UNIVERSITY OF CALIFORNIA
RIVERSIDE

Molecular Basket Weaving and Their Carrying Capacity in Water

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Briana L. Hickey

September 2023

Dissertation Committee:
Dr. Richard Hooley, Chairperson
Dr. Kevin Kou
Dr. Ana Bahamonde

The Dissertation of Briana L. Hickey is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGMENTS

This work is dedicated to my beautiful husband Phil, you are the best thing that has come out of graduate school. I love and appreciate everything about you and can't imagine how I would have finished this degree without you.

The test of this dissertation, in part or in full, is a reprint of the material as they appear in the following publications:

Chapter 2: Chen, J.; Hickey, B. L.; Wang, L.; Lee, J.; Gill, A. D.; Favero, A.; Pinalli, R.; Dalcanale, E.; Hooley, R. J.; Zhong, W. Selective Discrimination and Classification of G-Quadruplex Structures with a Host:Guest Sensing Array. *Nat. Chem.* **2021,** *13,* 488-495.

Chen, J.; Gill, A. D.; Hickey, B. L.; Gao, Z.; Cui, X.; E.; Hooley, R. J.; Zhong, W. Machine Learning Aids Classification and Discrimination of Noncanonical DNA Folding Motifs by an Arrayed Host:Guest Sensing System. *J. Am. Chem. Soc.* **2021,** *32,* 12791-12799.

Chen, J.; Hickey, B. L.; Raz, A. A. P.; Gao, Z.; Hooley, R. J.; Zhong, W. Sensing Base Modifications in Non-Canonically Folded DNA with an Optimized Host:Guest Sensing Array. *ACS Sens.* **2022**, 7, 2164–2169.

Gill, A. D.; Perez, L.; Salinas, I. N. Q.; Byers, S. R.; Liu, Y.; Hickey, B. L.; Zhong, W.; Hooley, R. J. Selective Array-based Sensing of Anabolic Steroids in Aqueous Solution by Host:Guest Reporter Complexes. *Chem. Eur. J.*, **2019**, *25*, 1740-1745.

Gill, A. D.; Hickey, B. L.; Zhong, W.; Hooley, R. J. Selective Sensing of THC and Related Metabolites in Biofluids by Host:Guest Arrays. *Chem. Commun.* **2020**, *56*, 4352 - 4355.

Chapter 3: Hickey, B. L.; Chen, J.; Zou, Y.; Gill, A. D.; Zhong, W.; Millar, J. G.; Hooley, R. J. Enantioselective Sensing of Insect Pheromones in Water. *Chem. Comm.*, **2021**, *57*, 13341 – 13344.

Chapter 4: Hickey, B. L.; Raz, A. A. P.; Chen, J.; Moreno J. L. Jr.; Hartman, J. D.; Zhong, W.; Hooley, R. J. Selective Anion Sensing in High Salt Water via a Remote Indicator Displacement Assay. *Chem. Commun.*, **2023**, *59*, 7819-7822.

The co-author Richard J. Hooley listed in these publications directed and supervised the research which forms the basis for this dissertation. All other co-authors listed in these publications contributed technical expertise.

ABSTRACT OF THE DISSERTATION

Molecular Basket Weaving and Their Carrying Capacity in Water

by

Briana L. Hickey

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, September 2023
Dr. Richard J. Hooley, Chairperson

The field of molecular recognition illustrates the ability of synthetic receptors to act as hosts for important target molecules. These hosts are capable of binding guests in a manner akin to larger biological systems. Synthetic receptors are ideal molecules for observing host:guest interactions as they possess tunable structures that can be tailored to the desired target. As the differences in many biological targets are relatively small, the use of an array-based system allows for greater discrimination of these slight variations. This approach uses multiple variably functionalized receptors that act as a "chemical nose" for target molecules in a variety of optical sensing applications. This array-based system enables the receptors to sense a broader range of relevant biomacromolecules. The differential sensing of these small molecules can be measured *via* fluorescence outputs which are deconvoluted using Principal Component Analysis and other statistical methods.

The use of water soluble deep cavitands allows for greater scope and tunability of their recognition abilities compared to other macrocyclic receptors. Much previous work has investigated deep cavitands with a negatively charged upper rim, so new structures that

incorporate a positive charge were synthesized as a direct comparison of functionality. The diversification of functional groups at the upper rim broadens the scope of target molecules that the cavitands can sense, while the derivatization of the lower rim allows for increased solubility as well as secondary effects of the binding interactions. Deep cavitand hosts exhibit this selective target recognition by utilizing a self-folding deep pocket that provides a defined cavity for binding.

This work has highlighted the importance of the synthetic variation of the receptors for sensing specific targets. The novel cationic cavitands remain neutral at the upper rim and only experience charge interactions at the lower rim of the structure. These positively charged receptors have provided an array-based system that is capable of enantioselective sensing of pheromone guests which have small differences in their structures that the array-based system was still able to discriminate between. Additionally, the discrimination of anions with the cationic cavitands showed a strong preference for iodide, despite large concentrations of competitive ions in solution. This secondary interaction at the lower rim of the cavitand structure causes a conformational change in the binding pocket of the cavitand, which influences the affinity of the bound guest for the cavitand. From the success with negatively charged guests, the sensing of phosphosugars has been investigated, and preliminary results show discrimination of fructose-6-phosphate and glucose-6-phosphate using the cavitand array.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter One: Molecular Recognition with Synthetic Receptors

1.1 Introduction

Molecular recognition events continuously occur in all biological systems and play essential roles in many cellular activities, including the monitoring of epigenetic markers.[1,2] These binding events take place in various aqueous environments between large biomacromolecular hosts, such as enzymes, antibodies, and proteins, and a wide assortment of small molecule targets.[3,4] A molecular recognition event is defined as the interaction of two or more molecules through dynamic non-covalent interactions.[5-8] This process of recognition is facilitated *via* the binding capabilities of these large host molecules, and directly relies on the physical and chemical properties that are exhibited by their binding sites and cavities.[5-9] The field of molecular recognition illustrates the ability of synthetic receptors to act as hosts for important target molecules in a manner akin to these larger biological systems. Through noncovalent interactions, these binding events can occur. Each host exploits a combination of these interactions which consist of hydrogen bonding, electrostatic, and hydrophobic interactions. Additionally, the implementation of aromatic systems to the macrocyclic host structures allows for an even more type of interactions between the host and guest, such as π-π, CH-π, and cation-π interactions (Figure 1.1).[5-13] The development of synthetic receptors aims to exploit these noncovalent interactions for greater recognition of a broad scope of guests.

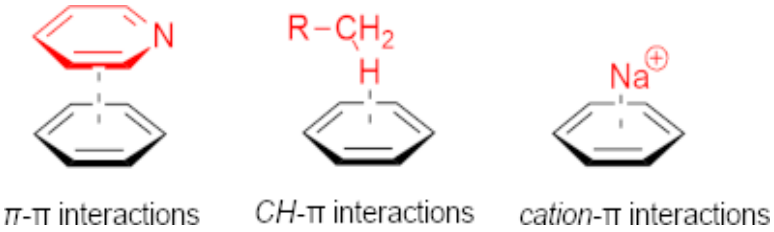


Figure 1.1. Types of interactions exhibited by aromatic rings that enable molecular recognition.

Early studies of cation-π interactions were focused on a type of synthetic host known as cyclophanes (Figure 1.2). These synthetic receptors were studied for their recognition of cationic small molecules.[14] The cationic guests showed greater affinity for binding with the π-systems than remaining free in solution. The C-H dipoles created by the aromatic rings of the host structure compound to form a region of negative electrostatic potential. With the face of the π-system being negatively charged, the cationic guests are inherently attracted to this surface.[14,15] Macrocyclic receptors can also employ CH-π interactions for the recognition of neutral species, where the CH groups of the guest indirectly interacts with the π-system, as opposed to the electrostatic nature exhibited by cation-π interactions.[5-13]

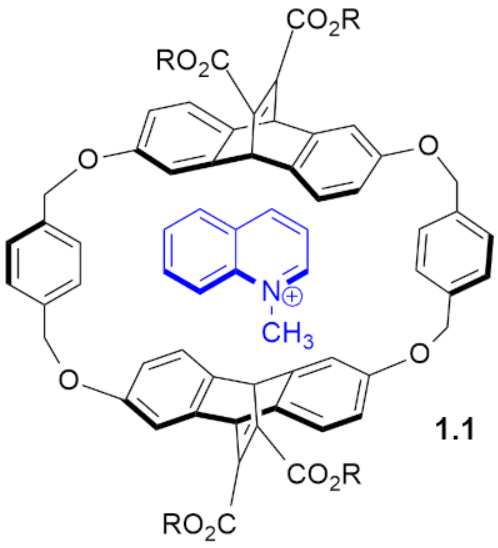


Figure 1.2. A cyclophane synthetic receptor, which shows high affinity for cationic guests *via* cation-π interactions.

Along with the solvent conditions of the environment, and their effect on the solubility of the receptor, there are many other factors to consider for the implementation of synthetic receptors for molecular recognition. Various other types of interactions must be considered as well, such as interactions between the host with itself, the host with the guest, and secondary interactions with other surrounding components in the environment. All these factors may increase competitive interactions that must be overcome for the binding event to occur. Free molecules are in a dynamic environment where they can interact with each other, as well as the solvent molecules. When these molecules are confined, their mobility is reduced, and their conformations and behaviors change, akin to a substrate free in solution versus bound inside an enzyme's active site. The binding cavity of the host will have a higher affinity for guests that are complementarity in both shape and size, as well as hydrophobicity.

The hydrophobic effect plays a key role for binding events that occur in aqueous media, as the targets are commonly non-polar molecules that are drawn to the confined binding spaces of these hosts to avoid further interaction with the surrounding aqueous environment. This effect is defined by the tendency of nonpolar molecules to exclude water and aggregate in aqueous solutions.[16-20] This process breaks the hydrogen bonds between water molecules and is endergonic in nature. The hydrogen bonds of the surrounding water molecules reorient around the hydrophobic guest, so as to minimize the disruption of the hydrogen bonding network overall.[20-24] This new system is highly structured and decreases the total entropy. The interaction between the hydrophobic molecules themselves in aqueous solution, however, is favorable. This effect produces a hydrophobic pocket in which the hydrophobic molecule exists, and both the size and shape of the guest molecule influences the strength of this effect. The larger the surface area and more nonpolar a hydrophobic molecule is, the greater the strength of the hydrophobic interactions it exhibits. The best example of this is protein folding, which implements hydrophobic interactions to decrease the surface area and the protein and limit its undesirable interactions with water.[17-21] The hydrophobic effect is considered one of the largest contributors and driving forces for the binding of small molecules in aqueous receptors, as it is a thermodynamically favored event. This effect is exploited by the receptor molecules at their binding sites, which provide a hydrophobic cavity that isolates the guest from the bulk solution, resulting in minimal disruption to the water molecules.

While the hydrophobic effect focuses on the ability of nonpolar molecules to dissolve in aqueous environments, another inter-related effect, known as the Hofmeister

effect, also plays a key role in guest solubility.[25-27] The Hofmeister effect describes the ability of salts to modulate the properties of the aqueous environment. The salt compounds present in the aqueous environment use their charged components to directly interact with the guest molecules, influencing their solubility properties.[28-32] Another important factor for molecular recognition is the environment of the binding pocket of the hosts. The solvation of the binding site greatly impacts the ability of the host to recognize and bind its substrate. This characteristic is known as wettability and is predominately controlled by the overall solubility of the host. To increase the solvation of the binding pockets, these host molecules are modified to include charged groups in their structures. These groups increase the overall solubility of the host, even the binding pocket, and affect guest binding. In conjunction with solvation of the binding pocket, its size and shape are also influential for guest recognition. In biology, enzymes are known to be flexible in their conformations, which allows them to express a high degree of selectivity for their substrate.[1-4] This dynamic quality enables the interactions between the host and guest molecules to adjust based on the size of the substrate and maximizes the strength of the covalent forces necessary for recognition. Generally, most synthetic receptors lack this flexibility and exist in rigid conformations that are unable to adjust their binding pocket according to each guest.

1.2 Molecular Recognition by Macrocyclic Synthetic Receptors

There are multiple characteristics possessed by macrocyclic receptors that make them ideal for molecular recognition, such as areas of the host structure that can be synthetically modified to increase solubility and recognition selectivity. In this regard

synthetic receptors are extremely robust and more widely applicable, due to their derivatization possibilities, as well as their potential for scalability, and cost-effective production. The relationship that these hosts have with selectivity and specificity is complicated, if the hosts are specifically tailored for their substrates, it hinders the scope of their selectivity. This broad scope of selectivity by synthetic receptors can be exploited to increase the variety of targets for recognition. Common synthetic receptors are based on two key components for their design, solubility in aqueous environments, and the inclusion of a defined binding cavity, which is critical for the ability of these molecules to bind desired targets with not only high affinity, but hopefully selectivity as well.

The field of macrocyclic synthetic receptors is quite broad, and includes many unique structures, which all have their own advantages and limitations. The binding events for these host:guest complexes are dependent on the solvent that they occur in. It is much simpler to sense targets in pure organic solvents than the complex aqueous environments that sustain life. The application of these receptors is much more complex in aqueous environments, where the interactions of ions, co-solutes, and competitive analytes must be accounted for as well. These additional factors influence the strength of the host:guest interactions, and even the host:guest complex formation. Binding affinities of these synthetic receptors have been measured, and a scale of their strength is used for comparison to biological hosts. These affinities can be measured in terms of K_a, the association constant for the formation of the host:guest complex, or K_d, the dissociation constant for the dissolution of the host:guest complex.[17] The binding of alkali metals to 18-crown-6 is the known standard for weak binding affinity of a synthetic host in water, with a K_a of 0.91 M^{-}

[1], while strong binding affinity is exhibited by cucurbit[n]urils which have an astoundingly high affinity of approximately $K_a = 10^{15}$ M^{-1} for tetramethyl ammonium guests.[34]

The study of many synthetic macrocyclic receptors has been broad, and these host molecules have shown promise with their molecular recognition performance in aqueous media. All of the various synthetic receptors detailed below are capable of recognition events in aqueous media, and more specifically have been used for the successful sensing of protein post translational modifications (PTMs). These post translational modifications occur at histone peptides and increase the diversity of proteins amongst living cells. They are useful markers for monitoring the epigenetic regulation of cells, which is a key indicator of disease progression, and often include rare residues which are difficult to individually recognize. The detection of the low concentrations these disease markers is made even more difficult in complex biological medias, such as blood, saliva, and urine.[1-4]

While most synthetic receptors possess multiple aromatic groups for their scaffold, there are a few examples of macrocyclic receptors that are not. For example, cyclodextrins are glycosidic macrocycles, with a sugar derived scaffold, that allows for them to have a hydrophobic cavity, but a hydrophilic exterior.[33] They are widely used to bind small molecules in aqueous environments but lack strong selectivity for their targets. Similarly, cucurbit[n]urils (CB[n]s) are comprised of glycouril subunits that are connected *via* methylene groups, to form larger oligomeric structures.[34,35] CB[n]s feature nonaromatic, urea-containing rings, that rely on hydrophobic, ion-dipole and hydrogen bonding interactions *via* the carbonyl groups positioned at the rims for recognition of guest molecules (Figure 1.3). The size of these receptors can vary by the number of repeating

glycouril units present, which increases the area of the binding cavity, while the length of the receptor remains constant. This scaffold produces a barrel-like structure with two identical portal regions at each rim. The lack of internal electron density, and the orientation of the functional groups and lone pairs towards the outer surface, makes the binding cavity of these structures is largely inert.[35,36]

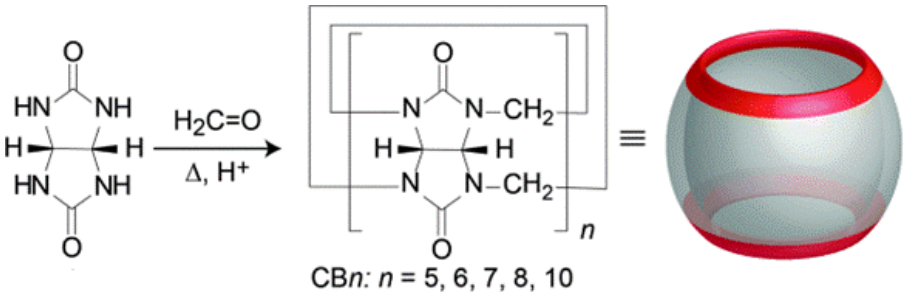


Figure 1.3. Formation of cucurbit[n]uril receptors from the condensation reaction of *n* glycoluril units and formaldehyde. Reproduced from ref. 35 with permission. Copyright 2015, American Chemical Society.

Their hydrophobic cavity is primed for the recognition and binding of suitably sized hydrophobic guests but shows a distinct preference for cationic targets over neutral and anionic ones. The best guests are those with both hydrophobic and cationic regions, such as trimethyllysine KMe_3 (K_d = 0.53 μM), where the ammonium cation can interact with the polar carbonyl groups and the hydrophobic alkyl groups are isolated in the binding cavity.[36-38] The solubility of these receptors varies greatly and is directly influenced by the number of repeating units, but can be enhanced in the presence of a guest molecule. Both CB[5] and CB[7] show moderate water solubility at concentrations of 20-30 mM. With CB[7], the affinity (K_a) of a bis-trimethylammonium ferrocene derivative was found to be approximately 10^{15} M^{-1}, even higher than the affinity of biotin-streptavidin, which is the

benchmark for high affinity binding in nature.[39] This high affinity is in part due to the poor interactions between water and the CB[n] cavity, which is due to the narrow portals at both ends and makes the binding cavity less appealing to water than remaining in the bulk solution.

One of the largest categories of synthetic receptors is known as cyclophanes,[14-16] the most notable of this group being the well-studied calix[n]arene[40,41] and the newer pillar[n]arene families.[34,42] The first studies of the binding interactions between aromatic-containing macrocycles, known as cyclophanes, and various guests, were conducted by Dougherty.[14-16] These structures, in their simplest form, are aromatic units that are bridged by a hydrocarbon chain. They feature six aromatic rings, linked *via* ether groups, resulting in a rigid aromatic surface (Figure 1.4). These structures were further modified with ester, and carboxylate functional groups, which increase the solubility of the host without influencing the binding cavity.

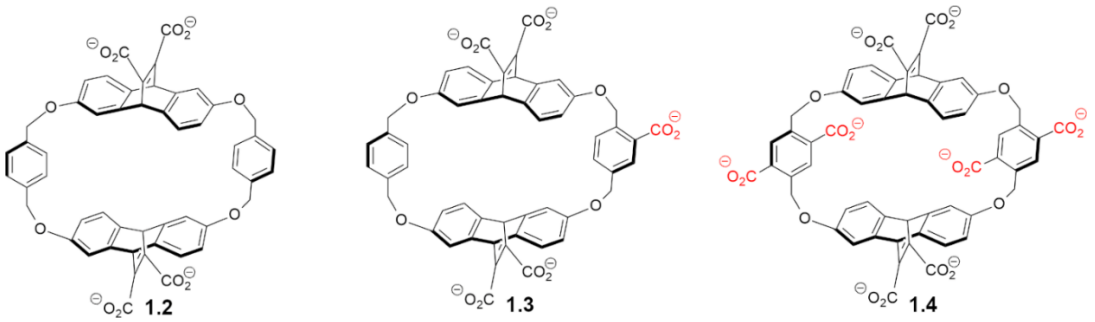


Figure 1.4. Cyclophane receptors derivatized with water-solubilizing corboxylate groups to study cation-π interactions in water.

Initial studies showed the preference of these hosts for positively charged ammonium guests *via* cation-π interactions, over hydrophobic interactions between the

host and non-polar guests. This emphasized the importance of cation-π interactions for the selectivity of the host and these cationic species, with only minimal influence from the hydrophobic effect on the overall binding affinity.[14-16] The aromatic rings of the cyclophane host are essential for the recognition of the charged molecules, as the C-H bond dipole of the aromatic system creates a region of negative electrostatic potential that the postively charged guests are drawn to. This is caused by the sp^2 carbon of the C-H bond, which is more electonegative than the hydrogen. This structural motif is also seen in biological hosts, and is known as the aromtic box, and many synthethic receptors mimic this quality to enhance their recognition abilites.

This notable preference for cation-π interactions by these macrocycles for quaternary ammonium salts, was further investigated by Waters, using a ß-hairpin peptide model, and methylated Lysine and Arginine amino acids.[44] The results of these studies illustrated a remarkable magnitude of interaction between the trimethyl lysine (KMe_3) guests, and a single cation-π interaction of the host, where a methyl group forms a direct contact with an aromatic ring of the host structure. This work spawned the synthesis of mercaptophanes, which are generated from aromatic compounds that are connected *via* disulfide linkages, and feature negatively charged exteriors. Mercaptophanes are formed by derivatization of the cyclophane scaffold, where water soluble aromatic components possessing thiol groups are used to generate the macrocycle (Figure 1.5).[45] This library of aromatic building blocks is known as a dynamic combinatorial library (DCL) and allows for the formation of the receptors with the highest affinity to self-assemble for each target of interest.

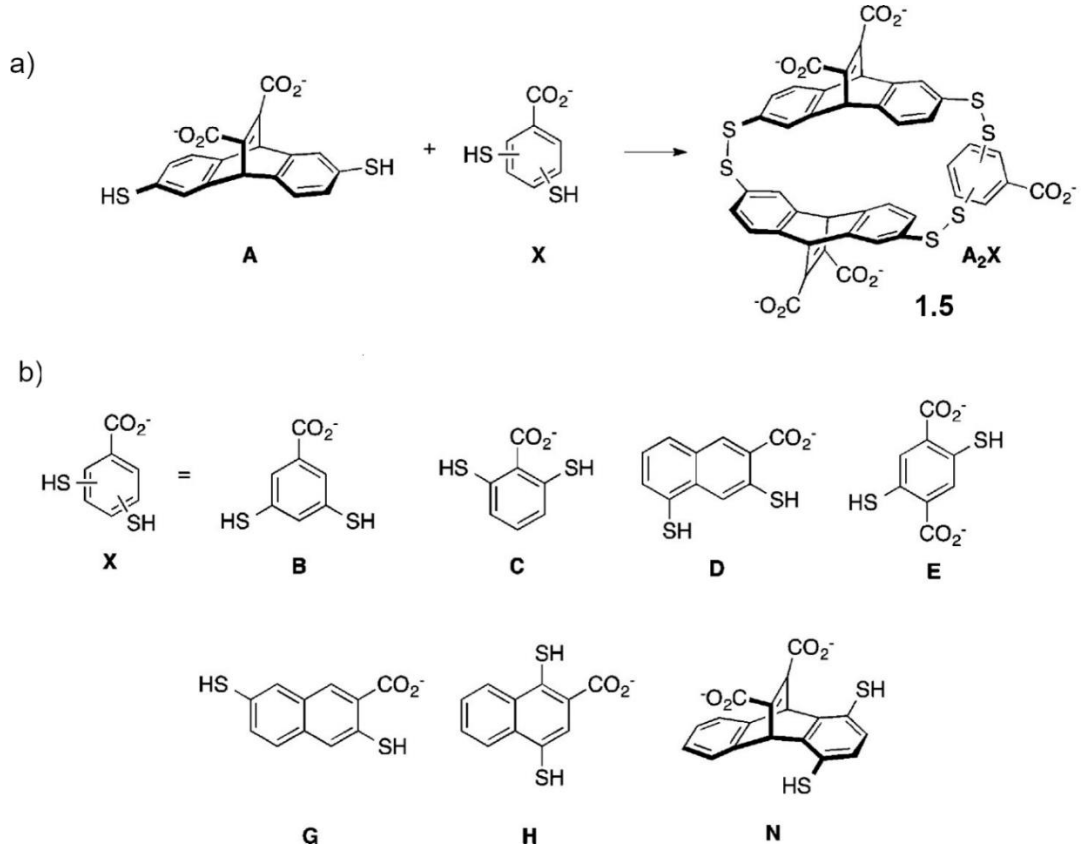


Figure 1.5. Development of a mercaptophane receptor using a dynamic combinatorial library method where a) is the self-assembly of the receptors, and b) is the library of thiol building blocks. Adapted from ref. 44 with permission. Copyright 2016, American Chemical Society.

The A_2X (**1.5**) receptor was shown to have a binding affinity of K_d = 2.6 μM for KMe_3 and expressed significant selectivity for the trimethylated lysine groups when compared to the demethylated KMe_2 (K_d = 6.3μM) and unmethylated Lys (K_d = 22 μM) guests. The negatively charged groups of the aromatic building blocks aid the binding of these PTMs, and the high selectivity for KMe_3 is a result of the increased cation-π interactions from the additional methyl group in comparison to KMe_2. There is also a higher solvation cost for the lower methylation states, as they are neutral in charge, and more hydrophobic in nature. When studies were performed to enhance the charge

localization at the binding pocket, it was shown to increase binding affinity for all methylation states of Lys by 10-fold but had negligible effects on selectivity.[46,47]

Calix[n]arenes (CXn) are aromatic macrocycles that are composed of repeating units of phenol rings that are linked by methylene groups in the meta positions. This scaffold allows for a variety of sizes for the overall host structure, and the number of rings can be adjusted to accommodate a plethora of target sizes. The most common calix[n]arenes are those where n = 4-8 for the number of phenol rings in the structure.[48,49] These receptors can also be further functionalized at the alcohol groups, which increases both the recognition abilities and solubility of the molecule, while the aromatic cavity facilitates the incorporation of hydrophobic guests. Calix[n]arenes exhibit a cone-like conformation with two entry portals at either end, and the structure is stabilized by the intramolecular hydrogen bonding between the phenol groups of the main scaffold (Figure 1.6). The lower portal is much narrower due to the hydrogen bonding interactions of the phenol groups at the lower rim, while the upper portal opens much wider. One of the most notable calix[n]arene derivatives is the para-sulfonatocalix[4]arene (CX4) **1.6**, which employs sulfonate groups around the upper rim of the structure.[50-52]

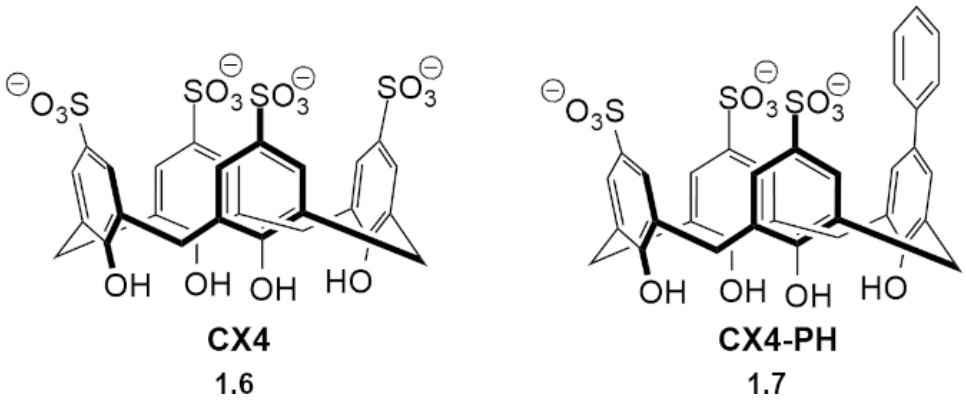


Figure 1.6. Calix[4]arene derivatives with fully sulfonated **1.6** or trisulfonated **1.7** upper rims.

The CX4 host is able to bind KMe_3 (K_d = 95μM) *via* interactions at both the binding pocket, and the sulfonate groups, using a combination of cation-π and electrostatic interactions. A derivative of CX4 was synthesized with a deeper binding pocket to analyze the effect of the size of the pocket on the binding affinity of the KMe_9 groups presents on longer histone peptide side chains. This CX4-PH receptor **1.7**, is trisulfonated at the upper rim of the structure and includes the addition of a phenyl ring. This derivative was shown to increase the binding affinity of KMe_3 by a factor of 2.5 (K_d = 2.4 mM), due to the increased contact between the KMe_3 and the host structure, and greatly improved selectivity by 150-fold because of the increased hydrophobicity of the binding pocket.[51,52]

Pillar[n]arenes are a relatively new type of synthetic receptor, with 1,4-dimethoxypillar[5]arene being first synthesized in 2008 by Ogoshi.[53,54] Similar to calix[n]arenes, pillar[n]arenes are also comprised of repeating units, they are linked by methylene groups in the para position, and these repeating units are hydroquinone groups instead of the phenyl rings used for calix[n]arenes. Pillar[n]arenes feature a pillar-like

structure that has two identical portals at either end. The inner tube-like surface of the pillar[n]arene structure is highly electron dense, due to the hydroquinone groups, which allows for greater preference of electron poor guests.[55-57] The internal cavity dimensions of the pillar[n]arene structures can be altered by increasing the number of repeating hydroquinone groups. The length of the pillar[n]arene structure can be increased as well, which is commonly done by linking hydrazide units through intramolecular hydrogen bonding, to induce the formation of the tubular structures. The functionalization of the alkoxy groups is used to increase host:guest interactions, as well as the solubility of the structure. This derivatization of pill[n]arenes is essential to confer water solubility and can be done completely, or selectively, at the alkoxy groups (Figure 1.7).

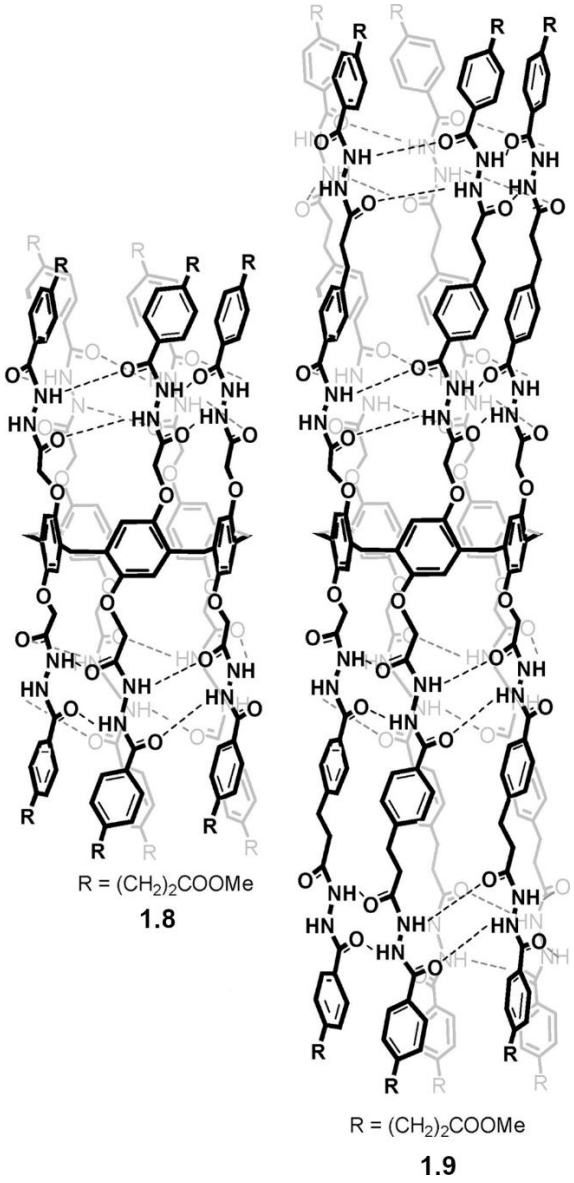


Figure 1.7. Pillar[5]arene receptors of varying lenghts, linked by intramolecular hydrogen bonding of the hydrazine groups, with functionalized upper and lower rims. Adapted from ref. 54 with permission. Copyright 2016, American Chemical Society.

Analysis of a water-soluble pillar[5]arenes **1.8** and **1.9**, has shown effective binding with positively charged guests. A greater selectivity for unmethylated Lysine over trimethylated KMe_3 guests was observed, as hydrogen bonding interactions plays a greater

role for the host:guest complexation of pillar[n]arenes. The circumference of the pillar-like shape of pillar[n]arenes limits the size of the guests that can fit in the binding pocket. For example KMe_3 is too large for the binding pocket of the pillar[5]arene hosts, and it is excluded from the aromatic core, while the much smaller Lysine guest is able to bind *via* electrostatic interactions with the carboxylate groups at either end of the cavity. Further derivatization of the pillar[n]arene scaffold with cationic ammonium groups was able to produce a positively charges pillar[5]arene host, which was fully soluble in water, and showed high affinity for anionic guests, such as sodium alkyl sulfonates.[57] Additionally, the tube-like nature of the pillar[n]arene structures allows them to be used as an artificial channel for molecular transport (Figure 1.8). When they are inserted into membranes, the pillar[5]arene structure can be used as a transmembrane water channel.[58,59]

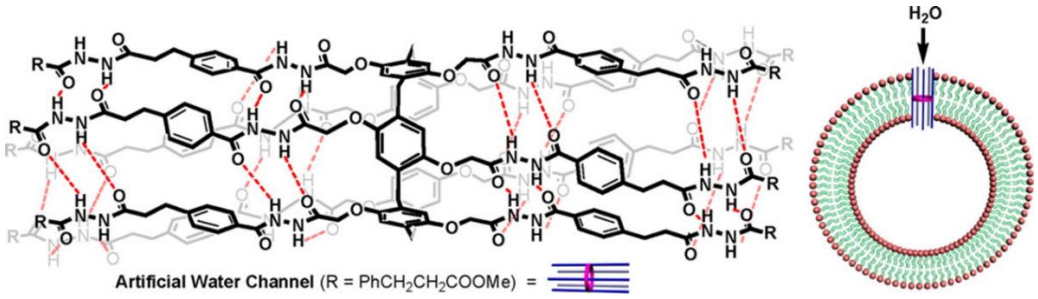


Figure 1.8. Pillar[5]arene receptors as acting as an artifical water channel across a cell memebrane. Reproduced from ref. 59 with permission. Copyright 2012, American Chemical Society.

While these synthetic receptors have shown promising results for the sensing of a handful of specific small molecules, they also each have limitations in their nature. One major limitation of synthetic receptors is their generally small, rigid, cavities which greatly hinders the target scope.[13,36,45,55] Many biologically interesting molecules are too large and

inappropriately shaped for said cavities. One major difference between biological hosts and synthetic hosts is size, for example, the large size of enzymes allows them to not only be more flexible, but to also have much more intricate binding sites. A more flexible and conformationally dynamic receptor would allow for a greater scope of recognition, as the host cavity would be able to adjust to best accommodate a variety of different sized guests. The targets of interest often exist as charged species. Notably, "soft" cations have been shown to be excellent guests, while negatively charged guests exhibit poor electrostatic interactions with the generally electron rich cavities of most synthetic receptors, which limits their recognition. All these factors reduce the selectivity and strength of the binding interactions of these receptor molecules and are areas of improvement for molecular recognition.

1.3 Use of Water-Soluble Deep Cavitands for Molecular Recognition

The resorcin[n]arene scaffold is a well-established framework for water-soluble cavitands and is composed of an aromatic cavity that shows excellent binding affinity for a variety of guests in aqueous environments. This concave structure is formed through the condensation of resorcinol and an alkyl aldehyde, through an electrophilic aromatic substitution reaction, resulting in the shallow bowl-like resorcin[n]arene base. The first derivatization of this resorcinarene scaffold was Cram's shallow cavitand,[60,61] which links the eight hydroxyl groups at the upper rim *via* methylene units, resulting in receptor **1.10**. This receptor has a shallow cavity, with a depth of 3.3-4.2Å. and is capable of binding small hydrophobic molecules (Figure 1.9). Using an approach to further deepen the binding cavity of the cavitand, Gibb connected the phenols of resorcin[4]arene groups (**1.11**), the

addition of these aromatic rings positioned them above the opening of the cavity, enlarging it to 380 Å^3 (Figure 1.9).[62]

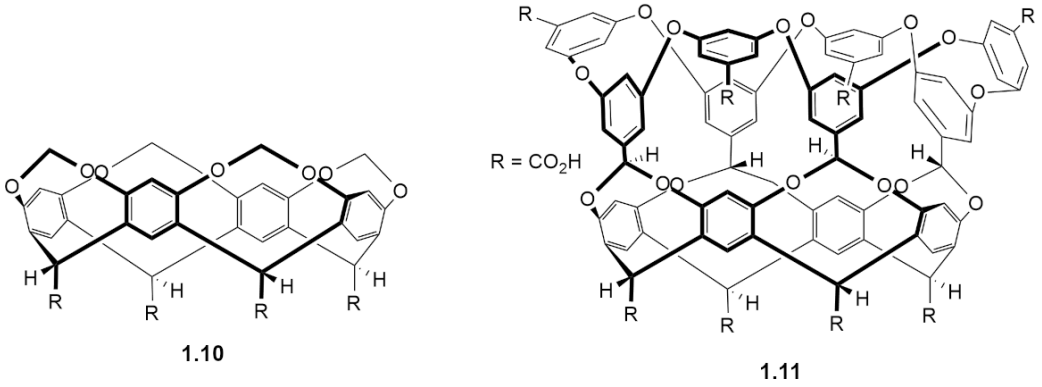


Figure 1.9. Resorcinarene based receptors by Cram (**1.10**) and an expanded scaffold with anionic groups at the upper and lower rims by Gibb (**1.11**).

These cavitand structures vary from other synthetic receptors in that they are bowl-shaped, resulting in only one portal and a closed lower rim. With this single-entry point, the solvation of the binding pocket is more difficult, and the cavity itself can be described as more hydrophobic than that of other synthetic receptors. The replacement of water from the binding pocket is much more difficult and proceeds through a more S_N1-like mechanism.[62] Cavitands have two regions that can be functionalized to increase both their solubility and recognition capabilities simultaneously. The upper rim modifications are adjacent to the binding pocket, which orient towards the target, and can also affect the host's interactions with its environment. The addition of alkyl pendant chains at the lower rim of the structure, also referred to as "feet", can be implemented by varying the alkyl aldehyde used in the initial condensation reaction. These modifications of the feet can

impart differing solubility properties while retaining an identical binding pocket at the upper rim and host cavity. When functional groups are added to both the upper rim and lower feet of the cavitand, the result is the addition of four, eight, or twelve total functional groups to the structure.

Gibb's initial expanded scaffold was poorly soluble in water, so both cationic and anionic derivatives were synthesized.[62] The addition of eight carboxylate groups by Gibb at both the upper and lower rims, resulted in a water-soluble deep cavitand, known as the octa-acid (OA) cavitand **1.11**, that was capable of binding various small molecules with high affinity, ranging from values of 10^3-10^6 M^{-1} for K_a.[62] Guests with hydrophilic groups showed the highest affinity, such as hexanoate. Binding affinity studies were done *via* ^{1}HNMR, and showed that amphiphilic guests, such as tetramethylammonium salts, formed a 1:1 complex with the OA host. It was observed that the alkyl groups were placed inside of the binding cavity, and the positively charged ammonium groups were oriented towards the upper rim, increasing their interactions with the negatively charged groups of the cavitand structure. When hydrophobic guests are in non-aqueous solutions, they prefer to extend to increase their surface area to minimize their own steric interactions, however in aqueous environments this extended conformation is energetically disfavored. When these guests are bound in the binding pocket of cavitands, they coil into helices to maximize their interactions with the host, but also to limit their contact with water.[63,64] When this characteristic is paired with the influence of the hydrophobic effect, these larger flexible guests can greatly compress themselves into the relatively smaller binding pocket of the cavitand.

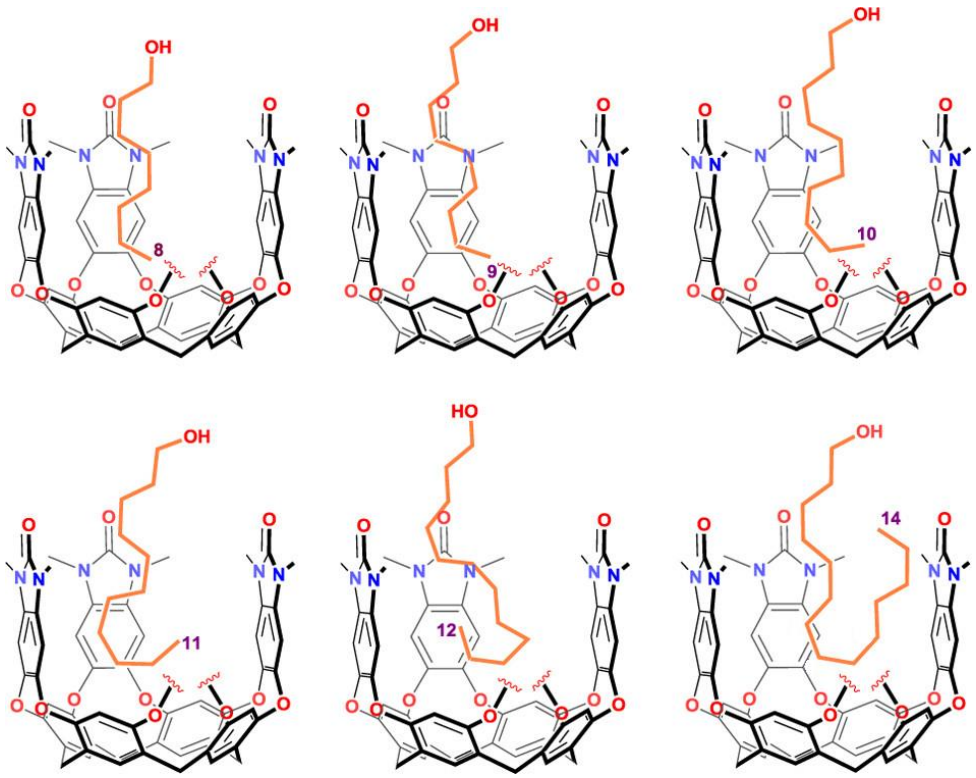


Figure 1.10. Complexation of cavitand hosts and n-alcohols of varying lengths, showing the helical conformation taken on by the guest molecule inside of the binding cavity of the host. Reproduced from ref. 66 with permission. Copyright 2014, American Chemical Society.

A study done with n-alcohols illustrates how the concave nature of deep cavitands allows for the guests to adopt these compressed conformations (Figure 1.10).[65,66] With the n-alcohols, guests with fewer than ten carbons in length were able to fit in the cavity without much coiling, but with the addition of more carbon atoms (C ≥10) the guests began to adopt a J-shaped conformation, where the long alkyl chain is buried in the hydrophobic pocket, and the hydroxyl group is placed towards the upper rim.

While this deeper cavity allows for binding of larger guests, it can also result in greater movement and flexibility of the cavity walls, as each aromatic ring is independently linked, and interactions between the rings occur *via* hydrogen bonding. This was an issue seen in the expanded deep cavitands that possessed these flexible walls, which allowed the cavitand to easily convert between two conformations.[60-64] These conformations are known as "kite" and "vase" and can be distinguished *via* ^{1}H NMR spectroscopy at the bridged methine protons (Figure 1.11). The vase conformation with C_{4v} symmetry, forms the deep cavity for binding with all the walls of the cavitand oriented up, and an observed peak at about 5.5 ppm, corresponding to the methine protons. The kite conformation exhibits C_{2v} symmetry, wherein the walls of the cavitand are flexed outward, resulting in a large, flat, surface and a methine peak that is shifted upfield at approximately 4 ppm due to an increase in shielding effects. This kite conformation exposes much of the aromatic surfaces of the cavitand to the aqueous environment. To minimize those interactions, a dimer known as a velcrand (D_{2d}) can form. When an appropriate guest is added to the system, the velcrand can disassemble, and reorient to the vase conformation and bind the guest molecule.[63,64]

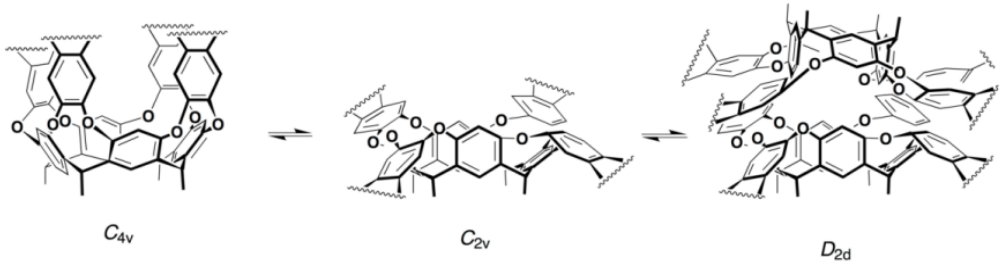


Figure 1.11. Visual representation of flexing by the wall of cavitands to produce vase, kite, and velcrand conformations respectively.

1.4 Sensing Applications of Synthetic Receptors *via* Assays and Larger Array-Based Systems

While most synthetic receptors show high affinity for specific guests, the selective recognition of structurally similar targets can be difficult for the host molecule to discriminate between alone. To maximize selectivity, different methods have been developed that couple the recognition event of the host with a measurable signal response. When these interactions are coupled with an optical output, an assay can be generated where each binding event creates a unique response. These experiments are more commonly known as indicator displacement assays (IDA), where the host forms a complex with an indicator molecule, which is then disrupted by the target molecule, in turn generating a signal response (Figure 1.12).[67,68] When the optical signal generated is a change in fluorescence, it is known as a fluorescence displacement assay (FDA). The indicator molecule is a fluorescent dye which can generate a change in the fluorescence emission of the system, *via* an enhancement or quenching of the overall signal.[69-73] Fluorophore binding must be strong enough to occur without the presence of the target molecule, but also sensitive enough to be displaced upon binding of the target.

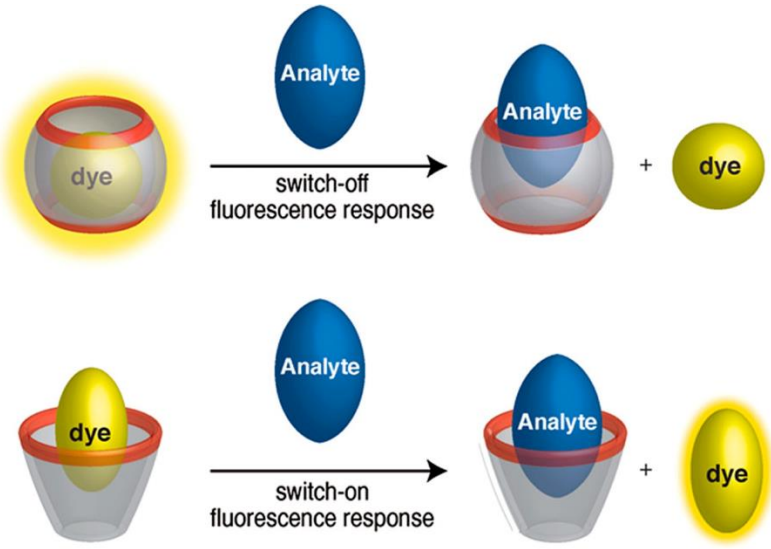


Figure 1.12. Cartoon representation of indicator displacement assays for sensing with macrocyclic hosts and a fluorescent dye, illustrating changes in the emission signal upon addition of the target molecule. Reproduced from ref. 73 with permission, Copyright 2014, American Chemical Society.

For successful IDAs the indicator molecule is reversible bound to the host, this allows for the differentiation of the equilibrium between the host and multiple guests. As the reporter motif is present in the bound indicator molecule, it does not need to be incorporated into the design of the receptor. A variety of known indicators molecules can be screened, or specific molecules can be tailored for these assays, resulting in an infinite combination of receptor:indicator pairs. For these receptor:indicator pairs to be implemented in sensing assays, they must each generate a detectable response upon binding. The recognition capabilities of the hosts can be coupled with real-time detection strategies of these signals to quantify binding affinities and guest selectivity.

The implementation of these assays allows for the analysis of the direct binding mechanism of a single molecule (Figure 1.13). By identifying these steps in a more simplistic manner, more information about the complex process of host:guest

complexation is revealed. The signal change that is monitored is modulated by the binding event of the guest to the host. For direct sensing the guest is directly bound to the receptor to produce this signal change.

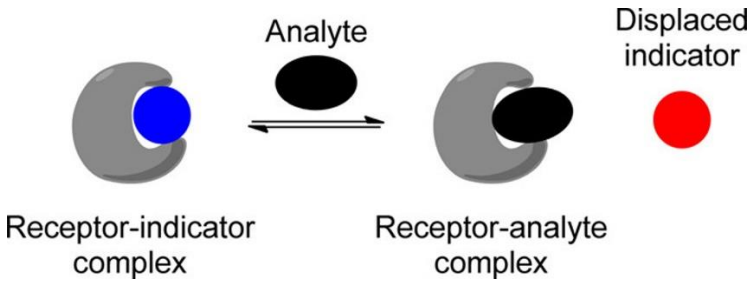


Figure 1.13. Illustration of a direct sensing mechanism utilized by indicator displacement assays. Reproduced from ref. 69 with permission. Copyright 2015, American Chemical Society.

This sensing mechanism is seen in enzymes and antibodies, which have a unique structural shape of their binding sites, that are compatible with only specific shaped substrates, meaning they will only interact with a limited number of targets. This concept of complementarity is influential for the design of synthetic receptors.[15,69,70] The successful discrimination of structurally similar targets is much more difficult, and the design of selective receptors for each of the subtle differences is impractical. One of the most efficient methods of implementing IDA for selectivity is templated synthesis. One such strategy is the self-assembly of multicomponent receptors, to quickly determine the best host for a specific target. This strategy is known as dynamic combinatorial chemistry (DCC) and allows for the efficient analysis of diverse host building blocks. These components can be combined *via* guest-templated synthesis, resulting in the formation of optimized receptors in the presence of the target guest. While this method eliminates the

need for involved and tedious synthesis of specialized receptor molecules for each desired target, a new challenge arises from the identification of good indicator and receptor pairs.

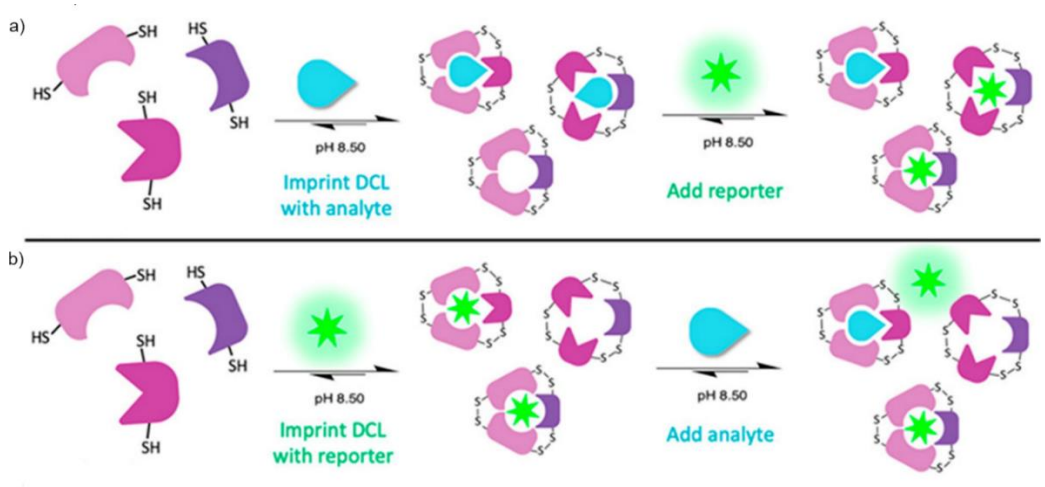


Figure 1.14. The imprint-and-report approach developed by Waters where a) shows the imprint method using the target analyte, and b) shows the imprint method using the reporter molecule. Reproduced from ref. 47 with permission. Copyright 2021, American Chemical Society.

Building upon this concept, the Waters group developed their "imprint-and-report" approach which greatly increased the throughput for the selection of suitable receptor:indicator combinations. They created large dynamic combinatorial libraries (DCL), that templated either the target or dye molecules with the building blocks of the host structures (Figure 1.14).[46,47] This generated a large data set of every distinct optical signal possible, and was subjected to statistical analysis, to determine which pairs use for further sensing. One advantage of this method is the dual paths of discovery for either receptor:indicator pairs, or analyte specific receptor pairs, based on the templation step.

The Waters group found indicator templation to be much more efficient, in respect to both time and materials, and therefore deemed this path more suitable for their sensing needs. Another benefit of this approach is that purification of the receptors was not necessary, allowing for their swift implementation in the impressive differentiation of all Arginine and Lysine methylation states of tested histone peptides.[44-47]

Generally, negatively charged and neutral targets interact with synthetic hosts predominately through electrostatic interactions, but the implementation of various indicator molecules can enhance their sensing capabilities. Most analytes targeted by differential sensing *via* macrocyclic receptors are small, nitrogen containing, "soft" cations. As previously described, these qualities make them the best guests for macrocyclic receptors. Alternatively, larger neutral guests, and negatively charged species, are far more difficult for these hosts to sense, especially in aqueous media. These elusive targets force the hosts to rely mainly on the electrostatic interactions of the host:guest complexes, generally resulting in weaker affinity. The indicator molecule remains present in the system after it is displaced, it can therefore interact with the host, guest, and host:guest complexes present throughout the binding event. Alternatively, non-cavity-based interactions of the indicator can also modulate fluorescence of the system, without any direct interaction with the host:guest complex. As a result of these secondary interactions, the indicator can act as the recognition motif, and the host as a signal modulator, through an indirect sensing mechanism.[71,72] In this regard, the host acts as a competitive sensing element with the indicator, and greatly increases the scope of target molecules capable of differentiation. This is where differential sensing comes in, which acts as a "chemical-nose", in the sense

that the targets are recognized by their combined response to all the receptors present in the array.[67] Instead of having each receptor specifically recognize one analyte, an array of receptors is used that recognize multiple analytes to various extents (Figure 1.15).

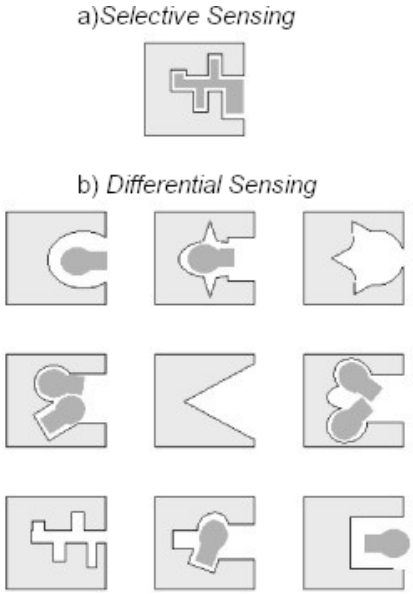


Figure 1.15. Depiction of analyte binding in host binding sites, where a) shows the highly specific binding event of a selective sensing mechanism, and b) shows a larger array of hosts interacting with one analyte with various responses illustrating differential sensing. Reproduced from ref. 67 with permission. Copyright 2001, Wiley.

This makes IDA more adaptable to array-based sensing, as a large diversity and cross-reactivity within the system is necessary for selective recognition. The strength of cross-reactivity is its ability to discriminate between complex mixtures of analytes composed of unknown components. This recognition is not possible with the direct sensing mechanisms, which focus on the enhancement of affinity and selectivity of a single target.

While the recognition of most guests relies heavily on interactions with the host, these qualities presented by indirect and differential sensing mechanisms allow for more challenging guests to be examined.

The data from these arrays produces a unique pattern for each analyte, which can be further analyzed using statical methods such as principal component analysis (PCA) and linear discriminant analysis (LDA).[69,74] Both methods produce scores plots with data represented for each analyte tested, which are generally displayed in a 2D coordinate system to illustrate discrimination. These methods apply the raw data to a matrix technique, which utilizes the eigenvectors to produce the axes of the coordinate graph. The eigenvalues represent a measurement of the level of discrimination of the data, and while the math is complicated, the analysis of the scores plot is relatively straightforward (Figure 1.16).[74] A good plot is seen when the repetitive data for an analyte is clustered closely, and minimal overlap between datasets from separate analytes is observed.

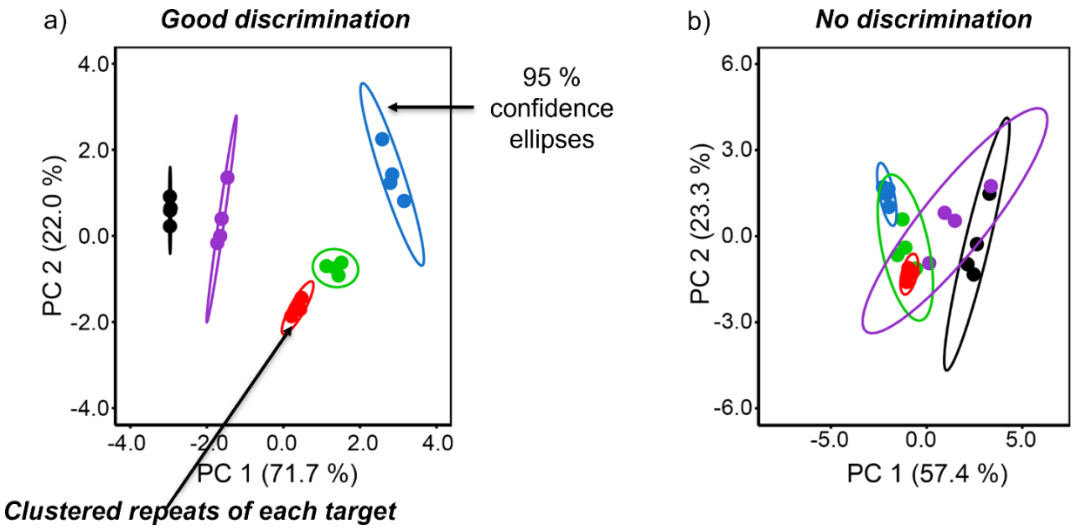


Figure 1.16. PCA scores plots illustrating, a) good discrimination showing no overlap between target dataset, and b) poor discrimination with distinct overlap amongst targets.

One drawback of this array-based system is the seemingly limitless combination of components that can be used for differentiation. These numerous pairings produce a large multidimensional data set, that is comprised of every unique signal response that is generated, regardless of how influential it is for discrimination. This enormity illustrates the limitations of the array-based system, as it is difficult and time consuming to extrapolate key data from the vast quantity available.[73] One method that can be implemented to efficiently quantify the most effective components from a large data pool, is machine learning.[75,76] This synergistic approach between synthetic and analytical techniques allows for a streamlined process from data collection to analysis for a broad scope of molecular recognition events.

1.5 References

1. Behr, J.-P. The Lock-and-Key Principle, 100 Years On; *John Wiley & Sons: Chichester, U.K.*, 2008.

2. Delves, P. J.; Martin, S. J.; Burton, D. R.; Roitt, I. M. Roitt's Essential Immunology. *John Wiley & Sons, Chichester,* 13th ed, 2017.

3. Beck, A.; Wurch, T.; Bailly, C.; Corvaia, N. Strategies and Challenges for the Next Generation of Therapeutic Antibodies. *Nat. Rev. Immunol.,* **2010**, *10*, 345− 352.

4. Bradbury, A.; Pluckthun, A. Reproducibility: Standardize Antibodies Used in Research. *Nature*, **2015**, *518*, 27−29.

5. Hunter, H.; Lawson, K. R.; Perkins, J.; Urch, C. J.; Aromatic Interactions. *J. Chem. Soc.*, **2001**, *5*, 651–669.

6. Ariga, K.; Ito, H.; Hill, J. P.; Tsukube, H. Molecular Recognition: From Solution Science to Nano/Materials Technology. *Chem. Soc. Rev.*, **2012**, *41*, 5800–5835.

7. Nishio, M.; Hirota, M.; Umezawa, Y. The CH-π Interactions: Evidence, Nature, and Consequences. *Wiley-VCH, New York*, 1998.

8. Goldman, S.; Bates, R. G. Focus on Cooperativity with Hydrogen Bonding. *J. Am. Chem. Soc.,* **2002**, *124*, 1476 – 1484.

9. Kearney, P. C.; Mizoue, L. S.; Kumpf, R. A.; Forman J. E.; McCurdy, A.; Dougherty, D. A. Molecular Recognition in Aqueous Media. *J Am Chem Soc.,* **1993**, *115,* 9907–9919.

10. Vincent, J. C. Specific Ion Interactions with Aromatic Rings in Aqueous Solutions. *Chem. Phys. Lett.*, **2015**, *638,* 1–8.

11. Ben-Amotz, D. Interfacial Solvation Thermodynamics. *J. Phys. Conden.*, **2016**, *28*, 414013-414022.

12. Tobias, D. J.; Stern, A. C.; Baer, M. D.; Levin, Y.; Mundy, C. J. Simulation and Theory of Ions at Atmospherically Relevant Aqueous liquid–air Interfaces. *Ann. Rev. Phys. Chem.*, **2013**, *64*, 339–359.

13. Pinalli, R.; Pedrini, A.; Dalcanale, E. Biochemical Sensing with Macrocyclic Receptors. *Chem. Soc. Rev.*, **2018**, *47*, 7006−7026.

14. Diederich, F.; Dick, K. A New water-soluble Macrocyclic Host of the Cyclophane Type. *J. Am. Chem. Soc*., **1984**, *106*, 8024–8036.

15. Petti, M. A.; Shepodd, T. J.; Dougherty, D. A. Design and Synthesis of a New Class of Hydrophobic Binding Sites. *Tetrahedron Lett.*, **1986**, *27*, 807–810.

16. Petti, M. A.; Shepodd, T. J.; Barrans, J. R. E.; Dougherty, D. A. "Hydrophobic" Binding of water-soluble Guests by high-symmetry Chiral Hosts. *J Am Chem Soc.*, **1988**, *110*, 6825–6840.

17. Smithrud, D. B.; Diederich, F. Strength of Molecular Complexation of Apolar Solutes in Water and in Organic Solvents is Predictable by Linear Free-energy Relationships. *J Am Chem Soc.*, **1990**, *112*, 339–343.

18. Blokzijl, W.; Engberts, J. B. F. N. Hydrophobic Effects: Opinions and Facts. *Angew. Chem. Int. Ed.*, **1993**, *32*, 1545–1579.

19. Ben-Amotz, D. Water-mediated Hydrophobic Interactions. *Annu. Rev. Phys. Chem.*, **2016**, *67*, 617–638.

20. Bakker, H. J. Water's Response to the Fear of Water. *Nature*, **2012**, *491*, 533–534.

21. Chandler, D. Oil in Troubled Waters. *Nature*, **2007**, *445*, 831–832.

22. Morita, T.; Westh, P.; Nishikawa, K.; Koga, Y. How Much Weaker are the Effects of Cations than those of Anions? The Effects of K^+ and Cs^+ on the Molecular Organization of Liquid H_2O. *J. Phys. Chem. B.*, **2014**, *118*, 8744 – 8749.

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

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

25. Gibb, C. L. D.; Gibb, B. C. Anion Binding to Hydrophobic Concavity is Central to the Salting-in Effects of Hofmeister Chaotropes. *J. Am. Chem. Soc.*, **2011**, *133*, 734 –7347.

26. Carnegie, R. S.; Gibb, C. L. D.; Gibb, B. C. Anion Complexation and The Hofmeister Effect. *Angew. Chem. Int. Ed.*, **2014**, *53*, 1149 –11500.

27. Jungwirth, P.; Cremer, P. S. Beyond Hofmeister. *Nat. Chem.*, **2014**, *6*, 261–263.

28. Fox, J. M. Interactions Between Hofmeister Anions and the Binding Pocket of a Protein. *J. Am. Chem. Soc.*, **2015**, *137*, 3859–3866.

29. Otten, D. E.; Shaffer, P. R.; Geissler, P. L; Saykally, R. J. Elucidating the Mechanism of Selective Ion Adsorption to the Liquid Water Surface. *Proc. Natl Acad. Sci.*, **2012**, *109*, 701–705.

30. Perera, P.; Wyche, M.; Loethen, Y.; Ben-Amotz, D. Solute-induced Perturbations of Solvent-shell Molecules Observed Using Multivariate Raman Curve Resolution. *J. Am. Chem. Soc.*, **2008**, *130*, 4576–4577.

31. Rembert, K. B. Molecular Mechanisms of ion-specific Effects on Proteins. *J. Am. Chem. Soc.*, **2012**, *134*, 10039–10046.

32. Pegram, L. M.; Record, M. T. J. Thermodynamic Origin of Hofmeister Ion Effects. *J. Phys. Chem. B.*, **2008**, *112*, 9428–9436.

33. Harries, D.; Rau, D. C.; Parsegian, V. A. Solutes Probe Hydration in Specific Association of Cyclodextrin and Adamantane. *J. Am. Chem. Soc.*, **2005**, *127*, 2184–2190.

34. Murray, J.; Kim, K.; Ogoshi, T.; Yao, W.; Gibb, B. C. The Aqueous Supramolecular Chemistry of Cucurbit[n]urils, Pillar[n]arenes, and Deep-Cavity Cavitands. *Chem. Soc. Rev.*, **2017**, *46*, 2479–2496.

35. Assaf, K. I.; Nau, W. M. Cucurbiturils: from Synthesis to High-affinity Binding and Analysis. *Chem. Soc. Rev.*, **2015**, *44*, 394–418.

36. Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.*, **2015**, *115*, 12320−12406.

37. Vinciguerra, B.; Cao, L.; Cannon, J. R.; Zavalij, P. Y.; Fenselau, C.; Isaacs, L. Synthesis and Self-assembly Processes of Monofunctionalized Cucurbit[7]uril. *J. Am. Chem. Soc.*, **2012**, *134*, 13133–13140.

38. Yao, L.; Sekutor, M.; Zavalij, P. Y.; Mlinarić-Majerski, K.; Glaser, R.; Isaacs, L. Cucurbit[7]uril Guest Pair with an Attomolar Dissociation Constant. *Angew. Chem., Int. Ed.*, **2014**, *53*, 988−993.

39. Lagona, J.; Mukhopadhyay, P.; Chakrabarti, S.; Isaacs, L. The Cucurbit[n]uril Family. *Angew. Chem., Int. Ed.*, **2005**, *44*, 4844–4870.

40. Daze, K. D.; Jones, C. E.; Lilgert, B. J.; Beshara, C. S.; Hof, F. Determining the Effects of Salt, Buffer, and Temperature on the Complexation of Methylated Ammonium Ions and Methyllysines by Sulfonated Calixarenes. *Can. J. Chem.*, **2013**, *91*, 1072–1076.

41. Daze, K. D.; Pinter, T.; Beshara, C. S.; Ibraheem, A.; Minaker, S. A.; Ma, M. C. F.; Courtemanche, R. J. M.; Campbell, R. E.; Hof, F. Supramolecular Hosts that Recognize Methyllysines and Disrupt the Interaction Between a Modified Histone Tail and its Epigenetic Reader Protein. *Chem. Sci.*, **2012**, *3*, 2695−2699.

42. Xue, W.; Zavalij. P. Y.; Isaacs, L. Pillar[n]MaxQ: A New High Affinity Host Family for Sequestration in Water. *Angew. Chem., Int. Ed.*, **2020**, *59*, 13313–13319.

43. Corbellini, F.; Fiammengo, R.; Timmerman, P.; Crego-Calama, M.; Versluis, K.; Heck, A. J. R.; Luyten, I.; Reinhoudt, D. N. Supramolecular Mimics of heme-protein Binding Sites. *J. Am. Chem. Soc.*, **2002**, *124*, 6569–6575.

44. Beaver, J. E.; Waters, M. L. Molecular Recognition of Lys and Arg Methylation. *ACS Chem. Biol.*, **2016**, *11*, 643–653.

45. James, L. I.; Beaver, J. E.; Rice, N. W.; Waters, M. L. A Synthetic Receptor for Asymmetric Dimethyl Arginine. *J. Am. Chem. Soc.*, **2013**, *135*, 6450-6455.

46. Li, J.; Nowak, P.; Otto, S. Dynamic Combinatorial Libraries: From Exploring Molecular Recognition to Systems Chemistry. *J. Am. Chem. Soc*., **2013**, *135*, 9222− 9239.

47. Harrison, E. E.; Carpenter, B. A.; St. Louis, L. E.; Mullins, A.G.; Waters, M. L. Development of "imprint-and-report" Dynamic Combinatorial Libraries for Differential Sensing Applications. *J. Am. Chem. Soc.*, **2021**, *143*, 14845− 14854.

48. Mutihac, L.; Lee, J. H.; Kim, J. S.; Vicens, J. Recognition of Amino Acids by Functionalized Calixarenes. *Chem. Soc. Rev.*, **2011**, *40*, 2777–2796.

49. Hof, F. Host–guest Chemistry that Directly Targets Lysine Methylation: Synthetic Host Molecules as Alternatives to Bio-reagents. *Chem. Commun*., **2016**, *52*, 10093-10108.

50. Beatty, M. A.; Borges-González, J.; Sinclair, N. J.; Pye, A. T.; Hof, F. Analyte-Driven Disassembly and Turn-On Fluorescent Sensing in Competitive Biological Media. *J. Am. Chem. Soc.*, **2018**, *140*, 3500-3504.

51. Selinger, A. J.; Cavallin, N. A.; Yanai, A.; Birol, I.; Hof, F. Template-Directed Synthesis of Bivalent, Broad-Spectrum Hosts for Neuromuscular Blocking Agents. *Angew. Chem. Int. Ed.*, **2022**, *61*, e202113235.

52. Shaurya, A.; Garnett, G. A. E.; Starke, M. J.; Grasdal, M. C.; Dewar, C. C.; Kliuchynskyi, A. Y.; Hof, F. An Easily Accessible, Lower Rim Substituted Calix[4]arene Selectively Binds N,N-dimethyllysine. *Org. Biomol. Chem.*, **2021**, *19*, 4691-4696.

53. Ogoshi, T.; Kanai, S.; Fujinami, S.; Yamagishi, T. A.; Nakamoto, Y. para-Bridged Symmetrical Pillar[5]arenes: Their Lewis Acid Catalyzed Synthesis and Host–Guest Property. *J. Am. Chem. Soc.*, **2008**, *130*, 5022–5023.

54. Ogoshi, T.; Yamagishi, T. A.; Nakamoto, Y. Pillar-Shaped Macrocyclic Hosts Pillar[n]arenes: New Key Players for Supramolecular Chemistry. *Chem. Rev.*, **2016**, *116*, 7937–8002.

55. Xue, M.; Yang, Y.; Chi, X.; Zhang, Z.; Huang, F. Pillararenes, A New Class of Macrocycles for Supramolecular Chemistry. *Acc. Chem. Res.*, **2012**, *45*, 1294–1308.

56. Strutt, N. L.; Zhang, H.; Schneebeli, S. T.; Stoddart, J. F. Functionalizing Pill[n]arenes. *Acc. Chem. Res.*, **2014**, *47*, 2631–2642.

57. Wang, Y.; Ping, G.; Li, C. Efficient Complexation Between Pillar[5]arenes and Neutral Guests: from Host–Guest Chemistry to Functional Materials. *Chem. Commun.*, **2016**, *52*, 9858–9872.

58. Wu, X.; Gao, L.; Hu, X. Y.; Wang, L. Supramolecular Drug Delivery Systems Based on Water-Soluble Pillar[n]arenes. *Chem. Rec.*, **2016**, *16*, 1216–1227.

59. Hu, X. B.; Chen, Z.; Tang, G.; Hou, J. L.; Li, Z. T. Single-molecular Artificial Transmembrane Water Channels. *J. Am. Chem. Soc.*, **2012**, *134*, 8384–8387.

60. Cram, D. J.; Choi, H. J.; Bryant, J. A.; Knobler, C. B. Host-guest Complexation. Solvophobic and Entropic Driving Forces for Forming Velcraplexes, which are 4-fold, Lock-key Dimers in Organic Media. *J. Am. Chem. Soc.*, **1992**, *114*, 7748−7765.

61. Moran, J. R.; Ericson, J. L.; Dalcanale, E.; Bryant, J. A.; Knobler, C. B.; Cram, D. J. Vases and Kites as Cavitands. *J. Am. Chem. Soc.*, **1991**, *113*, 5707− 5714.

62. Jordan, J. H.; Gibb, B. C. Molecular Containers Assembled through the Hydrophobic Effect. *Chem. Soc. Rev.*, **2015**, *44*, 547−585.

63. Rudkevich, D. M.; Hilmersson, G.; Rebek, J., Jr. Self-folding Cavitands. *J. Am. Chem. Soc.*, **1998**, *120*, 12216−12225.

64. Biros, S. M.; Rebek, J., Jr. Structure and Binding Properties of Water-Soluble Cavitands and Capsules. *Chem. Soc. Rev.*, **2007**, *36*, 93−104.

65. Hooley, R. J.; Van Anda, H. J.; Rebek, J., Jr. Extraction of Hydrophobic Species into a Water-Soluble Synthetic Receptor. *J. Am. Chem. Soc.*, **2007**, *129*, 13464− 13473.

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

67. Lavigne, J. J.; Anslyn, E. V. Sensing A Paradigm Shift in the Field of Molecular Recognition: From Selective to Differential Receptors. *Angew. Chem., Int. Ed.*, **2001**, *40*, 3118-3130.

68. Sedgwick, A. C.; Brewster, J. T.; Wu, T.; Feng, X.; Bull, S. D.;Qian, X.; Sessler, J. L.; James, T. D.; Anslyn, E. V.; Sun, X. Indicator displacement assays (IDAs): The Past, Present and Future. *Chem. Soc. Rev.*, **2021**, *50*, 9−38.

69. You, L.; Zha, D.; Anslyn, E. V. Recent Advances in Supra-molecular Analytical Chemistry Using Optical Sensing. *Chem. Rev.*, **2015**, *115*, 7840− 7892.

70. Dsouza, R. N.; Pischel, U.; Nau, W. M. Fluorescent Dyes and their Supramolecular host/guest Complexes with Macrocycles in Aqueous Solution. *Chem. Rev.*, **2011**, *111*, 7941-7980.

71. Guo, C.; Sedgwick, A. C.; Hirao, T.; Sessler, J. L. Supramolecular Fluorescent Sensors: An Historical Overview and Update. *Coord. Chem. Rev.*, **2021**, *427*, 213560-213591.

72. Stich, M. I. J.; Fischer, L. H.; Wolfbeis, O. S. An Overview on Nanoparticles Commonly Used in Fluorescent Bioimaging. *Chem. Soc. Rev.*, **2010**, *39*, 3102-3114.

73. Ghale, G.; Nau, W. M. Dynamically Analyte-responsive Macrocyclic Host-fluorophore Systems. *Acc. Chem. Res.*, **2014**, *47*, 2150-2159.

74. Stewart, S.; Ivy, M. A.; Anslyn, E. V. The Use of Principal Component Analysis and Discriminant Analysis in Differential Sensing Routines. *Chem. Soc. Rev.*, **2014**, *43*, 70−84.

75. Goecks, J.; Jalili, V.; Heiser, L. M.; Gray, J. W. How Machine Learning Will Transform Biomedicine. *Cell*, **2020**, *181*, 92-101.

76. Jorner, K.; Tomberg, A.; Bauer, C.; Sköld, C.; Norrby, P. O. Organic Reactivity from Mechanism to Machine Learning. *Nat. Rev. Chem.*, **2021**, *5*, 240-255.

Chapter 2: Molecular Recognition Using Water-Soluble Deep Cavitands

2.1: Properties of Water-Soluble Deep Cavitands

The application of macrocyclic synthetic receptors for differential sensing has shown selectivity that approaches that of biomacromolecule sensors, such as antibodies, in a more cost-effective and timely manner.[1-4] The culmination of all these factors allows synthetic hosts to have excellent binding affinity for "soft" cationic guests. As a result, other soft polarizable ions can act as direct competitors. While ideal guests have been shown to be charged, it is important that they are not too highly charged, as that means that before complexation, they are highly hydrated by the aqueous media. This causes the desolvation energy for these guests to be quite large, and the penalty upon binding will be too great to overcome, which is why heteroatom-rich, highly charged, and zwitterionic guests are only shown to have weak affinity with most current synthetic hosts. The highly variable nature of these aqueous environments is essential for comparison with biofluids, which themselves are a mixture of salts, proteins, and organic metabolites all present at various concentrations. A majority of host:guest sensing has focused on the derivatization of anionic hosts, and the complimentary binding of soft cationic guests, with minimal consideration as to the effects of other ions in solution.

As a result of these ionic conditions, hosts that employ charged functional groups to confer solubility may face some challenges. The charged groups result in an amphiphilic host that is prone to aggregation, which triggers the formation of micelles, where all the charged regions are pointed outwards to maximize their contact with water. This leaves all the hydrophobic groups pointed internally and renders the binding capabilities of the host

moot. There are many conditions that can cause micelle formation, such as the addition of certain guests when the host concentration is very low, but special attention needs to be paid to the formation of micelles due to electrostatic interactions with the oppositely charged ions that are present in solution. When these oppositely charged species interact with the charged host, the structure is rendered neutral in charge, which can impact the overall solubility of the host, and even result in its precipitation out of solution. Not only does the ionic content of the media used affect solubility, but the counter ions of the hosts themselves can also play a role in the solubility of the host:guest complex. While there are ways to limit the effect of the ions on the host, it may come at a cost for solubility. The addition of neutral polar functional groups to the host structure is one method to increase solubility in water while avoiding complications due to electrostatic interactions from charged groups.

Water-soluble deep cavitands are easily derivatized to incorporate functional groups that confer both solubility and increased binding affinity for a wide scope of target molecules. The resorcinarene scaffold that cavitands are based upon can be expanded *via* a variety of different methods to produce a deeper binding cavity. This was first demonstrated by Cram,[5,6] and later expanded by Rebek, which is the method used for the work outlined here.[7-10] Through a nucleophilic aromatic substitution rection with 1,2-difluoro-4,5-dinitrobenzene the octanitro cavitand (**2.1**) was produced, named for the eight nitro groups that decorate the upper rim, extending the cavity by the addition of four aryl rings linked to the phenolic groups of the resorcinarene scaffold. The octanitro cavitand can be reduced to the octaamino cavitand, through reduction of the nitro groups with tin

(II) chloride in the presence of hydrochloric acid (Figure 2.1). The octaamino cavitand (**2.2**), named for the eight reactive amine groups present at the upper rim, is primed for further functionalization due to presence of these amine groups.

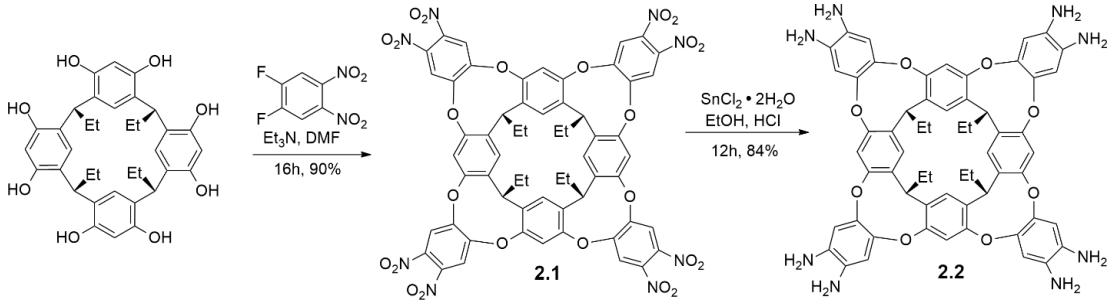


Figure 2.1. Expansion of the resorcin[4]arene scaffold towards the formation of a water-soluble deep cavitand.

One such functionalization is the conversion of the amino groups to amide groups through a Schotten-Baumann reaction of the amines with propionyl chloride, to result in the octaamide cavitand (Figure 2.2).[10] The octaamide cavitand (**2.3**) exhibits directional hydrogen bonding at its upper rim, arising from the intramolecular participation of both the oxygen and nitrogen atoms of the eight amide groups. The introduction of self-complementary hydrogen bonding groups at the upper rim of the cavitand can aid in the conformational stability of the host in water, based on which hydrogen bonding groups are used, and how they influence the self-folding properties of the cavitand.[11-13] The vase conformation is a 1:1 complex of the guest with the cavitand and is stabilized by the intramolecular hydrogen bonding between the carbonyl oxygen atom and amine group of the amide functionality present at the upper rim.

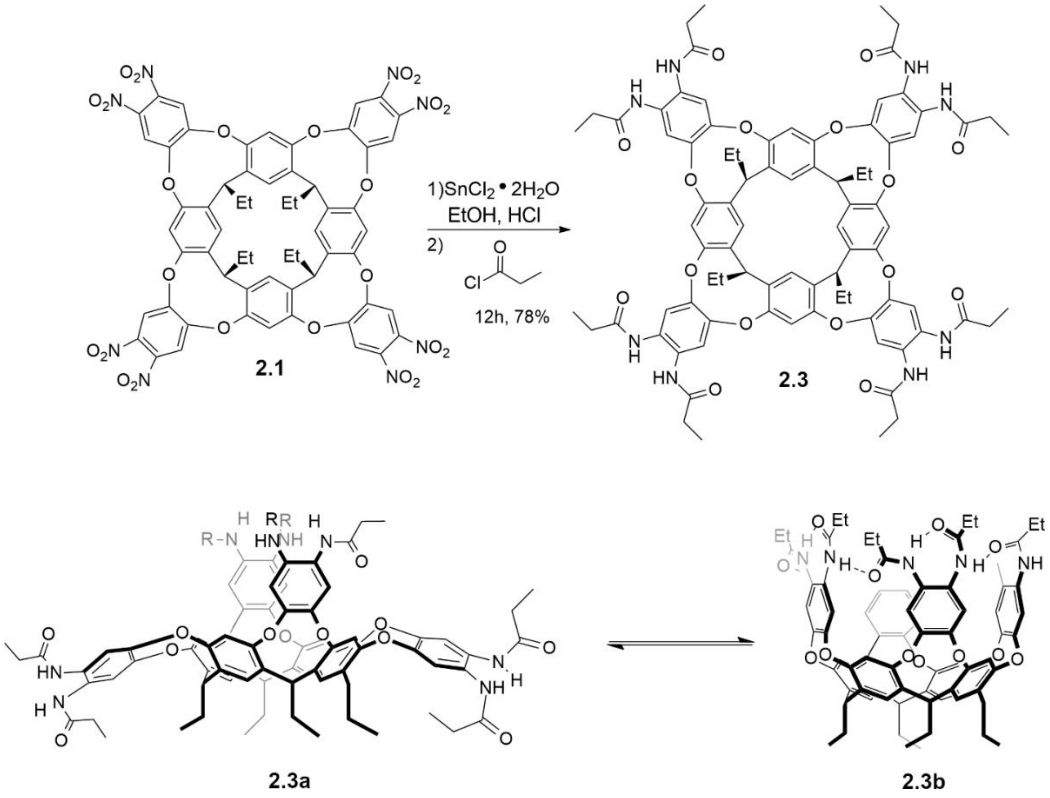


Figure 2.2. Functionalization of the cavitand upper rim from nitro groups (**2.1**) to amide groups (**2.3**) and the dynamic conformations of the octaamide cavitand, which is interchangeable between the open kite (**2.3a**) and folded vase (**2.3b**) conformations.

To synthesize a deep cavitand with greater kinetic stability, the octaamino cavitand can be functionalized to include a more rigid expanded benzimidazole scaffold (Figure 2.3). The inclusion of an ester moiety is accomplished *via* a reaction with the ethyl-3-ethoxy-3-iminopropionate HCl salt which results in the formation of the four benzimidazole rings at the upper rim and produces four protruding ester functional groups simultaneously. This tetraester cavitand (**2.5**) can then undergo hydrolysis, resulting in a negatively charged tetracarboxylate cavitand, **TCC** (**2.6**).[9,10] This additional expansion of the resorcin[4]arene scaffold with the benzimidazole rings, allows for the cavitand to

participate in intermolecular hydrogen bonding between four water molecules and the nitrogens in the benzimidazole ring. These interactions increase the stability of the vase conformation of the cavitand in water. In comparison to the octaamide cavitand (**2.3**), the benzimidazole cavitands are kinetically stable in water, and exhibit stronger binding constants, with slower guest exchange.[11-13] This is likely a result of the competing intermolecular hydrogen bonding exhibited by the octaamide cavitand in water.

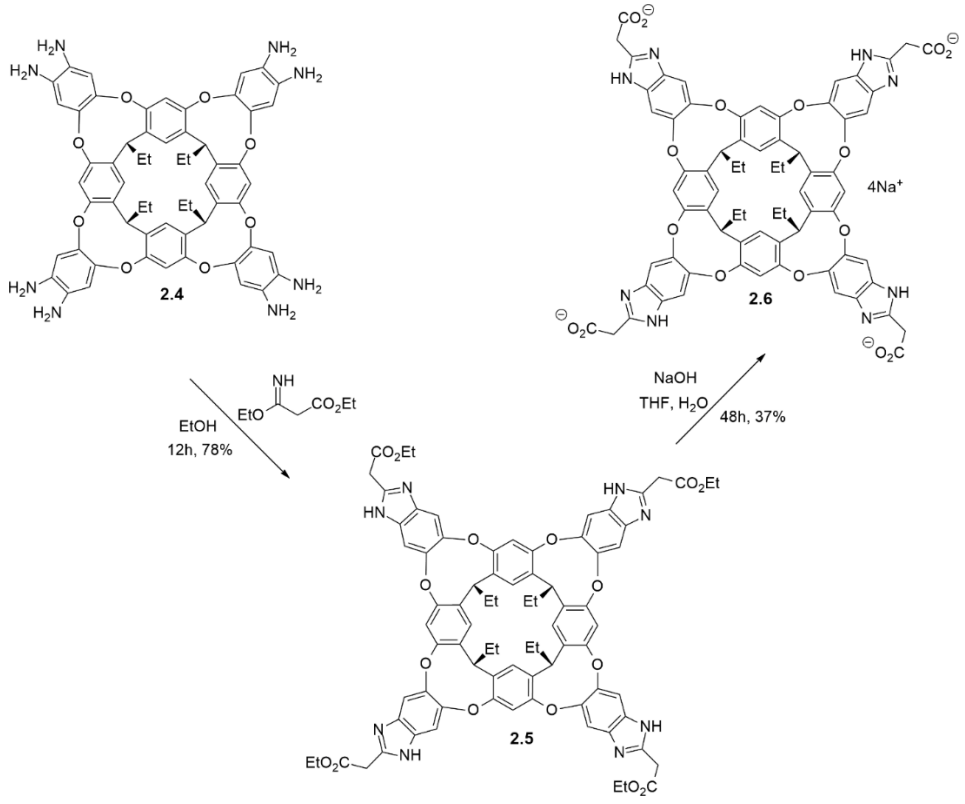


Figure 2.3. Synthesis of the **TCC** cavitand (**2.6**) from the tetraester (**2.5**) and octaamine (**2.4**) derivatives.

Another contributing factor for stability in the case of the **TCC** cavitand specifically, is the encapsulation of one molecule of THF during the saponification step, which enters the cavity as a more favorable guest than the surrounding water molecules, until it is displaced by a more desirable guest.[11] The tetracarboxylate cavitand exhibits high (~20 mM) water solubility due to the negatively charged upper rim and short ethyl feet, and readily extracts hydrocarbons into its hydrophobic pocket. **TCC** is able to bind a wide variety of guests, such as *n*-alkanes, which coil into a helical conformation in the cavity, and the guests can be observed rapidly tumbling on the NMR timescale, as well as cycloalkanes, that fit nicely into the pocket of the cavitand.[11] This deep cavitand is also able to selectively recognize molecules based on shape and size, as well as if they possess a positive charge at their surface, such as trimethylammonium salts.[12,13] This size-discrimination is illustrated through the poor binding affinity of the cavitand for triethylammonium salts, which bind much more weakly than those of the trimethyl equivalents. The triethylammonium moieties still possess a cationic charge that should exploit the cation-π interactions with the aromatic pocket, yet they are too bulky to fit efficiently inside the cavity. The binding affinity for trimethylammonium groups is stronger than those for hydrocarbons, yet in the presence of a long alkyl chain linked to a trimethylammonium group, the alkyl chain will show preference for binding inside the cavity over the positively charged group. The influence of the hydrophobic effect places the hydrophobic alkyl chain inside of the cavity, leaving the positively charged groups adjacent to the negatively charged upper rim of the tetracarboxylate cavitand, where they are free to exploit their electrostatic interactions.

The anionic **TCC** cavitand has been the most successful cavitand host for molecular recognition purposes in previous work.[17-25] Due to the orientation of the charged groups, which are localized at the upper rim of the host, the cavitand itself is lipophilic and prone to self-aggregation. This property has a large effect on the sensing of both charged and hydrophobic species, and greatly increases the target scope. Additionally, this feature was exploited in membrane bilayers, and other cellular environments, where the **TCC** host was able to act as a sensor.[14-16] Using the **TCC** host, as sensing system for modified histone H3 peptides was established. This system employed two different types of fluorescent indicators, fluorescein (**2.7**) and rhodamine isothiocyanate (**2.8**) derivatives with cholamine binding handles (Figure 2.4), that were tested *via* indicator displacement assays. This array system showed selective recognition of methylated lysine, with a strong preference for trimethylated modifications.[17,18]

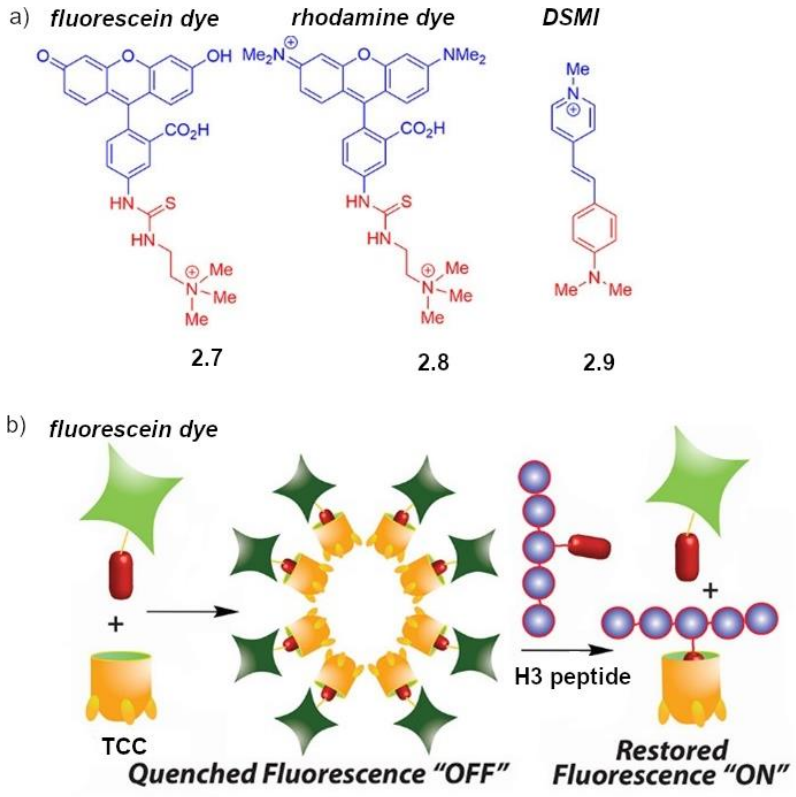


Figure 2.4. a) Fluorescent dye molecules that bind with the **TCC** host, cholamine derivatives of fluorescein (**2.7**) and rhodamine isothiocyanate (**2.8**), as well as a styrylpyridinium dye (**2.9**) and b) the sensing mechanism of the **TCC**:fluorescein complex with the H3 modified peptide guest.

The upper rim functional groups of the cavitand host are oriented towards the peptide backbone, allowing for secondary interactions with neighboring side chains. The sensing mechanisms of the **TCC** host are different with each dye molecule. Upon binding, the **TCC**:fluorescein pair results in drastic quenching of the fluorescent signal, due to the triggered aggregation of micelles.[17-20] When the modified peptide target is introduced to the system, the micelles are disrupted by the formation of the host:guest complex. The dye

molecule is displaced, and the fluorescence signal is recovered. These non-cavity-based interactions extend the target scope through an indirect sensing mechanism.

This indirect sensing mechanism was then applied to other types of peptide modifications, such as phosphorylation and citrullination, by colleagues Dr. Lizeth Perez and Adam Gill. Because the anionic **TCC** host can interact with the cationic peptide groups *via* charge-based interactions instead of solely cavity-based recognition, an IDA where the host cavity interacts with the indicator dye while the upper rim of the host interacts with the target peptide.[17-20] This combined sensing assay required the use of a new dye molecule which would form an initial host:dye complex that could then be disrupted by the target peptide. This new dye was based on a styrylpyridinium backbone (Figure 2.4a), and possesses a positively charged region, which was believed to express a binding orientation similar to the cationic peptides, with the positively charged region buried inside the binding pocket. This initial dye has a dimethylamino headgroup and is known as **DSMI** for short (**2.9**), more details about the synthesis and properties of these dyes will be discussed later in the chapter. Using this **TCC:DSMI** system, the phosphorylation of alcoholic side chains of serine, tyrosine, and threonine amino acids were selectivity recognized.[21] As only the anionic **TCC** host can interact with the cationic peptide targets, additional non-cavitand elements were incorporated to achieve the greatest amount of differentiation amongst the guests. Luckily, the **TCC** host shows affinity for heavy metal ions in solution at micromolar affinity due to the orientation of the carboxylate groups and their ability to chelate large metals. Introducing these metal ions to the **TCC**:dye complexes greatly diversify the fluorescence responses generated.[22] The metal ions can interact with the

host:dye complex or displace the dye from the host cavity, and the response of the host:dye:metal complexes vary based in the metal ion used. Notably, heavy metals quench the fluorescence, while filled shell metals do not.[22] The addition of only 10-50 μM of these metal salts expands the array without the necessity for further host or dye synthesis. Using a similar array system, the citrullination of arginine residues of the histone peptides showed successful discrimination with work done by Adam Gill.[23]

Water-soluble deep cavitands are excellent candidates for molecular recognition as they possess binding cavities that are more flexible when compared to other synthetic receptors. These larger resorcinarene cavitands possess a deep binding pocket, that is defined, yet breathable, which allows them to accommodate a larger range of guest molecules. This deep binding pocket creates a single-entry portal for guest binding, but when the large cavity is left empty, the cavitands are still capable of molecular recognition *via* guest templation. The properties of cavitands are affected not only by what is bound inside their cavity, but their external environment as well.

2.2 Expansion of Target Scope *via* Lower Rim Functionalization

The use of water-soluble deep cavitands allows for greater scope and tunability of their recognition abilities compared to other macrocyclic receptors. Deep cavitand hosts exhibit this selective target recognition by utilizing their self-folding, deep pocket, that provides a defined cavity for binding. This work focuses on the synthesis of novel deep cavitand structures to increase their target variability, as well as enhancing their solubility in aqueous environments. Past work has shown that these deep cavitand structures can bind

a wide variety of small molecule targets in both aqueous media and biological fluids, such as saliva and urine.[24,25] Much of this previous work has explored the capability of deep cavitands with a negatively charged upper rim, so new structures that incorporate a positive charge were synthesized as a direct comparison of functionality. Synthetic receptors show an overall bias for positively charged targets, but these derivatizations aim to achieve selectivity for neutral, and negatively charged guests with cationic hosts.

Functionalization of cavitands is difficult as the reaction must occur between four and eight times for complete transformation. Therefore, chemical reactions that lead to the desired product are often limited to those that can be achieved cleanly in full conversion, as the products cannot be purified by standard methods such as column chromatography. The lower rim of the cavitand is generally easier to derivatize as there are only four reactions that need to happen. To begin the lower rim transformations the ethyl feet are replaced with a longer alkyl chain that ends in an alcohol group (Figure 2.5). This is done by switching out the aldehyde used in the initial condensation step to dihydropyran, resulting in an OH-footed cavitand **2.10**. The solubility of the cavitand structure is greatly enhanced with the inclusion of these longer alkyl chains when they are capped with hydrophilic alcohol terminal groups. The OH groups are also useful as they are easily replaceable with other functional groups for further derivatization. To begin this process the alcohol groups are switched out with chlorides using thionyl chloride. This reaction is very effective and full conversion to the alkyl chlorides is achieved, producing the Cl-footed cavitand (**2.11**). Once the installation of the halides is completed, the amount of possible functionalization at the feet of the cavitand should increase exponentially.

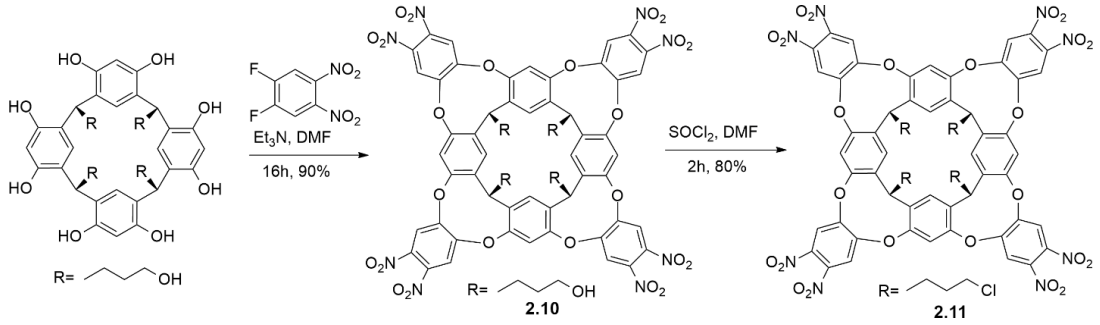


Figure 2.5. Synthesis of lower rim functionalized cavitand with various alkyl pendant groups.

Through a seemingly simple S_N2 reaction, the four terminal chloride atoms can be replaced with many desired nucleophiles. In this vein, numerous nucleophiles were selected and their substitution reactions were attempted. Through much trial and error, it was determined that the best nucleophiles for lower rim modifications were heterocyclic compounds, generally nitrogen-containing. The solubility of the Cl-footed cavitand became an issue, as the substitution reactions require a solvent that dissolves both the cavitand and the nucleophile. Many solid nucleophiles proved unsuccessful, as the addition of a solvent that was not the nucleophile itself complicated the reaction. The benzimidazole Cl-footed cavitand was expectedly less soluble in the polar aprotic solvents used for S_N2 reactions, so the focus was switched to the octamide Cl-footed cavitand, which exhibited much greater solubility in the necessary solvents. Not only is solubility throughout the reaction an important consideration, but the ability of the resulting substitution product to be isolated and purified is also a factor. While some polar aprotic solvents, such as DMSO, are capable of fully dissolving all the reactants, the resulting work up and isolation becomes

complicated. Therefore, nucleophiles that are liquids at room temperature showed the best results in the substitution reactions with the Cl-footed cavitands, as the reaction conditions could be carried out neat. The most successful of these nucleophiles were N-methylimidazole (NMI), and pyridine, which are both liquids with high boiling points, making them excellent candidates for this specific substitution reaction. Using these nucleophiles both the benzimidazole and octamide Cl-footed cavitand cationic derivatives were successfully synthesized (Figure 2.6). These new cavitands were easily isolated from the reaction mixture *via* precipitation, and all displayed promising solubility in water up to mM concentrations. As expected, the octamide derivatives were much more soluble in organic solvents and water than their benzimidazole based analogs. The use of these nucleophiles produced cationic cavitands that exist as a neutral salt with four chloride counter ions and have an overall 4^+ charge.

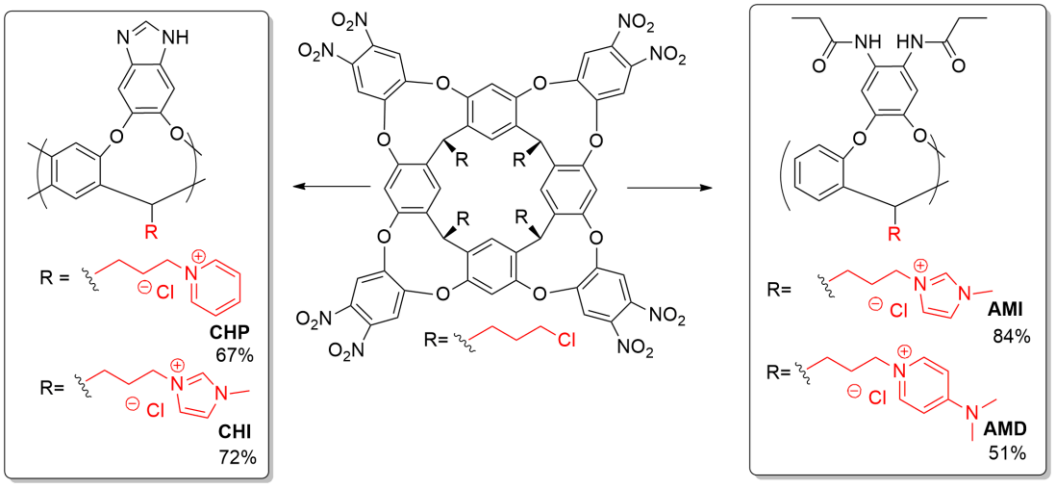


Figure 2.6. Successful cationic derivatives at the lower rims of the benzimidazole and octamide Cl-footed cavitands.

When other heteroatomic nucleophiles were attempted, the results were less successful. Despite identical reaction conditions, these non-nitrogen containing heteroatoms were poorer nucleophiles and full conversions to the cationic cavitands were not observed. These attempts included molecules such as oxazole, thiazole, and their methylated derivatives. While some of the nucleophiles showed partial reaction, with conversion of either two or three of the Cl groups, the reproducibility was variable. While generally less effective, some solid nucleophiles were successfully used without the use of solvent. This melt reaction was performed with solid samples of both the cavitand and nucleophiles. Upon heating the nucleophile was able to fully melt and solvate the cavitand to enable the substitution reaction. Again, the best reactants for this set of conditions were found to be nitrogen containing heterocycles, imidazole and dimethylamino pyridine (DMAP). These reaction conditions were only successful with the octamide Cl-footed cavitand. The octamide DMAP cavitand (**AMD**) is also cationic with similar properties concerning solubility as the NMI (**AMI** and **CHI**) and pyridine (**CHP**) footed cavitands described previously. However, the use of imidazole results in a cavitand that possesses a neutral charge at the lower rim, due to the NH group of the imidazole. This neutral cavitand species is capable of being attached to surfaces and can line the inside of capillary tubes. This static attachment allows the cavitand to be exploited for flow separation, which was attempted in collaboration with Ziting Gao (Zhong Group, UCR). Disappointingly this proved challenging, as the cavitand was not attached to the surface strongly enough and was easily displaced after the initial coating process.

With the success of these cationic derivatives, the synthesis of anionic lower-rimmed cavitands was explored. These reactions proved much more complex, as a simple substitution reaction could not install these anionic groups. A two-step route was proposed with the first step consisting of an S_N2 reaction using the Cl-footed cavitands to install functional groups at the lower rim that can be further reacted to be anionic in nature. This first step faced similar problems as with the attempts of the cationic derivatizations, in that the addition of a neutral group was challenging. Some attempts included the addition of thiol functional groups which could then be oxidized to a sulfone or linked with other sulfides to form disulfide bonds. Attempts with octanethiol were moderately successful with partial addition observed (Figure 2.7). The use of propanesultone as the nucleophile was less successful and no conversion to the sulfoxide was seen.

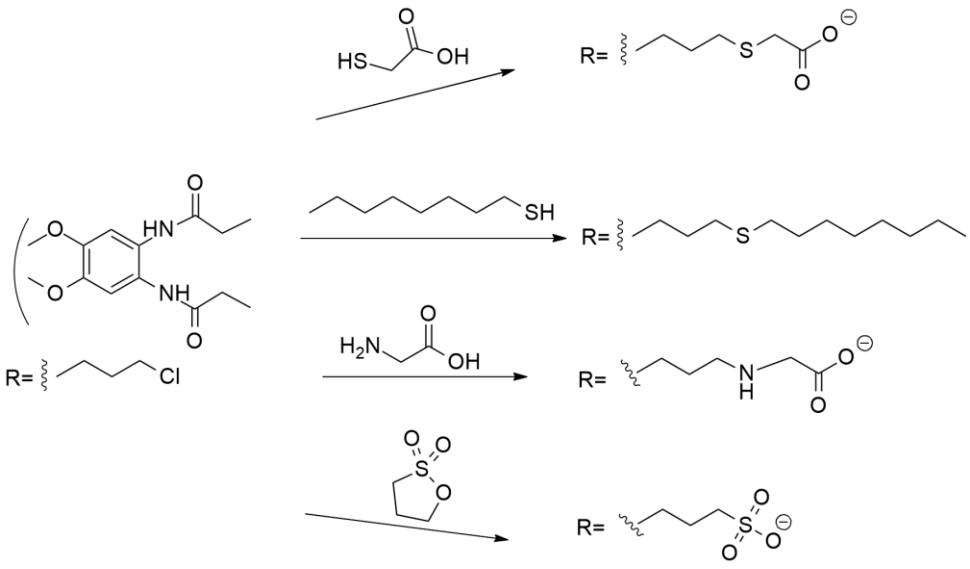


Figure 2.7. Attempts for lower rim derivatization with anionic or sulfide functional groups, performed with varying success.

As purely anionic groups proved difficult to add to the lower rim, the potential of a zwitterionic lower rim was explored (Figure 2.8). To do this, a suitable nucleophile was synthesized from ethyl bromoacetate and imidazole. This neutral molecule includes a nitrogen-containing heterocycle which was previously successful as a nucleophile in the formation of cationic cavitands, as well as an alkyl chain with an easily accessible ester group that can be hydrolyzed to a carboxylate ion (**2.12**). The addition of this small molecule to the octamide Cl-footed cavitand was successful after optimization of the reaction conditions (**2.13**). The hydrolysis of the ester portion of the feet proved much more challenging and despite multiple attempts, products arising from partially converted reactions were difficult to isolate, due to the zwitterionic and ionic liquid nature of the products.

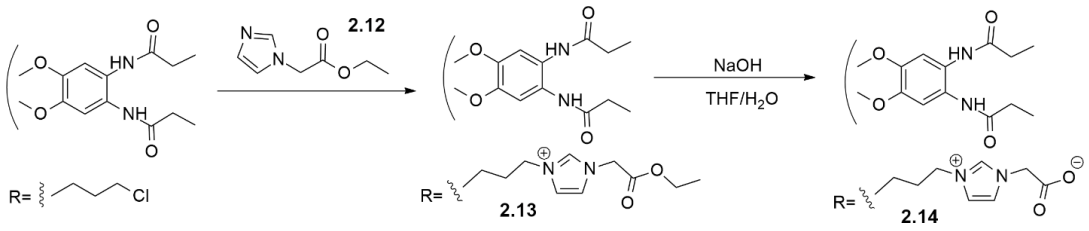


Figure 2.8. Synthetic route for the formation of a zwitterionic cavitand at the lower rim, starting from the octamide Cl-footed cavitand.

While complete reaction at all four lower rim positions is ideal, the partial reaction of even just one position would still be of interest for the structures' attachment to surfaces and analysis of the binding capabilities at interfaces. As the lower rim functionalization

does not directly affect the binding cavity of the cavitand structure, the incomplete functionalization of the lower rim groups is less deleterious to the overall functionality.

Applications of these newly synthesized cationic cavitands involve the recognition of negatively charged biomolecules, such as the many structures of DNA polyanions. With the previous success shown by the **TCC** cavitand to selectively recognize targets that do not bind in the host cavity, the new cationic cavitand hosts were implemented as fluorescence modulators and the indicator dyes were used as the recognition motif. This work was done by my collaborator, Junyi Chen (Zhong Group, UCR), and began with the investigation of highly structurally similar DNA G4 quadraplexes. Through the combination of both the hosts and dyes interacting with the DNA structures, the differential sensing of G4s with 23 different sequences, including parallel, antiparallel, and hybrid topologies, was achieved (Figure 2.9).[26] The previously mentioned styrylpyridinium dyes bind the G4 DNA strands with micromolar affinity, and the four cationic cavitand hosts, **AMI**, **CHI**, **CHP**, and **AMD**, also have strong affinities with these dyes, allowing for small changes in fluorescence emissions with the addition of the DNA guests. All the various styrylpyridinium dyes that were tested bind with the DNA and have differential emission enhancement when bound. Alternatively, the host:dye complexes can form ternary complexes with the DNA targets, the cationic cavitands can interact with the DNA backbone at their charged lower rims, while their binding pocket is occupied by the dye molecule.[26-28] These complex interactions are necessary for the high selectivity of the system, as discrimination between the DNA guests is not possible with the dyes alone, despite being the main reporters (Figure 2.9).

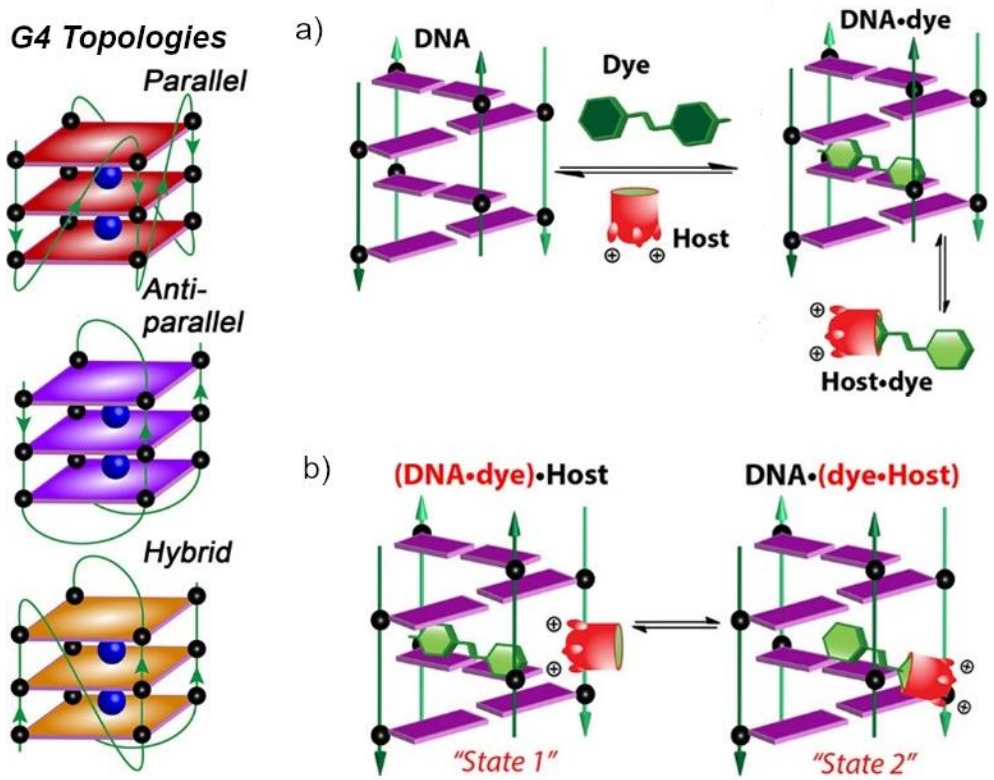


Figure 2.9. Sensing of different G4 DNA topologies and the a) direct sensing mechanism where the dye is the recognition motif and emission is changed upon its removal from the DNA structure, and b) the indirect sensing mechanisms of the possible ternary complexes of the host:dye:DNA.

Building upon this work, an improved array was developed with the novel cationic cavitands and a suite of four styrylpyridinium dyes, which was used to differentiate between nonclassical folded DNA structures, including hairpins, Hoogsteen triplexes, *i*-motifs, as well as imperfect G4s with vacancies or bulges.[27] From this large DNA pool, the most important array components were isolated by support vector machine (SVM) algorithms, and an optimal array with full discrimination of guests was produced.[27] Taking this discrimination one step further, the single base-modified structures of non-canonically

folded G4s was also explored. These structures differ only in the presence of a single oxidation of methylation modification in the guanine base, which results in subtle changes to their G4 folding.[28] These structural differences are often undetectable by traditional methods, such as circular dichroism, but the application of a cavitand array-based system showed the successful discrimination of these individual modifications.

These novel cationic cavitand structures were also implemented in the sensing of neutral hydrophobic molecules. The indirect sensing mechanisms were crucial for the selective recognition of biothiols[29] and steroids,[30] when paired with the **TCC** host and an assortment of heavy metal salts. The investigation of cannabinoids was performed by my colleague, Adam Gill, who saw the selective recognition of THC and its metabolites using a combination of direct guest binding and indirect sensing with the addition of metal salts.[31] This widespread variety of target molecules illustrates the abilities of water-soluble deep cavitands as high affinity hosts. Their accessible, tunable structures allow them to be functionalized to exploit multiple different recognition mechanisms for differential sensing when paired with appropriate indicator molecules.

2.3 Expansion of Target Scope *via* Upper Rim Modifications

While lower rim modification attempts were variably successful, similar attempts to modify the upper rim came with new challenges depending on the nucleophile and reaction conditions. The upper rim can either be derived with four or eight new functional groups, but more consideration as to which groups can be added is required, therefore making it harder to functionalize. Incomplete addition can result in adverse effects for the

dynamic nature of the self-folding process of the cavitand, and large bulky groups can block the binding cavity entirely. Additionally, incomplete reactions greatly reduce the overall solubility of the molecule, making these compounds difficult to purify. To minimize these problems, the benzimidazole scaffold was primarily used for upper rim modifications, as the addition of only four groups is necessary. The precedented synthesis of the **TCC** cavitand structure led to the design of various imidate salts to install four ester groups at the upper rim. These imidates can be synthesized in two ways, either in acidic or basic conditions, and allow for a variety of motifs to be added to the cavitand scaffold (Figure 2.10).

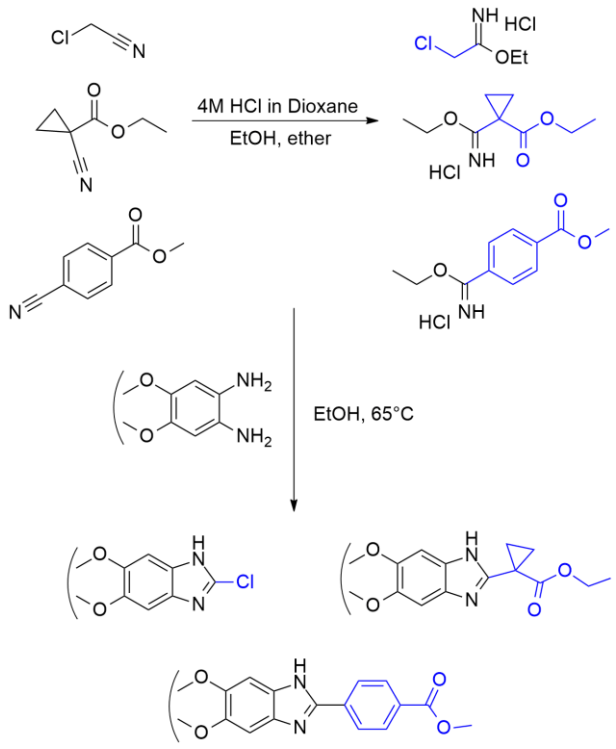


Figure 2.10. The synthesis of a) imidate salts with various functional groups, and b) the reaction of these imidates with the octaamine cavitand to produce newly functionalized cavitands.

The reaction of the octaamine cavitand **2.4**, with imidates creates a wide range of expanded benzimidazole cavitands with either four chloride atoms or ester motifs present at the upper rim (Figure 2.10). The final goal is the replacement of the chloride atoms with other functional groups, or the hydrolysis of the esters to create a diverse suite of carboxylated cavitands that have a negatively charged upper rim. The main drawback of this synthetic route is that the imidates themselves are difficult to synthesize, as they are highly reactive and both air and water-sensitive. The basic conditions for this reaction involve the use of sodium metal and the product is formed as a gooey liquid that is hard to isolate without decomposition. Acidic conditions are best performed with gaseous HCl, which is slightly impractical, therefore many attempts were done with various HCl sources with varying success. The synthesis of a chloro-imidate would allow for the addition of four Cl atoms at the upper rim, which would enable further functionalization; however, the resulting imidate placed the chloride atoms directly on the benzimidazole ring. The position of the Cl was not optimal for the desired substitution reactions and was therefore unsuccessful for upper rim functionalization. The attempts to synthesize different imidates were moderately successful as was their formation of the novel ester cavitands, but the hydrolysis reaction is where the biggest challenges have arisen.

The last attempts at upper rim modification were performed with the octamide scaffold. This scaffold is made directly from the octanitro cavitand **2.3**, and an acyl chloride in the previously described two-step process involving a reduction followed by a Schotten-Bauman reaction, to produce eight amide functional groups. Prior success was seen with propionyl chloride by Rebek,[10] and those optimized conditions were tested with other acid

chlorides to increase the diversity of the groups present in the upper rim groups (Figure 2.11). With the same intention as the attempts with the chloro-imidate, chloroacetyl chloride was used in these conditions to produce a novel octamide cavitand with eight chloride groups at the upper rim (**2.15**). From this structure an S_N2 rection with NMI was performed with the exact conditions used for the addition of cationic lower rims, and a cationic upper rim cavitand (**2.17**) was successfully synthesized. Sadly, these groups were too large and ultimately blocked the binding cavity, rendering this cavitand not conducive for binding. Since the main issue was the additional groups blocking the cavity, the next plan was to increase the distance between the cavity and the upper rim functionalization.

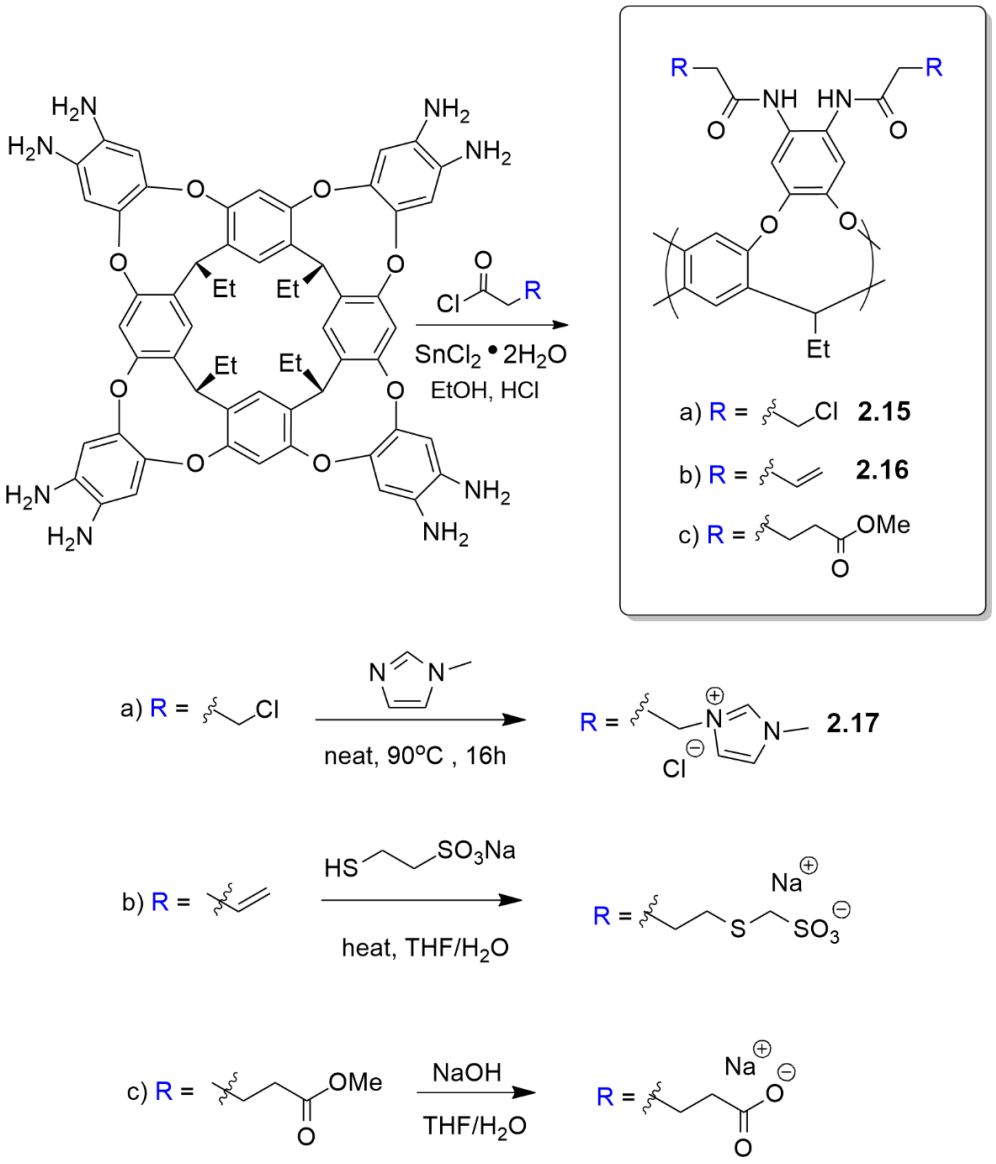


Figure 2.11. Synthesis of octamide cavitand derivatives with various functionalization at the upper rim.

The best results were obtained when using acryloyl chloride, which produced eight vinyl alkenes at the upper rim that can be used for further derivatization (**2.15**). Additional attempts were made with methyl 3-chloropropanoate, which would introduce eight ester

groups, which could again be further reacted to form a negatively charged carboxylate upper rim. These derivatives showed moderate success and were further pursued by colleagues in the lab.

2.4 Dye Synthesis and Characterization

Molecular recognition in water is the minimum standard for comparison to any biological analog, and binding in more complex media, in the presence of competitive analytes is the goal. In these more complicated aqueous environments, techniques such as indicator displacement assays are implemented to analyze the binding events of the host molecules more precisely. Commonly used reporter dyes, such as lucigenin, do not bind in cavitand structures. Therefore, a wide variety of other dye molecules were developed and tested for implementation *via* IDA. These indicator molecules include fluorescein isothiocyanate, rhodamine, and styrylpyridinium dye derivatives. The use of multiple dyes can help modulate the selectivity of the hosts based on the interactions of each host-dye pair. The binding environment also affects these reporter molecules, as they are both free and bound from the host, which produces various fluorescence responses.

Past work with the fluorescein and rhodamine derivatives showed excellent binding affinity with **TCC**. Upon binding, the dyes experience quenching of the overall fluorescence, and the reverse is observed once they are displaced from the binding cavity the reverse is observed.[24-26] These specific dyes and their derivatives are quite large; therefore, the exposed portion of the dye structure can induce aggregation of the host:dye complex, making the sensing capabilities complicated. Due to these and other factors, other

dyes were explored, namely those with a styrylpyridinium backbone. Initial work was performed with commercially available **DSMI** (Figure 2.12), which features a cationic pyridinium group connected to an aromatic ring *via* conjugation with a central alkene. Successful IDA with **DSMI** and a wide variety of cavitand hosts led to the development of more reporter molecules with this styrylpyridinium backbone. These dyes are all cationic, which allows them to have good affinity for the cavitand binding pocket and have excitation and emission wavelengths that are in a usable range for IDA.[32-34] They are easily synthesized through a Knoevenagel condensation reaction between the pyridinium salt and an aldehyde.

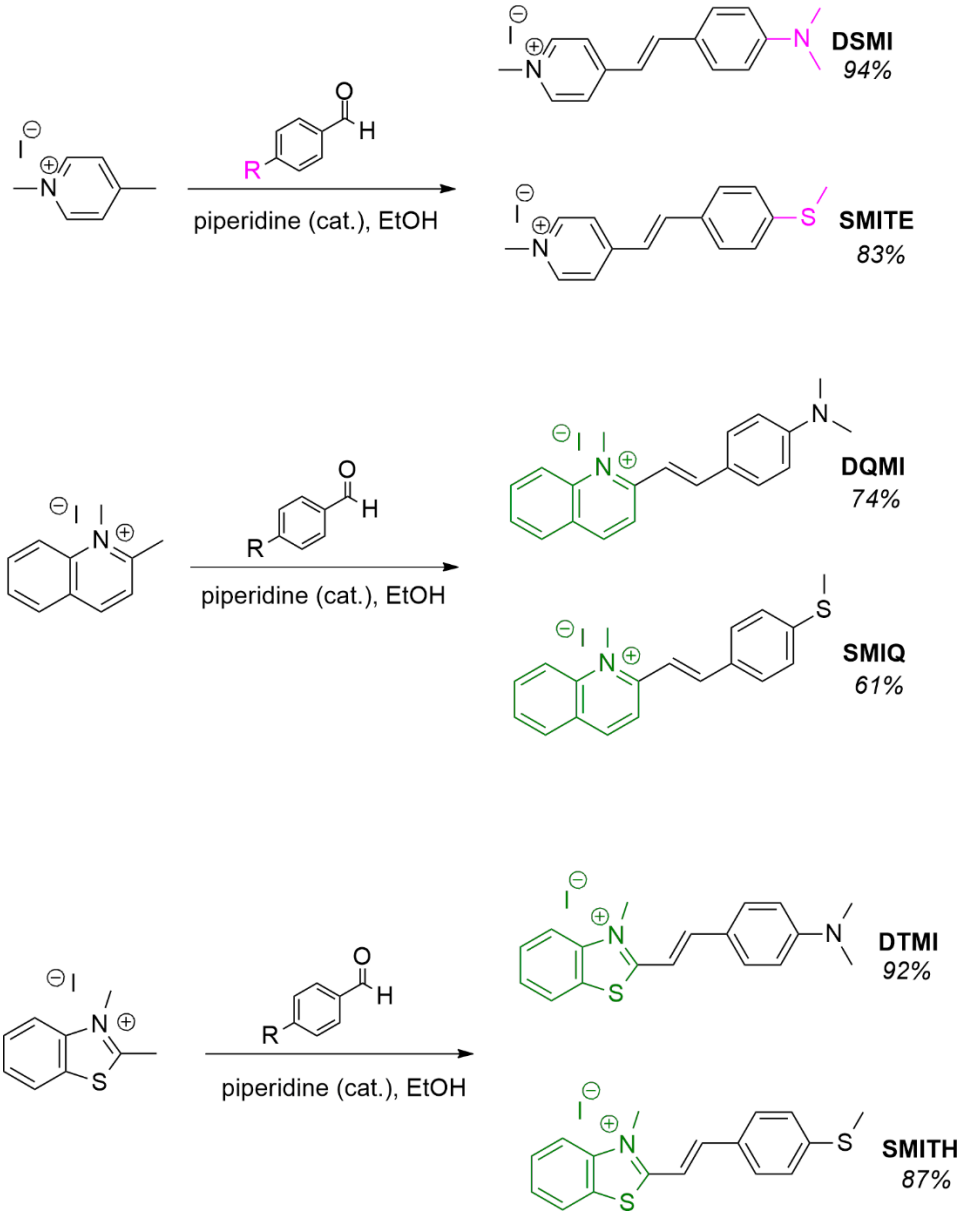


Figure 2.12. Successful synthetic routes for reporter dyes *via* Knoevenagel condensation reactions between pyridinium salts and dimethylamino or thiomethyl aldehydes.

The diversity of these dyes is twofold, as both the functional groups of the aldehyde and salt can be altered. Differences to the pyridinium salt include 1,4-dimethyl-pyridinium,

1,2-dimethylpyridinium, 1,2-quinoline, 1,2-benzothiazole, and 1,2-benzimidazole groups, and the aldehydes were focused to either dimethylamino or thiomethyl groups. While all combinations of these pyridinium salts and aldehydes were explored, the successful synthesis, by Alexie Raz, of a handful of these dyes was accomplished (Figure 2.12). The dyes with the thiomethyl groups produce much brighter fluorescence, and their wavelengths of excitation are higher than for the dimethylamino derivatives. When these cationic dyes are bound by the host, it was presumed that the charged region would prefer to be bound more deeply in the binding cavity, surprisingly the opposite was observed. Additionally, the thiomethyl derivatives show a greater binding affinity for the **TCC** cavitand, which makes them more difficult to displace when utilized in array applications. The binding affinity of the dyes for the cavitand hosts was measured using Supramolecular.org with a 1:1 model from the UV titration data. The calculated affinities are represented by K_d values, which is the concentration of the host:dye complex, selected values are seen in Table 2.1. The successful synthesis of this wide range of dye molecules produced a large suite of host:dye pairings that were conducive for array-based sensing systems.

K_d (μM)			
	TCC	**CHI**	**AMI**
DSMI	5.2	6.3	39.8
PSMI	10.1	18.6	66.1
SMIQ	11.6	3.7	13.3
DTMI	8.8	9.9	18.2

Table 2.1. Binding affinities measured in K_d for selected cavitand host and dye pairings.

These host:dye pairings were then used for the differential sensing of a variety of target molecules. Their application in IDA was exploited in both direct and indirect sensing mechanisms. These dyes can be displaced from the binding cavity of the hosts and can also interact with the targets molecules independently of the host:dye system. The observation of ternary complexes of the host:dye:guest components add an increased complexity to the molecular recognition capabilities of water-soluble deep cavitands.

2.5 References

1. Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.*, **2015**, *115*, 12320−12406.

2. Beaver, J.; Waters, M. L. Molecular Recognition of Lys and Arg Methylation. *ACS Chem. Biol.*, **2016**, *11*, 643− 653.

3. Pinalli, R.; Dalcanale, E. Supramolecular Sensing with Phosphonate Cavitands. *Acc. Chem. Res.*, **2013**, *46*, 399−411.

4. Ogoshi, T.; Yamagishi, T.; Nakamoto, Y. Pillar-Shaped Macrocyclic Hosts Pillar[n]arenes: New Key Players for Supra- molecular Chemistry. *Chem. Rev.*, **2016**, *116*, 7937−8002.

5. Cram, D. J.; Choi, H. J.; Bryant, J. A.; Knobler, C. B. Host-guest complexation. 62. Solvophobic and entropic driving forces for forming velcraplexes, which are 4-fold, lock-key dimers in organic media. *J. Am. Chem. Soc.*, **1992**, *114*, 7748−7765.

6. Moran, J. R.; Ericson, J. L.; Dalcanale, E.; Bryant, J. A.; Knobler, C. B.; Cram, D. J. Vases and Kites as Cavitands. *J. Am. Chem. Soc.*, **1991**,*113*, 5707− 5714.

7. Rudkevich, D. M.; Hilmersson, G.; Rebek, J., Jr. Self-folding Cavitands. *J. Am. Chem. Soc.*, **1998**, *120*, 12216−12225.

8. Hooley, R. J.; Rebek, J., Jr. Chemistry and Catalysis in Functional Cavitands. *Chem. Biol.*, **2009**, *16*, 255−264.

9. Biros, S. M.; Rebek, J., Jr. Structure and Binding Properties of Water-Soluble Cavitands and Capsules. *Chem. Soc. Rev.*, **2007**, *36*, 93−104.

10. Mosca, S.; Yu, Y.; Rebek, J., Jr. Preparative Scale and Convenient Synthesis of a Water-Soluble, Deep Cavitand. *Nat. Protoc.*, **2016**, *11*, 1371−1387.

11. Hooley, R. J.; Gavette, J. V.; Mettry, M.; Ajami, D.; Rebek, J., Jr. Unusual Orientation and Reactivity of Alkyl Halides in Water-Soluble Cavitands. *Chem. Sci.*, **2014**, *5*, 4382−4387.

12. Hof, F.; Trembleau, L.; Ullrich, E. C.; Rebek, J., Jr. Acetylcholine Recognition by a Deep, Biomimetic Pocket. *Angew. Chem., Int. Ed.*, **2003**, *42*, 3150−3153.

13. Hooley, R. J.; Van Anda, H. J.; Rebek, J., Jr. Extraction of Hydrophobic Species into a Water-Soluble Synthetic Receptor. *J. Am. Chem. Soc.*, **2007**, *129*, 13464−13473.

14. Perez, L.; Mettry, M.; Hinman, S. S.; Byers, S. R.; McKeating, K. S.; Caulkins, B. G.; Cheng, Q.; Hooley, R. J. Selective Protein Recognition in Supported Lipid Bilayer Arrays by Tailored, Dual-Mode Deep Cavitand Hosts. *Soft Matter*, **2017**, *13*, 3966-3974.

15. Perez, L.; Mettry, M.; Caulkins, B.G.; Mueller, L.J.; Hooley, R.J. Lipid Bilayer Environments Control Exchange Kinetics of Deep Cavitand Hosts and Enhance Disfavored Guest Conformations. *Chem. Sci.*, **2018**, *9*, 1836-1845.

16. Chen, J.; Tabaie, E. Z.; Hickey, B. L.; Gao, Z.; Raz, A. A. P.; Li, Z.; Wilson, E. H.; Hooley, R. J.; Zhong, W. Selective Molecular Recognition and Indicator Displacement Sensing of Neurotransmitters in Cellular Environments. *ACS Sens.*, **2023**, ASAP Article.

17. Daze, K. D.; Hof, F. The Cation−π Interaction at Protein−Protein Interaction Interfaces: Developing and Learning from Synthetic Mimics of Proteins That Bind Methylated Lysines. *Acc. Chem. Res.*, **2013**, *46*, 937−945.

18. Liu, Y.; Perez, L.; Mettry, M.; Easley, C. J.; Hooley, R. J.; Zhong, W. Self-Aggregating Deep Cavitand Acts as a Fluorescence Displacement Sensor for Lysine Methylation. *J. Am. Chem. Soc.*, **2016**, *138*, 10746−10749.

19. Liu, Y.; Perez, L.; Mettry, M.; Gill, A. D.; Byers, S. R.; Easley, C. J.; Bardeen, C. J.; Zhong, W.; Hooley, R. J. Site Selective Reading of Epigenetic Markers by a Dual-Mode Synthetic Receptor Array. *Chem. Sci.* **2017**, *8*, 3960-3970.

20. Liu, Y.; Perez, L.; Gill, A. D.; Mettry, M.; Li, L.; Wang, Y.; Hooley, R. J.; Zhong, W. Site-Selective Sensing of Histone Methylation Enzyme Activity via an Arrayed Supramolecular Tandem Assay. *J. Am. Chem. Soc.* **2017**, *139*, 10964-10967.

21. Liu, Y.; Lee, J.; Perez, L.; Gill, A. D.; Hooley, R. J.; Zhong, W. Selective Sensing of Phosphorylated Peptides and Monitoring Kinase and Phosphatase Activity with a Supramolecular Tandem Assay. *J. Am. Chem. Soc.* **2018**, *140*, 13869-13877.

22. Liu, Y.; Mettry, M.; Gill, A. D.; Perez, L.; Zhong, W.; Hooley, R. J. Selective Heavy Element Sensing with a Simple Host:Guest Fluorescent Array. *Anal. Chem.*, **2017**, *89*, 11113−11121.

23. Gill, A. D.; Hickey, B. L.; Wang, S.; Xue, M.; Zhong, W.; Hooley, R. J. Sensing of Citrulline Modifications in Histone Peptides by Deep Cavitand Hosts. *Chem. Commun.*, **2019**, *55*, 13259− 13262.

24. Hooley, R. J.; Zhong W. Applications of Synthetic Receptors in Bioanalysis and Drug Transport. *Bioconj. Chem.*, **2022**, *33*, 2245-2253.

25. Chen, J.; Hooley, R. J.; Zhong W. Combining Excellent Selectivity with Broad Target Scope: Biosensing with Arrayed Deep Cavitand Hosts. *Acc. Chem. Res.* **2022**, 55, 1035-1046.

26. Chen, J.; Hickey, B. L.; Wang, L.; Lee, J.; Gill, A. D.; Favero, A.; Pinalli, R.; Dalcanale, E.; Hooley, R. J.; Zhong, W. Selective Discrimination and Classification of G-Quadruplex Structures with a Host:Guest Sensing Array. *Nat. Chem.* **2021,** *13*, 488-495.

27. Chen, J.; Gill, A. D.; Hickey, B. L.; Gao, Z.; Cui, X.; E.; Hooley, R. J.; Zhong, W. Machine Learning Aids Classification and Discrimination of Noncanonical DNA Folding Motifs by an Arrayed Host:Guest Sensing System. *J. Am. Chem. Soc.* **2021,** *32*, 12791-12799.

28. Chen, J.; Hickey, B. L.; Raz, A. A. P.; Gao, Z.; Hooley, R. J.; Zhong, W. Sensing Base Modifications in Non-Canonically Folded DNA with an Optimized Host:Guest Sensing Array. *ACS Sens.* **2022**, 7, 2164–2169.

29. Liu, Y.; Duan, Y.; Gill, A. D.; Perez, L.; Jiang, Q.; Hooley, R. J.; Zhong, W. Metal-Assisted Selective Recognition of Biothiols by a Synthetic Receptor Array. *Chem. Commun.*, **2018,** *54*, 13147-13150.

30. Gill, A. D.; Perez, L.; Salinas, I. N. Q.; Byers, S. R.; Liu, Y.; Hickey, B. L.; Zhong, W.; Hooley, R. J. Selective Array-based Sensing of Anabolic Steroids in Aqueous Solution by Host:Guest Reporter Complexes. *Chem. Eur. J.*, **2019**, *25*, 1740-1745.

31. Gill, A. D.; Hickey, B. L.; Zhong, W.; Hooley, R. J. Selective Sensing of THC and Related Metabolites in Biofluids by Host:Guest Arrays. *Chem. Commun.* **2020**, *56*, 4352 - 4355.

32. Dsouza, R. N.; Pischel, U.; Nau, W. N. Fluorescent Dyes and Their Supramolecular Host/Guest Complexes with Macrocycles in Aqueous Solution. *Chem. Rev.*, **2011**, *111*, 7941–7980.

33. Fang. X.; Yongzan, Z.; Duan, Y.; Liu, Y.; Zhong, W. Recent Advances in Design of Fluorescence-Based Assays for High-Throughput Screening. *Anal. Chem.*, **2019**, *91*, 482−504.

34. You, L.; Zha, D.; Anslyn, E. V. Recent Advances in Supramolecular Analytical Chemistry Using Optical Sensing. *Chem. Rev.*, **2015**, *115*, 7840−7892.

Chapter 3: Diastereoselective Recognition of Pheromones by Water-Soluble Deep Cavitands

3.1 Introduction

Previous studies have found that positively charged and highly hydrophobic small molecules are the most suitable guests for recognition with synthetic receptors. While the sensing of racemic guests is relatively straightforward, the differentiation of stereoisomers and enantiomers is much harder to achieve, as the small structural changes of the target molecule that the recognition mechanism generally exploits, do not exist. Chiral recognition has been achieved with asymmetric metal-ligand cage complexes,[1-3] but these receptors still favor charged substrates. Other examples of selectivity between diastereomers exist with H-bonded organic capsules, but they lack solubility in water and perform much better in organic solvents.[4] The optical sensing of neutral, hydrophobic, chiral molecules is underexplored. The most notable example comes from Anslyn, who was able to determine enantiomeric excess for small molecule alkanols, amines, and saccharides.[5-7] This work was not focused on the recognition of the chiral centers of the molecules, and therefore could not discriminate between the different sizes of these small molecules. Similarly, Nau designed cucurbituril-based indicator displacement assays (IDA) that were able to determine enantiomeric excess, although this recognition was indirect, as the cucurbituril receptors did not bind the chiral species.[8] The stereospecific discrimination of guests is challenging for water-soluble cavitands, as they are not homochiral and have not shown diastereo- or enantiomeric preference for guest molecules. By employing cavitands that are capable of recognizing the alkyl components of the target,

the size and shape-based discrimination can be combined with the diastero- and enantioselective discrimination.

3.2 Pheromones as Targets and Initial Screening

One large area of biorelevant neutral molecules is pheromones which are used by insects for communication.[9] These small molecules are composed of hydrocarbon skeletons of varying sizes. They are robust in the functional group that they are decorated with and are generally hydrophobic. A desirable characteristic of interest is their homochiral nature, which allows for each isomer to play a different role biologically. For instance, fuscumol has been identified as an aggregation pheromone used by many families of the longhorn beetle.[10] A suite of five pheromone targets and their enantiomer, for a total of ten guests, were selected for investigation (**3.1**, **3.2**, **3.3**, **3.4**, **3.5**). Each of these target alkanols were either purchased or synthesized as single enantiomers, and their enantiopurity was determined *via* GC by the Millar group at UCR (Figure 3.1).[11-13] These compounds are excellent candidates for recognition with cavitands, as they are a good size and shape-matched for the binding cavity. Cavitands have been shown to be excellent hosts in the array-based discrimination of highly similar molecules with subtle differences in their structures.

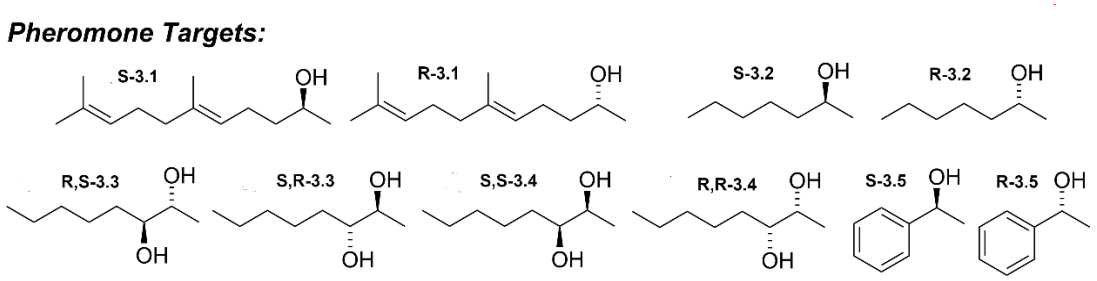


Figure 3.1. Pheromone targets selected for analysis.

Initial screening of target pheromones with various cavitand hosts showed that the alkanols are weakly binding guests. This result was expected, and further investigation with three cavitand hosts, **TCC** (**3.6**), **CHI** (**3.7**), and **AMI** (**3.8**), were carried out via indicator displacement assays (IDA). All three components for the IDAs, (*i.e.*, host, dye, and guest), are soluble in water, and tests were carried out in a buffered Tris solution (Figure 3.2). The IDAs were performed in a 96-well plate, and the change in fluorescence from host:dye complexes to host:guest complexes was measured. The effect of the indicator dyes used was also screened, and ultimately narrowed down to two similarly structured styrylpyridinium dyes, **DSMI** (**3.9**) and **SMITE** (**3.10**), that vary only at their terminal functional group. While the differences are subtle, (the presence of a dimethylamine versus a thiomethyl) their optical properties are similar to one another, including their excitation and emission wavelengths. All three cavitand hosts, and both dyes were used as components for the initial screening (Figure 3.2), to increase the number of suitable matches that could generate a measurable signal response from alkanol sensing in the host:guest system.

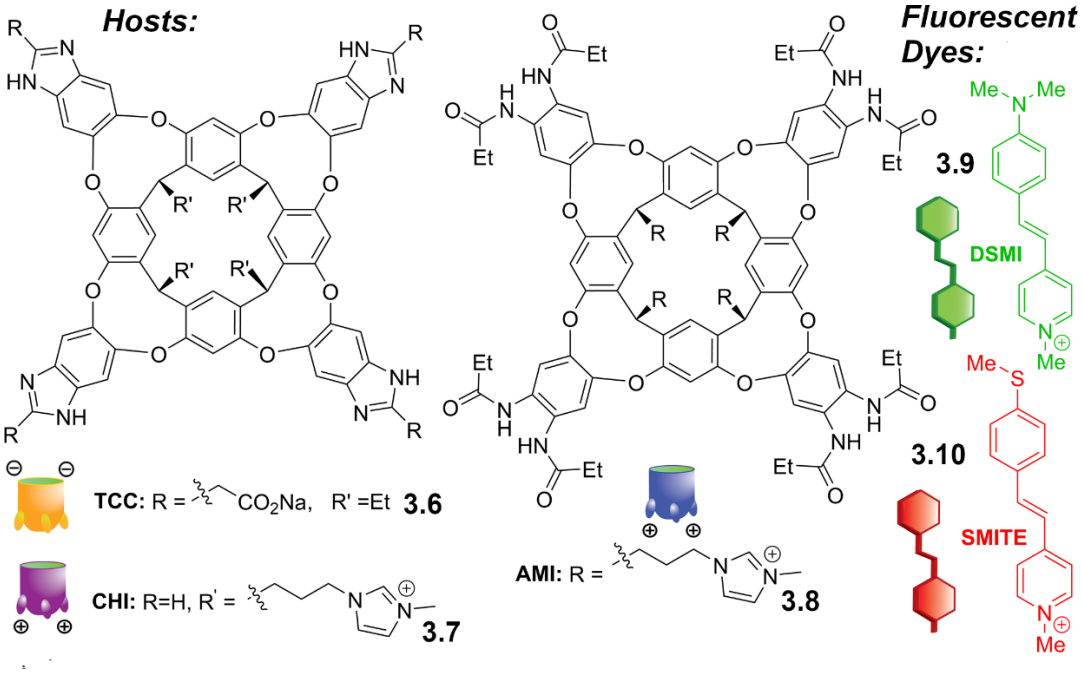


Figure 3.2. Host and dye components screened for application in array-based pheromone recognition.

When the cavitand receptors are coupled with reporter dyes in an IDA, the small changes in pheromone structure become more apparent for discrimination. The dye will initially bind with the cavitand, and its fluorescence properties will be quenched and turned off. The dye can be competitively displaced by the target pheromones, and the now free dye will regain its fluorescent properties, and turn on, generating a detectable difference of optical output (Figure 3.3). The signal response measured from these observed differences can then be subjected to multivariate analysis. These analysis techniques include things such as principal component analysis (PCA) and limited discrimination analysis (LDA), which organized the data into digestible graphs.

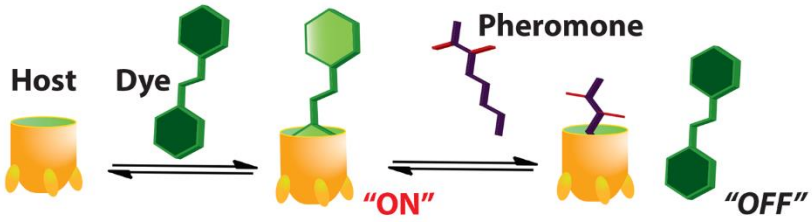


Figure 3.3. Indicator displacement sensing mechanism of the cavitand:dye:pheromone system.

This array system produces differentiated responses, but they were not as significant as was anticipated. The array was able to differentiate some of the more structurally different pheromones, such as **2.2** and **2.5**, from the rest of the targets, but the strongly similar pheromones showed poorer discrimination, and stereoselectivity was nonexistent. To enhance the changes in the optical response, heavy metal salts, such as $La(NO_3)_3$, $Ce(NO_3)_3$, and $UO_2(NO_3)_2$, were added to the array as additional components. In previous studies, the inclusion of heavy metal salts has shown to coordinate with the cavitand hosts and more finely modulates the signal responses generated.[14] This combination of host:dye:metal complexes did show a more diverse response when tested with the alkanol targets. Upon addition of the guests to this system, the fluorescence was noticeably lowered, indicating competitive binding of the guest and displacement of the dye from the binding cavity of the receptor. Initial analysis of the data from this large 24-component screen showed various results, ranging from large changes in the fluorescence response to more subtle and less relevant changes (Figure 3.4). All three of the cavitand hosts were able to sense each pheromone target, and both dyes produced measurable changes.

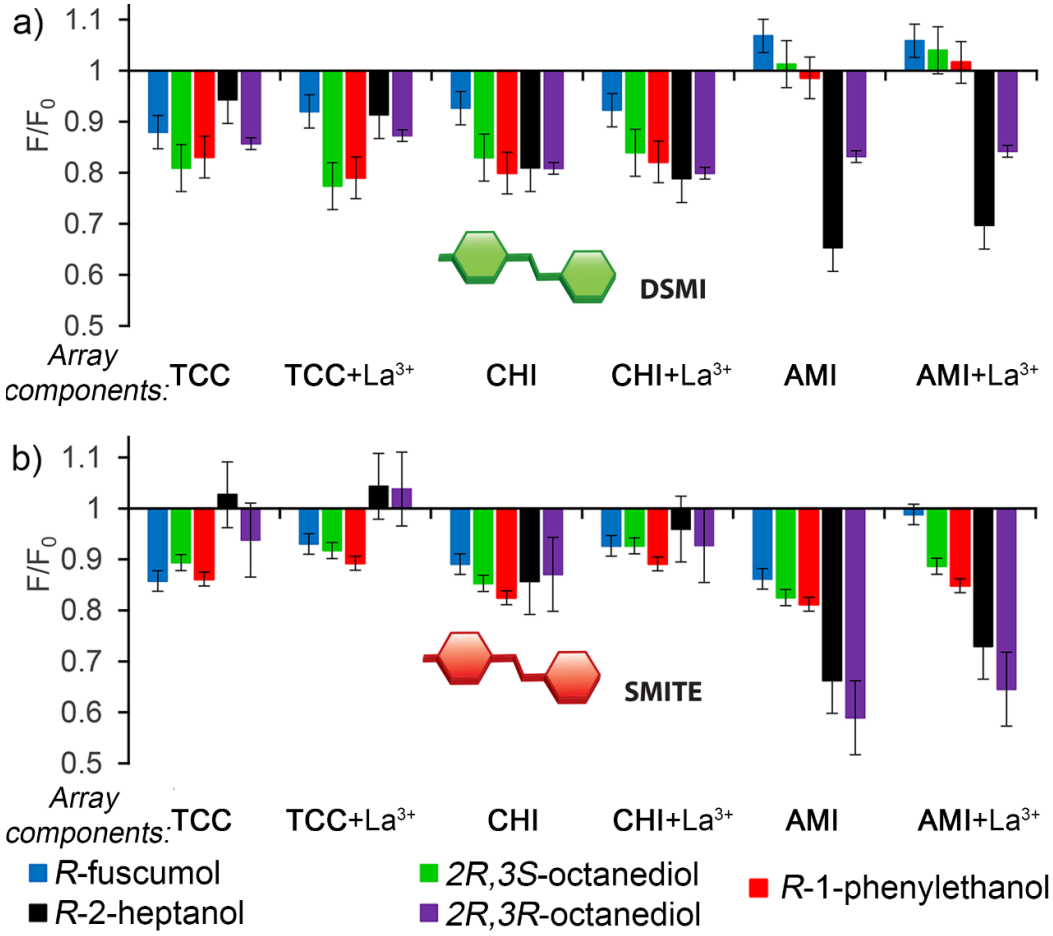


Figure 3.4. Initial optical detection of the relevant fluorescence responses shown as F/F_0, where F_0 is the Host•Dye•M^+ complexes and F is the measured change upon addition of pheromone targets with a) **DSMI** (**3.9**) dye and b) **SMITE** (**3.10**) dye in 20 mM Tris buffer, pH 7.4 [*Host*] =20 µM, [*Dye*] = 3.0 µM, [M^+] = 50 µM, and [*Pheromone*] = 50 µM.

The largest change in signal response was observed predominantly by the **AMI** host with either dye alone or with addition of the La^{3+} heavy metal salt. This large collection of data was subjected to PCA for discrimination, and the results were disappointing (Figure 3.5). Despite the large component pool, and measurable differences in fluorescence there was no diastereoselective discrimination observed. Some results were still promising, as

the discrimination of each pair of target isomers was accomplished by this array, so optimization of the components was the next step towards enhanced discrimination.

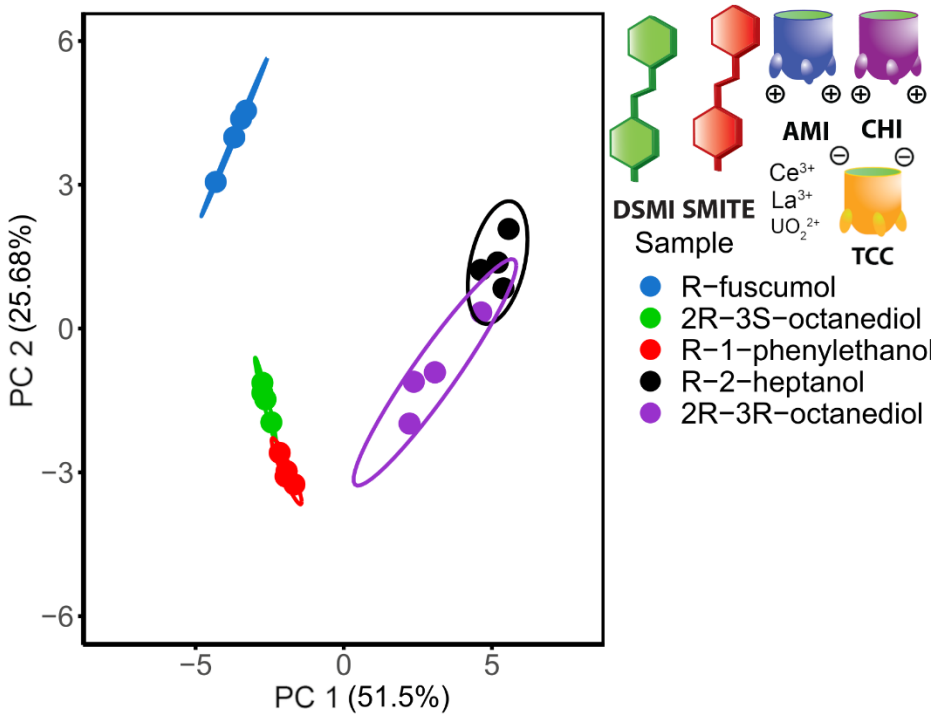


Figure 3.5. PCA scores plot derived from the data of the 24-component array, ellipses indicate 95% confidence intervals. 20 mM Tris buffer, pH 7.4 [*Host*] =20 µM, [*Dye*] = 3.0 µM, [M^+] = 50 µM, and [*Pheromone*] = 50 µM.

3.3 Data Optimization with Implementation of Machine Learning Algorithms

Although generally more data is seen as a positive, it became apparent that some combinations of the array components were ineffective for discrimination. The inclusion of these combinations in the data processing was damaging to the overall discrimination of the array. To optimize the data, the most impactful components needed to be identified and separated from the 24-component pool. As the sheer number of components for the array-

based system was unwieldy for effective and time efficient analysis by hand, machine learning algorithms were explored.[15-18] The benefits of machine learning algorithms lie in their ability to detect hidden patterns in large or noisy data sets. For the chosen method used, the data was treated with SVM-RFE (support vector machine recursive feature elimination) functions from the sklearn library in Python 3.9, which works to recursively remove non important features of the data.[19,20] In this manner each individual component was screened, in collaboration with Junyi Chen (Zhong group, UCR), and then ranked based on its relevance for discrimination (Figure 3.6).

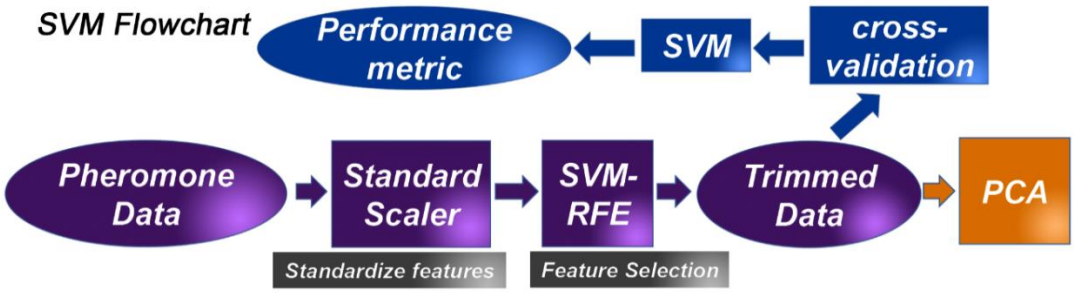


Figure 3.6. Machine learning optimized sensing flowchart illustrating the SVM-FRE process for determination of most relevant components.

From these results, the six highest ranked components were selected for their optimal discrimination properties. The top six components were heavily populated by the **AMI** and **CHI** cavitands, despite their relatively small changes in fluorescence when compared to the more noticeable differences of the **TCC** cavitand.

SVM-Optimized Array Components

Component	Rank
CHI+DSMI+Ce	1
AMI+DSMI+Ce	1
AMI+DSMI	2
CHI+SMITE	3
AMI+DSMI+La	4
AMI+SMITE+Ce	5

Sub-optimal Array Components

Component	Rank
TCC+SMITE+La	23
TCC+SMITE+Ce	22
AMI+DSMI+U	21
CHI+DSMI+U	20
CHI+DSMI+La	19
TCC+SMITE+U	18

Table 3.1. Tables of the six best (optimized) and worst (sub-optimized) array components as ranked by the SVM-FRE algorithm.

Additionally, the six lowest ranked components were also analyzed as a control study, of which were predominated by the **TCC** cavitand (Table 3.1). This low performance seen by the **TCC** host is surprising, as it is the most water-soluble of all the tested cavitands and displays the largest effect from heavy metals. Both sets were again processed with PCA with vastly different results (Figure 3.7). The optimized array showed discrimination of the stereoisomers, as well as the different diastereomers of the same compound, **3.3** and **3.4**. This also confirms that the inclusion of the cavitands is required, as the dye components alone gave no discrimination.

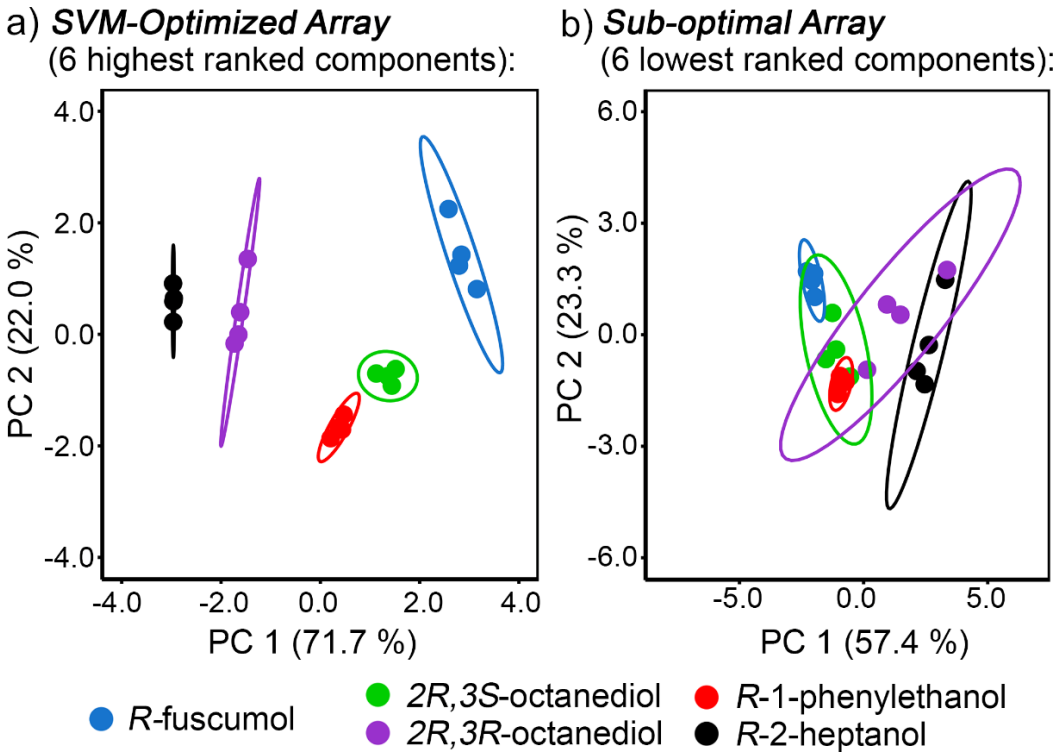


Figure 3.7. PCA scores plots derived from data using a) the optimized array with the 6 most important components, and b) the sub-optimal array using the 6 least important components, ellipses indicate 95% confidence intervals. 20 mM Tris buffer, pH 7.4, [*Host*] =20 µM, [*Dye*] = 3.0 µM, [M^+] = 50 µM, and [*Pheromone*] = 50 µM.

Further investigation into the recognition mechanism of these host:guest complexes was performed *via* ^{1}H NMR spectroscopy. The affinity of the alkanol guest for the **AMI** and **CHI** hosts was confirmed with guest titration experiments. The NMR spectra of select guests with the **AMI** and **CHI** hosts are indicative of rapid in/out exchange (Figure 3.8).

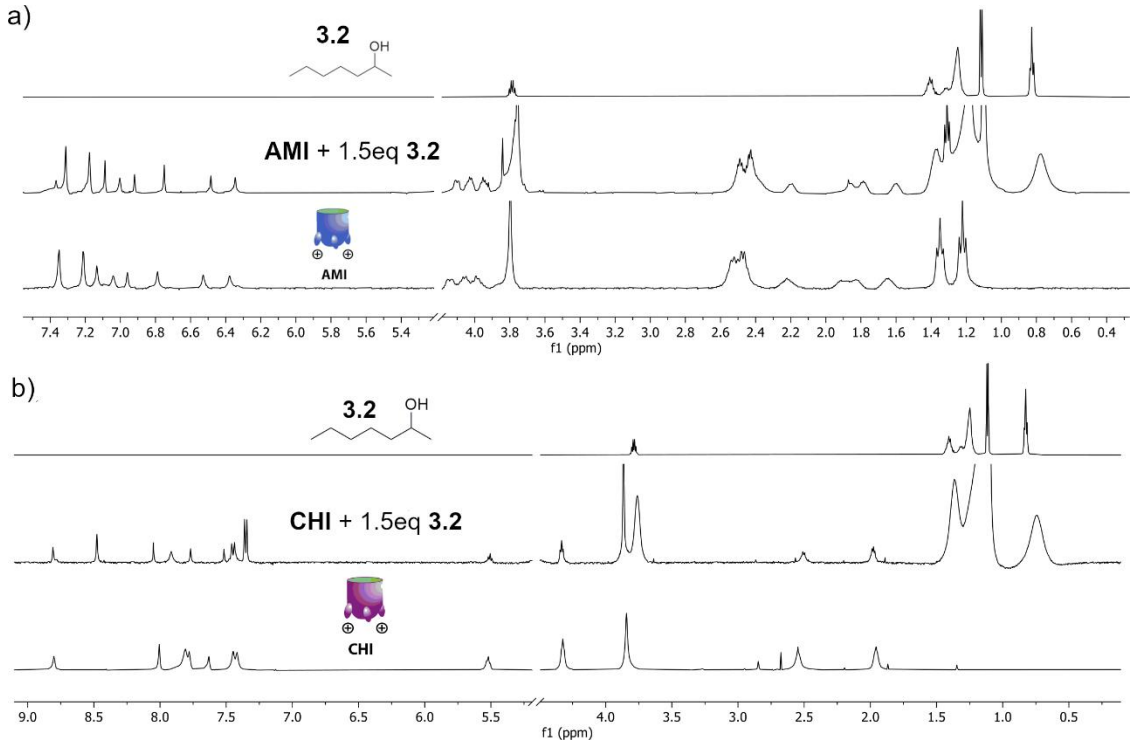


Figure 3.8. ^{1}H NMR spectra (400 MHz, D_2O, 298K) showing rapid in and out exchange of guest **3.2** with a) **AMI** host, where an upfield shift of the methyl group of guest **3.2** is observed in the aliphatic region with addition of **3.2** to the host and with b) **CHI** host where the same upfield shift of the methyl group of guest **3.2** is present in the aliphatic region, along with the sharpening of the aromatic peaks, with addition of **3.2** to the host.

The **TCC** host has a much higher affinity for the dye molecules, and tightly binds them, thereby inhibiting their displacement. Without a competitive binding mechanism, the alkanol guests cannot bind with the **TCC** host, and the dye remains bound, and no change in signal response is observed. Therefore, the affinity of the pheromone targets and reporter dyes is more evenly matched for the **AMI** and **CHI** cavitands. This balance between guest-dye recognition is a more useful combination for discrimination, as competitive displacement can occur. While the signal responses generated by these components are

visually much smaller, their application in the array-based system is more impactful overall.

3.4 Effects of Chiral Additives on Stereoselective Discrimination

While the application of these optimized components was able to show diastereoselective discrimination, enantioselective discrimination was still lacking. As the hosts themselves are homochiral, the addition of a chiral additive component was necessary. Multiple options were surveyed, including L-(+)-tartaric acid, $Eu(hfc)_3$, methyl-D-glucopyranoside, and sodium potassium tartrate. When paired with the heavy metal salt components already utilized in the array components, these additives were able to aid in the formation of larger chiral complexes. The heavy metals in the system can coordinate to the chiral additives by exploiting electrostatic interactions. Anionic additives were chosen for this specific quality, as their coordination with the positively charged metal species results in charged chiral complexes poised for interaction with the host:guest complexes of the system. After the addition of 50 µM of each additive to the array system, it was determined that sodium potassium tartrate provided the best discrimination of the pheromone targets with the other array components.

The introduction of additional array components meant that the previously optimized 6-component array required some adjustment for optimal discrimination. While the **TCC** cavitand was included in the initial screening experiments for the chiral additives, it was ultimately removed from the optimized component pool, as the previous machine learning process invalidated its utility in this case. The top six components for enantioselective discrimination were found to be broadly similar to those used for

diastereoselective discrimination, with some minor changes. The use of the reporter dye **SMITE** was not as effective, and the UO_2^{2+} heavy metal salt was switched in. Discrimination of enantiomers of the same molecule is much more nuanced, as they are chemically identical, varying only in their 3D conformation. The analysis of these subtle differences makes the use of PCA invalid, as the analysis of the data is reduced to only two dimensions, and a minimum of three dimensions are required for this method.[21] Therefore, supervised LDA was used, because it successfully analyzes the differences between only two data classes.

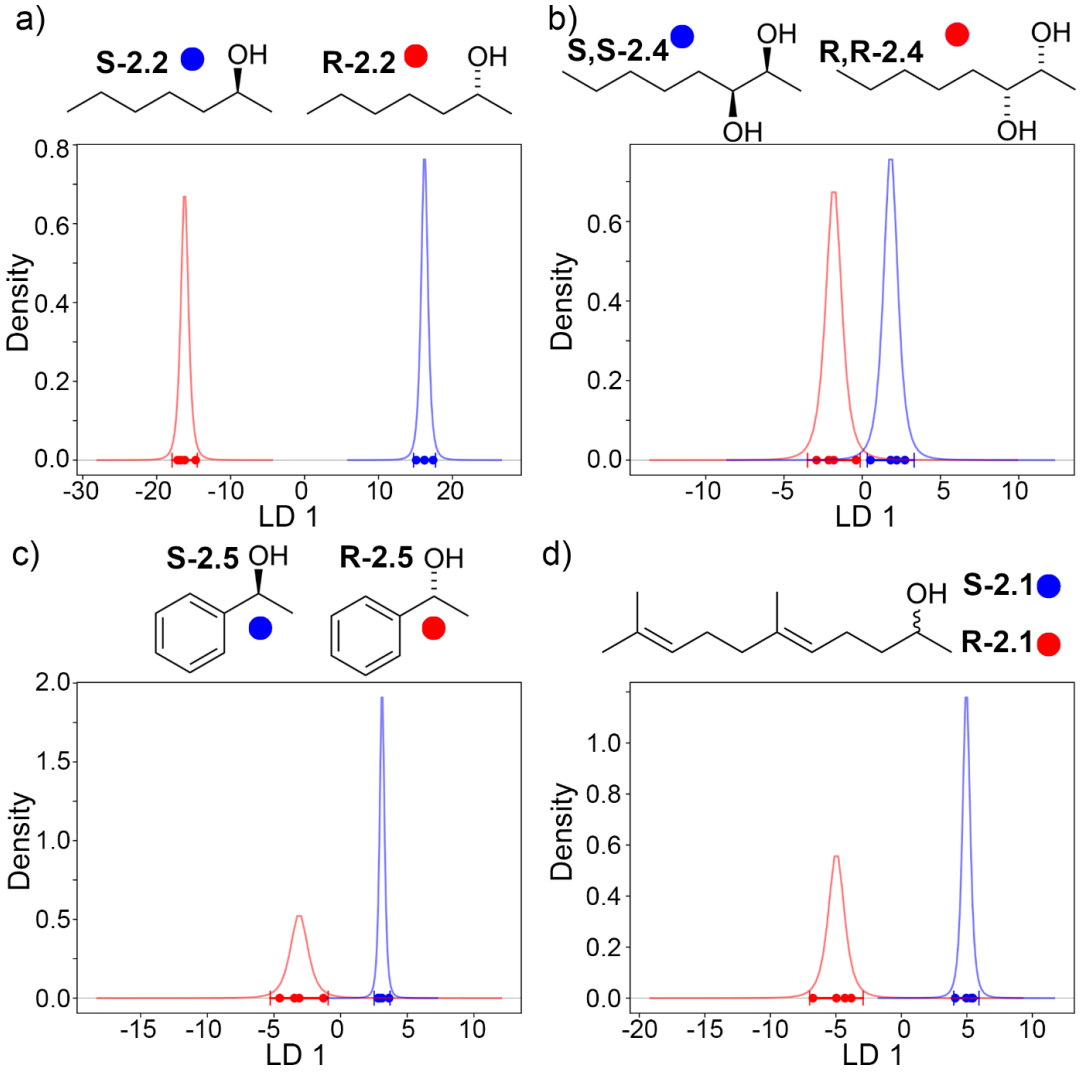


Figure 3.9. 1D LDA plots showing discrimination between pheromone enantiomers. 20 mM Tris buffer, pH 7.4 [*Host*] =20 µM, [*Dye*] = 3.0 µM, [M^+] = 50 µM, and [*Pheromone*] = 50 µM. Red/blue dots = datapoints, curve = t-distribution probability density, vertical markers = 95% confidence intervals.

With this new method of analysis, each pair of enantiomers were fully discriminated from each other with a 95% confidence interval. While all enantiomer pairs were fully discriminated from each other, the enantiomer pair of (±)-2-heptanol was best separated (Figure 3.9a). This is believed to be a result of the placement of the chiral center

of 2-heptanol at the upper rim of the cavitands when it is bound in their cavities. The observed weaker discrimination of other enantiomer pairs, such as (±)-2,3-octanediol (Figure 3.9b), is likely due to the positioning of their chiral centers further away from the upper rim of the cavitands due to poor size and shape match with the binding cavity, or a lower overall affinity with the hosts. Once again, the cavitand hosts were essential for enantioselective discrimination, as their exclusion from the array resulted in significant overlap and no discrimination. The addition of chiral additives enhances the discrimination properties of the cavitand array-based system, and their interactions are crucial for highly specific discrimination of diastereomeric guests by the macrocyclic receptors.

3.5 Conclusion

From these studies the stereoselective discrimination of pheromone guests was accomplished in aqueous environments. The alkanol targets themselves have no optical detection motif but can be implemented in an IDA with reporter dyes that can be competitively displaced. By combing the recognition abilities of water-soluble deep cavitand hosts and indicator dyes, an array-based sensing system was established. Extensive screening of multiple macrocyclic hosts, reporter dyes, and heavy metal salts were monitored for changes in fluorescence signal response upon addition of pheromone targets at μM concentrations of all components. The large data pool that was generated by the 24-component array was subjected to machine learning algorithms to highlight key features and optimize data analysis. The complete enantioselective discrimination of the multiple pairs of alkanol enantiomers was successful with the inclusion of chiral additives to the array system. The discriminatory abilities of the cavitand hosts were synergistically

enhanced with the formation of these larger chiral complexes and their interactions with the host:guest complexes.

3.6 References

1. Fiedler, D.; Leung, D. H.; Bergman, R. G.; Raymond, K. N. Selective Molecular Recognition, C-H Bond Activation, and Catalysis in Nanoscale Reaction Vessels. *Acc. Chem. Res.*, **2005**, *38*, 351-360.

2. Fiedler, D.; Leung, D. H.; Bergman, R. G.; Raymond, K. N. Enantioselective Guest Binding and Dynamic Resolution of Cationic Ruthenium Complexes by a Chiral Metal-ligand Assembly. *J. Am. Chem. Soc.*, **2004**, *126*, 3674-3675.

3. Zhu, Q. Y.; Zhou, L. P.; Cai, P. X.; Li, X. Z.; Zhou, J.; Sun, Q. F. Chiral Auxiliary and Induced Chiroptical Sensing with 5d/4f Lanthanide–organic Macrocycles. *Chem. Commun.*, **2020**, *56*, 2861-2864.

4. Rivera, J. M.; Martin, T.; Rebek, J. Jr. Chiral Softballs: Synthesis and Molecular Recognition Properties. *J. Am. Chem. Soc.*, **2001**, *123*, 5213-5220.

5. Lin, C. Y.; Lim, S.; Anslyn, E. V. Model Building Using Linear Free Energy Relationship Parameters–Eliminating Calibration Curves for Optical Analysis of Enantiomeric Excess. *J. Am. Chem. Soc.*, **2016**, *138*, 8045-8047.

6. Chapin, B. M.; Metola, P.; Vankayala, S. L.; Woodcock, H. L.; Mooibroek, T. J.; Lynch, V. M.; Larkin, J. D.; Anslyn, E. V. Disaggregation is a Mechanism for Emission Turn-On of ortho-Aminomethylphenylboronic Acid-Based Saccharide Sensors. *J. Am. Chem. Soc.*, **2017**, *139*, 5568-5578.

7. Herrera, B. T.; Pilicer, S. L.; Anslyn, E. V.; Joyce, L. A.; Wolf, C. Optical Analysis of Reaction Yield and Enantiomeric Excess: A New Paradigm Ready for Prime Time. *J. Am. Chem. Soc.*, **2018**, *140*, 10385-10401.

8. Bailey, D. M.; Hennig, A.; Uzunova, V. D.; Nau, W. M. Supramolecular Tandem Enzyme Assays for Multiparameter Sensor Arrays and Enantiomeric Excess Determination of Amino Acids. *Chem. Eur. J.*, **2008**, *14*, 6069-6077.

9. Hanks, L. M.; Millar, J. G.; Sex and Aggregation-Sex Pheromones of Cerambycid Beetles: Basic Science and Practical Applications. *J. Chem. Ecol.*, **2016**, *42*, 631-654.

10. Meier, L. R.; Zou, Y.; Mongold-Diers, J. A.; Millar, J. G.; Hanks, L. M. Pheromone composition and chemical ecology of six species of cerambycid beetles in the subfamily Lamiinae. *J. Chem. Ecol.*, **2020**, *46*, 30-39.

11. Wickham, J. D.; Millar, J. G.; Hanks, L. M.; Zou, Y.; Wong, J. C. H.; Harrison, R. D.; Chen, Y. Rapid Assessment of Cerambycid Beetle Biodiversity in a Tropical Rainforest in Yunnan Province, China, Using a Multicomponent Pheromone Lure. *Environ. Entomol.*, **2016**, *45*, 223-228.

12. Mitchell, R. F.; Hughes, D. T.; Luetje, C. W.; Millar, J. G.; Soriano-Agatón, F.; Hanks, L. M.; Robertson, H. M. Sequencing and Characterizing Odorant Receptors of the Cerambycid Beetle Megacyllene caryae. *Insect. Biochem. Mol. Biol.*, **2012**, *42*, 499-505.

13. Hughes, G. P.; Zou, Y.; Millar, J. G.; Ginzel, M. D. (S)-fuscumol and (S)-fuscumol Acetate Produced by a male Astyleiopus variegatus (Coleoptera: Cerambycidae). *Can. Entomol.*, **2013**, *145*, 327-332.

14. Liu, Y.; Mettry, M.; Gill, A. D.; Perez, L.; Zhong, W.; Hooley, R. J. Selective Heavy Element Sensing with a Simple Host:Guest Fluorescent Array. *Anal. Chem.*, **2017**, *89*, 11113-11121.

15. Dreher, S. D.; Krska, S. W. Chemistry Informer Libraries: Conception, Early Experience, and Role in the Future of Cheminformatics. *Acc. Chem. Res.*, **2021**, *54*, 1586-1596.

16. Shi, Y.; Prieto, P. L.; Zepel, T.; Grunert, S.; Hein, J. E. Automated Experimentation Powers Data Science in Chemistry. *Acc. Chem. Res.*, **2021**, *54*, 546-555.

17. Goecks, J.; Jalili, V.; Heiser, L. M.; Gray, J. W. How Machine Learning Will Transform Biomedicine. *Cell*, **2020**, *181*, 92-101.

18. Jorner, K.; Tomberg, A.; Bauer, C.; Sköld, C.; Norrby, P. O. Organic Reactivity from Mechanism to Machine Learning. *Nat. Rev. Chem.*, **2021**, *5*, 240-255.

19. Cortes, C.; Vapnik, V. Support-Vector Networks. *Machine Learning*, **1995**, *20*, 273-297.

20. Ivanciuc, O. Applications of Support Vector Machines in Chemistry. *Rev. Comput. Chem.*, **2007**, *23*, 291-400.

21. Stewart, S.; Ivy, M. A.; Anslyn, E. V. The Use of Principal Component Analysis and Discriminate Analysis in Differential Sensing Routines. *Chem. Soc. Rev.*, **2014**, *43*, 70-84.

Chapter 4: Selective Anion Recognition with Cationic Cavitands

4.1 Introduction

Molecular recognition of a wide range of neutral and positively charged targets has been explored with the use of macrocyclic receptors, while negatively charged guests are less explored. The diverse population of these receptors typically includes a defined cavity that is integral to the recognition of target molecules. The cavities of the hosts are typically electron rich and favor guests that are cationic or exhibit surface positive charges. Conversely, anions are not common guests, as they are generally less favorable binding targets. Most studies with macrocyclic hosts and anionic guests were conducted by the Gibb group,[1-3] who has reported numerous examples of anion binding in rigid hydrophobic cavitands. These results suggest that recognition is driven by favorable water expulsion from the host cavity and is further modulated by the Hofmeister effect.[4,5] The Hofmeister effect impacts the ability of ions to bind with the host as a function of hydration.[6-8] The larger an ion, the more weakly hydrated it is; therefore the desolvation energy is lower, resulting in more favorable binding. Conversely, the smaller ion is more strongly hydrated and as a result of its high polarizability it more strongly attracts the water molecules, resulting in a higher desolvation energy.

Other examples of anion binding with macrocyclic hosts exploit the use of direct hydrogen bonding groups. Urea derivatives,[9-11] metals, and the electron poor C-H bonds[12,13] found in triazine groups, incorporate into the hosts' structures. Ultimately the greatest affinity and selectively for anions has been observed with rigid macrocyclic hosts,[14-16] with cavities that provide optimal space-filling potential for the anionic guests.

More flexible receptors, such as water-soluble deep cavitands, are believed to be less effective for anion binding, as the recognition event must also be strong enough to trigger a conformational change in the host. In water especially, the desolvation penalties of the anions must first be overcome for binding to occur. Application in aqueous environments means that the receptors themselves must be water-soluble, at relevant concentrations for sensing. The poor solubility means that current studies with macrocyclic receptors are commonly performed in DMSO-water mixtures instead of pure water to ensure complete solubility of the host. Therefore, it is even less common for these studies to be performed in biorelevant media with the presence of competitive analytes that may reduce the selectivity of anion recognition.

Recent studies with water-soluble deep cavitands have shown their affinity for complex polyanions, such as complex DNA structures, in both water and other buffered environments, as discussed in Chapter 2. The use of deep cavitands as anion receptors is unusual, as they are flexible structures with electron-rich binding cavities. These qualities make the recognition mechanism for anionic guests with deep cavitands unconventional, as the lower rim is exploited for anion binding. Generally, the bowl-shaped cavity of these hosts is used for the recognition of targets, but in this case it would remain unoccupied. This binding mechanism leaves the receptor's cavity empty, and it can be filled with an indicator dye, while the lower rim of the cavitand is free to bind anionic guests. Upon binding of the anions, any change in the fluorescence signal of the indicator can be observed, resulting in optical detection of anions despite their lack of optical properties.

4.2 Selective Recognition of Anions

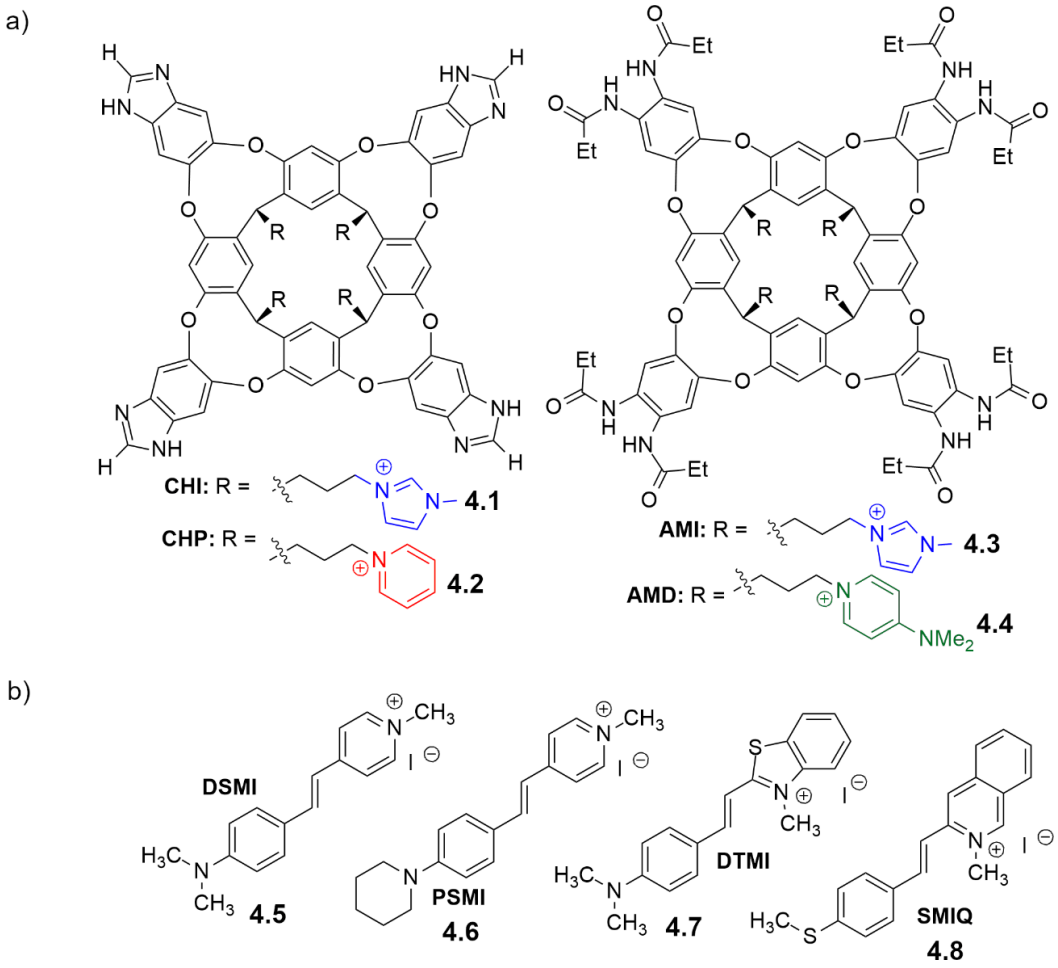


Figure 4.1. Structures of a) cationic cavitand hosts and b) styrylpyridinium indicator dyes used for initial screens of anion recognition.

To explore the anion recognition abilities of water-soluble deep cavitands, initial tests were performed with four cationic hosts with varying cationic groups at their lower rims (Figure 4.1), as well as their conformational states in aqueous environments. The diversity of cationic functional groups present at the lower rim of the cavitands are various nitrogen-containing heterocycles, such as *N*-methylimidazole (NMI), pyridine, and

dimethylamino pyridine (DMAP). These hosts were paired with three styrylpyridinium dyes, **DSMI**, **DTMI**, and **SMIQ**, which had previously showed good performance with the cationic hosts for the detection of DNA polyanions. These cationic dyes are all able to bind in the cavity of the cavitand hosts, and even induce reconfiguration of the octamide cavitands to the vase conformation.

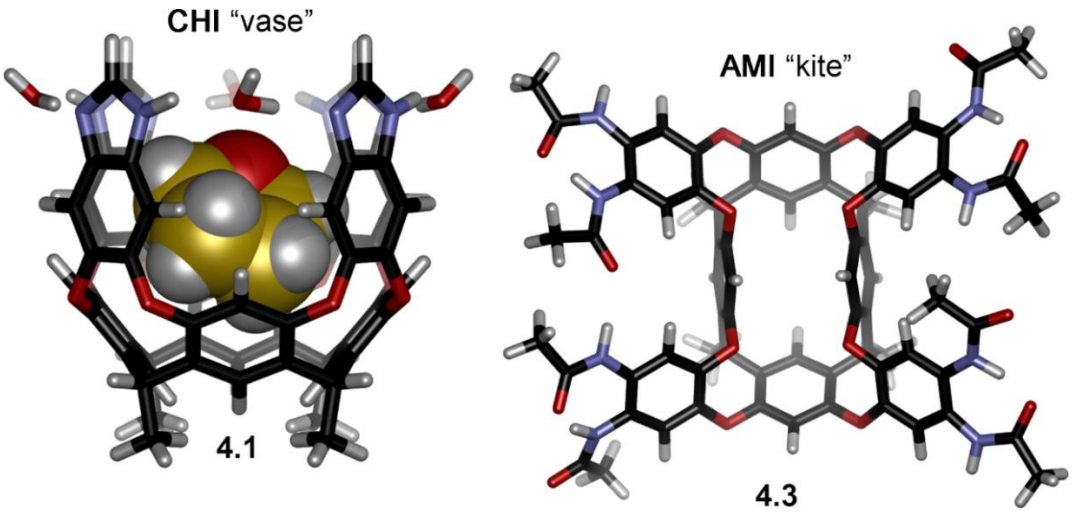


Figure 4.2. Minimized structures of the favored conformations of water-soluble deep cavitands showing the folded "vase" shape of **CHI** (with a THF molecule in the cavity, side view) and open "kite" structure of **AMI** (top view), in solution (lower rim groups truncated for clarity, SPARTAN, AM1 forcefield).

The two benzimidazole cavitands, **CHI** and **CHP**, are kinetically stable in water and have a readily accessible binding cavity (Figure 4.2). They are held in the folded vase conformation by four intercalated water molecules, which provide favorable hydrogen bonding interactions with the benzimidazole walls of the cavity. Alternatively, the octamide cavitands **AMI** and **AMD**, are dynamic in solution, and do not have an innate

binding cavity (Figure 4.2). This open kite formation of the cavitand can interconvert with the folded vase conformation in the presence of a cavity-filling guest. Upon binding with the hosts, the emission of all the dyes increases compared to the dyes that are free in solution.

The initial screening for the sensing of the anion guests was carried out in Tris buffer at pH 7.4, to avoid competition with anions that are present in other buffered solutions. The host:dye complexes, comprised of 5 μM of each host and dye, were titrated with increasing concentrations of various anions of interest from 0-200 mM, and the changes in emission were monitored. The fluorescence plots show the change in emission plotted as F/F_0, where F_0 is the initial emission value of the host:dye complex alone, and F is the observed change in emission upon addition of anionic guest. These plots were variable in the response observed for each anion, despite the small structural differences of both the hosts and dyes (Figure 4.3). By far the greatest changes occurred with the **AMI**:**DSMI** complex, which showed a significant drop in emission with the addition of iodide and monobasic phosphate, and an increase in emission with the addition of bromide.

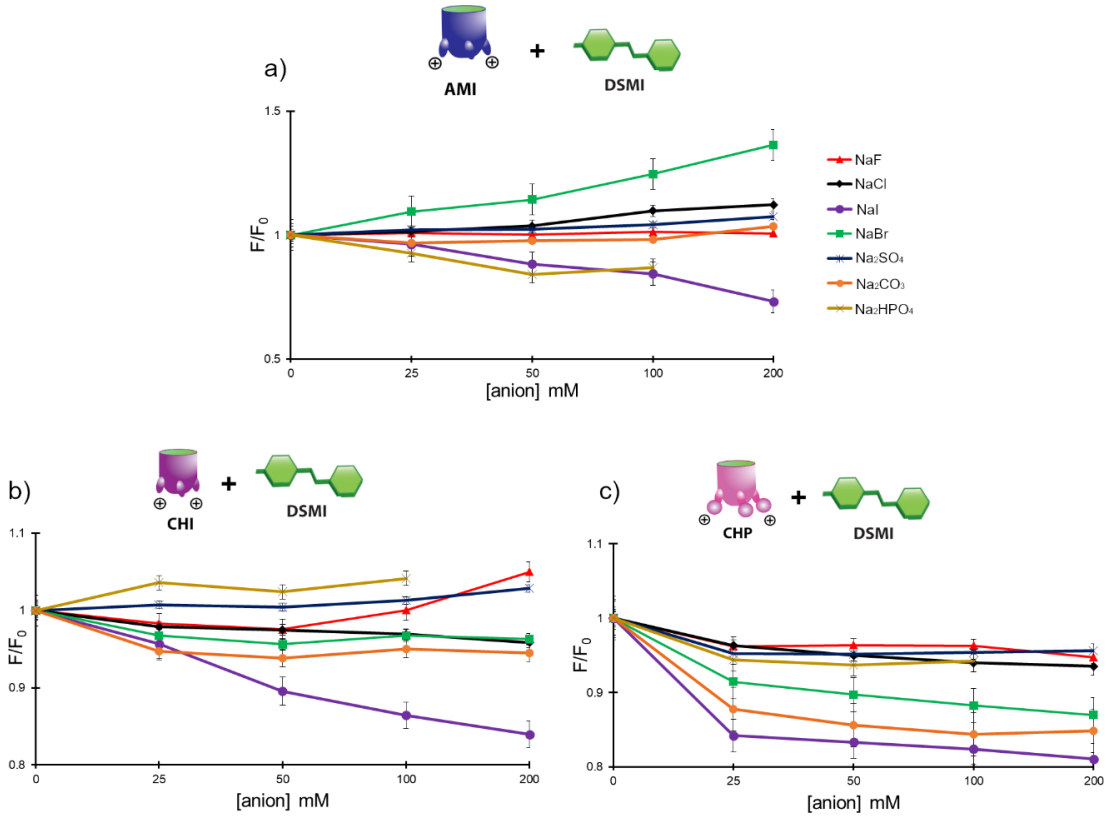


Figure 4.3. Relative fluorescence responses of the **Host•DSMI•X**$^-$ complex in 20 mM Tris buffer, pH 7.4. [**Host**] = 5 μM, [**DSMI**] = 5 μM. F_0 = fluorescence response of the **Host•DSMI** complex, F = fluorescence response of the **Host•DSMI•X**$^-$ complex, where a) is the F/F_0 response of **AMI•DSMI•X**$^-$, b) is the F/F_0 response of **CHI•DSMI•X**$^-$, and c) is the F/F_0 response of **CHP•DSMI•X**$^-$.

The results of the initial screen with **CHI** and **CHP** hosts had similar trends to the **AMI** host, but with far less intensity. The **CHI**:**DSMI** complex showed only an 18% drop in emission upon addition of iodide, even after the maximum concentration of 200 mM was reached. Notably, the **CHP**:**DSMI** complex was also seen to be mildly sensitive to the presence of carbonate and bromide, with slight decreases in emission observed. When the dye was changed to **DTMI** or **SMIQ**, the host:dye complexes resulted in less significant changes in emission, with the only notable observations for all three cavitand hosts being

with iodide. With iodide accounting for the biggest changes in emission, the same titrations were performed in the absence of the cavitand host as a control. Iodide is known to greatly quench dye emission, but no significant loss emission of the dyes was observed from these studies.

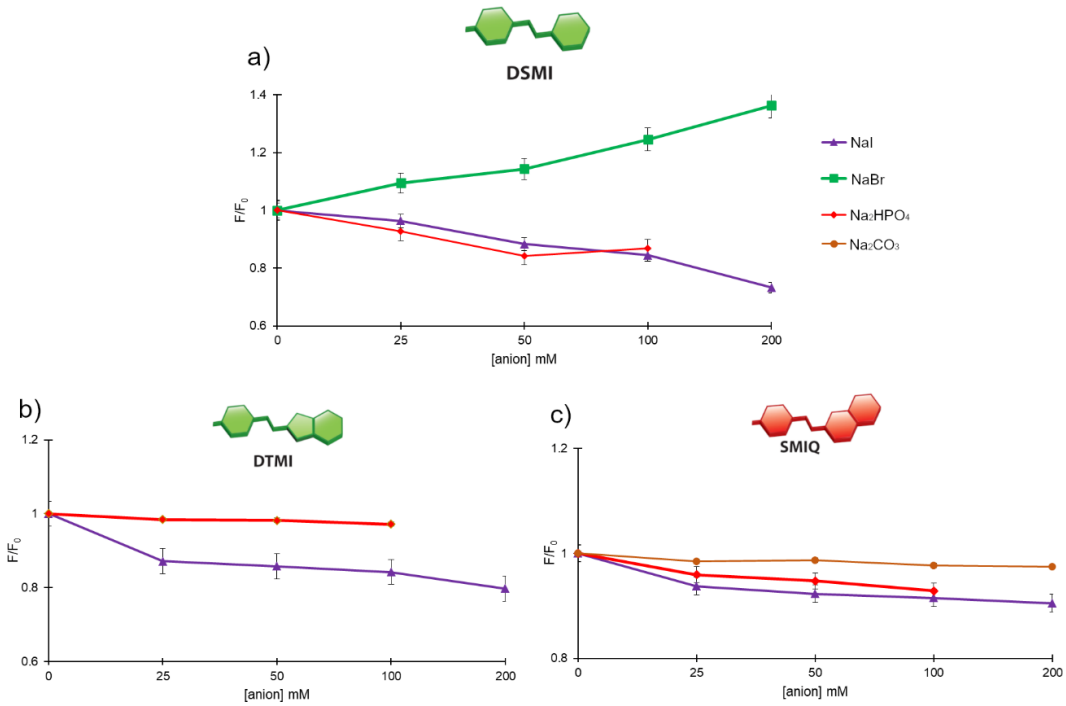


Figure 4.4. Relative fluorescence responses of the **AMI•dye•X**$^-$ complex in 20 mM Tris buffer, pH 7.4. [**AMI**] = 5 µM, [**dye**] = 5 µM, where a) is the F/F_0 response of **AMI•DSMI•X**$^-$, b) is the F/F_0 response of **AMI•DTMI•X**$^-$, and c) is the F/F_0 response of **AMI•SMIQ•X**$^-$.

These tests indicate that the flexible octamide cavitand scaffold is more responsive to the sensing of anions than the kinetically stable benzimidazole scaffold. This analysis is contradictory to the presumed outcomes, where more rigid host structures have affinity for anionic guests. To further support this theory, more tests were performed with another octamide cavitand **AMD**. Initial tests with **AMD**:**DSMI** complexes had similar results as

the **AMI**:**DSMI** complexes (Figure 4.5), with rapid and strong reduction in the fluorescence response observed upon addition of anions, specifically iodide.

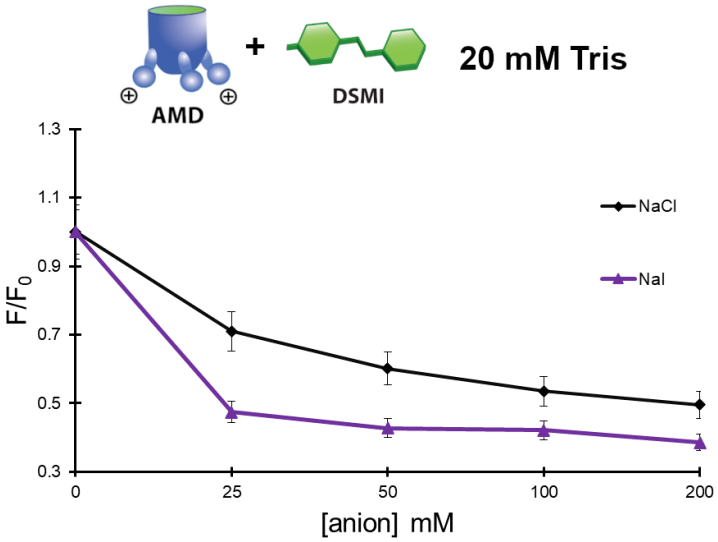


Figure 4.5. F/F_0 response responses of the **AMD•DSMI•X**$^-$ complex in 20 mM Tris buffer, pH 7.4. [**AMD**] = 5 μM, [**DSMI**] = 5 μM.

Ultimately these results were inconclusive, as further tests revealed the precipitation of the **AMD** host at greater concentrations, making it unsuitable for further anionic guest sensing. From this analysis, it was observed that the cationic hosts were selectively affected by the presence of iodide. As the **AMI** host is the most water-soluble at higher concentrations and was shown to be the most greatly affected by the addition iodide, it was used as the main host for further mechanistic investigation.

Further studies were performed *via* ^{1}H NMR analysis, by Alexie Raz, to confirm precisely where the anions were bound to the hosts. These experiments were carried out by titrating solutions of the select salts, NaCl, NaI, and NaBr, into a 1 mM solution of **AMI**

in D_2O. Upon addition of 5 mM excess of each anion, changes in the cavitand peaks were observed. While anion exchange was fast on the NMR timescale, peak shifts for the **AMI** cavitand reached saturation with NaI after the addition of 5mM. Rapid in/out exchange was also seen with NaBr, but no saturation was observed, even after addition of 40 mM, and no binding was seen with NaCl. The same peaks were shifted in the **AMI** spectra, regardless of the anion that was added to the host.

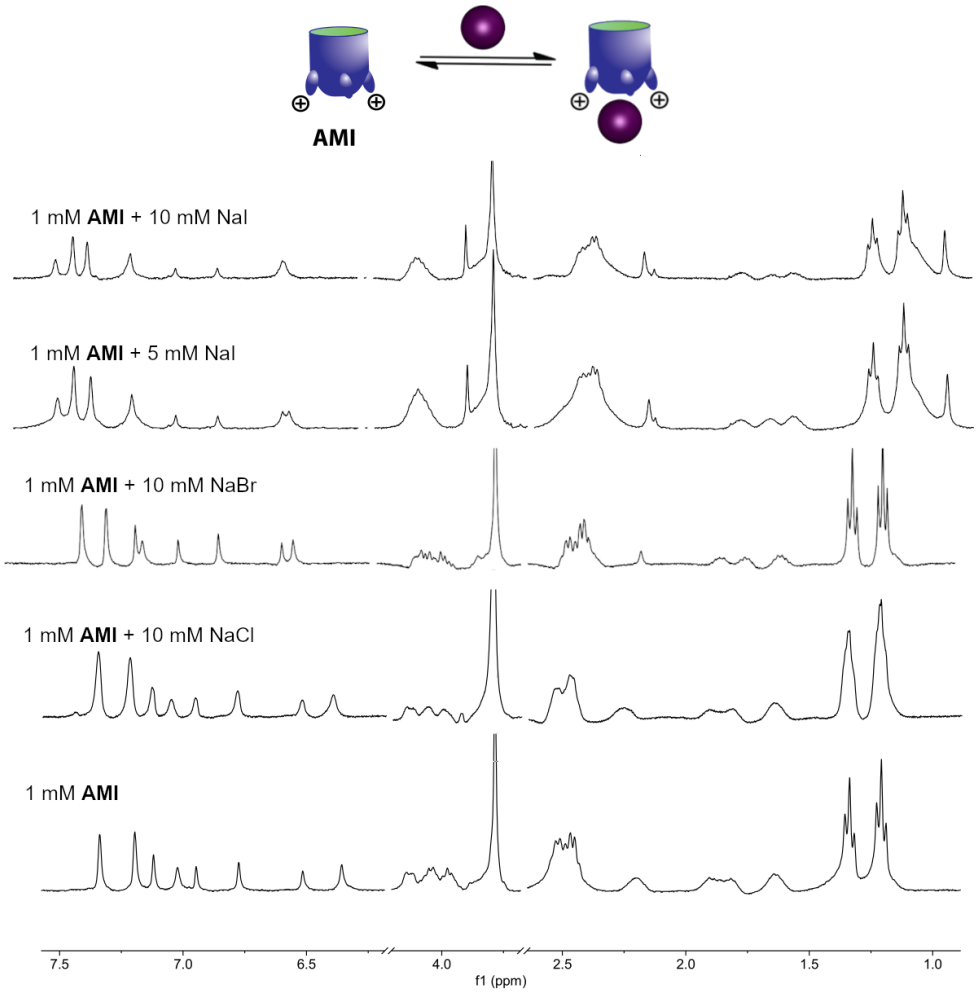


Figure 4.6. ^1H NMR spectra (600 MHz, D_2O, 298K) showing halide salts binding with **AMI** host.

The peaks with the greatest shift correspond to the lower rim aromatic protons of the cavitand, as all three imidazolium peaks are shifted downfield (Figure 4.7). The conformation of the **AMI** cavitand remains unchanged as the unfolded kite, as no new symmetry of the peaks is observed. Additional minor shifts are also observed for the methine protons, as well as the central CH_2 protons of the lower rim alkyl chain.

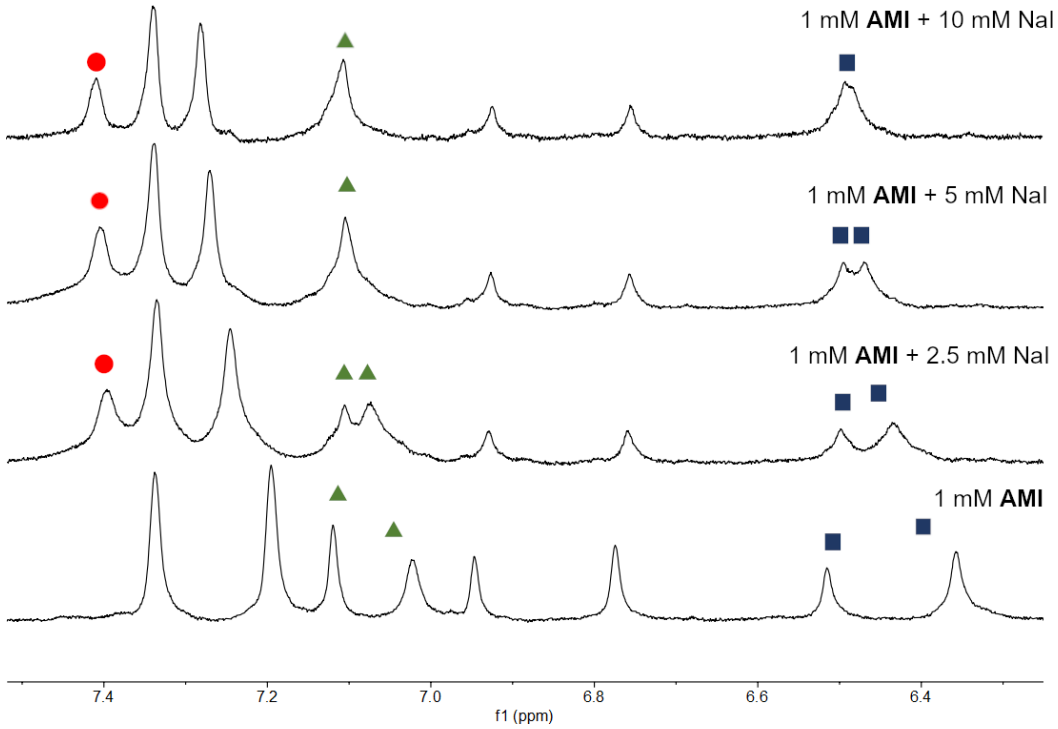


Figure 4.7. ^{1}H NMR spectra (600 MHz, D_2O, 298K) showing fast in and out exchange of NaI binding with **AMI** host with tracking of shifted peaks and emergence of new peaks shown.

To analyze why the **AMI** cavitand showed such strong selectivity for iodide, the **AMD** cavitand with its comparable binding cavity and kite conformation, was also tested *via* ^{1}H NMR titration experiments. While both **AMI** and **AMD** are octamide cavitands,

they differ in their lower rim functional groups. The **AMI** cavitand has four methyl imidazole groups present at its feet, while the **AMD** cavitand has four larger DMAP groups. Both hosts are cationic due to the four tetrasubstituted nitrogen heterocycles at the lower rim, which should allow binding to anions in a presumably similar manner.

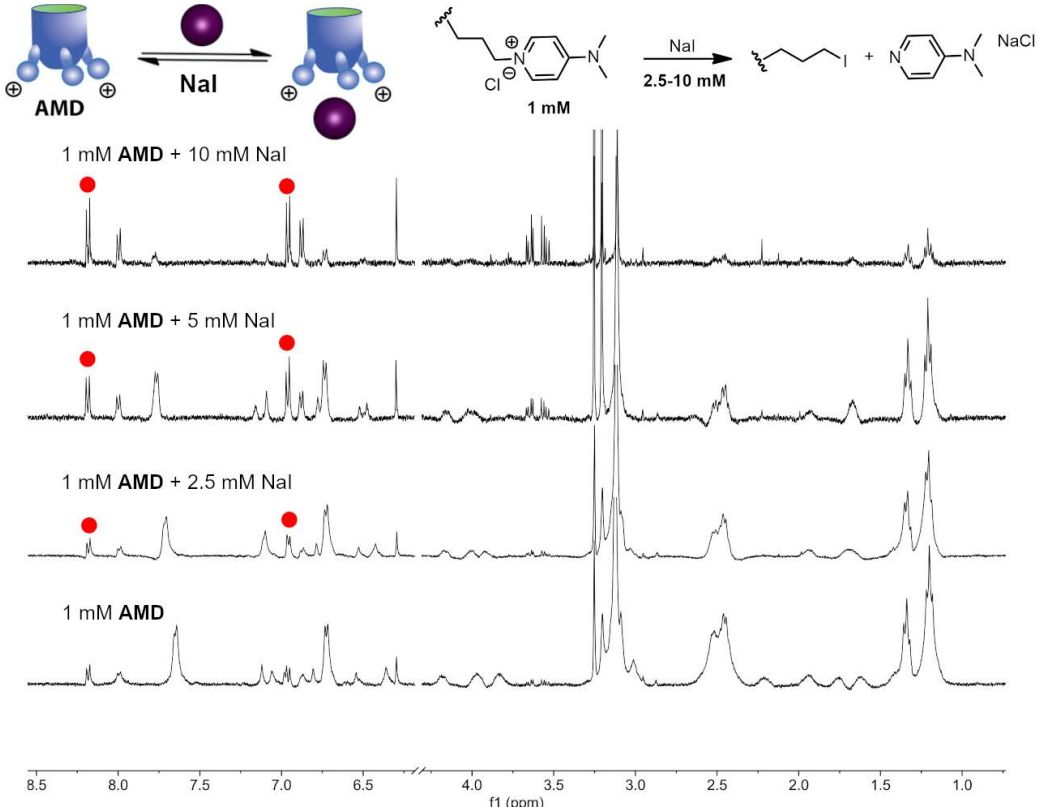


Figure 4.8. ^{1}H NMR spectra (600 MHz, D_2O, 298K) studying titrations of NaI with **AMD** showing cleavage of DMAP groups via substitution reaction with NaI, labeled peaks show emergence of free DMAP as the insoluble AM-iodide cavitand is formed.

The ^{1}H NMR titration experiments with **AMD** were carried out identically to those with **AMI**, but analysis of the **AMD** cavitand was complicated because the higher concentrations of the host resulted in the formation of a precipitate upon addition of the

anions. Precipitation was most significant with the addition of iodide, presumably due to the iodide participating in a side reaction with the **AMD** host, as the only signals present in the NMR spectra at higher iodide concentrations were of free DMAP (Figure 4.8). Solutions of the **AMD** cavitand in water are not prone to solvolysis at room temperature, but in the presence of iodide, the nucleophilic substitution of the cationic DMAP groups is observed. This side reaction suggests that **AMD** selectively binds iodide in a similar manner as **AMI**, and the increased effective concentration of iodide accelerates the substitution reaction. The replacement of the cationic DMAP groups with iodide results in the formation of the insoluble iodo-cavitand, and accounts for the free DMAP signals seen in the NMR spectra. An analogue for the **AMD** cavitand (**4.9**) was synthesized by Alexie Raz to study whether molecular recognition was necessary for the accelerated reaction rate. The titration experiments were repeated with this alkyl DMAP molecule (**4.9**), and no substitution reaction was observed (Figure 4.9). Therefore, molecular recognition with the **AMD** host is essential to enhance the reactivity of the solvated iodide nucleophile.

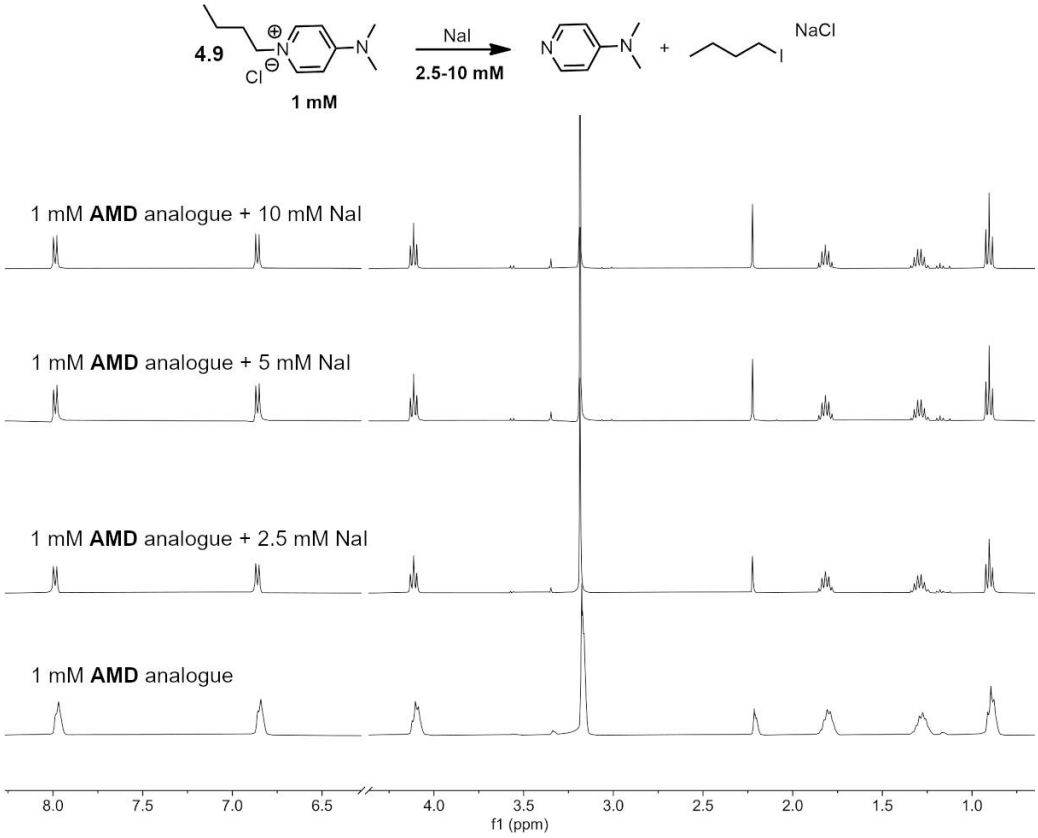


Figure 4.9. ^{1}H NMR spectra (600 MHz, D_2O, 298K) studying titrations of NaI with **AMD analogue** where no substitution reaction with NaI is observed.

To determine a measurable binding affinity (K_a) of the anions for the cavitand host, isothermal titration calorimetry (ITC) experiments were performed with the help of a collaborator, Junyi Chen. Solutions of three salts, NaCl, NaI, and NaBr, were titrated into a solution of 1 mM **AMI**, which were all prepared in ultrapure water. The data measured in relation to the heat absorbed or released upon binding was plotted, and the K_a values for each anion were calculated. The iodide unsurprisingly showed the highest binding affinity, $K_a = 4.5 \times 10^3\ M^{-1}$, whereas the binding affinity of bromide was much lower, $K_a = 381\ M^{-1}$, and chloride showed no measurable affinity for the cavitand. These results corroborated

the results previously observed from the fluorescence data. Additionally, it was determined from the ITC experiments that iodide binding was both entropically and enthalpically favorable.

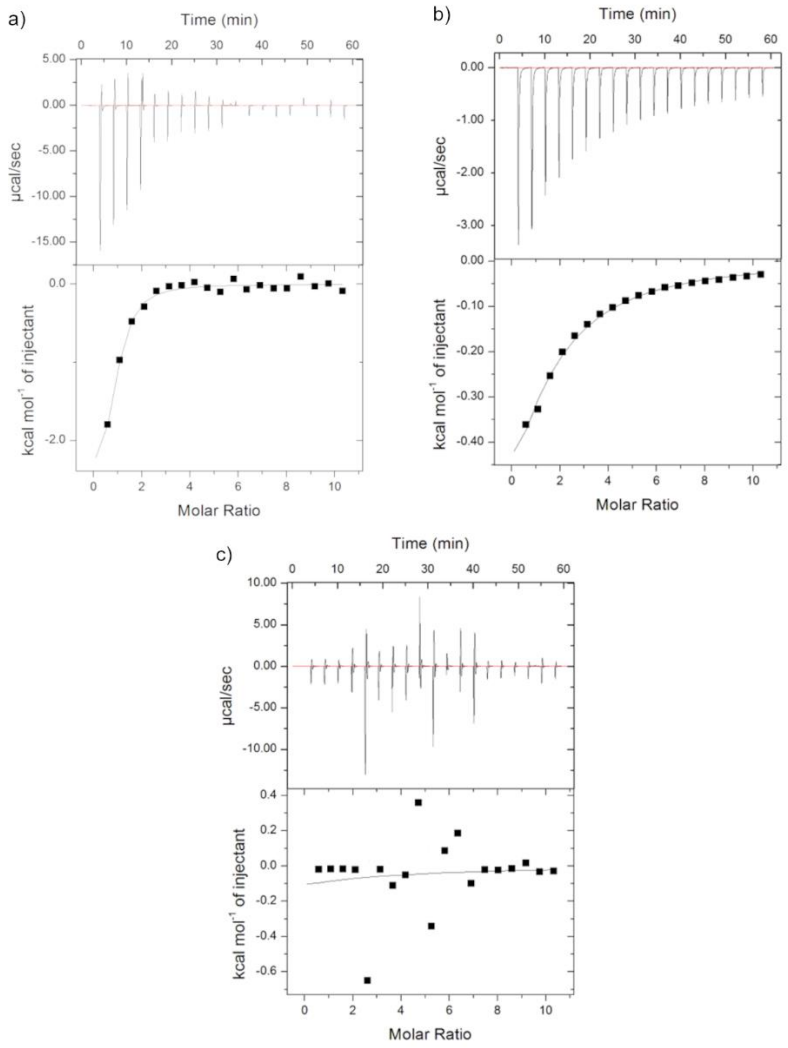


Figure 4.10. ITC titrations of increasing amounts of a) 50 mM NaI, and b) 50 mM NaBr, and c) 50 mM NaCl with 1 mM **AMI**, measured at 20 °C. Top trace: raw data for the ITC titration. Bottom trace: binding isotherm of the integrated calorimetric titration data. The heat of dilution, measured by the injection of titrant into H_2O, was subtracted for each titration to obtain the net reaction heat value.

The greater selectivity for iodide over chloride was also explored via DFT analysis by Prof. Joshua Hartman (UCR), who provided the optimal minimized structures of the **AMI**:anion complexes (Figure 4.11). These structures support the halides binding at the lower rim of the cavitand as expected, with two of the four imidazolium ions surrounding the anion. In solution, the rapid exchange between two of the imidazolium ions that coordinate to the anion would be expected, as no change in the lower rim symmetry is observed on the NMR timescale. From these structures it is apparent that the iodide anions more fully fill the "cavity" at the lower rim of the host, and therefore have closer contacts with the CH bonds of the imidazolium groups. The DFT binding affinities calculated by the r2SCAN-D3(BJ)/def2-SVP method[17,18] for iodide and chloride from these structures do not match the large difference in affinity derived from the experimental data. As the selectivity is driven by anion dehydration, the smaller and more highly solvated chloride anions have a much larger desolvation energy in water than the larger and less solvated iodide anions. While the anions do not need to be fully desolvated for binding to occur, as only the top half is isolated from water when bound, the lower desolvation penalty of iodide confers the binding selectivity by the hosts.

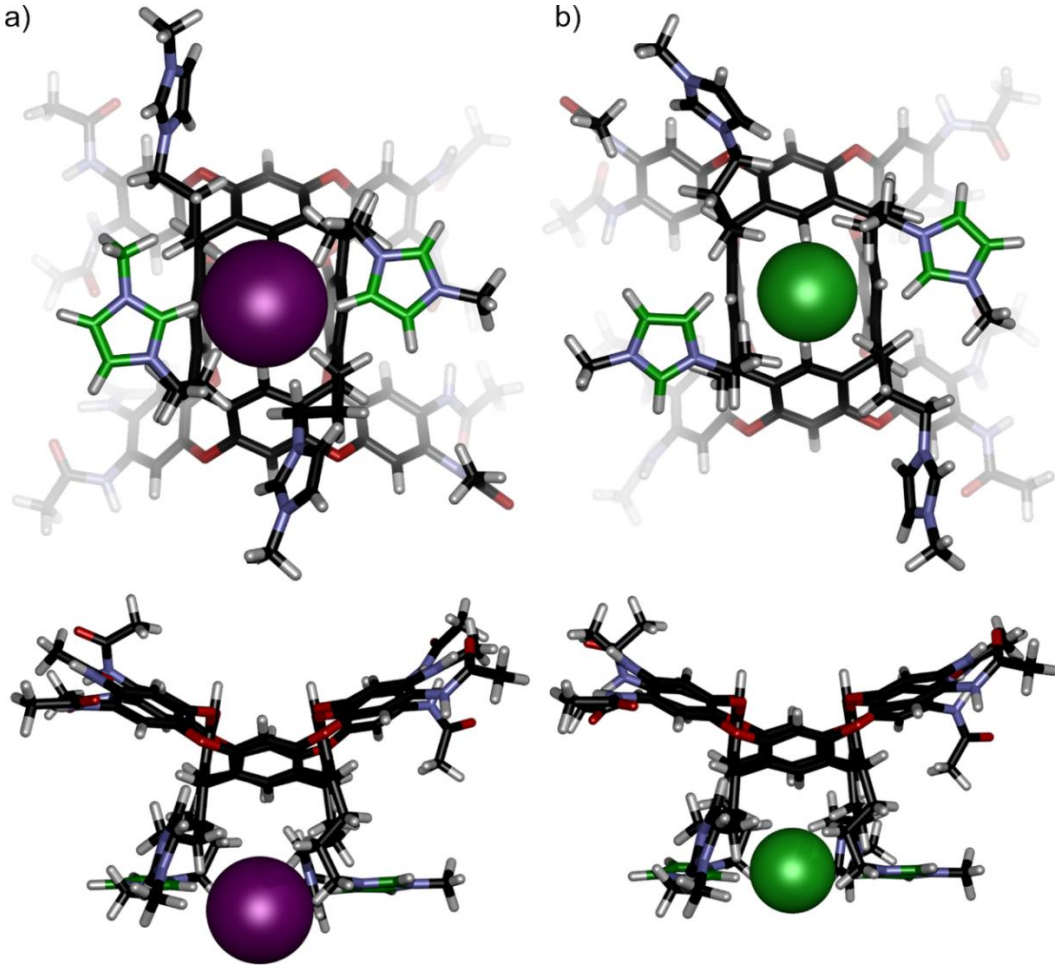


Figure 4.11. DFT-optimized structures of a) **AMI•I**$^{-}$; b) **AMI•Cl**$^{-}$ (r2SCAN-D3(BJ)/def2-SVP) viewed from the top and profile.

4.3 Selective Anion Binding in the Presence of Competitive Analytes

From the analysis of anion recognition in Tris buffer, it was indicated that IDA experiments in other aqueous media would also show anion selectivity. Therefore, the performance of these host:guest sensor combinations was tested in other aqueous environments, ultrapure water and 10X PBS buffer, both at roughly pH 7.4. The hosts used were narrowed down to just the **AMI** and **CHI** cavitands, and the indicator dyes were also

limited to **DSMI** only, as these complexes showed the most notable changes in optical response. The anion scope was narrowed to targets with the largest effect on the fluorescence changes from the initial screen; NaI, NaBr, $NaHPO_4$, and $NaCO_3$. Changing the aqueous environment from Tris buffer to ultrapure water had minimal effect on the sensing, as there are no high concentration of competitive analytes in either solution. The **AMI**:**DSMI** complexes still decreased the overall emission with addition of iodide and increased the emission with addition of bromide. The magnitude of the responses varied slightly, whereas the trends are broadly similar to those in Tris, the sensing is less effective is ultrapure water.

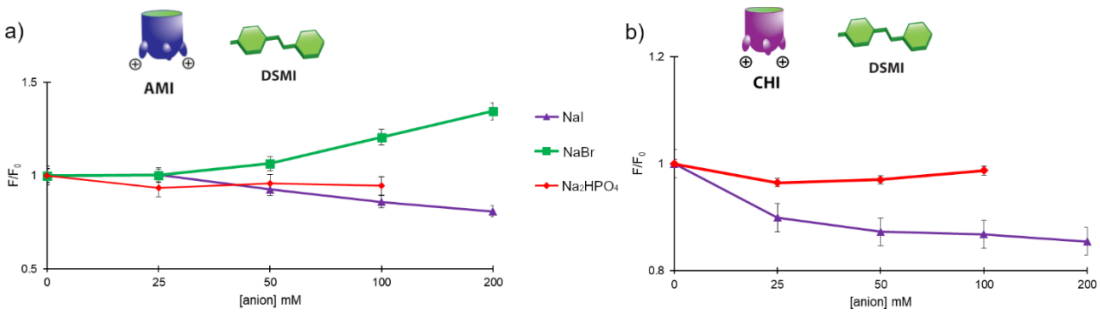


Figure 4.12. Relative fluorescence responses of the **Host•DSMI•X**$^-$ complex in ultrapure H_2O, pH 7.4. [**AMI**] = 5 µM, [**DSMI**] = 5 µM, where a) is the F/F_0 response of **AMI•DSMI•X**$^-$ and b) is the F/F_0 response of **CHI•DSMI•X**$^-$.

When the solvent was changed to 10X PBS buffer, the results were dramatically different, with a far more rapid decrease in fluorescence with the **AMI**:**DSMI** complex in the presence of iodide, and no measurable change upon addition of bromide. When the concentration of iodide was increased from 0-200 mM, a linear reduction in fluorescence with no saturation was seen in both Tris buffer and ultrapure water, while in 10X PBS

buffer a rapid drop in fluorescence was seen from titrations of 0-25 mM of iodide, followed by minimal changes in fluorescence with increasing iodide concentration. 10X PBS buffer is composed of 1.37 M NaCl, 27 mM KCl, 100 mM Na_2HPO_4, and 18 mM KH_2PO_4, and the competitive presence of phosphate and chloride ions enhances the sensing capabilities of the cavitand hosts for iodide.

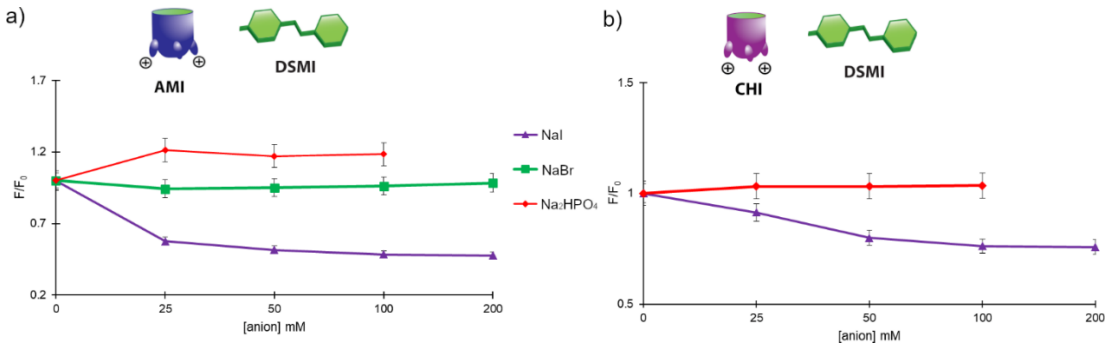


Figure 4.13. Relative fluorescence responses of the **Host•DSMI•X**$^-$ complex in 10X PBS buffer, pH 7.4. [**AMI**] = 5 μM, [**DSMI**] = 5 μM, where a) is the F/F_0 response of **AMI•DSMI•X**$^-$ and b) is the F/F_0 response of **CHI•DSMI•X**$^-$.

The conformational preferences of the **AMI** and **CHI** cavitands did influence the fluorescence response, as **CHI** has an identical lower rim as **AMI** with four imidazolium ions but showed similar changes in fluorescence in both ultrapure water and 10X PBS buffer (Figure 4.13). The limit of detection for iodide was calculated in 10X PBS buffer with the **AMI**:**DSMI** complex and was found to be 21 μM. While this LOD is relatively high, it is still impressive considering the molar concentrations of competitive ions present in the solution. Overall, these cationic hosts have shown excellent performance in selectively binding anions when paired with an indicator.

From these results many questions arose: why does selective anion binding at the base of the cavitand host cause such a drastic decrease in fluorescence, and why is the performance of the **AMI** cavitand so much better than that of the **CHI** cavitand when they possess identical recognition units? Most perplexingly, why does the addition of bromide result in an increase in emission, but only with the **AMI** host in ultrapure water or Tris buffer solutions? The most plausible theory relies on the different properties seen by each conformation of the cationic hosts. The **AMI** cavitand is an unfolded kite in the absence of guest, which deforms the resorcinarene scaffold and lower rim functional groups. With the addition of the indicator dye molecule, the cavity of the **AMI** host adopts the folded vase shape, which is C_{4v} symmetric. When the dye is bound, the vase shape of the **AMI** cavitand results in less flexibility of the lower rim groups to be able to adapt upon binding of the anionic guests.

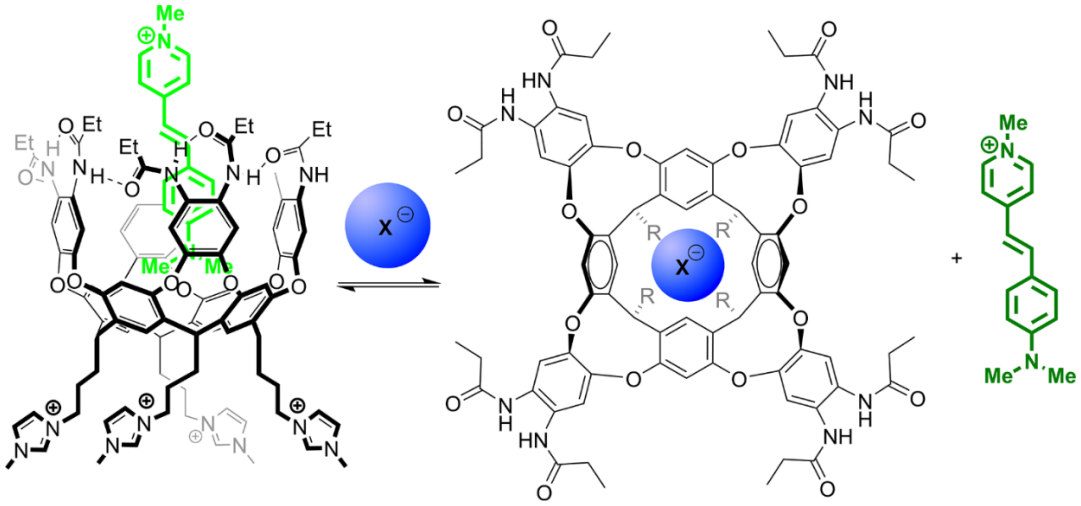


Figure 4.14. Mechanism of anion sensing with the **AMI** cavitand, where a conformational switch occurs upon anion binding, opening the host and disfavoring **DSMI** binding.

Binding of the anionic guests at the base of the cavitand should favor a reorganization of the cavitand, and disfavor dye binding in the cavity of the host, therefore causing a drop in overall emission. The selective recognition of iodide anions at the base can cause heavy atom quenching of the bound dye, and while some background quenching was seen with **DSMI**, it is not the sole cause of the magnitude of the quenching that is observed.

To further explore this relationship between the **AMI•DSMI•X**$^-$ complexes, fluorescence emission titrations with all three components were performed. The affinity of 0.5 µM **DSMI** dye to the cavitand host **AMI**, at increasing concentrations from 0-50 µM, was established in the presence of 50 mM solutions of NaBr and NaI in ultrapure water (Figure 4.15). The affinity of **DSMI** for **AMI** without any anionic guest was determined from Hill 1 Plot analysis to be K_d = 17 µM, and a noticeable drop in affinity was seen in the presence of iodide, with an affinity K_d = 60 µM. This decrease in affinity corroborates the proposed sensing mechanism, as the presence of a coordinating anion disfavors dye binding and expels the dye from the host cavity, resulting in lower emission. The analysis of the Hill 1 Plots also provides some insight for the emission increase with addition of NaBr, which is unique to the **AMI**:**DSMI** complex. The binding affinity upon bromide addition was found to be K_d = 31 µM, which is lower than the affinity of **DSMI** in water alone, but higher than the affinity of DSMI in the presence of iodide. The fluorescence response signal of the **AMI•DSMI•**Br$^-$ titration produced a graph that is sigmoidal (Figure 4.15 b), while the other complexes exhibited linear progression with a n = 1 binding modality. This change in shape indicates that multiple binding modes are present for **DSMI**

in the **AMI•DSMI•**Br^- complex, and while the Hill Plot fitting was not perfect, it does suggest a value of n > 1.

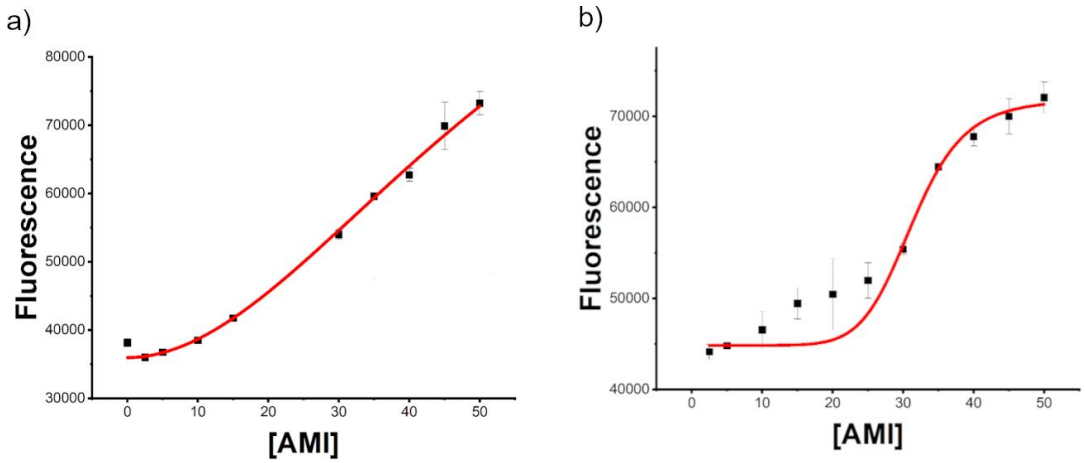


Figure 4.15. Hill 1 Plots of **DSMI** with increasing concentration of **AMI**, where a) shows binding with NaI, and b) shows binding with NaBr. The raw fluorescence of 0.5 µM **DSMI** + 0-50 µM **AMI** with 50 mM NaI or NaBr in ultrapure H_2O was collected and fitted with Hill 1 equation: *y = START + (END - START) * x^n / (k^n + x^n)* using Origin software. Error bars represent the standard deviation of 3 repeats.

One theory to explain this phenomenon is the formation of aggregates of the host and dye molecules. Water-soluble deep cavitands are precedented to undergo aggregation in high salt solutions, so the presence of **AMI** and **DSMI** aggregates in the presence of high concentrations of anions is plausible.[19] Due to the formation of these aggregates, an increase in fluorescence by **DSMI** is observed as an outlier predominately with bromide ions. The variable affinity of the halide ions for the cavity is related to the components present in the buffered solutions. With media that is composed of high concentrations of multiple salts, the affinity of the anionic guests is affected by competitive buffer complexation. As a result, iodide sensing is more effective in the presence of high salt 10X

PBS buffer than ultrapure water. The LOD of iodide in 10X PBS solution is 21 µM which is impressive considering the high concentrations of competitive ions. This is not due to high affinity of iodide for the host, but a greater change in fluorescence response from the **AMI**:**DSMI** complex.

To further investigate the effect of the affinity of the bound dye on anion selectivity, a slightly larger styrylpyridinium dye, **PSMI** was also tested, as it was previously shown to have a higher binding affinity with the **AMI** cavitand. The **PSMI** dye differs from **DSMI** structurally, as it possesses a larger piperidine functional group opposed to the diethylamino group present in **DSMI**. From ^{1}H NMR analysis of the **AMI**:dye complexes, both dyes bind in the cavity of the host. The addition of **DSMI** to the **AMI** solution shows rapid in/out exchange on the NMR timescale with no formation of a discrete Michaelis complex, apparent by the averaged peak shifts observed between the two complexes. In contrast, the addition of **PSMI** to the **AMI** solution shows slow exchange and peaks for both the free guest and host:guest complex are present.

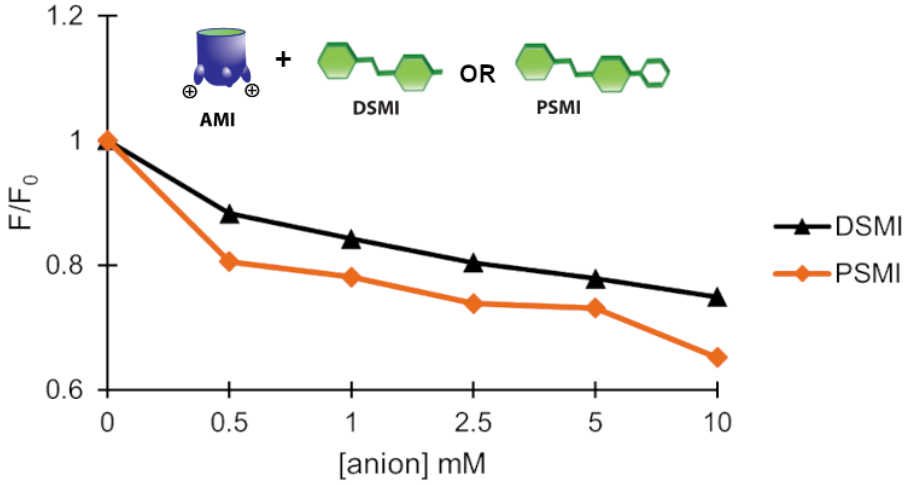


Figure 4.16. Relative fluorescence responses of the **AMI•dye•I**$^-$ complex in 10X PBS buffer, pH 7.4. [**AMI**] = 5 µM, [**dye**] = 5 µM, where a) is the F/F_0 response of **AMI•DSMI•I**$^-$ and b) is the F/F_0 response of **AMI•PSMI•I**$^-$.

Additionally, fluorescence titration experiments of iodide to the **PSMI**:**AMI** complex show a larger decrease in emission when compared to that of the **DSMI**:**AMI** complex (Figure 4.16). As the affinity of **PSMI** for the **AMI** cavitand is higher than that of **DSMI**, the conformational switch caused by the binding of iodide is more pronounced, causing a greater difference in the emission response that is generated by the displacement of the dye from the host cavity.

4.4 Conclusion

By exploiting multiple recognition motifs of these cationic cavitand hosts, the selective sensing of anions in various aqueous environments was achieved. The non-covalent recognition of the anionic analytes at the lower rim of the host, paired with the cavity-mediated binding of an indicator dye, was shown to enhance the sensing performance of the cavitand hosts without being solely dependent of the affinity on the

target analyte for the cavitand. These cationic hosts bind anionic guests in pure aqueous environments, with the strongest affinity seen with the **AMI**:**DSMI** complex for iodide with a binding affinity of 4500 M^{-1}. The recognition of these non-traditional targets occurred despite the lack of a rigid cavity or pre-coordinated electron acceptor groups in the system. All anion recognition took place at the lower rim of the cavitand hosts, where four freely rotating alkyl cationic groups are present. The combination of the lower rim recognition element, with an indicator dye bound in the host cavity allows for this molecular recognition process to be converted to an optical sensing platform. The selectivity of the host:dye system is high enough that μM concentrations of iodide can be detected in 10X PBS buffer, which is comprised of competitive salt solutions at molar concentrations. Greater quenching effects of this high iodide selectivity are observed with flexible octamide cavitand structures than with more rigid benzimidazole cavitands. The conformational change of the lower rim that is induced by anion binding displaces the bound dye from the host cavity, resulting in a greater effect on the fluorescence response. This unprecedented preference of anions for the flexible host structures is unique for water-soluble deep cavitands, as their dynamic conformations in aqueous media allows for a wider scope of target applications.

4.5 References

1. Jordan, J. H.; Gibb, C. L. D.; Wishard, A.; Pham, T.; Gibb, B. C. Ion−Hydrocarbon and/or Ion−Ion Interactions: Direct and Reverse Hofmeister Effects in a Synthetic Host. *J. Am. Chem. Soc.* **2018**, *140*, 4092-4099.

2. Aziz, H. R.; Yao, W.; Jordan, J. H.; Gibb, B. C. Dual Binding Modes of a Small Cavitand. *Supramol. Chem*., **2021**, *6*, 266-271.

3. Jordan, J. H.; Ashbaugh, H. S.; Mague, J. T.; Gibb, B. C. Buffer and Salt Effects in Aqueous Host− Guest Systems: Screening, Competitive Binding, or Both? *J. Am. Chem. Soc.*, **2021**, *143*, 18605−18616.

4. Sokkalingam, P.; Shraberg, J.; Rick, S. W.; Gibb, B. C. Binding Hydrated Anions with Hydrophobic Pockets. *J. Am. Chem. Soc.*, **2016**, *138*, 48−51.

5. Gibb, C. L.; Oertling, E. E.; Velaga, S.; Gibb, B. C. Thermodynamic Profiles of Salt Effects on a Host-Guest System: New Insight into the Hofmeister Effect. *J. Phys. Chem. B.*, **2015**, *119*, 5624−5638.

6. Flores, S. C.; Kherb, J.; Cremer, P. S. On the Mechanism of the Hofmeister Effect. *J. Phys. Chem.*, **2012**, *116*, 14408−14413.

7. Chen, X.; Flores, S. C.; Lim, S. M.; Zhang, Y.; Yang, T.; Kherb, J.; Cremer, P. S. Specific Anion Effects on Water Strucutre Adjacent to Protein Monolayers. *Langmuir*, **2010**, *26*, 16447−16454.

8. Bostrom, M.; Parsons, D. F.; Salis, A.; Ninham, B. W.; Monduzzi, M. Possible Origin of the Inverse and Direct Hofmeister Series for Lysozyme at Low and High Salt Concentration. *Langmuir*, **2011**, 27, 9504− 9511.

9. Wu, X.; Gale, P. A. Measuring Anion Transport Selectivity: A Cautionary Tale. *Chem. Comm.*, **2021**, *57*, 3979.

10. Docker, A.; Tse, Y. C.; Tay, H. M.; Taylor, A. J.; Zhang, Z.; Beer, P. D. Anti-Hofmeister Anion Selectivity via a Mechanical Bond Effect in Neutral Halogen-Bonding [2]Rotaxanes. *Angew. Chem. Int. Ed.*, **2022**, *61*, e2022145.

11. Docker, A.; Marques, I.; Kuhn, H.; Zhang, Z.; Félix, V.; Beer, P. D. Selective Potassium Chloride Recognition, Sensing, Extraction, and Transport Using a Chalcogen-Bonding Heteroditopic Receptor. *J. Am. Chem. Soc.*, **2022**, *144*, 14778-14789.

12. Parks, F. C.; Sheetz, E. G.; Stutsman, S. R.; Lutolli, A.; Debnath, S.; Raghavachari, K.; Flood, A. H. Revealing the Hidden Costs of Organization in Host–Guest Chemistry Using Chloride-Binding Foldamers and Their Solvent Dependence. *J. Am. Chem. Soc.*, **2022**, *144*, 1274-1287.

13. Liu, Y.; Zhao, W.; Chen C.-H.; Flood, A. H. Chloride Capture Using a C-H Hydrogen-Bonding Cage. *Science*, **2019**, *365*, 159-161.

14. Edwards, S. J.; Valkenier, H.; Busschaert, N.; Gale, P. A.; Davis, A. P. Hich-Affinity Anion Binding by Steroidal Squaramide Receptors. *Angew. Chem.*, **2015**, *127*, 4675-4679.

15. Butler, S.; Jolliffe, K. Anion Receptors for the Discrimination of ATP and ADP in Biological Media. *ChemPlusChem*, **2021**, *86*, 59-70.

16. Berry, S. N.; Qin, L.; Lewis, W.; Jolliffe, K. A. Conformationally Adaptable Macrocyclic Receptors for Ditopic Anions: Analysis of Chelate Cooperativity in Aqueous Containing Media. *Chem. Sci.*, **2020**, *11*, 7015-7022.

17. Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and Numerically Efficient r2SCAN Meta-Generalized Gradient Approximation. *J. Phys. Chem. Lett.*, **2020**, *11*, 8208-8215.

18. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem.*, **2009**, *113*, 6378-6396.

19. Liu, Y.; Perez, L.; Mettry, M.; Easly, C. J.; Hooley, R. J.; Zhong, W. Self-Aggregating Deep Cavitand Acts as a Fluorescence Displacement Sensor for Lysine Methylation *J. Am. Chem. Soc.*, **2016**, *138*, 10746–10749.

Chapter 5. Site-Selective Sensing of Phosphate Polyanions

5.1. Introduction

The solvation of both the host and guest is crucial to molecular recognition. This makes anions a slightly more difficult target, as their intrinsic properties are less compatible for binding in aqueous environments.[1] As such, most molecular recognition of more complex anionic guests has been carried out in organic solvents, or a combination of these solvents and water.[2] Additionally, an increase in the number of negatively charged regions of these anionic guests also makes them relatively large in size. This reduces the ability of the host to interact with these guests, due to weaker electrostatic interactions and lack of size and shape complementarity. Anions are much more highly solvated in water when compared to analogous cations due to their relative size and the negative charges they possess.[3,4] The energy of desolvation that must be overcome for binding to occur often comes at too high of a penalty, because anions are generally very soluble in water. The strongest binding of anions in water is seen with direct coordination with a metal cation which shows high affinity, but often poorly selectivity.[5,6] An increase in selectivity has been seen with the use of synthetic receptors that have been functionalized with positively charged groups to exploit electrostatic interactions with the guest molecules.[7,8] Again, this approach often results in poor selectivity, as the charged hosts interact with the competitive ions present in the environment. Hosts with an overall neutral structure are not affected by these setbacks, but also do not participate in the electrostatic interactions with the anionic guest, resulting in lower binding affinity.[5] The nature and concentration of any co-solutes present in the aqueous environment affect the binding of the anionic guests. The

Hofmeister effect is a useful scale to predict the general effects that these competitive ions can have on the system, such that certain salts act as kosmotropes, increasing the binding of hydrophobic guests, while others act as chaotropes which decreases binding.[3,9]

Recent studies with water-soluble deep cavitands have shown their affinity for a variety of anionic targets, such as a range of complex DNA structures,[10-14] as well as more simplistic anionic salts.[15] Anions are not common guests for macrocyclic receptors, as they are generally less favorable binding targets, therefore the diversity of these targets is relatively underexplored. The use of deep cavitands as anion receptors is unconventional, as the electron-rich binding cavities are uninvolved in the recognition process, and the cationic lower rim is responsible for anion binding. The ionic recognition of these analytes at the lower rim of the host is paired with the cavity-mediated binding of an indicator dye. The multiple recognition motifs of these cationic cavitand hosts were able to modulate the fluorescence signal, resulting in the selective differential sensing of anionic guests in a variety of aqueous environments.

5.2. Initial Screening of Simple Phosphate Polyanions

From the success with simple anionic salts, and to further understand the complexities of the binding mechanism of larger polyanionic targets, a group of intermediate sized guests were chosen for further analysis. This group of anionic guests was comprised of six different phosphosugars, which have multiple charged locations in their structures. Initial screening was performed with an array consisting of four cavitand hosts as well as four indicator dyes (Figure 5.1).

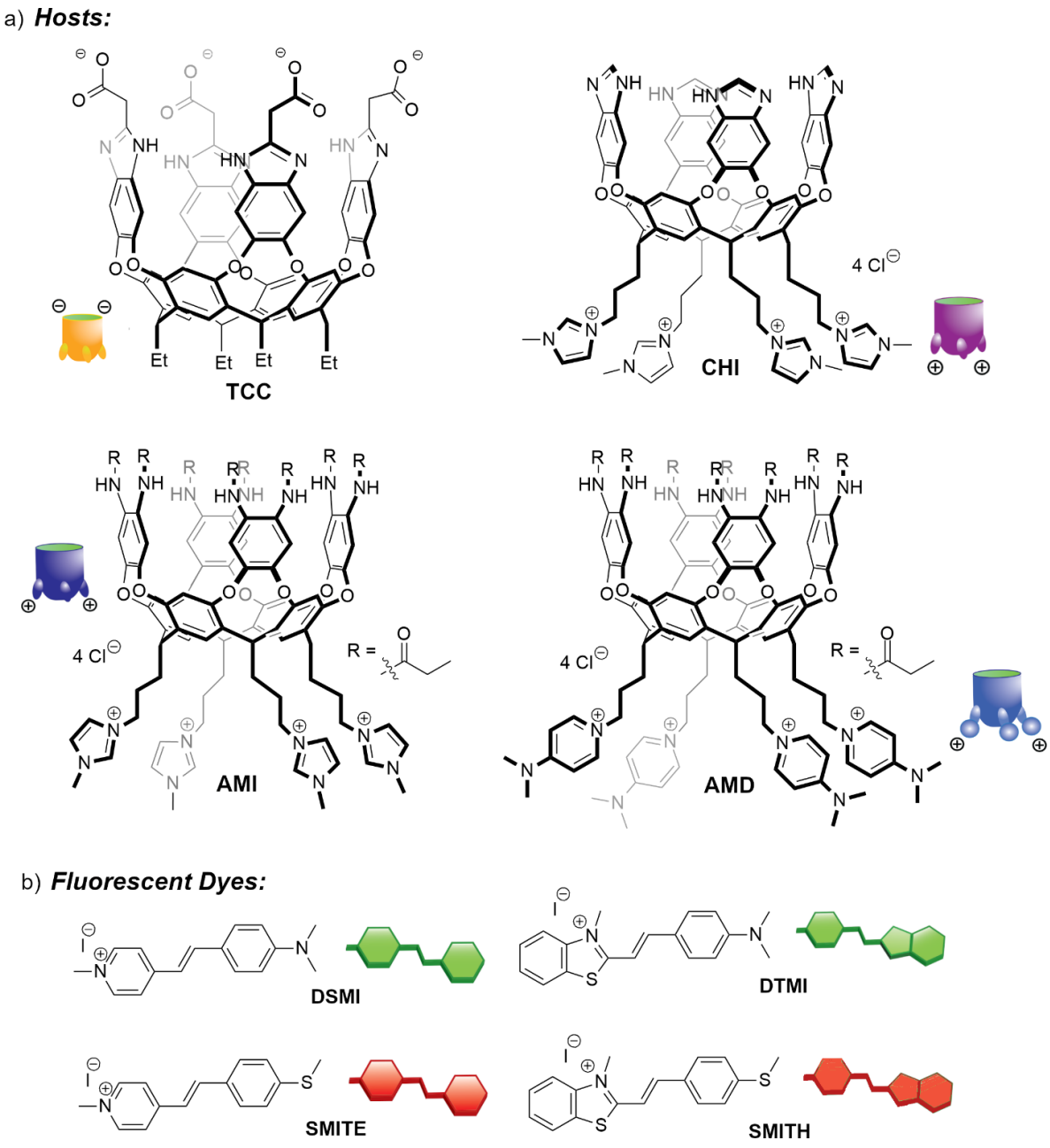


Figure 5.1. Structures of a) cavitand hosts and b) fluorescent dye molecules chosen for use in the initial array-based system.

The target molecules were limited to three phosphosugar molecules, AMP, ATP, and 6-phosphogluconic acid and the sensing was carried out in 20 mM Tris buffer at pH

7.4 (Figure 5.2). These structures vary in not only their size, but their overall charge as well, ranging from -2 – 4 at the neutral pH they were tested at, resulting in more obvious structural differences for their discrimination.

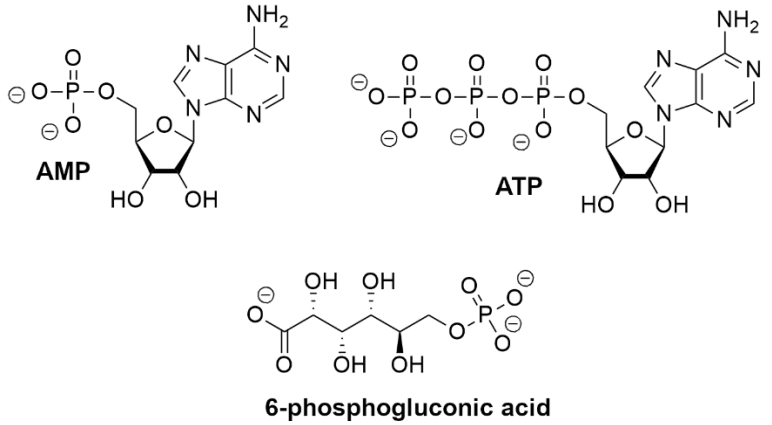


Figure 5.2. Structures of the initial phosphate anionic guest, which vary in their number of charged locations.

The analysis of the data from this 16-component array showed a wide range of differences in the emission signal, with some of the greatest differences observed for ATP. Interestingly, a different dye and cavitand complex was shown to produce this large change in fluorescence with the ATP guest. Notably, the **CHI:SMITH** and **AMI:DTMI** complexes both caused around a 1.5-fold increase in fluorescence upon the addition of 50 μM of ATP.

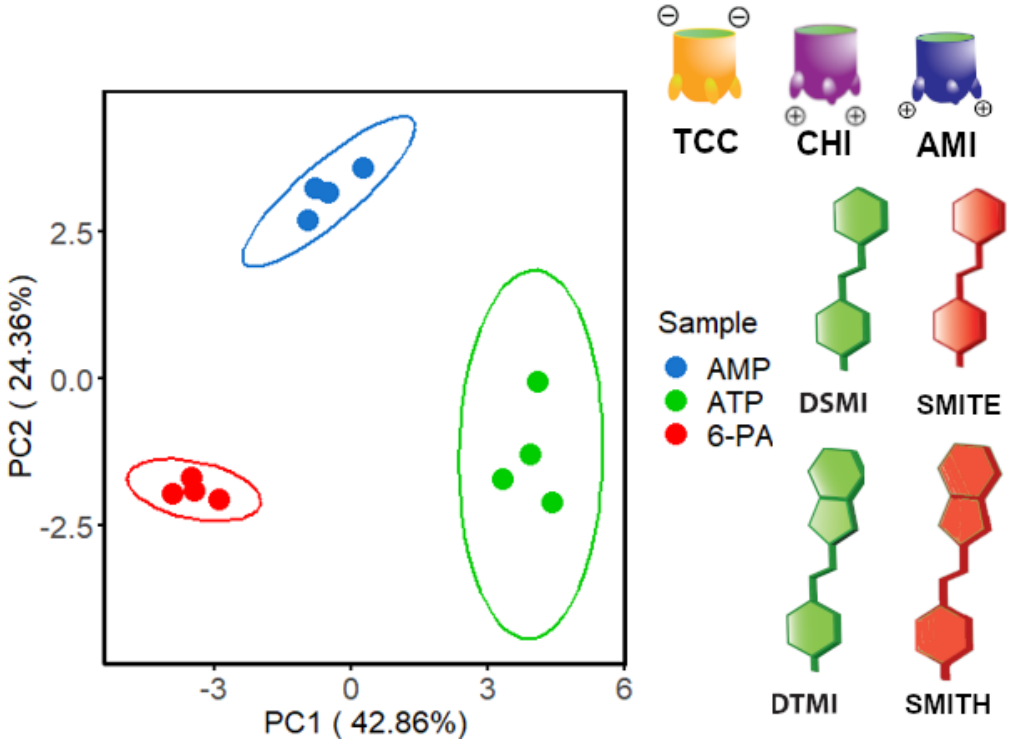


Figure 5.3. PCA scores plot derived from the data of the 8-component optimized array, ellipses indicate 95% confidence intervals. 20 mM Tris buffer, ph 7.4, [*Host*] = 5 µM, [*Dye*] = 5 µM, and [*Phosphosugar*] = 50 µM.

When the data of the entire 16-component array was subjected to statistical analysis, the PCA scores plot showed overlap between all three guests, but when the array was optimized by hand, an 8-component array was successful for their discrimination from each other (Figure 5.3). This optimized array removed the **AMD** cavitand and featured the host:dye complexes that expressed the greatest changes in emission across all three guests. From these results, more complex phosphosugar structures were selected to test the differential sensing of this array system.

5.3. Expansion of Target Scope to Complex Sugars

These more complex phosphosugars consisted of three highly similar glucose and fructose phosphate derivatives, glucose-1-phosphate (G-1-P), glucose-6-phosphate (G-6-P), and fructose-1-phosphate (F-6-P). These molecules vary only in the size of their sugar ring, 5 vs 6 membered, and the position of the phosphate group on the ring, 1 vs 6, while each structure has an overall -2 charge. (Figure 5.4).

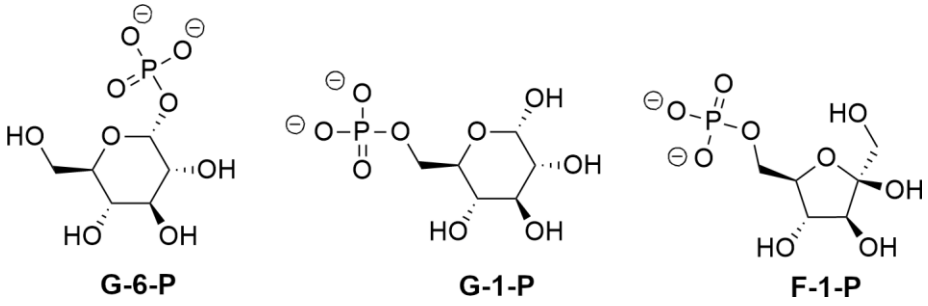


Figure 5.4. Structures of more complex phosphosugar guests with differences in ring size and charge location.

The previously optimized array components were used for analysis of these new phosphosugar guests, again in 20 mM Tris buffer at pH 7.4. Interestingly, the change in fluorescence seen upon addition of 50 μM of guest for all the host:dye complexes tested was decreased for all. Despite these obvious changes in the emission, statistical analysis of the fluorescence data produced a PCA scores plot with no discernible discrimination between these new targets (Figure 5.5).

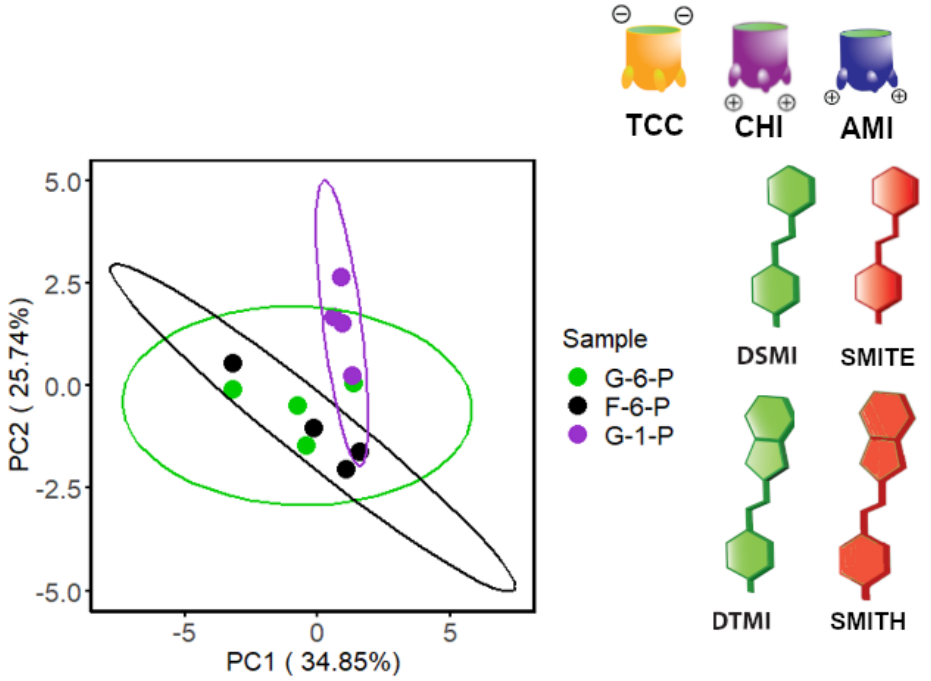


Figure 5.5. PCA scores plot derived from the data of the 8-component optimized array, ellipses indicate 95% confidence intervals. 20 mM Tris buffer, ph 7.4, [*Host*] = 5 µM, [*Dye*] = 5 µM, and [*Phosphosugar*] = 50 µM.

While disappointing, this lack of discrimination was not unexpected, as the similarities of these three guests are very high. Regardless, further optimization by hand was done, as some of the more dramatic changes in fluorescence signal led me to believe that discrimination of these guests could be achieved based on the size of the ring or the location of the phosphate group only.

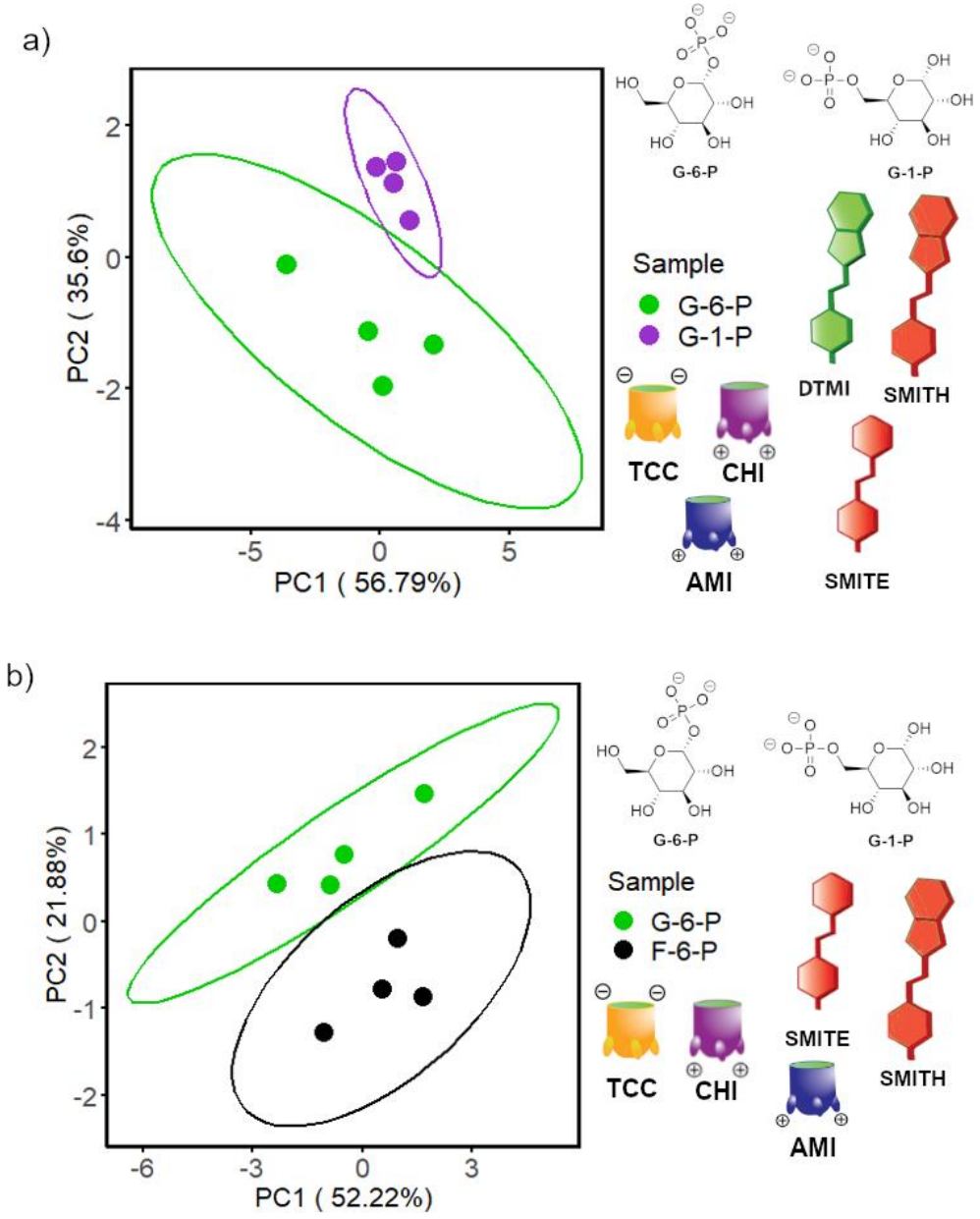


Figure 5.6. PCA scores plot derived from the data of the a) 5-component optimized array for the discrimination of G-1-P and G-6-P, and the b) 4-componenet optimized array for the discrimination of G-6-P and F-6-P, ellipses indicate 95% confidence intervals. 20 mM Tris buffer, pH 7.4, [*Host*] = 5 µM, [*Dye*] = 5 µM, and [*Phosphosugar*] = 50 µM.

From this an optimized 4-component array was able to discriminate between G-6-P and F-6-P, using combinations of the three cavitand host and two dyes, **SMITE** and **SMITH** (Figure 5.6b). Similarly, the discrimination of G-1-P and G-6-P was successfully seen with an optimized 5-component array consisting of the three cavitand hosts and three dyes, **SMITE**, **DTMI**, and **SMITH** (Figure 5.6a). Control studies were performed with the absence of the cavitand hosts, and it was observed that the molecular recognition element with the host is essential to the discrimination of these guests.

5.4. Future Work

While these preliminary results are promising, further testing must be done to understand the recognition mechanism used by the phosphosugar guests. It could be assumed that the mechanism is similar to that of the smaller anionic salts, which was discussed in chapter 4, but more analysis is needed. NMR analysis of the host:guest complexes can be used to determine the location of binding for the phosphosugar guest. If the guest is interacting at the binding pocket of the cavitand, then the lower rim interactions can be exploited by the addition of co-solutes to the system. If the phosphosugars are indeed interacting at the lower rims of the cavitands, at least those with a positive charge, then the recognition mechanism can be further optimized based on which dye is used as an indicator, as well as modulated by additional co-solutes. The array can be subjected to testing with other buffer conditions, such as PBS, or a change in the pH of the system can be tested to analyze the importance of the charge equilibria for selectivity. The addition of co-solutes of other ionic guests can also be observed, as we know the positively charged hosts show a bias for iodide.

5.5. References

1. Morita, T.; Westh, P.; Nishikawa, K.; Koga, Y. How Much Weaker are the Effects of Cations than those of Anions? The effects of K^+ and Cs^+ on the Molecular Organization of Liquid H_2O. *J. Phys. Chem. B.*, **2014**, *118*, 8744 – 8749.

2. Parks, F. C.; Sheetz, E. G.; Stutsman, S. R.; Lutolli, A.; Debnath, S.; Raghavachari, K.; Flood, A. H. Revealing the Hidden Costs of Organization in Host–Guest Chemistry Using Chloride-Binding Foldamers and Their Solvent Dependence. *J. Am. Chem. Soc.*, **2022**, *144*, 1274-1287.

3. Carnegie, R. S.; Gibb, C. L.; Gibb, B. C. Anion Complexation and the Hofmeister Effect. *Angew. Chem. Int. Ed.*, **2014**, *53*, 11498 – 11500.

4. Butler, S.; Jolliffe, K. Anion Receptors for the Discrimination of ATP and ADP in Biological Media. *ChemPlusChem*, **2021**, *86*, 59-70.

5. Langton, M. J.; Serpell, C. J.; Beer, P. D. Anion Recognition in Water: Recent Advances from a Supramolecular and Macromolecular Perspective. *Angew. Chem. Int. Ed.*, **2015**, *54*, 2–16.

6. Liu, Y.; Zhao, W.; Chen C.-H.; Flood, A. H. Chloride Capture Using a C-H Hydrogen-Bonding Cage. *Science*, **2019**, *365*, 159-161.

7. Berry, S. N.; Qin, L.; Lewis, W.; Jolliffe, K. A. Conformationally Adaptable Macrocyclic Receptors for Ditopic Anions: Analysis of Chelate Cooperativity in Aqueous Containing Media. *Chem. Sci.*, **2020**, *11*, 7015-7022.

8. Sokkalingam, P.; Shraberg, J.; Rick, S. W.; Gibb, B. C. Binding Hydrated Anions with Hydrophobic Pockets. *J. Am. Chem. Soc.*, **2016**, *138*, 48− 51.

9. Gibb, C. L.; Oertling, E. E.; Velaga, S.; Gibb, B. C. *Thermodynamic Profiles of Salt Effects on a Host-Guest System: New Insight into the Hofmeister Effect. J. Phys. Chem.*, **2015**, *119*, 5624−5638.

10. Chen, J.; Hooley, R. J.; Zhong W. Combining Excellent Selectivity with Broad Target Scope: Biosensing with Arrayed Deep Cavitand Hosts. *Acc. Chem. Res.*, **2022**, *55*, 1035-1046.

11. Chen, J.; Hickey, B. L.; Wang, L.; Lee, J.; Gill, A. D.; Favero, A.; Pinalli, R.; Dalcanale, E.; Hooley, R. J.; Zhong, W. Selective Discrimination and Classification of G-Quadruplex Structures with a Host:Guest Sensing Array. *Nat. Chem.*, **2021,** *13*, 488-495.

12. Chen, J.; Gill, A. D.; Hickey, B. L.; Gao, Z.; Cui, X.; E.; Hooley, R. J.; Zhong, W. Machine Learning Aids Classification and Discrimination of Noncanonical DNA Folding Motifs by an Arrayed Host:Guest Sensing System. *J. Am. Chem. Soc.*, **2021**, *32*, 12791-12799.

13. Chen, J.; Hickey, B. L.; Raz, A. A. P.; Gao, Z.; Hooley, R. J.; Zhong, W. Sensing Base Modifications in Non-Canonically Folded DNA with an Optimized Host:Guest Sensing Array. *ACS Sens.*, **2022**, *7*, 2164–2169.

14. Hickey, B. L.; Raz, A. A. P.; Chen, J.; Moreno J. L. Jr.; Hartman, J. D.; Zhong, W.; Hooley, R. J. Selective Anion Sensing in High Salt Water *via* a Remote Indicator Displacement Assay. *Chem. Commun.*, **2023**, *59*, 7819-7822.

Chapter 6. Experimental

6.1. General Information

^{1}H and ^{13}C NMR spectra were recorded on Bruker Avance Neo 400 MHz and 600 MHz NMR spectrometers. The spectrometers were automatically tuned and matched to the correct operation frequencies. Proton (^{1}H) and carbon (^{13}C) chemical shifts are reported in parts per million (δ) with respect to tetramethylsilane (TMS, δ=0), and referenced internally. Deuterated NMR solvents were obtained from Cambridge Isotope Laboratories (Andover, MA), and used without further purification. All NMR spectra were digitally processed (phase and baseline corrections, integration, peak analysis) using Bruker Topspin 1.3 and MestreNova by Mestrelab. All other materials were purchased from Sigma Aldrich (St. Louis, MO) or Fisher Scientific (Fairlawn, NJ), and were used as received. Solvents were dried through a commercial solvent purification system (Pure Process Technologies, Inc.). High resolution accurate mass spectral data were obtained on an Agilent 6545 QTOF LC/MS instrument. UV-Vis spectroscopy was performed on a Cary 60 Photospectrometer using the Varian Scans program to collect data. Fluorescence measurements were performed with a Bio-Tek Synergy HT Multi-Detection Microplate Reader, with the Ex/Em wavelengths at 485/528 nm or 485/600 nm. All isothermal titrations calorimetry (ITC) experiments were performed using a MicroCal iTC200 (GE Healthcare, Freiburg, Germany) with a stirring rate of 700 rpm. Research S.L. Molecular modelling (Hartree-Fock) was performed using SPARTAN ‘06.

6.2. Experimental for Chapter 2

See the following publications for full synthesis/characterization of previously published cavitand structures:

- Biros, S. M.; Ullrich E. C.; Hof F.; Trembleau, L.; Rebek, J. Jr. Kinetically Stable Complexes in Water: The Role of Hydration and Hydrophobicity. *J. Am. Chem. Soc.*, **2004**, *126*, 2870–2876.
- Mosca, S.; Yang, Y.; Rebek, J. Jr. Preparative Scale and Convenient Synthesis of a Water-Soluble, Deep Cavitand. *Nat. Proc.*, **2016**, *11*, 1371-1387.

General Procedure for NMR titrations:

Separate 1 mM solutions of **TCC** and **AMI** hosts were prepared in 300 µL D_2O and placed in an NMR tube. A ^1H NMR spectrum was taken. To the NMR tubes, selected dye molecules, **DSMI** and **PSMI**, (0.5 – 5.0 mM in D_2O) were titrated into the host solutions. A ^1H NMR spectrum was taken each time.

General procedure for fluorescence titrations and affinity measurements:

The fluorescence titrations were carried out by mixing 10 µL of the fluorescent dye (0.625 µM in water), and 10 µL of cavitand host (0 - 50 µM in water), then adding 80 µL of the incubation buffer (10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4) to bring the total volume up to 100 µL for each well in the 96-well plate, then incubating with mild agitation for 15 min at room temperature. The fluorescence signal (F) was recorded with the Ex/Em wavelengths at 485/600 or 485/528 nm for all fluorescent dyes. Fluorescence response (F/F_0) are normalized to the response of the dye alone in the absence of host, that is F_0 is defined as the fluorescence recorded for that concentration when [host] = 0 µM. Each

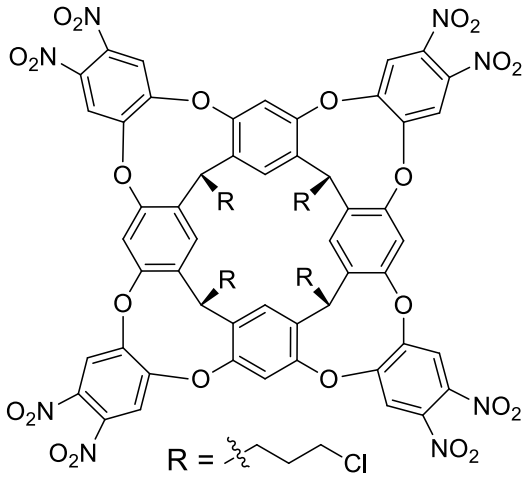
experimental condition was repeated in quadruplicate across four separate wells of the 96-well plate using identical sensor components, simultaneously collecting fluorescence signals for each target at once. The affinity of the dye for the host was determined by the Hill 1 function.

Hill 1 function is a more general form of Hill function. The equation is:

$$y = START + (END - START)\frac{x^n}{k^n + x^n}$$

x is the ligand concentration, *k* is half-maximal concentration constant, *n* is Hill coefficient. Job plots were obtained by measuring the fluorescence intensity of solutions containing dye and guest with the dye mole fraction (X_{Dye}) varied between 0 and 1.

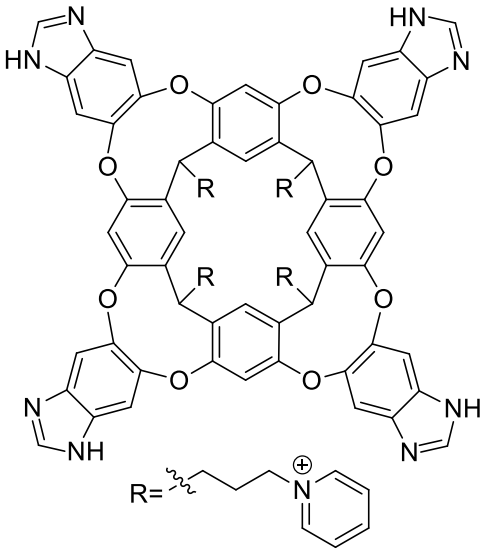
Synthesis of Chloro-amide cavitand (2.11):



The chloro-nitro cavitand was synthesized via an adaptation of Rebek's literature procedure.[2] Chloro-nitro cavitand (100 mg, 0.16 mmol) was placed in a round-bottomed flask with excess tin (II) chloride dihydrate (450 mg) and a stir bar. A 4:1 mixture of ethanol and concentrated HCl (4:1 mL) was added to the flask and the reaction was stirred at 75 °C overnight. The reaction mixture was cooled, and the solvent was removed by rotary

evaporation. The resulting residue was transferred to a flask and suspended with ethyl acetate (10 mL). A solution of potassium carbonate (2 g in 10 mL water) was added slowly to the until the mixture was shown to be basic by litmus. The mixture was stirred vigorously and propionyl chloride (3 x 0.2 mL) was added waiting 10 minutes between additions. The organic layer was collected, dried over anhydrous sodium sulfate, filtered through cotton, and the solvent evaporated. This crude product was used in the next step without purification.

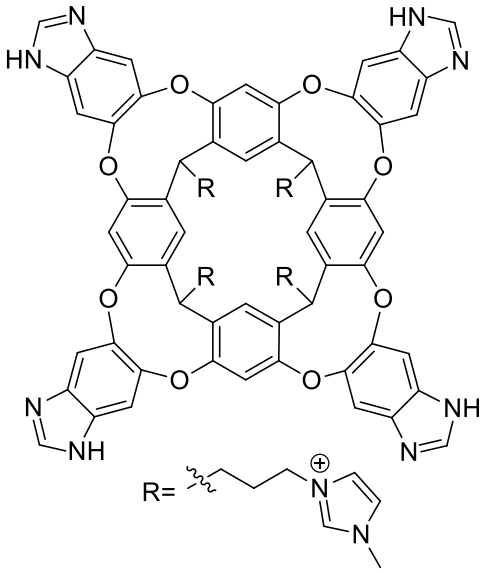
Synthesis of pyrimidine-benzimidazole cavitand (CHP):



Chloro-benzimidazole cavitand (75 mg, 0.059 mmol) was placed in a round-bottomed flask with excess pyridine (2 mL) and a stir bar and the reaction was stirred at 90 °C for 16 h. The reaction was cooled to room temperature, and cold acetone (2 mL) was added to form a pale-yellow precipitate which was filtered and collected. The solid was then refluxed in acetone (3 mL) for 16 h. The reaction was again cooled to room temperature and the solid was filtered and dried resulting in **CHP** cavitand (72 mg, 76% yield) as a pale-yellow solid.

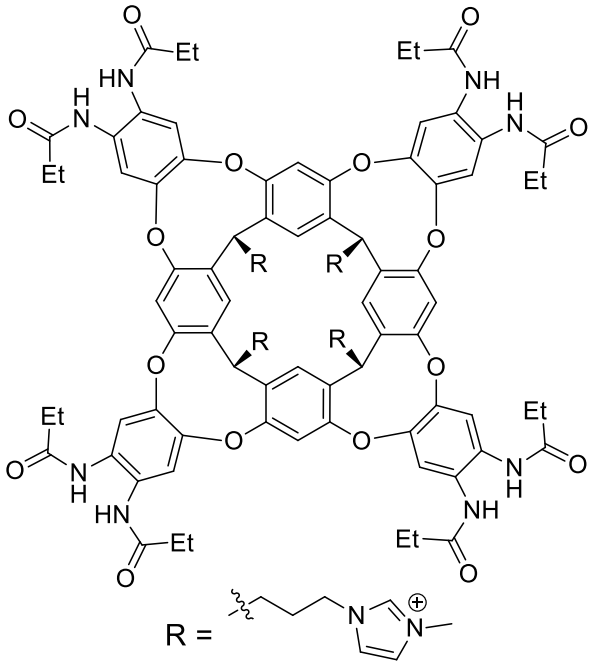
^{1}H NMR (400 MHz, DMSO) δ 9.35 (s, 3H), 8.66 (d, *J* = 4.9 Hz, 12H), 8.23 (d, *J* = 7.2 Hz, 7H), 8.07 (s, 1H), 7.99 (t, *J* = 7.8 Hz, 6H), 7.95 (s, 1H), 7.56 (dd, *J* = 7.6, 5.4 Hz, 8H), 5.46 (s, 1H), 4.86 (s, 3H), 1.93 – 1.89 (m, 3H).

Synthesis of N-methyl imidazole cavitand (CHI):



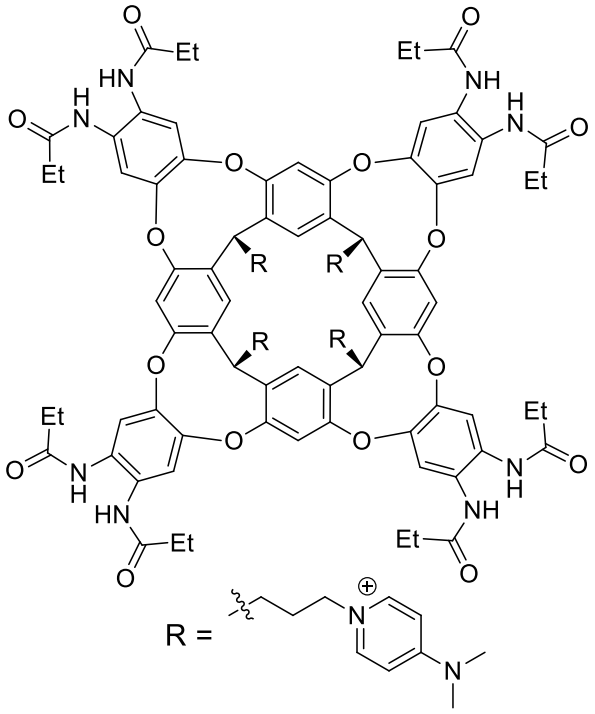
Chloro-benzimidazole cavitand (75 mg, 0.059 mmol) was placed in a round-bottomed flask with excess N-methyl-imidazole (2 mL) and a stir bar and the reaction was stirred at 90 °C for 16 h. The reaction was cooled to room temperature, and cold acetone (2 mL) was added to form a pale-yellow precipitate which was filtered and collected. The solid was then refluxed in acetone (3 mL) for 16 h. The reaction was again cooled and the solid was filtered and dried resulting in **CHI** cavitand (61 mg, 64% yield) as a pale-yellow solid. ^{1}H NMR (400 MHz, DMSO) δ 9.46 (s, 2H), 8.08 (s, 3H), 7.99 (s, 3H), 7.85 (s, 2H), 7.74 (s, 3H), 5.49 (s, 1H), 4.33 (s, 4H), 3.88 (s, 9H), 3.66 (s, 3H), 1.77 (s, 3H), 1.24 (s, 2H).

Synthesis of N-methyl imidazole-amide cavitand (AMI):



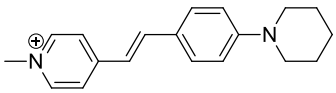
Crude chloro-amide cavitand (75 mg, 0.045 mmol) was placed in a round-bottomed flask with excess N-methyl-imidazole (2 mL) and a stir bar and the reaction was stirred at 90 °C for 16 h. The reaction was cooled and cold acetone (2 mL) was added to form a pale-yellow precipitate which was filtered and collected. The solid was then refluxed in acetone (3 mL) for 16 h. The reaction was again cooled and the solid was filtered and dried resulting in **AMI** cavitand (55 mg, 51% yield) as a pale yellow solid. ^{1}H NMR (400 MHz, D_2O) δ 7.35 (s, 2H), 7.21 (s, 2H), 7.13 (s, 1H), 7.04 (s, 1H), 6.96 (s, 1H), 6.79 (s, 1H), 6.53 (s, 1H), 6.38 (s, 1H), 4.13 (d, J = 7.9 Hz, 2H), 4.05 (d, J = 7.9 Hz, 2H), 3.99 (d, J = 6.8 Hz, 2H), 3.80 (s, 6H), 2.49 (m, 8H), 2.22 (m, 2H), 1.83 (m, 3H), 1.64 (s, 2H), 1.35 (t, J = 7.4 Hz, 6H), 1.22 (t, J = 7.4 Hz, 6H).

Synthesis of 4-dimethylaminopyridine-amide cavitand (AMD):



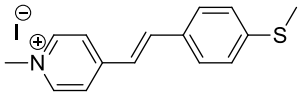
Chloro-amide cavitand (50 mg, 0.031 mmol) was placed in a pear-shaped flask and excess 4-diemthylaminopyridine (148 mg, 1.22 mmol) was placed on top, the reaction was melted at 130 °C for 16 h. The reaction was cooled and cold acetone (2 mL) was added to form a pale beige precipitate which was filtered and collected. The solid was then refluxed in acetone (3 mL) for 16 h. The reaction was again cooled and the solid was filtered and dried resulting in 4-dimethylaminopyrdine-amide cavitand (47 mg, 67% yield) as a beige solid. 1H NMR (400 MHz, Acetone) δ 8.45 (d, J = 6.8 Hz, 3H), 7.92 (s, 1H), 7.66 (s, 1H), 7.41 (s, 1H), 7.03 (d, J = 6.8 Hz, 4H), 4.58 (s, 2H), 3.24 (s, 2H), 3.20 (d, J = 6.0 Hz, 14H), 2.78 (s, 1H), 2.44 – 2.36 (m, 5H), 1.96 – 1.79 (m, 5H), 1.13 (t, J = 7.6 Hz, 10H). ESI MS: m/z $C_{120}H_{140}N_{16}O_{16}^{4+}$ calculated: 515.266, found: 515.493.

Synthesis of (E)-1-methyl-4-(4-(piperidin-1-yl) styryl) pyridin-1-ium iodide (PSMI):



1,4-Dimethylpyridinium iodide (125 mg, 0.529 mmol) and 4-(1-piperidinyl) benzaldehyde (100 mg, 0.529 mmol) were dissolved in ethanol (3 mL) inside a round bottom flask. While stirring, one drop of piperidine was added and the resulting solution was refluxed for 5 h. The reaction was cooled, then diluted with water (5 mL). The resulting precipitate was filtered, rinsed with water and cold ethanol, then dried under vacuum to yield (E)-1-methyl-4-(4-(piperidin-1-yl) styryl) pyridin-1-ium iodide (200 mg, 93% yield) as a dark red powder. ^{1}H NMR (400 MHz, D_2O) δ 8.49 (d, J = 5.6 Hz, 2H), 7.96 (d, J = 5.6 Hz, 2H), 7.76 (d, J = 14.5 Hz, 1H), 7.69 (d, J = 7.3 Hz, 2H), 7.22 (s, 1H), 7.18 (d, J = 7.6 Hz, 2H), 4.24 (s, 3H), 3.28 (m, 4H), 1.71 (m, 4H), 1.63 (d, J = 4.9 Hz, 2H). ^{13}C NMR (150 MHz, D_2O) δ 153.39, 153.46, 143.96, 140.74, 129.60, 123.15, 126.83, 123.15, 119.94, 117.35, 50.35, 46.67, 24.70, 23.43. ESI MS: m/z $C_{19}H_{23}N_2^+$ calculated 279.1856, found: 279.1862. Exc. λ_{max} = 410 nm, Em. λ_{max} = 615 nm.

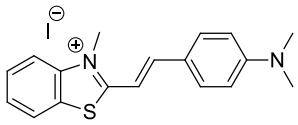
Synthesis of (E)-1-methyl-4-(4-(methylthio) styryl) pyridin-1-ium iodide (SMITE):



1,4-Dimethylpyridinium iodide (235 mg, 1.0 mmol) and 4-(methylthio) benzaldehyde (152 mg, 1.0 mmol) were dissolved in ethanol (5 mL) in a round bottom flask. While stirring, one drop of piperidine was added and the resulting solution was refluxed for 12 h. The reaction was cooled, then diluted with water (10 mL). The resulting precipitate was filtered, rinsed with water and cold ethanol, then dried under vacuum to yield (*E*)-1-methyl-

4-(4-(methylthio) styryl) pyridin-1-ium iodide (340 mg, 92% yield) as a dark yellow powder. ^{1}H NMR (400 MHz, DMSO) δ 8.83 (d, J = 6.8 Hz, 1H), 8.18 (d, J = 6.9 Hz, 1H), 7.97 (d, J = 16.4 Hz, 1H), 7.69 (d, J = 8.5 Hz, 1H), 7.47 (d, J = 16.4 Hz, 1H), 7.36 (d, J = 8.5 Hz, 1H), 4.24 (s, 3H), 2.54 (s, 3H).^{13}C NMR (100 MHz, DMSO) δ 153.03, 145.47, 142.33, 140.67, 131.95, 129.05, 126.20, 123.77, 122.60, 47.34, 14.64. ESI-MS: *m/z* $C_{15}H_{16}NS^{+}$ calculated: 242.0998, found: $(M)^{+}$ 242.1001. UV/Vis: Exc. λ_{max} = 385 nm, Em. λ_{max} = 545 nm.

Synthesis of (E)-2-(4-(dimethylamino)styryl)-3-methylbenzothiazol-3-ium iodide (DTMI):

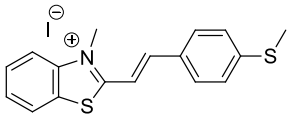


2-methylbenzothiazole (200µL, 1.60 mmol) was dissolved in ethanol (5 mL), iodomethane (1 mL) was added to the reaction mixture while stirring and the reaction was refluxed for 12 h. The solution was diluted with diethyl ether (10 mL) and the resulting precipitate was filtered, then rinsed with diethyl ether and dried under vacuum to yield 2,3-dimethylbenzothiazol-3-ium iodide (398 mg, 87%) as a white solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.43 (dd, J = 8.0, 1.3 Hz, 1H), 8.29 (dt, J = 8.5, 1.2 Hz, 1H), 7.90 (dt, J = 8.5, 1.3 Hz, 1H), 7.81 (dd, J = 8.0, 1.2 Hz, 1H), 4.20 (s, 3H), 3.17 (s, 3H).

2,3-dimethylbenzothiazol-3-ium iodide (290mg, 1.00 mmol) and 4-(dimethylamino)benzaldehyde (149 mg, 1.00 mmol) were dissolved in ethanol (5 mL) inside a round bottom flask. While stirring, one drop of piperidine was added and the resulting solution was refluxed for 12 h. The reaction was cooled, then diluted with water

(10 mL). The resulting precipitate was filtered, rinsed with water and cold ethanol, then dried under vacuum to yield (E)-2-(4-(dimethylamino)styryl)-3- methylbenzothiazol-3-ium iodide (386 mg, 92% yield) as a dark purple powder. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.32 (dd, J = 7.9, 1.5 Hz, 1H), 8.11 (dd, J = 7.7, 1.1, 1H), 8.09 (d, J = 15.3, 1H), 7.93 (d, J = 8.8 Hz, 2H), 7.80 (td, J = 7.9, 1.1 Hz, 1H), 7.70 (td, J = 7.7, 1.5 Hz, 1H), 7.65 (d, J = 15.3 Hz, 1H), 6.87 (d, J = 8.8 Hz, 2H), 4.24 (s, 3H), 3.13 (s, 6H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 150.61, 142.45, 133.31, 129.36, 127.94, 127.29, 124.28, 121.95, 116.43, 112.46, 106.76, 40.51, 36.01. ESI-MS: *m/z* $C_{18}H_{19}N_2S^+$ calculated: 295.4213, found: $(M)^+$ 295.1270. UV/Vis: Exc. λ_{max} = 510 nm, Em. λ_{max} = 600 nm.

Synthesis of (E)-3-methyl-2-(4-(methylthio) styryl) benzo[d]thiazol-3-ium iodide (SMITH):

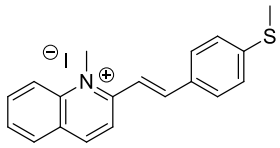


2-methylbenzothiazole (200µL, 1.60 mmol) was dissolved in ethanol (5 mL), iodomethane (1 mL) was added to the reaction mixture while stirring and the reaction was refluxed for 12 h. The solution was diluted with diethyl ether (10 mL) and the resulting precipitate was filtered, then rinsed with diethyl ether and dried under vacuum to yield 2,3-dimethylbenzothiazol-3-ium iodide (398 mg, 87%) as a white solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.43 (dd, J = 8.0, 1.3 Hz, 1H), 8.29 (dt, J = 8.5, 1.2 Hz, 1H), 7.90 (dt, J = 8.5, 1.3 Hz, 1H), 7.81 (dd, J = 8.0, 1.2 Hz, 1H), 4.20 (s, 3H), 3.17 (s, 3H).

2,3-dimethylbenzothiazol-3-ium iodide (290mg, 1.00 mmol) and 4-(methylthio) benzaldehyde (140 µL, 1.00 mmol) were dissolved in ethanol (5 mL) inside a round bottom

flask. While stirring, one drop of piperidine was added and the resulting solution was refluxed for 12 h. The reaction was cooled, then diluted with water (10 mL). The resulting precipitate was filtered, rinsed with water and cold ethanol, then dried under vacuum to yield (E)-3-methyl-2-(4-(methylthio) styryl) benzo[d]thiazol-3-ium iodide (341 mg, 80% yield) as a dark orange powder. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.42 (d, J = 15.6 Hz, 1H), 8.22 (m, 2H), 8.01 (m, 3H), 7.88 (td, J = 7.2, 1.3 Hz, 1H), 7.79 (td, J = 7.2, 1.0 Hz, 1H), 7.43 (d, J = 8.3 Hz, 2H), 4.35 (s, 3H), 2.58 (s, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 172.01, 148.25, 145.12, 142.11, 130.34, 130.25, 129.43, 128.42, 127.83, 125.53, 124.28, 116.87, 112.69, 36.40, 14.05. ESI-MS: *m/z* $C_{17}H_{16}NS_2^+$ calculated: 299.3254, found: $(M)^+$ 299.2941. UV/Vis: Exc. λ_{max} = 415 nm, Em. λ_{max} = 560 nm.

Synthesis of (E)-1-methyl-2-(4-(methylthio) styryl) quinolin-1-ium iodide (SMIQ):



2-methylquinoline (250 µL, 1.88 mmol) was dissolved in ethanol (3 mL), iodomethane (0.5 mL) was added to the reaction mixture while stirring and the reaction was refluxed for 12 h. The solution was diluted with diethyl ether (10 mL) and the resulting precipitate was filtered, then rinsed with diethyl ether and dried under vacuum to yield 1,2-dimethylquinolin-1-ium iodide (457 mg, 85%) as a yellow solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.09 (d, J = 8.5 Hz, 1H), 8.59 (d, J = 9.1 Hz, 1H), 8.40 (d, J = 8.0 Hz, 1H), 8.23 (t, J = 8.0 Hz, 1H), 8.12 (d, J = 8.6 Hz, 1H), 7.99 (t, J = 7.5 Hz, 1H), 4.44 (s, 3H), 3.07 (s, 3H).

1,2-dimethylquinolin-1-ium iodide (150 mg, 0.50 mmol) and 4-(methylthio) benzaldehyde (70 µL, 0.50 mmol) were dissolved in ethanol (5 mL) inside a round bottom flask. While stirring, one drop of piperidine was added and the resulting solution was refluxed for 12 h. The reaction was cooled, then diluted with water (10 mL). The resulting precipitate was filtered, rinsed with water and cold ethanol, then dried under vacuum then recrystallized with toluene to yield (E)-1-methyl-2-(4-(methylthio) styryl) quinolin-1-ium iodide (134 mg, 61% yield) as a dark purple powder. ^{1}H NMR (400 MHz, DMSO- d_6) δ 9.05 (d, J = 8.9 Hz, 1H), 8.57 (dd, J = 9.0, 6.1 Hz, 2H), 8.35 (dd, J = 8.1, 1.6 Hz, 1H), 8.25 – 8.14 (m, 2H), 8.00 – 7.92 (m, 3H), 7.90 (d, J = 11.9 Hz, 1H), 7.42 (d, J = 8.5 Hz, 2H), 4.56 (s, 3H), 2.57 (s, 3H). ^{13}C NMR (150 MHz, DMSO-d_6) δ 156.69, 147.13, 144.38, 144.09, 139.71, 135.35, 131.68, 130.55, 130.17, 129.44, 128.21, 126.03, 121.48, 119.82, 118.53, 40.40, 14.55. ESI-MS: *m/z* $C_{19}H_{18}N_2S^+$ calculated: 292.1154, found: $(M)^+$ 292.0802. UV/Vis: Exc. λ_{max} = 410 nm, Em. λ_{max} = 580 nm.

6.3. Experimental for Chapter 3

General Procedure for titrations:

Separate 1 mM solutions of each host **3.6**, **3.7**, and **3.8** were prepared in 300 µL D_2O and placed in an NMR tube. A ^{1}H NMR spectrum was taken. To the NMR tubes, racemic mixtures of selected pheromone guests, **3.2** and **3.4**, (0.5 – 4.0 mM in D_2O) were titrated into the host solutions. A ^{1}H NMR spectrum was taken each time.

General procedure for fluorescence measurements:

The fluorescence assays were carried out by mixing 10 μL of the fluorescent dye **3.9** or **3.10** (3 μM in water), 10 μL of cavitand host **3.6**, **3.7**, or **3.8** (20 μM in water), 10 μL metal salts (50 μM in water), and 10 μL of each pheromone guest molecule (50 μM in water), then adding 60 μL of the incubation buffer (Tris buffer HCl, pH 7.4, 20 mM) to bring the total volume up to 100 μL for each well in the 96-well plate, then incubating with mild agitation for 15 min at room temperature. The fluorescence signal (F) was recorded with the Ex/Em wavelengths at 485/600 or 485/528 nm for all fluorescent dyes. Fluorescence response (F/F_0) are normalized to the response of cavitand:dye:M^+ complex in the absence of guest, that is F_0 is defined as the fluorescence recorded for that concentration of host and guest when [pheromone] = 0 μM. Each experimental condition was repeated in quadruplicate across four separate wells of the 96-well plate using identical sensor components, simultaneously collecting fluorescence signals for each target at once.

The quadruplicate raw fluorescence data sets were subjected to Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). PCA and confidence ellipses were performed with RStudio (Version 1.2.5019), an integrated development environment (IDE) for R (version 3.6.1). 2D scores plots and confidence intervals, as well as biplots were graphed in RStudio using the packages ggplot2, ggpubr, ggfortify, devtools, and factoextra. 3D PCA plots were graphed in OriginPro. Supervised LDA was performed in Python 3.9. The raw fluorescence data sets of two guests were subjected to StandardScaler for standardization, then LinearDiscriminantAnalysis was used for classification, resulting in the transformed LD 1 scores. For each class of samples,

the probability density of t-distribution and 95% confidence intervals of LD 1 values were generated and calculated by scipy.stats.t. The cross validation was performed using RepeatedStratifiedKFold(n_splits=4, n_repeats=10) with LinearDiscriminantAnalysis as the classification estimator. Feature selection was performed with Python 3.9, using StandardScaler for data standardization.

Recursive feature elimination with cross-validation (RFECV) to select the number of features, Support Vector Machine (SVM) (kernel='linear') as the supervised classification estimator, RFECV(estimator=svm.SVC(kernel='linear'), step=1, cv=StratifiedKFold(n_splits=4,shuffle=True),scoring='accuracy',min_features_to_select= 1. Performance metrics for the classification evaluation were calculated by using RepeatedStratifiedKFold (n_splits=4, n_repeats=3) for cross validation. The correlation heatmap of selected features was computed using pandas.DataFrame.corr(method='pearson'). PCA was applied for orthogonal linear transformation and dimensionality reduction, and SVM decision region boundary of PCA plot was generated using plot_decision_regions. All other fluorescence data charts were created in Microsoft Excel, with values representing the mean of the quadruplicate responses and error bars indicating their standard deviation.

6.4. Experimental for Chapter 4

General Procedure for titrations:

Separate 1 mM solutions of hosts **4.3** and **4.4**, as well as the **AMD analogue**, were prepared in 300 µL D_2O and placed in an NMR tube. A ^1H NMR spectrum was taken. To

the NMR tubes, mixtures of selected anion guests, NaCl, NaBr, and NaI (1– 10 mM in D_2O) were titrated into the host solutions. A ^1H NMR spectrum was taken each time.

General procedure for fluorescence measurements:

The fluorescence assays were carried out by mixing 10 μL of the fluorescent dye **4.5**, **4.6**, **4.7** or **4.8** (5 μM in water), 10 μL of cavitand host **4.1**, **4.2**, **4.3**, or **4.4** (5 μM in water), and 10 μL of each anion guest molecule (25 - 200 mM in water), then adding 70 μL of the incubation buffer (Tris buffer HCl, pH 7.4, 20 mM, ultrapure H_2O, or 10X PBS buffer) to bring the total volume up to 100 μL for each well in the 96-well plate, then incubating with mild agitation for 15 min at room temperature. The fluorescence signal (F) was recorded with the Ex/Em wavelengths at 485/600 or 485/528 nm for all fluorescent dyes. Fluorescence response (F/F_0) are normalized to the response of cavitand:dye complex in the absence of guest, that is F_0 is defined as the fluorescence recorded for that concentration of host and guest when [anion] = 0 μM. Each experimental condition was repeated in quadruplicate across four separate wells of the 96-well plate using identical sensor components, simultaneously collecting fluorescence signals for each target at once.

The quadruplicate raw fluorescence data sets were subjected to Principal Component Analysis (PCA). PCA and confidence ellipses were performed with RStudio (Version 1.2.5019), an integrated development environment (IDE) for R (version 3.6.1). 2D scores plots and confidence intervals, as well as biplots were graphed in RStudio using the packages ggplot2, ggpubr, ggfortify, devtools, and factoextra.

Fluorescence response curves of dye in the presence of guest anions upon titration of hosts were obtained by using 0.5 μM dye, 0-50 μM cavitand, and 50 mM guest in

ultrapure H_2O. The binding affinities were achieved by fitting F/F_0 data by Hill 1 function in Growth/Sigmoidal category from Origin software.

General procedure for fluorescence titrations and affinity measurements:

The fluorescence titrations were carried out by mixing 10 µL of the fluorescent dye **4.5** (0.5 µM in water), 10 µL of cavitand host **4.3** (0 - 50 µM in water), an 10 µL of anion guest, NaCl, NaBr, NaI, (50 mM in water), then adding 70 µL of the incubation buffer (ultrapure H_2O, pH 7.4) to bring the total volume up to 100 µL for each well in the 96-well plate, then incubating with mild agitation for 15 min at room temperature. The fluorescence signal (F) was recorded with the Ex/Em wavelengths at 485/600. Fluorescence response (F/F_0) are normalized to the response of the dye alone in the absence of host, that is F_0 is defined as the fluorescence recorded for that concentration when [host] = 0 µM. Each experimental condition was repeated in quadruplicate across four separate wells of the 96-well plate using identical sensor components, simultaneously collecting fluorescence signals for each target at once. The affinity of the dye for the host was determined by the Hill 1 function.

Hill 1 function is a more general form of Hill function. The equation is:

$$y = START + (END - START)\frac{x^n}{k^n + x^n}$$

x is the ligand concentration, *k* is half-maximal concentration constant, *n* is Hill coefficient. Job plots were obtained by measuring the fluorescence intensity of solutions containing dye and guest with the dye mole fraction (X_{Dye}) varied between 0 and 1.

General procedure for isothermal titration calorimetry (ITC) measurements:

All ITC experiments were performed using a MicroCal with a stirring rate of 700 rpm. The baseline was stabilized prior to the experiment, and a pre-injection delay was set to 60 s. A stock solution of guest titrant, NaCl, NaBr, or NaI, dissolved in ultrapure H_2O was added in 2 µL aliquots to the cavitand host, **4.1**, **4.2**, **4.3**, or **4.4**, solution of 1 mM in ultrapure H_2O, respectively. All experiments were conducted at 20 °C. The heat of dilution, measured by the injection of titrant into H_2O, was subtracted for each titration to obtain the net reaction heat value. Curve fitting was performed by the MicroCal program using the One Set of Sites model.

General procedure for geometry optimizations:

Geometry optimizations were performed using the r^2SCAN meta-generalized gradient approximation (GGA) density functional with the D3(BJ) dispersion correction and the def2-SVP basis set. Solvent effects were accounted for in the geometry optimizations using the conductor-like polarizable continuum model, CPCM, with water as the solvent. Density fitting using the def2/J auxiliary basis was used to accelerate the calculation. All geometry optimizations were performed using the ORCA 5.0 software package.

6.5. Experimental for Chapter 5

The fluorescence assays were carried out by mixing 10 µL of the fluorescent dye **DSMI**, **DTMI**, **SMITE**, or **SMITH** (5 µM in water), 10 µL of cavitand host **TCC**, **AMI**, **CHI**, or **CHP** (5 µM in water), and 10 µL of each phosphosugar guest molecule (50 µM in water), then adding 70 µL of the incubation buffer (Tris buffer HCl, pH 7.4) to bring the total volume up to 100 µL for each well in the 96-well plate, then incubating with mild

agitation for 15 min at room temperature. The fluorescence signal (F) was recorded with the Ex/Em wavelengths at 485/600 or 485/528 nm for all fluorescent dyes. Fluorescence response (F/F_0) are normalized to the response of cavitand:dye complex in the absence of guest, that is F_0 is defined as the fluorescence recorded for that concentration of host and dye when [phosphosugar] = 0 μM. Each experimental condition was repeated in quadruplicate across four separate wells of the 96-well plate using identical sensor components, simultaneously collecting fluorescence signals for each target at once.

The quadruplicate raw fluorescence data sets were subjected to Principal Component Analysis (PCA). PCA and confidence ellipses were performed with RStudio (Version 1.2.5019), an integrated development environment (IDE) for R (version 3.6.1). 2D scores plots and confidence intervals, as well as biplots were graphed in RStudio using the packages ggplot2, ggpubr, ggfortify, devtools, and factoextra.

6.6. Selected Spectra for Chapter 2

Selected hosts and dyes are known, such as **TCC** and **DSMI**, and their spectra have been previously published, see references.[1,2] ^{1}H NMR spectra of the specific samples used are published here to illustrate their purity and ensure reproducibility.

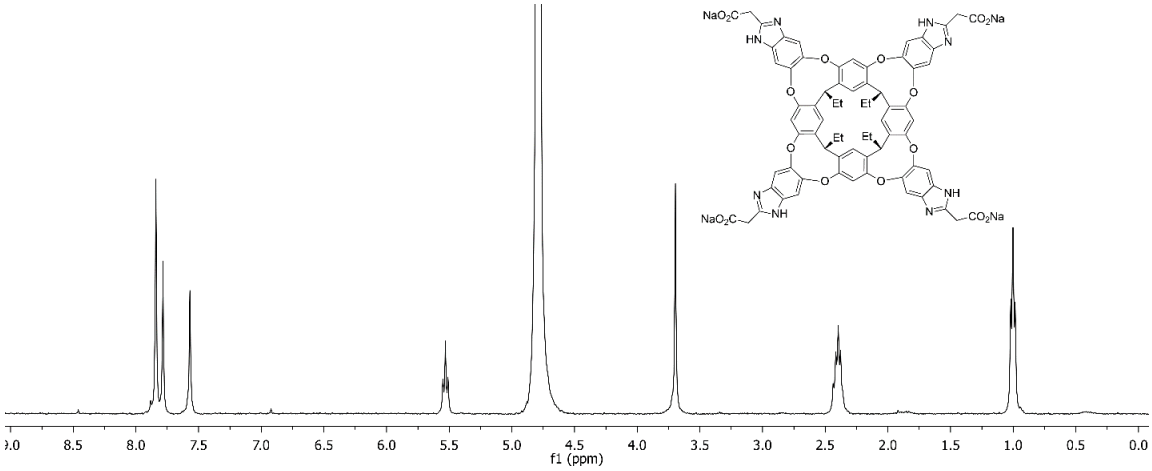


Figure 6.1. ^{1}H NMR spectrum of **TCC** cavitand (D_2O, 400 MHz, 298K).

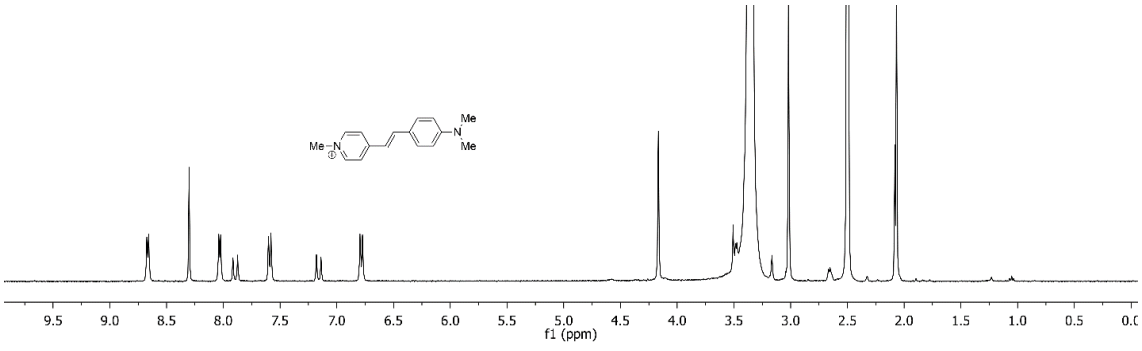


Figure 6.2. ^{1}H NMR spectrum of **DSMI** fluorophore (DMSO-d_6, 400 MHz, 298K).

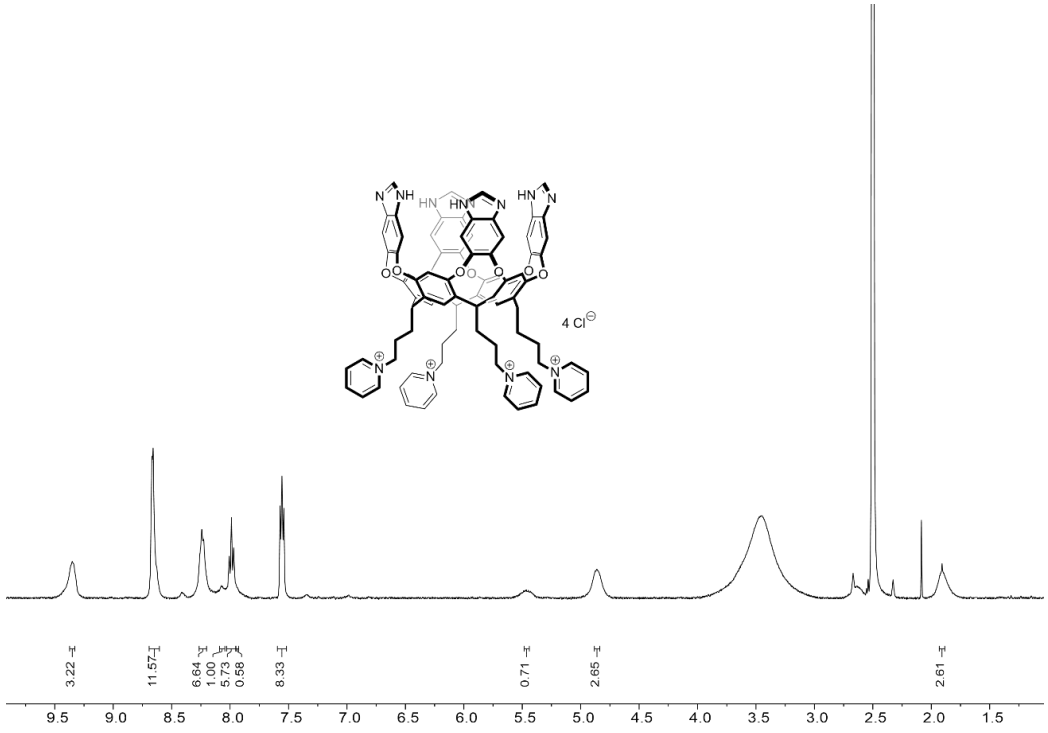


Figure 6.3. ^{1}H NMR spectrum of **CHP** cavitand (DMSO-d_6, 400 MHz, 298K).

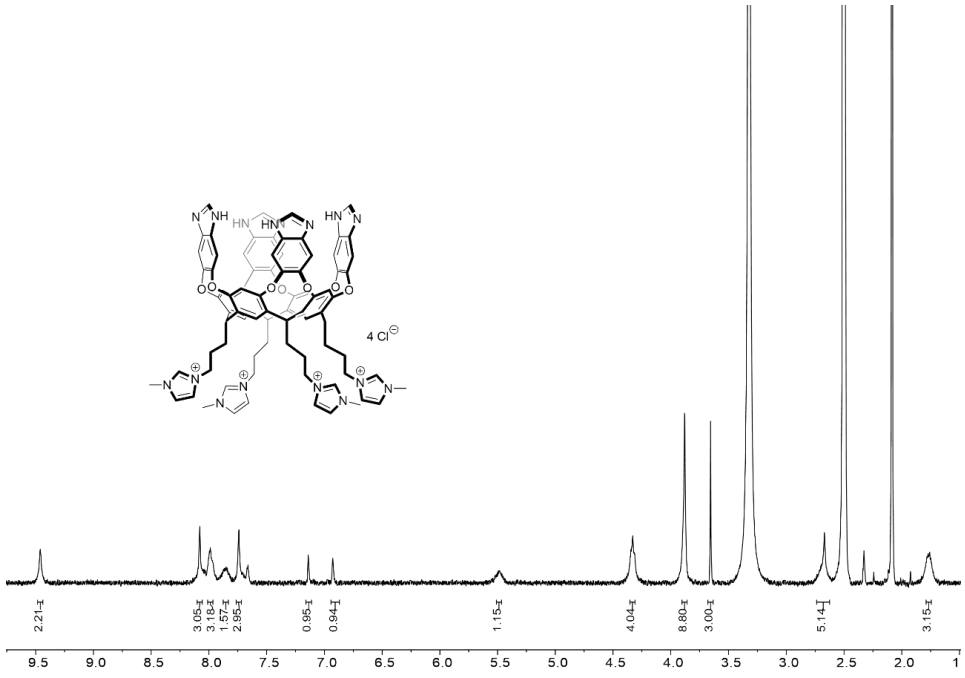


Figure 6.4. ^{1}H NMR spectrum of **CHI** cavitand (DMSO-d_6, 400 MHz, 298K).

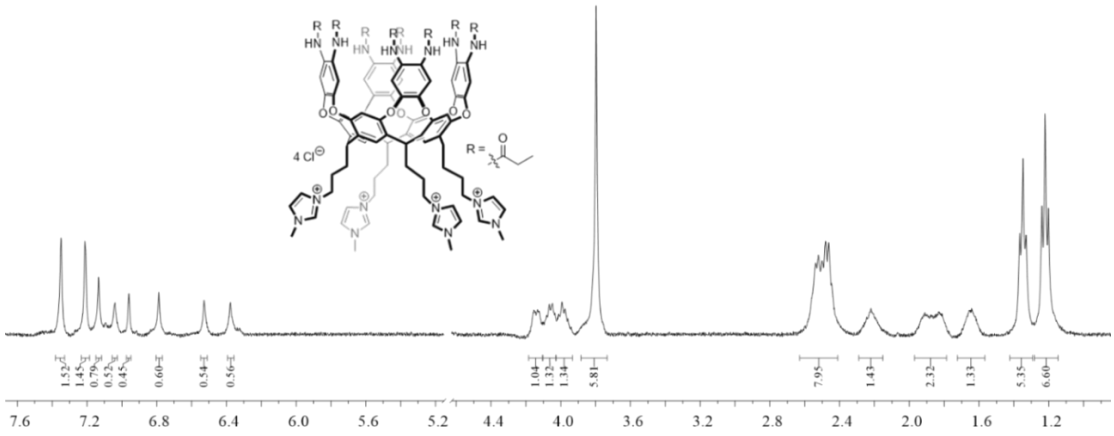


Figure 6.5. ^{1}H NMR spectrum of **AMI** cavitand (D_2O, 400 MHz, 298K).

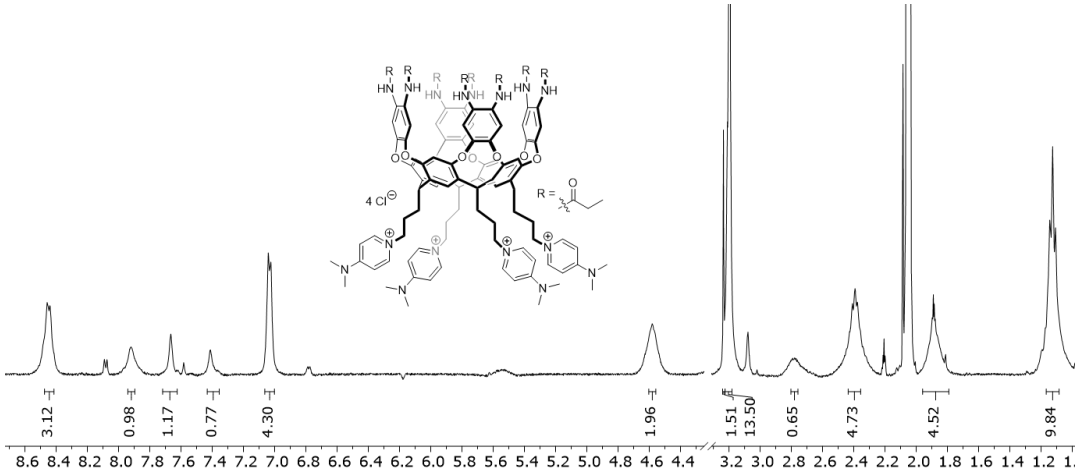


Figure 6.6. ^{1}H NMR spectrum of **AMD** cavitand (D_2O, 400 MHz, 298K).

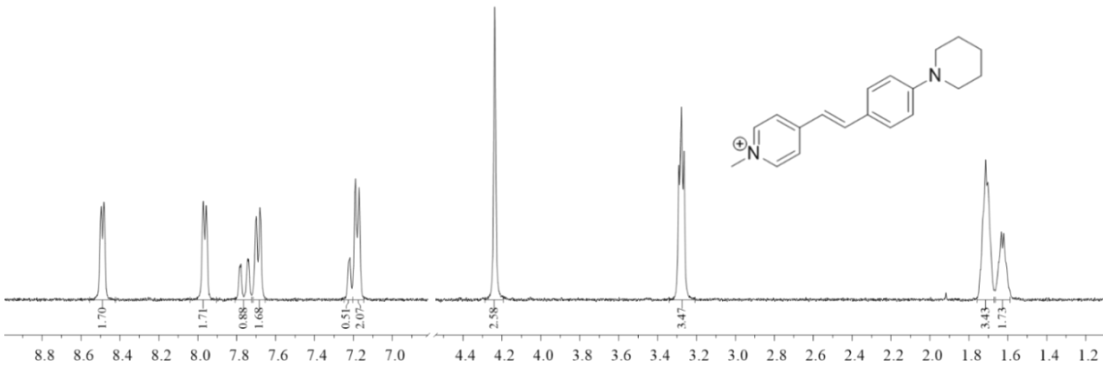


Figure 6.7. ^{1}H NMR spectrum of fluorophore **PSMI** (D_2O, 400 MHz, 298K).

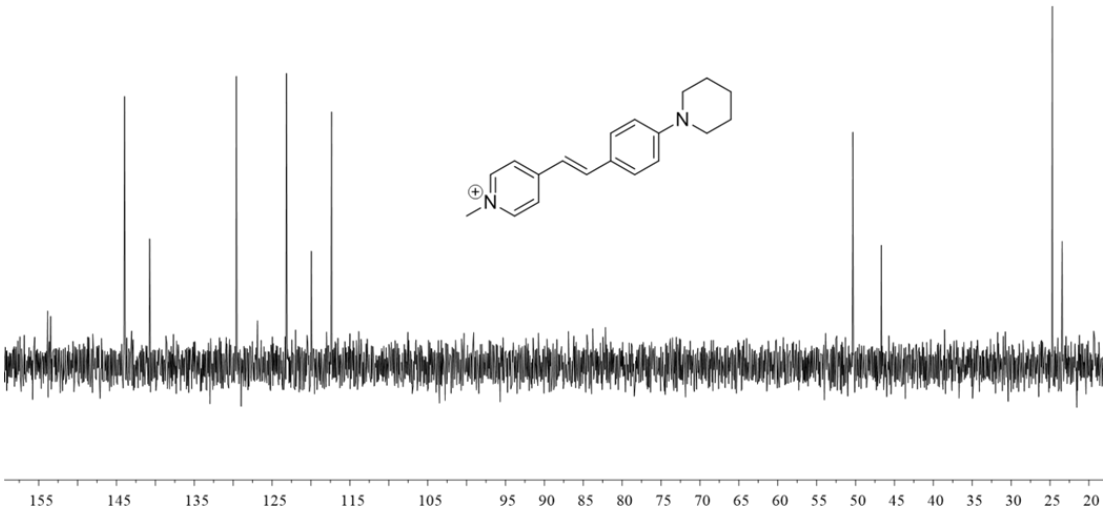


Figure 6.8. ^{13}C NMR spectrum of fluorophore **PSMI** (D_2O, 150 MHz, 298K).

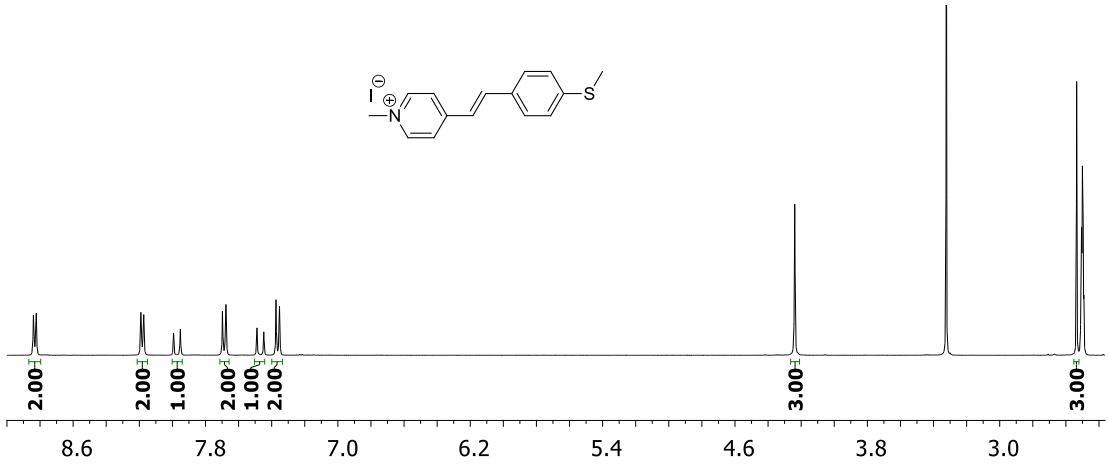


Figure 6.9. ^{1}H NMR spectrum of **SMITE** (DMSO-d_6, 400 MHz, 298K).

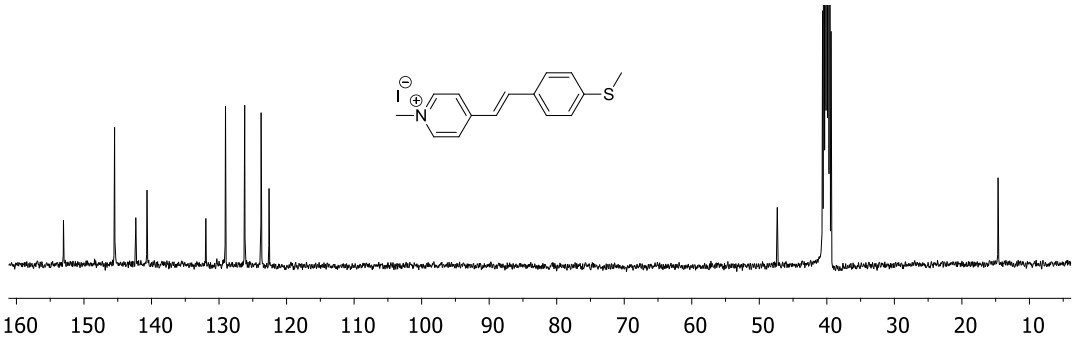


Figure 6.10. ^{13}C NMR of **SMITE** (DMSO-d_6, 100 MHz, 298K).

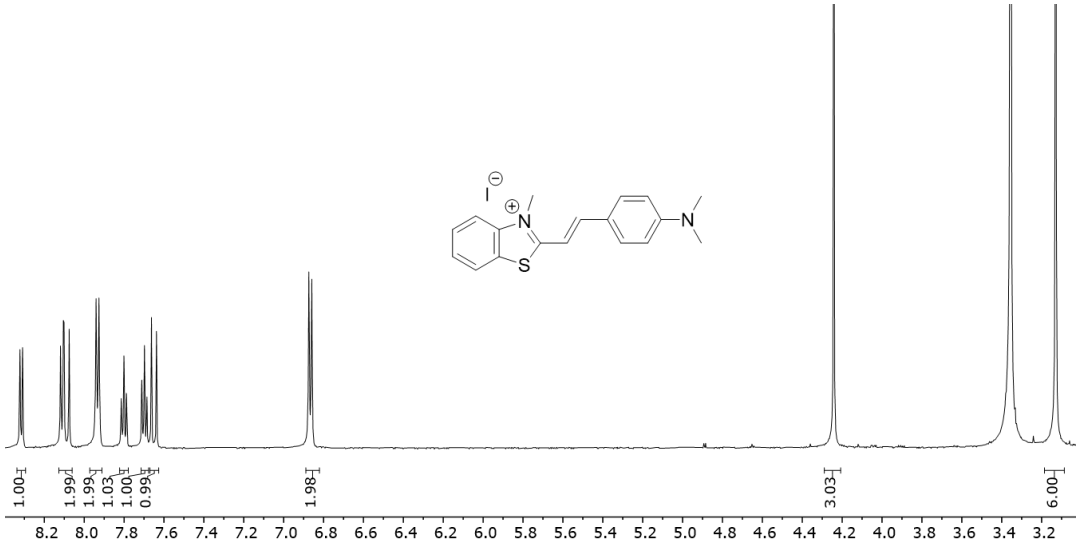


Figure 6.11. ^{1}H NMR spectrum of **DTMI** (DMSO-d_6, 400 MHz, 298K).

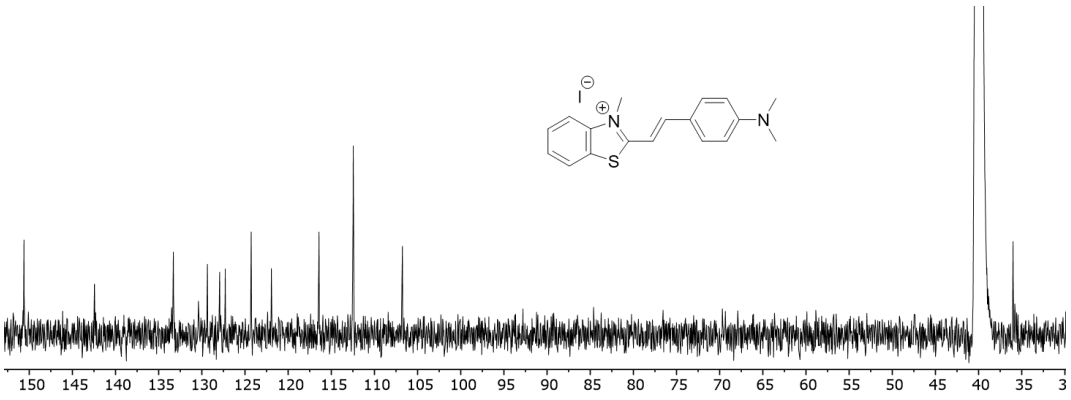


Figure 6.12. ^{13}C NMR of **DTMI** (DMSO-d_6, 100 MHz, 298K).

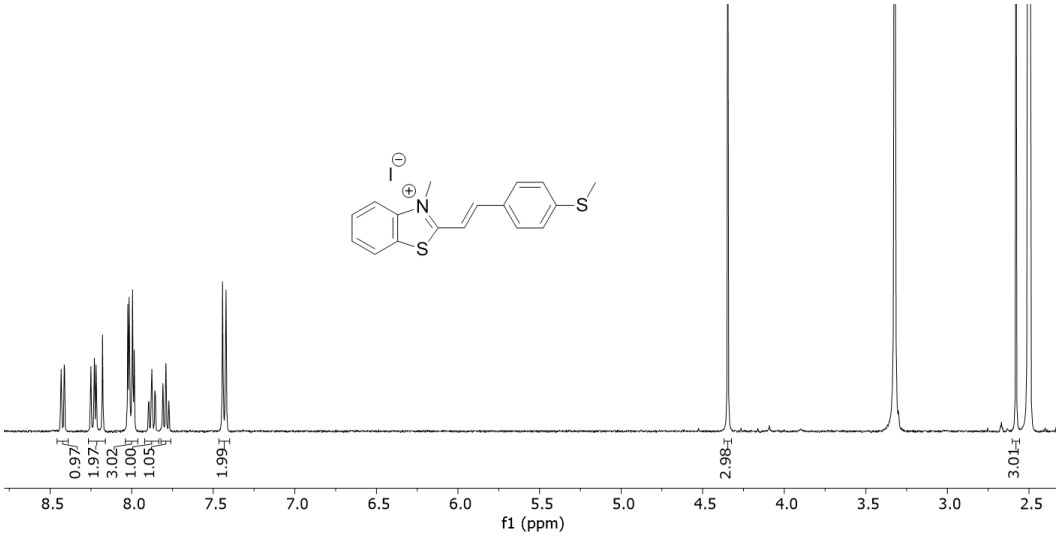


Figure 6.13. ^{1}H NMR spectrum of **SMITH** (DMSO-d_6, 400 MHz, 298K).

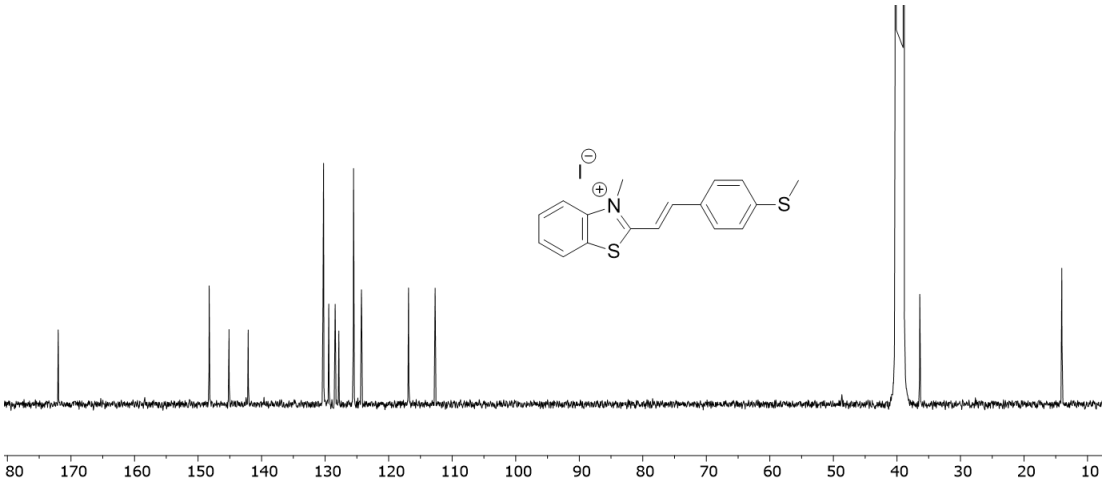


Figure 6.14. ^{13}C NMR of **SMITH** (DMSO-d_6, 100 MHz, 298K).

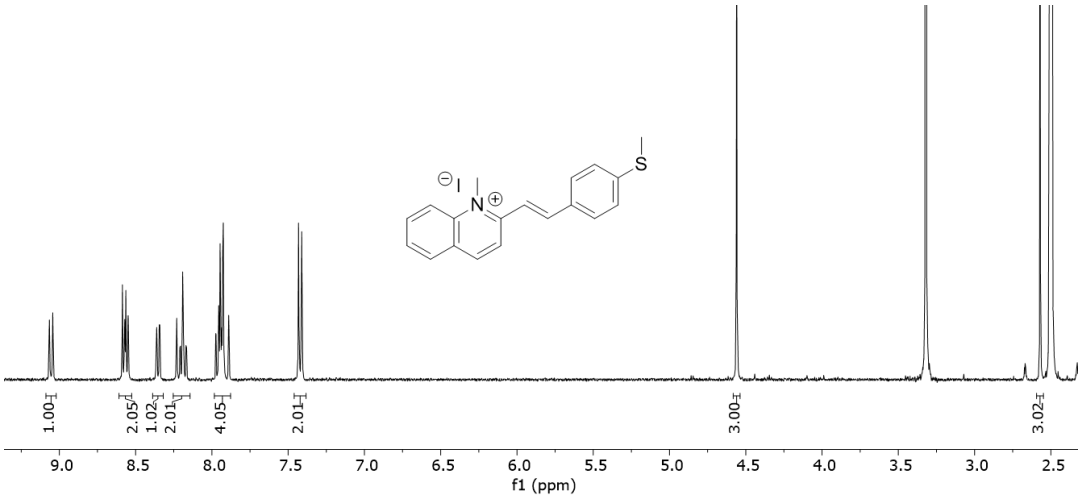


Figure 6.15. ^{1}H NMR spectrum of **2-SMIQ** (DMSO-d_6, 400 MHz, 298K).

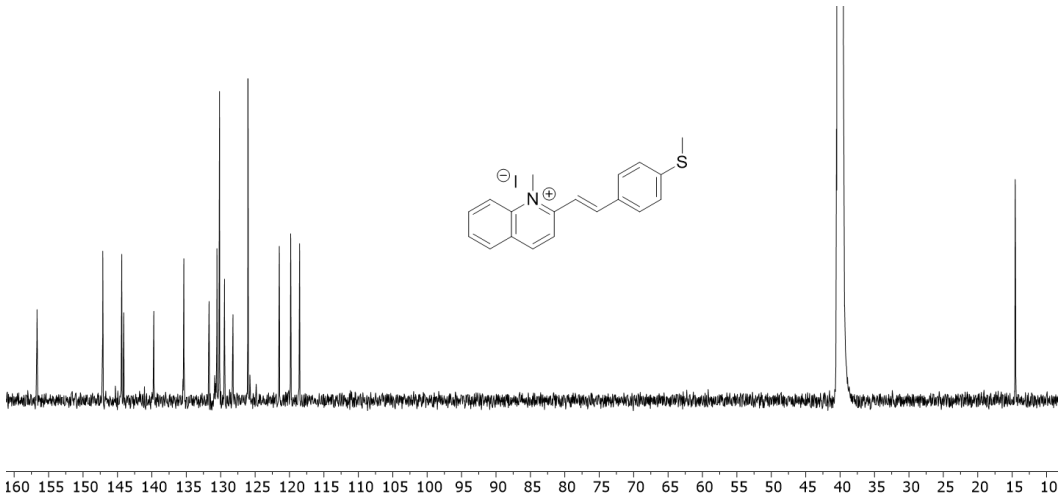


Figure 6.16. ^{13}C NMR of **2-SMIQ** (DMSO-d_6, 100 MHz, 298K).

NMR Analysis of Host:Guest Binding

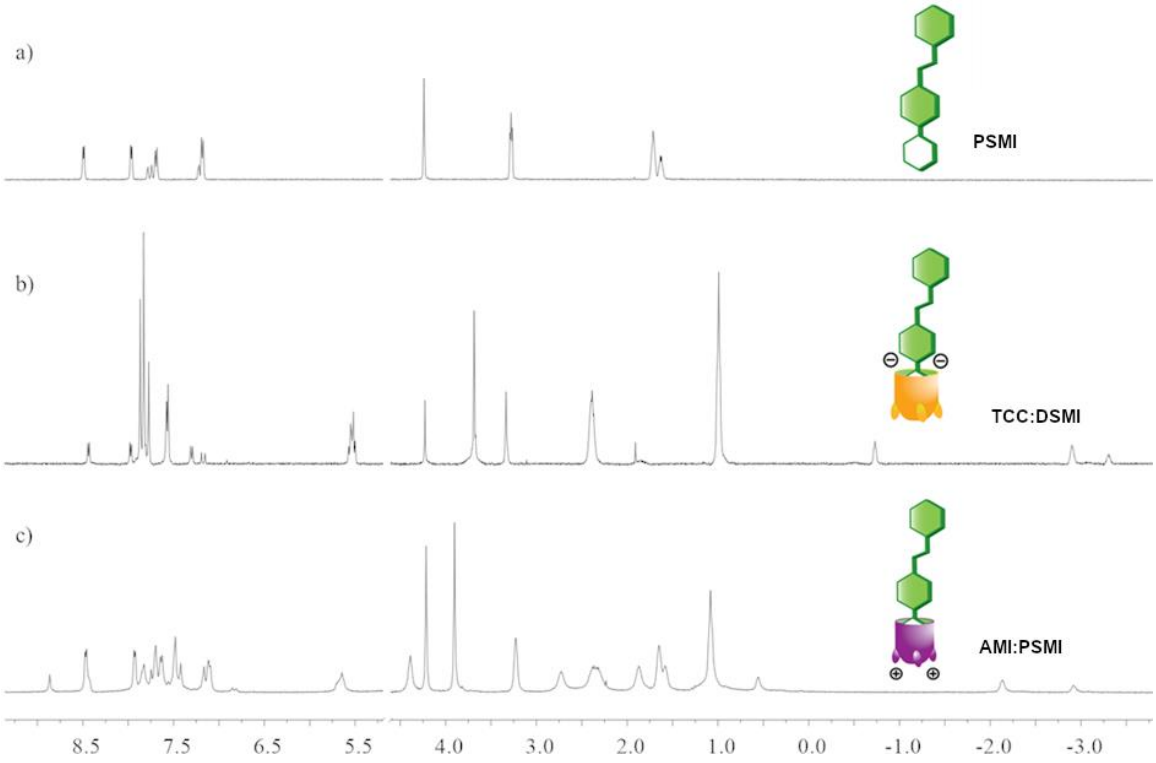


Figure 6.17. ^{1}H NMR spectra (D_2O, 400 MHz, 298 K) of a) guest **PSMI**; b) guest **PSMI** binding in host **TCC**; c) guest **PSMI** binding in host **AMI**.

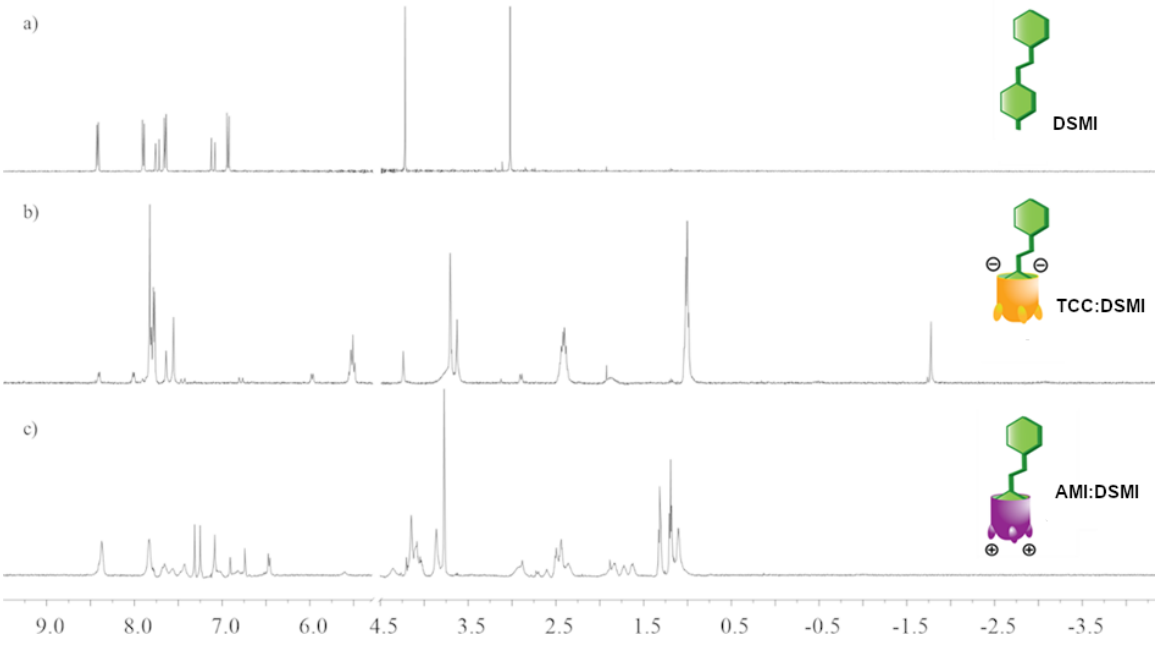


Figure 6.18. ^{1}H NMR spectra (D_2O, 400 MHz, 298 K) of a) guest **DSMI**; b) guest **DSMI** binding in host **TCC**; c) guest **DSMI** in rapid exchange with host **AMI**.

Fluorescence Titrations of Dyes and Hosts

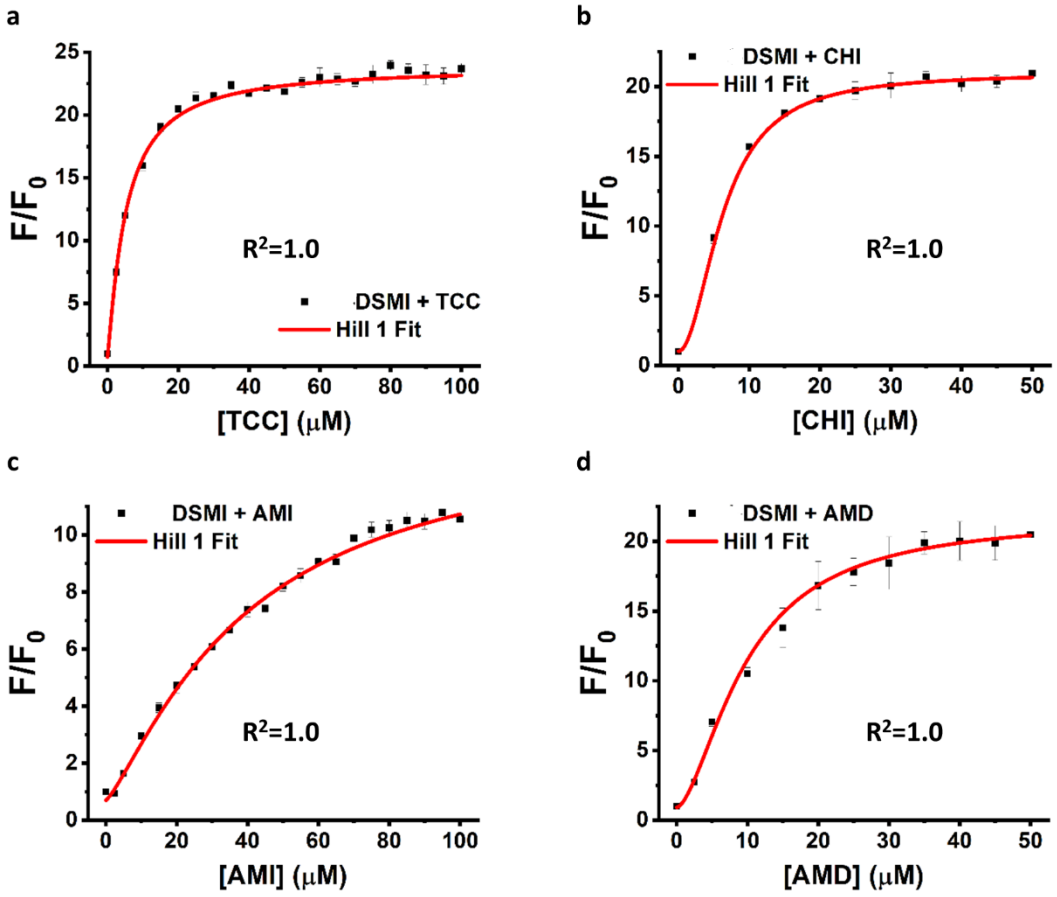


Figure 6.20. Affinity measurement of **DSMI** with cavitand hosts via fluorescence, where a) **DSMI**:**TCC**; b) **DSMI**:**CHI**; c) **DSMI**:**AMI;** d) **DSMI**:**AMD** complexes. [**DSMI**] = 0.625μM, Buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4.

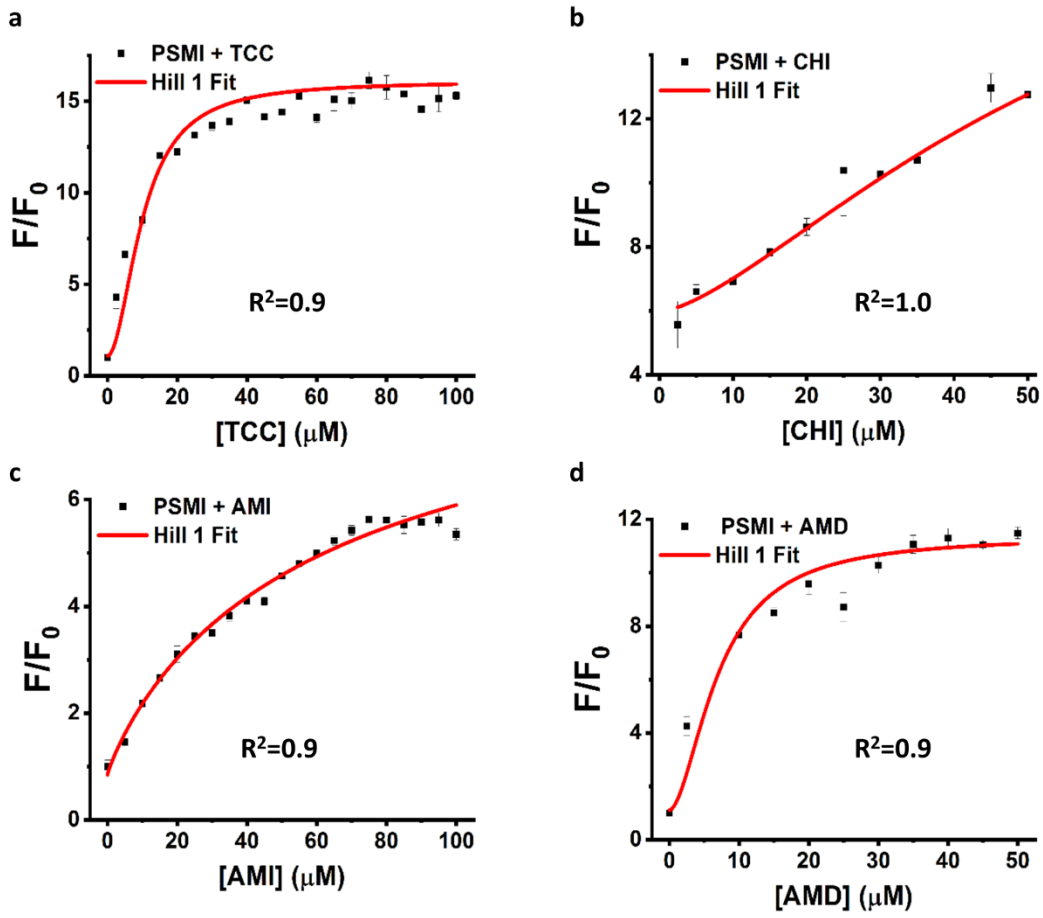


Figure 6.21. Affinity measurement of **PSMI** with cavitand hosts via fluorescence, where a) **PSMI**:**TCC**; b) **PSMI**:**CHI**; c) **PSMI**:**AMI;** d) **PSMI**:**AMD** complexes. [**PSMI**] = 0.625μM, Buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4.

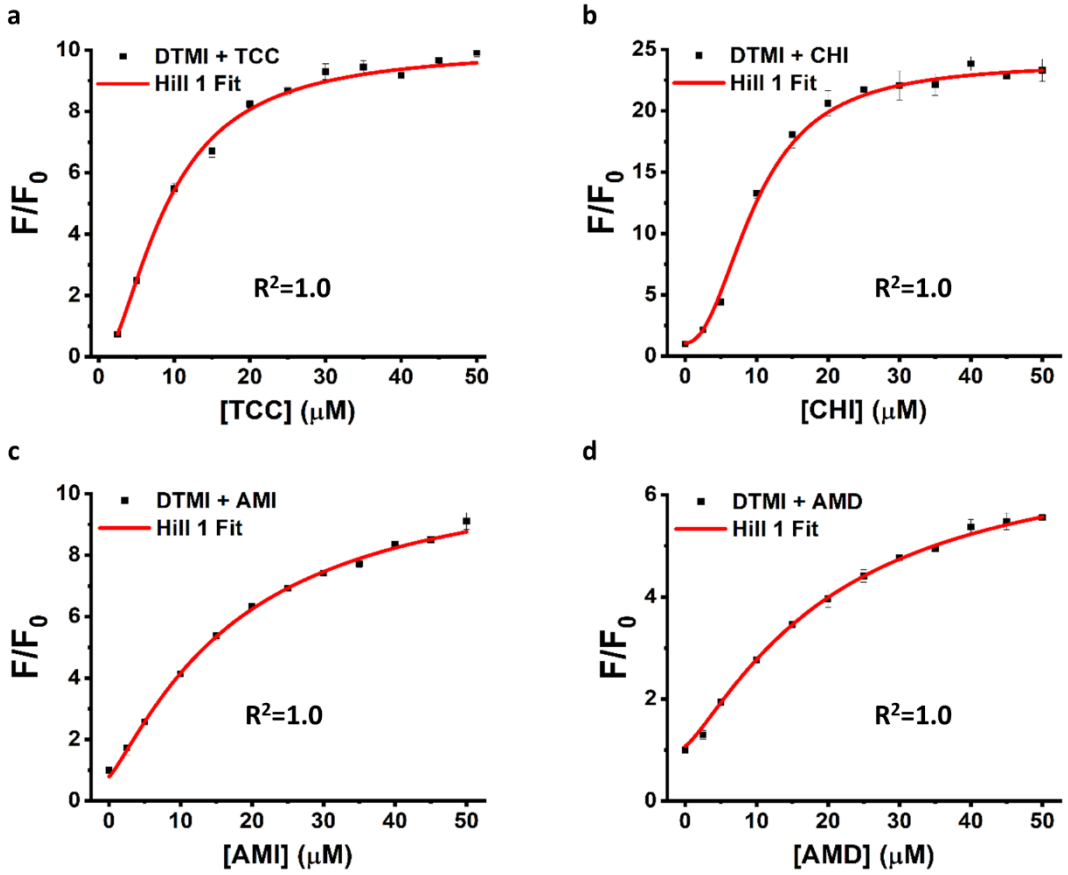


Figure 6.22. Affinity measurement of **DTMI** with cavitand hosts via fluorescence, where a) **DTMI**:**TCC**; b) **DTMI**:**CHI**; c) **DTMI**:**AMI;** d) **DTMI**:**AMD** complexes. [**DTMI**] = 0.625μM, Buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4, Ex/Em = 540nm/600 nm.

K_d (μM)	DSMI	PSMI	DTMI
TCC	5.2±2.0	10.1±7.6	8.9±0.6
CHI	6.4±2.6	57.3±33.0	9.9±0.5
AMI	39.8±6.2	66.1±37.5	18.2±3.1
AMD	9.8±0.9	7.1±3.5	20.9±3.3

Table 6.1. The binding affinities of the cavitand:dye complexes: **DSMI**/**PSMI**/**DTMI** with different hosts **TCC/CHI/AMI/AMD** were obtained using Hill 1 fitting of data from ITC data.

6.7. Selected Spectra and Characterization for Chapter 3

Octanediols *S,S*-3.4, and ***R,R*-3.4:**

(2*S*,3*S*)-2,3-Octanediol (***S,S*-3.4**) and (2*R*,3*R*)-2,3-octanediol (***R,R*-3.4**) were prepared as shown in Figure 6.23, following a similar route to the stereospecific syntheses of (2*R*,3*R*)- and (2*S*,3*S*)-2,3-hexanediols from (L)- and (D)-threonine.[3]

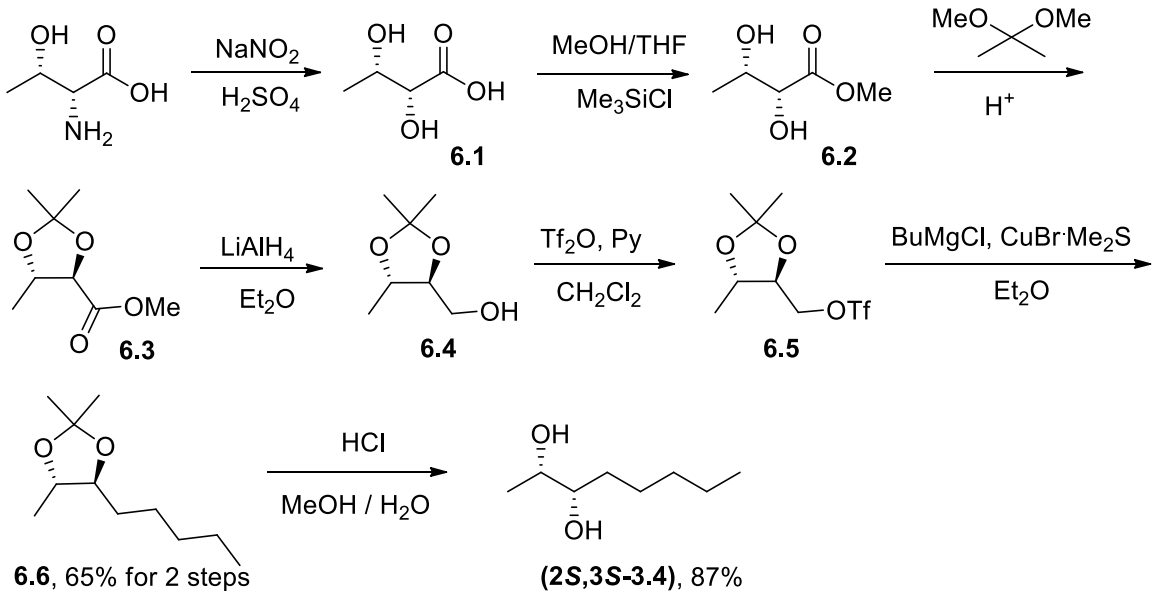


Figure 6.23. Synthetic route for pheromone targets (2*S*, 3*S*)-2,3-octanediol (***S,S*-3.4**) and (2*R*, 3*R*)-2,3-octanediol (***R,R*-3.4**).

Synthesis of (4*S*,5*S*)-2,2,4-Trimethyl-5-pentyl-[1,3]dioxolane (6.6):

(4*S*,5*S*)-4-hydroxymethyl-2,2,5-trimethyl-[1,3]-dioxolane **6.4** was prepared in 4 steps from (D)-threonine as described in previous literature.[4] Briefly, (D)-threonine was converted to the corresponding diol **6.1** by treatment with aqueous $NaNO_2$ and sulfuric acid, followed by sequential methylation of the carboxylic acid to give ester **6.2**, and ketalization of the vicinal diol with dimethoxypropane, yielding ester **6.3**. Reduction of ester **6.3** then gave (4*S*,5*S*)-4-hydroxymethyl-2,2,5-trimethyl-[1,3]-dioxolane **6.4**. Pyridine (1.6 mL, 20 mmol) and triflic anhydride (4.0 mL, 24 mmol) were added sequentially to a cold (-15 °C), stirred solution of (4*S*,5*S*)-4-hydroxymethyl-2,2,5-trimethyl-[1,3]-dioxolane **6.4** (2.92 g, 20 mmol) in dry CH_2Cl_2 (80 mL). The reaction was warmed to 0 ºC and stirred for 1 h, then diluted with hexane (160 mL) and filtered through a pad of Celite. The filtrate was concentrated in vacuo to afford crude triflate **6.5**, which was used immediately in the

next step without further purification. *n*-BuMgCl (20 mL, 2 M in THF, 40 mmol) was added to a suspension of $CuBr{\cdot}Me_2S$ (0.82 g, 4.0 mmol) in Et_2O (80 mL) at 0 °C followed by triflate **6.5** in Et_2O (40 mL). The mixture was stirred for 2.5 h while warming to room temperature, then poured into a solution of saturated aqueous NH_4Cl and $NH_3{\cdot}H_2O$ (9:1) and extracted with Et_2O. The ether extract was washed with water and brine, then dried over anhydrous Na_2SO_4. The crude product was purified by flash chromatography (hexane/EtOAc = 40/1) to give **6.6** as a colorless liquid (2.43 g, 65%). The ^{1}H NMR spectrum matched that previously reported.[5,6]

Synthesis of (2*S*,3*S*)-2,3-Octanediol (*S,S*-3.4): Ketal **6.6** (2.43 g, 13 mmol) was heated in a mixture of aqueous HCl (6 M, 10 mL) and MeOH (10 mL) at 60 °C overnight. The reaction mixture was poured into aqueous K_2CO_3 and extracted with EtOAc. The combined organic layers were washed with brine and dried over anhydrous Na_2SO_4. The crude product was purified by vacuum flash chromatography (hexane/EtOAc = 2/1, then EtOAc) to give ***S,S*-3.4** as a colorless liquid (1.67 g, 87%). The ^{1}H NMR spectrum matched that previously reported.[7,8]

Synthesis of (2*R*,3*R*)-2,3-Octanediol (*R,R*-3.4): This compound was prepared from (L)-threonine using the same series of reactions as described above for synthesis of its enantiomer, diol ***S,S*-3.4**.

Enantiopurity Determination:

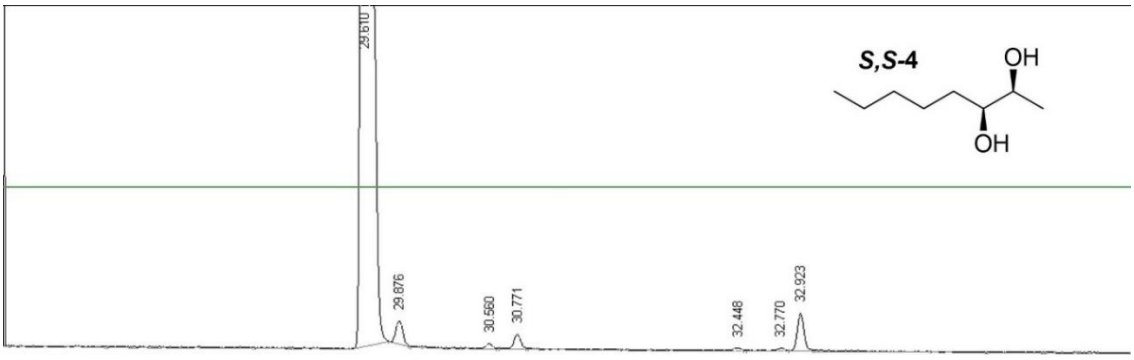


Figure 6.24. GC trace of ***S,S*-3.4** on a chiral stationary phase Cyclodex B column. ***S,S*-3.4** eluted at 29.61 min with an e.e. of 98.5%.

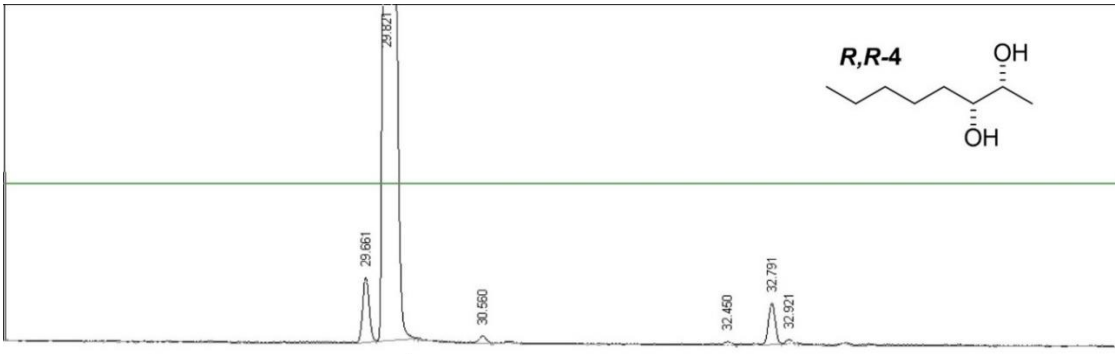


Figure 6.25. GC trace of ***R,R*-3.4** on a chiral stationary phase Cyclodex B column. ***R,R*-3.4** eluted at 29.88 min with an e.e. of 95.9%.

NMR Titrations
AMI (3.8) + 2-heptanol (3.2):

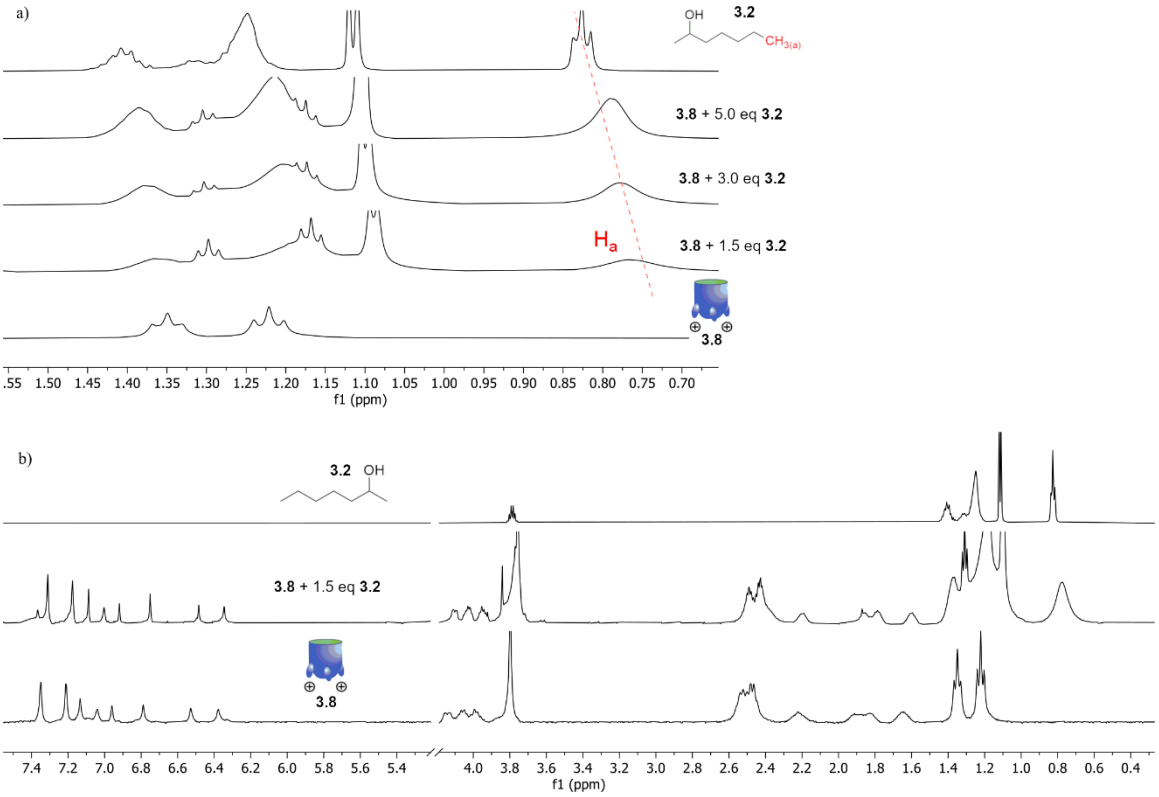


Figure 6.26. ^{1}H NMR spectra (D_2O, 400 MHz, 298K) showing rapid in and out exchange of guest **3.2** with host **3.8** where a) an upfield shift of the methyl group of **3.2** can be seen in the aliphatic region with addition of **3.2** to **3.8**; b) full spectra for the addition of **3.2** to **3.8**.

CHI (3.7) + 2-heptanol (3.2):

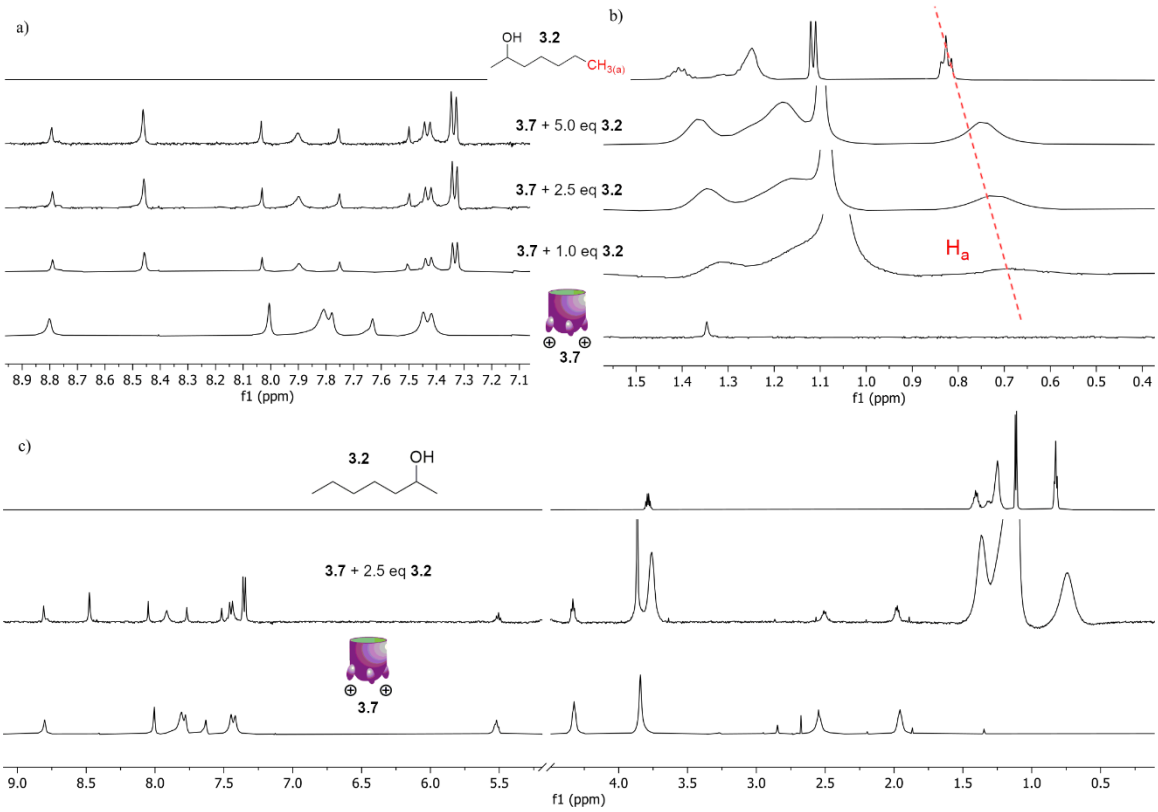


Figure 6.27. ^{1}H NMR spectra (D_2O, 400 MHz, 298K) showing rapid in and out exchange of guest **3.2** with host **3.7** where a) the aromatic peaks of the host **3.7** become sharper with addition of **3.2** and b) an upfield shift of the methyl group of **3.2** can be seen in the aliphatic region; c) full spectra for the addition of **3.2** to **3.7**.

AMI (3.8) + octanediol (3.4):

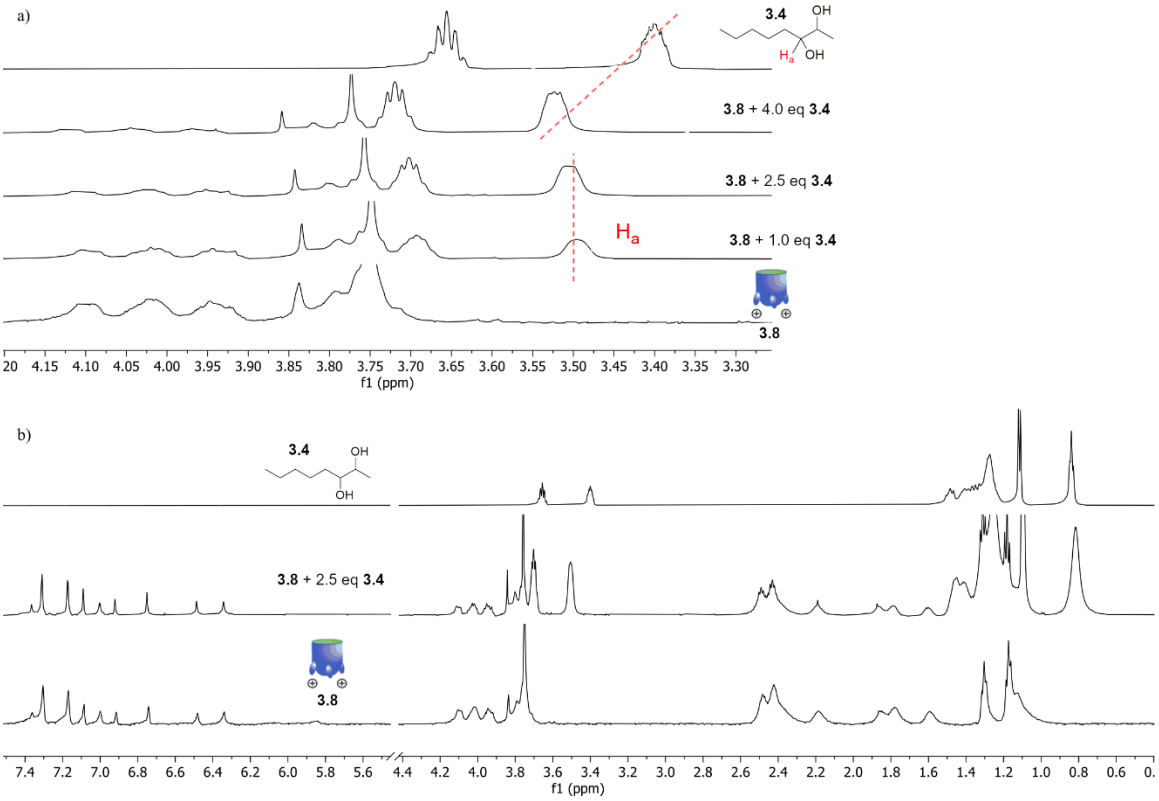


Figure 6.28. ^{1}H NMR spectra (D_2O, 400 MHz, 298K) showing rapid in and out exchange of guest **3.4** with host **3.8** where a) an upfield shift of the methine proton of **3.4** can be seen in the aliphatic region with addition of **3.4** to **3.8**; b) full spectra for the addition of **3.4** to **3.8**.

CHI (3.7) + octanediol (3.4):

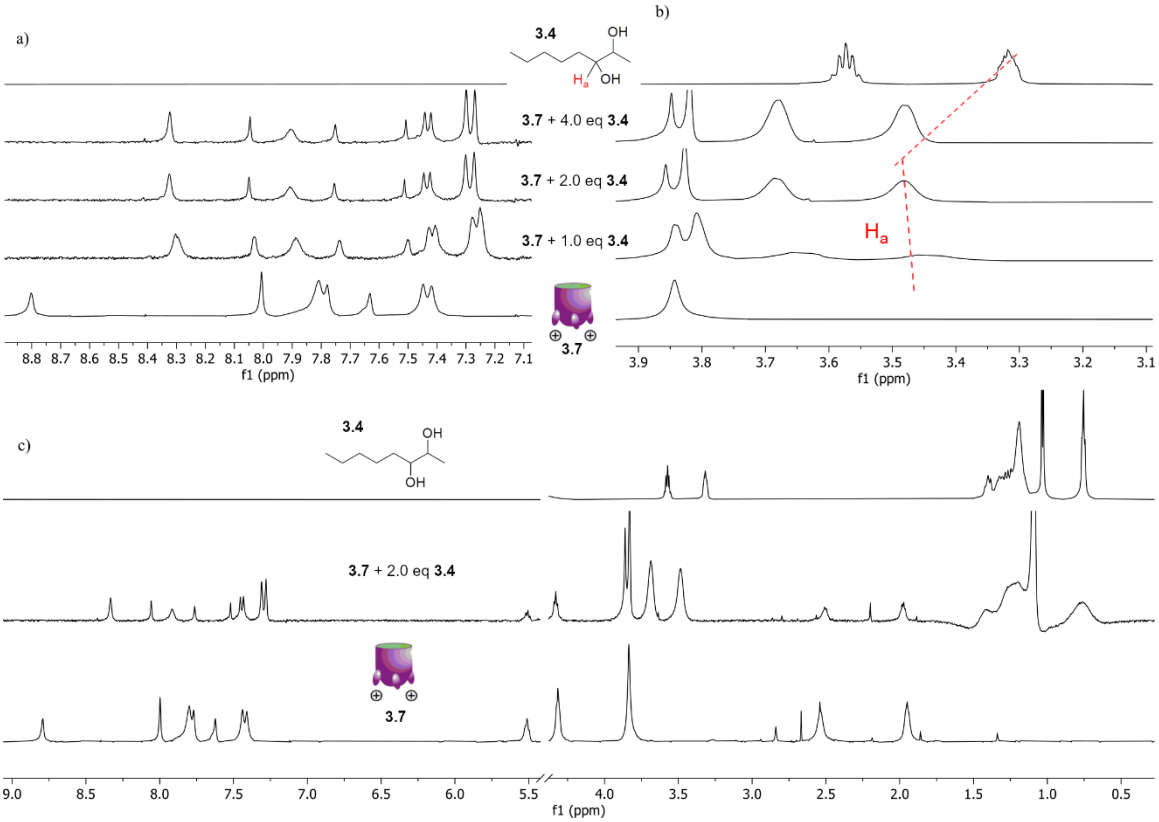
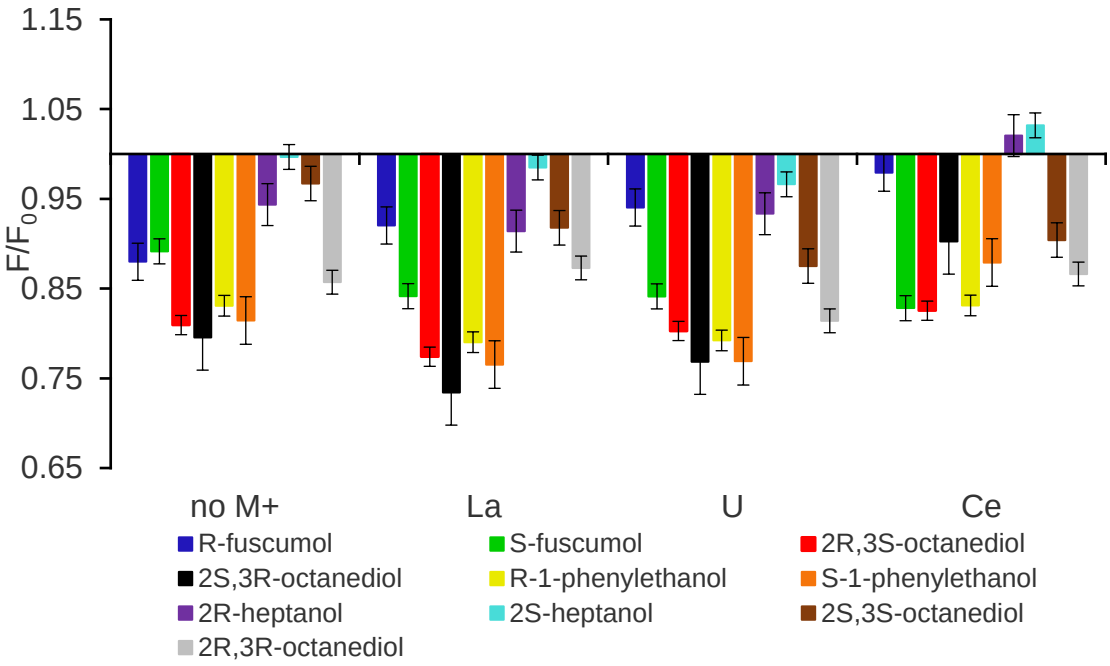
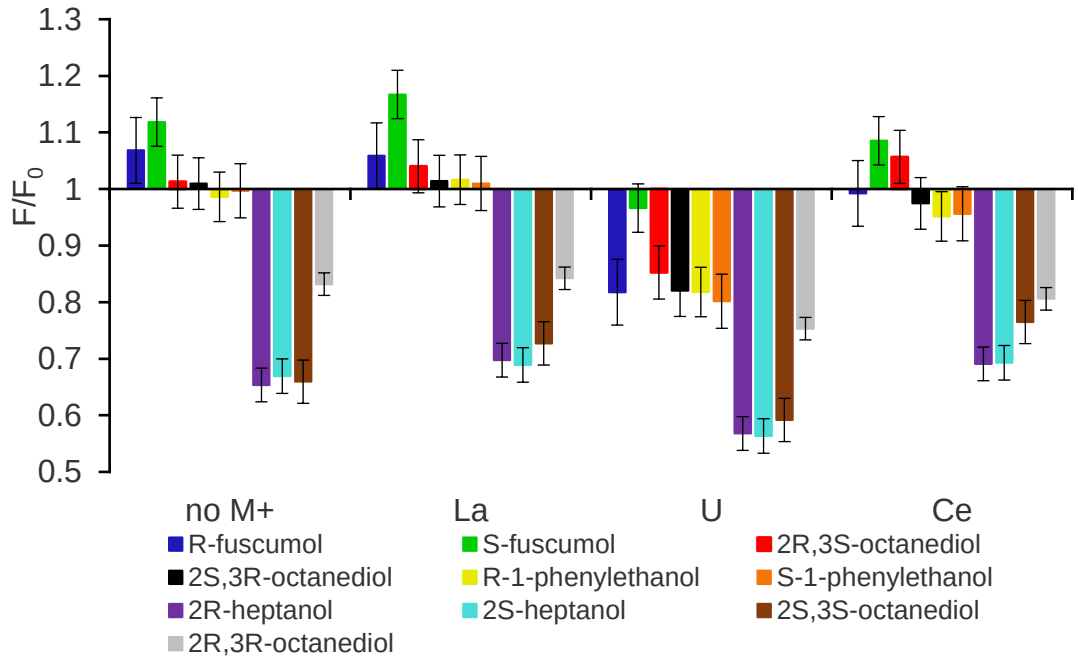


Figure 6.29. ^{1}H NMR spectra (D_2O, 400 MHz, 298K) showing rapid in and out exchange of guest **3.4** with host **3.7** where a) the aromatic peaks of the host **3.7** become sharper with addition of **3.4** and b) an upfield shift of the methine proton of **3.4** can be seen in the aliphatic region; c) full spectra for the addition of **3.4** to **3.7**.

Target additions to TCC (3.6) •DSMI (3.9) •Metal Combinations:



Target additions to AMI (3.8) •DSMI (3.9) •Metal Combinations:



Target additions to CHI (3.7) •DSMI (3.9) •Metal Combinations:

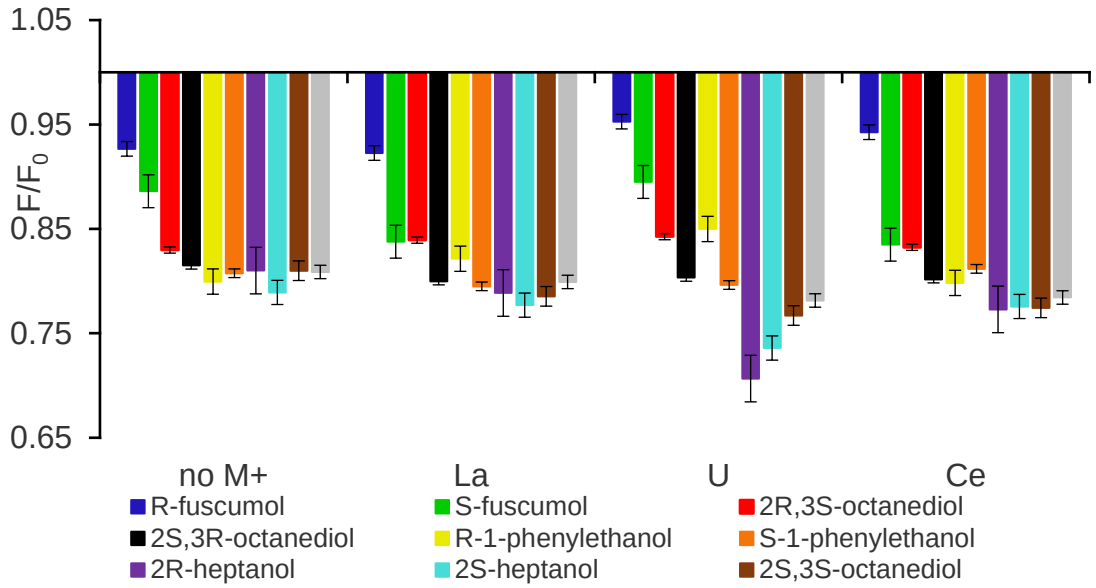
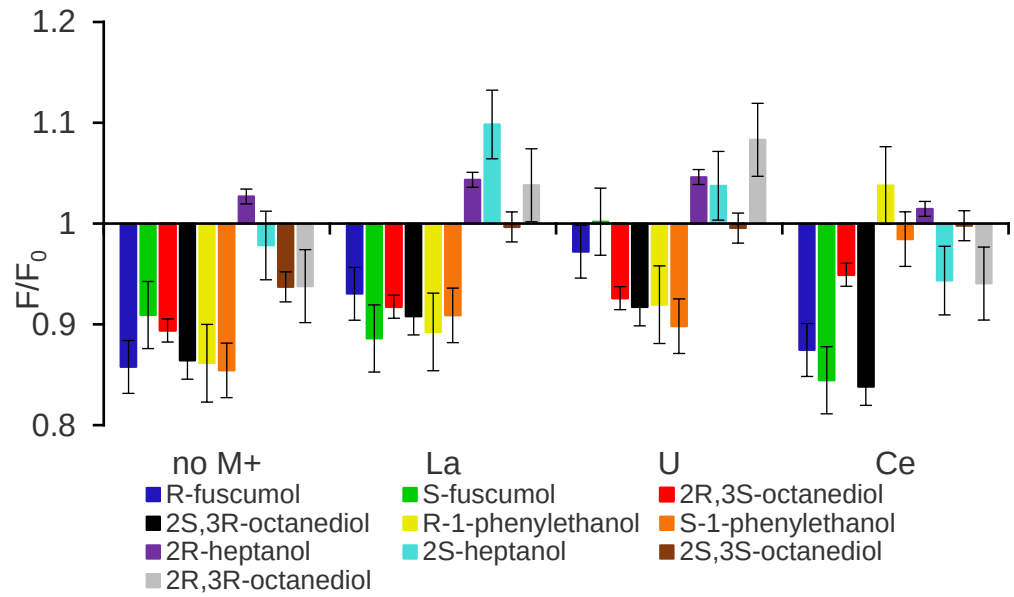
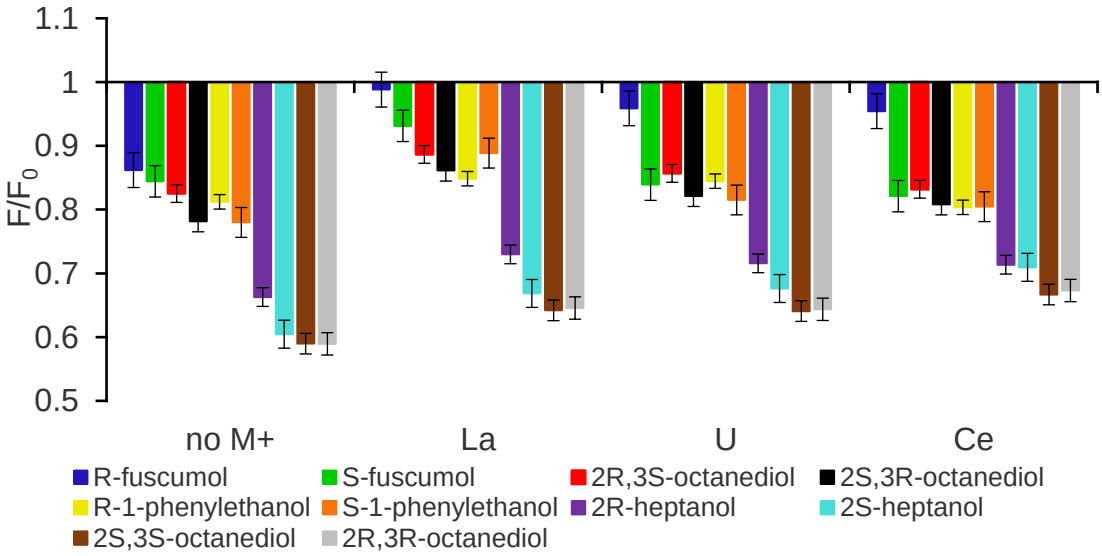


Figure 6.30. Relative fluorescence responses of the **Host•DSMI•M$^+$•Pheromone** complex in 20 mM Tris buffer, pH 7.4. [**Host**] = 20 μM, [**DSMI**] = 3.0 μM, [Metal] = 50 μM, [Pheromone] = 50 μM. F_0 = fluorescence response of the **Host•DSMI•M$^+$** complex, F = fluorescence response of the **Host•DSMI•M$^+$•Pheromone** complex.

Target additions to TCC (3.6) •SMITE (3.10) •Metal Combinations:



Target additions to AMI (3.8) •SMITE (3.10) •Metal Combinations:



Target additions to CHI (3.7)• SMITE (3.10) •Metal Combinations:

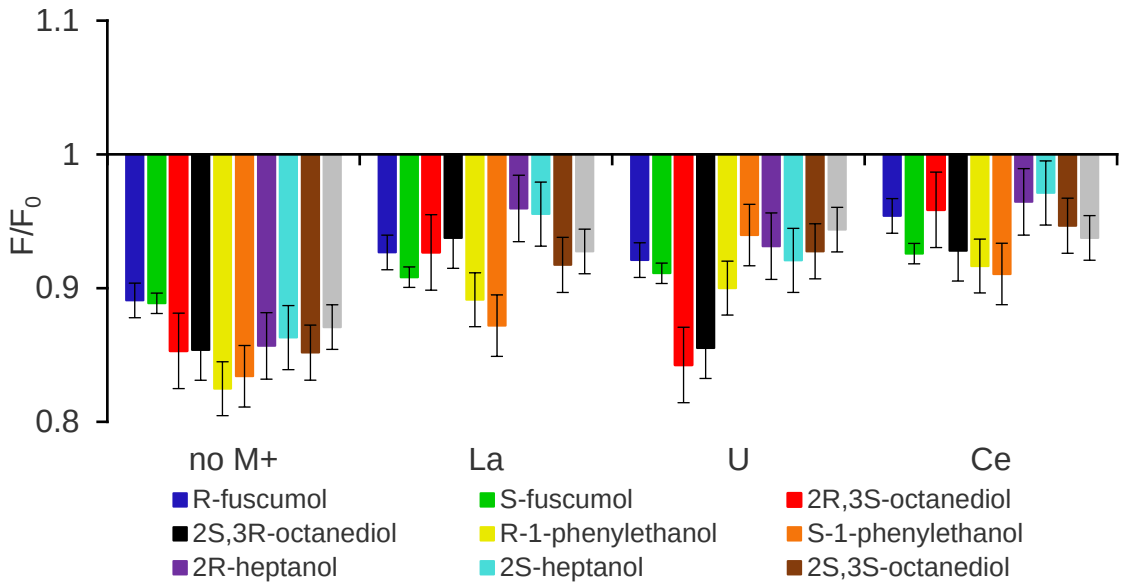


Figure 6.31. Relative fluorescence responses of the **Host•SMITE•M$^+$•Pheromone** complex in 20 mM Tris buffer, pH 7.4. [**Host**] = 20 μM, [**SMITE**] = 3.0 μM, [Metal] = 50 μM, [Pheromone] = 50 μM. F_0 = fluorescence response of the **Host•SMITE•M$^+$** complex, F = fluorescence response of the **Host•SMITE•M$^+$•Pheromone** complex.

Machine Learning Output Tables

Component	Rank	Component	Rank
TCC*DSMI*noM^{+}	5	CHI*SMITE*noM^{+}	1
TCC*DSMI*La^{3+}	10	CHI*SMITE*La^{3+}	2
TCC*DSMI*UO_2^{+}	6	CHI*SMITE*UO_2^{+}	11
TCC*DSMI*Ce^{3+}	12	CHI*SMITE*Ce^{3+}	8
TCC*SMITE*noM^{+}	13	AMI*DSMI*noM^{+}	1
TCC*SMITE*La^{3+}	19	AMI*DSMI*La^{3+}	1
TCC*SMITE*UO_2^{+}	14	AMI*DSMI*UO_2^{+}	17
TCC*SMITE*Ce^{3+}	18	AMI*DSMI*Ce^{3+}	1
CHI*DSMI*noM^{+}	9	AMI*SMITE*noM^{+}	4
CHI*DSMI*La^{3+}	15	AMI*SMITE*La^{3+}	3
CHI*DSMI*UO_2^{+}	16	AMI*SMITE*UO_2^{+}	7
CHI*DSMI*Ce^{3+}	1	AMI*SMITE*Ce^{3+}	1

Evaluation Metrics	**Score (standard deviation from 3 repeated runnings of the 4-fold cross validation)**
Accuracy	1.0000 (0.0000)
Sensitivity	1.0000 (0.0000)
Specificity	1.0000 (0.0000)
Precision	1.0000 (0.0000)
F1 Score	1.0000 (0.0000)
AUC	1.0000 (0.0000)

Table 6.2. Tables showing the SVM-RFE ranking of all 24-components used for statistical analysis with the top six components highlighted (top); Performance metrics of 3 repeated 4-fold cross validation with SVM as the estimator by using the 6 best features selected by SVM-RFE (bottom).

PCA Scores Plot using the Full 24-component Combined Array

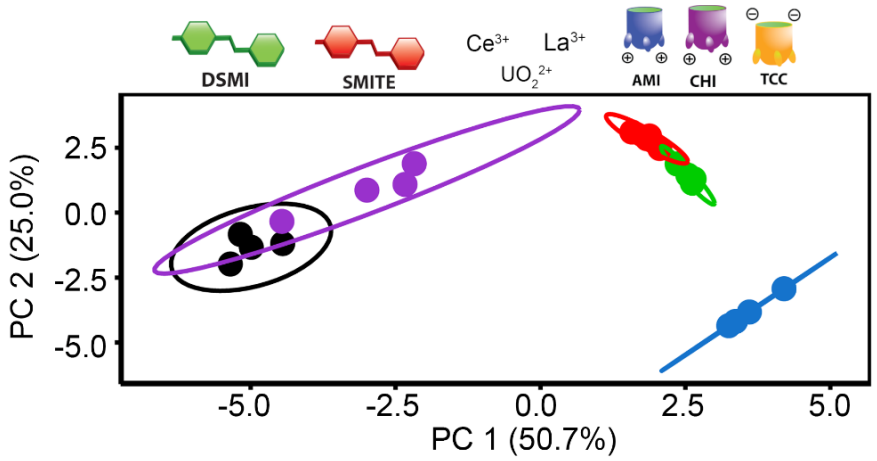


Figure 6.32. PCA scores plots with 95% confidence intervals for the full 24-factor **Host•DSMI•M**$^{+}$ and **Host•SMITE•M**$^{+}$ arrays in Tris buffer (obtained from statistical analysis of data in Figures S-10 and S-11). [**Host**] = 20 μM, [**DSMI**] and [**SMITE**] = 3 μM, [Metal] = 50 μM, [Pheromone] = 50 μM, [Tris] = 20 mM (pH 7.4).

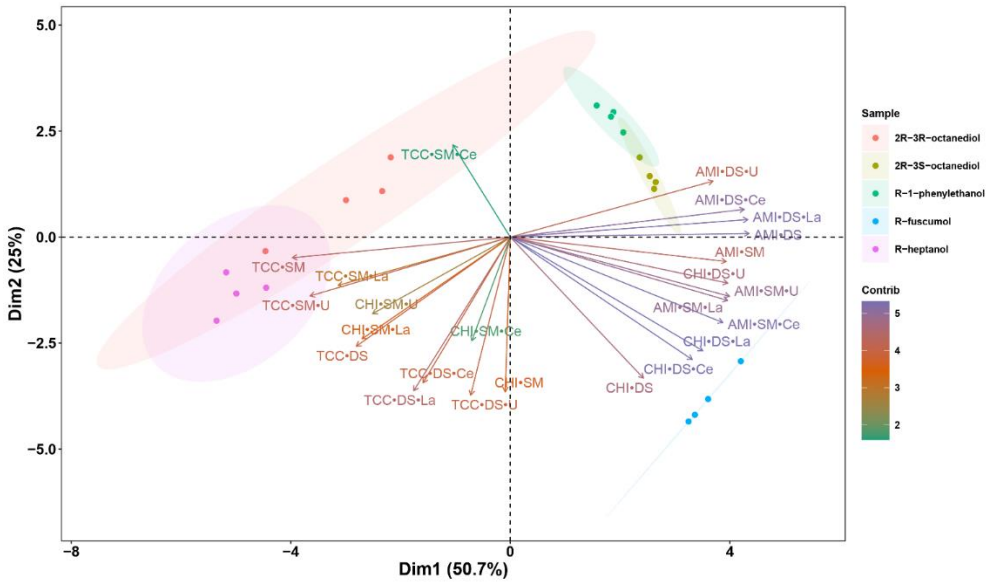


Figure 6.33. PCA biplot (combining both PCA scores plot and loading plot) using the full 24-component array system from Figure 6.32 with prcomp (x,center = TRUE, scale. = TRUE) as the PCA function. Loadings are gradient colored according to the contribution of each variable. Ellipses indicate 95% confidence intervals.

PCA Scores Plot using the Optimized Arrays

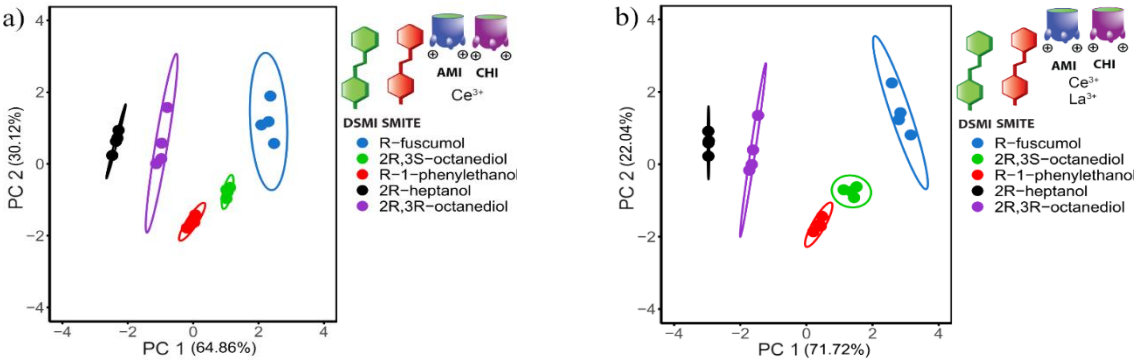


Figure 6.34. PCA scores plots with 95% confidence intervals for an optimized a) 4-factor array with **Host•DSMI** or **Host•SMITE** and either no metal or Ce^{3+} in Tris buffer and b) a 6-factor array with **Host•DSMI** or **Host•SMITE** and either no metal, La^{3+}, or Ce^{3+} (obtained from statistical analysis of selected data from Figures S-12 and S-13). [**Host**] = 20 μM, [**DSMI**] and [**SMITE**]= 3 μM, [Metal] = 50 μM, [Pheromone] = 50 μM, [Tris] = 20 mM (pH 7.4).

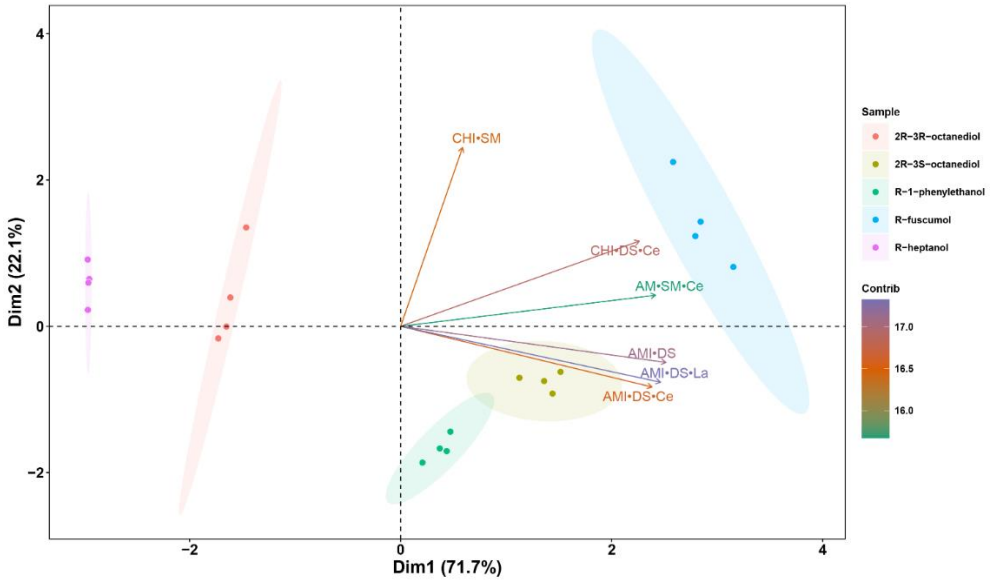


Figure 6.35. PCA biplot (combining both PCA scores plot and loading plot) using the 6-factor optimized array system from Figure 6.34b with princomp(x,cor=TRUE, scores = TRUE) as the PCA function. Loadings are gradient colored according to the contribution of each variable. Ellipses indicate 95% confidence intervals.

PCA Scores Plot using the Unoptimized Arrays

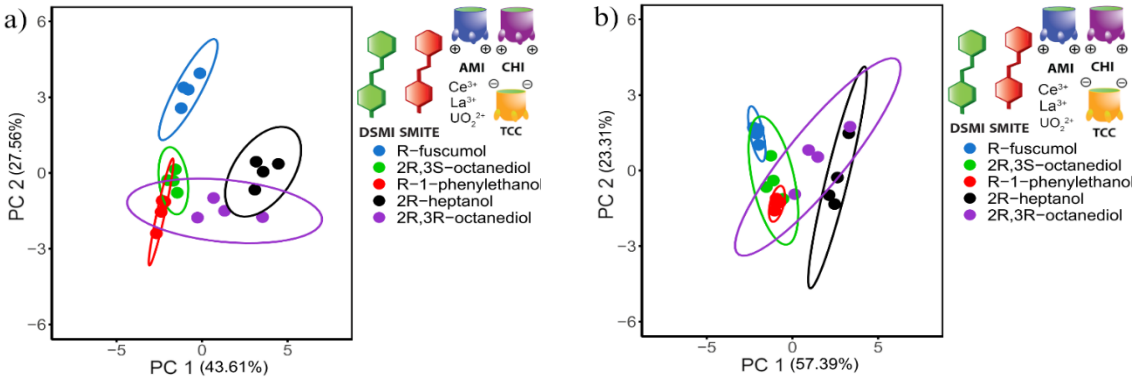


Figure 6.36. PCA scores plots with 95% confidence intervals for unoptimized arrays with **Host•DSMI** or **Host•SMITE** and either La^{3+}, UO_2^{2+}, or Ce^{3+} in Tris buffer using a) a 12-factor array and b) a 6-factor array (obtained from statistical analysis of selected data from Figures S-12 and S-13). [**Host**] = 20 μM, [**DSMI**] and [**SMITE**]= 3 μM, [Metal] = 50 μM, [Pheromone] = 50 μM, [Tris] = 20 mM (pH 7.4).

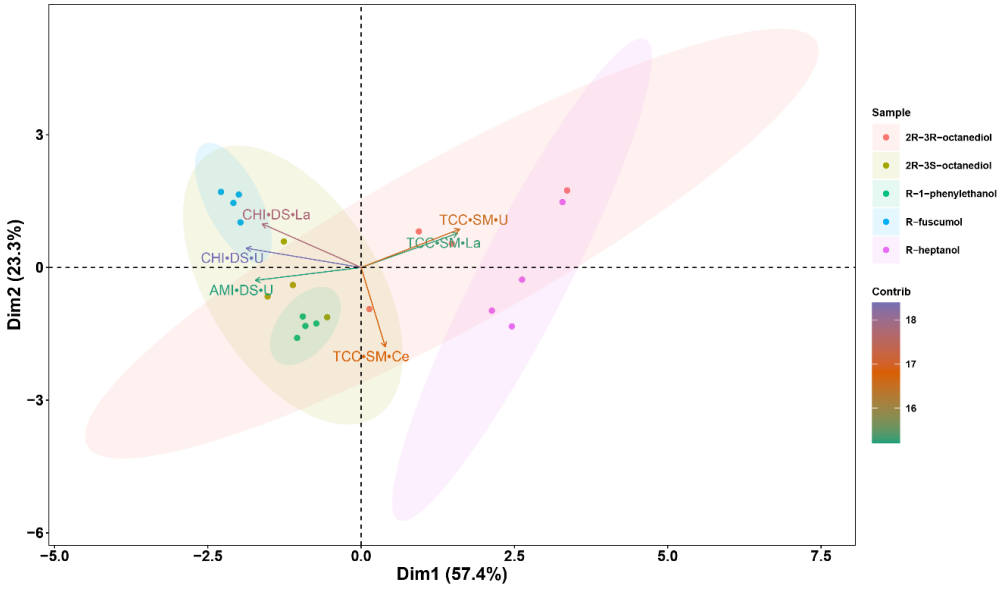
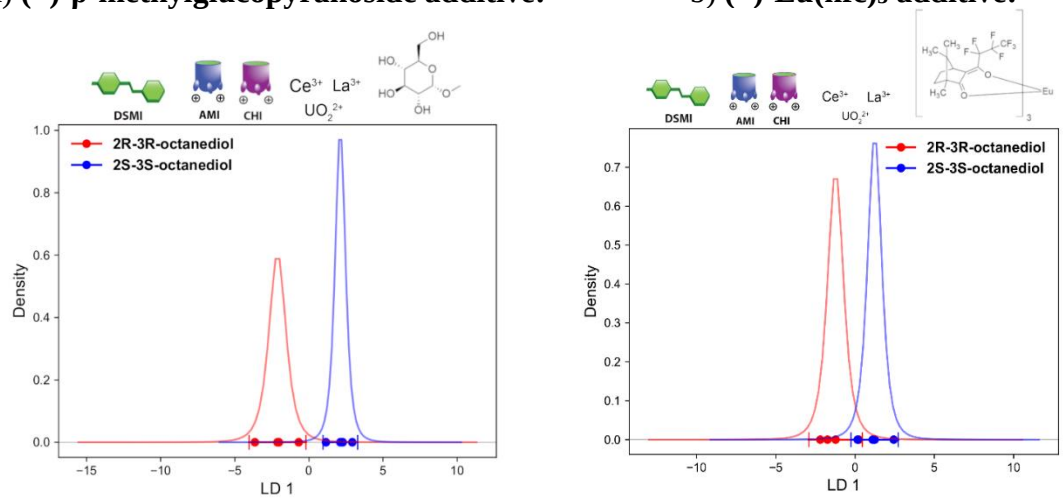


Figure 6.37. PCA biplot (combining both PCA scores plot and loading plot) using the 6-factor unoptimized array system from Figure 6.36b with princomp(x,cor=TRUE, scores = TRUE) as the PCA function. Loadings are gradient colored according to the contribution of each variable. Ellipses indicate 95% confidence intervals.

1D LDA Plots using other Chiral Additives:

a) **(+)-β-methylglucopyranoside additive:** b) **(+)-Eu(hfc)$_3$ additive:**



c) **L-(+)-tartaric acid additive:**

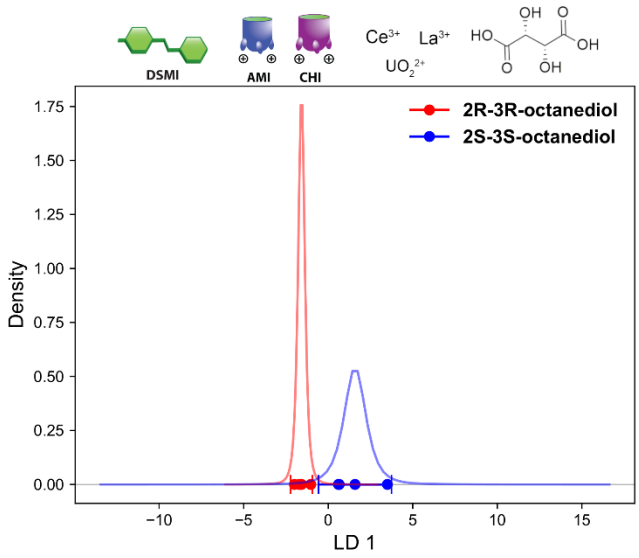


Figure 6.38. 1D LDA (x-axis=LD 1) plots for the 6-factor **Host•DSMI•Additive** sensor arrays in Tris buffer with a) chiral additive (+)-β-methylglucopyranoside, b) chiral additive Eu(hfc)$_3$, and c) chiral additive L-(+)tartaric acid were each separately tested. [**Host**] = 20 μM, [**DSMI**] = 3 μM, [Metal] = 50 μM, [Pheromone] = 50 μM, [Additive] = 50 μM, [Tris] = 20 mM (pH 7.4). Red/blue dots = datapoints, curve = probability density of student's t-distribution, vertical markers = 95% confidence intervals.

Cross Validation Scores of Pheromone Enantiomer Classification

Sample	R/S-2-heptanol	2R,3R-/2S,3S-octanediol	R/S-1-phenylethanol	R/S-fuscumol
Accuracy	1.0000 (0.0000)	0.8750 (0.2905)	0.9125 (0.1900)	0.9875 (0.0781)
Sensitivity	1.0000 (0.0000)	0.8750 (0.2905)	0.9125 (0.1900)	0.9875 (0.0781)
Specificity	1.0000 (0.0000)	0.8750 (0.2905)	0.9125 (0.1900)	0.9875 (0.0781)
Precision	1.0000 (0.0000)	0.8500 (0.3298)	0.8688 (0.2850)	0.9812 (0.1171)
F1 Score	1.0000 (0.0000)	0.8583 (0.3152)	0.8833 (0.2533)	0.9833 (0.1041)
AUC	1.0000 (0.0000)	1.0000 (0.0000)	1.0000 (0.0000)	1.0000 (0.0000)

Table 6.3. Performance metrics of pheromone enantiomers classification in **Figure 6.38** calculated by 10 repeated 4-fold cross validation with LDA as the estimator.

Discriminant Analysis Plots with Non-cavitand Control Arrays:

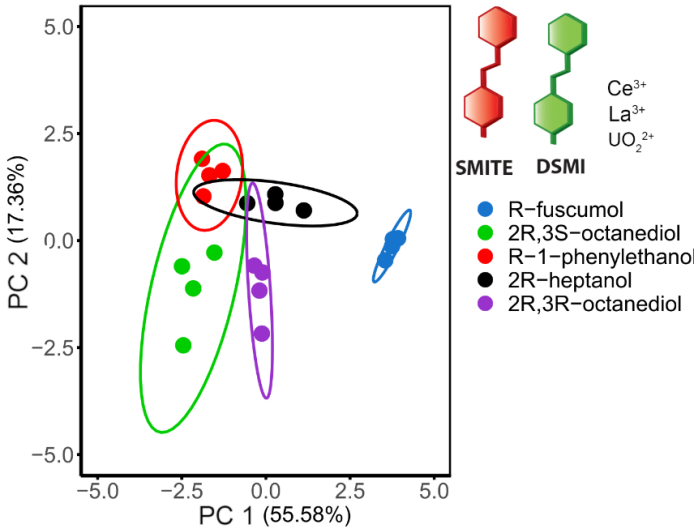
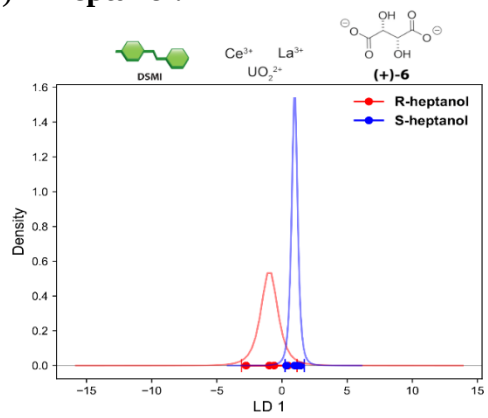


Figure 6.39. PCA scores plots with 95% confidence intervals for the control **DSMI•M⁺•Pheromone** and **SMITE•M⁺•Pheromone** sensor array in Tris buffer. [**DSMI**] and [**SMITE**] = 3 μM, [Metal] = 50 μM, [Pheromone] = 50 μM, [Tris] = 20 mM (pH 7.4).

a) **2-Heptanol:**



b) **2,3-octanediol:**

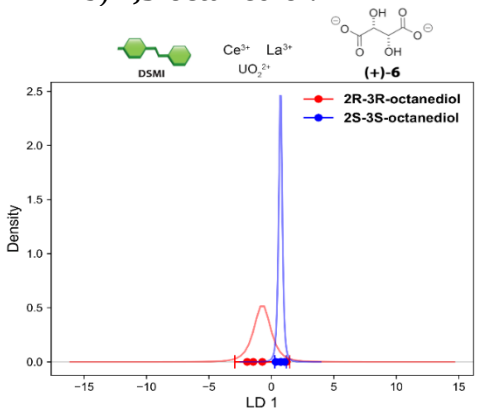


Figure 6.40. 1D LDA (x-axis=LD 1) plots for the control **DSMI•2•Additive** and **DSMI•4•Additive** sensor array in Tris buffer**.** [**DSMI**] = 3 mM, [Metal] = 50 mM, [Pheromone]= 50 mM, [Additive] = 50 μM, [Tris] = 20 mM (pH 7.4). Red/blue dots = datapoints, curve = probability density of student's t-distribution, vertical markers = 95% confidence intervals.

6.8. Selected Spectra for Chapter 4

Screening of Host•Dye•Anion Combinations in Tris

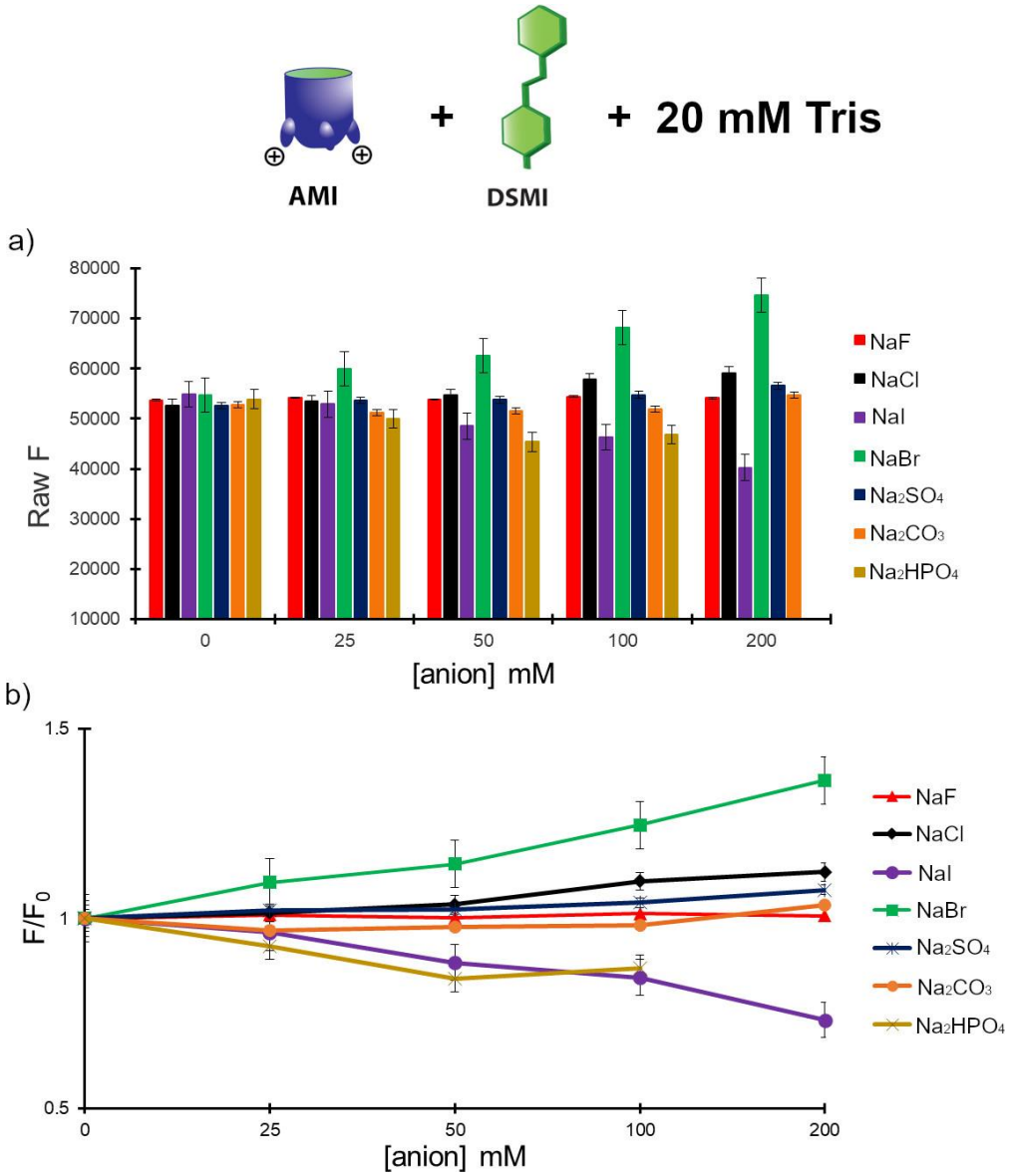


Figure 6.41. Relative fluorescence responses of the **AMI•DSMI•X**$^-$ complex in 20 mM Tris buffer, pH 7.4. [**AMI**] = 5 μM, [**DSMI**] = 5 μM. F_0 = fluorescence response of the **AMI•DSMI** complex, F = fluorescence response of the **AMI•DSMI•X**$^-$ complex, where a) is the raw fluorescence of **AMI•DSMI•X** and b) is the F/F_0 response of **AMI•DSMI•X**$^-$.

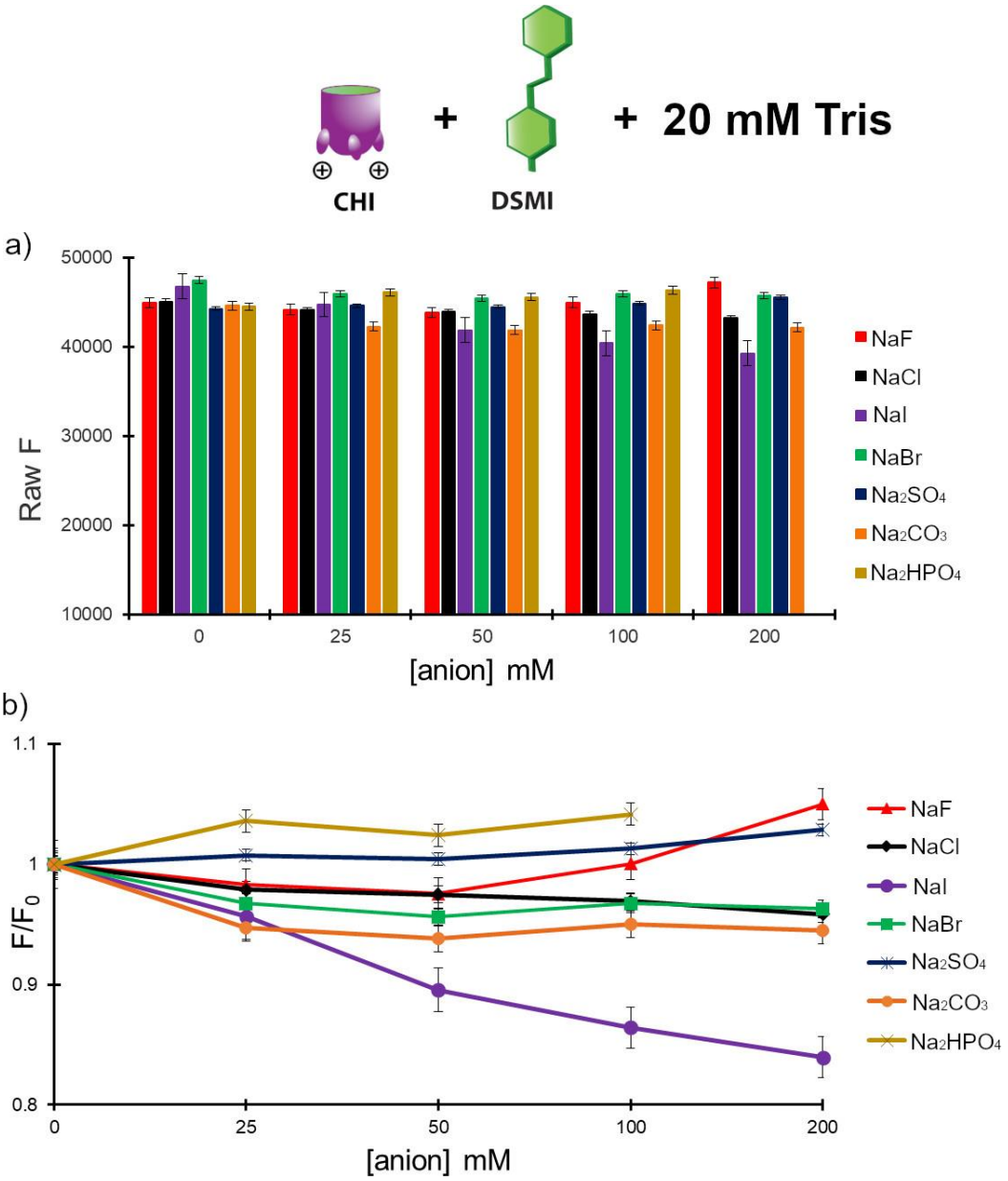


Figure 6.42. Relative fluorescence responses of the **CHI•DSMI•X**$^-$ complex in 20 mM Tris buffer, pH 7.4. [**CHI**] = 5 μM, [**DSMI**] = 5 μM. F_0 = fluorescence response of the **CHI•DSMI** complex, F = fluorescence response of the **CHI•DSMI•X**$^-$ complex where, a) is the raw fluorescence of **CHI•DSMI•X**$^-$and b) is the F/F_0 response of **CHI•DSMI•X**$^-$.

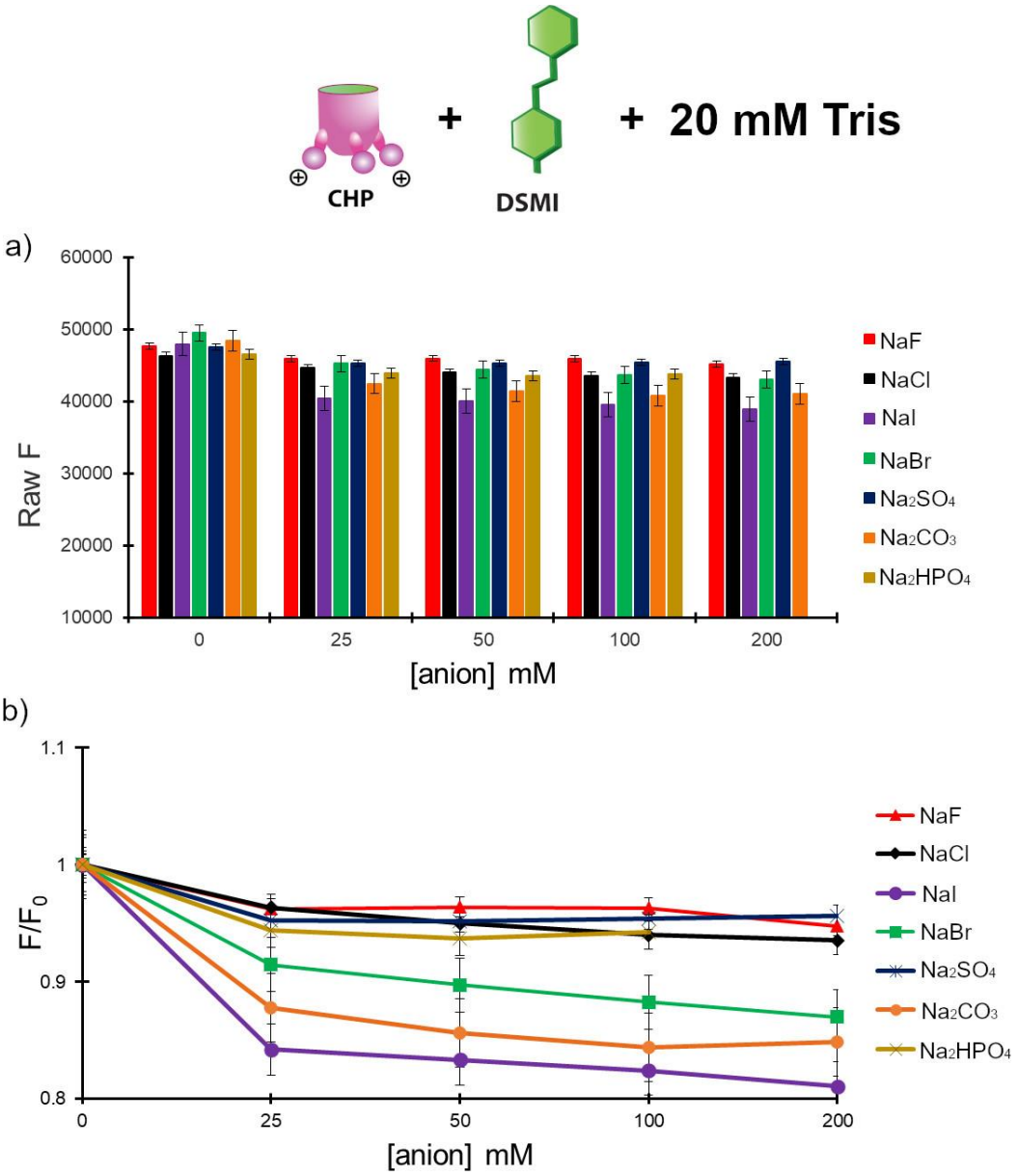


Figure 6.43. Relative fluorescence responses of the **CHP•DSMI•X**$^-$ complex in 20 mM Tris buffer, pH 7.4. [**CHP**] = 5 μM, [**DSMI**] = 5 μM. F_0 = fluorescence response of the **CHP•DSMI** complex, F = fluorescence response of the **CHP•DSMI•X**$^-$ complex where, a) is the raw fluorescence of **CHP•DSMI•X**$^-$and b) is the F/F_0 response of **CHP•DSMI•X**$^-$.

Selected Host•Dye•Anion Combinations in Tris Buffer

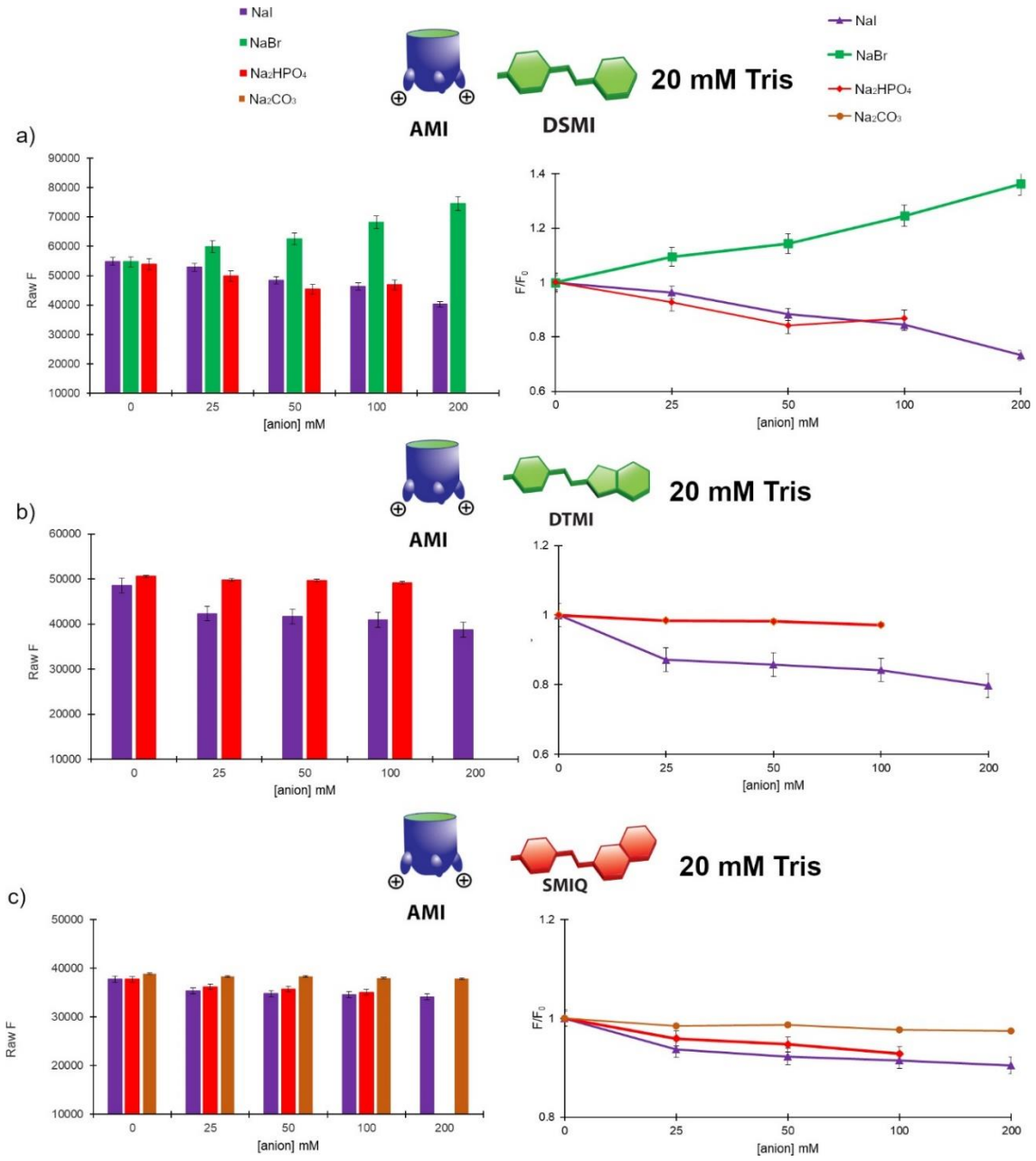


Figure 6.44. Relative fluorescence responses of the **AMI•Dye•X**$^-$ complex in 20mM Tris buffer, pH 7.4, [**AMI**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **AMI•Dye** complex, F = fluorescence response of the **AMI•Dye•X**$^-$ complex where a) is the response of **AMI•DSMI•X**$^-$, b) is the response of **AMI•DTMI•X**$^-$, and c) is the response of **AMI•SMIQ•X**$^-$.

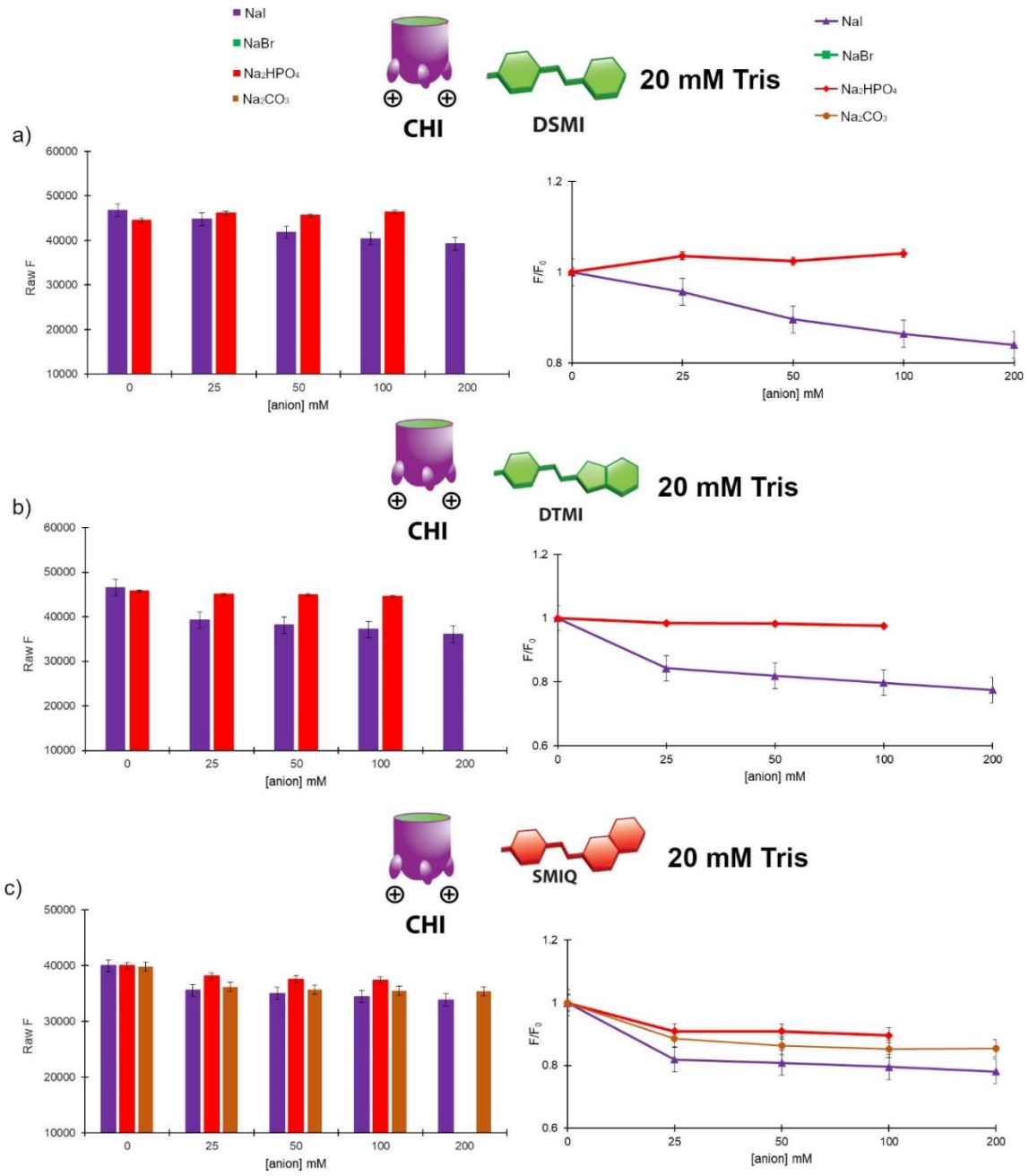


Figure 6.45. Relative fluorescence responses of the **CHI•Dye•X**$^-$ complex in 20mM Tris buffer, pH 7.4, [**CHI**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **CHI•Dye** complex, F = fluorescence response of the **CHI•Dye•X**$^-$ complex where a) is the response of **CHI•DSMI•X**$^-$, b) is the response of **CHI•DTMI•X**$^-$, and c) is the response of **CHI•SMIQ•X**$^-$.

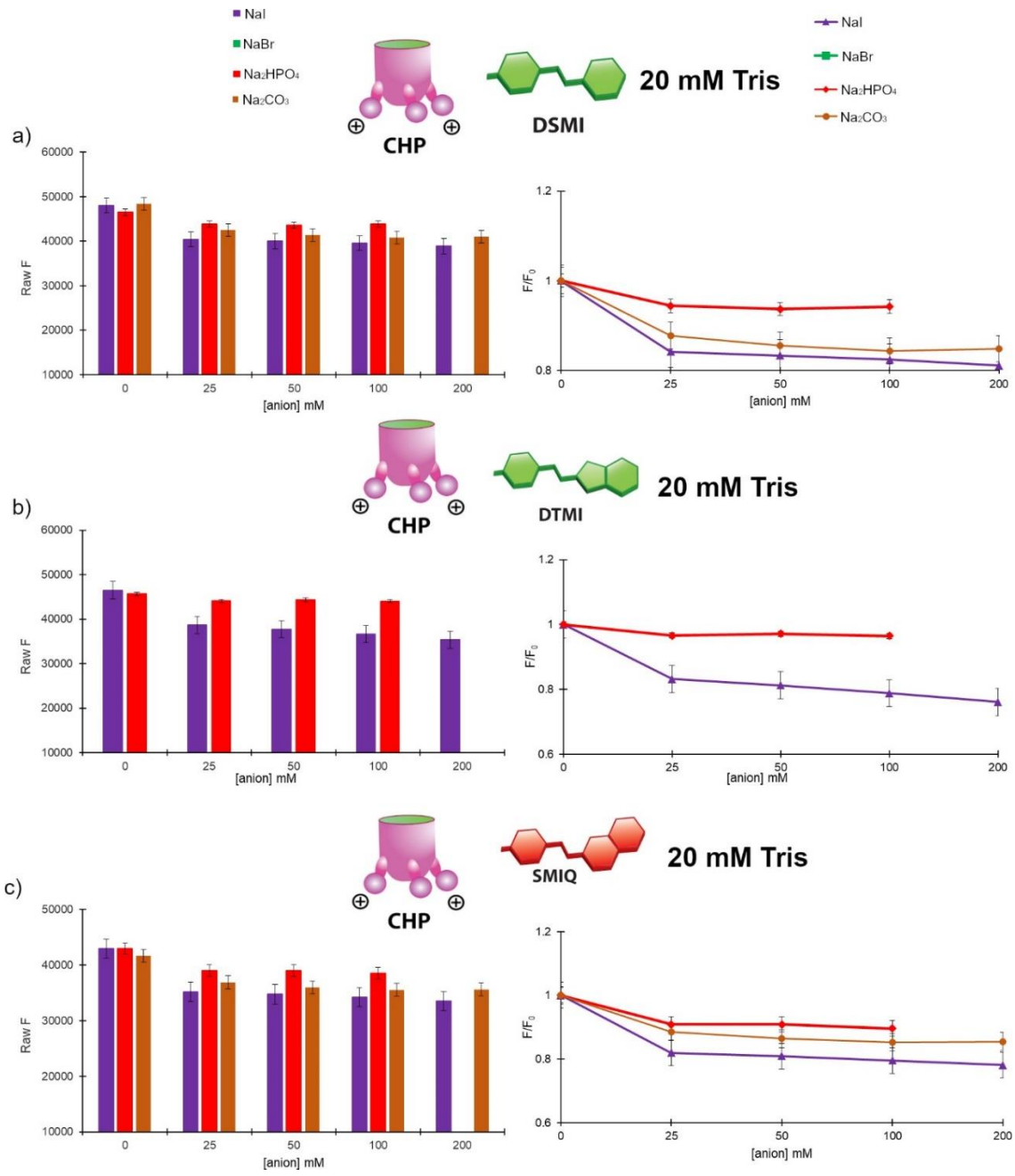


Figure 6.46. Relative fluorescence responses of the **CHP•Dye•X**$^-$ complex in 20mM Tris buffer, pH 7.4, [**CHP**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **CHP•Dye** complex, F = fluorescence response of the **CHP•Dye•X**$^-$ complex where a) is the response of **CHP•DSMI•X**$^-$, b) is the response of **CHP•DTMI•X**$^-$, and c) is the response of **CHP•SMIQ•X**$^-$.

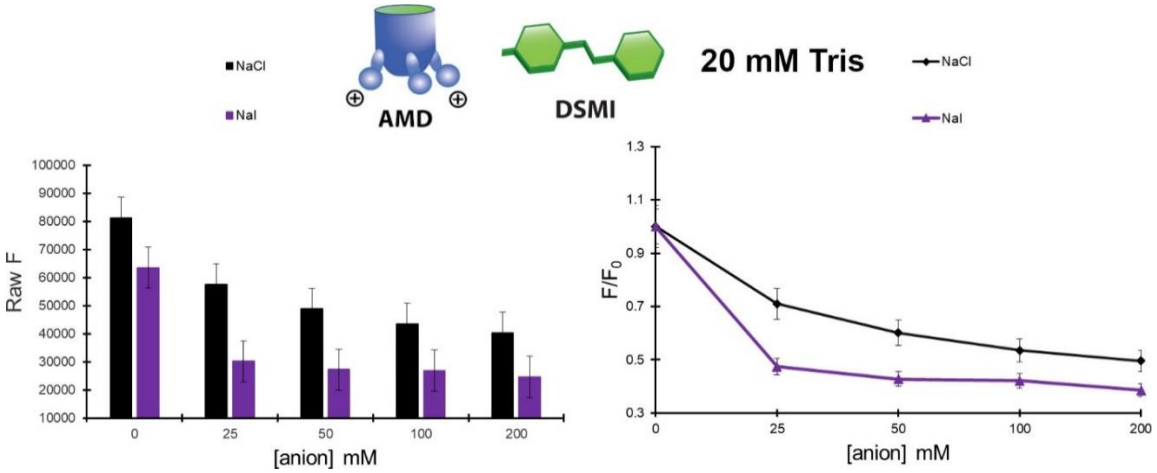


Figure 6.47. Relative fluorescence responses of the **AMD•DSMI•X**$^{-}$ complex in 20mM Tris buffer, pH 7.4, [**AMD**] = 5 μM, [**DSMI**] = 5 μM. F_0 = fluorescence response of the **AMD•DSMI** complex, F = fluorescence response of the **AMI•DSMI•X**$^{-}$ complex.

Selected Host•Dye•Anion Combinations in Ultrapure H_2O

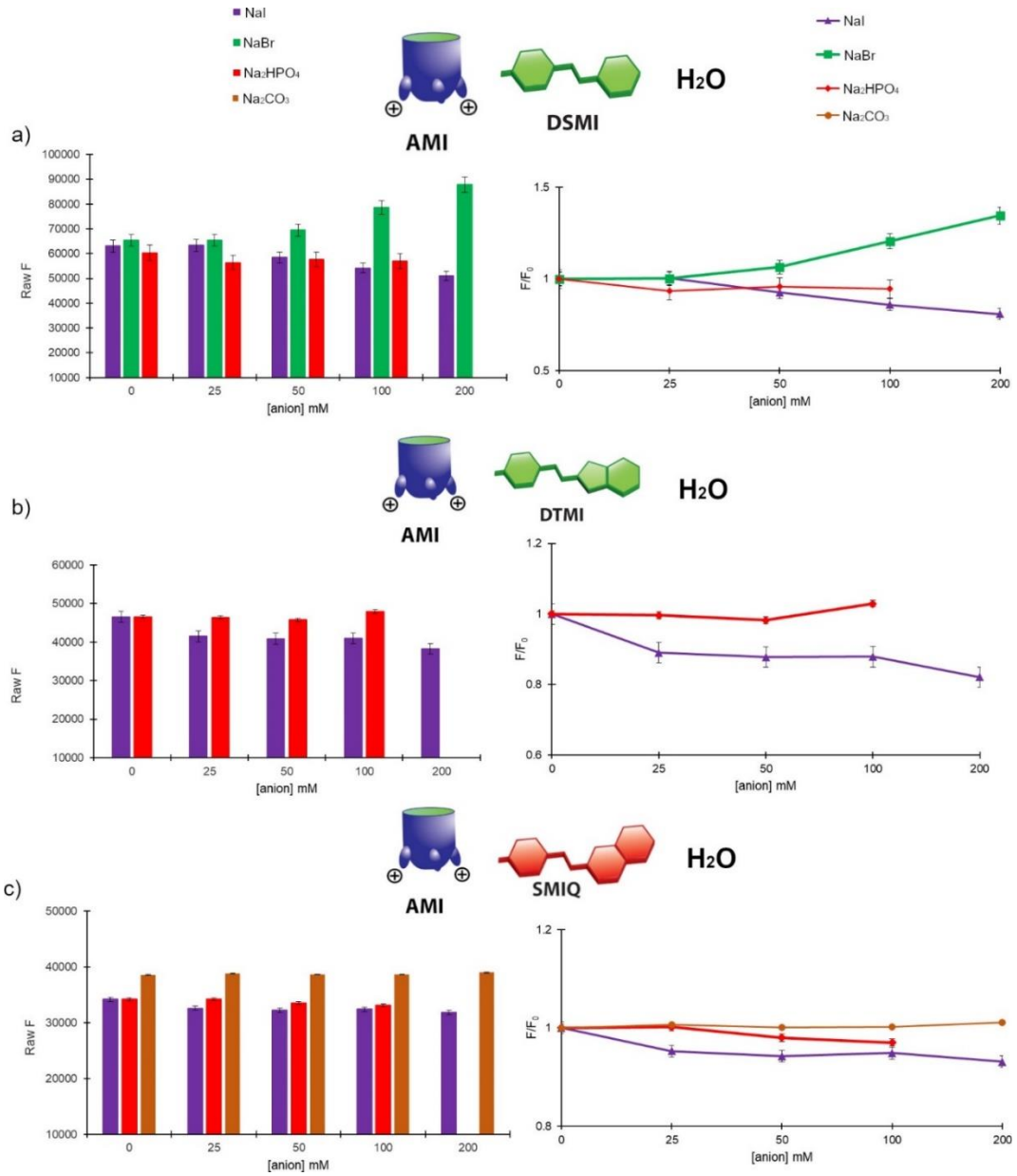


Figure 6.48. Relative fluorescence responses of the **AMI•Dye•X**$^-$ complex in ultrapure H_2O, pH 7.4, [**AMI**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **AMI•Dye** complex, F = fluorescence response of the **AMI•Dye•X**$^-$ complex where a) is the response of **AMI•DSMI•X**$^-$, b) is the response of **AMI•DTMI•X**$^-$, and c) is the response of **AMI•SMIQ•X**$^-$.

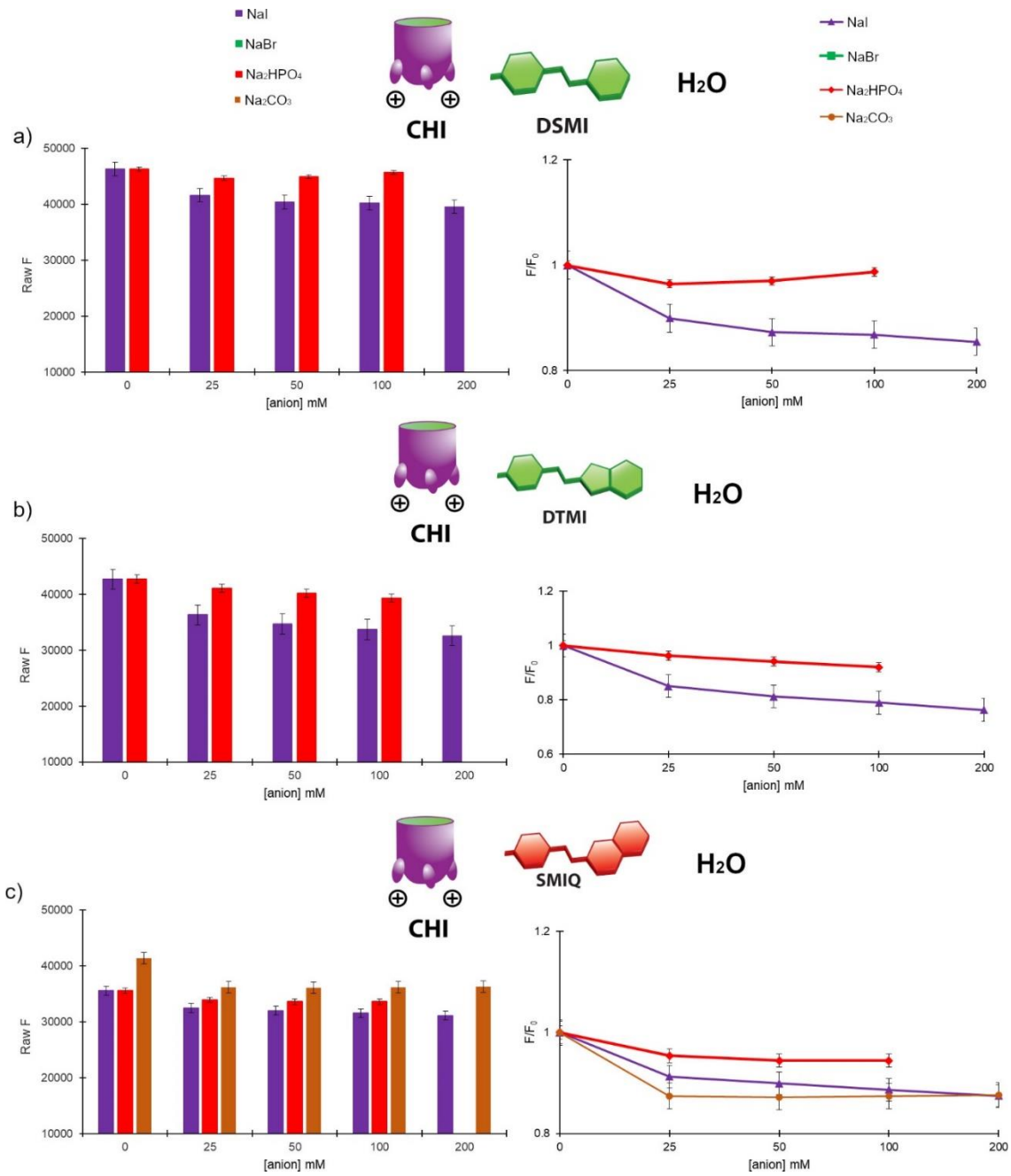


Figure 6.49. Relative fluorescence responses of the **CHI•Dye•X**$^-$ complex in ultrapure H_2O, pH 7.4, [**CHI**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **CHI•Dye** complex, F = fluorescence response of the **CHI•Dye•X**$^-$ complex where a) is the response of **CHI•DSMI•X**$^-$, b) is the response of **CHI•DTMI•X**$^-$, and c) is the response of **CHI•SMIQ•X**$^-$.

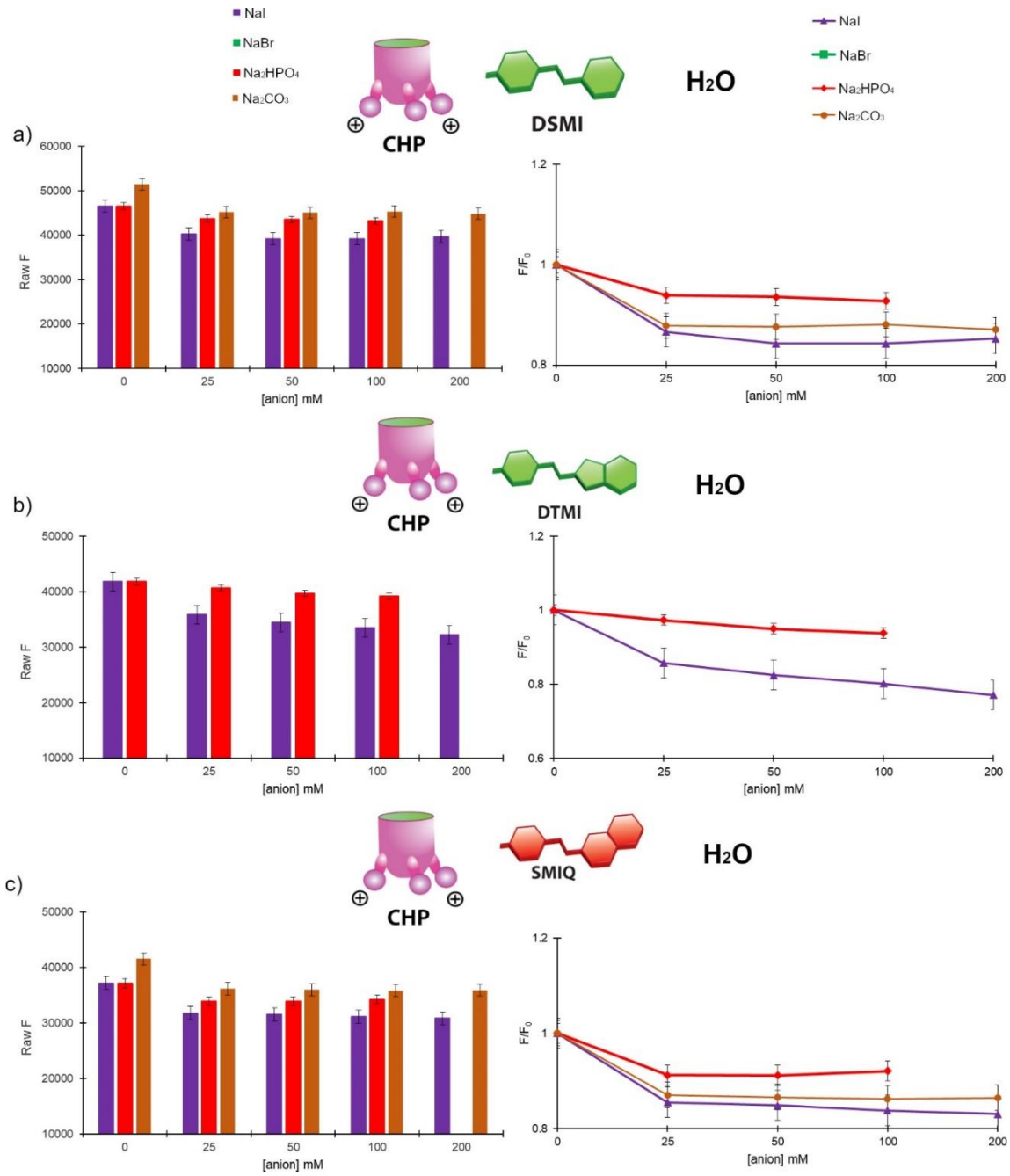


Figure 6.50. Relative fluorescence responses of the **CHP•Dye•X**$^-$ complex in ultrapure H_2O, pH 7.4, [**CHP**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **CHP•Dye** complex, F = fluorescence response of the **CHP•Dye•X**$^-$ complex where a) is the response of **CHP•DSMI•X**$^-$, b) is the response of **CHP•DTMI•X**$^-$, and c) is the response of **CHP•SMIQ•X**$^-$.

Selected Host•Dye•Anion Combinations in 10X PBS Buffer (1.37 M NaCl, 27 mM KCl, 80 mM NaH_2PO_4, and 20 mM KH_2PO_4)

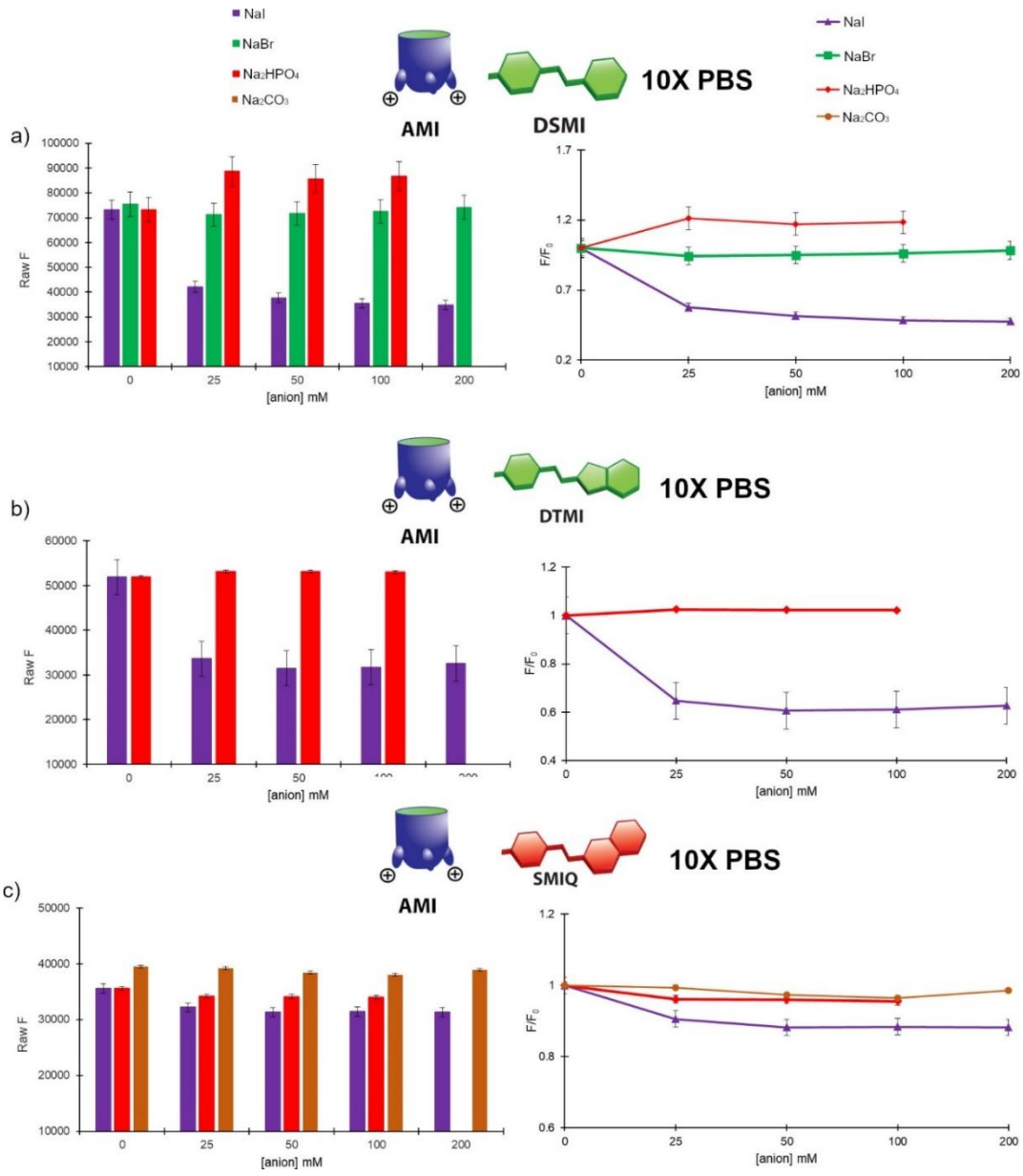


Figure 6.51. Relative fluorescence responses of the **AMI•Dye•X**$^-$ complex in 10X PBS, pH 7.4, [**AMI**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **AMI•Dye** complex, F = fluorescence response of the **AMI•Dye•X**$^-$ complex where a) is the response of **AMI•DSMI•X**$^-$, b) is the response of **AMI•DTMI•X**$^-$, and c) is the response of **AMI•SMIQ•X**$^-$.

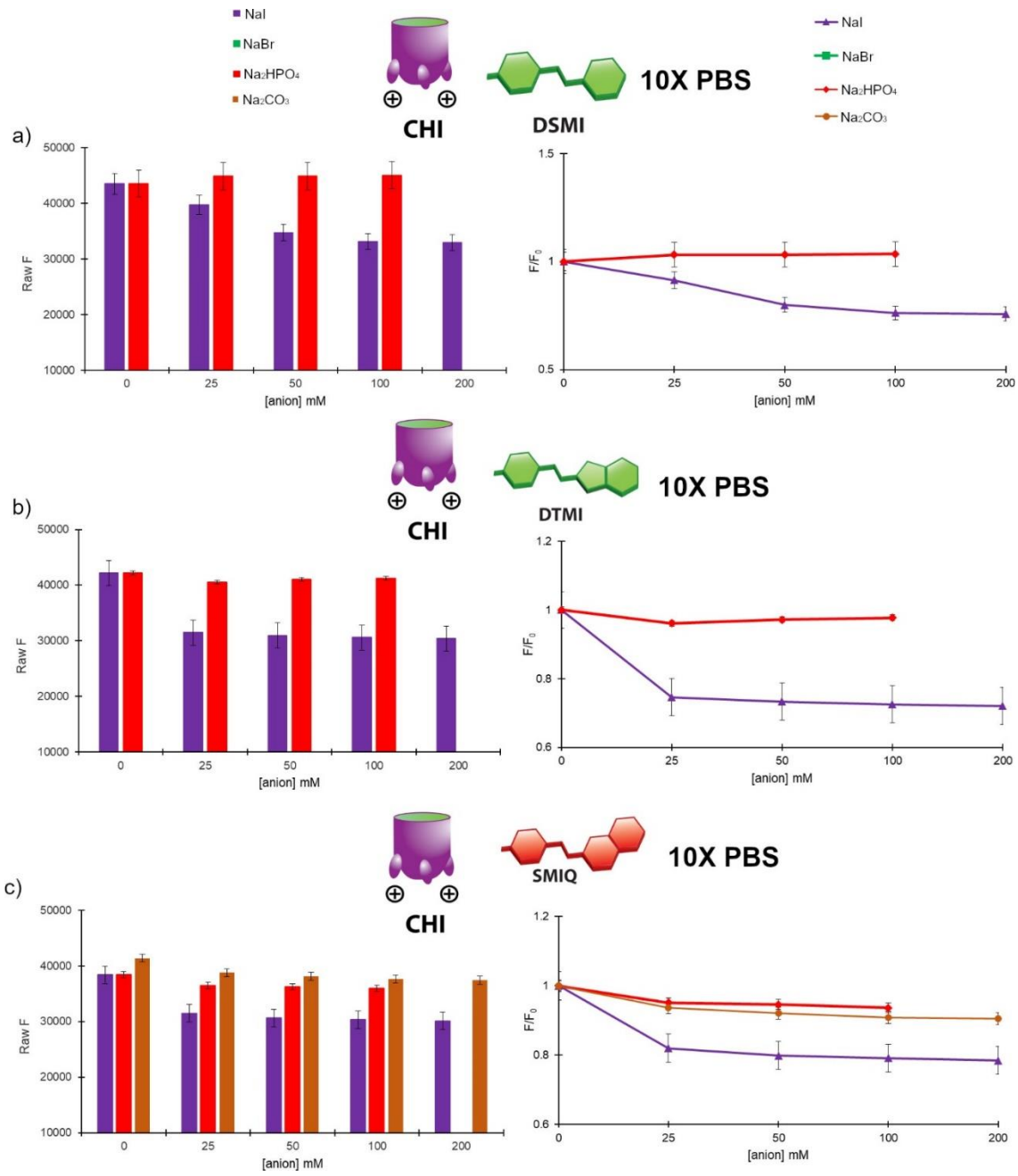


Figure 6.52. Relative fluorescence responses of the **CHI•Dye•X**$^-$ complex in 10X PBS, pH 7.4, [**CHI**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **CHI•Dye** complex, F = fluorescence response of the **CHI•Dye•X**$^-$ complex where a) is the response of **CHI•DSMI•X**$^-$, b) is the response of **CHI•DTMI•X**$^-$, and c) is the response of **CHI•SMIQ•X**$^-$.

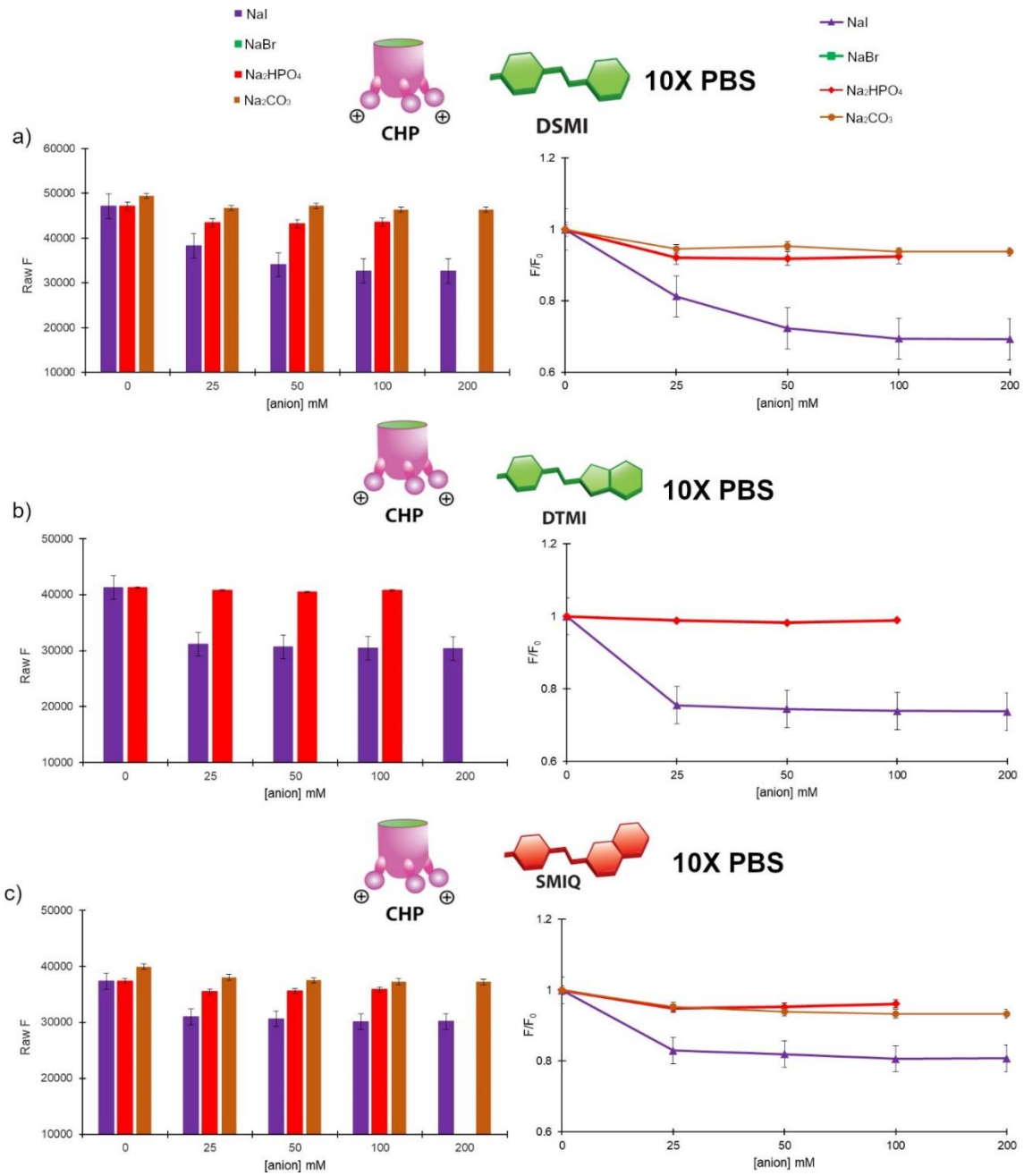


Figure 6.53. Relative fluorescence responses of the **CHP•Dye•X**$^-$ complex in 10X PBS, pH 7.4, [**CHP**] = 5 μM, [**Dye**] = 5 μM. F_0 = fluorescence response of the **CHP•Dye** complex, F = fluorescence response of the **CHP•Dye•X**$^-$ complex where a) is the response of **CHP•DSMI•X**$^-$, b) is the response of **CHP•DTMI•X**$^-$, and c) is the response of **CHP•SMIQ•X**$^-$.

Controls

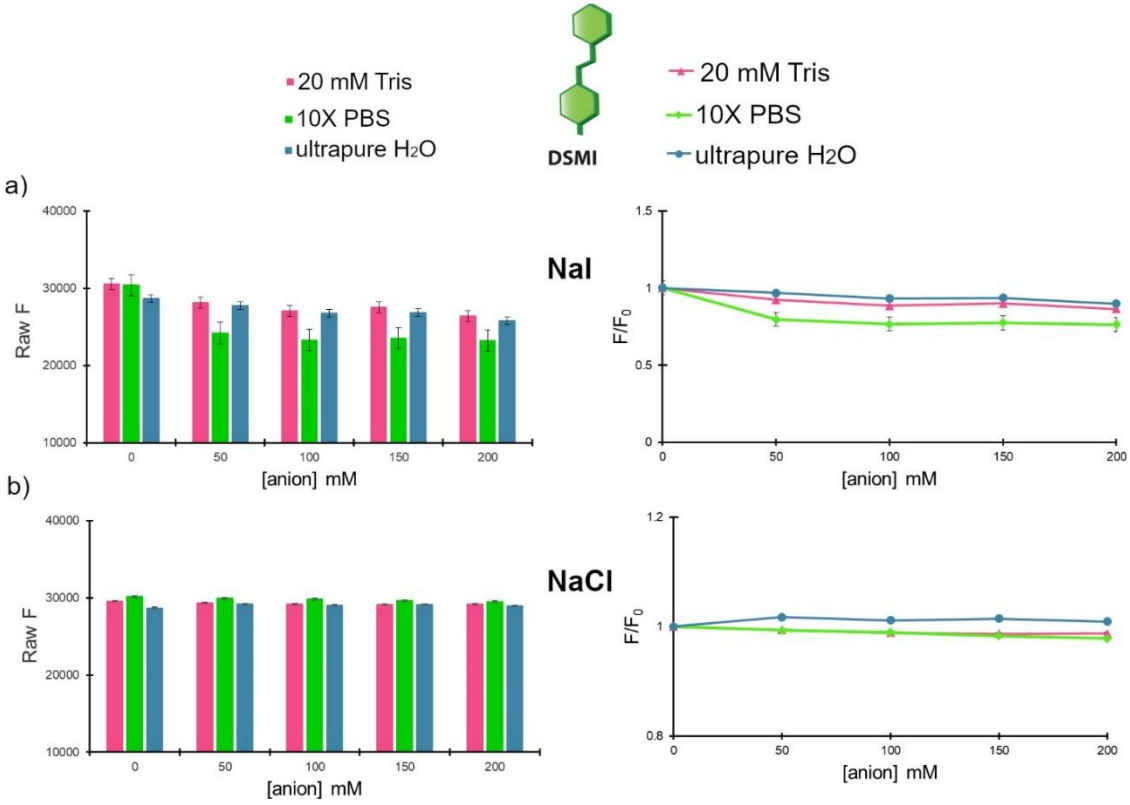


Figure 6.54. Relative fluorescence responses of the **Dye•X**$^-$ complex in 20 mM Tris, pH 7.4. [**Dye**] = 5 μM, F_0 = fluorescence response of the **Dye** only, F = fluorescence response of the **Dye•X**$^-$ complex where a) is the response of **DSMI•X**$^-$ where **X**$^-$ is NaI, and b) is the response of **DSMI•X**$^-$ where **X**$^-$ is NaCl.

NMR Analysis of AMI (4.3) •Anion Binding

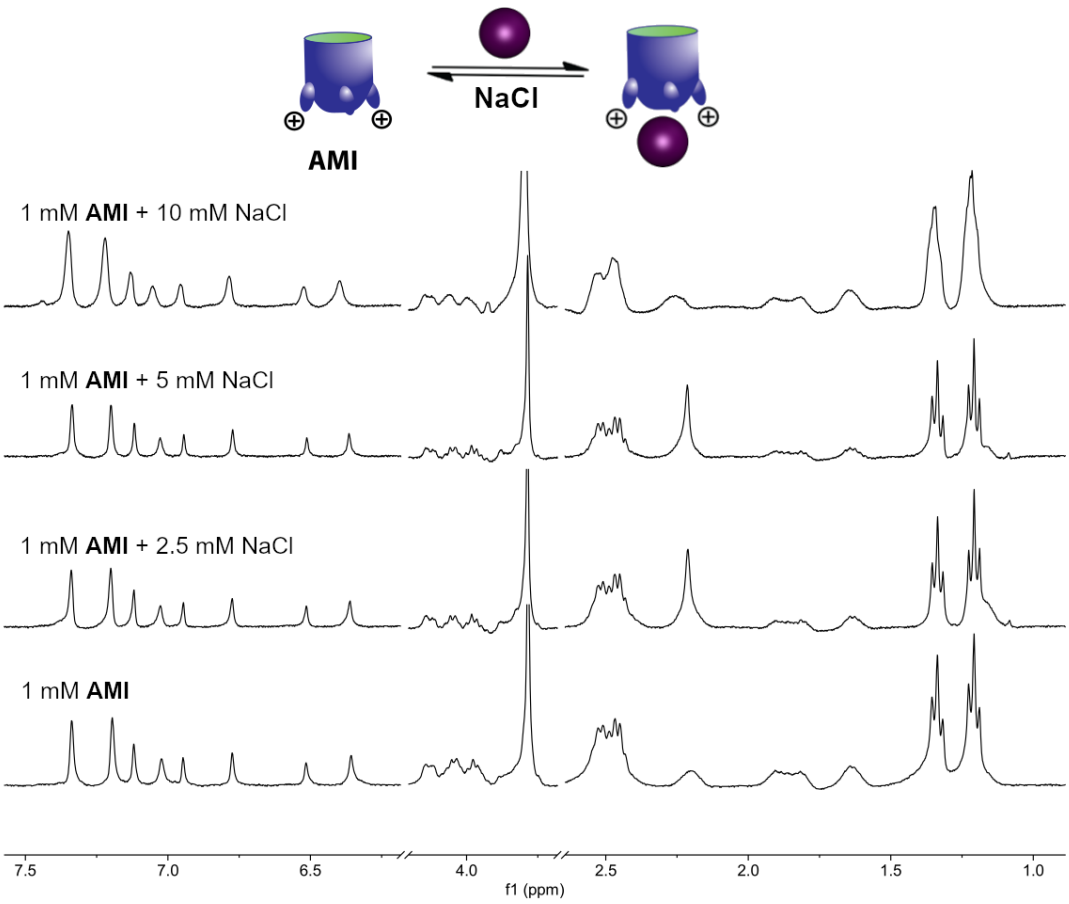


Figure 6.55. ^{1}H NMR spectra (D_2O, 600 MHz, 298 K) showing rapid in and out exchange of NaCl with **4.3**.

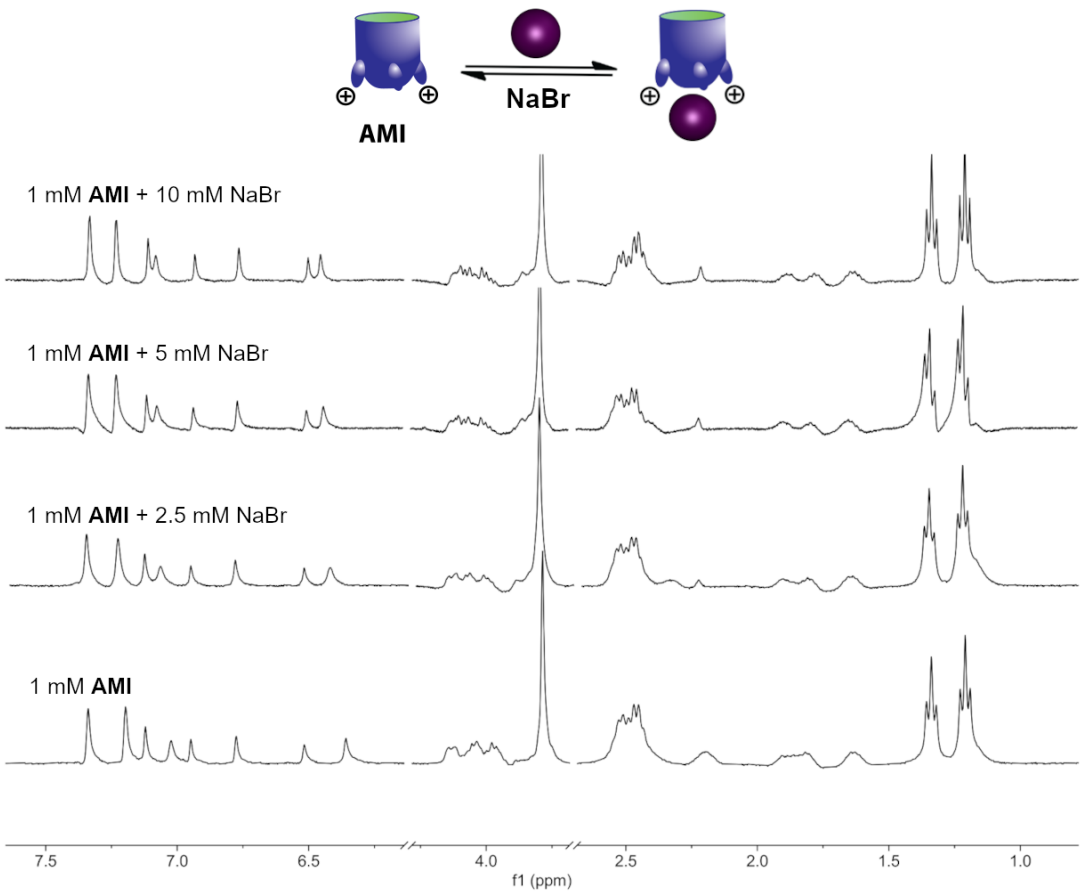


Figure 6.56. ^{1}H NMR spectra (D_2O, 600 MHz, 298 K) showing rapid in and out exchange of NaBr with **4.3**.

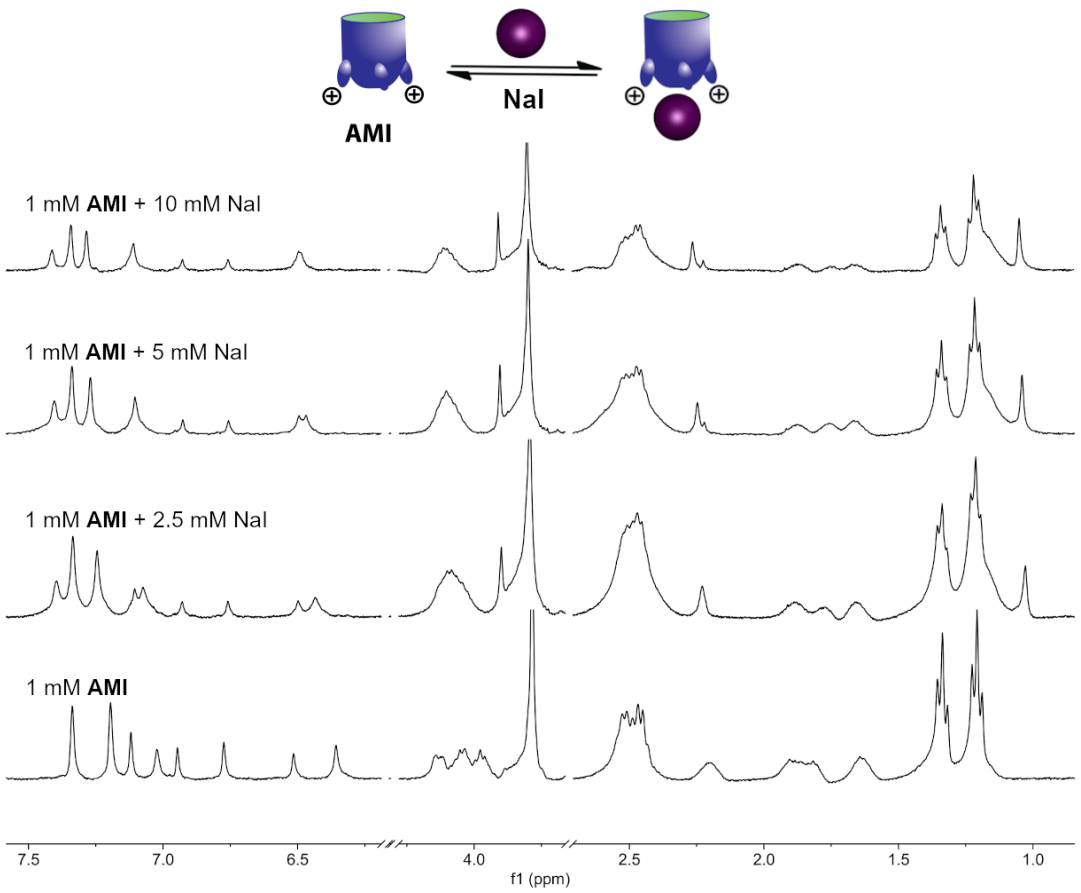


Figure 6.57. ^{1}H NMR spectra (D_2O, 600 MHz, 298 K) showing slow in and out exchange of NaI with **4.3**.

NMR Analysis of AMD (4.4) •Anion Binding

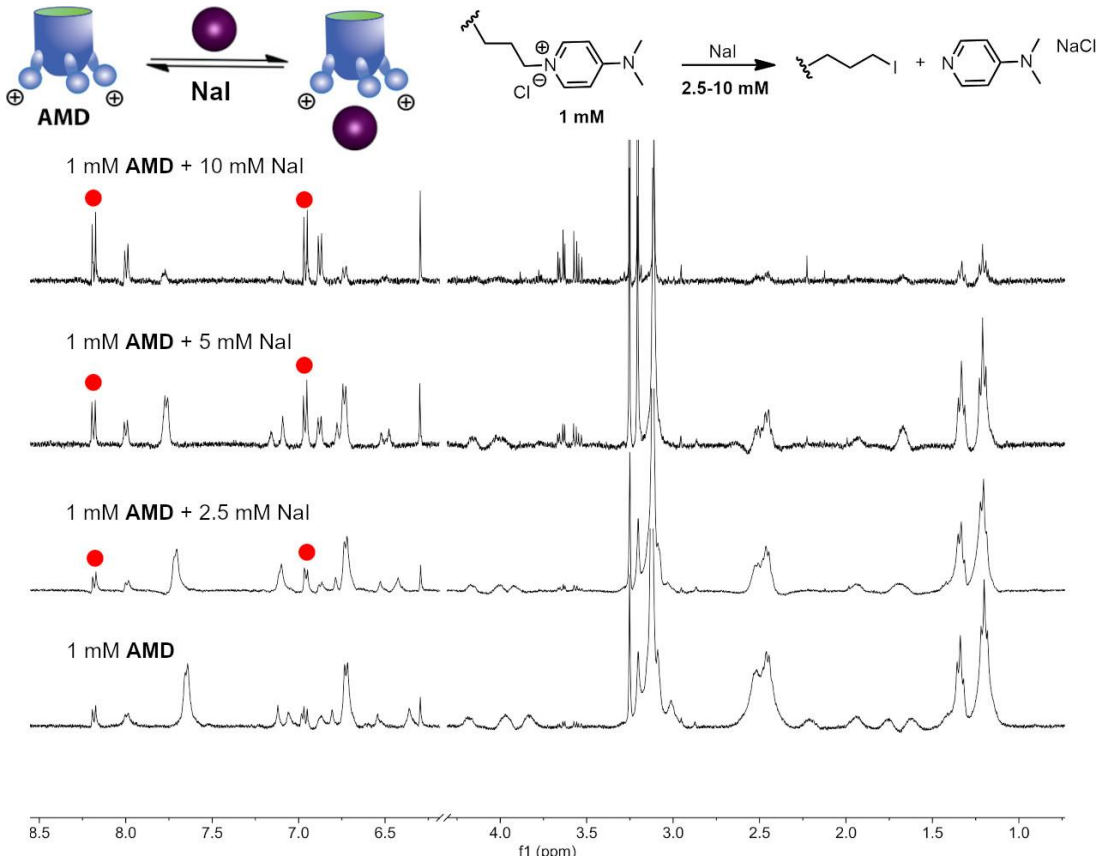


Figure 6.58. ^{1}H NMR spectra (D_2O, 600 MHz, 298 K) titrations of NaI with **4.4** showing cleavage of DMAP groups via substitution reaction with NaI, labeled peaks show emergence of free DMAP as the insoluble AM-Iodide cavitand is formed.

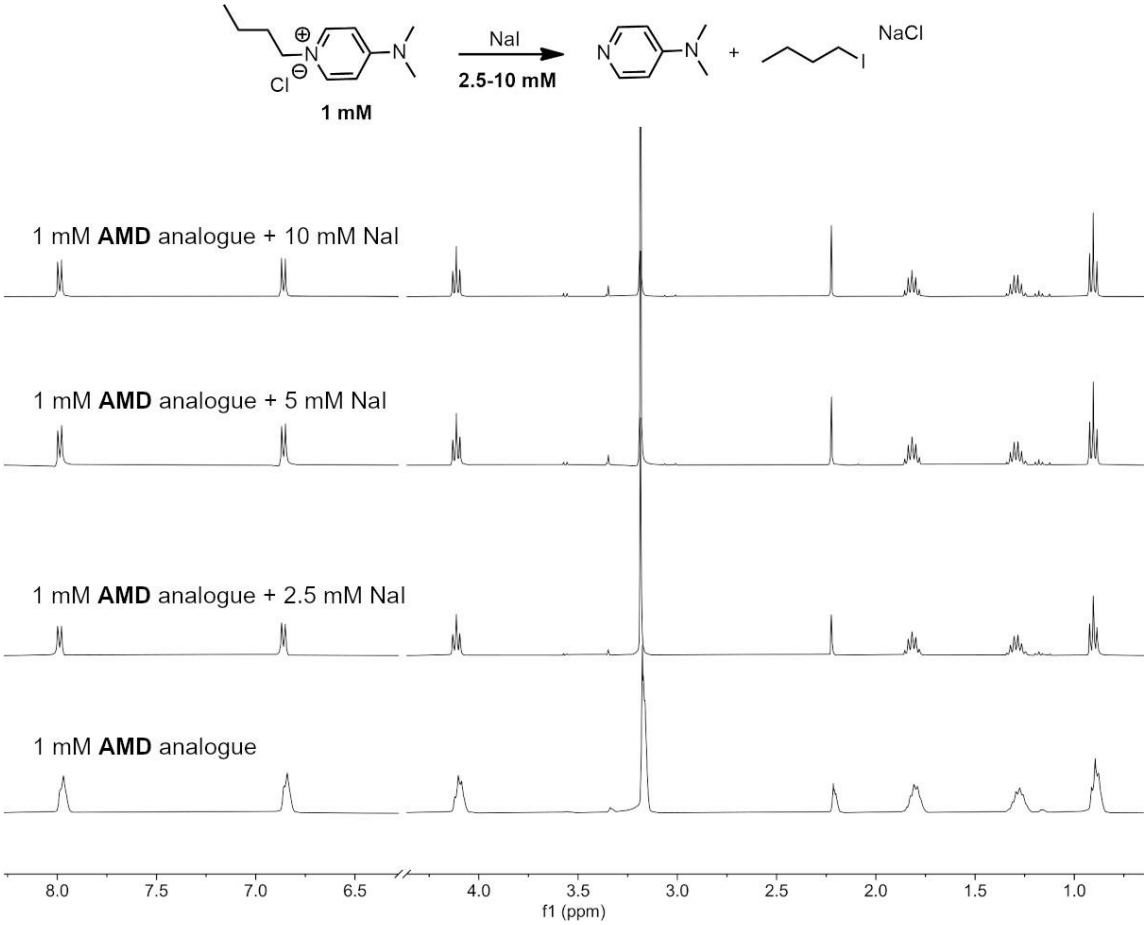


Figure 6.59. ^{1}H NMR spectra (D_2O, 600 MHz, 298 K) titrations of NaI with analogue **4.9** where no substitution reaction with NaI is observed.

Isothermal Calorimetry Analysis of AMI (4.4) •Anion Complexes

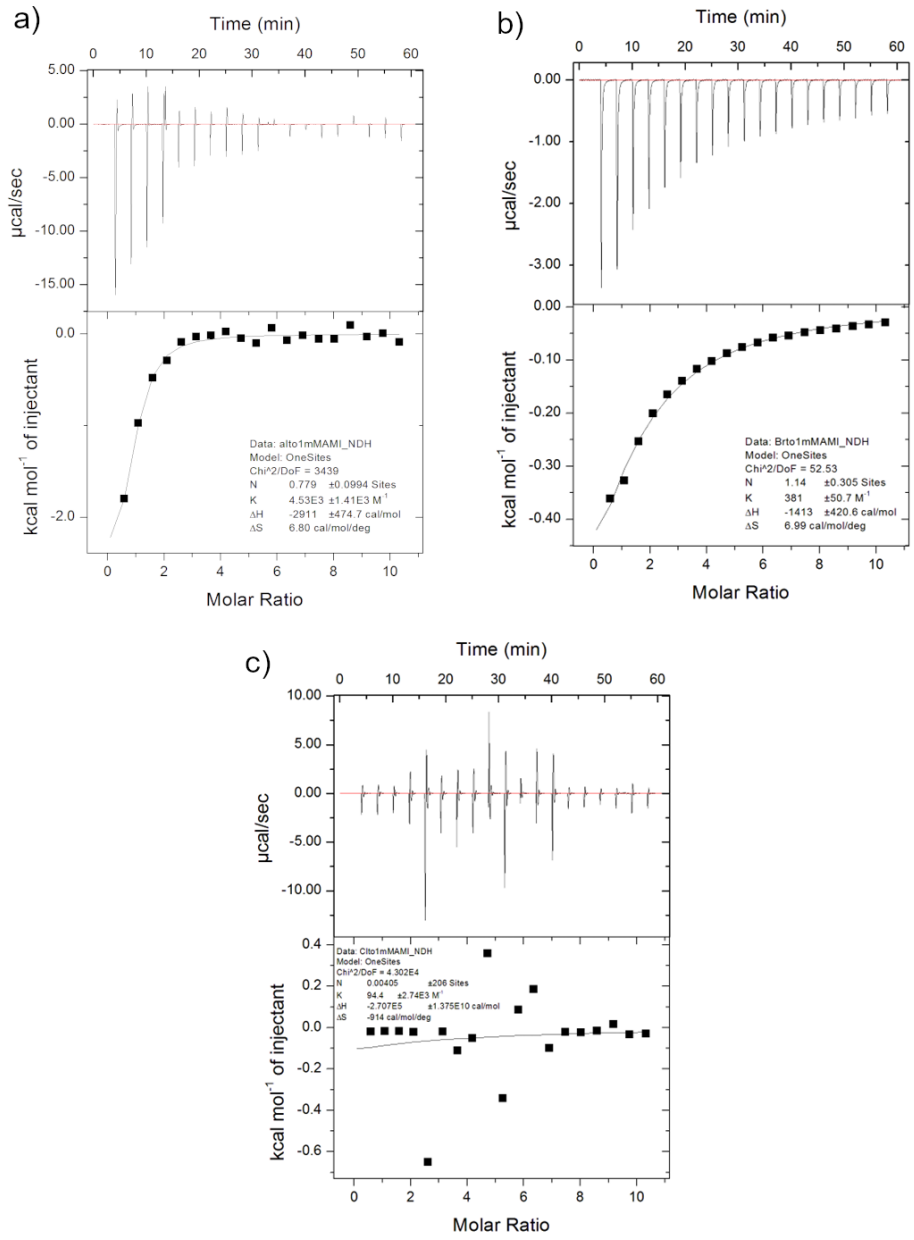


Figure 6.60. ITC titrations of increasing amounts of a) 50 mM NaI, and b) 50 mM NaBr, and c) 50 mM NaCl with 1 mM **4.4**, measured at 20 °C. The 1 mM solution of **4.4** was placed in the cell and 50 mM anion solution in the syringe. All solutions were diluted with ultrapure H_2O. Top trace: raw data for the ITC titration. Bottom trace: binding isotherm of the integrated calorimetric titration data. The heat of dilution, measured by the injection of titrant into H_2O, was subtracted for each titration to obtain the net reaction heat value.

Limit of Detection

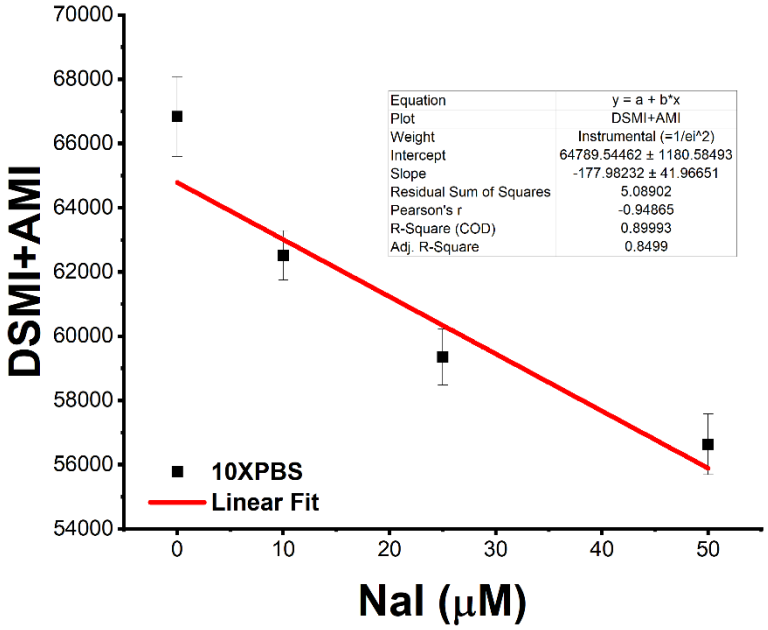


Figure 6.61. Limit of Detection (LOD) for **AMI•NaI** complex where [**AMI**] = 5 μM, [**NaI**] = 0-50 μM and [**DSMI**] = 0.5 μM in 10X PBS buffer. Error bars represent the standard deviation of 3 repeats. The calculation was done using the equation *LOD = 3 * (SD of blank) / (slope)* and was found to be 21 μM.

Hill 1 Plots

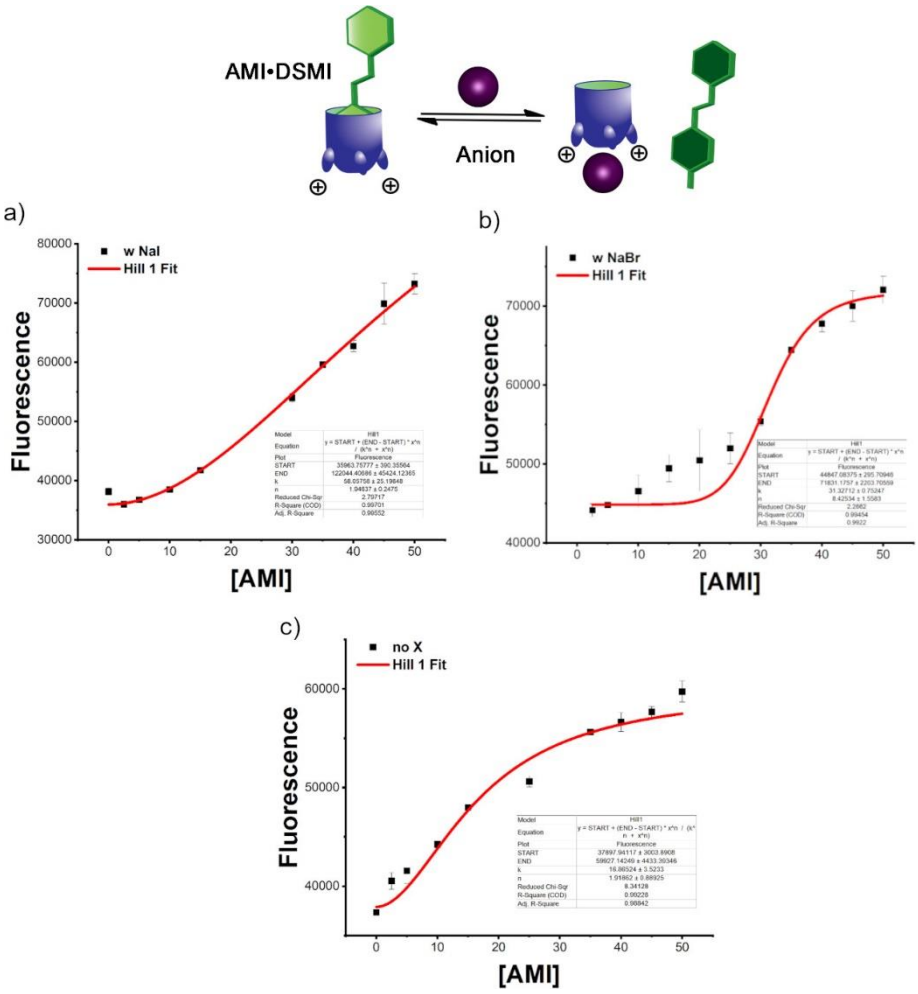


Figure 6.62. Hill 1 Plots of **DSMI** with increasing concentration of **AMI**, where a) shows binding with NaI, b) shows binding with NaBr, and c) is a control with no anion. The raw fluorescence of 0.5 μM **DSMI** + 0-50 μM **AMI** with 50 mM NaI, NaBr or no anion in ultrapure H_2O was collected and fitted with Hill 1 equation: *y = START + (END - START) * x^n / (k^n + x^n)* using Origin software. Error bars represent the standard deviation of 3 repeats.

Halide	k	n
No X	16.9 ± 3.5	1.9 ± 0.9
NaI	58.0 ± 25.2	1.9 ± 0.2
NaBr	31.3 ± 0.8	8.4 ± 1.6

Table 6.4. Summary of Hill 1 fitting of 0.5 μM **DSMI** + 0-50 μM **AMI** with 50 mM NaI, NaBr, or no anion in ultrapure H_2O.

6.9. Selected Spectra for Chapter 5

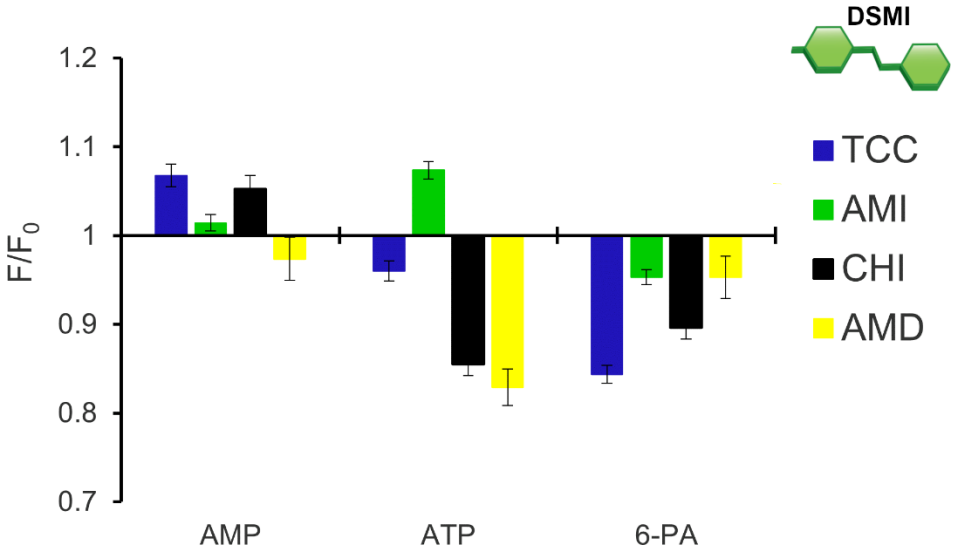


Figure 6.63. Relative fluorescence responses of the **Host•DSMI•Phosphosugar** complex in 20 mM Tris buffer, pH 7.4, [**Host**] = 5 μM, [**DSMI**] = 5 μM. F_0 = fluorescence response of the **Host•DSMI** complex, F = fluorescence response of the **Host•DSMI•Phosphosugar** complex.

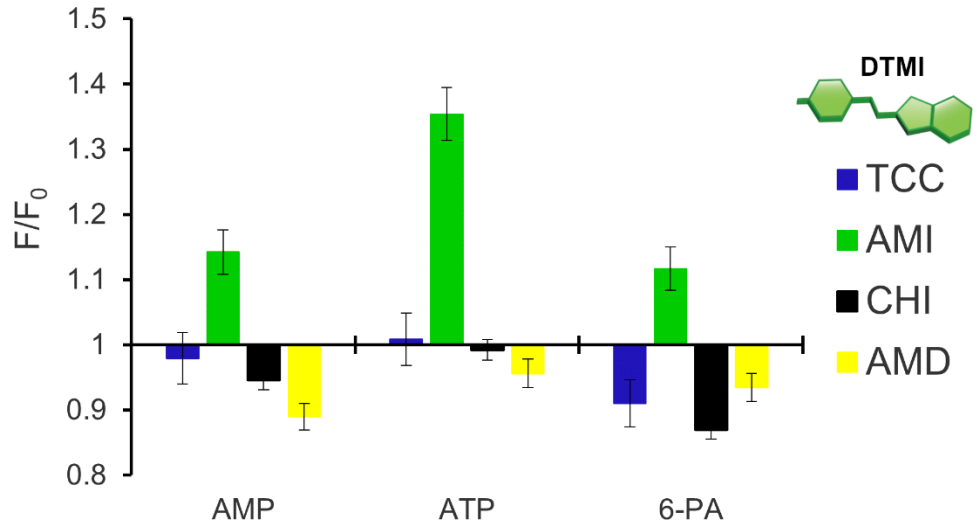


Figure 6.64. Relative fluorescence responses of the **Host•DTMI•Phosphosugar** complex in 20 mM Tris buffer, pH 7.4, [**Host**] = 5 μM, [**DTMI**] = 5 μM. F_0 = fluorescence response of the **Host•DTMI** complex, F = fluorescence response of the **Host•DTMI•Phosphosugar** complex.

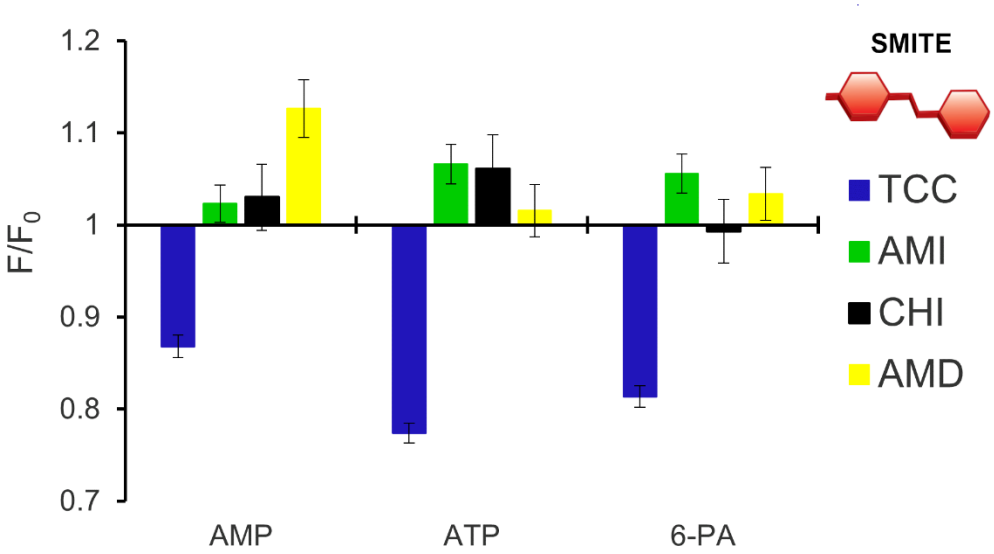


Figure 6.65. Relative fluorescence responses of the **Host•SMITE•Phosphosugar** complex in 20 mM Tris buffer, pH 7.4, [**Host**] = 5 μM, [**SMITE**] = 5 μM. F_0 = fluorescence response of the **Host•SMITE** complex, F = fluorescence response of the **Host•SMITE•Phosphosugar** complex.

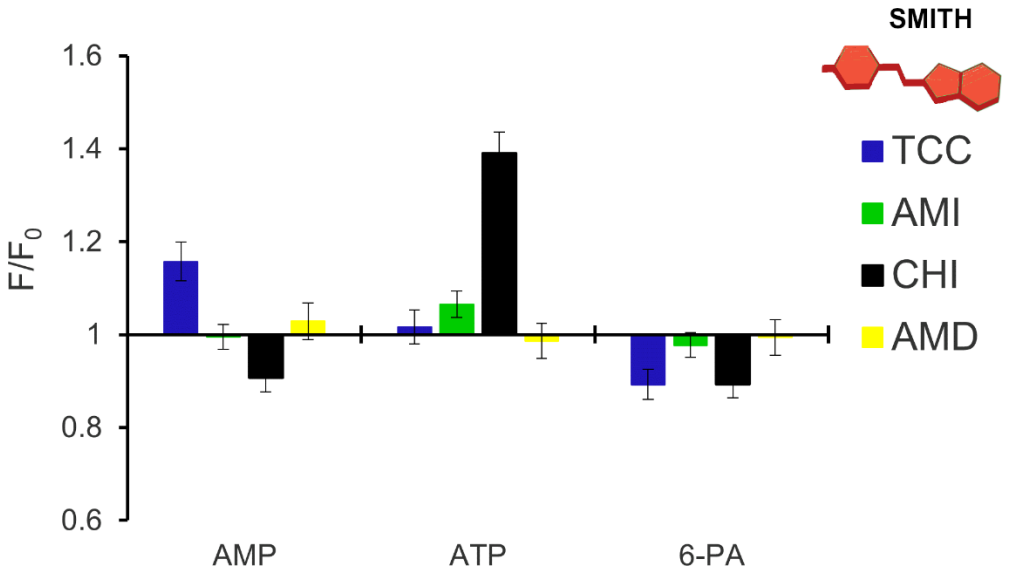


Figure 6.66. Relative fluorescence responses of the **Host•SMITH•Phosphosugar** complex in 20 mM Tris buffer, pH 7.4, [**Host**] = 5 μM, [**SMITH**] = 5 μM. F_0 = fluorescence response of the **Host•SMITH** complex, F = fluorescence response of the **Host•SMITH•Phosphosugar** complex.

Controls

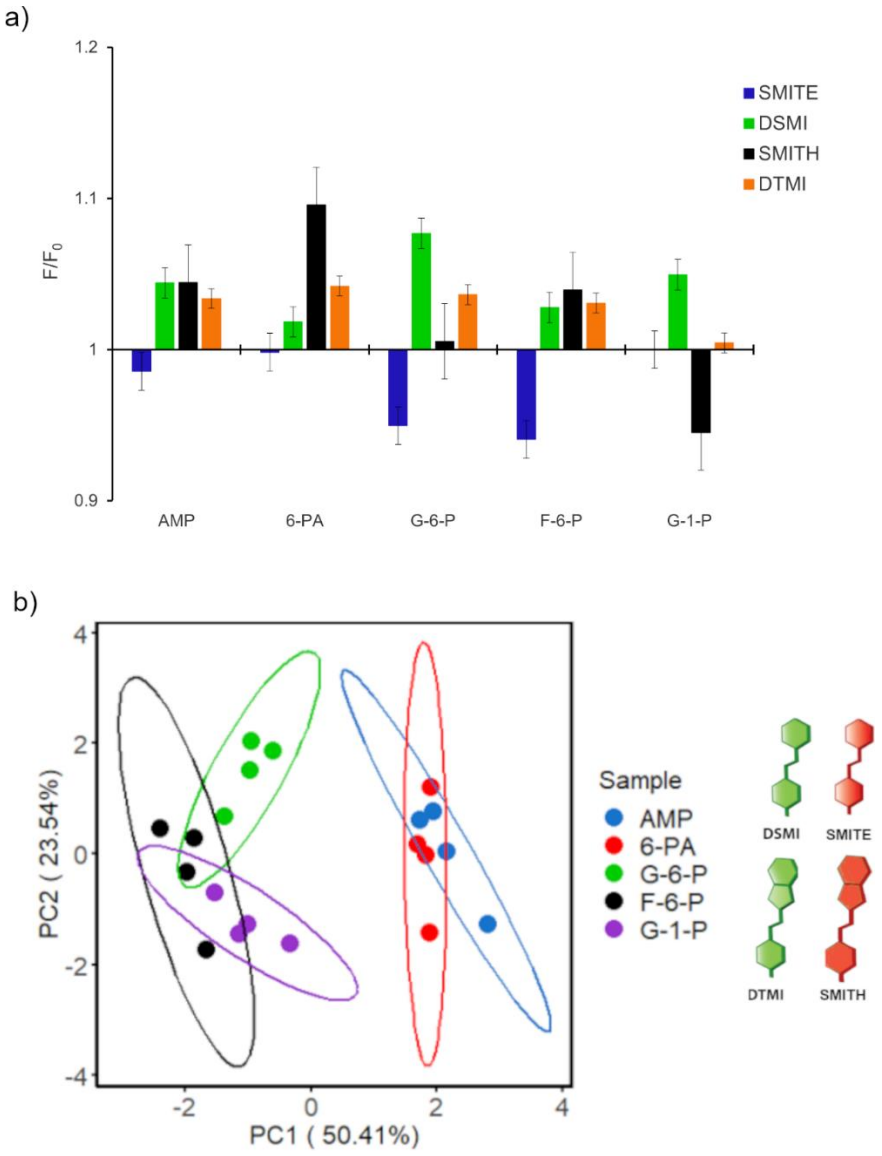


Figure 6.67. Control experiments for the a) relative fluorescence responses of the **Dye•Phosphosugar** complex in 20 mM Tris buffer, pH 7.4, [**Dye**] = 5 μM, [**Phosphosugar**] = 50 μM. F_0 = fluorescence response of the **Dye** alone, F = fluorescence response of the **Dye•Phosphosugar** complex, and b) the PCA scores plot for that fluorescence data.

6.10. References

1. Biros, S. M.; Ullrich E. C.; Hof F.; Trembleau, L.; Rebek, J. Jr. Kinetically Stable Complexes in Water: The Role of Hydration and Hydrophobicity. *J. Am. Chem. Soc*., **2004**, *126*, 2870–2876.

2. Mosca, S.; Yang, Y.; Rebek, J. Jr. Preparative Scale and Convenient Synthesis of a Water-Soluble, Deep Cavitand. *Nat. Proc.*, **2016**, *11*, 1371-1387.

3. Mitchell, R. F.; Hughes, D. T.; Luetje, C. W.; Millar, J. G.; Soriano-Agatón, F.; Hanks L. M.; Robertson, H. M. Sequencing and Characterizing Odorant Receptors of the Cerambycid Beetle *Megacyllene caryae*. *Insect Biochem. Mol. Biol.*, **2012**, *42*, 499-505.

4. Bianchi, P.; Roda, G.; Riva, S.; Danieli, B.; Zabelinskaja-Mackova, A.; Griengl, H. 3,4-O-Isopropylidene-2,7-di-O-p-tolylsulfonyl-alpha-L-xylo-3-heptulo-3,6-furanosononitrile. *Tetrahedron*, **2001**, *57*, 2213-2220.

5. Kang, S. K.; Lee, D. H.; Lee, J. M. Reduction of Organic Functional Groups Using Hypophosphites. *Bull. Korean Chem. Soc.*, **1990**, ***11***, 274-275.

6. Paolucci, C.; Mazzini, C.; Fava, A. Chiral Building Blocks from 1,4:3,6-dianhydrohexitols. iii. An Expeditious Enantiospecific Synthesis of the Geissman-Waiss Lactone. *J. Org. Chem.*, **1995**, ***60***, 169-175.

7. Bel-Rhlid, R.; Fauve, A.; Veschambre, H. Use of Biological Systems for the Preparation of Chiral Molecules. *J. Org. Chem.*, **1989**, *54*, 3221-3223.

8. Hall, D. R.; Cork, A.; Phythian, S. J.; Chittamuru, S.; Jayarama, B. K.; Venkatesha, M. G.; Sreedharan, K.; Vinod Kumar, P. K.; Seetharama H. G.; Naidu, R. The Coffee White Stemborer *Xylotrechus quadripes* (Coleoptera: Cerambycidae): Bioecology, Status and Management. *J. Chem. Ecol.*, **2006**, *32*, 195-219.