and compared. This work involves synthesis of the designed systems and their experimental characterization, wherein we explore molecular features that modulate Xe affinity.

CHAPTER 2 : Methods

2.1. Molecular dynamics simulation

Molecular dynamics simulation (MD simulation) is a widely used numerical method to model the atomistic movements and to obtain collective properties in molecular systems. Classical MD simulation assume each particle in the simulation follows Newton's equation,

$$m_i \times \frac{\partial x_i}{\partial t^2} = -\sum_j \frac{\partial U_{i,j}}{\partial x_i} \tag{2.1}$$

where m_i, x_i and $U_{i,j}$ are the mass, position and interaction potential energy involving particle j and particle i. Initial positions for each atom normally come from crystal structure or posited structure, and the potential each atom experience comes from a molecular force field. The interactions in the force field are composed of bonded interactions and non-bonded interactions. Bonded interaction in all the work presented in this thesis are composed of harmonic potentials to describe two atoms in equilibrium bond length, harmonic potential to describe the bond angle formed by three atoms, and cosine functions to specify the portion of the potential energy associated with dihedral angles formed by four atoms. Improper dihedral potentials are also used in the thesis, to constrain four atoms in a plane, such as the four heavy atoms in carboxylate group and the four atoms in phenol groups (three adjacent aromatic atoms and the oxygen atom). Non-bonded interactions are composed of two parts, electrostatic interactions and van der Waals interactions. Electrostatic interactions involve the coulomb interactions between partial charges assigned at each atom. The partial charges in each are derived to fit the gas phase electrostatic potential. For the charges used to calculate free energy change of mutating ions inside cryptophane, all partial charges are scaled by a factor of $\sqrt{\epsilon_{elec}}$ to account for the polarizability of the atoms [49, 50]. Here, ϵ_{elec} is the static dielectric constant of water. Van der Waals interactions are described by 6-12 Lennard-Jones interaction and the σ and ϵ are assigned based on the similar atom types except the paramters of Xe come from viscosity data[51].

In implementation of all simulations, 12Å was used as nonbonded cutoff distance, 10Å was used as switching distance from where the switching function begin to work and 14Å was used as pair list distance to indicate the maximium space to include the neighbouring atoms. The nonbonded pair list was evaluated every 10 steps, and every step is 2fs. Each system was divided into 2Å cubes. The SHAKE algorithm[52] was used to make all bonds between hydrogen and heavy atoms rigid and thus making 2fs timestep simulation correctly. Nonbonded interaction was evaluated every step and electrostatic interaction was evaluated every 2 steps. Particle Mesh Ewald (PME)[53] method was used to evaluate the full system periodic electrostatics with 1Å as grid space. All simulations and calculations are implemented in NPT ensemble, pressure is 1atm and temperature is 300K. In order to implement NPT ensemble simulation, Langevin dynamics method[54] was used to maintain the temperature with damping coefficient to be 1/ps and Langevin piston Nose-Hoover method[55] to maintain the pressure with oscillation period equal to 100fs and decay time equal to 50fs.

2.2. Free energy perturbation method

Helmholtz free energy (A) for an ensemble is related to the partition function (Q), displacement (x) and momentum (p_x) as,

$$A = -\frac{1}{\beta} \cdot lnQ = -\frac{1}{\beta} \cdot ln \int \int exp[-\beta \cdot H(x, p_x)] dx dp_x \tag{2.2}$$

where β is the $\frac{1}{RT}$ and H(x,p_x) is the Hamiltonian of the system. And free energy change between state 0 and state 1 will be,

$$\Delta A = -\frac{1}{\beta} ln(\frac{Q_1}{Q_0}) = -\frac{1}{\beta} ln \frac{\int \int exp[-\beta \cdot H_1(x, p_x)] dx dp_x}{\int \int exp[-\beta \cdot H_0(x, p_x)] dx dp_x} \tag{2.3}$$

And after we simplify the equation2.3, we have

$$\Delta A = -\frac{1}{\beta} ln \frac{\int \int exp[-\beta \cdot \Delta H(x, p_x)] exp[-\beta \cdot H_0(x, p_x)] dx dp_x}{\int \int exp[-\beta \cdot H_0(x, p_x)] dx dp_x} \tag{2.4}$$

Since the probability of finding a sample defined by x and p_x is,

$$P_0(x, p_x) = \frac{exp[-\beta \cdot H_0(x, p_x)]}{\int\int exp[-\beta \cdot H_0(x, p_x)]dxdp_x} \tag{2.5}$$

The free energy change from state 0 to state 1 can be represent as,

$$\Delta A = -\frac{1}{\beta} ln \int\int exp[-\beta \cdot \Delta H(x, p_x)] \cdot P_0(x, p_x) \tag{2.6}$$

And we can also represent free energy change of two states using below formula which is clearer in implementation of free energy perturbation.

$$\Delta A = -\frac{1}{\beta} ln \langle exp[-\beta \cdot \Delta H(x, p_x)] \rangle_0 \tag{2.7}$$

Since the kinetic energy term is unchanged for all cases considered here as the system changes state, in equation 2.7, the ΔH can be substitued by difference in potentials ΔU directly. Therefore, the free energy change from one state to another state can be calculated as: sampling at the initial state, plugging in the force field at the final state to calculate the energy difference, and taking the log after averaging the exponential term involving the energy difference using sampled configurations. However, the free energy change from equation 2.7 is hard to converge if the two states are substantially different, thus small perturbations are required. We divide the whole free energy change process into a series of smaller processes, with comparatively modest changes in the potential energy. The summation of these free energy differences will be the overall free energy change.

During an FEP process, the potential of the system in the intermediate states will be scaled linearly, $H_\lambda = (1\text{-}\lambda)H_0 + \lambda H_1$, where λ is used to represent the progress of the fep process, and λ=0 stands for the initial state and λ=1 stands for the final state. Due to in the systems investigated in this thesis are interaction of Xe and ions with the rest atoms in the system, therefore only the non-bonded interactions are perturbed. In the work of estimating Xe binding affinity of cryptophane, the initial state of Xe FEP process has Xe fully coupled

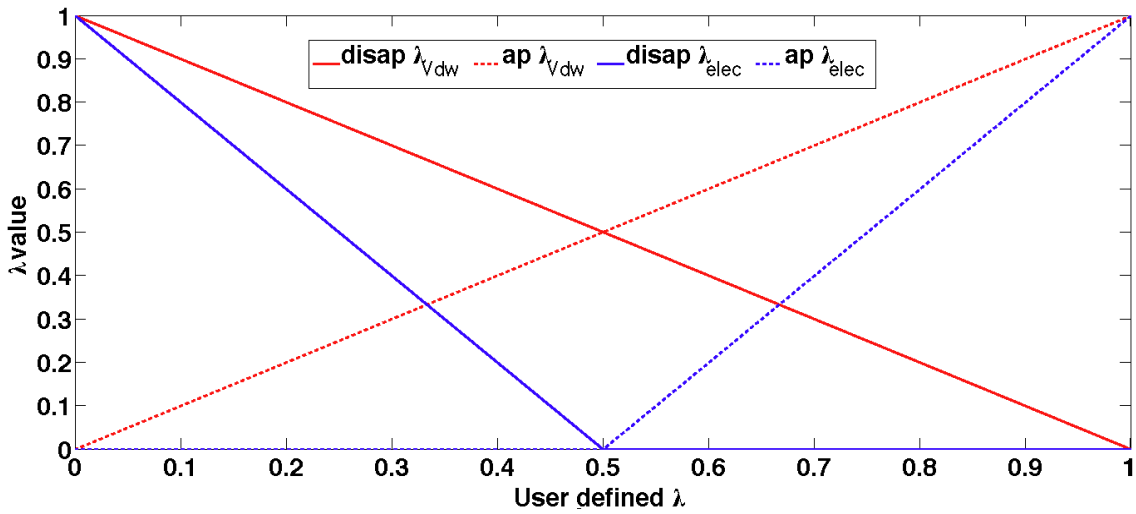


Figure 2.1: Strategy of scaling of λ during free energy perturbation process

to the system and the final state is Xe is completely decoupled. In the work of estimating relatively binding affinity of cryptophane to different ions, the initial state is one ion fully coupled to the system, the final state is the previous ion fully decoupled from the system and the new ion fully coupled to the system, and the intermediate states having two ions exist in the system but they don't interact with each other. In conditions of mutating ions, a new ion is appearing in the system and it maybe collide with existing atoms (atoms in the cryptophane and water molecules), leading to near singularities in simulation3.4. To avoid this, shifted LJ potentials are used to model the vdW interactions. In adddition, the electrostatic and vdW potentials are scaled differently as functions of λ so as to speed up convergence. As in figure 2.2, for the vdW interaction of ions with environment, which are Lennard-Jones potentials in implementation, the interaction of annihilated ("disappeared") ion with environment are scaled at the same rate as global λ; the interaction of introduced ("appeared") ion with environment are scaled up at the same speed as λ as well. For the electrostatic interaction of ions with environment, the speed of scaling down the interaction of annihilated ion with environment and scaling up the final ion with environment occur at a rate twice that of the global λ increase. This leads that the electrostatic interaction of disappearing ion becoming decoupled from system when $\lambda = 0.5$ and then the electrostatic

interaction of appearing ion start to couple to the system and fully couple when λ reaches 1.0.

2.3. Adaptive biasing force

Adaptive biasing force (ABF)[42–45] simulation is a widely used method to investigate host/guest interaction along one or multiple reaction coordinate.

$$A(r) = -\frac{1}{\beta} lng(r) \tag{2.8}$$

In principle, the potential of mean force (PMF) with respect to specific coordinate can be calculated from a distribution function (equ (2.8)) which can be collected from equilibrium simulation. However, it is hard to obtain accurate PMF at high energy points since those high energy configurations are rare and are unlikely to be visited during a simulation. The resulting radial distribution function approaches 0 for such configurations and makes the PMF meaningless at these points. In order to enhance sampling at high energy states along a chosen reaction coordinate, an external force can be applied to bias the simulation. Adaptive biasing force (ABF) simulations can sample efficiently at states we interested in and thereby obtain the free energy profiles by integrating the forces applied to the system along the coordinate. By collecting the forces at one specific general coordinate with the free energy, we would obtain,

$$\frac{dA\left(\xi\right)}{d\xi} = \langle \frac{\partial V\left(x\right)}{\partial \xi} \; - \; \frac{1}{\beta} \frac{\partial ln|J|}{\partial \xi} \rangle_{\xi^*} = -\langle F_\xi \rangle_{\xi^*} \tag{2.9}$$

where ξ is the general coordinate of interest, and |J| is the Jacobian determinant of transforming Cartesian coordinates to generalized coordinates which includes ξ. Herein ξ is the distance of two particles in the system, the forces applied to bias the simulation can also be represent as,

$$F_\xi = \frac{1}{2}(F_2 - F_1) \cdot x_{1,2} + \frac{2}{\beta\xi} \tag{2.10}$$

where F_1 and F_2 are the net forces on particle 1 and 2, $x_{1,2}$ is the unit vector connect particle 1 to particle 2.

In implementation, the whole range of ξ is divided into several windows and each small window is a independent simulation. In each window, the range of ξ is composed of small tiny bins and the biasing forces will be applied at each bin is assumed to the same. Prior to applying the biasing force, a certain number of samples are required for an intial estimate of the biasing force and avoiding "wrong" biasing due to samples are not representative.

2.4. Water map analysis

Water map analysis is a technique to explore the distribution of water molecules as well as the orientations of water molecules in a confined space[56]. Water map analysis is composed of two steps, the first step is the position clustering, which would generate the most populated positions that water is positioned inside the confined space among the obtained simulation samples. The second step is orientation clustering, which would yield orientations for those water molecules that are associated with a given position cluster. Prior to the position clustering, coordinates of all atoms in the simulation box are aligned. In position clustering, we considered all water molecules within the cryptophane cavity and count the number of all other water molecules that are closer than 1.0Å to each of them. The coordinates of the oxygen in the water molecule that has highest number of water molecules nearby is defined to be center of the first cluster. Following the same criteria, we can identify the second and third clusters. The selection of the number of clusters is guided by the average number of water molecules in each confined region, as determined using an unbiased simulation. In orientation clustering, the orientation of water molecules within 1.0Å of the cluster center are considered and compared. The rotation matrix of rotating each water molecule to the reference water molecule are calculated and converted to quanterion representation (q). In order to rotate water molecules to a specific water orientation, we rotate the water molecules with two steps, first is to rotated the dipole vector of water to the reference orientation, and the second is to rotate the hydrogen-hydrogen to

the reference orientation. And the rotation matrix will be the multiplication of these two matrices. The criteria of two water molecules have similar orientation is $|q_i \cdot q_j| > 0.9$, and we count the number of similar orientations each water molecule in cluster has, and the water molecule has highest number is the most representative orientation.

CHAPTER 3 : Xe Affinities of Water-Soluble Cryptophanes and the Role of Confined Water

This chapter is adapted from the article: Lu Gao, Wenhao Liu, One-Sun Lee, Ivan J. Dmochowski, and Jeffery G. Saven. "Xe affinities of water-soluble cryptophanes and the role of confined water." *Chemical Science* 6, no. 12 (2015): 7238-7248.. As a equal contribution author, Wenhao Liu's major contributions to this work include: verification of simulation results; error analysis in Xe binding affinity calculation; determination of distribution of water molecules around cryptophanes; verification of the results with a different force field and at different protonation states of cryptophane; and preparation of text and figures for publication.

3.1. Abstract

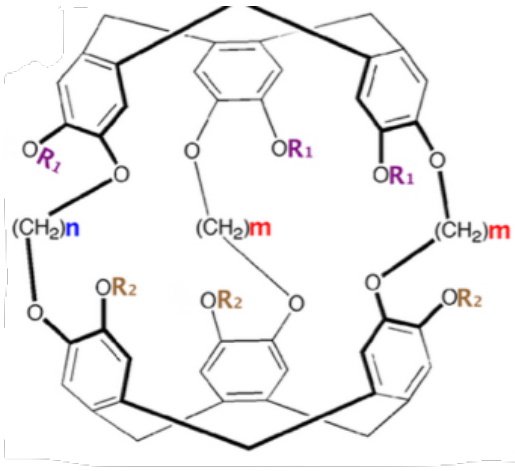
Given their relevance to drug design and chemical sensing, host-guest interactions are of broad interest in molecular science. Natural and synthetic host molecules provide vehicles for understanding selective molecular recognition in aqueous solution. Here, cryptophane-Xe host-guest systems are considered in aqueous media as a model molecular system that also has important applications. ^{129}Xe-cryptophane systems can be used in the creation of biosensors and powerful contrast agents for magnetic resonance imaging applications. Detailed molecular insight on the determinants of Xe affinity is difficult to obtain experimentally. Thus, molecular simulation and free energy perturbation methods were applied to estimate the affinities of Xe for six water-soluble cryptophanes. The calculated affinities correlated well with the previously measured experimental values. The simulations provided molecular insight on the differences in affinities and the roles of conformational fluctuations, solvent, and counter ions on Xe binding to these host molecules. Displacement of confined water from the host interior cavity is a key component of the binding equilibrium, and the average number of water molecules within the host cavity is correlated with the free energy of Xe binding to the different cryptophanes. The findings highlight roles for molecular

simulation and design in modulating the relative strengths of host-guest and host-solvent interactions.

3.2. Introduction

Host-guest interactions are of broad interest in chemistry and offer many potential applications, including the development of drugs, sensors and agents for molecular delivery. Synthetic host-guest systems also provide vehicles for probing the roles of host structure and solvent on the affinities for particular guest species. A variety of noncovalent features can play important roles in the binding of guest molecules, including complementarity of shapes and volumes, van der Waals interactions, hydrogen bonding, hydrophobic contacts, electrostatic interactions, and solvation of both the guest and host. Synthetic host molecules provide a means to dissect some of these interactions, but a detailed molecular picture is often difficult to achieve experimentally. Herein, molecular simulations are used to explore the binding of a rare gas guest, Xe, to cryptophane host molecules in aqueous environments. The findings provide a molecular perspective on the relative affinities of these versatile cryptophane host molecules and highlight the roles of solvent in the formation of the Xe-bound complex.

Cryptophanes are cage-like molecules with two cyclotriguaiacylene caps linked by three variable-length alkyl chains to form nearly spherical cavities of tunable internal volume (figure 3.1). The binding of small molecules to cryptophanes has been explored extensively in organic solvents. These host molecules bind a variety of neutral and charged guest molecules, including methane[17], halogenated alkane derivatives[19–21], alkyl-ammonium ions[57], and rare gases[22]. Water-soluble cryptophane variants have been created[58–60], which exhibit useful affinities for Xe and can be used in biosensors. The binding of Xe in aqueous media is particularly relevant for applications in magnetic resonance based detection and imaging (MRI), where the cryptophane can bind to this essentially inert element. Functionalization of the cage exterior improves cryptophane solubility while also enabling Xe targeting to biomolecules of interest and biosensing.[61]



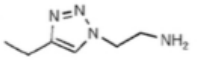
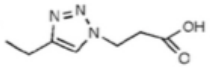
	R1	R2	m	n
TTEC		CH_3	2	2
TTPC		CH_3	2	2
TAAC	CH_2COOH	CH_3	2	2
m2n2	CH_2COOH	CH_2COOH	2	2
m2n3	CH_2COOH	CH_2COOH	2	3
m3n3	CH_2COOH	CH_2COOH	3	3

Figure 3.1: A) The structures of six water-soluble cryptophanes. B) Renderings of cryptophane molecules with Xe bound.

The ^{129}Xe isotope can be hyperpolarized, the basis of its applications in biosensing[62] and MRI[28? –30]. The resonance of the ^{129}Xe spin- nucleus is sensitive to the local chemical environment and presents a nearly 300 ppm range of chemical shifts when bound to cryptophane in water.[31] Changes in the chemical shift of Xe have been used to monitor binding to cryptophanes and to determine molecular affinities.[61]Delivering xenon to a specific biological target, however, can be difficult due the fact that the atom is nonreactive and its weak, non-specific affinity for proteins and other analytes.[63] Thus, cryptophanes provide a means to bind, localize, and "chemically functionalize " Xe in a variety of solvent environments.[62, 64, 65] Appropriately derivatized cryptophanes can be used to target Xe in aqueous environments.[35, 66, 67]

In aqueous media, the Xe affinities for cryptophane derivatives at neutral pH have been measured. These cage variants can be distinguished by the lengths of the bridging linkers between the two cyclotriveratrylene capping moieties and the side chain groups that confer solubility (Figure 1); m and n denote the number of carbon atoms in the three alkyl linkers. While cryptophanes *m2n2* and *m3n3* maintain three-fold symmetry with uniform ethyl and propyl linkers, *m2n3* has one longer (propyl) linker that breaks the symmetry. Three hexa-acid cryptophane derivatives have been reported with Xe association constants

in the range of Ka = 1000-6800 M-1 at rt.[58] Among these hexa-acid cages, the *m2n2* variant (a derivative of cryptophane-A)[67] has the greatest affinity, Ka = 0.68 × 104 M-1 (Figure 3.1). More recently, cryptophanes with only three ionizable side chains have been studied. At pH = 7.5, tris-(triazole ethylamine) cryptophane (TTEC) is triply protonated with Ka = 4.2 × 104 M-1; at pH = 2.5, TTEC is singly protonated with Ka = 3.4 × 104 M-1 (293 K).[36] These are the highest reported cryptophane-Xe affinities in aqueous media. Other soluble derivatives include a triacetic acid cryptophane-A derivative (TAAC, Ka = 3.3 × 104 M-1)[35] and tris-(triazole propionic acid) cryptophane-A derivative (TTPC, Ka = 1.7 × 104 M-1).[36] In addition, a Ru-coordinated cryptophane-111 derivative has a Xe affinity comparable to that of TAAC, Ka = 2.9 × 104 M-1.[68] A variety of distinct experimental methods have been used to determine these Xe affinities. Shifts in NMR spectral peaks have been used to estimate the affinities of the hexa-acid cryptophanes (*m2n2, m2n3, m3n3*) and the cryptophane-111 derivative.[34] , [68] Fluorescence quenching and isothermal titration calorimetry have been used to determine the affinities of TTEC, TAAC, TTPC.[33] , [69] , [70] Interestingly, the central cavities of *m2n2*, TTEC, TTPC, and TAAC are identical in chemical structure, but these molecules have appreciably different Xe affinities. A molecularly detailed, quantitative framework for understanding the relative affinities of this series of cryptophanes is desirable.

Herein, molecular simulations are used to estimate the association free energies involving Xe and six water-soluble cryptophanes (Figure 3.1). The molecular properties of these systems in an aqueous environment are also explored. Molecular simulations have been widely used to estimate affinities of host-guest molecular systems,[71] , [72] and such simulations have provided insight to the structural fluctuations and binding properties of cryptophanes.[73–75] Herein, a free energy perturbation methodology was applied to calculate the binding free energy of Xe to six different water-soluble cryptophane derivatives. The experimentally determined binding constants correspond to a range of binding free energies spanning only about 2 kcal/mol. Such subtle differences can be difficult to discern with free energy calculations. Nonetheless, the calculated results correlate well with measured experimental

values. The simulations also provide molecular insight into the relative binding affinities of different cryptophanes. The distribution of water molecules and counter ions (Na^+, Cl^-) within and near the cryptophane molecules is characterized. The average number of solvent molecules within the interior cavity is found to be correlated with the free energy of Xe binding.

3.3. Methods

3.3.1. Alchemical free energy perturbation.

An alchemical double decoupling method[40, 41] based upon free energy perturbation theory was used to estimate the binding free energy between Xe and cryptophane, $\Delta Gbinding$, in the presence of explicit solvent (water) and counter ions. The binding free energy was estimated by performing two simulations, one with the Xe atom in only water solvent (no cryptophane) and the other with the Xe atom within the interior of the solvated cryptophane. For each simulation, the potential energy interactions that couple the Xe atom to the remainder of the system gradually vanish; Xe is "decoupled " during the course of the calculation. The thermodynamic cycle for estimating $\Delta G_{binding}$ from the simulations is illustrated in Figure3.2. From the thermodynamic cycle, we have :

$$\Delta G_{binding} = \Delta G_1 - \Delta G_2 + \Delta G_{restrain} \tag{3.1}$$

Here, ΔG_1 and ΔG_2 are free energy changes of the two decoupling processes in which the Xe atom is converted into a noninteracting "ideal gas " particle. ΔG_1 is the free energy associated with removing a Xe atom from bulk solvent, as the Xe atom no longer interacts with its local environment. ΔG_2 is the free energy change associated with removing the potential energy interactions the Xe atom has with its environment, while the Xe nucleus is confined inside the host cryptophane; the Xe atom essentially vanishes from the interior of the cage. $\Delta G_{restrain}$ is the change in free energy associated with restraining the noninteracting Xe atom to the interior of the cryptophane.

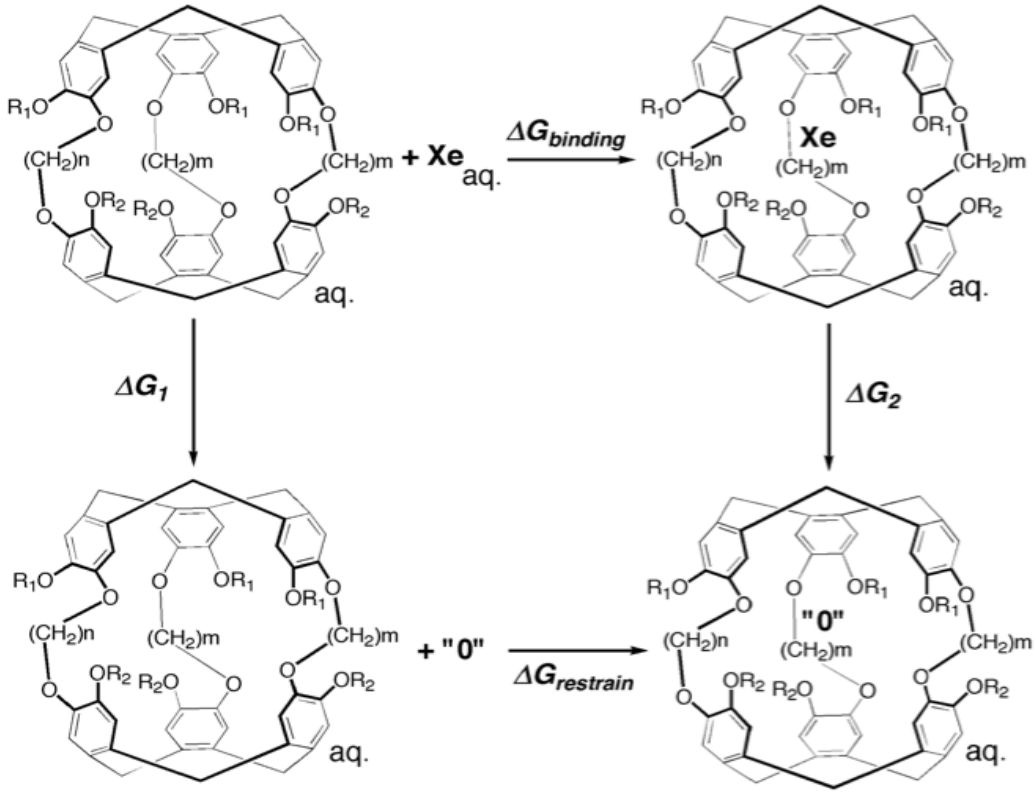


Figure 3.2: Thermodynamic cycle of double decoupling method. All simulations are performed in an aqueous environment. The state "0 " indicates the complete decoupling of Xe atom from its environment when potential energy terms involving the Xe atom vanish.

An alchemical free energy perturbation method[76–80] implemented in NAMD2[81] is used to calculate ΔG_1 and ΔG_2. The hybrid potential energy that comprises interactions between the Xe atom with the other atoms in its environment is determined by a decoupling parameter λ and is expressed as,

$$U_{Xe-X}(\lambda) = (1-\lambda) \cdot U_{Xe-X} \tag{3.2}$$

Here, $UXe\text{-}X$ contains all components of the systems potential energy dependent on the coordinates of the Xe atom. As λ is increased from 0 to 1, the interaction between the Xe atom and its environment vanishes: $U_{Xe-X}(\lambda = 0) = U_{Xe-X}$ and $U_{Xe-X}(\lambda = 1) = 0$. The coupling parameter is discretely incremented to ensure appropriate sampling. The Helmholtz free energy change from one value λi to its neighboring value $\lambda i{+}1$ can be expressed as,

$$\Delta A(\lambda_i \rightarrow \lambda_{i+1}) = -k_B T ln \langle e^{-\frac{U(\lambda_{i+1}) - U(\lambda_i)}{k_B T}} \rangle_i \tag{3.3}$$

The brackets represent an equilibrium ensemble configurational average with the Hamiltonian $H(\lambda_i)$. The Helmholtz free energy differences, $\Delta A1$ and $\Delta A2$, between state $\lambda = 0$ and state $\lambda = 1$ (Xe atom decoupled from its environment), are obtained using

$$\Delta A_n(0 \rightarrow 1) = \sum_{i=1}^{N} \Delta A_n(\lambda_i \rightarrow \lambda_{i+1}) \tag{3.4}$$

Simulations are performed for N values of λ, such that $\lambda 1 = 0$, $\lambda N{+}1 = 1$, and λi ¡ $\lambda i{+}1$. To obtain the estimate of the Gibbs free energy change for the isothermal-isobaric (NPT) ensemble, the work contributed by the volume change between the initial and end states is included.

$$\Delta G_n(0 \rightarrow 1) = \Delta A_n(0 \rightarrow 1) + P \cdot (V_{\lambda=1} - V_{\lambda=0}) \tag{3.5}$$

Where, P is the pressure, 1 atm or 1.01325 bar, and $V_{\lambda=0}$ and $V_{\lambda=1}$ are the volumes of the simulation cell at initial and end states, respectively. Typically, the PV term is negligible, and $\Delta G_n\,(0 \rightarrow 1) = \Delta A_n\,(0 \rightarrow 1)$ The free energy change associated with restraining the noninteracting Xe within the host molecule[40] is calculated using

$$\Delta G_{restrain} = -RTln(c_0 V_{restarin}) \tag{3.6}$$

Here, c_0 is the standard concentration $c_0 = 1$ mol/L or $c_0 = 1$ molecule/1660 Å3. $V_{restrain}$ is the volume accessible to the Xe atom when restrained inside the host molecule. An external isotropic potential *u(r)* is enforced on the Xe atom to confine it within the cryptophane cavity,

$$U(r) = \begin{cases} 0 & r \leq r_{restrain} \\ \frac{1}{2}k(r - r_{restrain})^2 & r > r_{restrain} \end{cases} \tag{3.7}$$

Here, r is the distance between Xe atom and the center of the cryptophane (the center of mass of the aromatic carbon atoms), $r_{restrain}$ was chosen to be 2.8 Å, and k is set to be 100 kcal/mol. After pooling all sampled configurations from all six Xe-cryptophane systems, this confining radius $r_{restrain}$ was determined as the maximum value of r in which Xe was present and fully coupled to its local environment (λ= 0). Since Xe was not observed to exit the cage in any of the simulations, the "restraining force " was essentially not applied. The value of $r_{restrain}$ does enter into the calculation of $\Delta G_{restrain}$

3.3.2. MD simulations.

Force Field Parameters. Structures and atomic coordinates of the water-soluble cryptophanes were based on the X-ray crystal structure of cryptophane-A.[82] The van der Waals (Lennard-Jones) parameters were taken from OPLS.[83] , [84] The van der Waals parameters for Xe atom were taken from measured viscosity data.[51] Initial structures of side chains were built using Spartan, and the geometries were optimized with the MP2 method and 6-31+G* basis set using Gaussian98.[85] Atomic charges of the backbone were adapted

from previous work,[74] and the charges for the side chains were obtained by fitting the electrostatic potential using the R.E.D. tools.[86] (See Figure S1 in Supporting Information (SI).) The bonded parameters were taken from AMBER94.[87] For the missing bond, angle and dihedral parameters, frequency calculations based on the optimized geometries were carried out, and the resulting Hessian matrices transformed into internal coordinates. Potential parameters appear in SI Table S1.

System Preparation. For the calculation of $\Delta G2$, a Xe atom was placed inside the cryptophane cavity, and the complex was solvated in a modified TIP3P[88] water box with periodic boundary conditions using the Solvate plugin implemented in VMD.[89] The initial periodic box size for each system was 36 $\times$36 $\times$36 Å3 and contained about 1200 water molecules. Ionic groups at the end of side-chains were modeled in their fully ionized forms: 3 sodium ions were added to TAAC and TTPCs systems and 3 chloride ions were added to TTEC system solvent box, resulting in a counter ion concentration of [NaCl] = 0.1 M in each case; 6 sodium ions were added to *m2n2*, *m2n3*, and *m3n3* systems, yielding [NaCl] = 0.2 M. The concentrations used in the simulations were targeted to be close to the experimental concentration (0.040 M), while maintaining the overall charge neutrality of the simulated system. Counter ions were added using the Autoionize[89] plug-in of VMD to render each system neutral. A cutoff in the potential energy was set to be 12 Å for non-bonded interactions. Conjugate gradient energy minimization was followed by a 10 ns equilibration with the NPT ensemble. A Langevin dynamics method was used to maintain the temperature of the system at 300 K with a coupling coefficient 1/ps, and pressure was maintained at 1 atm using the Langevin piston Nose-Hoover method with a piston period of 100 fs, a damping time constant of 50 fs, and piston temperature of 300 K. Full electrostatics was employed using the particle mesh Ewald method with a 1 Å grid spacing. Covalent bonds involving hydrogen atoms were held rigid using the SHAKE algorithm,[90] thus allowing a 2-fs time step. For calculating $\Delta G1$, a single Xe atom was simulated in water with no other solutes. All the molecular dynamics simulations were performed using NAMD2.[81]

3.3.3. Alchemical Free Energy Perturbation

(FEP). The values $\lambda = 0$ and $\lambda = 1$ represent states where Xe was present and absent, respectively. Nine intermediate, equally spaced values (windows) of λ were selected with a spacing of $\lambda i{+}1$ - $\lambda i = 0.1$. The free energy change with each increment of λ was less than 1 kcal/mol (1.6 kBT), enhancing sampling and overlap between neighboring windows.[79] The alchemical FEP calculation was performed for each window with an equal equilibration time of 1 ns followed by varying sampling time. The correlation time of the quantity was 1 ns or less for each window. A block average statistical analysis[91] was used to determine the length of the sampling time of each window. A blocks length was chosen to be the above correlation time (1 ns). The number of blocks was increased until the standard deviation of block averages was no longer dependent on the number of blocks and the calculated free energy change obtained using n blocks was within 0.01 kcal/mol of that obtained using the previous $n{-}1$ blocks. As λ increases, water molecules may enter the cryptophane interior. The residence time of water molecules at the end point was on average 0.1 ns and is shown in SI Table S2. Thus, the sampling time was selected to be much larger than this residence time of water molecules. As λ increased, longer sampling times were used to obtain converged results. The total sampling time ranged from about 3 ns for the first [0, 0.1] λ window to about 10 ns for the last [0.9, 1] λ window. To avoid the "end-point catastrophes " near $\lambda = 1$ where overlapping particles lead to singularities in the Lennard-Jones potential, a modified soft-core potential[92], [93] was used with a radius-shifting coefficient set at 5. For the $\Delta G2$ pathway, to confine the Xe atom inside the host molecule, an isotropic harmonic potential $u(r)$ (Eq. 7) was exerted using the COLVAR module,[94] and the radial coordinate r of the Xe atom was measured with respect to the center of mass of the aromatic carbon atoms of the host cryptophane molecule.

The statistical uncertainties in the free energies of association (Table 1) were determined using the standard deviation of the block averages for each increment in λ; the trajectory of each such increment was divided into blocks of 1 ns of sampled configurations, and the

	Experimental		Calculated
	K_a[a] ($M\times10^4$)	$\Delta G_{binding}$ (kcal/mol)	$\Delta G_{binding}$ (kcal/mol)
TTEC	3.40±0.10	-6.07±0.02	-5.97±0.09
TAAC	3.33±0.28	-6.06±0.05	-5.89±0.10
TTEC	1.70±0.17	-5.67±0.06	-5.78±0.15
TTEC	0.68±0.23	-5.13±0.24	-4.89±0.06
TTEC	0.22±0.04	-4.48±0.12	-4.55±0.10
TTEC	0.10±0.04	-4.02±0.30	-4.10±0.18

Table 3.1: The calculated and experimental binding free energies of Xe and water-soluble cryptophanes. [a] Previously reported values

variance of the free energy $\sigma 2$ was estimated using the values of the free energies estimated for each λ window: $\sigma^2[\Delta G_{0\to1}] = \sigma^2[\Delta G_{0\to0.1}] + \sigma^2[\Delta G_{0.1\to0.2}] + \ + \sigma^2[\Delta G_{0.9\to1}]$

3.4. Result

Binding free energy calculation results. Molecular dynamics simulations were used to estimate the free energies of binding (Figure 3.2 and Eq. 1) for the six cryptophanes presented in Table 1. To estimate $\Delta G1$, a single decoupling calculation was performed for a Xe atom solvated in water. To obtain the Gibbs free energy change for the NPT ensemble, the work contributed by the volume change between the initial and end states should be estimated. At constant pressure, the volume change of the periodic box for each decoupling pathway is less than 100 Å3, resulting in a negligible contribution due to pressure-volume work *(P* ΔV ¡ 1.510-3 kcal/mol). As a result, we take that $\Delta A1 = \Delta G1$ and $\Delta A2 = \Delta G2$. The calculated free energy of solvation of Xe is $\Delta G1$ = -0.40 ± 0.06 kcal/mol. This value is within 1 kcal/mol of the experimentally inferred value for transfer of Xe from a dilute gas to water,[95, 96], $\Delta G1$ = -1.34 kcal/mol, and the discrepancy may arise from the simple models used here for solvent and for the Xe atom. As $\Delta G1$ is the same for each cryptophane system considered and the major emphasis of this study is the relative affinities of the different cryptophanes, the experimentally inferred value of $\Delta G1$ was used in estimating $\Delta Gbinding$. $\Delta Grestrain$ is associated with

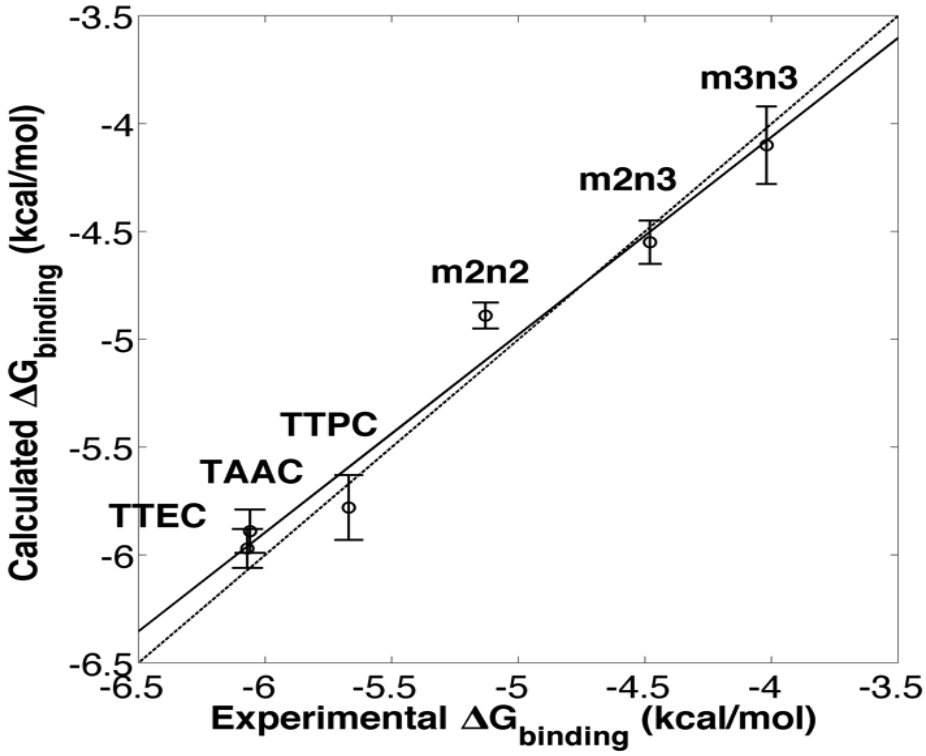


Figure 3.3: Experimental6, 20, 21, 24 and calculated free energies of Xe binding. The solid line is a linear least squares fit. The dotted line corresponds to equality of the experimental and calculated values of ΔG.

the confinement of the ideal (noninteracting or decoupled) Xe atom to the interior of the cryptophane, and for all the systems considered here $\Delta Grestrain$ = 1.72 kcal/mol (Eq. 6). To estimate $\Delta G2$, free energy perturbation (decoupling) calculations were performed for Xe associated with each of the six systems listed in Figure 1: TTEC, TAAC, TTPC, *m2n2*, *m2n3* and *m3n3*. Values of $\Delta G_{binding}$ were found to be invariant upon slight variation of $r_{restrain}$, e.g., for TAAC, $\Delta G_{binding}(r_{restrain} = 2.8\text{Å}) = 6.27 kcal/mol$ and $\Delta G_{binding}(r_{restrain} = 2.0\text{Å}) = 6.26 kcal/mol$. The resulting calculated binding free energies $\Delta Gbinding$ were in excellent agreement with the experimentally inferred affinities, having a correlation coefficient r = 0.98 (Table 1, Figure 3.3). Thus the simulations and free energy perturbation approach applied here is useful for quantitatively reproducing the relative ordering of the free energies of Xe binding to the cryptophane systems considered.

Host structural variation upon Xe binding. As shown in Table 2, the distance between the centers of mass of the two cyclotriguaiacylenes increased with the length of the linker. Due to the flexibility of dihedral angles of the propyl linkers in *m2n3* and *m3n3*, these interior distances as well as the sizes of the cages fluctuated more than in cryptophanes with only ethyl linkers. Thus, for cryptophanes with propyl linkers, the distances were such that the cage was substantially larger and less suited for Xe binding. While *m3n3* and *m2n3* had the largest interior cavities by these distance measures, the remaining cryptophanes had similar, well-defined interior cavity dimensions. Thus, differences in xenon affinities could not be resolved in terms of the cavity size and structural fluctuations alone.

Host-solvent interactions Each cryptophane was observed to accommodate water molecules within the interior cavity. When multiple waters were present within the cavity, these were often observed to hydrogen bond to each other (Figure 3.4). The radial distribution functions of water molecules as functions of the distance r from an oxygen atom and the center of mass of the aromatic carbons of the cryptophane cage are shown in Figure 3.5. The radial distribution function indicated low probability of water molecules over the range r = 3.0-6.5 Å. This was expected since the average distance of carbon atoms in the cyclotriguaiacylene

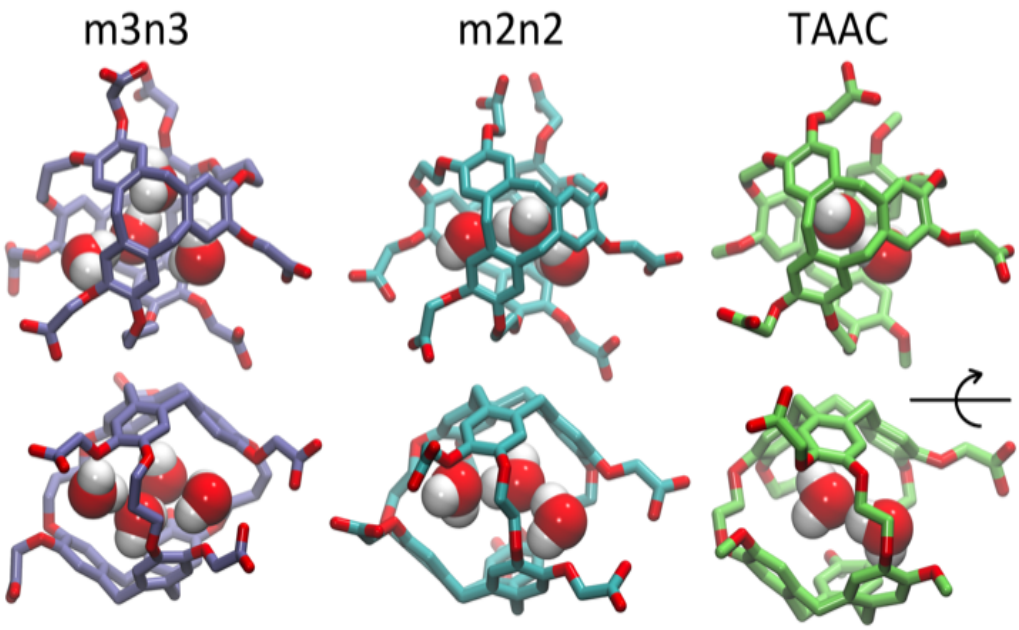


Figure 3.4: Configurations with water molecules inside cryptophanes m3n3, m2n2 and TAAC. In the lower row, each configuration is rotated 90.

units from the center of mass were in the range r = 4.5-5.5 Å; the steric interactions with the cryptophane prevents water from being observed at these values of r. Thus, oxygen atoms having r ¡ 4 Å corresponded to interior water molecules within the cavity. Each of the cryptophanes was free of interior water with Xe present in the cage. Figure 5A indicates no water within 6 Å to the center of the cage. However, when Xe is absent, water molecules occupy the cavity as shown in Figure 5B. The peaks near r = 1 Å arise due to the water molecules near the center of the cryptophane cavity. Representative interior water configrations are shown in Figure 4. The position of the interior peak is slightly displaced for *m3n3* to r = 1.5 Å, which is not unexpected due to the larger interior cavity. A second peak appears near r = 3 Å was also observed for the hexa-acid cryptophanes *m2n2*, *m2n3*, and *m3n3*. This peak corresponds to the water molecules at the 3 pores (Figure 4), where the hydrophilic side-chains and counter ions localize these water molecules. This second peak was not observed for tri-functionalized cryptophanes TTEC, TAAC and TTPC, presumably due in part to the substitution of three carboxylate groups with hydrophobic methyl side chains. The hexa-acid cryptophanes possess greater water occupancies within

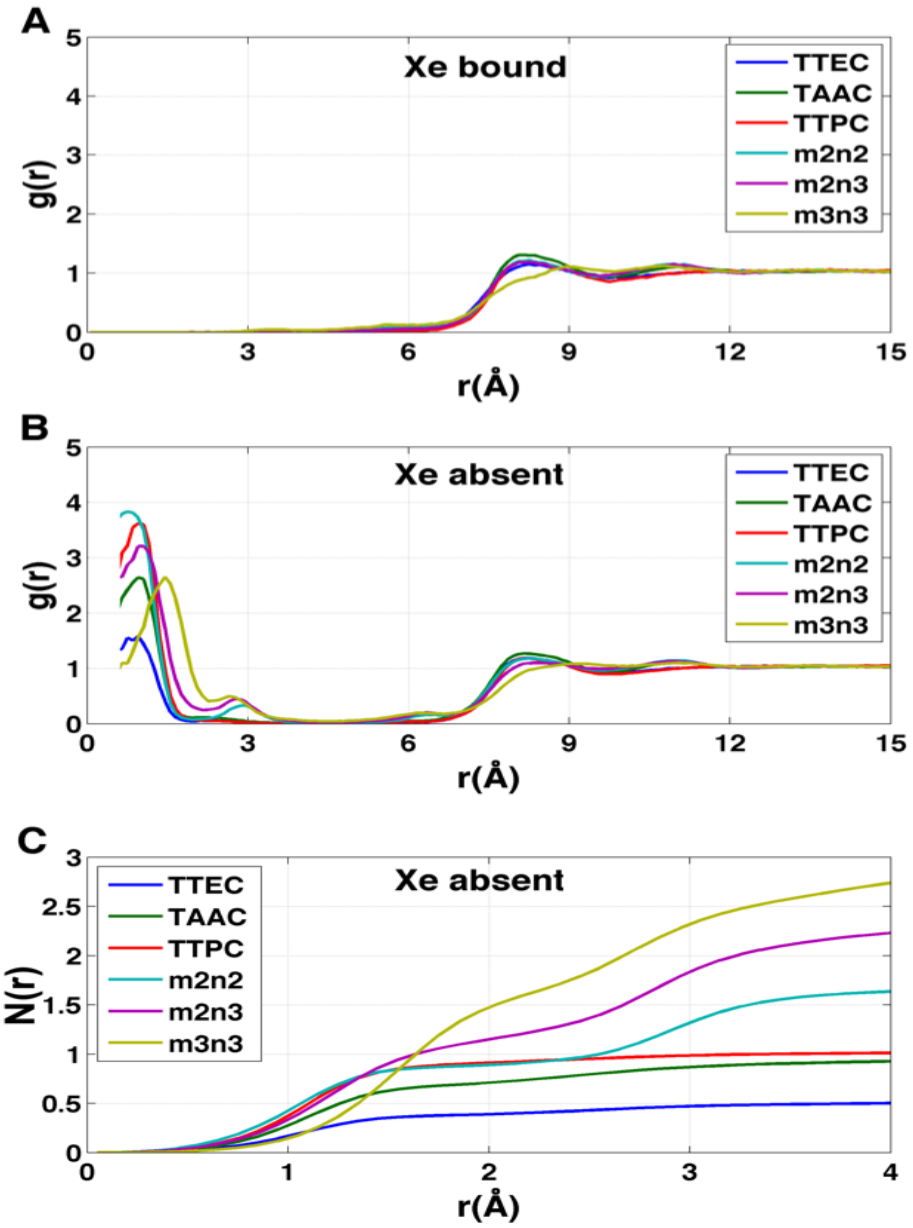


Figure 3.5: Distribution of water molecules in cryptophane interior. The distance r is measured between the cryptophane center of mass and a water oxygen nucleus. A) Radial distribution function g(r) for water when Xe is present in the interior (= 0). B) Radial distribution function g(r) for water when Xe is absent from the interior (= 1). C) Average number of water molecules N(r) within a distance r from the cryptophane center of mass in the absence of Xe. Curves are obtained from integrating the corresponding g(r).

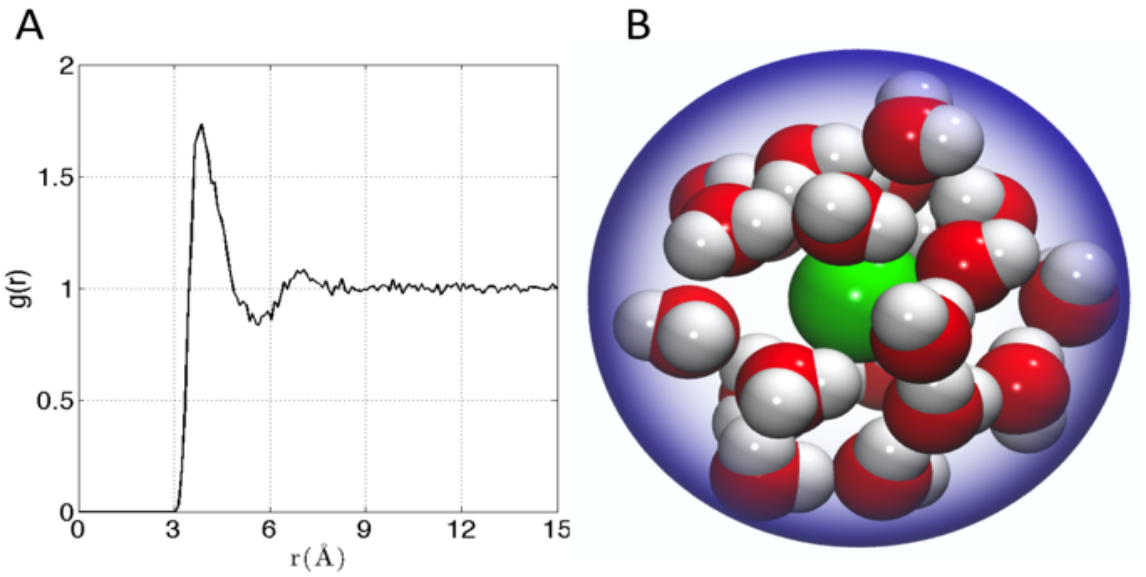


Figure 3.6: A) Water radial distribution function g(r) about Xe in water, where r is Xe-O distance. B) Sampled configuration depicting first solvation shell (waters within 6Å) of Xe in water.

the cavity than the tri-functionalized cryptophanes (Figure 5C). In contrast, the higher Xe affinity cryptophanes are empty (lacking water) for a large fraction of the sampled configurations (SI Figure S2). The average numbers of water molecules inside each cage are 0.50 for TTEC, 0.92 for TAAC, 1.01 for TTPC, 1.63 for *m2n2*, 2.23 for *m2n3*, and 2.73 for *m3n3*, respectively. This trend is also the order of decreasing Xe binding affinity for these cryptophanes.

The affinity of Xe for the cryptophanes is believed to result in part from release of the waters that form an ordered water structure that surrounds the Xe atom in aqueous solution.[]6 Xe in water was simulated to explore the water structure about the rare gas atom. The radial distribution function of water molecules about Xe in water is shown in Figure 3.6A. Integration of this distribution function yields 15-20 water molecules within the first solvation shell (r ¡ 5-6 Å from the Xe atom). Figure 3.6B depicts 20 water molecules surrounding Xe as its first solvation shell within 6 Å. These estimates for the numbers of water molecules released upon sequestration of the Xe in the cryptophane are consistent with those inferred from isothermal titration data, which suggested that 20 water molecules were released upon

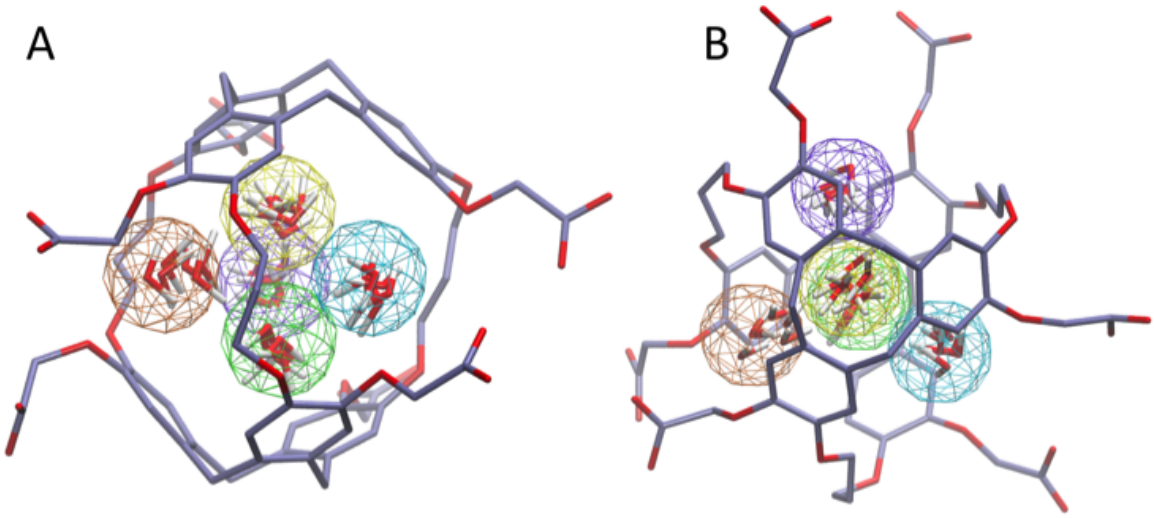


Figure 3.7: Water map clustering for hexa-acid cryptophane m3n3. A) Side view and B) top view differ by a rotation of 90. Five position-based water clusters are indicated by wireframe spheres with radius of 1 and are colored in order of decreasing population: orange, yellow, tan, green and purple. Within each cluster, coordinates of six representative water molecules (red) are rendered.

binding.[97]

Water map analysis. To investigate further the ordering and orientation of water molecules within the cryptophanes, a water map analysis was applied.[56]

In the absence of Xe, coordinates of all the water molecules within 4 Å of the center of the cage were collected from sampled configurations. All such coordinates were pooled, aligned and superimposed. The water molecule with the highest number of neighbors within 1 Å (O-O distance) was selected and taken as the center of a generated cluster. Such a cluster comprises water molecules from distinct sampled configurations in the molecular dynamics trajectory. The resulting water cluster was then excluded before a second cycle of cluster identification was performed. At most five water molecules were observed at the *m3n3* interior (SI, Figure S2), and the five most populated clusters of water molecules inside the cryptophane are shown in Figure 3.7. Within each position-based cluster, clusters of orientations of the water molecules were obtained by using a similar algorithm, where the criterion for association with a particular cluster was $|q_1 q_2| > 0.9$, where q is the quaternion

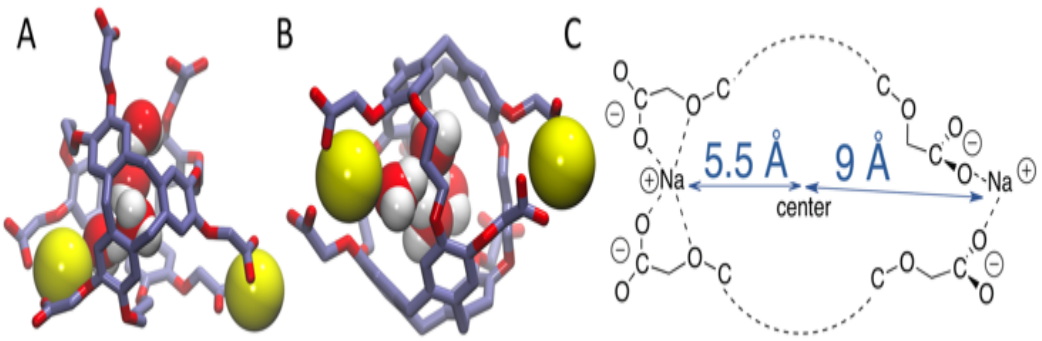


Figure 3.8: Na^+ counter ions (yellow) near the m3n3 cryptophane molecule in the absence of Xe ($\lambda = 1$). A) Top view. B) Side view (90o rotation of configuration in A). C) Schematic illustration of counter ion coordination by the side chains.

that rotates a water molecule to a fiducial reference water orientation. In addition to the two water clusters near the center of the cryptophane, three other water clusters near the pores surrounded by side-chains were identified. These three water clusters were stabilized by hydrogen bonding to hydrophilic side-chains of hexa-acid cryptophanes and coordination of Na+ ions (Figure 3.8). These cluster orientations showed that their oxygen atoms are directed outwardly from the interior of the cage to coordinate Na+ counter ions (Figure 3.8).

The other hexa-acid cryptophanes, *m2n3* and *m2n2*, have smaller interiors but also exhibited water clusters near the pore regions. For the tri-functionalized cryptophanes, TAAC, TTPC and TTEC, water molecule clusters appeared at the cavity center, other water clusters at the three pore regions also existed but with lower populations (Figure S4, Supporting Information); the hydrophobicity of the methyl side chains and the diminished presence of counter ions (Figure 9b) likely reduced the likelihood of persistently structured water near the pores of these cryptophanes.

Distribution of counter ions. The radial distributions of counter ions are shown in Figure 9. When no Xe atom was present in the hexa-acid cryptophanes (*m2n2*, *m2n3* and *m3n3*), peaks in the counter ion (Na+) distribution appeared near $r = 5.5$ Å and $r = 9$ Å. For *m2n2*, *m2n3* and *m3n3*, the peak near $r = 9$ Å corresponded to coordination of the

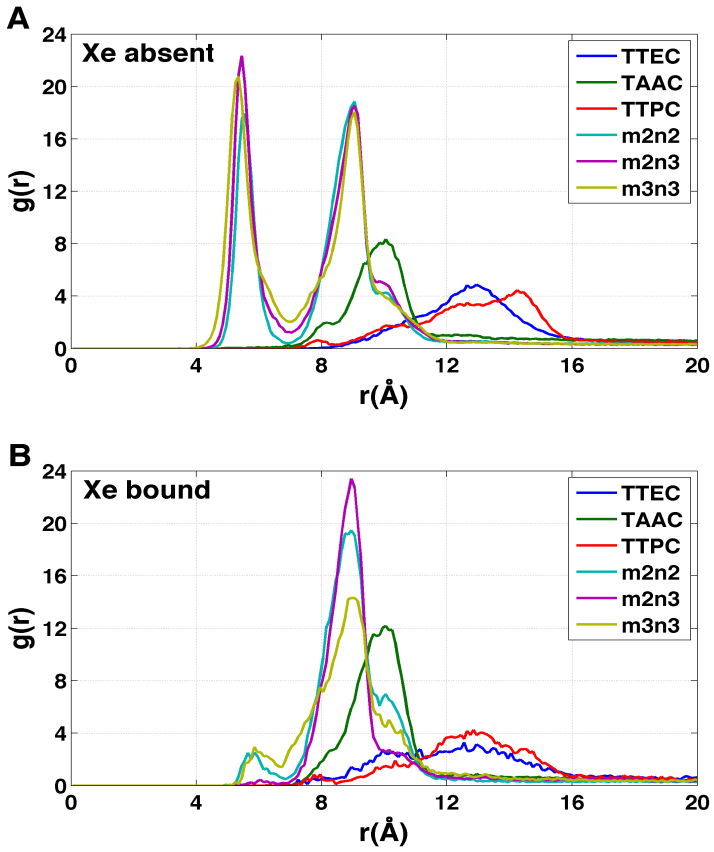


Figure 3.9: The radial distribution g(r) of counter ions relative to the center of mass of the cryptophane molecule. A) Xe atom is absent ($\lambda = 1$). B) Xe atom is present within the cryptophane ($\lambda = 0$).

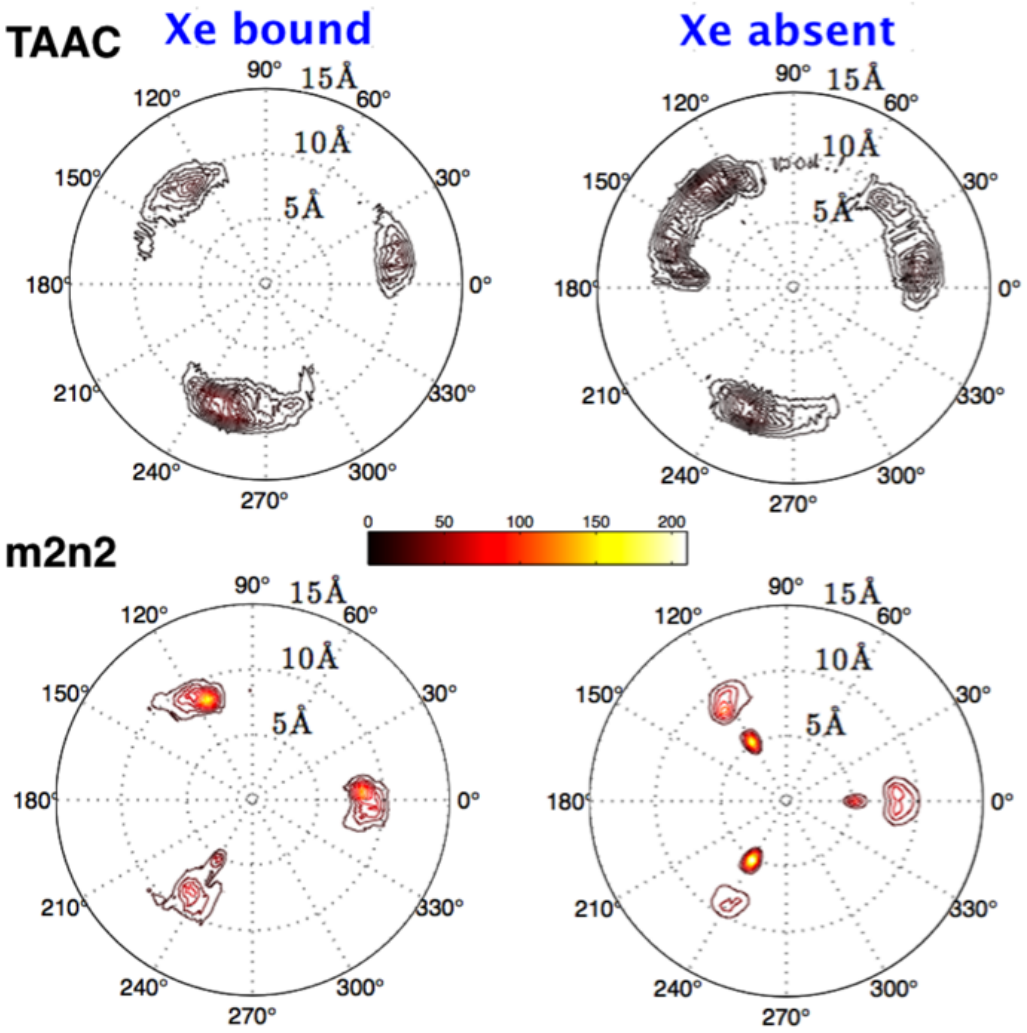


Figure 3.10: Polar contours for distribution of counter ions. The left and right columns are the polar distributions of counter ions when $\lambda = 0$ and $\lambda = 1$, respectively. The origin is set at the center of mass of the 36 aromatic carbon atoms of the cryptophane; the x-y plane is perpendicular to the vector connecting the two centers of mass of the top and bottom caps, and the x axis is the projection of the vector connecting the origin and the average position of two phenol oxygen atoms on the same linker. On the contour plot, the radial coordinate r is the distance between the origin and a counter ion; θ is the angle between the x axis and the projection of r on the x-y plane. Color scale bar is shown above for g(r, θ.

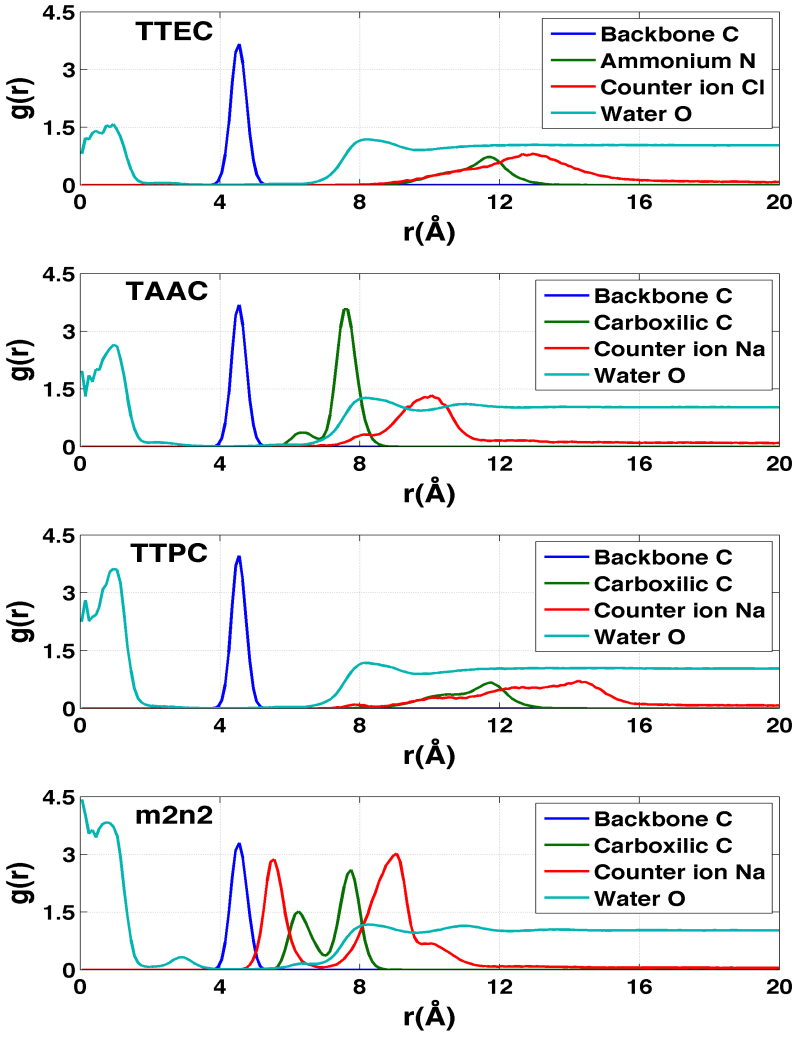


Figure 3.11: Radial distribution functions of atoms associated with simulations of the cryptophanes TTEC, TAAC, TTPC, and m2n2 in the absence of Xe. In each case, r is the distance of the indicated atom type to the center of mass of the cryptophane. Blue (backbone): the aromatic carbon atoms of the cryptophane. Green (ionizable groups in the side chains): for TTEC, the ammonium nitrogen atom N; for TAAC, TTPC and m2n2, the central carbon of the carboxylate group. Red (counter ions): Cl- for TTEC and Na+ for TAAC, TTPC, and m2n2. Cyan (water): the oxygen atoms of water molecules. The blue, green and red distribution functions have been rescaled by dividing by 100, 25 and 6, respectively.

Na+ counter ions with only the carboxylate oxygen atoms at the end of the side chains. The peak at $r = 5.5$ Å corresponded to configurations where Na+ was coordinated by both carboxylate and ether oxygen atoms of the side chains. The peak near $r = 5.5$ Å was not observed for the tri-functionalized cryptophanes TTEC, TAAC and TTPC.

For the hexa-acid cryptophanes, the presence of the Xe atom within the host molecule precluded localization of the ions at $r = 5.5$ Å due to van der Waals overlap, as is evidenced by the corresponding reduction of the peak near $r = 5.5$ Å in the pair distribution (Figure 3.9). On the other hand, for the tri-acid cryptophanes (TAAC, TTPC), the distribution of Na+ counter ions was less sensitive to the presence of Xe. This is presumably due to the reduced negative charge and three fewer coordinating carboxylate atoms in the tri-acid cryptophanes, which reduced the "tight binding" of Na+ counter ions.

The polar distributions of counter ions relative to the axis of three-fold symmetry of the TAAC and *m2n2* cryptophanes are presented in Figure 3.10. The counter ions have much less configurational freedom and have narrower distributions for the hexa-acid cryptophanes, such as *m2n2*. These hosts have six carboxylate side chains, which results in well-defined counter ion distributions about the cage. The peak located at $r = 5.5$ Å was lost upon binding of Xe. In TAAC, the counter ions were more broadly distributed and had much more configurational freedom about the cryptophane, as evidenced by the more diffuse distributions for the tri-acid TAAC compared to the hexa-acid *m2n2* cryptophane (Figure 3.10).

Distribution of solvent and counter ions. In the absence of Xe, the distributions of water molecules and counter ions were correlated (Figure 3.11). The tri-functionalized cryptophanes for the most part had the charged side chain functional groups distal from the center of the cavity. For the hexa-acid cryptophane *m2n2*, the carboxylate side chains had an additional peak near 6.25 Å which was associated with the presence of counter ions at 5.5 Å. This 5.5 Å peak in the counter ion distribution was accompanied by a peak at 3 Å for oxygen atoms of water molecules (Figure 3.11D). These water molecules were near

the pores of cryptophane and are in the first solvation shell for the sodium ion, having a Na^+ - O distance of 2.5 Å.[98] For the tri-functionalized cryptophanes TTEC, TAAC and TTPC, the water peaks at $r = 3$ Å vanished, and counter ions did not appear closer than $r = 7$ Å to the center of the cage. Compared to *m2n2*, TAAC does not anchor Na+. For TAAC (Figure 11B), counter ions are much less probable at 5.5 Å compared with *m2n2*. Because the charged groups of TTEC and TTPC (positive ammonium groups for TTEC and negative carboxylate groups for TTPC) are at the end of their long side-chains, the counter ions (Cl- for TTEC and Na+ for TTPC) are distributed further from the center of the cage ($r = 8 \sim 16$ Å).

Solvation of side chains of cryptophanes. As shown in Figure 3.12A, the long hydrophilic side chains of cryptophanes TTEC and TTPC both contain the triazole ring, but with oppositely charged terminal groups, ammonium and carboxylate respectively. Although the difference in their affinities is typically beyond the resolution of such free energy calculations, guided by the simulations we speculate on the origin of the different observed affinities. Due to the difference in side chains, the average number of interior water molecules when Xe is absent differs for the two cryptophanes: the water occupancy is nearly 1 water molecule for TTPC while an occupancy of only 0.5 water is observed for TTEC (Figure **??**C). The triazole ring helps to introduce and stabilize the water molecules near the cage pores. As shown in Figure 3.12B, a chain of four hydrogen-bonded water molecules is rendered which extends from the interior of the cage through the pore to the triazole side chain. The N(2) atom on the triazole ring is directed inwardly and hydrogen bonds with the fourth water molecule with a H-N2 distance of 2.4 Å. However, this orientation of triazole side chain along with a water molecules chain was less often observed for TTEC than TTPC. The positively charged ammonium group of TTEC can coordinate with the triazole N(2) atom, yielding a N-N distance of 3.4 Å, consistent with hydrogen bonding (Figure 3.12B). In this particular configuration, the nitrogen atoms of the triazole ring were oriented away from the entry channel. This configuration was not observed for TTPC. We plotted the radial distribution of ammonium N and carboxylic C about the triazole N(2)

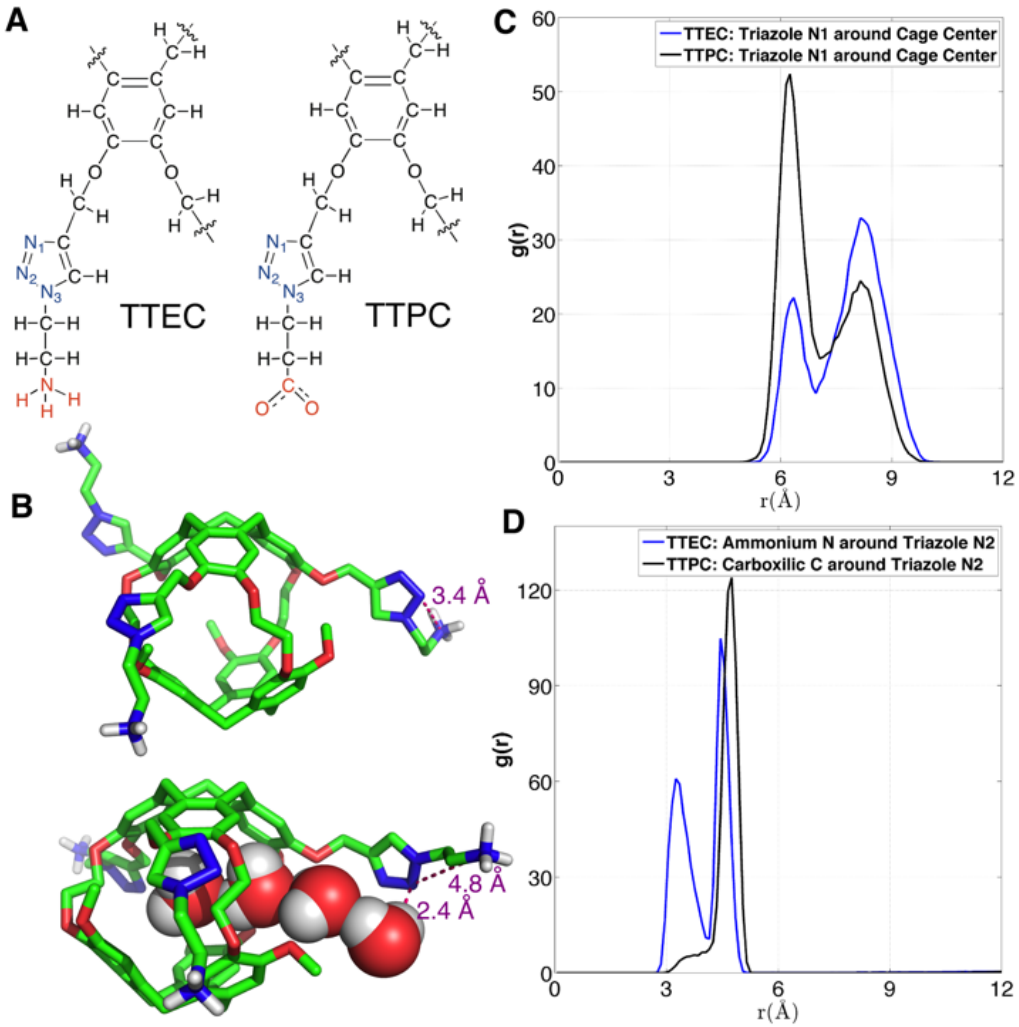


Figure 3.12: Triazole side chains and solvation of TTEC and TTPC. A) Structures of the triazole side chains of TTEC and TTPC. The triazole N atoms are marked blue, and their different terminal groups are marked red. B) Two configurations of triazole side chains in TTEC. C) Radial distribution function of triazole N1 atoms relative to the mass-weighted cryptopane centers in TTEC (blue) and TTPC (black). D) Radial distribution functions of ammonium N atoms (TTEC, blue) and carboxylate C atoms (TTPC, black) relative to the triazole N2 atoms.

for TTEC and TTPC, respectively (Figure 3.12D), and the additional peak at about 3.5 Å was observed for TTEC but not for TTPC. This additional peak of TTEC corresponds to the coordination state between its ammonium N and the triazole N(2). Furthermore, as shown in Figure 3.12C, the N(2) atom (and also other triazole N atoms) are closer to the cage center for TTPC. These observations from the simulations and the distribution functions are consistent with the following picture: 1) the triazole ring aids in stabilizing the water molecule chain that traverses the pores of the cryptophane cage; 2) the particular coordination state between the ammonium N and the triazole N(2) atom in TTEC leads to configurations with outwardly directed N atoms, which are less preferable for stabilizing chains of water molecules that enter the cryptophane pores; 3) the triazole side chain is more extended in TTPC, and its triazole N atoms are available to hydrogen bond to water molecules that form chains or that are poised to enter the cage.

3.5. Discussion

The simulation studies are insightful for understanding the xenon-binding properties of water-soluble cryptophanes. In addition to estimating free energies, the simulations provide a molecular perspective on cryptophane conformational fluctuations and local environment. The distribution of water and counter ions are addressed directly to aid understanding of their influence on Xe binding. The calculated affinities correlate well with the experimentally determined values (Figure 3.3), despite the fact that K_a values of various soluble cryptophanes for Xe only differ only by factors of 2-3 (Table 1).

Upon binding Xe to cryptophane, about 20 water molecules surrounding Xe (Figure 7B) within its first solvation shell are restored to bulk-solvent local environments. In addition, water molecules in the cryptophane cavity (Figure 4) are also displaced. The distributions for the numbers of interior water molecules varied for different cryptophanes (SI, Figure S2). The average numbers of water molecules inside the cage calculated from ∼20 ns equilibrium trajectories were 0.50 for TTEC, 0.92 for TAAC, 1.01 for TTPC, 1.63 for *m2n2*, 2.23 for *m2n3*, and 2.73 for *m3n3*, respectively. This trend was also the order of decreasing Xe

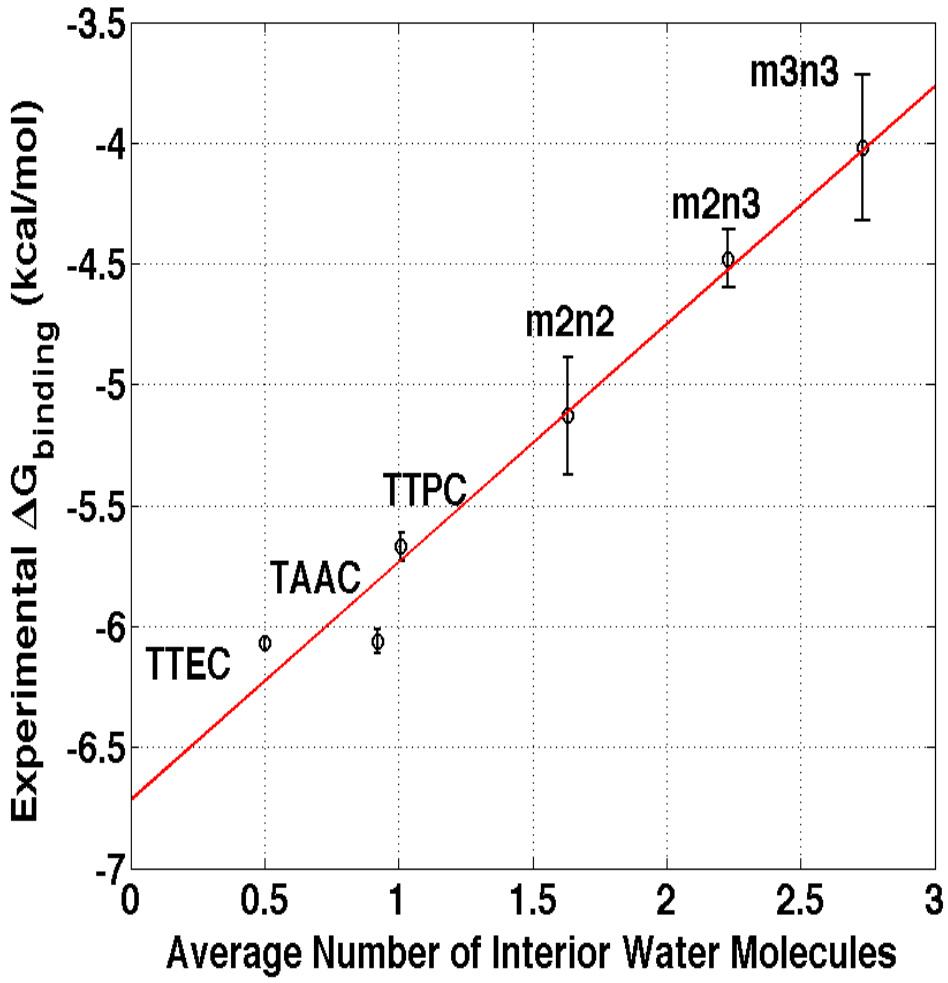


Figure 3.13: Correlation between the average number of interior water molecules and experimental free energies of Xe binding.6, 20, 21, 24 The red solid line is a linear least squares fit. The radius cutoff for identifying interior waters was set at 4Å relative to the center of mass of the 6 aromatic rings of cryptophane molecules.

affinity for these cryptophanes (Figure 3.13). Larger cryptophanes have more waters to displace upon Xe binding. The tri-functionalized cryptophanes are of the same size as *m2n2*, but they contain fewer water molecules on average due to their hydrophobic side-chains and lack of ability to anchor the counter ions in well-defined regions close to the cavities. Previous molecular dynamics simulation studies suggest that the number of water molecules in the cavity of *m3n3* (a derivative of cryptophane-E) ranged from 0 to 5, with an average number of 2.1,[74] which is consistent with the simulation results presented here. Experimental studies of such interior water could be insightful. Bound water molecules have not yet been observed using solution 1H NMR spectroscopy at room temperature. Recently, the first X-ray structure of a cryptophane-water complex was obtained, and the cage contained essentially one interior water molecule but with partial electron density occupying seven distinct cryptophane locations.[99]

Water is commonly ignored or not treated explicitly when designing aqueous host-guest systems. Water molecules in many binding sites must necessarily be displaced to accommodate the binding of a guest species.[72] The work required for such displacement is a component of the free energy of binding, and this component is expected to increase (lower binding affinity) with increasing numbers of water molecules to be displaced. For hydrophobic host cavities, this process is often accompanied by an entropy gain due to the release of the ordered water molecules sequestered inside a confined host environment into the bulk solvent. Accompanying this binding process is the formation of host-guest and water-water energetic (enthalpic) interactions. Similarly, for Xe to bind to a cryptophane, any water molecules in the cage must be displaced. For cryptophanes with ionic side-chains, anchored counter ions and some of their coordinating water molecules that are interior to the cryptophane must also be displaced upon binding. The results of this study highlight the importance of counter ions, and other studies have found that Xe-cryptophane binding is sensitive to the type of counter ion.[100] In addition to water within the cryptophane interior, water that partially occludes the Xe binding site is also expected to decrease Xe affinity; such partially ordered water can be part of the first solvation sphere of an anchored

counter ion or hydrogen bonded to a polar or ionizable moieties associated with the cryptophane. Modifications to the cryptophane that position polar functional groups and counter ion coordination sites further from the interior may reduce the number of ordered interior water molecules, increasing Xe affinity.

The simulation results indicate that the lowest occupancy of water molecules is associated with TTEC and that higher occupancies are observed for the hexa-acid cryptophanes. Among the molecules considered here, there is a linear correlation between free energy of association and the average number of interior water molecules (Figure 3.13). Although the calculations presented here use a simple model for water and do not include explicit polarizability of host and guest, the notion of increasing affinity with decreasing interior solvent occupation is in harmony with recent studies of cryptophane-E and its binding of tetrachloroethane in a variety of halogenated organic solvents.[21]

The findings herein suggest that molecular simulations and free energy calculations could be potentially useful in the design of cryptophane systems. For a set of cryptophane molecules under consideration, free energy calculations could be performed to vet the set and identify those having targeted affinities for Xe (and other guest molecules) prior to synthesis. The data in Figure 3.13 suggests that the average number of water molecules within the interior may be a strong indicator of Xe affinity, and the estimation of such water occupation of the cryptophane interior is a much less costly calculation than the explicit calculation of free energies of binding.

3.6. Conclusion

This study informs our understanding of the hierarchy of cryptophane binding affinities. The hexa-acid cryptophanes, *m2n2*, *m2n3*, *m3n3* have the same side chains but different sizes due to different lengths of linkers, and affinity decreases with increasing cage size. However, the cryptophanes TTEC, TAAC, TTPC and *m2n2* have identical cage cores, therefore their different affinities with Xe are a result of the differences in their side chains

and the host-solvent, host-counter ion and solvent-counter ion interactions. Even though the cryptophane interior is hydrophobic, it accommodates water molecules and small water clusters. The binding with Xe thus requires interior water molecules to be displaced. The simulation results show that the occupancy numbers of water molecules inside the cryptophane are correlated with the binding free energies (Figure 3.13). Furthermore, hexa-acid cryptophanes can chelate counter ions with two side chains. This "anchoring " is not observed in the cages with only three solubilizing moieties. The counter ions trapped by side chains of hexa-acid cryptophanes are displaced in order for a Xe atom to bind within the cryptophane molecule. Associated with these trapped counter ions are solvating water molecules inside the cryptophane cage, including near the pores. This results in the higher water occupancy observed for *m2n2* (1.6), compared to those of TTEC (0.6), TAAC (0.9), and TTPC (1.0). The results presented here exemplify how simulations can potentially be used to understand and to guide the design cryptophane molecules to possess both high aqueous solubility and high affinity for Xe.

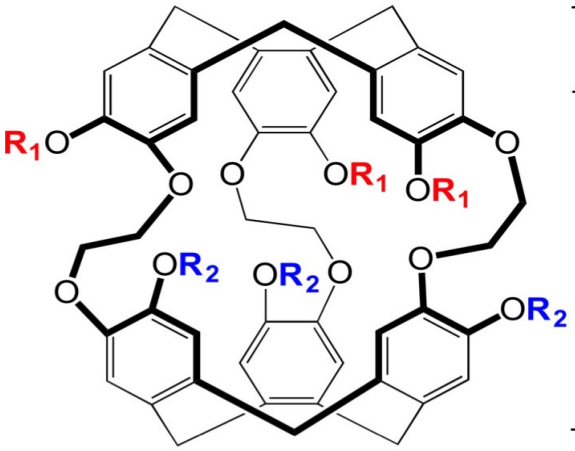
CHAPTER 4 : Free Energy of Xe Entry into a Cryptophane Interior

4.1. Abstract

Xe bound inside a cryptophane system is one kind of novel complex molecule that can work as biosensor. Among water-soluble cryptophanes, four cryptophanes with different side chains are computationally investigated in this work. The potential of mean force is calculated for the displacement of one Xe atom from cryptophane cage center using an adaptive force biasing method. The applied forces and integrated free energy profiles are calculated as functions of the distance Xe away from cryptophane center. All calculated potentials of mean force have similar shapes but different free energy minima, which leads to different Xe binding affinity. Additionally, the structures of the cryptophanes in process of Xe exit are analyzed. It is found that side chains will have to change dihedral angles so as to allow space for Xe to pass through. In addition, two linkers closest to Xe are also disturbed and altered, which leads to slightly increase of cryptophane interior volume. The energetic and conformational knowledge learned from this work will shed light on designing Xe-cryptophane systems.

4.2. Introduction

Host-guest interactions involving relatively complicated host molecules and guests with relatively simple structure are important in chemistry. Due to high affinities between host and guest involving non-bonded interactions, guest particles can be captured to bind with host molecules from solutions. During the encapsulation process, host molecules often need to adjust their conformations to ease this reaction, e.g. induced fit.[46–48] Cryptophane-Xe belongs to this type of interaction. To quantitatively understand the high binding affinity of cryptophanes to bind Xe and analyze the conformational adjustments of the cryptophane cavity to encapsulate Xe, adaptive biasing force (ABF)[42–45] simulation is implemented. ABF is a widely used method for sampling and calculation to investigate host/guest interactions and conformational changes along one or multiple reaction coordinate.



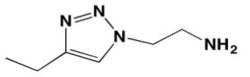
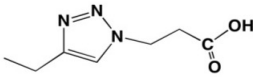
	R_1	R_2
TTEC	CH_3	
TTPC	CH_3	
TAAC	CH_3	CH_2COOH
m2n2	CH_2COOH	CH_2COOH

Figure 4.1: Structure representations of four cryptophanes: TTEC, TTPC, TAAC and m2n2.

Though the potential of mean force (PMF) with respect to specific coordinate can be calculated from radial distribution function (equ (4.1)) which can be collected from equilibrium simulation.

$$A(r) = -RTlng(r) \tag{4.1}$$

It is hard to obtain accurate PMF at high free-energy points, since those high energy samples are less likely to be visited during a normal simulation. The radial distribution function approaches 0 and makes estimation of PMF meaningless at these points. In order to sample configurations adequately at high energy states, many techniques have been created to efficiently solve this problem: one type is to manipulate the potential during the simulation, like umbrella sampling[101–103] and another is to provide external force to the system, like adaptive biasing force (ABF) method[42–45]. By applying the external force to bias the simulation, we can sample efficiently at all states we are interested in. Herein we will obtain the free energy profile of pulling Xe out of cavity. These simulations will not only reveal the molecular details but also provide us means to manipulate this host-guest binding reaction.

Cryptophane molecules are cage-like synthetic molecules with two cyclotriguaiacelene caps linked by three variable length alkyl chains to form nearly spherical cavities and can have

variable groups connected with its phenol oxygen as sidechain (Figure4.1). Due to its special and variable chemical structure, many small molecules and atoms are reported that can bind to its interior in organic solutions or water; examples include methane[104], Xe[17], and ionic species[23–25]. Among host-guest systems, Xe129-cryptophane system is an important applications[22, 26]. The 129Xe isotope can be hyperpolarized with regard to its spin states and used as a magnetic resonance imaging (MRI) contrast agent. The resonance of this nucleus is very sensitive to the environment and therefore shows a very broad range of chemical shift when bound to different molecular environments. Therefore, the Xe system is a good candidate for chemical sensor. But Xe alone is difficult to be incorporated into molecualr systems since it is inert, hard to dissolve in water, and has weak interactions with most molecules. Water soluble cryptophanes that have high Xe binding affinity provide a means to deliver and localize Xe-129 to function as chemical sensor, especially in biological systems[22, 26, 32].

Although experimental work has obtained a crystal structure of one Xe@cryptophane complex[105] and previous computational simulation work have revealed that water can occupy cryptophane interior[106, 107] and will be replaced by Xe[107], an atomistic picture of the conformation change of cryptophane molecules when Xe is encapsulated from solution into the cavity is still missing. This is quite interesting especially when we focus on four cryptophanes that have high binding affinities with Xe but have different sidechain structures (figure4.1B): m2n2, TTPC, TAAC and TTEC. Experiments have measured Xe binding affinities of those four cryptophanes[33, 34, 36, 69] which range from -5.13 kcal/mol to -6.07 kcal/mol. Herein, by taking use of adaptive biasing force simulation, we calculated free energy profile of Xe-cryptophane systems with respect of distance between center mass of Xe and center mass of cavity, and translated potential of mean force to binding affinity which showed to be close to experimental measurements. In addition to the energetic analysis, we also implemented conformational analyses on all cryptophane systems so as to discover the structural adjustment acquired as Xe enters the cryptophane cavities.

4.3. Methods

4.3.1. Molecular models.

Four cryptophane molecules (Figure 1: m2n2, TAAC, TTPC and TTEC) are simulated in this work. Structures of four molecules are generated by modifying the crystal structure of m2n2 in molfacture module in VMD**??** and optimized in NAMD**??**. In each simulated systems, 3353 TIP3P**??** water molecules are added which create 48Å ×48Å ×48Å water boxes with concentration of solute is ∼0.02M and 6, 3, 3 sodium ions and 3 chloride ions are added respectively to neutralize the charges. Parameters of bonded interactions used in simulations come from AMBER9428 force field. Additional required boned parameters come from Hessian matrices calculation based on optimized geometries from Gaussian9829. For non-bonded interaction, van der Waals interactions are modeled by Lennard-Jones potential and parameters of cryptophane molecules and Xe come from OPLS force field**??** while electrostatic interactions are modeled by coloumbic interaction with partial charges in cryptophane backbone come from developed charge set20 and the charges in sidechains come from R.E.D. server**??** where charges are fitted to reproduce quantum calculated electrostatic potential around each atom and atoms at junction position are edited so as to construct integer charge for overall molecule ultimately.

4.3.2. Molecular dynamics simulation.

NAMD[108] package is applied in this work to perform simulations. In prior to molecular dynamics simulations, 50000 steps of conjugate gradient energy minimization are performed to eliminate bad initial configurations after we solvate the cryptophane molecules. In this work, step size in simulation is 2 femtosecond. Then 20ns length of NPT ensemble simulations are done for each system where we use Langevin dynamics method [54] to maintain the system temperature to be 300 K and pressure to be 1.0atm. The coupling coefficient for temperature constrain is 1/ps and the period is 100.0 fs, decay constant is 50.0 fs and corresponded temperature is 300.0K for Langevin piston using Nose-Hoover method[55].

Electrostatic interactions are calculated through Particle Mesh Ewald (PME) method [53] with 1.0Å grid spacing. Lennard-Jones interactions are computed with cutoff is 12.0Å. SHAKE/RATTLE[52] algorithms are applied to constrain the covalent bonds linked with hydrogen to save computation. The last frames of trajectories generated from NPT simulations will be used for analysis and provide initial conformations to later ABF simulations.

$$\frac{dA(\xi)}{d\xi} = \left\langle \frac{\partial V(x)}{\partial \xi} - \frac{1}{\beta}\frac{\partial ln|J|}{\partial \xi} \right\rangle_{\xi^*} = -\langle F_\xi \rangle_{\xi^*} \tag{4.2}$$

4.3.3. Adaptive biasing force simulation

.In this work, ABF method is used to calculate the free energy profile of the cryptophane-Xe system with respect to the distance (ξ) between center mass of cryptophane backbone and center mass of Xe and generate trajectories that having ξ range from 0 to 15.0Å. In ABF method, one accumulated biasing force along the direction of ξ is applied to the system so as to cancel out net force along this direction thereby the system will do the diffusion and having uniform distribution with respect to ξ. The formula of estimating biased force from integration can be seen in equ(6.1), where A(ξ) is the free energy of the system at ξ, V(x) is the potential energy in original Cartesian coordinates, β is the temperature factor, |J | is the Jacobian determinant of transforming Cartesian coordinates to generalized coordinates which includes ξ. The bracket represent the ensemble average that having ξ constrained to ξ^*. In implementation, the range of ξ we are interested in is divided into small bins and the forces in each bin will be considered to be same, and in prior to apply the biasing force a threshold number of samples in each bin are required to be reached due to force fluctuate significantly initially otherwise will not representative for that ξ state and bias the system to wrong sampling. For all four cryptophane systems, overall range of ξ (0.0 ∼15.0Å) are divided into three windows (0.0 ∼5.0Å, 5.0 ∼10.0Å and 10.0 ∼15.0Å) and every one of them represent one separate ABF simulation and having ξ constrained to corresponding ξ distance range so as to assure ABF calculation converge sooner. The bin size is chosen to

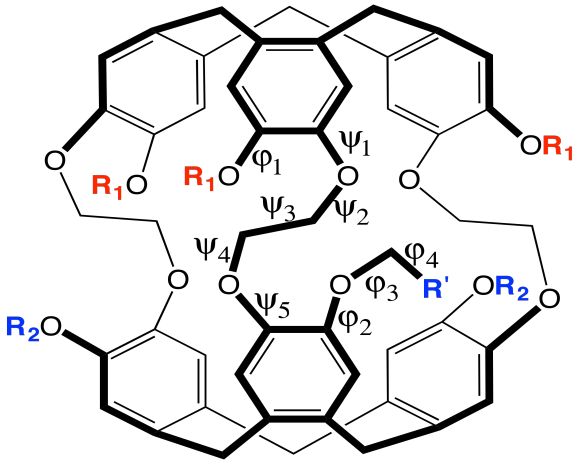


Figure 4.2: Representation of analyzed dihedrals.

be 0.05Å and the threshold in each bin is chosen to be 30000. In each ABF simulation run, 20 ∼40ns simulation have been done so as to assure the uniform sampling of ξ and thereby provide the accuracy of calculated biasing force and PMF.

4.3.4. Structure analysis.

Using the trajectories having ξ range from 0.0Å to 15.0Å, we analyze and summarize the changes of cryptophane structures when Xe is passing through by calculating the distributions of dihedrals in cryptophane among simulation samples. For cryptophanes TAAC, TTEC and TTPC, they are top-bot asymmetric with respect to the "top" and "bottom" of the cryptophane, since they have two sets of different sidechains. In each case, five dihedrals in one linker are different (figure 4.2A) and are analyzed separately. Additionally, the torsion of the methyl group and the first three torsions in each charged sidechain connected to the aromatic rings are also analyzed. For every dihedral, three sets of samples are used to calculate the corresponding distributions: equilibrium trajectory of cryptophane with no Xe inside cage, equilibrium trajectory of cryptophane with Xe occupied inside cage, and ABF trajectory with Xe at the energy barrier state. When Xe is at the energy barrier state, for dihedrals in sidechains, we consider the dihedral in sidechain that is closest to the Xe. For dihedrals in linkers, since Xe sits between two linkers, we only consider dihedrals in these

two linkers and differentiate them as "left " and "right ": when we view the cryptophane from top bowl that has methyl groups as sidechains to the bottom bowl that have ionic sidechain groups, the linker in the left side of Xe is defined to be "left " while the linker in the right side of Xe is defined to be "right " (figure 4.2B).

4.4. Results and Discussion

4.4.1. Potential of Mean Force results.

The free energy profile (potential of mean force, PMF) of cryptophane and Xe systems were calculated with respect to ξ, the distance between center mass of cryptophane and Xe (figure 4.3). Its clear that PMFs of the four cryptophanes have similar shapes: there are three significant states along the reaction coordinate (ξ), a deep global minima with Xe bounded at the center of cavity (ξ ~0.1Å), one energy maxima with Xe at the edge of cavity (ξ ~5.0Å) and a plateau with Xe is unbounded and isolated from cavity (ξ ~15.0Å). In each case, the PMF is calculated with respect to the state that with Xe unbounded. From table 1, it shows that the depths of the global minima (ΔW_{min}) which equal to the free energy difference between the Xe bound and unbound states of all four cryptophanes are TTEC (-11.47 kcal/mol) > TAAC (-11.45 kcal/mol) > TTPC (-11.25 kcal/mol) > m2n2 (-10.47 kcal/mol), their ordering matches that of the experimental measured binding affinities. The height of the energy barrier (ΔW_{max}) of all four cavities showed to between 2 and 3 kcal/mol. The m2n2 cryptophane, whose sidechains are all polar and hydrophilic, is the lowest while the rest three cryptophanes have relatively higher energy barrier. Additionally, there are also one small shoulder (ξ ~7.0Å) and one local minima (ξ ~8.3Å) outside cavity along the reaction coordinate (ξ). These suggest that Xe would still associate with cryptophane though Xe exited the cavity. For these four cryptophanes, the heights of their shoulders are different: m2n2 > TAAC > TTPC > TTEC, the more charged sidechain the Xe more difficult to get closer to cavity interior. When we translated the PMF data to Xe concentration, we find Xe would significantly accumulate at the local minima position and the concentration of it is 2 ~4 times as in bulk (figure 4.4). And the more complex

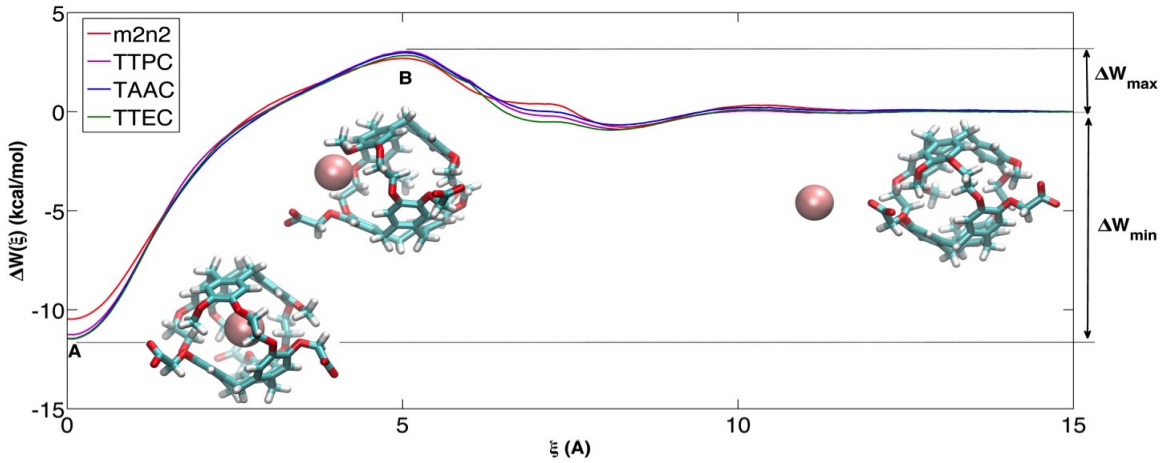


Figure 4.3: Calculated free energy profile.

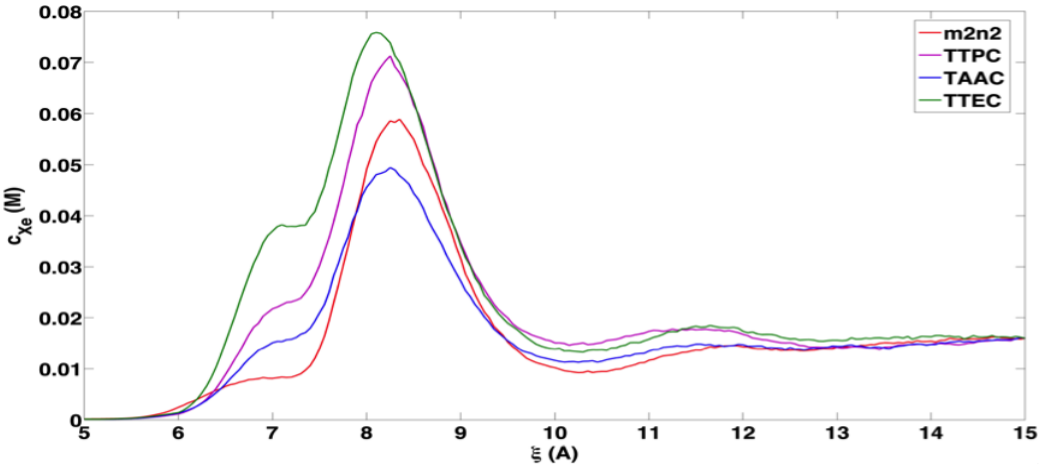


Figure 4.4: Concentration of Xe with respect to the reaction coordinate ξ

Cryptophane	ΔG_{exp} (kcal/mol)	ΔG_{cal} (kcal/mol)	ΔW_{min} (kcal/mol)	ΔW_{max} (kcal/mol)
TTEC[b]	-6.07±0.02	-6.57	-11.47	2.84
TAAC	-6.06±0.05	-6.52	-11.45	2.98
TTPC	-5.67±0.06	-6.37	-11.25	3.04
m2n2	-5.13±0.24	-5.69	-10.47	2.69

Table 4.1: The experimental and calculated binding affinities between Xe and cryptophanes, calculated PMF minima (ΔW_{min}) and calculated PMF maxima (ΔW_{max}).

the sidechains, the higher accumulated concentration of Xe: TTEC > TTPC > m2n2 > TAAC.

$$K_{association} = 4\pi \int_0^{r_d} r^2 e^{-\beta U(r)} dr \tag{4.3}$$

$$K_0 = \frac{1}{1660\text{Å}^3} K_{association} \tag{4.4}$$

$$\Delta G_{bind} = -RTlnK_0 \tag{4.5}$$

4.4.2. Binding affinity calculation.

The potential of mean force of Xe-cryptophane system can be used to estimate the equilibrium association constant by using equation 6.2. The standard state equilibrium constant and binding affinity can be calculated then from equations and equation 6.4. From Table 4.1, the binding affinities from calculation correlate well with experimental data, calculated values are all within 1 kcal/mol of the experimentally determined values. The trend with regard to the order of the binding affinities matches that of the experiments, though the calculations overestimate the affinities by 0.5 ∼0.7 kcal/mol (figure4.5). The deviation of computationally determined results compared with experimental measurements arise in part from the fact that parameters for Xe used in this work come from OPLS force field

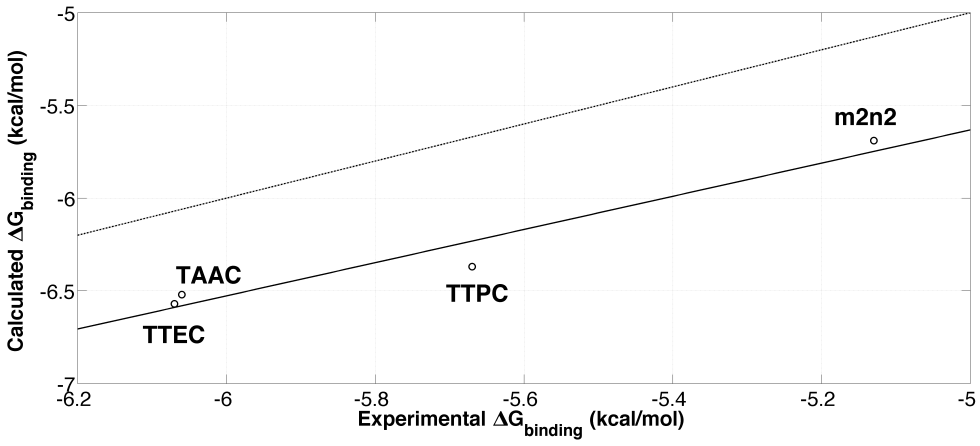


Figure 4.5: Experimental and calculated binding affinities.

which was derived from viscosity experiments directly. The solvation energy of Xe in water with this parameter set is calculated to be 1.75 kcal/mol, while experimentally it is measured to be 1.34 kcal/mol[95, 109]. Additionally, the approximations in equation 6.2 also contribute to the deviation: the equilibrium constant of cryptophane-Xe binding reaction comes from the integral of the exponential of the potential of mean force with respect of r, and at each value of r within the continous range of interest, we assume the space for Xe to occupy is ideal and use the $4\pi r^2 dr$ in integral, however, there are some space Xe are certainly unable to access. Furthermore, in derivition of euqation 6.2, the interaction of single Xe with environment is neglected and therefore not appeared in formula of estimating equilirbium constant of this reaction.

4.4.3. Trajectory analysis.

To reveal the molecular pictures of how cryptophane would change its structures when Xe exit, the interior space, linker and sidechain conformations in samples from equilibrium and ABF simulation are investigated.

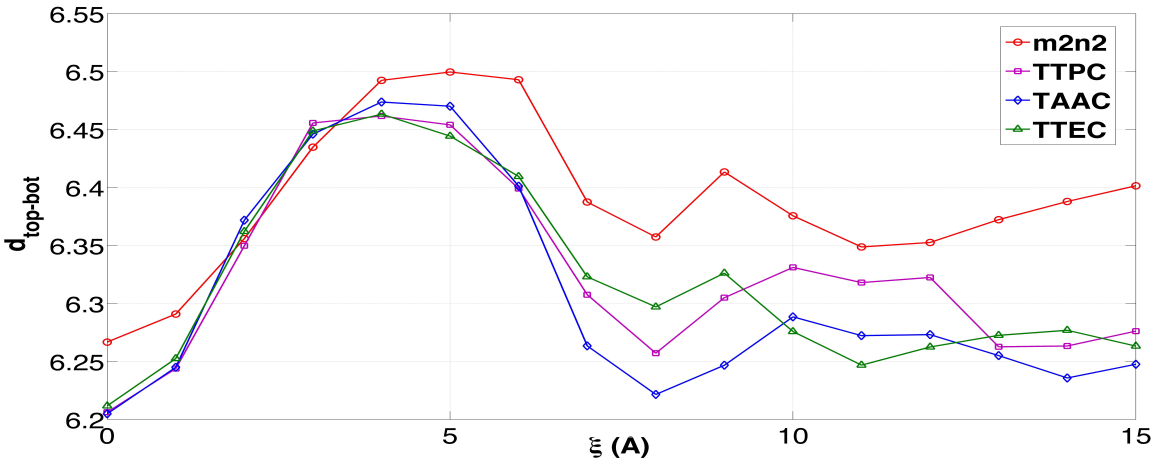


Figure 4.6: Top-bot distance of all four cryptophanes with Xe at different positions.

Top-bottom distance

Cryptophane molecules are approximately spherical, and they can be viewed as having top and bottom cyclotriguaiacylene caps that are linked with alkyl chains. The distance d(top-bot) which measures the distance between top benzene bowl and bottom bowl can be a good indicator of the approximate dimension of the interior cavity. The d(top-bot) data (figure 4.6) for all four cryptophanes showed similar trends: as ξ increases which correspond to Xe exiting the cavity, the d(top-bot) increases $\sim$0.3Å to a maximum value and after ξ reaches $\sim$5Å and continue increases till $\sim$8Å, d(top-bot) decreases $\sim$0.2Å. This trend can be explained as that when Xe exits the cryptophane, the cage would have to open its portal so as to spare space for Xe to pass through, where d(top-bot) increases. After Xe has exited, the cage would close and relax back to its normal interior dimensions. For all four cryptophanes, d(top-bot) is smallest when Xe bounded inside. This is due to the strong non-bonded interaction between Xe and aromatic atoms in cryptophane which makes the top and bottom bowl of cryptophane linked tightest. Since m2n2 has relatively lower Xe binding affinity, its d(top-bot) is largest among four cryptophanes. The d(top-bot) of Xe unbounded state also has small value but due to water molecules would enter the cavity and fill the cavity, the d(top-bot) for Xe unbounded state is slightly larger than the Xe

bounded state. Since there are several water molecules occupied inside m2n2 in average while other three cryptophanes only have less than one water molecule occupied inside, d(top-bot) of m2n2s unbounded state is also the largest among four cryptophanes. As ξ continues to increase to values larger than 8.0Å, Xe leaves its local minima state and water thereafter is more probable to enter the cavity and fill the cavity interior which is reflected in the increases of d(top-bot). As ξ continues to increases, d(top-bot) fluctuates with a much smaller magnitude.

Dihedral distribution analysis for linkers and sidechains

Previous top-bot distance calculation found that cryptophane would adjust its interior space so as to allow Xe to pass through, since the top-bot distance is related with the linkers who connect top and bottom aromatic bowls, we would assume the alkali linkers would be more stretched and be affected in prospects of torsions. Furthermore, its also desirable to investigate if the torsions of sidechain groups and how they are affected when Xe passes through a nearby pore.

4.4.4. Dihedrals in linkers.

From the results of TAAC (figure 4.7), we can see the distribution of all dihedrals in linkers are showed to be the same for conditions of both the Xe unbound state and Xe bound state. This is due to the few water molecules can enter the cavity and form a cluster so as to fill the cryptophane as Xe does. But in conditions of Xe present at the energy barrier state, we can see the distributions in the two linkers closest to Xe changed dramatically. For Xe bounded state and Xe unbounded state, all dihedrals except Ψ_3, can adopt multiple values, and the cryptophane linkers thereby have multipleconformations (figure 4.8A). The dihedral Ψ_3 has a trans conformation so as to avoid the repulsion between two phenol oxygens in the same linker. When Xe occupy one of the pore position, due to its steric effect, both the "left " and "right " linkers are all stretched so as to spare sufficient space for Xe to occupy. Since the "right " Ψ_1 and "left " Ψ_5 are most close to Xe, they all have to choose the 60 °

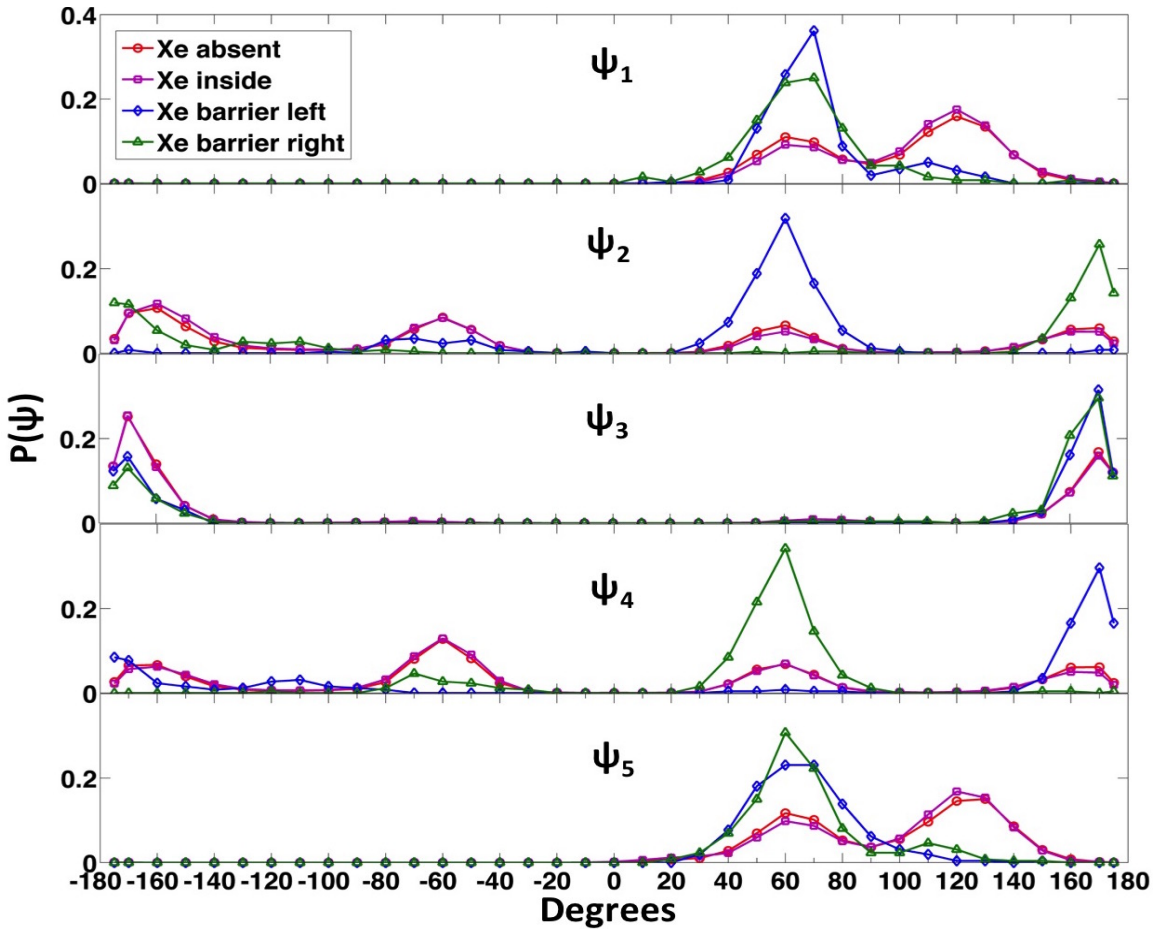


Figure 4.7: Distributions of dihedrals in TAACs linkers from equilibrium simulation and ABF simulation

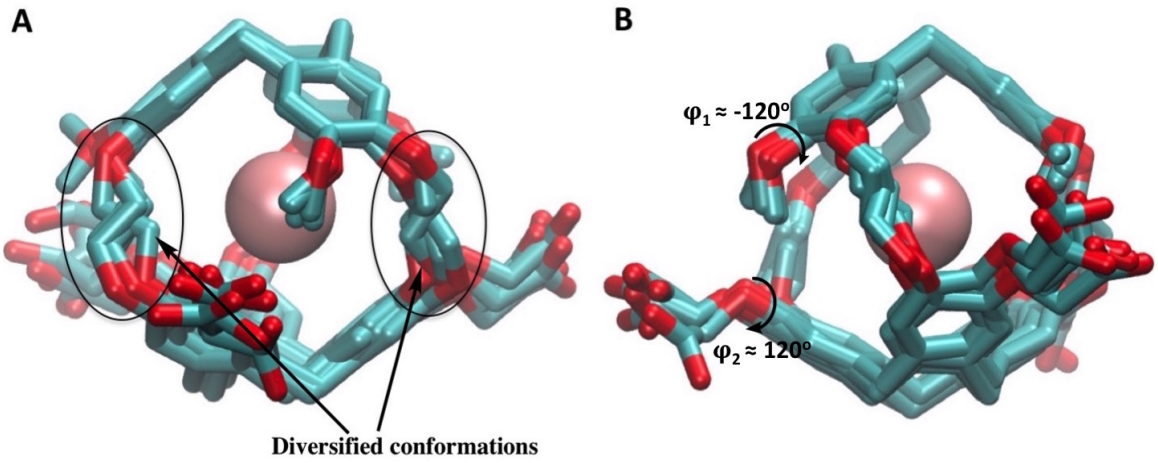


Figure 4.8: Representative structures of TAAC with Xe bounded inside cavity.

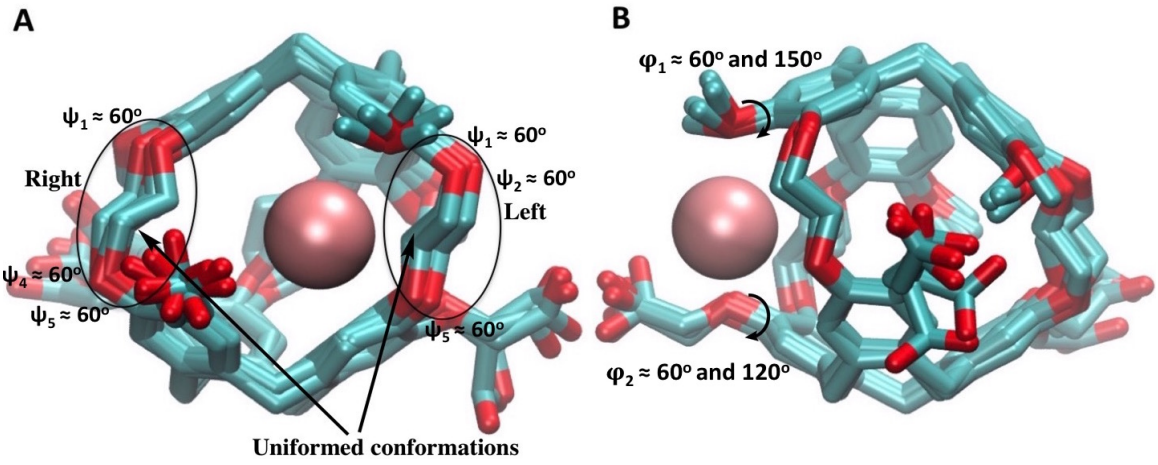


Figure 4.9: Representative structures of TAAC with Xe at the pore of cavity.

conformations. And because of the correlation of the and in the same linker, they all have to choose close values, this leads "left " Ψ_1 and "right " Ψ_5 would adopt 60 ° conformations as well. Similarly as steric effect and felt, the "right " Ψ_2 and "left " Ψ_4 would adopt the trans conformations with peak at ∼170 ° while the "left " Ψ_2 and "right " Ψ_4 would adopt the 60 ° conformations. As a consequence, we observed cryptophane linkers having their linkers adopted uniform conformations (figure4.9A).

Dihedrals in sidechains.

From result of the TAAC simulation (figure 4.10), similar to the analysis of dihedrals in linkers, we can see the distribution of dihedrals in sidechains is also the same for both conditions of Xe unbound state and Xe bound state. For dihedrals in the sidechains when Xe is present at pore position, however, the distribution of them changes dramatically as well. For Xe bound and unbound states, the methyl group is more hydrophobic than the acetic acid, therefore, it is much more preferable to adopt the orientation of pointing toward the interior of the pore where is un-accessible to water than pointing to the exterior of the pore where would immerse in the water (figure 7B), resulting in the major peak at -120 ° in dihedral distribution plot of φ_1 (figure 9). Since the pore space is limited and the methyl occupied the pore, the first torsion connect the acetic group would have to adopt

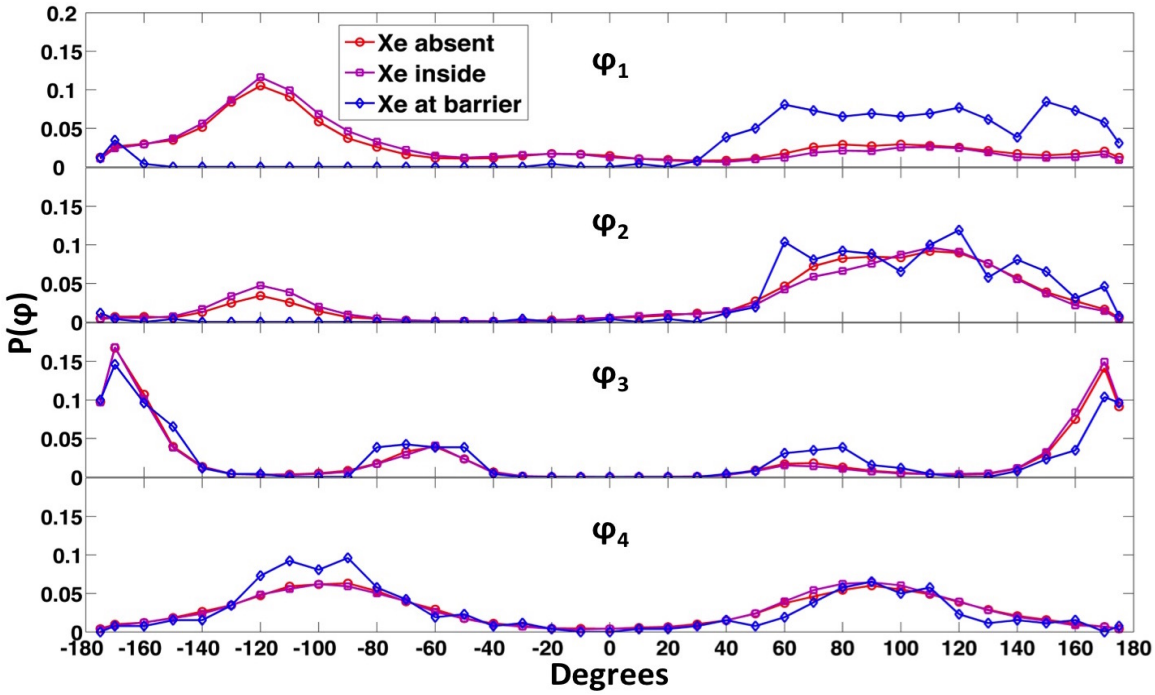


Figure 4.10: Distributions of dihedrals in TAACs sidechains from equilibrium simulation and ABF simulation.

the conformation with acetic group point to the exterior of the pore which leads to the major peak at 120 ° (figure 9). For Xe at the energy barrier state, due to Xe occupying the pore positions, both the methyl and acetic acid groups all have to point to the exterior of the pore to leave space to accommodate the Xe (figure 8B). Because of the presence of Xe, most of conformations that have methyl and acetic group point to the exterior of the pore (φ_1 range from 60 ° to 180 ° and φ_2 range from 60 ° to 160 °) are all possible for cryptophane to adopt except the conformations with unstable 1-4 dihedral interaction. For the distal dihedrals in the charged sidechain, since the torsion around φ_3 mostly prefer the trans conformations so as to avoid the repulsion between carboxyl and phenol, and this trans conformation will not collide with Xe, therefore its distribution is not affected by the presence of Xe (figure 9). For φ_4, because of the presence of Xe, the two conformations (-120 ° and 120 °) of it are no longer equivalent. In order to be further displaced from Xe, the carboxyl oxygen would prefers to be directed away from the pore, leading to φ_4 being more populated at -120 ° than at 120 °.

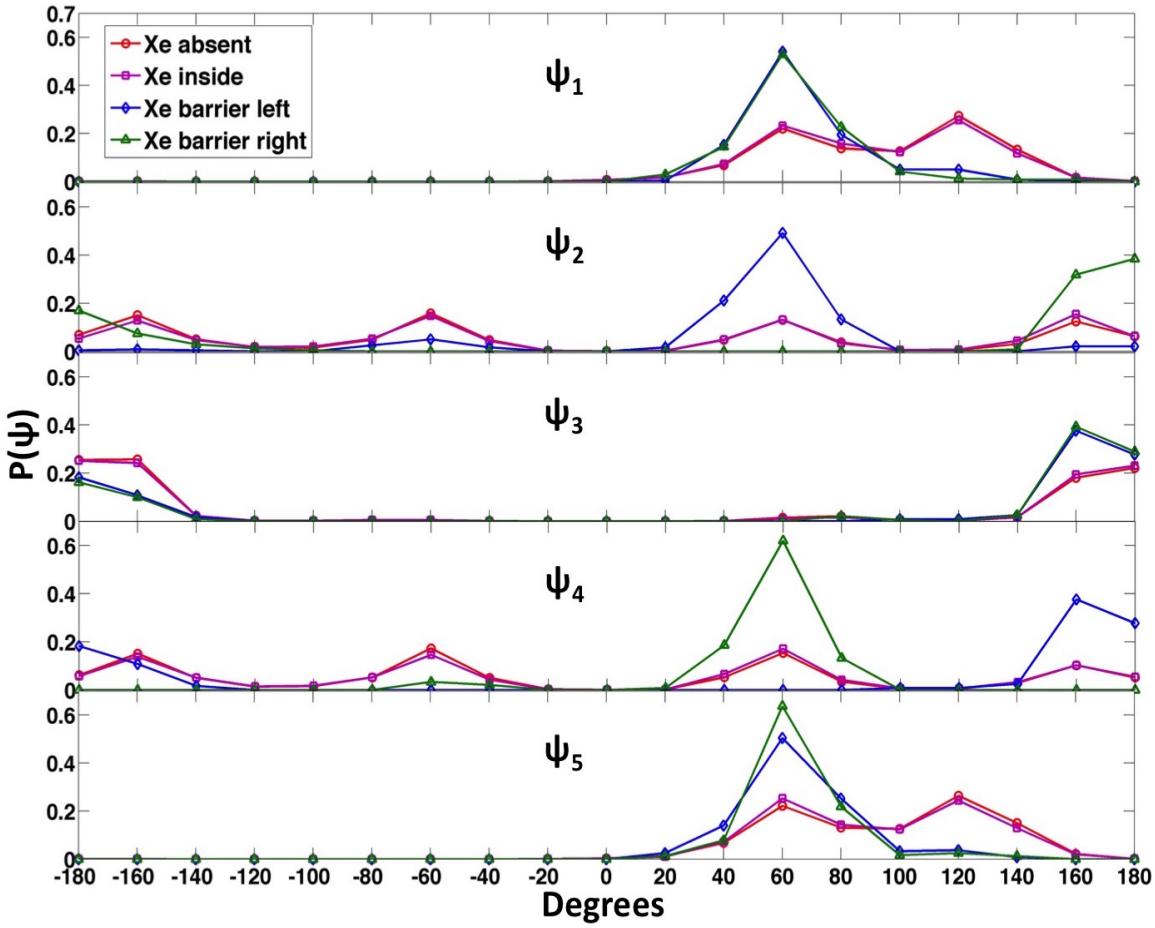


Figure 4.11: Distributions of dihedrals in TTPCs linkers from equilibrium simulation and ABF simulation.

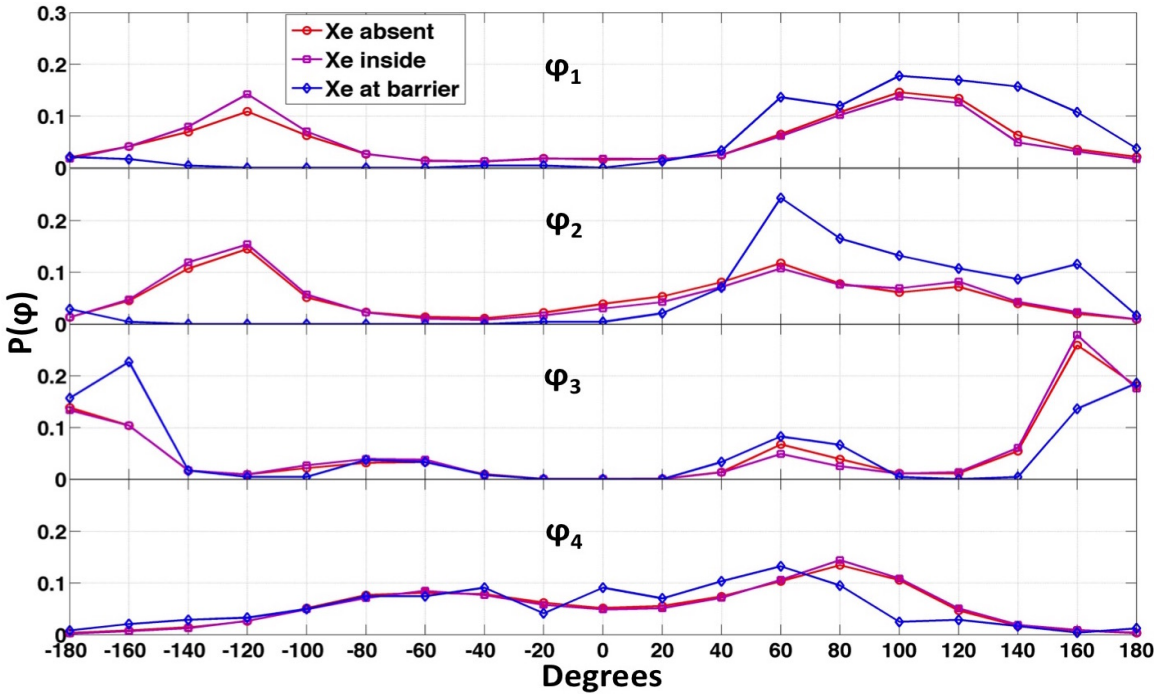


Figure 4.12: Distributions of dihedrals in TTPCs sidechains from equilibrium simulation and ABF simulation.

Dihedrals in linkers and sidechains of other three cryptophanes.

For TTPC and TTEC, who also have methyl group as one set of their sidechain with charged group as the other set of sidechain, the distribution in their linkers have the same features as TAAC (figure4.11 and figure4.13). In conditions of both Xe bounded and unbounded, due to either Xe or water can occupy and fill the cavity, distributions of linker dihedral are showed to be the same and all linker dihedral have multiple peaks except Ψ_3 adopted trans conformations. When Xe at the energy barrier state, due to the steric effect of Xes presence, both the "left " and "right " linkers would be stretched so as to spare space for Xe to occupy and therefore most samples at this state having their Ψ_1, Ψ_2, Ψ_3, Ψ_4 and Ψ_5 populated at specific value (uniform conformations). For the dihedrals in the sidechains, unlike in the case of TAAC, one set of sidechain methyl group is much more hydrophobic than the other, acetic acid, the other set of sidechain in TTPC and TTEC, triazole ethylamine and triazole propionic acid are also hydrophobic. And this leads that both sets of sidechains in

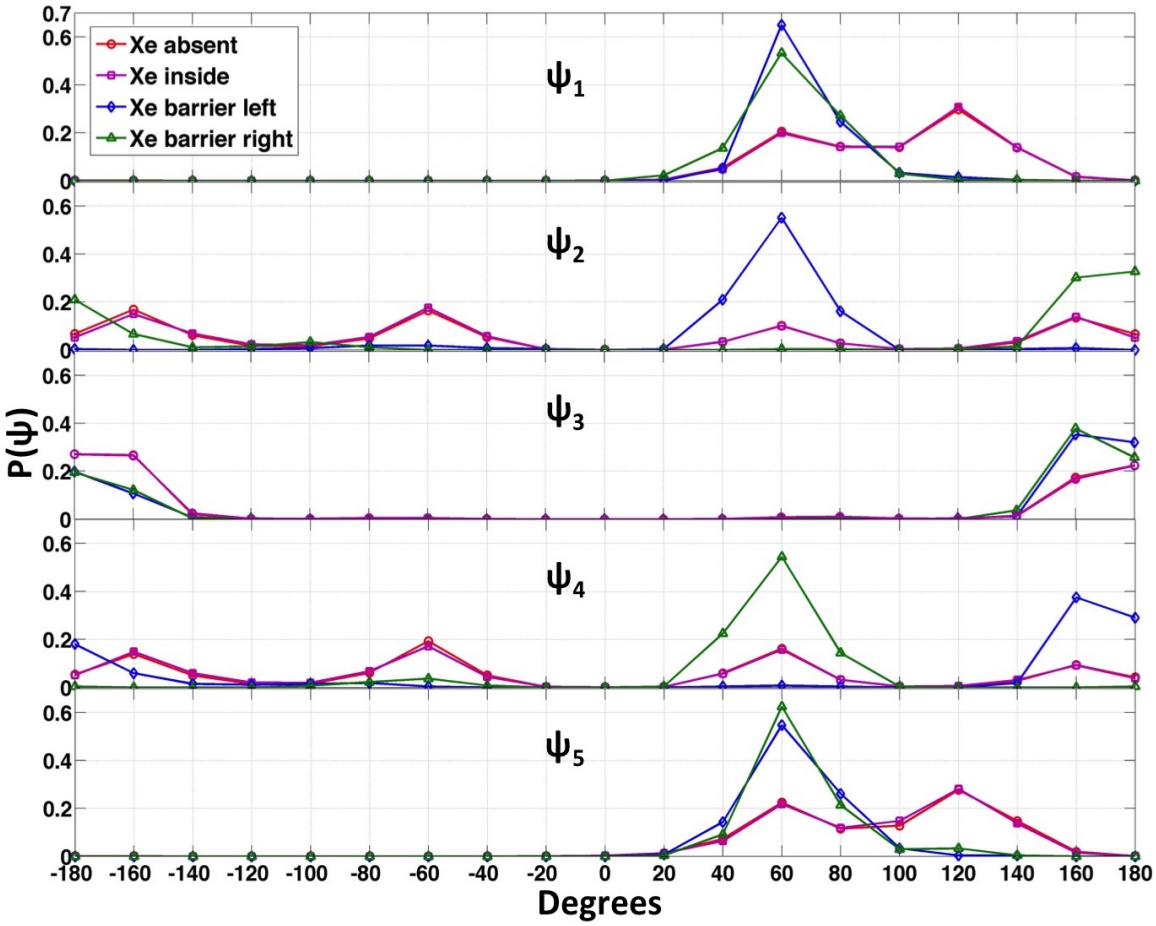


Figure 4.13: Distributions of dihedrals in TTECs linkers from equilibrium simulation and ABF simulation.

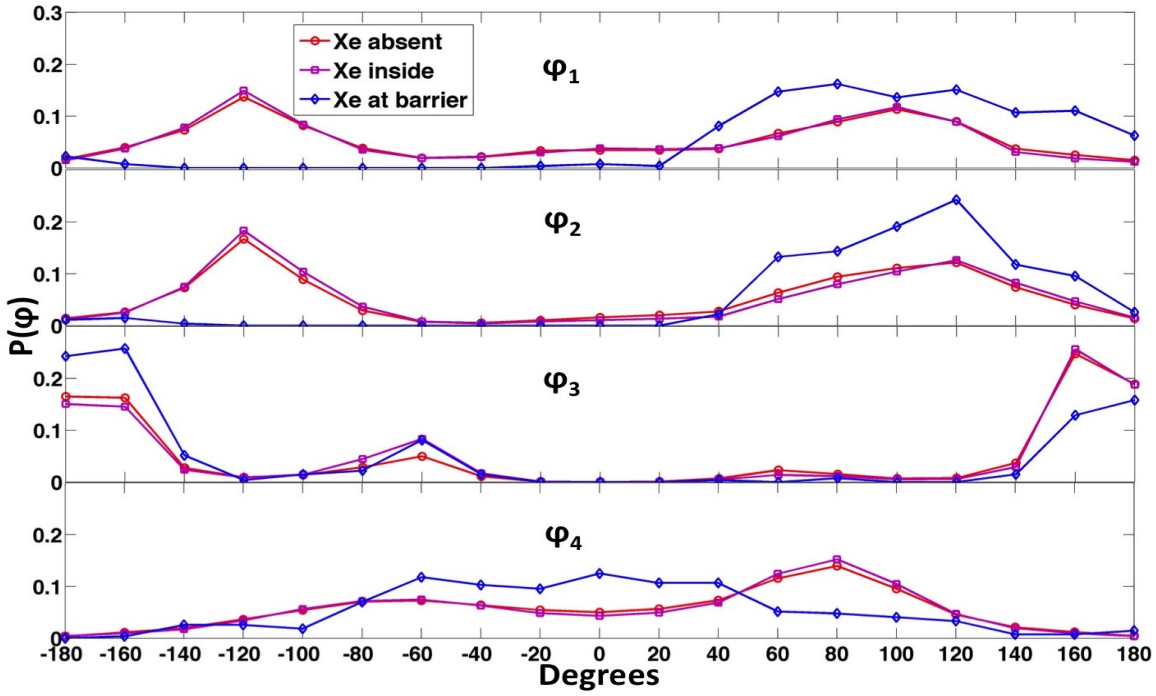


Figure 4.14: Distributions of dihedrals in TTECs sidechains from equilibrium simulation and ABF simulation.

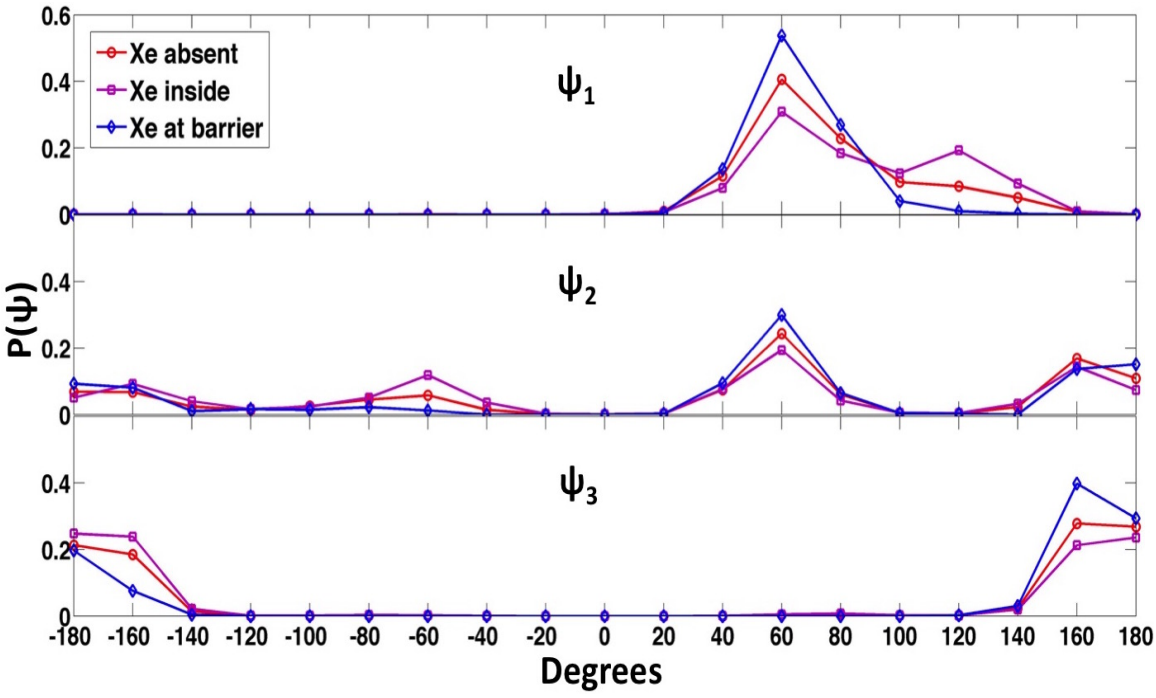


Figure 4.15: Distributions of dihedrals in m2n2s linkers from equilibrium simulation and ABF simulation.

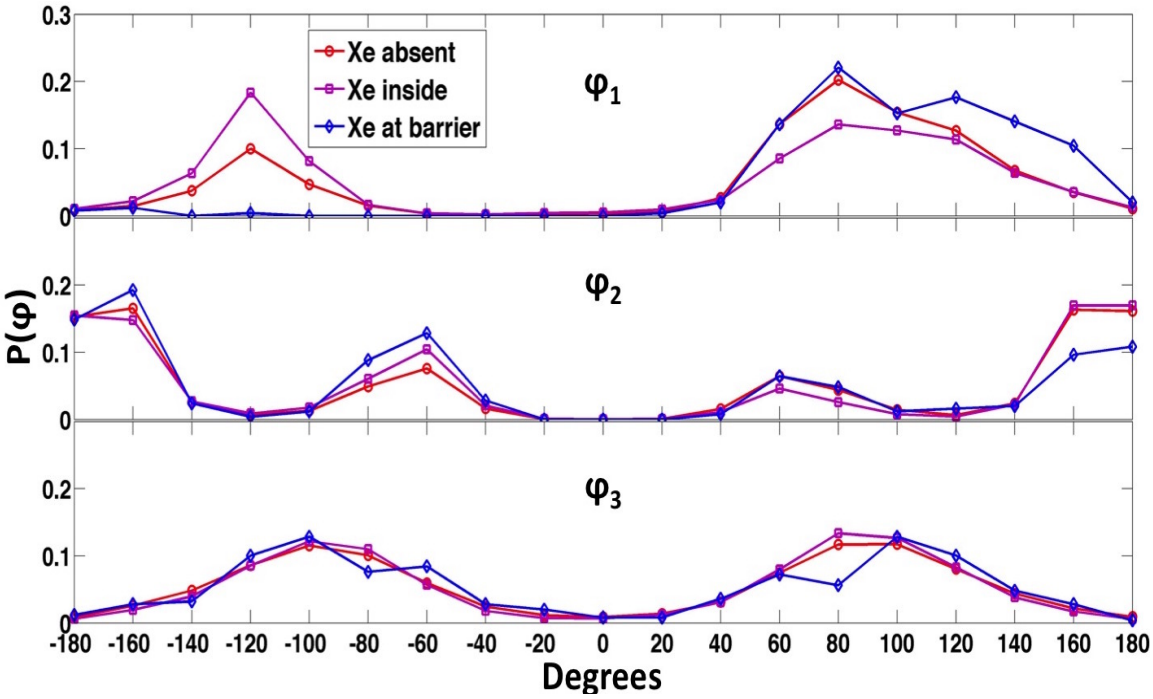


Figure 4.16: Distributions of dihedrals in m2n2s sidechains from equilibrium simulation and ABF simulation.

TTPC and TTPC all prefer to point to the interior of cavity pore, but because the space in the cavity pore is limited, therefore, φ_1 and φ_2 are observed to point to pore interior (-120 °) and exterior (120 °) with similar probability in both Xe bounded and Xe unbounded states (figure4.12 and figure4.14). When Xe is at energy barrier state, both sidechains would point to the pore exterior so as to accommodate Xe. In most of the configurations at the energy barrier, φ_1 and φ_2 have adopted positive values (figure 4.12). For m2n2, due to it is top and bottom symmetric, it only have 3 dihedrals in linker (Ψ_1, Ψ_2, and Ψ_3) and 3 dihedrals in sidechain (φ_1, φ_2, and φ_3) which are the three dihedrals connect acetic acid. And the symmetry also make the two linkers that are closest to Xe identical. Since the m2n2 cavity interior is much less hydrophobic than the other three cryptophanes, the dihedrals in linkers in Xe unbounded choose the conformations that having the interior space already stretched and fulfilled, $\Psi_1 \simeq 60$ ° and $\Psi_2 \simeq 60$ °, which is very close to the condition of Xe at energy barrier state (figure4.15). In condition of Xe bound inside, the φ_1 have siginificant probability to adopt the 120° conformation in addition to the 60°

one due to the cryptophane cavity become more hydrophobic and lead m2n2 close up its interior space a little bit(figure 4.15). At the energy barrier state, linkers closest to Xe adopt the structure allowing Xe to pass through. TTPC and TTEC have hydrophobic side chains, whereas m2n2 has acetic acid groups, and even in the unbound state, these acid groups would prefer to be directed toward solvent and away from the pore ($\varphi_1 \approx 120\ ^\circ$) and (figure4.16). Interestingly, when Xe is present inside the cage, m2n2 conformations where the side chain is directed toward the interior are also observed ($\varphi_1 \approx$ -120 °). When Xe is at energy barrier state, as observed in cases of other cryptophanes, the sidechain would have to point to the exterior ($\varphi_1 \approx 120\ ^\circ$) so as to spare space for Xe to pass through.

4.5. Conclusions

Through extensive ABF simulations, free energy profiles of the processes of Xe entering different cryptophane cavities are calculated. The potential of mean force is calculated as a function of the distance ξ between the center of mass of the cryptophane and the Xe nucleus, where ξ ranges from from 0.0 Å to 15.0 Å. From calculated free energy profile (PMF curve), it is shown that all systems share similar PMF's and that Xe will have to overcome a free energy barrier (∼3kcal/mol high) upon entering the cage and residing in a deep energy well (10.5-11.5 kcal/mol height). By making use of free energy profile with respect to ξ, binding affinity of Xe are calculated, and the estimated affnities are close to experimental measurements. Experimental binding affinities are positively correlated with the depth of the energy well that TTEC ¿ TAAC ¿ TTPC ¿ m2n2. Additionally, heights of energy barrier of three higher binding affinity cryptophane (TTEC, TTPC and TAAC), which also possess more hydrophobic siechains showed to be higher than m2n2; m2n2 has lowest binding affinity and less hydrophobic side chains. In addition to the free energy calculations, the trajectories generated from the ABF simulations are used to analyze the conformational changes of cryptophane as Xe exits the cavity interior and enters aqueous solvent. Analysis of dihedrals in linkers and sidechains revealed that all cryptophanes would adopt the similar conformations in which the linkers are twisted and stretched and

sidechains are pointed to the exterior of the cavity. These structures provide space for Xe to pass through. These structure adjustments, especially of the linkers of cryptophanes, lead the cryptophane interior volume increase as Xe pass through the portal compared with Xe bound inside and Xe isolated from cryptophane. Additionally, among the four cryptophanes investigated in this work, the structure adjustment of sidechain is different as Xe passes through the portal. For the cryptophane in which sidechains are both very hydrophobic (TTEC and TTPC), these sidechains all have equal probability to point to the interior of cavity portal so as to keep the cavity "dry "at the Xe bounded state and Xe unbound state. When Xe needs to pass through the portal, the two sidechains closest to Xe reorient and point to the exterior at the same time. For the cryptophane where one set of sidechains is much more hydrophobic than the other set (TAAC), the more hydrophobic sidechain (methyl) group that would point to the interior of portal when Xe is bounded inside and Xe is isolated from is directed toward the exterior of the portal when Xe needs to pass through portal. The other less hydrophobic sidechain (acetic acid) already points to the exterior of the portal for most samples when Xe is bound or unbound, and this side chain remains directed to the exterior when Xe passes through the portal. For the cryptophane m2n2 with both sets of sidechains are less hydrophobic (acetic acid), in unbound state the sidechains point to the exterior of the portal. When Xe is bound inside, the acetic acid become less probable to point to the exterior and more probable to point to interior of the portal. For m2n2 when Xe occupy the portal, the acetic acid would exclusively point to the exterior of the portal. These findings provide a molecular picture along the process of Xe exiting and entry cryptophanes from both energetics and conformational prospects. Taking use of the findings in this work, we would be able to put forward ideas to modify cryptophane structures to slow down or speed up the Xe exchange rate by substitute the hydrophobic sidechain to a harder or easier to rotate group. These would also have effect on the hydrophobicity of cryptophane interior which would probably result a higher or lower binding affinity cryptophane.

CHAPTER 5 : Molecular basis of Alkali cation affinity for water soluble cryptophane

5.1. Abstract

Cryptophane-1 has been reported to have high binding affinities with cations experimentally, especially with large size alkali metal ions (K^+, Rb^+ and Cs^+). In this work, we used molecular dynamics simulation to model ion-cryptophane systems and free energy perturbation to calculate relative and absolute binding affinities of cryptophane-1 to three cations, K^+, Rb^+ and Cs^+. The calculated relative binding affinities among ions are in good agreement with experimental measurements. As with other cryptophanes that allow water molecules to occupy the interior cavity, simulations of cryptophane-1 also revealed small numbers of interior water molecules. These water molecules form hydrogen-bonded clusters within the cavity. We find the binding process invoving these large metal cations can be considered as the metal ion substituting form some of water molecules inside cavity, and the cryptophane complex has ion coordinated with water inside. The water coordinated ion inside cryptophane has a similar water-ion coordination distance as that observed in bulk water. The reason that cryptophane has the highest binding affinity to Cs^+ is that Cs^+ has the lowest desolvation energy and the binding of the ion liberates more water molecules than that of K^+ or Rb^+.

5.2. Introduction

Host-guest interaction can be traced back to mid 1960s when the first crown ether molecule dibenzo-18-crown-6 was synthesized accidentally by Charles Pedersen to bound potassium in water[1]. Subsequently, a series of crown ether molecules were synthesized for encapsulating multiple sizes of cations[110] . Though five decades have passed, host-guest interactions are still being investigated, especially the interactions associated with encapsulating the metal ions. [111–113]

Among the metal cations, cesium-137 is a problematic radioative one that is a common

fission product during the nuclear reactions. It has a half life of about 30 years. Especially in light of the recent Fukushima Daiichi Nuclear Power Plant accident on March 11, 2011, where release of Cs-137 affected the water of hundreds of thousands of people [38, 39]. Due to its high solubility in water, it is diffcult to be remove by precipitation. Other methods have been developed for removing cesium ions. Inorganic materials can adsorp cesium and filter it, and these include potssium zinc hexacyanoferrate[114], polyacrylonitrilezeolite nanocomposite[115], titanate nanofibers and nanotubes[116], prussian blue encapsulated in alginate beads together with carbon nanotubes[117]. Organic adsorbents[118] or ion exchange methods may be used to remove cesium in water[119]. Non-aqueous solvents have been used instead of water to extract cesium from aqueous to hydrophobic solvent; these approaches include using calix[4]arene-bix(tert-octylbenzo-crown-6) ionic liquid[120] or using crown-like organic solvent[121]. With many of these systems, it is difficult to confer selectivity for Cs and to obtain molecular insight on the strength and selectivity of binding.

Cryptophane molecules, which are cage-like synthetic molecules that have two clclotriguaiacelene caps linked by alkyli linkers, can also be used to bind or encapsulate the cesium cations in water. Some cryptophane derivatives can be dissolved in organic solvent and then capture cesium from water[122]. A recently synthesized cryptophane having no sidechains (figure 5.1) was found to have an unusually high aqueous binding affinity with cesium cation compared to other alkali cations[25, 123, 124]. Although in experimental studies, binding of cesium by a cryptophane can be observed through NMR spectroscopy isothermal titration calorimetry(ITC) experiments, a molecular picture of the encapsulation process involving cryptophane and Cs^+ is missing. Therefore, molecular dynamics simulation is used herein to model one cryptophane(MM) system that has high binding affinity with the cesium cation. We seek to shed light on the atomistic reasons why the cesium cation has a higher binding affinity compared to the other alkali cations.

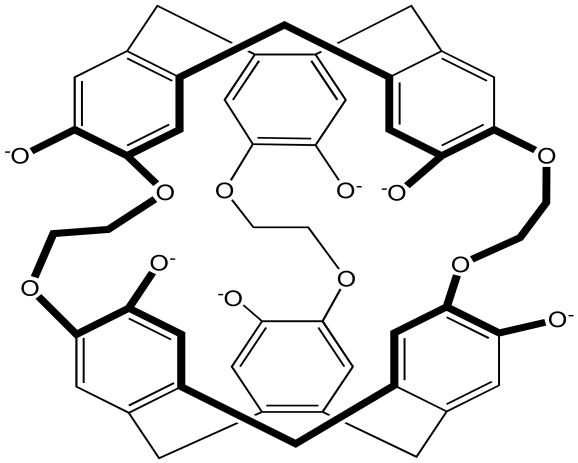


Figure 5.1: Strucure representation of cryptophane-1

5.3. Methods

5.3.1. Molecular models.

Cryptophane-1 is investigated in this work and its initial coordinates come from the crystal structure of m2n2[125] with all its sidechains removed using VMD[126]. Given that the cryptophane is soluble in basic solution and experiments of binding to different cations were done in pH ¿ 12 solutions, the molecule is simulated in an ionization state such that all the hydroxyl groups on the phenol rings are deprotonated. The molecule was optimized and simulated using NAMD[**?**]. In each system simulated, 1150 TIP3P[**?**] water molecules are added for solvation, yielding a water box with dimension of 36Åx36Åx36Å. Sodium ions are randomly added to netural the charges. Bond parameters for cryptophane come from AMBER94[**?**] force field. Non-bonded interactions composed of two parts, Lennard-Jones potantial and electrostatic interactions. Lennard-Jones paramters for cryptophane atoms come from OPLS force field[127, 128], while the parameters for ions come from newly developed force field[129]. The partial charges for each atom in cryptophane are generated in R. E. D. server[130] which fitted the charges in each atom so as to reproduce the electrostatic potential at each atom.

5.3.2. Molecular dynamics simulation.

All molecular dynamics simulations and below free energy perturbation calculations are carried out in NAMD[108]. In all simulations and calculations, in order to evaluate non-bonded interactions, 12Å was used as cutoff distance, 10Å was used as switching distance from where the switching function begin to work and 14Å was used as pairlist distance to indicate the maximium space to include the neighboring atoms. The nonbonded pairlist was evaluated every 10 steps, and every step is 2 fs. Each system was divided into 2Å cubes. The SHAKE algorithm[90] was used to make all bonds between hydrogen and heavy atoms rigid and thus making 2 fs timestep simulation correctly. Nonbonded interaction was evaluated every step and electrostatic interaction was evaluated every 2 steps. Particle mesh Ewald[53] method was used to evaluate the full system periodic electrostatics with 1Å as grid space. All simulations and calculations are implemented in NPT ensemble, pressure is 1 atm and temperature is 300 K. In order to implement NPT ensemble simulation, Langevin dynamics method was used to maintain the temperature with damping coefficient to be 1/ps and Langevin piston Nose-Hoover method[55] to maintain the pressure with oscillation period equal to 100 fs and decay time equal to 50 fs. Each system was minimized by using conjugate gradient and line search algorithm[131] for 50,000 steps (100ps) and then equilibrated for 10,000,000 steps (20ns). Configurations were sampled from the simulation trajectories every 1000 steps (2ps) using for analysis.

5.3.3. Free energy perturbation method

The free energy difference between two states (state 0 and state 1) can be estimated using equation 5.1.

$$\Delta A_{0\to 1} = -RTln\left\langle exp\{\tfrac{1}{RT}[H_0(x,p_x) - H_1(x,p_x)]\}\right\rangle_0 \tag{5.1}$$

This equation states that the free energy change can be obtained by averaging Hamiltonian change from state 0 to 1 for all the samples generated using state 0. However, this method is only applicable in the case that state 0 and state 1 significantly overlap with each other

in phase space and thereby can provide a easy-to-converge and accurate result. For cases that two states significantly differ from each other or have high free energy difference, perturbation to gradually transition from one state to another is a good solution to obtain the accurate free energy estimates. The Hamiltonian in whole process from state 0 to state 1 is then represented as a hybrid of initial state (0) and final state (1), and the progress of the perturbation can be indicated as λ which ranges from 0 (initial state) to 1 (final state): $\mathrm{H}_\lambda = 1 - \lambda)H_0 + \lambda H_1$. During implementation, this can be done by scaling the interaction of atoms involved in the two states and changing their interactions with other all atoms whose properties do not change with λ. To assure convergence, λ is chosen to increase gradually. Since the Gibbs free energy change (ΔG) of a process in NPT has the relationship with Helmholtz free energy change (ΔG) as equation 5.2, and the simulation box volume change along the simulation process is tiny (PΔV), this contribution can be neglected and $\Delta\mathrm{G} = \Delta\mathrm{A}$.

$$\Delta G_{0\to1} = \Delta A_{0\to1} + P(V_1 - V_0) \tag{5.2}$$

In this work, the free energy changes of two processes are calculated: mutate one ion $\mathrm{M_1}^+$ to another ion $\mathrm{M_2}^+$ in water and mutate one ion to another inside cryptophane (**??**). These can be incorporated in a thermodynamic cycle to estimate relative free energies of binding (figure 5.2). During the mutation process that occurs within the cryptophane, to avoid ions escaping the cavity. Ions are constrained to a spherical space inside cavity through a harmonic potential, and the radius beyond which potential would take effect is equal to the maxmium distance of the ion from center mass of cryptophane as observed in an unconstratined equilibrium simulation. The free energy perturbation processes starts with the ensemble ($\lambda = 0$) that having $\mathrm{M_1}^+$ fully coupled to all other atoms, then as λ increases, the interaction of $\mathrm{M_1}^+$ and all other atoms is scaled down gradually with $\mathrm{M_2}^+$ appearing and coupling to all other atoms. The interactions of M_2^+ with all other atoms are scaled up gradually, until $\lambda = 1$, at which point $\mathrm{M_2}^+$ is fully coupled to the system and $\mathrm{M_1}^+$ fully

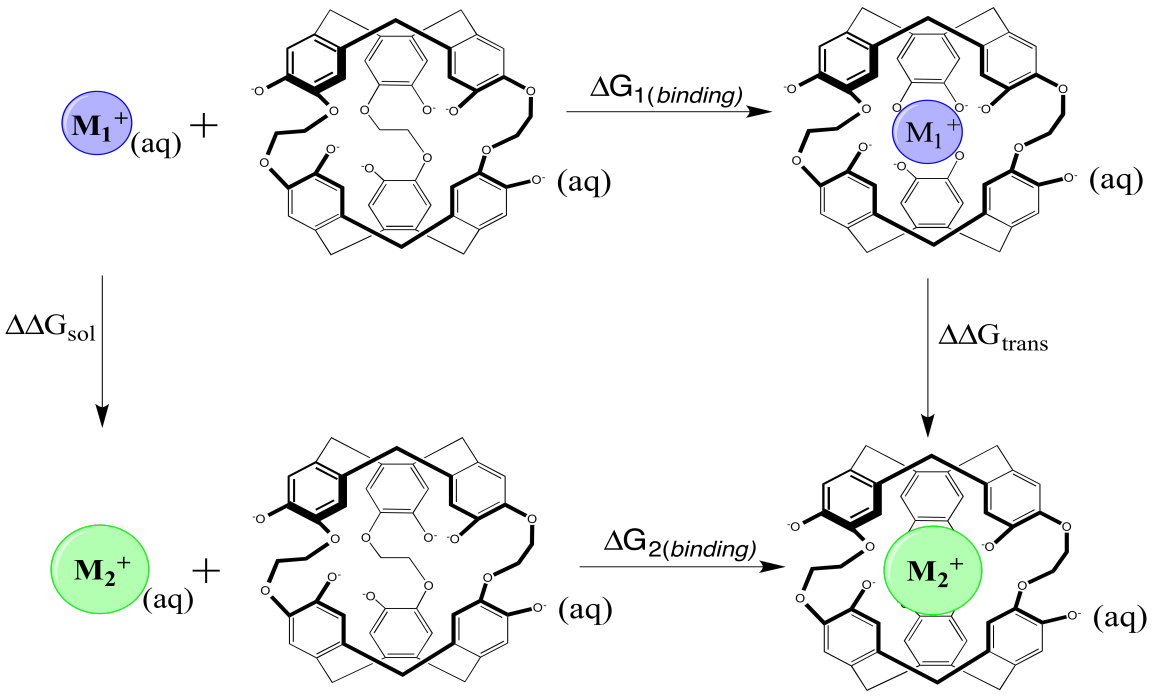


Figure 5.2: Free energy perturbation calculation strategy

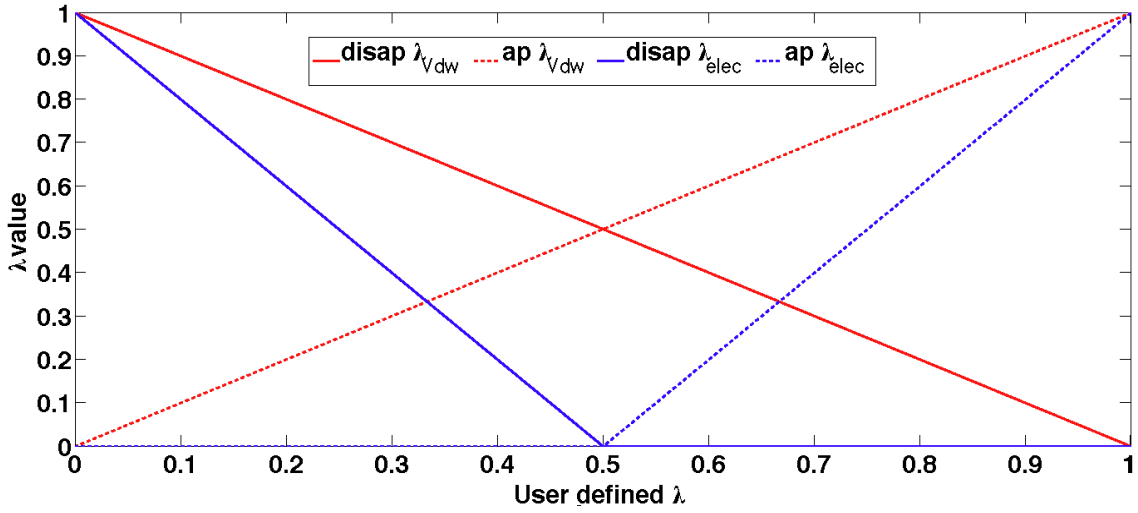


Figure 5.3: Scaling of λ during free energy perturbation process

decoupled. The "appearing" ion ($M_2{}^+$) could overlap with existing atoms in the system when it is introduced and cause unstable simulations and singularities in the potential [132]. A shifted Lennard-Jones potential feature[133] is used to avoid the $M_2{}^+$ overlapping with other atoms at intermediate values of λ. Additionally, in order to improve the convergence of the calculation at each λ, the van der Waals (vdW) interactions and electrostatic interactions for ions are scaled differently (figure 5.3). The whole process of mutating ions are divided into 20 windows with λ increases by 0.05 at each step. For the vdW interactions of ions with environment, which are Lennard-Jones potential in implementation, the interaction of the annihilated or "disappeared" ion ($M_1{}^+$) with the environment are scaled down at the same rate as λ; the interaction of "appeared" ion ($M_2{}^+$) with the environment are scaled up at the same rate as λ as well. For the electrostatic interaction of ions with environment, the speed of scaling down the interaction of disappeared ion ($M_1{}^+$) with environment and scaling up the appeared ion ($M_2{}^+$) with environment occur with slopes twice that of *lambda*. This electrostatic interaction of $M_1{}^+$ is decoupled from environment when $\lambda = 0.5$. The electrostatic interaction of the appearing $M_2{}^+$ start to couple to the environment at $\lambda = 0.5$ and is fully coupled when $\lambda = 1.0$. This can be expressed as follows and is illustrated in figure 5.3.

$$V_{elec}(\lambda) + V_{vdW}(\lambda) = (1-\lambda)V_{0,vdW} + \lambda V_{1,vdW} + max(0, 1-2\lambda)V_{0,elec} + max(0, 2\lambda-1)V_{1,elec} \tag{5.3}$$

In the process of estimating free energy change of mutating ions inside cryptophane, in order to consider the effects of polarizability in aqueous media, we scale charges of ions and each atom in cryptophane with a factor of $1/\sqrt{\epsilon_{el}}$, where ϵ_{el} is the electronic dielectirc constant of water. This modification is compatiable with the TIP3P water model has been found to be effective in other aqueous simulations [49, 50]. By subtracting the free energy changes in these two processes, we obatin the relative binding affinity difference of cryptophane to two individual ions (equation 5.4).

$$\Delta\Delta G_binding = \Delta G_{1(binding)} - \Delta G_{2(binding)} = \Delta\Delta G_{sol} - \Delta\Delta G_{mut} \tag{5.4}$$

5.4. Results and discussion

5.4.1. Relative binding affinity calculation results.

By using free energy perturbation method (FEP) and thermodynamic cycle, we obtained relative binding affinities between Cs^+ and K^+, and between Cs^+ and Rb^+. Our results confirm that cryptophane-1 has higher binding affinity to larger size ions, which follow the ordering: $Cs^+ > Rb^+ > K^+$. The results also match experimental measurements quantatively (5.1). The calculated free energy change of mutating ions in water matches solvation free energy difference of ions.[134] The solvation energy of smaller ions are larger in value than larger size ions, this is due in part to small ions have lower polarizability. Regarding the process of mutating ions inside cavity, it's interesting that though cryptophane-1 has much higher binding affinity with Cs^+ than smaller ions, the free energy change $\Delta\Delta G_{mut}$ of mutating K^+ or Rb^+ to Cs^+ is positive which suggests that the smaller ions (K^+ and Rb^+) within the occupied cryptophane are could be more stable than Cs^+ occupied cryptophane. For the smaller cations, water molecules can enter the cavity to coordinate each, and more water molecules can fit into the cavity than for Cs^+. Although Cs^+ is not as stable as other smaller ions to occupy the cryptophane interior relatively, the desolvation energy, or the energy Cs^+ need to acquire to escape from aqueous hydration is much lower (5 10kcal/mol) than other smaller size ions. This results in Cs^+ having the highest affinty to be encapsulated by cryptophane-1.

	Calculated Result			Experimental Result
	$\Delta\Delta G_{sol}$ (kcal/mol)	$\Delta\Delta G_{mut}$ (kcal/mol)	$\Delta\Delta G_{binding}$ (kcal/mol)	$\Delta\Delta G_{binding}$ (kcal/mol)
$K^+ \rightarrow Cs^+$	10.8 ±0.3	3.1±0.2	7.7±0.4	7.7±0.6
$Rb^+ \rightarrow Cs^+$	5.2 ±0.3	1.2±0.2	4.0±0.4	3.9±0.7

Table 5.1: Binding energy result table

Guest	top-bot distance (Å)
Na^+	6.19±0.13
K^+	6.23±0.17
Rb^+	6.21±0.18
Cs^+	6.03±0.22

Table 5.2: Dimension of cryptophane-1 in conditions of different guest ions: the distance between top aromatic cap and bottom aromatic cap.

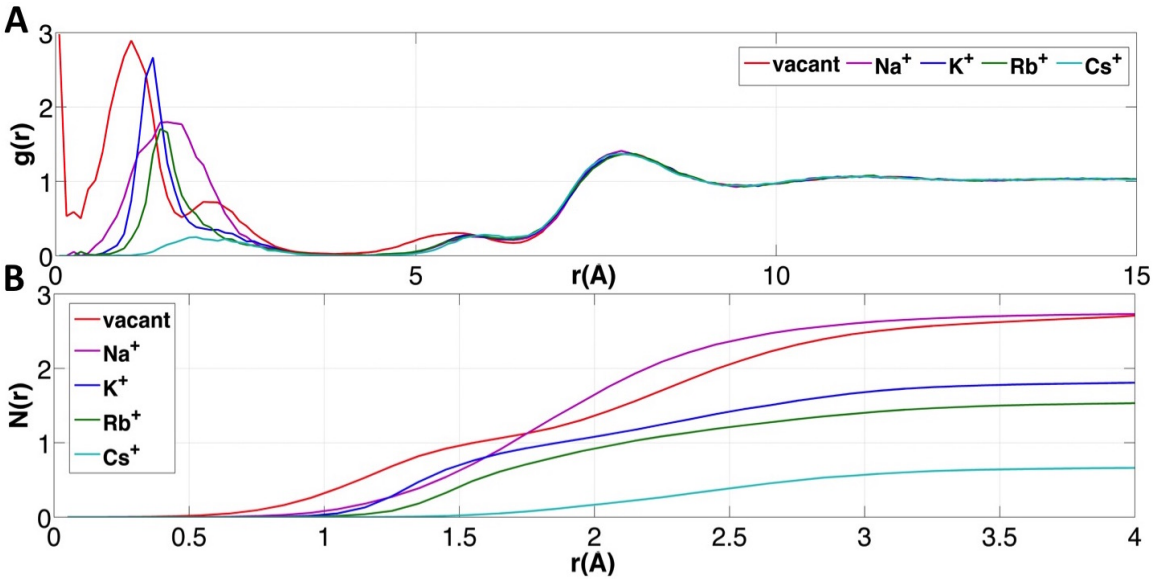


Figure 5.4: Radial distribution of water molecules around center mass of cryptophane

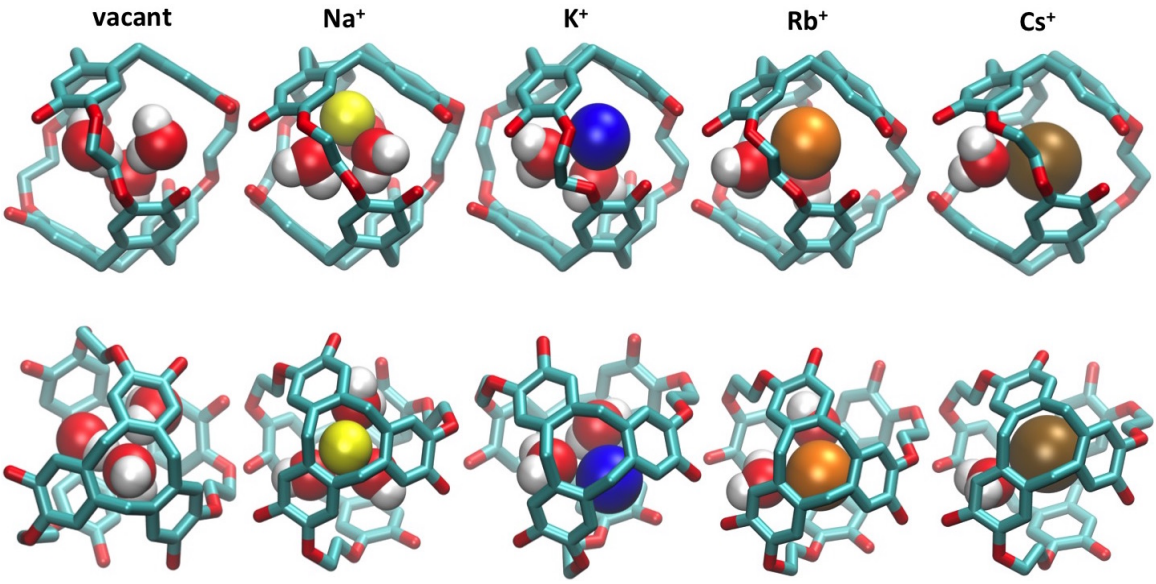


Figure 5.5: Representative structures of ion-water complex inside cryptophane cavity

5.4.2. Host-solvent interaction

To further explore the molecular aspects of cation binding, 20 ns equilibrium simulations are implemented for cryptophane-1 with no cation present and with Na^+, K^+, Rb^+ and Cs^+ occupied inside cavity. The radial pair distribution of water molecules (center of oxygen in water is used to identify the water position, figure5.4A) around cryptophane-1 and average number of water molecules within sphere of radius r relateive to center of cryptophane are calculated and analyzed so as to identify the position of water molecules and their occupancies. For the vacant cryptophane, water molecules would enter the cavity spontaneously as in other cryptophane solutions[106]. And they are populated at two locations, one is at the center of cryptophane cavity and the other is 1.1 Å away from cryptophane center (figure5.4A). On average, there are 2.7 water molecules inside cavity (figure5.4B). The molecular picture is that there is one water molecule at the center and two at the 1.1Å position and they bind each other to form a three molecule cluster via hydrogen bonding, the center water molecule provide its oxygen to coordinate with hydrogens from the two other water molecules (figure5.5). When Na^+ occupies the interior, there are 2.7 water molecules (figure5.4B) at 1.6Å away from center(figure5.4A). The Na^+ sits inside

one cap and coordinate with the all three oxygens from each water molecules and form a tetrahedral structure (figure5.5). This is due to Na^+ having high amount of solvation energy and requiring as many water molecules as possible to solvate. The cryptophane interior provides a fitted space to accommodate three water molecules at the same time. K^+ and Rb^+ also occupy the cryptophane interior, given the size of them are bigger than Na^+, the interior space of the cryptophane can no longer accommodate three water molecules. There are two water molecules occupy the cryptophane interior for most of time and there are 1.8 water molecules in case of K^+ and 1.5 water molecules inside cavity in case of Rb^+ in average. The ion-water complex has a structure where K^+ or Rb^+ are positioned close to the cavity center. Each cation coordinates with two water molecules, and the hydrogens of these two water molecules interact with phenol oxygen at the edge of cryptophane interior (figure5.5). Since Rb^+ is bigger than K^+, the water molecules in the Rb^+ system (1.5Å) are positioned further away from center of cryptophane than in K^+ system (1.3Å). For the case of Cs^+ occupying the cryptophane, there are only 0.7 water molecules (figure5.4B) at 1.9Å away from cryptophane center (figure5.4A). The structure of the cryptophane-ion-water complex has Cs^+ situated at the center of the cavity with one water oriented and use its oxygen to coordinate Cs^+ with its two hydrogens stablized by the phenol oxygen of cryptophane (figure5.5).

5.4.3. Guest-solvent interaction

Since water molecules are observed to enter the cavity to solvate ions in all conditions, the distribution and average number of water molecules about the ion inside cryptophane are calculated and analyzed. These are also examined for the single cation in bulk water for comparison.

5.4.4. Ion-solvent interaction in crytophane environment

Water molecules can enter the cavity so as to coordinate with ions (figure5.6), and as the size of alkali metal ions increases from Na^+ to Cs^+, the distances between ions and

	Peak of first solvation shell (Å)		Coordination number in solvation shell	
	In cryptophane	In water	In cryptophane	In water
Na^+	2.35	2.35	2.7	7.5
K^+	2.75	2.75	1.8	8.3
Rb^+	2.85	2.95	1.6	8.8
Cs^+	3.05	3.10	0.7	9.0

Table 5.3: Solvation results of ions in cryptophane cavity and in bulk water

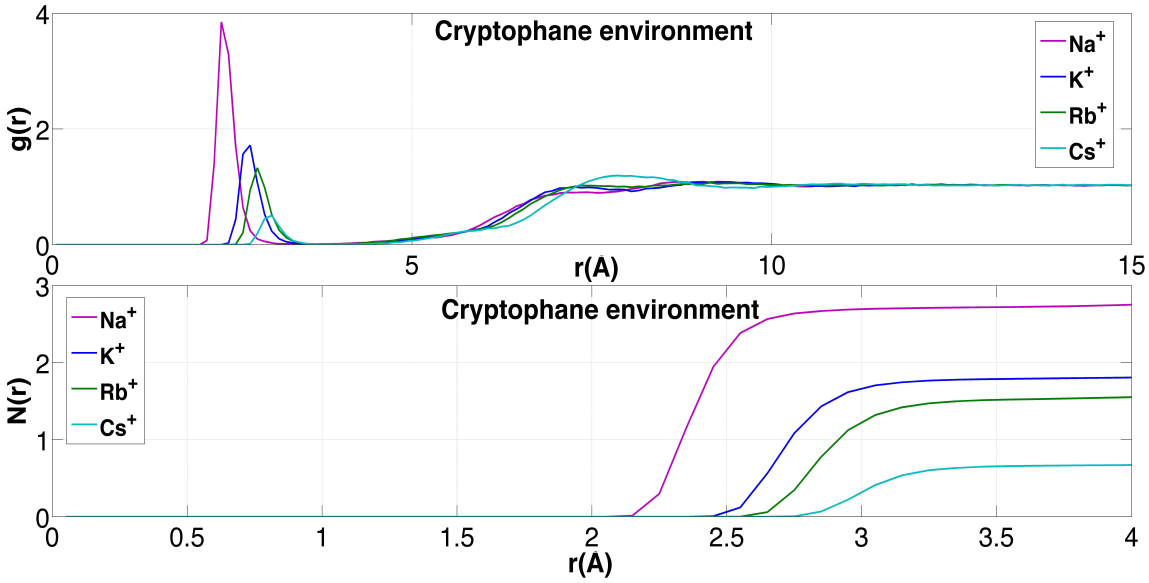


Figure 5.6: Radial distribution of water molecules around ions in conidtion of cryptophane encapsulate ion

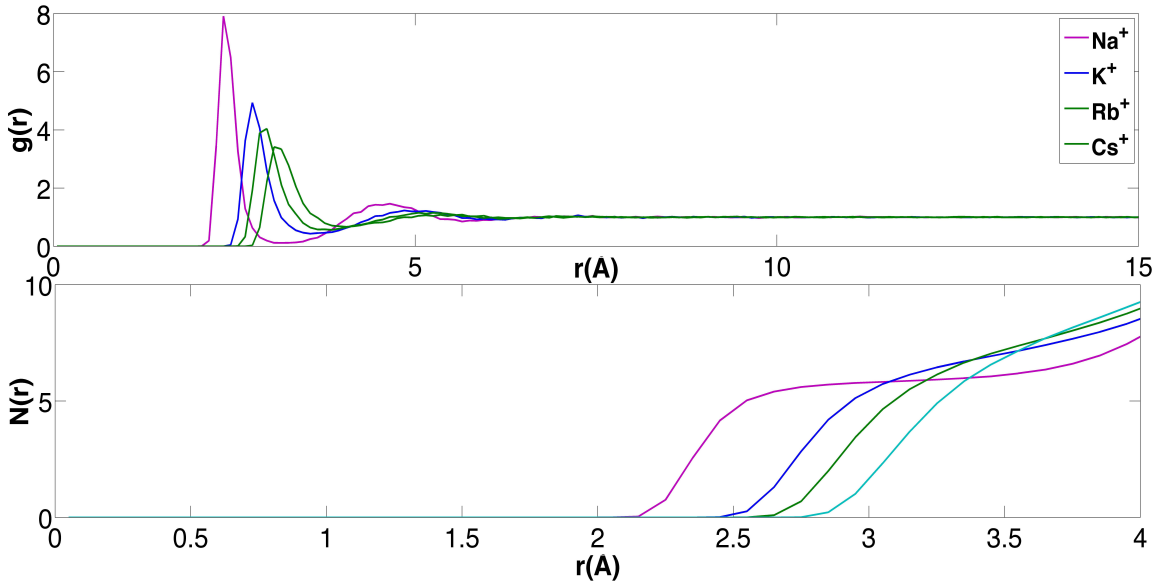


Figure 5.7: Radial distribution of water molecules around ions in bulk water

coordinating water molecules increase as well: 2.35Å, 2.75Å, 2.85Å and 3.05Å respectively (table5.3). Since metal ion and coordinated water molecules are encapsulated inside cavity and isolated from bulk water, and the $g(r)$ of water molecules vanish to 0 at 4.0Å, 4.0Å is used to define the radius of the first solvation shell of water molecules around ion. The coordination number of water molecules in this solvation shell in each system decrease from 2.7 to 0.7 as ion's radius increase from Na^+ to Cs^+ (table5.3), this is due to the interior space available for water molecules to occupy decreases as bigger ion occupy more cavity interior space.

5.4.5. Ion-solvent interaction in water environment

For a single ion in bulk water, the position of the water molecules to coordinate with ions also shifts further from ion as the size of alkali metal ions increases (figure5.7). These positions in all conditions equal the first coordination distances of cryptophane systems except for large size ions, Rb^+ and Cs^+, whose position of water molecules are slightly closer to ion within cryptophane. These positions also match previous simulation results[129]. Though the equilibrium distance between ion and water inside cryptophane match the result in

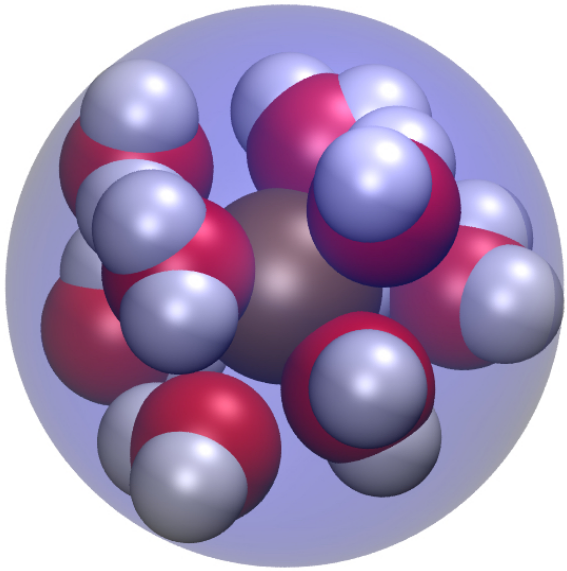


Figure 5.8: Solvation of Cs^+ in bulk water

bulk water, the coordination number are significantly different(table5.3). In bulk water, as many as 7 to 9 water molecules are disturbed and reoriented in presence of ions to solvate ion (table5.3) while very few water molecules in presence of cryptophane. As size of ions increase, though large ion Cs^+ do not interact with water as strongly as smaller ion, due to it has larger solvation shell, it has the highest coordination number. The vast difference of coordination number in conditions of cryptophane and bulk water suggest cryptophane-1 is a good host molecule which can replace these coordinating waters and release them to bulk. The effect is most pronounced for Cs^+, which has largest solvation coordination number change of the cations considered.

5.4.6. Water map analysis

In order to investigate more detailed atomistic picture of water molecules and ions inside the cryptophane cavity, a water map analysis are implemented in this work[56]. Water map analysis is composed of two steps. The first step is the position clustering, which would generate the most populated position that water or ion located inside cavity among sampled configurations. The second step is the orientation clustering, which would generate

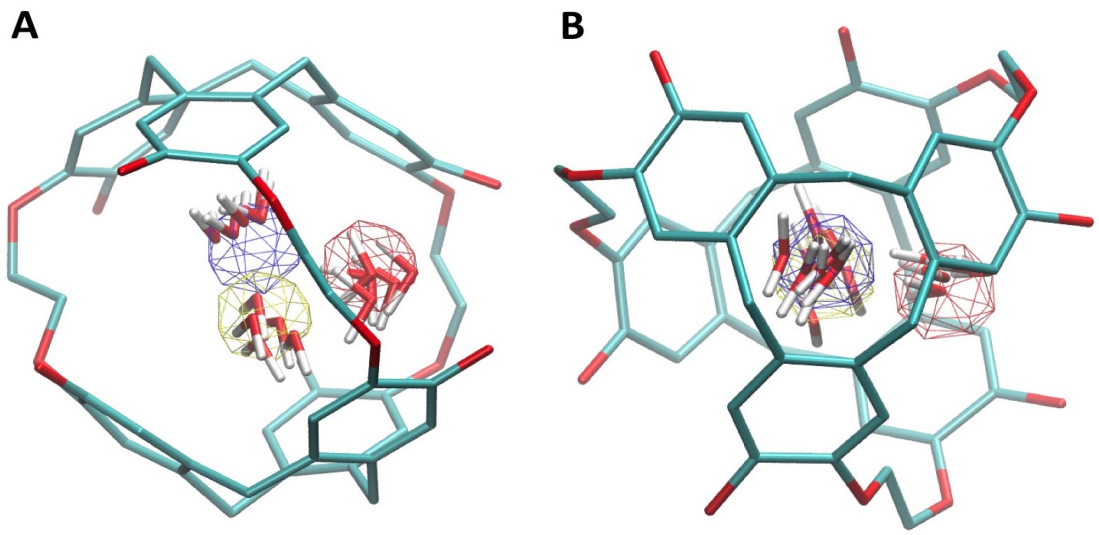


Figure 5.9: Water map clustering for vacant cryptophane. Left panel is the horizontal view and right panel is the perpendicular view. Red wire frame is the most populated cluster, the blue one is the second most populated and the yellow one is the thrid most populated frame. In each cluster, five water molecules with most populated orientation are displayed.

most likely orientations for those water molecules that appear nearby the position cluster adopted. Prior to the position clustering, coordinates of all atoms in the simulation box are aligned. In position clustering, we considered all water molecules that enter the cavity and count the number of all other water molecules that are closer than 1.0Å to each of them. And the coordinates of the oxygen in the water molecule that has highest number of water molecules nearby is defined to be center of the first cluster. Following the same criteria, we can identify the second and third clusters according to the average number of water molecules in each system. In orientation clustering, the orientation of water molecules within 1.0Å of the cluster center are considered and compared. The rotation matrix of rotating each water molecule to the reference water molecule are calculated and converted to quanterion representation (q). Then the criteria of two water molecules having similar orientation is $|q_i \cdot q_j| > 0.9$, and we count the number of similar orientations each water molecule in cluster has, and the water molecule has highest number is the most representative orientation. For the vacant cryptophane, three clusters are generated and the most populated cluster (red wireframe) are located near the center of cavity and near the pore of

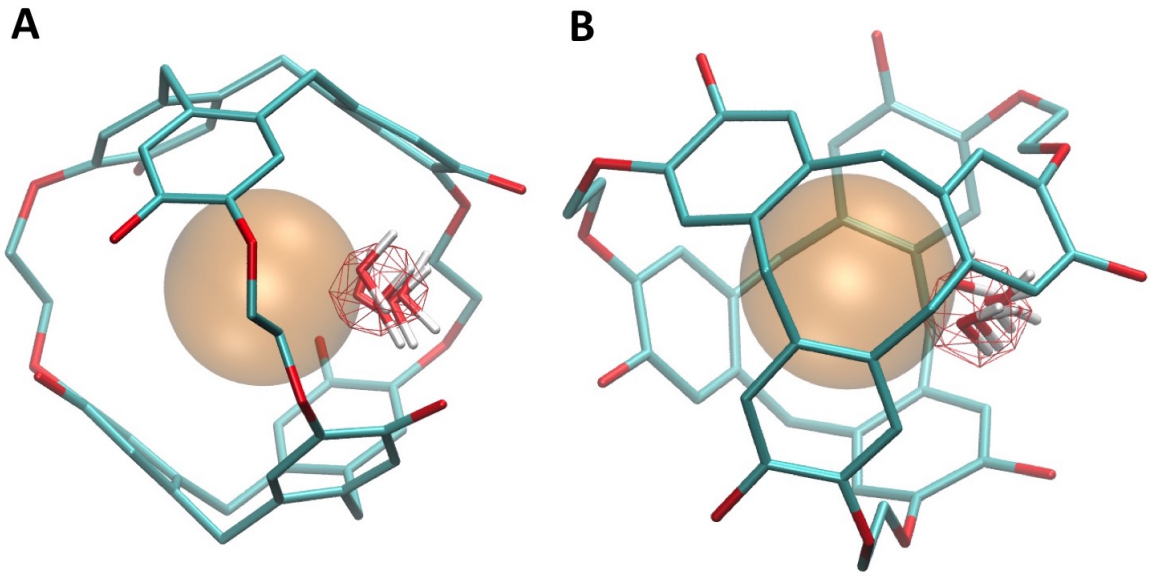


Figure 5.10: Water map clustering for cryptophane with Cs occupied inside

cryptophane(figure5.9). This cluster corresponds to the peak at 1.1Å in plot of g(r) of water around (figure5.4A). The second (blue wireframe) and third (yellow wireframe) clusters are symmetric in reference to center of cavity. These clusters are near the center of cryptophane cavity, and they correspond to the peak at 0.0Å plot of g(r) of water around (figure5.4A). Inside each cluster, all five most populated are showed to have highly uniform orientations, and the water molecules at near the pore adopt the orientation that having two hydrogens to coordinate with oxygen of the water molecules near the center of cavity. In condition of ions occupied inside, most populated positions of ions occupied are generated and displayed as well (figure5.10).

5.5. Conclusions

In this chapter, the molecular basis of cryptophane without hydrophobic sidechain group binding alkali metal ions are explored. Through free energy perturbation method, we distinguished the binding affinities for three relatively large size ions, K^+, Rb^+ and Cs^+, and the cryptophane has the highest binding affinity for Cs^+. The binding process of cryptophane-1 binding ions can be described as ion would subsitute part of the water culster inside cavity.

There are water molecules inside cryptophane cavity to solvate ions in all systems (Na^+, K^+, Rb^+ and Cs^+). It's found that the smaller the ions, the more water molecules would entry the cavity. The reason cryptophane-1 has the highest binding affinity on one hand is the comparatively low solvation energy for Cs^+, who has the largest radii and the highest polarizability and would perturb the structure of water slightest compared to other smaller ions. And on the other hand, the inteior volume of cryptophane-1 with Cs^+ occupied inside cavity is smaller than all other cryptophane-ion complexes suggest the gemotric fit cryptophane-1 to Cs^+ is also the key of the strong binding interaction.

CHAPTER 6 : Design high Xe binding affinity cryptophane

6.1. abstract

The host-guest interaction of Cryptophane-Xe molecular system has great application in magnetic resonance imaging (MRI) given the guest Xe has relatively high natural abundance of 26.4% and high polarization ability while cryptophane structures can be modified and bind a variety biomolecules. The cryptophane-Xe complex can act as bio-sensor to image the environment change based on the Xe chemical shift. Though several water-soluble cryptophanes have been synthesized and have high binding affinity ranging from ΔG = -4.10 to -5.97 kcal/mol, we hypothesised that we could obtain higher Xe binding affinity cryptophanes by using simulation-based studies. In previous work, we proposed two aspects to consider in designing higher affinity Xe binding cryptophanes. One is to increase the hydrophobicity of the sidechains based on our understanding that the water occupancy inside cryptophane cavity is correlated with binding affinity. The water occupancy should be impacted by the sidechains and their hydrophobicity. The second consideration is to reduce the interior volume of the cryptophane. From simulation studies of cryptophane structural adjustments, the cryptophane may have increased affinity with smaller interior volume. Previously considered cryptophanes with a linker of two methylenes are a starting point for such reduction of volume while still accomodating Xe. In this work, we use ABF method to estimate the Xe binding affinities of proposed cryptophanes and investigate their water occupancy and orientations of sidechains.

6.2. Introduction

Proton-based magnetic resonance imaging (MRI) is a widely used technique for the detection and diagnosis of various disease states. For applications involving MRI and biosensing, issues involving sensitivity can be overcome by using heteronuclei such as ^{13}C, ^{15}N, ^{3}H, and ^{129}Xe.[135] Of these, the ^{129}Xe isotope is particularly attractive for applications in biomolecular imaging, due its relatively high natural abundance of 26.4%, its solubility in

many organic and aqueous solvents[136], and its large, polarizable electron cloud. This latter feature imparts it great sensitivity to its molecular environment, resulting in a nearly 300 ppm 129Xe NMR chemical shift window.[137] Moreover, 129Xe can be hyperpolarized (hp) via spin exchange optical pumping,[138] which allows for an increase in the ^{129}Xe NMR signal by more than 10000-fold.[139] These advantages of using ^{129}Xe-based biosensors have been readily exploited via the design of xenon hosts that generate a unique ^{129}Xe NMR chemical shift upon association with their target. The best-studied of these xenon hosts are the small, cage-like molecules known as cryptophanes. There is currently a lack of knowledge as to what makes a potent host agent and high affinity ^{129}Xe binder. Recent simulation studies have suggested that the presence of water molecules in the hydrophobic cavity of the cryptophane is key in governing the free energy of association between cryptophane and xenon. Specifically, it was seen that cryptophane variants with the lowest average number of interior water molecules possessed the highest xenon binding affinities.[107] Thus, one could consider synthesizing a biosensor with higher affinity for xenon by reducing the propensity of water molecules to occupy the cage interior. One method by which this could be achieved is placing the ionized solubilizing moieties farther away from the cryptophane cage, where they would less readily attract water molecules to the cryptophane core. Previously, a tris-(triazole propionic acid) cryptophane-A derivative was to found to have a xenon association constant of 17 000 M-1 at 293K.[69] We explore modification of xenon affinity by using the longer hexanoic acid or undecanoic acid solubilizing groups. We also consider variation of aliphatic side chains not associated with solubilization.

6.3. Methods

6.3.1. Molecular models.

All cryptophanes considered are the derivitives of TAAC and TTPC[69]. For the derivitives of TAAC, three of them are with methyl sidechain get substituted by ethyl, propyl and isopropyl group (Figure 6.1). Two cryptophanes have smaller interior volumes (TAAC-221 and TAAC-111). Derivitives of TTPC with longer sidechains are also investigated (Figure

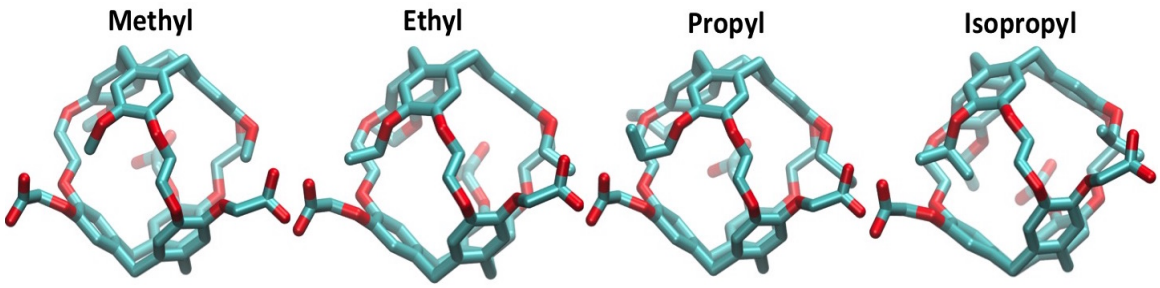


Figure 6.1: Renderings of TAAC and TAAC derivitives with methyl get substituted by ethyl, propyl and isopropyl

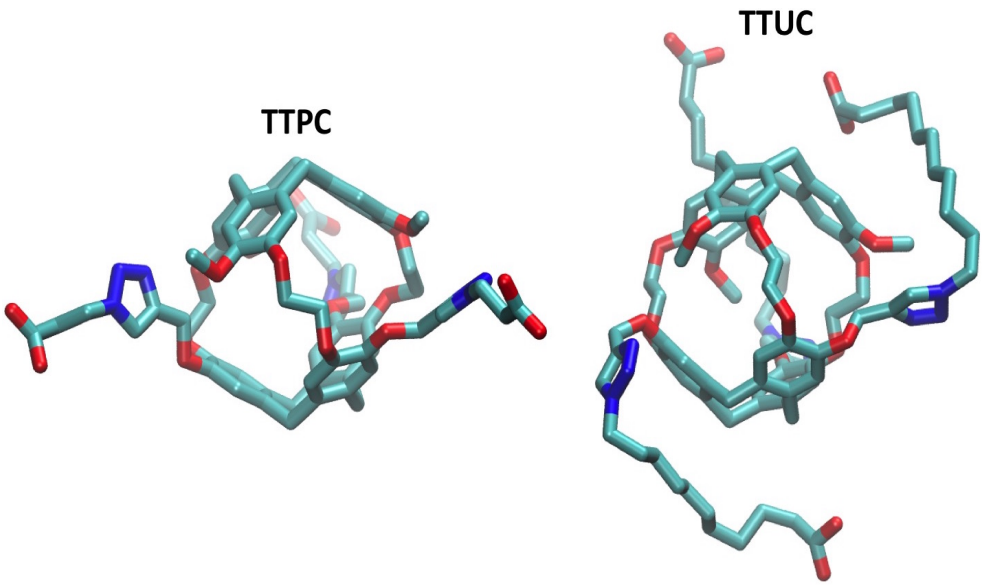


Figure 6.2: Renderings of TTPC and TTPC with 10 more carbons longer in triazole propionic acid sidechain

6.2). Structures of these molecules are generated by modifying the structure of TAAC and TTPC used in previous study[107]. The molfacture module in VMD is used to generate the structures, which are then energetically optimized in NAMD. In each simulated systems, 3353 TIP3P[140] water molecules are added which create 48Å ×48Å ×48Å water boxes with concentration of solute is ∼0.02M and 3 sodium ions are added to neutralize the charges in each system. Parameters of bonded interaction used in simulations come from the AMBER94[141] force field. Additional bonded parameters come from Hessian matrices calculated based on optimized geometries from Gaussian. For non-bonded interaction, van der Waals interactions are modeled by Lennard-Jones potentials and parameters of cryptophane molecules and Xe come from OPLS force field[127, 128] while electrostatic interactions are modeled by coulombic interaction with partial charges in cryptophane backbone come from developed charge set and the charges in sidechains come from R.E.D. server[130] where charges are fitted to reproduce quantum calculated electrostatic potential around each atom and atoms at junction position are edited so as to construct integer charge for the overall molecule.

6.3.2. Molecular dynamics simulation.

The NAMD[108] package is applied in this work to perform simulations. As in prior molecular dynamics simulations, 50,000 steps of conjugate gradient energy minimization are performed to eliminate high-energy initial configurations after solvation the cryptophane molecules. In this work, step size in the simulation is 2 femtosecond. The 20 ns trajectories in the NPT ensemble simulations are performed for each system, where we use Langevin dynamics method [54] to maintain the system temperature to be 300 K and pressure to be 1.0 atm. The coupling coefficient for the temperature constraint is 1/ps and the period is 100.0 fs, the decay constant is 50.0 fs, and the corresponding temperature is 300.0K for Langevin piston using the Nose-Hoover method. Electrostatic interactions are calculated through Particle Mesh Ewald (PME) method [53] with 1.0Å grid spacing. Lennard-Jones interactions are computed with cutoff is 12.0Å. SHAKE/RATTLE[52] algorithms are ap-

plied to constrain the covalent bonds linked with hydrogen. The last frames of trajectories generated from NPT simulations will be used for analysis and provide initial conformations to later ABF simulations.

$$\frac{dA(\xi)}{d\xi} = \left\langle \frac{\partial V(x)}{\partial \xi} - \frac{1}{\beta}\frac{\partial ln|J|}{\partial \xi} \right\rangle_{\xi^*} = -\langle F_\xi \rangle_{\xi^*} \tag{6.1}$$

6.3.3. Adaptive biasing force simulation

In this work, ABF method is used to calculate the free energy profile of the cryptophane-Xe system with respect to the distance (ξ) between center mass of cryptophane backbone and center mass of Xe and generate trajectories that having ξ range from 0 to 15.0Å. In ABF method, one accumulated biasing force along the direction of ξ is applied to the system so as to cancel out net force along this direction thereby the system will do the diffusion and having uniform distribution with respect to ξ. The formula of estimating biased force from integration can be seen in equ(6.1), where A(ξ) is the free energy of the system at ξ, V(x) is the potential energy in original Cartesian coordinates, β is the temperature factor, |J | is the Jacobian determinant of transforming Cartesian coordinates to generalized coordinates which includes ξ. The bracket represent the ensemble average that having ξ constrained to ξ^*. In implementation, the range of ξ we are interested in is divided into small bins and the forces in each bin will be considered to be same, and in prior to apply the biasing force a threshold number of samples in each bin are required to be reached due to force fluctuate significantly initially otherwise will not representative for that ξ state and bias the system to wrong sampling. For all four cryptophane systems, overall range of ξ (0.0 ∼15.0Å) are divided into three windows (0.0 ∼5.0Å, 5.0 ∼10.0Å and 10.0 ∼15.0Å) and every one of them represent one separate ABF simulation and having ξ constrained to corresponding ξ distance range so as to assure ABF calculation converge sooner. The bin size is chosen to be 0.05Å and the threshold in each bin is chosen to be 30000. In each ABF simulation run, 20 ∼40ns simulation have been done so as to assure the uniform sampling of ξ and thereby

provide the accuracy of calculated biasing force and PMF.

6.4. Results and discussions

6.4.1. Potential of Mean Force results.

TTAC deriviatives with sidechain get substituted. Free energy profile (potential of mean force, PMF) of cryptophane and Xe systems were calculated with respect to ξ, the distance between center mass of cryptophane and Xe(figure 6.3). Although those modfied cryptophanes have different sidechains, its clear that PMFs of the all derivatives have similar shapes: there are three significant states along the reaction coordinate (ξ), a deep global minima with Xe bounded at the center of cavity (ξ $\sim$0.1Å), one energy maxima with Xe at the edge of cavity (ξ $\sim$5.0Å) and a plateau with Xe is unbounded and isolated from cavity (ξ $\sim$15.0Å). In each case, the PMF is calculated with respect to the state that with Xe unbound. From table 6.1, it shows that the depths of the global minima (ΔWmin) which equal to the free energy difference between the Xe bound and unbound states of all cryptophanes are TAAC-ethyl (-11.72 kcal/mol) > TAAC (-11.35 kcal/mol) > TAAC-isopropyl (-11.00 kcal/mol) > TAAC-propyl (-10.92 kcal/mol). The height of the energy barrier (ΔWmax) of these cavities are between 1.4 to 3 kcal/mol and the ordering is

Cryptophane	ΔG_{cal} (kcal/mol)	ΔW_{min} (kcal/mol)	ΔW_{max} (kcal/mol)
TAAC	-6.44	-11.35	2.76
TAAC-ethyl	-6.96	-11.72	1.41
TAAC-propyl	-6.20	-10.92	1.84
TAAC-isopropyl	-6.26	-11.00	1.87
TTUC	-7.27	-12.19	2.09
TAAC-221	-7.15	-12.09	2.65
TAAC-111	-5.07	-10.48	12.31

Table 6.1: Calculated binding affinities of derivitives of cryptophane TAAC and TTPC to Xe, calculated PMF minima (ΔW_{min}) and calculated PMF maxima (ΔW_{max}).

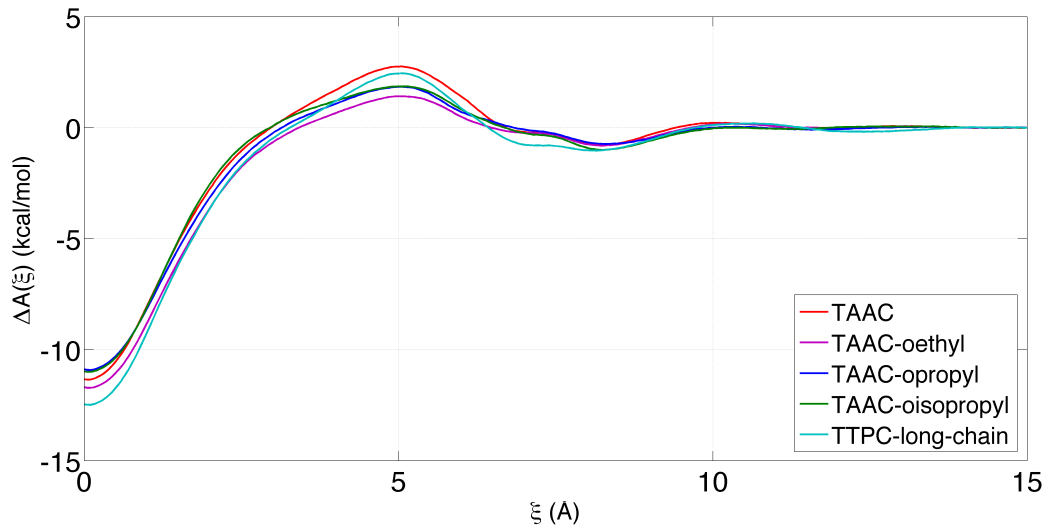


Figure 6.3: Free energy profile of Xe along the coordinate of entering cavity center (ξ)

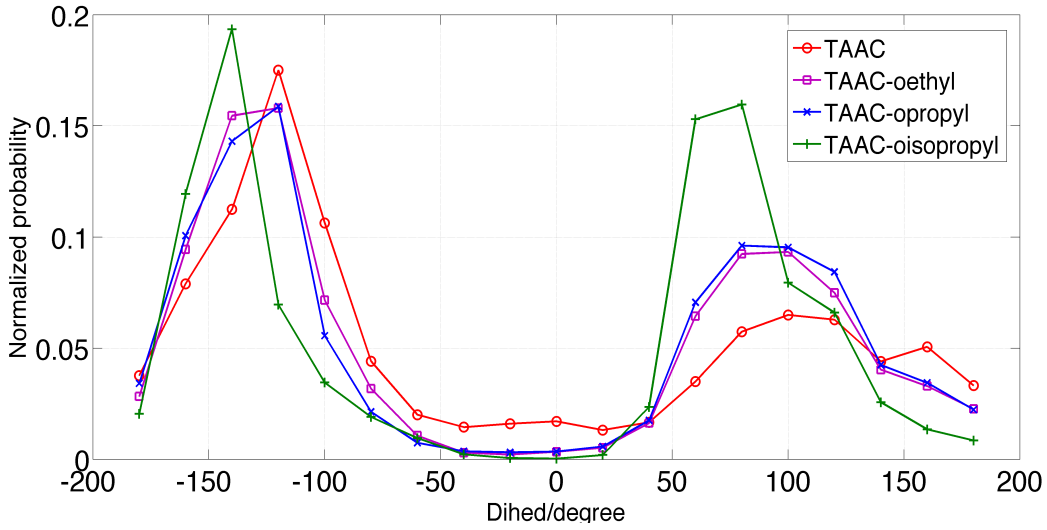


Figure 6.4: Dihedral distribution of the rotation of alkyl group for TAAC derivitives

TAAC > TAAC-isopropyl > TAAC-propyl > TAAC-ethyl. This result is a little surprising in that for all TAAC derivitives their sidechains are larger than TAAC and should have steric effect to keep Xe entering the cavity and the height for Xe to enter should be higher. Once the sidechain is larger than the methyl, the alky group can no longer always orient to the interior of the cavity. This means ethyl, propyl and isopropyl already orient to the exterior of the cryptophane cavity when Xe is not present to enter the cavity from the dihedral distribution anaylsis of the rotation to rotate alky group (Figure 6.4). Therefore, the structure adjustment for these derivatives from the prospect of sidechains are less than those required with TAAC, which leads to the lower energy barriers to accomodate Xe with these larger alkane sidechains. Additionally, there is also one small shoulder (ξ ~7.0Å) and one local minima (ξ ~8.3Å) outside cavity along the reaction coordinate (ξ). These suggest that Xe would still associate transiently with cryptophane with Xe exterior to the cavity.

TTPC with 11 carbon sidechain (TTUC) TTUC is also investigated and its PMF curve is showed as well (6.3). Given our previous work [107] suggest that having sidechain with charged group further away from cryptophane center could help to "dry" the cryptophane and increase the Xe binding affinity accordingly. Here a cryptophane with fatty acid in its triozale sidechain is considered, which is 10 carbons longer than the original TTPC. The PMF shows that this cryptophane also has the similar shape as other cryptophanes, but with a relatively much deeper ΔW_min and higher ΔW_max. In addition, we can observe the exterior local minima (ξ ~8.3Å) of this cryptophane is much broader than all other cryptophanes, which suggest the Xe have relatively strong interaction with the large sidechains.

TAAC with smaller interior volume. TAAC-221 and TAAC-111 are studied here to see if smaller volume cryptophane can posses higher Xe binding affinity. The cryptophane with each linker (TAAC-111) one methylene shorter than TAAC has significantly different shape compared to TAAC6.5, its energy barrier is as high as 12.31 kcal/mol for translocation

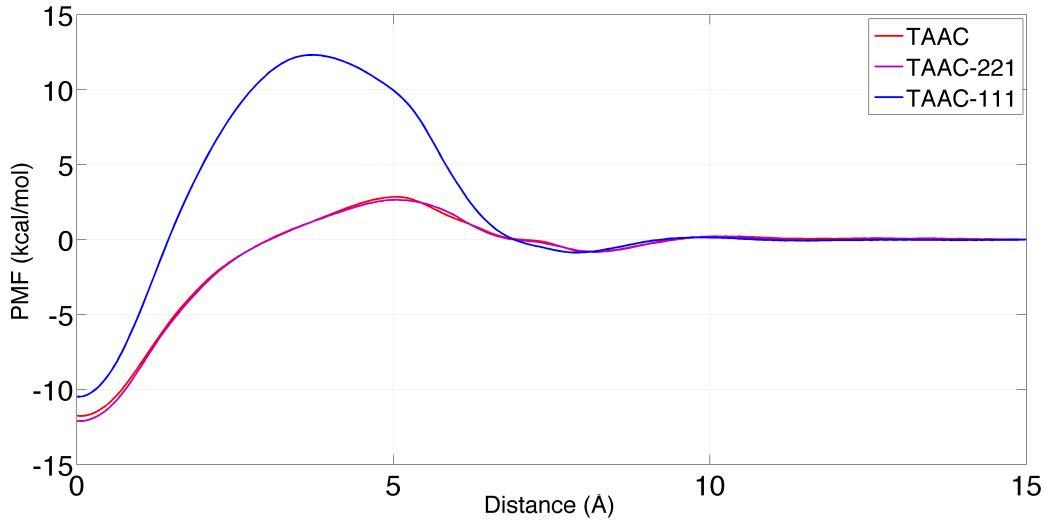


Figure 6.5: Potential of Mean Force of of TAAC derivitives with smaller volume

inside the cavity (ξ=3.7Å) which suggest Xe would extremely unlikely to enter the cavity. The energy barrier for cryptophane to accomondate Xe involves not only sidechain re-orientation but also a dramatic stretching of the linkers to open the portal to the cavity. The cryptophane TAAC-221 which has only one one-methylene linker has a PMF with very similar features as TAAC. TAAC-221 has deeper energy minima depth and similar energy barrier compared with TAAC. These suggest that this slight smaller interior volume of TAAC-221 could lead to enhanced affinity relative th TAAC.

$$K_{association} = 4\pi \int_0^{r_d} r^2 e^{-\beta U(r)} dr \tag{6.2}$$

$$K_0 = \frac{1}{1660\text{Å}^3} K_{association} \tag{6.3}$$

$$\Delta G_{bind} = -RTlnK_0 \tag{6.4}$$

6.4.2. Binding affinity calculation.

The potential of mean force of Xe-cryptophane system can be used to estimate the equilibrium association constant by using equation6.2. The standard state equilibrium constant and binding affinity can be calculated then from equations 6.3 and equation 6.4.

From Table 1, it is shown that the binding affinities are correlated with the energy minima (Δ W$_m in$): the deeper the minima, the higher the binding affinity to Xe. From the table 6.1, we can see the TAAC-oethyl and TAAC-221 are the two promising cavities that have higher binding affinity than TAAC, whose binding affinities are 0.52 kcal/mol and 0.71 kcal/mol higher. TTUC is also a promising cavity that has 0.83 kcal/mol higher Xe binding affinity. The reason that the methyl in TAAC subsitituted by ethyl in TAAC-oethyl is favorable for Xe binding is that the ethyl would impede water association with the cavity and keep the cryptophane interior dry while also accommodating Xe. As the alkyl sidechain gets larger the effect diminishes as the larger alkyl groups impinging on the guest Xe. TAAC-221 has reduced volume cavity, leading to a slight increase in Xe affinity. TTUC has improved affinity due to the fact that its charged acid (ionizable group) is much further away from cavity center than TTPC.

6.5. Conclusions

Through extensive ABF simulations, free energy profiles of processes of Xe entering several proposed cryptophane cavities are calculated for distances between Xe and the center of mass of the cryptophane (ξ) from 0.0 Å to 15.0 Å. From calculated free energy profiles (PMF curves), it is shown that all systems share the same shape that Xe will have to overcome one energy barrier to enter an energy well. For most cryptophanes, the energy barrier is relatively small (2-4 k$_B$T) while the energy well is deep (16-20 k$_B$T). TAAC-111 has the smallest interior volume, and its structure changes dramatically to encapsulate Xe, resulting in the largest energy barrier for Xe entry. By making use of free energy profile with respect to ξ, binding affinity of Xe are calculated for proposed cryptophanes. We

conclude that cryptophane with slightly smaller volume (one of the linker is one methylene shorter), a slightly larger alkyl group to substitute methyl group (ethyl) and longer charged sidechain would all enhance the Xe binding affinity relative to TAAC.

BIBLIOGRAPHY

[1] C. J. Pedersen. Cyclic polyethers and their complexes with metal salts. *J. Am. Chem. Soc.*, 89(10):2495–2496, May 1967.

[2] Donald J. Cram and Jane M. Cram. Host-Guest Chemistry. *Science*, 183(4127):803–809, 1974.

[3] Yun Yan, Jianbin Huang, and Ben Zhong Tang. Kinetic trapping a strategy for directing the self-assembly of unique functional nanostructures. *Chemical communications Chemcomm /*, 52(80):11870,11884, 2016.

[4] Guang Yang, Libin Wu, Guosong Chen, and Ming Jiang. Precise protein assembly of array structures. *Chem. Commun.*, 52:10595–10605, 2016.

[5] Fan Zhang, Yue Sun, Demei Tian, Weon Sup Shin, Jong Seung Kim, and Haibing Li. Selective molecular recognition on calixarene-functionalized 3d surfaces. *Chem. Commun.*, 52:12685–12693, 2016.

[6] Joan Teyssandier, Steven De Feyter, and Kunal S. Mali. Host-guest chemistry in two-dimensional supramolecular networks. *Chem. Commun.*, 52:11465–11487, 2016.

[7] Trevor Douglas and Mark Young. Host-guest encapsulation of materials by assembled virus protein cages. *Nature*, 393(6681):152, 1998.

[8] Patricia Horcajada, Christian Serre, María Vallet-Regí, Muriel Sebban, Francis Taulelle, and Gérard Férey. Metal–organic frameworks as efficient materials for drug delivery. *Angewandte chemie*, 118(36):6120–6124, 2006.

[9] Yincheng Chang, Chenxi Hou, Jingli Ren, Xiaoting Xin, Yuxin Pei, Yuchao Lu, Shoupeng Cao, and Zhichao Pei. Multifunctional supramolecular vesicles based on the complex of ferrocenecarboxylic acid capped pillar[5]arene and a galactose derivative for targeted drug delivery. *Chem. Commun.*, 52:9578–9581, 2016.

[10] Joo Hee Jang, Hoyeon Lee, Amit Sharma, Sang Min Lee, Tae Hoon Lee, Chulhun Kang, and Jong Seung Kim. Indomethacin-guided cancer selective prodrug conjugate activated by histone deacetylase and tumour-associated protease. *Chem. Commun.*, 52:9965–9968, 2016.

[11] Jianxiang Zhang and Peter X Ma. Polymeric core–shell assemblies mediated by host–guest interactions: versatile nanocarriers for drug delivery. *Angewandte Chemie International Edition*, 48(5):964–968, 2009.

[12] Roseita Esfand and Donald A Tomalia. Poly (amidoamine)(pamam) dendrimers: from biomimicry to drug delivery and biomedical applications. *Drug discovery today*, 6(8):427–436, 2001.

[13] Messaoud Benounis, Nicole Jaffrezic-Renault, J-P Dutasta, K Cherif, and Adnane Abdelghani. Study of a new evanescent wave optical fibre sensor for methane detection based on cryptophane molecules. *Sensors and Actuators B: Chemical*, 107(1):32–39, 2005.

[14] Cédric Boulart, Matthew C Mowlem, Douglas P Connelly, Jean-Pierre Dutasta, and Christopher R German. A novel, low-cost, high performance dissolved methane sensor for aqueous environments. *Optics express*, 16(17):12607–12617, 2008.

[15] Carsten Koopmans and Helmut Ritter. Color change of n-isopropylacrylamide copolymer bearing reichardts dye as optical sensor for lower critical solution temperature and for host- guest interaction with β-cyclodextrin. *Journal of the American Chemical Society*, 129(12):3502–3503, 2007.

[16] Franz L Dickert and Alexander Haunschild. Sensor materials for solvent vapor detection?donor–acceptor and host–guest interactions. *Advanced Materials*, 5(12):887–895, 1993.

[17] Kristin Bartik, Michel Luhmer, Jean Pierre Dutasta, André Collet, and Jacques Reisse. 129Xe and 1H NMR study of the reversible trapping of xenon by cryptophane-A in organic solution. *Journal of the American Chemical Society*, 120(4):784–791, 1998.

[18] Laurent Garel, Jean-Pierre Dutasta, and Andre Collet. Complexation of Methane and Chlorofluorocarbons by Cryptophane-A in Organic Solution. *Angewandte*, 32(8), 1993.

[19] J. Canceill, L. Lacombe, and A. Collet. *J. Am. Chem. Soc*, 107:6993–6996, 1985.

[20] J. Canceill, L. Lacombe, and A. Collet. *J. Am. Chem. Soc*, 108:4230–4232, 1986.

[21] G. Haberhauer, S. Woitschetzki, and H. Bandmann. *Nat. Commun*, 5:3542, 2014.

[22] D. R. Jacobson, N. S. Khan, R. Colle, R. Fitzgerald, L. Laureano-Perez, Y. B. Bai, and I. J. Dmochowski. *Proc. Natl. Acad. Sci. U. S. A*, 108:10969–10973, 2011.

[23] Thierry Brotin, Dominique Cavagnat, Patrick Berthault, Roland Montserret, and Thierry Buffeteau. Water-soluble molecular capsule for the complexation of cesium and thallium cations. *The journal of physical chemistry. B*, 116(35):10905–14, sep 2012.

[24] Thierry Brotin, Roland Montserret, Aude Bouchet, Dominique Cavagnat, Mathieu Linares, and Thierry Buffeteau. High affinity of water-soluble cryptophanes for cesium cations. *The Journal of organic chemistry*, 77(2):1198–201, jan 2012.

[25] Patrick Berthault Dominique Cavagnat Thierry Brotin, Sylvie Goncalves and Thierry Buffeteau. Influence of the cavity size of water-soluble cryptophanes on their binding properties for cesium and thallium cations. *J. Phys. Chem. B*, 117:12593–12601, 2013.

[26] M M Spence, S M Rubin, I E Dimitrov, E J Ruiz, D E Wemmer, a Pines, S Q Yao, F Tian, and P G Schultz. Functionalized xenon as a biosensor. *Proceedings of the National Academy of Sciences of the United States of America*, 98(19):10654–7, 2001.

[27] K. Ruppert, J. R. Brookeman, K. D. Hagspiel, B. Driehuys, and J. P. Mugler. *NMR Biomed.*, 13:220–228, 2000.

[28] G. Duhamel, P. Choquet, E. Grillon, L. Lamalle, J. L. Leviel, A. Ziegler, and A. Constantinesco. *Magn. Reson. Med*, 46:208–212, 2001.

[29] S. D. Swanson, M. S. Rosen, K. P. Coulter, R. C. Welsh, and T. E. Chupp. *Magn. Reson. Med*, 42:1137–1145, 1999.

[30] S. Korchak, W. Kilian, and L. Mitschang. *Chem. Commun*, 51:1721–1724, 2015.

[31] D. Raftery. In G. A. Webb and Elsevier, editors, *Annual Reports on Nmr Spectroscopy*, pages vol. 57, 205–270. Academic Press Inc, San Diego, 2006.

[32] Jennifer Chambers, Aru Hill, Julie Aaron, Zhaohui Han, David Christianson, Nicholas Kuzma, and Ivan Dmochowski. Cryptophane Xenon-129 Nuclear Magnetic Resonance Biosensors Targeting Human Carbonic Anhydrase. *Journal of the American Chemical Society*, 131(14):563–569, 2008.

[33] David R Jacobson, Najat S Khan, Ronald Collé, Ryan Fitzgerald, Lizbeth Laureano-Pérez, Yubin Bai, and Ivan J Dmochowski. Measurement of radon and xenon binding to a cryptophane molecular host. *Proceedings of the National Academy of Sciences of the United States of America*, 108(27):10969–10973, 2011.

[34] G. Huber, T. Brotin, L. Dubois, H. Desvaux, J. P. Dutasta, and P. Berthault. *J. Am. Chem. Soc.*, 128:6239–6246, 2006.

[35] P. A. Hill, Q. Wei, T. Troxler, and I. J. Dmochowski. J. am. chem. soc. 131:3069–3077, 2009.

[36] P. A. Hill, Q. Wei, R. G. Eckenhoff, and I. J. Dmochowski. *J. Am. Chem. Soc*, 129:11662–11662, 2007.

[37] Yubin Bai, P Aru Hill, and Ivan J Dmochowski. Utilizing a water-soluble cryptophane with fast xenon exchange rates for picomolar sensitivity nmr measurements. *Analytical chemistry*, 84(22):9935–9941, 2012.

[38] Nobuyuki Hamada and Haruyuki Ogino. Food safety regulations: what we learned from the fukushima nuclear accident. *Journal of environmental radioactivity*, 111:83–99, 2012.

[39] Nobuyuki Hamada, Haruyuki Ogino, and Yuki Fujimichi. Safety regulations of food and water implemented in the first year following the fukushima nuclear accident. *Journal of radiation research*, 53(5):641–671, 2012.

[40] M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon. *Biophys. J.*, 72:1047–1069, 1997.

[41] D. Hamelberg and J. A. McCammon. *J. Am. Chem. Soc*, 126:7683–7689, 2004.

[42] Eric Darve and Andrew Pohorille. Calculating free energies using average force. *The Journal of Chemical Physics*, 115(20):9169–9183, 2001.

[43] Eric Darve, David Rodriguez-Gomez, and Andrew Pohorille. Adaptive biasing force method for scalar and vector free energy calculations. *The Journal of Chemical Physics*, 128(14):144120, 2008.

[44] Jerome Hénin, Giacomo Fiorin, Christophe Chipot, and Michael L. Klein. Exploring multidimensional free energy landscapes using time-dependent biases on collective variables. *Journal of Chemical Theory and Computation*, 6(1):35–47, 2010.

[45] Jérôme Hénin and Christophe Chipot. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *The Journal of chemical physics*, 121(7):2904–2914, 2004.

[46] Jonathan W Steed, David R Turner, and Karl Wallace. *Core concepts in supramolecular chemistry and nanochemistry*. John Wiley & Sons, 2007.

[47] Michael D Pluth and Kenneth N Raymond. Reversible guest exchange mechanisms in supramolecular hostguest assemblies. *Chem. Soc. Rev.*, 36(2):161–171, 2007.

[48] Ferey G. Serre, C., Mellot-Draznieks, C., Surble, S., Auderbrand, N., Filinchu, Y. Role of Solvent-Host Interactions That Lead to Very Large Swelling of Hybrid Frameworks. *Science*, 315(5820):1828–1831, 2007.

[49] Igor Leontyev and Alexei Stuchebrukhov. Accounting for electronic polarization in non-polarizable force fields. *Physical Chemistry Chemical Physics*, 13(7):2613–2626, 2011.

[50] IV Leontyev and AA Stuchebrukhov. Electronic polarizability and the effective pair potentials of water. *Journal of chemical theory and computation*, 6(10):3153–3161, 2010.

[51] A. A. Clifford, P. Gray, and N. Platts. *Journal of the Chemical Society*, 73:381–382, 1977.

[52] Universite Libre De Bruxelles. Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *Journal of computational physics*, 23:327–341, 1977.

[53] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089, 1993.

[54] Scott E. Feller, Yuhong Zhang, Richard W. Pastor, and Bernard R. Brooks. Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics*, 103(11):4613, 1995.

[55] Glenn J Martyna, Douglas J Tobias, and Michael L Klein. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*, 101(5):4177–4189, 1994.

[56] T. Young, R. Abel, B. Kim, B. J. Berne, and R. A. Friesner. *Proc. Natl. Acad. Sci. U S A*, 104:808–813, 2007.

[57] L. Garel, B. Lozach, J. P. Dutasta, and A. Collet. *J. Am. Chem. Soc*, 115:11652–11653, 1993.

[58] J. Canceill, L. Lacombe, and A. Collet. *J. Chem. Soc. Chem. Commun*, pages 219–221, 1987.

[59] T. Traore, G. Clave, L. Delacour, N. Kotera, P. Y. Renard, A. Romieu, P. Berthault, C. Boutin, N. Tassali, and B. Rousseau. *Chem. Commun*, 47:9702–9704, 2011.

[60] M. Linares H. gren D. Cavagnat A. Bouchet, T. Brotin and T. Buffeteau. *Journal of Organic Chemistry*, 76:1372–1383, 2011.

[61] O. Taratula and I. J. Dmochowski. *Current Opinion in Chemical Biology*, 14:97–104, 2010.

[62] M. M. Spence, S. M. Rubin, I. E. Dimitrov, E. J. Ruiz, D. E. Wemmer, A. Pines, S. Q. Yao, F. Tian, and P. G. Schultz. *Proc. Natl. Acad. Sci. U S A*, 98:10654–10657, 2001.

[63] S. M. Rubin, S. Y. Lee, E. J. Ruiz, A. Pines, and J. Mol D. E. Wemmer. *J. Mol. Biol*, 322:425–440, 2002.

[64] T. Brotin and J. P. Dutasta. *Chemical Reviews*, 109:88–130, 2009.

[65] A. I. Joseph, S. H. Lapidus, C. M. Kane, and K. T. Holman. *Angewandte Chemie International Edition*, 54:1471–1475, 2015.

[66] B. S. Kim, Y. H. Ko, Y. Kim, H. J. Lee, N. Selvapalam, H. C. Lee, and K. Kim. *Chem. Commun*, pages 2756–2758, 2008.

[67] A. Collet, J. P. Dutasta, B. Lozach, and J. Canceill. *Top. Curr. Chem*, 165:103–129, 1993.

[68] R. M. Fairchild, A. I. Joseph, K. T. Holman, H. A. Fogarty, T. Brotin, J. P. Dutasta, C. Boutin, G. Huber, and P. Berthault. *J. Am. Chem. Soc*, 132:15505–15507, 2010.

[69] P Aru Hill, Qian Wei, Thomas Troxler, and Ivan J Dmochowski. Substituent effects on

xenon binding affinity and solution behavior of water-soluble cryptophanes. *Journal of the American Chemical Society*, 131(8):3069–77, 2009.

[70] P. Aru Hill, Qian Wei, Roderic G. Eckenhoff, and Ivan J. Dmochowski. Thermodynamics of xenon binding to cryptophane in water and human plasma. *Journal of the American Chemical Society*, 129(30):9262–9263, 2007.

[71] K. N. Houk, A. G. Leach, S. P. Kim, and X. Y. Zhang. *Angew. Chem. -Int. Edit*, 42:4872–4897, 2003.

[72] P. P. Wanjari, B. C. Gibb, and H. S. Ashbaugh. *J. Chem. Phys.*, 139(234502), 2013.

[73] J. CostanteCrassous, T. J. Marrone, J. M. Briggs, J. A. McCammon, and A. Collet. *J. Am. Chem. Soc*, 119:3818–3823, 1997.

[74] P. D. Kirchhoff, M. B. Bass, B. A. Hanks, J. M. Briggs, A. Collet, and J. A. McCammon. *J. Am. Chem. Soc*, 118:3237–3246, 1996.

[75] C. C. Roberts and C. e. A. Chang. *J. Chem. Theory Comput*, 9:2010–2019, 2013.

[76] C. Chipot and D. A. Pearlman. *Mol. Simul.*, 28:1–12, 2002.

[77] P. Kollman Chemical Reviews. *Chemical Reviews*, 93:2395–2417, 1993.

[78] R. W. Zwanzig. J. chem. phys. 22:1420–1426, 1954.

[79] D. L. Beveridge and F. M. Dicapua. *Annual Review of Biophysics and Biophysical Chemistry*, 18:431–492, 1989.

[80] T. P. Straatsma and J. A. McCammon. *Annu. Rev. Phys. Chem.*, 43:407–435, 1992.

[81] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. *J. Comput. Chem.*, 26:1781–1802, 2005.

[82] D. Cavagnat, T. Brotin, J. L. Bruneel, J. P. Dutasta, A. Thozet, M. Perrin, and F. Guillaume. *J. Phys. Chem. B*, 108:5572–5581, 2004.

[83] W. L. Jorgensen, D. S. Maxwell, and J. TiradoRives. *J. Am. Chem. Soc*, 118:11225–11236, 1996.

[84] W. L. Jorgensen and J. Tiradorives. *J. Am. Chem. Soc*, 110:1657–1666, 1988.

[85] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, E. S., J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski,

J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople. Gaussian 98 (revision a.10). *Gaussian Inc. Pittsburgh, PA,*, 2001.

[86] F. Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, and P. Cieplak. *Phys. Chem. Chem. Phys*, 12:7821–7839, 2010.

[87] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. *J. Am. Chem. Soc*, 118:2309–2309, 1996.

[88] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. J. chem. phys. 79:926–935, 1983.

[89] W. Humphrey, A. Dalke, and K. Schulten. J. mol. graph. 14:33–38, 1996.

[90] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. *J. Comput. Phys*, 23:327–341, 1977.

[91] D. Frenkel and B. Smit. Understanding molecular simulation, academic press, waltham, ma (usa), 2nd edn. 2001.

[92] T. C. Beutler, A. E. Mark, R. C. Vanschaik, P. R. Gerber, and W. F. Vangunsteren. *Chem. Phys. Lett*, 222:529–539, 1994.

[93] M. Zacharias, T. P. Straatsma, and J. A. McCammon. J. chem. phys. 100:9025–9031, 1994.

[94] J. Henin, G. Fiorin, C. Chipot, and M. L. Klein. J. chem. theory comput. 6:35–47, 2010.

[95] A. Bennaim and Y. Marcus. *J. Chem. Phys.*, 81:2016–2027, 1984.

[96] E. Wilhelm, R. Battino, and R. J. Wilcock. *Chemical Reviews*, 77:219–262, 1977.

[97] P. Scharlin, R. Battino, E. Silla, I. Tunon, and J. L. Pascual-Ahuir. *Pure Appl. Chem.*, 70:1895–1904, 1998.

[98] N. T. Skipper and G. W. Neilson. J. phys. condes. matter. 1:4141–4154, 1989.

[99] O. Taratula, P. A. Hill, N. S. Khan, P. J. Carroll, and I. J. Dmochowski. *Nat. Commun*, 1:148, 2010.

[100] P. Berthault, H. Desvaux, T. Wendlinger, M. Gyejacquot, A. Stopin, T. Brotin, J. P. Dutasta, and Y. Boulard. *Chem. -Eur. J*, 16:12941–12946, 2010.

[101] Peter Kollman. Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.*, 93:2395–2417, 1993.

[102] G.M. Torrie and J.P. Valleau. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199, 1977.

[103] Marc Souaille and Benoît Roux. Extension to the weighted histogram analysis method: Combining umbrella sampling with free energy calculations. *Computer Physics Communications*, 135(1):40–57, 2001.

[104] L. Garel, J. P. Dutasta, and A. Collet. *Angew. Chem*, 32:1169–1171, 1993.

[105] Olena Taratula, P. Aru Hill, Najat S. Khan, Patrick J. Carroll, and Ivan J. Dmochowski. Crystallographic observation of 'induced fit' in a cryptophane hostguest model system. *Nature Communications*, 1(9):148, 2010.

[106] P D Kirchhoff, M B Bass, B A Hanks, J M Briggs, Andre Collet, J A Mccammon, Ecole Normale, and Superieure De Lyon. Structural Fluctuations of a Cryptophane Host : A Molecular Dynamics Simulation. *Journal of the American Chemical Society*, (11):3237–3246, 1996.

[107] Lu Gao, Wenhao Liu, One-sun Lee, Ivan J Dmochowski, and Jeffery G Saven. Xe Affinities of Water- Soluble Cryptophanes and the Role of Confined Water. *Chemical science*, 6:7238–7248, 2015.

[108] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kalé, and Klaus Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.

[109] Emmerich Wilhelm, Rubln Battino, and Robert J Wilcock. Low-pressure solubility of gases in liquid water. *Chemical Reviews*, 77:219–262, 1977.

[110] Michelle E. Weber George W. Gokel, W. Matthew Leevy. Crown ethers: Sensors for ions and molecular scaffolds for materials and biological models. *Chem. Rev.*, 104:2723–2750, 2004.

[111] Richard H Fish and Gérard Jaouen. Bioorganometallic chemistry: structural diversity of organometallic complexes with bioligands and molecular recognition studies of several supramolecular hosts with biomolecules, alkali-metal ions, and organometallic pharmaceuticals. *Organometallics*, 22(11):2166–2177, 2003.

[112] Mingming Zhang, Donghua Xu, Xuzhou Yan, Jianzhuang Chen, Shengyi Dong, Bo Zheng, and Feihe Huang. Self-healing supramolecular gels formed by crown ether based host–guest interactions. *Angewandte Chemie*, 124(28):7117–7121, 2012.

[113] Peter D Southon, Lang Liu, Elizabeth A Fellows, David J Price, Gregory J Halder, Karena W Chapman, Boujemaa Moubaraki, Keith S Murray, Jean-François Létard, and Cameron J Kepert. Dynamic interplay between spin-crossover and host- guest function in a nanoporous metal- organic framework material. *Journal of the American Chemical Society*, 131(31):10998–11009, 2009.

[114] Chang-Ping Zhang, Ping Gu, Jun Zhao, Dong Zhang, and Yue Deng. Research on the treatment of liquid waste containing cesium by an adsorption-microfiltration process with potassium zinc hexacyanoferrate. *J Hazard Mater*, 167(1-3):1057–62, Aug 2009.

[115] Mohammad Moayed Mohammad Ghannadi-Maragheh Hossein Faghihian, Mozhgan Iravani. A novel polyacrylonitrile–zeolite nanocomposite to clean cs and sr from radioactive waste. *Environ Chem Lett*, 11:277–282, 2013.

[116] Dongjiang Yang, Sarina Sarina, Huaiyong Zhu, Hongwei Liu, Zhanfeng Zheng, Mengxia Xie, Suzanne V Smith, and Sridhar Komarneni. Capture of radioactive cesium and iodide ions from water by using titanate nanofibers and nanotubes. *Angewandte Chemie International Edition*, 50(45):10594–10598, 2011.

[117] Adavan Kiliyankil Vipin, Baiyang Hu, and Bunshi Fugetsu. Prussian blue caged in alginate/calcium beads as adsorbents for removal of cesium ions from contaminated water. *Journal of hazardous materials*, 258:93–101, 2013.

[118] Md Rabiul Awual, Tsuyoshi Yaita, Yuji Miyazaki, Daiju Matsumura, Hideaki Shiwaku, and Tomitsugu Taguchi. A reliable hybrid adsorbent for efficient radioactive cesium accumulation from contaminated wastewater. *Scientific reports*, 6:19937, 2016.

[119] M.F. Elkady M.M. Abd El-Latif. Kinetics study and thermodynamic behavior for removing cesium, cobalt and nickel ions from aqueous solution using nano-zirconium vanadate ion exchanger. *Desalination*, 271:41–54, 2011.

[120] Peter V. Bonnesen A. C. Buchanan III John D. Holbrey Nicholas J. Bridges Huimin Luo, Sheng Dai and Robin D. Rogers. Extraction of cesium ions from aqueous solutions using calix[4]arene-bis(tert-octylbenzo-crown-6) in ionic liquids. *Anal. Chem.*, 76:3078–3083, 2004.

[121] Joseph A. McDonough W. Jack McDowell, Gerald N. Case and Richard A. Bartsch. Selective extraction of cesium from acidic nitrate solutions with didodecylnap ht halenesulf onic acid synergized with bis(tert=butylbenzo)-2 1-crown-7. *Anal. Chem.*, 64:3013–3017, 1992.

[122] Torsten Rambusch Holger Stephan Karsten Gloe Christian E. O. Roesky, Edwin Weber and Matyas Czugler. A new cryptophane receptor featuring three endo-carboxylic acid groups: Synthesis, host behavior and structural study. *Chem. Eur. J.*, 9(5):1104–1112, 2003.

[123] Aude Bouchet Dominique Cavagnat Mathieu Linares Thierry Brotin, Roland Montser-

ret and Thierry Buffeteau. High affinity of water-soluble cryptophanes for cesium cations. *J. Org. Chem.*, 77:1198–1201, 2012.

[124] Patrick Berthault Roland Montserret Thierry Brotin, Dominique Cavagnat and Thierry Buffeteau. Water-soluble molecular capsule for the complexation of cesium and thallium cations. *J. Phys. Chem. B*, 116:10905–10914, 2012.

[125] Dominique Cavagnat, Thierry Brotin, Jean-Luc Bruneel, Jean-Pierre Dutasta, Alain Thozet, Monique Perrin, and François Guillaume. Raman microspectrometry as a new approach to the investigation of molecular recognition in solids: Chloroform-cryptophane complexes. *The Journal of Physical Chemistry B*, 108(18):5572–5581, 2004.

[126] W Humphrey, A Dalke, and K Schulten. Vmd: visual molecular dynamics. *J Mol Graph*, 14(1):33–8, 27–8, Feb 1996.

[127] William L Jorgensen, David S Maxwell, and Julian Tirado-rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. 7863(15):11225–11236, 1996.

[128] William L Jorgensen and Julian Tirado-Rives. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

[129] In Suk Joung and Thomas E. Cheatham. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *The Journal of Physical Chemistry B*, 112(30):9020–9041, 2008. PMID: 18593145.

[130] François-Yves Dupradeau, Adrien Pigache, Thomas Zaffran, Corentin Savineau, Rodolphe Lelong, Nicolas Grivel, Dimitri Lelong, Wilfried Rosanski, and Piotr Cieplak. The r.e.d. tools: advances in resp and esp charge derivation and force field library building. *Phys Chem Chem Phys*, 12(28):7821–39, Jul 2010.

[131] Mike C Payne, Michael P Teter, Douglas C Allan, TA Arias, and JD Joannopoulos. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Reviews of modern physics*, 64(4):1045, 1992.

[132] Thomas C Beutler, Alan E Mark, René C van Schaik, Paul R Gerber, and Wilfred F van Gunsteren. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chemical physics letters*, 222(6):529–539, 1994.

[133] Christophe Chipot and Andrew Pohorille. Calculating free energy differences using perturbation theory. *Free energy calculations*, pages 33–75, 2007.

[134] a. Ben-Naim and Y. Marcus. Solvation thermodynamics of nonionic solutes. *The Journal of Chemical Physics*, 81(1984):2016, 1984.

[135] A Viale, F Reineri, D Santelia, E Cerutti, S Ellena, R Gobetto, and S Aime. Hyperpolarized agents for advanced mri investigations. *The Quarterly Journal of Nuclear Medicine and Molecular Imaging*, 53(6):604, 2009.

[136] H Lawrence Clever. *Krypton, Xenon & Radon*, volume 2. Elsevier, 2013.

[137] Daniel Raftery. Xenon nmr spectroscopy. *Annual reports on NMR spectroscopy*, 57:205–270, 2006.

[138] Thad G Walker and William Happer. Spin-exchange optical pumping of noble-gas nuclei. *Reviews of Modern Physics*, 69(2):629, 1997.

[139] Boyd M Goodson. Nuclear magnetic resonance of laser-polarized noble gases in molecules, materials, and organisms. *Journal of Magnetic Resonance*, 155(2):157–216, 2002.

[140] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926, 1983.

[141] Jay W Ponder and David A Case. Force fields for protein simulations. *Advances in protein chemistry*, 66:27–85, 2003.