UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
"Scienza e Tecnologia dei Materiali"

CICLO XXXII

Biochemical and environmental sensing with cavitands

Coordinatore:
Prof. Enrico Dalcanale

Tutore:
Prof.ssa Roberta Pinalli

Dottorando: Alessia Favero

Alla mia famiglia

"Quasi tutti gli aspetti della vita sono regolati a livello molecolare, per cui, senza una chiara comprensione delle molecole, possiamo avere una cognizione solo molto sommaria della vita. "
Francis Crick

INDEX

APPENDIX A

1. DNA sequences and characterization

   1.1 DNA sequences
   1.2 Circular Dichroism (CD) corroboration of G4 topology
   1.3 CD analysis of *AG22* topology in Na^+/K^+ buffer
   1.4 Gel electrophoresis corroboration of unimolecolar G4 structures

2. NMR spectra of components used

   2.1 ^{1}H NMR spectrum of cavitand **1**
   2.2 ^{1}H NMR spectrum of cavitand **2**
   2.3 ^{1}H NMR spectrum of cavitand **3**
   2.4 ^{1}H NMR spectrum of cavitand **4**
   2.5 ^{1}H NMR spectrum of cavitand **5**
   2.6 ^{1}H NMR spectrum of **DSMI** dye
   2.7 ^{1}H NMR spectrum of **PSMI** dye

3. Minimized Host:Dye structures

4. Binding affinity measurements

   4.1 Fluorescence titrations of dyes and hosts **1-5**
   4.2 Isothermal calorimetric analysis of dye•***c-myc 2345*** binding
   4.3 Fluorescence titrations of dye•DNA
   4.4 Fluorescence response curves of host addition to **DMSI**•DNA complexes
   4.5 Fluorescence response curves of host addition to **PMSI**•DNA complexes
   4.6 Fluorescence response curves of DNA addition to **DMSI**•host complexes
   4.7 Fluorescence response curves of DNA addition to **PMSI**•host complexes

5. Limit of detection (LOD) for each host•dye complex

CHAPTER 4

Molecular recognition of aromatic aminoacids by deep cavitands

CHAPTER 5
Synthetic studies toward new benzene receptors for environmental sensing

CHAPTER 1

Supramolecular Sensing with Macrocycles

1.1 INTRODUCTION

Supramolecular sensing can be defined as the selective molecular recognition of target molecules (guests) by synthetic supramolecular receptors (hosts) exploiting weak interactions such as hydrogen bonds, cation-π, CH-π, π-π and Van der Waals interactions to produce a detectable change in a chemical or physical signal.[1] Designed to understand biological systems mimicking Nature's specific binding interactions, supramolecular receptors bears specific recognition functions, which allow the assembly of supramolecular aggregates, to bind selective substrates exploiting the lock and key principle[2] together with the concepts of shape recognition and binding site complementarity. These concepts get the maximum of expressiveness when incorporation of supramolecular receptors into chemical sensors gives rise to devices. Generally composed on a chemically sensitive layer able to attract and detect analyte molecules and a coupled transducer that transforms the established interaction into an analytical detectable signal (**Figure 1.1**), chemical sensors have to face selectivity, sensitivity and reliability demands.

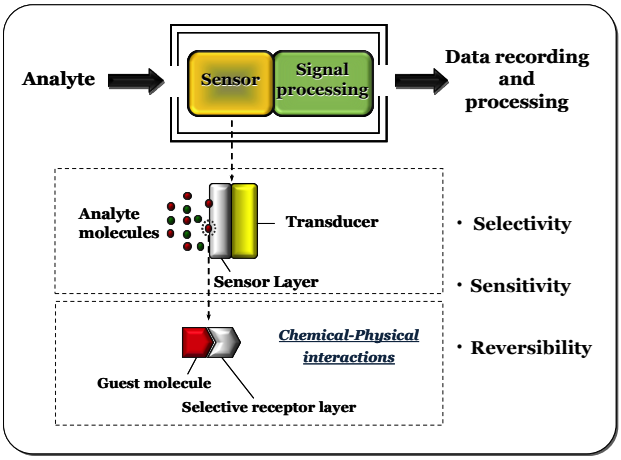


Figure 1.1. Working principles of supramolecular sensors.

Thus, the application of synthetic supramolecular receptors for molecular recognition in environmental, safety and biomedical applications must comply with the requirements of a sensor in terms of: (i) selectivity to

allow the detection of the target analyte in the presence of interferents; (ii) sensitivity to detect the desired analyte at very low concentrations; (iii) ruggedness to withstand various operating conditions, even unexpected ones.[3] Even the development of sensors including all the properties mentioned above remains the final goal, it realization is not possible. The opposition of concepts such as high sensitivity and selectivity, that required the establishment of strong interactions between analyte and receptor, versus reversibility, associated with weak interactions necessary to restore the sensor, meant that a compromise is compulsory. Hence, by preferring partial selectivity in favor of reversibility, has led to an inadequacy of the chemical sensors currently available, mainly affected by false positive and false negative responses. Therefore, the necessity to improve the recognition capacity of supramolecular sensors through a detailed understanding of the binding interactions between sensor and analyte, which must dominate over non-specific ones, has become predominant according to the class of guests to be detected.
At present, supramolecular receptors overtake antibodies in reproducibility, but not in selectivity. By combining the remarkable selectivity of single analyte sensors, differential arrays of synthetic receptors can be developed opening the possibility of compete with antibodies and aptamers for specificity in binding complex mixtures of bio- and environmental analytes, cellular imaging and quality control in manufacturing and processing.

The nature of the weak interactions involved in the recognition event is the field of study of Supramolecular Chemistry. Defined as “chemistry beyond the molecule” by Nobel Laureate Jean Marie Lehn,[4] bases its foundations on understanding and controlling weak interactions. These reversible interactions embrace a wide range of attractive and repulsive forces, in which hydrogen bond, cation-π, CH-π, π-π and Van der Waals interactions represent the most common ones.

Hydrogen bond. Historically, the most studied of these interactions is the hydrogen bond for its particular importance, since it is the basis of the existence of life. Acting as an important organizational element in the molecules of living things, the presence of these weak forces leads to highly thermodynamic stable structures, but still kinetically labile under

specific conditions (i.e. DNA double helix). Based on directional interactions, the strength of hydrogen bonds can span from a very weak to an almost covalent bond strength. Nobel Laureates Francis Crick and James Watson[5] gave a great insight into H-bond interactions, recognizing their pivotal role in the recognition of substrates by enzymes as well as for the maintenance of the double helix structure of DNA.

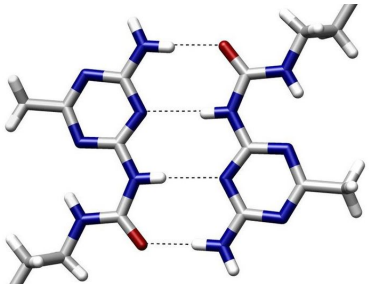


Figure 1.2. Visual example of intermolecular DADA hydrogen bonds in a self-assembled diaminotriazine dimer. Adapted from F.H. Beijer, H. Kooijman, A.L. Spek, R.P. Sijbesma, E.W. Meijer, *Angew. Chem. Int. Ed.* **1998**, *37*, 75.

π-π interaction. Similarly to H-bonds, π-π interactions play a decisive role in molecules assemblies, and once again DNA structure represents a prominent example (**Figure 1.3**). The interaction energies depend intimately on the orientations that the rings tend to adopt and are commonly divided into a columbic term and a van der Waals term; the first related to the relative charge distribution and the second to the contact surface area.[6] In spite of this, π-π interactions represent a significant example of a weak interaction especially in aqueous media where the effects of solvation tend to be significant.

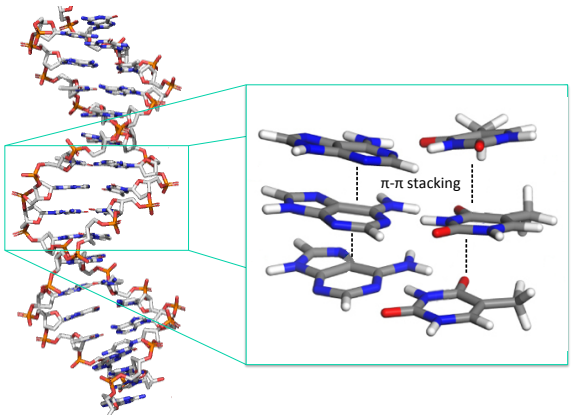


Figure 1.3. π-π interactions in DNA structure. Adapted from W.-Q. Deng, L. Sun, J.-D. Huang, S. Chai, S.-H. Wen, K.-L. Han, *Nat. Prot.* **2015**, *10*, 632.

Cation-π interaction. Cation-π interactions, considered electrostatic and containing a polarization component, have been successfully demonstrated in a variety of synthetic receptors as the binding of a cation to π-electrons, mainly belonging to aromatic rings.[7] Cation–π interactions are ubiquitous and are of prime importance in several fields, such as chemistry, material science, biology, and related areas.[8] Together with hydrogen bonding and π–π stacking interactions, cation–π interactions have emerged as forces of great importance in controlling the structure and function of macromolecules. The nature of this interaction is essentially of electrostatic origin deriving from a positively charged cation interacting with a negatively charged electron cloud of π systems. Hence, the strength of cation-π interaction depends both on the nature of cation and of the π acceptor. The significance and importance of the cation–π has been firmly recognized and appreciated in recent years.

Van der Waals interaction. The origin of van der Waals interactions lies in attractive or repulsive forces that are established between electronic clouds and nuclei of the atoms involved, which can be modified by the presence of atoms of nearby molecules and the surrounding environment.

Being based on these weak interactions, supramolecular systems are reversible, controllable and tunable as well as highly adaptable. Therefore, molecular receptors can be designed with respect to the target analyte by

exploiting the geometric and molecular complementarity of the interacting functional groups of the binding site leading to spontaneous assembly via thermodynamically-driven processes.

1.2 AMINO ACIDS RECOGNITION WITH MACROCYCLIC RECEPTORS

Medical diagnostics is one of the scientific fields in which the demand of sensors for the selective recognition of amino acids and other biological molecules is constantly increasing. The preventive healthcare requires highly performing devices for early detection of diseases to reduce mortality among the population and treatments costs. This demand leads chemists to design cheap and easy to use sensors capable to detect markers in biological fluids such as blood or urine, usually at very low concentration.
In this context, supramolecular chemistry offers a wide choice of synthetic receptors for biosensing: macrocycles such as cyclodextrins, calixarenes, cucurbiturils and cavitands have already demonstrated excellent molecular recognition properties for a large and diverse group of guests. The presence of a molecular-sized cavity that serves as a molecular recognition site for the analytes represents the common feature of these families of compounds. The main non-covalent interactions involved in the complexation events for these macrocycles are the above cited H-bonding, π-π stacking, cation-π and CH-π.

Cyclodextrins. Capable of complexing lipophilic molecules of appropriate shape and size, the hydrophobic cavity of cyclodextrins is a worthy representative of the macrocycles just described. Composed of D-glucose monomers, mostly in a variable number from 6 to 8, these macrocycles are commonly named α-, β- and γ-cyclodextrin, respectively (**Figure 1.4**). Countless studies have already been conducted in the last decades proving the excellent complexation ability of these macromolecules in sensing different types of guests[9] in aqueous media. The driving forces involved in complexation are represented by van der Waals forces, hydrophobic interactions and, to a lesser extent, hydrogen bonds.

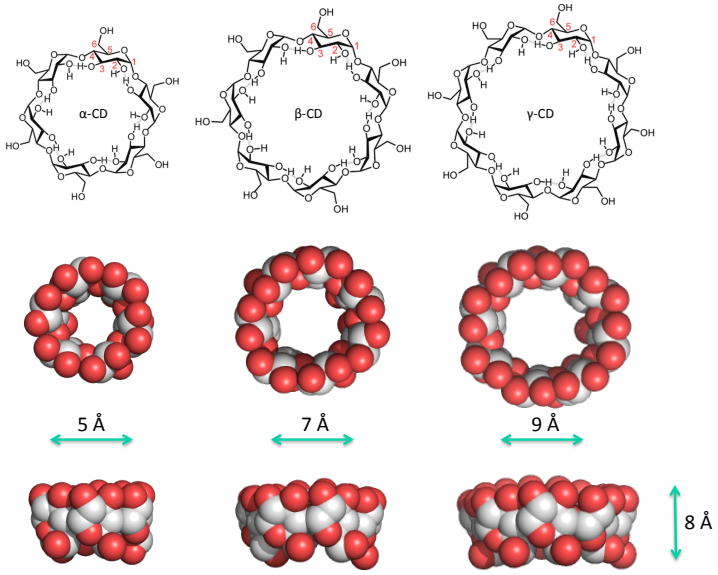


Figure 1.4. Structural and space-filling representation of α-, β- and γ-cyclodextrins.

Numerous examples of amino acids recognition by these receptors are also present in the literature.[10] In particular, the effectiveness of the complexation of L-lysine, L-phenylalanine and L-glutamic acid, by α- and β-cyclodextrin has been demonstrated.[11] The study underlines how the formation of 1:1 complexes is driven by supramolecular interactions like van der Waals interactions, hydrogen bonds, electrostatic forces and hydrophobic effects. Moreover, the release of two molecules of "high-energy water"[12] from the cavity boosts the complexing abilities of the macrocycles. Overall, the described receptor has showed *via* surface tension, pH, density, viscosity, and refractive index analyses an acceptable accuracy since the greater hydrophobic interaction between the -CH_2Ph group of L-Phe and the host hydrophobic cavity leads to the formation of a stronger inclusion, in which cavity size complementarity play a decisive role. Hence, the findings discussed demonstrate the suitability for significant different applications such as controlled transport and delivery of the guest in living cells, as well as the possibility of increasing the solubility and reactivity of the guest molecules without any chemical modification.

Constant development in pharmaceutical and therapeutic fields has led to investigate the interactions exploited in biological systems as an approach for the optimization of potential drugs. The interaction between β-cyclodextrin and amyloid β-peptide (Aβ), identified as a major feature of insoluble and neurotoxic amyloid fibrils formation in Alzheimer's disease, represents an exquisite example.[13] The capability of β-cyclodextrin to inhibit the conformation change of Aβ from random coil to a more organized β-sheet was demonstrated *via* NMR and Circular dichroism experiments. In fact, by preventing aggregation and fibril formation in the β-amyloid peptide, the accumulation of insoluble neurofibrillary cell in neurons can be countered. Downfield shifts (^{1}H NMR) together with cross-peaks (NOESY NMR) of Aβ aromatic protons of Phe19 and Phe20 with β-cyclodextrin further confirmed the inclusion of the amino acidic residues in the inner ring for the cavity of the receptor. Thus, the inhibition of these neurotoxic fibrils formation opens the way to further investigation in the pharmaceutical proprieties of β-cyclodextrin as a potential therapeutics for Alzheimer's disease.

Calix[n]arenes. Whereas the inclusion of some naturally occurring amino acids in natural cyclodextrins has been extensive studied, also the ability of water soluble calixarenes in biding a wide range of such biomolecules have already been reported.[14] Besides complementary shape, multiple non-covalent weak interactions are involved in the complexation mode of calix[4]arene, such as hydrogen bonding, π-π stacking, cation-π and CH-π interactions.[15]
Easily functionalized on the upper and lower rim to adapt their recognition properties to the desired class of analytes, calixarenes can be synthesized with a number of units ranging from 4 to 8.[16] In the concept of cavity complementary, the dimension of the cavity can also be modulated exploiting the different conformations available for these macrocycles.

A complete study presents the exquisite recognition properties of one of the main exponent of the calixarene family, the *p*-sulfonatocalix[4]arene.[17] In particular, the ability of complexing a series of natural occurring amino acids in buffered solutions was demonstrated. ^{1}H NMR and ITC experiments shows that the complexation event is enthalpically driven and, surprisingly, even small changes of the lower rim pattern can

influence the binding properties of the receptor apolar cavity. Among the studied amino acids, histidine turned out having the lowest affinity for the calixarene cavity: the conditions used in the study led to a partial protonation of the imidazole ring reducing the apolar character of this AA. On the contrary, L-leucine (L-Leu) and L-phenylalanine (L-Phe) resulted to have the highest affinity for the studied host. In the first case, the driving force for the inclusion lies in the CH-π interactions between the L-Leu side chain and the host cavity, while in the second case, the presence of the aromatic residue in L-Phe allows the establishment of π-π interactions.

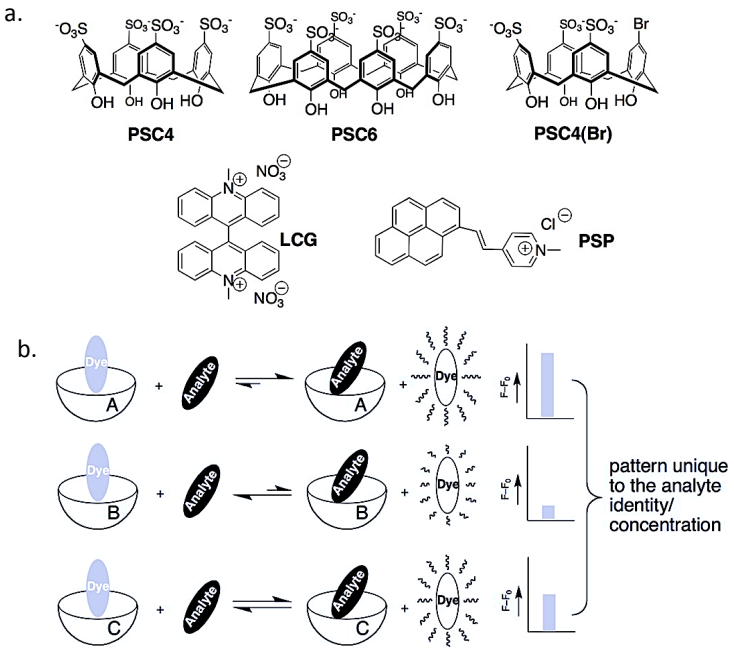


Figure 1.5. a) Supramolecular *p*-sulfonatocalix[n]arenes receptors and fluorescent dyes used in the sensor arrays; **b)** visual example of the receptors working principles. Adapted from S.A. Minaker, K.D. Daze, M.C.F. Ma, F. Hof, *J. Am.Chem. Soc*. **2012**, 134, 11674.

Combining *p*-sulfonatocalix[n]arenes of different cavity size, a sensor array for the discrimination of post-translational modified (PTM) lysines and arginines was developed (**Figure 1.5**).[18] Known to affect gene regulation, epigenetic inheritance and cancer, PTMs are considered useful targets for new cancer therapies. In this perspective, the possibility to identify and

discriminate closely related modified amino acids represents an attractive opportunity. To achieve this goal, a sensor based on a fluorescent indicator displacement was tested. By using a fluorescent dye bounding and quenching the *p*-sulfonatocalix[n]arenes hosts, a turn-on response was obtained after the addition of the analytes, which compete with the dye for the host's binding site. Operating in buffered conditions and/or varying the pH, unique sets of different fluorescence responses were obtained for different guests related to analyte affinities for each of the *p*-sulfonatocalix[n]arenes used. The richness of information obtainable combined to the simplicity of the analyses led to a cheap and easy to set up array sensor able to operate in homogeneous solution providing continuous data with a confidential level of discrimination up to 99%.

Taking inspiration from natural recognition processes occurring at lipid membrane interfaces, a sensing platform at the air-water interface was developed exploiting an interesting approach to fluorescent detection of amino acids and proteins using calix[n]arenes as receptors.[19] The platform is composed of a lipid monolayer doped with a small amount of molecular receptors, namely calix[4]arene tetraphosphonate, which thanks to its amphiphilic nature points its cavity toward water allowing the interaction with protein surfaces (**Figure 1.6**). The addition of basic amino acids like lysine and arginine to the doped monolayer expanded it upon binding. The monolayer was also demonstrated to attract peptides and proteins (e.g. Proteinase K, thrombin, BSA, chymotrypsin, trypsin, cytochrome c, histone H1, and ferritin) from the aqueous sub-phase complexing them in a cooperative fashion *via* multiple binding.

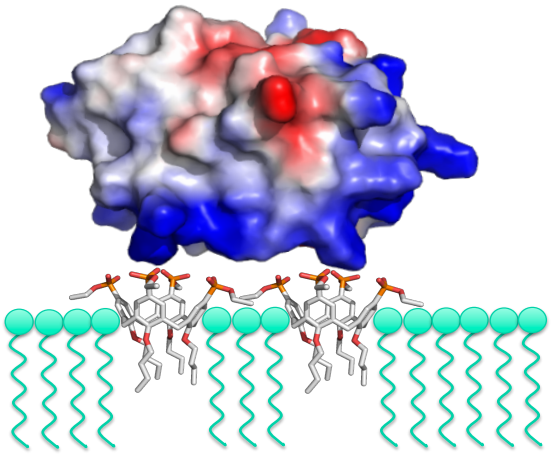


Figure 1.6. Schematic representation of cytochrome C (depicted in electrostatic surface potential) being bound by tetraphosphonate calix[4]arene receptor in the stearic acid monolayer. Adapted from R. Zadmard, T. Schrader, *J. Am. Chem. Soc.* **2005**, *127*, 904.

Cavitands. A subfamily of calixarenes is represented by cavitands. Similarly to calix[n]arenes, resorcinarene-based macrocycles can be functionalized at both the upper and lower rim to impart specific binding properties to the resulting receptor and to tune solubility. The nature of the bridging groups connecting the phenolic hydroxyls of the resorcinarene scaffold determines the shape and size of the cavity as well as the final complexing properties of the receptor. Multiple synergic supramolecular interactions such as hydrogen bonds,[20] π-π stacking,[21] CH-π[22] and cation-π[7] interactions, represent the main interactions responsible of cavitands recognition properties. In biochemical sensing, the most studied class of cavitands is represented by tetraphosphonate cavitands.

Tetraphosphonate cavitands (Tiiii) have demonstrated remarkable molecular recognition abilities mainly towards mono methyl ammonium salts.[23] The receptor selectivity derives from the presence of specific synergistic interaction modes such as N^{+}--O=P cation-dipole interactions, cation-π interactions between the acidic $^{+}N\text{-}CH_3$ group and the basic resorcinarene scaffold, and two hydrogen bonds between two adjacent P=O bridges and the nitrogen protons (**Figure 1.7**). The simultaneous presence of all these interactions determine the selectivity of Tiiii towards

mono-methylated species compared to non-, di- and tri-methylated ones, where cation–π or H-bonding interactions are not or less present.

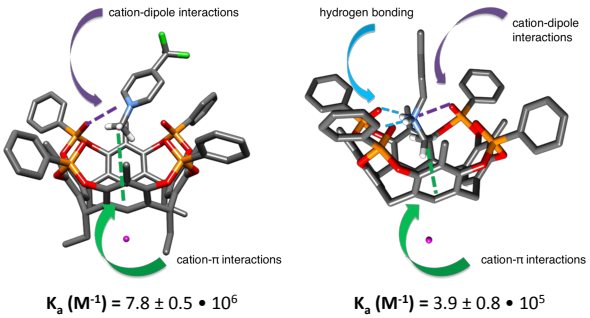


Figure 1.7. Non-covalent interactions involved in Tiiii receptor recognition with N-methyl pyridinium (left) and N-methyl ammonium salt (right). Adapted from E. Biavardi, C. Tudisco, F. Maffei, A. Motta, C. Massera, G.G. Condorelli, E. Dalcanale, *Proc. Natl. Acad. Sci. U. S. A*. **2012**, *109*, 2263.

An elegant example[24] of the selectivity of Tiiii towards monomethylated ammonium salt is the detection of sarcosine, a marker of aggressive prostate cancer (PC),[25] in urine samples.
Exploiting the demonstrated ability of Tiiii to recognize sarcosine from its non-methylated biochemical precursor glycine *via* crystal structure and ITC, a new sensor based on a cavitand-decorated Si surface was developed (**Figure 1.8**). By taking advantage of silicon to reducing or even eliminating nonspecific interactions that often mask the recognition events, a Tiiii@Si wafer was prepared *via* photochemical hydrosilylation and tested both in water and in urine. Complementary XPS and fluorescence guest displacement tests revealed an exquisite selectivity for sarcosine at the solid–liquid interface, even in presence of naturally occurred potential ionic interferents.

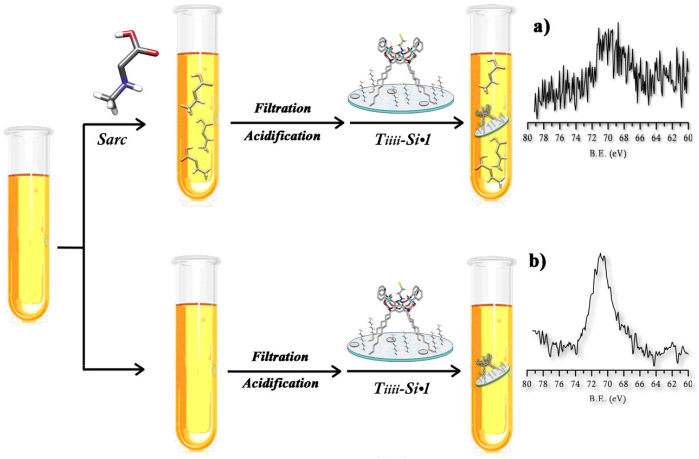


Figure 1.8. Sarcosine detection in human urine. **a)** XPS Br 3D spectrum of Tiiii@Si sliced dipped in sarcosine-added sample and **b)** in the control sample. Taken from E. Biavardi, C. Tudisco, F. Maffei, A. Motta, C. Massera, G.G. Condorelli, E. Dalcanale, *Proc. Natl. Acad. Sci. U. S. A*. **2012**, *109*, 2263.

A simple, portable and inexpensive device was then developed for accurate and reliable diagnoses of PC. By covalently decorating magnetic microspheres (MMB) functionalized with tetraphosphonate cavitands,[26] sarcosine detection was obtained with electrochemiluminescence (ECL) measurements at the µM-mM range leading to a Tiiii@MMB sensor proven to efficiently quantifying sarcosine in complex matrices at the concentration necessary for a diagnosis in healthy subjects and patients with PC.

Having demonstrated the possibility of using this macrocycle for the selective identification of biologically active compounds containing the H_2N^+-CH_3 group, such as synthetic drugs [27] and biomarkers [24] in biological fluids, a sensor capable of identifying the entire class of amphetamine, in particular the methylated homolog – methamphetamine – and recognizing the so-called "designer-drugs" was developed.
After studying the interaction modalities established by Tiiii with MDMA, cocaine, amphetamine, and 3-fluoromethamphetamine (3-FMA) hydrochloride salts in the formation of host-guest complexes by X-ray diffraction analysis on single crystals, the known interactions already described for AA were revealed. Then, thanks to the grafting of

tetraphosphonate cavitands on silicon microcantilevers (Si-MC),[28] the molecular recognition is translated into nanomechanical work produced by the host-guest interaction.[29]
The sensor, also tested on a real "road" sample, showed a nominal detection limit equal to tenths of ppm and the ability to selectively recognize "designer-drug" in the presence of excipients present in the sample, i.e. caffeine and glucose.

Cucurbits[n]urils. Similarly to cyclodextrins in their form, cucurbits[n]urils (n = 5–8 and 10; CB[n]) exhibit a barrel type cavity formed by glycoluril oligomers joined by methylene bridges (**Figure 1.9**).[30] The number of glycoluril groups determines not only the size of the CB[n] cavity, but also other chemical-physical proprieties such as the receptor solubility in water. Equipped with an apolar cavity able to host hydrophobic guests, the main supramolecular interactions exploited are the non-classical hydrophobic effect, due to the release of "high energy water" following by the inclusion of apolar organic residues,[12] ion-dipole interactions and hydrogen bonds.
The development of CB[n]s host-guest chemistry has been strongly empowered in recent years, resulted in significant applications in the detection of analytes or biologically relevant molecules in water.

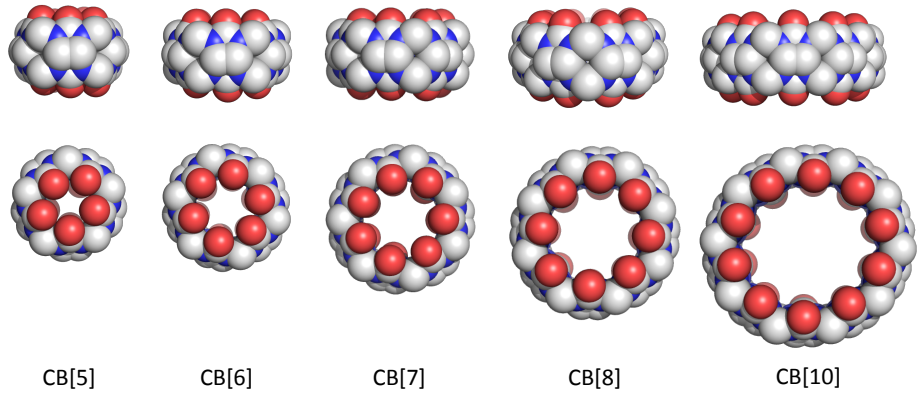


Figure 1.9. CBs structures and cavity increasing size related to glycoluril units.

An important example concerning the recognition of aromatic amino acids in aqueous media is represented by the selective recognition of phenylalanine (Phe) on an insulin residue.[31] In the study, CB[7] receptor, by

binding to the N-terminal aromatic residue Phe in human insulin, is able to selectively recognize insulin also in presence of other proteins that do not possess terminal aromatic residues (**Figure 1.10**). This approach to native proteins recognition differs from existing methods, such as the genetic engineering of polyhistidine (His-tag) or chemical modification of proteins with markers, since the protein recognition occurs by targeting the terminus without modifying the protein.

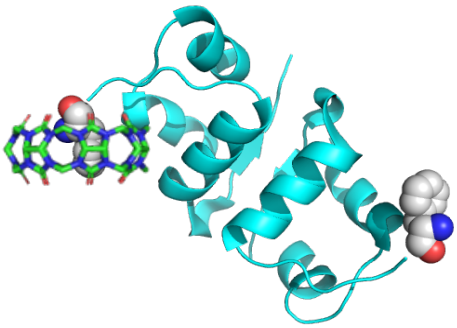


Figure 1.10. CB[7] binding on a Phenylalanine terminal residue on human insulin. Adapted from J.M. Chinai, A.B. Taylor, L.M. Ryno, N.D. Hargreaves, C.A. Morris, P.J. Hart, A.R. Urbach*, J. Am. Chem. Soc.* **2011**, *133*, 8810.

Recently, the use of CB[8] in the electrochemical detection of tryptophan (Trp) has been proposed.[32] Preliminary electrochemical experiments in solution were conducted[33] revealing Trp inclusion in the CB[8] cavity. Upon binding the highest E_p shift (about 100 mV) for the Trp response to increasing positive values was recorded along with a decrease in the oxidation current. Further studies using a functionalized glassy carbon electrode *via* deposition of CB[8] and Nafion from a homogeneous suspension, confirm the possibility of detecting Trp at the nanomolar level, even in the presence of interferents, making the sensor a valid candidate in the quantification of Trp in human serum for early diagnosis.

By providing new advantageous tools for chemical investigations on biologically important pathways, supramolecular receptor act as a vehicle for the understanding of biological interactions giving the ability to target, probe, detect, sense and learn how life works.

1.3 ENVIRONMENTAL GAS SENSING WITH MACROCYCLIC RECEPTORS

With the advent of industries and modernization in our lives, high living standards have been achieved but at the same time the phenomenon of environmental pollution has spread worldwide. Every day millions of chemicals are used or produced in the most varied fields such as industry, agriculture, pharmaceutics, cosmetics and transport. It follows that poor air quality of cities affects also the soil, contaminated by the same air-pollutants as well as heavy metals and pesticides. As a wistful cycle, soil pollutant ended in aquifers, harming the quality of the drinking water. By increasing mortality rates, the pollution produced by humans every day is therefore damaging not only the entire ecosystem but is becoming a serious danger to human health. Being exposed daily to environmental contaminants, pollutants and toxins, the world population has demonstrated the dangerousness of the phenomenon by undergoing an increase in cases of cancers of the gastrointestinal tract due to the consumption of contaminated water, just to give an example.[34] Thus, monitoring and detecting of all contaminants in the environment matrices in order to prevent serious diseases for humans and the planet has become more and more important.

Once again, valuable help comes from chemosensors for diagnosing and monitoring the environment as well as controlling industrial safety and quality.[35] Environmental monitoring represents an intriguing challenge for the chemical community, due to the stringent sensor requirements for sensitivity, selectivity and ruggedness, namely the ability of a sensor to maintain its performance under particular adverse operating conditions.[36] Compared to biological receptors, normally operating in aqueous environment where dispersion interactions largely cancel out in moving the analyte from water to the receptor site and where the entropic cost of binding is partly paid by water release, sensors for gaseous pollutants must faces general and non-specific dispersion interactions during the recognition event due to the moving of the analyte to the solid state. Furthermore as already pointed out, they are required to possess sensitivity at the ppb level, molecular selectivity towards the analyte of interest, and ability to maintain their performances even in adverse operating conditions over the presence of huge quantities of interferents.[37]

Taking inspiration from Nature and exploiting the concepts of complementarity of form and binding site, several molecular receptors have been designed. Consequently, numerous examples of molecular receptors based on cyclodextrins, calixarenes and cavitands for gas sensing are reported in literature.

Cyclodextrins. Among them, electrochemical sensors consisting of carbon-paste electrodes (CPE) functionalized with α-, β- or γ-cyclodextrin and screen-printed carbon-based electrodes (SPE) decorated with β-CD or derivatives, have been tested using voltammetry technique for the selective recognition of carcinogenic polycyclic aromatic hydrocarbons amino derivatives (APAHs) (**Figure 1.11**).[38] By taking advantage of the ease of anodic oxidation of the aromatic amines mediated by voltammetry and the possibility of in-situ preconcentration of analytes on the surface of the electrode chemically modified with CDs, it is possible to detect APAHs also at trace level. This is possible due to the preconcentration effect by the β-CD macrocycle, better than the α- and γ-CD, which led to the drafting of an analytical procedure for the determination of APAHs. The amino acids of polycyclic aromatic hydrocarbons are known chemical carcinogens commonly present in the environment,[39] they are metabolites of nitroderivatives of polycyclic aromatic hydrocarbons (NPAH) which are also carcinogenic[40] and because of that they can be used as biomarkers for monitoring human exposure to NPAH,[41] becoming a valuable tool for diagnosis and prevention.
On the same line, the monitoring of industrial emission of harmful volatile organic compounds (VOCs) play a key role in the prevention of human health and the environment. VOCs, aside from human-made and naturally occurring chemical compounds, are harmful organic chemical possessing "an initial boiling point less than or equal to 250 °C measured at a standard atmospheric pressure of 101.3 kPa"[42] commonly represented by halogenated and aromatic hydrocarbons such as Benzene, Toluene, Ethylbenzene and Xylenes (BTEX). Because of their low concentrations and slow-developing symptoms, these compound are typically not acutely toxic in short-term, but their correlation in cancer development requires adequate monitoring systems. One of them is represented by an electronic nose based on a suitably functionalized β-cyclodextrin deposited on reduced graphene oxide (RGO) transducer.[43] Using pyrene adamantane as

a linker to construct a supramolecular assembly between the non-covalent linked β-CD with RGO, a highly selective VOCs sensor was developed. The combination of host–guest inclusion combined to the high surface area and electrical conductivity of graphene, gave rise to a sensor capable to discriminate between the ten cancer-VOCs related biomarkers at ppb level.

Calix[n]arenes. The great versatility in tuning the functional groups on calix[n]arene remain one of the most attractive feature as host molecules. Recently, an organic vapor sensing properties of Langmuir–Blodgett (LB) thin films of calix[4]arene derivatives that contain different numbers of *tert*-butyl groups on their upper rims were reported (**Figure 1.12**).[44] By using a quartz crystal microbalance (QCM), the detection of chloroform, benzene, toluene and ethanol gas vapors was investigated resulting in a time-dependent reversible response of the calix@LB films towards VOCs vapors with an unexpected effect of the cavity size and number of calix[n]arene substituents. In fact, despite the removal of *tert*-butyl groups in the *para* position generally decrease the molecular interaction of calix[n]arene, the formation of the partially conical conformation in the receptor bearing three p-*tert*-butyl groups on the upper rim of the scaffold, leads to obtain excellent experimental results affecting the gas sensing efficiency of the receptor.

Cavitands. As demonstrated by the examples just discussed, supramolecular receptors reveal their complexing properties towards the recognition of specific analytes in a sublime way. By taking advantage of the incorporation of solvent molecules in the gas state into the cavity of resorcinarene-based macrocycles previously deposited on quartz microbalance-based sensors (QMB), it is possible to create microelectronic devices by converting the concentrations of bound analytes into electrical signals.[45] Thus, combining the supramolecular properties of cavitands with the principles of the transducer, the development of sensors for qualitative and quantitative detection of analytes in air, even at ppm level, lead to a new era of environmental sensing. A detailed description of the literature related to environmental sensing with cavitands is reported in Chapter 5.

1.4 REFERENCES

1. L. You, D. Zha, E.V. Anslyn, *Chem. Rev.* 2015, 115, 7840.
2. J.-P. Behr, *The Lock and Key principle. The state of the art – 100 years on,* Wiley Publishing, New York, **1994**.
3. R. Pinalli, E. Dalcanale*, Acc. Chem. Res.* **2013***, 46,* 399.
4. J.M. Lehn*, Pure Appl. Chem*. **1978**, *50*, 871.
5. J.D. Watson, F.H. Crick, *Nature* **1953**, 1*71*, 737.
6. W.B. Jennings, B.M. Farrell, J.F. Malone, *Acc. Chem. Res.* **2001**, *34*, 885.
7. D.A. Dougherty, *Acc. Chem. Res.* **2013**, *46*, 885.
8. A. Subha Mahadevi, G. Narahari Sastry, *Chem. Rev.* **2013**, *113*, 2100.
9. M.V Rekharsky, Y. Inoue, *Chem. Rev.* **1998**, *98*, 1875.
10. H.J. Schneider, F. Hacket, V. Rudiger, H. Ikeda, *Chem. Rev.* **1998**, *98*, 1755.
11. M.N. Roy, D. Ekka, S. Saha, M.C. Roy, Florea, *RCS Adv.* **2014**, *4*, 42383.
12. F. Biedermann, W.M. Nau, H.J. Schneider, *Angew. Chem. Int. Ed.* **2014**, *53*, 11158.
13. X.-R. Qin, H. Abe, H. Nakanishi, *Biochem. Biophys. Res. Comm.* **2002**, *297*, 1011.
14. P. Neri, J.L. Sessler, M.-X. Wang, *Calixarenes and Beyond*, Springer International Publishing, Netherlands, **2016**.
15. (a) L. Mutihac, J.H. Lee, J.S. Kim, J. Vicens, *Chem. Soc. Rev*. **2011**, *40*, 2777; (b) K. Ariga, H. Ito, J.P. Hill, H. Tsukube, *Chem. Soc. Rev*. **2012**, *41*, 5800.
16. (a) C.D. Gutsche, R. Muthukrishnan, *J. Org. Chem.* **1978**, *43*, 4905; (b) H. Kammerer, G. Happel, F. Caesar, *Macromol. Chem.* **1972**, *162*, 179; (c) M.-Z. Asfari, V. Böhmer, J. Harrowfield, J. Vicens, *Calixarenes 2001*, Springer International Publishing, Netherlands, **2001**; (d) C.D. Gutsche, *Calixarenes: An Introduction* - Monographs in Supramolecular Chemistry, The Royal Society of Chemistry, Cambridge, **2008**.
17. G. Arena, A. Casnati, A. Contino, A. Magrì, F. Sansone, D. Sciotto, R. Ungaro, *Org. Biomol. Chem.* **2006**, *4*, 243.
18. S.A. Minaker, K.D. Daze, M.C.F. Ma, F. Hof, *J. Am.Chem. Soc*. **2012**, 134, 11674.
19. R. Zadmard, T. Schrader, *J. Am. Chem. Soc.* **2005**, *127*, 904.
20. J. Rebek, *Angew. Chem. Int. Ed.* **1990**, *29*, 245.
21. H. Hunter, K.R. Lawson, J. Perkins, C.J. Urch, *J. Chem. Soc., Perkin Trans.* **2001**, *2*, 651.
22. M. Nishio, M. Hirota, Y. Umezawa, *The CH-π Interaction: Evidence, Nature, and Consequences*, Wiley-VCH, New York, **1998**.

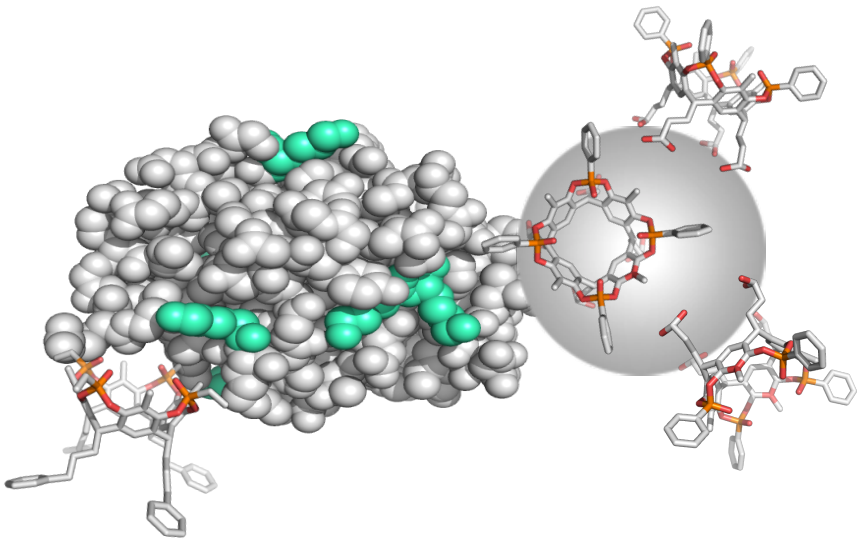
23. (a) R.M. Yebeutchou, E. Dalcanale, *J. Am. Chem. Soc.* **2009**, *131*, 2452; (b) D. Menozzi, E. Biavardi, C. Massera, F.P. Schmidtchen, A. Cornia, E. Dalcanale, *Supramol. Chem*. **2010**, *22*, 768; (c) E. Biavardi, M. Favazza, A. Motta, I.L. Fragalà, C. Massera, L. Prodi, M. Montalti, M. Melegari, G.G. Condorelli, E. Dalcanale, *J. Am. Chem. Soc.* **2009**, *131*, 7447; (d) M. Dionisio, G. Oliviero, D. Menozzi, S. Federici, R.M. Yebeutchou, F.P. Schmidtchen, E. Dalcanale, P. Bergese, *J. Am. Chem. Soc.* **2012**, *134*, 1392.
24. E. Biavardi, C. Tudisco, F. Maffei, A. Motta, C. Massera, G.G. Condorelli, E. Dalcanale, *Proc. Natl. Acad. Sci. U. S. A*. **2012**, *109*, 2263.
25. A. Sreekumar, L.M. Poisson, T.M. Rajendiran, A.P. Khan, Q. Cao, J. Yu, B. Laxman, R. Mehra, R.J. Lonigro, Y. Li, M.K. Nyati, A. Ahsan, S. Kalyana-Sundaram, B. Han, X. Cao, J. Byun, G.S. Omenn, D. Ghosh, S. Pennathur, D.C. Alexander, A. Berger, J.R. Shuster, J.T. Wei, S. Varambally, C. Beecher, A.M. Chinnaiyan, *Nature* **2009**, *457*, 910.
26. G. Valenti, E. Rampazzo, E. Biavardi, E. Villani, G. Fracasso, M. Marcaccio, F. Bertani, D. Ramarli, E. Dalcanale, F. Paolucci, L. Prodi, *Faraday Discuss.* **2015**, *185*, 299.
27. (a) E. Biavardi, S. Federici, C. Tudisco, D. Menozzi, C. Massera, A. Sottini, G.G. Condorelli, P. Bergese, E. Dalcanale, *Angew. Chem. Int. Ed.* **2014**, *53*, 9183; (b) E. Biavardi, F. Ugozzoli, C. Massera, *Chem. Commun*. **2015**, *51*, 3426.
28. M. Sepaniak, P. Datskos, N. Lavrik, C. Tipple, *Anal. Chem.* **2002**, *74*, 568A.
29. J. Fritz, *Analyst* **2008**, *133*, 855.
30. J. Lagona, P. Mukhopadhyay, S. Chakrabarti, L. Isaacs, *Angew. Chem. Int. Ed*. **2005**, *44*, 4844.
31. J.M. Chinai, A.B. Taylor, L.M. Ryno, N.D. Hargreaves, C.A. Morris, P.J. Hart, A.R. Urbach, *J. Am. Chem. Soc.* **2011**, *133*, 8810.
32. M. del Pozo, P. Hernandez, L. Hernandez, C. Quintana, *J. Mater. Chem.* **2011**, *21*, 13657.
33. (a) M. Bush, N. Bouley, A.R. Urbach, *J. Am. Chem. Soc.* **2005**, *127*, 14511; (b) P. Rajgariah, A.R. Urbach, *J. Inclusion Phenom. Macrocyclic Chem.* **2008**, *62*, 251.
34. K.S. Crump, H.A. Guess, *Annual review of public health* **1982**, *3*, 339.
35. V. Gubala, L.F. Harris, A.J. Ricco, M.X. Tan, D.E. Williams, *Anal. Chem.* **2012**, *84*, 487.
36. J. Janata, *Principles of Chemical Sensors*, Plenum Press, New York, **1989**.
37. H. Hierlemann, R. Gutierrez-Osuna, *Chem. Rev.* **2008**, *108*, 563.
38. A. Ferancovà, E. Korgovà, J. Labuda, J. Zima, J. Barek, *Electroanalysis* **2002**, *14*, 1668.

39. H. Egan, Environmental Carcinogens. Selected Methods of Analysis. Vol. 4, Some Aromatic Amines and Azodyes in the General and Industrial Environment, *IARC*, Lyon **1981**.
40. I.K. O'Neil, L. Fishbein, *Intern. J. Environ. Anal. Chem*. **1986**, *26*, 229.
41. J.C. Moreira, J. Barek, *Quimica Nova* **1995**, *18*, 362.
42. The VOC solvent emission directive (Directive 1999/13/EC), *EUR-Lex*, European Union Publications Office.
43. S. Nag, L. Duarte, E. Bertrand, V. Celton, M. Castro, V. Choudhary, P. Guegan, J.-F. Feller, *J. Mater. Chem. B* **2014**, *2*, 6571.
44. M. Ozmen, Z. Ozbek, S. Buyukcelebi, M. Bayrakci, S. Ertul, M. Ersoz, R. Capan, *Sensor. Actuat. B-Chem.* **2014**, *190*, 502.
45. F.L. Dickert, U.P.A. Bäumler, G.K. Zwissler, *Synthetic Metals* **1993**, *61*, 47.

CHAPTER 2

Aliphatic amino acids recognitions with tetraphosphonate cavitands

Post-translational modifications (PTMs) such as histone lysine and arginine residues methylation have a large influence on gene expression being a fundamental part of the histone code.
Synthetic supramolecular receptors able to selectively recognized different degree of methylation and isomeric modification states such as asymmetric dimethylarginine and symmetric dimethylarginine for histone code analyses are in highly demand.
In this chapter the recognition properties of a Tetraphosphonate cavitands (Tiiii) towards mono and di-methylated arginine were explored via NMR and ITC studies.
Moreover, the ability of the Tiiii cavitand in sequestering monomethylated amines exploiting the multivalency approach was investigated. In particular, Tiiii grafted onto ferromagnetic nanoparticles was used to selectively obtain mono-methylated lysine residues onto proteins during the protein methylation reaction.
Finally, the Tiiii cavitand was grafted onto gold nanoparticles and its molecular recognition properties towards sarcosine and N-phenethylamine was tested via NOE pumping and STD experiments.



2.1 INTRODUCTION

The study of histones, their assembly and modification by enzymatic regulators has a pivotal role in the epigenetic field.
Histones, a set of alkaline proteins around which negatively charged DNA wraps itself thanks to the presence of phosphonate groups, are composed of lysine and arginine residues for around one fifth of their amino acids.
By the H-bond assisted interactions between histones and DNA, the double helix of DNA, on average 5 cm long in each human chromosome, can be compressed in few nanometers in diameter inside the cell nucleus leading to histones self-aggregation. Wrapping DNA twice around a spool composed of two copies each of the basic core histones, namely H2A, H2B, H3, and H4, leads to the formation of assemblies called nucleosomes, structural components of the chromatin.

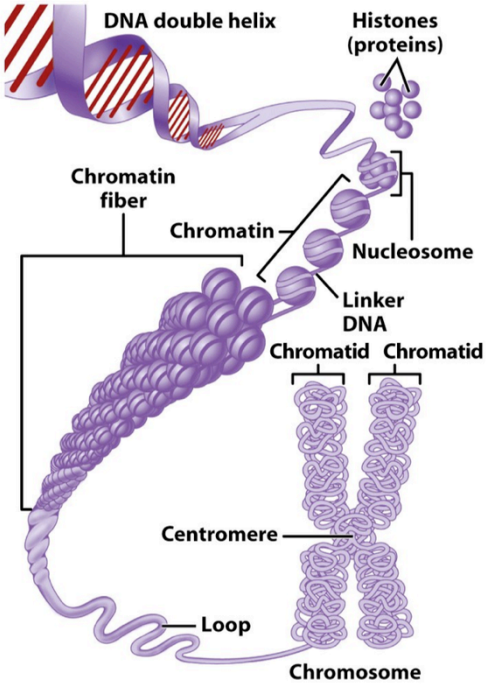


Figure 2.1. DNA at different packing levels forming nucleosome, chromatin and chromosome, respectively. Taken from *Principles of Anatomy and Physiology*, **2006**, ed. John Wiley and Sons.

Containing almost 2 meters of human genome, histones form a homogeneous and evolutionarily conserved group of structural proteins.
Numerous studies[1] have already highlighted the reciprocal impact between higher-order chromatin structure and gene activity, observing at several levels how specific histone modifications, namely post-translational modifications (PTMs), are intimately related to individual chromatin domains. Promoted by specific enzymes, PTMs commonly occurred on histones cationic N-terminal ($-NH_3^+$) tails[2] originating chemical modification of the amino acidic residues such as acetylation, methylation, phosphorylation and ubiquitation, just to name a few.
Not affecting the underlying DNA sequence, histone modifications represent a heritable change in phenotype. Histone code is able to dictate the chromatin environment and thus to regulate nuclear processes, such as gene transcription and regulation, replication, DNA repair, chromosome condensation (mitosis) and spermatogenesis (meiosis).[3]
As shown in **Table 2.1,** various PTMs are related to nearly all aspects of biological processes. For example, ubiquitination is implied in protein stability, cell division and immune response. The stability impairment of MTSS1, a suppressor of tumorous metastasis, by targeting it for ubiquitination, causes the decrease of MTSS1 expression, which has been observed to be associated with several types of aggressive forms of human cancers, such as breast and prostate cancers.[4]
Diversely, the impossibility to phosphorylate a mutated form of Thr-313 in (PTEN)-induced kinase 1 (PINK1), an important regulator of mitochondrial traffic, causes the suppression of PINK1 kinase activity and transcription, contributing to neurodegeneration of Parkinson pathogenesis and inducing neuronal cell death due to the inability of PINK1 in the regulation of mitochondrial trafficking in neurons.[5]
In particular, abnormal post-translational modifications to the histone code, influencing all the biological processes, lead to imbalances and dysfunctions related to the onset of human disorders such as tumors, cancers, diabetes, and neurologic and neurodegenerative diseases.[6]

	Phosphorylation	Glycosylation	S-Nitrosylation	Methylation	N-Acetylation	Palmitoylation	N-Myristoylation	Prenylation	Sumoylation	Ubiquitination
Apoptosis			✖		✖		✖			
Protein stability		✖	✖		✖				✖	✖
Protein-Protein interactions	✖	✖					✖	✖		
Protein-Membrane interactions						✖	✖	✖		
Protein Trafficking	✖	✖				✖				
Protein kinetic and thermodynamic		✖								
Protein activity		✖	✖							
Extracellular protein export							✖		✖	
Cell-signalling			✖			✖	✖		✖	
Transcription				✖	✖					
DNA-repair					✖				✖	
Cell cycle division	✖									✖
Immune response	✖									✖
Chromosome maintenance				✖						
Chromosome assembly									✖	

Table 2.1. PTMs and their related implication in biological function. Adapted from M. Audagnotto, M. Dal Peraro, *Comput. Struct. Biotechnol.* **2017**, *15*, 307.[7]

Among them, methylation of lysines and arginine emerged to be a complex and often elusive PTM. Arginine and lysine residues methylation can affect the regulation of transcription and repair functions.[3b]

In particular, lysine methylation involves the addition of one to three methyl groups on the lysine's ε-amine group, to form mono-, di- or tri-methyl lysine. Methylation on the tails of histone proteins, together with acetylation and phosphorylation, controls their interaction with other proteins, affects chromatin compaction and the regulation (up or down) of gene expression.

The methylation of arginine involves the addition of one or two methyl groups to the guanidinium group, forming mono- or di-methylarginine. This modification is known to affect RNA regulation and processing.[8]
Therefore, knowing how epigenetics influences genomic reprogramming during embryonic development represents the natural pursuance for new therapeutic approaches for anticancer research and regenerative medicine.[9]
For these reasons, extensive studies on the recognition on mono-methylated lysines and methylated arginines were conducted and reported in paragraph **2.2** and **2.3**, respectively.

2.2 LYSINE METHYLATION DEGREES

Because of its implications in epigenetics and chromosome maintenance, the study of the methylated lysine residues on histones is crucial for understanding the causes of activation or repression of genes during transcription and developing efficient therapies for diseases caused by these epimutations.[10]
In nature, methylation of the lysine side chains is catalyzed by lysine-methyltransferase (KMT) an enzyme of the histone methyltransferase class (HMT). Catalyzing the transfer of methyl groups from S-adenosylmethionine (SAM) to the lysine residues on histones H3 and H4 predominantly, KMTs can add multiple methyl groups to each lysine reaching pathological situations in which high level of methylated histones are generated. Being the major chromatin component, histones presenting a variable number of methylation states induce chromatin to remodel itself altering the gene expression. For instance, in histone H3, mono- and tri-methylation of lysine 4 (H3K4me and $H3K4me_3$) and mono-methylation of lysine 9 (H3K9me) are associated with gene activation[11] promoting the opening of chromatin and nucleosomes unwrapping, while $H3K9me_2$ and $H3K9me_3$ are implicated in gene repression[12] as they are associated with heterochromatin, the condensed part of chromatin which has no transcription activity but allows the packaging and ordering of DNA into nucleosomes.

The methylation states of various lysine residues have been extensively investigated, with a focus on the development of selective receptors for methylated lysine recognition, as they could be used to detect biomarkers associated to chronic pathological situations, like prostate cancer[13] and Alzheimer's disease.[14]
As far as molecular recognition is concerned, Nature has fulfilled a particular binding-motif for the recognition of the methylation states of lysine, called aromatic cage[15] (**Figure 2.2**). This latter is a pre-organized hydrophobic pocket surrounded by electron-rich aromatic residues that coordinate the methylated lysine, eventually with adjacent anionic residues.

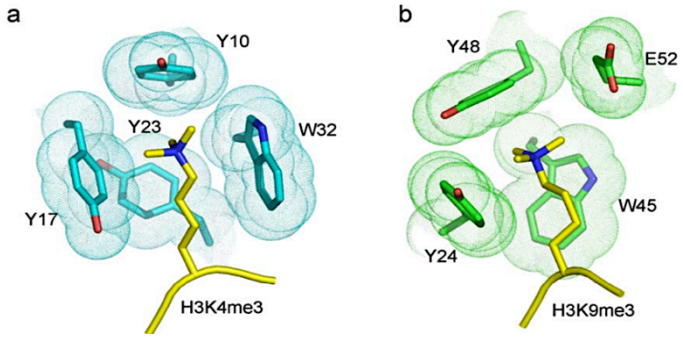


Figure 2.2. Aromatic pocket examples of (**a**) BPTF PHD finger[16] and (**b**) HP1 chromodomain[17] complexes with tri-methylated lysine H3K4me$_3$ and H3K9me$_3$, respectively. Adapted from P.R. Nielsen, D. Nietlispach, H.R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A.G. Murzin, N.V. Murzina, E.D. Laue, *Nature* **2002**, *416*, 103.[17]

Inspired by Nature, researchers tried to synthetically recreate this pre-organized cavity and in 2012 Hof[18], Nau[9e] and Crowley[19] independently published on this topic. The authors reported three studies in which a water soluble *p*-sulfonatocalix[4]arene was found to be able to bind tri-methylated lysine residues on proteins revealing a promising tool for detecting this histone methylation state.
Similarly, a sensor arrays composed by self-folding water soluble deep cavitands was proven to be able to recognize and detect tri-methylated lysines at different positions on peptide backbone even in complex mixtures.[20]

As for di-methylated lysine residues, in 2018 Crowley and co-workers demonstrated that the synthetic donut-shaped host cucurbit[7]uril (CB[7]) is able to complex di-methylated lysine residues onto a model protein. The complexation event and the involved interactions were demonstrated *via* NMR and X-ray crystallography.[21]

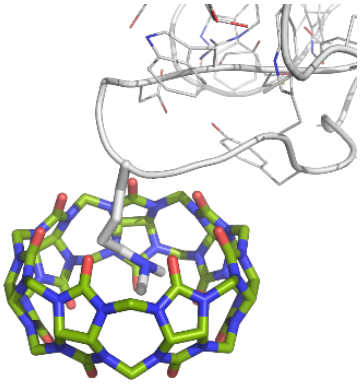


Figure 2.3. Complexation of CB[7] at K34me$_2$ in chemically di-methylated RSL. Taken from F. Guagnini, P.M. Antonik, M.L. Rennie, P. O'Byrne, A.R. Khan, R. Pinalli, E. Dalcanale, P.B. Crowley, *Angew. Chem. Int. Ed.* **2018**, *57*, 7126.[21]

The novelty of this research is enclosed in the possibility of studying the complexation of a synthetic receptor with a protein in the solid state as well as the capability of the CB[7] to recognize di-methylated lysine side chains by encapsulation of the di-methylammonium functional group together with entrapped water molecules.
The studies presented so far provided exquisite evidences on complexation of di-methylated and tri-methylated lysine side chains by synthetic receptors. Unfortunately, selectivity in recognizing different lysine methylated states has not been reported so far. In fact, as highlighted by Gamal-Eldin and Macartney,[22] the binding constants recorded in water for both calix[4] and CB[7] complexes are high for all the methylated lysine residues, hampering the possibilities to selectively recognize one methylated state over the other.
The only example present in the literature of a synthetic host able to recognize exclusively the mono-methylated lysine was presented by Dalcanale and co-workers in 2016, as reported in the next section.[23]

2.2.1 TETRAPHOSPHONATE CAVITANDS: THE LYSINE CASE

As already discussed in Chapter 1, the origin of the selectivity of tetraphosphonate cavitands (Tiiii) toward *N*-methylammonium salts derives from the presence of different synergic interaction modes: (i) $N^{+}\cdots O{=}P$ cation–dipole interactions; (ii) cation–π interactions between the $^{+}N–CH_3$ group with the π-basic cavity of the resorcinarene scaffold; (iii) two hydrogen bonds between two adjacent P=O bridges and the two nitrogen protons. Within the ammonium salt series, the simultaneous presence of all these interactions determine the selectivity of the receptor towards mono-methylated ammonium residues over di- and tri-methylated ones, for which the number of H-bond interactions decreased, while non-methylated ammonium ions are less favored in the binding because of the lack of cation–π interactions.[24]

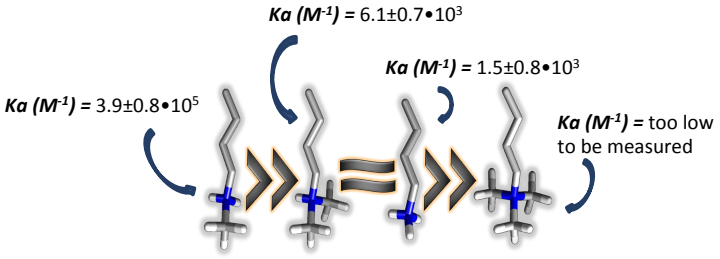


Figure 2.4. Methylated ammonium salts complexation trend. K_a are calculated *via* ITC.

In 2016 Dalcanale and co-workers conducted an extensive study *via* ^{1}H and ^{31}P NMR together with ITC and crystal structure analyses onto thirteen AAs using the Tiiii as receptor.[23] The Authors highlighted the non-covalent interactions involved in the binding event, the discrimination abilities of the receptor between different AAs as well as the solvent influence in the complexation selectivity.

In particular, the Authors focalized their attention onto the complexation ability of Tiiii towards mono-methyalted lysine. Complexation with *N*-methyl-L-lysine was found to be highly solvent dependent. By single crystal X-ray diffraction, the formation of a 2:1 (host:guest) complex in trifluoroethanol emerged (**Figure 2.5** *left*). The $-NH_3^{+}$ ammonium group of the lysine peptidic fragment forms H-bonds with the P=O groups of the

cavitand and a water molecule, whereas the methyl group on N_ε position is hosted in the cavity of an opposing cavitand. On the contrary, crystals grown from water resulted in a stoichiometric 1:1 complex, in which only the $^{+}NH_2{-}CH_3$ group is bound by the cavity (**Figure 2.5** *right*). The strong hydrophobic interaction between the *N*-methyl group of the guest and the π-basic cavity of the host allowed the ammonium hydrogens and the P=O groups to be closer, releasing the water molecules present in the cavity which were hydrogen bonded to the P=O groups.

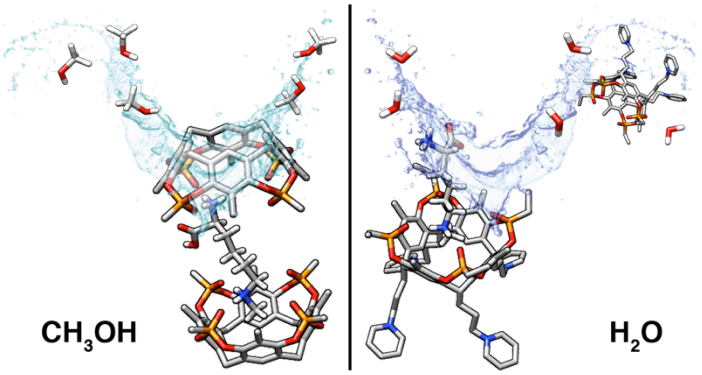


Figure 2.5. Crystal structures of Tiiii and *N*-methyl-L-lysine in MeOH (*left*) and water (*right*). Taken from R. Pinalli, G. Brancatelli, A. Pedrini, D. Menozzi, D. Hernandez, P. Ballester, S. Geremia, E. Dalcanale, *J. Am. Chem. Soc.* **2016**, *138*, 8569.[23]

These evidences were further confirmed *via* ITC, where two sigmoidal curves suggested the presence of more than one binding event when the analyses were carried out in MeOH. By analyzing the data, two different binding events were observed: the first one related to the formation of a 1:1 complex with an inflection point at the molar ratio of about 0.5 eq and a calculated K_a of $1.07 \bullet 10^6$ M^{-1}, and the second associable to the 2:1 complex obtained from the sigmoidal inflection point recorded near the molar ratio of 1 eq. In particular, the association constant of the second binding was found to be $1.15 \bullet 10^3$ M^{-1}, three orders of magnitude lower than the first one. In water, the titration data could be fitted to a simple 1:1 binding model. The calculated K_a in this case was $1.49 \bullet 10^3$ M^{-1}. The thermodynamic parameters obtained *via* ITC showed that, upon moving from methanol to water, the binding event changes from an

enthalpy−entropy driven process to an enthalpy driven-entropy opposed process.
These results matched up with those obtained through ^{1}H and ^{31}P NMR titrations both in CD_3OD and D_2O, and the calculated association constants were perfectly in agreement. The overall pictures emerging from these studies pointed out the crucial role of the mono-methylated ammonium ion. Its presence inside the cavity acts as a "hook" to boost both cation-dipole and H-bonding interactions, and it revealed to be necessary to ensure complexation in water.

The specificity demonstrated by tetraphosphonate cavitands in detecting mono-methylated lysine in water and PBS solutions, inspired Dalcanale and co-workers to transfer this specificity into detecting lysine mono-methylation in human histone H3 tails.[25]
Non-plasmonic Surface Enhanced Raman Scattering (SERS) was selected for the signal transduction coupling the cavitand receptors to resonators made of SiO_2/TiO_2 core/shell (T-rex) colloids. The Authors analyzed a library of five different peptides, identified as PC, P1, P2, P3 and P123, consisting of a sequence of 15 amino acids *N*-terminated with a 9-fluorenylmethyloxycarbonyl (FMOC) moiety as the labelling agent. Three lysine residues are present on each peptide in positions 4, 9 and 14 (Lys4, Lys9, and Lys14). Peptide PC presented no methylated lysine residues, and it was used as negative control. Peptides P1, P2 and P3 presented one mono-methylated lysine residue in positions 4, 9 and 14, respectively, while P123 displayed all the three lysine residues to be mono-methylated. The study showed that the complexation of the peptides on the T-rex surface functionalized with the cavitand receptors is dependent on the number of mono-methylated lysines present in the H3 histone tail. Peptide PC was not retained onto the T-rex surface while P123 is retained more strongly than the other mono-methylated histones thanks to the multiple binding modes.

2.2.2 REDUCTIVE AMINATION ON PROTEINS

Taking inspiration from the studies presented above and conducted in our research group, we decided to test the ability of the Tiiii cavitand in recognizing mono-methylated lysine residues moving from AAs or peptides directly onto proteins. To this purpose, proteins bearing mono-methylated lysine residues are necessary. Commonly procedures present in literature[26] do not lead to obtain exclusively mono-methylation of lysines, but on the contrary, are known to achieve near complete di-methylation of protein amino groups under non-denaturing conditions.
These conditions are represented by a reductive amination carried out in aqueous media and in presence of formaldehyde and a reducing agent, namely sodium cyanoborohydride or dimethylamino borane complex (DMAB). The mechanism of the reductive amination follows two steps: first the formation of the imine and then the subsequent conversion to the methylated species upon reduction. The formation of the di-methylated lysines is favored over the formation of the mono-methylated ones. In fact, after mono-alkylation, the resulting secondary amine is considerably more nucleophilic than a primary amine. Thus, free formaldehyde preferentially reacts with mono-methylated lysines, favoring the formation of the di-methylated species.

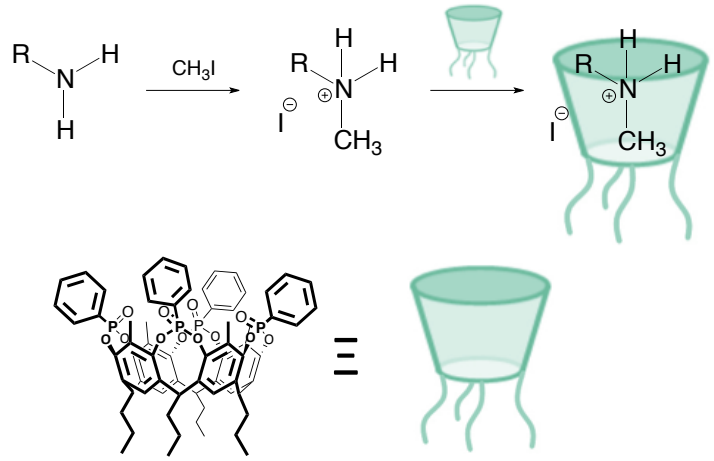
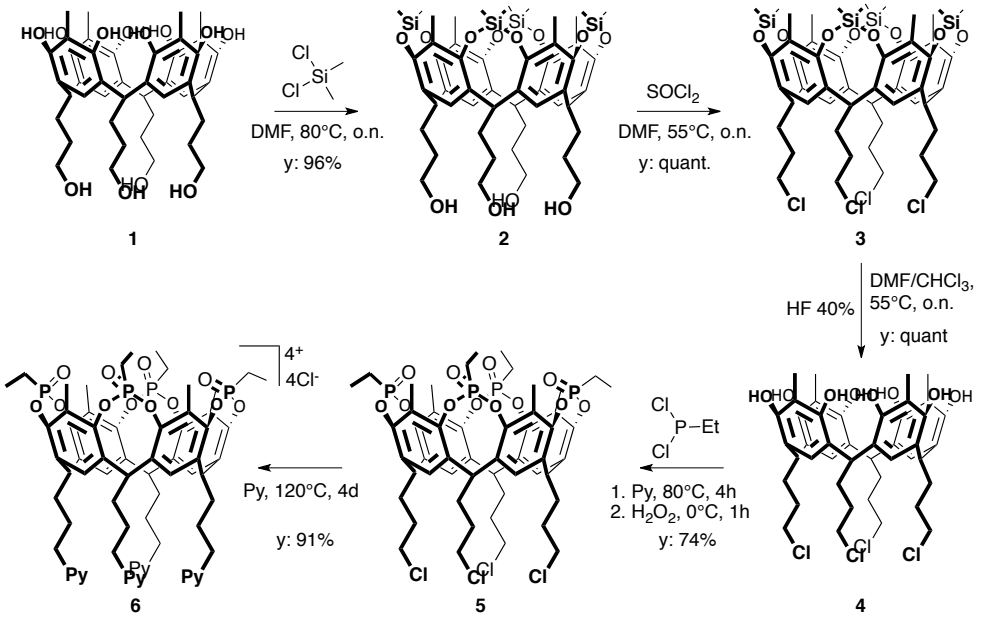


Figure 2.6. Schematic representation of *N*-mono-methylation reaction on primary amine using Tiiii cavitand as sequestrating agent.

To better control the per-methylation reaction on the protein, the ability of tetraphosphonate cavitand to act as captor of mono-methyl amines was explored, employing the receptor as sequestrating agent to prevent the further methylation of the lysine residues, as already proven by Yebeutchou and Dalcanale in 2009 (**Figure 2.6**).[27]

Hen Egg White Lysozyme (HEWL), which presents 5 lysine residues and one N-terminus, was employed as protein model system, while a water soluble tetraphosphonate cavitand was synthesized as receptor to perform the studies in water. This receptor, bearing four pyridinium moieties at the lower rim that provide water solubility, was obtained in five steps from resorcinarene **1**, with 65% overall yield, following a literature procedure (**Scheme 2.1**).[23]



Scheme 2.1. Synthetic pathway for water soluble cavitand **6**.

A selective protection of the phenolic groups *via* bridging reaction with dichlorodimethylsilane was firstly performed affording cavitand **2**. The subsequent chlorination reaction with $SOCl_2$ in presence of a catalytic amount of DMF afforded cavitand **3** in quantitative yield. The deprotection

of siloxane bridging groups with hydrofluoric acid gave resorcinarene **4**. Chlorinated resorcinarene **4** was reacted with dichloroethylphosphine in the presence of pyridine to give exclusively the tetraphosphonito cavitand, presenting the lone-pair on the P groups pointing inward the cavity. The subsequent *in situ* oxidation with hydrogen peroxide, which proceeded with retention of configuration at phosphorous center, led to cavitand **5** which presents all the four P=O groups pointing inward the cavity, as confirmed by ^{31}P NMR (see inset in **figure 2.7**). The final desired product was obtained *via* nucleophilic substitution of the four chlorides with pyridine[23,28] affording the water-soluble cavitand **6**, which was characterized by ^{31}P and ^{1}H NMR spectroscopy and HR-ESI spectrometry.

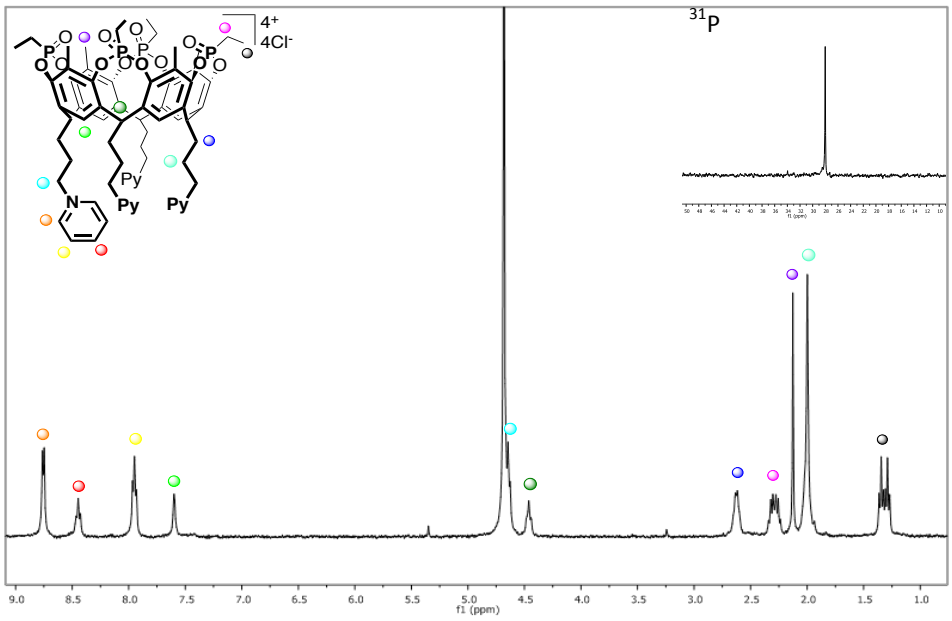


Figure 2.7. ^{1}H (400 MHz) and ^{31}P NMR (162 MHz) spectra of water soluble cavitand **6**. D_2O, 25 °C, 400 MHz.

In the ^{1}H NMR, the signals of the heteroaromatic pyridine protons in positions *ortho*, *para* and *meta* are present at 8.76 ppm, 8.44 ppm and 7.95 ppm, respectively (orange, red and yellow dots). The diagnostic methine triplet of the resorcinarene scaffold, displayed with a green dot, is present at 4.48 ppm, while the ethylene moiety attached to the

phosphorus atom is characterized by the four signals, due to the P-H coupling, centered at 2.29 ppm and 1.32 ppm, displayed with a pink and black dot respectively. Furthermore, the ^{31}P spectrum shows a single resonance at 28.02 ppm, indicating that only the isomer bearing all the four P=O groups oriented inward the cavity is present (inset).

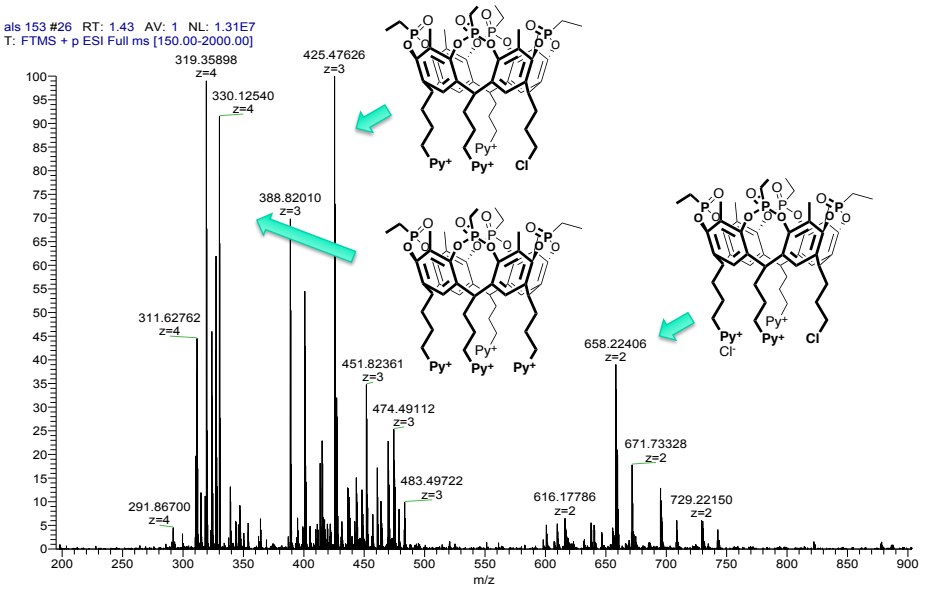


Figure 2.8. HR-ESI spectrum of water soluble cavitand **6** and its trisubstituted form.

The HR-ESI mass spectrum on the same sample further confirms the presence of the desired cavitand **6**, characterized by the presence of the tetracharged peak $[M-4Cl]^{4+}$ at m/z 330.12540 Da, together with the cavitand presenting only three pyridinium groups at the lower rim. This latter is present with two peaks: one related to the tri-charged species $[M-3Cl]^{3+}$ at m/z 425.47626 Da and the other related to the bis-charged species $[M-2Cl]^{2+}$ at m/z 658.22406 Da. No purification was attempted since also the cavitand presenting only three pyridinium groups resulted water soluble.

Cavitand **6** was then used as a mixture in the subsequent methylation experiments to test the ability of the receptor in preventing the per-methylation of the lysine residues. Reductive amination was thus

performed on Lysozyme. The reaction conditions envisaged the use of a stoichiometry amount of formaldehyde and cavitand with respect to the reactive amines present on the protein, thus 6 equivalents of both formaldehyde and cavitand.
The used protocol foresaw the addition of formaldehyde and cavitand to a 1.3 mM lysozyme solution in 0.1 M phosphate buffer (pH 5.7). This pH was selected since it was found that at pH higher than 6 di-methylation is preferred over mono-methylation. This is due to the strong correlation between lysines protonation/deprotonation kinetics and pK_a of the residues and the buffer used.[29] The addition of the reducing agent (15 eq) was performed after 1 h during which the reaction mixture was left to equilibrate. Finally, the reaction was left to proceed with gentle mixing for 6 hrs prior to be quenched with 0.1 M TRIS-HCl (pH 7.5), intensively dialysed against H_2O to get rid of the salts, concentrated and stored at 4 °C. For every trial, a blank experiment in which the cavitand was not added to the HEWL solution was carried out.

Several attempts to obtain the mono-methylation of the lysine residues on the lysozyme were performed changing type and/or amount of reducing agents, equivalents of formaldehyde, and reaction time, as reported in **table 2.2**. The obtained products were analyzed by HR-ESI spectrometry in order to firstly quantify the amount of methyl groups introduced onto the protein (Kme in **table 2.2**) disregarding the degree of methylation. From the obtained data, it was evident that changing parameters such as amount of formaldehyde and amount and type of reducing agent resulted to strongly affect the reaction, while the presence of the cavitand seems to have some effect only for entries **1** and **4**. In general, DMAB seems to inhibit the methylation reaction on the protein.

Batch	Sample	Tiiii (eq)	Red. Agent (eq)		Formaldehyde (eq)	Time	m/z	Kme
1	1[a]	0	$NaBH_3CN$	15	6	6 h	14477	12
	1	6	$NaBH_3CN$	15	6	6 h	14424	9
2	2[a]	0	DMAB	15	6	6 h	14346	3
	2	6	DMAB	15	6	6 h	14361	4
3	3[a]	0	DMAB	30	12	6 h	14388	6
	3	6	DMAB	30	12	6 h	14382	6
4	4[a]	0	DMAB	30	12	6 h	14370	5
	4	6	DMAB	30	12	6 h	14331	2
5	5[a]	0	$NaBH_3CN$	15	6	4 h	14388	6
	5	6	$NaBH_3CN$	15	6	4 h	14375	5

Table 2.2. Preliminary studies on mono-methylation reaction. [a] *blank experiment.*

In batch **1**, the reaction was carried out following a literature protocol[26a] using the same stoichiometry of formaldehyde with respect to reactive amine leading to a per-methylated Lysozyme, also in the presence of cavitand **6**. These results suggested to use the milder reducing agent DMAB instead of $NaBH_3CN$ (batch **2**). In this case, an incomplete methylation of the amine residues on the protein was obtained, as evinced from the ESI-MS analyses. For this reason, in the next experiment (batch **3**) the double of the reactants in comparison with batch **2** was used to boost the conversion. In this case, the ESI-MS showed that 6 methyl groups were added to the protein. This result being promising, batch **3** conditions were replicated in batch **4.** As shown in **table 2.2**, in this case the blank experiments resulted in 5 methyl groups added to the protein (sample 4[a]), while the experiment in which the cavitand was present (sample 4), resulted in the addition of only two methyl groups. Thus, it was not possible to replicate the results obtained in batch **3**.

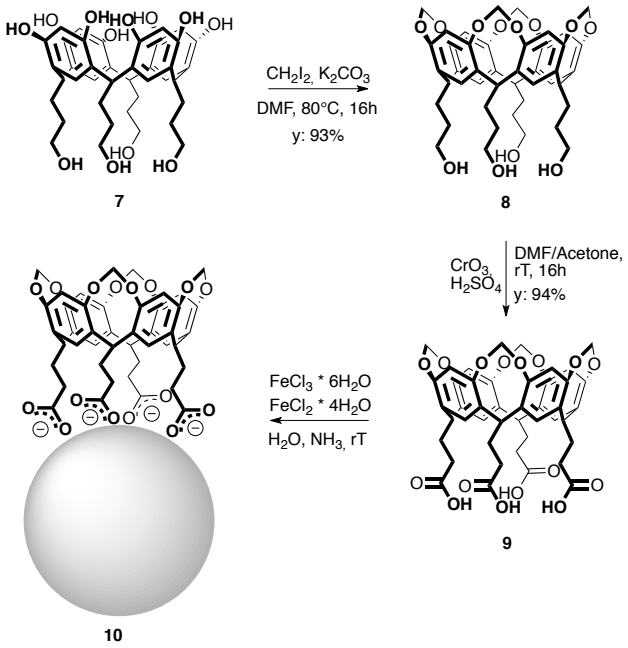
Batch **5** was set to test if and how the reaction time could influence the outcome of the reaction. By decreasing the reaction time from 6 hrs to 4

hrs maintaining the same ratio of reactants of batch **1** the grade of methylation was halved for both experiment 5[a] and 5.
Other experiments were then performed to find the best reaction conditions and reproducibility, but no significant results were obtained in term of reproducibility. Most of all, the presence of the cavitand seemed randomly affect the reaction outcomes.
The inefficiency of the cavitand to control the methylation degree of the amino groups can be imputable to the relatively low K_a to bind the mono methylated residue in water. For this reason, it was decided to exploit a multivalency approach by functionalizing biocompatible magnetite nanoparticles with tetrasphosphonate cavitands. Multivalent binding events lead to higher binding affinities and complex dynamic phenomena between the interacting partners. Such effects are also called cooperative, which refers both to the effects of true multivalency (chelate cooperativity) and to the mutual influence that neighboring binding sites exert on each other (allosteric cooperativity).[30] Many biological processes such as molecular recognition, catalysis, regulatory processes and signal transduction rely on the cooperative effects associated with the presence of multivalent interactions. Consequently, cooperativity is intensely sought for in synthetic systems, because of the huge potential it offers for advanced molecular receptors. The design and use of multivalent interactions in solution and at biological and artificial materials interfaces (e.g. at lipid membranes, or at self-assembled monolayer (SAM) systems, or nanoparticles) are of particular importance. At interfaces, multivalent interactions lead towards functions and applications that cannot be achieved otherwise. In fact, multivalency greatly increases the sensitivity of the particle-substrate interaction to external conditions, resulting in an ultra-sensitive and highly non-linear dependence of the binding strength on parameters such as temperature, pH or receptor concentration.

2.2.3. BIOCOMPATIBLE FERROMAGNETIC NANOPARTICLES

To maximize the binding action of Tiiii receptors towards the mono-methylated lysines during the reductive amination, it was decided to graft the receptor onto the surface of iron NPs (FeNPs) in order to have a greater number of cavitand molecules confined in the proximity of the reactive amino residue exploiting the multivalency feature of the system.
The choice of FeNPs lies in their unique physical and chemical properties extensively exploited in an eclectic range of disciplines, which include heterogeneous and homogenous catalysis, magnetic fluids, data storage, magnetic resonance imaging (MRI), environmental remediation such as water decontamination, and especially biomedicine and therapy applications, which represent important fields in medicine in general and in cancer therapy in particular.[31]
Moreover, the biocompatibility provided by Fe_3O_4 nanoparticulate systems has made possible their use with biological systems such as proteins. The ease in the functionalization of the FeNPs' surface exploiting small molecules like carboxylic acids moieties that can chelate FeNPs, made them even more attractive.

To this purpose, a model cavitand, namely tetramethylene bridged cavitand **9** (MeCav), functionalized at the lower rim with four carboxylic groups was synthetized to be grafted onto FeNPs to test the feasibility of the proposed approach. In **Scheme 2.2** the synthetic pathway is presented.



Scheme 2.2. Two-steps synthesis of the control system **10**.

Starting from the hydroxyl footed resorcin[4]arene **7**, the first synthetic step was the bridging reaction of the phenolic groups of the resorcinarene scaffold with diiodomethane in presence of K_2CO_3 as base and DMF as solvent. The obtained cavitand **8** was then oxidized with chromic anhydride and sulphuric acid (Jones reagent), leading to cavitand **9** bearing four carboxylic groups at the lower rim. All the obtained products were characterized through ^{1}H NMR and ESI-MS. FT-IR and thermogravimetric analysis (TGA) were also used to characterized MeCav **9**.

The ^{1}H NMR analysis of MeCav **9** reported in **Figure 2.9** shows the presence at 5.71 and 4.45 ppm of the diagnostic doublets belonging to the H_{out} (orange dot) and H_{in} (red dot) of the methylene bridges, respectively. The methine protons present as a triplet at 4.78 ppm (purple dot) together with two singles at 7.73 (light green) and 6.64 ppm (green) are diagnostic for the resorcinarene scaffold. The methylene next to the carboxylic acid is present as a triplet at 2.35 ppm (cyan dot), while the other protons of the alkyl chain are present as a multiplet at 2.68 (blue dot).

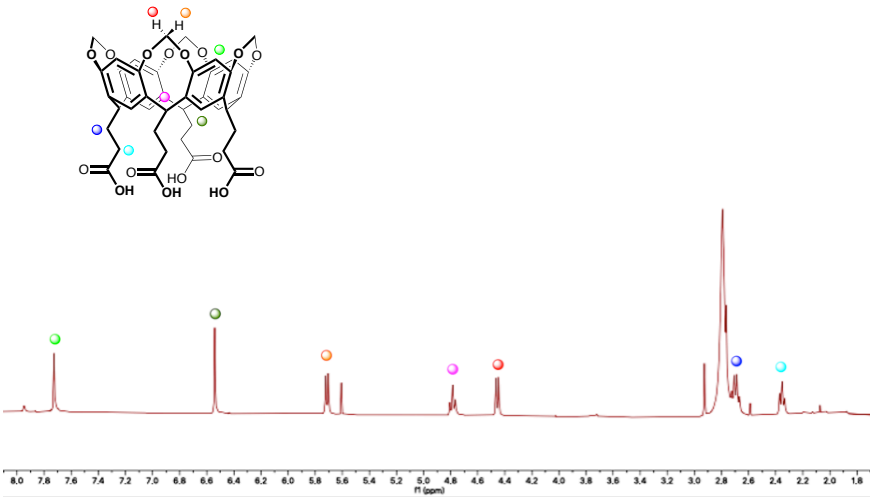


Figure 2.9. ^{1}H NMR spectrum of cavitand **9**. Acetone-d_6, 25 °C, 400 MHz.

Complete characterization of receptor **9** was performed by high resolution ESI mass spectrometry. The analysis confirms the nature of MeCav **9** thanks to the presence of the mono-charged ion $[M-H]^-$ at m/z 823.2247 Da and the di-charged ion $[M-2H]^{2-}$ at m/z 411.1085 Da (**Figure 2.10**).

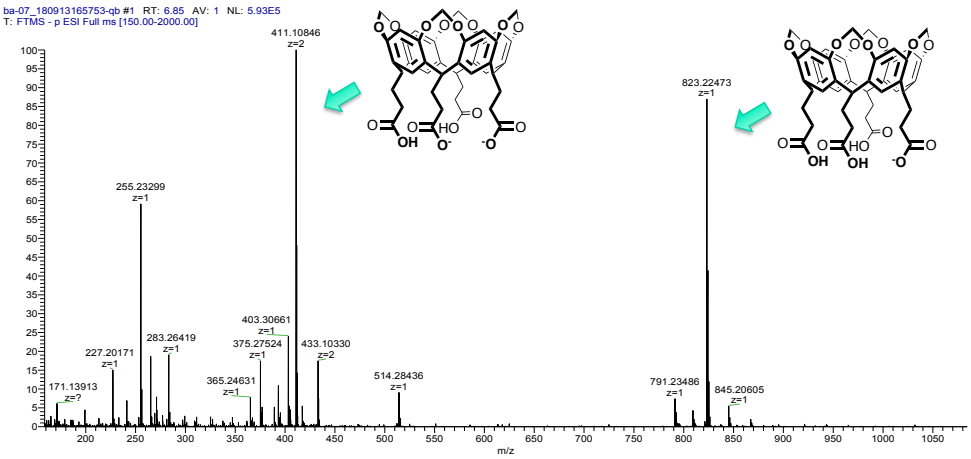


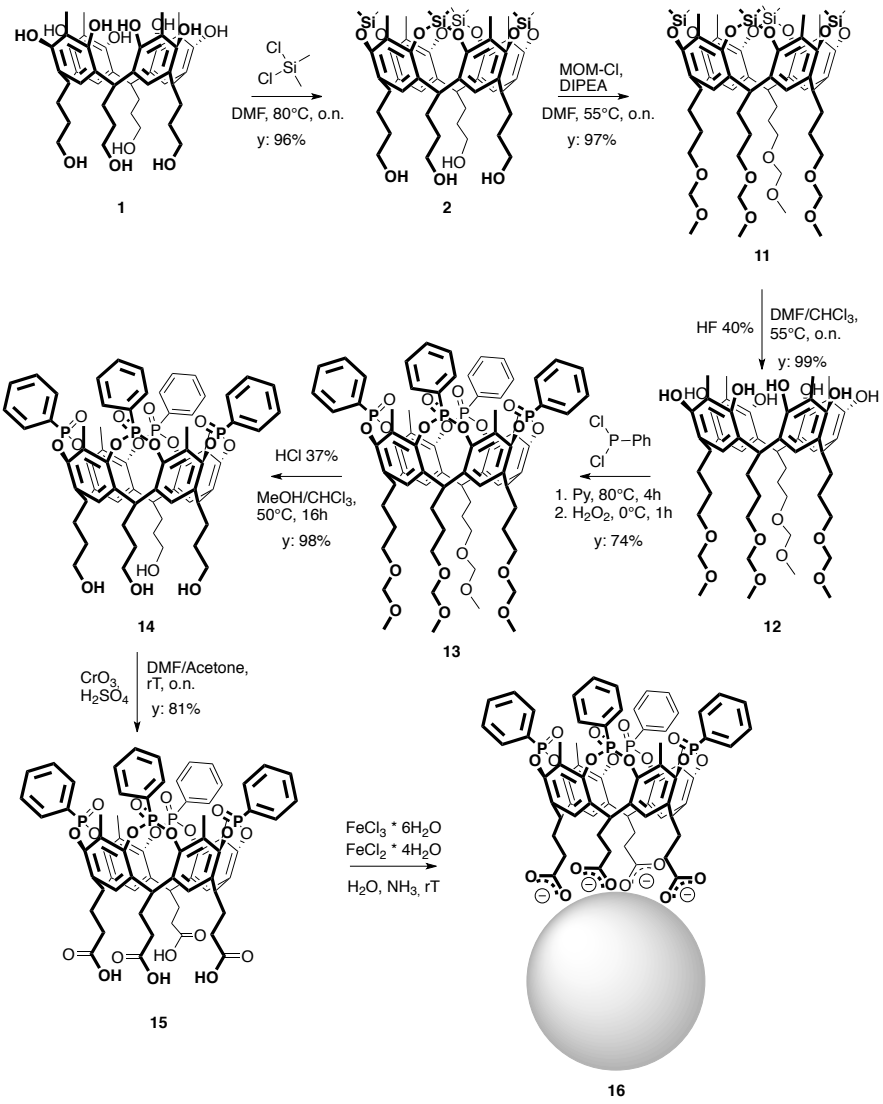
Figure 2.10. HR-ESI spectrum of cavitand **9**.

FT-IR and TGA analyses are reported in **paragraph 2.2.4** dedicated to the characterizations of the NPs.

The functionalization of the FeNPs with cavitand **9** was performed *via in situ* co-precipitation of $FeCl_3$ and $FeCl_2$ in presence of the receptor acting as surfactant to stabilize the FeNPs surface. With this method, Fe_3O_4 magnetite nanoparticles are obtained. In particular, **9** was suspended in water, then ammonia was added to allow the deprotonation of the carboxylic acid groups fostering the solubilization of the receptor in H_2O. A basic reaction environment is also necessary to boost the formation of nanoparticles. Rapidly and subsequently, a water solution of $FeCl_3 \bullet 6H_2O$ followed by a water solution of $FeCl_2 \bullet 4H_2O$ were added to the cavitand basic solution. The obtained suspension was centrifuged and washed with H_2O. MeCav@FeNPs **10** were finally separated by magnetic decantation and dried with an airflow. Full characterization of MeCav@FeNPs **10** is reported in the dedicated **paragraph 2.2.4**.

The steadiness of the MeCav@FeNPs system to the mono-methylation reaction conditions was then tested. The stability test was carried out in a buffer solution of NaH_2PO_4/Na_2HPO_4 at pH 5.7 to which formalin and the functionalized FeNPs were added. The suspension was left to equilibrate at room temperature for 1 h, and subsequently the reducing agent $NaBH_3CN$ was added. The suspension was placed on an oscillating platform for 16 hrs, after which it was centrifuged and washed with H_2O. Thermogravimetric analyses of MeCav@FeNPs and ^1H NMR of the supernatant were performed to confirm the stability of the system. The analyses results are reported in Appendix A.

Once proved the feasibility of the adopted synthetic approach, the target tetraphosphonate cavitand **15** (POCav) equipped at the lower rim with four carboxylic units was prepared in six steps with a 54% overall yield, starting from the readily available hydroxyl footed resorcinarene **1** as reported in **Scheme 2.3**.



Scheme 2.3. Synthetic pathway of POCav@FeNps system **16**.

The first synthetic step was the protection of the phenolic OHs in the presence of dichlorodimethylsilane, followed by the protection of the aliphatic OHs with chloromethylmethyl ether to obtain the fully protected cavitand **11**. This step is necessary since the subsequent bridging reaction with dichlorophenylphosphine partially chlorinates the hydroxyl groups. Deprotection of the phenolic hydroxyls was performed with an aqueous

solution of HF (40% wt) and resorcinarene **12** was obtained in almost quantitative yield. Resorcinarene **12** was then reacted with dichlorophenylphosphine in the presence of pyridine to obtain the tetraphosphonito cavitand. The tetraphosphonito intermediate was then oxidized *in situ* with oxygen peroxide to obtain cavitand **13** which shows all four P=O groups pointing inward the cavity. The desired POCav **15** was obtained after the deprotection of the hydroxyl groups at the lower rim of **13** in acidic methanol, followed by their oxidation using the Jones reagent in the same conditions employed for cavitand **9**. Cavitand **15** was characterized via NMR, ESI-MS, FT-IR and TGA.

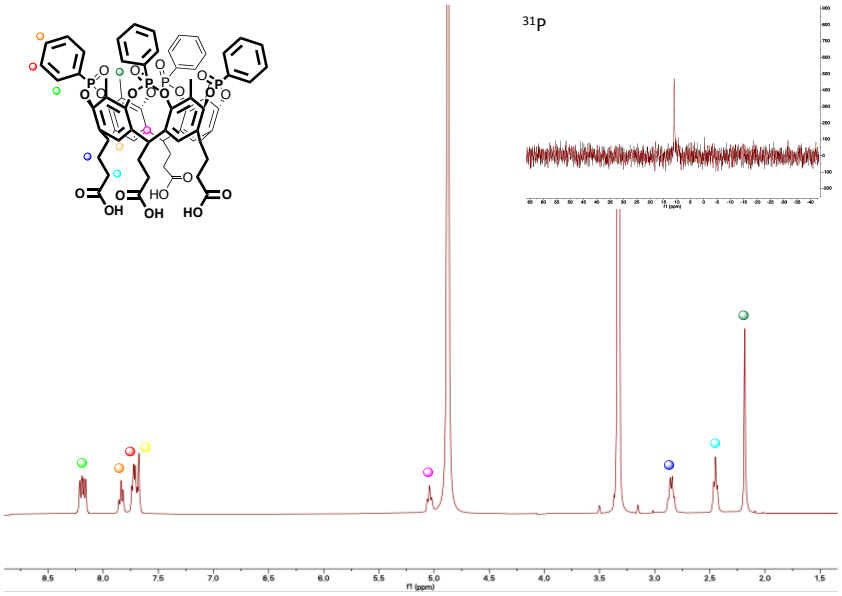


Figure 2.11. ^{1}H (400 MHz) and inset of ^{31}P (162 MHz) NMR spectra of cavitand **15.** CD_3OD, 25 °C, 400 MHz.

NMR (**Figure 2.11**) and HR-ESI (**Figure 2.12**) analyses confirmed the nature of the desired compound. As shown in the spectra below, the protonic resonances of **15** present all the characteristic peaks. In particular, displayed with a pink dot, the diagnostic methine protons rose as a triplet at 5.04 ppm, while in the aromatic region, from 7.67 to 8.20 ppm, the signals related to resorcinarene aromatic protons H_{down} and the phenyl groups attached to the phosphorus atoms are present, labelled with a

yellow, orange, red and light green dot, respectively. The protons of the methylene group adjacent to the carboxylic acid are present at around 2.5 ppm (cyan dot). ^{31}P NMR (inset) further confirms the presence of only one isomer with all the four P=O groups pointing inward the cavity, displaying only one peak at 11.03 ppm (inset).

The mass analysis confirms the nature of **15** as reported in **Figure 2.12**. The characteristic peaks of the molecule are present as mono-charged ion $[M-H]^-$ at m/z 1319.2542 Da and di-charged ion $[M-2H]^{2-}$ at m/z 659.1240 Da.

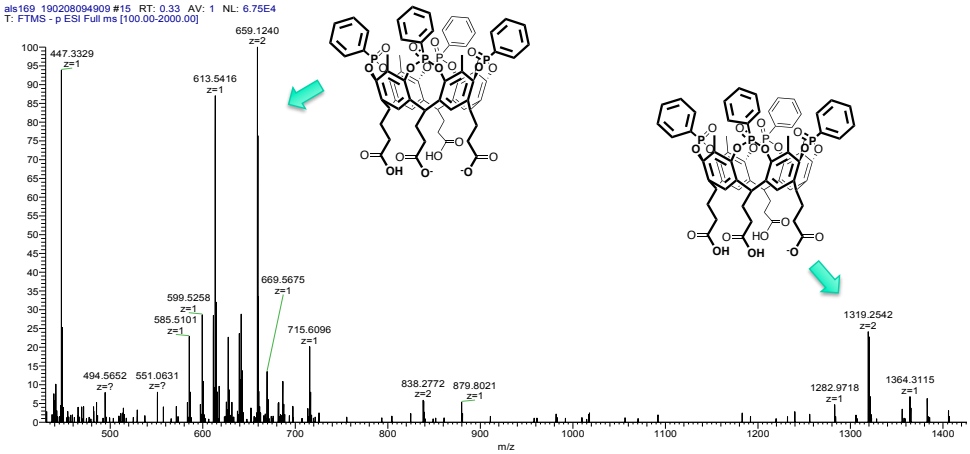


Figure 2.12. HR-ESI spectrum of cavitand **15**.

FT-IR and TGA analyses are reported in **paragraph 2.2.4** dedicated to the characterizations of the NPs.

POCav@FeNPs system **16** was synthetized *via* co-precipitation of $FeCl_3$•$6H_2O$ and $FeCl_2$•$4H_2O$ in presence of POCav **15** under basic conditions according to the procedure previously described. In particular, cavitand **15** was suspended in H_2O and NH_3 was added to allow the deprotonation of the carboxylic acids at the lower rim of the resorcinarene scaffold, favoring the solubilization of the cavitand in H_2O. Then $FeCl_3$•$6H_2O$ and $FeCl_2$•$4H_2O$ were rapidly added to the receptor solution and the obtained nanoparticles were separated by magnetic decantation and dried with an airflow. A full characterization of POCav@FeNPs **16** is reported in **paragraph 2.2.4**.

Also in this case, a stability test of POCav@FeNPs system to the mono-methylation reaction conditions was performed using the same protocol reported for MeCav@FeNPs **10**. The stability test was carried out in a buffer solution of NaH_2PO_4/Na_2HPO_4 at pH 5.7 to which formalin and the functionalized FeNPs **16** were added. The suspension was left to equilibrate at room temperature for 1 h, and subsequently 15 eq of the reducing agent $NaBH_3CN$ were added. The suspension was placed on an oscillating platform for 16 hrs, after which it was centrifuged and washed with H_2O. Thermogravimetric analyses of POCav@FeNPs and ^1H NMR of the supernatant were performed, confirming the stability of the system (see Appendix A).

2.2.4 CHARACTERIZATION OF THE FUNCTIONALIZED FeNPs

Bare and functionalized FeNPs were characterized by thermogravimetric analysis (TGA), X-ray photoelectron spectroscopy (XPS), and FT-IR.

Bare ferromagnetic NPs were prepared using a 1:2 ratio of $FeCl_2 \bullet 4H_2O$ and $FeCl_3 \bullet 6H_2O$ to obtain Fe_3O_4 magnetite NPs by co-precipitation of Fe(II) and Fe(III) in a solution of H_2O containing few drops of NH_3 0.7 M ammonia is required to achieve basic conditions ($9 < pH < 14$) necessary to obtain the formation of nanoparticles.[32]

$$Fe^{2+} + 2Fe^{3+} + 8OH^- \rightarrow Fe_3O_4 + 4H_2O$$

Scheme 2.4. Chemical reaction for the synthesis of magnetite via co-precipitation.

The nanoparticles thus obtained were washed with $HClO_4$ 2M, collected by centrifugation and subsequently dried with an airflow. Bare FeNPs were characterized by TGA and XPS. Moreover, they were characterized via powder X-ray diffractometry (XRD) to determine the particle size.

In particular, the diffraction peaks present in the XRD spectrum perfectly reflect the pattern of standard Fe_3O_4. Through the Debye-Scherrer equation the average size of the nanoparticles was calculated, which appears to be in a range between 5 and 15 nm, in line with the data

reported in the literature.[33] No preferential orientation was observed, since the ratio between the simulated peaks relative to the space group of the magnetite and the experimental ones appeared identical and superimposable, highlighting how the crystalline material has grown uniformly during the synthetic process.

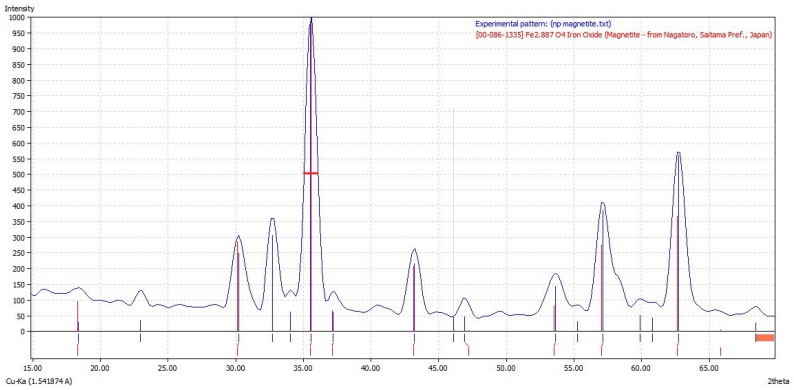


Figure 2.13. XRD spectrum of bare FeNPs.

FT-IR spectrum presents the characteristic peak of the OH stretch at around 3000 cm^{-1} (**Figure 2.15 a**).
TGA thermogram in **Figure 2.14** shows two weight losses: one of 2.3% at 89 °C due to H_2O evaporation and a second one of 2.0% at 272 °C related to DMF evaporation. FeNPs demonstrated to be stable up to 800 °C.

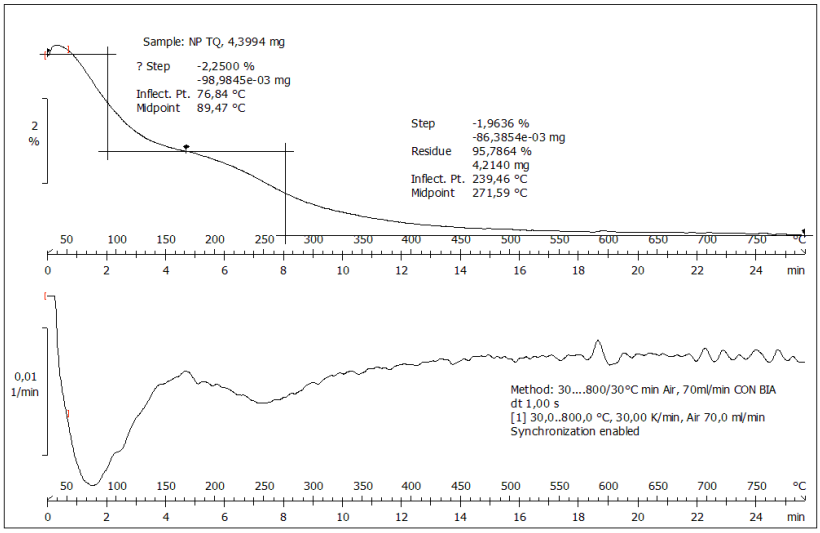


Figure 2.14. TGA thermogram of bare FeNPs.

The functionalized FeNPs, namely MeCav@FeNPs **10** and POCav@FeNPs **16**, were characterized as well by FT-IR to prove the effective grafting of cavitands **9** and **15** onto the surface of FeNPs. A comparison between the FT-IR spectra of the bare FeNPs, cavitands **9** and **15**, and the corresponding functionalized FeNPs **10** and **16** is reported in **Figure 2.15**.

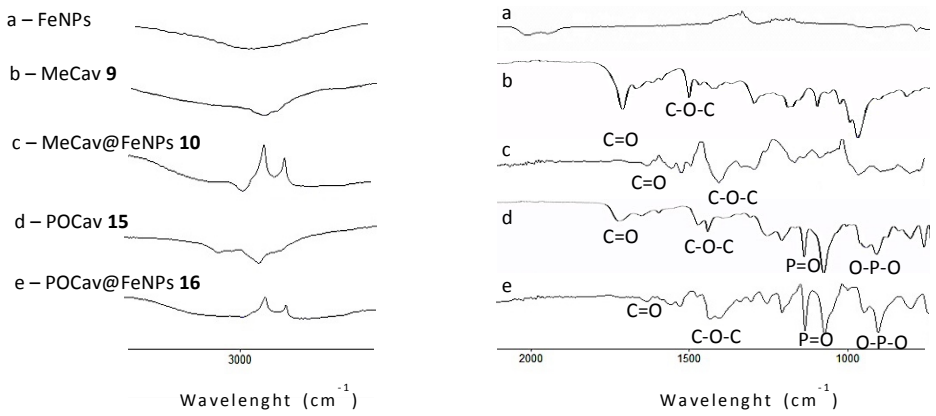


Figure 2.15. FT-IR spectral regions in the 3000 cm^{-1} (*left*) and 2000–800 cm^{-1} (*right*) ranges of (a) bare FeNPs, (b) MeCav **9**, (c) MeCav@FeNPs **10**, (d) POCav **15**, and (e) POCav@FeNPs **16**.

FT-IR spectrum of MeCav@FeNPs **10** (trace c) shows the presence of C=O and C-O-C stretches of the carboxylic groups decorating the lower rim of the cavitand. The peaks are shifted with respect to cavitand **9** (trace b) since carboxylic ions chelating the NPs' surface are now present, thus confirming the occurred functionalization of the NPs.
FT-IR spectrum of POCav@FeNPs **16** shows the presence of two strong bands at 1150 and 950 cm^{-1} related to P-O-C stretches of POCav **15**, and two peaks at 1205 cm^{-1} and 902 cm^{-1} due to the P=O and O-P-O stretching vibrations, respectively, thus confirming the presence of **15** onto the NPs' surface. Moreover, C=O and C-O-C stretches, slightly shifted respect to the pure cavitand (trace d), are present at 1650 cm^{-1}, 1580 cm^{-1} and 1450 cm^{-1}, thus confirming the functionalization of the NPs's surface with **15**.

Thermogravimetric analysis (TGA) was used to determine the degree of functionalization of the NPs surface.
Firstly, TGA analysis was carried out on MeCav **9**, showing that the cavitand is stable up to about 400 °C. Two important deflections, at 409 °C corresponding to a mass loss of 47.5% and at 458 °C corresponding to a mass loss of 42.2%, both attributable to the cavitand degradation, are present in the thermogram (see Appendix A).

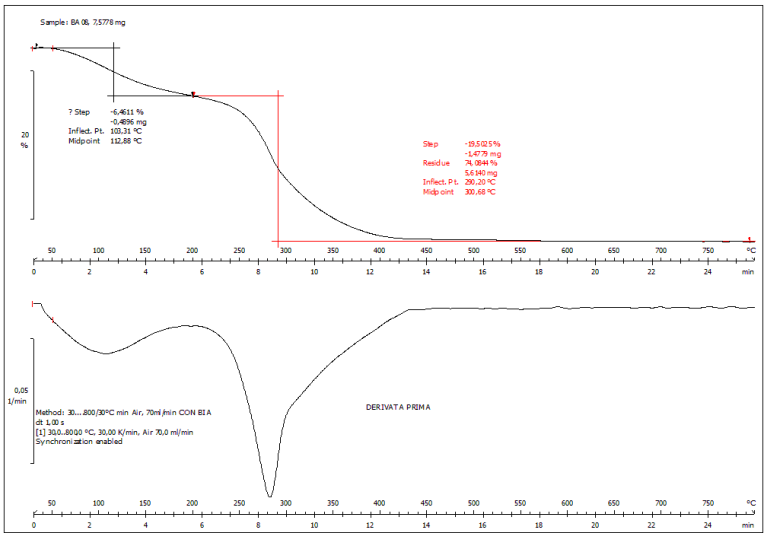


Figure 2.16. TGA thermogram of MeCav@FeNPs **10.**

The system MeCav@FeNPs **10** was investigated using the same technique. An initial weight loss of 6.5% was recorded at 113 °C, attributed to the evaporation of the solvent in the first layer of the NPs,[34] and a second loss of 19.5% was observed at 301 °C, probably due to the degradation of cavitand **9** (**Figure 2.16**) since no other degradation peaks were observed. Thus, a functionalization estimated at approximately 19% w/w of the FeNPs surface was revealed.
TGA thermograms of MeCav **9** and MeCav@FeNPs **10** after the stability test were reported in Appendix A.

Similarly, also POCav **15** showed in its TGA thermogram to be stable up to about 400 °C. Two important deflections, at 402 °C corresponding to a mass loss of 53.7% and at 438 °C corresponding to a mass loss of 36.4%, both attributable to cavitand degradation, are present in the thermogram (see Appendix C). POCav@FeNPs **16** thermogram, was compared with the previous one of POCav **15**, showing a small weight decrease at 119 °C related to the solvent and a weight loss equal to 18.5% at 304 °C due to the degradation of the grafted POCav **15**. Also in this case, the degradation of the cavitand grafted onto the NPs' surface occurred at lower temperature with respect to the cavitand alone. These data revealed a functionalization estimated at approximately 18% w/w (**Figure 2.17**).

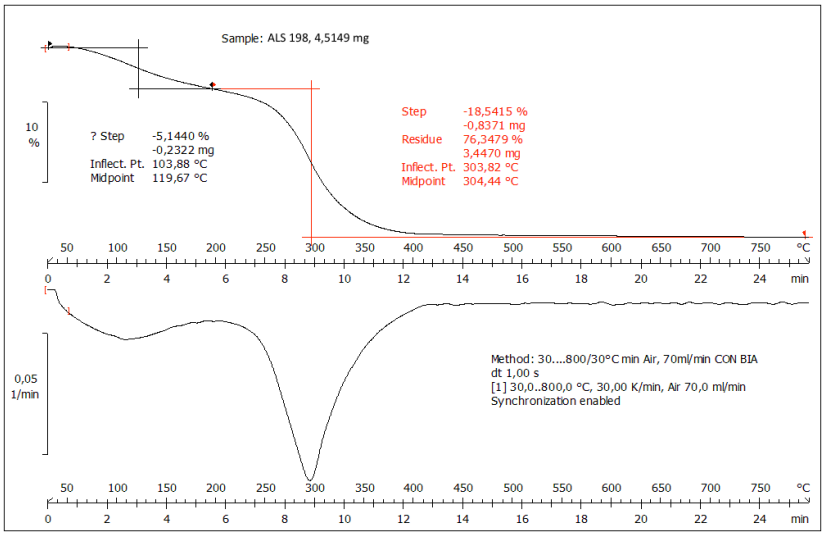


Figure 2.17. TGA thermogram of POCav@FeNPs **16.**

XPS analyses were performed on bare FeNPs (**Figure 2.18**), MeCav **9** (**Figure 2.19**) and MeCav@FeNPs **10** (**Figure 2.20**), POCav **15** (**Figure 2.21**) and POCav@FeNPs **16** (**Figure 2.22**) to further confirm the presence of the cavitands onto the NPs' surface. The analyses were carried out by Dr. Roberto Verucchi at the Institute of Materials for Electronics and Magnetism in Trento, who is kindly acknowledged.
The presence of iron, carbon, oxygen and sodium, without significant binding energy (BE) shifts, was detected in bare FeNPs XPS spectra. Stoichiometry analysis of the data led to 11.01%, 50.96% and 38.03% of atomic percentage contribution for Fe, C and O, respectively.

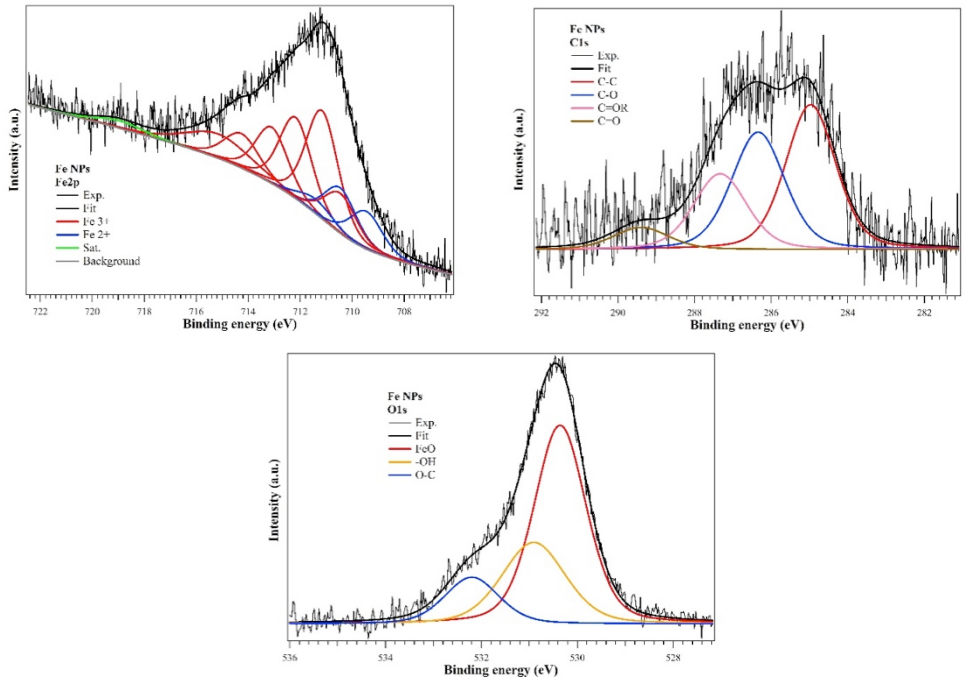


Figure 2.18. XPS spectra of the bare FeNPs. Fe 2p (*upper right*) C 1s (*upper left)* and O 1s (*below*).

Even if Fe 2p core level interpretation is still a matter of debate[35] due to the presence of a multiplet structure of core 2p vacancy levels in transition metals, a well established and accepted core level interpretation has been achieved for iron[36] exploiting theoretical models.[37] The presence of specific fingerprints in Fe 2p 3/2 core level analysis allowed to identify the

presence of specific oxides. In Fe 2p spectrum (**Figure 2.18**), the main peak located around 711eV represents a value typical for Fe_3O_4 and Fe^{3+} oxides (FeOOH). A satellite at 718.96eV (+ 8eV with respect to the main peak) is instead typical of Fe^{3+} compounds but not of Fe_3O_4.[36a] Thus, starting from Fe_3O_4 model, Fe 2p 3/2 peak was deconvoluted considering Fe^{3+} and Fe^{2+} multiplet splitting. The Fe^{3+}/Fe^{2+} ratio was then calculated to be 3.85, far from the expected 2, confirming the presence of Fe^{3+} species, probably FeOOH. In fact, from the Fe/O stoichiometry ratio, bare FeNPs resulted to be composed by 62% of Fe_3O_4 and 38% of FeOOH.
The presence of carbon in the sample, due to adventitious oxygen and carbon,[38] is demonstrated by C-C, C-O, C=OR and C=O species in C 1s and C-O/C=O contributions in O 1s core levels (**Figure 2.18**).

The presence of carbon and oxygen in atomic percentage of 72.4% and 27.6% was observed in the XPS spectra of MeCav **9**. The calculated C/O ratio equal to 2.6 appeared to be in agreement with the theoretic 2.75. C 1s core level showed the presence of C-C bonds both for aliphatic and aromatic residues of the resorcinarene scaffold together with O-C-O and C=O bonds related to the carboxylic acid moieties at the lower rim (**Figure 2.19** - right). This latter presented a BE peak at 289.6eV at a lower value than expected[39] probably due to formation of H bonds with the alkyl chains. O 1s core level, composed by C=O, C-O-C and C-OH species, showed the typical BEs in line with the number of emitting oxygen atoms.

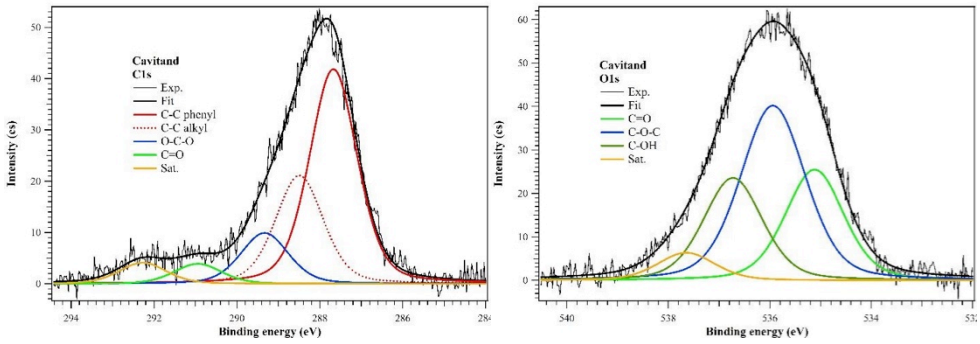


Figure 2.19. XPS spectra of the MeCav **9**. C 1s (*right)* and O 1s (*left*).

In contrast with bare FeNPs, the functionalized MeCav@FeNPs **10** resulted to be mainly composed by Fe_3O_4, with a good Fe^{2+}/Fe^{3+} ratio of 1.44 with

respect to the theoretical value of 2. The Fe/O stoichiometry ratio obtained (0.68 calculated; 0.75 theoretical) allowed us to say that less defects are present on the surface of the functionalized FeNPs **10** than the bare ones. This is probably due to the presence of MeCav **9** that, acting as a surfactant, prevent the formation of adventitious carbon on the FeNPs surface. Moreover, this latter, being present in smaller quantities (about 50%) with respect to the bare NPs, indicates that cavitand **9** is grafted on about 50% of the free surface, which cannot be covered by contaminants. The XPS analysis confirmed the presence of $-COO^-$ groups, due to the bond established between the iron of the NP and the carboxylate group of the cavitand, as evidenced both by the signal referred to C 1s with a BE peak at 289.5eV as the one referred to O 1s at 534.7eV (**Figure 2.20**). Finally, the comparison of Fe/C ratio concedes the presence of cavitand molecules not only on the surface of the iron nanoparticle but also in their first layer. This result is imputable to the co-precipitation synthetic technique used for the formation of the studied system.

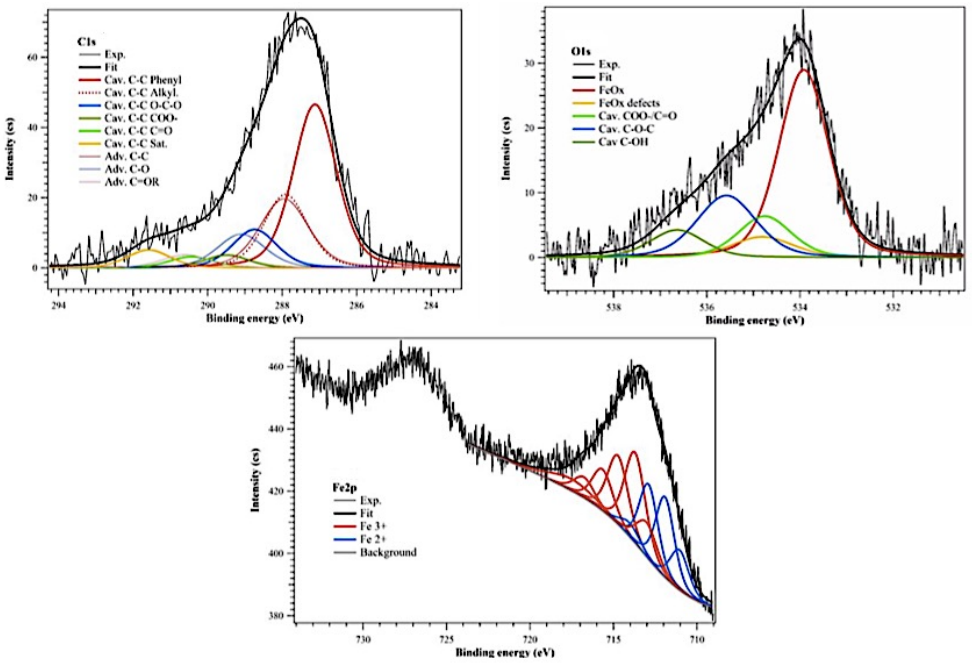


Figure 2.20 XPS spectra of MeCav@FeNPs **10**. C 1s (*upper right*), O 1s (*upper left*) and Fe 2p (*below*).

POCav **15** showed in its XPS spectra (**Figure 2.21**) the presence of carbon, oxygen and phosphorus in atomic percentage of 76.24%, 19.33% and 4.43% in very good agreement with the expected 73.91%, 21.74% and 4.35% values. All core levels show a ≈+3.2eV BE shifts, typical of insulating materials. C 1s core level presents C-C bonds in phenyls, alkyl and meso positions, besides C-P and C-OOH bonds with a specific weight corresponding to the number of emitting carbon atoms and very close to the expected values. Similarly to MeCav **9**, a weak satellite peak at very high BE imputable to a shake-up transition typical of π-conjugated systems[40] is present together with a lower BE in C-OOH peak related to the formation of H-H bonding among the four alkyl chains ending with carboxylic groups. O 1s core level is composed by four components belonging to P=O, C=O, O-P-O and C-OH species, with BEs typical of P-O bonds.[41] Finally, P 2p core level showed a single doublet located at about 133.6eV (corrected value), in good agreement with similar organic compounds.[41,42]

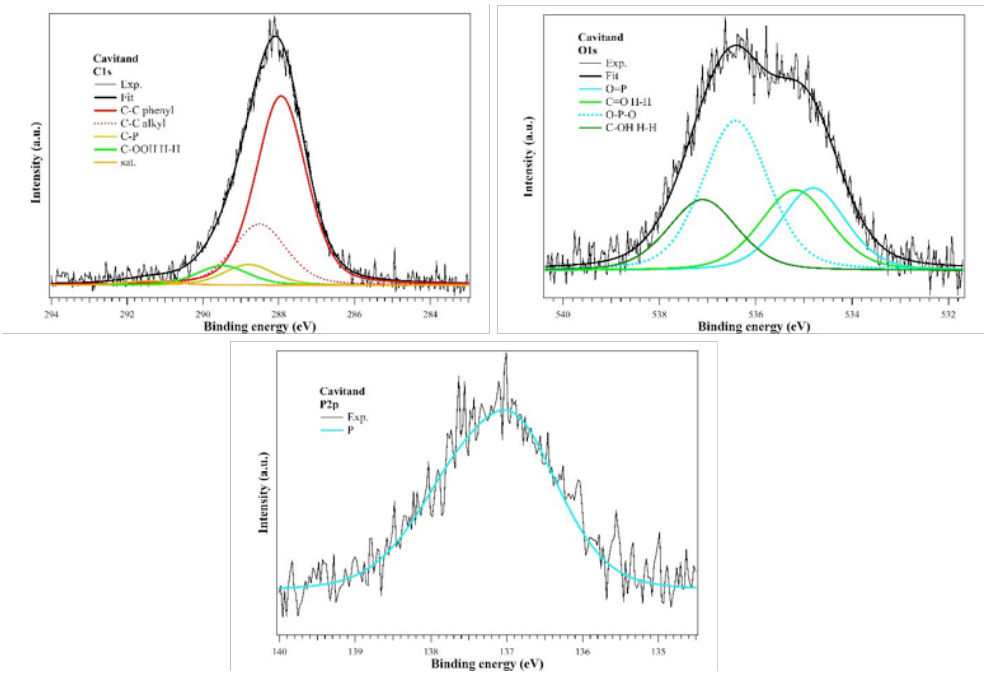
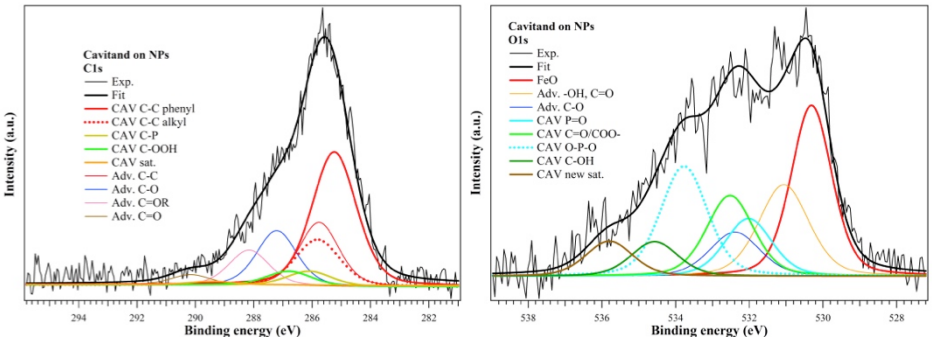


Figure 2.21 XPS spectra of POCav **15**. C 1s (*upper right*), O 1s (*upper left*) and Fe 2p (*below*).

From XPS analyses on POCav@FeNPs **16** (**Figure 2.22**) it can be immediately observed the typical fingerprints of both FeNPs and POCav **15** in C 1s and O 1s core levels, suggesting the successfully functionalization of the inorganics counterpart with the organic one. In addition, a core levels BE shift of about +0.6eV, probably due to the insulating properties of POCav **15** partially suppressed by the iron oxide NPs, can be observed.
Being composed by 60% of Fe_3O_4 and 40% of FeOOH, POCav@FeNPs **16** possessed a more stable composition with respect to MeCav@FeNPs **10**.
Even if a highly reduced Fe2p signal was obtained, similar iron oxide properties previously found for bare FeNPs were observed.
Typical fingerprints of POCav **15** are present. C 1s core level, composed by the same components of the isolated POCav **15**, presents a slightly large BE (+0.2eV) related to COOH group suggesting the breaking of some H-H bonds. This data is supported by a significant increase of 532eV (corrected BE) peak in O 1s core level related to the formation of C-OO^- groups that demonstrated the formation of bonds between iron oxide NPs and the carboxylic groups of POCav **15**. Analyzing the relationship between P and the carbon- and oxygen-based species, the concentration ratio between P and Fe was calculated to be 0.61, comparable with the theoretical one, and with P% = 2.17 and Fe% = 3.58. From the comparison with MeCav@FeNPs **10**, in which the ratio between C% and Fe% is 45.53 / 9.07 = 5.02, in the system POCav@FeNPs **16** the ratio between C% and Fe% is 41.67 / 3.58 = 11.6. Having duplicated the value of the C/Fe (%) ratio, a more efficient functionalization in term of coverage of the FeNPs with the POCav **15** was obtained and demonstrated.



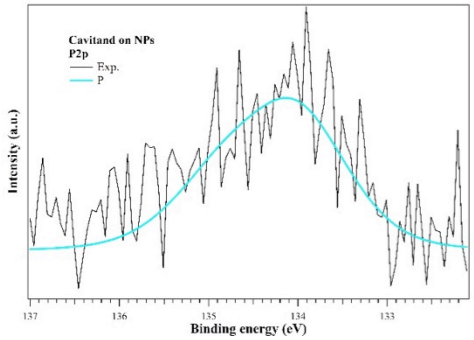


Figure 2.22 XPS spectra of POCav@FeNPs **16**. C 1s (*upper right*), O 1s (*upper left*), P 2p (*below*).

On the basis of the overall XPS, TGA and FT-IR results, it could be concluded that cavitands **9** and **15** have been successfully grafted on FeNPs. FeNPs are functionalized with an average number of three cavitands molecules, which are grafted onto the FeNPs with three or less carboxylic groups.

2.2.5. PROTEIN METHYLATION WITH FUNCTIONALIZED FeNPs

MeCav@FeNPs **10** and POCav@FeNPs **16** systems were used to investigate the selective monomethylation of the lysine residues on the surface of Lysozyme exploiting the multivalency approach. The reaction was carried out in the same experimental conditions illustrated in **paragraph 2.2.2**, modifying the amount of formaldehyde and reducing agent in order to find the best reaction conditions.
Formally, to 1.3 mM HEWL in 0.1 M phosphate buffer pH 5.7, 6 eq of CH_2O were added, followed by 6 eq of FeNPs. The reaction was left to equilibrate at rT. After 1 h, 15 eq of $NaBH_3CN$ were added and the reaction was left to proceed for 6 hrs at rT. The reaction was quenched with 0.1 M TRIS buffer pH 7.5, and left to equilibrate for 30 minutes.
The obtained products were analyzed by HR-ESI spectrometry in order to quantify the amount of methyl groups introduced onto the protein (Kme in **table 2.3**).

Three different systems were used for each batch: (a) a blank experiment in which only the protein is present, (b) the control system, namely MeCav@FeNPS **10**, able to bind mono-methylated ammonium salts only through CH_3-π interactions, and then less efficient with respect to the tetraphosphonate cavitand and (c) the working system POCav@FeNPS **16** which offers multiple synergic interactions with the monomethylated species.
In **table 2.3** preliminary experiments are reported.

Batch	Sample	FeNPs (eq)*	$NaBH_3CN$ (eq)	Formaldehyde (eq)	Time	m/z	Kme
1	1[a]	0	15	5	6 h	14268	1
	1[b]	6	15	5	6 h	14275	2
	1[c]	6	15	5	6 h	14269	1
2	2[a]	0	30	12	6 h	Protein denaturation	
	2[b]	6	30	12	6 h		
	2[c]	6	30	12	6 h		
3	3[a]	0	15	6	6 h	14516	12
	3[b]	6	15	6	6 h	14491	12
	3[c]	6	15	6	6 h	14448	10

Table 2.3. Mono-methylation reaction conditions. [a] *blank experiment;* [b] *control experiment with MeCav@FeNPs* ***10****;* [c] *experiment with working system POCav@FeNPs* ***16****.*
[*] *The amount of FeNPs was calculated on the MeCav or POCav average covering the FeNPs surface.*

A first attempt (batch **1**) was performed using 5 eq of formaldehyde with respect to the 6 reactive lysines residues present on the protein in order to enhance mono-methylation. As shown in **table 2.3**, we obtained the insertion of maximum 2 methyl groups on the protein, without any significant difference between the blank, the control and the working system.

In the next experiment, the amount of formaldehyde was then doubled with respect to the amount of amino groups (batch **2**) to boost the protein methylation. Unfortunately, denaturation of the protein by cross-linking *via* cysteine residues occurred, formig high molecular weight aggregates which did not allow the resolution of the peaks present in the HR-ESI spectrum.
Given the results obtained, in batch **3** the equivalents of formaldehyde were used in a stoichiometric amount with respect to the amino residues present on the HEWL protein. In this case, the number of methyl groups inserted onto the protein ranged from 10, when POCav@FeNPS was used, to 12 for both the blank and the control experiments. Even if the number of methyl groups inserted is less when the PO cavitand is present, it is not possible to assert that the tetrasphosphonate receptor exploited its sequestrating action towards mono-methylated lysines, because a part of the amino acid residues resulted to be per-methylated.

Further experiments were set varying the reaction time (**Table 2.4**).

Batch	Sample	FeNPs* (eq)	$NaBH_3CN$ (eq)	Formaldehyde (eq)	Time	m/z	Kme
4	4[a]	0	15	6	8 h	14499	14
	4[b]	6	15	6	8 h	14524	16
	4[c]	6	15	6	8 h	14444	10

Table 2.4. Mono-methylation reaction conditions increasing the reaction time. [a] *blank experiment;* [b] *control experiment with MeCav@FeNPs* ***10****;* [c] *experiment with working system POCav@FeNPs* ***16****.* [*] *The amount of FeNPs was calculated on the MeCav or POCav average covering the FeNPs surface.*

A new batch (batch **4**) was prepared using the same amount of reagents used in batch **3**, (i.e. 6 eq of formalehyde, 6 eq of functionalized FeNPs and 15 eq of $NaBH_3CN$ to a 1.3 mM HEWL in 0.1 M phosphate buffer pH 5.7), but increasing the reaction time from 6 to 8 hrs. The aim was to boost the reducing activity of $NaBH_3CN$ in order to faster consume the unreacted formaldehyde avoiding protein per-methylation. As previously seen, the reaction mixtures were quenched by addition of a 0.1 M TRIS pH 7.5 and

analyzed *via* HR-ESI in H_2O/ACN in direct infusion under acidic conditions (0.2% formic acid).

In this case, the HR-MS analyses showed that the presence of the POCav@FeNPs **16** system reduces the amount of methyl groups added to the protein with respect to the blank and the control experiments, as highlighted in **Figure 2.23**.

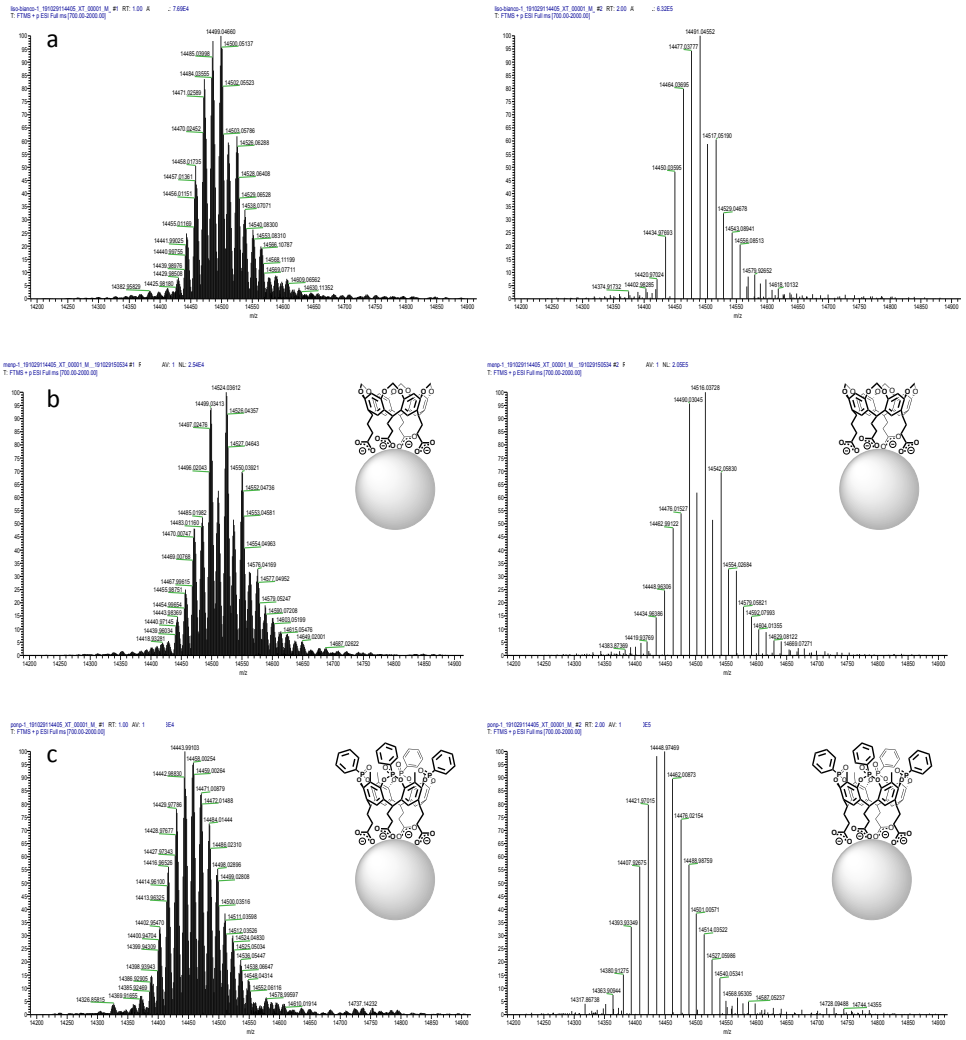


Figure 2.23. HR-ESI Isotopic profile (*left*) and monoisotopic deconvolution (*right*) of a) native Lysozyme, b) Lysozyme + MeCav@FeNPs **10** and c) Lysozyme + POCav@FeNPs **16**.

In **Figure 2.23** the HR-ESI mass spectra are reported to show how in the presence of the working system POCav@FeNPs **16**, the molecular weight distribution is placed towards lower Da.
Considering the blank experiment with native Lysozyme (**Figure 2.23 a**), an average number of 14 methyl residues were added to the protein. The value was obtained by simply subtracting the MW of native Lysozyme, i.e. 14302.2 Da, to the highest deconvoluted mono-isotopic peak calculated from the real spectrum, i.e. 14491.04 Da.
Similarly, in the HR-ESI spectrum of the reaction in which the control system MeCav@FeNPs **10** was used, the addition of 16 methyl groups was calculated. The presence of POCav@FeNPs **16** seems to allow a better control of the methylation reaction, has highlighted by the HR-MS results reported in **Figure 2.23** c. In fact, the addition of only 10 methyl groups was detected.
These attractive results prompted us to further test the ability of the POCav@FeNPs system by replicating the experiment and analyzing the reaction mixture *via* HR-ESI after enzymatic digestion with Trypsin, a serine protease enzyme that cleaves peptide bonds in proteins at the carboxyl side of lysine and arginine amino acidic residues (see Appendix A). The enzymatic digestion was exploited to understand on which lysine residue the methylation occurred and which is the real average of methylation degree for every lysine. The enzymatic digestion process and the subsequent HR-ESI MS analyses were performed in the laboratory of Prof. Sidoli at the Albert Einstein College of Medicine in New York.

Three more set of experiments were performed (**Table 2.5**):

- batch **5** was carried out using the same conditions of batch **4** to verify the reproducibility of the reaction;
- in batch **6** the eq of receptor were increased while the eq of $NaBH_3CN$ and formaldehyde were maintained as in batch **5**;
- batch **7** is a duplicate of batch **6** for reproducibility.

For each batch, experiments employing the tetraphosphonate cavitand **6**, functionalized at the lower rim with four pyridinium moieties, and bare FeNPs were carried out as well to further test the validity of the multivalency approach and the possible influence of FeNPs as heterogeneous metal systems, respectively.

In general, for each experiment 6 eq of CH_2O were added to 1.3 mM HEWL in 0.1 M phosphate buffer at pH 5.7, followed by 6 eq of bare FeNPs (experiment b), or Tiiii **6** (experiment c), or functionalized FeNPs (experiment d and e). The reaction was left to equilibrate for 1 h at rT. Subsequently, 15 eq of $NaBH_3CN$ were added and the reaction was left to proceed for 8 hrs at rT. The reaction was quenched with 0.1 M TRIS buffer pH 7.5 and left to equilibrate for 30 minutes. After enzymatic digestion, the samples were analyzed *via* HR-ESI (see Appendix A).
The reaction conditions are reported in **Table 2.5**.

Batch	Sample	FeNPs* or Tiiii 6 (eq)	$NaBH_3CN$ (eq)	Formaldehyde (eq)	Time
5	5[a]	0	15	6	8 h
	5[b]	6	15	6	8 h
	5[c]	6	15	6	8 h
	5[d]	6	15	6	8 h
	5[e]	6	15	6	8 h
6 and 7	6[a] and 7[a]	0	15	6	8 h
	6[b] and 7[b]	45	15	6	8 h
	6[c] and 7[c]	45	15	6	8 h
	6[d] and 7[d]	45	15	6	8 h
	6[e] and 7[e]	45	15	6	8 h

Table 2.5. Mono-methylation reaction conditions. [a] *blank experiment;* [b] *control experiment with bare FeNPs;* [c] *experiment with Tiiii* ***6****;* [d] *control experiment with MeCav@FeNPs* ***10****;* [e] *experiment with working system POCav@FeNPs* ***16****.* [*] *The amount of FeNPs was calculated on the MeCav or POCav average covering the FeNPs surface.*

The HR-ESI MS results are reported in Appendix A.

The HR ESI-MS analyses highlight that the presence of free or anchored cavitand does not significantly affect the methylation degree of lysine residues, which resulted to be either no, mono- or di-methylated. Rather, the different methylation degree observed for the 6 lysine residues was found to be mostly dependent on the relative surface accessibility[43] (**Figure 2.24**).

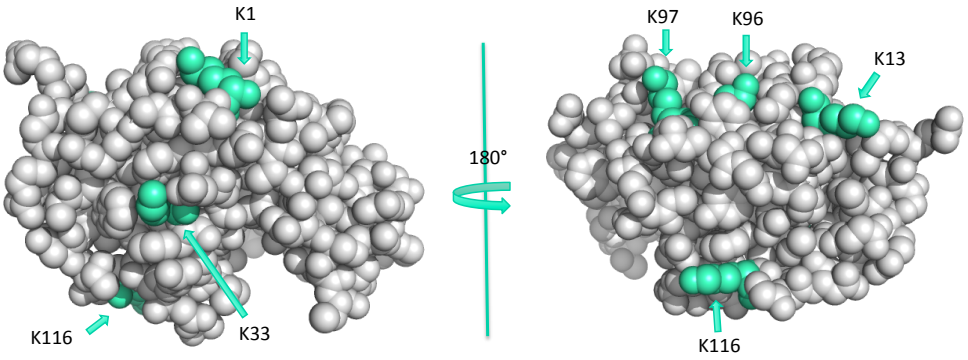


Figure 2.24. Lysine residues accessibility in Lysozyme structure.

In particular, lysine K33 is more exposed to the reaction environment and resulted to be present both as K33me and K33me$_2$. On the contrary, lysine K96, which is less exposed to the reaction environment, resulted to be mainly not methylated.[44]
In the light of these results we decided to test the efficiency of the reductive amination approach, together with ability of the cavitand to drive the reaction towards the formation of the mono-methylated species, on a simplest system, namely the lysine amino acid. Studies are ongoing.

2.3. ARGININE METHYLATION

Among amino acids, Arginine (Arg) represents a unique example with its guanidinium group that potentially possesses five hydrogen bond donors favorable for interactions with biological hydrogen bond acceptors.[45] As demonstrated in protein-DNA complexes, Arg residues are

the most frequent hydrogen bond donors to thymine, adenine and guanine bases over phosphate groups.[46]

Because of the richness of Arg residues in biological molecules, methylation of arginine side chains is quantitatively one of the most extensive PTMs reactions in mammalian cells.[47] In Nature, three types of methylated arginine residues were found: (i) N^G-monomethylarginine (MMA), (ii) N^G,N'^G-dimethylarginine (Symmetric Di-Methyl Arginine – SDMA), in which one methyl group is placed on each of the terminal guanidinium nitrogens and (iii) the most prevalent N^G,N^G-dimethylarginine (Asymmetric Di-Methyl Arginine – ADMA) possessing two methyl groups on one of the terminal nitrogen atoms.

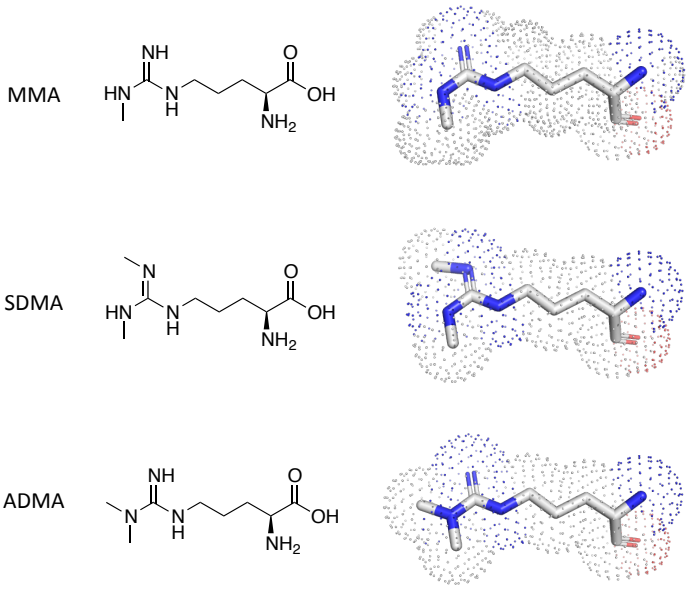


Figure 2.25. Arginine methylation states.

Similarly to lysine methylation, methylation on arginine residues plays a key role in gene regulation, influencing biological processes such as transcription, cell signaling, mRNA translation, DNA damage signaling, receptor trafficking, protein stability and pre-mRNA splicing by activating or repressing histone marks.[48] These effects are partially due to the interaction of modified arginine residues with the methyl-arginine-binding pocket (or "aromatic cage") in Tudor domains that regulates protein-

protein interactions. Thus, modification of arginine residues in proteins can readily modulate their binding interactions and regulate their physiological functions.

In particular, arginine methylation compromises genetic regulation and transcription activities by decreasing the number of possible hydrogen bonding and by altering the shape of the arginine side chain, leading to the development of neurodevelopmental diseases like multiple sclerosis, cancer, cardiovascular diseases and to death.[49] (**Figure 2.26**)

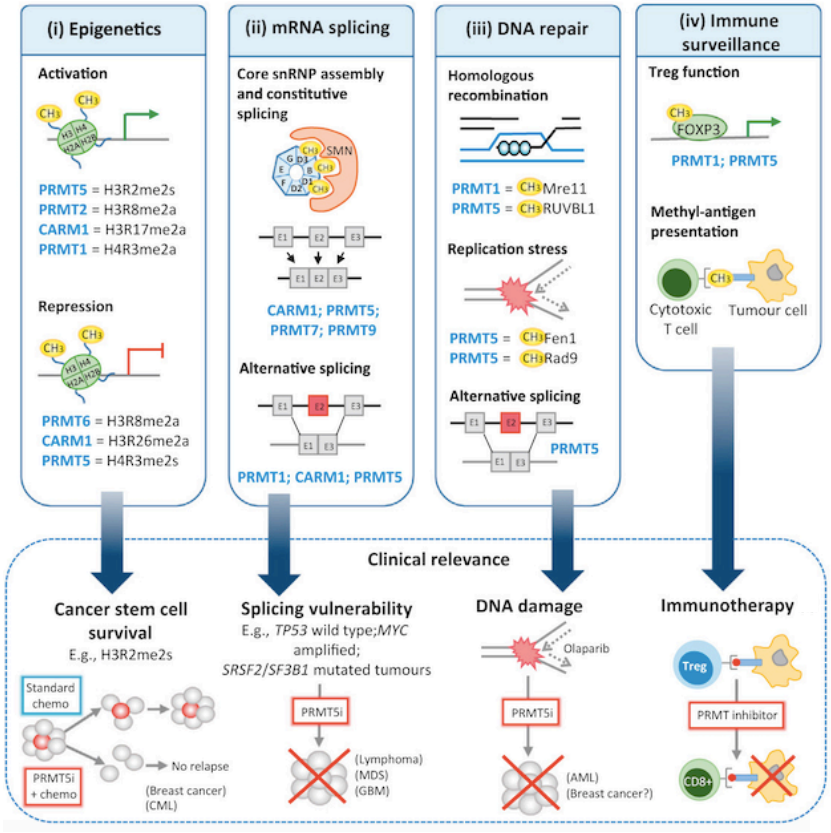


Figure 2.26. Clinical relevance of arginine methylation (*below*) and protein arginine methyltransferases (PRMTs) involvement (*above*). Adapted from J. Jarrols, C.C. Davies, *Trends in Mol. Med*. **2019**, *25*, 993.[50]

2.3.1 BINDING STUDIES WITH TETRAPHOSPHONATE CAVITANDS

Firstly reported about 60 years ago, arginine methylation catalyzed by protein arginine methyltransferase (PRMT) is still a particularly unexplored field of biochemical research,[51] despite the recognition of modified arginine residues can boost the elucidation of arginine methylation function and the role of mutations in diseases paving the way for the discover of new therapeutic approaches.

Taking advantage of the already demonstrated complexation properties of Tiiii cavitands towards methylated lysine,[23] a solid state study of the interactions involved in the host-guest binding between Tiiii and arginine has recently been published.[52]

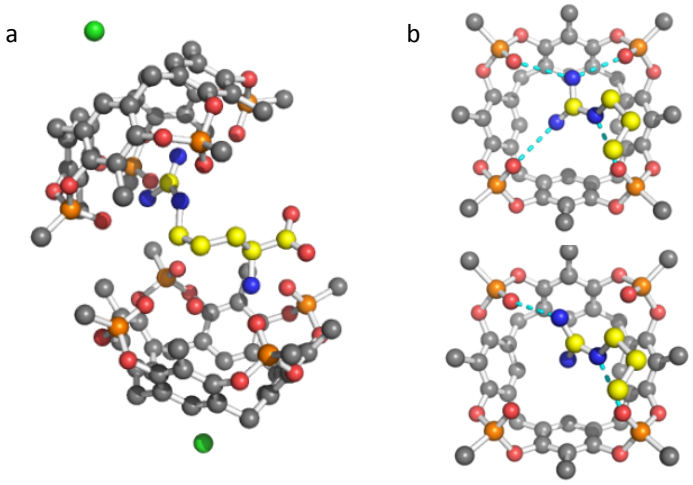
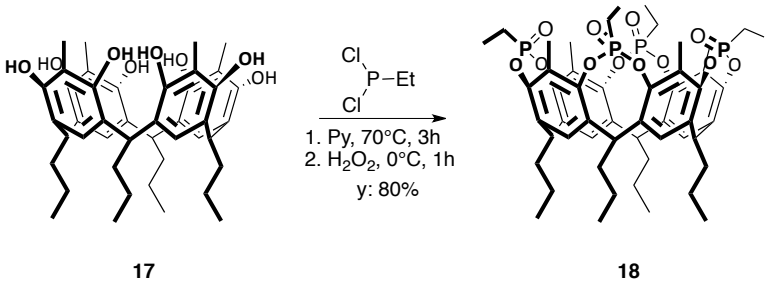


Figure 2.27. Crystal structure of the Arg@Tiiii complexes. Crystals were growth in TFE. Taken from G. Brancatelli, E. Dalcanale, R. Pinalli, S. Geremia, *Molecules* **2018**, *23*, 3368.[52]

The analyses of the diffraction pattern revealed that a 2:1 complex (2Tiiii@1Arg) was obtained in trifluoroethanol (TFE) thanks to the presence of H-bond interactions between the guanidinium group and the P=O groups, and a particular H-bond interaction between the carboxylic acid group of arginine and a P=O group, mediated by a water molecule (**Figure 2.27 a**). Thanks to its electron-rich cavity, Tiiii forms stable complexes with

arginine in two different conformations: one in which the guanidinium group interacts with all the four phosphonate moieties (**Figure 2.27 b**, top) and the other in which it interacts only with two opposite P=O groups (**Figure 2.27 b**, bottom). In the latter case, a NH_2 group is located in the center of the cavity.

In the light of the complexing capabilities of this synthetic receptor also towards arginine in the solid state, a comprehensive study on the recognition properties of Tiiii was performed in solution via ITC and NMR experiments with a series of methylated arginine (see Appendix B). Consequently, tetrasphosphonate cavitands **6** (for the experiments performed in water) and **18** (for the experiments performed in methanol) were synthetized according to the synthetic pathway reported in **Scheme 2.1** (**paragraph 2.2.2**) and **Scheme 2.5**, respectively.
Tiiii cavitand **18** was synthesized in one step starting from the corresponding resorcinarene **17**, which was directly bridged with dichloroethylphosphine in pyridine. The intermediate with four P(III) bridges was oxidized *in situ* to P(V) with hydrogen peroxide retaining its configuration.



17 **18**

Scheme 2.5. Synthetic pathway for tetraphosphonate cavitand **18**.

The following arginine derivatives were studied as guests: Arginine (**Arg**), N^G-monomethylarginine (**MMA**), N^G,N'^G-dimethylarginine (**SDMA**) and N^G,N^G-dimethylarginine (**ADMA**). N_α-Acetyl-L-arginine (**Ac-Arg**) was used only in NMR titration experiments performed in deuterated methanol. Guests were used as hydrochloride or acetate salts. Hosts and guests molecular structures are reported in **Figure 2.28**.

ITC titrations were performed in two solvents, namely methanol and water, at 25 °C on a MicroCal PEAQ-ITC instrument with four different guests (**Arg**, **MMA**, **SDMA** and **ADMA**). To evaluate the strength of the binding events, initial titrations were carried out in methanol. In all the cases, the host solution (conc. *circa* 3.2 mM) was put in the reaction cell and titrated with a ten times more concentrated guest solution in which non- and methylated arginines were used as hydrochloride salts. Experimental titration curves were analyzed and stoichiometry (*n*), association constant (K_a), enthalpy (ΔH) and entropy (ΔS) values were reported as the average of three titrations for every guest. To account for non-specific heats of dilution, each guest was titrated into pure solvent (blank titrations). In all cases, the signal from blank titrations was subtracted from the binding isotherm before curve fitting with single-site models.

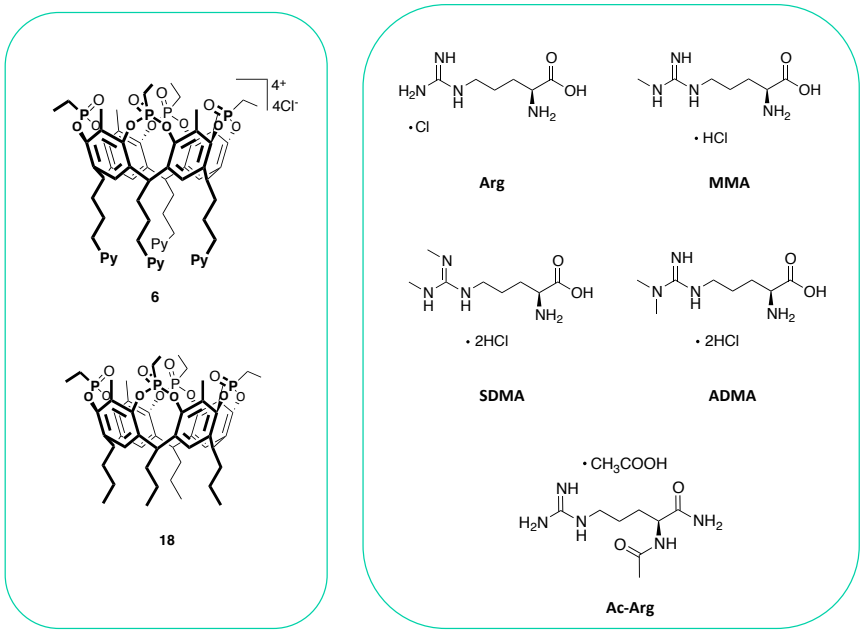


Figure 2.28. Hosts (*left*) and guests (*right*) used in the complexation experiments *via* ITC and NMR.

As shown in **Table 2.6**, the inclusion process in methanol is both enthalpy and entropy driven for all the analyzed guests, although the enthalpy

contribution is preponderant for **MMA** and **ADMA** guests, with an entropy contribution almost neglectable. Moreover, for these two amino acids a stoichiometry close to 0.5 was observed, as an indication of the presence of two binding sites on the guest molecule. Taken together, these results are in agreement with the formation of 2:1 complexes in which both the amino acid and the residue are complexed by Tiiii molecule, as already observed for methylated lysine.[23] Interestingly, if in the lysine case two distinct binding phenomena were observed, in the complexation of **MMA** and **ADMA** the two events appear to have a similar thermodynamic signature. Differently from the solid state studies reported by Geremia and co-workers[52], the complexation stoichiometry observed in the case of the arginine is 1.

The biggest K_a value was recorded for **MMA** (1.8 10^4 M^{-1}), while the K_a of **Arg**, **SDMA** and **ADMA** were at the same order of magnitude (10^3 M^{-1}), with slightly higher values for **ADMA** and **SDMA**.

	n	K_a (M^{-1})	ΔH (kJ/mol)	ΔG (kJ/mol)	TΔS(kJ/mol)
Arg	1.01	5.2 (± 2.6) x 10^3	- 12.9 (± 0.4)	- 21.1 (± 1.3)	8.1 (± 1.7)
MMA	0.566 (± 0.008)	1.8 (± 0.3) x 10^4	- 23.72 (± 0.06)	- 24.1 (± 0.7)	0.6 (± 0.4)
SDMA	0.85 (± 0.04)	8.4 (± 0.1) x 10^3	- 10.7 (± 0.8)	- 22.45 (± 0.03)	11.8 (± 0.8)
ADMA	0.68 (± 0.02)	8.0 (± 0.5) x 10^3	- 22.9 (± 0.2)	- 22.3 (± 0.1)	0.6 (± 0.3)

Table 2.6. ITC studies in methanol. *n = number of sites*

Because of the nature of the guests, ITC experiments were performed in water with the water-soluble cavitand **6** and the same set of AAs. These experiments immediately showed that molecular recognition events are highly solvent dependent. The heats involved in arginines complexation in water are one order of magnitude lower that the ones detected in methanol. This trend was already observed in previous studies on mono-methylated lysine complexation,[23] which demonstrated a 1:1 binding in water guided by the *N*-methylated ammonium residue with a drop of the binding constant of three order of magnitude (from circa 10^6 M^{-1} in methanol to 10^3 M^{-1}in water). In the arginine case, the binding constants

are already in the 10^3-10^4 M^{-1} range in methanol, and a similar drop moving to water makes them undetectable with this technique. In fact, the binding curves obtained by peaks integration do not present a clear sigmoidal curve, hampering the fitting of the data. Any attempts to increase the concentration to achieve higher *c* values (c = n × [**6**]K_a) were frustrated by the limited solubility of the cavitand. However, from initial injection heat values is possible to deduce a qualitative trend for the tested AAs: a major enthalpy contribution for **MMA** complexation in water was reflected in a first injection at about – 5.2 kJ/mol, followed by – 2.5 kJ/mol for **SDMA** and – 0.8 kJ/mol for **Arg**. Almost neglectable heat released was observed for **ADMA** (less than 0.2 kJ/mol).

For the nature of the expected binding constants for arginines complexation with Tiiii, NMR titrations were performed at room temperature in CD_3OD and D_2O monitoring the shift of the guest protons at the ^1H NMR and the cavitand phosphorus at the ^{31}P NMR, both diagnostic for the complexation event.
As in the case of ITC experiments, cavitand **18** was employed for NMR titrations in CD_3OD while cavitand **6** was used for NMR titrations in D_2O. N_α-Acetyl-L-arginine (**Ac-Arg**) was only used in NMR titration in CD_3OD instead of **Arg**, to avoid the formation of a 2:1 complex. The other guests were used as hydrochloride or acetate salts (**Figure 2.27**).
The NMR tubes were filled with 0.6 mL of a solution containing the cavitand. The titrant solution was prepared by dissolving the guest (10 fold more concentrated than the receptor) in 0.3 mL of the above solution, thus the concentration of the receptor remained constant during the titration. Portions of the titrant (2 - 30 µL) were added by syringe to the NMR tube and the ^1H and ^{31}P NMR spectra were recorded after each addition.

As example, the ^{1}H NMR titration of **18** with incremental amount of **MMA** in CD_3OD is reported in **Figure 2.29**.

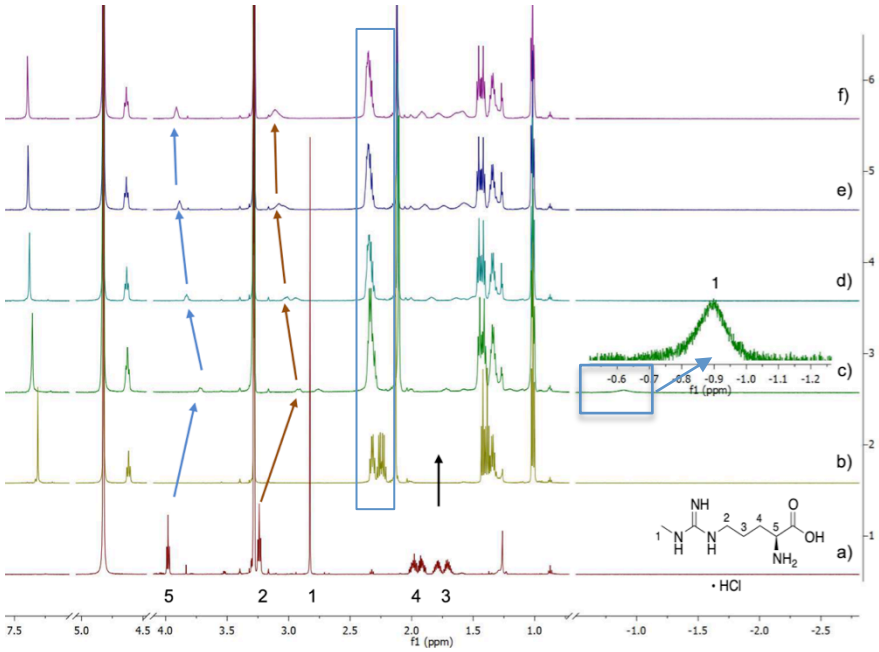


Figure 2.29. ^{1}H NMR (600 MHz, CD_3OD, 298K) spectra acquired during the titration of cavitand **18** with incremental amounts of **MMA**. a) **MMA**, b) **18**, c) **18** + 0.2 eq of **MMA**, d) **18** + 0.5 eq of **MMA**, e) **18** + 1.0 eq of **MMA**, f) **18** + 1.5 eq of **MMA**.

An initial upfield shift of **MMA** signals assigned to protons 1, 2 and 5 was observed upon addition of 0.2 eq of the titrant to the solution of host **18**, as expected for a shielding effect caused by the inclusion of the guest in the cavity. In particular, the $^{+}$N-CH_3 resonance (protons labelled as 1) showed an upfield shift of more than 4 ppm, moving from 2.83 ppm to -1.61 ppm (**Figure 2.29 c**) and the signals of the CH_2-ethylene residues attached to the phosphorus atom of **18** underwent a downfield shift indicating both the effective inclusion and interaction of the methylated guanidinium group with the host cavity.
Initially, the signal related to the protons $^{+}$N-CH_3 appears enlarged due to their complexation by **18** with which exhibit an exchange rate still comparable to the NMR time scale. Subsequently, the exchange rate of the methyl protons increases becoming faster than the NMR time scale. Thus,

the signal related to $^{+}$N-CH_3 can be no longer detected both for the free form and the one complexed by **18**.
Further confirmation of the host-guest interaction is provided by the ^{31}P resonance of the four phosphonate groups of the host that experienced a downfield shift of 1.75 ppm (from 25.17 ppm to 26.92 ppm, **Figure 2.30**).

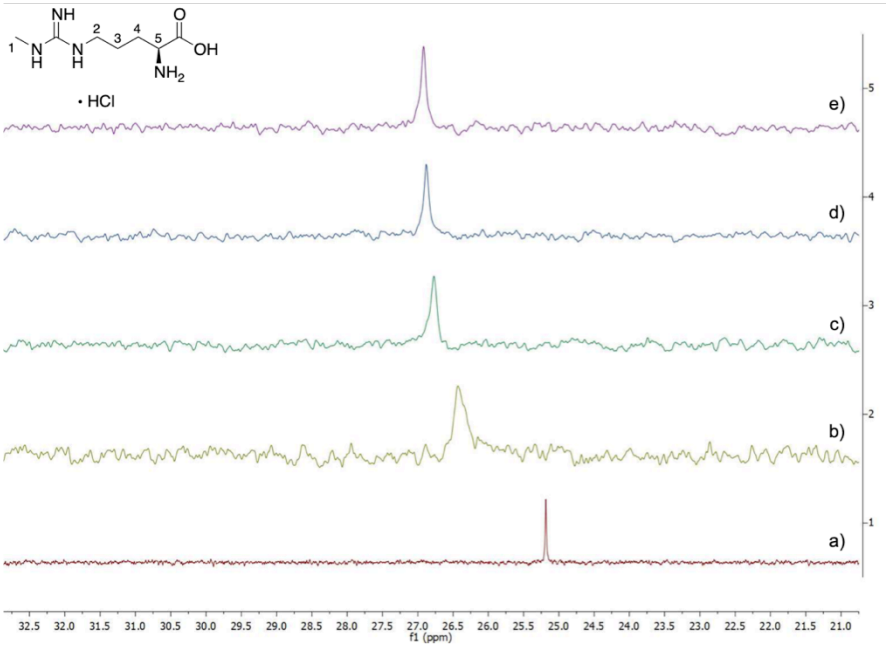


Figure 2.30. ^{31}P NMR (243 MHz, CD_3OD, 298K) spectra acquired during the titration of cavitand **18** with incremental amounts of **MMA**. a) **18**, b) **18** + 0.5 eq of **MMA**, c) **18** + 1.0 eq of **MMA**, d) **18** + 1.5 eq of **MMA**.

For **Ac-Arg**, **SDMA** and **ADMA**, an upfield shift of the ^{1}H NMR signal related to the methyl group was observed, even if less pronounced. The same holds for the ^{31}P NMR (see Appendix C).

To complete the NMR studies, titration experiments were performed also in D_2O for a better understanding of the binding event in the biological medium. Overall, the protons assigned to the guests experienced minimal shifts. The largest upfield shift was observed for the $^{+}$N-CH_3 signal in **MMA** and **SDMA** ($\Delta\delta$ < -1 ppm) while the protons of the other amino acidic guests do not show significant changes. The same behavior was observed

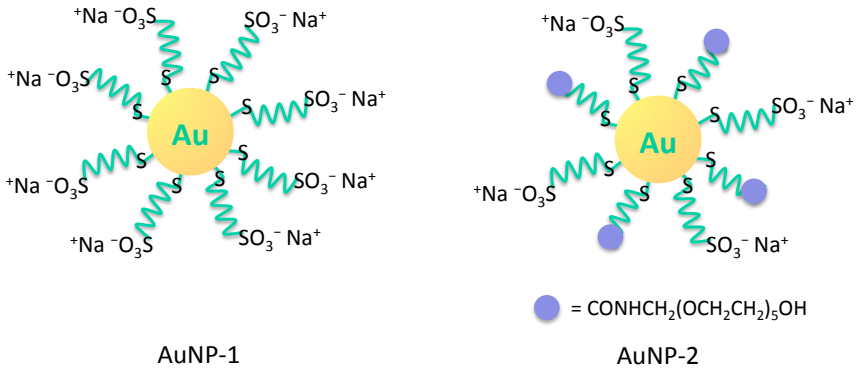
in the ^{31}P NMR titration experiments, where no shift for the P signal was detected for any guests, suggesting that the interactions between the tested arginine derivatives and cavitand **6** in water are too weak to be recorded also by NMR.

These preliminary studies have shown that methylated arginines are recognized by the tetraphosphonate cavitand less effectively with respect to the mono-methylated lysine. In fact, in the case of *N*-methyl lysine a K_a in the order of 10^6 mol^{-1} was calculated in methanol, while for mono-methylated arginine the K_a was found to be in the order of 10^4 mol^{-1}, and for di-methylated arginines in the order of 10^3 mol^{-1}. As expected, moving from methanol to water caused a further drop in the K_a reveling that despite the presence of an *N*-methylated group no efficency in the complexation was obtained.
Solid state studies to understand the interactions involved in the binding event are ongoing, in particular to understand if the methylated guanidinium group is too large to effectively interact with the cavity of the cavitand receptor.

2.4. TETRAPHOSPHONATE CAVITANDS@AuNPs

Monolayer protected gold nanoparticles have emerged as molecular systems with promising applications in the field of biology, medicine, and chemistry. Taking advantage of the ease functionalization of their surface with organic ligands, the interest in the development of supramolecular receptors[53] has increased because of the possibility of these systems to form transient pocket-like binding sites similar to the ones found in enzymes or proteins. Defined binding pockets to boost the affinity to the target of interest, and therefore increasing the analyte detection limit, can be obtained by grafting rigid structures to the nanoparticle surface such as tetrasphosphonate cavitands, which act as pre-synthetized molecularly-precise binding pocket. In collaboration with Prof. Fabrizio Mancin, University of Padova, the possibility to functionalize gold nanoparticles with tetrasphosphonate cavitands was explored.

Exploiting negatively charged sulfonate-functionalized AuNPs (**Figure 2.31**), cavitand **6** was successfully grafted on the AuNPs surface *via* the Tiiii's cationic pyridinium feet *via* a self-assembly process driven by ionic interactions.



AuNP-1 AuNP-2

Figure 2.31. Schematic representation of sulfonate AuNP-1 and AuNP-2.

The obtained supramolecular receptor systems were used for NMR-chemosensing experiments *via* NOE pumping.

2.4.1. NOE PUMPING

NOE (Nuclear Overhauser Effect) Pumping is an analytical method for identification of compounds with binding affinity for macromolecules. The basis of this technique relays on the fact that the translational diffusion coefficient of a small guest changes substantially upon binding to the macromolecule receptor. Affinity NMR uses diffusion editing to "filter out" signals from nonbinding components, and thus, the bound guest will be selected from the mixture.[54]

In particular, Prof. Mancin applied the NOE-Pumping techniques to gold nanoparticles. This method requires the use of a nanoparticles with receptor ability and a combination of different NMR experiments. The procedure involves the application of a pulse of π/2 which magnetizes all the molecules present in solution, followed by a diffusion filter (**Figure**

2.32). This latter is a sequence of pulses and gradients of the magnetic field that allows the elimination of the signals of the small molecules that quickly spread in solution. The application of a sequence of pulses from the transient allows the transfer of magnetization from nanoparticles to interacting molecules by NOE effect returning an NMR spectrum in which only the signals of the nanoparticles and molecules that have undergone the transfer of magnetization are present.
The elimination of the signals related to the AuNPs to obtain a spectrum in which only the signals of the analytes interacting with the supramolecular receptors, can be obtained by applying an additional filter, namely CPMGz.

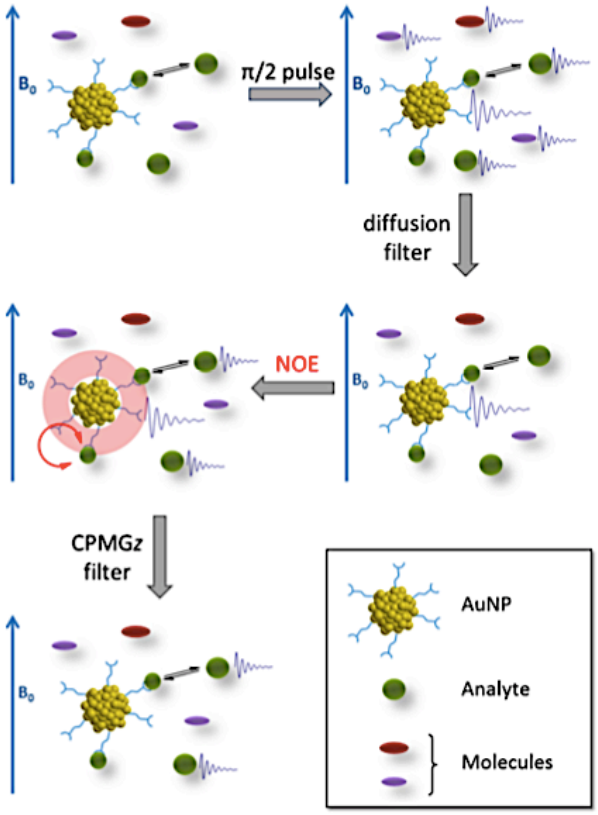


Figure 2.32. NMR Chemosensing acquisition scheme.

The optimal working conditions are those in which the guest is in a rapid exchange regime remaining tied to the receptor for a long enough time to

be able to receive the magnetization, to then be released in solution, thus allowing other molecules target to be magnetized.[55]

2.4.2. NOE PUMPING EXPERIMENTS

Having demonstrated the possibility of using Tiiii for the selective identification of biologically active compounds containing the H_2N^+-CH_3 group, Tiiii@AuNPs systems were initially tested for the detection of sarcosine. This latter, a marker of aggressive prostate cancer[56] in urine samples, is known to be efficiently recognized by Tiiii cavitand. [57]
Experimentally, NMR spectra were recorded on a solution at a known concentration in which gold nanoparticles and cavitand were directly mixed in the NMR tube together with the analyte.
Proton NMR subspectra of sarcosine, cavitand and sulfonate nanoparticles AuNP-1, both alone and combined, are reported in **Figure 2.33**.

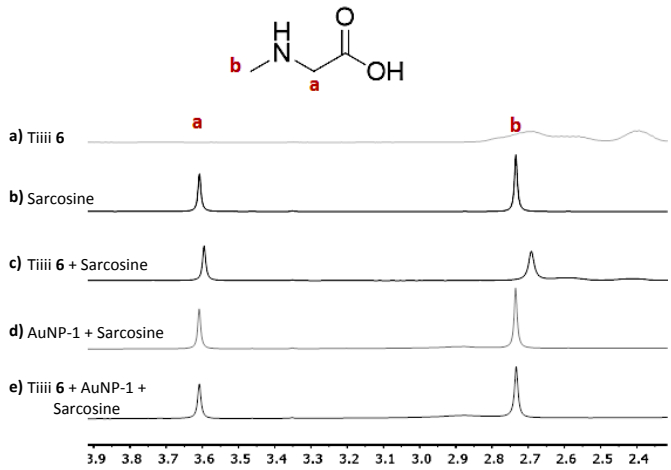


Figure 2.33. ^{1}H NMR spectra comparison (500 MHz). **a**) 1 mM Tiiii **6** in D_2O; **b**) 1 mM Sarcosine in 10 mM phosphate buffer; **c**) 0.125 mM Tiiii **6** + 1 mM Sarcosine in 10 mM phosphate buffer; **d**) 2 mM AuNP-1 + 1 mM Sarcosine in 10 mM phosphate buffer; **e**) 0.125 mM Tiiii **6** + 2 mM AuNP-1 + 1 mM Sarcosine in 10 mM phosphate buffer.

The initial experiments conducted with sarcosine as guest indicating that an interaction between the Tiiii cavitand **6** and the guest occurred. As

proof of this, shifted and broadened signals of the analyte were detected when sarcosine and Tiiii **6** where mixed (**Figure 2.33 c**).
Surprisingly, the addition of the AuNP-1 to the Tiiii:sarcosine solution led to sarcosine signals very close to the ones in the free form, even if no interaction between the AuNP-1 and the analyte was detected in the pertinent control experiment.
Apparently, the grafting of the cavitand on the particle surface decreases its affinity for sarcosine disfavoring the binding.
Standard NMR chemosensing experiments were then carried out (**Figure 2.34**) comparing proton spectra with spectra recorded after a diffusion filter was applied to remove the signals of the fast diffusing species (**Figure 2.34 c** and **f**).

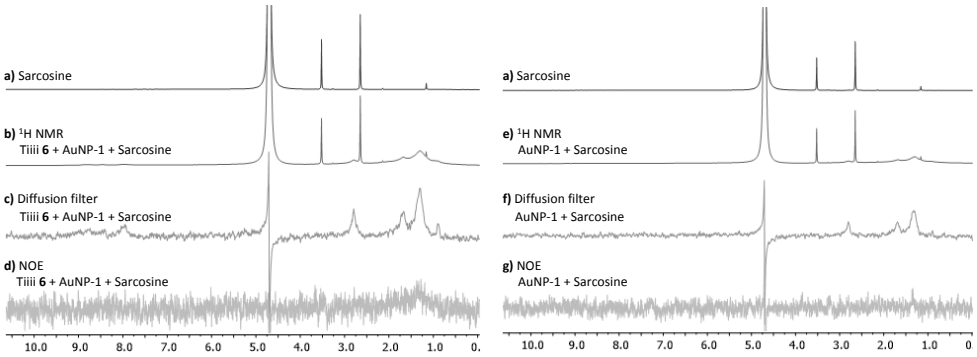


Figure 2.34. ^{1}H NMR spectra comparison (500 MHz). **a**) 1 mM Sarcosine in 10 mM phosphate buffer; **b**) 0.125 mM Tiiii **6** + 2 mM AuNP-1 + 1 mM Sarcosine in 10 mM phosphate buffer; **c**) diffusion filter of b; **d**) NOE pumping of b; **e**) 2 mM AuNP-1 + 1 mM Sarcosine in 10 mM phosphate buffer; **f**) diffusion filter of e; **g**) NOE pumping of e.

The ionic interactions between the four pyridinium feet with the sulfonato group on the AuNPs were demonstrated to be strong enough to ensure the binding of the cavitand to the nanoparticle monolayer, as demonstrated by the cavitand and nanoparticles signals retained after diffusion filter application. However, no detectable signals of the analyte were observed after NOE pumping experiments.

Similar results (**Figure 2.35**) were obtained with *N*-methyl-phenethylamine (MePEA) as guest. This latter was chosen because similar

in structure to catecholamine metabolites found to be related to neuroblastoma, a malignant tumor.[58]

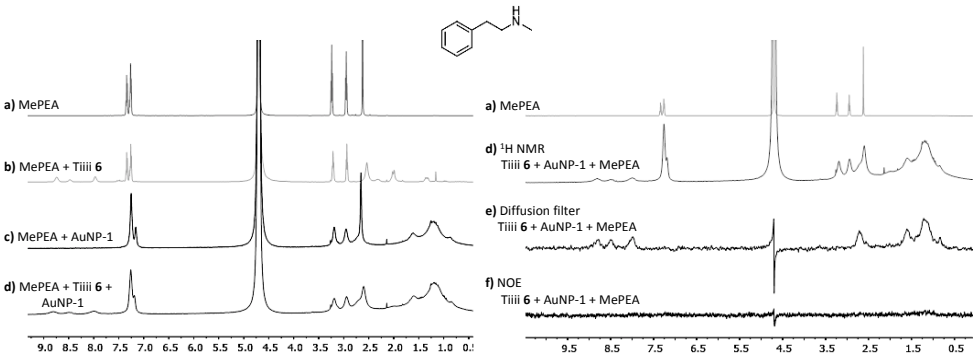


Figure 2.35. ^{1}H NMR spectra comparison. **a**) 1 mM MePEA in 10 mM phosphate buffer; **b**) 1 mM MePEA + 0.125 mM Tiiii **6** in 10 mM phosphate buffer; **c**) 1 mM MePEA + 2 mM AuNP-1 in 10 mM phosphate buffer; **d**) 1 mM MePEA + 0.125 mM Tiiii **6** + 2 mM AuNP-1 in 10 mM phosphate buffer; **e**) diffusion filter of d; **f**) NOE pumping of d.

2.4.3. STD EXPERIMENTS

The STD (*Saturation Transfer Difference*) method is based on the saturation of the magnetization of a macromolecule through the selective irradiation of a narrow region of its ^{1}H NMR spectrum.[59] The saturation, propagating by spin diffusion, is able to magnetize the analytes interacting with the macromolecule for a sufficiently long time. These analytes, experiencing a negative NOE, undergo a lowering of the intensity of their signals. By subtracting those signals from an off-resonance NMR spectrum used as reference, the analytes can be detected.
STD experiment allowed to obtain more intense signals compared to NOE pumping, since saturation can be maintained for longer times than the mixing times characteristic of the NOE experiments.[60]

Experimentally, saturation of both the alkyl chains of the AuNPs thiol monolayer and the pyridinium rings of Tiiii **6** was tested. Differently from the NOE pumping experiments, saturation of the alkyl chains of the nanoparticle monolayer at 641 Hz (in green in **Figure 2.36**) resulted to be

more efficient than the saturation at 4405 Hz of the pyridinium moieties in cavitand **6** (in red in **Figure 2.36**), with a signal to noise ratio that allows analyte detection.
STD experiments with sarcosine (**Figure 2.36** left) and *N*-methyl-phenethylamine (**Figure 2.36** right) are reported.

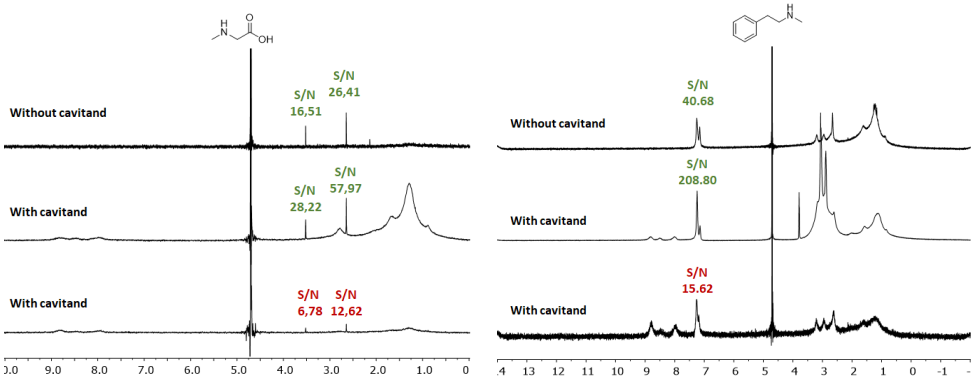


Figure 2.36. STD experiments of sarcosine (left) and MePEA (right). Saturation was applied in the AuNP monolayer (green) and on the pyridinium feet of cavitand **6** (red).

In conclusion, the preliminary experiments conducted with the Tiiii@AuNPs systems reveled that a strong interaction between the cavitand and the nanoparticle is established allowing the magnetization transfer to the analyte bound by the Tiiii receptor. Unlike NOE pumping experiments, STD experiments showed a greater sensitivity and analyte signals were detected. Other studies are ongoing to confirm the results.

2.5 EXPERIMENTAL SECTION

Resorcinarene [C_3H_6OH; Me] (1).
To a solution of 2-methylresorcinol (15 g, 0.12 mol) in MeOH (90 mL), 2,3-Dihydrofuran (9.15 mL, 0.12 mol) and a 37% solution of HCl (22.5 mL) were added dropwise over 30 min at 0° C. After the addition, the reaction mixture was stirred at 50° C for a week. The reaction was quenched with water, filtered, dried under vacuum and recrystallized three times from MeOH. The final product **1** was obtained as a pale yellow powder (7.9 g, 34%).**^{1}H NMR** (DMSO-d_6, 400 MHz): δ ppm = 8.66 (s, 8H, **OH**), 7.29 (s, 4H, **ArH_{down}**), 4.19 (t, 4H, -**CH**-, J = 9.80 Hz), 3.44 (m, 8H, -**CH_2**OH), 2.25 (m, 8H, CH**CH_2**-), 1.94 (s, 12H, Ar**CH_3**), 1.35 (m, 8H, -**CH_2**CH_2OH); **ESI-MS:** m/z 777.41 $[M+H]^+$.

Tetra(dimethylsiloxy)-bridged cavitand (2).
Tetrahydroxy-footed resorcinarene **1** (850 mg, 1.09 mmol) was dissolved in 10 mL of dry pyridine. The solution was cooled to 0° C and $(CH_3)_2SiCl_2$ (2.37 mL, 19.7 mmol) was added. The mixture was stirred at room temperature for 20 min and heated at 100° C for 3 hrs. The reaction was cooled down to room temperature and quenched with MeOH (20 mL). The obtained precipitate was filtered, washed with methanol and dried, to give product **2** as an off-white powder (1.07 g, 98%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ ppm = 7.47 (s, 4H, Ar**H_{down}**), 4.54 (t, 4H, -**CH**-, J = 7.9 Hz), 3.66 (bt, 8H, -**CH_2**OH), 2.45 (m, 8H, CH**CH_2**-), 1.87 (s, 12H, Ar**CH_3**), 1.50-1.46 (m, 8H, -**CH_2**CH_2OH), 0.46 (s, 12H, Si**CH_{3out}**), -0.73 (s, 12H, Si**CH_{3in}**). **ESI-MS:** m/z 1002.50 $[M+H]^+$.

Tetrachloride-footed dimethylsiloxy cavitand (3).
To a solution of **2** (376 mg, 0.38 mmol) in 10 mL of DMF, thionyl chloride (327 µL, 4.5 mmol) was added. The reaction was heated at 55 °C overnight. The solvent was evaporated and the residue was dissolved in DCM, washed with water and dried over $MgSO_4$. Product **3** was obtained as brownish solid (408 mg, quantitative yield). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ ppm = 7.19 (s, 4H, Ar**H_{down}**), 4.64 (t, 4H, -**CH**-, J = 6.0 Hz), 3.63 (t, 8H, -**CH_2**Cl, J = 4.8 Hz), 2.39 (m, 8H, CH**CH_2**-), 1.93 (s, 12H, Ar**CH_3**), 1.80 (m, 8H, -**CH_2**CH_2Cl), 0.54 (s, 12H, Si**CH_{3out}**), -0.67 (s, 12H, Si**CH_{3in}**). **ESI-MS:** m/z 1075.49 $[M+H]^+$; m/z 1094.36 $[M+NH_4]^+$.

Resorcinarene [C_3H_6Cl; Me] (4).
An aqueous 40% HF solution (454 µL, 4.17 mmol) was added to **3** (408 mg, 0.38 mmol) dissolved in 4 mL of a 3:1 mixture of DMF/$CHCl_3$. The reaction was heated at 55 °C overnight, quenched with water and filtered. Resorcinarene **4** was obtained as a yellowish solid (320 mg, quantitative yield). **^{1}H NMR** (Acetone-d_6, 300 MHz): δ ppm = 8.05 (s, 8H, **OH**), 7.48 (s, 4H, Ar**H$_{down}$**), 4.43 (t, 4H, -**CH**-, J = 7.8 Hz), 3.65 (t, 8H, -**CH$_2$**Cl, J = 6.5 Hz), 2.46 (m, 8H, CH**CH$_2$**-), 1.76 (m, 8H, -**CH$_2$**CH_2Cl). **ESI-MS:** m/z 847.45 $[M-H]^-$.

POEt Cavitand [C_3H_6Cl; Me] (5).
Resorcinarene **4** (323 mg, 0.38 mmol) was dissolved in 8 mL of dry pyridine and dichloroethyphosphine (174 µL, 1.67 mmol) was added under Ar atmosphere. The solution was stirred at 80 °C for 4 hrs. The mixture was then cooled to 0 °C, and 2 mL of 30% H_2O_2 were added. The reaction was stirred at room temperature for 1 h and then quenched with 100 mL of water. The obtained precipitate was filtered, washed with water and dried. The solid was dissolved in DCM and precipitated with hexane. After filtration, cavitand **5** was obtained as white solid (322 mg, 74%). **^{1}H NMR** ($CDCl_3$, 400 MHz): δ ppm = 7.13 (s, 4H, Ar**H$_{down}$**), 4.63 (t, 4H, -**CH**-, J = 7.3 Hz), 3.69 (t, 8H, -**CH$_2$**Cl, J = 6.2 Hz), 2.46 (m, 8H, CH**CH$_2$**-), 2.20 (m, 20H, P(O)**CH$_2$**CH_3 + Ar**CH$_3$**), 1.85 (m, 8H, -**CH$_2$**CH_2Cl), 1.48 (m, 12H, P(O)**CH$_2$**CH_3). **^{31}P NMR** ($CDCl_3$, 162 MHz): δ ppm = 23.03 (s, **P**=O). **ESI-MS:** m/z 1163.79 $[M+NH_4]^+$; m/z 1169.75 $[M+Na]^+$; m/z 1185.93 $[M+K]^+$.

POEt Cavitand [$C_3H_6Py^+Cl^-$; Me] (6).
Cavitand **5** (322 mg, 0.28 mmol) was dissolved in dry pyridine and stirred at 120 °C for 3 days. After cooling to room temperature, the solvent was removed in vacuum and the obtained precipitate was dispersed in acetone. The mixture was stirred for 4 hrs. The obtained solid was filtered and washed with acetone. The final product **6** was obtained as a pale yellow powder (373 mg, 91%). **^{1}H NMR** (D_2O, 400 MHz): δ ppm = 8.76 (d, 8H, Py**H$_o$**, J = 5.8 Hz), 8.44 (t, 4H, Py**H$_p$**, J = 7.7 Hz), 7.95 (t, 8H, Py**H$_m$**, J = 6.8 Hz), 7.53 (s, 4H, Ar**H$_{down}$**), 4.64 (t, 8H, -**CH$_2$**Py^+, J = 7.2 Hz), 4.48 (t, 8H, -**CH**-, J = 6.2 Hz), 2.64 (m, 8H, CH**CH$_2$**-), 2.29 (dt, 8H, P(O)**CH$_2$**CH_3, J = 7.3 Hz, J_{P-H} = 26.0 Hz), 1.97 (m, 20H, Ar**CH$_3$** + -**CH$_2$**CH_2Py^+) 1.32 (dt, 12H, P(O)CH_2**CH$_3$**, J = 7.6 Hz, J_{P-H} = 22.4 Hz). **^{31}P NMR** (D_2O, 162 MHz): δ ppm = 28.23 (s, **P**=O).

HR-ESI: calculated for $C_{72}H_{84}N_4O_{12}P_4$ m/z 330.1254 $[M-4Cl]^{4+}$, found 330.1256.

Recorcinarene [C_3H_6OH; H] (7).
To a solution of resorcinol (15 g, 0.14 mol) in MeOH (90 mL), 2,3-Dihydrofuran (10.3 mL, 0.14 mol) and a 37% solution of HCl (22.5 mL) were added dropwise over 30 min at 0° C. After the addition, the reaction mixture was stirred at 50° C for a week. The reaction was quenched with water, filtered, dried under vacuum and recrystallized three times from MeOH. The final product **7** was obtained as a pale yellow powder (8.4 g, 33%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ ppm = 8.88 (s, 8H, OH), 7.21 (s, 4H, **ArH_{up}**), 6.13 (s, 4H, **ArH_{down}**), 4.18 (t, 4H, -**CH**-, J = 7.92 Hz), 3.40 (m, 8H, -**CH_2**OH), 2.07 (m, 8H, CH**CH_2**-), 1.31 (m, 8H, **-CH_2**CH_2OH); **ESI-MS**: m/z 743.39 $[M+Na]^+$; m/z 759.38 $[M+K]^+$.

Methylene-bridged Cavitand (8).
Resorcinarene **7** (300 mg, 0.42 mmol) was dissolved in anhydrous DMF (7 mL). K_2CO_3 (574 mg, 4.16 mmol) and CH_2I_2 (0.84 mL, 16.65 mmol) were added and the solution was stirred at 80° C for 16 hrs. The solvent was removed and the crude was extracted with DCM/H_2O. The organic phase was filtered through Na_2SO_4 and dried under vacuum. The final product **8** was obtained as a ochre powder (298 mg, 93%). **^{1}H NMR** (DMSO-d_6, 400 MHz): δ ppm = 7.59 (s, 4H, **ArH_{up}**), 6.48 (s, 4H, **ArH_{down}**), 5.68 (d, 4H, ArO**CH_{2out}**O**,** J = 7.60 Hz), 4.54 (t, 4H, -**CH**-, J = 8.00 Hz), 4.42 (t, 4H, -**OH**, J = 4.60 Hz), 4.36 (d, 4H, ArO**CH_{2in}**O, J = 7.60 Hz), 3.49 (m, 8H, -**CH_2**OH); 2.38 (m, 8H, CH**CH_2**-), 1.42 (m, 8H, -**CH_2**CH_2OH). **ESI-MS**: m/z 770.39 $[M+H]^+$.

TetraCOOH - Methyl bridged cavitand (9).
CrO_3 (310 mg, 3.12 mmol) was added to a solution of cavitand **8** (150 mg, 0.195 mmol) in 10 mL of a DMF-acetone (1:1) mixture. Then, H_2SO_4 was added until pH 2 was reached. After 5 hrs, the solvent was removed and the precipitate washed with acidic water (pH 2) and centrifuged three times. Product **9** was obtained, after been dried under vacuum, as a buff-colored powder. (152 mg, 94%). **^{1}H NMR** (Acetone-d_6, 400 MHz): δ ppm = 7.72 (s, 4H, **ArH_{up}**), 6.54 (s, 4H, **ArH_{down}**), 5.71 (d, 4H, ArO**CH_{2out}**O**,** J = 7.48 Hz), 4.79 (t, 4H, -**CH**-, J = 8.16 Hz), 4.45 (d, 4H, ArO**CH_{2in}**O**,** J = 7.52 Hz), 2.69 (m, 8H, CH**CH_2**-), 2.35 (t, 8H, -**CH_2**COOH, J = 7.42 Hz). **^{13}C NMR** (Acetone-d_6,

100 MHz): δ ppm = 205.25, 173.36, 155.22, 138.12, 121.98, 166.77, 99.43, 36.72, 32.43, 24.98. **HR-ESI**: m/z calculated for $C_{40}H_{39}O_{16}$ 823.2244 $[M-H]^-$, found 823.2247; m/z calculated for $C_{40}H_{38}O_{16}$ 411.1085 $[M-2H]^{2-}$, found 411.1085.

Tetra(dimethylsiloxy)-bridged MOM protected cavitand (11).
To a solution of cavitand **2** (1.07 g, 1.07 mmol) in 5 mL of dry DMF, N,N-diisopropylethylamine (2.80 mL, 16.03 mmol) and chloromethyl methyl ether (812 µL, 10.68 mmol) were added. The mixture was stirred at 40° C for 16 hrs. The reaction was cooled down to room temperature and the solvent was evaporated under reduced pressure. The obtained precipitate was washed with water and dried, to give a beige product (1.24 g, 98%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ ppm = 7.17 (s, 4H, Ar**H**$_{up}$), 4.63 (s, 8H, -O**CH**$_2$O-), 4.62 (t, 4H, -**CH**-, J = 8.2 Hz), 3.58 (t, 8H, -**CH**$_2$OMOM, J = 6.3 Hz), 3.38 (s, 12H, O**CH**$_3$), 2.29 (q, 8H, CH**CH**$_2$-, J = 7.1 Hz), 1.90 (s, 12H, Ar**CH**$_3$), 1.58 (q, 8H, -**CH**$_2$CH$_2$OMOM, J = 7.1 Hz), 0.51 (s, 12H, Si**CH**$_{3out}$), -0.69 (s, 12H, Si**CH**$_{3in}$). **ESI-MS:** m/z 1177.53 $[M+H]^+$.

Resorcinarene [C_3H_6OMOM; Me] (12).
To a solution of cavitand **11** (1.24 g, 1.05 mmol) in 15 mL of DMF, an aqueous 40% HF solution (1.2 mL, 11.54 mmol) was added. The suspension was stirred overnight at 50° C. The solvent was removed and addition of water to the reaction mixture promoted the precipitation the product. The solid was filtered, washed with water and dried to give product **12** as white solid (890 mg, 99%). **^{1}H NMR** (DMSO-d_6, 300 MHz): δ ppm = 8.68 (s, 8H, Ar**OH**), 7.29 (s, 4H, Ar**H**$_{up}$), 4.53 (s, 8H, -O**CH**$_2$O-), 4.22 (t, 4H, -**CH**-, J = 7.7 Hz), 3.48 (t, 8H, -**CH**$_2$OMOM, J = 6.4 Hz), 3.33 (s, 12H, O**CH**$_3$), 2.28 (m, 8H, CH**CH**$_2$-), 1.94 (s, 12H, Ar**CH**$_3$), 1.43 (m, 8H, -**CH**$_2$CH$_2$OMOM). **ESI-MS:** m/z 953.51 $[M+H]^+$.

POPh Cavitand [C_3H_6OMOM; Me] (13).
To a solution of resorcinarene **12** (890 mg, 1.05 mmol) in 18 mL of dry pyridine, dichloroethylphosphine (478 µL, 4.60 mmol) was added under argon atmosphere. The solution was stirred at 80° C for 3 hrs. The mixture was cooled to 0° C and 5 mL of 35% aqueous H_2O_2 were added. The reaction was stirred at room temperature for 1h and then quenched with 100 mL of water. The precipitate was filtered, washed with water and

dried. The solid was dissolved in DCM and precipitated with hexane. After filtration, cavitand **13** was obtained as a white solid (1.13 g, 75%).
^{1}H NMR ($CDCl_3$, 300 MHz): δ ppm = 8.13 (m, 8H, P(O)Ar**H_o**, J = 7.1 Hz), 7.64 (m, 4H, P(O)Ar**H_p**, J = 9.5 Hz), 7.57 (m, 8H, P(O)Ar**H_m**, J = 4.9 Hz), 7.20 (s, 4H, Ar**H_{down}**), 4.90 (t, 4H, -**CH**-, J = 7.0 Hz), 4.84 (s, 8H, -O**CH_2**O-), 3.69 (t, 8H, -**CH_2**OMOM, J = 5.9 Hz), 3.50 (s, 12H, O**CH_3**), 2.46 (m, 8H, CH**CH_2**-), 2.29 (s, 12H, Ar**CH_3**), 1.70 (m, 8H, -**CH_2**CH_2OMOM). **^{31}P NMR** ($CDCl_3$, 162 MHz): δ ppm = 7.13 (s, P=O). **ESI-MS:** m/z 1464.36 $[M+Na]^+$.

POPh Cavitand [C_3H_6OH; Me] (14).
Cavitand **13** (530 g, 0.37 mmol) was dissolved in 7 mL of MeOH and 3.5 mL of $CHCl_3$. A 37% solution of HCl (1 mL) was added and the solution was stirred overnight at 40° C. After evaporation of the solvent, the crude was dissolved in DCM and precipitated with hexane. The solid was filtered and dried to give cavitand **14** as an off white solid (468 mg, Quant). **^{1}H NMR** (CD_3OD, 400 MHz): δ ppm = 8.13 (m, 8H, P(O)Ar**H_o**, J = 7.4 Hz), 7.76 (m, 4H, P(O)Ar**H_p**, J = 7.2 Hz), 7.65 (m, 8H, P(O)Ar**H_m**, J = 5.0 Hz), 7.61 (s, 4H, Ar**H_{down}**), 4.90 (t, 4H, -**CH**-, J = 7.3 Hz), 3.75 (t, 8H, -**CH_2**OH, J = 6.2 Hz), 2.60 (m, 8H, CH**CH_2**-), 2.28 (s, 12H, Ar**CH_3**), 1.66 (m, 8H, -**CH_2**CH_2OH). **^{31}P NMR** ($CDCl_3$, 162 MHz): δ ppm = 8.88 (s, P=O). **ESI-MS:** m/z 1297.81 $[M+K]^+$.

POPh Cavitand [C_2H_4COOH; Me] (15).
CrO_3 (540 mg, 4.28 mmol) was added to a solution of cavitand **14** (339 mg, 0.27 mmol) in 30 mL of DMF-acetone (2:1) mixture. Then, H_2SO_4 was added until pH 2 was reached. After 16 hrs, the solvent was removed and the precipitate washed with acidic water (pH 2) and centrifuged three times. The product was obtained, after been dried under vacuum, as a pale green powder. (318 mg, 89%). **^{1}H NMR** (CD_3OD, 400 MHz): δ ppm = 8.16 (m, 8H, P(O)Ar**H_o**), 7.82 (m, 4H, P(O)Ar**H_p**), 7.71 (m, 8H, P(O)Ar**H_m**), 7.64 (s, 4H, Ar**H_{down}**), 5.04 (bt, 4H, -**CH**-), 2.86 (m, 8H, -**CH_2**COOH), 2.45 (m, 8H, CH**CH_2**-), 2.18 (s, 12H, Ar**CH_3**). **^{31}P NMR** (CD_3OD, 162 MHz): δ ppm = 10.2 (s, P=O). **^{13}C NMR** (CD_3OD, 100 MHz): δ ppm = 175.32, 145.37, 145.26, 134.74, 134.23, 131.26, 131.15, 129.09, 128.92, 125.66, 125.23, 120.26, 36.31, 31.54, 25.84. **HR-ESI:** m/z calculated for $C_{68}H_{59}O_{10}P_4^-$ 1319.2556 $[M-H]^-$, found 1319.2536.

Synthesis of magnetic nanoparticles.
Iron magnetite nanoparticles were synthesized by alkaline co-precipitation of Fe^{3+} and Fe^{2+} according to a literature procedure.[28]
In two separate vials, $FeCl_2 \cdot 4H_2O$ (18 mg, 0.09 mmol) was dissolved in 0.09 mL of H_2O while $FeCl_3 \cdot 6H_2O$ (50 mg, 0.18 mmol) was dissolved in 0.18 mL of H_2O. A solution of NH_3 30% (0.1 mL) was added to the second solution and immediately after also the Fe^{2+} solution was added. The resulting suspension was left to decant at room temperature and washed twice with distilled water. The obtained magnetic nanoparticles (FeNPs) were isolated from the solvent by magnetic decantation. **ATR-FTIR**: 3000 cm^{-1} (OH stretching). **TGA** (air, 30 °C – 800 °C, flow rate: 30 °C/min): 2.3% w/w loss at 89 °C; 2.0% w/w loss at 272 °C. **XPS**: see paragraph 2.2.4.

Functionalized FeNPs (10 and 16).
In separate beakers, $FeCl_2 \cdot 4H_2O$ (335 mg, 1.79 mmol) was dissolved in 1.7 mL of H_2O and $FeCl_3 \cdot 6H_2O$ (923 mg, 3.41 mmol) was dissolved in 3.4 mL of H_2O. Cavitand **9** (100 mg, 0.12 mmol) or Cavitand **15** (100 mg, 0.08 mmol) was dissolved in a mixture of NH_3 30% (2 mL) and water (40.8 mL) and the Fe^{3+} and Fe^{2+} solutions were added sequentially. The resulting suspension was decanted at room temperature, washed with distilled water and separated by centrifugation. The obtained functionalized magnetic nanoparticles (MeCav@FeNPs **10** or POCav@FeNPs **16**) were isolated from the solvent by magnetic decantation.
MeCav@FeNPs **10**. **ATR-FTIR**: 3000 cm^{-1} (CH and OH stretching), 1650 cm^{-1} (C=O stretching), 1495 cm^{-1} (C=C stretching), 1290 cm^{-1} and 1090 cm^{-1} (C-O-C stretching), 960 cm^{-1} (C-C stretching). **TGA** (air, 30 °C – 800 °C, flow rate: 30 °C/min): 6.5% w/w loss at 113 °C; 19.5% w/w loss at 301 °C. **XPS**: see paragraph 2.2.4.
POCav@FeNPs **16**. **ATR-FTIR**: 3000 cm^{-1} (CH and OH stretching), 1650 cm^{-1} (C=O stretching), 1495 cm^{-1} (C=C stretching), 1450 cm^{-1} (C-O-C stretching), 1205 cm^{-1} (P=O stretching), 960 cm^{-1} (C-C stretching), 902 cm^{-1} (O-P-O stretching). **TGA** (air, 30 °C – 800 °C, flow rate: 30 °C/min): 5.1% w/w loss at 119 °C; 18.5% w/w loss at 304 °C. **XPS**: see paragraph 2.2.4.

Resorcinarene [C_3H_7; Me] (17).
To a solution of 2-methylresorcinol (15 g, 0.12 mol) in MeOH (90 mL), *n*-butyraldehyde (10.9 mL, 0.12 mol) and a 37% solution of HCl (22.5 mL)

were added dropwise over 30 min at 0° C. After the addition, the reaction mixture was stirred at 50° C for a week. The reaction was quenched with water, filtered, dried under vacuum and recrystallized three times from MeOH. The final product **17** was obtained as a pale yellow powder (6.8 g, 35%). **^{1}H NMR** (Acetone-d_6, 400 MHz): δ ppm = 8.51 (s, 8H, **OH**), 7.49 (s, 4H, **ArH_{down}**), 4.41 (t, 4H, -**CH**-, J = 7.8 Hz), 2.28 (m, 8H, CH**CH_2**-), 2.04 (s, 12H, Ar**CH_3**), 1.35 (m, 8H, -**CH_2**CH_3), 0.98 (t, 12H, -CH_2**CH_3**, J = 7.3 Hz); **ESI-MS:** m/z 711.23 $[M-H]^-$.

POEt Cavitand [C_3H_7; Me] (18).

To a solution of resorcinarene **17** (1 g, 1.41 mmol) in 20 mL of dry pyridine, dichloroethylphosphine (0.6 mL, 5.92 mmol) was added. The solution was stirred at 70 °C for 3h, then the mixture was cooled to 0 °C and 35% aqueous H_2O_2 (15 mL) was added. The mixture was stirred for 1 h, quenched with water and filtered. The obtained white precipitate was collected and dried under vacuum (1.14 g, 80%). **^{1}H NMR** ($CDCl_3$, 400 MHz): δ ppm = 7.61 (s, 4H, Ar**H_{down}**), 4.58 (t, 4H, -C**H**-, J = 7.5 Hz), 2.46 (m, 8H, CH**CH_2**-), 2.28 (dq, 8H, P(O)**CH_2**CH_3, J_{H-P} = 18.8 Hz, J = 7.6 Hz), 2.13 (s, 12H, Ar**CH_3**), 1.48 (dt, 12H, P(O)CH_2**CH_3**, J_{H-P} = 21.7 Hz, J = 7.6 Hz), 1.39 (m, 8H, -**CH_2**CH_3), 1.05 (t, 12H, -CH_2**CH_3**, J = 7.3 Hz). **^{31}P NMR** ($CDCl_3$, 162 MHz): δ ppm = 25.9 (s, P=O). **ESI-MS:** m/z 1026.6 $[M+NH_4]^+$, 1031.5 $[M+Na]^+$.

2.6 REFERENCES

1. (a) C. Tse, T. Sera, A.P. Wolffe, J.C. Hansen, *Mol. Cell. Biol.* **1998**, *18*, 4629; (b) L.M. Carruthers, J.C. Hansen, *J. Biol. Chem.* **2000**, *275*, 37285; (c) X. Wang, C. He, S.C. Moore, J. Ausio, *J. Biol. Chem.* **2001**, *276*, 12764; (d) S. Chambeyron, W.A. Bickmore, *Genes. Dev*. **2004**, *18*, 1119; (e) A.E. Wiblin, W. Cui, A.J. Clark, W.A. Bickmore, *J. Cell. Sci.* **2005**, *118*, 3861.
2. (a) C.L. Peterson, M.A. Laniel, *Curr. Biol.* **2004**, *14*, 546; (b) R. Margueron, P. Trojer, D. Reinberg, *Curr. Opin. Genet. Dev.* **2005**, *15*, 163.
3. (a) N. Song, J. Liu, S. An, T. Nishino, Y. Hishikawa, T. Koji*, Acta Histochem. Cytoc.* **2011***, 44,* 183; (b) T. Kouzarides, *Cell* **2007***, 128*, 693.
4. J. Zhong S. Shaik, L. Wan, A.E. Tron, Z. Wang, L. Sun, H. Inuzuka, W. Wei, *Oncotarget* **2014**, *4*, 2339.
5. D. Matenia, C. Hempp, T. Timm, A. Eikhof, E.M. Mandelkow, *J. Biol. Chem.* **2012**, *287*, 8174.
6. (a) H. Jin, R.C. Zangar, *Biomarker Insights* **2009**, *4*, 191; (b) K. Morino, K.F. Petersen, S. Dufour, D. Befroy, J. Frattini, N. Shatzkes, S. Neschen, M.F. White, S. Bilz, S. Sono, M. Pypaert, G.I. Shulman, *J. Clin. Invest.* **2005**, *115*, 3587; (c) C.X. Gong, F. Liu, I. Grundkeiqbal, K. Iqbal, *J. Neural. Transm.* **2005**, *112*, 813.
7. M. Audagnotto, M. Dal Peraro, *Comput. Struct. Biotechnol.* **2017**, *15*, 307.
8. C.N. Pang, E. Gasteiger, M.R. Wilkins, *BMC genomics* **2010**, *11*, 92.
9. M.A. Surani, K. Hayashi, P. Hajkova, *Cell* **2007**, *128*, 747.
10. K. Banno, I. Kisu, M. Yanokura, K. Tsuji, K. Masuda, A. Ueki, Y. Kobayashi, W. Yamagami, H. Nomura, E. Tominaga, N. Susumu, D. Aoki, *Int. J. Oncol.* **2012**, *41*, 793.
11. (a) B.E. Bernstein, M. Kamal, K. Lindblad-Toh, S. Bekiranov, D.K. Bailey, D.J. Huebert, S. McMahon, E.K. Karlsson, E.J. Kulbokass III, T.R. Gingeras, S.L. Schreiber, E.S. Lander, *Cell* **2005**, *120*, 169; (b) E.V. Benevolenskaya, *Biochem. Cell Biol.* **2007**, *85*, 435.
12. (a) T.A. Volpe, C. Kidner, I.M. Hall, G. Teng, S.I.S. Grewal, R.A. Martienssen, *Science* **2002**, *297*, 1833. (b) A. Barski, S. Cuddapah, K. Cui, T.-Y. Roh, D.E. Schones, Z. Wang, G. Wei, I. Chepelev, K. Zhao, *Cell* **2007**, *129*, 823; (c) J.A. Rosenfeld, Z. Wang, D.E. Schones, K. Zhao, R. DeSalle, M.Q. Zhang, *BMC Genomics* **2009**, *10*, 143.
13. (a) R.A. Varier, H.T. Timmers, *Biochim. Biophys. Acta.* **2011**, *1815*, 75; (b) J. Ellinger, P. Kahl, J. von der Gathen, S. Rogenhofer, L.C. Heukamp, I. Gütgemann, B. Walter, F. Hofstädter, R. Büttner, S.C. Müller, P.J. Bastian, A. von Ruecker, *The Prostate* **2010**, *70*, 61.

14. S.N. Thomas, K.E. Funk, Y. Wan, Z. Liao, P. Davies, J. Kuret, A.J. Yang, *Acta Neuropathol.* **2012**, *123*, 105.
15. K.D. Daze, F. Hof, *Acc. Chem. Res.* **2013**, *46*, 937.
16. H. Li, S. Ilin, W. Wang, E.M. Duncan, J. Wysocka, C.D. Allis, D.J. Patel, *Nature* **2006**, *442*, 91.
17. P.R. Nielsen, D. Nietlispach, H.R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A.G. Murzin, N.V. Murzina, E.D. Laue, *Nature* **2002**, *416*, 103.
18. S.A. Minaker, K.D. Daze, M.C.F. Ma, F. Hof, *J. Am. Chem. Soc.* **2012**, *134*, 11674.
19. R.E. McGovern, H. Fernandes, A.R. Khan, N.P. Power, P.B. Crowely, *Nat. Chem.* **2012**, *4*, 527.
20. Y. Liu, L. Perez, A.D. Gill, M. Mettry, L. Li, Y. Wang, R.J. Hooley, W. Zhong, *J. Am. Chem. Soc.* **2017**, *139*, 10964.
21. F. Guagnini, P.M. Antonik, M.L. Rennie, P. O'Byrne, A.R. Khan, R. Pinalli, E. Dalcanale, P.B. Crowley, *Angew. Chem. Int. Ed.* **2018**, *57*, 7126.
22. M.A. Gamal-Eldin, D.H. Macartney, *Org. Biomol. Chem*. **2013**, *11*, 488.
23. R. Pinalli, G. Brancatelli, A. Pedrini, D. Menozzi, D. Hernandez, P. Ballester, S. Geremia, E. Dalcanale, *J. Am. Chem. Soc.* **2016**, *138*, 8569.
24. M. Dionisio, G. Oliviero, D. Menozzi, S. Federici, R.M. Yebeutchou, F.P. Schmidtchen, E. Dalcanale, P. Bergese, *J. Am. Chem. Soc.* **2012**, *134*, 2392.
25. N. Bontempi, E. Biavardi, D. Bordiga, G. Candiani, I. Alessandri, P. Bergese, E. Dalcanale *Nanoscale* **2017**, *9*, 8639.
26. (a) S.T. Larda, M.P. Bokoch, F. Evanics, R.S. Prosser, *J. Biomol. NMR* **2012**, *54*, 199; (b) F. Guagnini, PhD Thesis, **2019**.
27. R.M. Yebeutchou, E. Dalcanale, *J. Am. Chem. Soc.* **2009**,*131*, 2452.
28. K.-D. Zhang, D. Ajami, J.V. Gavette, J. Rebek Jr, *J. Am. Chem. Soc.* **2014**, *136*, 5264.
29. V. Borisenko, N. Sansom, G.A. Woolley, *Biophys J.* **2000**, *78*, 1335.
30. C.A. Hunter, H.L. Anderson, *Angew. Chem. Int. Ed.* **2009**, *48*, 7488.
31. (a) D. Faivre, M. Bennet, *Nature* **2016**, *535*, 235; (b) A.-H. Lu, W. Schmidt, N. Matoussevitch, H. Bönnermann, B. Spliethoff, B. Tesche, E. Bill, W. Kiefer, F. Schüth, *Angew. Chem. Int. Ed.* **2004**, *116*, 4403; (c) S. Chikazumi, S. Taketomi, M. Ukita, M. Mizukami, H. Miyajima, M. Setogawa, Y. Kurihara, *J. Magn. Mater.* **1987**, *65*, 245; (d) G. Reiss, A. Hütten, *Nat. Mater.* **2005**, *4*, 725; (e) Y. Hu, S. Mignani, J.P. Majoral, M. Shen, X. Shi, *Chem. Soc. Rev.* **2018**, *47*, 1874; (f) M. Takafuji, S. Ide, H. Ihara, Z. Xu, *Chem. Mater*. **2004**, *1610*, 1977.
32. A. Ali, H. Zafar, M. Zia, I. ul Haq, A.R. Phull, J.S. Ali, A. Hussain, *Nanotechnol. Sci. Appl.* **2016**, *9*, 49.
33. S.F. Chin, M. Makha, C.L. Raston, M. Saunders, *Chem Comm*. **2007**, *19*, 1948.

34. A. Ulu, S.A.A. Noma, S. Koytepe, B. Ates, *Appl. Biochem. Biotech*. **2019**, *187*, 938.
35. (a) G.S. Parkinson, *Surf. Sci. Rep.* **2016**, *71*, 272; (b) T. Yamashita, P. Hayes, *J. Electron Spectros. Relat. Phenomena* **2006**, *152*, 6; (c) E. Paparazzo, *J. Electron Spectros. Relat. Phenomena* **2006**, *154*, 38; (d) T. Yamashita, P. Hayes, *J. Electron Spectros. Relat. Phenomena* **2006**, *154*, 41.
36. (a) A.P. Grosvenor, B.A. Kobe, M.C. Biesinger, N.S. McIntyre, *Surf. Interface Anal.* **2004**, *36*, 1564; (b) M.C. Biesinger, B.P. Payne, A.P. Grosvenor, L.W.M. Lau, A.R. Gerson, R.S.C. Smart, *Appl. Surf. Sci.* **2011**, *257*, 2717.
37. R.P. Gupta, S.K. Sen, *Phys. Rev. B* **1975**, *12*, 15.
38. (a) T. Yamashita, P. Hayes, *Appl. Surf. Sci.* **2008**, *254*, 2441; (b) N.A. Zubir, C. Yacou, J. Motuzas, X. Zhang, J.C. Diniz Da Costa, *Sci. Rep.* **2014**, *4*, 1.
39. (a) S. Garcia-Gil, A. Arnau, A. Garcia-Lekue, *Surf. Sci.* **2013**, *613*, 102; (b) Y. Liu, Z. Xu, J. Zhang, J. Cheng, M. Miao, D. Zhang, *Dye. Pigment.* **2019**, *170*, 107586.
40. I. Bidermane, J. Lüder, S. Boudet, T. Zhang, S. Ahmadi, C. Grazioli, M. Bouvet, J. Rusz, B. Sanyal, O. Eriksson, B. Brena, C. Puglia, N. Witkowski, *J. Chem. Phys.* **2013**, *138*, 234701.
41. W. Luo, D.Y. Zemlyanov, C.A. Milligan, Y. Du, L. Yang, Y. Wu, P.D. Ye, *Nanotechnology* **2016**, *27*, 434002.
42. A. Ghiami, M. Timpel, M.V. Nardi, A. Chiappini, P. Nozar, A. Quaranta, R. Verucchi, *J. Phys. Chem. C* **2020**, *124*, 6732.
43. B. Lee, F.M. Richards, *J. Mol. Bio.* **1971**, *55*, 379.
44. E. Westhof, D. Altschuh, D. Moras, A.C. Bloomer, A. Mondragon, A. Klug, M.H.V. van Regenmortel, *Nature* **1984**, *311*, 123.
45. J. Fuhrmann, K.W. Clancy, P.R. Thompson, *Chem. Rev.* **2015**, *155*, 5413.
46. N.M. Luscombe, R.A. Laskowski, J.M. Thornton, *Nucleic Acids Res.* **2001**, *29*, 2860.
47. (a) J. Najbauer, B.A. Johnson, A.L. Young, D.W. Aswad, *J. Biol. Chem.* **1993**, *268*, 10501; (b) S.C. Larsen, K.B. Sylvestersen, A. Mund, D. Lyon, M. Mullari, M.V. Madsen, J.A. Daniel, L.J. Jensen, M.L. Nielsen, *Science Signaling* **2016**, *9*, rs9.
48. R.S. Blanc, S. Richard, *Molecular Cell* **2017**, *65*, 8.
49. J.M. Aletta, J.C. Hu, *Biotechnol. Annu. Rev.* **2008**, *14*, 203.
50. J. Jarrols, C.C. Davies, *Trends in Mol. Med*. **2019**, *25*, 993.
51. (a) G.E. Rogers, D.H. Simmonds, *Nature* **1958**; *182*, 186; (b) G.E. Rogers, *Nature* **1962**; *194*, 1149.
52. G. Brancatelli, E. Dalcanale, R. Pinalli, S. Geremia, *Molecules* **2018**, *23*, 3368.
53. (a) D. Astruc, M.C. Daniel, J. Ruiz, *Chem. Commun.* **2004**, *23*, 2637; (b) U. Drechster, B. Erdogan, V.M. Rotello, *Chem. Eur. J.* **2004**, *10*, 5570.
54. A. Chen, M.J. Shapiro, *J. Am. Chem. Soc.* **1998**, *120*, 10258.

55. M.-V. Salvia, F. Ramadori, S. Springhetti, M. Diez-Castellnou, B. Perrone, F. Rastrelli, F. Mancin, *J. Am. Chem. Soc.* **2015**, *137*, 886.
56. A. Sreekumar, L.M. Poisson, T.M. Rajendiran, A.P. Khan, Q. Cao, J. Yu, B. Laxman, R. Mehra, R.J. Lonigro, Y. Li, M.K. Nyati, A. Ahsan, S. Kalyana-Sundaram, B. Han, X. Cao, J. Byun, G.S. Omenn, D. Ghosh, S. Pennathur, D.C. Alexander, A. Berger, J.R. Shuster, J.T. Wei, S. Varambally, C. Beecher, A.M. Chinnaiyan, *Nature* **2009**, *457*, 910.
57. E. Biavardi, C. Tudisco, F. Maffei, A. Motta, C. Massera, G.G. Condorelli, E. Dalcanale, *Proc. Natl. Acad. Sci. U. S. A*. **2012**, *109*, 2263.
58. I.R. Verly, A.B. van Kuilenburg, N.G. Abeling, S.M. Goorden, M. Fiocco, F.M. Vaz, M.M. van Noesel, C.M. Zwaan, G.L. Kaspers, J.H. Merks, H.N. Caron, G.A. Tytgat, *Eur. J. Cancer* **2017**, *72*, 235.
59. T.D.W. Claridge, *High resolution NMR techniques in Organic Chemistry*, **2016**, Elsevier ed., Nederland.
60. B. Perrone, S. Springhetti, F. Ramadori, F. Rastrelli, F. Mancin, *J. Am. Chem. Soc.* **2013**, *135*, 11768.

APPENDIX A

*1. TGA thermogram of MeCav **9***

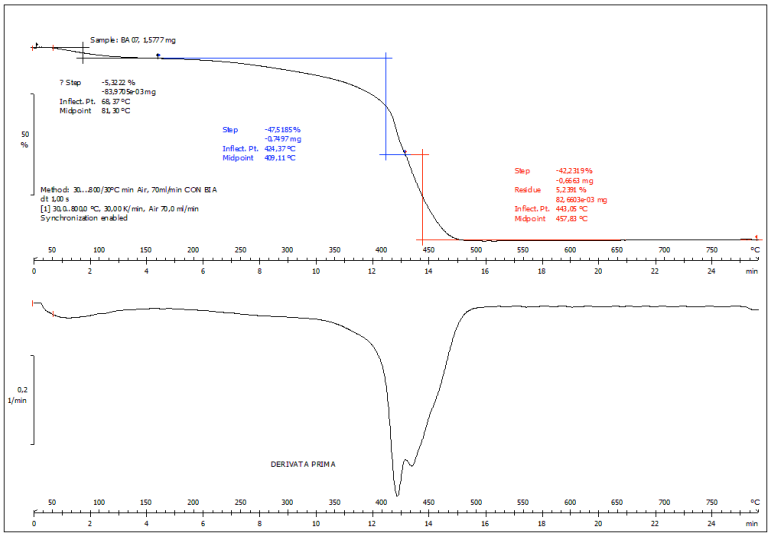


Figure A.1. TGA thermogram of MeCav **9**.

*2. TGA thermogram of MeCav@FeNPs **10** after the stability test*

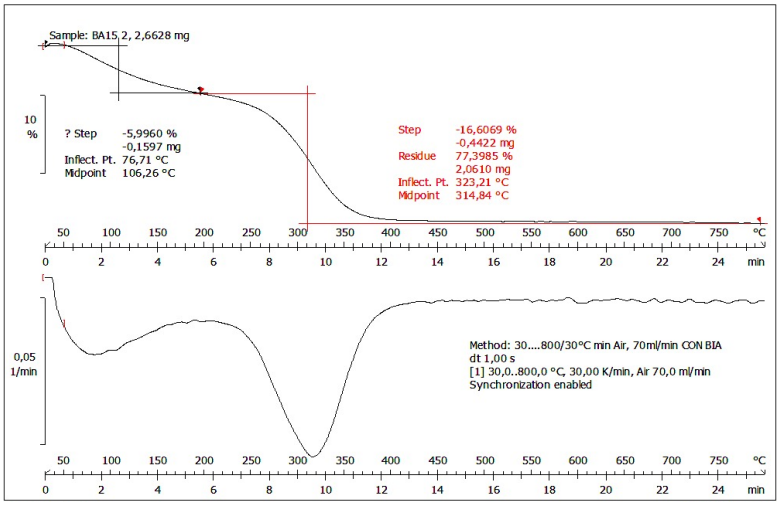


Figure A.2. TGA thermogram of MeCav@FeNPs **10** after the stability test.

*3. TGA thermogram of POCav **15***

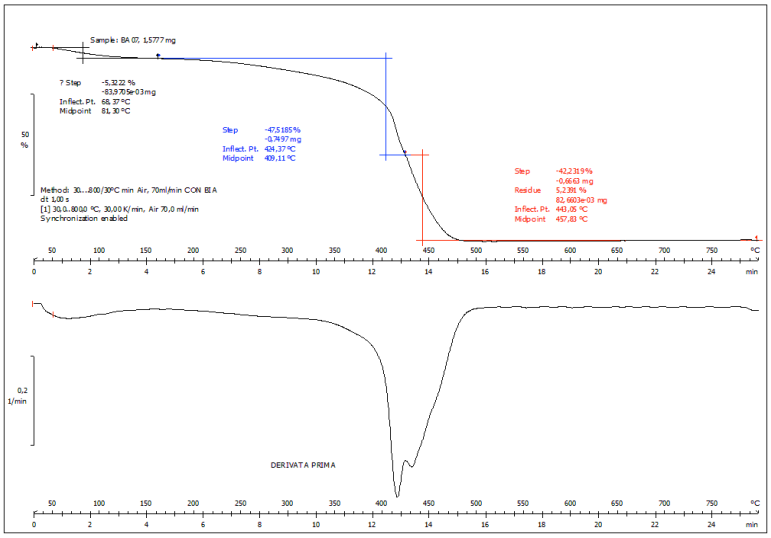


Figure A.3. TGA thermogram of POCav **15**.

*4. TGA thermogram of POCav@FeNPs **16** after the stability test*

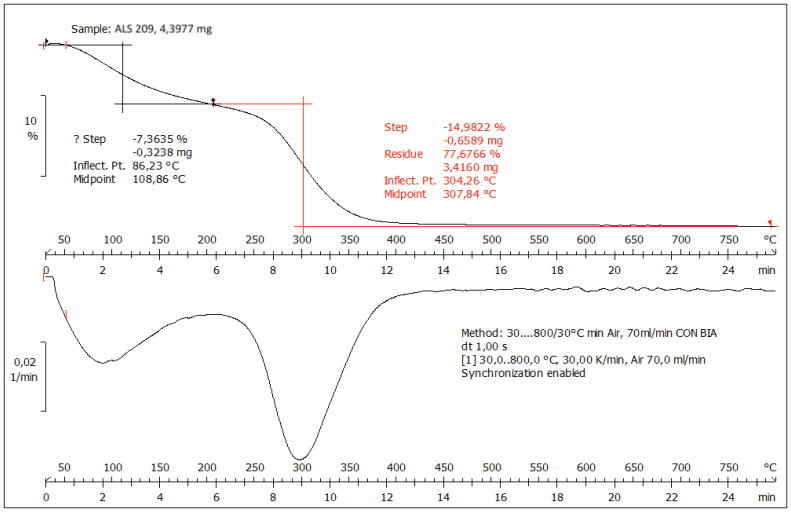


Figure A.4. TGA thermogram of POCav@FeNPs **16** after the stability test.

APPENDIX B

Mass spectrometric peptide mapping after enzymatic digestion with Trypsin of chemical modified HEWL.
The data presented are a comparison between three batches (**5**, **6** and **7**) of reductive amination of lysines residues.

Lys: blank experiment with only native Lysozyme
FeNP: control experiment with bare FeNPs
Tiiii: experiment carried out adding Tiiii receptor **6**
MeCav@FeNPs: control system
POCav@FeNPs: working system

Amino acidic sequence (FASTA) of HEWL:
KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT QATNRNTDGS TDYGILQINS RWWCNDGRTP GSRNLCNIPC SALLSSDITA SVNCAKKIVS DGNGMNAWVA WRNRCKGTDV QAWIRGCRL

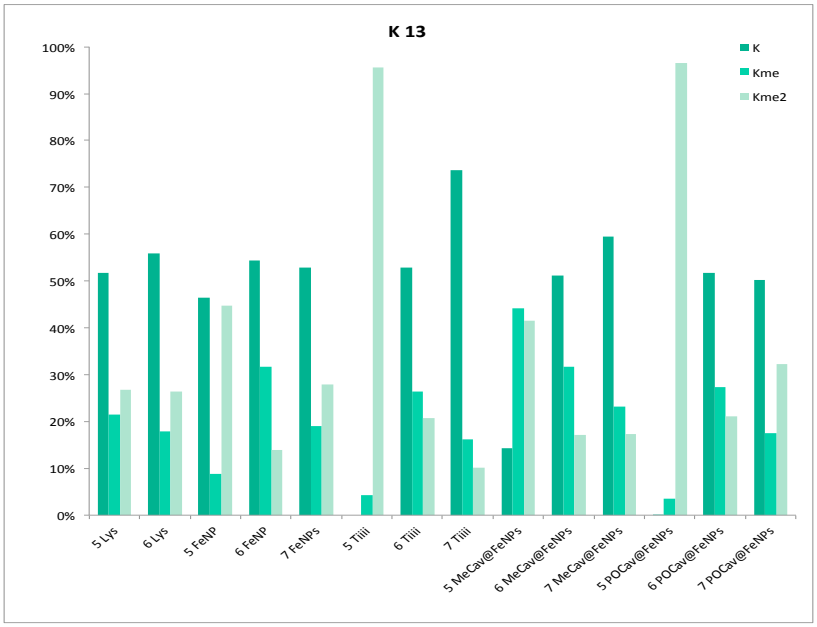


Figure B.1. Mass spectrometric methylation mapping of residue K13.

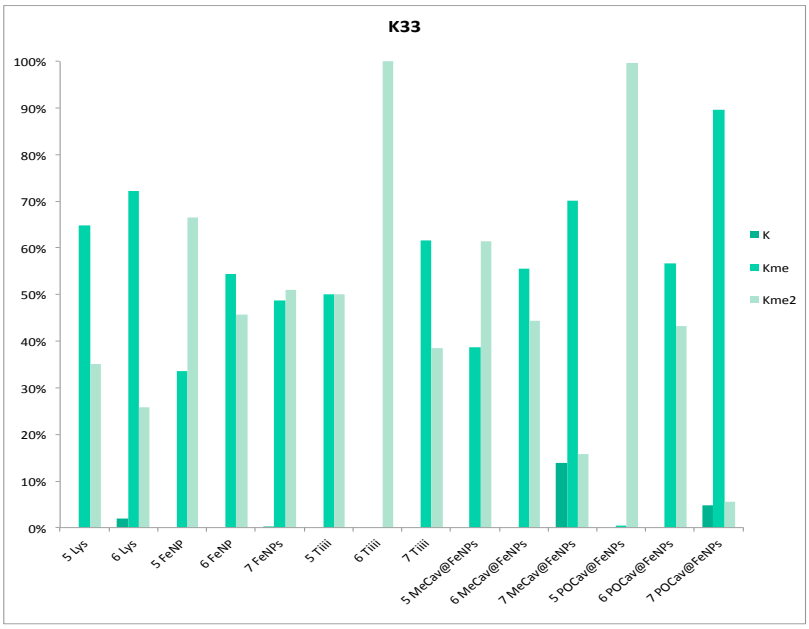


Figure B.2. Mass spectrometric methylation mapping of residue K33.

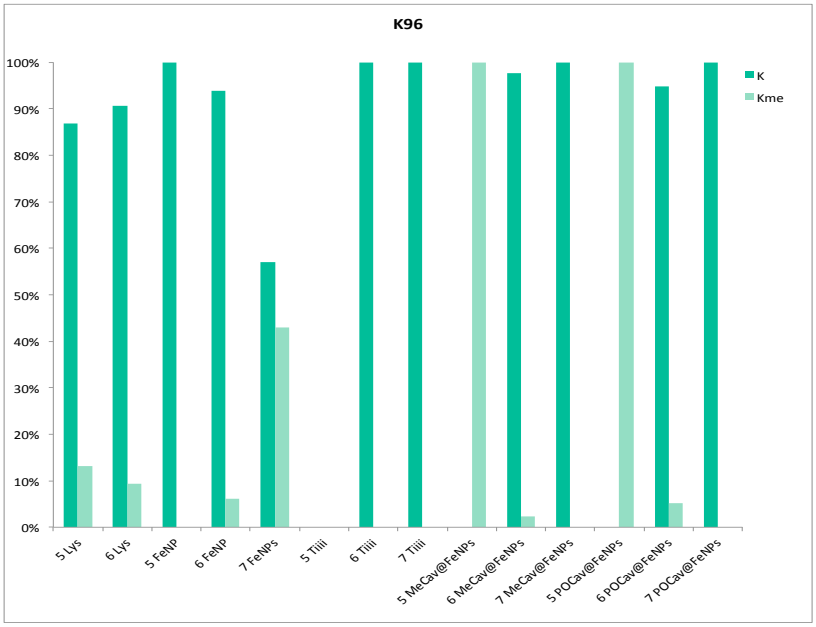


Figure B.3. Mass spectrometric methylation mapping of residue K96.

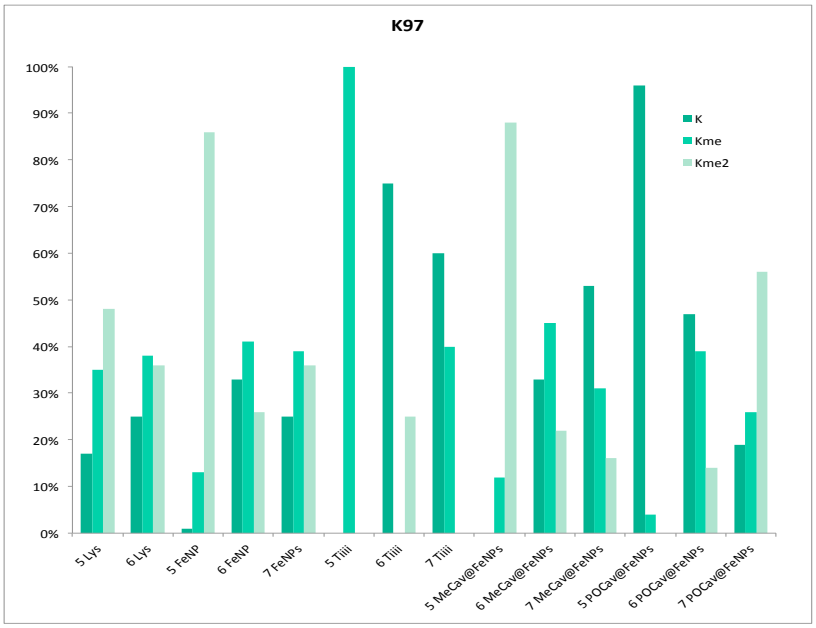


Figure B.4. Mass spectrometric methylation mapping of residue K97.

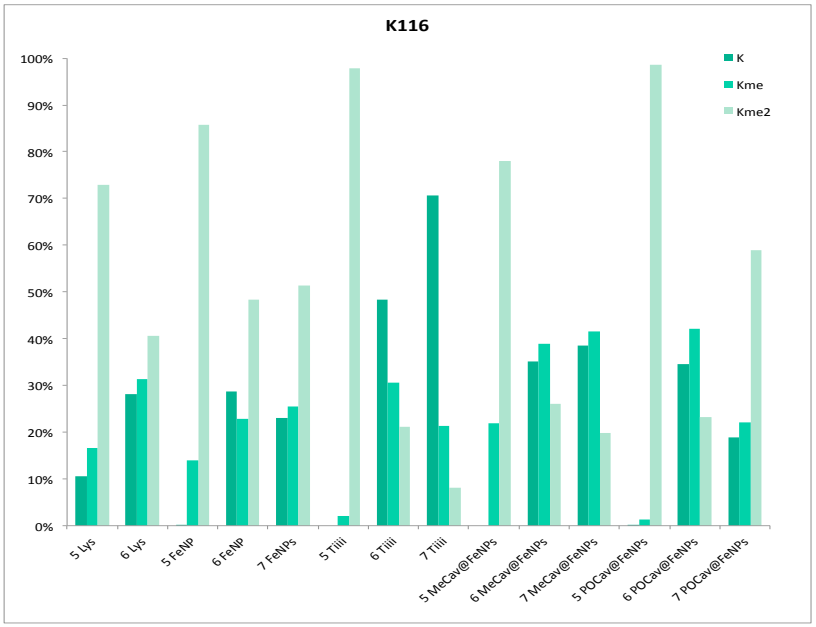


Figure B.5. Mass spectrometric methylation mapping of residue K116.

APPENDIX C

1. ITC titrations in MeOH

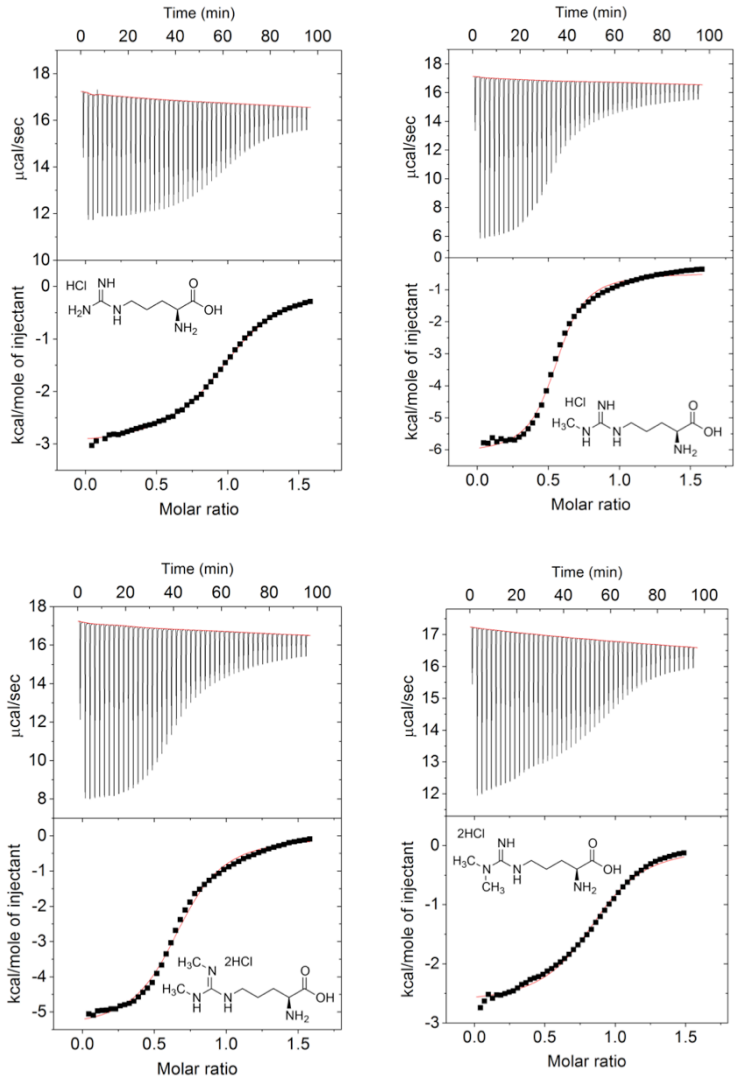


Figure C.1. ITC experiments of **18** and a) **Arg**, b) **MMA**, c) **SDMA**, d) **ADMA**. All experiments were conducted in MeOH; [cavitand] = 3.25 mM; [***R***.HCl] = 25 mM.

2. ITC titrations in water

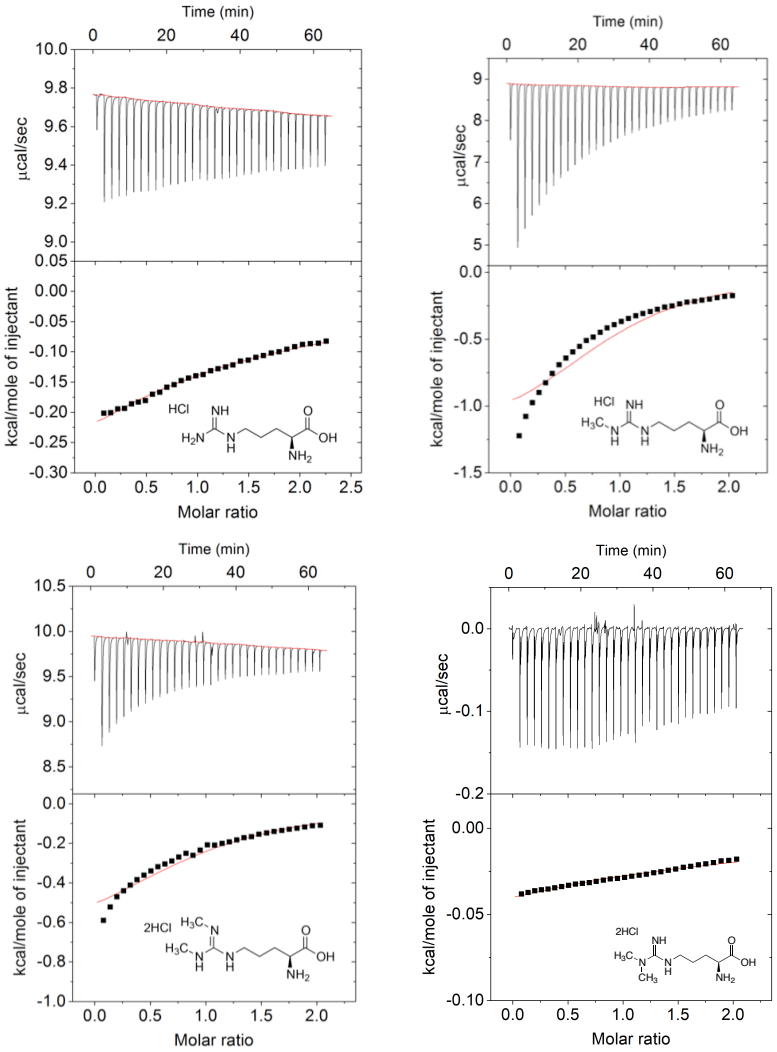


Figure C.2. ITC experiments of **6** and a) **Arg**, b) **MMA**, c) **SDMA**, d) **ADMA**. All experiments were conducted in H_2O; [cavitand] = 3.6 mM; [***R***.HCl] = 36 mM.

3. NMR titrations in CD_3OD

Cavitand **18** + **Ac-Arg** (^{1}H NMR)

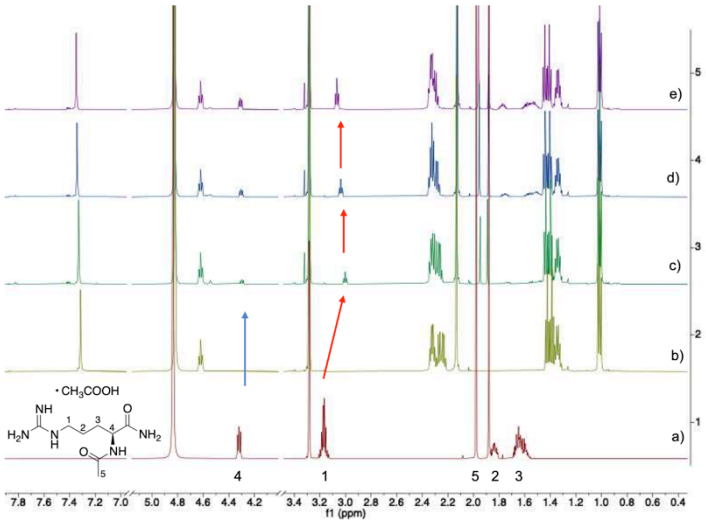


Figure C.3. ^{1}H NMR (600 MHz, CD_3OD, 298K) of a) **Ac-Arg**, b) **18**, c) **18** + 0.2 eq of **Ac-Arg**, d) **18** + 0.5 eq of **Ac-Arg**, e) **18** + 1.0 eq of **Ac-Arg**, f) **18** + 1.5 eq of **Ac-Arg**.

Cavitand **8** + **Ac-Arg** (^{31}P NMR)

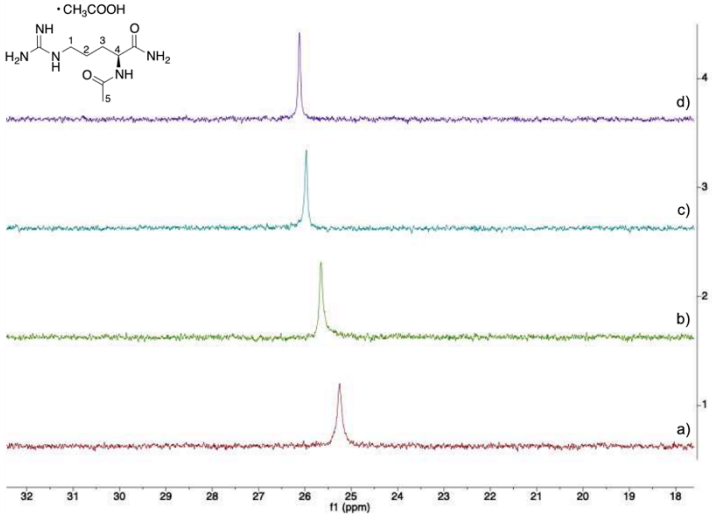


Figure C.4. ^{31}P NMR (243 MHz, CD_3OD, 298K) of a) **18**, b) **18** + 0.5 eq of **Ac-Arg**, c) **18** + 1.0 eq of **Ac-Arg**, d) **18** + 1.5 eq of **Ac-Arg**.

Cavitand **18** + **MMA** (^{1}H NMR)

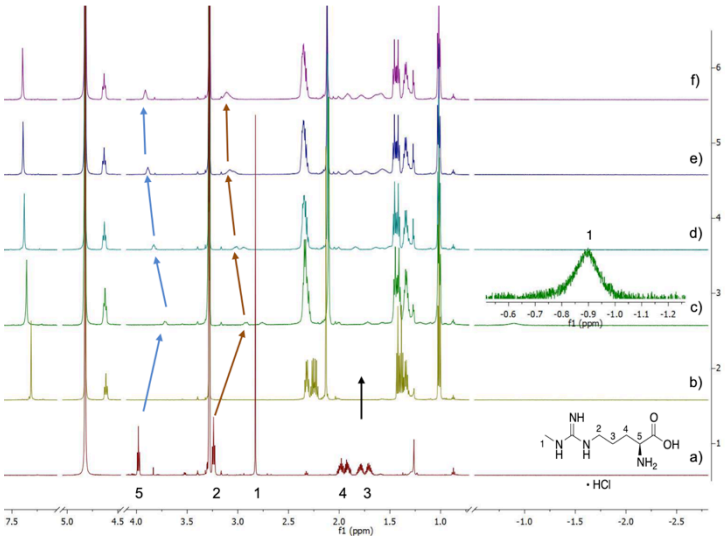


Figure C.5. ^{1}H NMR (600 MHz, CD_3OD, 298K) of a) **MMA**, b) **18**, c) **18** + 0.2 eq of **MMA**, d) **18** + 0.5 eq of **MMA**, e) **18** + 1.0 eq of **MMA**, f) **18** + 1.5 eq of **MMA**.

Cavitand **18** + **MMA** (^{31}P NMR)

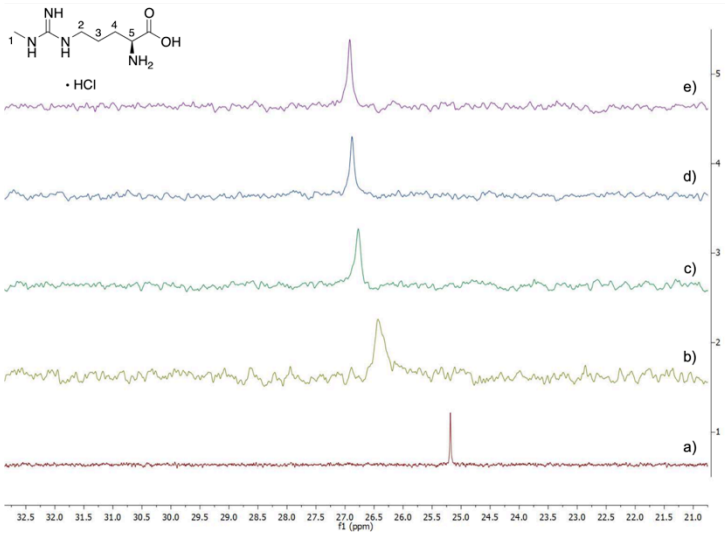


Figure C.6. ^{31}P NMR (243 MHz, CD_3OD, 298K) of a) **18**, b) **18** + 0.5 eq of **MMA**, c) **18** + 1.0 eq of **MMA**, d) **18** + 1.5 eq of **MMA.**

Cavitand **18** + **SDMA** (^{1}H NMR)

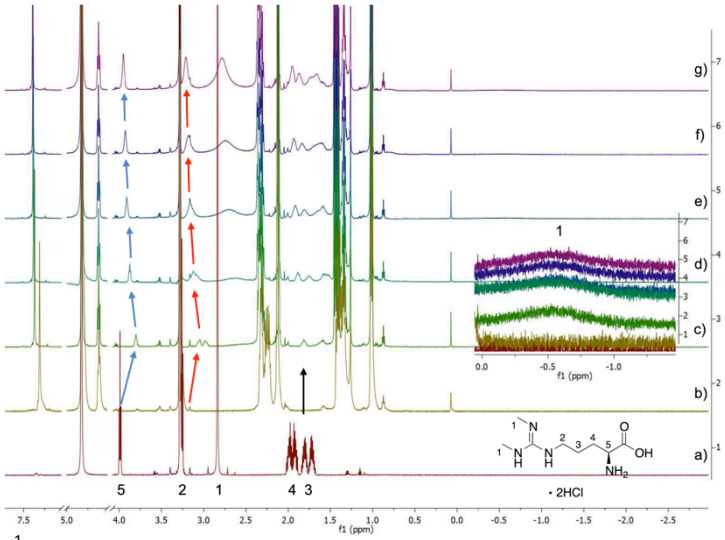


Figure C.7. ^{1}H NMR (600 MHz, CD_3OD, 298K) a) **SDMA**, b) **18**, c) **18** + 0.2 eq of **SDMA**, d) **18** + 0.5 eq of **SDMA**, e) **18** + 1.0 eq of **SDMA**, f) **18** + 1.5 eq of **SDMA**.

Cavitand **18** + **SDMA** (^{31}P NMR)

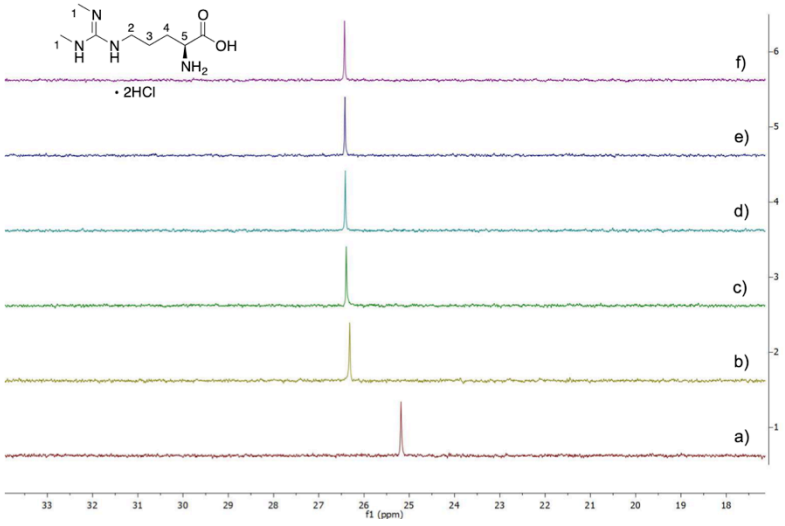


Figure C.8. ^{31}P NMR (243 MHz, CD_3OD, 298K) of a) **18**, b) **18** + 0.5 eq of **SDMA**, c) **18** + 1.0 eq of **SDMA**, d) **18** + 1.5 eq of **SDMA.**

Cavitand **18** + **ADMA** (^{1}H NMR)

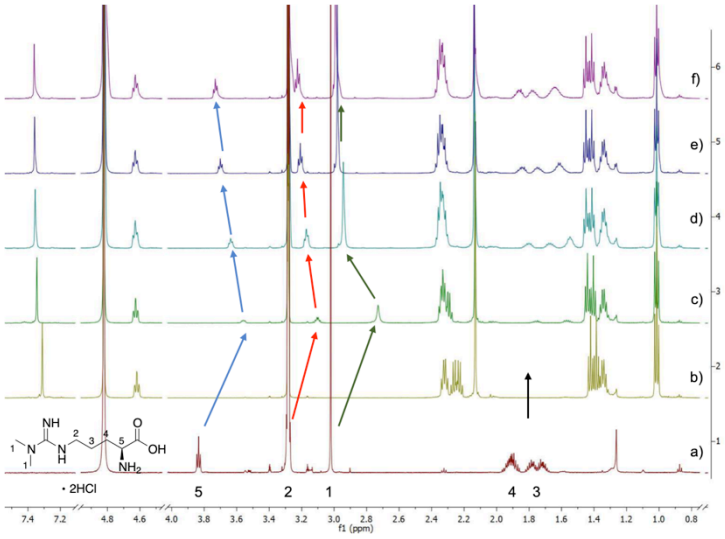


Figure C.9. ^{1}H NMR (600 MHz, CD_3OD, 298K) a) **ADMA**, b) **18**, c) **18** + 0.2 eq of **ADMA**, d) **18** + 0.5 eq of **ADMA**, e) **18** + 1.0 eq of **ADMA**, f) **18** + 1.5 eq of **ADMA**.

Cavitand **18** + **ADMA** (^{31}P NMR)

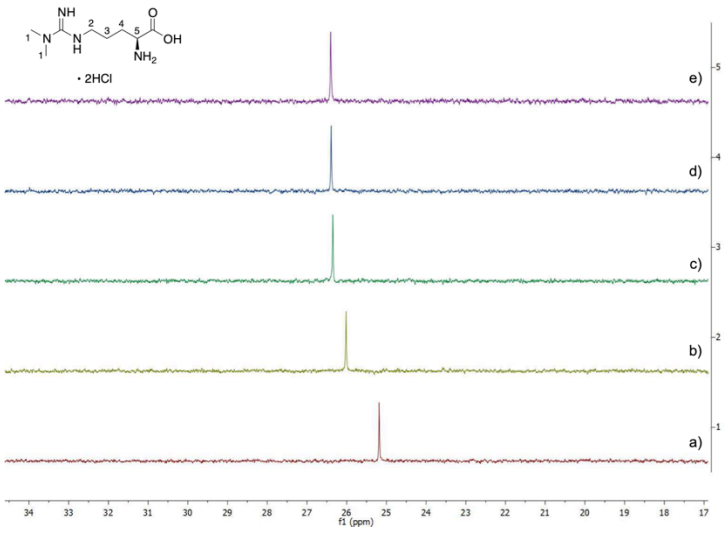


Figure C.10. ^{31}P NMR (243 MHz, CD_3OD, 298K) of a) **18**, b) **18** + 0.5 eq of **ADMA**, c) **18** + 1.0 eq of **ADMA**, d) **18** + 1.5 eq of **ADMA.**

4. NMR titrations in D_2O

Cavitand **6** + **Arg** (^{1}H NMR)

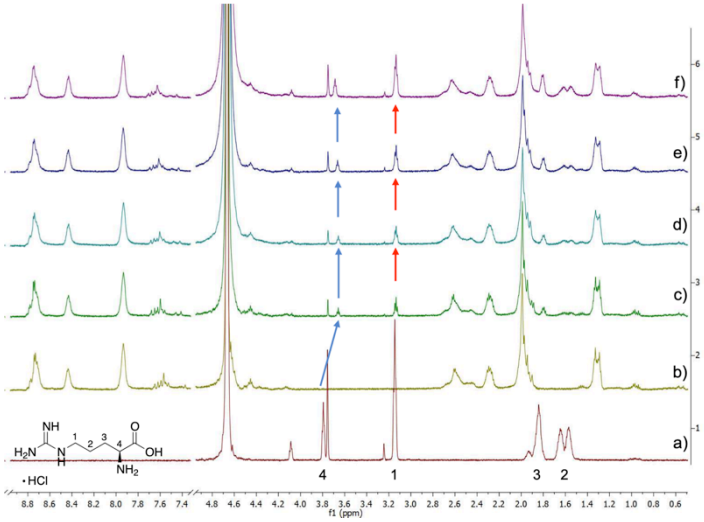


Figure C.11. ^{1}H NMR (600 MHz, D_2O, 298K) of a) **Arg**, b) **6**, c) **6** + 0.2 eq of **Arg**, d) **6** + 0.5 eq of **Arg**, e) **6** + 1.0 eq of **Arg**, f) **6** + 1.5 eq of **Arg**.

Cavitand **6** + **Arg** (^{31}P NMR)

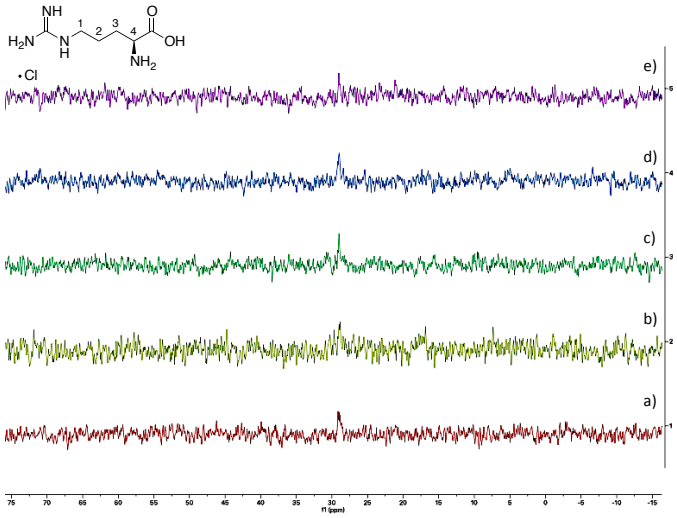


Figure C.12. ^{31}P NMR (243 MHz, D_2O, 298K) of a) **6**, b) **6** + 0.5 eq of **Arg**, c) **6** + 1.0 eq of **Arg**, d) **6** + 1.5 eq of **Arg**.

Cavitand **6** + **MMA** (^{1}H NMR)

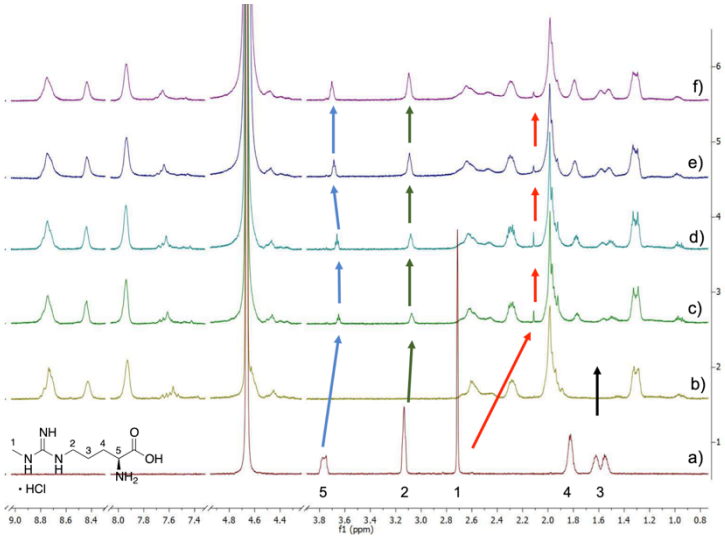


Figure C.13. ^{1}H NMR (600 MHz, D_2O, 298K) of a) **MMA**, b) **6**, c) **6** + 0.2 eq of **MMA**, d) **6** + 0.5 eq of **MMA**, e) **6** + 1.0 eq of **MMA**, f) **6** + 1.5 eq of **MMA**.

Cavitand **6** + **MMA** (^{31}P NMR)

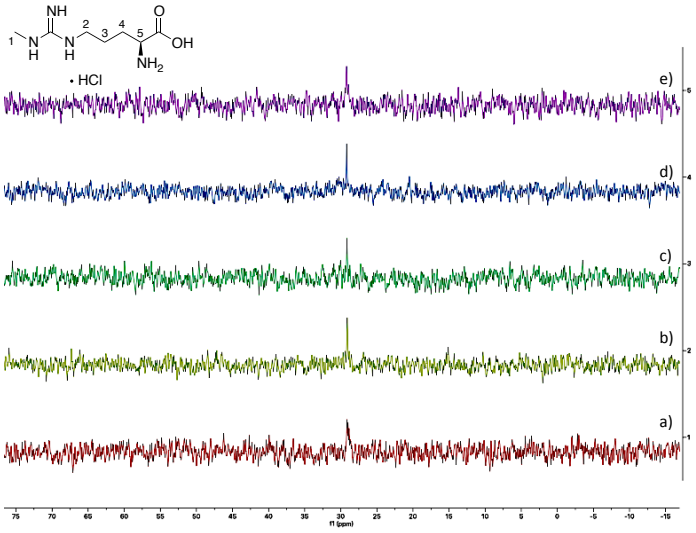


Figure C.14. ^{31}P NMR (243 MHz, D_2O, 298K) of a) **6**, b) **6** + 0.5 eq of **MMA**, c) **6** + 1.0 eq of **MMA**, d) **6** + 1.5 eq of **MMA**.

Cavitand **6** + **SDMA** (^{1}H NMR)

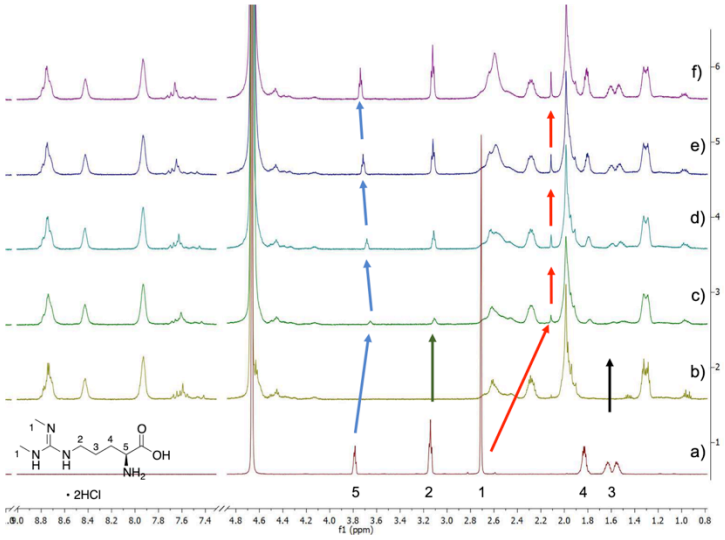


Figure C.15. ^{1}H NMR (600 MHz, D_2O, 298K) of a) **SDMA**, b) **6**, c) **6** + 0.2 eq of **SDMA**, d) **6** + 0.5 eq of **SDMA**, e) **6** + 1.0 eq of **SDMA**, f) **6** + 1.5 eq of **SDMA**.

Cavitand **6** + **SDMA** (^{31}P NMR)

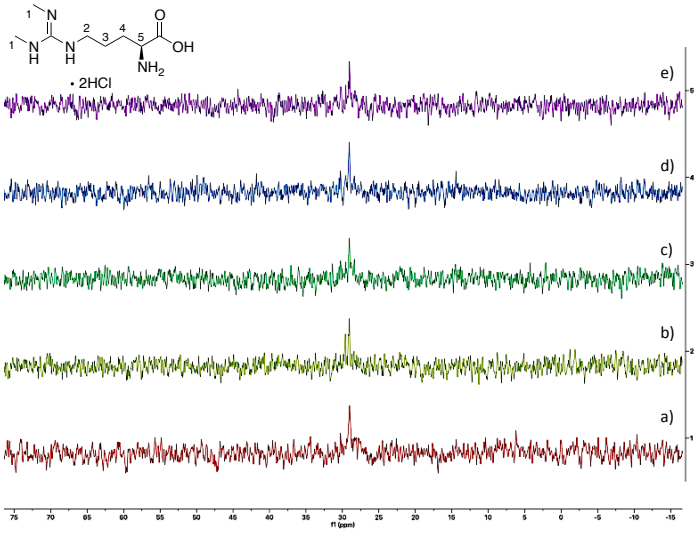


Figure C.16. ^{31}P NMR (243 MHz, D_2O, 298K) of a) **6**, b) **6** + 0.5 eq of **SDMA**, c) **6** + 1.0 eq of **SDMA**, d) **6** + 1.5 eq of **SDMA**.

Cavitand **6** + **ADMA** (^{1}H NMR)

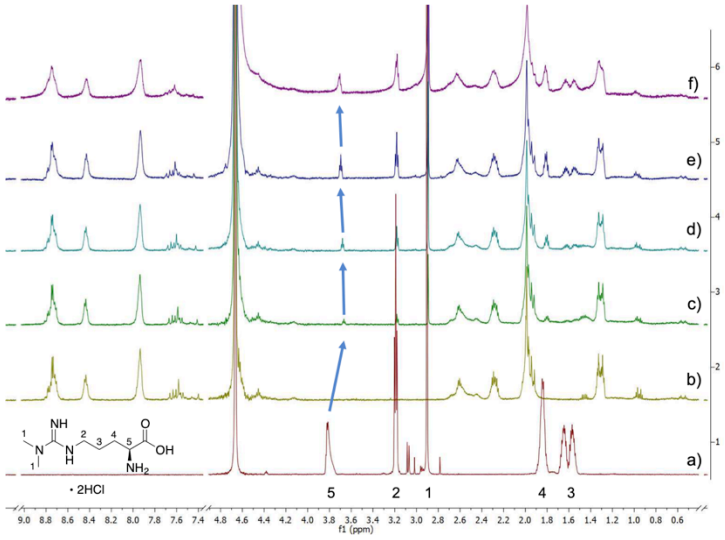


Figure C.17. ^{1}H NMR (600 MHz, D_2O, 298K) of a) **ADMA**, b) **6**, c) **6** + 0.2 eq of **ADMA**, d) **6** + 0.5 eq of **ADMA**, e) **6** + 1.0 eq of **ADMA**, f) **6** + 1.5 eq of **ADMA**.

Cavitand **6** + **ADMA** (^{31}P NMR)

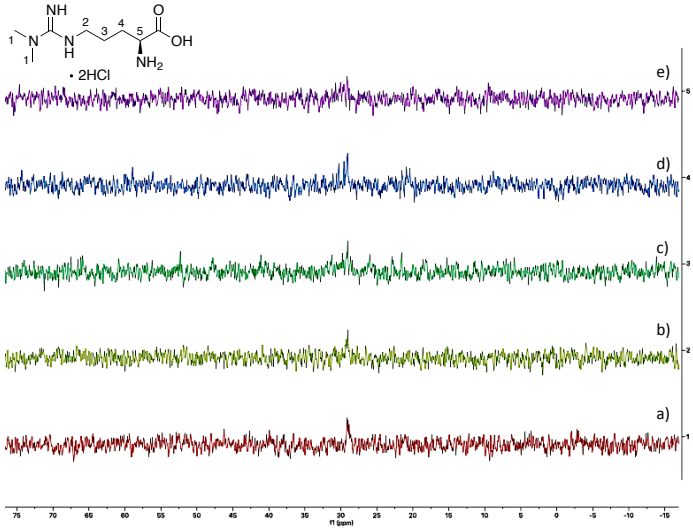


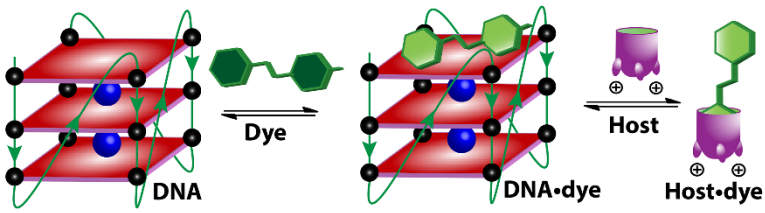
Figure C.18. ^{31}P NMR (243 MHz, D_2O, 298K) of a) **6**, b) **6** + 0.5 eq of **ADMA**, c) **6** + 1.0 eq of **ADMA**, d) **6** + 1.5 eq of **ADMA**.

CHAPTER 3

Selective discrimination and classification of G-Quadruplex structures with arrayed Host:Guest sensors*

*Part of this chapter has been submitted to Nature Chemistry: J. Chen, B.L. Hickey, L. Wang, J. Lee, A.D. Gill, A. Favero, R. Pinalli, E. Dalcanale, R.J. Hooley, W. Zhong, it is presently **under 2nd revision**.

An arrayed suite of synthetic hosts and dyes is capable of fluorescence detection of oligonucleotide secondary structures. By using cationic dyes that show affinity for both DNA G-quadruplexes and the synthetic hosts, multiple recognition mechanisms can be exploited to create a unique sensing fingerprint, consisting of different fluorescence enhancements in the presence of select DNA strands. Multivariate analysis of these sensing fingerprints enables discrimination between highly similar G4 structures of identical length and topological type. The selectivity is excellent: different G4s that display the same folding topology can be easily differentiated by the number of G-quartets and sequence differences at the 3' or 5' ends. Most importantly, the array is capable of both differentiation and classification of the G4 structures at the same time, in a simple, non-invasive manner. This sensing method does not require the discovery and synthesis of specific G4 binding ligands, but employs a simple multicomponent approach to ensure wide applicability. Cation-driven conformational changes in G4 folding topology can be detected by the arrayed sensor. It is even capable of detecting changes in G4 folding pattern in different types of complex media, including fetal bovine serum, and can selectively detect changing concentrations of G4s in mixtures of multiple nucleotides.



3.1 INTRODUCTION

The cellular functions of nucleic acids are dependent not only on the nucleotide sequence, but also on the secondary structural architecture. Nucleotide strands can form complex 3D structures, which control interactions with proteins and small molecule ligands, expanding their function beyond encoding and transferring genetic information in biological systems.[1] For example, uniquely folded DNA such as G-quadruplexes have been found in patient-derived cancer tissues and viral genomes, making them vital targets for the development of new therapeutic interventions.[2] In addition, synthetic DNA and RNA nanostructures have a wide variety of applications in materials science, synthetic biology, and nanomedicine.[3,4] As such, there is a need for methods that can quickly and easily assess the structure of nucleotides of various types, not just the sequence, which is now routine. Although DNA folding directed by Watson–Crick base-pairing can be predicted,[5] it remains challenging to analyze non-canonical structures (formed through other types of H- bonding, such as Hoogsteen base pairings and hydrophobic base stacking) from sequences alone. Determination of nucleic acid 3D structures can be achieved in different ways: complete structural analysis requires X-ray crystallography and/or multidimensional NMR spectroscopy,[6,7] whereas simple grouping into secondary structural types is possible with Circular Dichroism (CD) spectroscopy.[8] These techniques have their challenges, in that they are either too time-consuming and detailed for rapid analysis (X-Ray, NMR), or provide too little information (CD). A rapid, simple method that can selectively identify, differentiate and classify nucleotide secondary structures is highly desirable, but also very challenging, and applying this sensor in complex biological media towards sensing changes in nucleotide 3D structure is even more so.

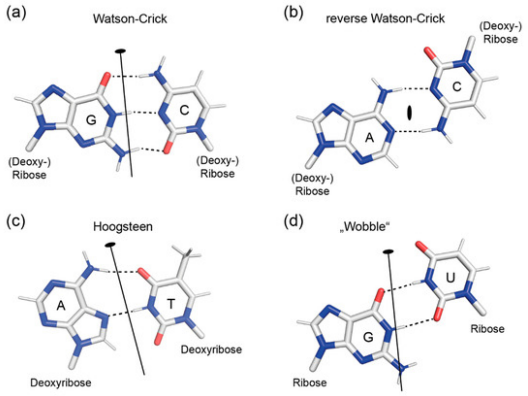


Figure 3.1. G:C (a) and reverse A:C (b) Watson–Crick base pairs, A:T Hoogsteen (c) and G:U wobble (d) base pairs. Taken from U. Heinemann, Y. Roske, *Symmetry* **2020**, *12*, 737.

There have been some exquisite examples of this type of detection used to analyze oligonucleotide secondary structures, most notably for RNA. This can be achieved with synthetic, derivatized nucleotides: covalent incorporation of dye molecules into individual bases, followed by chemical synthesis of the desired, labeled oligonucleotide allows optical methods to be used to differentiate nucleic acid structures and monitor induced changes in oligonucleotide structure.[9] The small changes in structure necessitate the use of differential sensing, which was popularized by Anslyn for small molecule analytes,[10] and has been used by others to classify the 2D and 3D structure of derivatized RNAs.[11] By incorporating dyes at multiple different positions on the RNA backbone and adding multiple small molecule effectors, array-based sensing can be exploited to classify different structural motifs and conformational changes, using multivariate analysis tools such as Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA).[12] Some beautiful examples of RNA classification have been shown, but they are not an effective method for monitoring native, *unmodified* oligonucleotide structures. That requires an exogenous fluorophore, which introduces new challenges: the reporter must show selective responses to nucleotides, and selectivity for different structural motif, which is not simple. Some excellent work has been performed that uses fluorescence displacement assays to characterize different types of DNA-binding dyes or ligands, with DNA structures that show quite large structural differences (duplex, i-motif, G4, etc.).[13]

The greatest challenge for any sensor is to distinguish between highly similar structures in a complex mixture. DNA G-Quadruplexes (G4s) are an excellent example of a challenging recognition target. They are a common non-canonical DNA structural motif involved in gene expression and genome stability, as well as being important building blocks in nucleic acid-based nanostructures.[14,15] G4s involve stacked guanine quartets (**Figure 3.2**), assembled *via* Hoogsteen base-pair contacts and intercalated Na^+ or K^+ ions linked by loop nucleotides.[16] Diverse G4 motifs can fold into a number of different secondary structures, with variations in G-quartet stacking, orientation of the G strands, and loop locations.[17] G-quadruplexes are generally classified into three different structural types, based on the directionality of the four G strands forming the G-quartet: parallel, antiparallel or hybrid (illustrated in **Figure 3.2** and **Figure 3.3**). These structural motifs are a challenging and enticing target for structural analysis, as the overall size and the internal structure are relatively conserved.

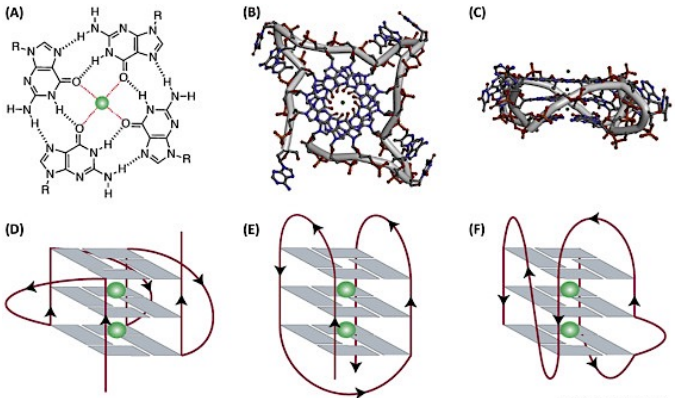


Figure 3.2. (A) G4 structure formed by Hoogsteen hydrogen-bonded guanines with cations (in green); Top view (B) and side view (C) of human telomeric G4s (PDB file: 1KF1); Schematic representation of G4s strand directions: parallel (D), anti-parallel (E) and hybrid (F). Taken from J.Spiegel, S. Adhikari, S. Balasubramanian, *Trends Chem.* **2020**, *2*, 123.

As in dsDNA, G4 structures have grooves, the dimensions of which are determined by the overall topology and the nature of the loops. The variations lie in the strand orientation, glycosidic angles and the directionality of the external loops. These changes are not easy to detect:

they are, after all, just oligomers of phosphorylated sugars oriented in different directions. CD spectroscopy can disclose the *types* of topology, (i.e. parallel, antiparallel, or hybrid), but cannot recognize small structural differences in G4s of the same folding type. Although computational prediction and genomic mapping have revealed hundreds of thousands of sequence motifs that could form G4 structures,[18,19] conformational studies have only been conducted on a few of these sequences. This has involved the use of small molecule ligands, which recognize G4s mainly through π-π stacking with the G tetrads, plus binding to the grooves or loops.[20,21] Some ligands are highly selective,[21] but they are not common, and not easily applied in sensing applications.

Here we show that small-molecule host:guest sensor arrays can be used to sense structural differences in unmodified DNA G-quadruplex structures, allowing both their differentiation and classification *via* simple fluorescence measurements and pattern recognition. Supramolecular probes are perfectly suited to array-based pattern recognition due to their simple synthesis, ease of use, and rapid access to multiple variables,[22] which is essential in detecting small changes in structure. Host:guest sensor arrays have been used to monitor enzyme reactions[23,24] and analyze peptide and protein modifications.[25,26] They have rarely been applied to nucleotide sensing, probably because synthetic hosts often show poor affinity for oligonucleotides.[27] The solution to this challenge lies in indirect sensing: by combining an arrayed suite of host molecules with multiple dye candidates that can bind *both* the hosts and the target oligonucleotides, small changes in the target structure can be selectively detected. This also removes the need for the discovery and synthesis of specific ligand targets for each oligonucleotide type, which allows wider scope in sensing applications.

3.2 RESULTS AND DISCUSSION

Pattern recognition-based sensing of oligonucleotide structure requires multiple different hosts and dyes that have variable affinity for each other and/or with the target oligonucleotide system, as well as easily variable fluorescence responses upon binding. The basic design principle is shown in **Figure 3.3 a**. The sensor components consist of cationic dye molecules that can bind to DNA structures, as well as a series of synthetic host molecules that can also competitively bind these dyes while modulating their fluorescence response. The fluorophores are the styrylpyridinium dyes **DSMI** and **PSMI**,[28,29] which are good candidates for DNA ligands, as they are cationic, yet water-soluble, and are similar in structure to known oligonucleotide ligands.[30] They vary only in the size of headgroup (dimethylamino vs piperidinyl) and have highly similar fluorescence properties (emission maxima at 595 nm and 615 nm for **DSMI** and **PSMI** respectively). The slightly varying size and shape should confer small differences in affinity for nucleotides, while maintaining similar detection ranges.

The second component of the array is provided by water-soluble host molecules that can bind to the two fluorophores and modulate their emission. To allow maximal differentiation between targets, we employed a series of 5 different cavitand hosts **1-5** (**Figure 3.3 b**), all of which should show affinity for the dyes, allowing a competition between the G4 and the host. These sensor components can then be combined in an array for pattern recognition-based fluorescence analysis. Tetraanionic host **1**[31] is known to bind cationic dyes (including **DSMI**[28] and **PSMI**[29]), and has been extensively used in array-based biosensing applications.[32,33] In addition, we synthesized three cationic variants. Octamide cavitand **2**[29] displays a similar cavity size to **1**, but displays imidazolium groups at the lower rim for water solubility. Similarly, the benzimidazole cavitands **3** and **4**[34] are derivatives of **1** that display cationic groups at the lower rim. Finally, the shallow phosphonate cavitand **5**[35] was used, as this host is well-known to bind N- methylammonium and N-methylpyridinium salts.[36]

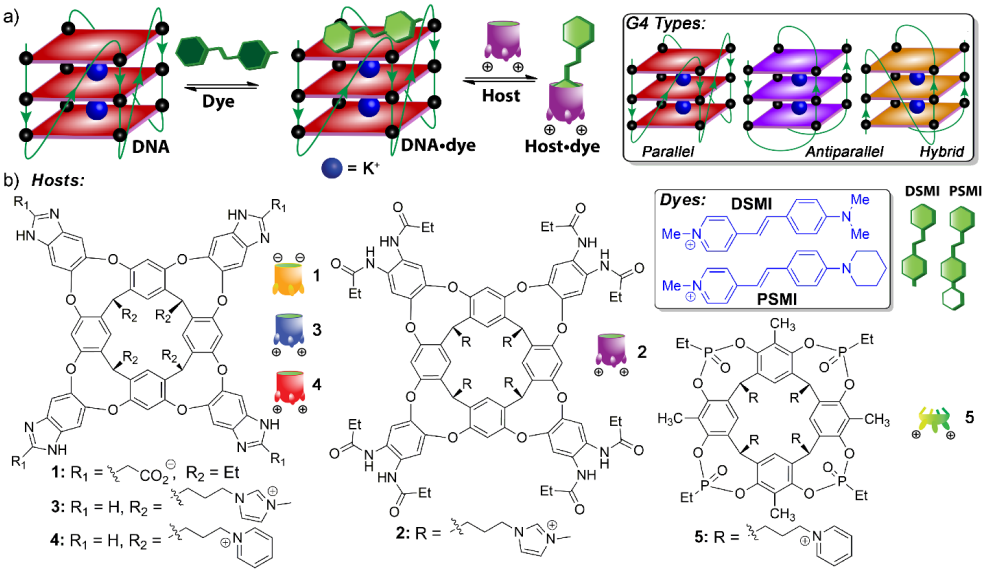


Figure 3.3. A host:guest sensor system for label-free classification and differentiation of G-quadruplex structures. (a) Illustration of the supramolecular sensing concept, and representation of the common G4 structural motifs targeted; (b) Structures of hosts **1**-**5** and dyes used.

The initial test was to determine whether the concept was correct: could the dyes show variable response to G-quadruplexes in the presence of the different hosts? The two dyes were added to a suite of DNA structures, followed by increasing concentrations of hosts **1-5**, and the fluorescence response curves recorded. Eight different DNA structures were initially tested, four single-stranded 20-mer oligonucleotide DNAs (A20, G20, T20 and C20), and four representative G4 structures containing 15-24 bases. Among them, *c-kit 1* and *c-myc 2345* are parallel G4 quadruplexes, *bcl-2 2345* forms a hybrid structure, and *TBA* exists in an antiparallel conformation.[37] Gel electrophoresis confirmed that the G4 strands all formed unimolecular folded structures and their folding topologies were all validated by CD (see Appendix **A**).

Both **PSMI** and **DSMI** (0.625 μM each) showed detectable fluorescence enhancements upon addition to all the DNA structures ([DNA] = 0.1 μM, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4). More specifically, titrating 0-20 μM **DSMI** or **PSMI** dye to *c-myc 2345* or A20 showed that the fluorescence increased up to 3-6 fold within a narrow concentration range

(from 0.08 to 1.25 µM dye) (**Figure A.17**). The fluorescence enhancements of the individual dyes with different G4 structures were not substantially different, so simple addition of the two dyes by themselves to the DNAs is not sufficient for selective sensing of different structures. However, upon addition of the cavitands (**1-5**, 0-8 µM) to the dye:DNA mixtures (0.625 µM dye, 0.1 µM DNA), significant variations in emission properties can be immediately observed (see **Figure 3.4 a-c** and **Figure A.18-19**).

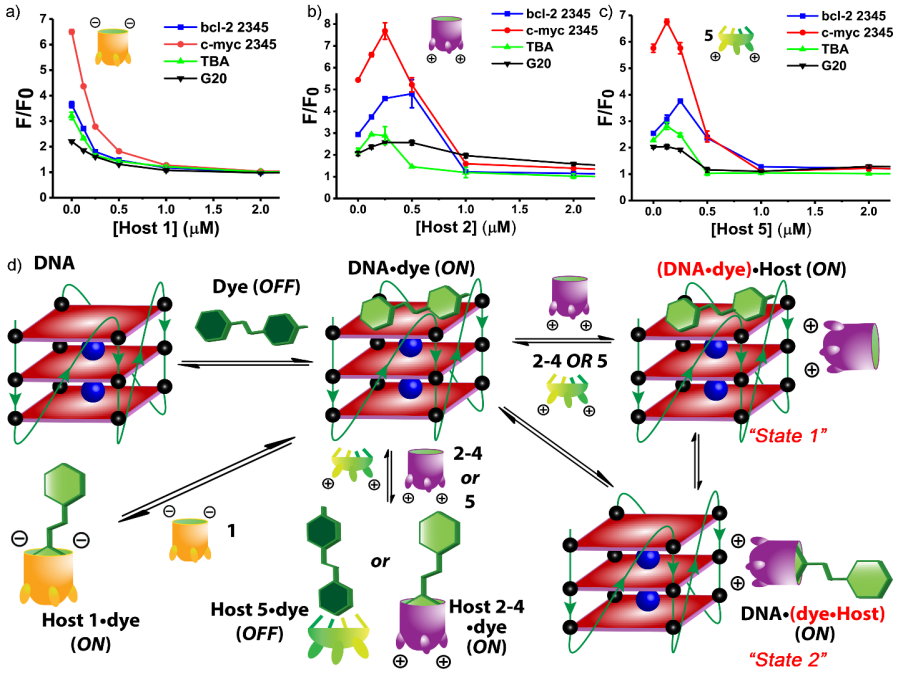


Figure 3.4. Host:guest complexes as sensors for DNA targets.
Fluorescence response curves of four **DSMI**•DNA complexes upon titration of hosts (a) **1**; (b) **2**; (c) **5**. [**DSMI**] = 0.625 µM, [DNA] = 0.1 µM, [**host**] = 0-2 µM, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4, Ex/Em = 485nm/605nm. F_0 = emission at [DNA] = 0. (d) Illustration of the various equilibria present when the hosts and dyes are exposed to DNA G-quadruplexes.

The responses are quite complex, and vary depending on the host. The variable emission profiles can be categorized into three broad types: the responses for anionic host **1**, shallow host **5**, and cationic hosts **2-4**. The response for host **1** is quite simple, as an initial lowering of emission is observed ([**1**] = 0-1 µM, **Figure 3.4 a** and **Figure A.18-19**), followed by an

increase in emission with increasing host concentration beyond 1 μM. To account for the effects of the DNA on the emission profiles, two types of fluorescence plots are shown for each host: the raw fluorescence counts (shown in Appendix A), and plots normalized to the response of cavitand:dye in the absence of DNA (**Figure 3.4 a-c** and Appendix A). The normalized fluorescence removes the effect of the host:dye emission on the signal, and focuses solely on the effects of changing DNA target.
In the case of anionic host **1**, the increase in emission with higher [**1**] after the initial decrease is solely due to **dye•1** interaction. The profiles for hosts **2-5** are different, however (**Figure 3.4 b** and **c**). When the cationic deep cavitands **2-4** are added to the DNA•dye complexes, an initial *increase* in emission is observed at very low [**2-4**] = (0-0.25 μM) followed by a fluorescence decrease. After 1 μM host is added, no effects on the *normalized* emission are seen, although the raw fluorescence increases due to formation of **host•dye** complexes with hosts **2-4**. In contrast, phosphonate host **5** causes a slight quenching of the dye upon **host•dye** complex formation (**Figure A.18-19**).[38] The fluorescence response characteristics for the two different dyes are similar, but not identical (**Figure A.18-19**), which provides a simple variable in the arrayed sensor: small changes in size of the dye can cause small, reproducible changes in emission signal.

In an attempt to rationalize these effects, we determined the binding affinities for **PSMI** and **DSMI** for the five hosts, and for a representative parallel G4 target *c-myc-2345* (**Figure A.14-16**). The **dye•host** affinities were measured by fluorescence emission titrations and calculated *via* the BindFit fitting program,[39] and are shown in **Table 3.1**. This data shows that both dyes bind strongly, yet variably to each host, with dissociation constants K_d ranging from 2.8 μM (**DSMI•5**) to 72.3 μM (**DSMI•2**). As expected, the dyes also binds to the DNA G4. In this case, ITC was used to measure the affinity, as the emission changes were too small for accurate analysis via fluorescence titrations. The accurate assessment of the affinity values and the specific location of binding site were complicated by the low, multivalent affinity of both dyes for the *c-myc-2345* G4, but simple fitting using the OneSites model gave approximate binding affinities of the order of K_d = 100-250 μM for both **DSMI** and **PSMI**. The smaller **DSMI** showed a greater multivalency than the larger **PSMI**, albeit with a lower

overall affinity. In each case, the affinity of the dye for the G4 was lower than the host:dye affinity, corroborating the theory that the hosts extract the dyes from the DNA upon addition.

Dye	K_d (1) μM[a]	K_d (2) μM[a]	K_d (3) μM[a]	K_d (4) μM[a]	K_d (5) μM[a]	K_d (*c-myc 2345*) μM[b]
DSMI	5.3± 0.4	72.3± 2.4	6.1± 1.3	20.3 ± 1.5	2.8±1.4	247±68
PSMI	8.2±0.9	51.7±2.8	8.6±1.3	5.8±1.22	3.3±1.8	112±27

Table 3.1. Binding Affinities Between Dyes and Hosts or DNA. ITC were performed in 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. [a] Measured by fluorescence emission titrations, with the [dye] = 0.625 μM for hosts **1-4**, and 10 μM for host 5. [b] Measured by ITC titrations at 25 °C, with increasing amounts of **DSMI** added to 10μM *c-myc 2345* or **PSMI** to 40μM *c-myc 2345*.

These results illustrate the possible sensing mechanisms in this system, as shown in abbreviated form in **Figure 3.4 d**. Cavitand **1** competitively binds the fluorophore, removing it from the DNA and causing an initial decrease in fluorescence. As an excess of host is added, the enhancement from forming the **host•dye** complex causes an overall emission increase. Cationic hosts **2-5** are slightly different: the hosts competitively bind the dyes (causing either an increase in emission with **2-4**, or a decrease with **5**), and at high concentrations this prevents any **dye•DNA** binding. At low concentrations, the arced fluorescence plots indicate that a second mechanism is present, presumably due to the cationic hosts also interacting with the DNA. Three states can be in equilibrium: **DNA•dye**, **host•dye** and **DNA•dye•host**, whereby the host causes an additional enhancement in fluorescence of the dye. This is consistent for all cationic hosts **2-5**, and only occurs at low [**host**]. It is not completely clear why the "arcing" in the fluorescence titrations occurs for the cationic hosts, however, as detailed structural information is not apparent from the fluorescence response. Some plausible options are that either the **dye•host** complex associates with the DNA ("state 2", **Figure 3.4 d**), causing an increase in emission, or the added host interacts with the **dye•DNA** complex ("state 1"), or both. The competitive binding of the dyes to both the DNA and the hosts is also observed by inverting the titration order (i.e. adding increasing [DNA] to **host•dye** complexes, see **Figure**

A.20-21). In this case, the dye fluorescence increases due to a greater [**DNA•dye**] complex. Interestingly, for cationic dyes **2-5**, there is a lag at low [DNA], where no signal increase occurs. The signal changes vary depending on the nature of the host, dye and DNA as before.

While there are multiple mechanistic possibilities for the emission differences with variable hosts and guests, the important conclusion for differential sensing purposes is that the responses are highly sensitive to DNA structure, and are perfectly suited for array-based sensing. As such, we constructed four different **host•dye** arrays of varying sizes. The specific concentrations of **host•dye** used here were determined by their responses from the host titration experiments above, and chosen to give the optimal response differences (see Experimental section and **Figure 3**.**5**). The **Type 1** array (6 component) uses **DSMI** as dye with hosts **1-5**, along with no host, and the **Type 2** array (6 component) uses **PSMI** as dye with hosts **1-5**, along with no host. Two larger arrays were also used - the 10-component array uses both **DSMI** and **PSMI** as dye with hosts **1-5**, and the 12-component array consists of the both **Type 1** and **Type 2** arrays. The relevant host concentrations are described in **Figure 3.5** and Methods.

Exposing the eight different DNA targets to the full 12-component array gives the fluorescence response plot in **Figure 3.5 a**, which illustrates the highly variable responses of which the array is capable. These responses were subjected to Principal Component Analysis, and the scores plot is shown in **Figure 3.5 b**. We can see from this plot that the array can fully discriminate all 8 strands tested, with all the repeated measurements included within the 95% confidence ellipses, and no overlap detected between these ellipses. Importantly, as well as distinguishing large differences between structured/unstructured DNA strands, the sensor array could discriminate between two different parallel G4 strands of identical length, *c-kit1* and *c-myc 2345*, indicating the array is sensitive to both sequence and structure. Another notable phenomenon is that the G20 strand is well separated from the unstructured A/T/C20 strands, as well as the G4 strands, which is presumably due to the formation of higher-order structures *in situ*.[40] The limit of detection was excellent, and the array can work within a wide DNA concentration range between 3-500 nM (see **Tables A.2-3**).

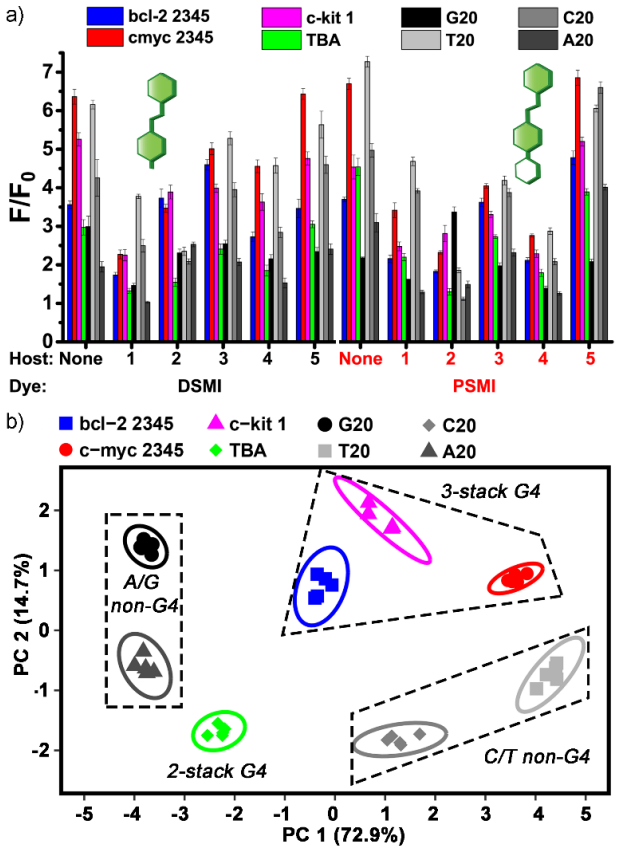


Figure 3.5. Selective sensing of DNA structures.
(a) Fluorescence **F** responses upon addition of the eight DNA strands to the **host•dye** sensor components, with F_0 = emission at [DNA] = 0; (b) PCA scores plot generated from the data in part (a), using the 12-component array (the combination of Type 1 and Type 2 arrays). [**Dye**] = 0.625 µM in both arrays, with Type 1 Array using **DSMI** and [**1**, **4**, or **5**] = 0.25 µM or [**2** or **3**] = 0.50 µM, or no cavitand, and Type 2 Array using **PSMI** and [**1**, **3**, **or 5**] = 0.25 µM, [**2**] = 1.0 µM, or [**4**] = 0.50 µM, or no cavitand. [**DNA**] = 0.1 µM, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4, Ex/Em = 485nm/605nm. Ellipses indicate 95% confidence.

In addition to the discrimination of different DNA structures shown by the array, classification of those structures into similar groupings is possible. The unstructured 20 nucleotides (nt) strands were poorly classified, as might be expected. Importantly, the classification of the G4 strands was more successful. The hybrid (*bcl-2 2345*) and parallel (*c-myc 2345, c-kit1*) G4 structures were well separated by the first principal

component (PC 1), yet clustered together in the scores plot, which matches well with the similar sizes of these structures: these three G4s are all 22 or 23 nt long and contain 3 stacks of G-quartets. The most notable difference was the antiparallel strand *TBA*, which is far shorter than the other G4s, with only 15 nt, and has only 2 stacks of G-quartets, indicating that the size of the G4 structures is important for discrimination, as well as the secondary structure and topology.
The combination of dyes and hosts in the array is essential - if the two dyes are used in the absence of host, minimal discrimination between the nucleotides is seen (**Figure A.25**). To determine which hosts had the greatest effect on G4 differentiation, we systematically removed array components, guided by the loadings of these variables on PC 1 and PC 2, and examined the grouping effect (**Figure A.22-27**). For the simple sensing shown in **Figure 3.5**, good discrimination can be achieved with 6 component arrays, using a single dye (either **PSMI** or **DSMI**, although **PSMI** is more effective) with the 5 hosts. Reducing the number of hosts decreased the discrimination power, although the combination of **DSMI/PSMI** and cavitand **2** was surprisingly effective (**Figure A.26 b**).

The data in **Figure 3.5** illustrates the abilities of the host:guest array: the different molecular recognition events, combined with varying fluorescence response mechanisms, allow both differentiation of nucleotide structures and preliminary classification of similar structures into families. The next, far more stringent test was to see if the sensor was capable of discriminating between a series of different yet structurally similar G4 structures, and whether it could also provide a classification of them. We applied the array to a suite of 23 different G-quadruplexes with strand length ranging from 15 to 26 nt, all grouped into the three family types. Six parallel (*c-kit1, c-kit2, pu22, c-myc 2345, EAD4 and PS5.M*), eight antiparallel (*TBA, Bom19, 2KF8, 442A3, TTT-L13, 2KKA, TA2, and 6FTU*) and nine hybrid structures (*AG22, bcl-2 2345, TP3, wtTel23, wtTel24, wtTel26, Tel 26, Telo24 and H24*) were tested (for full sequences see **Table A.1**). The topology of these strands were confirmed by CD analysis, and each strand was confirmed to be a unimolecular G4 by gel electrophoresis.

As most of the 23 G4 targets are highly similar in size and global structure, we employed the full 12-component array to deliver the most powerful discrimination possible (see **Figure A.28** for full fluorescence response plots). The fluorescence data was collected in the same manner as with the smaller data set above, and the responses analyzed by a series of multivariate analysis tools (**Figure 3.6**). For clarity, the data shown in **Figure 3.6 a** shows the average fluorescence change from 5 repeated measurements on each G4 as a single plot point. The full suite of data points is shown in the Supplementary Information, including error ellipses, but the simplified plot in **Figure 3.6 a** illustrates the results nicely. The clear outcome of the PCA plots is that the sensor can easily discriminate and classify the parallel G4 structures (six different G4s, illustrated in red) from all the other entries. The discrimination is remarkable: three of the G4s (*c-kit1, pu22, c-myc 2345*) have exactly the same strand length (22 nts) and the same topology, yet are fully separated on the scores plot. The classification of the hybrid and antiparallel G4 structures is, at first glance, less impressive, but closer inspection leads to some interesting observations. The nine hybrid G4 structures are quite closely related in structure: they all show 3 stacks of G-quartets, and are of relatively similar lengths (22-26 nts). In fact, the grouping of the hybrid structures is quite clear, and shown in blue in **Figure 3.6 a**, grouped around the centerpoint of the PCA scores plot. The most interesting point is the behavior of the antiparallel strands (shown in green): while the hybrid and parallel G4s are all properly grouped, there is significant variation in the positions of the various antiparallel strands, with four G4s (*TTT-L3, 2KKA, 442 A3* and *2KF8*) co-locating with the hybrid structures, and the other four structures in two separate groups (*6FTU/TA2* and *TBA/Bom19*).

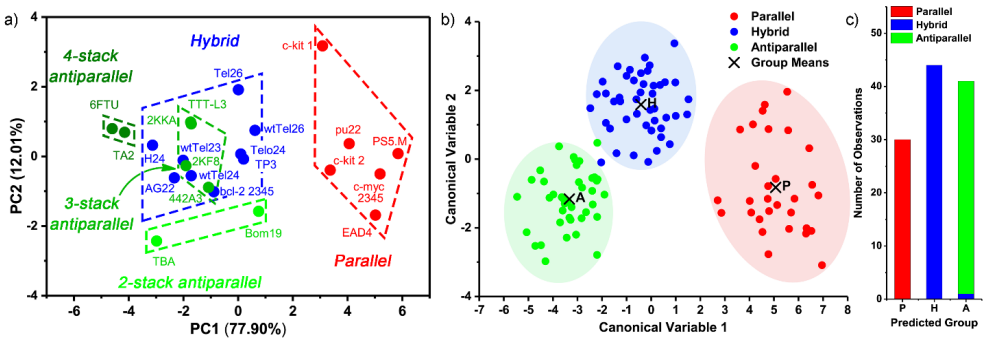


Figure 3.6. Classification and discrimination of a suite of 23 G-quadruplex structures. (a) PCA scores plot from the fluorescence responses of 23 G4 strands to the sensing array. Each point is the average of 5 repeated measurements; (b) Canonical Discriminant Analysis (CDA) scores plot of the individual responses of 115 samples classified by 12 characteristics, and grouped based on topology; (c) Tabulation of the CDA results to illustrate the classification error rate of 1/115 samples. Sensor conditions are identical to those described in **Figure 3.5**. 12-component array used.

The challenges in differentiating the hybrid and antiparallel structures can be easily explained by their structural similarities. While all four G strands in a parallel G4 are in the same 5'→3' orientation, with all glycosidic angles displaying an *anti* conformation, the hybrid and antiparallel G4s both contain one or two of the four G strands in an orientation opposite to the others and have both the *anti* and *syn* guanines. They are even, in certain literature cases, considered as a single topology group.[16,18] The discrimination between certain hybrid and antiparallel strands in the scores plot is, in fact, dependent on their *length*. The antiparallel and hybrid G4s that co-locate close to each other are all of similar lengths (22-26 nts) and all display *three* G-quartet stacks in the folded structure. The outlier antiparallel structures vary in the number of G-quartet stacks, either 2 (*TBA* and *Bom19*) or 4 (*6FTU* and *TA2*), and these larger variations in structure are fully separated from the 3-stack G-quartets (either antiparallel or hybrid). In addition, some impressive discrimination effects can be seen for these G4s. Although *2KKA* has high sequence similarity with *2KF8, AG22*, *wtTel23* and *wtTel24* (varying only by the absence of one T or one A on the 3' or 5' end), it could be well separated from the other strands in the PC 2 axis. Similar differentiation is also observed between *Tel26* and *wtTel26*, as well as between *Telo24* and *H24*. When an additional component was added to the PCA (i.e. PC 3,

accounting for 4.46% of the overall variation in the dataset) even greater separation between the hybrid and antiparallel structures with the same number of G-quartet stacks could be obtained: *442 A3* was well separated from *bcl-2 2345* (**Figure A.32**). These results illustrate that the array can differentiate these G4 structures based on multiple secondary structural features, including topology, number of G-quartet stacks, and the underlying base sequence. Dissecting the contribution of individual array elements to the differentiation shows that changing the host is most important for differentiating the structures by their topology, while the different dyes contribute more towards revealing differences in stack numbers and sequences (see **Figures A.29-31** for PCA plots with varying array elements). This illustrates the power of the array concept: by introducing (or removing) individual array components, the sensor can be tailored for specific targets, and is not restricted to one type of DNA structure, but can conceivably be applied to more varied applications in DNA/RNA structural analysis.

To further validate the G4 classifications, the fluorescence responses were also subjected to Canonical Discrimination Analysis (CDA) (**Figure 3.6 b**).[12] Whereas PCA is an unsupervised classification tool that finds the greatest variance between samples using the fluorescence signals and clusters the samples with a smaller variance, CDA is supervised in classification: the class of each sample (i.e. the topology of each G4) is included to maximize class discrimination. While the unsupervised PCA confirms differentiation of different G4 sequences, the supervised CDA analyzes whether the array can classify the G4s by topology. In this case, the raw data measurements with the full 12-element array (including all 5 repeats) for all the 23 G4 structures were used as the input (a total of 115 samples classified by 12 characteristics), together with their topology information. The results are extremely impressive: the Canonical Scores plot (**Figure 3.6 b**) clearly shows that the targets are robustly classified into the three expected topology groups, parallel, hybrid and antiparallel (shown in red, blue and green respectively in **Figure 3.6 b**), with 114 out of 115 samples being assigned to the correct type of topology. Only 1 sample in the hybrid group was wrongly assigned to antiparallel (**Figure 3.6 c**), which corresponds to an error rate in classification of 0.7%. In addition, k-fold (k=8) cross-validation analysis, in which 1/8 of the observations were removed from

the data set and treated as unknowns to perform the prediction test, led to an average precision rate of 98.21%. This analysis corroborates the PCA results well, and validates the power of the sensor array in classifying highly similar G4 sequences by topology type.

To further illustrate the power of the sensing array, we performed the discrimination tests in more complex media (**Figures A.34-41**). The sensor was completely tolerant to the presence of small saccharides, with no loss of performance in the presence of 20 μM lactose. Most impressively, the sensor remained effective in the presence of 5% fetal bovine serum (FBS). Full discrimination between the initial 8 targets (A/G/T/C20, *c-kit1*, *c-myc 2345*, *bcl-2 2345*) was possible with the **Type 1** array (**Figures A.35**). Considering the complex nature of FBS, containing a variety of interfering molecules, this selectivity is impressive. Greater serum concentrations reduced the effectiveness of the sensing, as would be expected. Cationic proteins such as lysozyme were the greatest interferents (**Figures A.36-39**), and the array could only tolerate 0.1 μM lysozyme before losing selectivity.

Finally, we tested the possibility of detecting structural changes upon addition of an external effector, as well as whether the array was capable of detecting changes in [G4] in mixtures of DNA (**Figures 3.7**). As G-quadruplexes use alkali metal cations as structural components, changing the nature of the cation (from K^+ to Na^+) can sometimes cause a change in G-quadruplex topology. Notably, *AG22* displays a hybrid quadruplex structure in potassium phosphate buffer, but switches to an antiparallel structure in sodium phosphate.[16,41] As such, we repeated the fluorescence measurements for a series of G-quadruplexes in either Na or K phosphate buffer, using the compressed **Type 1** array with **DSMI** as dye. In this case, five G4 structures were chosen: *AG22*, and *TBA*, *c-myc 2345*, *c-kit1*, *bcl-2 2345* as controls. CD measurements confirmed that the control G4s maintained their folding topology upon cation switch, whereas *AG22* indeed changed topology as reported (**Figures A.4**). The sensor results are shown in **Figures 3.7 a**: in the PCA scores plot in K^+ buffer, the data points corresponding to *AG22* are in close proximity to those of the hybrid G4 control of *bcl-2 2345*, with the 95% confidence ellipses overlapping significantly. The parallel and antiparallel controls are fully distinguished. When the same measurements are made in Na^+-containing buffer, the

changes become obvious: no longer is *AG22* co-located with the hybrid *bcl-2 2345*, but moves into close proximity with the antiparallel *TBA*. These observations correlate with the CD spectra of these strands: while the intensity of the characteristic CD peaks for topology determination in the spectra of *TBA* and *c-myc 2345* remained comparable when switching from the K^+ buffer to the Na^+ buffer, the peaks for *c-kit1* and *bcl-2 2345* dropped noticeably, although the topology type was not changed. While the sensor is sensitive to changes in buffer, it is still fully capable of sensing the hybrid→antiparallel structural change of *AG22*, which was one of the most challenging conformational variations to detect among our 23 G4 candidates in identical buffer conditions.

The sensor can detect changes in [G4] concentration in a mixture of nucleotides. The *c-myc 2345* G4 was combined with the A20 ssDNA strand in varying proportions, and added to a solution of the compressed **Type 1** array with **DSMI** as the dye. The fluorescence responses (**Figure A.42**) were subjected to PCA, the scores plot shown in **Figure 3.7 b**. As the proportion of A20/*c-myc 2345* varies, the signals on the scores plot move towards the relevant signals for 100% A20/*c-myc 2345* (shown in blue and red respectively), with the changes dominated by the position on PC 1. Even in the presence of competing nucleotides, the selectivity of the sensor for the folded structure is retained.

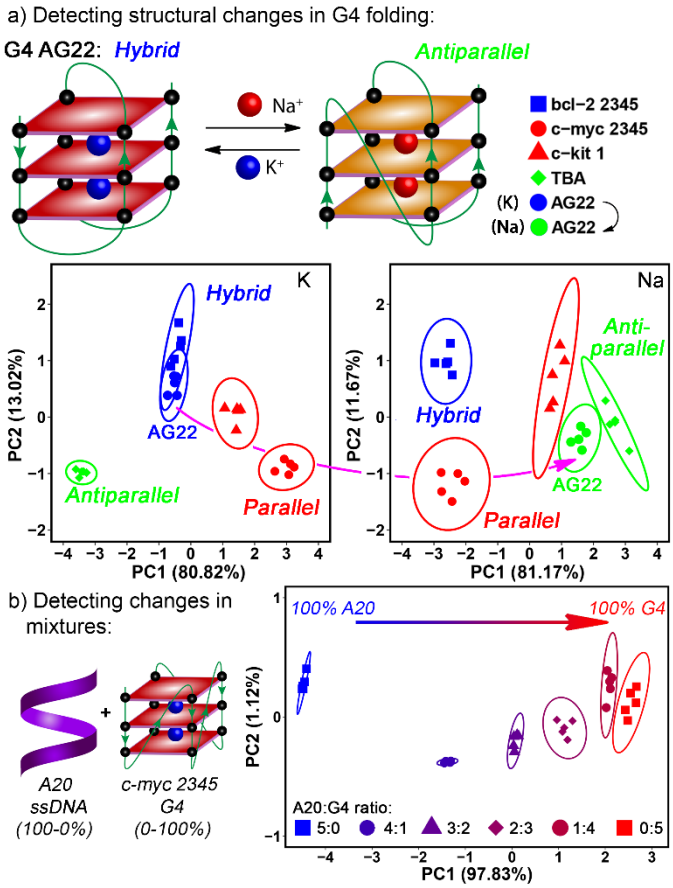


Figure 3.7. Sensing structural topology switching of *AG22* from hybrid to antiparallel with different structural cations. (a) PCA scores plot from the fluorescence responses of 5 G4 strands, *AG22* and 4 controls ([DNA] = 0.1 µM) in either 10 mM KH_2PO_4/K_2HPO_4 buffer or 10 mM NaH_2PO_4/Na_2HPO_4 buffer; (b) PCA scores plot from the fluorescence responses of solution containing various molar ratio of *c-myc 23455*:polyA20 ([total DNA] = 0.5 µM) in 10 mM KH_2PO_4/K_2HPO_4. **Type 1** sensing array was used for both. 1 mM EDTA, pH 7.4, was present in all buffer. Ex/Em = 485nm/605nm.

3.3 CONCLUSIONS

Here we have shown that an arrayed suite of synthetic hosts and dyes is capable of sensing oligonucleotide secondary structures. By combining cationic, DNA-binding dyes with a set of synthetic host molecules that also show affinity for the dyes, multiple recognition mechanisms can be exploited to create a unique sensing fingerprint consisting of different fluorescence enhancements in the presence of select DNA strands. Multivariate analysis of these sensing fingerprints enables discrimination between DNA structures. This discrimination is successful for strands that are as different in structure as single stranded A20 and G20, but is also possible for G-quadruplex strands that vary only slightly in structure and size.

Most importantly, the array is able to both *differentiate and classify* the G4 structures at the same time. Simply differentiating between different G4 types has limited utility by itself: the ability to differentiate between different G4 structures while also enabling their classification is essential for detecting the presence of specific secondary structural motifs as disease markers, for example. Whereas CD spectroscopy can quickly identify folding topology, it is incapable of distinguishing between different G4 structures of the same folding type. Mass spectrometry and sequencing techniques can easily detect differences in size, but are poor at determining the secondary structure in solution. Our array is capable of both types of analysis at the same time, in a simple, non- invasive manner, which requires little sample and does not require discovery and synthesis of specific ligand targets for each oligonucleotide type. The selectivity is excellent: different G4s that display the same folding topology can be easily differentiated by the number of G-quartets and sequence differences at the 3’ or 5’ ends (e.g. *Tel26* vs. *wtTel 26; Telo24* vs. *H24; 2KK8* vs*. 2KF8*). It can even detect changes in G4 folding pattern in different types of complex media, in the presence of interfering small molecules and in mixtures of nucleotides. The resultant data could be employed in computational modeling to reveal close correlation between sequence and structure, paving the way to accurate prediction of nucleic acid folding based on its sequence and their fluorescence responses in our array.

3.4 EXPERIMENTAL SECTION

General. Cavitands **1**,[31] **2**,[29] **3**,[34] and **4**,[34] as well as phosphonate cavitand **5**[35] and **PSMI** fluorophore[29] were synthesized according to literature procedures. ^{1}H and ^{13}C spectra were recorded on a Bruker Avance NEO 400 MHz. Deuterated NMR solvents were obtained from Cambridge Isotope Laboratories, Inc., Andover, MA, and used without further purification. All other materials, including *trans*-4-[4-(dimethylamino)-styryl]-1-methyl-pyridinium iodide (**DSMI**) were obtained from Aldrich Chemical Company (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.). Oligonucleotides were purchased from Integrated DNA Technologies (IDT) with standard desalting and no further purification, the sequence and topology information of which were given in **Table A.1**. All DNA solutions were prepared in 10 mM K_2HPO_4/KH_2PO_4 or Na_2HPO_4/NaH_2PO_4 buffer at pH 7.4, both containing 1 mM EDTA (referred as the K^+ or Na^+ buffer in the text); and their concentrations were determined by NanoDrop 2000 (Thermo Fisher Scientific) using the corresponding molar extinction coefficients provided by IDT after background subtraction. Before fluorescence or CD measurement, the DNAs were diluted to 1 μM with the K^+ or Na^+ buffer; and re-annealed to form the most stable folding topology, in which the DNA solution was heated at 95 °C for 5 min and then kept at 4° C for overnight. Fluorescence measurements were performed with a Perkin Elmer Wallac 1420 Victor 2 Microplate Reader (PerkinElmer) with the Ex/Em wavelengths at 485/605 nm. Principal Component Analysis (PCA) and confidence ellipses were performed with RStudio (Version 1.2.5019), an integrated development environment (IDE) for R (version 3.6.1). Canonical Discriminant Analysis was conducted with OriginPro 2018. Classification performance was evaluated by k-fold (k = 8) cross validation using Python 3, with Linear Discriminant Analysis (LDA) as the supervised classification model.

Fluorescence measurements. In general, the fluorescence assay was carried out by mixing 10 μL of the fluorescent guest (6.25 μM **DSMI** or **PSMI** in water), 10 μL of the cavitand (in water) or simply 10 μL water for the no cavitand sample, 70 μL of the incubation buffer, and 10 μL of 1 μM

DNA in the 96-well plate, resulting in a final total volume around 100 µL in 10 mM K_2HPO_4/KH_2PO_4 (or K_2HPO_4/KH_2PO_4) and 1 mM EDTA at pH 7.4. The mixture was incubated with mild shaking for 15 min at room temperature, before the fluorescence signal (F) was recorded in a Perkin Elmer Wallac 1420 Victor 2 Microplate Reader (PerkinElmer) with the Ex/Em wavelengths at 485/605 nm.

Array constituents. Type 1 Array: hosts **1**-**5** or no host, **DSMI** dye. [**DSMI**] = 0.625 µM, [**1**, **4**, or **5**] = 0.25 µM or [**2** or **3**] = 0.50 µM. **Type 2 Array:** hosts **1**-**5** or no host, **PSMI** dye. [**PSMI**] = 0.625 µM, [**1**, **3**, **or 5**] = 0.25 µM, [**2**] = 1.0 µM, or [**4**] = 0.50 µM. **10-component array:** Combination of Type 1 and Type 2 arrays, not including "no host". **12-component array:** Combination of Type 1 and Type 2 arrays.

Circular Dichroism (CD). CD spectra were recorded on a Jasco J-815 CD spectrophotometer over a wavelength range of 200 nm–350 nm at room temperature, with a band width of 1 nm and a data pitch of 0.5 nm. The instrument scanning speed was set at 100 nm/min, with a response time of 1 s. 10 µM of 200 µL oligonucleotide solution prepared in the K^+ buffer or Na^+ buffer was pipetted into a quartz cell with a path length of 0.1 cm. The CD spectra were presented with baseline-correction in which the background signal from the buffer was subtracted.

Gel Electrophoresis. The quality of the DNA solution was inspected by native gel electrophoresis using a precast gradient (4%-20%) PAGE gel. 10 µL of 1 µM DNA solution was loaded to the gel, after being denatured at 95 °C for 5 min, cooled on ice for 10 min and then at room temperature for 30 min. The gel was run at 120 V for 90 min at room temperature in 1×TBE buffer, and stained with SYBR Gold (1:10000 dilution) before imaged using the UV transilluminator (SPECTROLINE).

Binding Affinity Measurements: Fluorescence Titrations. For Host **1-4**, the fluorescence titration curves were obtained by adding 0-45 µM host **1-4** into the solution that contained 0.625 µM **DSMI** or **PSMI** in the K^+ buffer. For host **5**, the titration was carried in the same manner, by using 0-10 µM host **5** and 10 µM **DSMI** or **PSMI**, because the dye fluorescence was only slightly quenched by the host. Fluorescence was recorded after 15 min of

mixing in the plate reader using the instrument setting described above. The binding affinity calculations were performed at https://supramolecular.org.[39] The UV 1:1 filter within the Bindfit function was used along with the Nelder-Mead fitting method,[39] with the "subtract initial value" option selected.

Binding Affinity Measurements: ITC Titrations. All ITC experiments were performed using a MicroCal iTC200 (GE Healthcare, Freiburg, Germany) with a stirring rate of 800 rpm. The baseline was stabilized prior to the experiment, and a pre-injection delay was set to 60 s. A stock solution of **DSMI** or **PSMI** at 3 mM in K^+ buffer was added in 2 µL aliquots to the *c-myc 2345* solution of 10 and 40 µM (also in K^+ buffer), respectively. **DSMI** binds about 2 times more weakly than **PSMI**, and thus requires a higher dye:DNA ratio to reach a plateau. All experiments were conducted at 25 °C. The heat of dilution, measured by the injection of titrant into the buffer solution, 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.

3.5 REFERENCES

1. P. Belmont, J.-F. Constant, M. Demeunynck, *Chem. Soc. Rev.* **2001**, *30*, 70.
2. S. Balasubramanian, L.H. Hurley, S. Neidle, *Nat. Rev. Drug Discov.* **2011**, *10*, 261.
3. N.C. Seeman, H.F. Sleiman, *Nat. Rev. Mater.* **2017**, *3*, 17068.
4. A.V. Pinheiro, D. Han, W.M. Shih, H. Yan, *Nat. Nanotechnol.* **2011**, *6*, 763.
5. L. Jaeger, A. Chworos, *Curr. Opin. Struct. Biol.* **2006**, *16*, 531.
6. F.R. Winnerdy, B. Bakalar, A. Maity, J.J. Vandana, Y. Mechulam, E. Schmitt, A.T. Phan, *Nucleic Acids Res.* **2019**, *47*, 8272.
7. G.F. Salgado, C. Cazenave, A. Kerkour, J.-L. Mergny, *Chem. Sci.* **2015**, *6*, 3314.
8. R. del Villar-Guerra, J.O. Trent, J.B. Chaires, *Angew. Chem. Int. Ed.* **2018**, *57*, 7171.
9. C.S. Eubanks, J.E. Forte, G.J. Kapral, A.E. Hargrove, *J. Am. Chem. Soc.* **2017**, *139*, 409.
10. L. You, D. Zha, E.V. Anslyn, *Chem. Rev.* **2015**, *115*, 7840.
11. C.S. Eubanks, B. Zhao, N.N. Patwardhan, R.D. Thompson, Q. Zhang, A.E. Hargrove, *J. Am. Chem. Soc.* **2019**, *141*, 5692.
12. S. Stewart, M.A. Ivy, E.V. Anslyn, *Chem. Soc. Rev.* **2014**, *43*, 70.
13. R. del Villar-Guerra, J.O. Trent, J.B. Chaires, *Nucleic Acids Res.* **2018**, *46*, e41.
14. J.L. Huppert, *Chem. Soc. Rev.* **2008**, *37*, 1375.
15. M.L. Bochman, K. Paeschke, V.A. Zakian, *Nat. Rev. Genet.* **2012**, *13*, 770.
16. S. Burge, G.N. Parkinson, P. Hazel, A.K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402.
17. N.G. Dolinnaya, A.M. Ogloblina, M.G. Yakubovskaya, *Biochemistry* **2016**, *81*, 1602.
18. C.K. Kwok, C.J. Merrick, *Trends Biotechnol.* **2017**, *35*, 997.
19. E. Puig Lombardi, A. Londoño-Vallejo, *Nucleic Acids Res.* **2019**, *48*, 1.
20. M. Zuffo, A. Guédin, E.-D. Leriche, F. Doria, V. Pirota, V. Gabelica, J.-L. Mergny, M. Freccero, *Nucleic Acids Res.* **2018**, *46*, e115.
21. K.M. Felsenstein, L.B. Saunders, J.K. Simmons, E. Leon, D.R. Calabrese, S. Zhang, A. Michalowski, P. Gareiss, B.A. Mock, J.S. Schneekloth, *ACS Chem. Biol.* **2016**, *11*, 139.
22. R. Pinalli, A. Pedrini, E. Dalcanale, *Chem. Soc. Rev.* **2018**, *47*, 7006.
23. R.N. Dsouza, A. Hennig, W.M. Nau, *Chem. Eur. J.* **2012**, *18*, 3444.
24. A. Hennig, H. Bakirci, W.M. Nau, *Nat. Methods* **2007**, *4*, 629
25. B.C. Peacor, C.M. Ramsay, M.L. Waters, *Chem. Sci.* **2017**, *8*, 1422.
26. S.A. Minaker, K.D. Daze, M.C.F. Ma, F. Hof, *J. Am. Chem. Soc.* **2012**, *134*, 11674.
27. M. Florea, W.M. Nau, W. M. *Org. Biomol. Chem.* **2010**, *8*, 1033.

28. Y. Liu, M. Mettry, A.D. Gill, L. Perez, W. Zhong, R.J. Hooley, *Anal. Chem*. **2017**, *89*, 11113.
29. A.D. Gill, B.L. Hickey, S. Wang, M. Xue, W. Zhong, R.J. Hooley, *Chem. Comm.* **2019**, *55*, 13259.
30. P. Murat, Y. Singh, E. Defrancq, *Chem. Soc. Rev.* **2011**, *40*, 5293.
31. S.M. Biros, E.C. Ullrich, F. Hof, L. Trembleau, J. Rebek, *J. Am. Chem. Soc.* **2004**, *126*, 2870.
32. Y. Liu, L. Perez, A.D. Gill, M. Mettry, L. Li, Y. Wang, R.J. Hooley, W. Zhong, *J. Am. Chem. Soc.* **2017**, *139*, 10964.
33. Y. Liu, J. Lee, L. Perez, A.D. Gill, R.J. Hooley, W. Zhong, *J. Am. Chem. Soc.* **2018**, *140*, 13869.
34. S. Mosca, Y. Yu, J. Rebek, *Nat. Protoc.* **2016**, *11*, 1371.
35. R. Pinalli, G. Brancatelli, A. Pedrini, D. Menozzi, D. Hernández, P. Ballester, S. Geremia, E. Dalcanale, *J. Am. Chem. Soc.* **2016**, *138*, 8569.
36. D. Menozzi, E. Biavardi, C. Massera, F.-P., Schmidtchen, A. Cornia, E. Dalcanale, *Supramol. Chem.* **2010**, *22*, 768.
37. Q. Yang, J. Xiang, S. Yang, Q. Li, Q. Zhou, A. Guan, X. Zhang, H. Zhang, Y. Tang, G. Xu, *Nucleic Acids Res.* **2010**, *38*, 1022.
38. A.E. Früh, F. Artoni, R. Brighenti, E. Dalcanale, *Chem. Mater.* **2017**, *29*, 7450.
39. P. Thordarson, *Chem. Soc. Rev.* **2011**, *40*, 1305.
40. A. Sengar, B. Heddi, A.T. Phan, *Biochemistry* **2014**, *53*, 7718.
41. A. Ambrus, D. Chen, J. Dai, T. Bialis, R.A. Jones, D. Yang, *Nucleic Acids Res.* **2006**, *34*, 2723.

APPENDIX A

1. DNA sequences and characterization

1.1 DNA sequences

Name	Sequence	G4 type	Bases	Ref.
c-kit 2	CGGGCGGGCGCGAGGGAGGG	parallel	20	1
EAD4	CTGGGTTGGGTTGGGTTGGGA	parallel	21	2
c-myc 2345	TGAGGGTGGGGAGGGTGGGGAA	parallel	22	1,3,4
c-kit 1	AGGGAGGGCGCTGGGAGGAGGG	parallel	22	3,5
pu22	TGAGGGTGGGTAGGGTGGGTAA	parallel	22	1,6
PS5.M	GTGGGTCATTGTGGGTGGGTGTGG	parallel	24	7
AG 22	AGGGTTAGGGTTAGGGTTAGGG	Na^+: antiparallel K^+: hybrid	22	1,3
blc-2 2345	GGGCGCGGGAGGAATTGGGCGGG	hybrid	23	3,8
wtTel23	TAGGGTTAGGGTTAGGGTTAGGG	hybrid	23	9
TP3	TGGGGGCCGAGGCGGGGCTTGGG	hybrid	23	10
wtTel24	TAGGGTTAGGGTTAGGGTTAGGGT	hybrid	24	9
Telo24	TTAGGGTTAGGGTTAGGGTTAGGG	hybrid	24	11
H24	TTGGGTTAGGGTTAGGGTTAGGGA	hybrid	24	3
wtTel26	TTAGGGTTAGGGTTAGGGTTAGGGTT	hybrid	26	9
Tel26	AAAGGGTTAGGGTTAGGGTTAGGGAA	hybrid	26	12
TBA	GGTTGGTGTGGTTGG	antiparellel	15	1,3
Bon19	TAGGTTAGGTTAGGTTAGG	antiparellel	19	14
2KF8	GGGTTAGGGTTAGGGTTAGGGT	antiparellel	22	15
442-A3	GGGTTTTGGGTTTTGGGTTGGG	antiparellel	22	16
TTT-L13	AGGGTTTGGGTTAGGGTTTGGGT	antiparellel	23	13
2KKA	AGGGTTAGGGTTAGGGTTAGGGT	antiparellel	23	15
TA2	GGGGTTGGGGTGTGGGGTTGGGG	antiparellel	23	2
6FTU	GGGGGAGGGGTACAGGGGTACAGGGG	antiparellel	26	17

Table A.1. G4 sequences and topology types as reported in K^+ buffer (10mM K_2HPO_4/KH_2PO_4, 1mM EDTA, pH 7.4) except for AG22, the topology of which in both K+ and Na+ buffer (10 mM Na_2HPO_4/NaH_2PO_4) was reported.

1.2 Circular Dichroism (CD) corroboration of G4 topology

Parallel G4 sequences

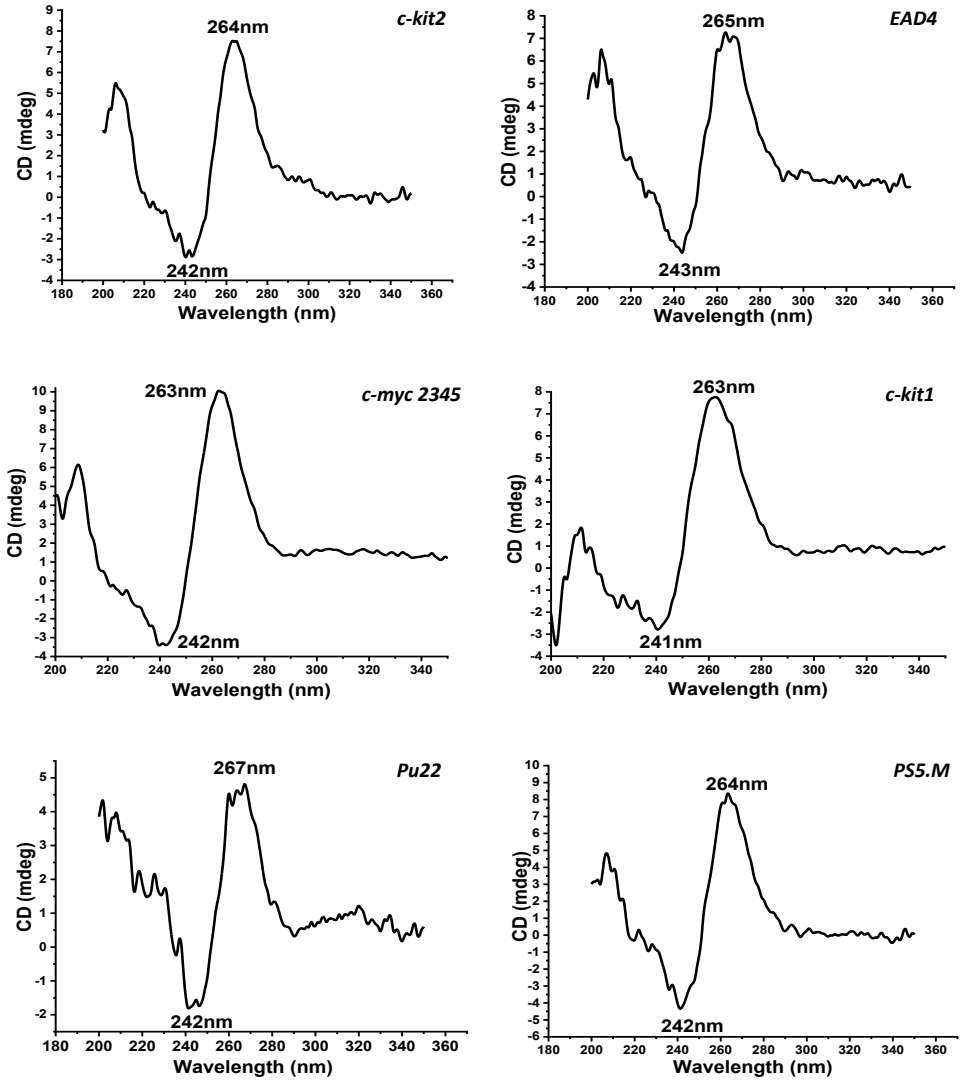
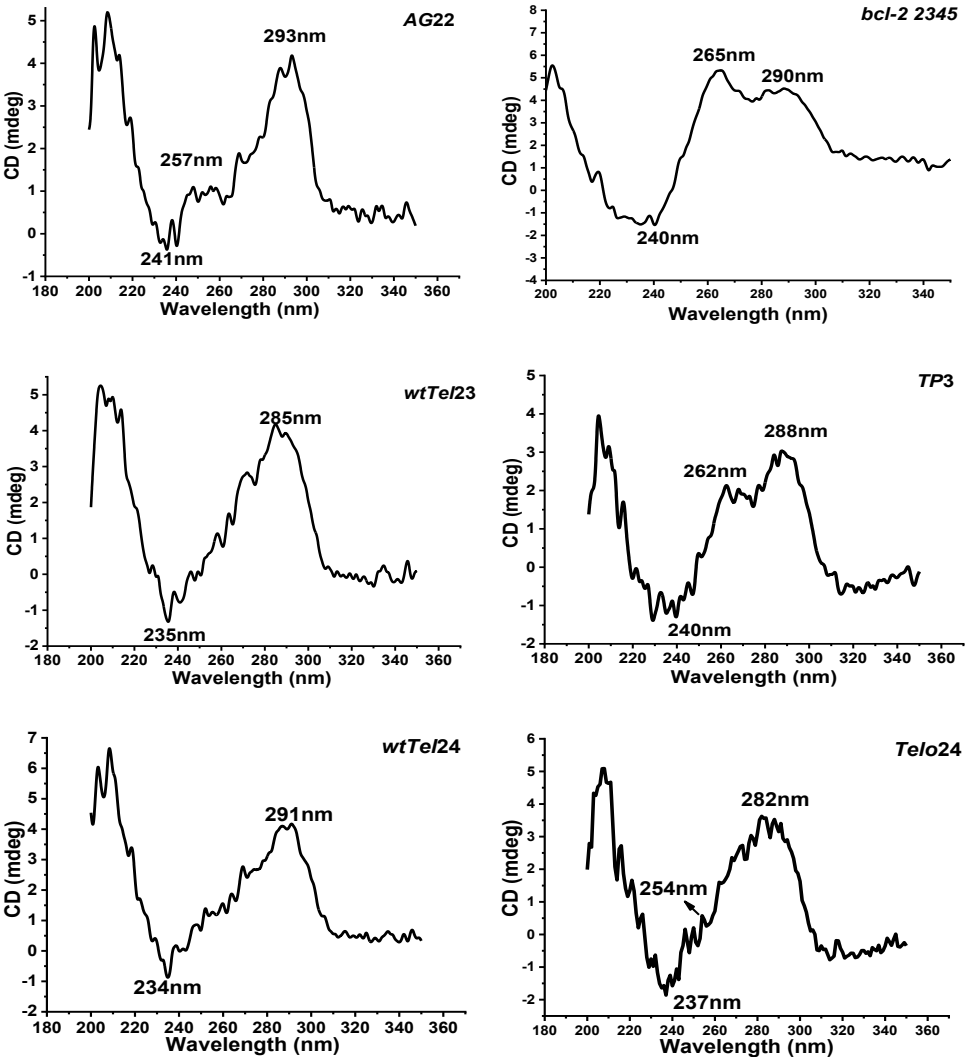


Figure A.1. CD spectra with baseline correction of 10μM parallel G4 DNA in K^+ buffer (10mM K_2HPO_4/KH_2PO_4, 1mM EDTA, pH 7.4). DNA was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.

Hybrid G4 sequences



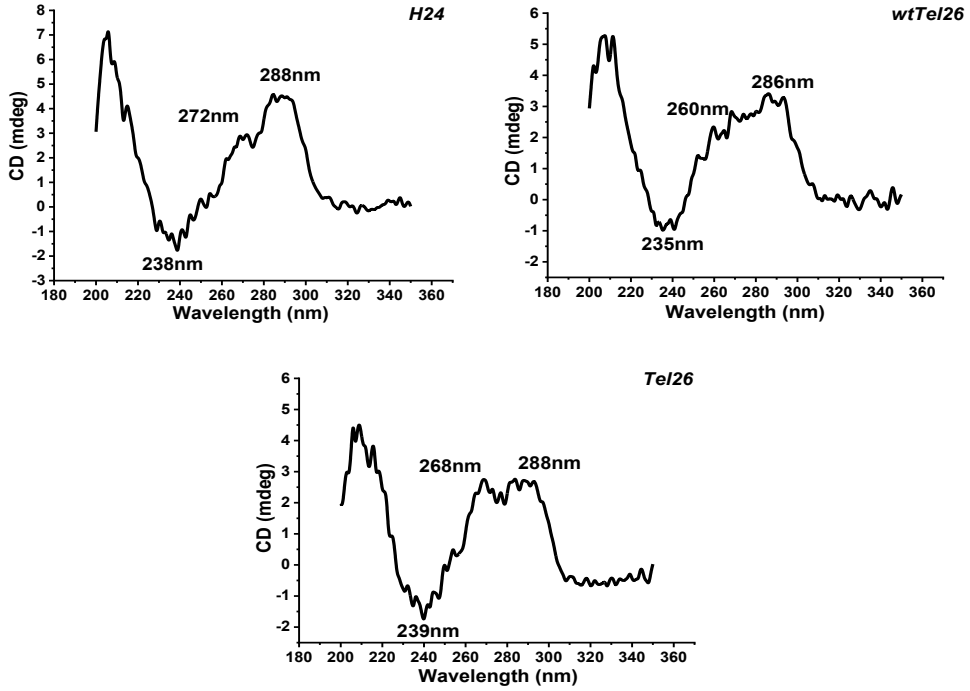
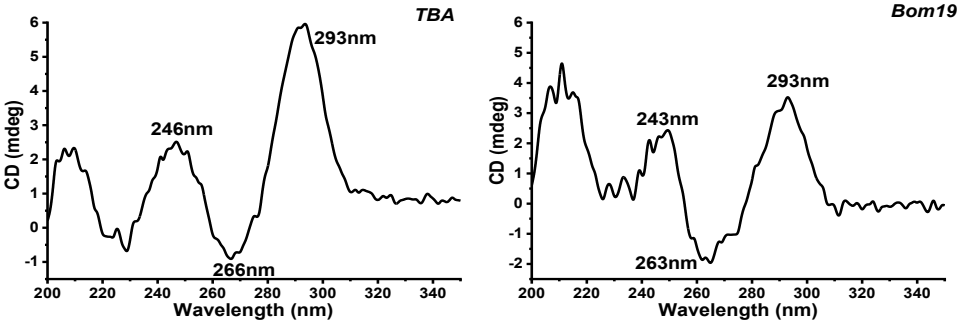


Figure A.2. CD spectra with baseline-correction of 10 µM hybrid G4 DNA in K^+ buffer (10mM K_2HPO_4/KH_2PO_4, 1mM EDTA, pH 7.4). DNA was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.

Antiparallel G4 sequences



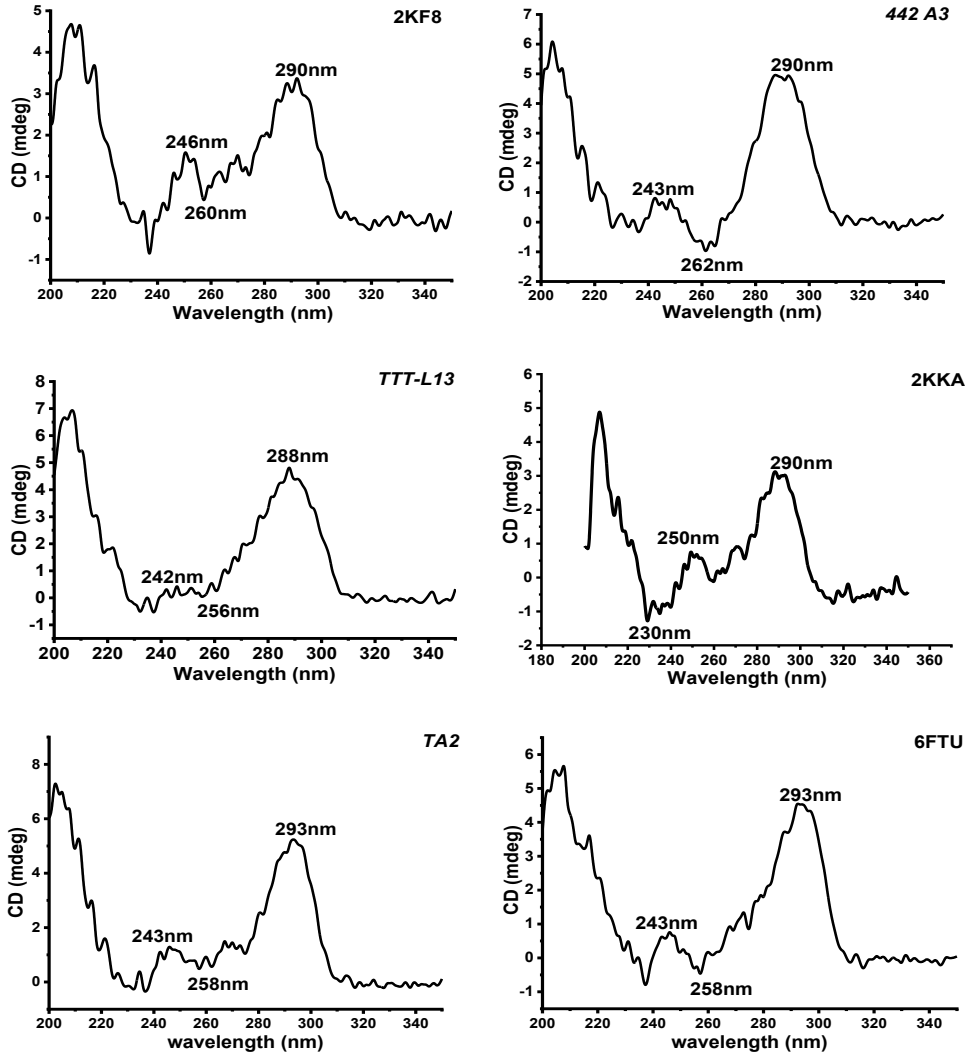


Figure A.3. CD spectra with baseline-correction of 10 µM antiparallel G4 DNA in K^+ buffer (10mM K_2HPO_4/KH_2PO_4, 1mM EDTA, pH 7.4). DNA was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.

1.3 CD analysis of AG22 topology in Na^+/K^+ buffer

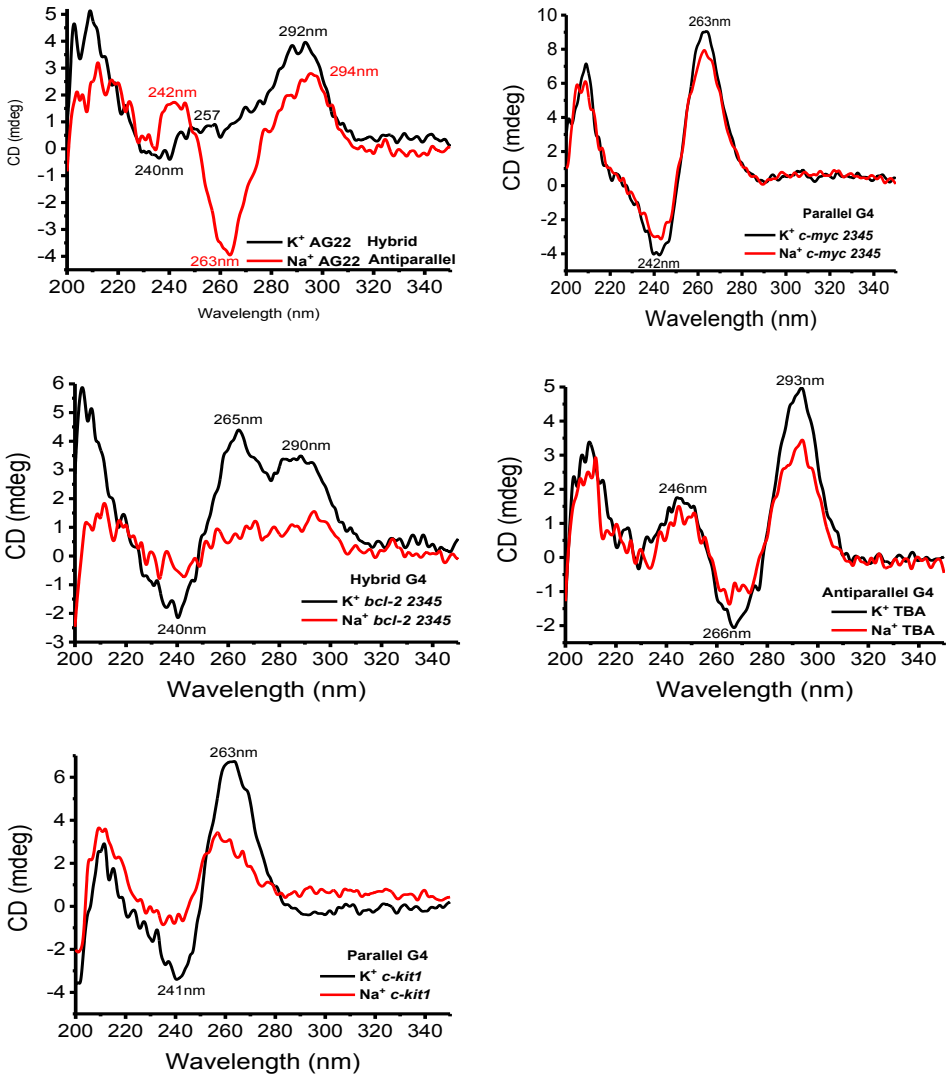


Figure A.4. CD spectra with baseline-correction of 10 µM *AG22, c-myc 2345*, *bcl-2 2345*, *TBA* and *c-kit1* in K^+ (10 mM K_2HPO_4/KH_2PO_4) (black line) or Na^+ buffer (10 mM Na_2HPO_4/NaH_2PO_4), both with 1 mM EDTA, pH 7.4.
DNA was denatured at 95 °C 5 min, then re-annealed at 4 °C overnight before experiment.

1.4 Gel electrophoresis corroboration of unimolecular G4 structures

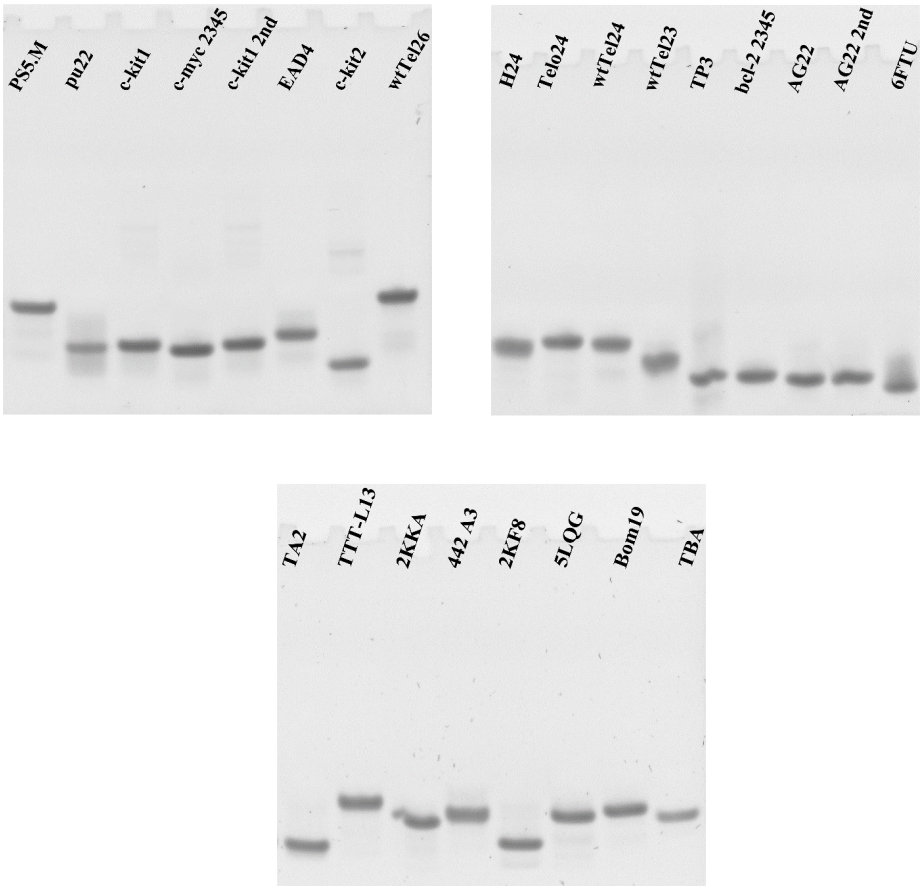


Figure A.5. The gradient native polyacrylamide gel electrophoresis (PAGE) gel (4%-20%) results of DNA G-quadruplex sequences. (c-kit1 2nd is the repeat of c-kit1, and AG22 2nd is the repeat of AG22)

2. NMR spectra of components used

All hosts and guests are known and their spectra have been previously published (see experimental section for the references).
^{1}H NMR spectra of the specific sample used are here reported to illustrate their purity and ensure reproducibility.
Some spectra were taken in DMSO-d_6 to maximize the signal:nose ratio.

*2.1 ^{1}H NMR spectrum of cavitand **1***

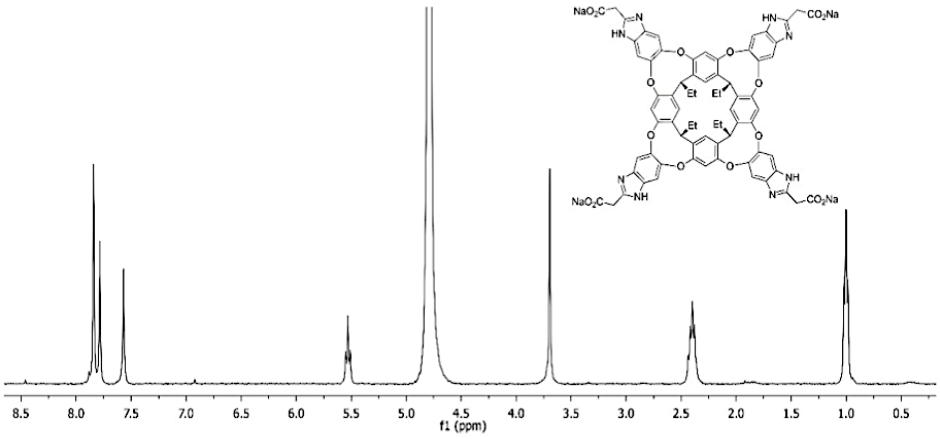


Figure A.6. ^{1}H NMR spectrum of cavitand **1** in D_2O, 25 °C, 400 MHz.

*2.2 ^{1}H NMR spectrum of cavitand **2***

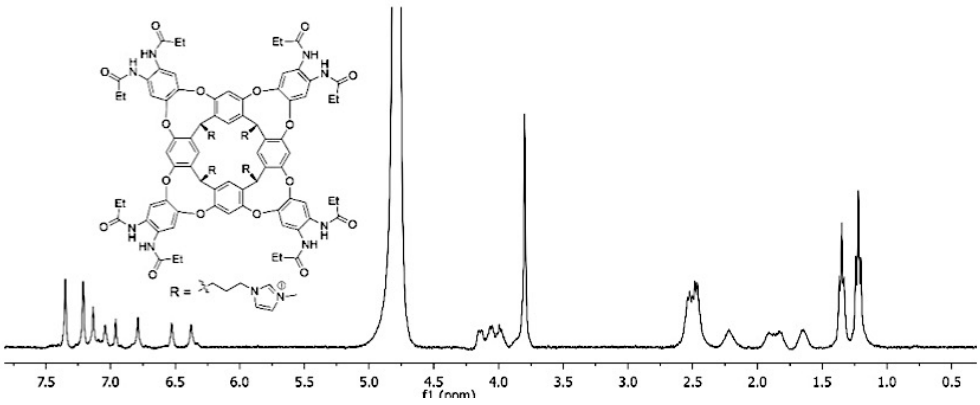


Figure A.7. ^{1}H NMR spectrum of cavitand **2** in D_2O, 25 °C, 400 MHz.

*2.3 ^{1}H NMR spectrum of cavitand **3***

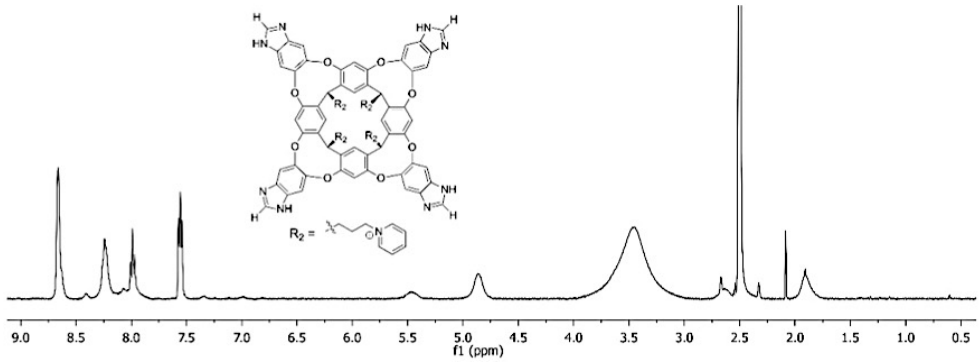


Figure A.8. ^{1}H NMR spectrum of cavitand **3** in DMSO-d_6, 25 °C, 400 MHz.

*2.4 ^{1}H NMR spectrum of cavitand **4***

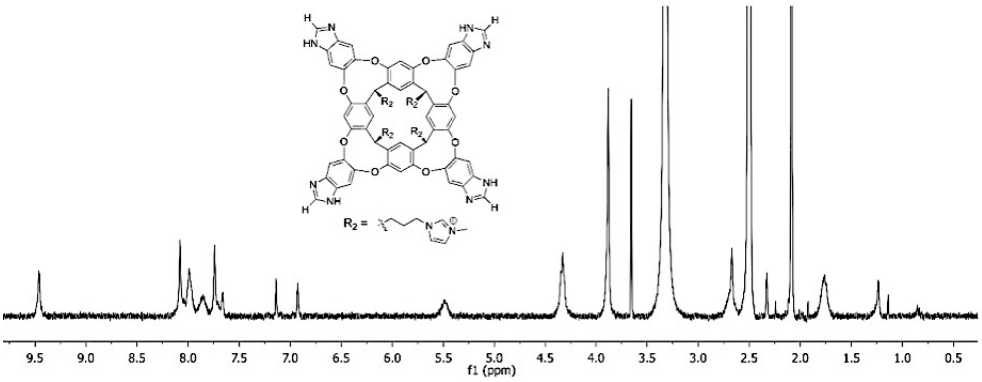


Figure A.9. ^{1}H NMR spectrum of cavitand **4** in DMSO-d_6, 25 °C, 400 MHz.

*2.5 ^{1}H NMR spectrum of cavitand **5***

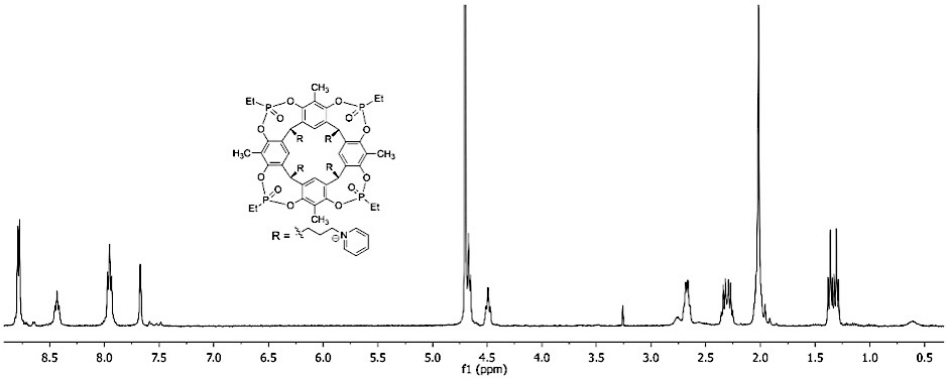


Figure A.10. ^{1}H NMR spectrum of cavitand **5** in D_2O, 25 °C, 400 MHz.

*2.6 ^{1}H NMR spectrum of **DMSI** dye*

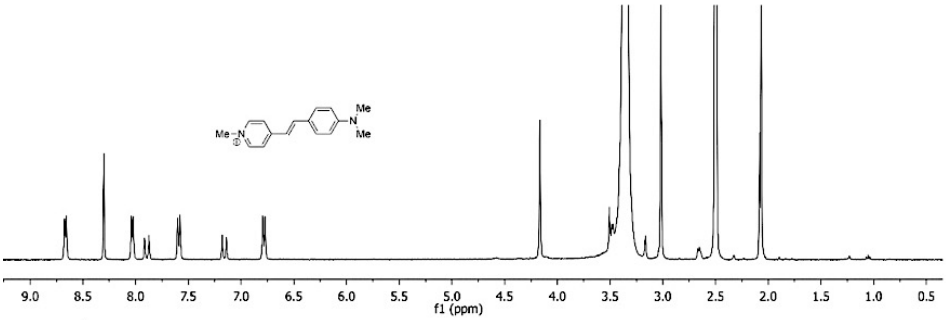


Figure A.11. ^{1}H NMR spectrum of **DMSI** dye in DMSO-d_6, 25 °C, 400 MHz.

*2.7 ^{1}H NMR spectrum of **PMSI** dye*

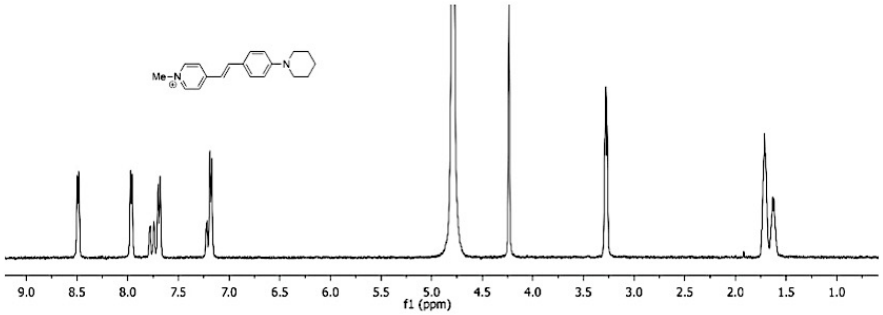


Figure A.12. ^{1}H NMR spectrum of **PMSI** dye in D_2O, 25 °C, 400 MHz.

3. Minimized Host:Dye structures

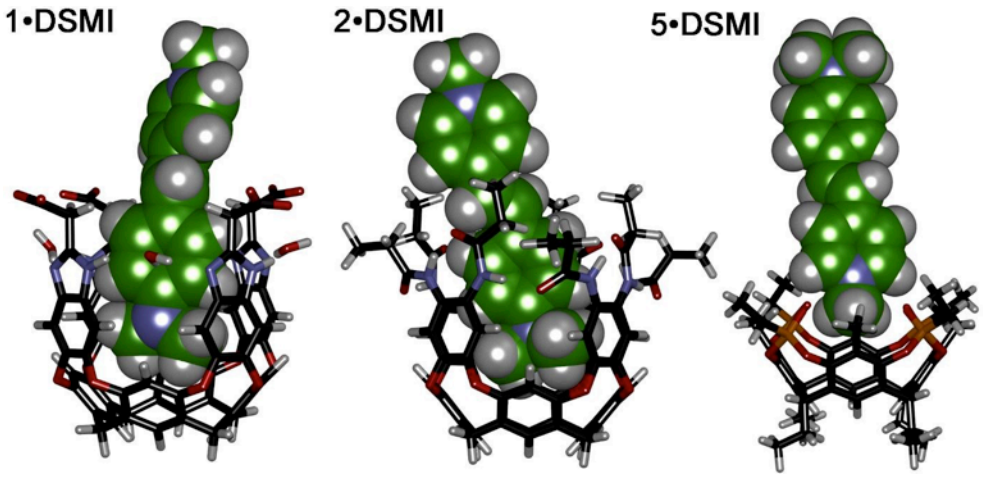
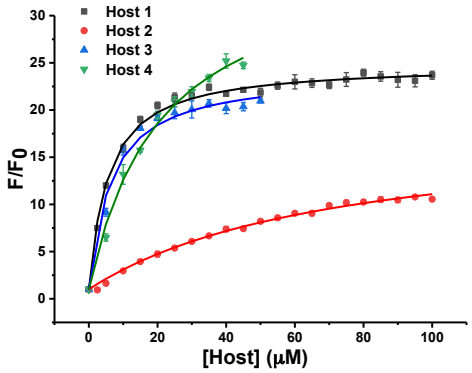


Figure A.13. Minimized structures of hosts **1**, **2** and **5** in complex with **DMSI** dye (SPARTAN, Hartree-Fock). Lower rim substituents were removed for clarity.

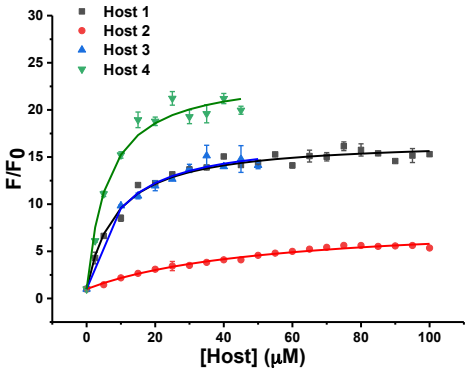
4. Binding affinity measurements

*4.1 Fluorescence titrations of dyes and hosts **1-5***

a) **DSMI** + host **1-4**



b) **PSMI** + host **1-4**



c) **DSMI/PSMI** + host **5**

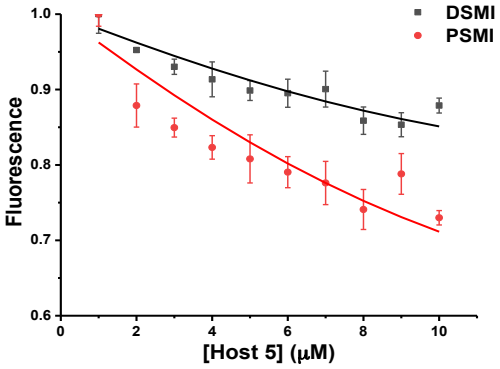


Figure A.14. Affinity measurement of dye: **DSMI**/**PSMI** with host **1-5** via fluorescence. a) **DSMI** + host **1-4**; b) **PSMI** + host **1-4**; c) **DSMI**/**PSMI** + host **5**. For hosts **1-4**, [Dye] = 0.625 μM; for host **5**, [Dye] = 10 μM. Buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Ex/Em = 485nm/605nm.

*4.2 Isothermal calorimetry analysis of dye•**c-myc 2345** binding*

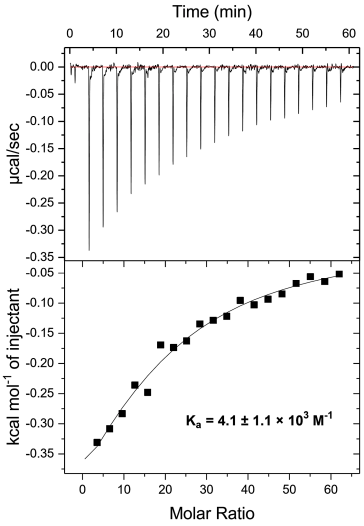


Figure A.15. ITC titrations of increasing amounts of **DSMI** with 10 µM *c-myc 2345*, measured at 25 °C. The 10 µM *c-myc 2345* was placed in the cell and the 3mM **DSMI** in the syringe. Both solutions were diluted with buffer 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Top trace: raw data for the ITC titration. Bottom trace: binding isotherm of the integrated calorimetric titration data. The heat of dilution, measured by the injection of titrant into the buffer solution, was subtracted for each titration to obtain the net reaction heat value. *c-myc 2345* was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.

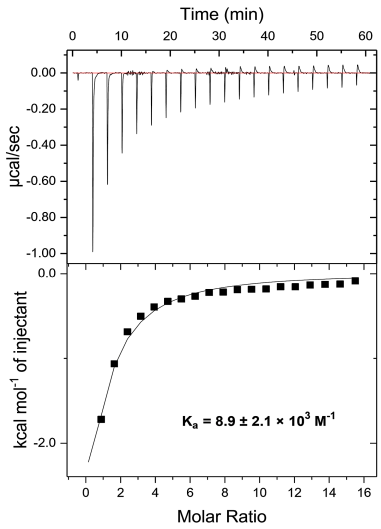


Figure A.16. ITC titrations of increasing amounts of **PSMI** with 40 µM *c-myc 2345*, measured at 25 °C. The 40 µM *c-myc 2345* was placed in the cell and the 3mM **PSMI** in the syringe. Both solutions were diluted with 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Top trace: raw data for the ITC titration. Bottom trace: binding isotherm of the integrated calorimetric titration data. The heat of dilution, measured by the injection of titrant into the buffer solution, was subtracted for each titration to obtain the net reaction heat value. *c-myc 2345* was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.

4.3 Fluorescence titrations of dye•DNA

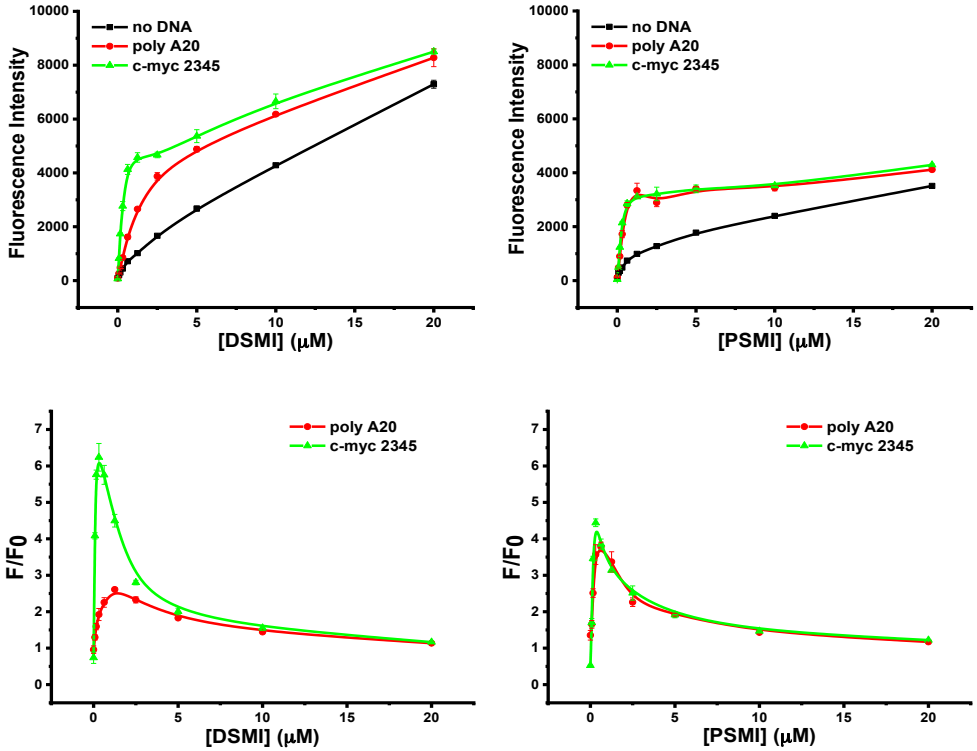
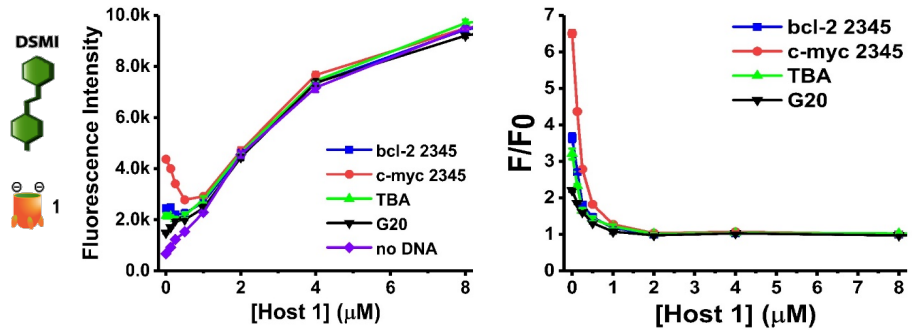


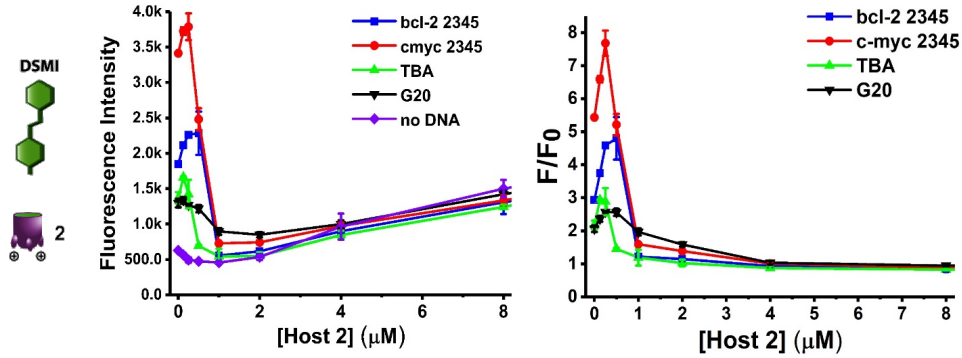
Figure A.17. Fluorescence response curves of *polyA20* or *c-myc 2345* with increasing concentration (0-20 μM) of **DSMI**/**PSMI**. Top: plots using the raw fluorescence counts; Bottom: plots using the fluorescence normalized against that of the dye (F_0 being the dye fluorescence in the absence of DNA). [**DSMI**/**PSMI**] = 0-20 μM, [DNA] = 0.1 μM, 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Ex/Em = 485nm/605nm.
DNA was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.

*4.4 Fluorescence response curves of host addition to **DMSI**•DNA complexes*

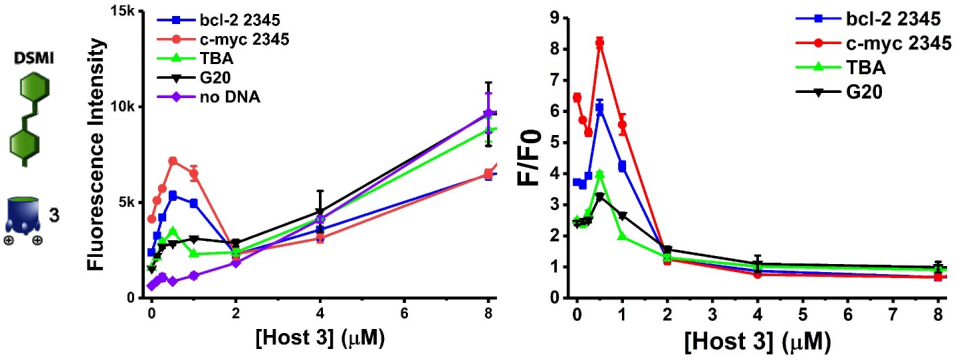
a) Host **1**



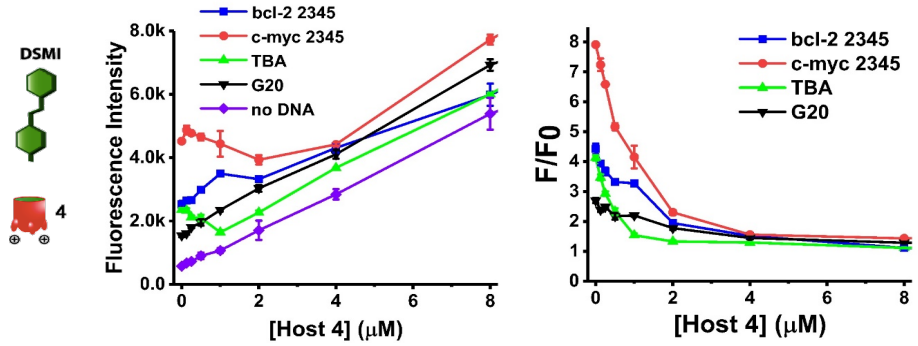
b) Host **2**



c) Host **3**



d) Host **4**



e) Host **5**

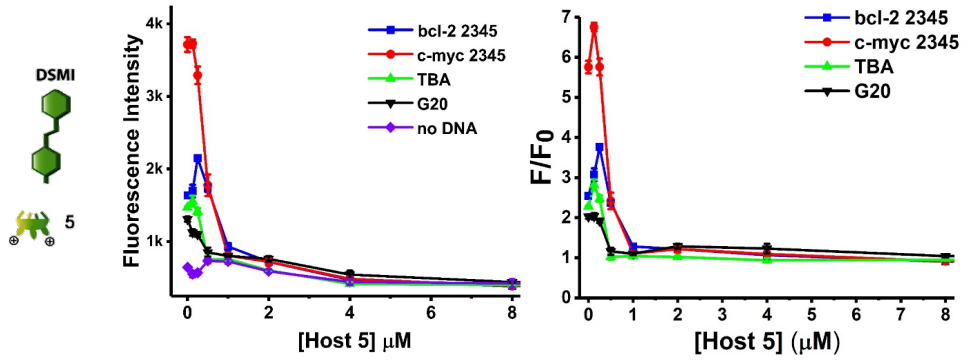
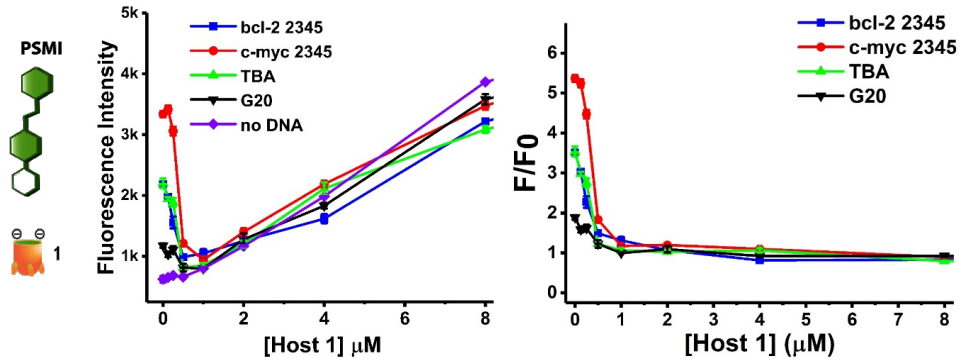


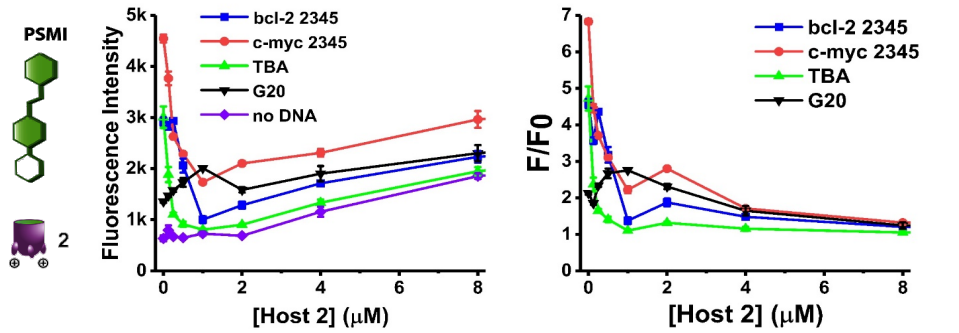
Figure A.18. Fluorescence response curves of **DSMI**•DNA complexes upon titration of hosts 1-5. Left: the raw fluorescence counts (**DSMI** + Host **1-5**); Right: plots normalized to the response of cavitand:**DSMI** in the absence of DNA (F_0). [**DSMI**] = 0.625 μM, [DNA] = 0.1 μM, [Host] = 0-8 μM, 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4.
Ex/Em = 485nm/605nm.
DNA was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.

*4.5 Fluorescence response curves of host addition to **PMSI**•DNA complexes*

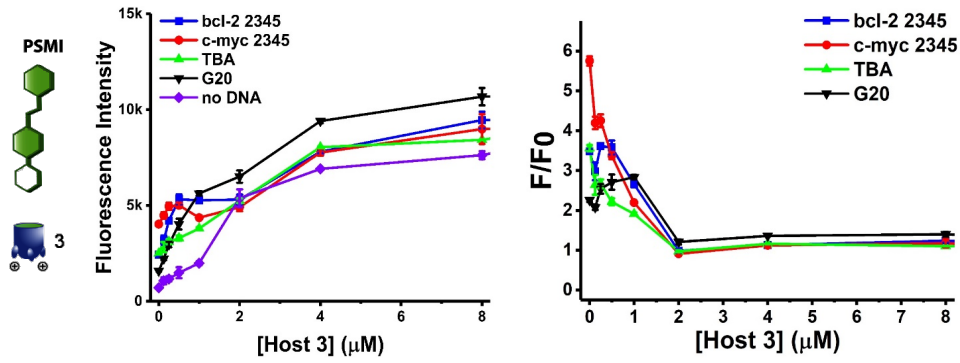
a) Host **1**



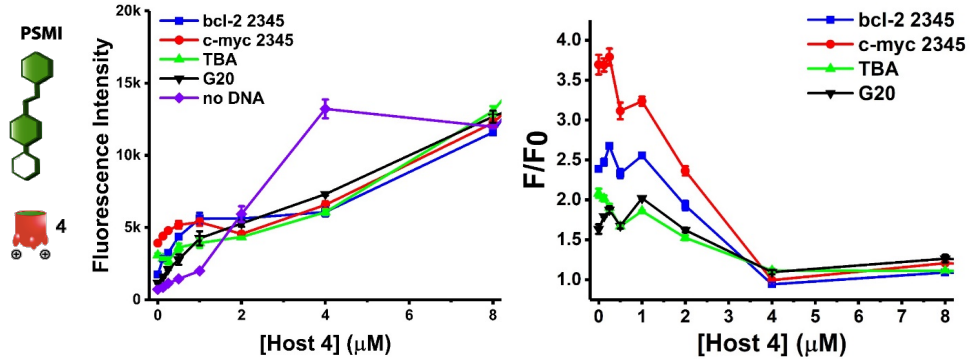
b) Host **2**



c) Host **3**



d) Host **4**



e) Host **5**

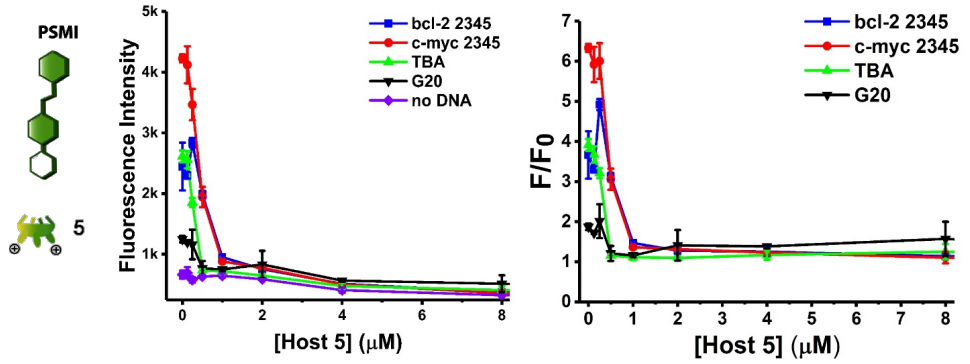
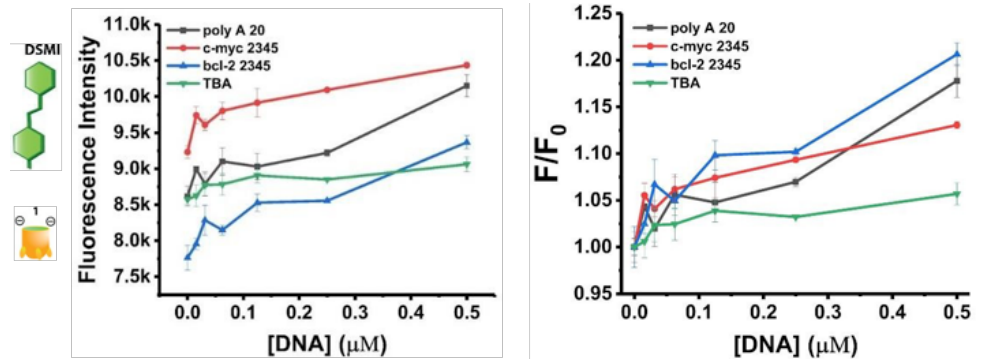


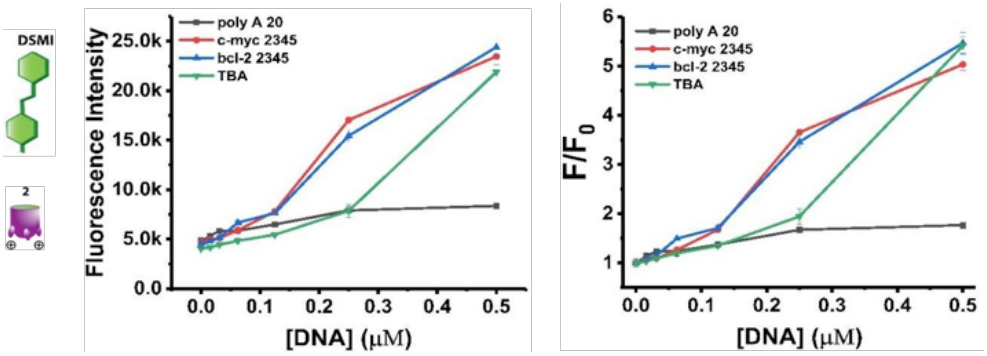
Figure A.19. Fluorescence response curves of **PSMI**•DNA complexes upon titration of Hosts **1-5**. Left: the raw fluorescence counts (**PSMI** + Host **1-5**); Right: plots normalized to the response of cavitand:**PSMI** (in the absence of DNA). [**PSMI**] = 0.625 μM, [DNA] = 0.1 μM, [Host] = 0-8 μM, 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4.
Ex/Em = 485nm/605nm.
DNA was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.

*4.6 Fluorescence response curves of DNA addition to **DMSI**•host complexes*

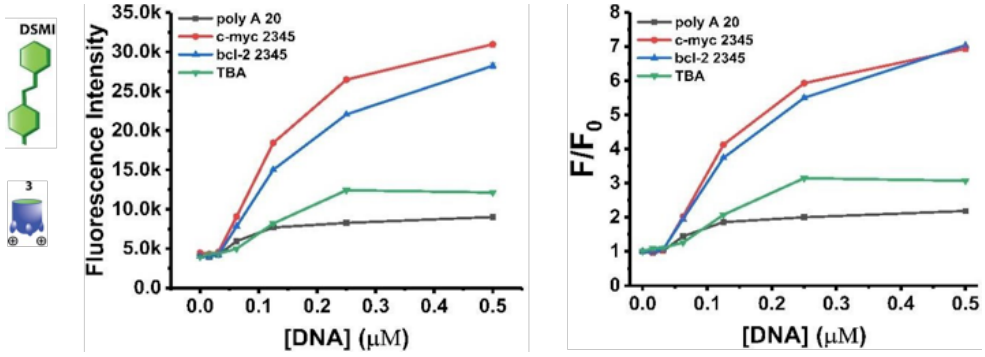
a) Host **1**



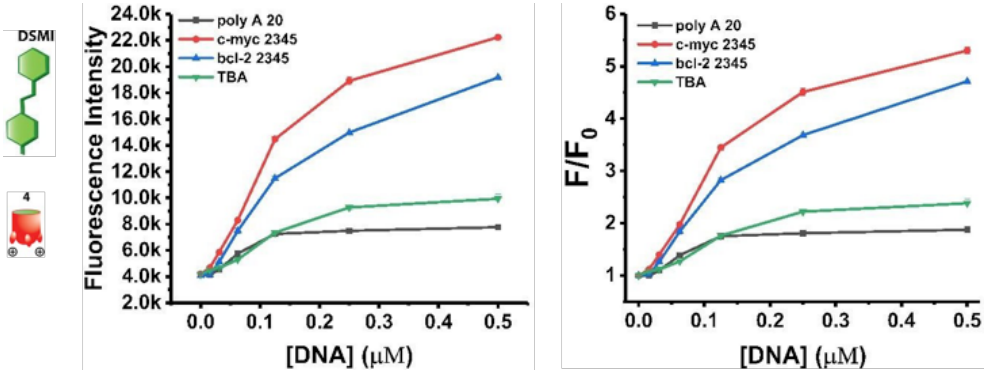
b) Host **2**



c) Host **3**



d) Host **4**



e) Host **5**

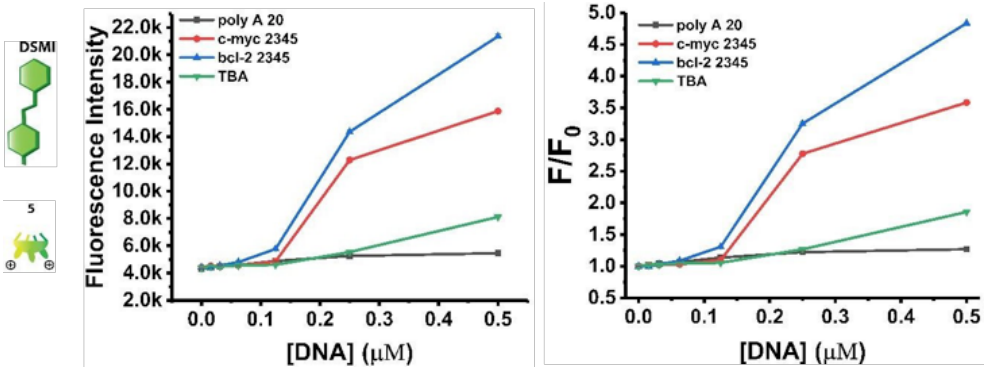
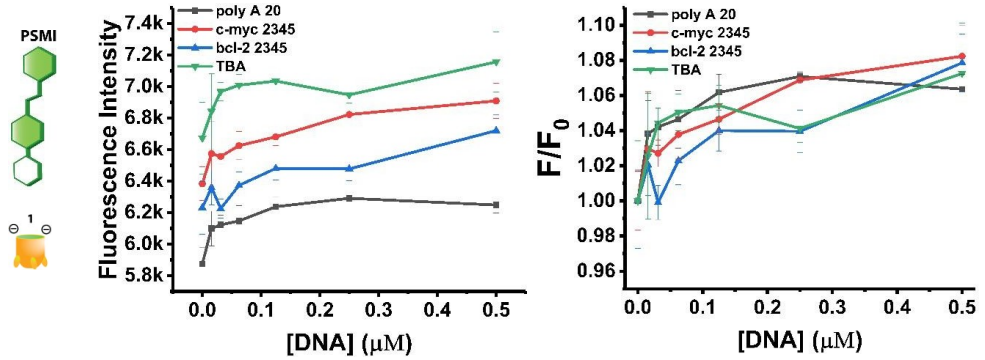


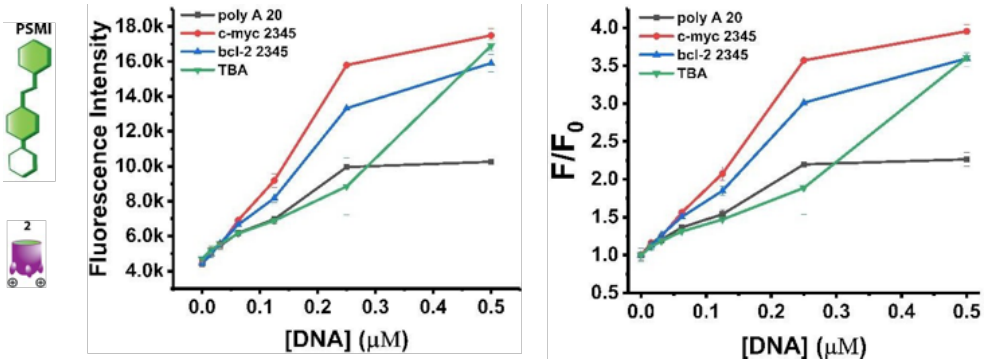
Figure A.20. Fluorescence response curves of **DSMI**•host complexes upon titration of DNA. Left: the raw fluorescence counts (**DSMI** + Host **1-5**); Right: plots normalized to the response of cavitand:**DSMI** in the absence of DNA (F_0). [**DSMI**] = 0.625 μM, [Host] = 1 μM, [DNA] = 0 - 0.5 μM, 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4.
DNA was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.
This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

*4.7 Fluorescence response curves of DNA addition to **PMSI**•host complexes*

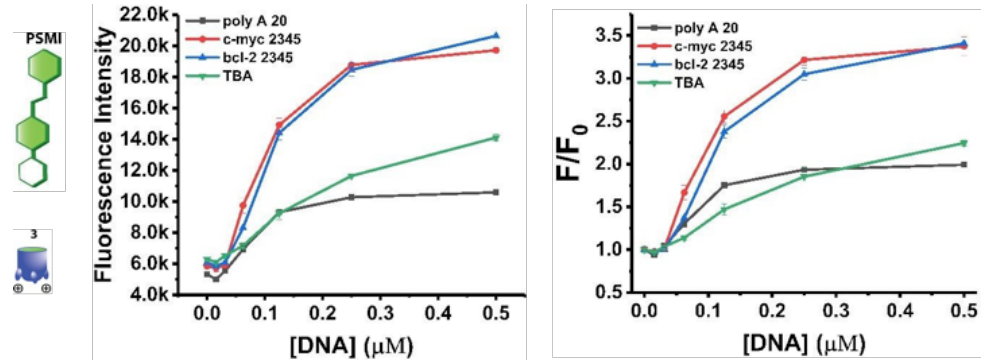
a) Host **1**



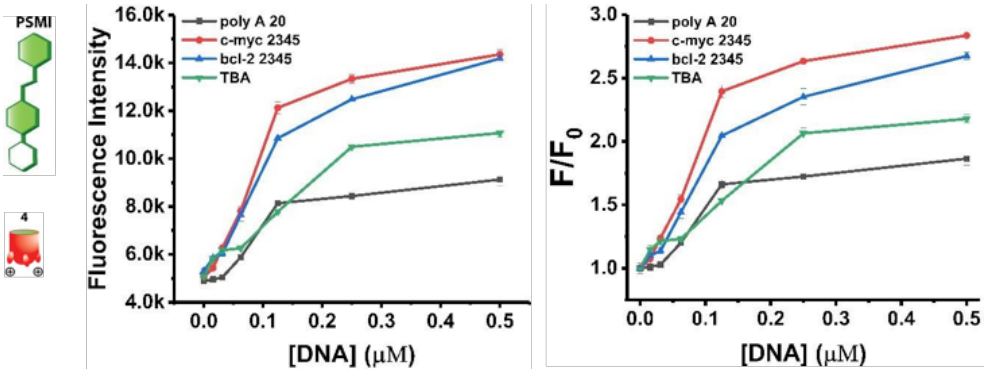
b) Host **2**



c) Host **3**



d) Host **4**



e) Host **5**

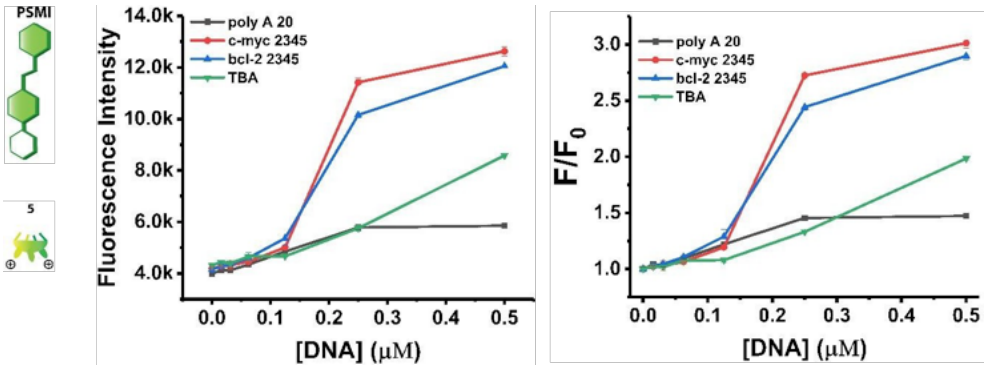


Figure A.21. Fluorescence response curves of **PSMI**•host complexes upon titration of DNA. Left: the raw fluorescence counts (**PSMI** + Host **1-5**); Right: plots normalized to the response of cavitand:**PSMI** in the absence of DNA (F_0). [**PSMI**] = 0.625 μM, [Host] = 1 μM, [DNA] = 0 - 0.5 μM, 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4.
DNA was denatured at 95 °C 5 min, then re-annealing at 4 °C overnight before experiment.
This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

*5. Limit of detection (LOD) for each **host•dye** complex*

*5.1 DNA detection with **DMSI***

LOD/DMSI	*polyA 20*	*c-myc 2345*
Host 2	0.00779 ± 0.00720 µM	0.00616 ± 0.00555 µM
Host 3	0.02480 ± 0.00070 µM	0.02985 ± 0.00186 µM
Host 4	0.01925 ± 0.00026 µM	0.00983 ± 0.00020 µM
Host 5	0.01173 ± 0.00005 µM	0.00476 ± 0.00005 µM
No host	0.01368 ± 0.00042 µM	0.00357 ± 0.00002 µM

Table A.2. LOD calculated from the titration curves shown in **Figure A.20** using 3σ method.

*5.2 DNA detection with **PMSI***

LOD/PMSI	*polyA 20*	*c-myc 2345*
Host 2	0.03125 ± 0.00206 µM	0.02000 ± 0.00084 µM
Host 3	0.02878 ± 0.00038 µM	0.03348 ± 0.00061 µM
Host 4	0.03151 ± 0.00031 µM	0.01986 ± 0.00082 µM
Host 5	0.03469 ± 0.00074 µM	0.03585 ± 0.00053 µM
No host	0.00307 ± 0.00002 µM	0.02120 ± 0.00095 µM

Table A.3. LOD calculated from the titration curves shown in **Figure A.21** using 3σ method.

6. Array Analysis for differentiation of 8 DNAs

6.1 PCA results with Type 1 or Type 2 array

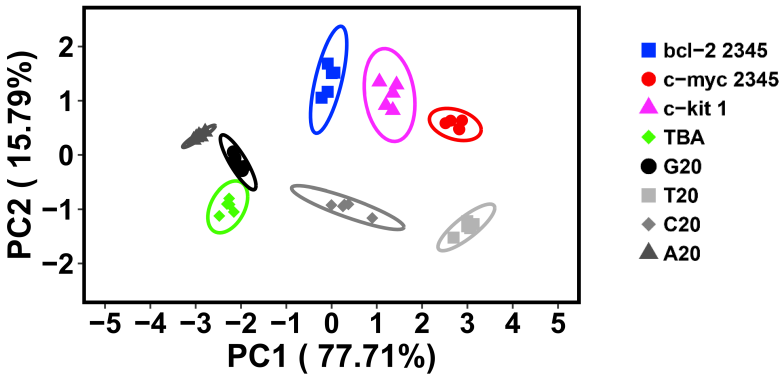


Figure A.22. PCA scores plot for selective sensing of 8 DNA structures using **Type 1 Array** (hosts **1**-**5** or no host, **DMSI** dye): [**DSMI**] = 0.625 µM, [**1**, **4**, or **5**] = 0.25 µM, [**2** or **3**] = 0.50 µM, [DNA] = 0.1 µM, buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Ex/Em = 485nm/605nm. Ellipses indicate 95% confidence.

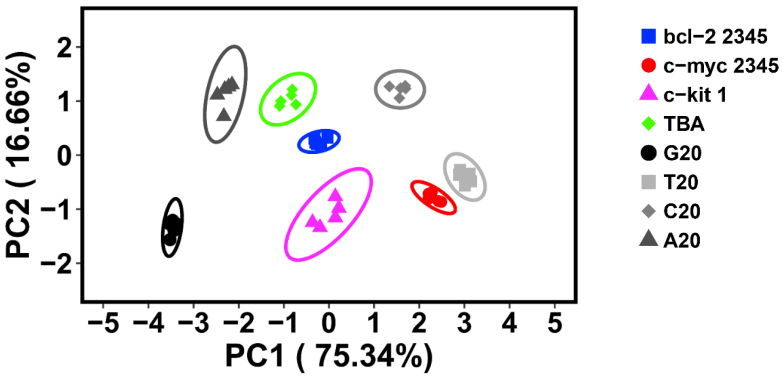


Figure A.23. PCA scores plot for selective sensing of 8 DNA structures using **Type 2 Array** (hosts **1**-**5** or no host, **DMSI** dye): [**PSMI**] = 0.625 µM, [**1**, **3**, or **5**] = 0.25 µM, [**2**] = 1.0 µM, [**4**] = 0.50 µM, [DNA] = 0.1 µM, buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Ex/Em = 485nm/605nm. Ellipses indicate 95% confidence.

6.2 PCA results with 10 host:dye array elements

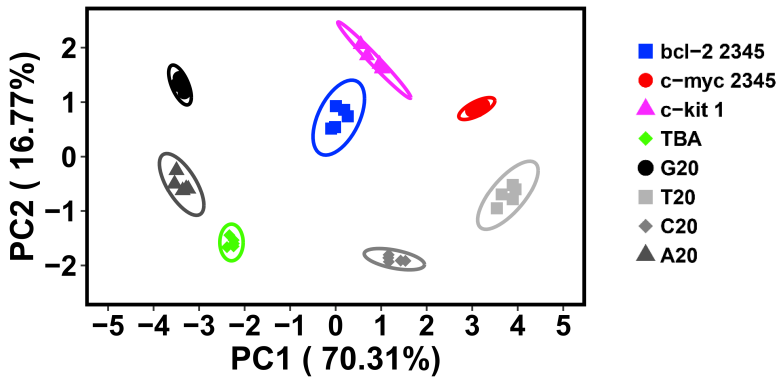


Figure A.24. PCA scores plot for selective sensing of 8 DNA structures using **10 elements sensing array** (the combination of Type 1 and Type 2 arrays, but not including "no host"). **Type 1 Array:** Conditions reported in **Figure A.22**. **Type 2 Array:** Conditions reported in **Figure A.23**. [DNA] = 0.1 µM, buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Ex/Em = 485nm/605nm. Ellipses indicate 95% confidence.

6.3 PCA results using only 2 dyes as array elements

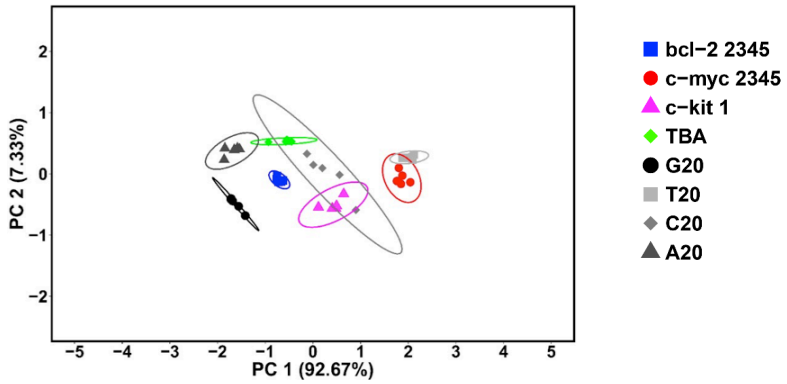
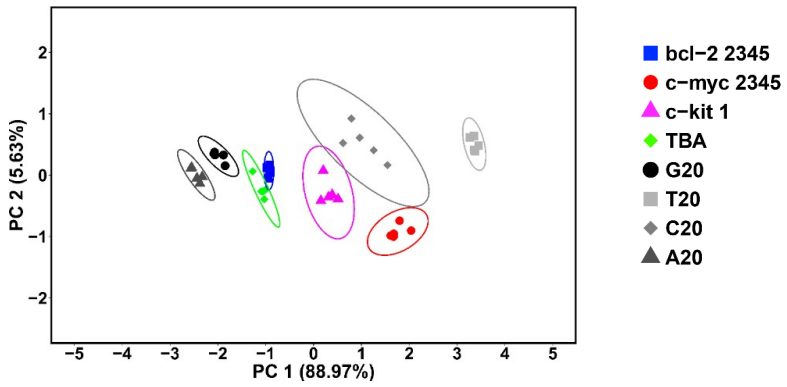


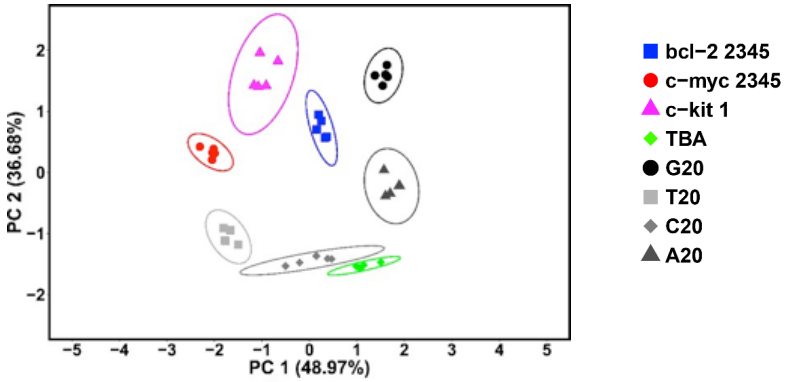
Figure A.25. PCA scores plot for selective sensing of 8 DNA structure using the **2-component array** (only **DMSI** or **PSMI** dye). **2-component Array:** [**DSMI** or **PSMI**] = 0.625 µM. Other sensor conditions are identical to those described in **Figure A.22**. Ellipses indicate 95% confidence.

6.4 PCA results using 4 array elements

a) Host **1**



b) Host **2**



c) Host **4**

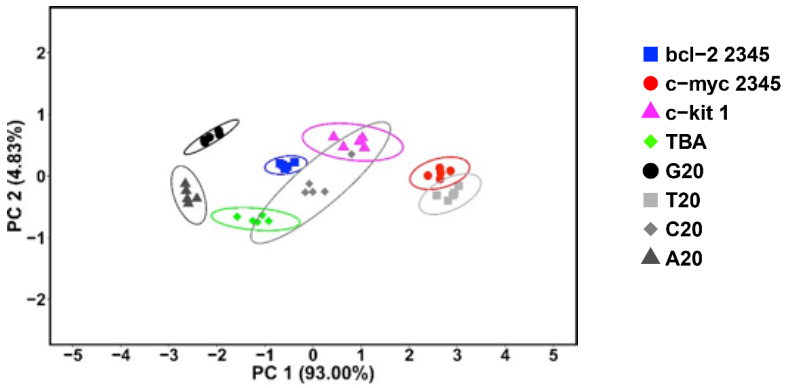


Figure A.26. PCA scores plot for selective sensing of 8 DNA structure using the **4-component array**. (a) **DSMI** or **PSMI** + host **1** or no host from **Type 1** and **Type 2** arrays; (b) **DSMI** or **PSMI** + host **2** or no host from **Type 1** and **Type 2** arrays; (c) **DSMI** or **PSMI** + host **4** or no host from **Type 1** and **Type 2** arrays. Other sensor conditions are identical to those described in **Figure A.22**. Ellipses indicate 95% confidence.

6.5 PCA results using 6 array elements

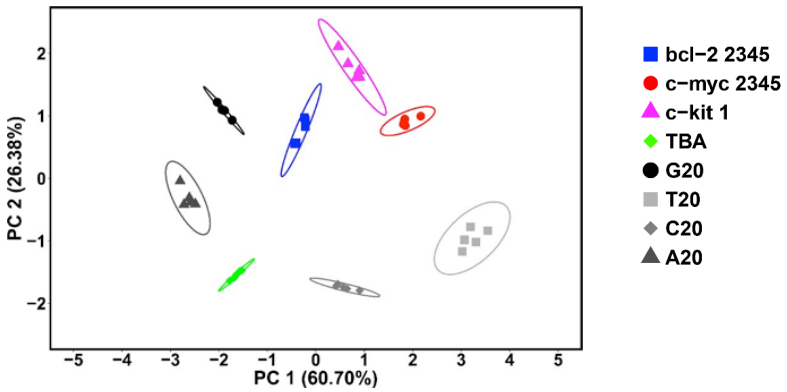
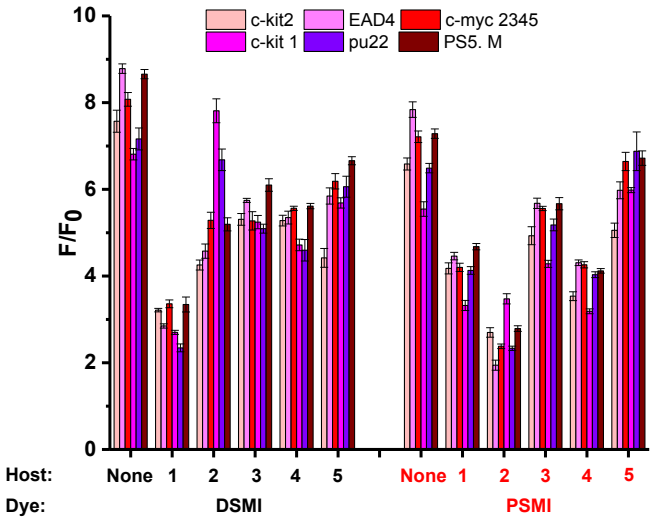


Figure A.27. PCA scores plot for selective sensing of 8 DNA structure using the **6-component array. DSMI** or **PSMI** + host **1**, **2** or **4** from **Type 1** and **Type 2** arrays. Other sensor conditions are identical to those described in **Figure A.22**. Ellipses indicate 95% confidence.

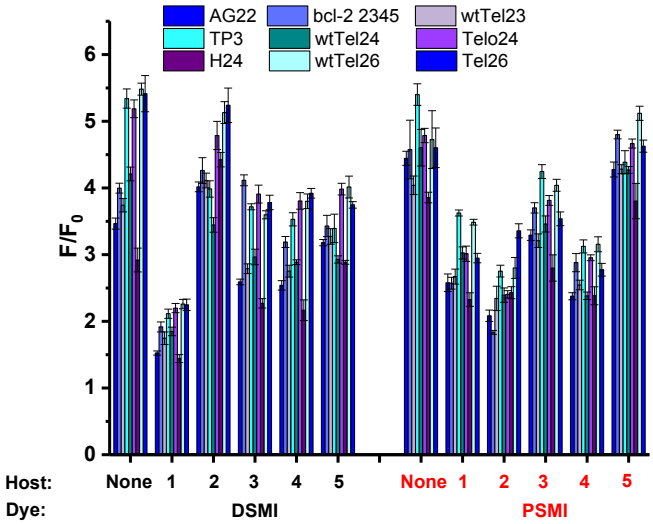
7. Array analysis with selected elements for differentiation of 23 G4 DNA

7.1 Bar plots for array signals from G4 with different topologies

Parallel G4



Hybrid G4



Antiparallel G4

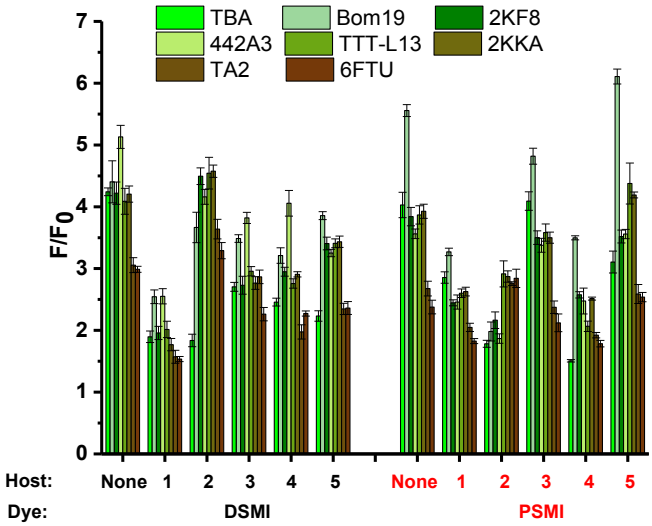


Figure A.28. Full fluorescence response plots of parallel, hybrid and antiparallel G4 sequences obtained with the **12-component array** (the combination of **Type 1** and **Type 2** arrays). **Type 1 Array:** Conditions reported in **Figure A.22**. **Type 2 Array:** Conditions reported in **Figure A.23**. [DNA] = 0.1 µM, buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

7.2 PCA score plots for differentiation of all 23 G4

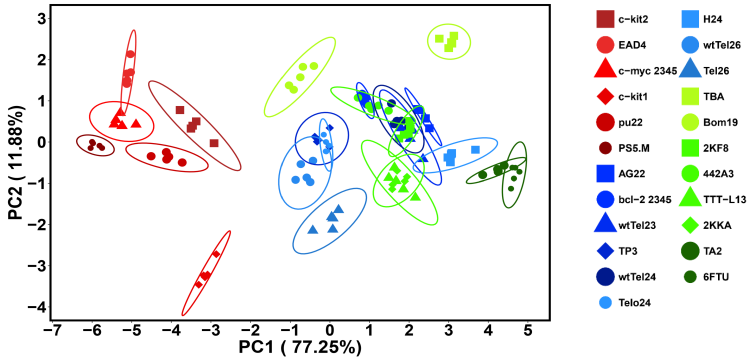


Figure A.29. G-quadruplex classification with **12-components array** (combination of **Type 1** and **Type 2** array). **Type 1 Array:** Conditions reported in **Figure A.22**. **Type 2 Array:** Conditions reported in **Figure A.23**. [DNA] = 0.1 µM, buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Ex/Em = 485nm/605nm. Ellipses indicate 95% confidence.
PCA scores plot generated from the data in **Figure A.28**, by treating each repeat as one individual sample, and processed in RStudio.

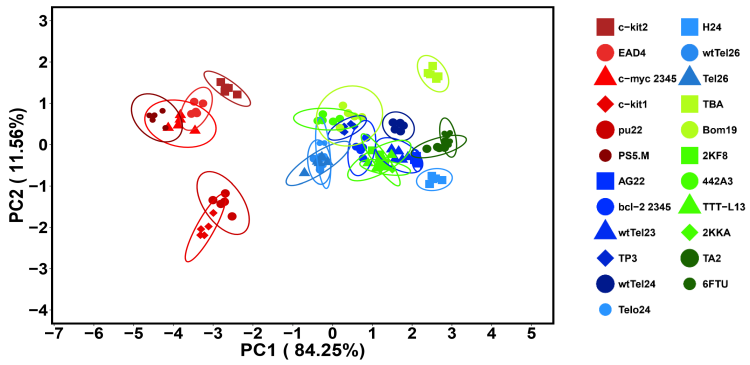


Figure A.30. G-quadruplex classification with only the **Type 1 Array**.
PCA scores plot generated from the data in **Figure A.22**, by treating each repeat as one individual sample, processed in RStudio. Ellipses indicate 95% confidence.

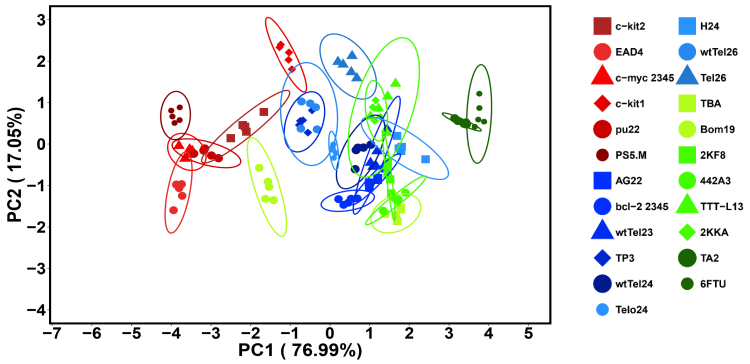


Figure A.31. G-quadruplex classification with only the **Type 2 Array**. PCA scores plot generated from the data in **Figure A.23**, by treating each repeat as one individual sample, processed in RStudio. Ellipses indicate 95% confidence.

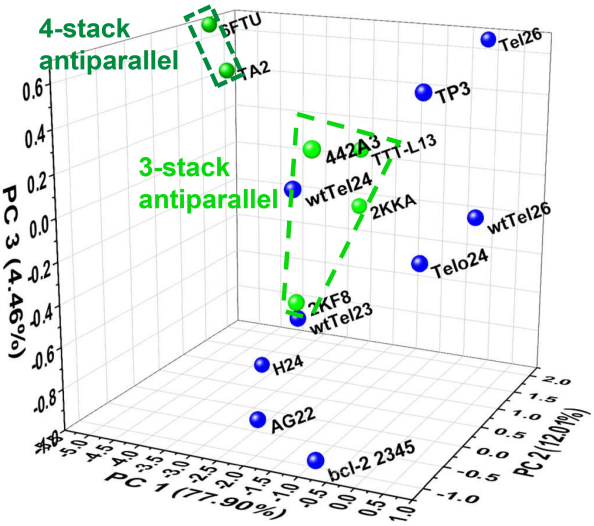
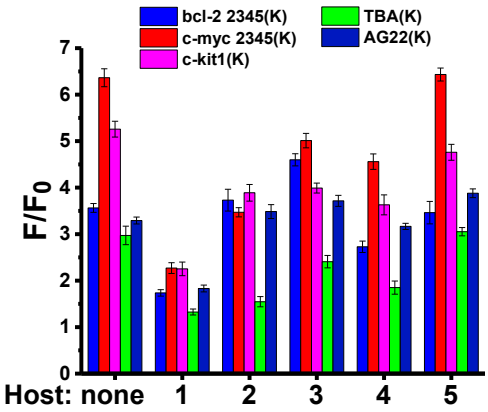


Figure A.32. 3D PCA scores plot generated by subtracting the averaged value of the 5 repeats for each hybrid and antiparallel G4 structure obtained with the **12-component array** (the combination of **Type 1** and **Type 2** arrays) as shown in **Figure A.28** for PCA. **Type 1 Array:** Conditions reported in **Figure A.22**. **Type 2 Array:** Conditions reported in **Figure A.23**. [DNA] = 0.1 µM, buffer: 10 mM KH_2PO_4/K_2HPO_4, 1 mM EDTA, pH 7.4. Ex/Em = 485nm/605nm. Ellipses indicate 95% confidence.

8. Array analysis for topology switch in K^+ and Na^+ Buffer

a) K^+ buffer



b) Na^+ Buffer

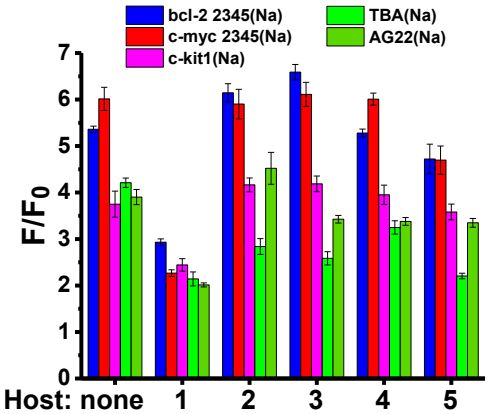


Figure A.33. Fluorescence response plots for structural switching of *AG22* detected by the host:guest sensor array in a) 10 mM KH_2PO_4/K_2HPO_4 buffer; or b) 10 mM NaH_2PO_4/Na_2HPO_4 buffer, using the **Type 1** sensing **array** (condition reported in **Figure A.22**). [DNA] = 0.1 µM, [buffer] = 10 mM, 1 mM EDTA, pH 7.4. Ex/Em = 485nm/605nm.

9. Array analysis in complex media

9.1 Fetal Bovine Serum (FBS)

Bar Plots

Type 1 array **Type 2 array**

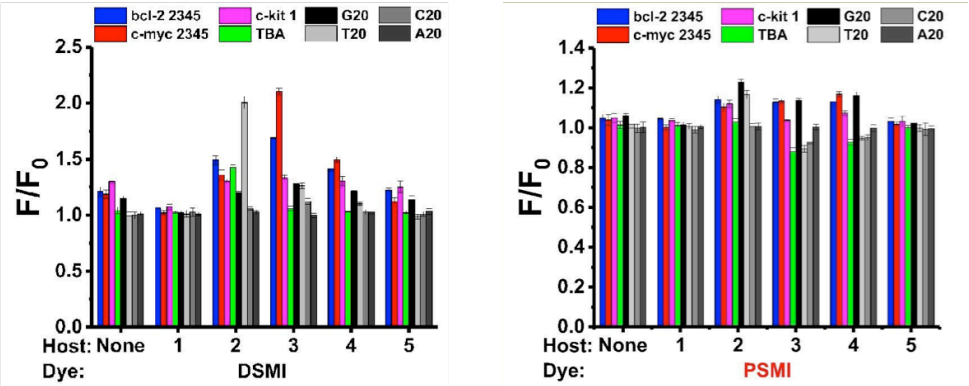
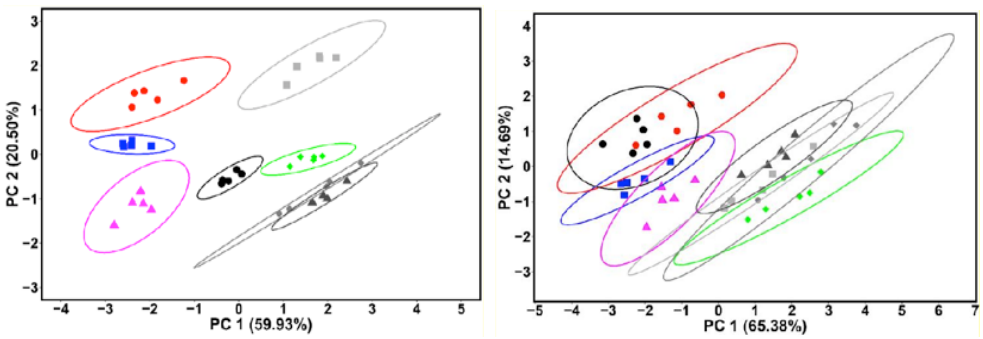


Figure A.34. Fluorescence responses upon addition of the 8 DNA strands to the **host•dye** sensor components in solution containing 5% FBS. Left: **Type 1** array. Right: **Type 2** array. [DNA] = 0.1 µM, 5% FBS, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4.
This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

PCA score Plots

Type 1 array **Type 2 array**



12-components array (Type 1 + Type 2 array)

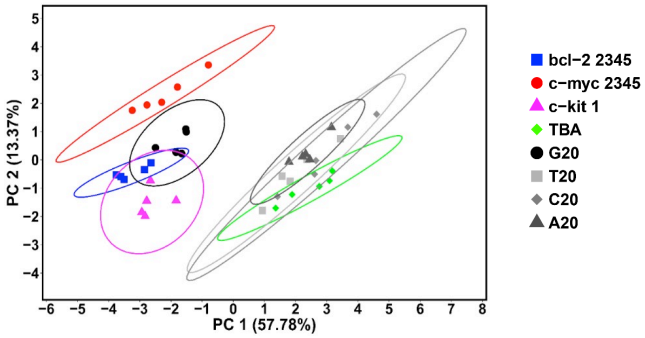


Figure A.35. PCA scores plot generated from the data for DNA sensing array in solutions containing 5% FBS in **Figure A.34**. **Type 1** array, **Type 2** array or the combination of **Type 1** and **Type 2** arrays were used. [DNA] = 0.1 μM, 5% FBS, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

9.2 Lysozyme (0.1 μM)

Bar Plot

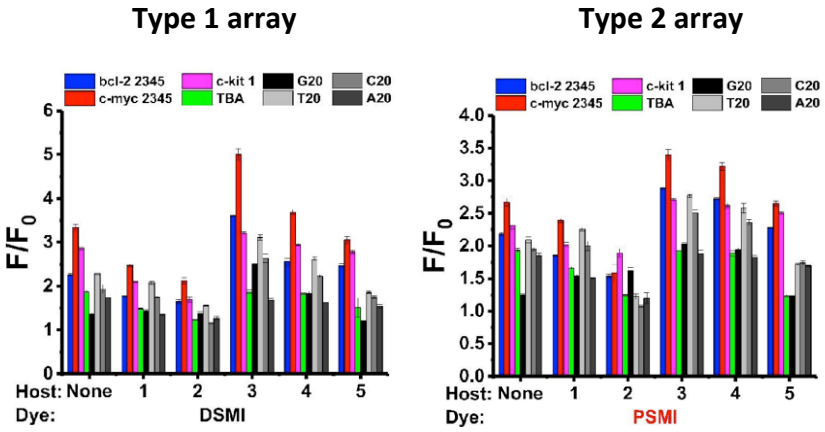


Figure A.36. Fluorescence responses upon addition of the 8 DNA strands to the **host•dye** sensor components in solution containing 0.1 μM Lysozyme. Left: **Type 1** array. Right: **Type 2** array. [DNA] = 0.1 μM, 0.1 μM Lysozyme, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

PCA score Plots

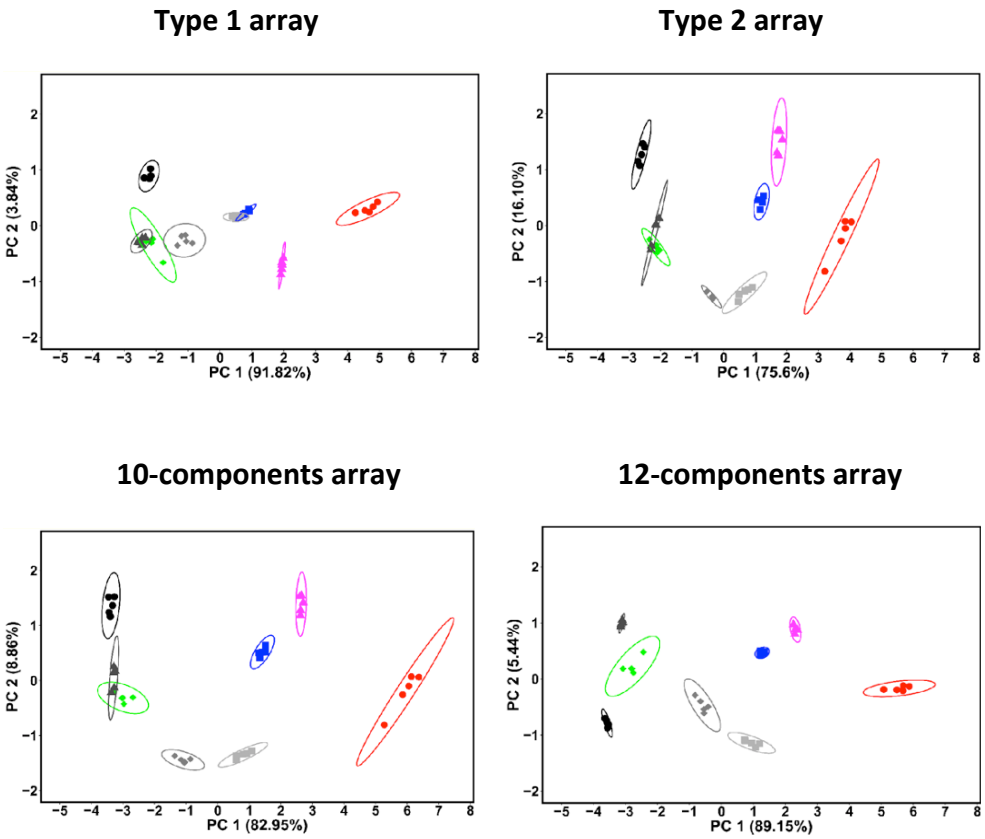


Figure A.37. PCA scores plot generated from the data for DNA sensing array with the presence of 0.1 µM Lysozyme in **Figure A.36**. **Type 1** array, **Type 2** array, **10-components** array (the combination of **Type 1** and **Type 2** arrays not including "host **2**") or **12-components** array (the combination of **Type 1** and **Type 2** arrays) were used. [DNA] = 0.1 µM, 0.1 µM Lysozyme, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

9.3 Lysozyme (1 μM)

Bar Plot

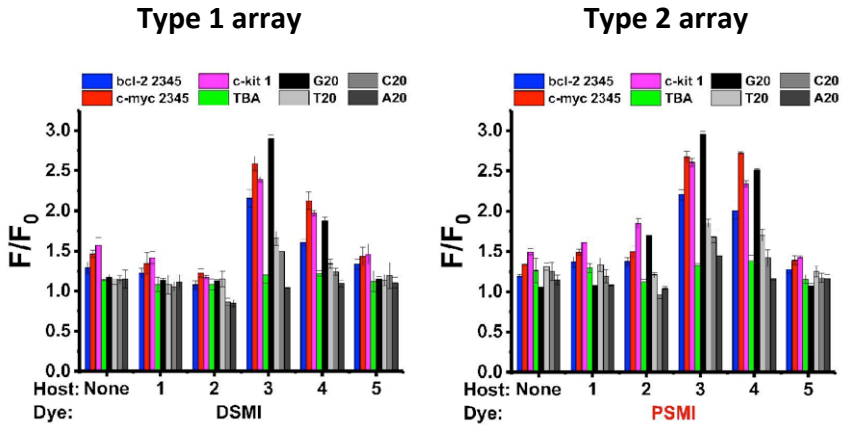


Figure A.38. Fluorescence responses upon addition of the 8 DNA strands to the **host•dye** sensor components in solution containing 1 μM Lysozyme. Left: **Type 1** array. Right: **Type 2** array. [DNA] = 0.1 μM, 1 μM Lysozyme, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

PCA score Plots

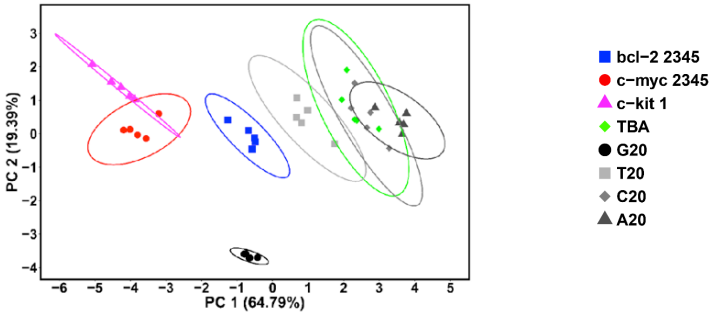


Figure A.39. PCA scores plot generated from the data for DNA sensing array with the presence of 1 μM Lysozyme in **Figure A.38** using the **12-components** array (combination of **Type 1** and **Type 2** arrays). [DNA] = 0.1 μM, 1 μM Lysozyme, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

9.4 Lactose (20 µM)

Bar Plot

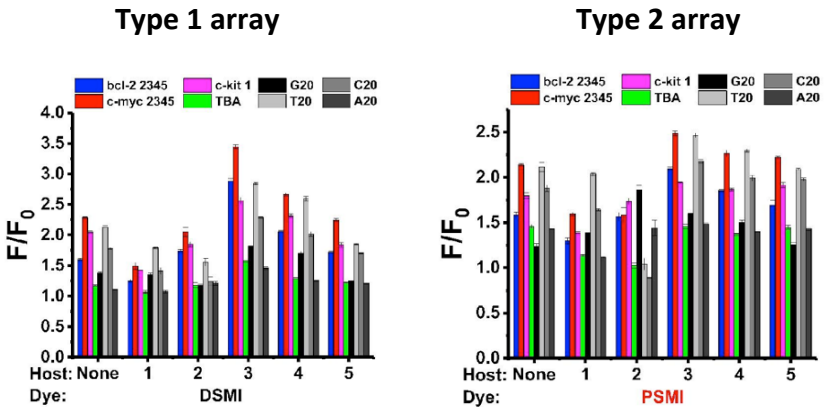
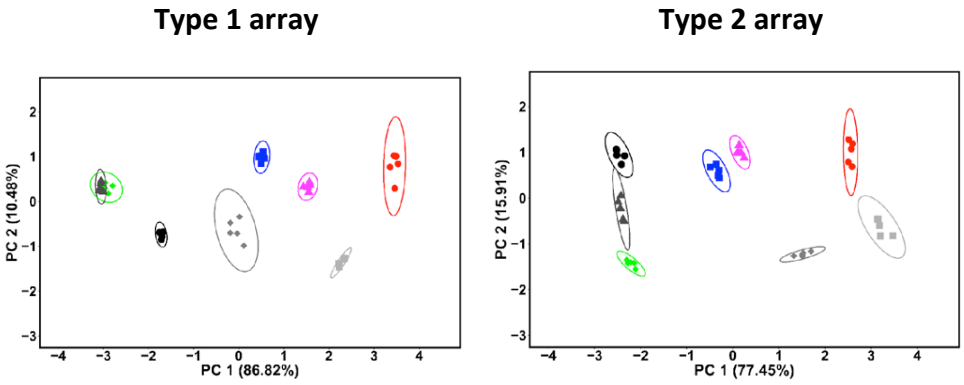


Figure A.40. Fluorescence responses upon addition of the 8 DNA strands to the **host•dye** sensor components in solution containing 20 µM Lactose. Left: **Type 1** array. Right: **Type 2** array. [DNA] = 0.1 µM, 20 µM Lactose, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

PCA score Plots



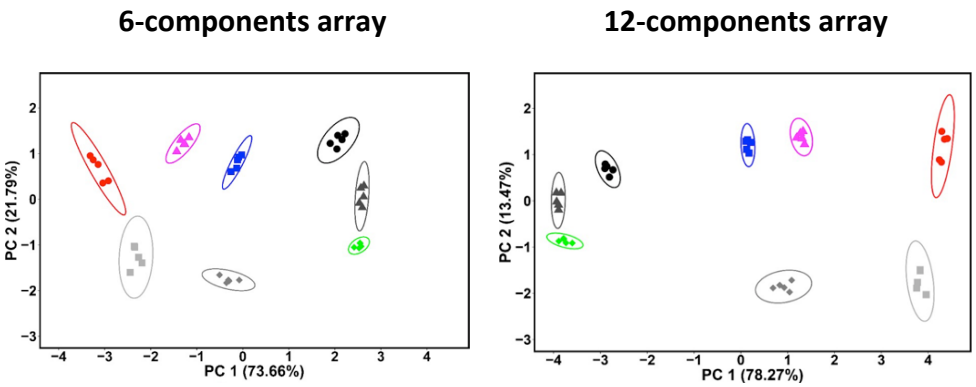


Figure A.41. PCA scores plot generated from the data for DNA sensing array with the presence of 20 μM Lactose in **Figure A.40**. **Type 1** array, **Type 2** array, **6-components** array (**DMSI** or **PSMI** + hosts **2**, **4** or no host from **Type 1** and **Type 2** array) or the combination of **Type 1** and **Type 2** arrays (**12-components** array) were used. [DNA] = 0.1 μM, 20 μM Lactose, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

10. Array-based recognition of the mixture of non-structural DNA and G4

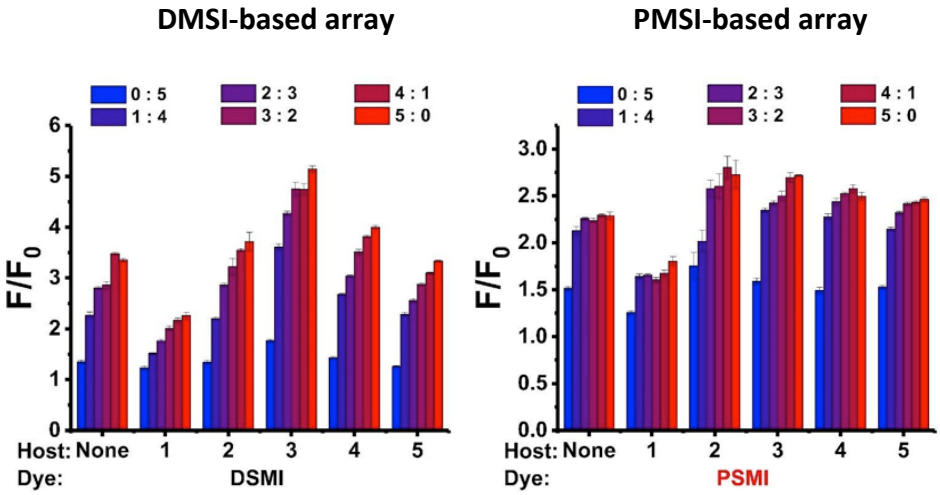


Figure A.42. Array-based analysis of the mixture of *polyA* and *c-myc* 2345. Fluorescence responses for the DMSI-based array elements (left) and the PMSI-based array elements (right) to the solutions. Total [DNA] = 0.5 μM, 10 mM KH_2PO_4/K_2HPO_4 buffer, 1 mM EDTA, pH 7.4. This set of data was acquired in BioTek Synergy HT microplate reader with the filter set of Ex/Em = 485nm/600nm.

[DNA]/μM	1	2	3	4	5	6	Control
polyA 20	0.5	0.4	0.3	0.2	0.1	0	0
c-myc 2345	0	0.1	0.2	0.3	0.4	0.5	0
c-myc 2345 to *polyA 20* mole ratio	0:5	1:4	2:3	3:2	4:1	5:0	-

Table A.4. Component concentrations in each tested solution.

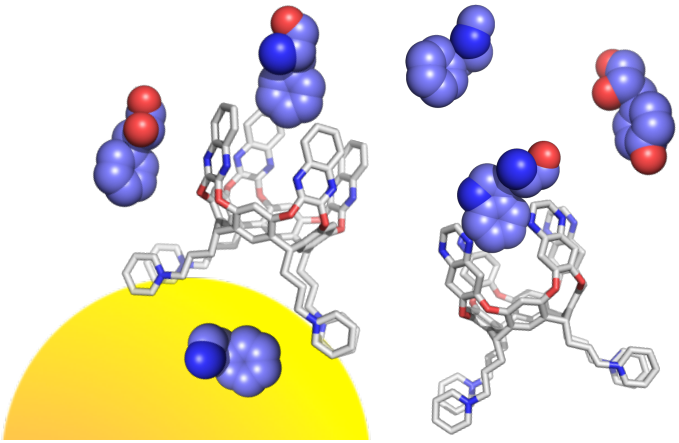
11. Appendix References

1. M. Jin, X. Liu, X. Zhang, L. Wang, T. Bing, N. Zhang, Y. Zhang, D. Shangguan, *ACS Appl. Mater. Interfaces* **2018**, *10*, 25166.
2. X. Cheng, X. Liu, T. Bing, Z. Cao, D. Shangguan, *Biochemistry* **2010**, *48*, 7817.
3. Q. Yang, J. Xiang, S. Yang, Q. Li, Q. Zhou, A. Guan, X. Zhang, H. Zhang, Y. Tang, G. Xu, *Nucleic Acids Res.* **2010**, *38*, 1022.
4. D. Lin, X. Fei, Y. Gu, C. Wang, Y. Tang, R. Li, J. Zhou, *Analyst* **2015**, *140*, 5772.
5. J. Ferreira, T. Santos, P. Pereira, M.C. Corvo, J.A. Queiroz, F. Sousa, C. Cruz, *Analyst* **2017**, *142*, 2982
6. P. Agrawal, E. Hatzakis, K. Guo, M. Carver, D. Yang, *Nucleic Acids Res.* **2013**, *41*, 10584.
7. H. Peng, A.M. Newbigging, Z. Wang, J. Tao, W. Deng, X.C. Le, H. Zhang, *Anal. Chem.* **2018**, *90*, 190.
8. Y. Gu, D. Lin, Y. Tang, X. Fei, C. Wang, B. Zhang, J. Zhou, *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* **2018**, *191*, 180.
9. C. Lin, G. Wu, K. Wang, B. Onel, S. Sakai, Y. Shao, D. Yang, *Angew. Chem. Int. Ed.* **2018**, *57*, 10888.
10. A. Sengar, J.J. Vandana, V.S. Chambers, M. Di Antonio, F.R. Winnerdy, S. Balasubramanian, A.T. Phan, *Nucleic Acids Res.* **2019**, *47*, 1564.
11. Y. Ma, Y. Tsushima, M. Sakuma, S. Sasaki, K. Iida, S. Okabe, H.Seimiya, T. Hirokawa, K. Nagasawa, *Org. Biomol. Chem.* **2018**, *16*, 7375.
12. G. Wu, L. Chen, W. Liu, D. Yang, *Molecules* **2019**, *24*, 1578.
13. F. Hao, Y. Ma, Y. Guan, *Molecules* **2019**, *24*, 1863.
14. S. Amrane, R.W.L. Ang, Z.M. Tan, C. Li, J.K.C. Lim, J.M.W. Lim, K.W. Lim, A.T. Phan, *Nucleic Acids Res.* **2009**, *37*, 931.
15. R. del Villar-Guerra, J.O. Trent, J.B. Chaires, *Angew. Chem. Int. Ed.* **2018**, *57*, 7171.
16. M. Cheng, Y. Cheng, J. Hao, G. Jia, J. Zhou, J.-L. Mergny, C. Li, *Nucleic Acids Res.* **2018**, *46*, 9264.
17. A. Guédin, L.Y. Lin, S. Armane, L. Lacroix, J.-L. Mergny, S. Thore, L.A. Yatsunyk, *Nucleic Acids Res.* **2018**, *46*, 5297

CHAPTER 4

Molecular recognition of aromatic amino acids by deep cavitands

Amino acids (AAs) represent an ideal playground for testing complexation ability and selectivity of synthetic receptors, due to their biological relevance and chemical diversity. The recognition of aromatic AAs was explored by using water soluble tetraquinoxaline cavitands (QxCav) and benzopyrazine cavitands (BzPyCav). ^{1}H NMR experiments, ITC analyses and solid state studies were performed in water to probe the ability of these cavitands in recognizing L-Phenylalanine and L-Tryptophan. The capability of the cavitand receptors to recognize aromatic guests was explored as well at the solid/liquid interphase by grafting QxCav onto the surface of gold nanoparticles.



4.1. INTRODUCTION

Molecular recognition in aqueous media and the application of water soluble synthetic molecular receptors for the detection of biologically important analytes, inspired supramolecular chemists to create unique synthetic receptors based on macrocyclic structures such as cavitands, calixarenes, cyclodextrins and cucurbiturils, capable of binding a cornucopia of different biological targets. However, receptors for aromatic and essential amino acids (AAs) such as phenylalanine (Phe) and tryptophan (Trp) are lacking. They are important biomedical markers since their overabundance can cause Phenylketonuria[1] or Hartnup disorder.[2]

Phenylketonuria. Considered to be among the most spread rare metabolic diseases in the world, affecting about 50,000 people and 1 child out of about 15,000 births in Europe, phenylketonuria (PKU) represents one of the challenges to early diagnosis.[3] Causing serious and irreversible intellectual disability together with autism, motor deficits, convulsions and psychiatric disorders, PKU is a congenital disease of phenylalanine metabolism and its conversion to tyrosine.

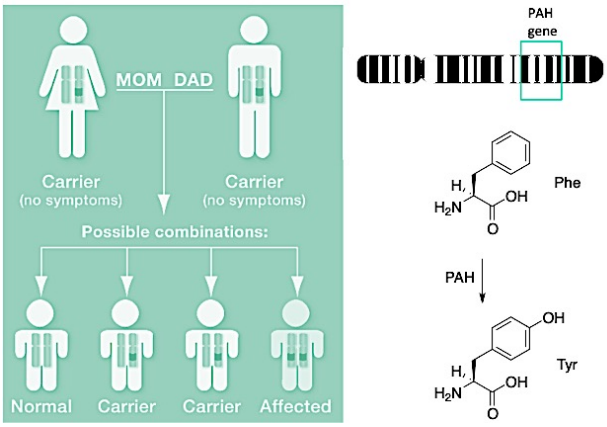


Figure 4.1. Autosomal recessive inheritance (*left*), chromosome 12 and related PAH gene (*upper right*), conversion of Phe to Tyr by PAH enzyme (*bottom right*).

PKU is caused by mutations of the gene for phenylalanine hydroxylase (PAH), found on chromosome 12, which converts Phe to other essential compounds in the body, in particular tyrosine (**Figure 4.1**). Tyrosine is necessary for the production of neurotransmitters like epinephrine, norepinephrine, and dopamine.
A lack in the catabolic pathway of Phe due to the reduced catalytic activity of the enzyme PAH results in an accumulation of the amino acid and consequently serious health problems related to it.
As a recessive hereditary disease, PKU should be diagnosed in the neonatal stages to prevent the onset of symptoms. Thanks to special screening, PKU can be identified and treated by special low-Phe diet. However, in order to maintain the level of concentration of the amino acid below the thresholds set by the national health systems patients are in need of constant monitoring.

Hartnup disease*.* Named after the surname of the first known family affected, Hartnup disease is a hereditary disorder caused by a genetic mutation of an enzyme (B^0AT1) operating in the transport of neutral AAs, that derive from proteins' metabolism, inside the renal and intestinal cells.[4]
Hartnup, like PKU, is an autosomal recessive disease. Presenting symptoms similar to pellagra as pellagrous rash, Hartnup disorder differed from it for the inability to assimilate neutral AAs leading to generalized nutritional defects and, in some cases, neuropsychiatric abnormalities and hepatic and myocardial failure. Under normal conditions, neutral amino acids such as Trp are absorbed in the intestine by the specific membrane transporter. A genetic defect of the enzyme B^0AT1 results in reduced absorption of these amino acids. Consequently, the whole organism is not able to synthesize other important biological molecules, such as niacin from Trp, causing a huge imbalance in the entire human system.

Relevant examples of suitable synthetic receptors for the aromatic AA recognition like cyclodextrins, calixarenes and cucurbiturils have been illustrated in paragraph **1.2.** Here we want to exploit the exquisite molecular recognition properties of quinoxaline cavitand towards aromatic molecules in water taking in account the binding energy of

supramolecular host–guest systems in water is largely driven by hydrophobic effect. Thermodynamic constraints of supramolecular host–guest chemistry in water needs a different approach compared with the well-studied interactions of supramolecular systems in organic solvents. An optimal balance between hydrophilic and hydrophobic interactions needs to be reached in order to ensure sufficient solubility in water of the host while warding off water molecules surrounding the guest in order to allow additional secondary interactions.[5] One of the main challenges in supramolecular chemistry is the design and synthesis of synthetic receptors that have both a high affinity and a high selectivity for guests in aqueous environment to overcome the competitive influence of the water.[6]

4.1.1. QUINOXALINE CAVITANDS

Firstly synthetized by Cram,[7] tetraquinoxaline cavitands (QxCav) are conveniently prepared by bridging the phenolic OH groups of a resorcinarene scaffold with 1,2-dichloroquinoxaline walls *via* a nucleophilic aromatic substitution. Such functionalization leads to the formation of a deep, hydrophobic and electron rich cavity suitable for aromatic guests complexation thanks to the presence of multiple π-π and CH-π interactions. Beyond complexation properties, the conformational flexibility of the QxCav system is demonstrated by the ability of the quinoxaline walls to arrange themselves either in axial or equatorial positions.

Hence, QxCav can reversibly switch between two different spatially well-defined conformations: a so called "*vase*" form (C_{4v} symmetry), in which the cavitand exhibits a closed deep cavity, and a "*kite*" conformation (C_{2v} symmetry), in which the quinoxaline flaps are completely open on the level of the molecule, providing an extended flat surface (**Figure 4.2**). Intermediate structures between these two conformers are less stable.

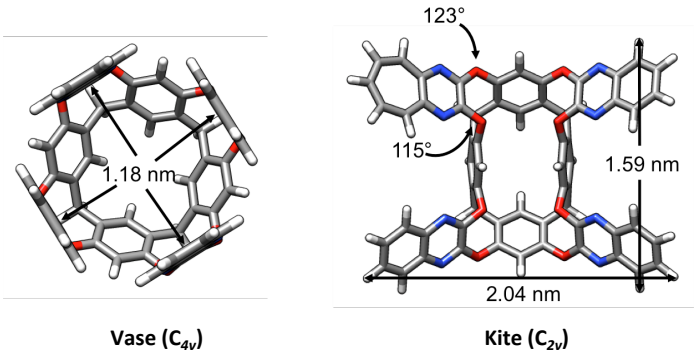


Figure 4.2. QxCav structures: *vase* conformer (*left*) and *kite* conformer (*right*).

Extensive studies pointed out that the behavior of these molecules and their equilibrium between *vase-kite* conformations is due to several factors. In particular, this conformational interconversion is temperature,[7a] pH[8]and solvent dependent.[9]

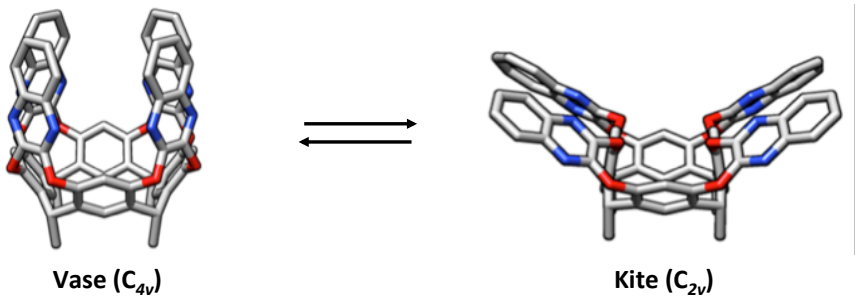


Figure 4.3. *Vase-kite* switching by controlling temperature, pH and solvent polarity.

The *vase* conformation is prevalently present above 25 °C while below -60 °C, the expanded *kite* conformer is the only present in solution as consequence of the different solvent interactions of the two conformers.[10] In fact, the more extended surface of the *kite* structure is able to interact with a higher number of solvent molecules than the more confined surface of the *vase* form. For both structures, these interactions are enthalpically favored but entropically opposed. At low temperatures (below -60 °C), the extended conformation is preferred as the entropy contribution is reduced and overridden by the enthalpy gain. On the contrary, at higher temperature, the solvation entropic

term becomes unfavorable because of a reduced solvent translational freedom, favoring the *vase* form.

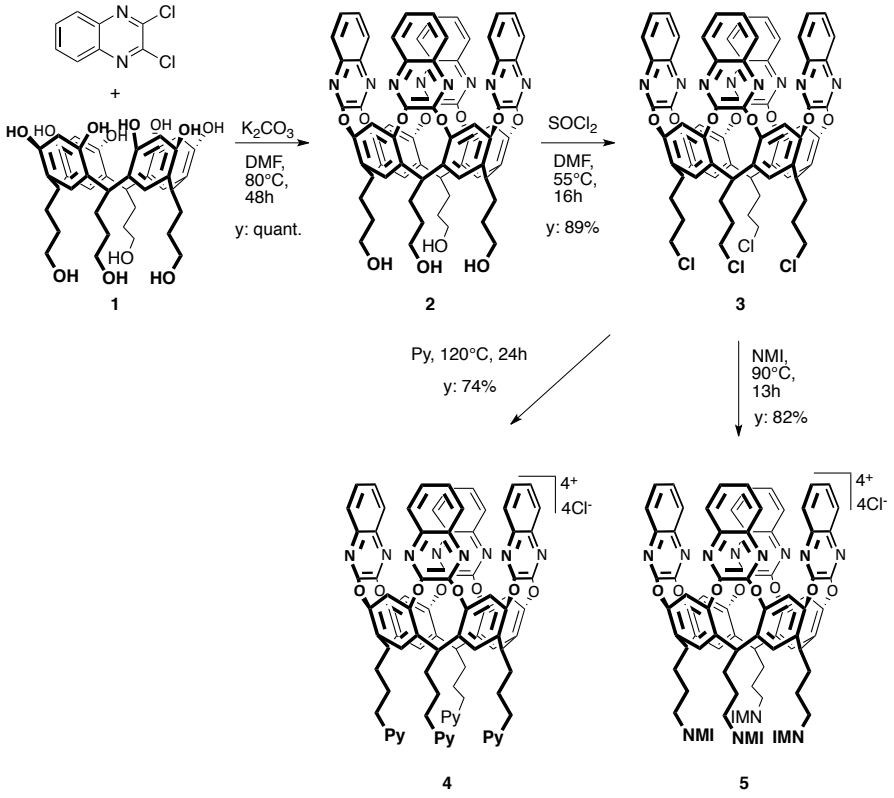
The conformational switch of QxCav upon pH variations highlights the perspective of exploiting pH as an effective method to exert dynamic control.[8] In fact, the addition of strong organic acid, such as trifluoroacetic acid (TFA), leads to the conversion from the *vase* to the *kite* form because of the protonation of the quinoxaline nitrogen and the establishment of Coulomb repulsion between the walls. The peculiarity of the system relies in the reversibility of the processes. In fact, the addition of a base to the system allows the deprotonation of the quinoxaline units making the interconversion *vase-kite* totally reversible.

A remarkable solvent-dependent switching behavior was also pictured out in a detailed study.[9] The Authors demonstrated that two different factors contribute to control the *vase-kite* switch: cavity solvation and solvent hydrogen-bonding acidity. Chlorinated, apolar and small benzene-based aromatic solvents efficiently stabilize the QxCav in the *vase* conformation. In particular, benzene is efficiently hosted inside the QxCav cavity since its dimensions are nearly complementary to the cavity size. Bulkier solvents like mesitylene do not fit inside the cavity and consequently the lack of inner solvation effect destabilizes the close *vase* geometry in favor of the *kite* one.

The *kite* form is favored by weak hydrogen-bonding interactions between the solvent and the basic nitrogen atoms of the quinoxaline walls. In the open conformation, the cavity nitrogen atoms are much more exposed to the solvent and thus much more prone to establish such weak interactions with solvent molecules with respect to the closed conformation. Consequently, these hydrogen-bonding interactions shift the equilibrium towards the more extended *kite* geometry.

4.2. WATER SOLUBLE TETRAQUINOXALINE CAVITANDS

QxCav is chosen to be tested as aromatic AAs receptor given its remarkable molecular recognition properties towards aromatic hydrocarbons both at the solid/air[11] and solid/liquid[12] interfaces.
Water solubility of the receptor is required to be used for the molecular recognition of amino acids present in biological fluids. To this purpose, the QxCav is functionalized with pyridinium and imidazolinium feet. These functional groups were already proven to facilitate water solubility of sibling cavitands.[13]
The functionalization of the QxCav lower rim has been carried out in 3 steps, starting from lower feet functionalized resorcinarene **1**. (**Scheme 4.1**)



Scheme 4.1. Synthesis of water soluble cavitands **4** and **5**.

After the insertion of the quinoxaline walls on the resorcinarene scaffold **1** *via* a nucleophilic aromatic substitution in DMF and in the presence of K_2CO_3 as base to obtain **2**, a chlorination reaction of the aliphatic OH groups of **2** with $SOCl_2$ was performed. The common intermediate **3** was obtained as an off white powder after column chromatography. Two different groups were separately introduced at the lower rim of cavitand **3**, namely pyridinium and imidazolinium. Using freshly distilled pyridine or N-methyl imidazole both as solvent and as reagent, target molecules **4**, in the first case, and **5**, in the second case, were obtained *via* nucleophilic substitution with an overall yield of 88% and 90%, respectively.

The synthesized water soluble cavitands 4QxPy **4** and 4QxNMI **5**, together with all the previous intermediates, were analyzed and characterized by ^{1}H NMR spectroscopy and HR-ESI mass spectrometry. For cavitand **4**, the characterizations pointed out that also the cavitand bearing 3 pyridine feet was obtained. Since this last one resulted water soluble as well, we decided to avoid a complex separation.
The effect of solvents on the *vase-kite* conformation of both cavitands **4** and **5** was analyzed *via* ^{1}H NMR in D_2O and deuterated DMSO. The latter was selected because it is one of the most commonly used solvent in biological science.[14] Although it is a protein denaturant at high concentration or at high temperature,[15] DMSO is usually used to crystallize proteins in order to study their mechanism of enzyme function and their conformational stability.[16]
The ^{1}H NMR experiment results for 4QxPy **4** are shown in **Figure 4.4**.

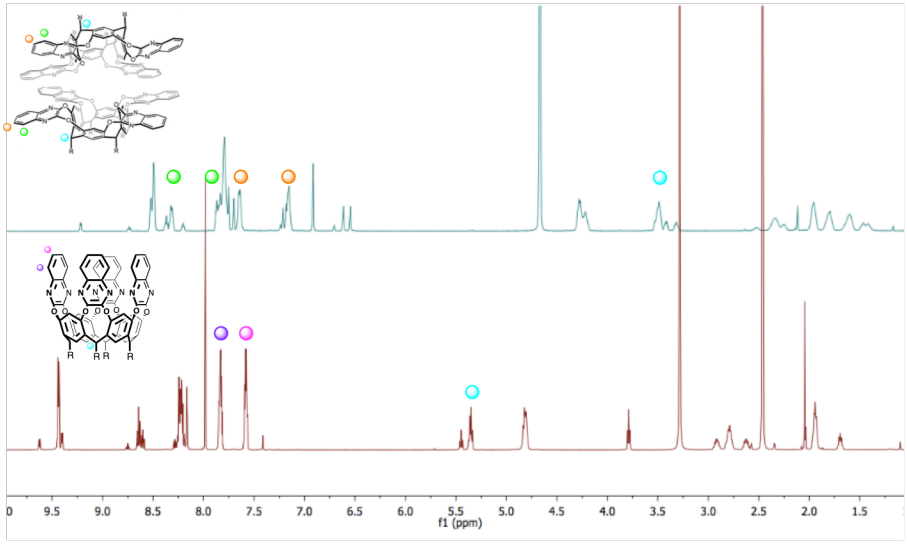


Figure 4.4. Effect of solvents on the vase-kite form of 4QxPy**.** ^{1}H NMR of 4QxPy in D_2O (*top*) and DMSO-d_6 (*bottom*) at 25 °C, 400 MHz.

While in deuterated DMSO (**Figure 4.4** bottom) the cavitand is present in its closed *vase* conformation, D_2O causes the opening of the cavitand into the dimeric *kite* conformation. The interconversion is clearly highlighted by the shift of the diagnostic methine signal (cyan dot) moving from 5.48 ppm in DMSO-d_6 to 3.61 ppm in D_2O.[7c] Moreover, in D_2O the rise of new signals in the aromatic part of the spectrum are diagnostic of the formation of a dimer, named velcrand.[7d] The dimerization buries the hydrophobic walls in the center of the dimer, leaving the hydrophilic feet exposed to the solvent. This drives the cavitands to move from the *vase* to the *kite* conformer in water.

Velcrand formation was further confirmed by HR-ESI mass spectrometry (**Figure 4.5**) as shown by the presence of the molecular ion of the dimeric complex at 1500.42305 Da related to the di-charged species. The other peaks are related to the tetra-functionalized 4QxPy **4** ([M-4Cl/4]$^{4+}$ at 368.39034 Da) and the tri-functionalized 4QxPy ([M-3Cl/3]$^{3+}$ at 476.16184 Da and [M-2Cl/2]$^{2+}$ at 732.72842 Da). Adduct with water molecules are also present as minor intensity peaks.

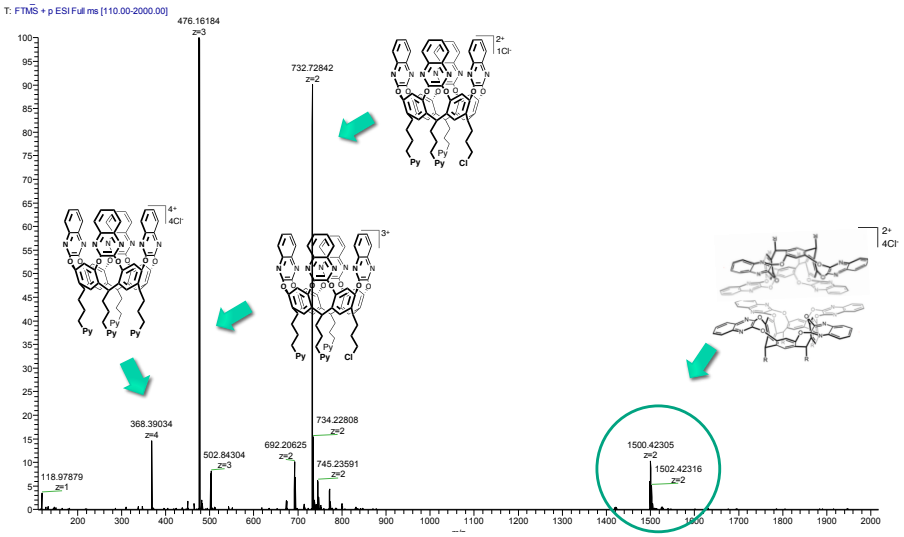


Figure 4.5. HR-ESI mass spectrum of 4QxPy in water.

A study of solvent and pH dependence in the *vase-kite* switching of these molecules was carried out to predict their behavior in biological aqueous solutions at room temperature.

^{1}H NMR titration *vase-to-kite* was performed by dissolving 4QxPy **4** in DMSO-d_6 and adding 50 µL aliquots of deuterated water (in **Figure 4.6** only the most relevant spectra are reported). The addition of D_2O during the analysis destabilized the *vase* form to the point that, above 1:1 ratio between the two solvents, the *kite* conformation became dominant leading to the dimer, as demonstrated by the shift of the methine signal (cyan circle - from 5.48 ppm to 3.61 ppm) to upper field and the rising of the signals assigned to dimer formation both in the aromatic and aliphatic part (highlighted by arrows and circles, respectively). As a result, the *kite* cavitand dimerizes to reduce the exposure of the large hydrophobic surfaces to the polar solvent mixture.

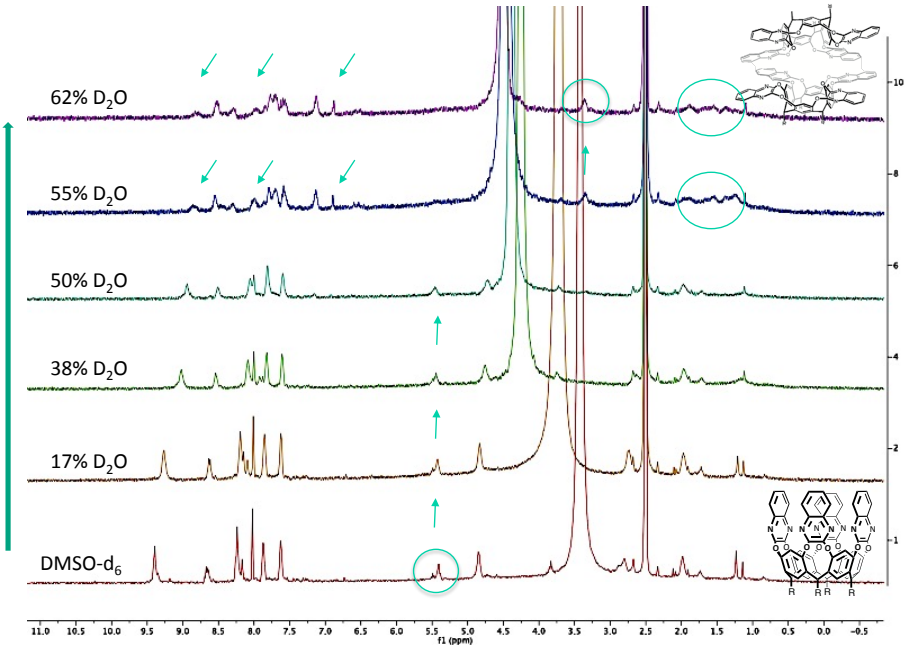


Figure 4.6. ^{1}H NMR titration to study the *vase*-to-*kite* switch of cavitand **4** at 25 °C, 400 MHz.Solvent composition is reported on the left.

In order to avoid protein denaturation, DMSO must be used in a concentration of maximum 10% v/v, unfortunately too low to guarantee the presence of the quinoxaline cavitand in a *vase* conformation.
Hence, further studies were conducted in order to find a more suitable medium that allows the cavitand to maintain its *vase* conformation and a free cavity for the binding event with suitable guests.
Typically used as an organic additive, MeOH can cause protein denaturation at high concentration, just like DMSO, but in some cases protein secondary structures were find to tolerate higher concentrations without compromising the nature of the proteins. For example, the rmethuG-CSF protein maintains its tertiary and secondary structures in a MeOH concentration range from 0 to 20% v/v and denaturation occurs with a MeOH concentration increasing by 30-70% v/v.[17] Myoglobin, instead, retains a secondary structure similar to the native protein up to 35-40% v/v MeOH concentration, showing a slight decrease in the helical secondary structure in solutions containing 50% methanol.[18]

In this case, a UV-Vis titration was performed as convenient technique for monitoring the cavitands conformational state (see Appendix A).[19] Adding methanol to a solution of 4QxPy cavitand in water revealed the receptor only in *vase* form within at least 80% of methanol in the mixture (**Figure 4.7**) discarding the possibility of using this solvent for the recognition of aromatic AAs on proteins surface without causing their denaturation.

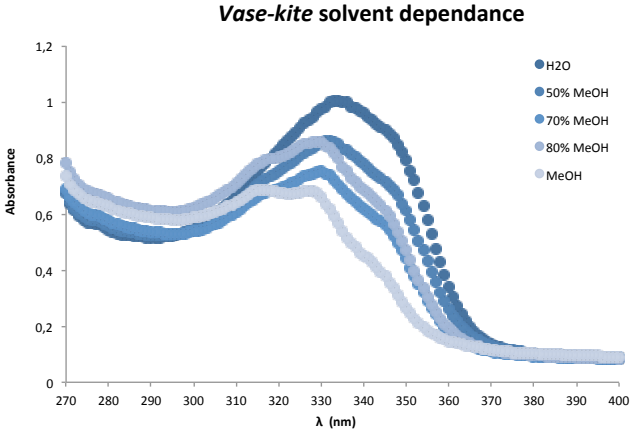


Figure 4.7. 4QxPy **4** water to methanol UV-Vis titration.

For investigation completeness and because both 4QxPy **4** and 4QxNMI **5** have exhibited an acid pH in solution during the tests, a water solution of the 4QxPy receptor was titrated with a base to verify if the change of pH affects the interconversion in water.
The titration (see Appendix A) was carried out taking into account the average isoelectric point of the proteins and the pH value at which the proteins undergo denaturation. This experiment confirmed that in water the kite conformation is the preferred one, even at pH 9.5.
These results, prompted us to exploit a guest-induced *kite-vase* switching as recently demonstrated. Studies conducted by Rebek and co-workers[13a-c,20] have demonstrated that the addition of hydrophobic guests in the presence of water soluble cavitands forces the switch from the dimeric *kite* conformation to the *vase* one improving host-guest aggregation. This latter, favored by the release of water molecules in the bulk, was firstly noted when amphiphilic and hydrophobic guests,

such as normal alkanes, were used giving stoichiometric complexes[13b,20] (**Figure 4.8**) in which long chain hydrocarbons adopt a folded conformation to fill the cavity.

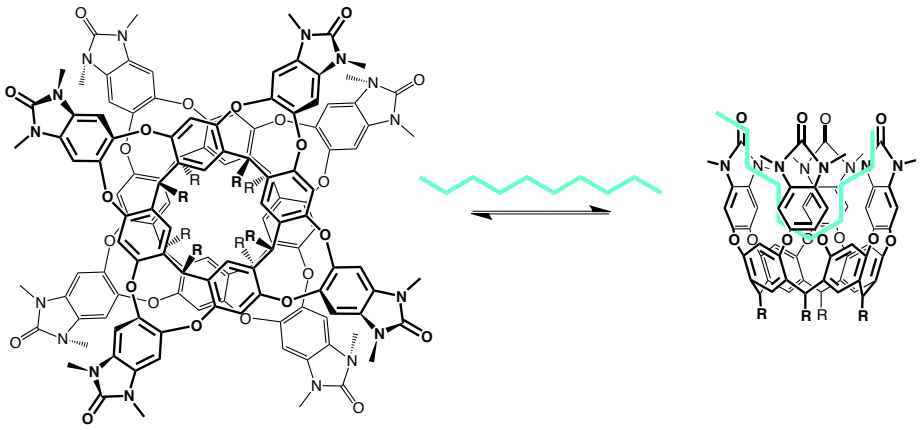


Figure 4.8. Schematic structure of a guest-induced switching.

Taking advantage of these studies, a similar approach was designed for 4QxPy and 4QxNMI cavitands using aromatic AAs as guests. Binding studies *via* NMR and ITC (see Appendix B), and crystallization trials were performed.

Firstly, studies were performed in water solution and the complexation abilities of 4QxPy and 4QxNMI cavitands were tested upon addition of increasing amount of guests in a NMR tube. A guest-induced switching was expected after the addition of 2 eq of guest, necessary to form a 1:1 complex deriving from the dissociation of the velcrand. Unfortunately, no complex formation or *kite-vase* switch was recorded by NMR experiments, probably because of the weakness of interactions involved in the complexation event resulting in the impossibility to investigate the recognition properties of these receptors with NMR technique.

In parallel, ITC titrations were performed using a Nano ITC (TA instrument) placing 1 mL of a 1 mM solution of receptor **4** dissolved in 0.1 M NaOAc pH 4.6 in the cell and titrating it with 300 µL of a 10 mM solution of the guests, namely Phe and Trp, dissolved in the same buffer.

Although evolution of heat was observed after initial injections, no satisfactory sigmoidal curves were provided from the titrations. As for the NMR results, a possible explanation is that the strength of the π-π interactions between the eight quinoxaline flaps of the velcrand resulted to be too strong to allow a velcrand dissociation followed by a *kite-vase* interconversion driven by guest inclusion.

The inconclusive results obtained in solution combined with the search for a deeper insight in the complexation processes of water soluble 4QxCav, led us to orient our research towards solid state investigations.
Co-crystallization trials were carried out at NUI University in Galway - Ireland, testing as guests aromatic AAs in the free form and on the surface of a model protein to verify the binding properties of these receptors. Different crystallization approaches were tested, both in a mixture of DMSO/H_2O (99%, 90% and 80% of DMSO content) and in pure water.
In the first approach, the cavitand was dissolved in DMSO, while the guests, namely L-phenylalanine and L-tryptophan, were dissolved in H_2O. Co-crystallizations trials were initiated with a one- and tenfold excess of the guests optimizing the conditions for the growth of crystals by addition of different buffers and salts exploiting a sparse matrix screen (Jena Bioscience JCSG++) using Oryx 8 robot (Douglas instrument).
(For the methodology description see Appendix C – paragraph 4).
The sitting drop vapor diffusion method was used for the crystallization trials at 20 °C preparing the drops by combining 0.15 µL volumes of 4QxPy or 4QxNMI (10 mM in DMSO), Phe or Trp (10mM or 20 mM in H_2O) and the reservoir solution, in a 96-well plate. Control drops were prepared by substituting the addition of receptors **4** or **5** with 0.15 µL of pure DMSO.

The use of DMSO in the experiment had completely change the processes involved in the classical vapor diffusion crystallization allowing in first instance the growth of salt crystals present in the buffer or of the AAs alone, in first instance, in most of the buffers used. However, after one and a half month, Phe (**Figure 4.9**) and Trp (**Figure 4.10**) co-crystals with 4QxPy started to grow [conditions of the

experiment: 50% PEG 400 + 100 mM NaOAc pH 4.5 + 200 mM Li_2SO_4]. Unfortunately, the shape of the obtained crystals was considered not promising to deserve analyses at the Synchrotone.

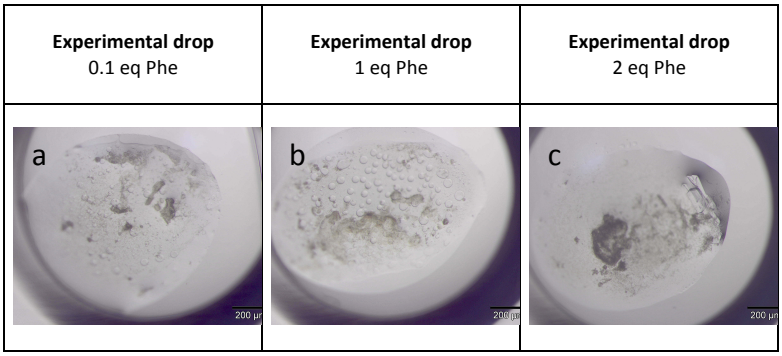


Figure 4.9. Crystals obtained in a preliminary screening. a) 10 mM 4QxPy and 1 mM Phe; b) 10 mM 4QxPy and 10 mM Phe; c) 10 mM 4QxPy and 20 mM Phe; are combined with a reservoir solution containing 50% PEG 400, sodium acetate (100 mM) pH 4.5, lithium sulfate (200 mM).

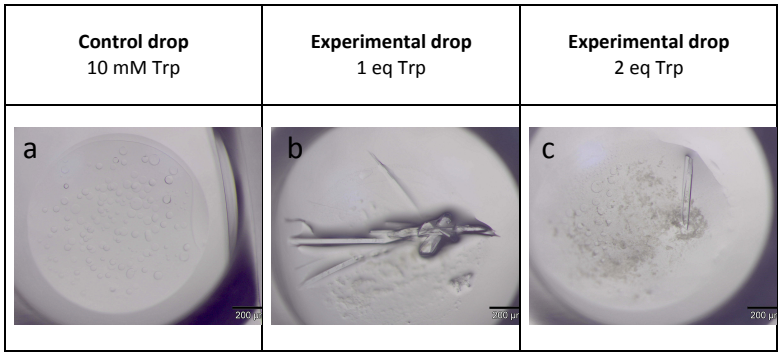


Figure 4.10. Crystals obtained in a preliminary screening. a) Control drop with 10 mM Trp; b) 10 mM 4QxPy and 10 mM Trp; c) 10 mM 4QxPy and 20 mM Trp; are combined with a reservoir solution containing 50% PEG 400, sodium acetate (100 mM) pH 4.5, lithium

To verify if the interaction between the host and the guest is stronger enough to induce the *kite-vase* switching, a second batch of experiments was arranged using pure water as medium.
Once again preliminary experiments were set exploiting the sitting drop technique and exploring different guest concentrations, as well as different reservoir solutions. In this case, both the receptor as AAs were

dissolved in water and employed in the same volumes described before. However, even after several round of optimization, no successful results were obtained.

4.3. AuNP FUNCTIONALIZATION and NOE PUMPING EXPERIMENTS

In collaboration with Professor Fabrizio Mancin, University of Padova, the possibility to functionalize gold nanoparticles with 4QxPy and 4QxNMI receptors was explored. AuNPs decorated with charged groups can be functionalized *via* electrostatic interactions with different molecules. To be functionalized, the AuNPs must be completely dissolved in the solvent, and the only solvent able to dissolve them is water. Since in water our quinoxaline cavitand is present in the *kite* form, we decided to exploit the guest-induced switching of the quinoxaline receptors at the solid-liquid interface. Taking advantage of gold nanoparticles bearing on the surface negative groups able to interact with the pyridinium or imidazolinium moieties present at the lower rim of the cavitand, 4Qx@AuNP systems were developed.
Using a modified Brust-Schiffrin procedure, monolayer-protected AuNPs were synthesized. This procedure allows for obtaining particles of regular size with a nucleus of about 180 Au atoms and a final diameter of approximately 2 nm, to which around 70 thiols are linked.[21]
The protocol foresaw the presence of tetraoctylammonium bromide (TOABr) as phase transfer agent to transfer $HAuCl_4$ from aqueous phase to toluene. The Au(III) was reduced to Au(I) by adding dioctylamine (DOA). DOA allowed for the AuNPs size control by surrounding the NPs surface. Then Au(I) was reduced to Au(0) using sodium borohydride ($NaBH_4$) as reducing agent. The DOA was finally displaced from the NPs surface by adding a methanol solution of 1-undecanethiol functionalized at the terminus with a sulphonate group. The thiol moiety is bound to the gold surface, while the sulphonate groups are exposed to the solvent (AuNP-1, **Figure 4.11**). The sulphonate groups are essential for the binding of 4QxPy or 4QXNMI to the AuNPs surface *via* electrostatic interactions. The obtained negatively charged AuNP-1 were purified by size exclusion chromatography on Sephadex resin and characterized by

^{1}H NMR, TGA and TEM. To enhance the solubility of the functionalized AuNPs, mixed AuNP-2 (**Figure 4.11**) were synthesized as well. In this case, two different thiol chains were used: the sulphonate functionalized one and a polyethylene glycol-derived thiol. The obtained NPs were purified and characterized as before.

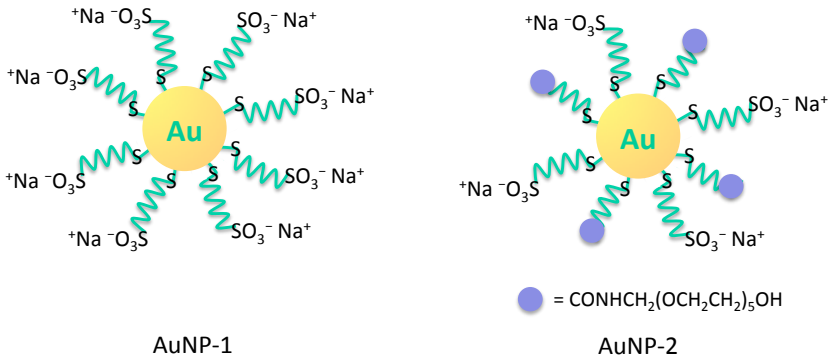


Figure 4.11. Schematic representation of thiol-functionalized AuNP-1 and AuNP-2.

The grafting of cavitands **4** and **5** onto the functionalized AuNP-1 and AuNP-2 were confirmed by fluorescence analysis. The cavitand on its own is fluorescent, while the AuNPs normally act as quencher when interacting with fluorescence species. By subsequently addition (2 µL) of a 0.5 mM 4QxPy cavitand solution in D_2O to a 40 µM solution of AuNP-1 and AuNP-2, fluorescence spectra were recorded (**Figure 4.12**).

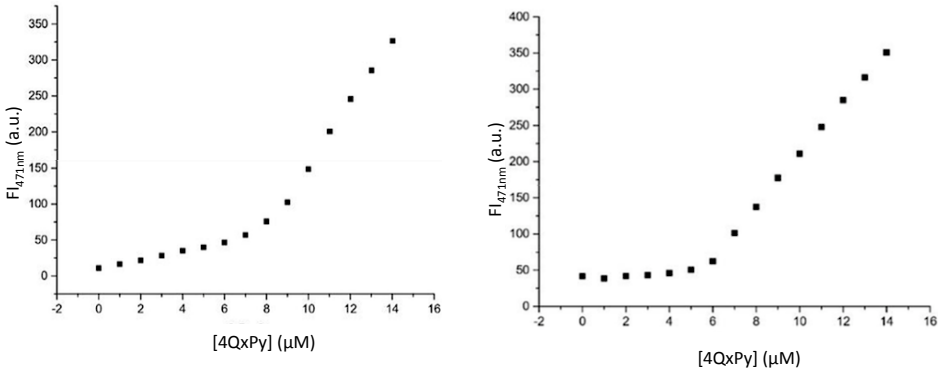


Figure 4.12. 4QxPy-AuNPs titrations adding 1 µM of cavitand to 40 µM AuNP-1 (left) and AuNP-2 (right) solutions in D_2O. Each point is the average of three measures.

The spectra report the fluorescence intensities *versus* the cavitand concentration, measured by exciting the system at 371 nm and monitoring the related emission at 471 nm.
Cavitand fluorescence is initially turned off by the interaction with the AuNPs acting as a quencher. Once all the binding sites on the AuNPs surface are saturated, a significant increase in the fluorescence proportional to the free cavitand additions can be clearly seen. Association constants between 4QxPy and AuNP-1/AuNP-2 were calculated by fitting on a 1:1 binding model, showing very strong interactions. The calculated K_a resulted $2.7 \bullet 10^6$ M^{-1} and $5.6 \bullet 10^7$ M^{-1} for AuNP-1 and AuNP-2, respectively. In both cases the dissociation constants resulted to be in the nanomolar scale. The higher K_a for AuNP-2 could be due to the fact that the sulphonate groups resulted to be more distant from each other allowing a better organization of the charges orientation and thus a better interaction with the cavitand.
In line with expectations, good cooperativity typical of ionic couple interactions was demonstrated. Moreover, these results showed that the concentration of the binding sites calculated by fluorescence measuraments, is about a quarter of the thiols concentration, indicating that essentially all the sulphonate groups present on the AuNPs are able to establish an ionic couple interaction with a positively charged group of 4QxPy. Similar results were obtained with cavitand **5**.

The obtained AuNPs were then tested as receptors for NMR chemosensing. The group. of Prof. Mancin is particularly interested in detecting the family of the catecholamine metabolites. Recently, quantification of a class of catecholamine metabolites in the urine of affected individuals has been demonstrated to be a valid method for the diagnosis and prognosis of neuroblastoma (NBL),[22] a malignant tumor that originates from the cells of the neural crest of the sympathetic nervous system that control involuntary activities such as breathing and heart rate, and which constitutes 10% of cases of pediatric neoplasms.[23]
In particular, a group of eight metabolic derivatives of the amino acid tyrosine was demonstrated to ensure the highest sensitivity to diagnose NBL patients, namely: dopamine (DA), epinephrine (E), norepinephrine (NE), 3-methoxytyramine (3MT), metanephrine (MN), normetanephrine

(NMN), homovanillic acid (HVA) and vanillylmandelic acid (VMA) (**Figure 4.13**).

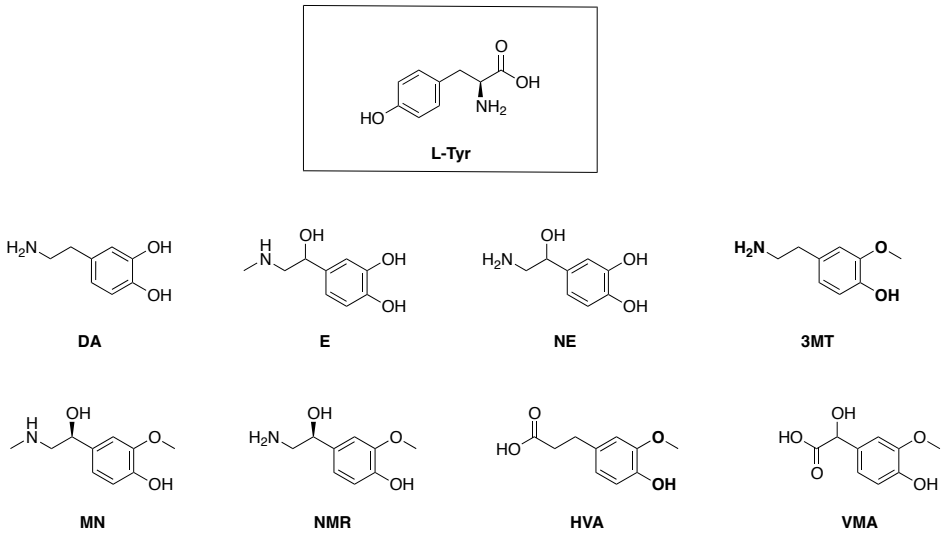


Figure 4.13. Tyrosine metabolic derivatives for neuroblastoma diagnosis.

Preliminary tests were conducted using analytes similar in the structure to the catecholamine metabolites, namely tyramine (TA), phenylpropionic acid (PPA), phloretic acid (PHLOR), phenethylamine (PEA) and some methylated derivatives of the latter, i.e. *N*-methyl-phenethylamine and *N,N*-dimethyl-phenethylamine (**Figure 4.14**). All these analytes are aromatic molecules, so potentially suitable guests for the quinoxaline cavitands. In addition to these analytes, experiments were carried out also with benzonitrile, already known to be efficiently complexed by the quinoxaline receptor.[24]

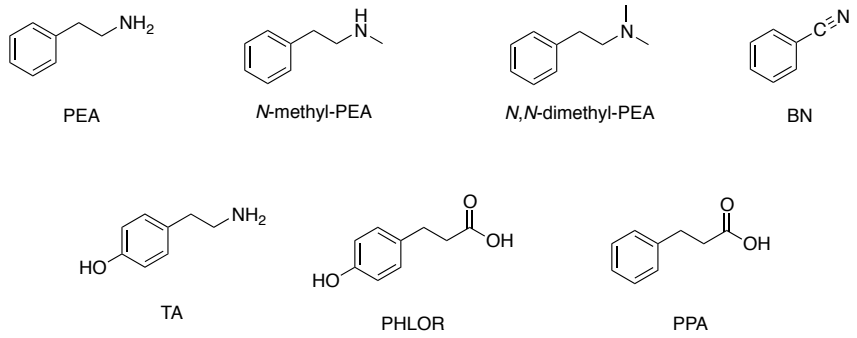


Figure 4.14. Analytes used in the study. Phenethylamine (PEA), *N*-methyl-PEA, *N,N*-dimethyl-PEA, Benzonitrile (BN), Tyramine (TA), phloretic acid (PHLOR) and phenylpropionic acid (PPA).

The tests were performed using the NOE pumping method described in **paragraph 2.4.1**. Introducing into a NMR tube the AuNPs, the cavitand and the analyte at a known concentration in deuterated water, the analyses were carried out. The pH was stabilized using 10 mM HEPES as buffer. In the case of AuNP-1, the addition of the cavitand causes an immediate precipitation. Probably, the cavitand in its *kite* conformation dimerizes acting as a cross-linker between different NPs, causes the precipitation of the AuNPs (**Figure 4.15**).

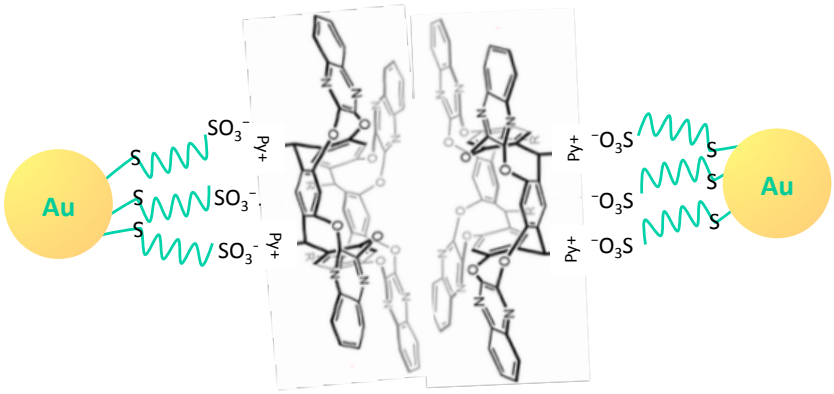


Figure 4.15. *Kite* dimerization driven aggregation and precipitation of AuNP-1.

Differently, the AuNP-2 particles, containing a mixed composition of neutral and negatively charged thiol branches, did not precipitate after

the addition of the cavitand. Thus, this system was used in the following tests, using as guests phenylethylamine (PEA), methylated PEA derivatives *N*-methyl-PEA and *N,N*-dimethyl-PEA, phenylpropanoic acid (PPA), phloretic acid (PHLOR) and benzonitrile (BN).

The first guests studied were phenylethylamine (PEA) and its methylated derivatives. The presence of a strong interaction between the functionalized AuNP-2 and PEA is highlighted by an evident broadening of the PEA aromatic signals (**Figure 4.16** - **b, d** and **f**). However, in the NOE-pumping subspectra (**Figure 4.16** - **c, e** and **g**), also without the addition of the cavitand, a decrease in the signal intensity was observed. This phenomenon can be due to: i) no interaction between the functionalized AuNPs and the analyte since the QxCav is still in the *kite* form and there is no guest-induced switching; ii) the interaction established between the functionalized AuNPs and the analyte is too strong to allow a sufficiently rapid exchange necessary to transfer the magnetization to a significant fraction of the molecules present in the solution.
Using mono and di-methylated phenethylamine similar results were obtained.

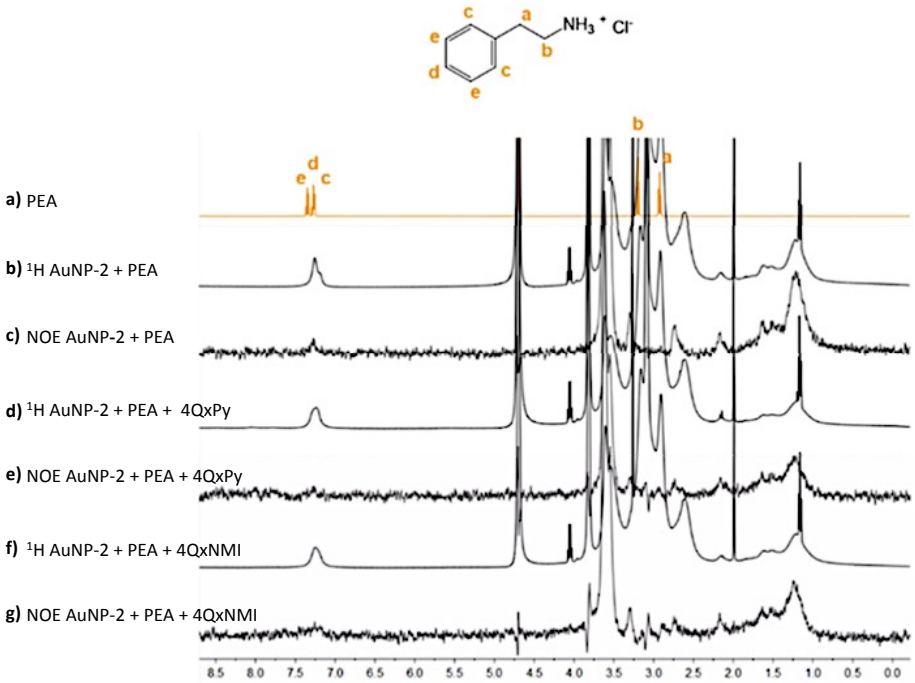


Figure 4.16. ^{1}H NMR spectra comparison. **a**) 1 mM PEA in D_2O; **b**) 2 mM AuNP-2 + 1 mM PEA in 10 mM HEPES/D_2O; **c**) NOE pumping of b; **d**) 2 mM AuNP-2 + 1 mM PEA + 1mM 4QxPy in 10 mM HEPES/D_2O; **e)** NOE pumping of d; **f**) 2 mM AuNP-2 + 1 mM PEA + 1mM 4QxNMI in 10 mM HEPES/D_2O; **g**) NOE pumping of f. NOE spectra intensity is increased by a factor of 10.

The experiments conducted with phenylpropanoic acid (PPA) as a guest confirmed the presence of a slow exchange regime. As a matter of fact, in this case there was a smaller broadening of the PPA signals in the presence of AuNP-2 (**Figure 4.17** - **b**) with respect to PEA, indicating that an interaction between the NPs and the analyte occurred. The lower widening of the analyte peaks caused the signals to be more intense in the NOE spectrum (**Figure 4.17** - **c**). The addition of 4QxPy and 4QxNMI to the PPA-AuNP-2 system led to a further broadening of the analyte signals (**Figure 4.17** - **d** and **f** respectively) suggesting the presence of a stronger interaction with the analyte when cavitands are present in solution. Nevertheless, a decrease in the intensity of the signals in the NOE pumping spectra was observed due to the establishment of a too

slow exchange regime that does not allow the acquisition of sufficient magnetization by the analyte (**Figure 4.17** - **e** and **g**).

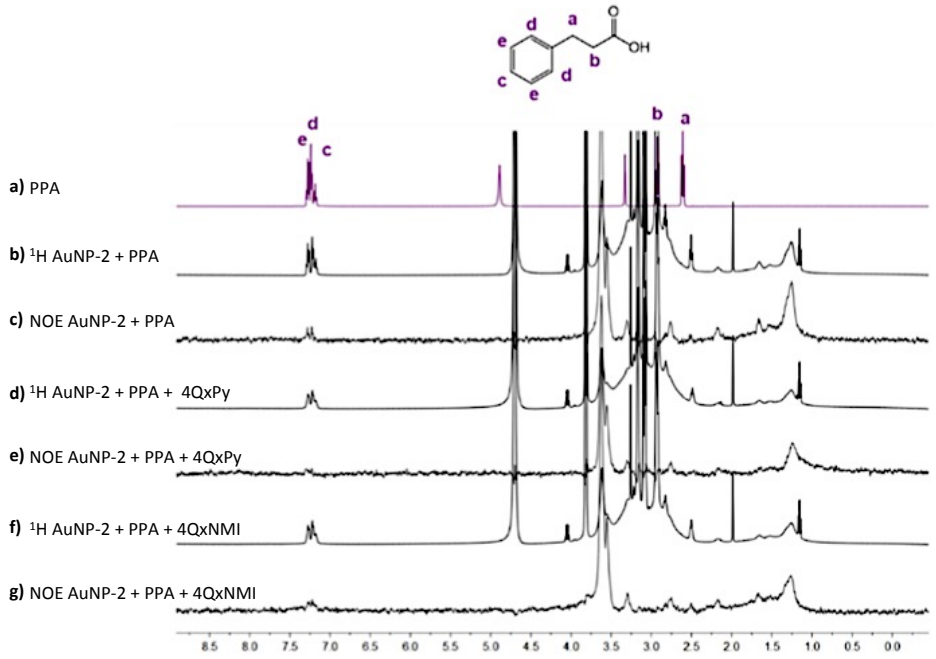


Figure 4.17. ^{1}H NMR spectra comparison. **a**) 1 mM PPA in CD_3OD; **b**) 2 mM AuNP-2 + 1 mM PPA in 10 mM HEPES/D_2O; **c**) NOE pumping of b; **d**) 2 mM AuNP-2 + 1 mM PPA + 1mM 4QxPy in 10 mM HEPES/D_2O; **e)** NOE pumping of d; **f**) 2 mM AuNP-2 + 1 mM PPA + 1mM 4QxNMI in 10 mM HEPES/D_2O; **g**) NOE pumping of f. NOE spectra intensity is increased by a factor of 10.

Finally, interactions with phloretic acid (PHLOR) were studied. The NOE spectrum of the AuNP-2 with PHLOR showed the peaks of the analyte, indicating an interaction between the AuNPs and PHLOR (**Figure 4.18** - **c**). The addition of the 4QxPy to the system showed an enlargement of the analyte signals (**Figure 4.18** - **d**) indicating that the analyte actually interacts with the receptors functionalizing the AuNPs surface and the established interaction is stronger than the one observed with AuNP-2 alone. On the contrary, in presence of 4QxNMI receptor the broadening of the signals was not so evident (**Figure 4.18** - **f**). Also in this case a decrease in the intensity of the signals in the NOE pumping spectra was observed due to the establishment of a too slow exchange regime that

does not allow the acquisition of sufficient magnetization by the analyte.

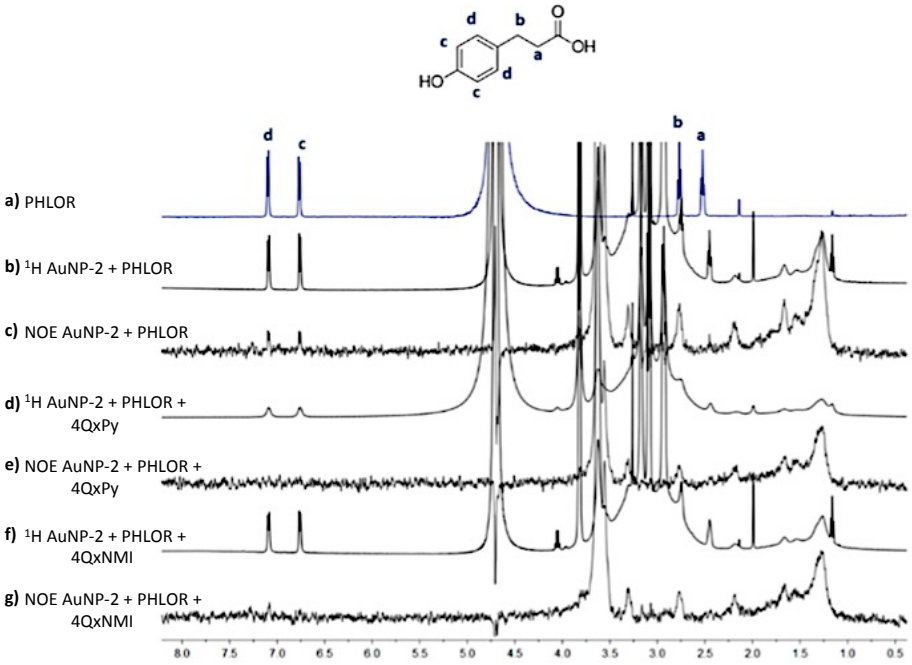


Figure 4.18. ^{1}H NMR spectra comparison. **a**) 1 mM PHLOR in D_2O; **b**) 2 mM AuNP-2 + 1 mM PHLOR in 10 mM HEPES/D_2O; **c**) NOE pumping of b; **d**) 2 mM AuNP-2 + 1 mM PHLOR + 1mM 4QxPy in 10 mM HEPES/D_2O; **e)** NOE pumping of d; **f**) 2 mM AuNP-2 + 1 mM PHLOR + 1mM 4QxNMI in 10 mM HEPES/D_2O; **g**) NOE pumping of f. NOE spectra intensity is increased by a factor of 10.

Although the interactions between quinoxaline cavitands and the guests PEA, PPA and PHLOR have been extensively studied and described verifying their feasibility, no significant widening of the analyte peaks were found in the proton spectra both in the case of AuNP-2 alone and in the case of 4Qx@AuNP-2. Further experiments were conducted also with benzonitrile, known to be complexed particularly well by the class of cavitands under investigation.[24] Even if a significant widening of the signals in the proton spectra indicated that the analyte interacts with the 4Qx@AuNPs nanoconjugates, no detection of the analyte was observed.

The results obtained from the analyses carried out lead to the conclusion that the negatively charged nanoparticle is itself an excellent receptor for phenethylamine derivatives. Two hypotheses can be formulated for the lack of complexation by the grafted cavitands: i) the conformational preference of the cavitands for the dimeric *kite* form suppress their ability to interact with analytes, so no guest-induced *kite-vase* switch; ii) there is the possibility that the analytes can induce the *kite-vase* switching and thus being encapsulated into the cavity. Since the NOE is efficient only for very short distances, the impossibility to detect them *via* NOE pumping can be due to the fact that the encapsulated analyte is relatively far from the protons of the AuNP monolayer coverage. The first hypothesis seems to be the most probable.

4.4. BENZOPYRAZINE CAVITANDS

The lack of results in using QxCav as molecular receptor in water, both alone and supported on AuNPs, prompted us to look at different synthetic supramolecular receptors for the recognition of aromatic amino acids in water. Recently, a new class of water soluble deep cavitands structurally similar to QxCav, namely benzopyrazine cavitands, has demonstrated molecular recognition properties towards aliphatic guests in water.[25]

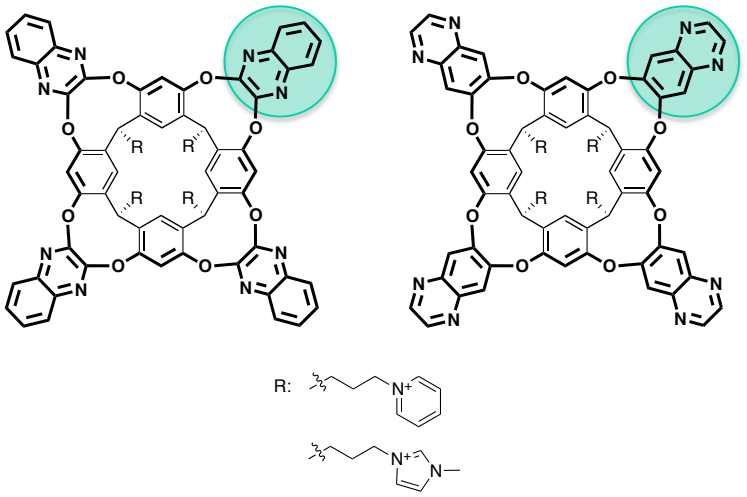


Figure 4.19. Comparison between the chemical structures of water soluble 4Qx cavitands (left) and Benzopyrazinic Cavitands (right).

The main feature that distinguishes benzopyrazine cavitands (BzPyCav) from quinoxaline ones derives from the presence of nitrogen atoms on the outer ring of the walls instead of the inner ring (**Figure 4.19**). Despite this small change, benzopyrazinic cavitands possess the same structural dynamic properties as QxCav. Both cavitands have been demonstrated to adopt different isomeric conformations depending on the solvent: the *vase* conformation is preferred in organic media under standard conditions while in water exclusively *kite* or dimeric *kite* conformations were found to be the stable forms (**Figure 4.20**). Although the formation of velcrands in water is favored, the outer location of the nitrogen atoms in BzPyCav allows a mild repulsion

between the flaps preventing strong interactions between them, as observed for QxCav, guaranteeing the optimal conditions to exploit guest-induced complexations.

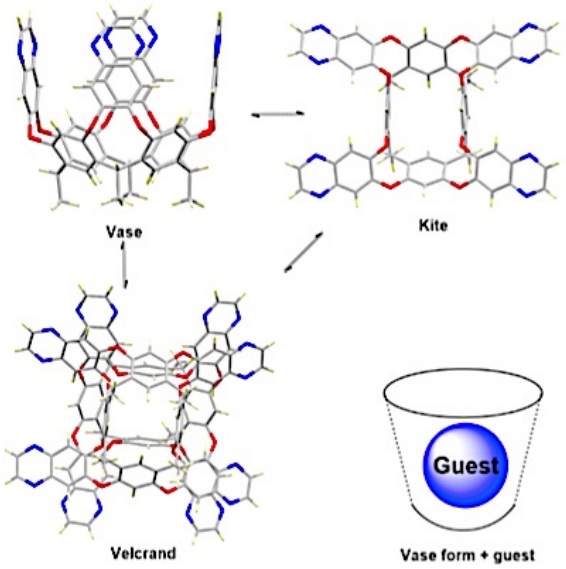


Figure 4.20. Isomeric conformations of BzPyCav (*vase*, *kite* and *velcrand*) and representation of guest-induced binding showing BzPyCav in its *vase* form.

Latterly, the cavitand interconversion between the *vase* and the *kite* form was demonstrated in different deuterated solvents (DMSO-d_6, CD_3OD and D_2O), also upon addition of cyclic hydrophobic guests. In fact, guest-induced switch and consequently guest complexation can be achieved controlling the guest hydrophobicity, over the solvent size and polarity.

Interestingly, the presence of N-donor atoms near the top of the quinoxaline flaps allowed for coordination with metals. Rebek and co-workers recently reported that the benzopyrazine walls can chelate two palladium(II) ions stabilizing the *vase* form and enhancing guest binding.[26] The successful rigidification of the walls with Pd(II) was shown by synthetizing both water soluble and organic soluble cavitands, revealing that structural dynamic conformations of BzPyCav can be modulated by metal coordination stabilizing the cavitand *vase* form regardless of the solvent used. Host-guest complexes (1:1) of the metal-

coordinated BzPyCav with different guests, such as alkanes, alcohols, acids and diols (up to C12), were prepared and studied *via* ^{1}H NMR spectroscopy.

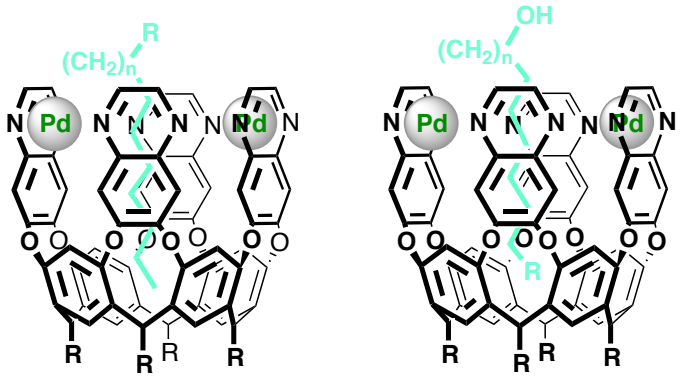


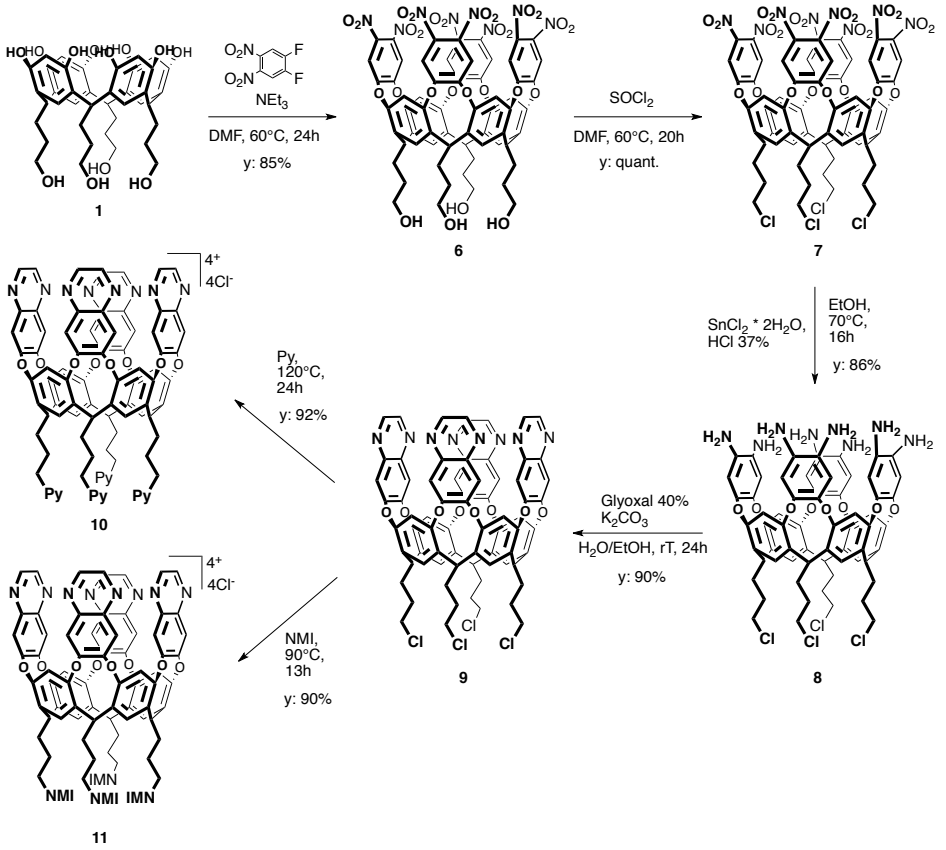
Figure 4.21. Examples of BzPyCav complexes with n-alkanes, n-alcohols and diols. In the case of n-alcohols, the hydrophilic OH group is exposed to water.

These promising results motivated us to investigate BzPyCav as molecular receptor for Phe and Trp in biological media.

Thus, BxPyCav were synthetized according to literature procedures with slight modifications (**Scheme 4.2**). The functionalization of the resorcinarene scaffold **1** with 1,2-difluoro-4,5-dinitro benzene in DMF and in the presence of EtN_3, led to octanitro intermediate **6**. In order to allow the introduction of the hydrophilic residues at the lower rim, a chlorination reaction of OH moieties was performed using $SOCl_2$ as chlorinating agent. Cavitand **7** was subsequently reduced to the octaamino form **8** by $SnCl_2 \bullet H_2O$, HCl 37% and EtOH. The reaction with K_2CO_3 and glyoxal in a H_2O/EtOH mixture for 24 hrs, converted cavitand **8** into cyclized benzopyrazine cavitand **9** in excellent yield without further purifications. Both pyridinium and imidazolinium moieties were inserted at the lower rim to favor the water solubility of the cavitand. Compound **9** was either refluxed in pyridine at 120 °C for 24 hours isolating the benzopyrazinic cavitand **10** or in freshly distilled N-methyl imidazole at 90 °C for 13 hours to obtain BzPyCav **11**.

The synthesized BzPy cavitands, together with all the previous intermediates, were analyzed and characterized by ^{1}H NMR spectroscopy and MALDI-TOF or HR-ESI mass spectrometry. Since

BzPyCav **11** was already reported in the literature,[25,26] all the characterization spectra here reported are related to the novel BzPyCav **10**.



Scheme 4.2. Synthetic pathway for the water soluble benzopyrazinic cavitands.

In **figure 4.22** the ^{1}H NMR spectrum of **10** in DMSO-d_6 is reported. The success of the reaction is highlighted by the presence, in the aromatic part of the spectrum, of the protons associated with the pyridinium moieties (in coral, green and light blue), the benzopyrazine walls (in blue, and pink) and resorcinarene scaffold protons (red and yellow). The diagnostic signal of the methine triplet is present at 5.51 ppm (dark green) indicating the intake of the cavitand **10** in its *vase* conformation, as expected being the solvent DMSO. In the aliphatic region, the signals

related to the alkyl chains are highlighted in orange, violet and aquamarine.

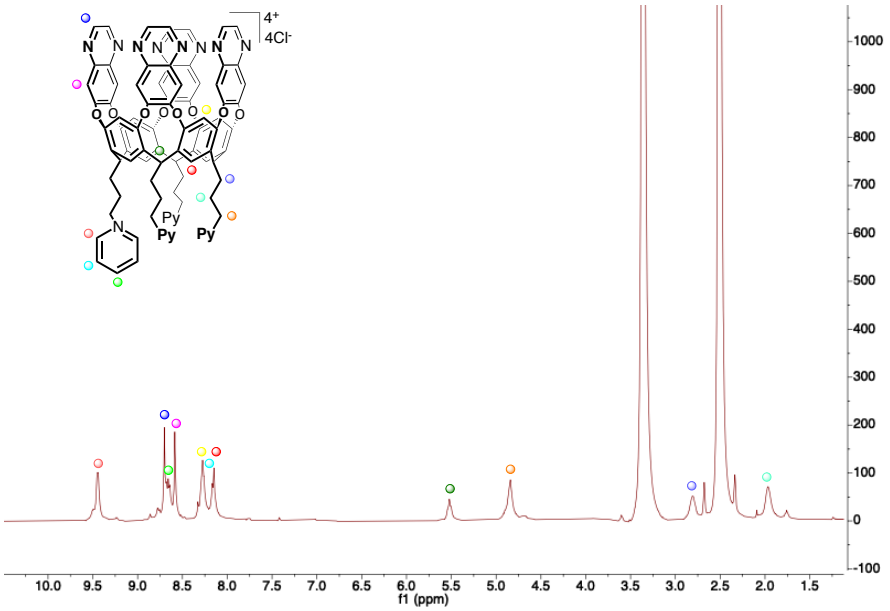


Figure 4.22. ^{1}H NMR of BzPy **10** in DMSO-d_6 at 25 °C, 400 MHz.

As previously described, this class of cavitands, similarly to QxCav, are present as velcrands when dissolved in water. The comparison between ^{1}H NMR spectra of BzPyCav **10** in DMSO-d_6 and D_2O is reported in **figure 4.23**.

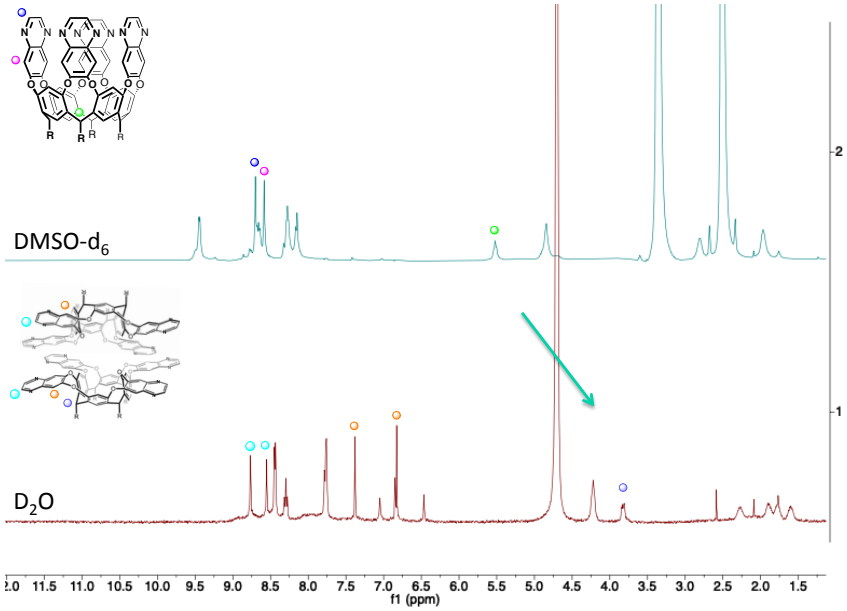


Figure 4.23. ^{1}H NMR spectra of BzPy **10** in DMSO-d_6 (upper) and D_2O (bottom) at 25 °C, 400 MHz.

The *vase* and dimeric *kite* conformations display substantially different ^{1}H NMR spectra. A shift of the diagnostic methine signal from 5.51 ppm (*vase*) to 3.81 ppm (*kite*) was observed changing the solvent from DMSO-d_6 to deuterated water. The shift together with the appearance of new peaks both in the aliphatic part related to the alkyl chains and in the aromatic part of the spectrum related to the benzopyrazine walls, are indicative of the formation of dimeric velcrands. The same behavior was also observed for BzPyCav **11**, bearing imidazolinium moieties at the lower rim.

Taking inspiration from the above mentioned studies on guest-induced binding with hydrophobic guests, the ability of cavitand **10** to switch from a *kite* form to a *vase* one by complexing phenylalanine and tryptophan in water was tested. Preliminary analyses with Phe and Trp were conducted *via* ^{1}H NMR by titrating a 1 mM solution of BzPyCav in D_2O with a 16 mM solution of guest.
After preparing a 1 mM solution of **10** in 0.9 mL of D_2O, 0.6 mL were transferred into an NMR tube. In the remaining 0.3 mL, 0.78 mg of

phenylalanine (F) were added to obtain a 16 mM solution. Four small aliquots of the Phe solution were then added to the NMR tube and the NMR spectrum was recorded after every addition. Each addition of 29 µL corresponded to 0.5 equivalents with respect to receptor **10**. After 2 additions, 1.0 eq of the guest was added and 1:1 complex stoichiometry was reached. As reported in the ^{1}H NMR spectra in **figure 4.24**, no detectable shifts related to inclusion of the guest in the cavity were observed. So, in order to force the guest inclusion, two further additions were added leading to a final ratio of 1:2 (host:guest). Again though, the aromatic signals of Phe, which were expected to be upfield shifted to the negative region upon complexation, remain present in the same position. In addition, the receptor retained its *kite* velcrand conformational arrangement, as suggested by the methine proton constantly at 3.81 ppm together with the presence of doubled peaks in the aromatic part and the velcrand alkyl chains resonance in the upfield region, making the cavity unable for the binding and decreasing the chances of capturing the guest.

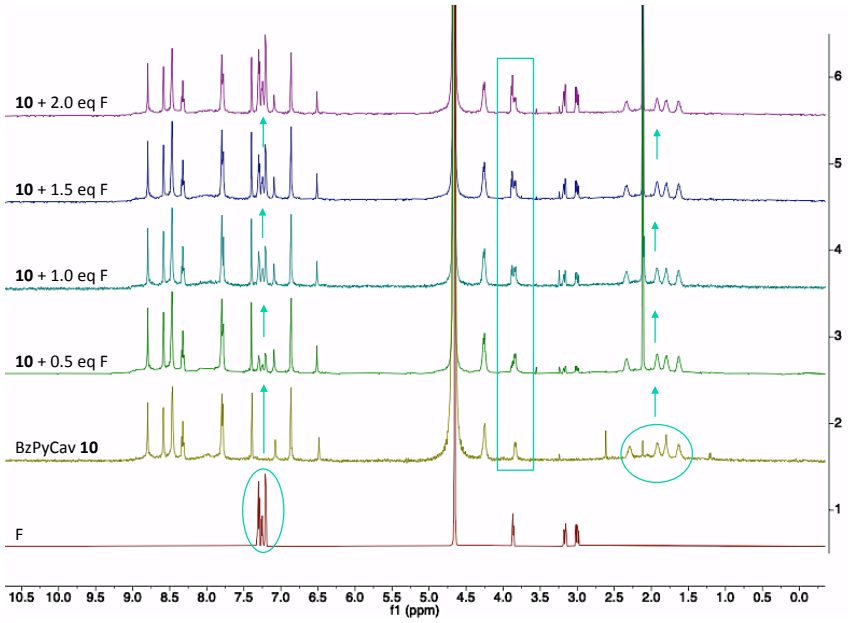


Figure 4.24. ^{1}H NMR titration of BzPyCav **10** with phenylalanine in D_2O, 25 °C, 600 MHz.

To verify the capability of the receptor to switch to its *vase* form upon complexation of slightly more hydrophobic guests, a titration with Trp (W) was performed. As before, a 1 mM solution of BzPyCav **10** was prepared and partially transferred in an NMR tube. In the remaining part, 1 mg of the guest was dissolved to obtain a 16 mM solution. Four aliquots of 28 µL, each of 0.5 equivalents with respect to receptor **10**, were added to the NMR tube recording the spectra after every addition reaching a final complex ratio 1:2 (host:guest). As seen previously in the titration with Phe, also in this case, the cavitand is present as velcrand in water as evinced by the presence of the peaks related to the velcrand protons both in the aliphatic and aromatic regions of the spectra (**figure 4.25**). Once again, the diagnostic methine protons do not undergo any shift.

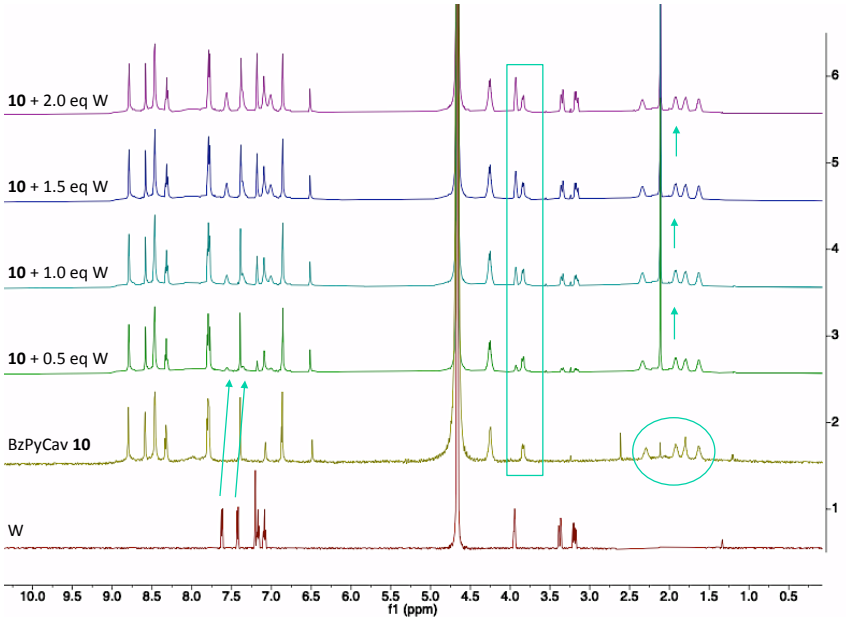
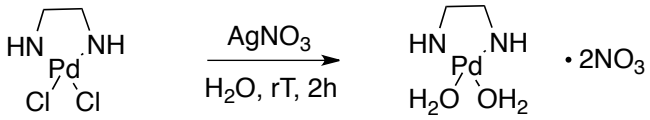


Figure 4.25. ^{1}H NMR titration of BzPyCav **10** with tryptophan in D_2O, 25 °C, 600 MHz.

Differently from the previous titration with Phe, the additions of tryptophan to receptor **10** cause small shifts of the aromatic protons of the amino acid to upper fields once reached 1:1 complex formation. This weak host-guest interaction is consistent with the greater hydrophobic character of Trp compared to Phe. However, the small

entity of the recorded shifts in the protonic spectrum cannot be considered as an indication of Trp recognition by cavitand **10**.

The possibility of forcing the closure of the cavitand walls upon chelation of benzopyrazine nitrogen atoms with a metal was also explored. Freshly synthetized palladium complex $Pd(en)(H_2O)_2 \cdot 2NO_3$,[27] was then added in different proportions to induce the closing of the aromatic walls.



Scheme 4.3. Synthesis of $Pd(en)(H_2O)_2 \cdot 2NO_3$ complex.

An ^1H NMR titration was performed adding $Pd(en)(NO_3)_2$ ranging from 1 eq to 5 eq to an NMR tube containing 0.6 mL of a 1 mM solution of BzPyCav **10** or BzPyCav **11**. After 1 hour of sonication, the spectra were recorded. Being chelated by two nitrogen atoms on different walls, the minimum amount of Pd(II) complex required is 2 eq, even if a more defined *kite-vase* interconversion is observed when 3 eq were added.
As demonstrated also by Rebek,[26] in D_2O the presence of the Pd(II) complex leads to a dynamic behavior between the free velcrand and the metal coordinated *vase* conformation, but it does not lead to a complete closure of the cavity and to a stable *vase* conformation in the absence of a guest. This could be due to the high instability of the empty *vase* conformer in water as well as the hydrophobicity of the cavity in which water hardly enters. Moreover, the possible coordination and subsequent dissociation of the Pd(II) complex with the benzopyrazinic nitrogen can lead to a dynamic equilibrium between the velcrand form and the *vase* conformation.

For the above reasons, ^1H NMR titrations in presence of Phe and Trp together with $Pd(en)(NO_3)_2$ complex were performed with both BzPyCav **10** and **11**. In this case, a partial closure of the BzPyCav walls by Pd(II) occurs, but a complete closure caused by the complexation of the guest in the cavity was not observed.

In **Figure 4.26** the ^{1}H NMR spectra of BzPyCav **10** and **11** with Phe (F) were reported. Evidence consistent with the coordination of Pd(en)$(NO_3)_2$ complex to the benzopyrazinc nitrogen atoms are represented by the doublet at 3.03 ppm and further confirmed by the decreased intensities of the resonances related to the aromatic and aliphatic protons of the velcrand. The singlet at 2.60 ppm accounts for the free Pd(en)$(NO_3)_2$ complex, which increases with the addition of Pd(II) complex equivalents. A complete conversion from velcrand to *vase* and the complexation of the guest Phe, was not observed for either BzPyCav **10** or **11**.

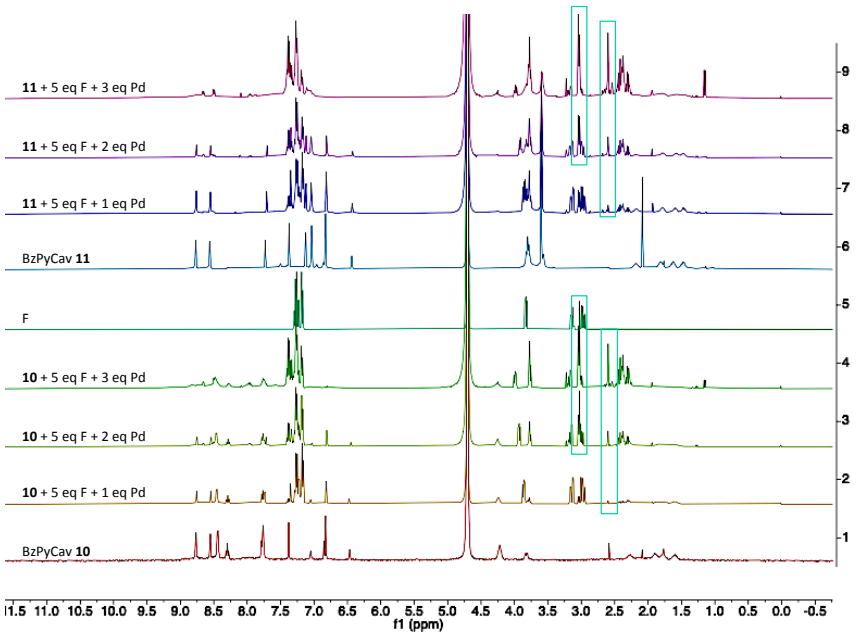


Figure 4.26. Velcrand-*vase* titration of BzPyCav **10** and **11** with Pd(en)$(NO_3)_2$ complex and F. Increasing signals of free and complexed Pd(en)$(NO_3)_2$ are highlighted by cyan rectangular. D_2O, 25 °C, 600 MHz.

Similarly, in the ^{1}H NMR titration of BzPyCav **10** and **11** with Trp (W), no complexation event was detected (**Figure 4.27**). Again, the presence of the coordinated Pd(en)$(NO_3)_2$ complex is proved by the signal at 3.22 ppm, while the peak of the free Pd(II) complex is seen to increase at 2.60 ppm. Even if the resonances of the aromatic and aliphatic protons

of the velcrand appeared to be not more present in the spectra upon the addition of 3 equivalents of Pd(II), no shifts of the diagnostic methine triplet were detected. Thus, a complete switch to *vase* conformation was not achieved.

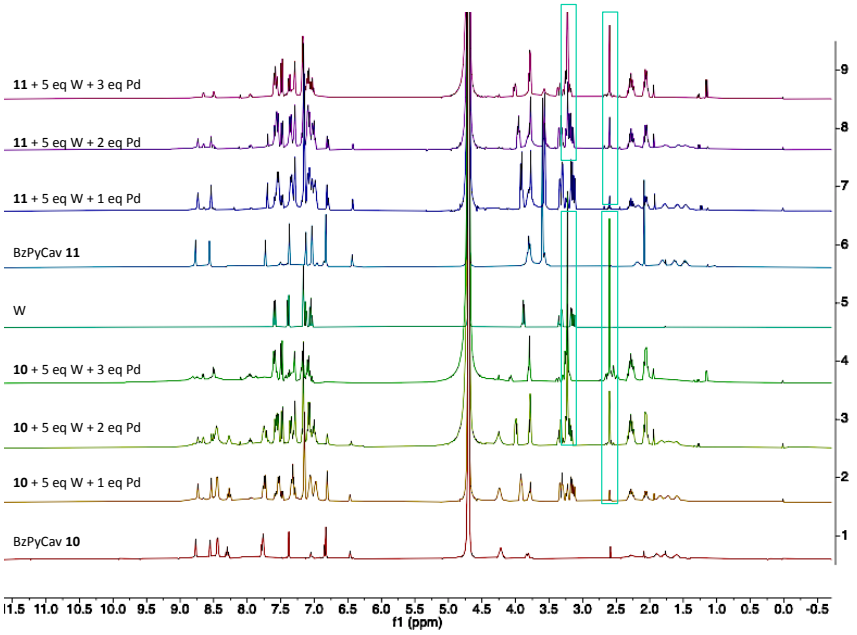


Figure 4.27. Velcrand-*vase* titration of BzPyCav **10** and **11** with Pd(en)$(NO_3)_2$ complex and W. Increasing signals of free and complexed Pd(en)$(NO_3)_2$ are highlighted by cyan rectangular. D_2O, 25 °C, 600 MHz.

Even in presence of a chelating Pd(II) complex, the poor affinity of the studied guests with the host cavity led to an incomplete velcrand-*vase* guest-induced switching excluding the possibility of using these receptors for the selective recognition of aromatic amino acids in water.

4.5. SELF-FOLDING WATER SOLUBLE CAVITAND

Another water soluble deep-cavity cavitand was investigated as potential molecular receptor for aromatic amino acid, namely tetracarboxylate cavitand **16**, shown in **Figure 4.28** and already reported by Hooley and Rebek (see **Chapter 3**).[28] The receptor was synthetized following the literature procedures and characterized as usual (see Appendix C).

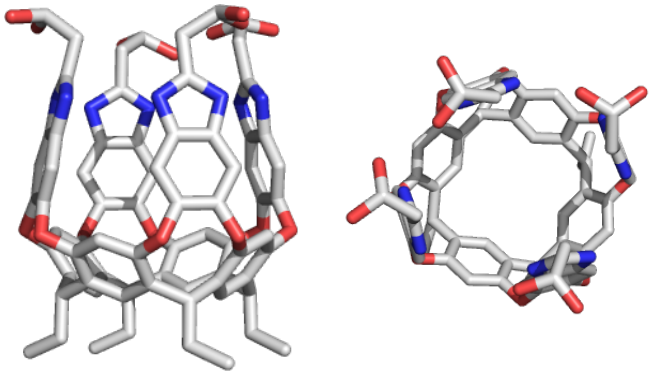


Figure 4.28. Side view (left) and front view (right) of the self-folding deep cavitand **16**.

Binding studies *via* ^{1}H NMR titrations with Phe and Trp were carried out together with ITC measurements. However, no successful data were obtained (see Appendix C).
In parallel, solid state host-guest complexations were performed at NUI Galway – Ireland, in the laboratory of Professor Crowley. Unfortunately, crystals obtained from the trials were judged not sufficient promising for further investigation by X-ray diffraction analysis (see Appendix C).

4.6 EXPERIMENTAL SECTION

Recorcinarene [C_3H_6OH; H] (1).
To a solution of resorcinol (15 g, 13.62 mmol) in 90 mL of MeOH, 2,3-dihydrofuran (10.3 mL, 13.62 mmol) and 37% HCl (22.5 mL) were added dropwise over 30 min at 0 °C. After additions, the reaction mixture was stirred at 50 °C for a week. The solvent was concentrated to 50 mL and the obtained precipitate was filtered. Water was added to the solution and the new precipitate was filtered, washed with water and dried under vacuum. Recrystallization of the precipitates from MeOH gave product **1** as a pale yellow powder (22 g, 56%).
^{1}H NMR (DMSO-d_6, 300 MHz): δ ppm = 8.88 (s, 8H, O**H**), 7.21 (s, 4H, Ar**H$_{up}$**), 6.13 (s, 4H, Ar**H$_{down}$**), 4.18 (t, 4H, -C**H**-, J= 7.92 Hz), 3.40 (m, 8H, -C**H$_2$**OH), 2.07 (m, 8H, CHC**H$_2$**-), 1.31 (m, 8H, -C**H$_2$**CH_2OH). **ESI-MS**: m/z 743.39 $[M+Na]^+$; 759.38 $[M+K]^+$.

Tetraquinoxaline Cavitand (2).
Resorcinarene **1** (500 mg, 0.69 mmol) was dissolved in dry DMF (10 mL). K_2CO_3 (767 mg, 5.55 mmol) and 2,3-dichloroquinoxaline (580 mg, 2.91 mmol) were added and the solution was stirred at 80 °C for 48 hrs. The solvent was removed and the crude was washed with H_2O and filtered. The final product **2** was obtained as a white powder (850 mg, quant.).
^{1}H NMR (DMSO-d_6, 300 MHz): δ ppm = 7.93 (s, 4H, Ar**H$_{up}$**), 7.86 (m, 8H, Ar**H$_{qx}$**), 7.82 (s, 4H, Ar**H$_{down}$**), 7.59 (m, 8H, Ar**H$_{qx}$**), 5.49 (t, 4H, -C**H**-, J = 8.04 Hz), 4.51 (t, 4H, -O**H**, J = 6.12 Hz), 3.54 (m, 8H, -C**H$_2$**OH); 3.32 (m, 8H, CHC**H$_2$**-), 1.47 (m, 8H, -C**H$_2$**CH_2OH). **ESI-MS**: m/z 1225.73 $[M+H]^+$.

4QxCl cavitand (3).
To a solution of **2** (200 mg, 0.16 mmol) in DMF (8 mL), thionyl chloride (284 µL, 3.92 mmol) was added. After the addition, the reaction mixture was stirred at 55 °C for 16 hrs. The reaction was quenched with water and filtered. The crude was purified by silica gel column chromatography (DCM/MeOH 98:2) to obtain **3** as a white powder (189 mg, 89%). **^{1}H NMR** ($CDCl_3$, 300 MHz): δ ppm = 8.22 (s, 4H, Ar**H$_{up}$**), 8.07 (s, 4H, Ar**H$_{down}$**), 7.81 (m, 8H, Ar**H$_{qx}$**), 7.49 (m, 8H, Ar**H$_{qx}$**), 5.69 (t, 4H, -C**H**-, J = 7.83 Hz), 3.75 (m, 8H, -C**H$_2$**Cl), 2.51 (m, 8H, CHC**H$_2$**-), 1.94 (m, 8H, -C**H$_2$**CH_2Cl). **MALDI-MS**: m/z 1297.2740 $[M+H]^+$.

4QxPy cavitand (4).

4QxCl cavitand **3** (185 mg, 0.14 mmol) was dissolved in 10 mL of dry pyridine, and the solution was heated at 120 °C for 24 hrs under an Ar atmosphere. Subsequently, the mixture was cooled to rT and the resulting precipitate was filtered and washed with acetone (3×30 mL). The obtained solid was dissolved in acetone (30 mL) and heated at 60 °C for 4 hrs. The precipitate was then collected by filtration and dried under vacuum to give **4** as a light beige solid (170 mg, 74 %).

^{1}H NMR (DMSO-d_6, 400 MHz): δ ppm = 9.57 (d, 8H, Py**H$_o$**, J = 3.0 Hz), 8.75 (m, 4H, Py**H$_p$**), 8.38 (s, 4H, Ar**H$_{down}$**), 8.32 (m, 8H, Py**H$_m$**), 8.12 (s, 4H, Ar**H$_{up}$**), 7.97 (m, 8H, Ar**Qx**), 7.72 (m, 8H, Ar**Qx**), 5.48 (t, 4H, -C**H**-, J = 6.3 Hz), 4.96 (t, 8H, -C**H$_2$**Py, J = 6.0 Hz), 2.80 (m, 8H, CHC**H$_2$**-), 2.17 (m, 8H, -C**H$_2$**CH$_2$Py). **^{13}C NMR** (DMSO-d_6, 600 MHz): δ ppm = 152.30, 146.17, 145.65, 139.46, 136.13, 130.17, 129.16, 128.68, 128.14, 126.91, 118.59, 61.82, 45.60, 34.77, 29.17. **HR-ESI**: m/z calculated for $C_{87}H_{67}Cl_2N_{11}O_8^{2+}$: 732.7256 $[M\text{-}2Cl]^{2+}$, found 732.7284; m/z calculated for $C_{92}H_{72}ClN_{12}O_8^{3+}$: 502.8434 $[M\text{-}3Cl]^{3+}$, found 502.8430; m/z calculated for $C_{87}H_{67}ClN_{11}O_8^{3+}$: 476.1616 $[M\text{-}3Cl]^{3+}$, found 476.1618; m/z calculated for $C_{92}H_{72}N_{12}O_8^{4+}$: 368.3902 $[M\text{-}4Cl]^{4+}$, found 368.3903.

4QxNMI cavitand (5).

10 mL of freshly distilled 1-methylimidazole were added by syringe to 4QxCl cavitand **3** (100 mg, 0.08 mmol) under an inert atmosphere. The mixture was stirred vigorously at rT for 30 min to ensure complete dissolution and then immersed in an oil bath preheated to 90 °C. After 13 hrs, the mixture was cooled to rT and stirred for 10 min at 0 °C. The solution was draw up through in a syringe equipped with a syringe filter-disc (PTFE, 0.2 m, 25 mm) and slowly added to 60 mL of 0 °C Et_2O, cooled to 0°C, under vigorous stirring and a N_2 flow. After 1 h, the solid was allowed to settle at 0 °C. The supernatant was removed and the precipitate was resuspended in 50 mL of Et_2O, stirred briefly and allowed to settle for 30 hrs. The solid was collected by centrifugation (speed 4000 *rpm*, 10 min, rT) after washing the residue with 40 mL of Et_2O (x 3 times). The collected solid was suspended in a mixture of H_2O/ACN (2:1), heated to 75 °C and then ACN was removed by evaporation. The target compound **4** was obtained, after lyophilisation in vacuum, as a pale beige powder (106 mg, 82%).

^{1}H NMR (DMSO-d_6, 400 MHz): δ ppm = 9.46 (s, 4H, -NC**H**N-), 8.26 (s, 4H, NCH_3C**H**-), 8.06 (s, 4H, -C**H**N^+), 8.04 (s, 4H, Ar**H_{down}**), 7.78 (m, 8H, Ar**Qx**), 7.64 (s, 4H, Ar**H_{up}**), 7.62 (m, 8H, Ar**Qx**), 5.43 (t, 4H, -C**H**-, J= 6.7 Hz), 4.35 (m, 8H, -C**H_2**NMI), 3.91 (s, 12H, ArN-C**H_3**), 2.81 (m, 8H, CHC**H_2**-), 1.82 (m, 8H, -C**H_2**CH_2NMI). **^{13}C NMR** (DMSO-d_6, 600 MHz): δ ppm = 152.26, 148.38, 139.36, 138.48, 137.14, 136.12, 128.07, 124.81, 123.97, 122.92, 122.39, 49.53, 36.65, 36.27, 28.79. **HR-ESI**: m/z calculated for $C_{88}H_{76}Cl_2N_{16}O_8^{2+}$: 778.2685 $[M-2Cl]^{2+}$, found 778.2705; m/z calculated for $C_{88}H_{76}ClN_{16}O_8^{2+}$: 506.5235 $[M-3Cl]^{3+}$, found 506.5236; m/z calculated for $C_{88}H_{76}N_{16}O_8^{2+}$: 371.1503 $[M-4Cl]^{4+}$, found 371.1506.

Tetrahydroxypropyl-octanitrocavitand (6).
Resorcinarene **1** (100 mg, 0.14 mmol) and 1,2-difluoro-4,5-dinitrobenzene (113 mg, 0.55 mmol) were suspended in 3 mL of anhydrous DMF and 310 µL of triethylamine were added dropwise. The mixture was heated to 65 °C for 16 hrs. The solution was cooled down to rT and methanol was added to inducing precipitation of a yellow solid. The solid was sonicated for 5 min and mixed vigorously at 65 °C for 1 h. After, the precipitate was allowed to settle at rT for 30 min. The insoluble portion was collected by filtration to afford octanitrocavitand **6** as a yellow solid (162 mg, 85%).
^{1}H NMR (DMSO-d_6, 400 MHz): δ ppm = 8.80 (s, 8H, Ar**H**), 8.22 (s, 4H, Ar**H_{up}**), 7.87 (s, 4H, Ar**H_{down}**), 5.52 (bt, 4H, -C**H**-), 4.49 (bt, 4H, O**H**), 3.49 (m, 8H, -C**H_2**OH), 2.41 (m, 8H, CHC**H_2**-), 1.40 (m, 8H, -C**H_2**CH_2OH).

Tetrachloropropyl-octanitrocavitand (7).
Cavitand **6** (124 mg, 0.09 mmol) was suspended in 3 mL of anhydrous DMF and 157 µL of thionyl chloride were added dropwise. The mixture was heated to 60 °C for 20 hrs. The solution was cooled to rT and the solvent was removed in vacuum. After, methanol was added to induce precipitation of a yellow solid. The solid was mixed vigorously for 1h at rT. The insoluble portion was collected by filtration to afford tetrachloropropyl-octanitrocavitand **7** as a yellow solid in quantitative yield (131 mg, quant.).
^{1}H NMR (DMSO-d_6, 400 MHz): δ ppm = 8.82 (s, 8H, Ar**H**), 8.26 (s, 4H, Ar**H_{up}**), 7.83 (s, 4H, Ar**H_{down}**), 5.58 (bt, 4H, -C**H**-), 3.70 (bt, 8H, -C**H_2**Cl), 2.48 (m, 8H, CHC**H_2**-), 1.72 (m, 8H, -C**H_2**CH_2Cl).

Tetrachloropropyl-octaaminocavitand (8).
13 mL of concentrated 37% HCl and $SnCl_2 \cdot 2H_2O$ (1.52 g, 6.76 mmol) were added to a suspension of octanitro cavitand **7** (131 mg, 0.09 mmol) in 10 mL of ethanol. The mixture was stirred for 16 hrs at 70 °C. After cooling to rT, the resulting precipitate was collected by centrifugation and washed with a 1:1 H_2O:EtOH mixture to give a white solid. The solid was suspended in 50 mL of ethyl acetate and 40 mL of NH_3 (28 % in H_2O) and the mixture was stirred for 30 min at room temperature. The organic layer was separated, dried over Na_2SO_4 and concentrated to afford the octaamino cavitand **8** (98 mg, 90%).
The compound showed gradual decomposition in contact to air and therefore was immediately used in the next step.

Tetrachloropropyl-benzopyrazinecavitand (9).
Tetrachloropropyl-octaaminocavitand **8** (230 mg, 0.19 mmol) was added to a solution of K_2CO_3 (210 mg, 1.52 mmol) dissolved in 40 mL of a H_2O:EtOH mixture 3:7 under vigorous stirring for 10 min at rT. Glyoxal 40% (184 µL, 1.52 mmol) was added in one portion and the reaction mixture was stirred at rT for 48 hrs. The brownish solid obtained was filtered and washed thoroughly with water (180 mg, 90%).
^{1}H NMR ($CDCl_3$, 600 MHz): δ ppm = 8.74 (s, 8H, **BzPy**), 7.81 (s, 8H, **BzPy**), 7.18 (s, 4H, Ar**H$_{up}$**), 6.79 (s, 4H, Ar**H$_{down}$**), 4.22 (bt, 4H, -C**H**-), 3.53 (t, 8H, -C**H$_2$**Cl, J = 6.0 Hz), 2.20 (m, 8H, CHC**H$_2$**-), 1.73 (m, 8H, -C**H$_2$**CH_2Cl). **MALDI-MS**: m/z 1297.3315 $[M+H]^+$.

Tetrapyridiniumpropyl-BzPy cavitand (10).
A solution of cavitand **9** (60 mg, 0.05 mmol) in 6 mL of dry pyridine was heated at 120 °C for 24 hrs under an Ar atmosphere. Subsequently, the mixture was cooled to room temperature and the resulting precipitate was filtered and washed with acetone (3×10 mL). The solid was then suspended in acetone (15 mL) and heated at 60 °C for 4 hrs. The obtained precipitate was collected by filtration and dried under vacuum to give **10** as a light beige solid (69 mg, 92 %).
^{1}H NMR (DMSO-d_6, 400 MHz): δ ppm = 9.45 (d, 8H, Py**H$_o$**, J = 2.0 Hz), 8.70 (s, 8H, **BzPy**), 8.65 (m, 4H, Py**H$_p$**), 8.58 (s, 8H, **BzPy**), 8.28 (s, 4H, Ar**H$_{up}$**), 8.28 (m, 8H, Py**H$_m$**), 8.14 (s, 4H, Ar**H$_{down}$**), 5.52 (bt, 4H, -C**H**-), 4.84 (bt, 8H, -C**H$_2$**Py), 2.81 (m, 8H, CHC**H$_2$**-), 1.97 (m, 8H, -C**H$_2$**CH_2Py).

^{13}C NMR (DMSO-d_6, 600 MHz): δ ppm = 154.85, 154.09, 146.09, 145.73, 145.48, 141.18, 135.71, 128.67, 126.26, 123.58, 116.97, 61.76, 34.27, 29.37, 28.59. **HR-ESI**: m/z calculated for $C_{92}H_{72}Cl_2N_{12}O_8^{2+}$: 772.2467 [M-2Cl]$^{2+}$, found 772.2501; m/z calculated for $C_{92}H_{72}ClN_{12}O_8^{3+}$: 502.8434 [M-3Cl]$^{3+}$, found 502.8442; m/z calculated for $C_{87}H_{67}ClN_{11}O_8^{3+}$: 476.4960 [M-3Cl]$^{3+}$, found 476.4968; m/z calculated for $C_{92}H_{72}N_{12}O_8^{4+}$: 368.3902 [M-4Cl]$^{4+}$, found 368.3910.

Tetraimidazoliniumpropyl-BzPy cavitand (11).
In a flame-dried flask, cavitand **9** (60 mg, 0.05 mmol) was suspended under an Ar atmosphere in 6 mL of freshly distilled 1-methylimidazole. The mixture was stirred vigorously at rT for 1 h to achieve a homogeneous fine suspension and then at 90 °C for 16 hrs. The mixture was then cooled to room temperature and the solid was allowed to settle at 0 °C after the addition of Et_2O. The solid was collected by centrifugation and suspended in acetone, triturated and heated under reflux while stirring vigorously for 4 hrs. The product was collected by vacuum filtration and washed with acetone. After drying under high vacuum, the product was obtained as a pale yellow powder (68 mg, 90%).
^{1}H NMR (DMSO-d_6, 400 MHz): δ ppm = 9.48 (s, 4H, -NC**H**N-), 8.70 (s, 8H, **BzPy**), 8.60 (s, 8H, **BzPy**), 8.22 (s, 4H, -N(CH_3)C**H**CHN-), 8.15 (s, 4H, -N(CH_3)CHC**H**N-), 8.10 (s, 4H, Ar**H$_{up}$**), 7.79 (s, 4H, Ar**H$_{down}$**), 5.56 (bt, 4H, -C**H**-), 4.36 (bt, 8H, -C**H$_2$**NMI), 3.91 (s, 12H, N-C**H$_3$**), 2.80 (m, 8H, CHC**H$_2$**-), 1.81 (m, 8H, -C**H$_2$**CH_2NMI). **^{13}C NMR** (DMSO-d_6, 600 MHz): δ ppm = 154.63, 154.23, 149.25, 141.28, 137.22, 135.61, 125.47, 124.04, 123.62, 123.10, 115.57, 53.37, 40.54, 36.38, 34.14, 24.42. **HR-ESI**: m/z calculated for $C_{88}H_{76}Cl_2N_{16}O_8^{2+}$: 778.2685 [M-2Cl]$^{2+}$, found 778.2722; m/z calculated for $C_{88}H_{76}ClN_{16}O_8^{3+}$: 506.5235 [M-3Cl]$^{3+}$, found 506.5244; m/z calculated for $C_{88}H_{76}N_{16}O_8^{4+}$: 371.4011 [M-4Cl]$^{4+}$, found 371.4020.

Recorcinarene [C_2H_5; H] (12).
To a solution of resorcinol (15 g, 0.14 mol) in MeOH (90 mL), propionaldehyde (10 mL, 0.14 mol) and 37% HCl (22.5 mL) were added dropwise over 30 min at 0 °C. After the addition, the reaction mixture was stirred at 50 °C for a week. The reaction was quenched with water,

filtered, dried under vacuum and recrystallized three times from MeOH. The final product **12** was obtained as a pale beige powder (8 g, 38%).
^{1}H NMR (DMSO-d_6, 300 MHz): δ ppm = 8.85 (s, 8H, O**H**), 7.17 (s, 4H, Ar**H_{up}**), 6.08 (s, 4H, Ar**H_{down}**), 4.03 (t, 4H, -**CH**-, J= 7.59 Hz), 2.05 (m, 8H, CH**CH_2**-), 0.73 (t, 12H, -CH_2**CH_3**, J= 6.93 Hz); **ESI-MS:** m/z 602.07 $[M+H]^+$; m/z 623.87 $[M+Na]^+$.

Octanitro cavitand (13).[13a]
To a solution of resorcinarene **12** (250 mg, 0.41 mmol) and 1,2-difluoro-4,5-dinitrobenzene (341 mg, 1.64 mmol) in 5 mL of dry DMF, NEt_3 was added (0.58 mL, 4.10 mmol) and the mixture was heated to 70 °C for 24 hrs. The solvent was removed in vacuum and the residue was treated with MeOH (15 mL). The resulting precipitate was filtered off and washed with MeOH (3 x 8 mL). The precipitate was suspended in MeOH (15 mL) and heated at reflux for 1 h. The residual precipitate was filtered and washed with MeOH (3 x 5 mL) and Et_2O (3 x 2 mL) to give compound **13** as a yellow solid (360 mg, 69%).
^{1}H NMR (DMSO-d_6, 300 MHz): δ ppm = 8.85 (s, 8H, Ar**H**), 8.26 (s, 4H, Ar**H_{down}**), 7.83 (s, 4H, Ar**H_{up}**), 5.41 (bm, 4H, -**CH**-), 2.44 (bm, 8H, CH**CH_2**-), 0.92 (t, 12H, -CH_2**CH_3**); **MALDI-TOF:** m/z calculated 1256.2223 $[M+H]^+$; m/z 1279.2115 $[M+Na]^+$.

Octaamino cavitand (14).[29]
50 mL of concentrated HCl and $SnCl_2 \cdot 2H_2O$ (672 mg, 2.98 mmol) were added to a suspension of octanitro cavitand **13** (50 mg, 0.04 mmol) in 4 mL of ethanol. The mixture was stirred for 16 hrs at 70 °C. After cooling, the resulting precipitate was collected by centrifugation and washed with a 1:1 H_2O:EtOH mixture to give a white solid. The solid was suspended in 20 mL of ethyl acetate and 15 mL of NH_3 (28 % in H_2O) and the mixture was stirred for 30 min at room temperature. The organic layer was separated, dried over Na_2SO_4 and concentrated to afford octaamino cavitand **14** (40 mg, 98%).
This compound showed gradual decomposition in contact to air and therefore was rapidly used in the next step.

Tetraester cavitand (15).
Octaamine cavitand **14** (40 mg, 0.04 mmol) and ethyl-3-ethoxy-imminopropionate (93 mg, 0.48 mmol) were combined in anhydrous EtOH (2 mL, stored over 4 Å molecular sieves) and heated at reflux overnight. The mixture was cooled to room temperature, concentrated to dryness and taken up in DCM. The suspension was filtered and washed with copious DCM, followed by copious water. The solid was dried under high vacuum for 48 hrs to yield the product as a buff-colored solid (48 mg, 85%).
^{1}H NMR (THF-d_8/D_2O, 400 MHz): δ ppm = 7.81 (s, 4H, $\mathbf{H_{Ar}}$), 7.68 (s, 4H, Ar$\mathbf{H_{up}}$), 7.34 (s, 4H, Ar$\mathbf{H_{down}}$), 5.60 (t, 4H, -**CH**-, J = 8.2 Hz), 4.12 (q, 8H, COO$\mathbf{CH_2}$-, J = 7.2 Hz), 2.29 (m, 8H, CH$\mathbf{CH_2}$-), 1.15 (t, 12H, -COOCH$_2\mathbf{CH_3}$, J = 7.0 Hz), 0.96 (t, 12H, CHCH$_2\mathbf{CH_3}$, J = 8.4 Hz). **MALDI-TOF:** m/z 1401.5235 $[M+H]^+$; m/z 1423.5019 $[M+Na]^+$.

Tetrasodium salt cavitand (16).
Cavitand tetraester **15** (47 mg, 0.03 mmol) was dissolved in a mixture of THF (4.2 mL) and water (2.8 mL), and NaOH (45 mg, 1.11 mmol) was added as a solid. The mixture was stirred for 1 h, and a precipitate occurred. Additional water (0.25 mL) gave a clear solution and stirring was continued at rT for 2 days. THF was slowly removed leading to the formation of a precipitate. The formed precipitate was collected by filtration, washed with DCM and dried under high vacuum to give product **16** as a buff-colored solid (36 mg, 77%).
^{1}H NMR (D_2O, 600 MHz): δ ppm= 7.72 (s, 8H, $\mathbf{H_{Ar}}$), 7.67 (s, 4H, Ar$\mathbf{H_{up}}$), 7.45 (s, 4H, Ar$\mathbf{H_{down}}$), 5.41 (t, 4H, -**CH**-, J = 8.5 Hz), 3.58 (s, 8H, -$\mathbf{CH_2}$COO$^-$ Na^+), 2.28 (m, 8H, CH$\mathbf{CH_2}$-), 0.88 (t, 12H, CHCH$_2\mathbf{CH_3}$, J = 7.1 Hz). **HR-ESI:** m/z calculated for $C_{72}H_{55}N_8O_{16}^{3-}$: 1287.3741 $[M-4Na+3H]^-$, found 1287.3717; m/z calculated for $C_{72}H_{54}N_8O_{16}^{2-}$: 643.1834 $[M-4Na^++2H^+]^{2-}$, found 643.1830; m/z calculated for $C_{72}H_{53}N_8O_{16}^{3-}$: 428.4532 $[M-4Na^++H^+]^{3-}$, found 428.4527.

4.7 REFERENCES

1. N.A. Hafid, J. Christodoulou, *Transl. Pediatr*. **2015**, *4*, 304.
2. M.D. Milne, *Biochem J.* **1969**, *111*, 3P.61.
3. R.A. Williams, C.DS. Marmotte, J.R. Burnett, *Clin. Biochem. Rev.* **2008**, *29*, 31.
4. (a) R. Kleta, E. Romeo, Z. Ristic, T. Ohura, C. Stuart, M. Arcos-Burgos, M.H. Dave, C.A. Wagner, S.R.M. Camargo, S. Inoue, N. Matsuura, A. Helip-Wooley, D. Bochenhauer, R. Warth, I. Bernardini, G. Visser, T. Eggermann, P. Lee, A. Chairoungdua, P. Jutabha, E. Babu, S. Nilwarangkoon, N. Anzai, Y. Kanai, F. Verrey, W.A. Gahl, A. Koizumi, *Nat. Gen.* **2004**, *36*, 999; (b) H.F. Seow, S. Bröer, A. Bröer, C.G. Bailey, S.J. Potter, J.A. Cavanaugh, J.E.J. Rasko, *Nat. Gen.* **2004**, *36*, 1003.
5. S. van Dun, C. Ottmann, L.-G. Milroy, L. Brunsveld*, J. Am. Chem. Soc.* **2017**, *139*, 13960.
6. G.V. Oshovsky, D.N. Reinhoudt, W. Verboom, *Angew. Chem. Int. Ed.* **2007**, *46*, 2366.
7. (a) J.R. Moran, S. Karbach, D.J. Cram, *J. Am. Chem. Soc.* **1982**, *104*, 5826; (b) J.A. Bryant, J.L. Ericson, D.J. Cram, *J. Am. Chem. Soc.* **1990**, *112*, 1254; (c) J.R. Moran, J.L. Ericson, E. Dalcanale, J.A. Bryant, C.B. Knobler, D.J. Cram, *J. Am. Chem. Soc.* **1991**, *113*, 5707; (d) D.J. Cram, H.J. Choi, J.A. Bryant, C.B. Knobler, *J. Am. Chem. Soc.* **1992**, *114*, 7748.
8. P.J. Skinner, A.G. Cheetham, A. Beeby, V. Gramlich, F. Diederich, *Helv. Chim. Acta* **2001**, *84*, 2146.
9. P. Roncucci, L. Pirondini, G. Paderni, C. Massera, E. Dalcanale, V.A. Azov, F. Diederich, *Chem. Eur. J.* **2006**, *12*, 4775.
10. V.A. Azov, B. Jaun, F. Diederich, *Helv. Chim. Acta* **2004**, *87*, 449.
11. J.W. Trzciński , R. Pinalli, N. Riboni, A. Pedrini, F. Bianchi, S. Zampolli, I.Elmi, C. Massera, F. Ugozzoli, E. Dalcanale, *ACS Sensors.* **2017**, *2*, 590.
12. M. Giannetto, A. Pedrini, S. Fortunati, D. Brando, S. Milano, C. Massera, R. Tatti, R. Verucchi, M. Careri, E. Dalcanale, R. Pinalli, *Sens. Actuators B Chem.* **2018**, *276*, 340.
13. (a) K-.D. Zhang, D. Ajami, J. Rebek, *J. Am. Chem. Soc*. **2013**, *135*, 18064; (b) K-.D. Zhang, D. Ajami, J.V. Gavette, J. Rebek, *J. Am. Chem. Soc*. **2014**, *136*, 5264; (c) N.-W. Wu, J. Rebek, *J. Am. Chem. Soc*. **2016**, *138*, 7512; (d) R. Pinalli, G. Brancatelli, A. Pedrini, D. Menozzi, D. Hernàndez, P. Ballester, S. Geremia, E. Dalcanale, *J. Am. Chem. Soc*. **2016**, *138*, 8569; (e) S. Mosca, Y. Yu, J. Rebek, *Nat. Protoc*. **2016**, *11*, 1371.

14. S.W. Jacob, E.E. Rosenbaum, D.C. Wood, Dimethyl Sulfoxide vol. 1, Marcel Dekker, New York, **1971**.
15. (a) T. Arakawa, J.F. Carpenter, Y. Kita, J.H. Crowe, *Cryobiology* **1990**, *27*, 401; (b) S.J. Baxter, G.H. Lathe, *Biochem. Pharmacol.* **1971**, *20*, 1079.
16. (a) K. Hamaguchi, *J. Biochem.* **1964**, *56*, 441; (b) P. Maurel, *J. Biol. Chem.* **1978**, *253*, 1677; (c) A. Miegel, K.I. Sano, K. Yamamoto, K. Maeda, Y. Maeda, H. Taniguchi, M. Yao, S. Wakatsuki, *FEBS Lett.* **1996**, *394*, 201.
17. K. Yamazaki, T. Iwura, R. Ishikawa, Y. Ozaki, *J. Biochem.* **2006**, *140*, 49.
18. K.R. Babu, D.J. Douglas, *Biochemistry* **2000**, *39*, 14702.
19. (a) M. Frei, F. Diederich, R. Tremont, T. Rodriguez, L. Echegoyen, *Helv. Chim. Acta* **2006**, *89*, 2040; (b) M. Torelli, I. Domenichelli, A. Pedrini, F. Guagnini, R. Pinalli, F. Terenziani, F. Artoni, R. Brighenti, E. Dalcanale, *Syn. Lett*. **2018**, *29*, 2503.
20. K-.D. Zhang, D. Ajami, J.V. Gavette, J. Rebek, *Chem. Commun*. **2014**, *50*, 4895.
21. F. Manea, C. Bindoli, S. Polizzi, L. Lay, P. Scrimin, *Langmuir* **2008**, *24*, 4120.
22. I.R. Verly, A.B. van Kuilenburg, N.G. Abeling, S.M. Goorden, M. Fiocco, F.M. Vaz, M.M. van Noesel, C.M. Zwaan, G.L. Kaspers, J.H. Merks, H.N. Caron, G.A. Tytgat, *Eur. J. Cancer* **2017**, *72*, 235.
23. N.C. Colon – MD, D.H. Chung – MD, *Adv. Pediatr.* **2011**, *58*, 297.
24. A. Aprile, F. Ciuchi, R. Pinalli, E. Dalcanale, P. Pagliusi, *J. Phys. Chem. Lett.* **2016**, *7*, 3022.
25. F.-U. Rahman, H.-N. Feng and Y. Yu, *Org. Chem. Front.* **2019**, *6*, 998.
26. F.-U. Rahman, Y.-s. Li, I.D. Petsalakis, G. Theodorakopoulos, J. Rebek Jr., Y. Yu, *PNAS* **2019**, *116*, 17648.
27. L.M. Mesquita, J. Anhäuser, D. Bellaire, S. Becker, A. Lützen, S. Kubik, *Org. Lett.* **2019**, *21*, 6442.
28. (a) S.M. Biros, E.C. Ullrich, F. Hof, L. Trembleau, J. Rebek, *J. Am. Chem. Soc*. **2004**, *126*, 2870; (b) R.J. Hooley, H.J. Van Anda, J. Rebek, *J. Am. Chem. Soc.* **2007**, *129*, 13464; (c) Y. Liu, P. Liao, Q. Cheng, R.J. Hooley, *J. Am. Chem. Soc.* **2010**, *132*, 10383; (d) Y.-J. Ghang, J.J. Lloyd, M.P. Moehlig, J.K. Arguelles, M. Mettry, X. Zhang, R.R. Julian, Q. Cheng, R.J. Hooley, *Langmuir* **2014**, *30*, 10161.
29. B. Soberats, E. Sanna, G. Martorell, C. Rotger, A. Costa, *Org. Lett*. **2014**, *16*, 840.

APPENDIX A

1. UV-Vis spectra of QxCav in vase and kite form

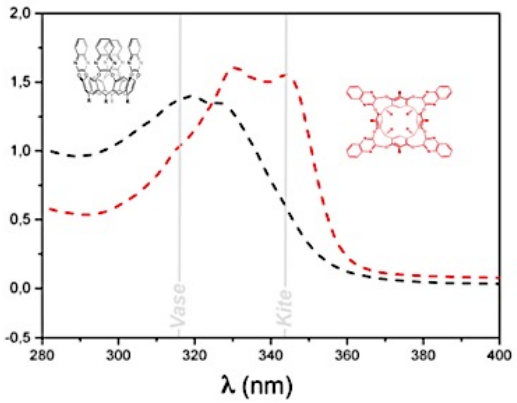


Figure A.1. UV-Vis spectra of QxCav in *vase* form (black) and in *kite* conformation (red) after TFA addition. The spectra were recorded using $CHCl_3$ as solvent.

2. Kite-Vase switch upon pH titration

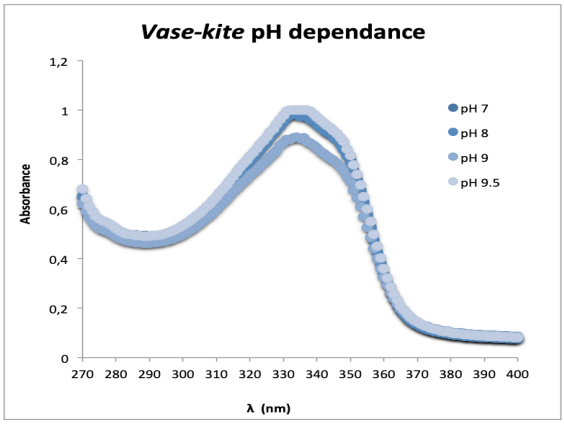


Figure A.2. UV-Vis titration of 4QxPy **4** in water.

APPENDIX B

1. ^{1}H NMR titration of 4QxPy with Phe at 25 °C

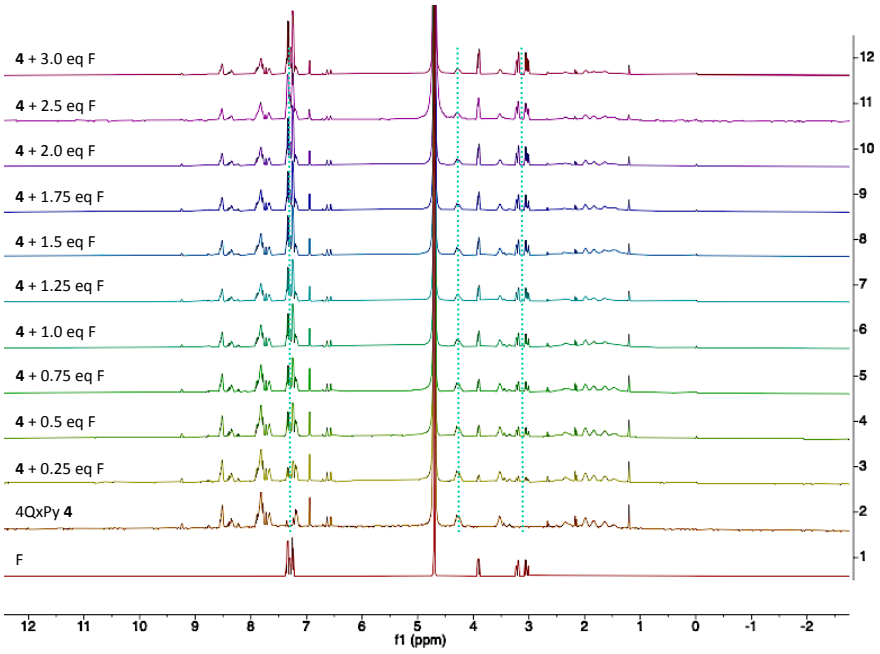


Figure B.1. ^{1}H NMR (D_2O, 25 °C, 600 MHz.) of Phenylalanine (**F**), 4QxPy cavitand **4** and subsequent addition of 0.25 eq. of F to **4**.

2. ^{1}H NMR titration of 4QxPy with Phe at 5 °C

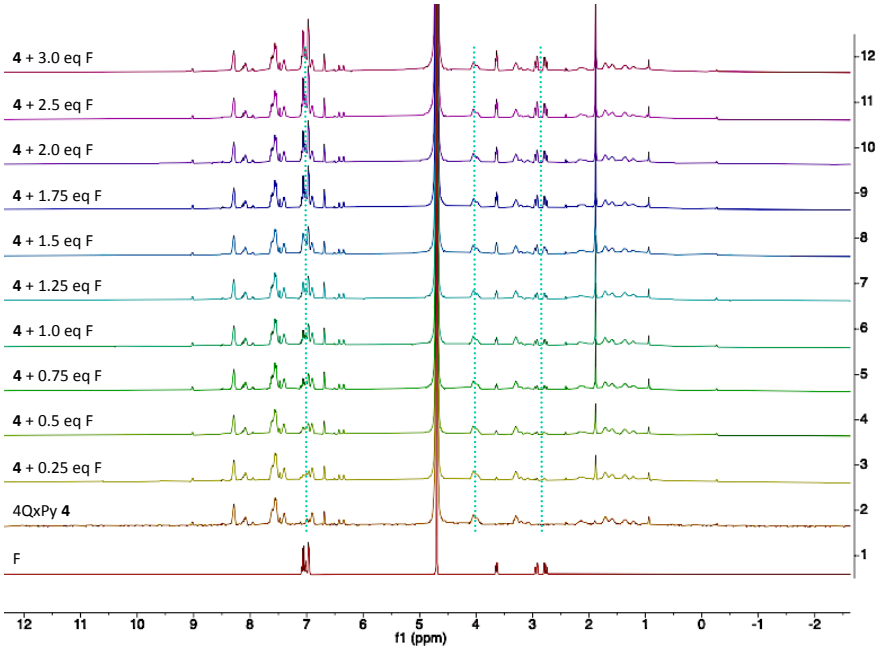


Figure B.2. ^{1}H NMR (D_2O, 5 °C, 600 MHz.) of Phenylalanine (**F**), 4QxPy cavitand **4** and subsequent addition of 0.25 eq. of F to **4**.

3. ITC titrations in NaOAc pH 4.6

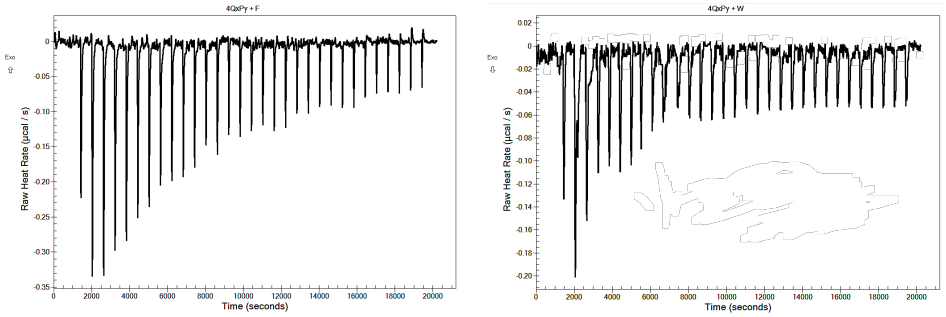
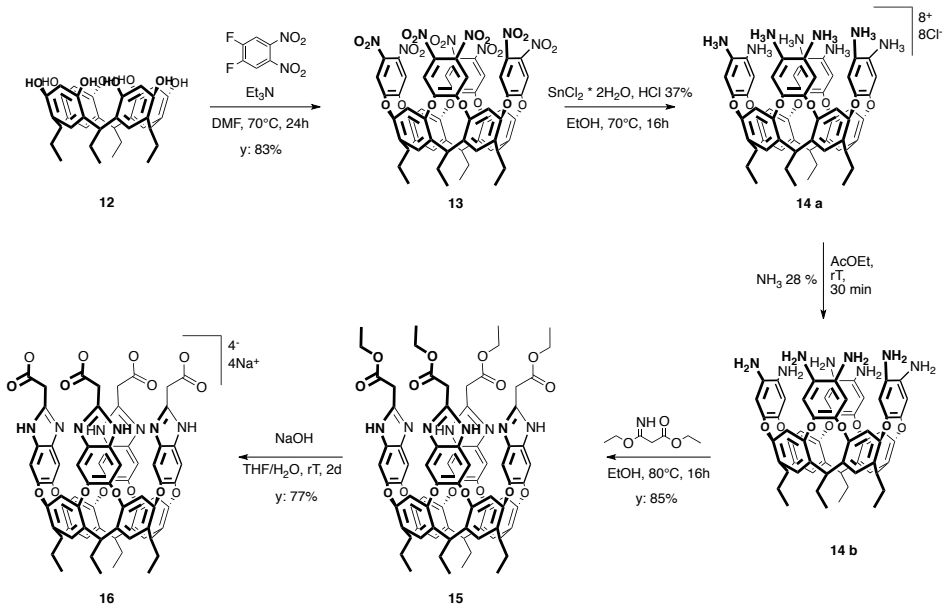


Figure B.3. ITC experiments of **4** with Phenylalanine (*right*) and Trypthophan (*left*). All experiments were conducted in NaOAc pH 4.6; [cavitand] = 1 mM; [AA] = 10 mM.

APPENDIX C

*1. Synthesis of deep cavitand **16***

The water soluble self-folding cavitand **16** was synthesized following literature procedures[1-3] (**Scheme C.1)**. The following synthetic pathway allowed us to obtain **16** as tetrasodium salt in four steps starting from resorcinarene **12**, with 54% overall yield.



Scheme C.1. Four-steps synthesis of cavitand **16**.

The resorcinarene scaffold **12** was initially bridged with 1,2-difluoro-4,5-dinitrobenzene, affording cavitand **13** in good yields. Octanitro cavitand was subsequently reduced with $SnCl_2 \bullet H_2O$ in the presence of HCl 37% to obtain octaamine cavitand **14 a**. Since the amino groups present on cavitand **14 a** are particularly instable, they were protected as ammonium chloride salt. The amino groups were deprotected immediately before the next step by treatment with a diluted solution of ammonia in ethyl acetate and subsequent liquid extraction. The free amino cavitand **14 b** was then directly employed in the condensation

reaction with ethyl 3-ethoxy-3-iminopropanoate providing tetrabenzimidazole ester cavitand **15**. Cavitand **15** was thereafter hydrolyzed under standard basic conditions to give the target compound **16** as tetrasodium salt in 77% yield.

*2. Characterization of **16***

All the products were fully characterized by NMR spectroscopy and ESI-MS spectrometry. In **figure C.1** the ^{1}H NMR of the target compound is reported.

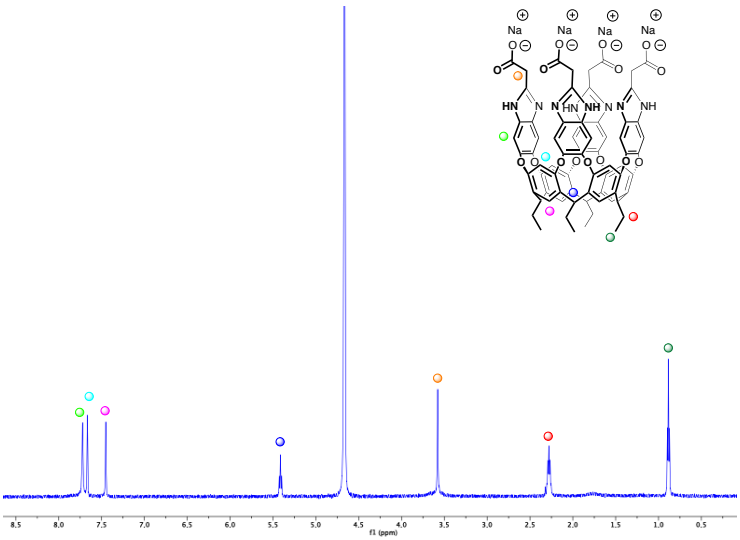


Figure C.1. ^{1}H NMR of cavitand **16** in D_2O, 25 °C, 600 MHz.

The ^{1}H NMR spectrum shows three singlets at 7.72, 7.67 and 7.45 ppm related to the benzoimidazole protons (green dot) and the aromatic protons of the resorcinarene scaffold (light blue and pink dots). At 5.41 ppm the signal of the diagnostic methine triplet (blue dot) confirms the *vase* conformation of the scaffold. In the aliphatic part of the spectrum, a multiplet and a triplet, at 2.28 and 0.88 ppm respectively, defined the alkyl chain of the resorcinarene. The presence of only one singlet at 3.58 ppm further confirms the success of the saponification of the ester.

The HR-ESI mass analysis confirms the nature of **16** as reported in the spectrum in **Figure C.2**. The characteristic peaks of the molecule are present as mono-charged ion at 1287.3717 Da (peak $[M-4Na^{+}+3H^{+}]^{-}$), di-charged ion at 643.1830 Da (peak $[M-4Na^{+}+2H^{+}]^{2-}$) and tri-charged ion at 428.4527 Da (peak $[M-4Na^{+}+H^{+}]^{3-}$).

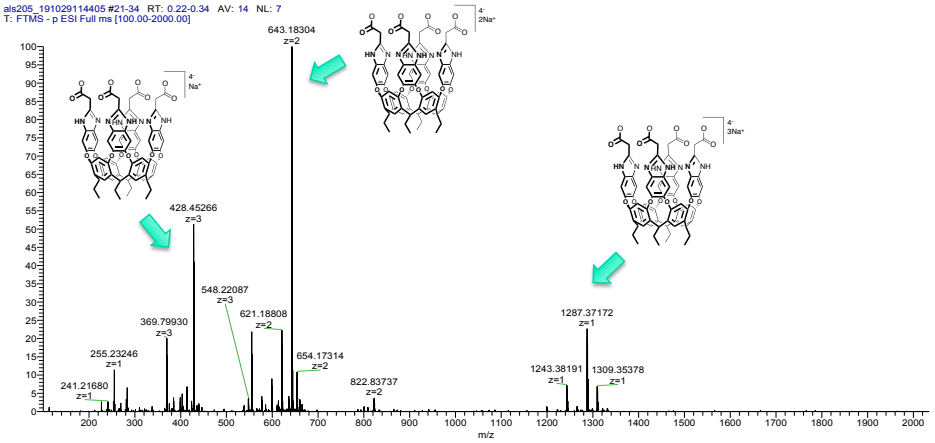


Figure C.2. HR-ESI of water soluble deep cavitand **16**.

3. Binding studies with Phe and Trp

3.1 ^{1}H NMR studies

For a better understanding of the complexing ability in water of this water soluble cavitand towards phenylalanine and tryptophan, ^{1}H NMR titrations in D_2O were performed. In **figure C.3** the experiments with Phe as guest are reported. To a 1 mM solution of **16** (1.25 mg, 0.9 µmol), subsequent portions of 0.25 eq of a 14 mM guest solution (0.68 mg, 4.12 µmol) were added and the ^{1}H NMR spectra recorded. After reaching 2 eq, addition were to 0.5 eq. As can be clearly seen from the spectra, no inclusion processes occurred, since no evident shift was observed both for the host and the guest.

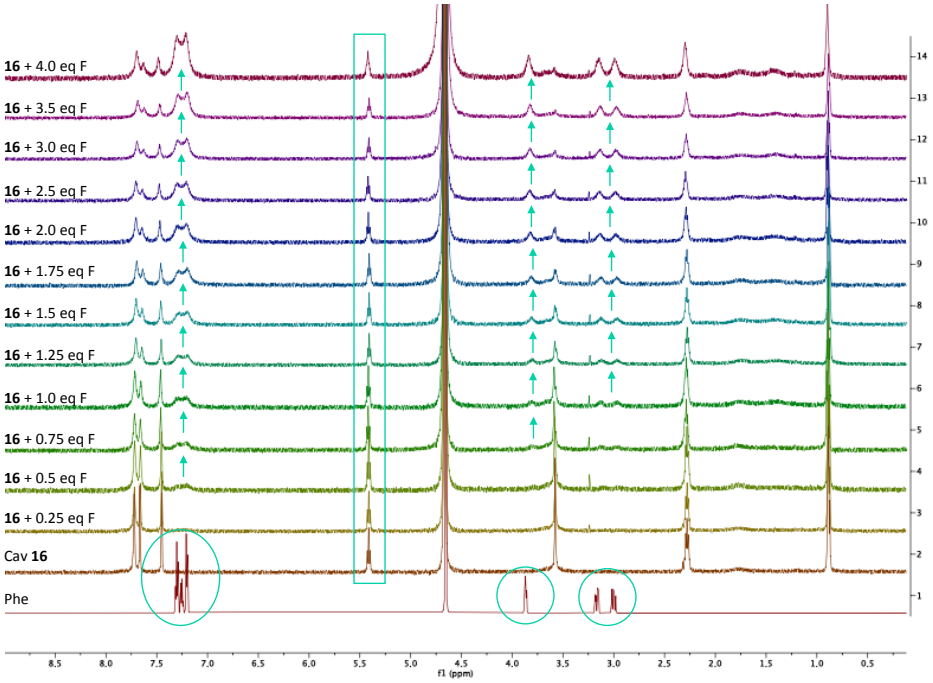


Figure C.3. ^{1}H NMR titration in D_2O with cavitand **16** and phenylalanine, 25 °C, 600 MHz.

3.2 ITC studies

The ITC experiments were carried out to quantify the thermodynamic parameters Δ*H*, K_a, Δ*G* and Δ*S* associated with the inclusion processes.
The experiments were performed in 0.1 M NaOAc buffer at pH 4.6, adding a 20 mM solution of Phe or Trp into the cell containing 1 mM of cavitand **16**. This buffer was chosen not only to maintain constant the pH during the titration experiments, but also for its low ionization heat that gives minimal additional heat contributions.
In **Figure C.4** ITC curves are reported showing that no valuable results were obtained. In the case of Phe (**Figure C.4** right), the low heat released prevented any fitting of the data. This, combined with the ^{1}H NMR titration, allows us to affirm that, if present, the host-guest interactions that are formed are very weak. In Trp titration experiments, the injection heat values suggest a major enthalpy contribution for this

guest. This can be due to the more hydrophobic nature of Trp in respect to Phe. However, also in this case, the not clear sigmoidal curve prevented to fit the data.

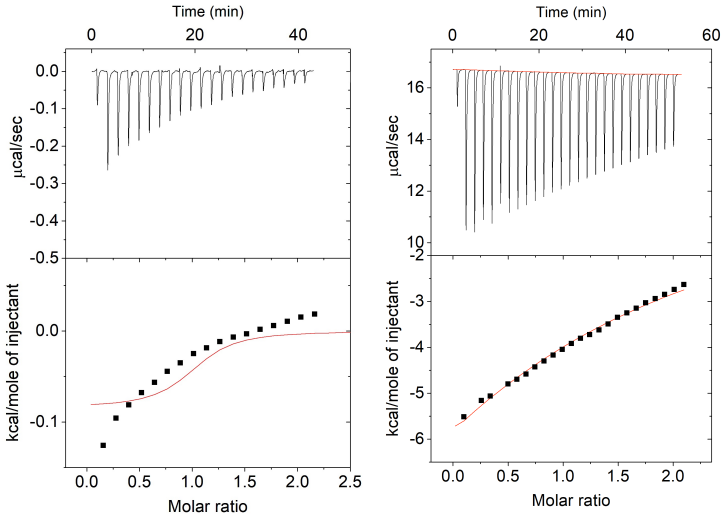


Figure C.4. ITC titrations of 1 mM **16** in 0.1 M NaOAc pH 4.6 with 20 mM Phe (*right*) and Trp (*left*), both in 0.1 M NaOAc pH 4.6.

In conclusion, liquid state analyses revealed the receptor **16** to be inefficient in complexing Phe and Trp.

4. Solid state studies

Systematic studies to elucidate the complexation event of **16** with Phe and Trp at the solid state were carried out in the laboratory of Professor Crowley, at the National University of Ireland (NUI) in Galway.
Co-crystallization trials were performed employing both a sparse matrix screen (Jena Bioscience JCSG++) for initial screening of crystallization conditions using the Oryx 8 robot (Douglas instrument) and in manual screens using the hanging or sitting drop techniques.

Hanging and Sitting drop. The hanging drop and sitting drop vapor diffusion techniques are known to be among the most efficient methods for the crystallization of macromolecules, in particular proteins. The principle of crystallization is bringing the macromolecule, commonly proteins, to supersaturation introducing it into a precipitating agent able to promote the nucleation of crystals in the solution. Experimentally, a drop composed of a mixture of protein, ligand and precipitants is placed in the vicinity of a liquid reservoir of pure precipitants with which established vapor equilibration. The mixed drop is placed on a microbridge in the case of the sitting drop technique (**Figure C.5** left) or a siliconized cover slide for the hanging drop (**Figure C.5** right). In both cases the wells were sealed to allow the initiation of the equilibration process. Containing a lower concentration of precipitants than the reservoir, water vapor leaves the drop and eventually ends up in the reservoir until equilibrium is achieved. As water leaves the drop, the sample undergoes an increase in relative supersaturation, ideally in the crystal nucleation zone.

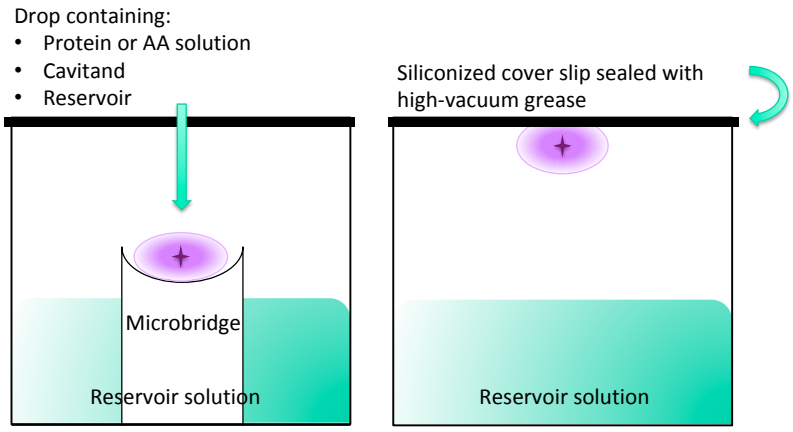


Figure C.5. Sitting drop (*left*) and hanging drop (*right*) methods.

Robot crystallization. Exploiting the sitting drop method, preliminary trials were performed using a sparse matrix screen JCSG++ with the Oryx 8 robot. A 96-well plate was prefilled with the commercial JCSG++ crystallization screening kit. Protein or AAs stock solutions were prepared along with the cavitand one. The instrument was set to

disperse 15 µL of each solution (i.e. AA, cavitand and reservoir) through three independent channels so the solutions do not mix in the tip but in the drop after they were dispensed. Screening for sitting drop was carried out using 3-drop plates where: i) a control drop, in which only the AA or the cavitand was mixed with the reservoir; and ii) two experimental drops containing a mixture of cavitand, AA and reservoir solution at different and arbitrarily concentrations. Control drops were always set to avoid false positive results, i.e. the growth of crystals of the host or the guest alone.
The results obtained from these preliminary tests with cavitand **16** showed the growth of crystals in JCSG++ screen condition C7 [10% PEG 3000 + 100 mM NaOAc pH 4.5 + 200 mM ZnOAc] for tryptophan and E7 [10% Isopropanol + 100 mM MES pH 6.5 + 200 mM ZnOAc] for phenylalanine.
These conditions (Jena screen JGSC++) were replicated manually changing also the amount of salts and precipitants in order to find the best crystallization conditions using the hanging drop technique.

Manual crystallization. Trials were set using hanging drops method by vapor diffusion technique in a 24-well plate. The manual experiments were set using 3 µL drops containing 1 µL of reservoir solutions, 1 µL of the receptor **16** and 1 µL of the guest, both at a known concentration. To avoid false positive results, control drops were always set as follows: 1 µL of reservoir solutions, 1 µL of AA or host **16** at the same concentration as the experimental drop and exchanging one of them with 1 µL of water to maintain equal the final volume of the drop. In particular, the addition of water results important to maintain constant the vapor diffusion equilibrium between the drop and the reservoir.
The guest concentration was sought in 1-40 mM range for both Phe and Trp. Proper concentration of host **16** was found to be 1 mM since fast precipitation occurred above 5 mM leading to an amorphous precipitate. Reservoir solutions of conditions C7 and E7 were explored varying the concentrations of precipitants and salt but crystals grown only under the conditions reported above.

As can be clearly seen in **figure C.6 a**, co-crystallization trials with phenylalanine leads to the formation of crystals only in the

experimental drop allowing to believe that an interaction between the AA and cavitand **16** has settled since no crystals are present in the control drops in which only the AA or the receptor are present. On the contrary, the presence of crystals also in the tryptophan control drop (**C.6 b**) suggests that crystals grown in the experimental drop are unlikely belong to the Trp@Cav **16** complex, even if the morphology of the crystal is different from that of the control drop.
Test carried out using only the receptor in the same reservoir solution were performed to exclude that the crystals present in the experimental drops are relative to the receptor only. In this case, no crystal formation was observed.

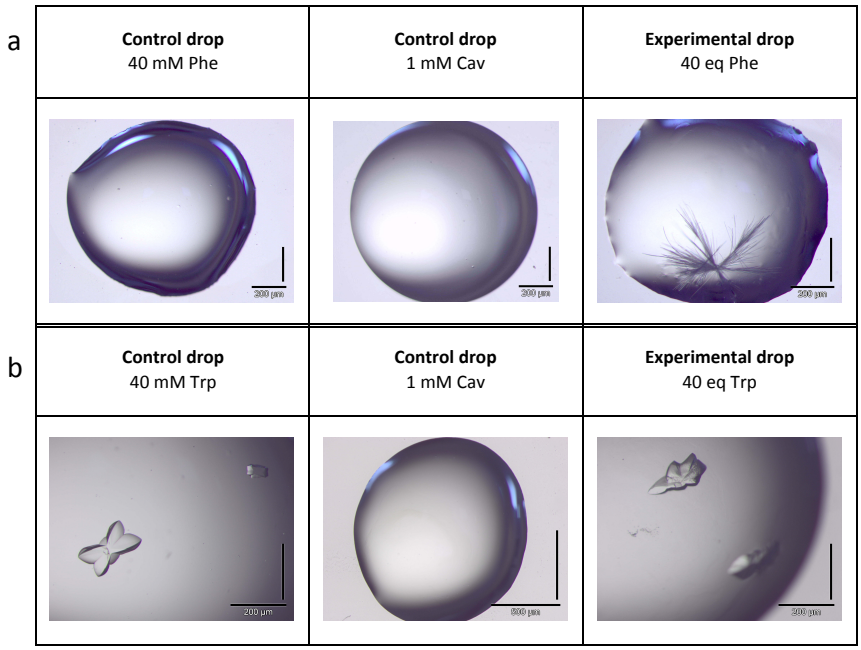


Figure C.6. Images obtained by optical microscopy of crystals grown in conditions E7 with Phe (**a**) and C7 with Trp (**b**).

In all cases, for having a clear response, the nature of these crystals has to be determinate by X-ray diffraction. Unfortunately, in the opinion of Prof. Crowley the crystals were not sufficient promising for further investigation through X-ray diffraction.

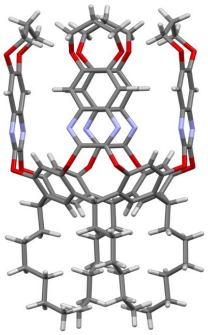
5. Appendix References

1. S.M. Biros, E.C. Ullrich, F. Hof, L. Trembleau, J. Rebek Jr., *J. Am. Chem. Soc*. **2004**, *126*, 2870.
2. K.-D. Zhang, D. Ajami, J. Rebek, *J. Am. Chem. Soc.* **2013**, *135*, 18064.
3. B. Soberats, E. Sanna, G. Martorell, C. Rotger, A. Costa, *Org. Lett*. **2014**, *16*, 840.
4. V.M. Bolanos-Garcia, N.E. Chayen, *Prog. Biophys. Mol. Biol*. **2009**, *101*, 3.

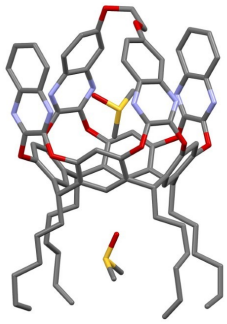
CHAPTER 5

Synthetic studies toward new benzene receptors for environmental sensing

Highly selective carcinogenic benzene detection is socially relevant and technologically challenging, due to the concurring requirements of high selectivity and extreme sensitivity.
Recently, we have developed a MEMs based device in which the specifically designed EtQxBox cavitand acts at the same time as highly sensitive preconcentrator for BTEX and GC-like separation phase, allowing for the selective desorption of benzene over TEX by applying a smart temperature program on the EtQxBox mesh. EtQxBox, obtained by covalently linking together the four quinoxaline walls, strongly retains BTEX to the point that BTEX can be exhaustively desorbed only at temperatures higher than 250 °C, thus reducing the lifetime of the device. Along this line we synthesized a new quinoxaline-based cavitand, where only two quinoxaline walls are linked together to reduce the binding strength of our receptor towards BTEX, allowing their decomplexation at lower temperatures. Similarly, synthetic studies for the development of quinoxaline-based deep cavitands were performed with the intent of maximizing the affinity of the receptor towards the sole benzene.



DioxepaneQxCav



EtQxAC

5.1 INTRODUCTION

Benzene is an ubiquitous atmospheric pollutant deriving mainly from anthropogenic sources, in particular combustion. It is a component of gasoline, vehicle exhaust gases, industrial emissions and tobacco smoke. Historically, it has been used as a solvent in industry and consumer products and now, even if its use as solvent is banned in most countries, benzene is still obtained in large volumes to be used as a chemical intermediate.

Benzene was officially included in the list of carcinogens to human in 1979 by the IARC (International Agency for Research on Cancer). A single exposure at a concentration of 20,000 ppm of commercial benzene is fatal for humans in just 10 minutes. However, even prolonged exposure to lower concentration levels can cause acute poisoning resulting in inflammation of the respiratory tract, pulmonary hemorrhage and cerebral edema,[1] alongside with blood cancers such as acute non-lymphocytic leukemia and acute myeloid leukemia.[2]

For the above mentioned reasons, the concentration of benzene in air is one of the values that is constantly measured by air quality control systems. The risks that occur after long exposures to particulates and toxic molecules such as benzene or NO_x are usually underestimated. Nowadays, real-time measurements systems for the detection of benzene are performed by bulky, expensive high-end systems based on advanced laboratory equipment, which require continuous service to operate.[3]

Recently, our research group has reported a stand-alone system based on molecular recognition for the selective detection of volatile organic compounds (VOCs).[4] (**Figure 5.1 b**)
The quinoxaline cavitand receptor (QxCav) was rationally designed to exploit multiple host-guest interactions involved in the binding event. Thanks to the specific interactions such as CH-π and π-π interactions, the receptor cavity is able to discriminate aromatic hydrocarbons against aliphatic ones in a selective manner both in the gas phase[5] and in the solid state.[6]

The novelty of this portable, low-cost and stand alone device for the detection of BTEX (Benzene, Toluene and Xylenes), at sub-ppb level lays in the combination of selectivity and sensitivity achieved combining two key components and technologies: a preconcentration unit (**Figure 5.1 a**, right) and GC separation microcolumn (**Figure 5.1 a**, left). The silicon MEMs preconcentrator, filled with a mesh of QxCav receptor powder, traps the target analytes by sampling large volumes of high flow air. After the given air volume is reached, a heat pulse to the MEMs cartridge is applied, releasing the entrapped BTEX molecules. The released analytes are consequently separated with the GC microcolumn using ambient air as carrier gas. As they exit the mini-GC column outlet, a miniaturized photoionization detector (PID) is used for the analytes quantification.

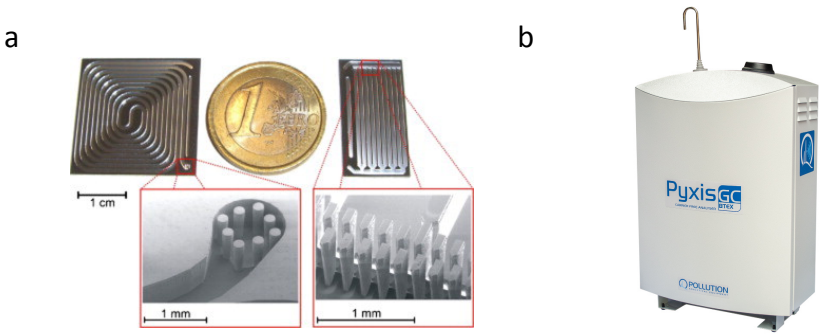


Figure 5.1. (a) Micromachined separation (left) and preconcentration (right) silica columns; (b) the commercially available device **Pyxis-GC BTEX.**

QxCav receptors have been used both in water quality control, exploiting hydrophobic effect[7], as well as in gas-solid interface for air monitoring, using mass[8] and surface plasmon resonance transducers.[9]
However, with these approaches it was impossible to obtain discrimination between similar aromatic VOCs, in particular, by selectively releasing benzene before the other aromatics.
Therefore, a new conformationally rigid cavitand was synthesized intended only for the benzene recognition. Developing a MEMs based device in which the specifically designed EtQxBox cavitand (**Figure 5.2 a**) acts at the same time as highly sensitive preconcentrator for BTEX and GC-like separation phase, the selective desorption of benzene over TEX is allowed. The desired selective desorption of benzene is achieved by applying a

smart temperature program on the EtQxBox mesh, which starts releasing benzene at lower temperatures than TEX, as predicted by the calculated binding energies. The sensor performances have been experimentally validated and ppb level sensitivity toward the carcinogenic and aromatic target benzene was demonstrated, as required for environmental benzene exposure monitoring in industrial applications and outdoor environment.[10]

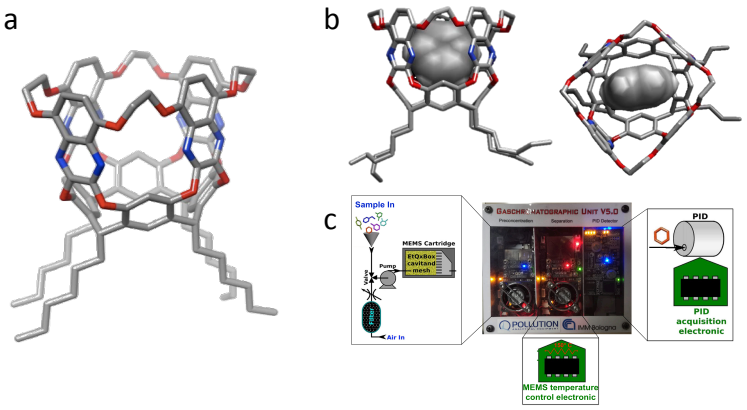


Figure 5.2. (a) EtQxBox structure, (b) crystal structure of EtQxBox in complex with benzene and (c) the ultimate device with EtQxBox cavitand mesh. Adapted from reference 10.

The selectivity of these QxCav macrocycles to specific classes of analytes can be increased by introducing functional groups capable of giving specific interactions with the class of analytes of interest or by varying the rigidity of the system. For example, the introduction of a COOH group at the upper rim of the quinoxaline cavitand allowed to increase the selectivity towards nitro aromatic compounds. As shown in **Figure 5.3**, cavitand 3QxCOOH was synthetized with an open cavity bearing three quinoxalinic walls and a carboxylic group bridged between two oxygen residues of the resorcinol, while the cavitand 4QxCOOH bridges three 1,4-diazanaphatalene moieties and a quinoxaline group functionalized in position 6 with a COOH residues. These cavitands were used as pre-concentrating cartridges and SPME fibres for the in air detection of nitroaromatic taggants.[11] It has been experimentally demonstrated that the first macrocycle undergoes decarboxylation already after the first use, due to the high desorption temperature, while the second receptor reveal

a good stability and a high selectivity towards the class of nitroaromatic explosives.

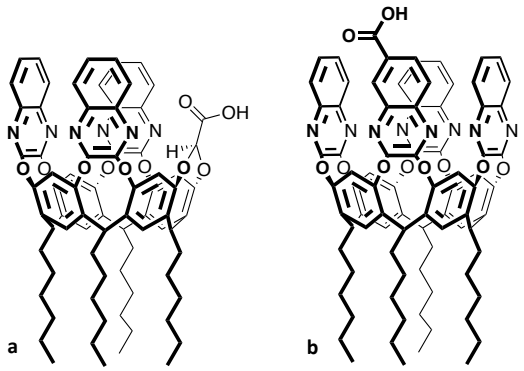


Figure 5.3. Structure of 3QxCOOH and 4QxCOOH cavitands used for detection of nitroaromatic compounds.

Along this line, it immediate appears that also the functionalization of the quinoxaline walls can be exploited in order to modulate and modify the selectivity of these macrocyclic receptors. In this respect, further studied were pursued in order to increase even more the selectivity of the receptor towards benzene. The previously mentioned EtQxBox, in fact, strongly retains BTEX, to the point that BTEX can be exhaustively desorbed only at temperatures higher than 250 °C, thus reducing the lifetime of the device. By reducing the dimension of the entrance of the cavity we are aiming to reverse the complexation trend of EtQxBox in favour of benzene over the TEX.

Computer simulations on 3D models (**Figure 5.4**) demonstrated that a new rigidified quinoxaline cavitand bearing the ethylendioxy bridges at the upper rim of the cavity rather than laterally, would lead to a narrow cage accessible only by benzene.

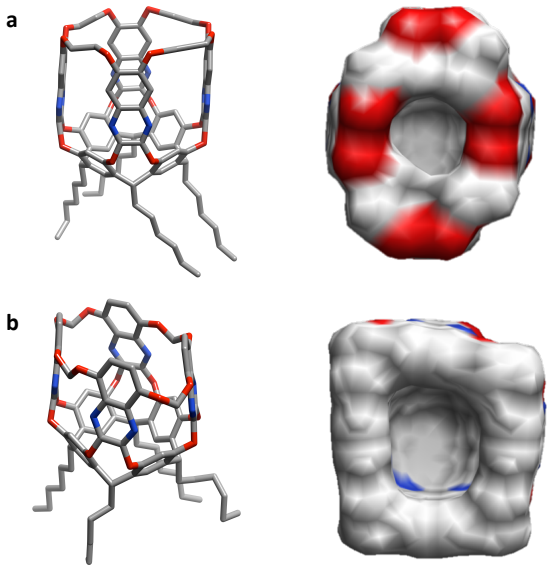


Figure 5.4. Comparison between the 3D model of EtQxCage (a) and the molecular structure of EtQxBox (b).

Previous attempts were performed to reduce the dimension of the entrance of the cavity *via* a stepwise approach.
Using a similar pathway to the one earlier employed for the EtQxBox, a quinoxaline-bridging-unit bearing eight methoxy groups in position 7,8 was synthetized. Then, after bridging the quinoxaline on the resorcinarene scaffold, the hydroxyl groups were deprotected and linked with ethylenedioxy bridges. Interestingly, a new isomeric deep cavitand (deep QxCav) was formed, as confirmed by the crystal structure. Instead of bridging the nearby OHs on different quinoxaline, the reaction led to the formation of a third new ring upon the quinoxaline units, resulting from the bridging of the OHs on the same quinoxaline rings. The formation of six-member rings is favoured over the bridging among quinoxaline walls (**Figure 5.5**).

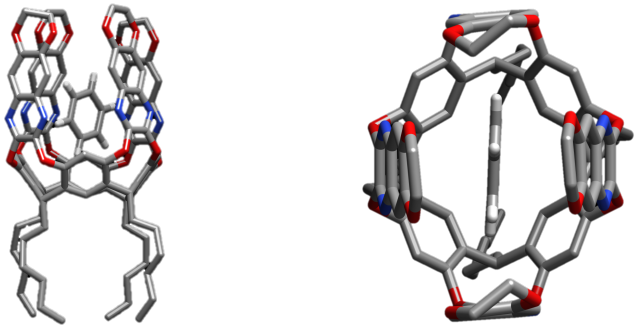
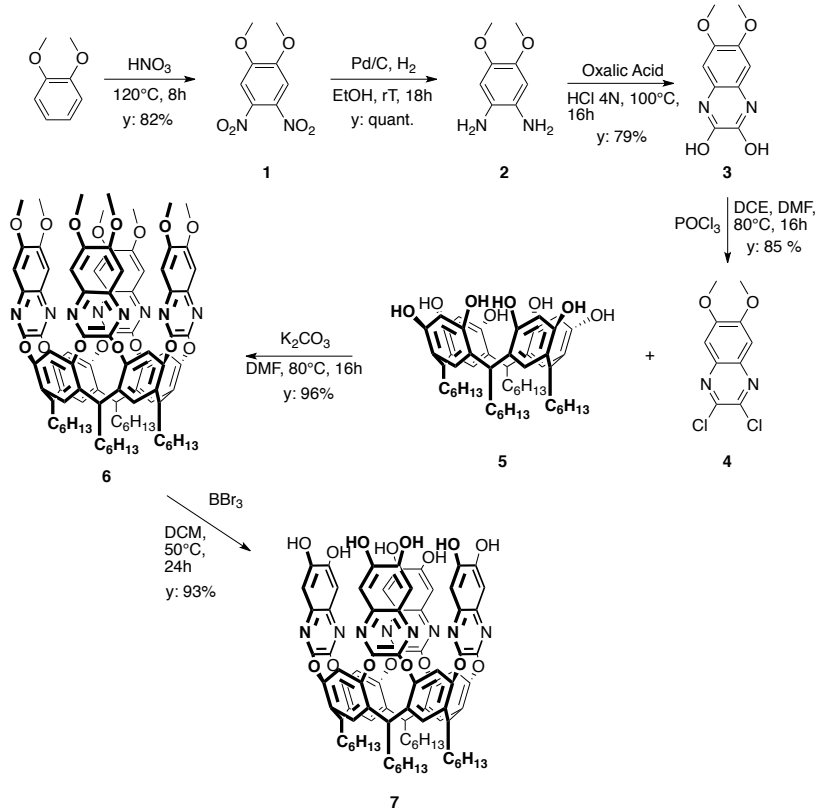


Figure 5.5. Molecular structure of the Deep QxCav with benzene as a guest.

5.2 TETRABRIDGED DEEP QUINOXALINE CAVITANDS

According to the Baldwin rules[12] the intramolecular closure leading to 5 and 6 member rings are favored because they are the best compromise between ring strain and the chances of ring formation.
We decided to test the same reaction with shorter and longer bridging units to evaluate the feasibility of the rigidification reaction at the upper rim versus the formation of 5- and particularly 7-term rings. In general, the formation of 5-member cycles is favoured over the corresponding acyclic reaction, while this is less the case with the intramolecular cyclization leading to 7-member rings.
The synthetic path followed involves a stepwise approach with cavitand **7** as common intermediate. (**Scheme 5.1**)

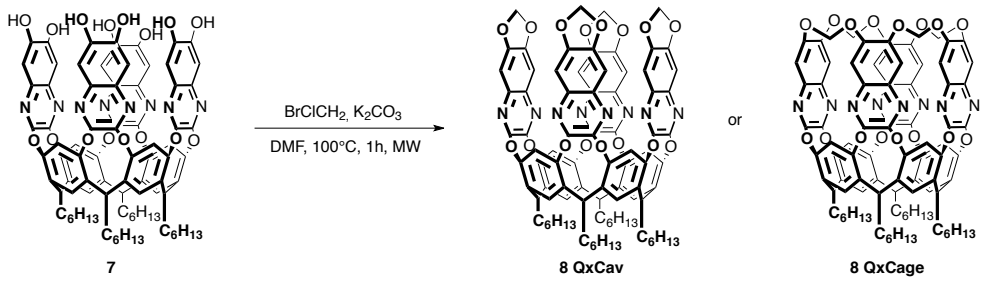


Scheme 5.1. Synthesis of the common intermediate for the upper rim rigidification of cavitands.

The synthesis of the quinoxaline unit began with a nitration reaction of 1,2-dimethoxy benzene. The reaction, carried out in 65% HNO_3, at 120 °C for 8 hours, produced **1** and 1,4-dimethoxy-2-nitrobenzene as a co-product. Reduction of the nitro groups with a hydrogen using 10% Pd/C as a catalyst, led to compound **2** in quantitative yield. Given its poor stability, **2** was immediately used in the subsequent cyclization reaction. By reacting **2** with oxalic acid in an aqueous solution of 4N HCl, **3** was prepared. Chlorination reaction of the intermediate **3** in anhydrous DCE as a solvent, $POCl_3$ as chlorinating agent and DMF as a catalyst, produced 1,8-dimethoxy-4,5-dichloroquinoxaline **4** in 93% yield. The bridging of the resorcinarene scaffold **5** with **4**, in anhydrous DMF in the presence of a base necessary to ensure the deprotonation of the phenolic OH groups of the resorcinarene, gave cavitand **6** in 80% yield. The deprotection of the eight OCH_3 groups at the upper rim with BBr_3 in DCM led to the desired intermediate cavitand **7**.

5.2.1 METHYLENE LINKER VERSUS 5-MEMBER RING AT THE UPPER RIM

The first investigated linker was the methylene one, theoretically leading to a narrower cavity entrance than the one modeled in **Figure 5.4 a**.



Scheme 5.2. Possible outcome of the bridging reaction with a $BrClCH_2$ linker.

Thus, using $BrClCH_2$ as linker, the reaction was tested in a microwave reactor, with DMF as solvent and K_2CO_3 as base. The choice of $BrClCH_2$ as reagent lies in its nature: it contains two halogen atoms with different

reactivity. Bromine, as better leaving group, is more reactive, starting the bridging reaction, with chlorine reacting in the second step. The work up of the reaction gave a single product in 14% yield. ^{1}H NMR and MALDI-TOF analyses (**Figures 5.6** and **5.7**) proved the formation of a single molecule.

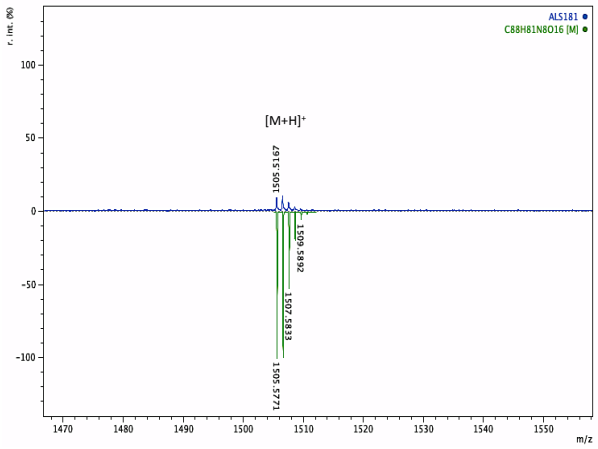


Figure 5.6. MALDI-TOF analysis of the purified product deriving from the bridging reaction with $BrClCH_2$.

As shown, the 1505.5167 mass corresponds to the protonated product **8** either as **QxCav** or **QxCage**, since the two compounds are constitutional isomers.

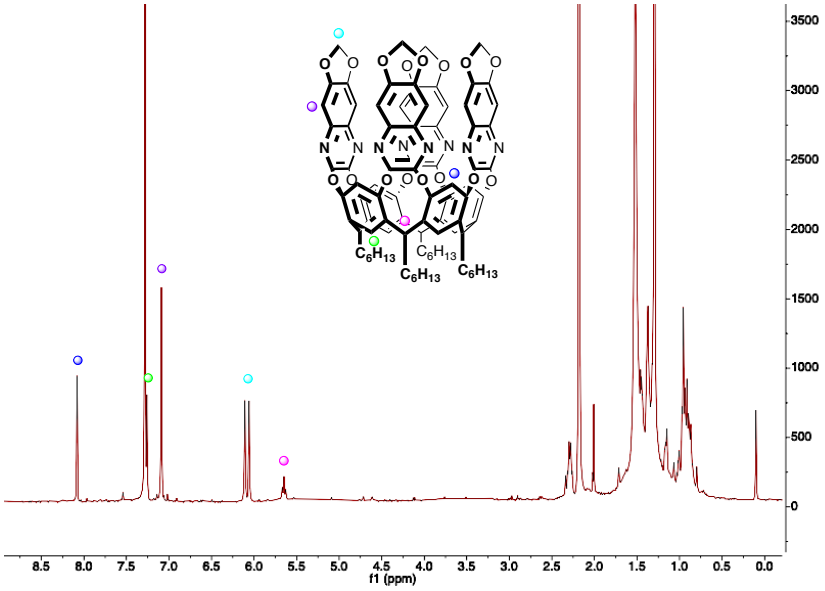


Figure 5.7. ^{1}H NMR of the purified compound deriving from the bridging reaction with $BrClCH_2$. $CDCl_3$, 25 °C, 400 MHz.

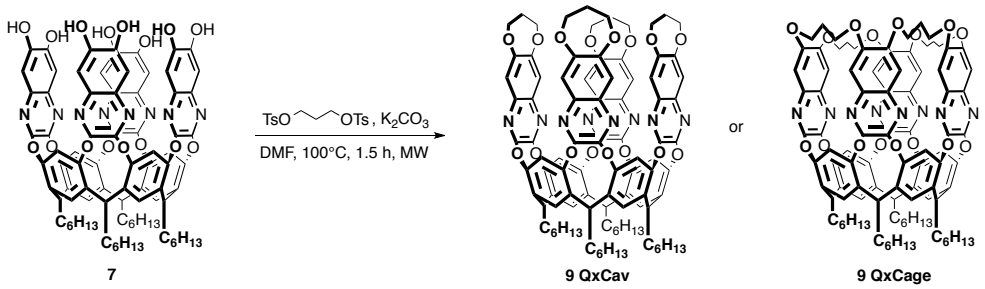
By analysing the NMR spectrum, the C_{4v} symmetry of both cavitands does not allow to distinguish between them on the basis of resonances number. Neither the four quinoxaline signals at 7.09 ppm, nor the resorcinarene protons at 8.06 and 7.24 ppm and the methine triplet at 5.59 ppm are diagnostic of the formation of one of the two compounds.

The only potentially diagnostic signals are the ones related to the methylene linker. The eight protons of the upper rim methylene units display two peaks at 6.12 and 6.09 ppm, indexed with a light blue dot in the spectrum. The appearance of two peaks indicates that the inward and outward facing protons are not equivalent, which however is common to both structures.

As in the case of deep QxCav (**Figure 5.5**), only the solid-state analysis can reveal the correct structure of the cavitand. Nevertheless, according to the results shown in the next paragraph, the **8 QxCav** structure is the most probable one.

5.2.2 PROPYLENE LINKER VERSUS 7-MEMBER RING AT THE UPPER RIM

It is known that cyclization reactions occur mainly when 5 or 6 atoms rings are formed, while the 7-member rings are less likely to be formed. Therefore we moved to 1,3-propanediol di-p-tosylate as a linker in the bridging reaction to try to prepare **9 QxCage** (**Scheme 5.3**).



Scheme 5.3. Possible overcome of the reaction with 1,3-propanediol di-p-tosylate.

Maintaining the reaction conditions as the previous experiment, DMF was used as solvent and K_2CO_3 as deprotonating agent. The reaction yielded a single product in 22% yield.
The MALDI-TOF spectrum is shown in **Figure 5.8**: the experimental $[M+X]^+$ ions in the upper part are respectively the proton (green), sodium (orange) and potassium (red) adducts; below the corresponding theoretical expected masses, which nicely match the found ones both as values and isotopic distribution.

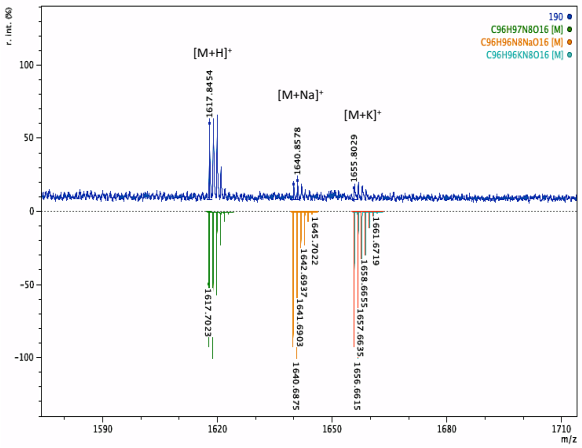


Figure 5.8. MALDI-TOF analysis of the isolated cavitand deriving from the bridging reaction with 1,3-Propanediol di-p-tosylate.

^{1}H NMR analysis confirms that the bridging reaction has occurred, but it does not prove which of the two C_{4v} symmetry cavitands **9** is formed. Signals of the aromatic protons of the resorcinarene scaffold are present at 8.01 and 7.15 ppm associable to the upper rim and lower rim protons, respectively, while, indicated with a purple dot, quinoxaline protons appear at 7.38 ppm. The methine proton present at 5.32 ppm (pink dot) demonstrates the *vase* conformation adopted by the cavitand after the bridging reaction, as shown also for the methylene-linked cavitand **8**. Finally, the peaks at 4.23 and 2.22 ppm indicate the presence of the propylene linkers. Unfortunately, also in this case the latter resonances are not diagnostic to distinguish between **9 QxCav** and **9 QxCage**.

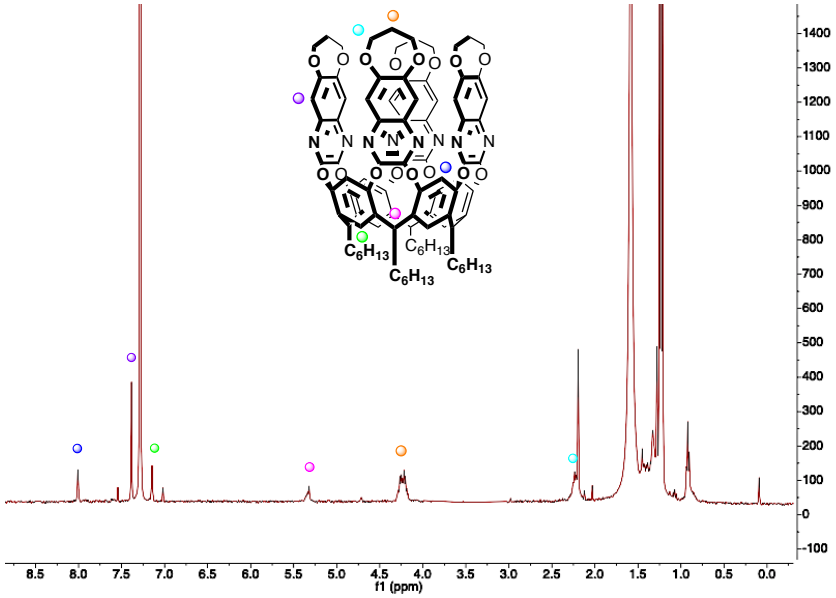


Figure 5.9. ^{1}H NMR of the cavitand from the bridging reaction with 1,3-Propanediol di-p-tosylate. $CDCl_3$, 25 °C, 400 MHz.

The solid-state structure of the synthesized cavitand was finally revealed by X-ray analysis on single crystals grown from a chloroform solution, which confirmed the formation of the isomer **9 QxCav** (**Figure 5.10**).
9 QxCav crystallizes in the space group *C2/c*; its asymmetric unit consists of half of the cavitand while the whole molecule is generated by rotation around a 2-fold axis. The four quinoxaline moieties bearing an additional 7-membered ring each form the walls of the cavitand. The distance between the mean-planes passing through the four C23 atoms (**Figure 5.10 c**) and the eight resorcinarene oxygen atoms is of 8.851(5) Å. Two of the walls parallel to each other are slightly inclined towards the inside of the cavity, as can be inferred by the distances C23A••• C23A^{i} and C23B•••C23B^{i} of 7.898(4) and 8.101(4) Å, respectively. (i = 1-x, y, ½-z)

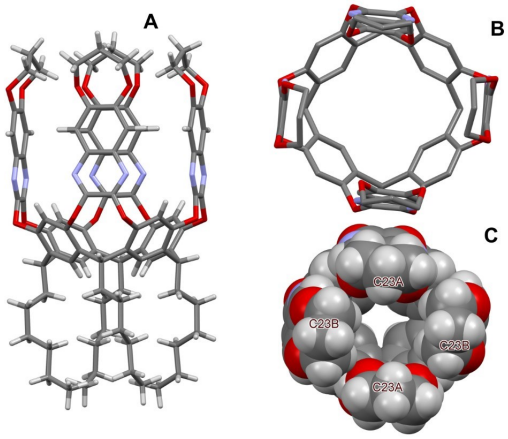
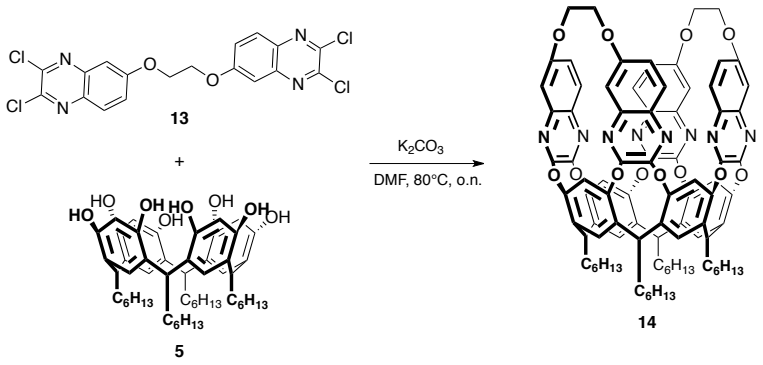


Figure 5.10. Side (A) and top (B) view of the molecular structure of **9 QxCav**. H atoms (B) and alkyl chains (B,C) have been omitted for clarity. Atoms displaying the same label are related by the symmetry operation 1-x, y, ½-z.

5.3 DIBRIDGED QUINOXALINE CAVITANDS

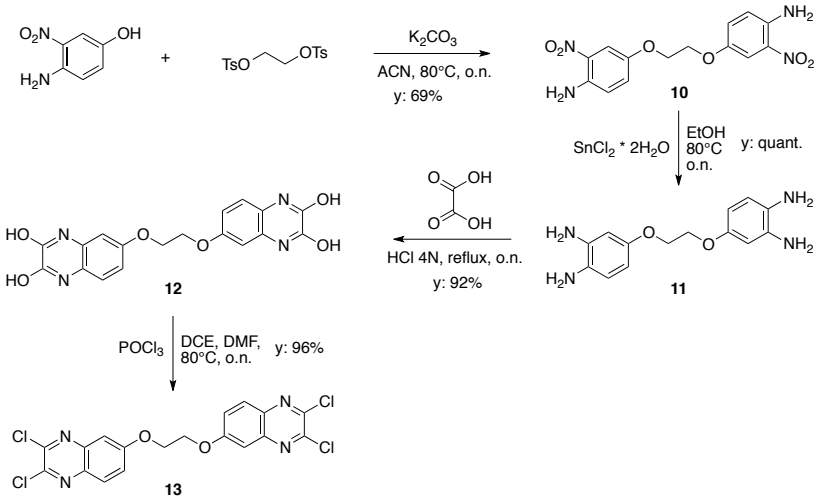
Taking into account the previous studies on this class of macrocyclic receptors, a new convergent synthesis was designed in order to obtain the di-bridged cavitand **14**, featuring two ethylendioxy tethers at the upper rim instead of four. In this way the vase-kite interconversion can be at least partially reduced with respect to the parent **QxCav**, favouring the desired vase form.

Scheme 5.4 describes the new synthons required for this convergent approach.



Scheme 5.4. Designed approach to obtain di-bridged cavitand **14**.

The benefit of this approach is the elimination of the final bridging reaction, which proved to be elusive in the tetra-bridged cases discussed above. The preparation of bis-quinoxaline **13** moves the bridging reaction at the beginning of the synthesis, where it can be controlled. Two molecules of **13** are then directly installed on resorcinarene **5** to give the desired cavitand **14**.



Scheme 5.5. Synthetic pathway for synthon **13**.

The quinoxaline building block **13** was achieved *via* a multistep synthesis with an overall yield of 61%, as shown in **Scheme 5.5** above.
The first synthetic step involves the bridging of two molecules of 3-amino-2-nitro phenol with ethylene di(p-toluenesulfonate) to give the linked diamino-dinitro intermediate **10**. For the reduction of the diamino compound **10**, different pairs of catalysts/reducing agents were tested as well as different reaction conditions. $SnCl_2$ turned out to be the best reducing agent. Because of the high reactivity of the amino groups towards oxidation, it was decided to skip the purification step, keeping the reaction mixture in an inert atmosphere, and subsequently adding oxalic acid to the crude **11,** leading to heterocyclic compound **12** *via* condensation reaction. Compound **12** was obtained in analytically pure form after precipitation from water. The final step, the chlorination of **12**, was performed using $POCl_3$ in DCE to give synthon **13**.
In parallel, resorcinarene **5** was obtained from the standard condensation of resorcinol and heptaldehyde under acidic conditions.

The formation of the quinoxaline cavitand of interest was attempted using two equivalents of quinoxaline **13**, in presence of K_2CO_3, i.e. the standard quinoxaline bridging procedure which typically proceeds in high yield.

In the present case, only trace amounts of a cavitand (evidenced in the MALDI-TOF spectrum of the crude) were produced, insufficient to be purified.
A $[M+H]^+$ molecular ion compatible with the di-bridged cavitand was detected in small amounts (**Figure 5.11** below)

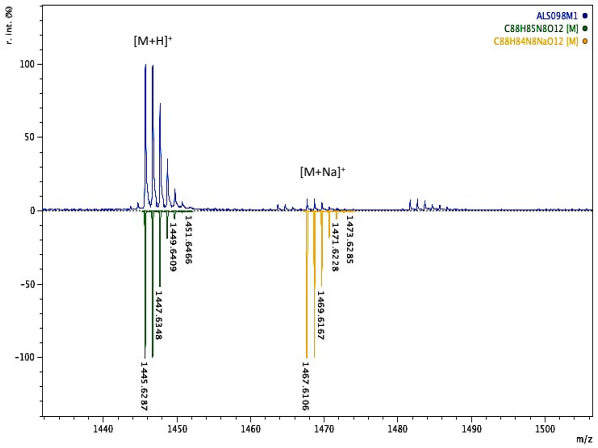
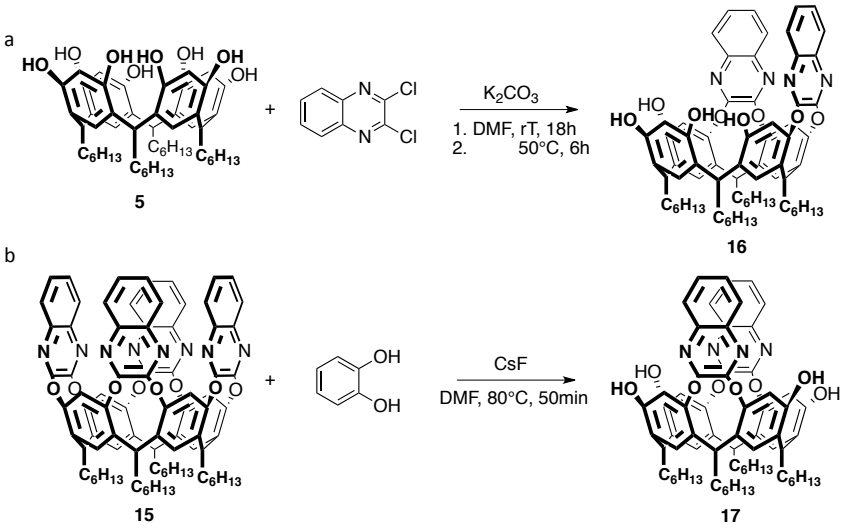


Figure 5.11. MALDI-TOF spectrum of the crude of cavitand **14** synthesis.

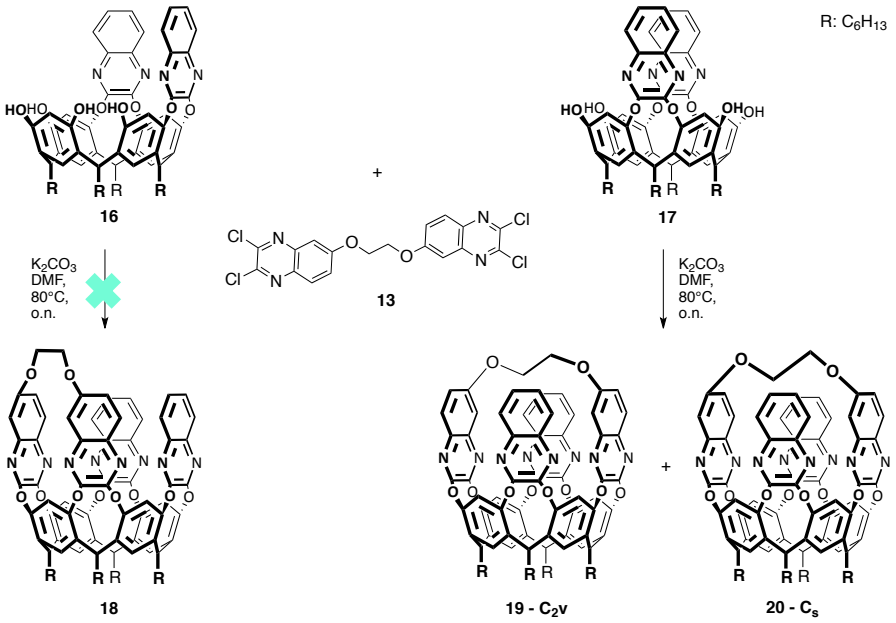
The difficulties found to bridge these two synthons together, suggested to simplify the system by introducing a single synthon **13**, instead of two, on an already di-bridged cavitand.

As reported in the literature, partially bridged cavitands can be prepared from the corresponding resorcinarene in two different ways depending on the number of quinoxaline walls desired on the scaffold. The first method (**a** in **Scheme 5.6**) for the synthesis of doubly-bridged cavitands, implies the use of 2 eq. of 2,3-dichloroquinoxaline, resulting in a mixture of several products from which the desired resorcinarene **AB-2QxCav 16** can be isolated as major product.[13] The other method (**b** in **Scheme 5.6**) provides mainly **AC-2QxCav 17** by selective removal of two diazanaphtalene walls from **QxCav 15**, *via* excision reaction with a proper amount of catechol under basic conditions.[14]



Scheme 5.6.Synthesis of di-quinoxaline resorcinarenes **16** and **17**.

Bis-quinoxaline **13** was then reacted with partially bridged cavitands **16** and **17**. In both cases, the bridging reaction was performed in DMSO, in the presence of potassium carbonate. Different strategies were pursued, since it is known that the use of microwave irradiation for these bridging reactions increases yields and reduces reaction times. For **AB-2QxCav 16** the starting material **16** was recovered in all cases (**Scheme 5.7**, left).
By performing the same reaction with **AC-2QxCav 17** two different isomers were isolated (**Scheme 5.7**, right) after extensive column chromatography and several preparative TLCs, due to the very close R_f.



Scheme 5.7. Bridging reactions carried out on **AB-2QxCav 16** and **AC-2QxCav 17**.

The two products turned out to be: (a) tetraquinoxaline cavitand **19 – C_2v** in which the quinoxaline linker crosses the cavity providing a disymmetric cavitand with C_2v symmetry (one C_2 axis and a σ plane containing the C_2 axis) and (b) tetraquinoxaline cavitand **20 – C_s** bearing the linker on the same side of the scaffold affording an asymmetric cavitand with symmetry C_s (only a σ plane).

MALDI-TOF analyses proved that **19 – C_2v** and **20 – C_s** have the predicted molecular weight and are isomers (**Figure 5.12**). On the left, MALDI-TOF spectrum of **19 – C_2v** is reported showing a protonated product at 1387.6254 $[M+H]^+$ and the $[M+NH_4]^+$ molecular ion at 1409.6063, in line with the calculated value and isotopic distribution. Similarly, MALDI-TOF spectrum of **20 – C_s** on the right, $[M+H]^+$ and $[M+NH_4]^+$ molecular ion peaks are present at 1387.6268 and 1409.6017 Da.

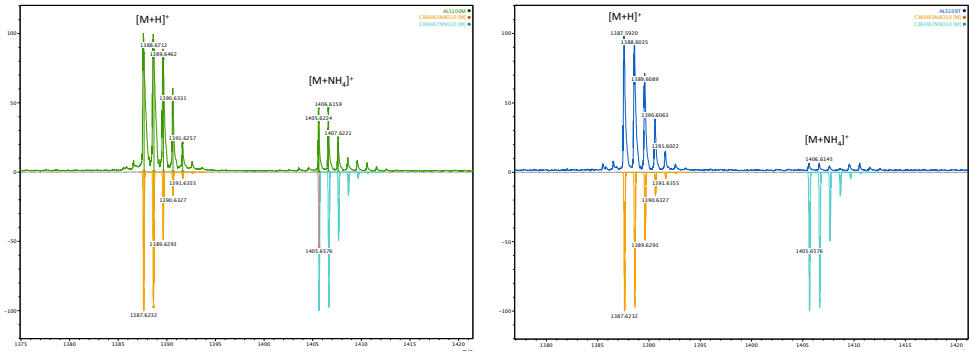


Figure 5.12. MALDI-TOF spectra of the isomers **19 – C_{2v}** (left) and **20 – C_s** (right).

The distinction of the two isomers was possible following the structure determination of **20 – C_s** by X-ray analysis on single crystals grown from DMSO, and was confirmed by ^{1}H NMR.

The molecular structure of the disymmetric linker bridged on cavitand **20 – C_s** is shown in **Figure 5.13**. **20 – C_s** crystallizes in the space group *C2/c* and the asymmetric unit comprises half of the molecule of the cavitand and two half molecule of DMSO. Also in this case the whole molecules are generated by rotation around a 2-fold axis. As a result, the ethylendioxy bridge is disordered over two equivalent positions, as are the two DMSO molecules which are located in the cavity and among the four alkyl feet.

The quinoxaline walls bridged by the ethylendioxy moiety are bent towards the inside of the cavity. This is evidenced by the C19B•••C20B^i (and the equivalent C20B•••C19B^i) distances of 4.714(5) Å (i = 1-x, y, ½-z) and by an inclination angle of 67.60(8) °. This latter is the angle between the mean planes passing though the walls and the mean plane passing though the eight O1/O2 oxygen atoms.

By comparison, the correspondent values for the other pair of quinoxaline walls are 10.377(4) Å for the C19A•••C20A^i distance and 89.49(7) ° for the inclination angle. Due to the geometry of the ethylendioxy bridge which leans more towards one of the two lateral walls, the opening of the cavity is asymmetric, as clearly shown by the space filling view in **Figure 5.13 c**.

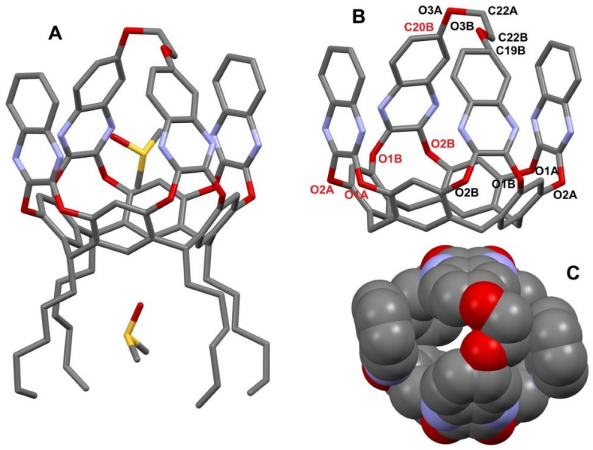


Figure 5.13. Molecular structure of **20 – C_s** (A). Detail of the cavity with partial labelling scheme (B) and space filling view of the cavity (C). H atoms have been omitted for clarity. Labels coloured in red refers to atoms generated by symmetry (1-x, y, ½-z).

The different symmetry of the two cavitands is reflected in the number of resonances of the methine protons of the resorcinarene skeleton in the ^{1}H NMR spectrum. Cavitand **19 – C_2v** presents two methine resonances at 5.6-5.7 ppm (pink dot - **Figure 5.14**) in 1:1 ratio, in agreement with the symmetry of its structure, while **20 – C_s** shows three resonances in 1:1:2 ratio (pink dot – **Figure 5.15**) in the same ppm range. The chemical shift of these resonances is diagnostic of the vase conformation for both cavitands in solution, also confirmed in the solid state for **20 – C_s** by crystal structure.

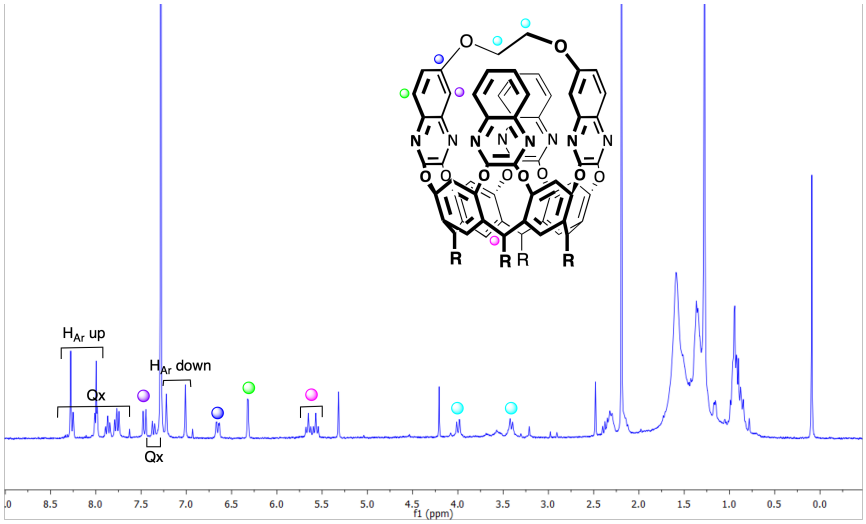


Figure 5.14. ^{1}H NMR spectrum of compound **19 – C_2v** in $CDCl_3$, 25 °C, 400 MHz.

In the low field region, from 6.0 to 8.3 ppm, signals related to the aromatic part are present. In particular, the resorcinarene aromatic protons are split in four singlets of which two are related to the benzene rings placed under the linker, while the quinoxaline protons are split in (a) the protons associated to non-functionalized 1,4-diazanaphalene walls and (b) the protons in *ortho* and *meta* positions with respect to the linker, highlighted respectively in blue, green and purple. The ethylene linker protons are present as doublets at 3.4 and 4.0 ppm respectively, showing the presence of a geminal coupling (**Figure 5.14**).
In the case of the C_s symmetry of compound **20 – C_s**, the ethylene bridge, indicated by light blue dots, appears as a pair of multiplets between 3.3 and 3.6 ppm (**Figure 5.15**). The resorcinarene aromatic protons are split in four singlets as for the **19 – C_2v** case, as expected by symmetry considerations.

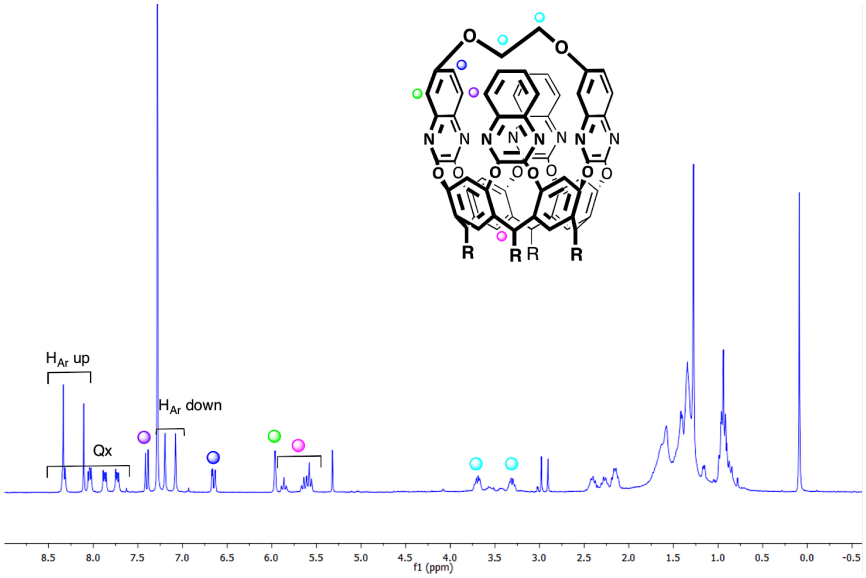
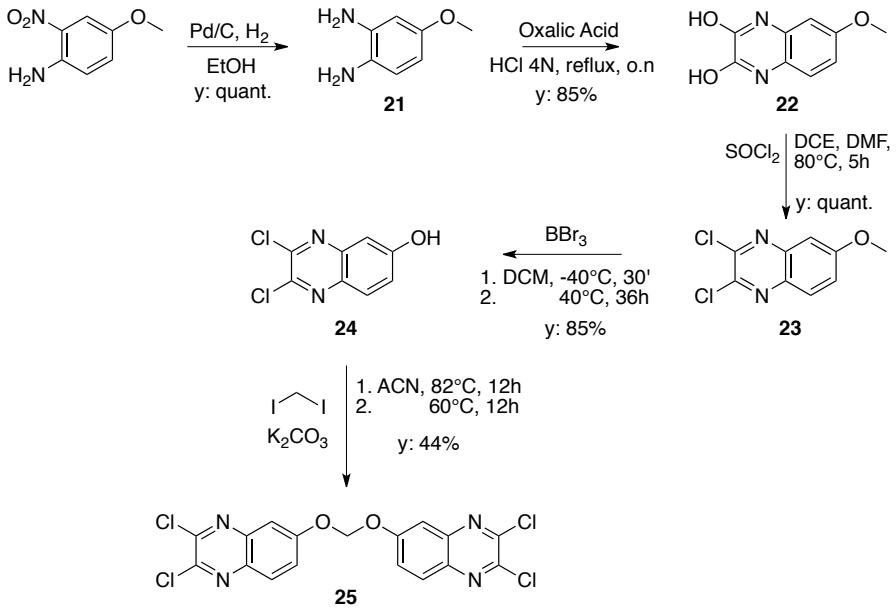


Figure 5.15. ^{1}H NMR spectrum of compound **20 – C_s** in $CDCl_3$, 25 °C, 400 MHz.

Taken together, these results provide important insight into the design of a linker of appropriate length. Both isomers are the result of a first attach of the dimeric compound **13** on **AC-2QxCav 17**, followed by reaction of the remaining portion either in a parallel or antiparallel mode on the remaining OHs. The presence of both AC isomers **19 – C_2v** and **20 – C_s** is further evidence that the tether is too long for the desired AB junction.

5.4 METHYLENEDIOXY VERSUS ETHYLENEDIOXY LINKER

The results of the bridging reactions described in the previous paragraph indicate that the ethylendioxy linker is too long for the AB-bridging reaction. Therefore we decided to maintain the same synthetic strategy, reducing the linker length by one carbon atom. Accordingly, a new methylene-diquinoxaline moiety was designed following the synthetic pathway presented below.



Scheme 5.8. Synthesis of methylendioxy linked bis-quinoxaline **25**.

The synthesis of chlorinated bis-quinoxaline **25** was carried out in five steps. The first one consisted in the reduction of the nitro group of 2-amino-5-methoxy nitrobenzene in amino group by hydrogenation catalysed with 10% Pd/C, to obtain 4-methoxy-1,2-diamino benzene **21**. Subsequently, the cyclization of **21** was performed with oxalic acid in an acid environment to form the second aromatic cycle via condensation. Thus, the desired product **22** was obtained with an 85% yield. Compound **22** was then chlorinated in anhydrous 1,2-dichloroethane with $POCl_3$ in the

presence of a catalytic amount of anhydrous DMF to give 2,3-dichloro-6-methoxyquinoxaline **23** in quantitative yield.
Product **25** was obtained by deprotection of the hydroxyl group in **23** using BBr_3 to give **24**, which was subsequently employed for the linking reaction with diiodomethane under basic conditions. Compound **25** was obtained in 44% yield after purification.

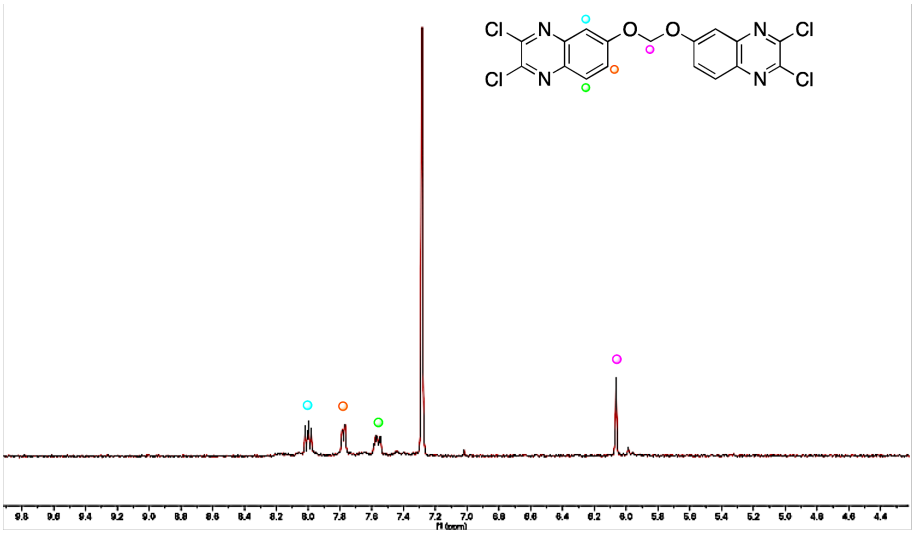


Figure 5.16. ^{1}H NMR spectrum of compound **25.** $CDCl_3$, 25 °C, 400 MHz.

Figure 5.16 shows the ^{1}H NMR spectrum of **25** in $CDCl_3$, in which the three low field multiples belong to the aromatic protons, while the singlet displayed with a pink dot is the resonance of the protons of the methylene linker.

Bis-quinoxaline **25** was reacted both with **AB-2QxCav 16** and **AC-2QxCav 17** to form the corresponding cavitands.
However, in all cases the degradation of the methylene spacer, was the only reaction observed. The expected cavitands were not detected, even in trace amounts. It turned out that the methylendioxy linker is prone to decomposition which hampers its use in the synthesis of the corresponding cavitands. This last result demonstrates that simple alkyl linkers are better suited than alkoxy ones for this type of bridging reactions, even if more difficult to prepare.

5.5 CONCLUSIONS

Four new quinoxaline cavitands possessing a narrow, rigid cage, potentially accessible only by benzene, have been designed. Two of them, namely **8 QxCage** and **9 QxCage**, require a four-fold bridging reaction to be formed. Their preparation turned out to be elusive. They did not form because the bridging reaction among the four quinoxaline walls is disfavoured with respect to the ring closure to form 5 and 7-member rings at the upper rim to give **8 QxCav** and **9 QxCav**. The determination of the correct structure in these cases required solving the crystal structure, since NMR and MS do not suffice to determine the molecular structure.

Failing the first tightening strategy, partial rigidification of the cavity was attempted by bridging resorcinarenes with two bis-quinoxaline synthons. In this way, the connection between the quinoxaline walls is attained before their instalment on the resorcinarene skeleton. Only traces of the desired product were obtained suggesting further investigations.

A single bis-quinoxaline synthon **13** was reacted with both **AB-2QxCav 16** and **AC-2QxCav 17** to simplify the system. In this way only one di-bridged wall can be introduced, either in the AB or AC mode. With distal **AC-2QxCav 17** two isomeric cavitands **19 – C_{2v}** and **20 – C_s** were obtained. The determination of their molecular structures was accomplished by solving the crystal structure of **20 – C_s** followed by that of **19 – C_{2v}** via ^{1}H NMR comparison. The length of the ethylendioxy linker prevents the vicinal bridging reaction on **AB-2QxCav 16** leading to the recovery of the unreacted reagents.

The outcome of the above bridging reaction suggested to shorten the length of the linker to one carbon atom. Bis-quinoxaline synthon **25** possessing a methylendioxy spacer was therefore synthetized and reacted with **AB-2QxCav 16**. Unfortunately, degradation of the methylene spacer was the only reaction observed.

Overall, the results of this synthetic work indicate that the best potential synthon for a rigidly blocked cavitand is a bis-quinoxaline molecule with an purely alkyl ethylene spacer.

5.6 EXPERIMENTAL SECTION

1,2-dimethoxy-4,5-dinitro benzene (1).
An aqueous solution of HNO_3 65% (5 mL) was added dropwise to solid 1,2-dimethoxy benzene (500 mg, 3.62 mmol). After 1 hr at rT, a yellow precipitate started to form and the reaction was heated at 120 ^oC and stirred for 8 hrs. The suspension was cooled down and poured in ice-cooled water, filtered and dried under vacuum. Yellow product **1** was obtained by recrystallization from glacial acetic acid (677 mg, 82%).
^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.38 (s, 2H, Ar**H**), 4.07 (s, 6H, -O**CH_3**). **ESI-MS:** m/z 229.15 $[M+H]^+$.

1,2-dimethoxy-4,5-diamino benzene (2).
Compound **1** (675 mg, 2.96 mmol) was dissolved in absolute ethanol (10 mL) and a catalytic amount of palladium on charcoal (10%, w/w) was added. The suspension was introduced in a PARR reactor and was put under an hydrogen atmosphere (5 bar) at room temperature for 24 hrs. The crude was filtered through a celite pad, washed with ethanol and concentrated under reduced pressure obtaining **2** as a brown solid (496 mg, quant.).
^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 6.42 (s, 2H, Ar**H**), 3.84 (s, 6H, -O**CH_3**), 3.28 (bs, 4H, Ar**NH_2**). **ESI-MS:** m/z 169.27 $[M+H]^+$.

6,7-dimethoxyquinoxaline-2,3-dione (3).
Compound **2** (495 mg, 2.93 mmol) was dissolved in 5 mL of HCl 4 N. Oxalic acid (316 mg, 3.51 mmol) was added and the mixture was refluxed for 16 hrs at 100 °C. After cooling to rT, the precipitate obtained was filtered, washed with water and dried under vacuum to give **3** in pure form (514 mg, 79%).
^{1}H NMR (DMSO-d_6, 400 MHz): δ ppm = 11.78 (s, 2H, -O**H**) 6.77 (s, 2H, Ar**H**), 3.71 (s, 6H, -O**CH_3**). **ESI-MS:** m/z 223.24 $[M+H]^+$.

2,3-dicloro-6,7-dimethoxyquinoxaline (4).
To 6,7-dimethoxyquinoxaline-2,3-dione **3** (514 mg, 2.31 mmol) dissolved in 4 mL of 1,2-dichloroethane containing a catalytic amount of dry DMF, $POCl_3$ (3.46 mL, 37.01 mmol) was added and the reaction was stirred at 80^{o}C for 16 hrs. After solvent removal under high vacuum, the crude was

triturated in MeOH and filtered giving the target compound **4** (509 mg, 85%).
^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.28 (s, 2H, Ar**H**), 4.07 (s, 6H, -O**CH_3**). **ESI-MS:** m/z 259.74 $[M+H]^+$.

Resorcinarene [C_6H_{13}; H] (5).
Resorcinol (15.0 g, 136 mmol) was dissolved in an ice-cooled round-bottom flask containing MeOH (150 mL). Subsequently, heptaldehyde (19 mL, 136 mmol) and 22 mL of HCl 37% were added dropwise. After the additions, the reaction mixture was stirred at 50 °C for 5 days. The crude was quenched with water, filtered, dried under vacuum and recrystallized three times from MeOH. The desired product **5** was obtained as a fine beige powder (16.79 mg, 60%).
^{1}H NMR (Acetone-d_6, 400 MHz): δ ppm = 8.46 (s, 8H, Ar**OH**), 7.57 (s, 4H, Ar**H_{up}**), 6.25 (s, 4H, Ar**H_{down}**), 4.32 (t, 4H, -**CH**-, J= 7.8 Hz), 2.30 (m, 8H, CH**CH_2**-), 1.36 (m, 32H, -**CH_2**-), 0.90 (t, 12H, -**CH_3**, J= 6.6 Hz). **ESI-MS:** m/z 825.31 $[M+H]^+$; m/z 847.56 $[M+Na]^+$.

Octamethoxy quinoxaline cavitand (6).
In an oven-dried microwave vessel resorcinarene **5** (200 mg, 0.24 mmol), quinoxaline **4** (264 mg, 1.02 mmol) and dry K_2CO_3 (536 mg, 3.88 mmol) were added to 6 mL of dry DMF. The reaction was put under microwave irradiation at 120 ^oC for 2hrs, extracted with DCM/H_2O and dried over Na_2SO_4. After solvent removal under reduced pressure, the crude was purified by flash chromatography (SiO_2; DCM/MeOH 98:2) affording octamethoxy cavitand **6** (361 mg, 96%).
^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.59 (s, 4H, Ar**H_{up}**), 7.22 (s, 8H, **Qx**), 7.02 (s, 4H, Ar**H_{down}**), 4.59 (bt, 4H, -**CH**-), 3.97 (s, 24H, ArO**CH_3**), 2.12 (m, 8H, CH**CH_2**-), 1.27 (m, 32H, -**CH_2**-), 0.86 (t, 12H, -**CH_3**, J= 6.6 Hz). **MALDI-TOF:** m/z 1569.7906 $[M+H]^+$.

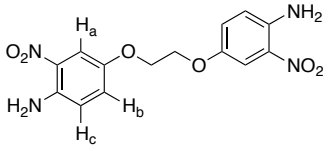
Octahydroxy quinoxaline cavitand (7).
In a cooling bath, cavitand **6** (300 mg, 0.19 mmol) was dissolved in 10 mL of dry DCM and BBr_3 (3.91 mL, 22.98 mmol) was added dropwise under Ar atmosphere. After stirring the reaction mixture at 50^{o}C for 24 hrs and cooling down to room temperature, 5 mL of 1N HCl were added to quench

the reaction. The precipitate obtained was filtrated, washed with H_2O and dried under vacuum, obtaining the final product (257 mg, 93%).
^{1}H NMR (DMSO-d_6, 400 MHz): δ ppm = 9.94 (s, 8H, ArO**H**), 7.85 (s, 4H, Ar**H_{up}**), 7.71 (s, 4H, Ar**H_{down}**), 7.08 (s, 8H, **Qx**), 5.39 (bt, 4H, -**CH**-), 2.38 (m, 8H, CH**CH_2**-), 1.38 (m, 8H, -**CH_2**CH_3), 1.25 (m, 24H, -**CH_2**-), 0.85 (t, 12H, -**CH_3**, J= 6.3 Hz). **MALDI-TOF:** m/z 1457.6612 $[M+H]^+$.

1,3-dioxolane-quinoxaline cavitand (8).
Cavitand **7** (50 mg, 0.03 mmol) was dissolved in 1 mL of dry DMF in an oven-dried microwave vessel and K_2CO_3 (125 mg, 0.90 mmol) was added under Ar atmosphere. After the solution turned pale pink, $BrClCH_2$ (67 µL, 1.03 mmol) was added and the reaction was put under microwave irradiation at 100 °C for 1 h. The crude was precipitate in H_2O, washed and filtered. Product **8** was obtained *via* preparative TLC (SiO_2, DCM/MeOH 98:2) as a yellowish powder (7 mg, 14%).
^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 8.06 (s, 4H, Ar**H_{up}**), 7.24 (s, 4H, Ar**H_{down}**), 7.09 (s, 8H, **Qx**), 6.12 (s, 4H, ArO**CH_{2in}**O), 6.09 (s, 4H, ArO**CH_{2out}**O), 5.59 (bt, 4H, -**CH**-), 2.28 (m, 8H, CH**CH_2**-), 1.34 (m, 8H, -**CH_2**CH_3), 1.28 (m, 24H, -**CH_2**-), 0.93 (bt, 12H, -**CH_3**). **MALDI-TOF:** m/z 1505.5167 $[M+H]^+$.

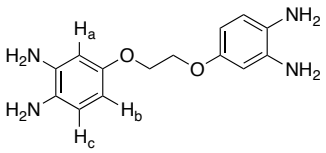
1,4-dioxepane-quinoxaline cavitand (9).
Cavitand **7** (50 mg, 0.03 mmol) was dissolved in 2 mL of dry DMF in an oven-dried microwave vessel and K_2CO_3 (125 mg, 0.90 mmol) was added under an Ar atmosphere. After the solution turned pale pink, 1,3-Propanediol di-*p*-tosylate (396 mg, 1.03 mmol) was added and the reaction was put under microwave irradiation at 100 °C for 1.5 hrs. The crude was precipitate in H_2O, washed and filtered. The target molecule **9** was obtained *via* preparative TLC (SiO_2, DCM/MeOH 98:2) as a yellowish powder (12 mg, 22%).
^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 8.01 (s, 4H, Ar**H_{up}**), 7.38 (s, 8H, **Qx**), 7.15 (s, 4H, Ar**H_{down}**), 5.32 (bt, 4H, -**CH**-), 4.23 (m, 16H, ArO**CH_2**CH_3-), 2.38 (m, 8H, CH**CH_2**-), 2.22 (m, 8H, ArOCH_2**CH_3**-), 1.38 (m, 8H, -**CH_2**CH_3), 1.32 (m, 24H, -**CH_2**-), 0.92 (bt, 12H, -**CH_3**). **MALDI-TOF:** m/z 1617.8454 $[M+H]^+$; m/z 1640.8578 $[M+Na]^+$; m/z 1655.8029 $[M+K]^+$.

4,4'-(ethane-1,2-diylbis(oxy))bis(2-nitroaniline) (10).



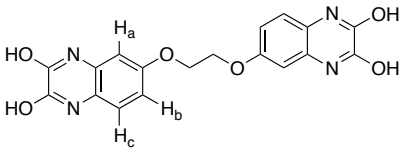
To a solution of 4-hydroxy-2-nitroaniline (500 mg, 3.24 mmol) in dry ACN (15 mL) was added K_2CO_3 (1.345 g, 9.73 mmol). After 5 min, ethylene di(p-toluene sulfonate) was added (1.803 g, 4.87 mmol) and the reaction mixture was stirred at 80°C for 24 h. The solvent was removed and water was added. The crude was sonicated for 15 min and filtered through a Buchner funnel to afford a red solid. The target compound was obtained after a recrystallization from DCM as a dark red powder (374 mg, 69%).
^{1}H NMR ($CDCl_3$, 300 MHz): δ ppm = 7.63 (d, 2H, **H_a**, J = 2.8 Hz), 7.13 (dd, 2H, **H_b**, J_o = 6.0 Hz, J_m = 3.0 Hz), 6.77 (d, 2H, **H_c**, J = 9.0 Hz), 4.29 (s, 4H; -O**CH_2CH_2**O-). **ESI-MS:** m/z 357.19 $[M+Na]^+$; m/z 373.26 $[M+K]^+$.

4,4'-(ethane-1,2-diylbis(oxy))bis(benzene-1,2-diamine) (11).



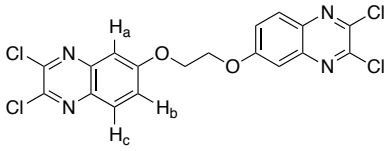
To a solution of compound **10** (374 mg, 1.12 mmol) in EtOH (40 mL) was added $SnCl_2$ * $2H_2O$ (2.52 g, 11.16 mmol) in one portion and the solution was stirred at 80°C for 17 h. The solvent was evaporated in vacuum to give a dark solid, which was maintained under inert atmosphere and carried to the next step without further purification (307 mg, quant.).

6,6'-(ethane-1,2-diylbis(oxy))bis(quinoxaline-2,3-diol) (12).



To a solution of compound **11** (307 mg, 1.12 mmol) in HCl 4N (18 mL) was added oxalic acid (242 mg, 2.69 mmol) and the mixture was stirred at 110°C for 16 h. The reaction mixture was filtered through a Buchner funnel, washed with water and the product was dried in vacuum to give a dark purple solid (394 mg, 92%).
^{1}H NMR (DMSO-d_6, 300 MHz): δ ppm = 11.83 (d, 2H, O**H,** J = 16 Hz), 7.05 (d, 2H, **H_a,** J = 8.7 Hz), 6.78 (dd, 2H, **H_b,** J_o = 11.3 Hz, J_m = 2.6 Hz), 6.72 (s, 2H, **H_c,** J = 2.5 Hz), 4.25 (s, 4H, -O**CH_2CH_2**O-).

1,2-bis((2,3-dichloroquinoxalin-6-yl)oxy)ethane (13).



To a solution of **12** (200 mg, 0.52 mmol) in dry DCE (30 mL) were added dry DMF (5 drops) and $POCl_3$ (1.95 mL, 20.90 mmol). The mixture was stirred at 80°C for 16 h, then cooled at room temperature and the solvent was removed under vacuum. The crude was suspended in MeOH and filtered through a Buchner funnel to give a brownish solid (490 mg, 96%).
^{1}H NMR ($CDCl_3$, 300 MHz): δ ppm = 7.93 (d, 2H, **H_a,** J = 9.2 Hz), 7.49 (dd, 2H, **H_b,** J_o = 9.3 Hz, J_m = 2.7 Hz), 7.38 (d, 2H, **H_c,** J = 2.8 Hz), 4.55 (s, 4H; -O**CH_2CH_2**O-). **ESI-MS:** m/z 456.97 $[M+H]^+$.

Attempted synthesis of di(ethylene-linked quinoxaline) cavitand (14).
After dissolving resorcinarene **5** (55 mg, 0.06 mmol) in dry DMF (5 mL), K_2CO_3 (146 mg, 1.05 mmol) and quinoxaline **13** (60 mg, 0.13 mmol) were added. and the reaction mixture was stirred at 80 °C for 16 hrs. After solvent removal, the crude was washed with DCM/H_2O, dried over Na_2SO_4

and DCM removed. A beige solid was obtained after precipitation in DCM/Hex (x3 times) and trituration in Et_2O (x3 times). The crude was analyzed via MALDI-TOF. The cavitand is present only in trace amounts.
MALDI-TOF: m/z 1446.30 $[M+H]^+$.

4QxCav (15).
To a solution of resorcinarene **5** (1.0 g, 1.21 mmol) dissolved in dry DMF, 2,3- dichloroquinoxaline (1.06 g, 5.33 mmol) and K_2CO_3 (2.51 g, 18.2 mmol) were added and the reaction heated at 80 °C for 24 hrs. Subsequently, the solvent was removed and the precipitate was triturated in H_2O, filtered, washed with water, and dried to afford the pure product as an off white powder (1.61 g, quant.).
^{1}H NMR ($CDCl_3$, 300 MHz): δ ppm = 8.17 (s, 4H, Ar**H$_{up}$**), 7.79 (m, 8H, Qx**H**) 7.49 (m, 8H, Qx**H**), 7.22 (s, 4H, Ar**H$_{down}$**), 5.54 (t, 4H, J = 7.8 Hz, -C**H**-), 2.29 (m, 8H, CHC**H$_2$**-), 1.42 (m, 32H, -C**H$_2$**-), 0.96 (t, 12H, J = 6.7 Hz, -**CH$_3$**).
MALDI-TOF: m/z 1329.6133 $[M+H]^+$.

AB-di quinoxaline resorcinarene (16).
To a degassed solution of **5** (1.24 g, 1.50 mmol) in 20 mL of DMSO, K_2CO_3 (208 mg, 1.50 mmol) and 2,3-dichloroquinoxaline (598 mg, 3.00 mmol) were added and the mixture was stirred for 1h at rT. More K_2CO_3 (415 mg, 3.00 mmol) was added, and stirring was continued for 18h at rT and 6h at 50 °C. After cooling, 30 mL of H_2O were added to the brown solution and the pH was adjusted to 7 by addition of 1N HCl. The pink precipitate formed was isolated by filtration, washed with H_2O and dried in vacuum. Cavitand **16** was recovered as a white powder after flash column chromatography purification (SiO_2, DCM/AcOEt 85:15) (637 mg, 19.6%).
^{1}H NMR (Acetone-d_6, 300 MHz): δ ppm = 8.79 (bs, 4H, ArO**H**), 8.36 (s, 1H, Ar**H$_{up}$**), 8.10 (d, 2H, Qx**H**, J= 7.5 Hz), 7.95 (s, 1H, Ar**H$_{up}$**), 7.83 (m, 4H, Qx**H**), 7.75 (s, 2H, Ar**H$_{up}$**), 7.68 (d, 2H, Qx**H**, J= 6.96 Hz), 7.50 (s, 1H, Ar**H$_{down}$**), 7.15 (s, 2H, Ar**H$_{down}$**), 6.20 (s, 1H, Ar**H$_{down}$**), 5.65 (t, 2H, -C**H**- below Qx's, J= 8.1 Hz), 4.28 (t, 2H, -C**H**- below OH's, J= 7.9 Hz), 2.44 (m, 4H, CHC**H$_2$**- below Qx's), 2.29 (m, 4H, CHC**H$_2$**- below OH's), 1.36 (m, 32H, -C**H$_2$**-), 0.89 (m, 12H, -C**H$_3$**). **ESI-MS:** m/z 1077.5752 $[M+H]^+$; m/z 1099.5563 $[M+Na]^+$.

AC-di quinoxaline resorcinarene cavitand (17).
To a solution of cavitand **15** (500 mg, 0.38 mmol) in DMF, CsF (1.154 g, 7.60 mmol) and catechol (133 mg, 1.20 mmol) were added. The suspension was heated at 80 °C and stirred for 1 h, then the reaction was poured into 500 mL of ice-cold brine. The precipitate was filtered, washed with water, and dried. The crude product was purified by flash column chromatography (SiO_2, gradient from DCM/AcOEt 95:5 to DCM/AcOEt 85:15) affording compound **17** as an off-white solid (219 mg, 54%).
^{1}H NMR (Acetone-d_6, 300 MHz): δ ppm = 9.03 (bs, 4H, ArO**H**), 7.77 (s, 4H, Ar**H$_{down}$**), 7.54 (m, 4H, Qx**H**), 7.24 (m, 4H, Qx**H**), 7.16 (s, 4H, Ar**H$_{up}$**), 5.52 (t, 2H, -C**H**- below Qx's, J= 8.0 Hz), 4.47 (t, 2H, -C**H**- below OH's, J= 7.6 Hz), 2.41 (m, 8H, CHC**H$_2$**-), 1.36 (m, 32H, -C**H$_2$**-), 0.92 (m, 12H, -C**H$_3$**). **ESI-MS:** m/z 1077.5741 $[M+H]^+$; m/z 1099.5561 $[M+Na]^+$.

19 – C_2v and **20 – C_s cavitands.**
AC-diquinoxaline cavitand **17** (53 mg, 0.05 mmol) was dissolved in 70 mL of DMF. After the addition of K_2CO_3 (41 mg, 0.3 mmol) and quinoxaline **13** (23 mg, 0.05 mmol), the reaction was stirred at 80 °C for 16h. The solvent was removed in vacuum, H_2O was added to the crude, sonicated and filtrated. Purification by flash column chromatography over silica gel and preparative TLCs using DCM/MeOH 98:2 afforded pure **20** (3 mg, 12%) followed by **19** (1.3 mg, 5%).

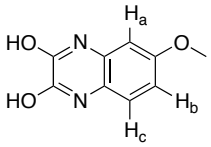
19. ^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 8.27 (s, 2H, Ar**H$_{up}$**), 8.25 (m, 2H, Qx**H**), 8.00 (m, 2H, Qx**H**), 7.98 (s, 2H, Ar**H$_{up}$**), 7.87 (t, 2H, Qx**H**, J= 7.03 Hz), 7.77 (t, 2H, Qx**H**, J= 6.90 Hz), 7.46 (d, 2H, Qx**H**, J= 9.12 Hz), 7.36 (d, 2H, Qx**H**, J= 8.07 Hz), 7.22 (s, 2H, Ar**H$_{down}$**), 7.01 (s, 2H, Ar**H$_{down}$**), 6.65 (dd, 2H, Qx**H**, J= 6.78), 6.32 (d, 2H, Qx**H**, J= 2.55 Hz), 5.65 (t, 2H, -C**H**- below linker, J= 8.11 Hz), 5.57 (t, 2H, -C**H**- below Qx's, J= 8.20 Hz), 3.99 (m, 2H, -OC**H$_2$**C**H$_2$**O-), 3.40 (m, 2H, -OC**H$_2$**C**H$_2$**O-), 2.33 (m, 4H, CHC**H$_2$**- below linker), 2.18 (m, 4H, CHC**H$_2$**- below Qx's), 1.35 (m, 32H, -C**H$_2$**-), 0.94 (m, 12H, -C**H$_3$**). **^{13}C NMR** ($CDCl_3$, 400 MHz): δ ppm = 154.11, 153.63, 152.85, 152.70, 152.57, 152.53, 151.59, 150.74, 140.50, 140.19, 139.45, 135.71, 135.53, 135.42, 135.11, 129.47, 129.13, 128.77, 128.48, 128.25, 124.06, 123.21, 122.84, 119.67, 109.24, 65.03, 34.41, 34.14, 33.99, 33.58, 32.00, 31.95, 31.84, 31.55, 30.72, 29.43, 29.37, 28.09, 27.94, 22.73, 22.68, 14.12, 14.07. **MALDI-TOF:** m/z 1387.6254 $[M+H]^+$; m/z 1409.6063 $[M+NH_4]^+$.

20. **^{1}H NMR** ($CDCl_3$, 400 MHz): δ ppm = 8.38 (s, 2H, Ar**H$_{up}$**), 8.33 (m, 2H, Qx**H**), 8.12 (s, 2H, Ar**H$_{up}$**), 8.03 (m, 2H, Qx**H**), 7.87 (m, 2H, Qx**H**), 7.73 (m, 2H, Qx**H**), 7.40 (d, 2H, Qx**H**, J= 9.21 Hz), 7.20 (s, 2H, Ar**H$_{down}$**), 7.08 (s, 2H, Ar**H$_{down}$**), 6.65 (dd, 2H, Qx**H**, J_1= 9.21 Hz, J_2= 2.58), 5.96 (d, 2H, Qx**H**, J= 2.58 Hz), 5.86 (t, 1H, -C**H**-, J= 8.25 Hz), 5.64 (t, 1H, -C**H**-) + 5.58 (t, 1H, -C**H**-, J= 8.32 Hz), 3.69 (m, 2H, -OC**H$_2$**C**H$_2$**O-), 3.30 (m, 2H, -OC**H$_2$**C**H$_2$**O-), 2.39 (m, 2H, CHC**H$_2$**-), 2.27 (m, 2H, CHC**H$_2$**-), 2.15 (m, 4H, CHC**H$_2$**-), 1.41 (m, 32H, -C**H$_2$**-), 0.99 (m, 12H, -C**H$_3$**). **^{13}C NMR** ($CDCl_3$, 400 MHz): δ ppm = 154.06, 153.60, 152.87, 152.66, 152.54, 152.52, 151.56, 150.72, 140.48, 140.16, 139.47, 135.78, 135.59, 135.49, 135.02, 129.47, 129.10, 128.76, 128.46, 128.26, 124.00, 123.20, 122.82, 119.64, 109.20, 64.98, 34.46, 34.12, 33.96, 33.57, 31.98, 31.93, 31.81, 31.52, 30.69, 29.41, 29.34, 28.04, 27.90, 22.69, 22.65, 14.10, 14.06. **MALDI-TOF:** m/z 1387.6268 $[M+H]^+$; m/z 1409.6017 $[M+NH_4]^+$.

4-Methoxy-1,2-benzenediamine (21).

To a solution of 4-methoxy-2-nitroaniline (1.01 g, 6.01 mmol) in EtOH (42 mL) and EtOAc (42 mL) 10% Pd/C (0.12 g) in small portions was added and the solution was hydrogenated in a PARR apparatus at room temperature for 5 hrs. The reaction mixture was filtered through a small pad of Celite, washed with EtOH, the solvent was evaporated in vacuum to give a dark solid, which was used in the next step without further purification (818 mg).

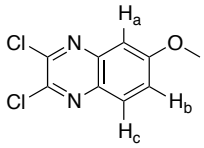
6-methoxyquinoxaline-2,3-diol (22).



To a solution of 4-methoxy-1,2-benzenediamine **21** (818 mg, 5.92 mmol) in HCl 4N (10 mL) was added a second solution of oxalic acid (586 mg, 6.51 mmol) in HCl 4N (10 mL) and the mixture was stirred at 100°C for 16 h. The reaction mixture was filtered through a Buchner funnel, washed with water and the product was dried in vacuum to give a grey solid (966 mg, 85%).

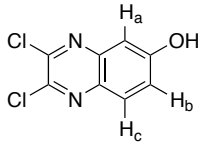
^{1}H NMR (DMSO-d_6, 300 MHz): δ ppm = 11.81 (bs, 2H, O**H**), 7.05 (d, 1H, **H_a**, J= 8.6 Hz), 6.71 (dd, 1H, **H_b**, J_o= 12.1 Hz, J_m= 2.5 Hz), 6.69 (s, 1H, **H_c**), 3.72 (s, 3H, -O**CH_3**). **^{13}C NMR** (75 MHz, DMSO-d_6): δ ppm = 155.94, 154.98, 126.90, 123.73, 119.90, 116.55, 110.03, 100.54, 55.88. **ESI-MS:** m/z 192.97 $[M+H]^+$.

2,3-dichloro-6-methoxyquinoxaline (23).



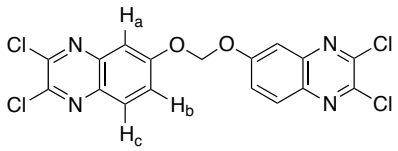
To a solution of **22** (965 mg, 5.02 mmol) in dry DCE (10 mL) dry DMF (1 mL) and $SOCl_2$ (4.37 mL, 60.26 mmol) were added. The mixture was stirred at 80°C for 3 h, then cooled at room temperature. The solvent was removed in vacuum to give a brownish solid (1.15 g, quant.).
^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.92 (d, 1H, **H_a**, J= 9.6 Hz), 7.46 (dd, 1H, **H_b**, J_o= 9.3 Hz, J_m= 2.8 Hz), 6.69 (d, 1H, **H_c**, J= 2.4 Hz), 3.98 (s, 3H, -O**CH_3**). **^{13}C NMR** ($CDCl_3$, 75 MHz): δ ppm = 161.79, 145.57, 142.43, 142.38, 136.60, 129.07, 124.34, 105.88, 55.98. **ESI-MS:** m/z 229.16 $[M+H]^+$.

2,3-dichloroquinoxalin-6-ol (24).



BBr_3 (6.28 mL, 36.54 mmol) was added to a solution of 2,3-dichloro-6-methoxyquinoxaline **23** (558 mg, 2.44 mmol) in 15 mL of dry DCM under Ar. The reaction mixture was stirred at 40°C for 16 hrs, cooled and hydrolysed with 6 mL of acidic MeOH. After the addition of water to the mixture, the product was collected via Buchner filtration as a brown solid (446 mg, 85%).
^{1}H NMR (Acetone-d_6, 300 MHz): δ ppm = 9.94 (s, 1H, O**H**), 7.94 (dd, 1H, **H_b**, J_o= 9.1 Hz, J_m= 3.4 Hz), 7.55 (d, 1H, **H_c**, J= 9.1 Hz), 7.31 (bs, 1H, **H_a**). **ESI-MS:** m/z 216.15 $[M+H]^+$; m/z 233.14 $[M+NH_4]^+$.

bis((2,3-dichloroquinoxalin-6-yl)oxy)methane (25).



Compound **24** (100 mg, 0.47 mmol) was dissolved in 2 mL of dry ACN and K_2CO_3 (193 mg, 1.40 mmol) was added, followed by diiodomethane (23 µL, 0.28 mmol). The reaction was stirred at 80 °C for 12h and at 60 °C for other 12h. After cooling, the solvent was partially removed under reduced pressure and water was added to induce precipitation. The obtained solid was collected via filtration affording compound **25** (46 mg, 44%).
^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 8.00 (dd, 2H, **H_b**, J_o= 9.2 Hz, J_m= 4.6 Hz), 7.77 (bs, 2H, **H_a**), 7.56 (d, 2H, **H_a**, J= 9.1 Hz), 6.06 (s, 2H, -OC**H_2**O-). **ESI-MS:** m/z 443.51 $[M+H]^+$; m/z 481.58 $[M+K]^+$.

X-ray crystallography

The crystal structures of compounds **9 QxCav** and **20 – C_s** were determined by X-ray diffraction methods. Intensity data and cell parameters were recorded at 190(2) K on a Bruker D8 Venture Photon ll diffractometer (Cu K_α radiation λ = 1.54178 Å).
The raw frame data were processed using the programs SAINT and SADABS to yield the reflection data files.[15] All structures were solved by Direct Methods using the SIR97 program[16] and refined on F_o^2 by full-matrix least-squares procedures, using the SHELXL-2014/7 program[17] in the WinGX suite v.2014.1.[18]
All non-hydrogen atoms were refined with anisotropic atomic displacements, unless in cases where disorder was present (solvent, alkyl chains, etc. in case of **20 – C_s**). The hydrogen atoms were included in the refinement at idealized geometry (C-H 0.95, 0.98, 0.99 and 1.00 Å for aromatic methyl, methylenic and methane H atoms, respectively) and refined "riding" on the correspondent parent atoms. The H atoms of the two disordered DMSO molecules in **20 – C_s** were neither found in the difference Fourier's map, nor calculated.
The weighting schemes used in the last cycle of refinement were $w = 1/[\sigma^2F_o^2 + (0.0721P)^2 + 3.5407P]$ and $w = 1/[\sigma^2F_o^2 + (0.1073P)^2 + 0.716348P]$, where $P = (F_o^2 + 2F_c^2)/3$, for **9 QxCav** and **20 – C_s**, respectively.
For **9 QxCav** the calculated molar mass, density and absorption coefficient include six disordered chloroform molecules per cell which do not appear in the final files because of the refinements carried out with data subjected to SQUEEZE.[19]
Crystal data and experimental details for data collection and structure refinement are reported in **Table 5.1**.

Compound	**9 QxCav**	**20 – C_s**
Empirical formula	$C_{96}H_{96}N_8O_{16}$ • 1.5 $CHCl_3$	$C_{86}H_{82}N_8O_{10}$ • 2 DMSO
M	1796.85	1543.85
Crys. Syst.	Monoclinic	Monoclinic
Space group	*C2/c*	*C2/c*
a/Å	21.549(3)	19.875(2)
b/Å	20.585(3)	17.471(2)
c/Å	21.327(3)	24.811(2)
α/°	-	-
β/°	110.703(10)	113.162(3)
γ/°	-	-
v/Å^3	11381.3(2)	7920.8(14)
Z	4	4
ρ/g cm^{-3}	1.349	1.295
μ/mm^{-1}	1.951	1.168
F(000)	3772	3272
Total reflections	37151	25382
Unique reflections (R_{int})	6299 (0.1038)	3939 (0.0681)
Observed reflections $[F_o>4\sigma(F_o)]$	4137	2861
GOF on F^2 [a]	1.014	1.006
$R_{indices}$ $[F_o>4\sigma(F_o)]$ [b] R_1, wR_2	0.0541, 0.1298	0.0879, 0.2183
Largest diff. peak and hole (eÅ^{-3})	0.249, -0.491	0.948, -0.914

[a] Goodness-of-fit S = $[\Sigma w(F_o^2 + F_c^2)^2/(n-p)]$ ½, where n is the number of reflections and p the number of parameters; [b] $R_1 = \Sigma \; || F_o | - | F_c || / \Sigma | F_o |$, $wR_2 = [\Sigma[w(F_o^2 + F_c^2)^2]/ \Sigma[w(F_o^2)^2]]^{½}$

Table 5.1. Crystal data and structure refinement information for compounds **9 QxCav** and **20 – C_s**.

5.7 REFERENCES

1. IARC *Monogr Eval Carcinog Risks Hum*, **1979**, *1*, 15.
2. IARC *Monogr Eval Carcinog Risks Hum*, **2012**, *100F*, 249.
3. (a) E.H. Daughtrey Jr., K.D. Oliver, J.R. Adams, K.G. Kronmiller, W.A. Lonneman, W.A. McClenny, *J. Environ. Monit.* **2001**, *3*, 166; (b) M. Marc, M. Tobiszewski, B. Zabiegała, M. de La Guardia, J. Namiesnik, *Anal. Chim. Acta* **2015**, *853*, 116.
4. (a) S. Zampolli, P. Betti, I. Elmi, E. Dalcanale, *Chem. Comm*. **2007**, *27*, 2790; (b) S. Zampolli, I. Elmi, F. Mancarella, P. Betti, E. Dalcanale, G.C. Cardinali, M. Severi, *Sens. Actuators B* **2009**, *141*, 322.
5. M. Vincenti, E. Dalcanale, P. Soncini, G. Guglielmetti, *J. Am. Chem. Soc.* **1990**, *112*, 445.
6. P. Soncini, S. Bonsignore, E. Dalcanale, F. Ugozzoli, *J. Org. Chem.* **1992**, *57*, 4608.
7. (a) E. Dalcanale, G. Costantini, P. Soncini, *J. Inclusion Phenom. Mol. Recognit. Chem.* **1992**, *13*, 87; (b) M. Ferrari, V. Ferrari, D. Marioli, A. Taroni, M. Suman, E. Dalcanale, E. *IEEE Trans. Instrum. Meas.* **2006**, *55*, 828.
8. J. Hartmann, P. Hauptmann, S. Levi, E. Dalcanale, *Sens. Actuators B* **1996**, *35*, 154.
9. E.B. Feresenbet, M. Busi, F. Ugozzoli, E. Dalcanale, D.K. Shenoy, *Sens. Lett.* **2004**, *2*, 186.
10. J.W. Trzciński, R. Pinalli, N. Riboni, A. Pedrini, F. Bianchi, S. Zampolli, I. Elmi, C. Massera, F. Ugozzoli, E. Dalcanale, *ACS Sens.* **2017**, *2*, 590.
11. F. Bianchi, A. Bedini, N. Riboni, R. Pinalli, A. Gregori, L. Sidisky, E. Dalcanale, M. Careri, *Anal. Chem.* **2014**, *86*, 10646.
12. K. Gilmore, I.V. Alabugin, *Chem. Rev.* **2011**, *111*, 6513.
13. V.A. Azov, P.J. Skinner, Y. Yamakoshi, P. Seiler, V. Gramlich, F. Diederich, *Helv. Chim. Acta* **2003**, *86*, 3648.
14. P.P. Castro, G. Zhao, G.A. Masangkay, C. Hernandez, L.M. Gutierrez-Tunstad, *Org. Lett* **2004**, *6*, 333.
15. (a) SADABS Bruker AXS, Madison, WI, USA, **2004**; (b) SAINT, Software Users Guide, Version 6.0, *Bruker Analytical X-ray Systems*, Madison, WI, USA, **1999**; (c) G.M. Sheldrick, *SADABS v2.03: Area-Detector Absorption Correction*, University of Göttingen, Germany, **1999**.
16. A. Altomare, M.C. Burla, M. Camalli, G.L. Cascarano, C. Giacovazzo, A. Guagliardi, A.G.G. Moliterni, G. Polidori, R. Spagna, *J. Appl. Cryst.* **1999**, *32*, 115.
17. G.M. Sheldrick, *Acta Crystallogr. A* **2008**, *A64*, 112.
18. L.J. Farrugia, *J. Appl. Cryst*. **2012**, *45*, 849.

19. P.V.D Sluis, A.L. Spek, *Acta Crystallogr. A* **1990**, *46*, 194.

Materials

All commercially obtained reagents were used as received. Unless stated otherwise, reactions were conducted in flame-dried glassware under an atmosphere of argon using anhydrous solvents (either freshly distilled, passed through activated alumina columns or stored on molecular sieves for at least 48 hrs).
Silica column chromatography was performed using silica gel 60 (Merck 230-400 mesh), or silica gel 60 (Merck 70-230 mesh). The reactions progress was verified by thin layer chromatography (Merck Analytical Chromatography F254®).

Methods

- **NMR Measurements.** ^{1}H NMR spectra were obtained using a Bruker AVANCE 300 (300 MHz), Bruker AVANCE 400 (400 MHz) or JEOL ECZ600R (600 MHz) spectrometer. All chemical shifts (δ) were reported in ppm relative to the proton resonances resulting from the incomplete deuteration of the NMR solvents. ^{13}C NMR spectra were obtained using a Bruker AVANCE 300 (75 MHz), Bruker AVANCE 400 (100 MHz) or JEOL ECZ600R (150 MHz) spectrometer. All chemical shifts (δ) were reported in ppm relative to the carbon resonances of the NMR solvents. ^{31}P NMR spectra were obtained using a Bruker AVANCE 400 (162 MHz) or JEOL ECZ600R (243 MHz) spectrometer. All chemical shifts (δ) were recorded in ppm relative to external 85% H_3PO_4 at 0.00 ppm.
  The terms s, d, t, bt and m indicate singlet, doublet, triplet, broad triplet and multiplet peaks, respectively. The coupling constants (J) were expressed in Hz.
- **MS Measurements.** Electrospray ionization ESI-MS experiments were performed on a Infusion Water Acquity Ultra Performance LC HO6UPS-823M spectrometer equipped with an electrospray interface both in positive and negative modes. MALDI was performed on an AB SCIEX MALDI TOF-TOF 4800 Plus. Exact masses were determined using a LTQ ORBITRAP XL Thermo spectrometer

equipped with an electrospray interface both in positive and negative modes.

- **UV-Vis Measurements.** UV-Vis spectra were collected using a Thermo Scientific Evolution 260 Bio spectrophotometer equipped with a Peltier water cooled cell changer device, using matched quartz cells of 1 cm path length.
- **ITC Experiments.** ITC titrations were performed on a MicroCal PEAQ-ITC (Malvern Paranalytic). Experimental titration curves were analyzed using MicroCal Origin 5.0 program. The association constant values (K_a) and the enthalpy values (ΔH) were calculated as the average of three titrations for every guest. The titrations were performed at 25 °C. Unless stated otherwise, the host solution was put in the reaction cell and titrated with a ten-fold higher concentration of guest solution. The heat released upon binding was tracked against time. To quantify unspecific heats of dilution, each guest was also titrated into pure solvent (blank titration). In all cases, the signal from blank titrations was subtracted from the binding signal before curve fitting.
- **XPS Measurements.** XPS spectra were run with a VSW HA100 analyzer interfaced to a PSP support unit equipped with a monochromatic Mg-K_α X-ray source at 1253.6 eV with a total resolution energy of 0.86 eV. In all spectra, XPS binding energy scale was calibrated by placing the C 1s peak due to hydrocarbon moieties and 'adventitious' carbon at 285.0 eV. The elemental composition of the samples was evaluated by estimating the integrated area of each component corrected for the corresponding photo-ionization cross-section.
- **ATR.** Infrared absorption spectroscopy analyses were performed with a Bruker FT-IR LUMOS instrument using powder samples. The background was subtracted from every spectra recorded.
- **TGA.** Thermogravimetric analyses were performed on a TGA/DSC 1 instrument equipped with a Mettler Toledo gas controller GC10 (pure air/nitrogen). Analyses were carried out in air with a temperature ramp from 30 °C to 800 °C and a flow rate of 30 °C/min.
- **PXRD.** Powder X-ray diffractometric spectra were measured at room temperature with a diffractometer X'TRA (THERMO) using a

Bragg-Brentano geometry θ-θ and Cu-Kα radiation (40kV-30mA), equipped with a solid state Si(Li) detector. The samples were placed in a silica zero-background sample holder. The spectra were collected in a 2θ 3-40° range with steps of 0.05° (2 seconds per step).

- **XRD.** Crystal structures were determined by X-ray diffractometry using a Bruker D8 Venture Photon ll diffractometer equipped with a Cu-Kα radiation (λ = 1.54178 Å) recording the intensity data and cell parameters at 190(2) K.

ABBREVIATIONS

3-FMA	3-fluoromethamphetamine
3MT	3-methoxytyramine
AA	Amino acid
Aβ	Amyloid β-peptide
ADMA	Asymmetric dimethyl arginine
APAHs	Polycyclic aromatic hydrocarbons amino derivatives
AuNP	Gold nanoparticle
BE	Binding energy
BN	Benzonitrile
BTEX	Benzene, toluene, ethylbenzene, xylenes
CB	Cucurbituril
CD	Cyclodextrin (Chapter 1); Circular Dichroism (Chapter 3)
CDA	Canonical Discrimination Analysis
CPE	Carbon-paste electrodes
DA	Dopamine
DCE	1,2-Dichloroethene
DCM	Dichloromethane
DMAB	Dimethylamino borane complex
DMF	Dimethylformamide
DMSI	4-(4-(dimethylamino)styryl)-1-methylpyridinium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOA	Dioctylamine
E	Epinephrine
ECL	Electrochemiluminescence
EDTA	(Ethylenedinitrilo)tetraacetic acid
EtOH	Ethanol
F	Phenylalanine
FeNP	Iron nanoparticle
FBS	Fetal bovine serum

FMOC	9-fluorenylmethyloxycarbonyl
FT-IR	Fourier-transform infrared spectroscopy
G4	DNA G-Quadruplexes
GC	Gas chromatography
HCA	Hierarchical Cluster Analysis
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HEWL	Hen Egg White Lysozyme
HMT	Histone methyltransferase
HR-ESI	High resolution electrospray ionization
HVA	Homovanillic acid
IDE	Integrated development environment
ITC	Isothermal titration calorimetry
KMT	Lysine methyltransferase
LB	Langmuir–Blodgett thin films
LDA	Linear Discriminant Analysis
LOD	Limit of detection
MALDI	Matrix assisted laser desorption/ionization
MEMs	Microelectromechanical systems
MeOH	Methanol
MMA	Mono-methyl arginine
MMB	Magnetic microspheres
MN	Metanephrine
NaOAc	Sodium acetate
NBL	Neuroblastoma
NE	Norepinephrine
NEt_3	Triethylamine
NMN	Normetanephrine
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NP	Nanoparticle
NPAH	Nitroderivatives of polycyclic aromatic hydrocarbons
PAGE	Polyacrylamide gel electrophoresis

PC	Prostate cancer
PCA	Principal Component Analysis
PAH	Phenylalanine hydroxylase
PEA	Phenethylamine
PEG	Poly ethylene glycol
PHLOR	Phloretic acid
PKU	Phenylketonuria
PID	Photoionization detector
PPA	Phenylpropanoic acid
PRMT	Protein arginine methyltransferase
PSMI	1-methyl-4-(4-(piperidinyl)styryl)pyridinium
PTMs	Post-translational modifications
PXRD	Powder X-ray diffractometry
QCM	Quartz crystal microbalance
QMB	Quartz microbalance-based sensors
QxCav	Tetraquinoxaline cavitand
RGO	Reduced grapheme oxide
RNA	Ribonucleic acid
SAM	S-adenosylmethionine
SERS	Surface Enhanced Raman Scattering
Si-MC	Silicon microcantilevers
SDMA	Symmetric dimethyl arginine
SPE	Screen-printed carbon-based electrodes
SPME	Solid phase microextraction
STD	Saturation Transfer Difference
TA	Tyramine
TEM	Transmission electron microscopy
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
TGA	Thermogravimetric analysis
Tiiii	Tetraphosphonate cavitand
TLC	Thin layer chromatography

TOABr	Tetraoctylammonium bromide
TOF	*Time of flight*
TRIS	2-Amino-2-(hydroxymethyl)-1,3-propanediol
UV-Vis	Ultraviolet–visible spectrophotometry
VMA	Vanillylmandelic acid
VOCs	Volatile organic compounds
W	Tryptophan
XPS	X-ray photoelectron spectroscopy