UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
"SCIENZA E TECNOLOGIA DEI MATERIALI"

CICLO XXXV

Synthetic receptors for the detection of proteins and the enrichment of methylated and acetylated peptides

Coordinatore:

Prof. Enrico Dalcanale

Tutore:

Prof. Roberta Pinalli

Dottorando: Martina Orlandini

Anni Accademici 2019/2020-2022/2023

Abstract

This thesis deals with the preparation of synthetic supramolecular receptors for biological applications. After a brief introduction, in Chapter 2 the use of tetraphosphonate cavitands (Tiiii) grafted onto biocompatible ferromagnetic nanoparticles (FeNPs) for the complexation of mono-methylated lysines present as residues in histone tails is presented. The grafting process was performed through a co-precipitation method in presence of $FeCl_2 \cdot 4H_2O$ and $FeCl_3 \cdot 6H_2O$. Before the complexation experiments, commercial animal histones (e.g., calf thymus) were subjected to proteolytic digestion as routinely performed in proteomics experiments to facilitate separation, detection *via* mass spectrometry and spectra identification. In addition, by using peptides as input material the enrichment is significantly more efficient, and in this way, residues otherwise hidden by protein folding are exposed to recognition by the molecular receptors. This peptide mixture was incubated with the functionalized nanoparticles to capture mainly the peptide containing the mono-methylated residues. The extracted enriched mixture was analysed *via* High Resolution ESI mass spectrometry in collaboration with the research group of Prof. S. Sidoli, College of Medicine (New York). In particular, we focused on H3_3_8 fragment, and the results were analyzed in term of enrichment percentage for each modification present in the fragments. The developed tool proceeds to be selective for the recognition of mono-methylated lysine residues and the results were confirmed by performing two control experiments using bare NPs and NPs functionalized with receptors not selective for the target guest.

In Chapter 3, we focused on another important biomarker in epigenetics, namely acetylated lysine (Kac). The work was performed at the ICIQ institute, Tarragona (Spain), under the supervision of Prof. P. Ballester. To this aim, water-soluble calix[4]pyrroles were considered. To investigate the effect of hydrophobicity in the recognition event, calix[4]pyrroles characterized by the presence of cavities different in depth were synthetized. The performances of the obtained molecular receptors toward the recognition of Kac were tested by ^{1}H NMR analyses. The best binding properties were displayed by the shallow receptor proving that both hydrophobicity and shape complementarity are fundamental in the binding process. Moreover, solvation effect related to the guest could represent the driving force of the binding. To move from binding studies in solution to surface, calix[4]pyrroles presenting four acidic moieties at the bottom rim were synthetized and anchored onto FeNPs *via* co-precipitation method. After fully characterization of the grafted system, the same will be used to recognized acetylated lysine residues in histone proteins.

In Chapter 4, the design and synthesis of a host-guest templated DNA duplex exploiting the complexation properties of CB[8] are presented. The reported system, developed in collaboration with the research group of Prof. Alessandro Bertucci at University of Parma, is based on the formation

of an heteroternary complex between CB[8] and two oligonucleotides derivatized with a fluorophore and a quencher, respectively. Firstly, a model system based on the functionalization of both fluorophore and quencher with a PEG-N_3 chain was tested to mimic the conditions of the final system with the oligonucleotides. The anthracene-derivative fluorophore proved to be the best guest for the formation of the heteroternary complex as resulted in a variation of the fluorescence profile upon complexation with CB[8] and the quencher. Then, preliminary click reaction trials were carried out with azide-terminal oligonucleotides and analysed by ESI-MS resulting in the observation of the anthracene fluorophore-DNA product but not that viologen-DNA one. The low reactivity under CuAAC click conditions of methyl-viologen salts and derivatives and the dependence of the yield on the counterions were identified as the probable cause.

Finally, in Chapter 5 the design of a fluorescent sensor for protein detection based on aptamer conjugated tetraphosphonate cavitands is reported. With this aim, a tetraphosphonate cavitand functionalized at the upper rim with four alkyl chains bearing a terminal alkyne unit was successfully synthetised. The system is based on the different affinity of the protein-hindered or unhindered receptor cavity toward a fluorescent guest, namely *trans*-4-[4-(dimethylamino) styryl]-1-methylpyridinium iodide (DMSI). The absence of the target allows the formation of the cavitand-DSMI complex and results in a fluorescence quenching, while in presence of the target the formation of the complex is precluded since the macrocycle cavity is sterically occluded and the fluorescence profile remain unaltered. Also in this case, a model system in which the cavitand was synthetized with four PEG-alkyne units was synthesized and tested in binding with DSMI. The formation of the complex with the unhindered receptor was confirmed through fluorescence analyses as the progressive addition of the dye resulted in the quenching of the dye emission. ITC reported an entropically and enthalpically driven 1:1 binding model with a K_a value in the order of 10^3 M^{-1}. Several click reaction trials were performed with azide-functionalized DNA strands, but cavitand-DNA product formation was not observed. Future perspective will concern the performance of click reaction in different conditions.

Table of content

Chapter 1

Exploiting supramolecular chemistry in biology

1.1 Supramolecular chemistry in biological field

Supramolecular Chemistry relies on the study of molecular recognition events exploiting the binding of guest molecules by highly specific molecular host systems (Figure 1). This process is possible thanks to the synergistic combination of a well-organized host structure, high shape complementary between host and guest components, and a precise control of the weak interactions involved in the binding event, such as: hydrogen bond, cation- π, CH- π, Van der Waals, and π- π interactions.[1] The implementation of these properties results fundamental for the creation of supramolecular architectures with high complexity degree to be exploited in the development of new advanced technologies and functional materials.[2,3,4]

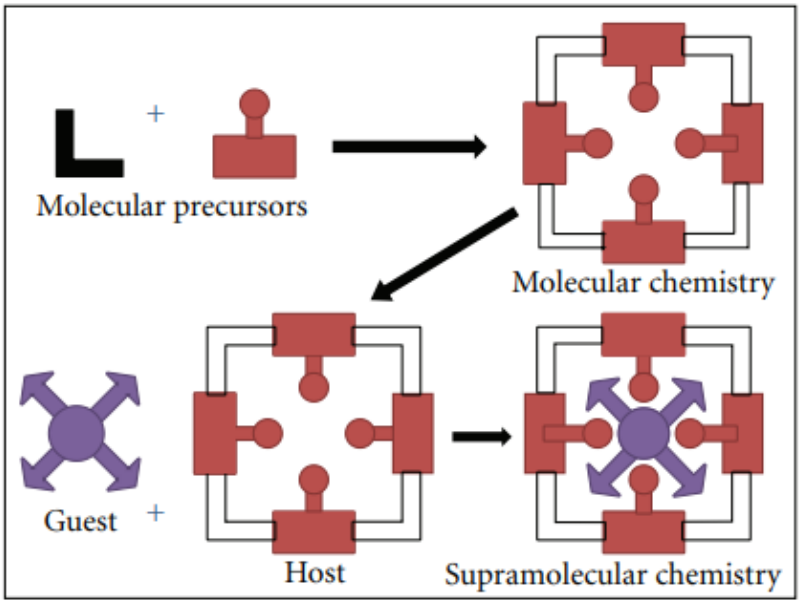


Figure 1. Supramolecular chemistry representation.[5]

Supramolecular Chemistry concept is inspired by nature and biological systems, such as proteins, lipids, and oligonucleotides that are able to manipulate and control biological processes, influencing the functioning or the conformation of proteins, exploiting the interactions between molecules. Synthetic systems reporting the same biological complexity are missing. Thus, supramolecular architectures formed by well-organized multimolecular complexes based on the self-assembly of synthetic molecules become even more interesting tools.[6] The increased knowledge in supramolecular assemblies has led to a fine tuning of the final architectures by shape, composition and functionalities modulation to define a properly self-organized system, bridging the gap between common synthetic compounds and biological systems in term of size and well-defined superstructure (Figure 2).

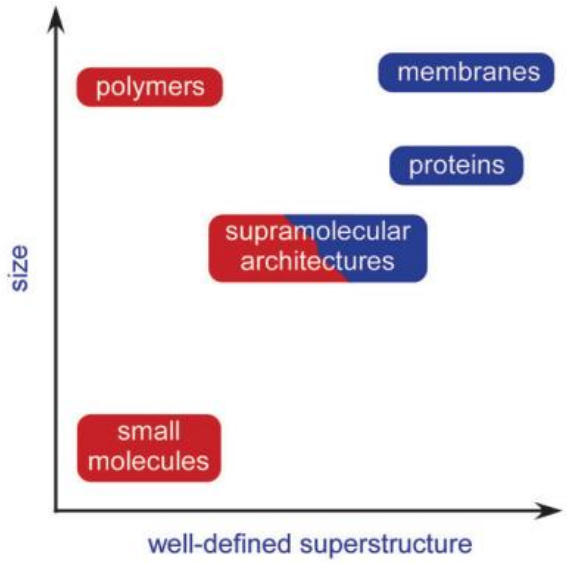


Figure 2. Supramolecular chemistry as self-assembling architectures able to bridge the gap between synthetic molecules and biological systems.[6]

These supramolecular features opened the field to interesting applications in highly interdisciplinary areas both in solution and in solid state. However, the development of supramolecular systems presenting both high affinity and high selectivity toward biological guests, suitable to find application in the study of biological processes remains a challenge because of the strictly demand to make the system performing in water medium.

1.2 Molecular recognition in water

Water represents the medium in which the biological events take place, including protein-ligand binding, enzyme-catalysed recognition and signalling or aggregation of lipids. Starting from these natural systems performances, Supramolecular Chemistry re-elaborates the molecular recognition and the self-organization concepts in order to apply the use of synthetic receptors to biological applications.

As mentioned above, the supramolecular concept bases its foundations on a wide range of non-covalent interactions, such as hydrogen bond, cation- π, CH- π, Van der Waals, and π- π interactions. In detail, hydrogen bonds and π-π interactions are responsible for the formation of thermodynamically highly stable and kinetically labile molecular assemblies (i.e., DNA double helix). Further, π-π interactions, which are particularly interesting in water due to the insignificant effect plays by solvation, are strictly dependent on the orientation adopted by the aromatic rings in solution.

Cation-π interaction is an electrostatic interaction between a cation and the polarizable π electron cloud of an aromatic ring. Together with hydrogen bond and π-π interactions contributes to the control of the supramolecular system functions and structures. Finally, Van der Waals interactions result from the attraction or repulsion between the electron clouds belonging to the nuclei atoms involved.

However, in water, the most exploited weak interaction is hydrogen bond supported by "Hydrophobic effect" that represents a dominant contribution for self-assembly in water for both supramolecular and biological ligands, and describes the tendency of non-polar molecules to aggregate in aqueous solution.[7] In biology, the hydrophobic effect is strictly connected with the structure or topography of the binding pocket, a cavity on the surface or in the interior of a protein or a macromolecule able to bind specific ligands, and with the free energy involved in the binding event. In particular, the network of water molecules inside the pocket can form a structure that has less favourable free energy than the bulk water (Figure 3). Free energy is represented by the contributions derived from free energy of bulk water and those associated to water molecules positioned near the non-polar surfaces (not necessarily in direct contact). When the hydrophobic aggregation occurs, the ligand and receptor hydrophobic surfaces approach one to the other resulting in the release of the high energy free water molecules present into the binding pocket in the lower free-energy bulk water.

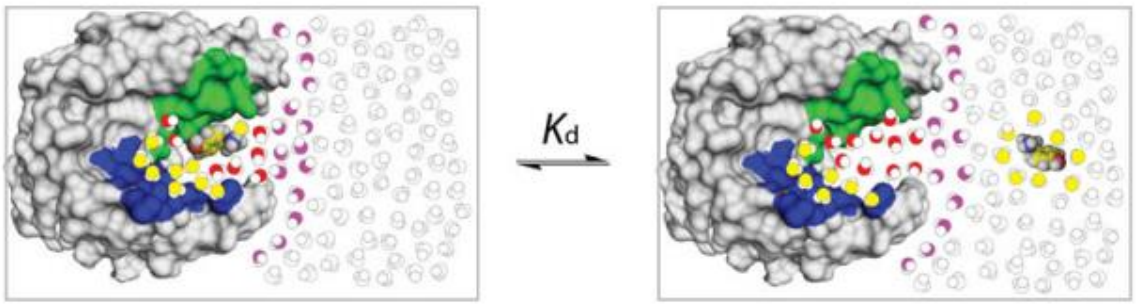


Figure 3. Coloured representation of the supramolecular binding pocket and the network of water molecules involved in the binding. With green and blue colours are illustrated the hydrophobic and polar portions of the cavity surface, respectively. While water molecules are designed as coloured spheres: white represents molecules of water that have free energies close to water in the bulk, while in yellow and red are represented the molecules of water that are less favorable in entropy and in enthalpy compared with bulk water, respectively.[7]

Interesting studies were conducted in supramolecular recognition field in organic solvents;[8,9,10] however, the interactions involved have no analogy with the uniquely described interactions that characterize water medium, such as hydrogen bonds and hydrophobic effect, that involve the release of high energy water from the pocket.

Moving the supramolecular approach to this challenge biological field, the development of innovative supramolecular systems is pivotal to combine the necessity of new dynamic, selective and reversible tools for biological applications.[11,12]

1.3 Moving from molecular recognition in water solution to more elaborated systems for biological sensing

To date, many researchers have developed a wide variety of different synthetic receptors to bind amino acids, small peptides or proteins, in water solution.[13,14,15] However, the availability of a wide range of modular synthetic receptors for efficient recognition of biological guest is still missing in the real applications.[16] Thus, the development of more elaborated supramolecular systems suitable to overcome this gap is becoming of increasingly interest. To date, two important elaborated supramolecular architecture approaches, in which the present thesis will deal in, are sparkling lot of interest in the recent landscape of supramolecular research. The first one concerns the conjugation of the synthetic receptors with biological molecules, resulting on the modulation and control of biological assemblies, while the second one deals with the implementation of solid supports functionalized with the receptor for the study of the binding event at interface level. In particular, in the latter approach, an added positive contribution could be represented by the use of the solid support (i.e., gold, silicon, silica or iron-based nanoparticles) that helps the pre-organization of the supramolecular system. Moreover, it can be recovered and re-used in the sensing event, making the system attractive for real applications. In both approaches, the presence of a pre-organized cavity able to establish supramolecular interactions with a suitable guest is pivotal for molecular recognition. Some recently and interesting studies on these approaches are reported below.

In 2012, Dalcanale and *co-workers,* developed a sensor based on tetraphosphonate cavitands as molecular receptors for the specific detection of sarcosine, a marker of aggressive prostate cancer (PC), in urine samples (Figure 4).[17] The authors exploited the use of silicon surfaces decorated with tetraphosphonate cavitands *via* photochemical hydrosilylation, to recognized selectively sarcosine in water and urine. From the studies performed by XPS and fluorescence guest displacement, an excellent selectivity for sarcosine at the solid–liquid interface was obtained even in presence of naturally ionic interferents.

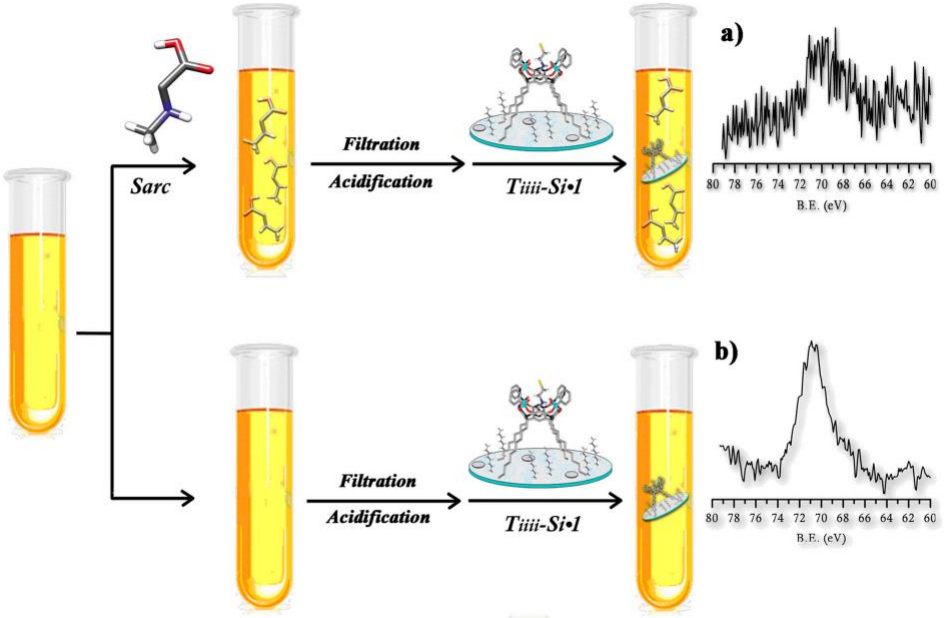


Figure 4. Scheme of sarcosine detection procedure in human urine sample. a) XPS spectrum of Tiiii@Si in sarcosine-added sample and b) control sample.[17]

In 2015, Jayawickramarajah and *co-workers* reported the first example of a covalent CB-DNA conjugate as a modular and responsive system to biologically/clinically relevant stimuli *via* its split DNA aptamer architecture (Figure 5).

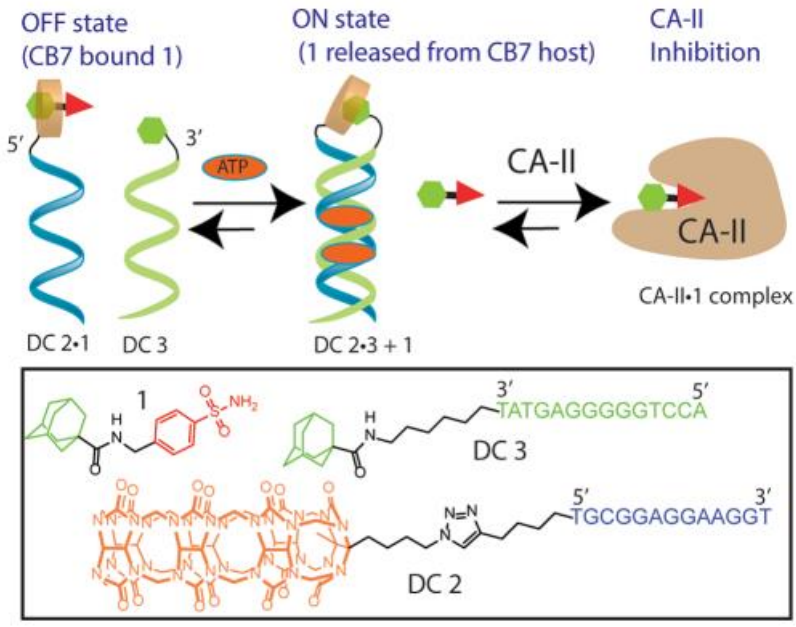


Figure 5. Representation of the covalent CB-DNA conjugate system as a modular and responsive system for the CA-II Inhibition.[18]

The system design is formed by i) two self-assembling DNA small molecule chimeras aptamers named DC 2 and DC 3, functionalized with CB7 and adamantane headgroup, respectively, ii)

carbonic anhydrase II (CA-II) inhibitor (1), containing an adamantane headgroup (for strong CB7 binding), and iii) 5'-triphosphate (ATP). The pair of aptamers units are non-complementary to each other; thus, the duplex formation is possible only in presence of ATP due to its recognition by the binding aptamer sequences. Initially, the presence of CA II inhibitor **1** molecule resulted in the formation of the DC 2·**1** host-guest complex. However, in this situation **1** is found to be too sterically hindered for an optimal CA-II inhibition. Only after the ATP input the two strands were found to be in closed proximity, providing an attractive displacement approach that caused the disruption of the DC 2·**1** host-guest complex, the release of **1** from the CB7 cavity and the establishment of the new host-guest complex between CB7 and adamantane headgroups (CD 2·3). Thereby, the release of molecule **1** in solution allows the optimal inhibition of CA-II.

Thus, depending on the properties of the biological systems, a fine tuning of the supramolecular receptor, combined with the possibility in bioconjugation and bio-orthogonal assembly in water medium, can allow to the creation of ideal platforms suitable for the study, manipulation and control of biological processes and assemblies, opening the field to the development of relevant molecular systems exploitable in biological field applications.

1.4 References

1. Lehn, J. -M. Supramolecular chemistry — Molecular information and the design of supramolecular materials. *Makromol. Chemie. Macromol. Symp.* **69**, 1–17 (1993).

2. Yashima, E. *et al.* Supramolecular Helical Systems: Helical Assemblies of Small Molecules, Foldamers, and Polymers with Chiral Amplification and Their Functions. *Chem. Rev.* **116**, 13752–13990 (2016).

3. Wu, J., Zhao, L., Guo, M. & Tang, J. Constructing supramolecular grids: From 4f square to 3d-4f grid. *Chem. Commun.* **51**, 17317–17320 (2015).

4. He, Y. *et al.* Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452**, 198–201 (2008).

5. Kumar, D., Sharma, D., Singh, G., Singh, M. & Rathore, M. S. Lipoidal Soft Hybrid Biocarriers of Supramolecular Construction for Drug Delivery. *ISRN Pharm.*, 1–14 (2012).

6. Uhlenheuer, D. A., Petkau, K. & Brunsveld, L. Combining supramolecular chemistry with biology. *Chem. Soc. Rev.* **39**, 2817–2826 (2010).

7. Snyder, P. W., Lockett, M. R., Moustakas, D. T. & Whitesides, G. M. Is it the shape of the cavity, or the shape of the water in the cavity?. *Eur. Phys. J. Spec. Top.* **223**, 853–891 (2014).

8. Sierra, A. F. *et al.* Optical Supramolecular Sensing of Creatinine. *J. Am. Chem. Soc.* **142**, 4276–4284 (2020).

9. Bontempi, N. *et al.* Probing lysine mono-methylation in histone H3 tail peptides with an abiotic receptor coupled to a non-plasmonic resonator. *Nanoscale* **9**, 8639–8646 (2017).

10. Biavardi, E. *et al.* Fully reversible guest exchange in tetraphosphonate cavitand complexes probed by fluorescence spectroscopy. *Chem. Commun.* 1638–1640 (2008)

11. Oshovsky, G. V., Reinhoudt, D. N. & Verboom, W. Supramolecular chemistry in water. *Angew. Chemie - Int. Ed.* **46**, 2366–2393 (2007).

12. Voskuhl, J. & Ravoo, B. J. Molecular recognition of bilayer vesicles. *Chem. Soc. Rev.* **38**, 495–505 (2009).

13. Pinalli, R. *et al.* The Origin of Selectivity in the Complexation of N-Methyl Amino Acids by Tetraphosphonate Cavitands. *J. Am. Chem. Soc.* **138**, 8569–8580 (2016).

14. Guagnini, F. *et al.* Cucurbit[7]uril-Dimethyllysine Recognition in a Model Protein. *Angew. Chemie - Int. Ed.* **57**, 7126–7130 (2018).

15. Martins, J. N., Lima, J. C. & Basílio, N. Selective Recognition of Amino Acids and Peptides by Small Supramolecular Receptors. *Molecules* **26**, 106 (2021).

16. Pinalli, R. & Dalcanale, E. Supramolecular sensing with phosphonate cavitands. *Acc. Chem. Res.* **46**, 399–411 (2013).

17. Biavardi, E. *et al.* Exclusive recognition of sarcosine in water and urine by a cavitand-functionalized silicon surface. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 2263–2268 (2012).

18. Zhou, X. *et al.* Host-guest tethered DNA transducer: ATP fueled release of a protein inhibitor from cucurbit[7]uril. *J. Am. Chem. Soc.* **139**, 13916–13921 (2017).

Chapter 2

Cavitand decorated magnetic nanoparticles in epigenetics: methylated histones recognition

Acronyms and abbreviations

ACN	Acetonitrile
DCM	Dichloromethane
DCC	N.N'-Dicyclohexyilcarbodiimide
DCU	Dicyclohexylurea
DIPEA	N,N'-diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethylsulphoxyde
ESI-MS	Electrospray Ionization - Mass Spectrometry
ESI-Orbitrap-MS	Electrospray Ionization - Orbitrap - Mass Spectrometry
EtOH	Ethanol
FTIR	Fourier Transform Infrared Spectrometry
HRTEM	High Resolution Transmission Electron Microscopy
$iPrPCl_2$	Dichloroisopropylphosphine
LD-TOF	Laser Desorption Time of Flight Mass Spectrometry
MeOH	Methanol
Me_2SiCl_2	Dichlorodimethylsilane
MNP	Magnetic nanoparticle
MOMCl	Chloromethyl-methylether
$N(Et)_3$	Triethylamine
SDS-PAGE	Sodium Dodecyl Sulphate – Poly Acrylamide Gel Electrophoresis
STEM	Scanning Transmission Electron Microscopy
HAADF-STEM	High-Angle Annular Dark-Field - Transmission Electron Microscopy
TCA	Trichloroacetic acid
TEA	Triethylamine
TEM	Transmission Electron Microscopy
TLC	Thin- Layer Chromatography
TGA	Thermogravimetric Analysis
XRD	X-Ray Diffraction
Py	Pyridine

2.1 Introduction

2.1.1 Epigenetics

Epigenetics is defined as the inheritable changes that are not coded in the DNA sequence. Epigenetics regulation is normally tuned by DNA modifications, non-coding RNAs and post-translational modifications on chromatin bound proteins. The eukaryotic genetic information encoded in DNA is organized in subunits called nucleosomes, the basic repeating element of chromatin. Chromatin is a polymeric nucleoprotein which consists of around 147 base pairs of superhelical DNA that is wrapped around an octamer of basic proteins, called histones (Figure 1).

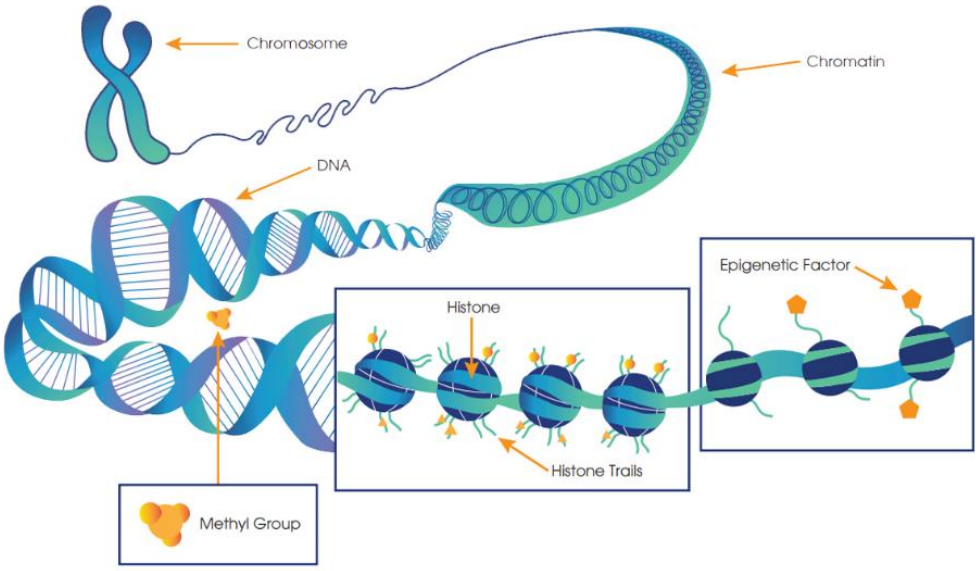


Figure 1. DNA at different packing levels. Taken from *Above the Genome: How Epigenetics Influences Cancer and Obesity*, ***2018***, www.technologynetworks.com.

Histones exist in 5 family, named H1, H2A, H2B, H3 and H4. They are small proteins, highly basic, with a globular domain.[1] All the core histone genes are essential. For example, histone H3 and H4 flank the dyad axis of the structure, bind to the terminal segments of the DNA that enter and leave the nucleosome, and are the most highly conserved histones, suggesting that they play an important role in chromatin formation.[2] Histones are formed by two amino- and carboxyl- terminal unstructured extensions, the first one largest than the second, that protrude from the core of the nucleosome structure, forming the "histones tails". As shown in the X-ray crystal structure reported in the Figure 2, approximately the 75% of the core histones mass is composed by "histone fold domains", which act as spool onto which the nucleosome DNA is wrapped. The core histones are highly conserved

proteins even from yeast to human, and they are pivotal for the maintenance of genetic material. The remaining 25-30% of the histones mass consists of structurally undefined tail domains.[3]

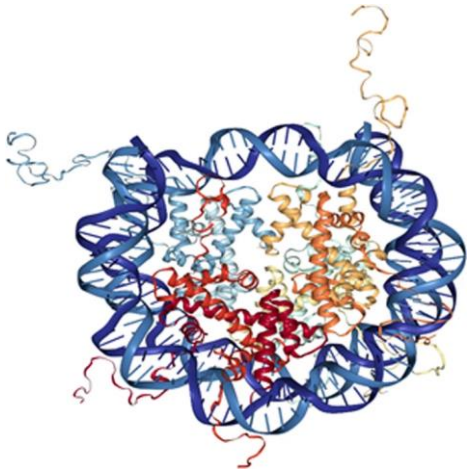


Figure 2. X-ray crystal structure of a nucleosome.[3]

These histone tails are the target of myriad of covalent modifications, called Post-Traslational Modification (PTMs), which take place either during or after the ribosomal synthesis, and include the addiction and removal of chemical groups from the side chains of specific amino acids. To date, most publication focus on biologically significant modifications that occur within the amino-terminal tails of histones. In fact, these basic stretches of amino-acids that protrude from the nucleosome core are believed to act as a platform accommodating the binding of chromatin-associated proteins to the modified residues.[2] PTMs, which are summarized in Figure 3, exist in a broad range. The most studied modifications include: lysine acetylation, ubiquitination and methylation (mono-, di- and tri-methylation), arginine methylation (mono methylation, and asymmetric and symmetric dimethylation), and serine and threonine phosphorylation.[4] These diverse modifications in chromatin play an important role in regulating many aspects of cell function, and, moreover, different types of PTMs can affect some types of interaction between histone-DNA or histone-histone.[5]

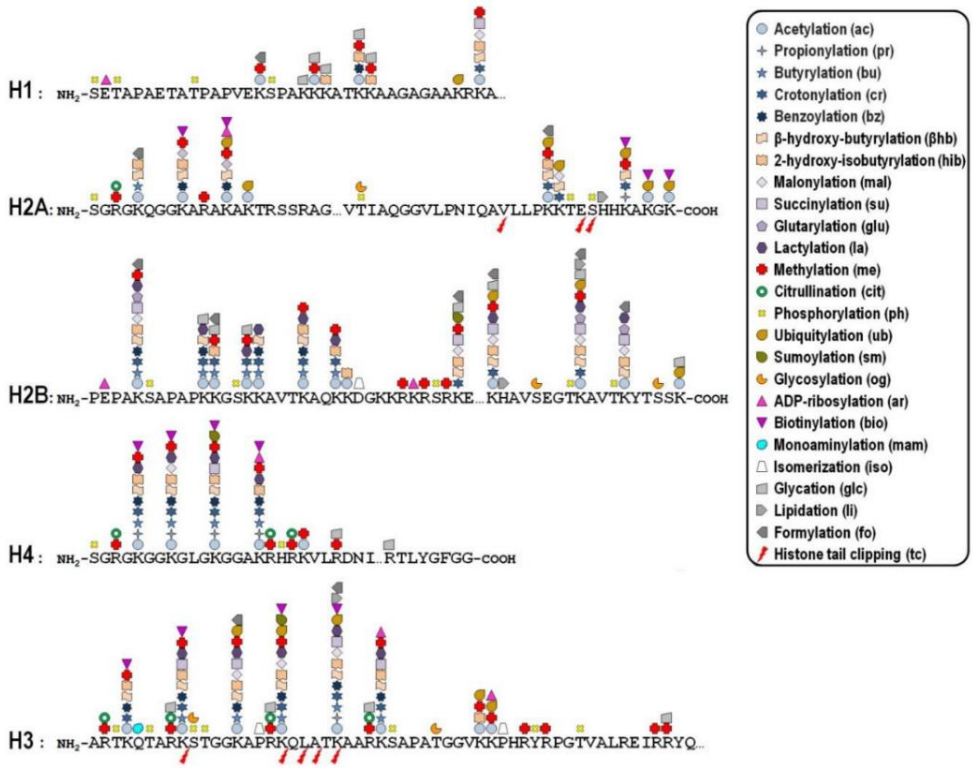


Figure 3. Post-translational modifications identified on histone proteins.[6]

In addition, dysregulation of PTMs can generate pathological effects and it is related to some human diseases like cancer, diabetes, and neurodegeneration.[7,8,9] In fact, different subsets of these PTMs can be: i) associated with therapeutic proteins, ii) involved in the regulation of intracellular processes, including gene expression and signal transduction (e.g., PTMs as phosphorylation and acetylation), iii) involved in product stability and biological activity (e.g., PTMs as phosphorylation), and iv) related to extracellular proteins (e.g., PTMs as glycosylation, carboxylation).[10] For these reasons, they can be used as crucial epigenetic marks or language.

One of the most important PTMs is the side chain methylation of lysine that results in transcriptional activation or silencing. Histone methyltransferases (HMTs) and demethylases (HDMs) enzymes dynamically regulate the process acting as activators/inhibitors[1] and carrying out fine-tuning of the methylation state (Figure 4).[6] These classes of enzymes contain a distinctive catalytic domain that transfers one, two, or three methyl groups derived from the S-adenosyl-L-methionine (SAM) donor to lysine. In addition, these two types of enzymes display a high grade of target specificity, being sensitive to the degree of methylation. For example, within HMT enzymes Suv39H1 univocally targets H3K9 for tri-methylation, while SET7/9 can only target mono-methylate H3K4, in *Neurospora crassa*. Considering HDM enzymes, JMJD2A can reverse H3Kme3 to H3K9me2 but it is inactive in achieving H3K9me1.[6] However, the main sites of methylation are histones H3 and H4, including: H3-K4, H3-K9, H3-K27, H3-K36, H3-K79, and H4-K20. In particular, the methylation in

histone H3 at K4 and K9 residues is the best example of histone methylation-mediated transcriptional regulation.[3]

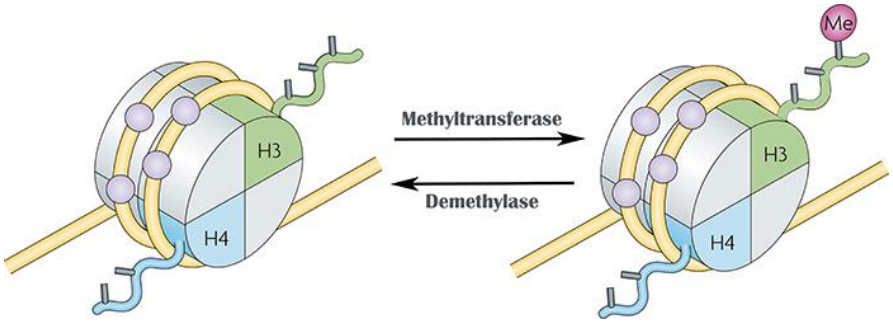


Figure 4. Methylation- demethylation equilibrium promoted by HMT and HDM enzymes.

Methylation of lysine residues can affect in different way the biological processes with diverse consequences. The addiction of the methyl group increases the basicity and hydrophobicity of the histone tails and has an irreversible effect on the structure and function of chromatin on gene activation[11] and on transcriptional repression.[12] These epigenetics disorders can give rise to significant human diseases. For example, recently it was found that lysine methylation contributes to the regulation of tau metabolism. In fact, the abnormal PTM and accumulation of the microtubule-associated protein tau has been implicated in the pathogenesis of Alzheimer's disease (AD).[13] Additionally, the monomethylation of lysine 20 in H4 is associated to the transcriptional repression in L3MBTL1, lethal 3 malignant brain tumor 1.[14]

In this context, during the past few decades, immunological methods were commonly used for the biochemical markers detection since antibodies offer unrivalled specificity toward big molecules like proteins[15] and have an indispensable application for the high-throughput identification of major disease biomarkers and drug targets characterized by epigenetics. However, they have some drawbacks since they are often expensive, unstable and show a poor batch-to-batch reproducibility. In this context, alternative solutions must be considered and the supramolecular chemistry can offer a wide choice of synthetic receptors that can outpace antibodies mainly in reproducibility, chemical stability and cost effectivness.[16]

2.1.2 Tetraphosphorous-bridged cavitands (Tiii)

In the supramolecular chemistry field, the use of resorcinarene-based macrocycles plays an important role in the field of molecular recognition, since they present a concave hydrophobic surface and an enforced cavity of molecular dimension[17] for the recognition of small molecules.

In the design of cavitands, the functionalization of the lower rim with suitable groups is functional to surface grafting or solubilization of the cavitand in different solvents, while the choice of the bridging groups connecting the phenolic hydroxyls of the resorcinarene scaffold is pivotal to determine shape,

dimension, and complexation properties of the resulting cavity. Particularly, phosphonate cavitands present one to four H-bonding acceptor P=O groups at the upper rim of the cavity to enhance the molecular recognition properties (Figure 5).

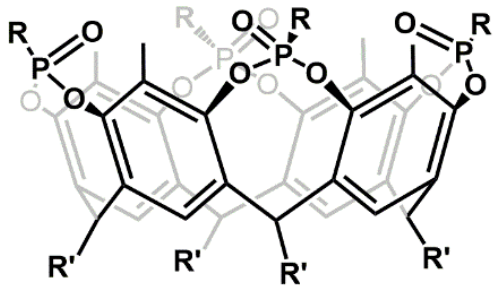


Figure 5. Tetraphosphonate cavitand.

The presence of P(V) stereocenters brings configurational properties into play, since the relative orientation of the P=O groups with respect to the cavity (inward or outward) determines the number of possible stereoisomers (Figure 6).[18] The nomenclature reported in Figure 6 indicates with the capital letter the number and nature of the bridges (T= tetra), with the lower-case letters the in-out stereochemistry (i = in; o = out) and with R_1, R_2 and R_3 the substituents at the lower rim, the apical position and on the P(V) groups, respectively.[18]

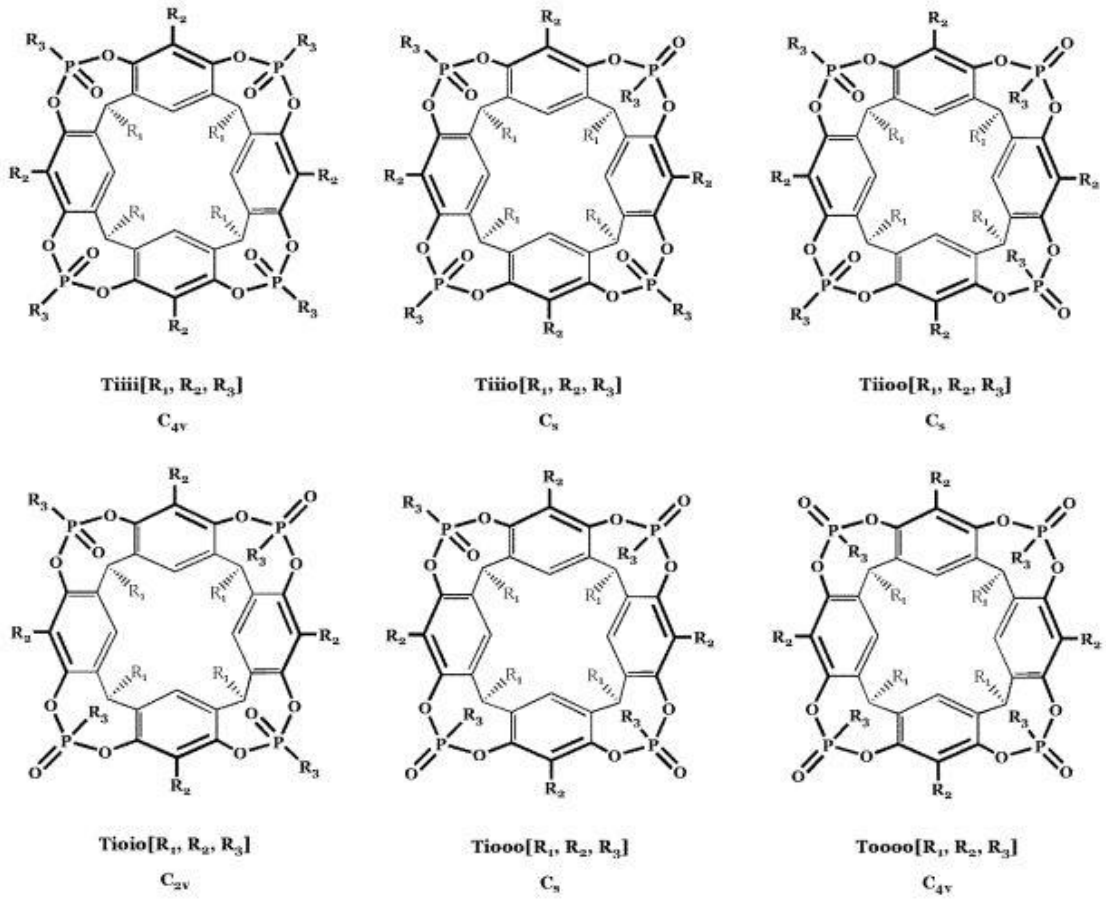


Figure 6. Isomers of tetraphosphonate cavitand.

Tetraphosphonate cavitands show notable molecular recognition properties towards N-methyl ammonium salts both in liquid and at the solid – liquid interface. The main specific and synergistic interactions involved in the recognition are three: $N^+\cdots O{=}P$ cation–dipole interactions; cation–π interactions between the acidic $^+N–CH_3$ group and the π basic cavity; two simultaneous hydrogen bonds between two adjacent P=O bridges and the two nitrogen protons (Figure 7).

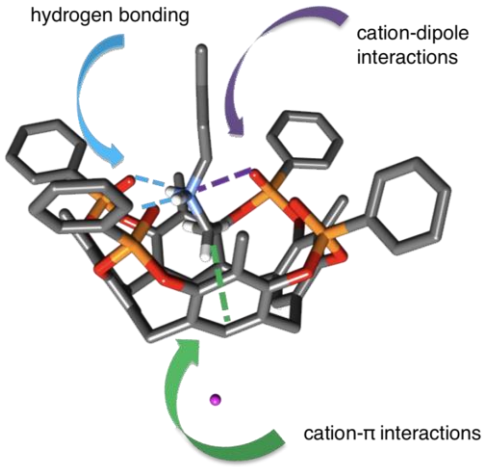


Figure 7. Interactions involved in tetraphosphonate cavitand with N-methyl ammonium salt.

In the ammonium salt series, the tetraphosphonate cavitand receptor prefers the mono-methylated species over the di- and tri-methylated ones as a result of the number of H-bonds formed, while the non-methylated ammonium ions are less complexed by the lack of cation–π interactions (Figure 8). This peculiar affinity of Tiiii toward the $H_2N^+ – CH_3$ group makes these receptors particularly interesting for the detection of a broad range of biologically active compounds containing this residue.[16]

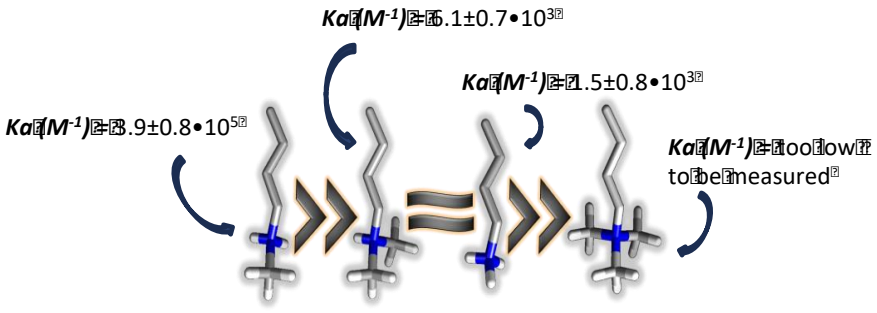


Figure 8. Methylated ammonium salts complexation trend. K_a calculated, in MeOH at 25°C, *via* ITC.[19]

In 2016 Pinalli and co-workers reported an interesting study on the molecular recognition properties exhibited by tetraphosphonate cavitands (Tiiii) toward amino acids. They focused on the recognition of N_ε-Methyl-Lysine and the study, conducted in solution through ^1H, ^{31}P NMR spectroscopy and ITC analysis, and at the solid state *via* X-ray diffraction analysis, revealed that the binding event is

solvent dependent. By single crystal X-ray diffraction, the formation of a 2:1 (host:guest) complex in methanol was elucidated (Figure 9 *left*). The -NH_3^+ ammonium group of the lysine fragment forms H-bonds with the P=O groups of the cavitand and a water molecule, while the methyl group on N_ε position is hosted in the cavity of an opposite cavitand. On the contrary, in water a stoichiometric 1:1 complex is formed and only the $^+NH_2$-CH_3 group is complexed by the cavity, leaving the alkyl chain completely exposed to water (Figure 9 *right*). The guest interacts with the cavity through the already mentioned synergistic interactions, namely H-bonding, cation-π and cation – dipole, and the strong hydrophobic interaction between the N-methyl group of the guest and the electron rich π-basic cavity of the host allowed the ammonium hydrogen and the P=O groups to be closer, releasing the water molecules present in the cavity, which were hydrogen bonded to the phosphonate groups.[20]

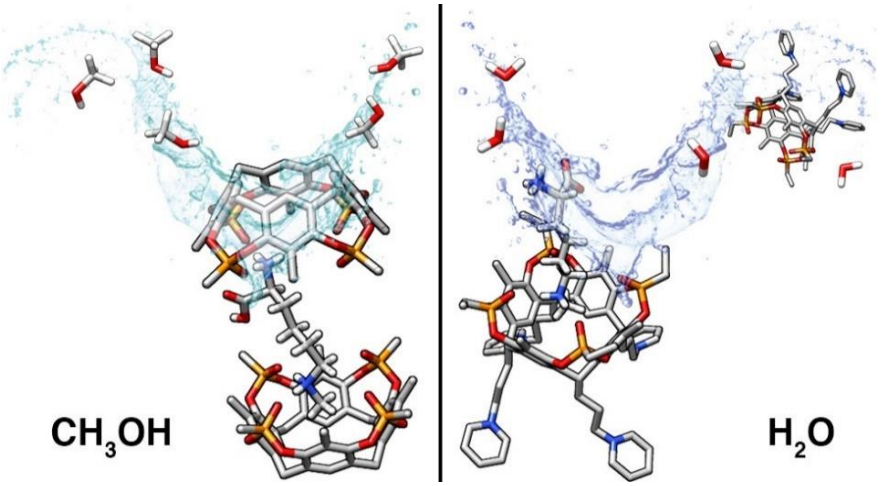


Figure 9. Crystal structures of Tiiii and N_ε- methyl lysine in methanol (*left*) and water (*right*).[20]

ITC analyses for the complexation in methanol solution are in complete agreement with the formation of the 2:1 complex observed in single crystal XRD analysis, and it returned a binding constant value of $1.07 \times 10^6\ M^{-1}$ for the first binding and $1.45 \times 10^3\ M^{-1}$ for the second one. In the water medium the titration data fit to a simple 1:1 binding model and the K_a value is three orders of magnitude lower than that in methanol. The thermodynamic parameters showed that moving from methanol to water the binding event changes from an enthalpy-entropy driven process to an enthalpy driven-entropy opposed process. Moreover, the system was found to be stable in physiological conditions, as the use of Phosphate Buffered Saline (PBS) does not significantly alter the thermodynamic signature of the process. In conclusion, these studies point out the pivotal role played by the monomethylated ammonium ion in water: inside the cavity it acts as a "hook" to reinforce both cation-dipole and H-bonding interactions, and its presence is necessary to ensure complexation in water.

The results obtained in detecting mono-methylated lysine in water and PBS solution inspired Dalcanale and co-workers to transfer this specificity in a more complex system as a human histone H3 tails, which have relevant function in epigenetic regulation (Figure 10).

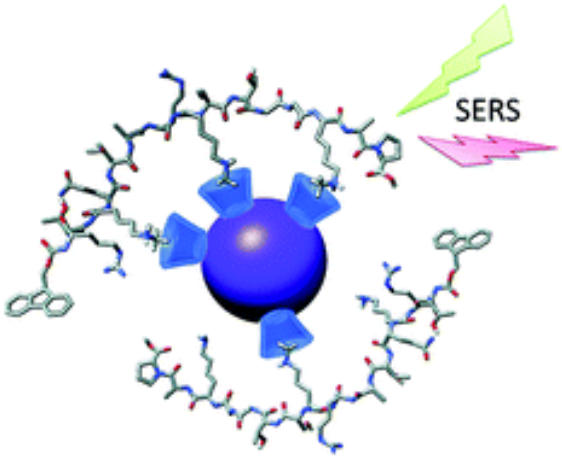


Figure 10. Representation of the supramolecular design using T-rex beads decorated with tetraphosphonate cavitands for the recognition of mono methylated histone H3 tail peptide.[21]

The molecular recognition experiment was carried out using a library of 5 peptides: PC, P1, P2, P3 and P123. Each peptide consists of a sequence of 15 amino acids N-terminated with a FMOC moiety as a labelling agent and bears three lysine residues at position 4 (Lys4), 9 (Lys9) and 14 (Lys14). PC was taken as negative control since no one of its lysine residue is methylated. P1, P2 and P3 bear monomethylated lysines in different positions: in P1 the methylation is located on Lys4, in P2 on Lys9, and in P3 is on Lys14. Peptide P123 displays all the three lysine residues monomethylated. The experiment was conducted in ethanol solution to unfold the peptide chain and expose to the solution also the monomethylated residues otherwise not accessible in aqueous buffers since hided into the folded chain. The peptide was complexed through the Lys-NMe^+ moieties by Tiiii that was anchored on SiO_2/TiO_2 core/shell (T-rex) beads. To this purpose, Tiiii cavitands were decorated at the lower rim with one carboxylic group able to interact with the T-rex beads. Non-Plasmonic Surface Enhanced Raman Scattering (SERS) was selected for the signal transduction, and the triple monomethylated histone P123 resulted to be preferentially recognized thanks to the surface multivalency.[21] Such a results pave the way to improve the enrichment and identification of mono-methylated residues in mixtures of short peptides produced by proteolytic digestion of larger proteins, like histones.

2.1.3 Multivalency and Ferromagnetic Nanoparticles (FeNPs)

Multivalency is a key principle in nature for achieving strong but also reversible chemical interactions between two species. It refers to the simultaneous interaction between multiple functionalities on one entity and complementary functionalities on another, which can be referred also as host and guest.[22] It plays an important role in supramolecular chemistry since it can be used to build a controlled, selective, and directional self-assembly of increasingly complex structures or to achieve a targeted chemical nano-structuring on surface structures (Appendix, 2.5.1).

The design and use of multivalent interactions in solution and at interfaces are of particular importance. At interfaces, multivalent interactions lead towards functions and applications that cannot be achieved otherwise. To exploit multivalency at the interface, supramolecular receptors can be grafted onto the surface of nanoparticles. Magnetic nanoparticles based on iron oxides have been studied since 1981 and have developed much consideration in the scientific community. Due to their unique features, iron and oxygen can combine to form ~16 iron oxides compounds. The most common types of iron oxides in nature are magnetite (Fe_3O_4), maghemite (γ-Fe_2O_3), and hematite (α-Fe_2O_3). Among these, magnetite and maghemite contain single domains of about 5-20 nm in diameter, and especially magnetite exhibits an inimitable form of magnetism. The choice of using iron nanoparticles is related not only to their magnetic properties, but also to their biocompatibility, chemical stability, and the possibility to functionalize their surface with suitable coatings or molecules. Moreover, FeNPs are inexpensive and play an important role in biological field, agriculture and environment. Thanks to their approval to be used in humans by Food and Drug Administration (FDA) they are used in *in-vivo* diagnosis, magnetic genes therapies,[23] drug delivery[24] and in biomedical application for protein immobilization, such as diagnostic Magnetic Resonance Imaging (MRI).[25,26] Moreover, they can be used in biosensing to detect biomolecules of interest, and in this perspective some important requirements are necessary such as particle uniform size and specific magnetic properties.[27] Different physical, chemical, and biological methods have been developed to achieve proper control of particle size, polydispersity, shape, crystallinity, and magnetic properties, as depicted in Figure 11.

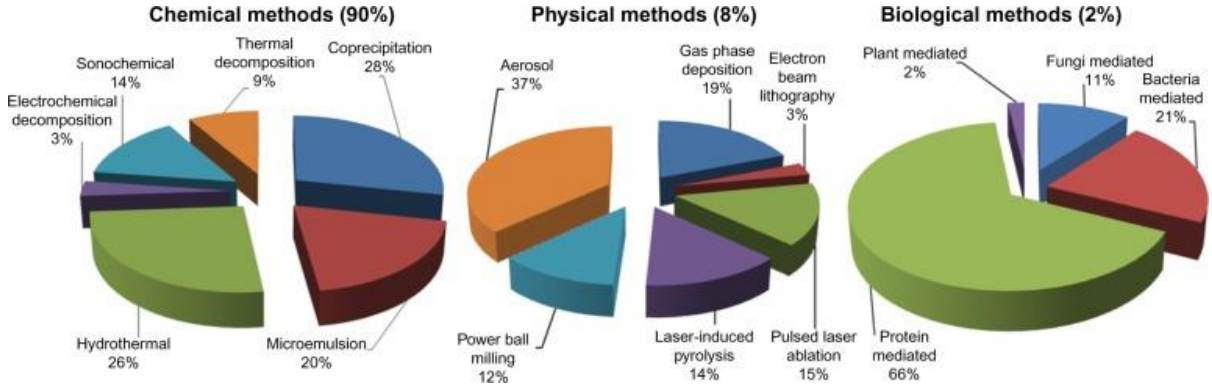


Figure 11. A comparison of the NPs preparation by three different way.[28]

Physical methods come with elaborate procedures, which suffer from the inability to size the particles in the nanometer range. As for biological methods, they ensure low cost, reproducibility, and high yield but are time-consuming. In contrast, chemical preparation methodologies are the most used since they are simpler, inexpensive and reproducible,[29] with high control and tunability of the NPs properties. Within this method, co-precipitation is the most convenient and cheapest way because iron oxides can be synthesized through a direct co-precipitation of Fe^{2+} and Fe^{3+} by the addiction of a base.[30] However some drawbacks have to be considered, in particular barley sufficient particle size

control, broad size distribution and various resulting phases of the oxyhydroxide iron nanoparticles, like akagenite (β- FeOOH), goethite (α-FeOOH), and magnetite (Fe_3O_4). [31]

In literature, several studies reporting the preparation of ferromagnetic nanoparticles grafted with molecular receptors are present. For example, A. Trabolsi *et al.* showed the coating of iron oxide (γ-Fe_2O_3) nanoparticles with cucurbit[7]uril (CB[7]) by microwave heating. The grafting of the CB[7] onto the NPs is due to the cooperative binding of the seven macrocycle carbonyl groups to the surface of the FeNPs (Figure 12).[32]

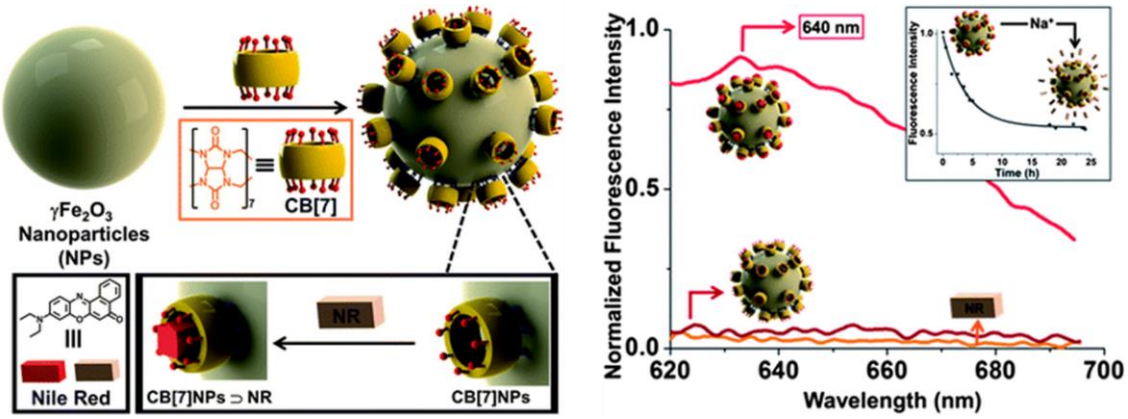


Figure 12. Representation of FeNPs grafting with CB[7] (*left*). Fluorescence emission spectra of CB[7]NPs@NR (red trace), NPs (brown trace)and NR (orange trace) in water (*right*).[32]

In this study, the Authors proposed the use of CB[7] as "nanocontainer" for the controlled release of anticancer drugs, anaesthetics, or agricultural fungicides. In particular, they demonstrated that Nile Red (NR) dye can be complexed into the cavity of the CB and release it into HCT116 cells. The intracellular delivery was monitored through fluorescence measurements since the fluorescence of NR is completely quenched in water while in hydrophobic environments, like the CB[7] cavity, an emission in the red region is present.

In 2007, M. Saunders *et al.* reported a study that involved the use of sulfonato-calixarenes macrocycles as stabilizing agents for magnetite (Fe_3O_4) nanoparticles.[33] They employed a rapid *in-situ* co-precipitation method that allowed the formation of nanoparticles with a good colloidal stability at physiological pH, and roughly spherical in shape. SQUID analysis showed super-magnetism properties of the obtained nanoparticles with high saturation magnetic moment at room temperature. Since the *p*-sulfonato-calixarene can potentially be applied in biomedical applications, the system can find application in contrast agents and drug delivery.

In previous works conducted in our research group, it was demonstrated the possibility of obtaining cavitand-functionalized FeNPs through a co-precipitative technique. The coating was possible thanks to the formation of non-covalent interactions between the iron of the NP and the COOH units

decorating the lower rim of the cavitand. In basic conditions, the carboxylic ion can coordinate the iron atoms behaving as a monodentate or a bidentate ligand, or it can coordinate two different iron atoms (Figure 13).[34]

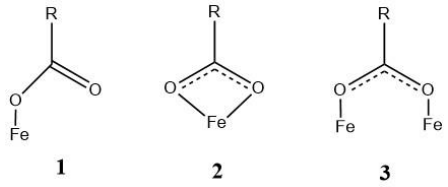


Figure 13. A comparison of the three possible chelation modes of a carboxylate on iron.

2.1.4 Aim of the project

In this chapter, the development of a system capable of selectively recognized N-monomethylated lysine residues present on histone protein tails is presented. The designed system consists of FeNPs functionalised with tetraphosphonate cavitands, exploiting the ability of these receptors in selectively recognize monomethylated ammonium ions with the multivalency effect (Figure 14). To this purpose, receptors **7** and **9** functionalized at the lower rim with four carboxylic groups able to interact with the FeNPs surface, were synthesised. Cavitand **9** functionalized with four methyl bridges at its upper rim is used as control cavitand since it is inefficient in recognizing effectively monomethylated ammonium salt.[19]

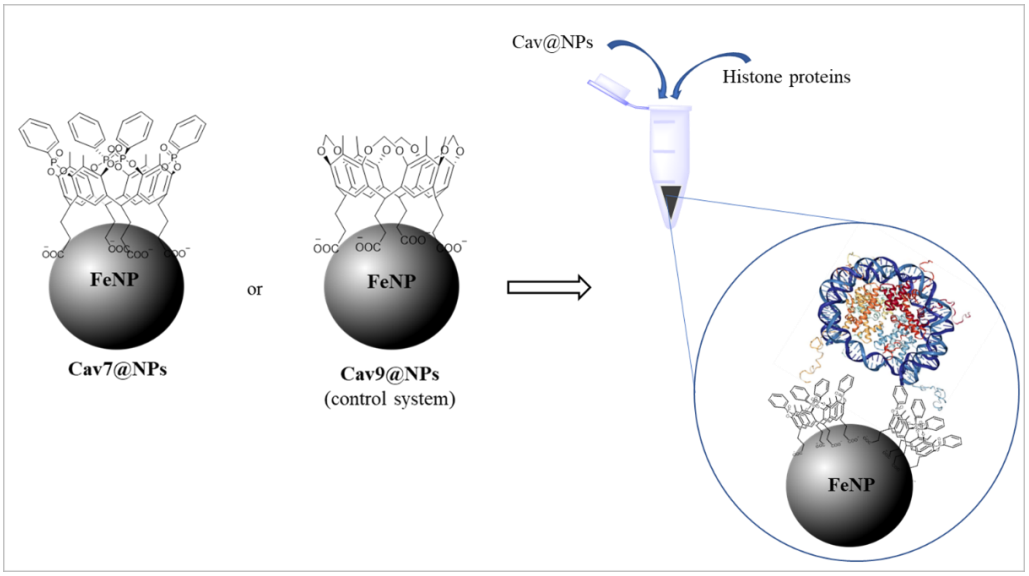
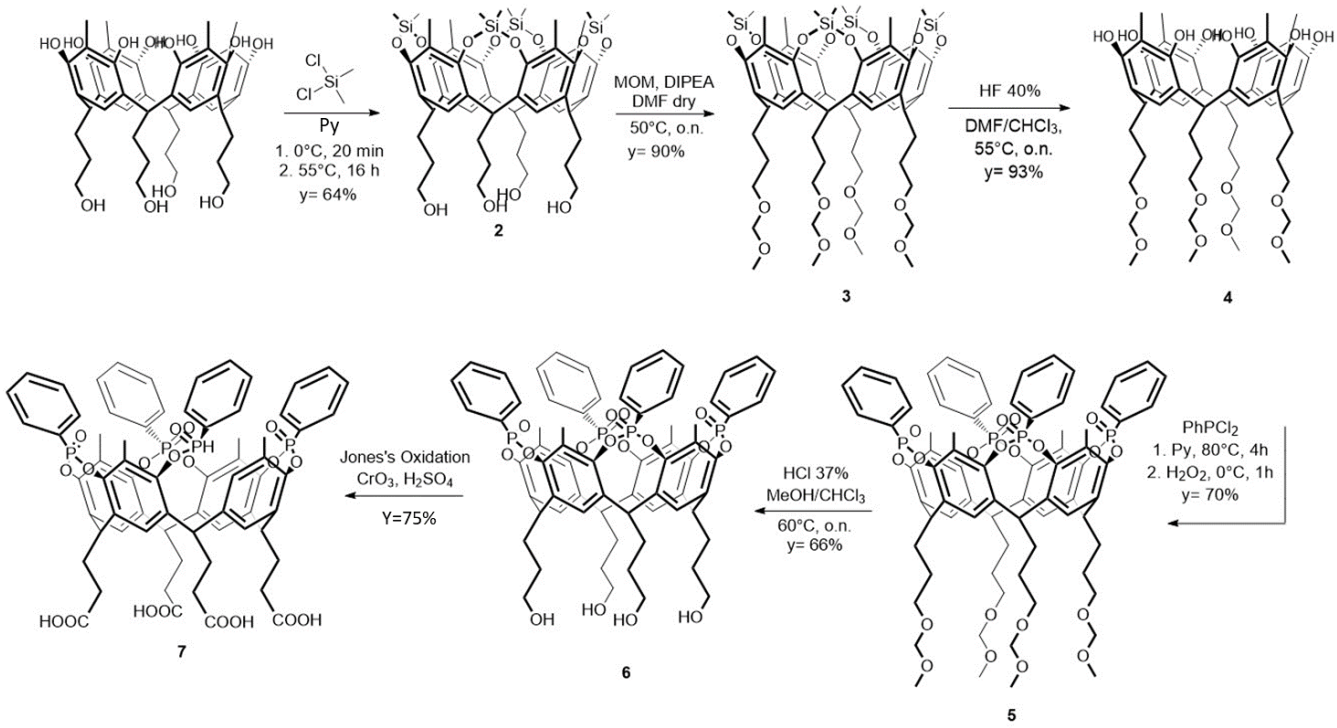


Figure 14. Representation of the aim of the project. Functionalized FeNPs for the recognition of mono-methylated lysine residues present on histone proteins, exploiting the multivalency effect.

2.2 Results and Discussion

2.2.1 Preparation and characterization of Cavitand 7 and Cavitand 9

In Scheme 1, the synthesis of the tetraphosphonate cavitand **7** bearing four -COOH moieties at the bottom rim is reported. The presence of carboxylates allows for the grafting of the receptor on the NPs surface.



Scheme 1. Synthetic pathway for cavitand **7** preparation.

The target compound was prepared in six steps with a 54% overall yield, starting from the readily available hydroxyl footed resorcinarene **1** (Scheme 1). Firstly, steps of protection/deprotection relative to both the phenolic and aliphatic OHs were required, since the subsequent bridging reaction with dichlorophenylphosphine partially chlorinates the hydroxyl groups. Phenolic OHs were protected using dichlorodimethylsilane as reactant in pyridine as solvent and base, obtaining cavitand **2** in 90% yield. The aliphatic OHs at the lower rim were reacted with chloromethylmethyl ether in dimethylformamide (DMF), to obtain the fully protected cavitand **3**. Deprotection of the phenolic hydroxyls was performed with an aqueous solution of HF (40% wt), and resorcinarene **4** was obtained in almost quantitative yield. Subsequently, resorcinarene **4** was reacted with dichlorophenylphosphine in presence of pyridine as solvent and base to obtain the tetraphosphonito cavitand. This step allows the formation exclusively of the tetraphosphonito cavitand presenting the lone pair on the P groups pointing inward the cavity. The following *in situ* oxidation with hydrogen

peroxide, which proceeded with retention of configuration at phosphorous center, led to cavitand **5** with all the four P=O groups pointing inward the cavity. The deprotection of the hydroxyl groups at the lower rim of **5** was performed in acidic methanol and allowed to obtain compound **6** in good yields. The four OHs were then oxidized to carboxylic acids to obtain target compound **7** via Jones oxidation that involved the use of chromic anhydride as oxidant and sulfuric acid as activator in a DMF/acetone (1/1) solution. Cavitand **7** was fully characterized *via* NMR, ESI-MS, FT-IR and TGA. ^{1}H NMR spectrum of cavitand **7** confirmed the structure of the desired compound, thanks to the presence of diagnostic peaks labelled with colored dots (Figure 15). In particular, the green dot at 4.88 ppm indicates the diagnostic triplet of the methine protons. In the aromatic region, from 7.69 to 8.13 ppm, the signals related to the phenyl groups attached to the phosphorus atoms and the resorcinarene scaffold (red, pink, brown and purple dots, respectively) are present. In the aliphatic region it is possible to distinguish the methylene protons of the alkyl chains (blue and yellow dots), and the apical methyl group at 2.10 ppm (light blue dot). ^{31}P NMR spectrum (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 7.86 ppm.

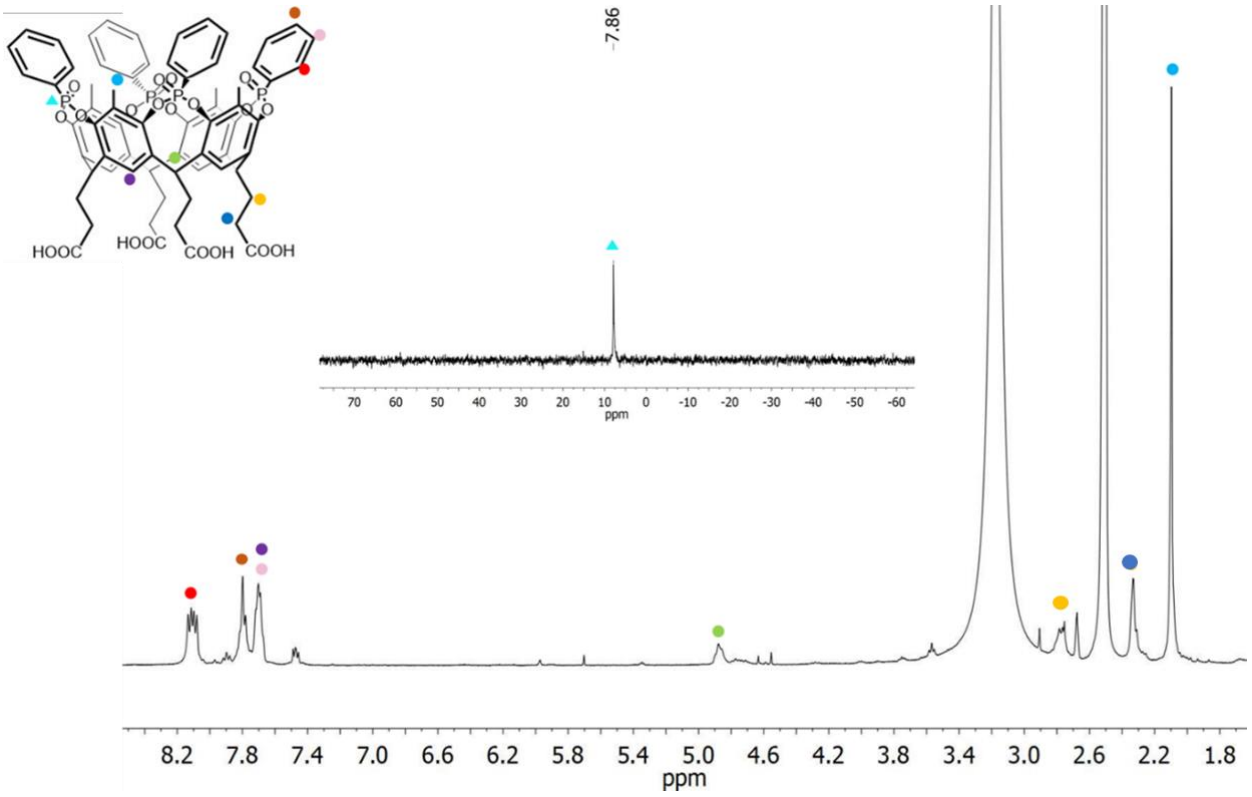
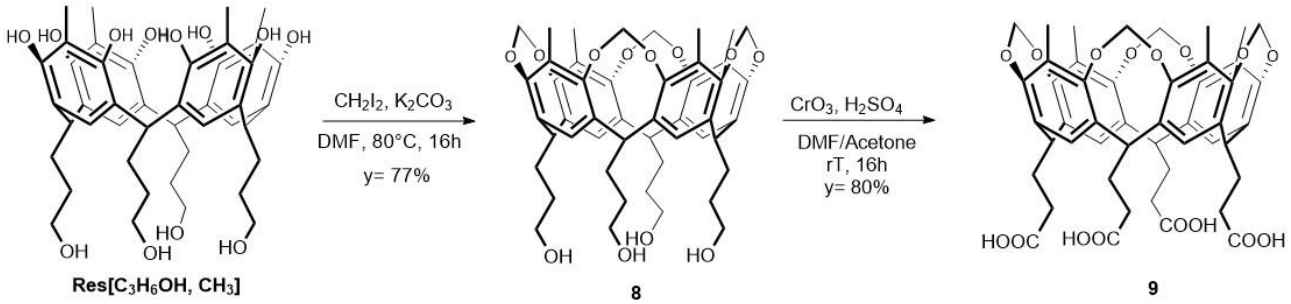


Figure 15. ^{1}H NMR of cavitand **7**, in DMSO-d_6, 400MHz, 80 °C. Inset: ^{31}P NMR.

The FT-IR analysis showed a diagnostic band at 1714 cm^{-1} related to the stretching of the C=O, while the signals at 1068 cm^{-1} and 900 cm^{-1} are associated to the stretching vibrations of P=O and O-P-O, respectively. Finally, a band related to the aliphatic C-H and acid O-H groups at 2927 cm^{-1} is present

(Appendix, Figure 31). The TGA thermogram, carried out in oxygen with a temperature ramp between 25 and 900 °C and flow rate of 20 °C/min, showed that the cavitand is stable up to 200°C. The first weight loss (0.4%) present at 29°C is due to the residual solvent (Appendix, Figure 32).

As reference for blank experiments, the control cavitand **9** bearing for methylene groups at the upper rim was synthesized (Scheme 2). This cavitand is structurally similar to Tiiii but it does not present any biding affinity towards methyl ammonium salts since it lacks specific interactions.[19] As illustrated in Scheme 2, firstly hydroxyl footed resorcin[4]arene was reacted with diiodomethane in anhydrous DMF, using K_2CO_3 as a base for selective deprotonation of the aromatic OHs, obtaining compound **8**. As before, the aliphatic hydroxyls at the bottom rim of the cavitand were oxidized to carboxylic acid via Jones oxidation in presence of chromic anhydride as oxidant, sulfuric acid 96% as activator, and a solution of DMF/acetone 1/1 as solvent. Target compound **9** bearing four carboxylic groups at the lower rim was obtained in good yield and characterized through ^{1}H NMR and ATR-FTIR spectroscopy, ESI-MS spectrometry, and TGA.



Scheme 2. Synthetic pathway for cavitand **9** preparation.

The ^{1}H NMR spectrum of cavitand **9**, reported in Figure 16, showed at 5.93 and 4.31 ppm the presence of the diagnostic doublets belonging to the H_{out} and H_{in} of the methylene bridges, assigned with a green and pale blue dot, respectively. The triplet at 4.86 ppm (orange dot), related to the methine protons, together with two singlets at 7.60 (red dot) and 1.99 ppm (blue) are diagnostic for the resorcinarene scaffold. The signal of the methylene hydrogens next to the carboxylic acids is present as a triplet at 2.36 ppm (brown dot), while for the other alkyl chain protons a multiplet is present at 2.70 (pink dot).

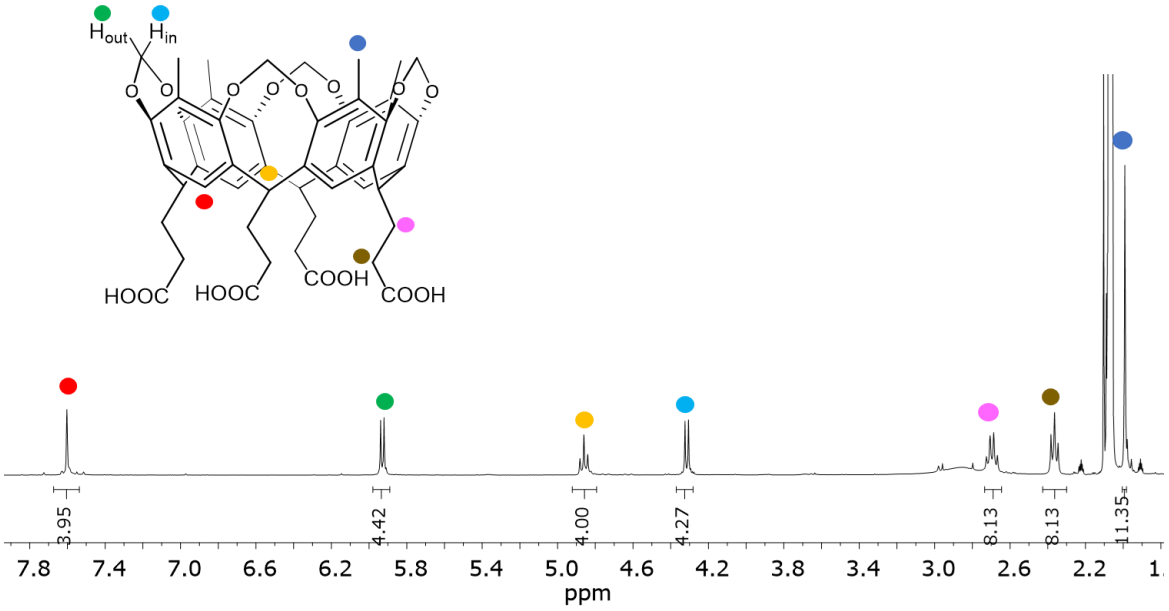


Figure 16. ^{1}H NMR of cavitand **9**, in acetone-d_6, 400MHz, 25°C.

FT-IR displayed at 1704 cm^{-1} the band related to the C=O stretching and at 1428 cm^{-1} the signal related to the stretching of the C-O-C bonds. Finally, at 2925 cm^{-1} the stretching band of the aliphatic C-H of the chains and the acidic O-H groups is present (Appendix, Figure 29).

The thermogravimetric analysis, carried out in the same conditions reported for the Cavitand **7**, revealed that compound **9** is stable up to 300 °C in oxygen. The thermogram (Appendix, Figure 30) showed a first weight loss (8.6%) recorded at 117 °C, due to the presence of residual solvent. The degradation of the cavitand occurs in three steps characterized by three deflections at 387, 459 and 470 °C corresponding to weight loss of 44, 27 and 13% respectively.

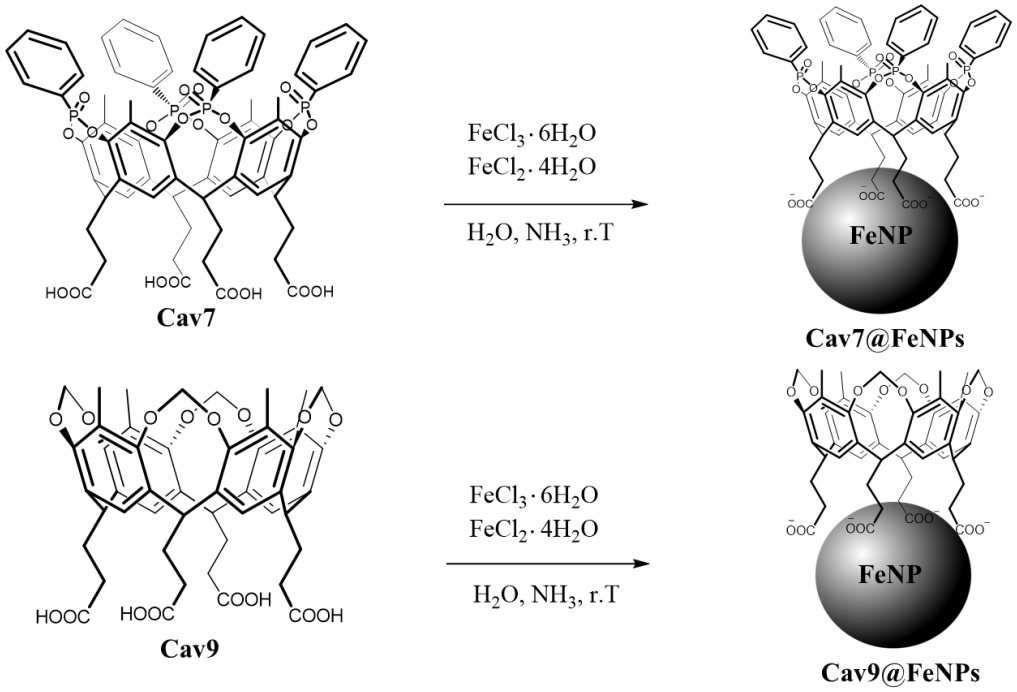
2.2.2 Preparation of bare and functionalized FeNPs

Non-functionalized ferromagnetic nanoparticles (FeNPs)were used as reference to uncover the effects of non-specific interactions during the complexation experiments. The pristine FeNPs were obtained using the co-precipitation method involving the preparation of an aqueous solution of NH_3 (30% aq.) and two aqueous solutions of $FeCl_3$ and $FeCl_2$ in a 2:1 molar ratio (Scheme 3). The latters were added in rapid succession to the first one, keeping the system under manual agitation. The appearance of firstly a dark colouration and then a black precipitate displaying magnetic properties indicated the successful formation of the particles.

Scheme 3. Chemical reaction for the synthesis of magnetite *via* co-precipitation.

The functionalization of FeNPs with cavitand **7** and **9** was performed *via in situ* co-precipitation of $FeCl_3$ and $FeCl_2$ in presence of the receptor, which acts as surfactant to stabilize the FeNPs surface

(Scheme 4). As general protocol, the cavitand was suspended in water, and subsequently ammonia (30% in water) was added to deprotonate the carboxylic acid groups allowing the solubilization of the receptor. To this basic mixture, a water solution of $FeCl_3 \cdot 6H_2O$ (0.98 M) and a water solution of $FeCl_2 \cdot 4H_2O$ (0.95 M) were added. The obtained suspension was centrifuged and washed with H_2O. The so obtained functionalized NPs were finally separated by magnetic decantation and dried with airflow.



Scheme 4. Synthesis of FeNPs functionalized with cavitand **7** and **9**.

Both bare and functionalized FeNPs were characterized through TGA, FT-IR and TEM analysis.

In Figure 17 the comparison between the FT-IR of pristine FeNPs, Cav**7**, Cav**9**, Cav**7**@FeNPs and Cav**9**@FeNPs is reported. Bare FeNPs spectrum is similar to other spectra reported in literature for such particles.[35] In particular, the diagnostic bands are at 551 cm^{-1}, assigned to the stretching of the Fe-O bonds, and the OH stretch at around 3000 cm^{-1}. FT-IR spectrum of Cavitand**9**@FeNPs showed at 1394 cm^{-1} and 1241 cm^{-1} the stretching bands of the C=O and C-O-C bonds decorating the lower rim of the cavitand. The peaks are shifted respect to cavitand **9** since carboxylic ions chelating the NPs' surface are now present, thus confirming the occurred functionalization of the NPs. At 2988 cm^{-1} two bands associated to the stretching of the aliphatic C-H bonds of the alkyl chains are present, while at 556 cm^{-1} it is possible to distinguish the stretching of the Fe-O bonds, as observed in the IR spectrum of non-functionalized FeNPs. For Cavitand**7**@FeNPs system, besides the presence of the band at 567 cm^{-1} associated to the stretching of the Fe-O bonds of the magnetic nanoparticles. The FT-IR spectrum showed diagnostic bands related to the cavitand at 1645 cm^{-1}, 1069 cm^{-1} and 898 cm^{-1}

[1], corresponding to C=O, P=O and P-O-P bonds stretching, respectively, thus confirming the presence of the receptor onto the FeNPs' surface.

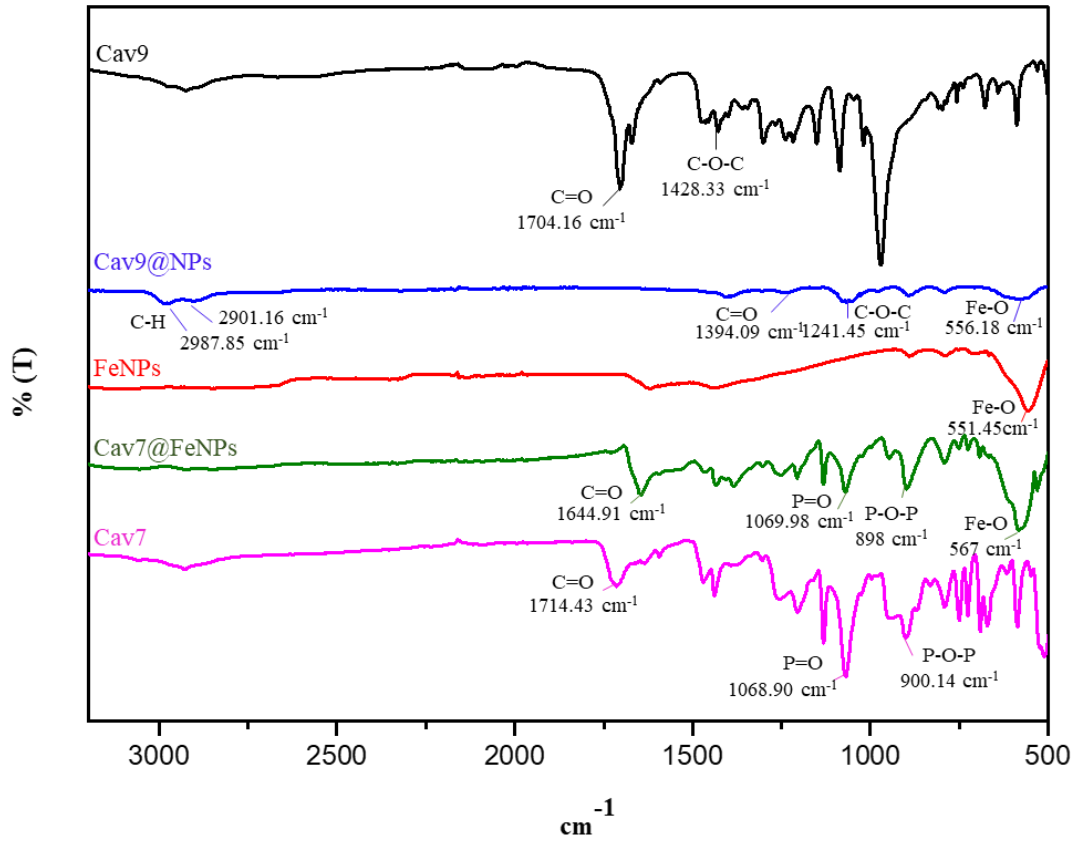


Figure 17. Comparison between ATR-FTIR spectra of Cav**9**, Cav**9**@FeNPs, FeNPs, Cav**7**@FeNPs and Cav**7** (*from top to bottom*).

The TGA was used to determine the grade of functionalization of the NPs surfaces, and the obtained results are reported in Table 1. TGA of Cav**9**@FeNPs (Appendix, Figure 33) was carried out with a temperature ramp between 25 and 800 °C, and an oxygen flow rate of 20 °C/min. The thermogram showed a first deflection around 81°C, corresponding to a loss in weight of 0.55% of the total, attributable to the residual solvent. The second and most important loss in weight is between 200 and 550 °C, with a significant weight loss of 13.6% at 298°C, associated to the degradation of the cavitand.

The TGA thermogram of Cav**7**@FeNPs (Appendix, Figure 34) was carried out in the same conditions of Cav**9**@FeNPs. Also in this case, the first weight loss of 0.1% is related to the evaporation of the solvent, while the degradation of the cavitand is characterized by two deflections, in accordance with the thermogram of cavitand **9**, located at 259 °C and 388 °C, with mass losses of 4.3% and 12.4%, respectively. From the TGA analysis, a mass percentage of 16.7% of cavitand onto the NP can be estimated.

Table 1. Values obtained *via* TGA analysis.

Sample	Temperature deflection (°C)	% loss weight (first deflection)	% loss weight (second deflection)	Functionalization grade (w%)
Cav**7**@FeNP	259; 388	4.3	12.4	16.7
Cav**9**@FeNP	298	13.6	/	13.6

The XRD diffraction pattern of the non-functionalized FeNPs perfectly reflected the pattern of standard Fe_3O_4, present as the predominant crystalline phase (Figure 18).

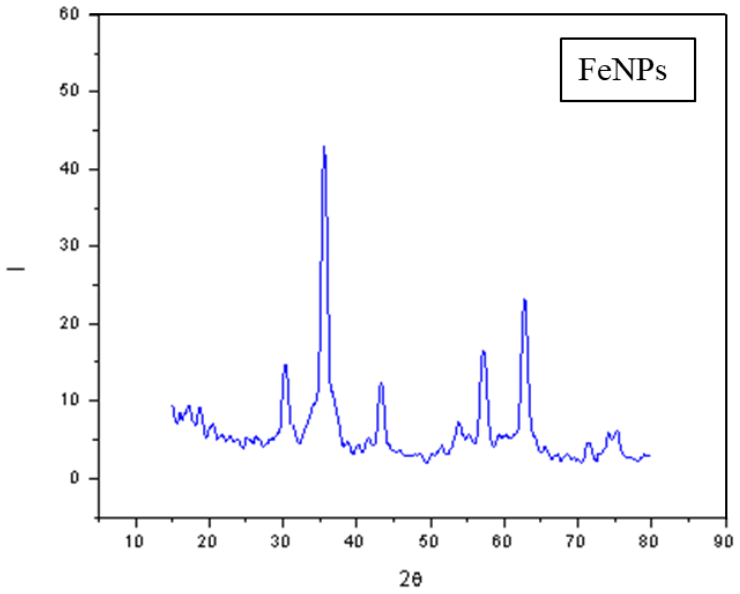


Figure 18. XRD pattern of bare FeNPs.

Both the XRD diffractogram of the functionalized nanoparticles Cav**7**@FeNPs and Cav**9**@FeNPs revealed the simultaneous presence of two crystalline phases: magnetite and goethite (Figure 19),[36] suggesting that the co-precipitation process to obtain functionalized NPs induces the formation of goethite as crystalline phase, in addition to the expected magnetite phase.

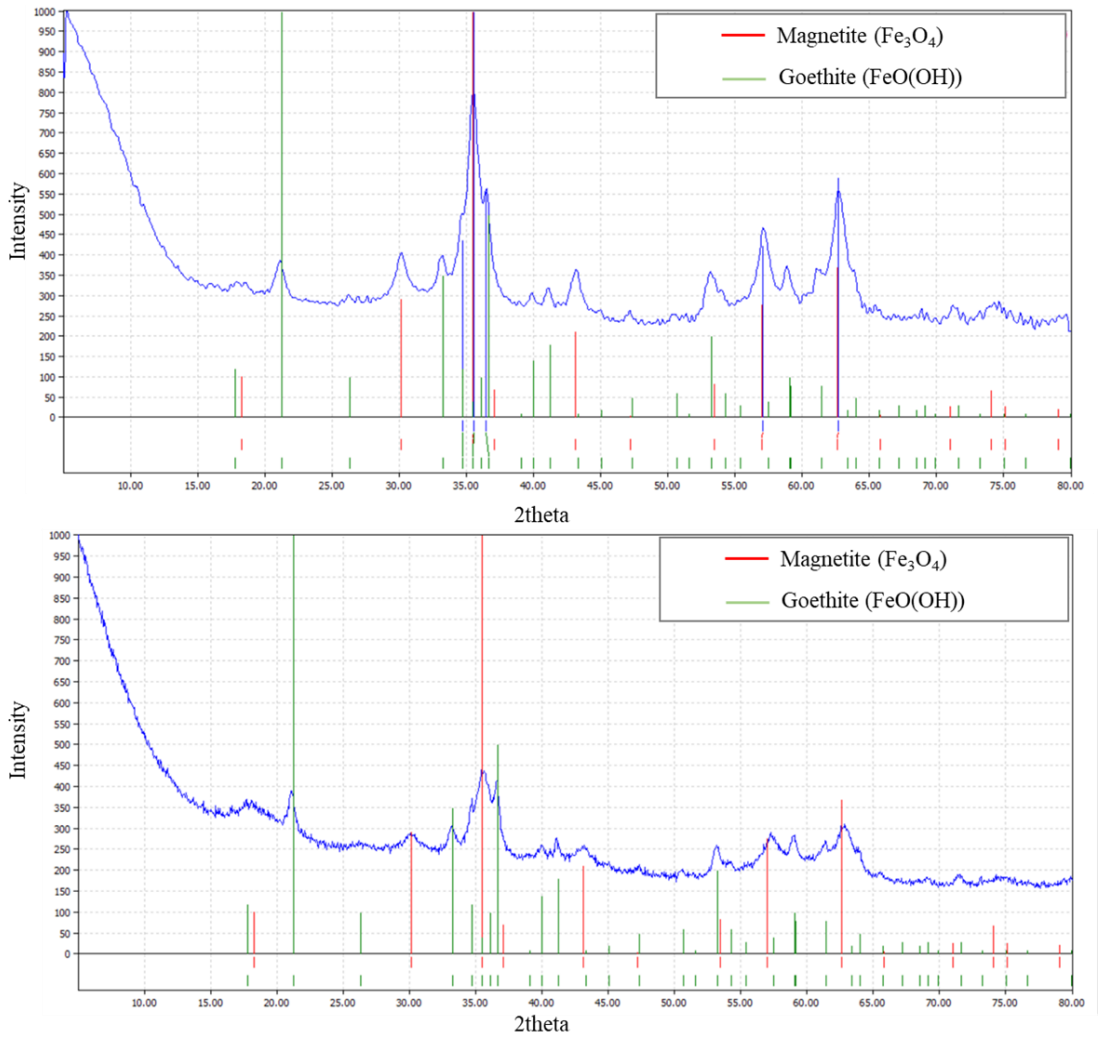


Figure 19. XRD patterns of Cav**7**@FeNps (*top*) and Cav**9**@FeNPs (*bottom*).

To determine shape and size, the synthetized magnetic nanoparticles were characterized using tunneling electron microscopy (TEM), high-resolution transmission electron microscopy (HR-TEM) and High-Angle Annular Dark-Field imaging STEM mode (HAADF-STEM). The followed protocol for the TEM analysis foresaw the suspension of the FeNPs in propanol and subsequent sonication. The obtained suspension was then dripped onto a Cu TEM grid with a continuous carbon film and finally dried with an infrared lamp. Unfortunately, the presence of an amorphous carbon film prevents the possibility to distinguish the presence of an organic layer onto their surface.

The TEM images of the pristine FeNPs collected in conventional TEM and HAADF-STEM, showed the tendency of these NPs to aggregate in large agglomerates, making difficult an accurate study of the single nanoparticle (Figure 20). However, an irregular and often faceted particle shape can be deduced, in agreement with literature.[27,33] The estimated size is in the range of 10 – 20 nm.

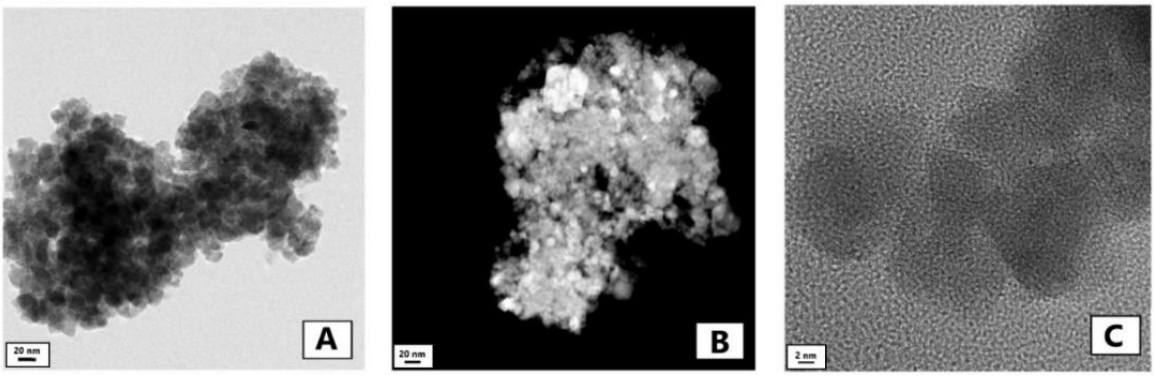


Figure 20. Images of: TEM (**A**), STEM (**B**) and HR-TEM (**C**) of bare FeNPs. Dr. Laura Lazzarini from CNR-IMEM Parma is warmly acknowledged.

As for Cavitand**9**@FeNPs, TEM and HAADF-STEM images revealed that the particles are smaller in size respect to the non-functionalized ones and presented two different shapes (Figure 21). In particular, along with the spherical particles attributable to magnetite phase, there are also crystalline rods of highly variable length and diameter (red circle, Figure 21, B), which are associated with goethite phase, in agreement with XRD analysis.[27]

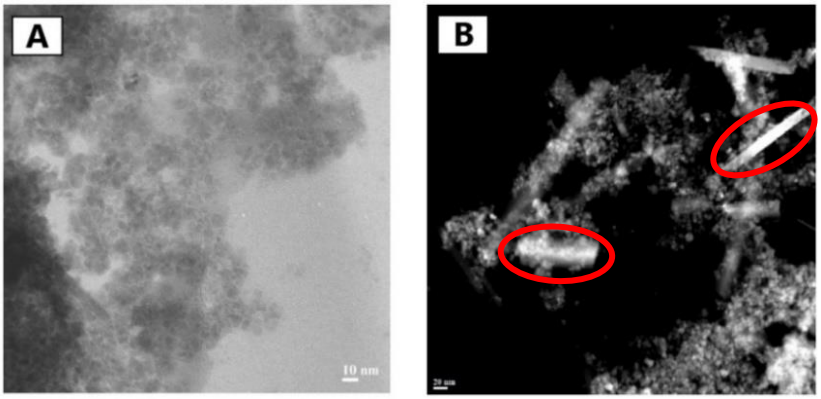


Figure 21. A) Images of TEM; **B)** STEM of cavitand**9**@FeNPs. The red circles highlight the rods related to goethite crystalline phase.

With regard to Cav**7**@FeNPs, TEM images showed that they prefer to form small agglomerates of a few units, as depicted in Figure 22. Crystalline rods related to goethite phase are still visible, but occasionally and of smaller in size respect to the Cav**9**@FeNPs (red circle, Figure 22, A, *top*), in agreement with XRD diffraction pattern. The collected HR-TEM analyses confirmed the XRD results and the predominance of the magnetite phase.

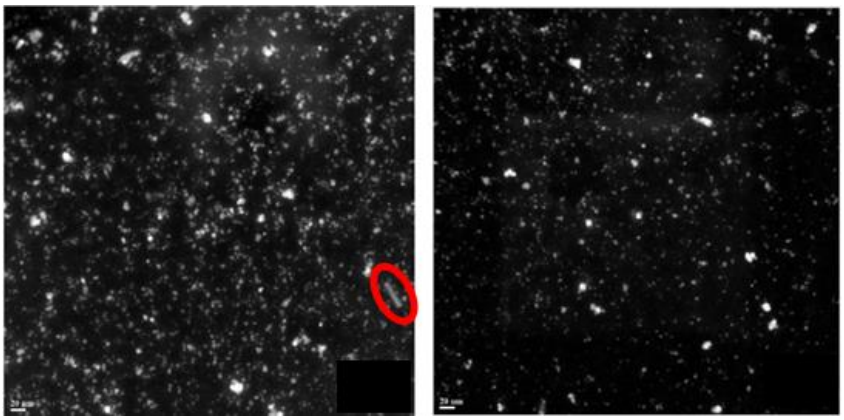


Figure 22. HAADF-STEM images of Cav**7**@FeNPs. The red circle highlights the crystalline rod related to goethite phase.

Z potential measurements were performed for both pristine and functionalised FeNPs by preparing a suspension of nanoparticles in water at controlled pH. The suspension was homogenized under ultrasound for 2 minutes before analysis.

In Figure 23 the obtained results are reported, comparing the uncoated FeNPs sample in pure water (**A**) and basic water (**B**)

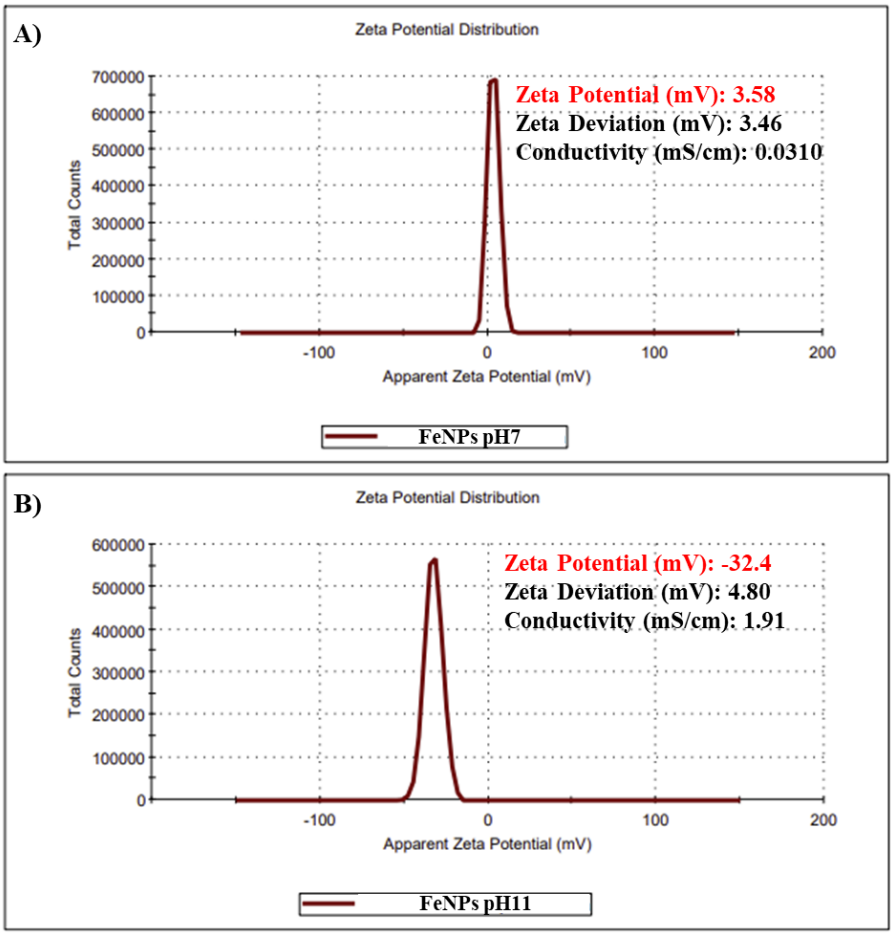


Figure 23. Zeta potential (n = 3) at room temperature of pristine FeNPs in **A)** water (pH=7) and **B)** basic water (pH=11).

Generally, the Zeta Potential of all dispersed systems such as nanoparticles changes as function of the pH. In aqueous systems Fe-OH groups covering the ferromagnetic nanoparticles surface are hydrated and their typical isoelectric point (pH where the Zeta Potential is zero) is related to a pH value around 6.8.[37] Thus, below this value the net charge for magnetite is usually positive and above is negative. In according to literature, in neutral water our bare FeNPs showed a positive value of ζ = 3.58 mV, which can be explained by the presence of water molecules absorbed on magnetite phase surface that allow the protonation of the surface Fe-OH groups generating small positive charges.[38] The increase of the pH from 7 to 11 leads the deprotonation of the groups obtaining a negative value of ζ = -32.4 mV, and thus indicating a prominent anionic character of the surface.

For the coated FeNPs, the analysis was performed in basic water (pH= 11), the same medium that will be used in the extraction protocol in Cav@FeNPs recognition experiments, and the results reported in Figure 24. A value of ζ = -38.7mV and -35.2 mV was calculated for Cav**9**@FeNPs (Figure 24 A) and Cav**7**@FeNPs (Figure 24 B), respectively, very similar to the one obtained at pH 11 for the bare NPs.

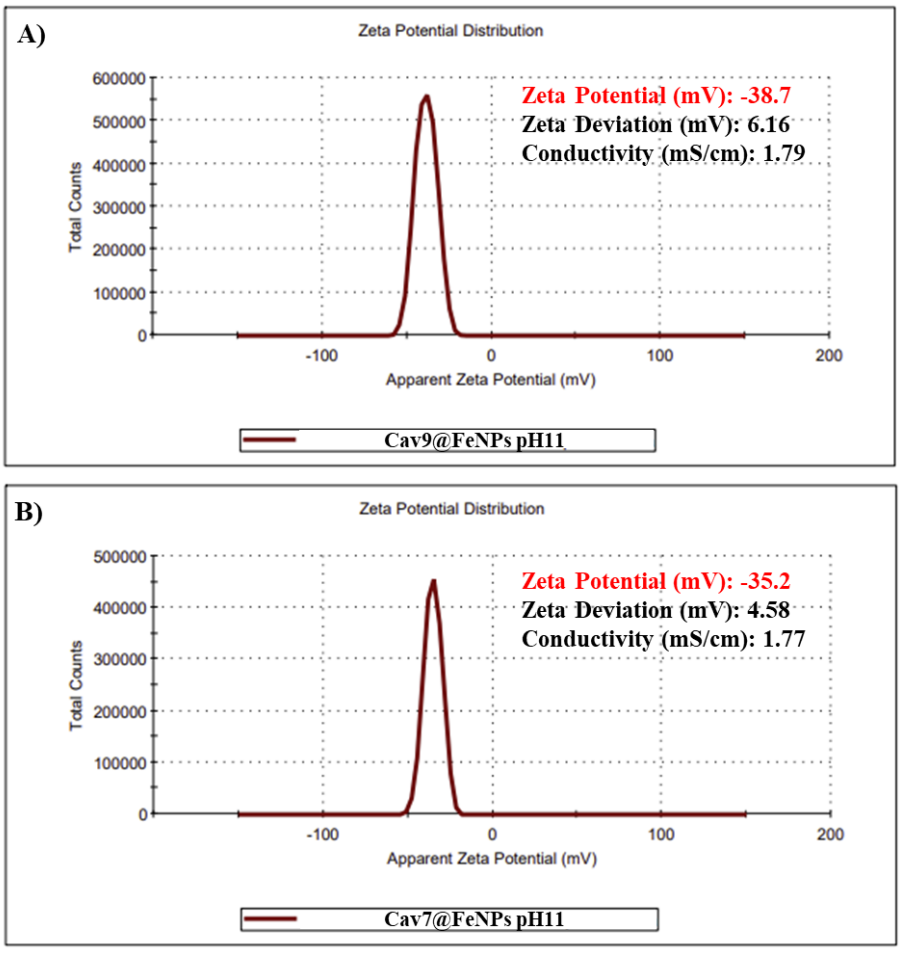


Figure 24. Zeta potential (n = 3) at room temperature of **A)** Cav9@FeNPs in basic water (pH=11), and **B)** Cav7@FeNPs, in basic water (pH=11).

We can conclude that, as expected, at pH 11 the addition of the cavitand-based coatings on the nanoparticle surface does not show any impact on the anionic character of FeNPs. In fact, the four negative charges relative to the carboxylic groups at the bottom rim of the cavitands are exploited for the grafting on the NPs surface through chelation on the iron atoms on the surface and thus, they do not contribute to the final value of Zeta Potential. Therefore, the negative values obtained are associated to the remaining Fe-O^- groups on the NPs, since their surface is not fully functionalized by the receptors, but it reaches just a value around 13-16 % of functionalization, in according to TGA analysis.

2.2.3 Extraction of histones from Saccharomyces cerevisiae

The first attempt to recognize modified lysine in histone proteins was performed using a culture of yeast, *Saccharomyces Cerevisiae*, in collaboration with the research group of Prof. R. Ferrari at Parma University. Compared to human cells, the histone methylation system in yeast is significantly simplified and comprehend only six lysine residues that can be modified by four methyltransferase (Set1p, Set2p, Set5p, and Dot1p) and four demethylases enzymes (Jhd1p, Jhd2p, Rph1p, and Gis1p). These lysine residues belong essentially to the H3 and H4 histone families (Figure 25).[39]

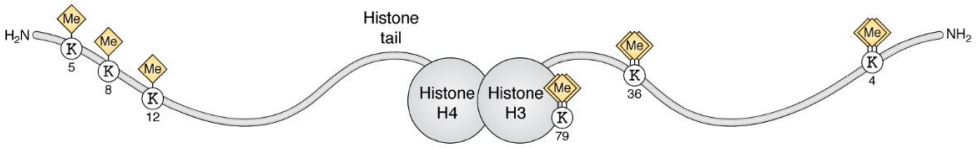


Figure 25. Methylation sites in *Saccharomyces cerevisiae* histones.[39]

To purify and separate histone proteins, numerous biochemical techniques have been developed. We followed a standard protocol present in literature,[40] based on acid extraction of histones from *Saccharomyces cerevisiae* chromatin (Figure 26). Firstly, cell wall degradation through enzymatic digestion with Zymolase 20T was performed to obtain the spheroblasts. The subsequent addiction of IGEPAL® CA-630 detergent allowed the removal of plasmatic membrane to extract the free DNA proteins. The last step foresaw the solubilization of the proteins in acid (H_2SO_4 0.4 N), followed by precipitation in trichloroacetic acid to enrich in histones content the protein precipitate.

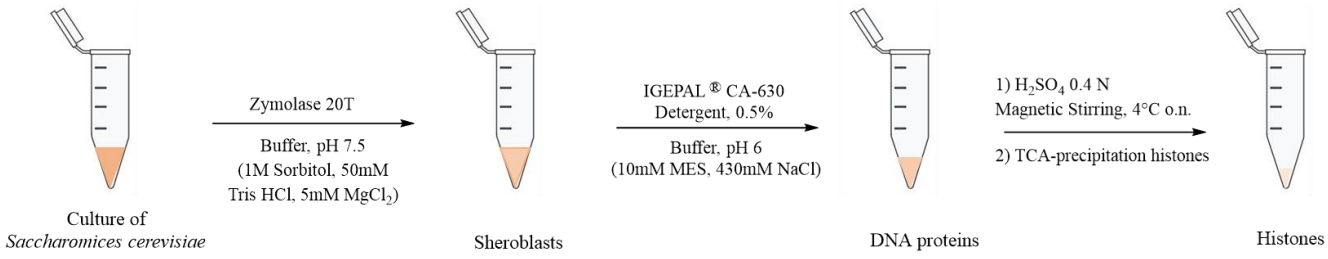


Figure 26. Histones extraction protocol steps.

The total mass of extracted protein was determined through a *Bradford* assay, a fast and accurate spectroscopic quantitative analytical procedure used to measure the concentration of proteins in a solution. The *Bradford* assay is a colorimetric test based on the use of Coomassie Brilliant Blue (CBB) G-250 as dye (Figure 27). In the presence of a protein, CBB forms a dye–protein complex, and hypso shift of the absorbance is recorded (Figure 27). The amount of the complex present in solution correlates to protein concentration and it can be estimated measuring the absorbance at 595 nm (blue form). In our case, the Bradford assay returned a total mass of extracted protein equal to 171 μg.

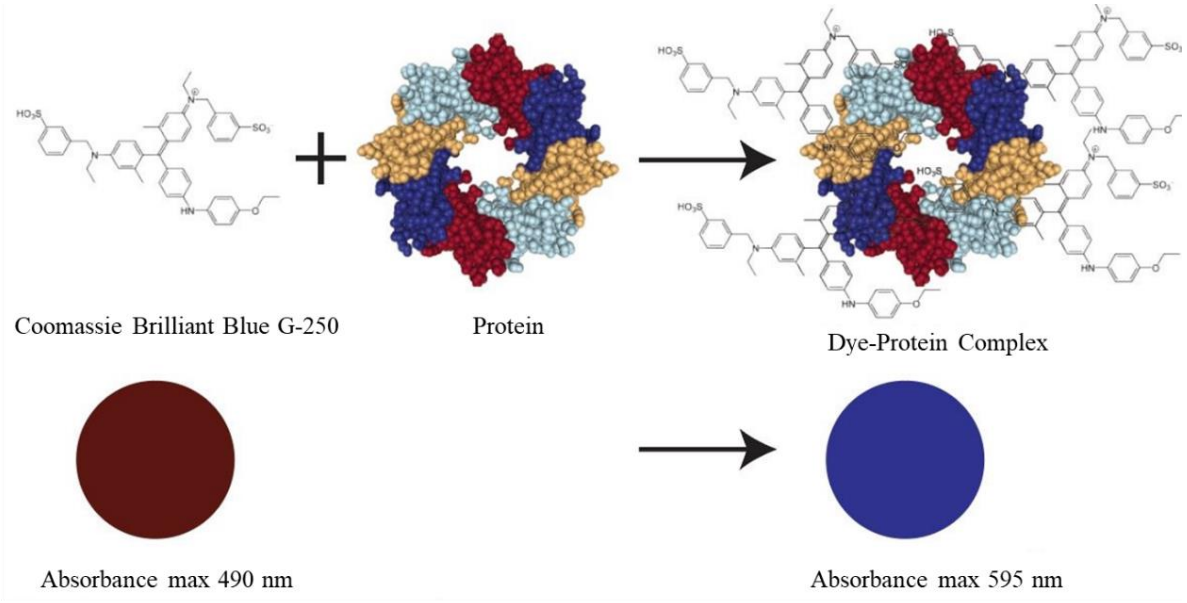


Figure 27. Representation of Bradford assay protein protocol.

The enrichment in histones content was confirmed through SDS-PAGE analysis, a widely used technique for qualitative analysis of protein mixture. SDS-PAGE foresees the use of sodium dodecyl sulphate (SDS) as anionic detergent that binds the protein breaking mainly the non-covalent interactions. Moreover, the addiction of 2-mercaptoethanol as reducing agent causes the reduction of the disulphide bonds present between the cysteine residues. In this linear conformation, the proteins present an overall negative charge proportional to their mass. For the separation of the protein, electrophoresis in polyacrylamide gel as support is employed. Polyacrylamide is used since it is chemically inert, and its porosity can be controlled varying the percentage of acrylamide. With this technique, SDS-protein complexes migrate to the anode and are separated according to their molecular weight. Smaller proteins migrate more rapidly through the gel while the largest are slower. A coloured tracer (blue bromophenol) is used to follow the electrophoretic course. A reference protein mixture is used to distinguish the different proteins based on their molecular weight. The results obtained by the SDS-PAGE analysis are reported in Figure 28. In our analysis, the SDS-protein

sample to be analysed was charged on the gel support in two different aliquots of 5 uL and 10 uL. The most intense spots in the two samples are those related to the four types of histones, as expected. The technique allowed the quantification of the histone protein content around 38.8 μg.

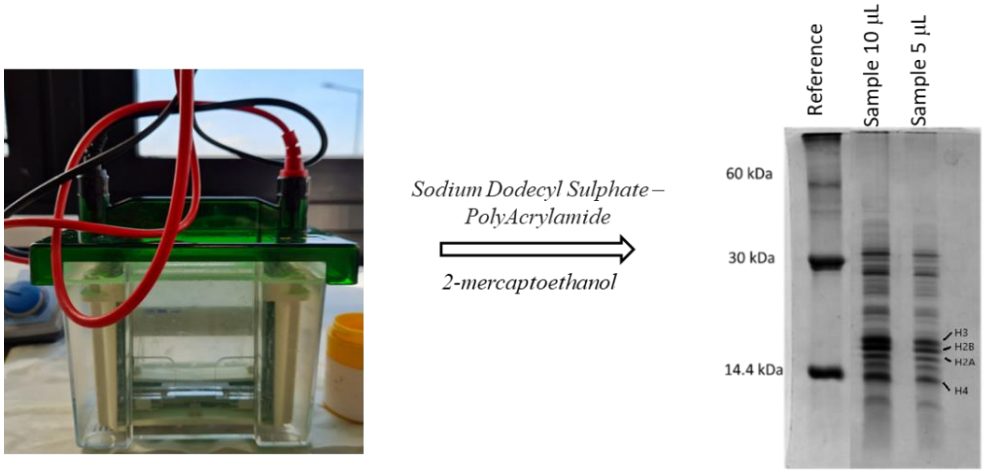
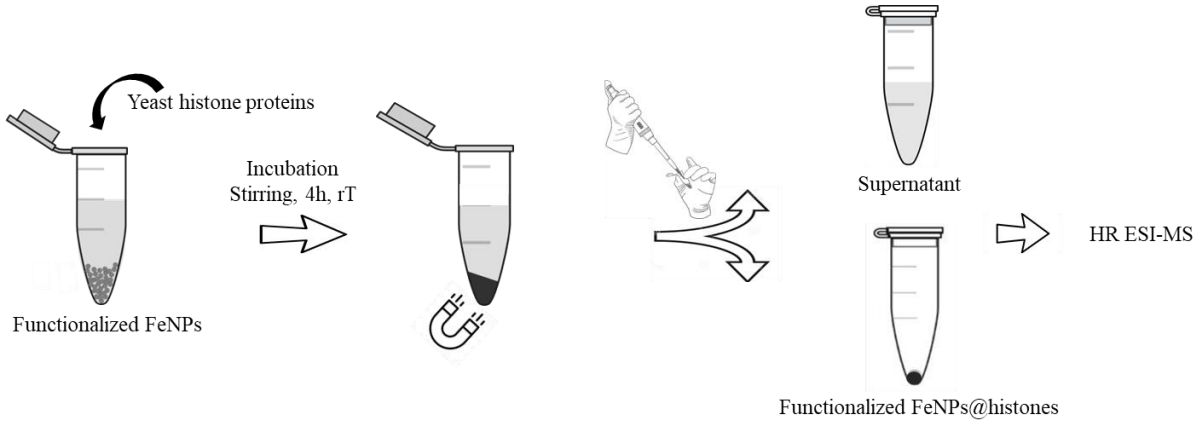


Figure 28. SDS-PAGE assay for histones content characterization.

2.2.4 Recognition experiments using yeast histones and NPs

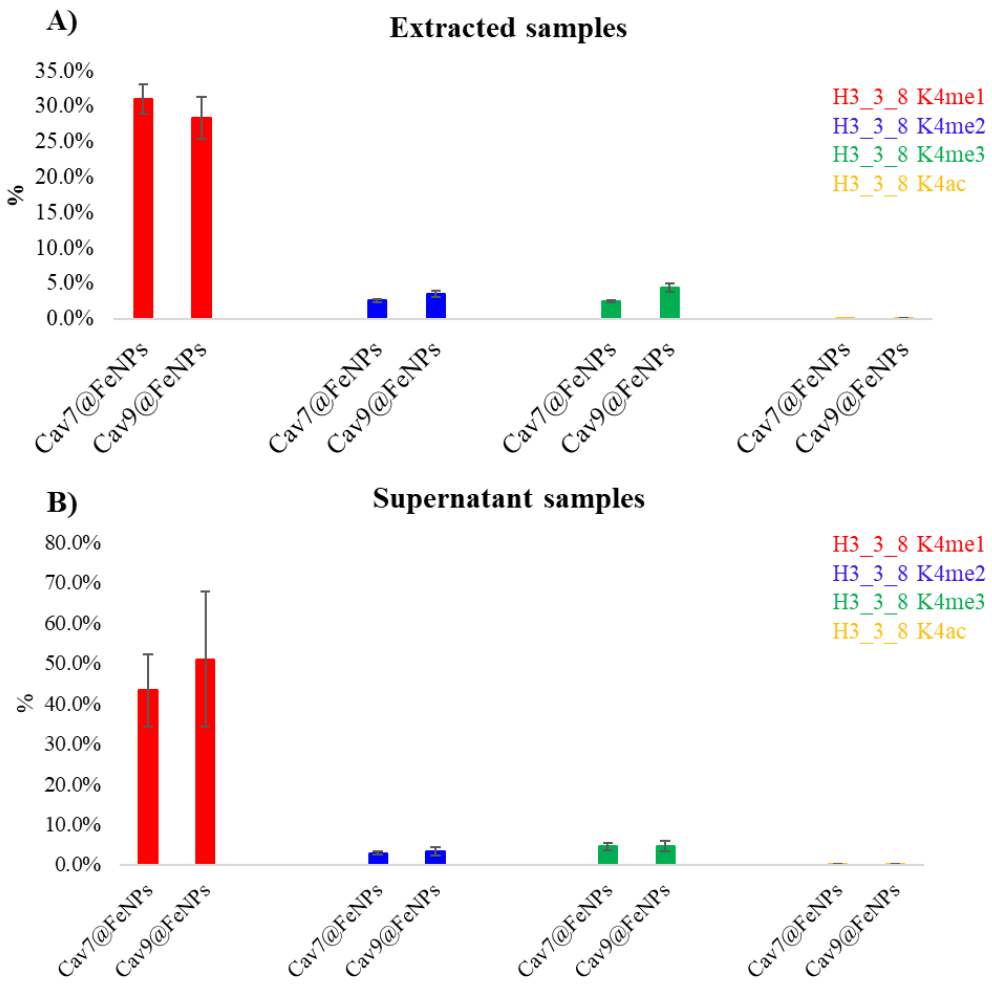
Recognition tests were carried out using as target guests the isolated histones proteins from yeast cells, that include mono-, di- and trimethylated lysine residues in different positions along the peptide chain, and as molecular receptors the systems Cav**7**@FeNPs and Cav**9**@FeNPs, the latter as control (Scheme 5). For both systems, three different replicates were carried out by incubating the histone proteins and the functionalised FeNPs, at a molar ratio of 50:1, in a EtOH/MeOH/HCl 37% (10:1:1, pH=5) mixture for 4 hours, at room temperature. The solvent mixture was chosen to ensure the solubility of the histones and to denature proteins so that modifications on the peptide chain are exposed to recognition. The experiments were performed at pH 5 to ensure the protonation of the amino residues on the histone tail to maximize the interaction with the receptors, while avoiding protonation of the carboxylates at the lower rim of the cavitands, which would cause the release of the receptors from the FeNPs surface. Once the incubation period was over, the supernatant was separated from the magnetic solid residue with the help of a magnet and lyophilised. The solid residue was washed and dried under nitrogen flow. The obtained supernatant and solid samples were analysed in collaboration with the research group of Prof. S. Sidoli (Albert Einstein College of Medicine, New York) through HR ESI-MS analysis.



Scheme 5. Representation of recognition experiment protocol pathway using functionalized FeNPs with yeast histones.

Before performing HR ESI-MS analysis, in Prof. S. Sidoli research group the samples were digested using Trypsin, an enzyme able to cut the histone chains after each arginine unit, allowing the fragmentation of the protein. Moreover, to get rid of possible solid residues, before injection all the samples were purified onto a chromatographic C18 column. The obtained chromatograms were compared to a database containing all the possible and usual fragmentations, allowing the identification of the peaks.

In particular, we focused the attention on the fragment H3_3_8 that is the most exposed to methylation modification. H3_3_8 fragment bears different modifications at the lysine residue, namely mono-, di, and tri-methylation, as well as acetylation. In Graph 1 we reported the results as enrichment % value obtained from mass analyses. As it is possible to note, the obtained results showed that both in the extracted samples (**A**) and in the supernatant samples (**B**) the same percentage of analyte was present for both Cav**7**@FeNPs and the control system Cav**9**@FeNPs. Unfortunately, this means that the system does not show any recognition properties for the mono-methylated species within the methylated residues (red, blue, and green bars) as well as the acetylated one (orange bar).



Graph 1. Enrichment % of modified lysine residues in **A)** extracted samples, and **B)** in supernatant samples.

We speculated that the reason of the poor selectivity is related to two main problems. Firstly, we believe that even if we try to unfold the histone protein using a mixture of alcohols, the majority of the proteins still remain folded, thus not all the lysine residues are exposed out of the protein surface hindering the interaction between the lysine residues and the cavitands. Secondly, a not efficient extraction of the complexed histones from the NPs. In fact, these samples were treated with a solvent mixture of water and acetonitrile for the digestion step. We suppose that this treatment was not sufficient to extract the histones from the cavitands grafted on the NPs. Indeed, it is known that an efficient decomplexation of a primary ammonium salt from the cavitand is performed using a basic solution, able to deprotonate the ammonium moiety allowing the release of the analyte.[41]

2.2.5 Recognition of digested histones from calf thymus

In light of the results obtained previously, we decided to perform an experiment in which a proteolytic digestion of the histone proteins occurs before the enrichment step, using an endopeptidase enzyme (clostripain) named Arg-C, able to cut protein chains at the C-terminus of arginine residues, including the sites next to proline (Figure 29, A). It is worth to note that in presence of Arg-C, cleavage can

also occur at lysine residues. In this way, small peptides are produced allowing the uncovering of target methylated residues that can be hidden by protein folding and not exposed to the environment (Figure 29, B).

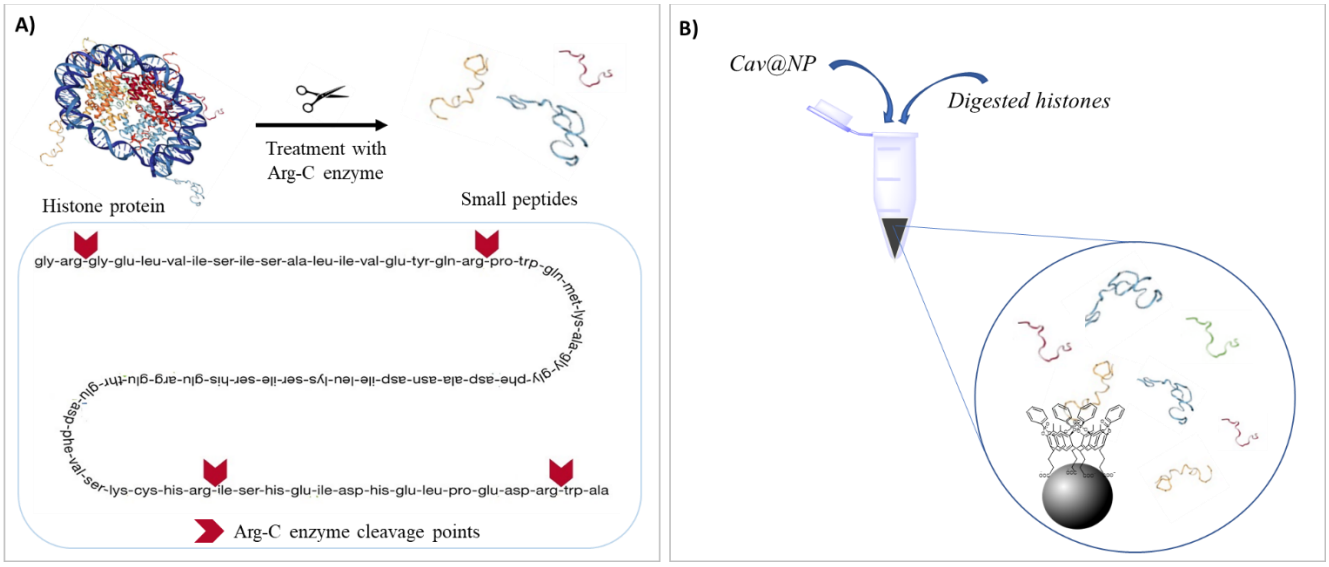
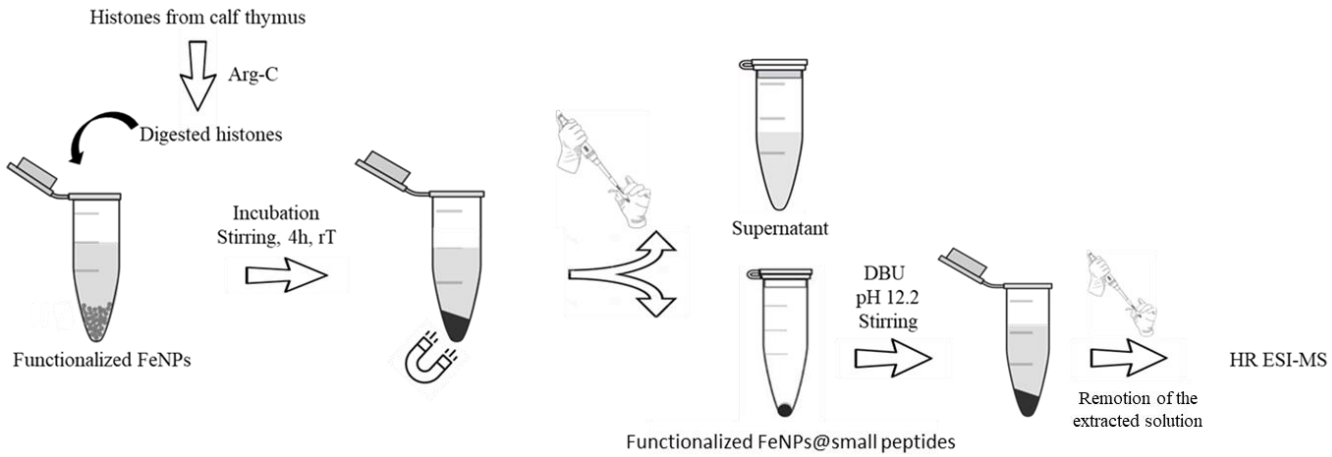


Figure 29. Representation of **A)** Arg-C enzyme cleaving the histone sequences, **B)** Recognition of the small peptides by Cav@FeNPs.

For this new experiment, we decided to use a commercial batch of histone, derived from calf thymus, and the followed steps are reported in Scheme 6: we started with the histone digestion, then incubation with the NPs and finally the ESI-MS tests of the obtained samples, as detailed below.



Scheme 6. Representation of recognition experiment protocol. Before the exposure to the functionalized NPs, histones from calf thymus were subjected to proteolytic digestion.

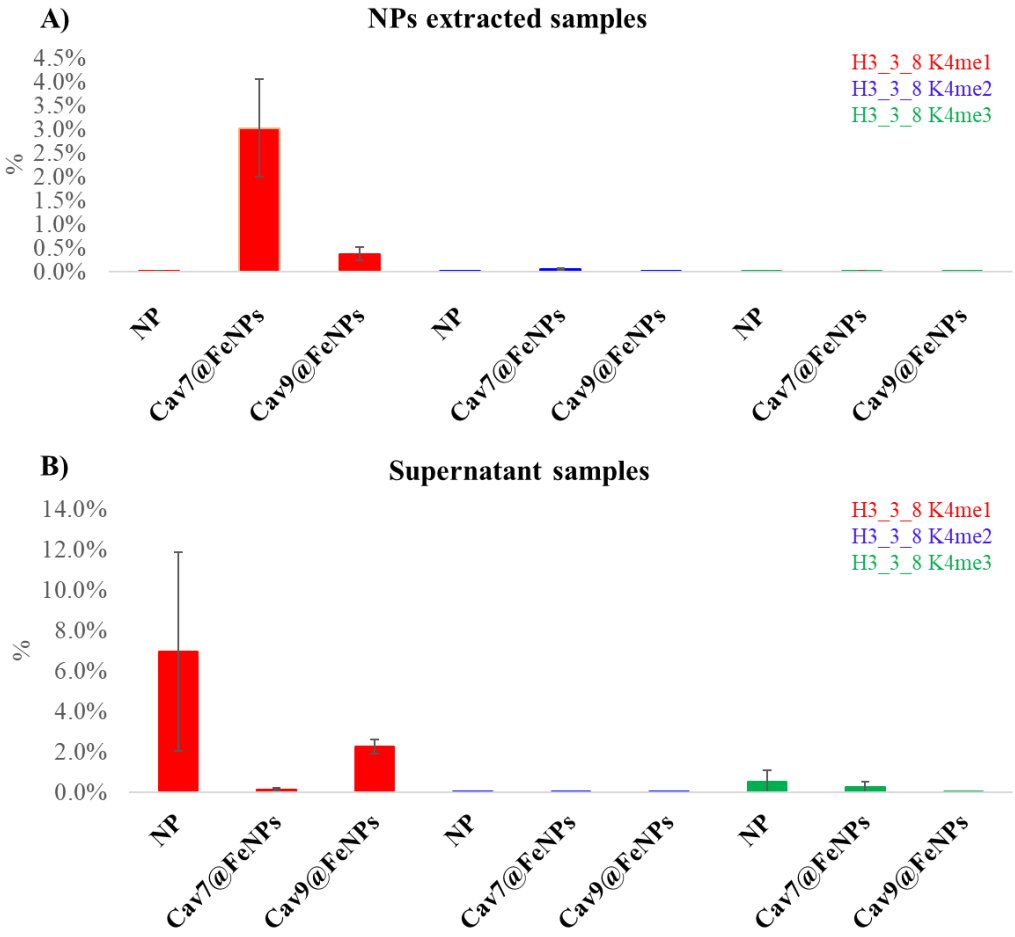
For the digestion step, the histone batch was solubilized in a water-based incubation buffer (pH= 7.82), composed by TRIS-HCl buffer (50 mM), $CaCl_2$ (5 mM), EDTA (2 mM) and DTT (50 mM). Arg-C, suspended in the same buffer, was then added. After incubation at 37°C for 14.5 hours, the

obtained digested histones were heated at 95 °C for 10 minutes to quench the reaction. The sample was then lyophilized and stored at -20°C.

The tests (5 replicates for each batch) were carried out suspending bare NPs, and Cav@FeNPs in milliQ water (4.5 mg and 6.8 mg respectively, in 100 μL of milliQ water) and, then 100 μL of the digested histones solution in the buffer prepared before were added. Both bare NPs and Cav**9**@FeNPs are used as control experiment. The incubation was performed at room temperature, for 4 hours. Once the incubation period was over, the supernatant was removed and lyophilized.

The solid residue composed by the FeNPs was subjected to a basic treatment to extract the complexed peptide fragments by deprotonation of the ammonium salt. In particular, 1,8-diazobicyclo[5.4.0]undec-7-ene (DBU) was dissolved in water (pH 12.2), added to the NPs previously suspended in water, and the obtained mixture stirred using a vortex mixer to homogenize the resulting suspension. Finally, the FeNPs were let to precipitate with the help of a magnet. The procedure was repeated twice to be sure to extract all the complexed small peptides. Subsequently, both the obtained extracted samples were mixed, lyophilized, and sent to Prof. S. Sidoli for HR ESI-MS analysis.

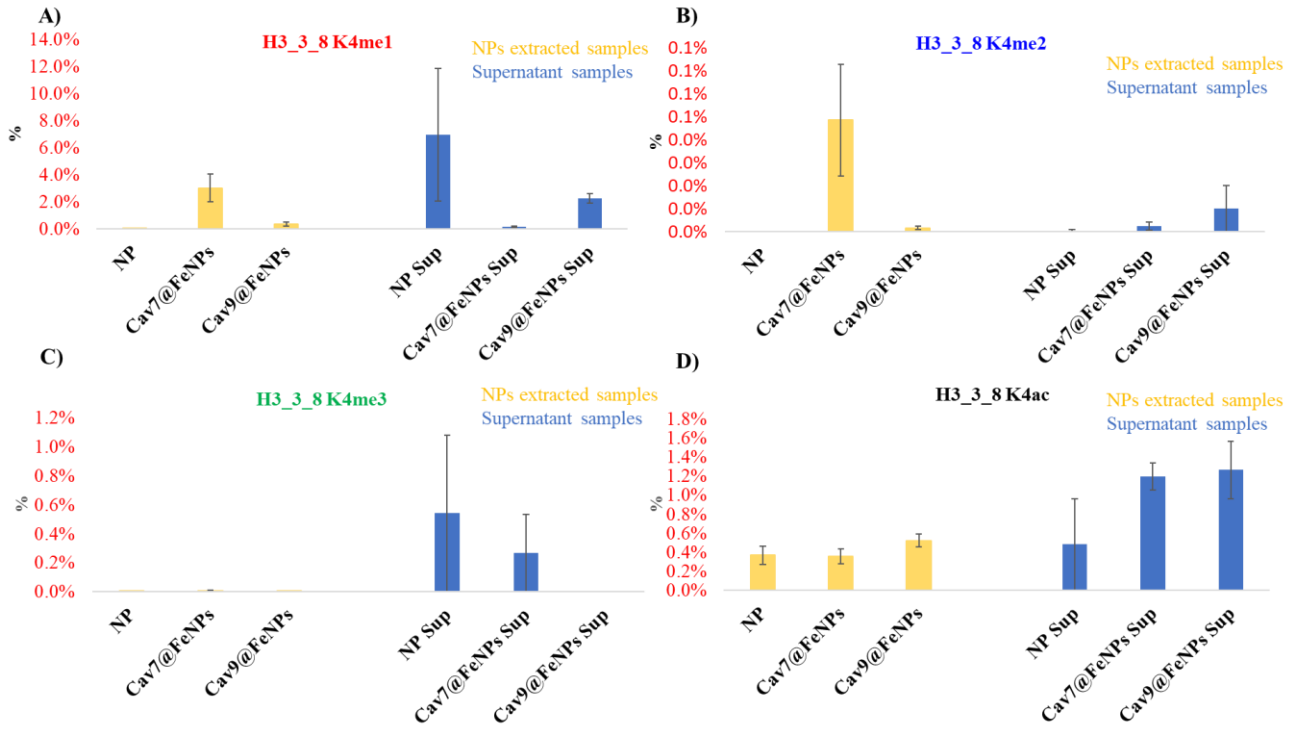
Once more, we focused our attention on the histone family H3 derived peptide TKQTAR (H3_3_8), which is known to be the most exposed to methylation modifications.[39] Graph 2 illustrates the results obtained for bare NPs, Cav**7**@FeNPs and Cav**9**@FeNPs as enrichment % toward the mono-, di- and trimethylated lysine residues present on the H3_3_8 peptide.



Graph 2. Enrichment % of methylated lysine residues in **A)** NPs extracted samples; **B)** supernatants samples.

As it is visible in Graph 2 A), this time the system is able to recognize selectively the mono-methylated lysine residues (red histograms) over di- and trimethylated ones (blue and green histograms). Moreover, comparing the analysed systems, Cav**7**@FeNPs showed the best performances respect to the control systems, namely bare NPs and Cav**9**@FeNPs. It is worth to note that also the control Cav**9**@FeNPs exhibits a slight recognition ability toward mono-methylated residue, even if significantly less respect to the Cav**7**@FeNPs. This behaviour could be related to pure hydrophobic effect involved in the binding event. In presence of a hydrophobic cavity, hydrophobic molecules tend to enter into it to isolate themselves from water contact. As for pristine FeNPs, as expected they did not show any recognition ability toward all the three methylated residues. In fact, the presence of H3_3_8 K4 in the extracted samples of bare NPs is almost negligible (Graph 2, A) for all the methylation state, while it is predominant in the supernatant sample (Graph 2, B).

In Graph 3 A), B) and C), we reported separately the results obtained for each methylation state of K4 present in H3_3_8, comparing the NPs extracted samples and the supernatant samples for all the sensing systems used. In Graph 3 D), the results obtained for H3_3_8 K4ac are reported.



Graph 3. Comparison of enrichment % in NPs extracted samples (NP, Cav**7**@FeNPs, Cav**9**@FeNPs) and supernatant samples (NP Sup, Cav**7**@FeNPs Sup, Cav**9**@FeNPs Sup) in: **A)** mono-methylated lysine residue; **B)** di-methylated lysine residue; **C)** tri-methylated lysine residue; **D)** acetylated lysine residue. Replicates: NPs n = 2; Cav**7**@FeNPs n = 5; Cav**9**@FeNPs n = 5.

As already highlighted in Graph 2, in Graph 3, A) it is evident that Cav**7**@FeNPs showed the best recognition properties toward H3_3_8 Kme1 respect to the other sensing systems used, namely NP and Cav**9**@FeNPs, and the presence of H3_3_8 K4me1 in the supernatant is almost negligible (blue bar), while it is not for the two reference systems. Moreover, for Cav**7**@FeNPs the % of enrichment calculated for H3_3_8K4me1 is more than one order of magnitude higher respect H3_3_8 K4me2 and H3_3_8 K4me3 (Graph 2 B) and C), respectively). As regards to H3_3_8 Kme2 (Graph 3, B), the performances of Cav**7**@FeNPs are higher respect to the two control experiments, while considering H3_3_8 K4me3 (Graph 3, C), all the sensing systems resulted inefficient in complexation, since the majority of this residue is present in the supernatant samples. Generally, in a biological sample the % values of di- and tri-methylated residues are lower respect to the mono-methylated ones and often under the limit of detection of the instrument, leading to not reliable results and high errors in mass analysis, as evident in Graph 3 B) for NP and in Graph 3C) for Cav9@FeNPs. In Graph 3 D), the analysis results obtained for acetylated K4 are reported. In this case, all the tested systems performed similar, with an equal distribution of H3_3_8 K4ac between NPs extracted and supernatant samples. This is a confirm that the target system Cav**7**@FeNPs is selective in complexing mono-methylated lysine residues even in presence of other modifications.

2.3 Conclusions

In this chapter the focus was on the preparation of FeNPs functionalized with cavitand receptors, to exploit the multivalency effect for the selective recognition of N-monomethylated lysine residues presents on histone proteins. The prior digestion of the histone protein samples to avoid folding of the long peptide chains and to make the modifications accessible to the sensing systems resulted fundamental. The selected Cav**7**@FeNPs sensing system displayed the best performances in recognizing mono-methylated lysine residues, as compared to the two control systems, namely bare NPs and Cav**9**@FeNPs. In particular, bare FeNPs did not demonstrate, as expected, recognition properties towards any peptide modifications, while Cav**9**@FeNPs demonstrated a slight binding ability toward mono- and di-methylated lysine residues, in any case significantly less respect to Cav**7**@FeNPs, and it can be mainly attributed to a pure hydrophobic effect. Future developments of the project will concern the preparation of sensing systems based on the heterogeneous functionalization of FeNPs with three different synthetic receptors: Tiiii cavitand, cucurbit[7]uril and sulphonato-calixarene, selective towards the mono-, di- and tri-methylated lysine, respectively ,with the aim of enriching the mixture in all the methylation degree, important aspect in the fields of proteomics, chromatin biology and epigenetics.

2.4 Experimental section

2.4.1 Reagent, Chemicals and Instruments used

All reagents and solvent were purchased from certified commercial sources and used as received, without further purification. For the synthesis, all solvents were dried and distilled according to standard procedures known in the literature.[42]

Thin layer chromatography was performed using TLC Analytical Chromatography F254®, Merck.

^{1}H and ^{31}P NMR spectra were recorded on Bruker 400MHz AVANCE or Jeol 600MHz using DMSO-d_6, Acetone-d_6, CD_3OD and $CDCl_3$ as solvents. All chemical shifts (δ) were reported in parts for million (ppm) relative to proton resonances resulting from incomplete deuteration of NMR solvents. The abbreviations: s, d, t, m and sb indicated the spectrum peaks referred to: singlet, doublet, triplet, multiplet and broad singlet, respectively. The coupling constant (J) are expressed in Hz.

ESI-MS characterization experiments were performed on a Waters ACQUILITY Ultra Performance LC HO6UPS-823M with ESI source ionization (electrospray ionization) in positive modality.

IR-ATR analyses were performed using Bruker FTIR LUMOS.

TGA analysis have been achieved using TGA/DSC 1 with gas controller GC10 (pure air/nitrogen), Mettler Toledo.

TEM analyses were performed using TEM JEOL JEM-2200FS with Schottky field emission gun. Operating voltage: 80 kV, 200 kV; Point resolution: 0.183 nm (TEM) and 0.132 nm (STEM); two High Angle Annular Dark Field detectors (Z-contrast); column Omega filter.

Zeta Potential analysis were performed using Malvern Instrument, Zetasizer Ver. 7.12, using a DTS 1070 disposable cuvette.

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 HCl 37% (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 $[M+H]^+$.

Cavitand (2)

In a Shlenk reactor, under anhydrous conditions, 3.00 g (3.86 mmol) of **Res [C_3H_6OH; Me]** were dissolved in 10 mL of anhydrous pyridine. The solution was cooled down to 0 °C through an ice bath, and 8.36 mL (69.53 mmol) of $(CH_3)_2SiCl_2$ were added under magnetic stirring. The reaction mixture was kept at 0°C under magnetic stirring for about 20 minutes, and then heated to 100 °C for 3 hours. The reaction was cooled down to room temperature and the solvent was removed under high vacuum using a trap. The solid obtained was suspended in methanol, filtered, washed with methanol and finally dried. Cavitand **2** was obtained pure, as a white solid, in 64% yield. **^{1}H NMR** ($CDCl_3$, 400 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 $[M+H]^+$.

Cavitand (3)

To a solution of 3.00 g (1.00 mmol) of Cavitand **2** in 12 mL of anhydrous DMF, diisopropylethylamine (2.79 mL, 15.99 mmol) and chloromethyl methyl ether (0.79 mL, 10.5 mmol) were added under nitrogen. The suspension was stirred at 40 °C overnight. Then, the reaction was cooled down to room temperature and the solvent evaporated under reduced pressure. The obtained precipitate was suspended in water, filtered, washed with water and dried, to give cavitand **3**, as light brown solid (yield 90%). **^{1}H NMR** ($CDCl_3$, 400 MHz): δ ppm = 7.17 (s, 4H, Ar**H**$_{down}$), 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_2OMOM, 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] (4)

In a Teflon flask, to a solution of cavitand **3** (1.06 g, 0.90 mmol) in 15 mL of DMF, HF 40 % (1.00 mL, 9.62 mmol) was added. The suspension was stirred overnight at 50 °C. Then, the solvent was

removed under reduced pressure and the addition of water promoted the precipitation of the product. The obtained solid was filtered, washed with water, and dried to give product **4** as a white solid, in 93% yield. **^{1}H NMR** (DMSO-d_6, 400 MHz): δ ppm = 8.68 (s, 8H, Ar**OH**), 7.29 (s, 4H, Ar**H**$_{down}$), 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 $[M+H]^+$.

Tiii Cavitand (5)

To a solution of resorcinarene **4** (0.80 g, 0.84 mmol) in 16 mL of anhydrous pyridine, dichlorophenyl phosphine (0.49 mL, 3.61 mmol) was added under nitrogen atmosphere. The solution was stirred at 80 °C for 3 hours. Then, the mixture was cooled down to 0 °C with an ice bath and 2 mL of 35% aqueous H_2O_2 were added. The reaction was stirred at room temperature for 1 hour, under magnetic stirring and nitrogen flux. The crude was poured into cold water and a white,a milky solution was obtained and kept refrigerated at 2 °C overnight. The formed precipitate was filtered, washed with water, and dried. The obtained solid was dissolved in DMC and precipitate with hexane to give pure cavitand **5** as a white solid with a yield of 71%. **^{1}H NMR** ($CDCl_3$, 400 MHz): δ ppm = 8.13 (m, 8H, P(O)Ar**H$_o$**), 7.64 (m, 4H, P(O)Ar**H$_p$**), 7.57 (m, 8H, P(O)Ar**H$_m$**), 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$_2$OMOM). **^{31}P NMR** ($CDCl_3$, 162 MHz): δ ppm = 7.13 (s, P=O). **ESI-MS:** m/z 1464.36 $[M+Na]^+$.

Tiii Cavitand (6)

Cavitand **5** (0.86 g, 0.60 mmol) was dissolved in 6 mL of $CHCl_3$ and 12 mL of MeOH. HCl 37% (1.37 mL) was added and the solution was stirred overnight at 40 °C. The solvent was evaporated obtaining a white solid that was triturated with water, acetone and finally filtered and dried to give cavitand **6**, in 86% yield. **^{1}H NMR** (MeOD-d_4, 400 MHz): δ ppm = 8.13 (m, 8H, P(O)Ar**H$_o$**), 7.76 (m, 4H, P(O)Ar**H$_p$**), 7.65 (m, 8H, P(O)Ar**H$_m$**), 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$_2$OH). **^{31}P NMR** ($CDCl_3$, 162 MHz): δ ppm = 8.88 (s, P=O). **ESI-MS:** m/z 1297.81 $[M+K]^+$.

Tiii Cavitand (7)

To a solution of Cavitand **6** (0.50 g, 0.40 mmol) in 12 mL of DMF and 12 mL of acetone, CrO_3 (0.63 g, 6.34 mmol) was added under magnetic stirring. Then, H_2SO_4 (96% aqueous) was added until pH 2 was reached and the solution was stirred at room temperature for 5 hours. The dark green solution was concentrate evaporating the solvent in high vacuum trap, then acidic water was added

maintaining pH 2, and the obtained suspension was centrifugated (METTERE RPM E MIN). The solid residue was washed many times with water till it reached a pale green colour. Cavitand **7** was obtained in 55% yield. **^{1}H NMR** (MeOD-d_4, 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** (MeOD-d_4, 162 MHz): δ ppm = 10.2 (s, P=O). **^{13}C NMR** (MeOD-d_4, 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-MS:** m/z calculated for $C_{68}H_{59}O_{10}P_4^-$ 1319.2556 $[M-H]^-$, found 1319.2536. **ATR-FTIR**: 2927 cm^{-1} C-H (aliphatic) and O-H; 1714 cm^{-1} C=O; 1068 cm^{-1} P=O; 900 cm^{-1} O-P-O. **TGA** (air, 25°C - 900 °C, flow rate: 20 °C/min): 0.4 % w/w loss at 29 °C; 9.5 % w/w loss at 266 °C; 20.6 % w/w loss at 441 °C; 7.6 % w/w loss at 523 °C; 16.1 % w/w loss at 574 °C; 26.7 % w/w loss at 710 °C.

Cavitand (8)

Resorcinarene **1** (0.60 g, 0.77 mmol) was dissolved in 14.5 mL of anhydrous DMF. K_2CO_3 (1.07 g, 7.73 mmol) and CH_2I_2 (1.56 mL, 30.86 mmol) were added, and the solution was stirred at 80 °C for 5 hours. The reaction was cooled down to room temperature and the crude was quenched in 200 mL of water weakly acidified with HCl (37% aqueous). The obtained orange precipitate was filtered and dried to give cavitand **8** with a quantitative yield, without the need of further purification. **^{1}H NMR** (DMSO-d_6, 400 MHz): δ ppm= 7.44 (s, 4H, Ar**H**), 5.88 (d, 4H, ArOC**H_{2out}**O, J = 7.48 Hz), 4.60 (t, 4H, -C**H**-, J = 8.18 Hz), 4.47 (bs, 4H, -CH_2O**H**), 4.20 (d, 4H, ArOC**H_{2in}**O, J = 7.44 Hz), 3.50 (m, 8H, -C**H_2**OH), 2.36 (m, 8H, CHC**H_2**-), 1.90 (s, 12H, ArC**H_3**), 1.43 (m, 8H, -C**H_2**CH_2OH).**ESI-MS**: m/z 770.39 $[M+H]^+$.

Cavitand (9)

To a solution of cavitand **8** (0.15 g, 0.18 mmol) in 6 mL of DMF and 6 mL of acetone, CrO_3 (0.29 g, 3.88 mmol) was added under magnetic stirring. Then, H_2SO_4 (96% aqueous) was added until pH 2 was reached and the solution was stirred at room temperature for 5 hours. The dark green solution was concentrated evaporating the solvent in high vacuum trap, then acidic water was added maintaining pH 2 and, the obtained suspension was centrifugated (10 minutes). The solid residue was washed many times with water till it reached a pale green colour. The final product **9** was obtained as a pale green powder in 80% yield. **^{1}H NMR** (Acetone-d_6, 400 MHz): δ ppm = 7.60 (s, 4H, Ar**H**), 5.93 (d, 4H, ArOC**H_{2out}**O, J = 7.20 Hz), 4.86 (t, 4H, -C**H**-, J = 8.30 Hz), 4.47 (bs, 4H, -CH_2O**H**), 4.31 (d, 4H, ArOC**H_{2in}**O, J = 7.32 Hz), 2.70 (m, 8H, CHC**H_2**-), 2.36 (t, 8H, -C**H_2**COOH, J = 7.52 Hz), 1.99 (s, 12H, ArC**H_3**). **HR-ESI-MS**: 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**.** **ATR-FTIR**: 2925 cm^{-1} C-H

(aliphatic) and O-H; 1704 cm^{-1} C=O; 1428 cm^{-1} C-O-C. **TGA** (air, 25°C - 800 °C, flow rate: 20 °C/min): 8.6 % w/w loss at 117 °C; 44.7% w/w loss at 388 °C; 13.1 % w/w loss at 471 °C.

Synthesis of ferromagnetic nanoparticles (FeNPs)

Iron magnetite nanoparticles were synthesized by alkaline co-precipitation of Fe^{3+} and Fe^{2+}. In two separate vials, $FeCl_3 \cdot 6H_2O$ (0.73 g, 2.70 mmol) and $FeCl_2 \cdot 4H_2O$ (0.27 g, 1.36 mmol) were solubilized in 3 mL and 1.5 mL of water, respectively. In a flask, 2.2 mL of NH_3 (30% aqueous) were added to 50 mL of water and, keeping the system under manual stirring, the two solutions of $FeCl_3$ and $FeCl_2$, were added in quick succession. A dark coloured suspension was immediately observed and left to decantation at room temperature for 2 hours. The supernatant was filtered off and the solid residue washed twice with water. The obtained FeNPs were isolated from the solvent by magnetic decantation and dried under nitrogen flux. **ATR-FTIR**: 3000 cm^{-1} OH, 551 cm^{-1} Fe-O.

Functionalized FeNPs

Considering the reagents conditions reported in Table 2, functionalized FeNPs were prepared following the protocol below.

In two different flasks, $FeCl_2 \cdot 4H_2O$ and $FeCl_3 \cdot 6H_2O$ was dissolved in water. Separately, the cavitand was dissolved in a previously prepared solution of NH_3, DMF and H_2O. To this solution kept under manual stirring, the two Fe^{3+} and Fe^{2+} solutions were added sequentially. The formation of a dark precipitate was visible. After magnetic decantation, the supernatant was filtered off and the solid residue washed twice with water and dried under nitrogen flux.

Table 2. Reagents conditions for the preparation of functionalized FeNPs.

Sample	Cavitand solution	$FeCl_2 \cdot 4H_2O$ solution	$FeCl_3 \cdot 6H_2O$ solution
Cav9@FeNPs	0.060 g, 0.068 mmol in 1.5 mL NH_3 + 30 mL H_2O	0.289 g, 1.43 mmol in 1.5 mL H_2O	0.79 g, 2.93 mmol in 3 mL H_2O
Cav7@FeNPs	0.10 g, 0.076 mmol in 3 mL NH_3 + 8 mL DMF+ 40.8 mL H_2O	0.34 g, 1.69 mmol in 1.7 mL H_2O	0.92 g, 3.41 mmol in 3.4 mL H_2O

Characterization of functionalised Cavitand9@FeNPs

ATR-FTIR: 2988 cm^{-1} C-H (aliphatic); 1394 cm^{-1} C=O; 1241 cm^{-1} C-O-C; 556 cm^{-1} Fe-O.

TGA (air, 25 °C - 800 °C, flow rate: 20 °C/min): 0.6 % w/w loss at 81 °C; 13.6 % w/w loss at 298 °C.

Characterization of functionalised Cavitand9@FeNPs

ATR-FTIR: 3056 cm^{-1} C-H (aliphatic); 1645 cm^{-1} C=O; 1069 cm^{-1} P=O; 898 cm^{-1} O-P-O; 567 cm^{-1} Fe-O.

TGA (air, 25°C - 800 °C, flow rate: 20 °C/min): 0.1% w/w loss at 29 °C; 4.3 % w/w loss at 259 °C; 12.4% w/w loss at 388 °C.

Extraction of histones from *Saccharomyces Cerevisiae*

The protein mixture enriched in histones content was obtained using a protocol reported in literature, that involved the use of a yeast culture of the species *Saccharomyces cerevisiae*. The protocol is based on 4 steps:

- Cell preparation: growing of the cell culture at 30 °C. The pellet was collected and wash with water.
- Separation of spheroblasts by incubating the pellet at 35 °C in presence of Zymolase.
- Extraction of histones: incubation of the spheroblasts with IGEPAL® CA-630 detergent and solubilization in 0.4N H_2SO_4, followed by precipitation adding TCA.
- Characterization by SDS-PAGE for the histone content quantification and by Bradford assay for the quantification of the total protein content:
  obtained histone content: 38.8 µg
  obtained total protein content: 171 µg

Activation buffer preparation for Arg-C (final volume in Milli-Q water: 5 mL; pH= 7.82)

- TRIS-HCl buffer: 4.01 mL, final concentration: 50 mM
- CaCl2: 2.78 mg, final concentration 5 mM
- EDTA: 2.92 mg, final concentration 2 mM
- DTT (dithiothreitol): 3.86 mg, final concentration 5 mM

Digestion of histones from calf thymus

The protocol is based on 4 steps:

- 500 µg of histones (from calf thymus, Type II-A, lyophilized powder) were solubilized in 0.5 mL of water-based incubation buffer (pH=8).
- 10 µg of commercial Arg-C enzyme were suspended in 20 µL of incubation buffer and added to the previous solution.
- The system was heated at 37 °C for 14.5 hours using a thermal cyclator.
- The digestion was quenched heating the system at 95 °C for 10minute.

Zeta Potential. Preparation of the sample

Z potential measurements were performed for both pristine and functionalised FeNPs by preparing a suspension of nanoparticles both in water (3.5 mg in 2 mL of milliQ water, pH= 7) and basic water (3.5 mg in 2 mL of DBU solution, pH= 11. The suspension was homogenized under ultrasound for 2 minutes before analysis.

2.5 Appendix

2.5.1 Multivalency

Multivalency is a key principle in nature for achieving strong but also reversible chemical interactions between two species. It refers to the simultaneous interaction between multiple functionalities on one entity and complementary functionalities on another that, can be reffered also as host and guest.[22] Interactions between an *m*-valent receptor and an *n*-valent ligand (*m,n* >1) are considered to be multivalent. If the receptor has multiple identical binding sites, they are homomultivalent. If they have multiple binding sites, which are different from each other, they are heteromultivalent. All interactions involving more than one host-guest interactions are considered multivalent. Multivalency play an important role in supramolecular chemistry since it can be used to build a controlled, selective, and directional self-assembly of increasingly complex structures or to achieve a targeted chemical nanostructuring on surface structures. There are many possible pathways that multivalent host and guest entities can follow. In the case of interactions between several multivalent units with complementary functionalities, the association modes intramolecular or intermolecular are both possible (Figure 28).

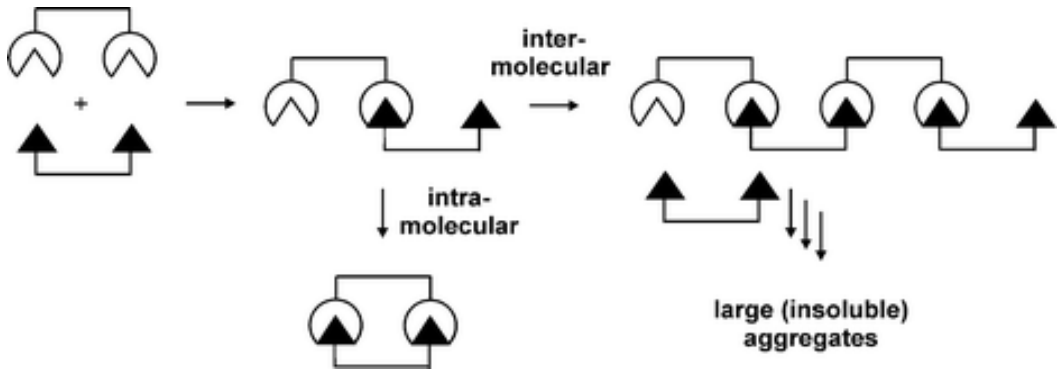


Figure 28. Multivalent *vs.* intermolecular binding.

A multivalent interaction normally gives rise to structures with defined geometries and binding motifs while the intermolecular binding leads to the formation of insoluble polymeric aggregates.[43] In this perspective, to achieve the multivalent binding is required a certain degree of system design. The critical aspects mainly concern architecture (size and shape) of the interacting entities that have a

strong influence on the mode binding and must be suitable for the establishment of multivalent bond. The nature of molecules prone to formation of multivalent bonds varies from combinations of small and relatively rigid three-dimensional entities, such as articles and dendrimers, to self-assembled systems in which host or guest functionalities are organised in a highly directional two-dimensional plane. If the entities architecture involved is suitable for the establishment of the multivalent bond, the nature of the interaction is dictated by the enthalpy and entropy variation that accompanied the binding event. By thermodynamic, the idea is that the multivalency is mainly governed by entropy and enthalpy is assumed to be proportional to the number of interactions.[43]

Another important parameter in multivalent binding is the effect of lengths and flexibility of the linkers between interacting sites that is consider by the effective concentration parameter (C_{eff}). It is based on the probability of interaction between two reactive or complementary interlinked entities and more generally is expressed using effective molarity (EM), Equation 1.

$$EM = \left(\frac{K_n}{bK_i^n} \right)^{1/(n-1)}$$

Equation 1. Effective molarity formula.

C_{eff} is based on concentrations calculated from physical geometries of complexes while, EM quantifies the advantage for intra- over intermolecular interaction[44] according to equation 1, in which K_n is the association constant for the *n*-variant interaction that is related to the intrinsic association constant K_i, and *b* is a scaling factor incorporating statistical factors determining the numbers of possible association and dissociation pathways in the subsequent interaction steps. Multivalency is having great importance not only for medicine[45] and biochemistry[46], but also for the design of new functional molecules in supramolecular chemistry and material sciences. In fact, synthetic multivalent architectures, using noncovalent bonding interactions can provide well-defined system for studying multivalency in nature and building blocks for the construction of nanoscale materials.

2.5.2 IR and TGA characterization of cavitand 7 and 9.

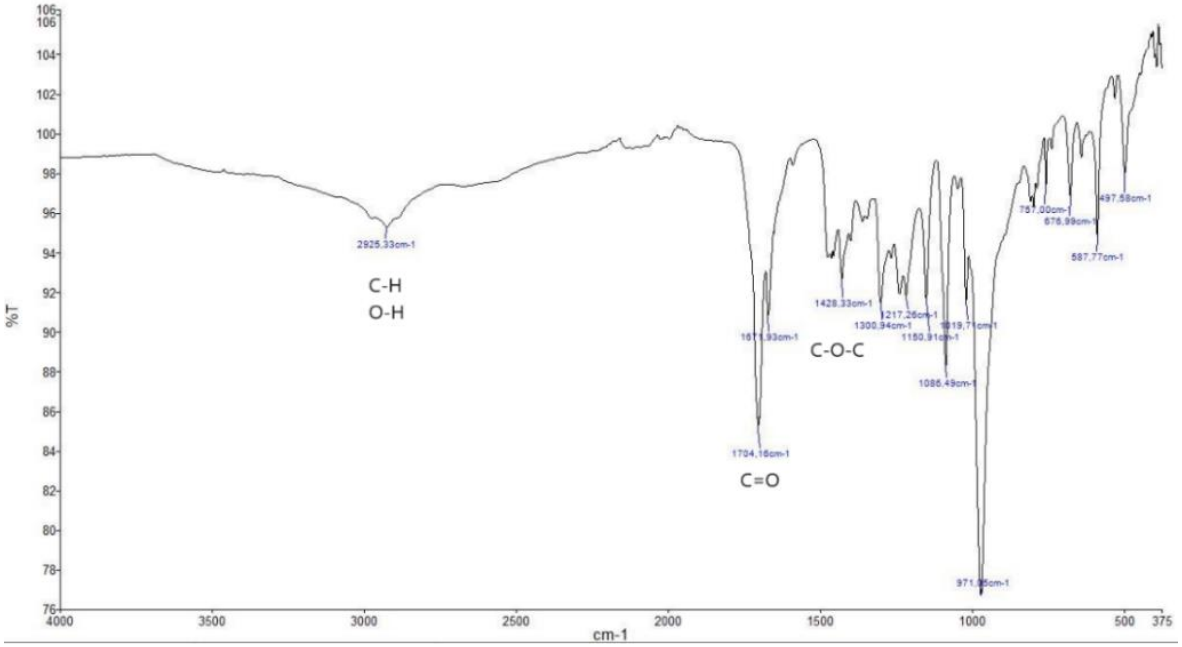


Figure 29. ATR-FTIR spectrum of cavitand **9**.

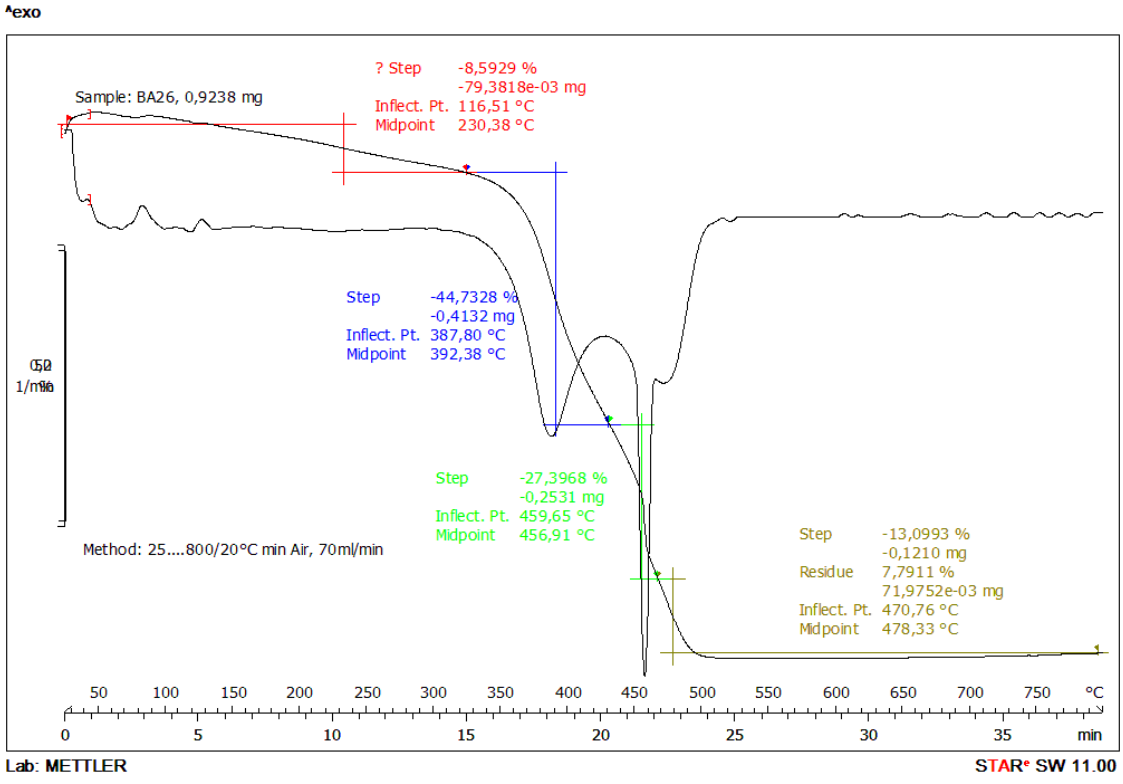


Figure 30. TGA analysis of cavitand **9**.

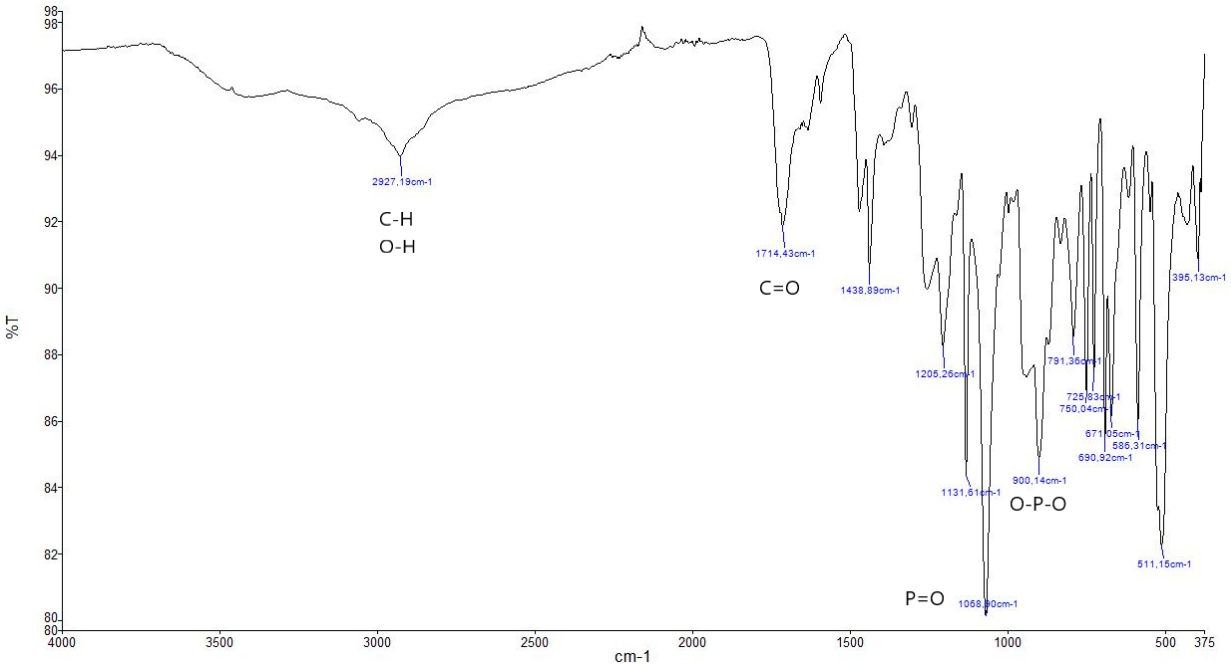


Figure 31. ATR-FTIR spectrum of cavitand **7**.

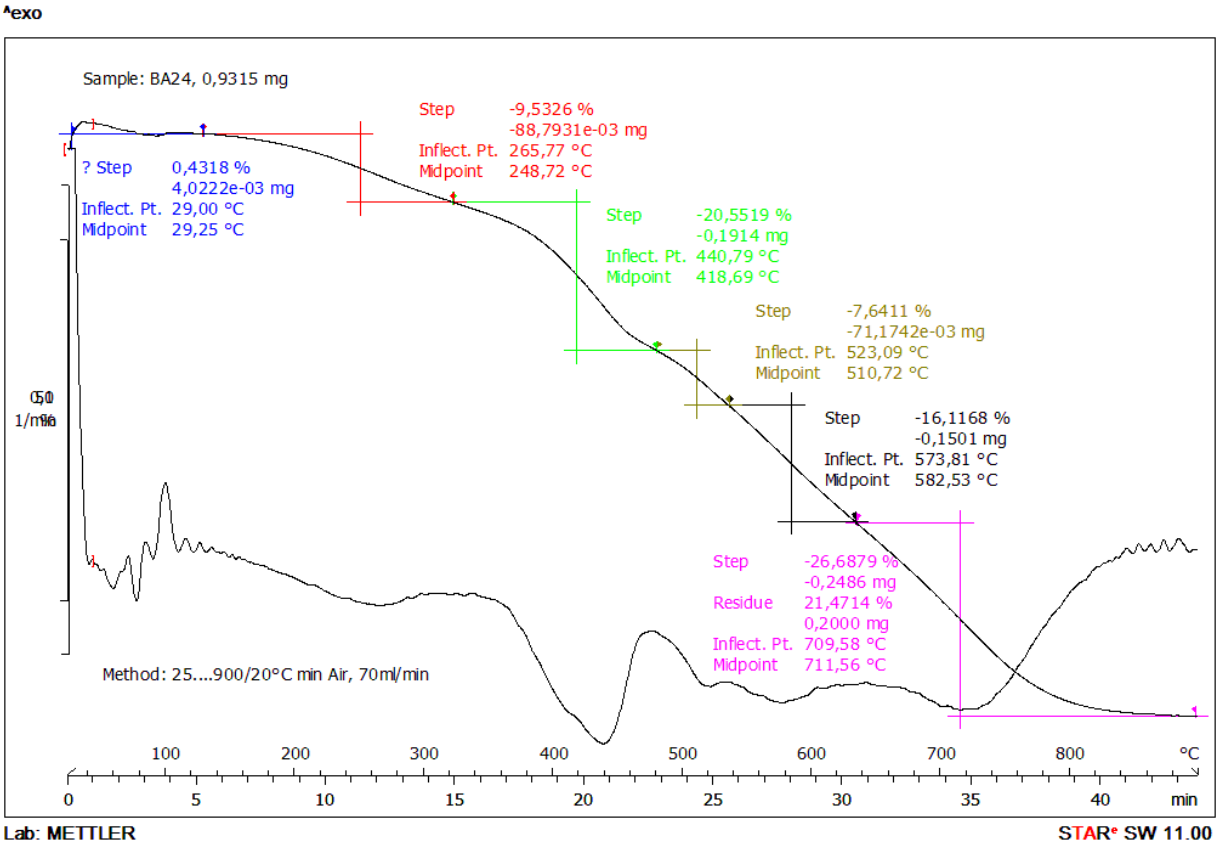


Figure 32. TGA analysis of cavitand **7**.

2.5.3. TGA characterization of Cav7@FeNPs and Cav9@FeNPs

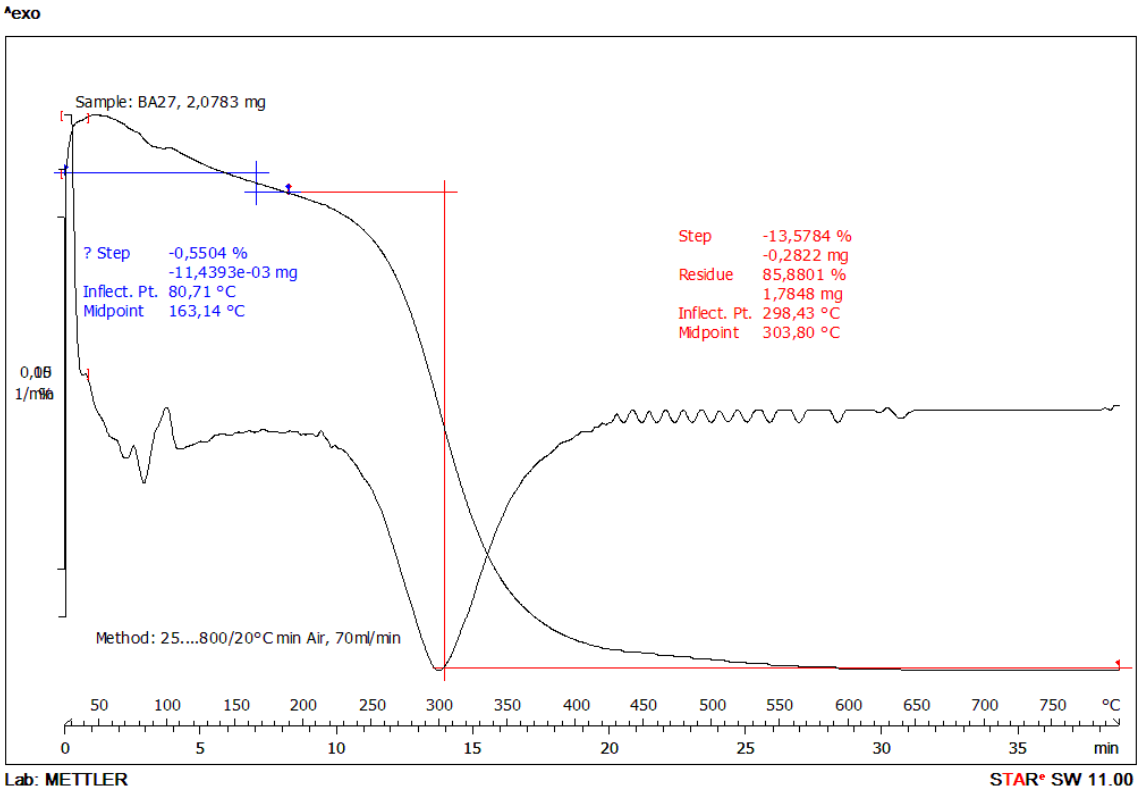


Figure 33. TGA analysis of Cav**9**@FeNPs.

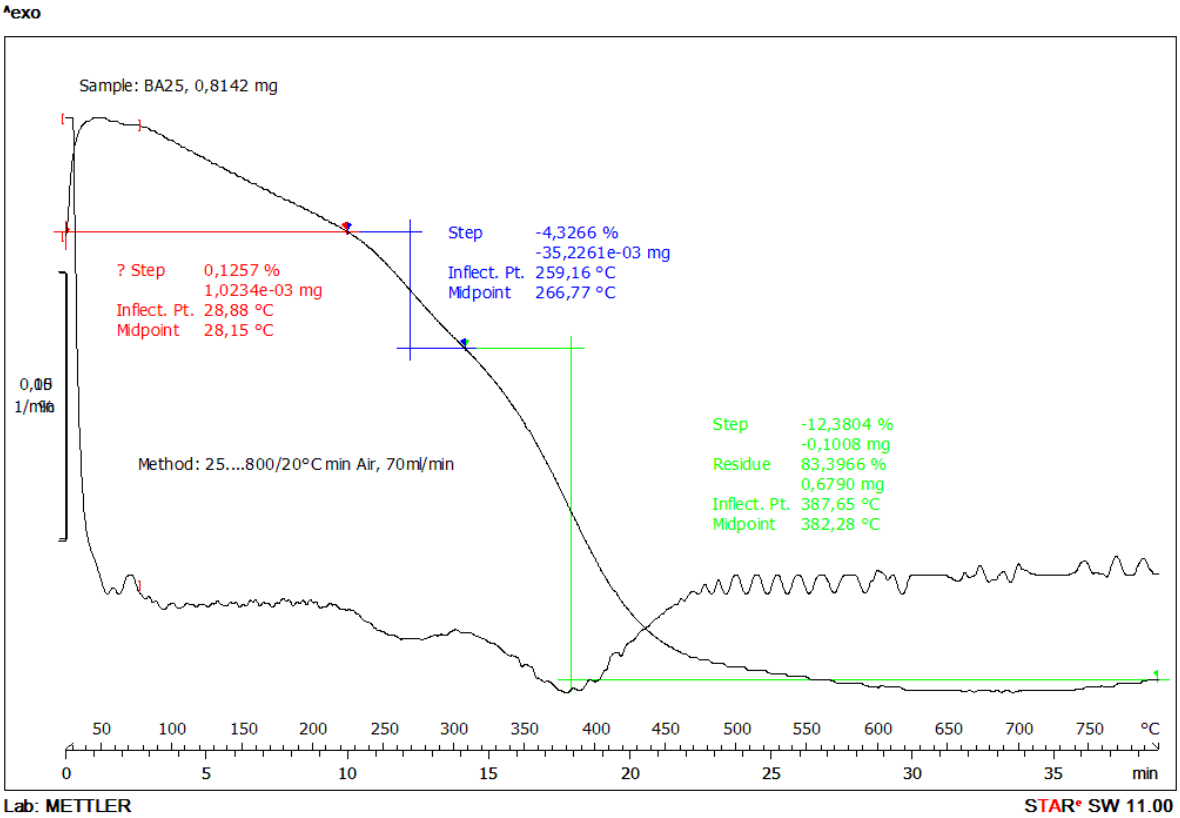


Figure 34. TGA analysis of Cav**7**@FeNPs.

2.6 References

1. Fan, J., Krautkramer, K. A., Feldman, J. L. & Denu, J. M. Metabolic regulation of histone post-translational modifications. *ACS Chem. Biol.* **10**, 95–108 (2015).
2. Hyland, E. M. *et al.* Insights into the Role of Histone H3 and Histone H4 Core Modifiable Residues in Saccharomyces cerevisiae . *Mol. Cell. Biol.* **25**, 11193–11193 (2005).
3. Amazi, S. H. R. & Llahverdi, Evaluation of post-translational modifications in histone proteins: A review on histone modification defects in developmental and neurological disorders. *J. Biosci.* **45**, 3–9 (2020).
4. Garcia, B. A., Shabanowitz, J. & Hunt, D. F. Characterization of histones and their post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.* **11**, 66–73 (2007).
5. Tessarz, P. & Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Publ. Gr.* **9**, 703-708 (2014).
6. Cavalieri, V. The expanding constellation of histone post-translational modifications in the epigenetic landscape. *Genes (Basel).* **12**, 1596 (2021).
7. Lorton, B. M. & Shechter, D. Cellular consequences of arginine methylation. *Cell. Mol. Life Sci.* **76**, 2933–2956 (2019).
8. Leroy, G. *et al.* A quantitative atlas of histone modification signatures from human cancer cells. *Epigenetics and Chromatin* **6**, 1–14 (2013).
9. Miao, F. *et al.* Evaluating the role of epigenetic histone modifications in the metabolic memory of type 1 diabetes. *Diabetes* **63**, 1748–1762 (2014).
10. Walsh, G. & Jefferis, R. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* **24**, 1241–1252 (2006).
11. Barski, A. *et al.* Resource High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129**, 823–837 (2007).
12. Egger, G., Liang, G., Aparicio, A. & Jones, P. A. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429**, 457–463 (2004).
13. Thomas, S. N. *et al.* Dual modification of Alzheimer's disease PHF-tau protein by lysine methylation and ubiquitylation: A mass spectrometry approach. *Acta Neuropathol.* **123**, 105–117 (2012).
14. Kalakonda, N. *et al.* Histone H4 lysine 20 monomethylation promotes transcriptional repression by L3MBTL1. *Oncogene* **27**, 4293–4304 (2008).
15. *www.myuptodate.com*.
16. Pinalli, R., Pedrini, A. & Dalcanale, E. Biochemical sensing with macrocyclic receptors. *Chem. Soc. Rev.* **47**, 7006–7026 (2018).
17. D. J. Cram. Cavitands: organic hosts with enforced cavities. *Science* **219**, 1177–1183 (1983).
18. Pinalli, R. & Dalcanale, E. Supramolecular sensing with phosphonate cavitands. *Acc. Chem. Res.* **46**, 399–411 (2013).
19. Dionisio, M. *et al.* Nanomechanical recognition of N -methylammonium salts. *J. Am. Chem. Soc.* **134**, 2392–2398 (2012).

20. Pinalli, R. *et al.* The Origin of Selectivity in the Complexation of N-Methyl Amino Acids by Tetraphosphonate Cavitands. *J. Am. Chem. Soc.* **138**, 8569–8580 (2016).

21. Bontempi, N. *et al.* Probing lysine mono-methylation in histone H3 tail peptides with an abiotic receptor coupled to a non-plasmonic resonator. *Nanoscale* **9**, 8639–8646 (2017).

22. Haag, R. Multivalency as a chemical organization and action principle. *Beilstein J. Org. Chem.* **11**, 848–849 (2015).

23. Price, P. M., Mahmoud, W. E., Al-ghamdi, A. A. & Bronstein, L. M. Magnetic Drug Delivery : Where the Field Is Going. *Frontiers in Chemistry* **6**, 1–7 (2018).

24. Polyak, B. & Friedman, G. Magnetic targeting for site-specific drug delivery: applications and clinical potential. *Expert Opinio on Drug Delivery* **6(1)**, 53-70 (2009).

25. Yin, X. *et al.* Large T 1 contrast enhancement using superparamagnetic nanoparticles in ultra-low field MRI. *Scientific Reports* **8(1)**, 11863 (2018).

26. Ferna, I., Lechuga-vieco, A. V, Enr, J. A., Morales, M. P. & Herranz, F. One-Step Fast Synthesis of Nanoparticles for MRI: Coating Chemistry as the Key Variable Determining Positive or Negative Contrast. *Langmuir* **33,39**, 10239–10247 (2017).

27. Matos, M. Microemulsion Synthesis of Superparamagnetic Nanoparticles for Bioapplications. *Int. J. Mol. Sci* **22(1)**, 427 (2021).

28. Zia, M., Phull, A. R. & Ali, J. S. Challenges of Iron Oxide Nanoparticles. *Powder Technol.* **7**, 49–67 (2016).

29. Hasany, S. F., Ahmed, I., Rajan, J. & Rehman, A. Systematic Review of the Preparation Techniques of Iron Oxide Magnetic Nanoparticles. *Nanoscience and Nanotechnology* **2(6)**, 148–158 (2012).

30. Wu, S. *et al.* Fe 3 O 4 magnetic nanoparticles synthesis from tailings by ultrasonic chemical. *Mater. Lett.* **65**, 1882–1884 (2011).

31. Jos, J. *et al.* One-minute and green synthesis of magnetic iron oxide nanoparticles assisted by design of experiments and high energy ultrasound : Application to biosensing and immunoprecipitation. *Materials Science & Engineering* **123**, 112023 (2021).

32. Benyettou, F. *et al.* Toward theranostic nanoparticles: CB[7] -functionalized iron oxide for drug delivery and MRI. *J. Mater. Chem. B* **1**, 5076–5082 (2013).

33. Chin, S. F., Makha, M., Raston, L. & Saunders, M. Magnetite ferrofluids stabilized by sulfonato-calixarenes *Chem. Commun.* **19**, 1948-1950 (2007).

34. Ayob, A., Alias, S., Dahalan, F. A. & Santiagoo, R. Kinetic removal of Cr^{6+} by carboxymethyl cellulose-stabilized nano zerovalent iron particles *Macedonian Journal of Chemistry and Chemical Enineering* **34(2)**, 295-308 (2015).

35. Khalil, M. I. Co-precipitation in aqueous solution synthesis of magnetite nanoparticles using iron(III) salts as precursors. *Arab. J. Chem.* **8**, 279–284 (2015).

36. Mäkie, P., Persson, P. & österlund, L. Adsorption of trimethyl phosphate and triethyl phosphate on dry and water pre-covered hematite, maghemite, and goethite nanoparticles. *J. Colloid Interface Sci.* **392**, 349–358 (2013).

37. Alfredo Reyes Villegas, V. *et al.* Synthesis and characterization of magnetite nanoparticles for photocatalysis of nitrobenzene. *J. Saudi Chem. Soc.* **24**, 223–235 (2020).

38. Vasić, K. *et al.* Structural and magnetic characteristics of carboxymethyl dextran coated magnetic nanoparticles: From characterization to immobilization application. *React. Funct. Polym.* **148**, 104481 (2020).

39. Separovich, R. J. & Wilkins, M. R. Ready, SET, Go: Post-translational regulation of the histone lysine methylation network in budding yeast. *J. Biol. Chem.* **297**, 100939 (2021).

40. Shechter, D., Dormann, H. L., Allis, C. D. & Hake, S. B. Extraction, purification and analysis of histones. *Nat. Protoc.* **2**, 1445–1457 (2007).

41. Yebeutchou, R. M. & Dalcanale, E. Highly selective monomethylation of primary amines through host-guest product sequestration. *J. Am. Chem. Soc.* **131**, 2452–2453 (2009).

42. Williams, D. B. G. & Lawton, M. Drying of organic solvents: Quantitative evaluation of the efficiency of several desiccants. *J. Org. Chem.* **75**, 8351–8354 (2010).

43. Mulder, A. & Reinhoudt, D. N. Multivalency in supramolecular chemistry and nanofabrication. *Org. Biomol. Chem*. **2**, 3409–3424 (2004).

44. Huskens, J. *et al.* A model for describing the thermodynamics of multivalent host-guest interactions at interfaces. *J. Am. Chem. Soc.* **126**, 6784–6797 (2004).

45. Mammen, M., Choi, S. K. & Whitesides, G. M. Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chemie - Int. Ed.* **37**, 2754–2794 (1998).

46. Vance, D., Martin, J., Patke, S. & Kane, R. S. The design of polyvalent scaffolds for targeted delivery. *Adv. Drug Deliv. Rev.* **61**, 931–939 (2009).

Chapter 3

Calix[4]pyrroles for the recognition of acetylated lysine residues in peptides

* The work reported in this chapter was carried out under the supervision of Prof. P. Ballester, The Institute of Chemical Research of Catalonia (ICIQ), Tarragona, Spain.

Acronyms and abbreviations

ACN	Acetonitrile
THF	Tetrahydrofuran
DCM	Dichloromethane
DMF	Dimethylformamide
DMSO	Dimethylsulphoxyde
ESI-MS	Electrospray Ionization - Mass Spectrometry
ESI-Orbitrap-MS	Electrospray Ionization - Orbitrap - Mass Spectrometry
EtOH	Ethanol
FTIR	Fourier Transform Infrared Spectrometry
MeOH	Methanol
MNP	Magnetic nanoparticle
TLC	Thin- Layer Chromatography
TGA	*Thermogravimetric Analysis*
XRD	*X-Ray Diffraction*
Py	*Pyridine*

3.1 Introduction

3.1.1 Histone acetylated lysine residues in epigenetics

As already deeply discussed in Chapter 2, post-translational modifications (PTMs) of histone proteins, such as acetylation, methylation, phosphorylation, and ubiquitylation, play essential roles in regulating chromatin dynamics. Many PTMs occur on histone proteins, regulating protein–protein interactions, localization, stability, and enzymatic activities of proteins involved in different cellular processes.[1,2] In Chapter 2, we presented the development of a tool for the enrichment of peptides bearing mono-methylated lysine residues. The tool was composed by ferromagnetic nanoparticles coated with tetraphosphonate cavitands, able to selectively complex the target guest. Here we present the design, synthesis, and characterization of molecular receptors for the binding of acetylated lysines. Acetylation is one of the most abundant PTM throughout the cell, and it plays an important role in gene transcription due to the weak electrostatic interactions between phosphate groups of DNA and histones[3] or adjacent nucleosomes.[4] More than 80% of human proteins show an acetyl group at the *α*-position of the first amino acid, that is transferred from acetyl-coenzyme A (Acetyl-CoA) to the primary amine located on the ε-lysine side chain. The acetylation process can occur spontaneously or *via* enzyme-mediated opposing mechanism of two class of enzymes named: histone acetyltransferases (HATs) and deacetylases (HDACs), Figure 1.

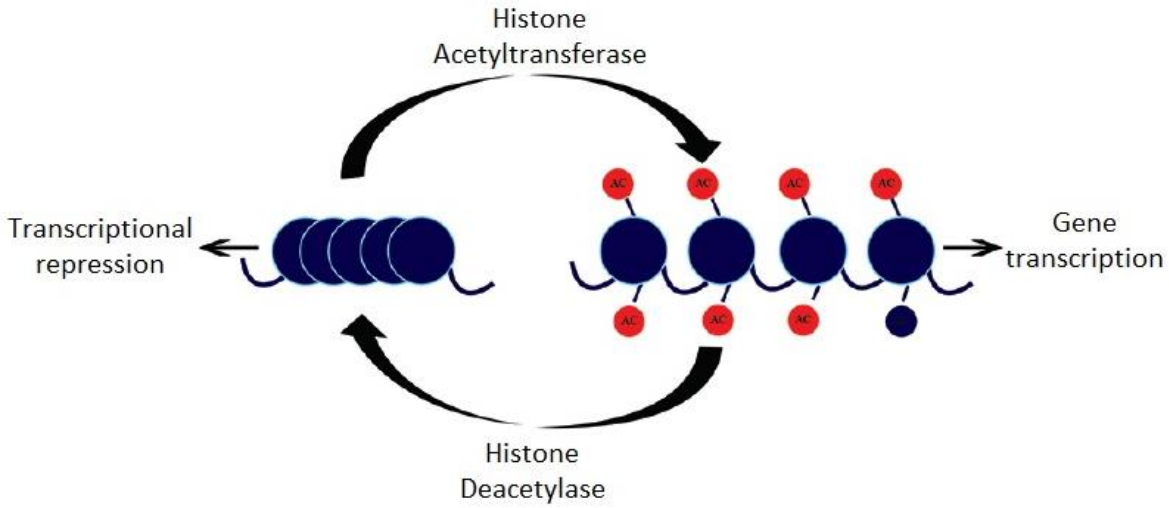


Figure 1. Representation of Acetylase/Deacetylase mechanism regulate by: HAT and HDACs.

HATs enzymes can be classified in three major categories: GNAT (GCN5-related N-acetyltransferases), p300/CBP (p300/CREB-binding protein), and MYST (MOZ, Ybf2, Sas2, and Tip60), as depicted in Figure 2A. They all exhibit different action mechanisms exploiting Acetyl-CoA as a cofactor. P300/CBP uses an organized and rapid mechanism named “hot and run” (Theorell-

Chance), where the protein substrate weakly interacts with the p300 surface, allowing the lysine residue to pass through the enzyme tunnel to receive the acetyl group, followed by rapid protein dissociation. Differently from the standard sequential ternary mechanism, no stable ternary complex formation is foreseen in this case.[5;6] With regard to GNAT and MYST, they exploit the same mechanism but based on a lysine-amine deprotonation using an active site glutamate.[7;8] This allows the acetyl-CoA carbonyl to be attached by nucleophiles forming a transitory tetrahedral intermediate, which leads to the formation of acetyl-lysine and coenzyme A.

Looking at the deacetylation processes, 18 human enzymes were identified for the catalysis of lysine deacetylation,[9;10] named HDAC inhibitors and silent information regulator (Sirtuin) enzymes (Figure 2B).[11] They involved the removal of acetyl groups from histones tails forming a compacted and transcriptionally repressed chromatin structure. These enzymes are grouped in four classes in relation on their sequence homology. Class I, II and IV, comprehending HDACs 1-11, are localized in the nucleus or in the cytoplasm and resulted in gene repression.[11] Class III includes Sirtuins proteins 1-7 and shows a NAD^+-dependent HDACs mechanism of action.[12] SIRT5, SIRT6 and SIRT7 exhibit a limited deacetylase property, SIRT4 has no activity, while SIRT1, SIRT2 and SIRT3,[13;14;15;16;17;18] which are present both in the nucleus and in the cytoplasm, display a strong deacetylase activity that exploits nicotinamide dinucleotide (NAD^+) as a cofactor for catalytic activity.

61 bromodomains (BDs) were found in 46 bromodomain-containing proteins (BCPs) of human proteome behaving as “readers” of lysine acetylated residues in specific sites on core histones or on the N-terminal tails of histones. As elucidated by X-ray crystal diffraction analysis (Figure 2C), the structure presents four left-handed α-helices linked by two loops forming a pocket for the recognition of the acetylated lysine. BDs have a wide variety of roles, mainly in chromatin regulation and function. For example, BRD4 act as transcriptional coactivators, while GCN5 bromodomain is involved in the re-modelling of chromatin[19] or state the sequential histone acetylation events.[20]

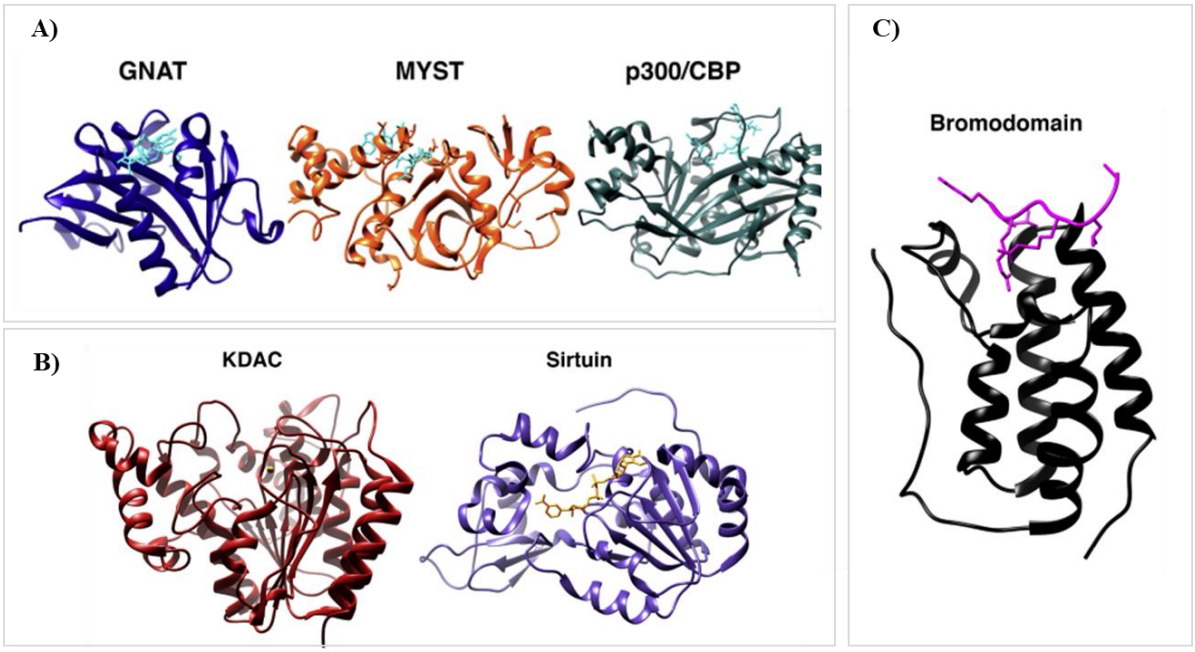


Figure 2. Structures of **A)** HAT domains (GNAT, MYST and p300/CBP); **B)** HDAC and Sirtuin inhibitors; **C)** acetylation reader domain (Bromodomain)

To date, acetylation has emerged as a crucial PTM and acetylome studies have reported the existence of thousands acetylation sites that can affect in different ways biological processes with diverse consequences such as aging and the development of several diseases including leukemia (Figure 3).[21,22]

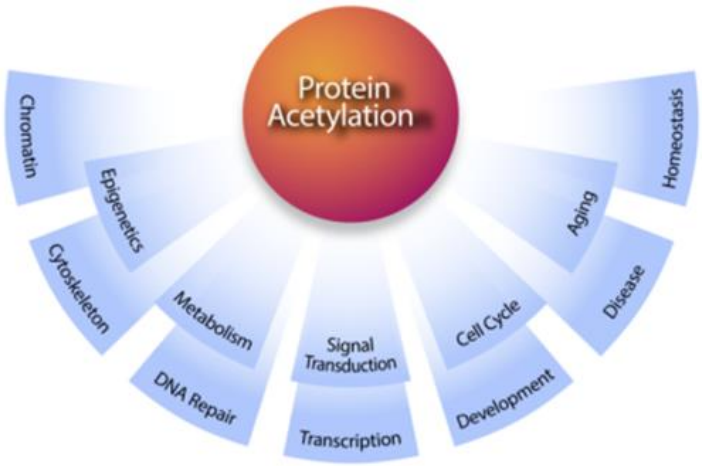


Figure 3. Library of protein acetylation regulation functions

Mass spectrometry-based proteomics is currently the gold standard for the large-scale identification and quantification of protein acetylation. Acetylome studies need rigorous optimization to reduce the background due to the large amounts of information obtained by the proteomic methods from a single sample.[23] To increase the quality of the mass spectra, an enrichment step is always necessary before performing mass spectrometry analysis. The use of commercially available pan-methyl antibodies for

immunoprecipitation (IP) experiments followed by mass spectrometry analyses have been proposed. Target modified proteins are usually immunoprecipitated from complex solutions, such as cell lysates, to be able to isolate and measure a specific protein. Immunoprecipitation (IP) is fast and relatively easy in comparison to affinity chromatography which is time-consuming and involves cycles of bind and washing. Unfortunately, IP is limited by the availability of antibodies that recognize the target protein. For this reason, alternative to the use of immunoprecipitation to enrich proteins' PTMs for mass spectrometry analysis is quite urgent.

3.1.2 Calix[4]pyrrole receptors and their ability in binding acetyl group

Calix[4]pyrrole are macrocycle compounds characterized by four modified pyrrole units held together by sp^3 hybridized carbon bridges (*meso*-carbons). They are usually colourless and highly conformationally flexible molecule. In fact, the relative orientation of the pyrrole units allows them to adopt a variety of flexible conformations in solution (Figure 4).

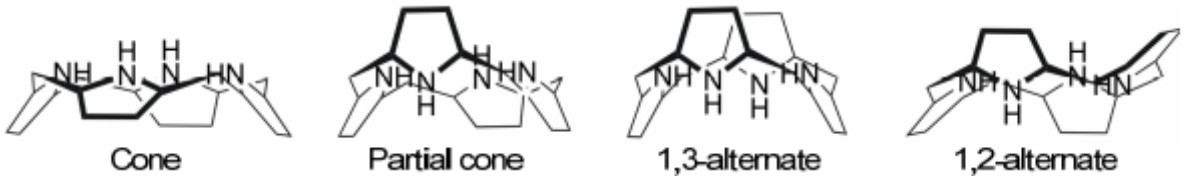


Figure 4. Representation of the possible conformations of calix[4]pyrrole in solution.

In non-polar solvents, calix[4]pyrrole assumes an alternate conformation designed as *1,2- or 1,3-alternate*, while in polar solvents (acetonitrile, methanol or acetone) they adopt a *cone* or *partial-cone* conformation.[24] Thanks to the presence of the four pyrrole moieties, the cavity core can act as a H-bond donating agent toward anions or other polar molecules (Figure 5). The addition of these guest induces a change in the conformation of the calix, from the less polar *1,3-alternate* conformation to the *cone* one, where the NH groups pointing inward the cavity, due to the instauration of simultaneous hydrogen bonding between the guest and the NH functionalities.

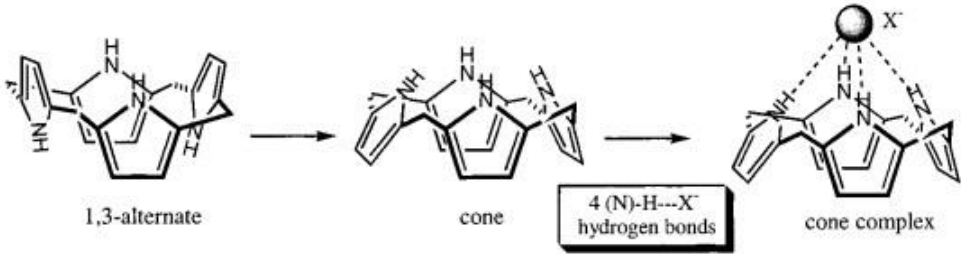


Figure 5. Representation of the change in conformation of the calix[4]pyrrole after the addition of an anion guest.[25]

Calix[4]pyrroles can be functionalized at *meso*- positions with different groups. If phenyl groups are present as substituents (Figure 6), aryl-extended calix[4]pyrrole derivatives are formed, presenting

four configurational isomers. According to the orientation of the aryl functionalities, the following nomenclature is used: *ααββ*, *αβαβ*, *αααβ*, and *αααα*, where α or β indicates the aryl substituent pointing up or down, respectively, considering to the four *meso*-carbon plane (Figure 6A).[26]

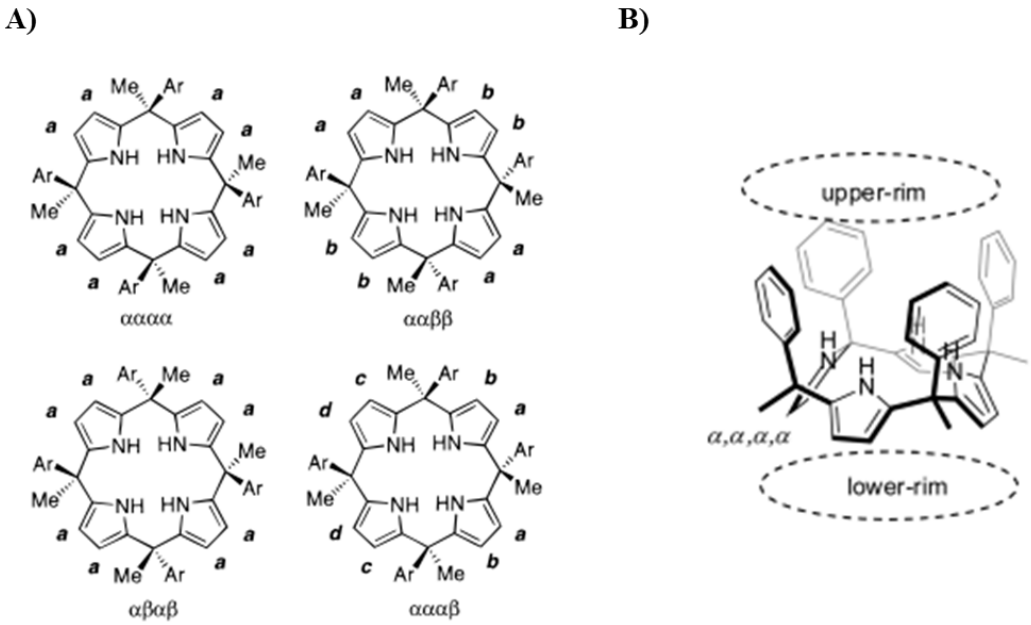


Figure 6. Representation of **A)** the four configurational isomers of calix[4]pyrrole; **B)** the *αααα-isomer* Aryl extended calix[4]pyrrole in cone conformation.[26;27]

The *cone*- conformation of the *αααα*-isomer (Figure 6B) results in a deep cavity showing unique properties towards the binding of anions and electron-rich molecules, thanks to the presence of synergistic hydrogen bonding, π-π, CH-π, and hydrophobic interactions. This molecular receptor can be exploited in numerous application such as the self-assembly of molecular dimeric capsules[28] and the development of functional materials.[29]

Aryl-extended calix[4]pyrroles are synthesised by an acid catalysed condensation reaction between a pyrrole and an aryl ketone (e.g. acetophenone), generating a mixture of the four different configurational isomers and open oligomers.

In general, the yield of this condensation reaction is low because of the possible formation of the several configurational isomers described above. Thus, to overcome this issue, these calix[4]pyrroles are usually prepared using methyltrialkylammonium chloride salts as templating agents (Figure 7). The role of this template is not completely understood. However, the calix[4]pyrrole cavity in its *cone* conformation can accommodate this salt thus favouring the cyclo-condensation reaction of linear oligomers and the stereoselectivity of the process toward the *αααα*-isomer.[27]

Figure 7. Acid catalysed condensation reaction between pyrrole and ketone to give Aryl-extended calix[4]pyrrole exploiting the use of methyltrialkylammonium chloride salts as template.[27]

In order to tune the recognition properties of the aryl-extended calix[4]pyrroles, both the upper and the bottom rim can be functionalized. The upper position can be exploited for the construction of more elaborate structures or capsules architectures,[30,31,32] while several functional groups can be attached at the lower rim to increase solubility in different media or allowing grafting on surfaces.

An interesting example of more elaborated water-soluble calix[4]pyrroles was reported by Ballester and co-workers. They described the preparation of a Super Aryl Extended (SAE)-calix[4]pyrrole, modifying the upper rim of the *αααα*-isomer of para-tetraiodo-*meso*-phenyl calix[4]pyrrole by installing ethynyl-aryl substituents (Figure 8).[31]

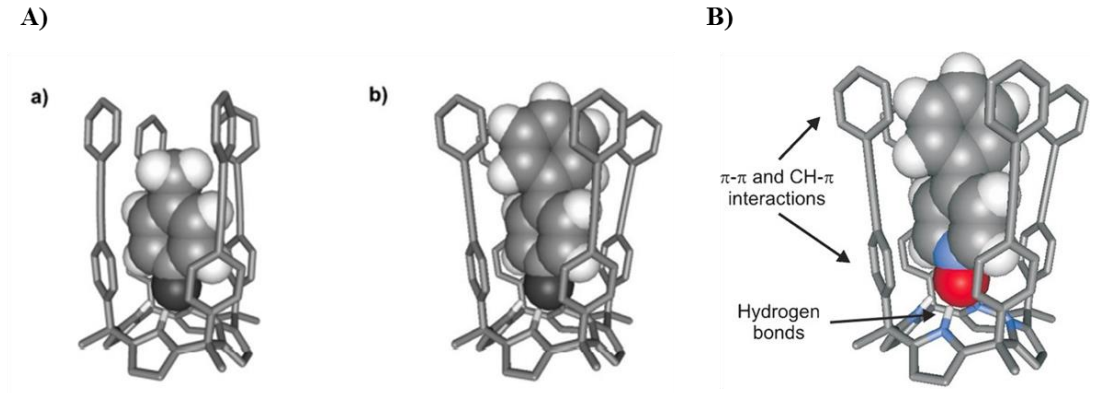


Figure 8. A) Molecular modelling of SAE-calix[4]pyrrole complexing 4-methyl pyridine-N-oxide (a) and 4-phenyl-pyridine-N-oxide (b); **B)** Representation of the interactions involved in SAE-calix[4]pyrrole binding 4-phenyl-pyridine-N-oxide.[31]

The authors illustrated the ability of these receptors in the inclusion of electron-rich neutral molecules, and ions, such as N-oxides, important antiviral drugs (Figure 8A).[33] The binding was possible since the pyridyl ring of the guest can establish π–π and CH–π interactions with the *meso*-phenyl substituents of the aromatic cavity, and the oxygen atom can form four hydrogen bonds with the NH functionalities stabilizing the cone conformation (Figure 8B). From ITC analysis, the binding process resulted to be enthalpically driven, with an association constant larger than 10^4 M^{-1}.

Secondary carboxamides are another interesting class of guests investigated in the Ballester research group using aryl-extended calix[4]pyrroles as hosts.[34,35] In biology, these guests determine the folding and the interactions involved in proteins and peptides. In nature, they can exist in both *cis* and *trans* isomers due to the restriction in rotation induced by the pseudo-double-bond character of the amido C-N bond (Figure 9).[35]

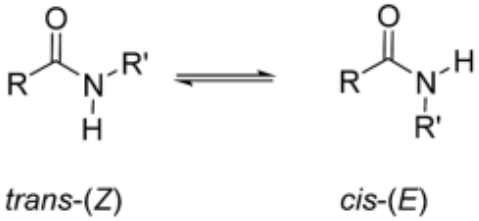


Figure 9. *Cis* and *trans* equilibrium of carboxyamide species.

The most significant factor that influences the ratio of the two relative isomers is the steric hindrance of the alkyl group(s) attached to the nitrogen. In N-alkylformamides for example, the *cis*-conformer percentage increases with the bulk of the N-substituent, due to steric interactions between the alkyl substituent and the carbonyl oxygen. While, the predominance of the *trans*-conformer is observed when the formyl hydrogen atom is replaced by an alkyl group due to the steric interaction between the chain and the alkyl group.[36]

In 2018, Ballester and *co-workers* demonstrated the ability of calix[4]pyrroles in selectively binding N-alkyl formamide and acetamide guests in their *cis*-conformation, in neutral water solution at mM concentration (Figure 10).[35] The binding is possible due to hydrophobic effect and the synergistic formation of interactions such as hydrogen-bonding, NH–π and CH–π, achieving a binding constant value around 10^3 M^{-1}.

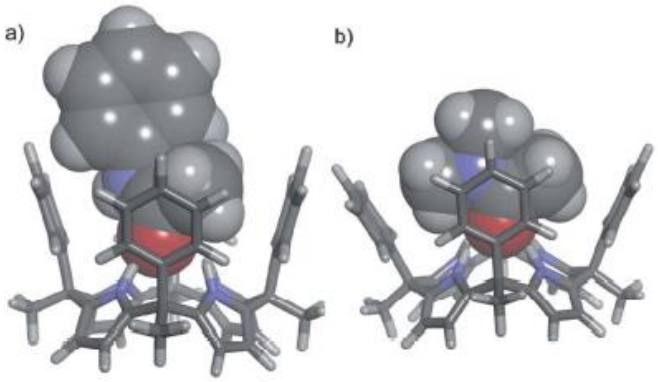


Figure 10. SCIGRESS modelling of aryl extended calix[4]pyrrole binding **a)** *cis*-N-phenyl-acetamide and **b)** *trans*- N-phenyl-acetamide.[35]

Moreover, the ability of a suitable functionalized calix[4]pyrrole to recognize a secondary amide such as creatine, a metabolic waste product related to muscle activity and associated with renal, muscular

and thyroid dysfunctions, and its derivatives was recently demonstrated.[37] To this aim, the Authors synthesized a calix[4]pyrrole phosphonate-cavitand functionalized at the upper rim with a dansyl chromophore as signalling unit , exploiting an indicator displacement assay of an inert guest dye or a black hole quencher (BHQ). The system performances were tested in organic solvent through ^{1}H and ^{31}P NMR, ITC, UV/vis absorption and fluorescence, reporting a binding constant value around (2−4) $\times 10^{5}$ M^{-1} and a turn on of the fluorescence, after the displacement of the fluorescent guest with the target, at micromolar concentrations, compatible with both healthy and sick patients (Figure 11).

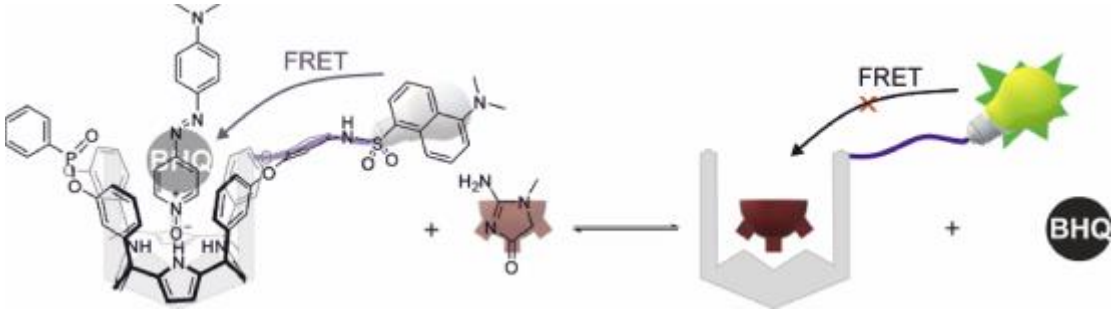


Figure 11. Representation of creatinine sensing in phosphonate-calix[4]pyrrole inducing a competitive indicator displacement assays.[37]

3.1.3 Aim of the project

In light of these results, the aim of this project focused on exploiting water soluble calix[4]pyrroles as molecular receptors for acetylated lysines. This work was conducted at the ICIQ, in the research laboratories of Professor Pablo Ballester. The project was developed in distinct steps (Figure 12). Initially, the abilities of different types of water-soluble calix[4]pyrrole in recognizing acetylated lysines were compared. The best performing receptor was suitably functionalized for the subsequent grafting on the surface of FeNPs. The grafting is necessary to exploit the multivalency effect in the recognition of the Kac residues present in animal histones, and subsequent enrichment of this modification.

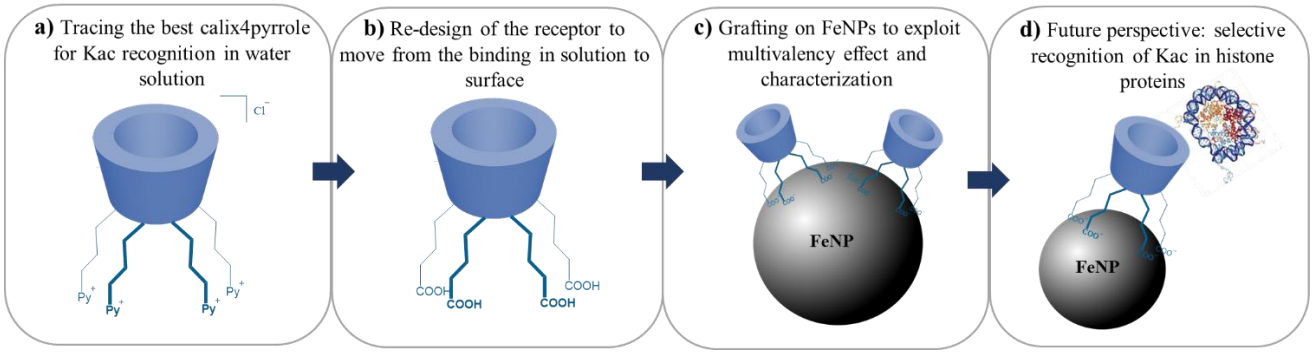


Figure 12. Representation of the aim in four steps. **a)** Find out the best calix4pyrrole; **b)** Re-design of the receptor in order to move from binding studies in solution to surface; **c)** Grafting on the FeNPs and characterization; **d)** Future perspective. Selective recognition of Kac in histone proteins.

3.2 Results and discussion

In Figure 13**A)**, the designed calix[4]pyrrole receptors for the binding of Kac in water are illustrated. All the receptors are functionalized at the lower rim with 4 pyridinium groups to allow water solubility. In details, receptor **H1** is a tetra-pyridinium aryl-extended calix[4]pyrrole, while **H2** is an octa-pyridinium super aryl-extended calix[4] pyrrole presenting a deeper aromatic cavity equipped with 4 para-ethynyl-aryl substituents, further functionalized with 4 pyridinium groups in *para* position. Finally, **H3** presents a cavity depth intermediate between **H1** and **H2**. **H2** and **H3** were designed to enhance the hydrophobic effect in the binding. **H1** and **H2** receptors were already available as pure in the research laboratory, thus they were directly studied in the binding with Kac target. As concern **H3**, its synthetic preparation was attempted.

In Figure 1**B)** the target guest Kac is depicted as both *cis* and *trans* isomers. As already stated, secondary amides such as Kac can exist in the two conformations. Typically, in secondary amide *cis* isomer reaches 10-20 %, a percentage value that increase with the steric hindrance of the N-substituent.[34] In the case of acetylated lysine, the *cis*/*trans* ratio is 1:99. As already discussed, calix[4]pyrroles preferentially bind the *cis* isomer of secondary amides.[35]

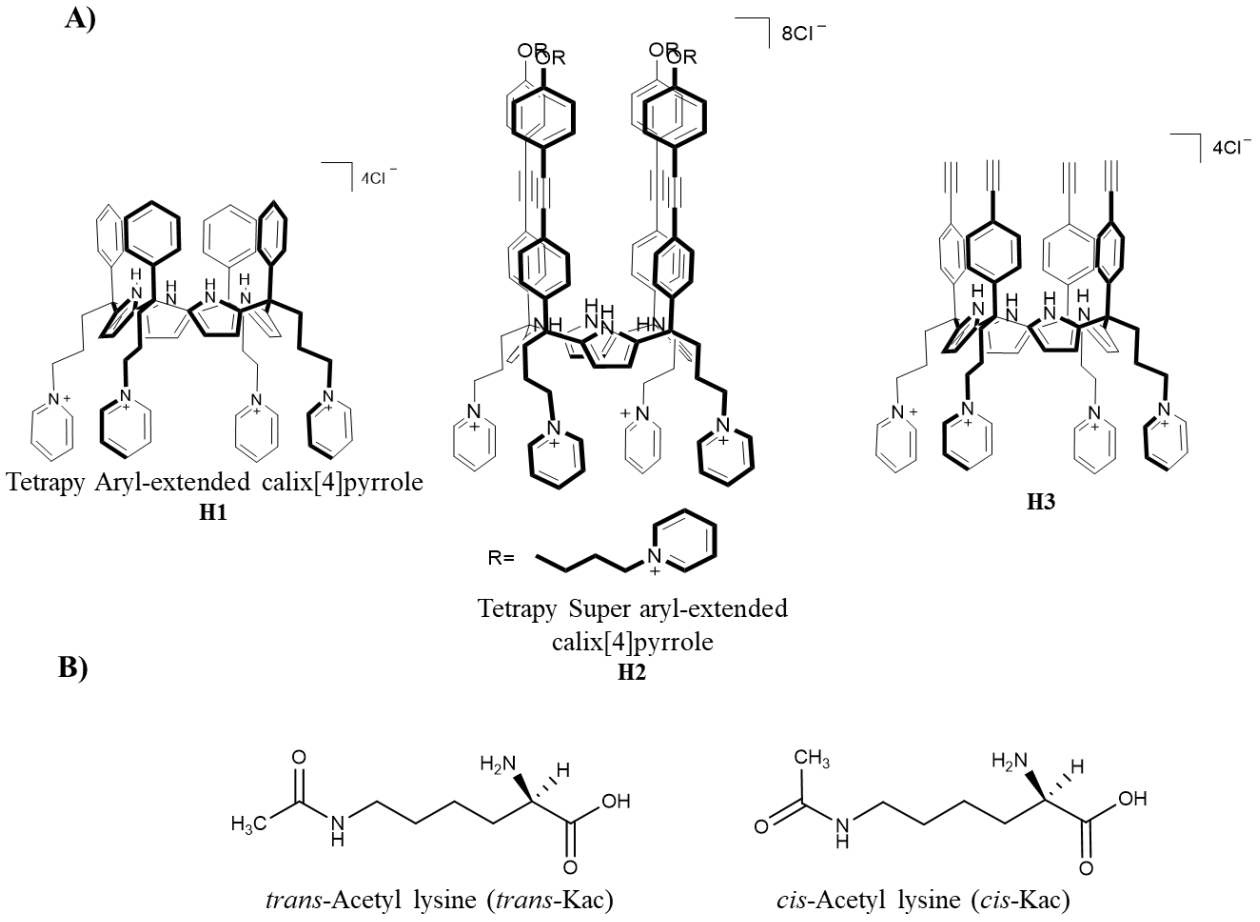


Figure 13. Representation of **A)** calix[4]pyrroles involved in the study; **B)** *cis-trans* conformations of the acetyl lysine guest.

3.2.1 ^{1}H NMR titration study involving H1 receptor

The recognition abilities of **H1** receptors toward Kac were firstly tested through ^{1}H NMR analysis, titrating incremental amounts of a deuterated water solution of Kac [mM] into a D_2O solution of **H1** receptor [mM].

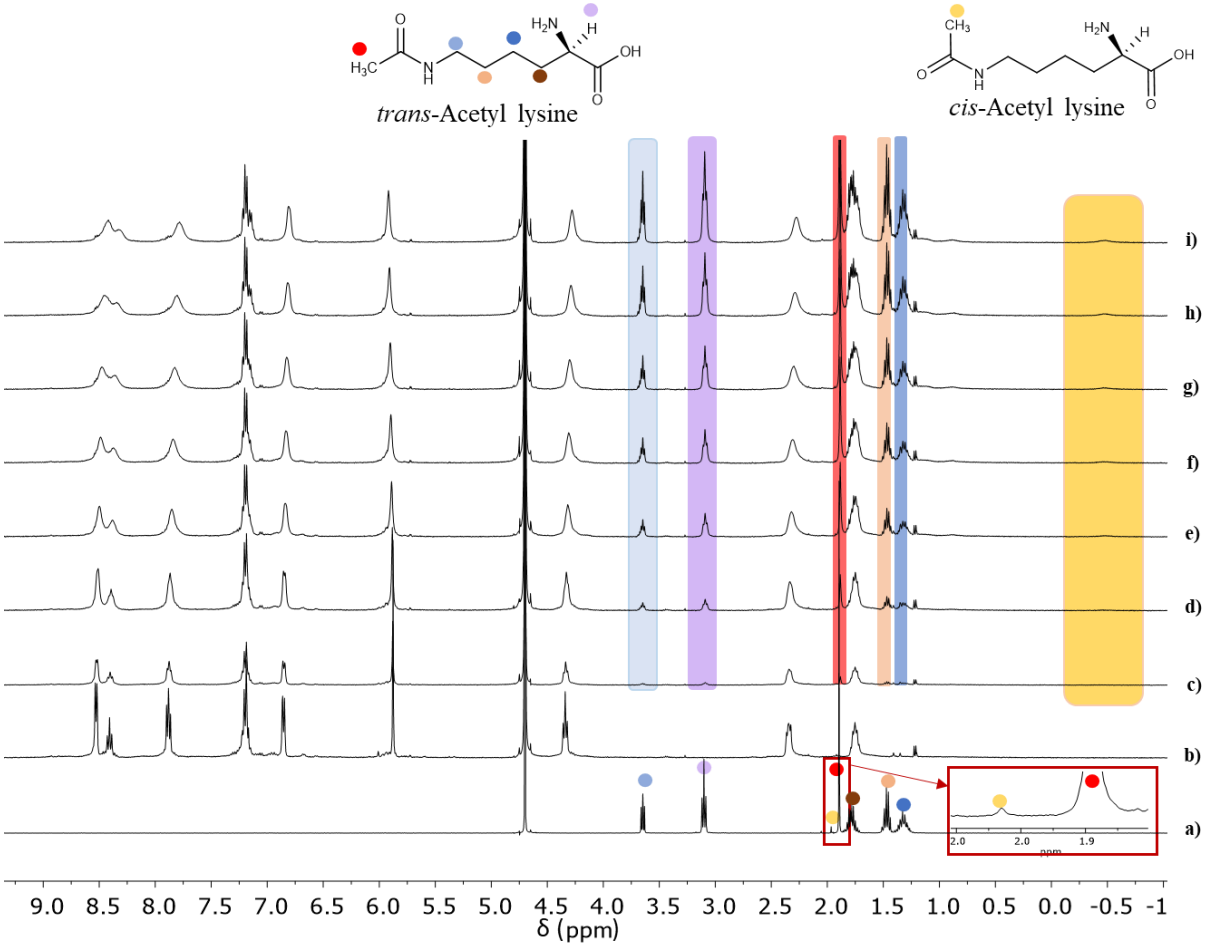


Figure 14. ^{1}H NMR (400 MHz, D_2O, 298K) spectra acquired during the titration of incremental amounts of Kac into a solution of **H1**. **a)** Kac, **b)** H1, 3 mM, **c)** H1 + 0.25 eq of Kac, **d)** H1 + 0.5 eq of Kac, **e)** H1 + 1.0 eq of Kac, **f)** H1 + 1.5 eq of Kac, **g)** H1 + 2 eq Kac, **h)** H1 + 3 eq Kac, **i)** H1 + 5 eq Kac.

Since Kac can assume two conformations (*cis* and *trans*) depending on the position of the CH_3 group respect to the alkyl chain, in the ^{1}H NMR spectrum of the molecule (Figure 14a), two peaks related to this methyl group (yellow and red marks, in the red square) are visible. They present different intensities: the smallest one is related to the *cis* conformer (yellow dot), while the peak with the highest intensity is associated to the *trans* conformer (red dot). The *cis*/*trans* ratio calculated from the integral values of these two signals is ~1:99. As expected, during the titration the methyl group of the *trans*-conformer did not experience any chemical shift change (red rectangle) since it is not recognized by the receptor, and the signals related to its chain remain unaffected as well (pink dot and rectangle). For the *cis* isomer, a broad signal related to the methyl group appeared at -0.48 ppm, and the peak intensity increased progressively with the incremental amount of added guest (yellow

rectangle, Figure 14 c-i). Unfortunately, because of the signal broadening it was not possible a precise association through NOESY NMR experiment (Appendix, Figure 26). Thus, the assignation of this signal as *cis* isomer peak was made based on previous work performed at the Ballester group.[35] The diagnostic signals related to the receptor **H1** produced very small changes in the chemical shift, but rather a broadening of the signals probably related to slow exchange in the NMR timescale. The binding constant value, K_a, was directly calculated from ^{1}H NMR spectra using a theoretical binding model that considers the existence of an equilibrium between the two isomers in solution, and the exclusive formation of a 1:1 complex between **H1** and the *cis*-conformer. The calculated K_a resulted in the 10^3 - 10^4 M^{-1} range.

3.2.2 ^{1}H NMR titration study involving H2 receptor

The same titration was performed using the deeper receptor **H2** (Figure 15).

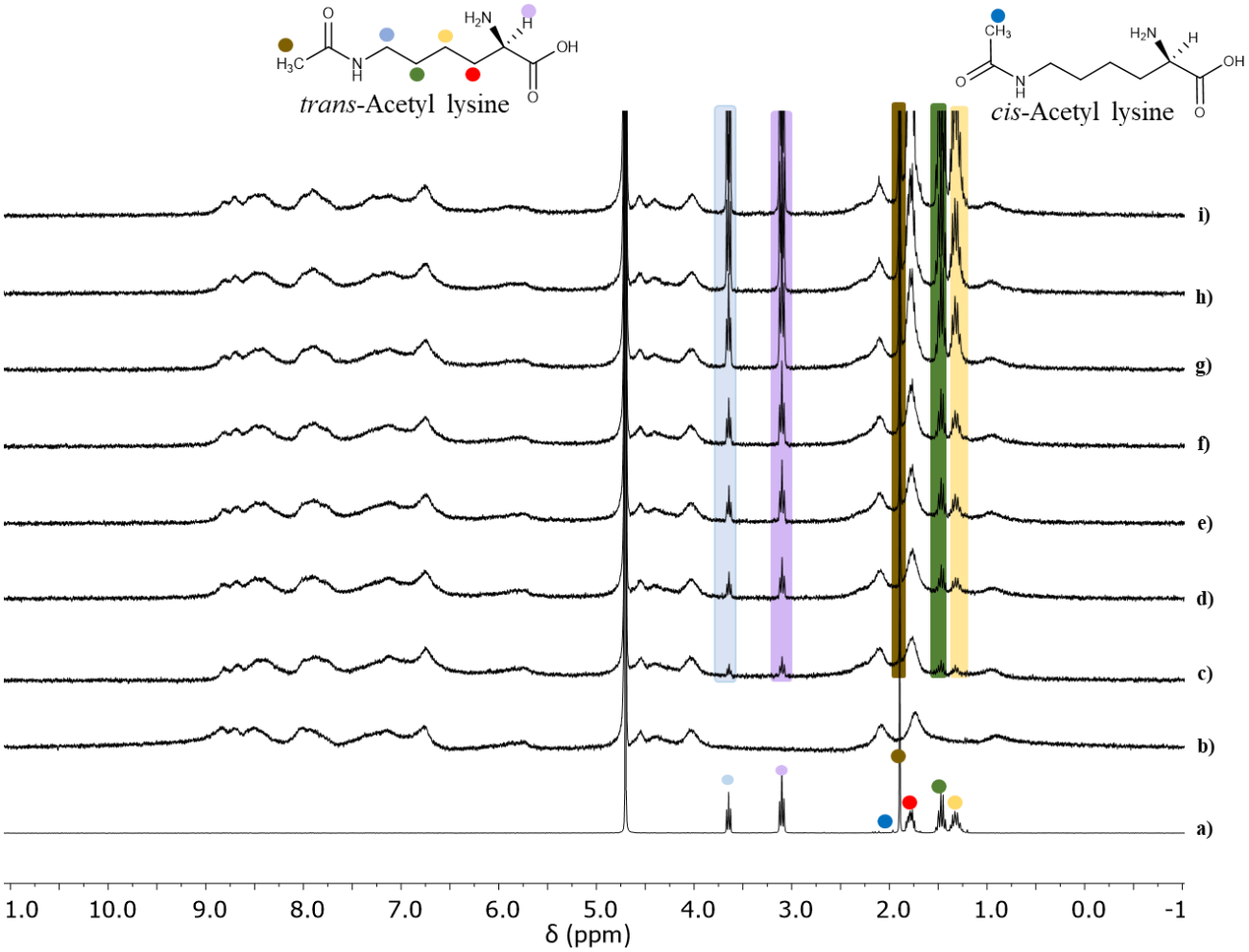
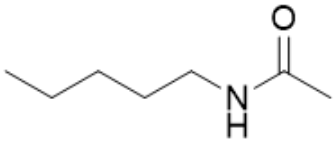


Figure 15. ^{1}H NMR (400 MHz, D_2O, 298K) spectra acquired during the titration of incremental amounts of a deuterated water solution of Kac [mM] into a D_2O solution of **H2** with. **a)** Kac, **b)** H2, 1.52 mM **c)** H2 + 0.25 eq of Kac, **d)** H2 + 0.5 eq of Kac, **e)** H2 + 0.75 eq of Kac, **f)** H2 + 1 eq of Kac, **g)** H2 + 2 eq Kac, **h)** H2 + 5 eq Kac, **i)** H2 + 12 eq Kac.

As depicted in Figure 15, the NMR spectrum of **H2** (Figure **15b**), appears to be very broad because of the presence of eight positive charges, four at the upper rim and four at the lower rim of the receptor, which causes aggregation in the water medium. The addition of the guest should break the aggregation resulting in a sharpening of the signals. As shown in Figure 15 c-d, the receptor signals remained very broad, suggesting that the guest is not complexed by the host. This is also supported by the absence of any shift of the guest peaks.

We speculated that the preference of Kac for **H1** and not **H2** could be related to solvation effect and geometry reasons. In general, to be complexed with an appreciable K_a value, a guest has to be complementary with the host cavity in terms of size, shape and chemical environment. Considering acetylated lysine, to be effectively complexed the acetyl C=O functionality has to point inward the host cavity to interact with the NH groups, while the aminoacidic part should point outward the cavity to remain solvated in water, thus not paying a negative contribution in terms of enthalpy loss for desolvation.

To confirm this thesis, we tested the complexation ability of **H2** towards pentyl acetamide (PA). We selected PA as control experiment since it is a secondary acetamide like Kac and presents the same alkyl chain length, without the terminal aminoacidic group (Figure 16). The absence of this group will exclude the negative enthalpy contribution due to desolvation during the complexation. Being a secondary amide, PA can exist in both *cis* and *trans* isomer, with the *trans* one as major population. The ^{1}H NMR spectrum of PA (Figure 17a) revealed that the *cis/trans* ratio is 20/80, calculated from the integral values of the two methyl peaks (grey and yellow dots for the *cis* and the *trans* isomer, respectively).



Pentylacetamide (PA)

Figure 16. Pentyl acetamide guest structure

The control experiment was performed through a ^{1}H NMR titration (Figure 17) in water medium, adding PA [mM] to **H2** [mM].

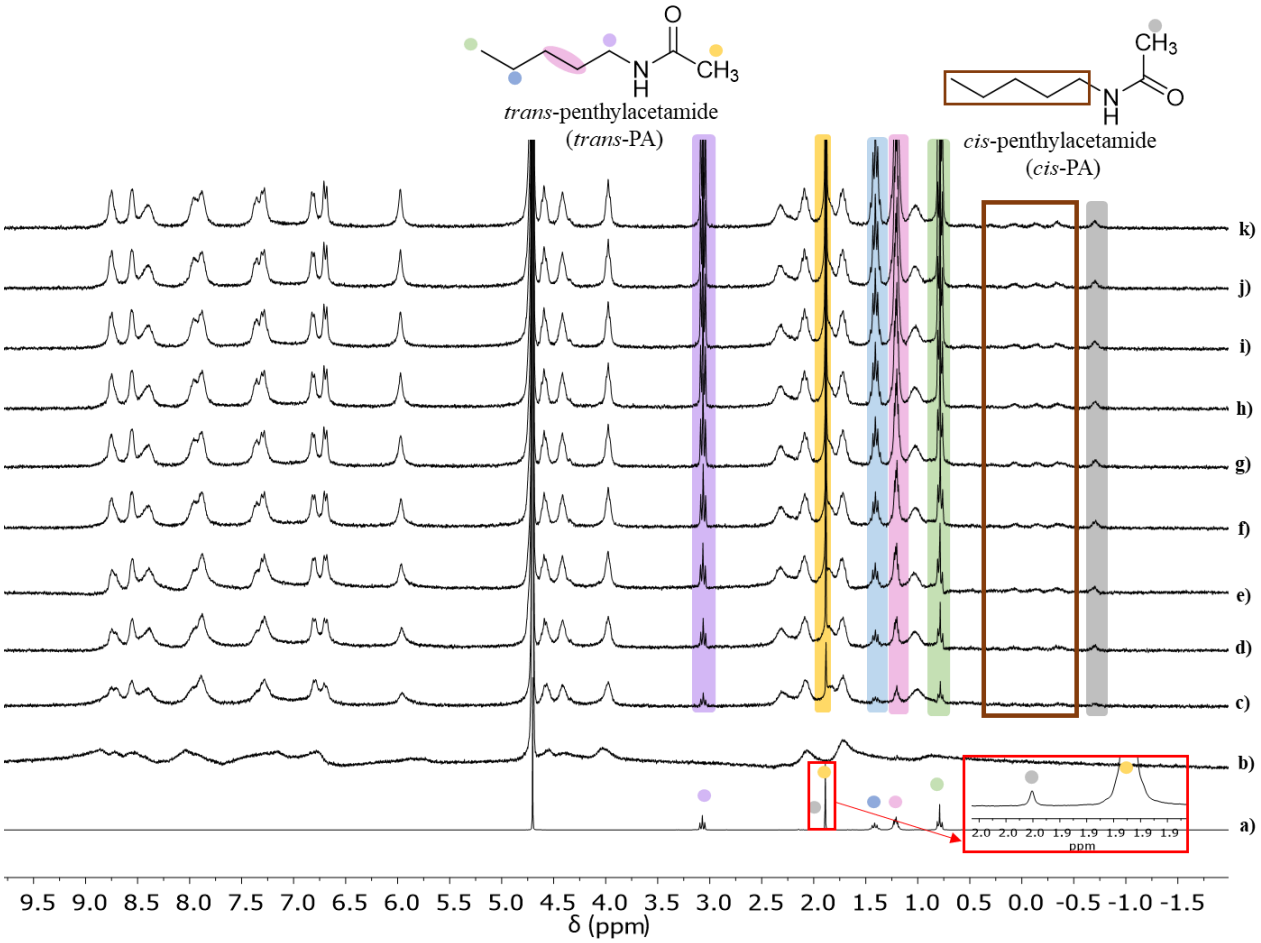


Figure 17. ^{1}H NMR (400 MHz, D_2O, 298K) spectra acquired during the titration of incremental amounts of PA into **H2** with. **a)** PA, **b)** H2, 1.5 mM **c)** H2 + 0.25 eq of PA, **d)** H2 + 0.5 eq of PA, **e)** H2 + 0.75 eq of PA, **f)** H2 + 1 eq of PA, **g)** H2 + 0.5 eq of PA, **h)** H2 + 2 eq of PA, **i)** H2 + 2.5 eq of PA, **j)** H2 + 3 eq of PA, **k)** H2 + 4 eq of PA.

As evident from Figure 17, **H2** is able to bind PA in its *cis* conformation.[35] With the progressive addition of the guest to the host solution (Figure 17 c-k), a new set of signals related to the alkyl chain and the acetyl protons of the *cis* conformer (brown and grey rectangles, respectively) appeared in the upfield region, between -0.6 and 0.5 ppm. The complexation event is in slow exchange at the NMR timescale, and the broadening of the peaks hampered the NOESY NMR analysis to elucidate peak association (Appendix, Figure 27). The protons of the *trans* isomer did not experience any shift (blue, pink and green rectangle). As for the host, a sharpening of the peaks can be observed during the titration, due to the break of the aggregates after complexation. The binding constant value, K_a, was directly calculated from ^{1}H NMR spectra using the theoretical binding model that considers the existence of the equilibrium between the two isomers in solution and the exclusive formation of a 1:1 complex of the receptor with the *cis*-conformer and resulted in the range of 10^3-10^4 M^{-1}. Taken together these results demonstrate that not only shape complementarity between the guest alkyl chain and the depth of the cavity receptor is necessary for complexation, but also desolvation plays a crucial role in the binding event. The opposite enthalpic and entropic contributions have at least to

compensate; if the desolvation enthalpic cost is higher respect to the entropic gain, the binding is disfavoured.

In light of these results, the preparation of calix[4]pyrrole **H3** (Figure 18A) was attempted. **H3** presents an aromatic cavity functionalized on the top with four alkynes and with a depth intermediate between of **H1** and **H2** cavities. As confirmed by the modelling *via* SCIGRESS (Figure 18B), the cavity of **H3** can accommodate Kac, with the acetyl C=O pointing towards the bottom of the cavity, involved in H bonds with the four pyrrole NHs, while the aminoacidic moiety is pointing outside the cavity.

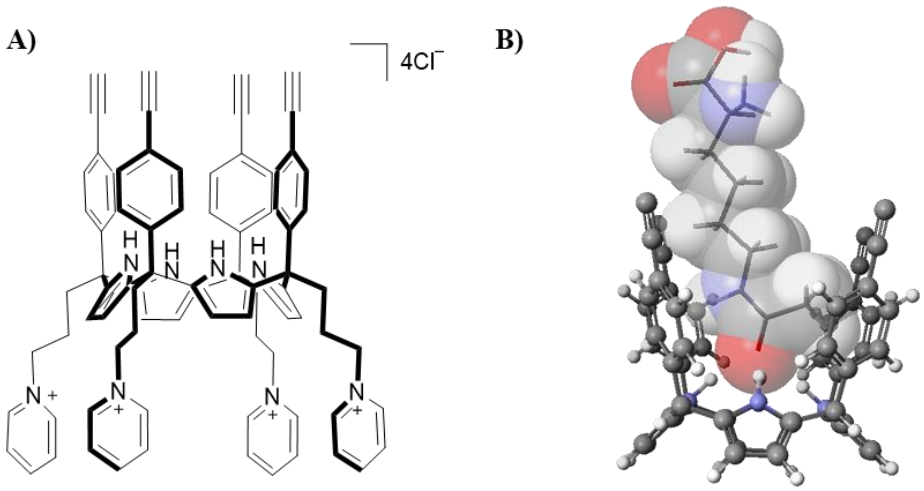
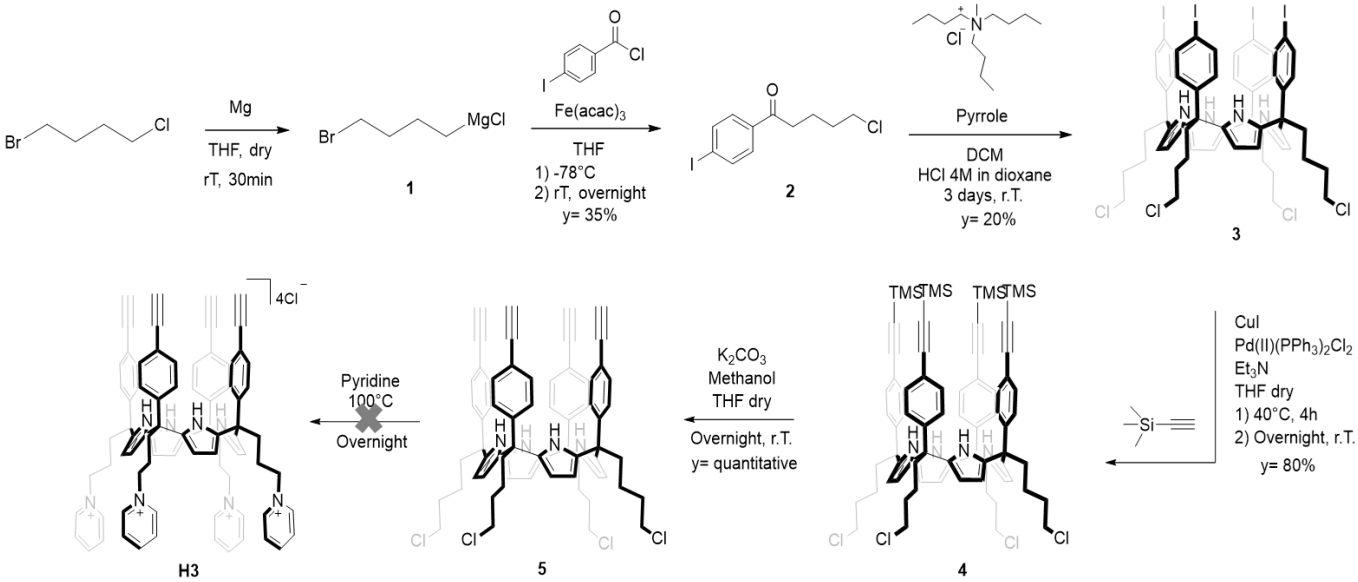


Figure 18. Representation of **A) H3** molecular structure and, **B)** modelling of **H3** receptor binding acetyl lysine guest (from SCIGRESS software)

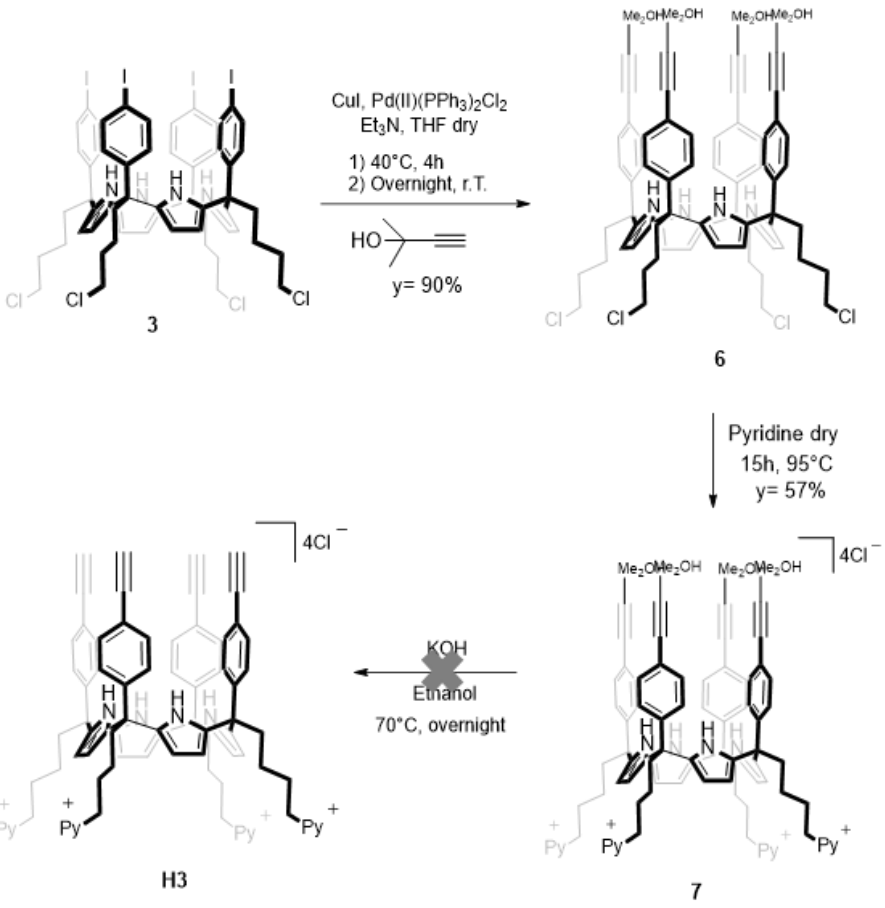
The first synthetic attempt for the preparation of **H3** is illustrated in Scheme 5.



Scheme 5. First synthetic pathway for the preparation of **H3**

The target compound **H3** was prepared in six steps starting from the commercially available 1-bromo-4-chlorobutane. The first two steps were carried out consequently and involved the *in situ* preparation of the Grignard compound **1**, followed by the addiction of 4-iodobenzoyl chloride to obtain ketone **2**. The synthesis was performed in dry conditions and with distilled THF as solvent. Intermediate **2** was obtained pure in 35% yield after purification by flash chromatography. Then, an acid mediated (HCl, 4M in dioxane) cyclo-condensation reaction between **2** and pyrrole in dry DCM as solvent was carried out, allowing the formation of tetraiodo calix[4]pyrrole **3** as αααα-isomer, in which the receptor adopts a cone conformation, with all the NHs groups pointing inward the cavity. The reaction resulted very slow; in fact after 3 days at room temperature, the crude reaction mixture consisted of different configurational isomers including the desired αααα-isomer.[32] The addition of methyl trialkylammonium chloride salt as template was necessary to increase the stereoselectivity of the reaction. Compound **3** was obtained in 20% yield after purification by firstly flash chromatography and then crystallization, the latter performed dissolving slowly **3** in the minimum amount of DCM and adding some drops of ACN.The solution was then cooled in the fridge for 24 hours to facilitate the precipitation of the desired pure compound.

To functionalize the upper rim of the receptor with the alkyne functionalities, a Sonogashira coupling reaction was performed in anhydrous THF under standard conditions, using trimethylsilylacetylene as protecting-group, CuI and $Pd(PPh_3)_2Cl_2$ as catalysts, and Et_3N as a base. Intermediate **4** was obtained in 80% yield after purification by flash chromatography. Deprotection of **4** was performed with K_2CO_3 in methanol and THF as solvents, allowing the formation of compound **5** in almost quantitative yield. Finally, a nucleophilic substitution of the chloride groups at the bottom rim with pyridine was attempted to obtain the desired compound **H3**, but unfortunately it was not possible to get the desired product **H3**. After overnight reaction, a brown-coloured precipitate was recovered and filtered off, which resulted insoluble in common organic solvents, DMSO and water. Even after several attempts in different conditions, it was not possible to obtain **H3**. We supposed that the alkyne groups at the upper rim are instable under the high temperature of the final reaction. Thus, it was decided to test a different approach, as illustrated in Scheme 6.



Scheme 6. Second synthetic approach for the preparation of **H3**

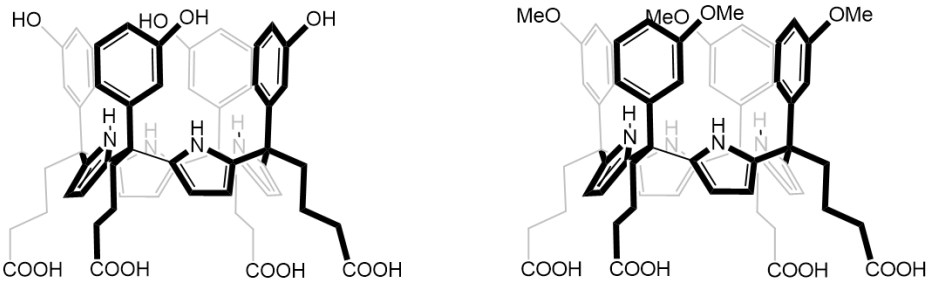
Starting from intermediate **3**, the Sonogashira coupling reaction was performed in the same conditions of compound **4** but, differently from the previous attempt, in this case 2-methyl-3-butyn-2-ol was selected as alkyne. This reagent is characterized by the presence of a more hydrophilic protecting group in order to reach a higher yield. The desired compound **6** was obtained pure in 90 % yield. Before deprotection of the alkynes groups, calix[4]pyrrole **6** was functionalized at the lower rim with four pyridinium moieties through a nucleophilic substitution in distilled pyridine obtaining compound **7** in 57% yield. The last step foresaw the deprotection of the alkynes in basic conditions using KOH and ethanol as solvent. Unfortunately, also in this case it was not possible to recover the desired product, but an insoluble precipitate was recovered, as in the previous case. We concluded that the calix presenting the alkyne groups are too instable during the reaction conditions and polymerization reactions between the alkyne functionalities could take place.

In light of these results, it was decided to focus only on **H1** as receptor, which showed the best affinity for Kac.

3.2.3 Moving from binding studies in water solution to surface

In order to exploit the multivalency effect in the recognition of Kac, **H1** has to be grafted on the surface of FeNPs. For this purpose, a re-design of **H1** receptor was necessary to insert at the bottom rim of the calix[4]pyrrole four carboxylic groups suitable for FeNPs grafting, as already discussed in Chapter 2 for resorcinarene-based cavitands.

To this aim, two receptors were considered, namely **H4** and **H5** (Figure 19). Both the calix[4]pyrroles present four carboxylic functionalities at the bottom rim, but they differ for the *meta*- substituents on the aromatic cavity: **H4** presents four hydroxy groups, while **H5** displays four methoxy groups.



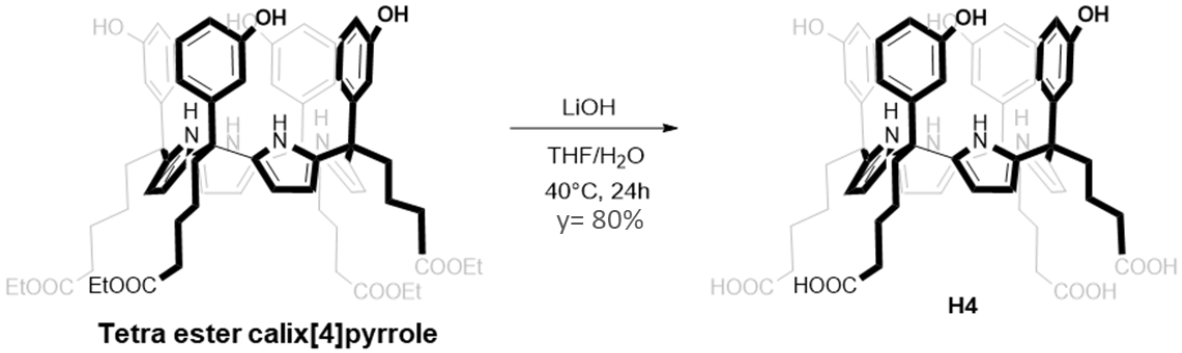
H4 **H5**

Figure 19. H4 and **H5** structures

These two calix[4]pyrroles were selected since the intermediate tetra-ester presenting four OHs as substituents on the cavity phenyls (Scheme 2) was already available in the Ballester's group. We speculated that the presence of the *meta*- substituents does not influence the final depth of the cavity receptor but can affect the complexation properties. For this reason, we decided to introduce also methoxy groups as substituents (**H5**) to investigate the effect of these groups on the binding properties of the cavity.

3.2.4 Preparation of H4

Calix[4]pyrrole **H4** was prepared performing an hydrolysis reaction starting from intermediate tetra-ester calix[4]pyrrole, already prepared as pure in the research group (Scheme 2).



Tetra ester calix[4]pyrrole **H4**

Scheme 2. Synthetic pathway for the preparation of **H4**

The hydrolysis of the ethyl ester groups was performed in presence of LiOH as base and a mixture of THF/Water 1:1 as solvent, to solubilized both the calix[4]pyrrole and LiOH. The reaction mixture was stirred at 40°C for 24 hours. The precipitate formed was filter off and washed several times with water to obtain compound **H4** in 80% yield without any further purification. The product was characterized by ^{1}H NMR (Figure 20), ESI-MS Orbitrap, FT-IR and TGA.

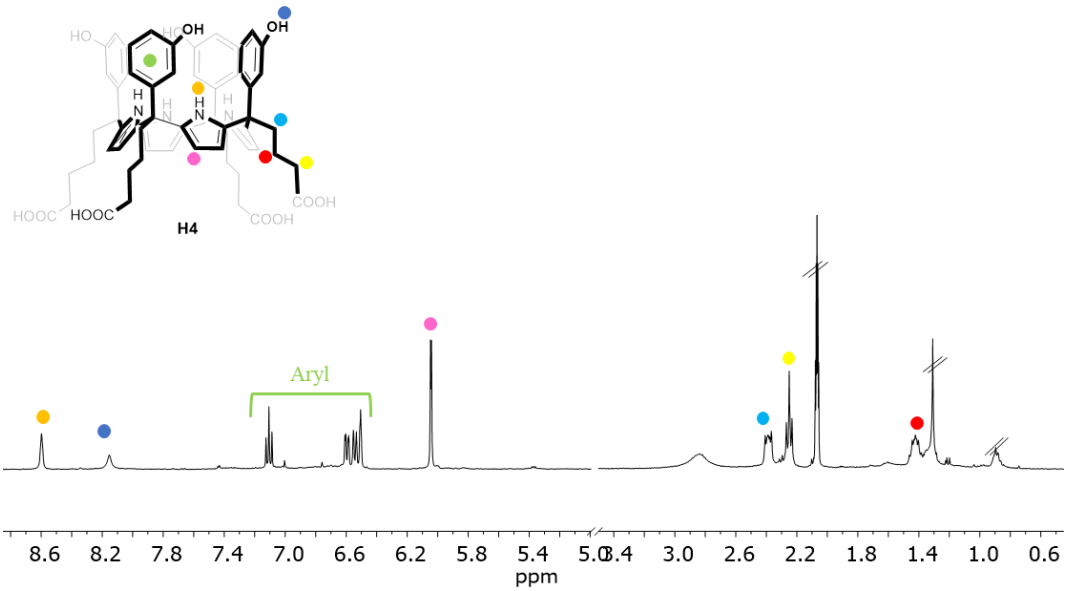


Figure 20. ^{1}H NMR spectrum of **H4**, in Acetone-d6, 400MHz, 25°C

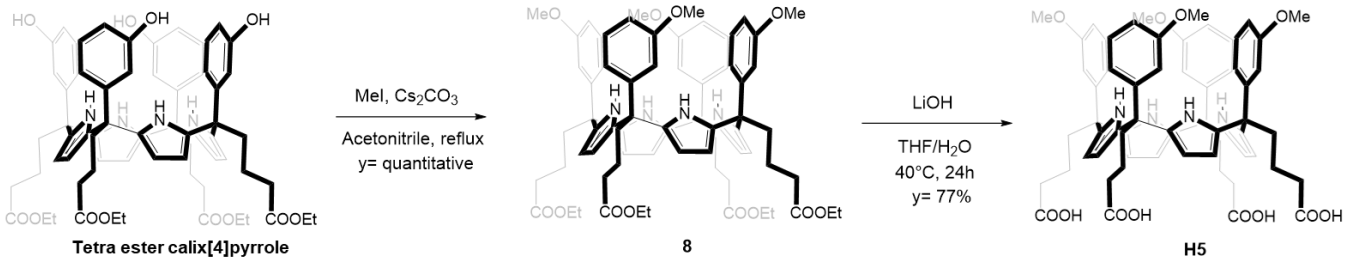
From the ^{1}H NMR spectrum reported in Figure 20, it is possible to observe the signals related to the NH and OH groups at 8.6 and 8.15 ppm (orange and blue dots), respectively. In the spectral region between 6.5 and 7.3 ppm the aryl signals are visible, and at 6.04 ppm the diagnostic signal assigned to β-pyrrole protons is present (pink dot). Finally, in the aliphatic region, it is possible to find the methylene protons of the alkyl chains at 2.39, 2.26 and 1.42 ppm (pale blue, yellow and red dots, respectively).

FT-IR displays at 1703.35 cm^{-1} the band related to the C=O stretching. At 2921.57 and 2852 cm^{-1} appear the stretching bands related to the aliphatic C-H of the chains and at 3395.59 cm^{-1} the N-Hs stretching. At 1204.31 cm^{-1} the C-N stretching of the pyrrole units are visible and, finally, at 1054.77 cm^{-1} the C-O stertching (Appendix, Figure 28).

The TGA thermogram, carried out in oxygen with a temperature ramp between 25 and 900 °C and flow rate of 20 °C/min showed that the receptor is stable up to 200 °C, and the degradation was indicated by three deflections respectively at 265, 430 and 630 °C, corresponding to weight losses of 30 and 45 %, respectively (Appendix, Figure 31).

3.2.5 Preparation of H5

H5 presenting the methoxy groups in *meta*- position of the aryls in the upper rim of the receptor was prepared following the synthetic pathway reported in Scheme 3.



Scheme 3. Synthetic pathway for **H5** preparation

As for **H4**, the synthesis of **H5** was started from the available intermediate tetra-ester calix[4]pyrrole. The introduction of four methoxy groups at the upper rim of the receptor was performed using MeI as methylating agent, Cs_2CO_3 as a base, and acetonitrile as solvent, under reflux for 1 hour. The reaction was quenched by liquid-liquid extraction in diethyl ether/water, allowing to obtain **8** in quantitative yield without any further purification. The last step involved the hydrolysis of the ester functionalities at the bottom rim using LiOH, in THF/water (1:1) as solvent mixture, at 40°C for 24 hours. The reaction was quenched by evaporating the solvent under reduced pressure; the obtained crude was dissolved in water, filtered off and washed several times with water allowing to isolate the product in 77% yield without any further purification. **H5** was characterized by ^{1}H NMR, ESI-MS Orbitrap, FT-IR and TGA.

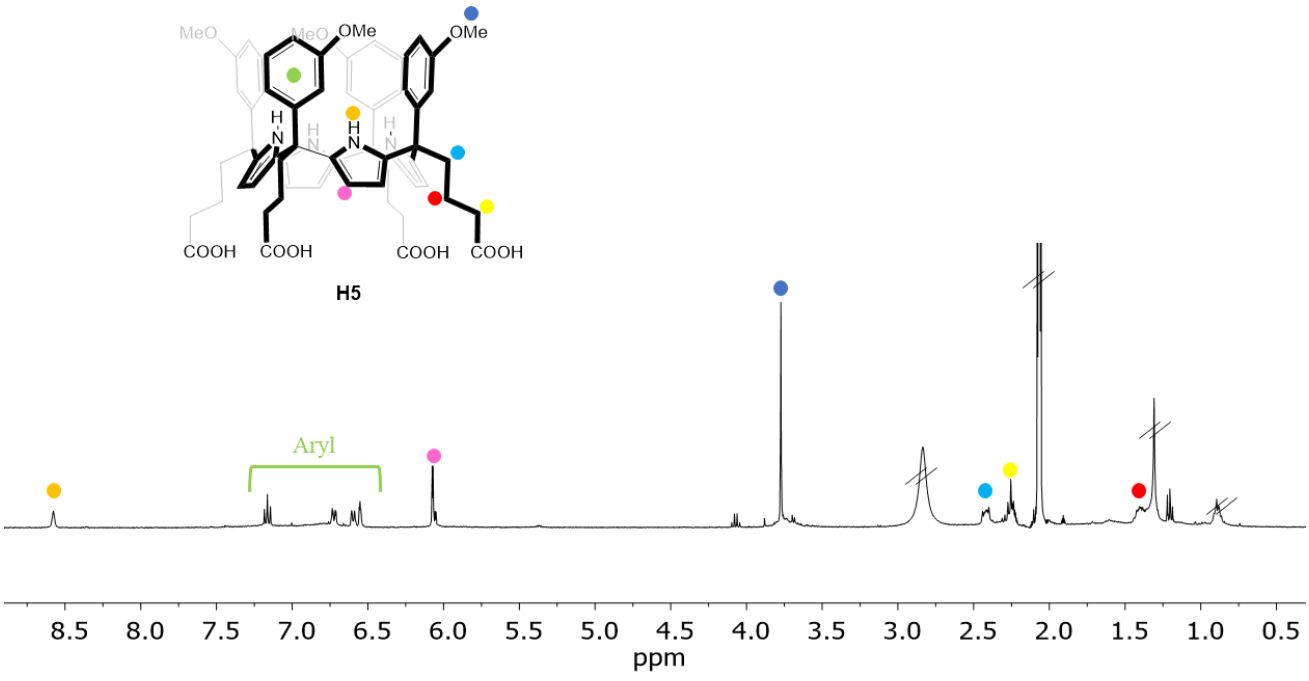


Figure 21. ^{1}H NMR spectrum of **H5**, in Acetone-d_6, 400MHz, 25°C

From the ^{1}H NMR spectrum reported in Figure 21, it is possible to observe the signal related to the NH groups at 8.58 ppm (orange dot). In the region between 7.2 and 6.5 ppm the aryl signals are visible, and at 6.07 and 3.78 ppm the diagnostic peaks assigned to β-pyrrole protons (pink dot) and methoxy group (blue dot), respectively. Finally, in the aliphatic region, it is possible to find the methylene protons of the alkyl chains at 2.42, 2.26 and 1.40 ppm (pale blue, yellow and red dots), respectively.

FT-IR display at 1703.29 cm^{-1} the band related to the C=O stretching and at 1429.14 cm^{-1} the band related to the C-O-C bonds. At 2921 and 2851.56 cm^{-1} appear the stretching bands related to the aliphatic C-H of the chains and at 3490 cm^{-1} the N-Hs stretching. At 1245.29 cm^{-1} the band of the C-N stretching of the pyrrole units is visible and, finally, at 1043 cm^{-1} the C-O stertching band (Appendix, Figure 29).

The TGA thermogram, carried out in oxygen with a temperature ramp between 25 and 900 °C and flow rate of 20 °C/min showed that the receptor is stable up to 200 °C, and the degradation was indicated by three deflections respectively at 266, 451 and 596 °C, corresponding to weight losses of 51 and 33 %, respectively (Appendix, Figure 30).

3.2.6 Recognition ability of H4 and H5 toward Kac guest through ^{1}H NMR

To investigate if steric hindrance or electronic effects of the substituents in the *meta*- position of the cavity aryl groups could affect the complexation ability of the receptor, an NMR titration was performed using Kac as guest. Deuterated methanol was used as solvent for both guest [mM] and hosts [mM] since the receptors are not water soluble.

The ^{1}H NMR titration using **H4** as receptor is reported in Figure 22.

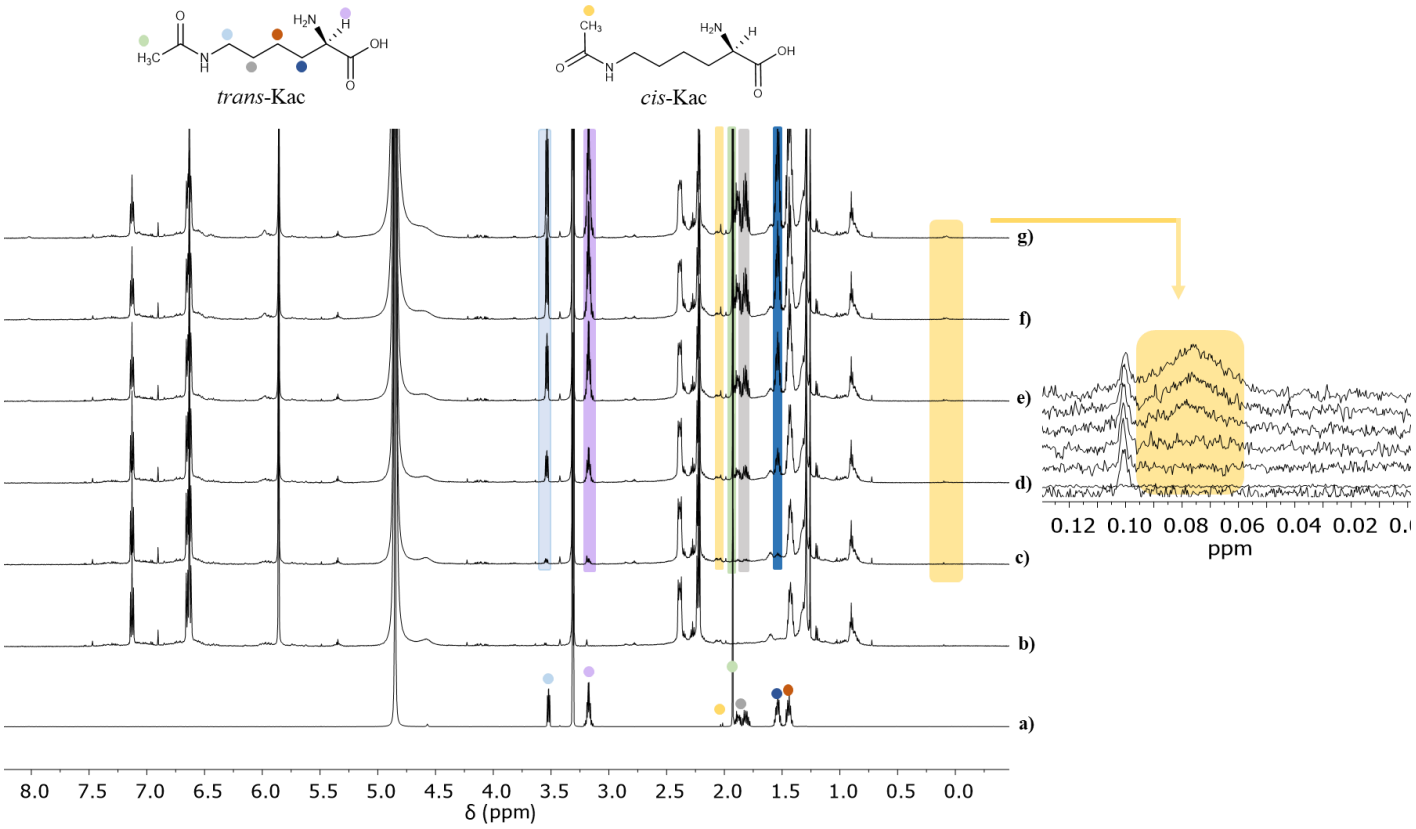


Figure 22. ^{1}H NMR (600 MHz, MeOD-d_4, 298K) spectra acquired during the titration of incremental amounts of a deuterated methanol solution of Kac in **H4**. **a)** Kac, **b)** H4, 3.2 mM **c)** H4 + 0.25 eq of Kac, **d)** H4 + 0.5 eq of Kac, **e)** H4 + 1 eq of Kac, **f)** H4 + 1.5 eq of Kac, **g)** H4 + 2 eq of Kac.

From the ^{1}H NMR results, it is possible to note that **H4** is able to bind the target guest since an upfield shift from 2.02 to 0.08 ppm of the signal related to the methyl group of the guest in its *cis* conformation (yellow mark) is present. As for **H1**, no signal shift of the *trans*-conformer was recorded, confirming that this isomer is not complexed by the calix[4]pyrrole receptor. As before, the binding constant was directly calculated from ^{1}H NMR spectra using the same theoretical binding model that considers the existence of an equilibrium between the two isomers in solution, and the exclusive formation of a 1:1 complex between the receptor and the *cis*-conformer. The value of the binding constant resulted $K_a = 10^3\ M^{-1}$.

The ^{1}H NMR titration using **H5** is reported in Figure 23.

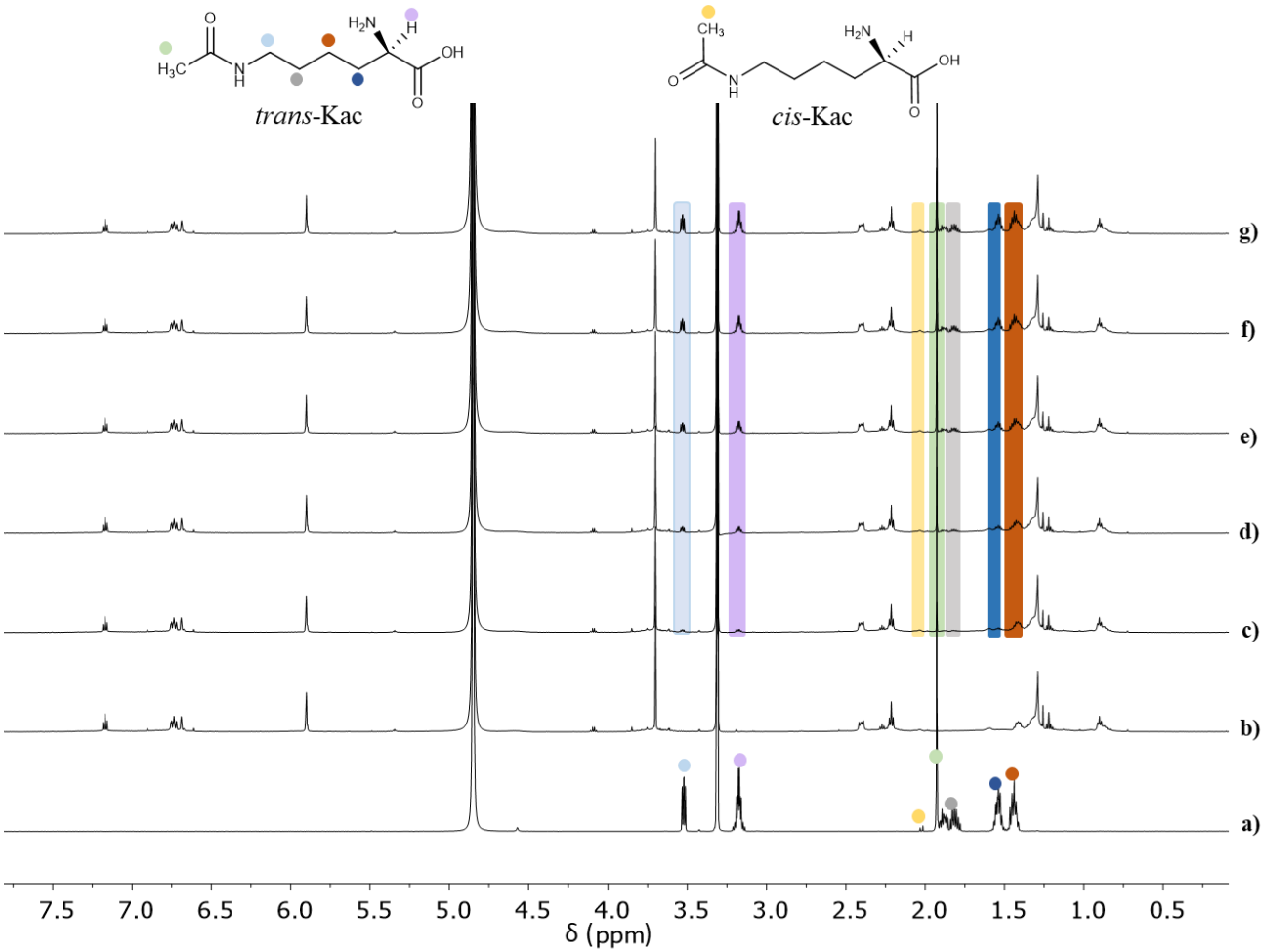


Figure 23. ^{1}H NMR (600 MHz, MeOD-d_4, 298K) spectra acquired during the titration of incremental amounts of Kac in **H5** with. **a)** Kac, **b)** H5, 3 mM, **c)** H5 + 0.25 eq of Kac, **d)** H5 + 0.5 eq of Kac, **e)** H5 + 1 eq of Kac, **f)** H5 + 1.5 eq of Kac, **g)** H5 + 2 eq of Kac.

In this case, no shifts of both *trans* and *cis* conformer signals were observed, proving that the guest is not complexed by the receptor.

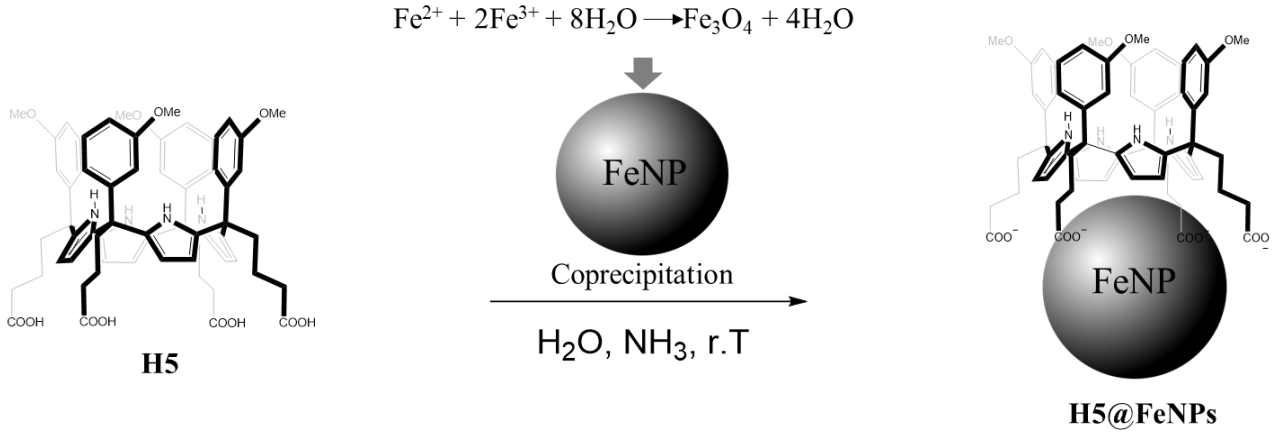
We speculated that the different affinity of the guest for the two-receptor cavity could be related to electronic reasons. The OH substituents can increase the electron density of the cavity both by inductive and mesomeric effect, while the methoxy groups can have only an inductive effect. Moreover, OHs can significantly contribute to the binding since they are both hydrogen bond donor and acceptor, and can establish H-bonding interactions with the aminoacidic moiety of the guest, which is oriented outside the cavity.

Despite the opposite results, it was decided to graft onto the NPs both receptors: **H5** will be used as control experiment in the enrichment tests to exclude unspecific interactions of Kac with the NPs surface, while **H4** will be used as working receptor.

3.2.7 Preliminary test of H5 grafting onto FeNPs and characterization

With the aim of optimizing the grafting protocol of the receptors, firstly the anchorage of **H5** on FeNPs surface was performed.

The followed protocol is the same described in Chapter 2, namely a co-precipitation method (Scheme 4).



Scheme 4. Synthesis of FeNPs functionalized with **H5** receptor.

The system **H5**@FeNPs was prepared firstly suspending **H5** in a basic water solution (NH_3 30% aq.) to allow the deprotonation of the carboxylic groups. Methanol was then added to solubilize the receptor. Separately, $FeCl_3(H_2O)_6$ and $FeCl_2(H_2O)_4$ were solubilized in water, and the so obtained solutions were poured in fast succession in the basic solution containing **H5**. After a vigorously manually stirring of the suspension, the solution turned to black colour, indicating the nanoparticles formation. The obtained nanoparticles were let to precipitate using a magnet, the supernatant removed, and the NPs were washed with water. Finally, the **H5**@FeNPs system was dried under vacuum and characterized by ATR-IR and TGA. The FT-IR spectrum was compared to the one of the bare NPs and **H5** to assess the presence of the calix[4]pyrrole on the NPs surface. TGA was also used to determine the grade of FeNPs functionalization, and it was carried out with a temperature ramp between 25 and 800 °C, and an oxygen flow rate of 20 °C/min (Appendix, Figure 32). The obtained thermogram showed a first deflection around 89°C, corresponding to a loss in weight of 6 % of the total, attributable to the residual solvent. The second and most important loss in weight is between 150 and 500 °C, with a significant loss weight of 21% at 330°C, associated to the degradation of the cavitand. The degree of functionalization obtained is about 21%.

In Figure 24, a comparison between the FT-IR spectra of **H5**@FeNPs, **H5** and bare NPs is reported. Considering the spectrum of **H5**@FeNPs at 1405.83 cm^{-1} the stretching band of the C=O bond decorating the lower rim of the calix[4]pyrrole is visible. The peak is shifted respect to the FT-IR spectrum of **H5** since in **H5**@FeNPs carboxylic ions chelating the NPs' surface are now present. At

2919 cm^{-1} a broad band associated to the aliphatic C-H groups of the alkyl chains is present, while at 1244 cm^{-1} the C-N stretching band is observable. As evident from the comparison with the IR spectrum of non-functionalized FeNPs, the stretching of the Fe-O bonds is visible at 552 cm^{-1}. All these results confirm the successful grafting of **H5** on the FeNPs.

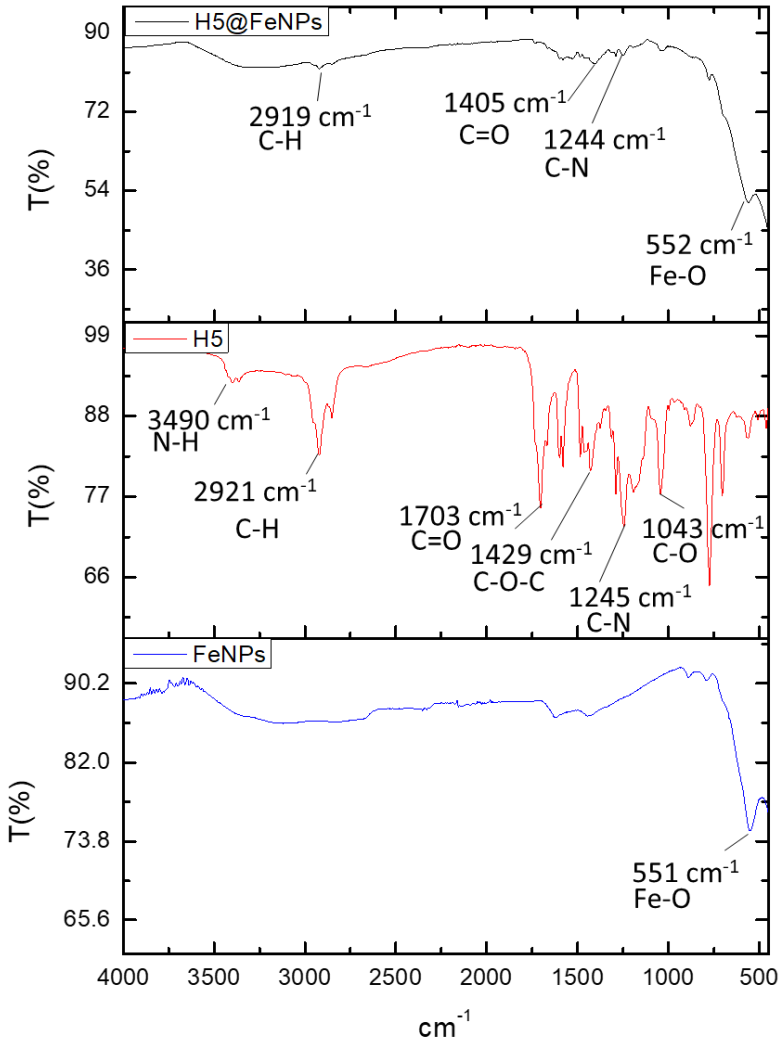


Figure 24. Comparison between ATR-FTIR spectra of **H5**@FeNPs, **H5**, and pristine FeNPs (*from top to bottom*).

Both the XRD diffraction patterns of pristine FeNPs (Figure 25A) and the functionalized nanoparticles **H5**@FeNPs (Figure 25B) perfectly reflected the pattern of standard Fe_3O_4. This iron oxide is present as the predominant crystalline phase, suggesting that the co-precipitation process to obtain functionalized NPs induces the formation of magnetite phase, as expected.

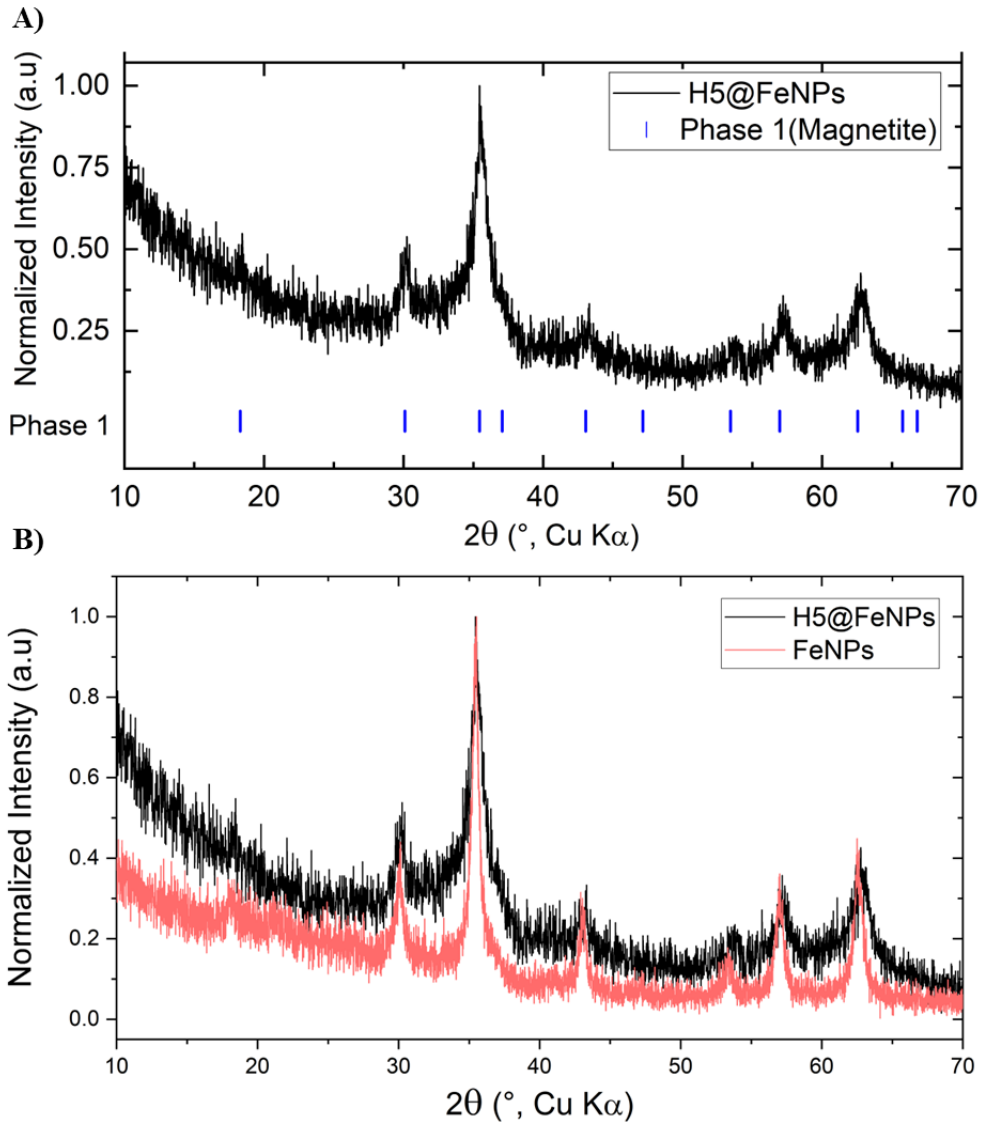


Figure 25. XRD diffraction pattern of **A)** functionalized FeNPs. and **B)** functionalized **H5**@FeNPs (*black pattern*) compared to the pristine FeNPs (*red pattern*). The blue lines in A) indicate the theoretical position of magnetite signals.

Future works will foresee the grafting of **H4** receptor onto FeNPs surface *via* co-precipitation method, and the two NPs systems will be used in enrichment experiments for the recognition of Kac residues in digested histone proteins. In particular, **H5**@FeNPs will be used as negative control together with bare NPs.

3.3 Conclusions

In this work the ability of different calix[4]pyrroles in binding acetylated lysine were tested. Kac represents an important target residue in epigenetic field since it is responsible as post translational modification of numerous human diseases.

At the beginning, three types of calix[4]pyrrole were conceived to be studied in the recognition event, namely **H1**, **H2** and **H3**.The three receptors differ in terms of cavity depth. Receptors **H1** and **H2** were already available as pure products in the research laboratory, while, unfortunately, it was not possible to synthesize host **H3**; thus only the recognition abilities of **H1** and **H2** towards Kac were assessed by NMR titration. The best performances were obtained for host **H1**, which complexed the Kac *cis* isomer with a $K_a \approx 10^3\ M^{-1}$. **H2**, which possesses a deeper cavity, failed in complexing Kac. This inefficiency is mainly due to a negative contribution in terms of enthalpy loss for the desolvation of the aminoacidic group to fit the cavity, not compensated by an entropy gain. In fact, in the binding process not only geometric compatibility between host and guest is important, but also enthalpy-entropy compensation plays a crucial role.

For FeNPs grafting, receptors **H4** and **H5** functionalized at the lower rim with four carboxylic groups were successfully synthetized. These two receptors differ from **H1** for the presence of *meta* substituents on the cavity aryl groups, while the cavity depth, essential for a successful complexation, is identical. These two hosts were considered since the tetra-ester intermediate was already available in the lab. Titration experiments were performed as well to assess their binding properties and to investigate eventual electronic effects of the two different substituents. As a result, receptor **H4**, presenting four OHs as *meta* substituents, was suited for Kac binding, while **H5**, presenting four methoxy groups on the cavity aryl groups, resulted inefficient in Kac complexation.

A first attempt of grafting **H5** on FeNPs was performed exploiting the co-precipitation method. The characterization by IR-ATR, TGA and XRD confirmed the presence of the receptors onto the FeNPs surface. Further characterization of the surface morphology and the estimation of the functionalized NPs' dimension will be performed *via* TEM analyses. At the same time, the grafting of receptors **H4** onto the NPs will be pursued and the obtained system, fully characterized, will be then tested in enrichment experiments, using **H5**@FeNPs and bare NPs as negative control.

3.4 Experimental section

3.4.1 Reagent, Chemicals and Instruments

All reagents and solvent were purchased from certified commercial sources and used as received, without further purification. For the synthesis, all solvents were dried and distilled according to standard procedures known in the literature.[38]

Thin layer chromatography was performed using TLC Analytical Chromatography F254®, Merck.

^{1}H and ^{31}P NMR spectra were recorded on Bruker 400MHz AVANCE using DMSO-d_6, D_2O, CD_3OD and $CDCl_3$ as solvents. All chemical shifts (δ) were reported in parts for million (ppm) relative to proton resonances resulting from incomplete deuteration of NMR solvents. The abbreviations: s, d, t, m and sb indicated the spectrum peaks referred to: singlet, doublet, triplet, multiplet and broad singlet, respectively. The coupling constant (J) are expressed in Hz.

ESI-MS characterization experiments were performed on a Waters ACQUILITY Ultra Performance LC HO6UPS-823M with ESI source ionization (electrospray ionization) in positive modality.

IR-ATR analyses were performed using Bruker FTIR LUMOS.

TGA analysis have been achieved using TGA/DSC 1 with gas controller GC10 (pure air/nitrogen), Mettler Toledo.

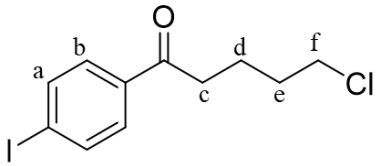
The PXRD patterns of the samples were collected using Ni-filtered Cu K radiation (λ_{Ka1} = 1.5406 Å, λ_{Ka2}= 1.5444 Å), on a Rigaku SmartLab XE diffractometer equipped with a HyPix-3000 detector. The data were processed with SmartLab Studio II (by Rigaku). PXRD patterns were collected in Bragg-Brentano geometry in the 2θ range 10-80°, placing the sample on a silicon zero background specimen holder.

Compound 1

In a dry Schlenk (kept in the hoven for 24 hours), magnesium (600 mg, 24.7 mmol) was dissolved in 8 mL of distilled THF. The suspension was sonicated for 10 minutes. Then, a solution of 1-Bromo-4-chlorobutane (4 g, 23.3 mmol) in 24 mL of distilled THF was slowly added to the mixture (dropping time: 1 hour, under magnetic stirring). After the addition, the mixture was stirred at room temperature until complete solubilization (40 minutes). The solution containing the Grignard product was immediately used for the following step, without any further purification.

Compound 2

4-iodobenzoylchloride (6.22 g, 23.3 mmol) and tris(acetylacetonate)Iron(III) (412 mg, 1.17 mmol) were solubilized in 24 mL of dry THF. The solution was kept at -78°C using a mixture of dry ice and acetone. Then, the Grignard compound **1** was slowly added to the solution with a cannula. After 2 minutes a colour shift from orange to dark brown was observed. The mixture was reacted overnight at -78°C under magnetic stirring, keeping the flask covered with aluminum foil. The crude was quenched by slowly adding 40 mL of water and then, after evaporation of the THF under reduce pressure, extracted by adding 40 mL of DCM. The organic phase was washed with HCl 1N (2 x 40 mL) and water (2 x 40 mL), dried with sodium sulphate and the solvent evaporated under reduced pressure. The obtained red oil was purified by column chromatography (eluent DCM/Hexane 4:6). Compound **2** was obtained pure, as a white solid, in 23% of yield.



^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.86 (d, 2H, **H_a**, J= 8.6 Hz), 7.69 (d, 2H, **H_b**, J= 8.58 Hz), 3.61 (t, 2H, **H_c**, J= 6.19 Hz), 3 (d, 2H, **H_f**, J= 6.82 Hz), 1.91 (m, 4H, **H_d-H_e**).

Compound 3

In a 50 mL round bottomed flask equipped with a magnetic stir bar, ketone **2** (1.76g, 4.74 mmol) and tributylmethylammonium chloride as template agent (3.35g, 14.2 mmol) were solubilized in 14 mL of dry DCM, keeping the system under magnetic stirring at room temperature for 15 minutes. Then, distilled pyrrole (0.318 g, 4.74 mmol) and subsequent HCl (4M in dioxane, 4mL, 14.2 mmol) were added drop by drop to the solution. The colour changed in 10 minutes from transparent to yellow, and finally to dark brown. The mixture was reacted under stirring at room temperature for 3 days, keeping the flask covered with aluminium foil. To quench the reaction, $NaHCO_3$ was added to the organic phase, which was then extracted and washed several times with water. Then, the organic phase was dried with sodium sulphate and the solvent evaporated under reduced pressure. The obtained crude was purified by column chromatography (eluent DCM/Hexane 4:6) and then crystallized dissolving

the solid in DCM, adding slowly some drops of ACN and keeping in the fridge for 24. The obtained withe crystals were filtrated and dry under vacuum. The final product **3** was obtained as a pale white powder in 14% of yield.

^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.56 (d, 2H, Ar***H***, J= 8.58 Hz), 6.87 (d, 2H, Ar***H***, J= 8.59 Hz), 5.87 (d, 8H, $\mathbf{H_{\beta Pr}}$, J= 2.7 Hz), 3.51 (m, 8H, -C***H*****$_2$**Cl), 2.27 (m, 8H, -CC***H*****$_2$**CH_2), 1.76 (m, 8H, -C CH_2C***H*****$_2$**-), 1.41 (m, 8H, -C***H*****$_2$**CH_2Cl).

Compound 4

In dry conditions, compound **3** (50 mg, 0.0336 mmol), $Pd(PPh_3)_2Cl_2$ (2.5 mg, 0.00336 mmol) and copper iodide (1 mg, 0.00336 mmol) were solubilized in 5 mL of distilled THF and 5 mL of distilled Et_3N. Trimethylsilyl acetylene (140 μL, 1.008 mmol) was added. The mixture was reacted under magnetic stirring, firstly at 40°C for 3 hours observing a change in colour from dark orange to brown, and then overnight at room temperature. The solvent was evaporated under reduced pressure and the solid dissolved in DCM-H_2O. The organic phase was extracted and washed several times with water, then it was dried with sodium sulphate, and, after filtration, the solvent evaporated under reduced pressure. The obtained crude was purified by column chromatography (gradient eluent: DCM/Hexane 4:6 first and then DCM) and crystallized dissolving the solid in the minimum amount of DCM, slowly adding some drops of methanol, and keeping the solution in the fridge for 24 hours. The obtained crystals were filter off and dry under vacuum. The final product **4** was obtained as a pale-yellow powder in 90% of yield.

^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.48 (bs, 4H, N***H***), 7.36 (d, 2H, Ar***H***, J= 8.47 Hz), 7.05 (d, 2H, Ar***H***, J=8.51 Hz), 5.84 (d, 2H, $\boldsymbol{H_{\beta Pr}}$, J= 2.65 Hz), 3.50 (t, 8H, -C***H*****$_2$**Cl, J= 6.53 Hz), 2.29 (m, 8H, -CC***H*****$_2$**CH_2), 1.75 (m, 8H, -CCH_2C***H*****$_2$**-), 1.38 (m, 8H, -C***H*****$_2$**CH_2Cl), 0.28 (s, 36H, Si(C***H*****$_3$**$)_3$).

Compound 5

In a 50 mL round bottomed flask, compound **4** (40 mg, 0.0306 mmol) was dissolved in 20 mL of methanol and 3 mL of THF and sonicated until the complete solubilization. K_2CO_3 (40 mg, 0.275 mmol) was added, and the mixture was reacted at room temperature under magnetic stirring overnight. The solvent was evaporated under reduced pressure, and the crude extracted with DCM/H_2O. The organic phase was washed several times with water, dried with sodium sulphate, and, after filtration, the solvent was evaporated under reduced pressure. Compound **5** was obtained in quantitative yield without any further purification.

^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.37 (d, 8H, Ar***H***, J= 8.41 Hz), 7.09 (d, 8H, Ar***H***, J= 8.24 Hz), 5.87 (bd, 8H, $\boldsymbol{H_{\beta Pr}}$), 3.52 (t, 8H, -C***H*****$_2$**Cl, J= 6.50 Hz), 3.10 (s, 8H, -C$\boldsymbol{H}_{\text{alkyne}}$), 2.30 (m, 8H, -CC***H*****$_2$**CH_2-), 1.77 (m, 8H, -CCH_2C***H*****$_2$**-), 1.42 (m, 8H, -C***H*****$_2$**CH_2Cl).

Compound 6

In dry conditions, compound **3** (50 mg, 0.0336 mmol), $Pd(PPh_3)_2Cl_2$ (2.5 mg, 0.00336 mmol) and copper iodide (1 mg, 0.00336 mmol) were dissolved in distilled THF and Et_3N in 1:1 ratio. 2-Methyl-3-butyn-2-ol (100 μL, 1.008 mmol) was added. The mixture was reacted under magnetic stirring, firstly at 40°C for 3 hours observing a change in colour from dark orange to brown, and then overnight at room temperature. The solvent was evaporated under reduced pressure and the solid extracted in DCM/H_2O, washing the organic phase several times with water. The organic phase was dried with sodium sulphate and the solvent evaporated under reduced pressure. The obtained crude was purified by column chromatography (ethyl acetate as eluent) and crystallized dissolving the solid in the minimum amount of DCM, slowly adding some drops of ethyl acetate and methanol, and keeping the solution in the fridge for 24 hours. The obtained crystals were filter off and dry under vacuum. The final product **6** was obtained as a pale-yellow powder in 90% of yield.

^{1}H NMR ($CDCl_3$, 400 MHz): δ ppm = 7.28 (bd, 8H, Ar***H***), 7.02 (bd, 8H, Ar***H***), 5.88 (d, 8H, $\boldsymbol{H_{\beta Pr}}$, J= 2.62 Hz), 3.50 (t, 8H, -C$\boldsymbol{H_2}$Cl, J= 6.53 Hz), 2.30 (m, 8H, -CC$\boldsymbol{H_2}CH_2$-), 1.74 (m, 8H, -CCH_2C$\boldsymbol{H_2}$-), 1.66 (s, 24H, -C(C$\boldsymbol{H_3}$)$_2$), 1.37 (m, 8H, -C$\boldsymbol{H_2}CH_2$Cl).

Compound 7

In a 100 mL round bottomed flask, compound **6** was solubilized in distilled pyridine. The solution was reacted under magnetic stirring, at 95°C, for 2 days. A yellow precipitate was observed. The solid was isolated, purified by trituration in DCM, filtered off and washed with DCM and hexane. Compound **7** was obtained pure in 57% of yield.

^{1}H NMR (D_2O, 400 MHz): δ ppm = 8.53 (m, 8H, Ar**H**), 8.53 (t, 4H, Ar$\mathbf{H_{Py}}$), 7.98 (t, 8H, Ar$\mathbf{H_{Py}}$), 7.35 (m, 8H, Ar**H**), 6.90 (m, 8H, Ar$\mathbf{H_{Py}}$), 5.90 (bd, 8H, $\boldsymbol{H_{\beta Pr}}$), 4.48 (t, 8H, -C$\boldsymbol{H_2}$Cl, J= 6.81 Hz), 2.35 (m, 8H, -CC$\boldsymbol{H_2}CH_2$-), 1.82 (m, 8H, -CCH_2C$\boldsymbol{H_2}$-), 1.67 (s, 24H, -C(C$\boldsymbol{H_3}$)$_2$), 0.99 (m, 8H, -C$\boldsymbol{H_2}CH_2$Cl).

Compound H4

The pure intermediate tetra-ester calix[4]pyrrole (40 mg, 0.0334 mmol) was dissolved in 3 mL of THF. Separately, LiOH (16 mg, 0.668 mmol) was dissolved in 1 mL of water and sonicated until complete solubilization and added to the first solution. The mixture was reacted at 40°C under magnetic stirring for 24 hours. Then, THF was removed under reduced pressure and 10 mL of HCl 1M were added, observing the formation of an orange precipitate that was filtered off and washed with water several times. **H4** was obtained in quantitative yield, without any further purification.

^{1}H NMR (Acetone-d_6, 400 MHz): δ ppm = 8.60 (s, 4H, N***H***), 8.16 (s, 4H, O***H***), 7.11 (m, 4H, Ar***H***), 6.59 (m, 4H, Ar***H***), 6.54 (m, 4H, Ar***H***), 6.51 (m, 4H, Ar***H***), 6.04 (d, 8H, Ar$\mathbf{H_{\beta Pr}}$, J= 2.62 Hz), 2.39 (m, 8H, -C$\boldsymbol{H_2}CH_2COOH$), 2.25 (t, 8H, -C$\boldsymbol{H_2}$COOH, J= 7.35 Hz), 1.42 (m, 8H, -CC$\boldsymbol{H_2}CH_2$-). **ATR-**

FTIR: 3395.59 cm^{-1} N-H; 2921.57 cm^{-1} C-H (aliphatic); 1703.35 cm^{-1} C=O; 1204.31 cm^{-1} C-N; 1054.77 cm^{-1} C-O. **TGA** (air, 25°C - 900 °C, flow rate: 20 °C/min): 30.43 % w/w loss at 265 °C; 20.40 % w/w loss at 430 °C; 45.56 % w/w loss at 630 °C.

Compound 8

The pure intermediate tetra-ester calix[4]pyrrole (46 mg, 0.0384 mmol) was dissolved in 40 mL of acetonitrile. Cs_2CO_3 (62 mg, 0.192 mmol) and iodomethane (60 μL 1.0752 mmol) were sequentially added. The mixture was reacted under reflux for 4 hours, under magnetic stirring. The solvent was removed under reduced pressure and the crude was dissolved in diethyl ether/water. The organic phase was washed several times with water, dried with sodium sulphate and the solvent evaporated under reduced pressure. Compound **8** was obtained pure in quantitative yield without any further purification.

^{1}H NMR (Acetone-d_6, 400 MHz): δ ppm = 8.58 (s, 4H, N***H***), 7.17 (m, 4H, Ar***H***), 6.73 (m, 4H, Ar***H***), 6.59 (m, 4H, Ar***H***), 6.55 (m, 4H, Ar***H***), 6.06 (d, 8H, Ar$\boldsymbol{H}_{\beta Pr}$, , J= 2.66 Hz), 4.07 (q, 8H, -COOC***H₂***CH_3, J= 7.1 Hz), 3.78 (s, 12H, -OC***H₃***), 2.41 (m, 8H, -C***H₂***CH_2COOH), 2.25 (t, 8H, -C***H₂***COOEt, J= 7.26 Hz), 1.39 (m, 8H, -CC***H₂***CH_2-), 1.21 (t, 12H, -COOCH_2C***H₃***, J= 7.12 Hz).

Compound H5

Compound **8** (47 mg, 0.0374 mmol) was dissolved in 3 mL of THF. Separately, LiOH (18 mg, 0.748 mmol) was dissolved in 1 mL of water and sonicated until it was completely solubilized and added to the first solution. The mixture was reacted at 60°C under magnetic stirring for 24 hours. Then, THF was removed under reduced pressure and 10 mL of HCl 1M were added, observing the formation of an orange precipitate that was filtered off and washed with water several times. **H5** was obtained in 78 % yield, without any further purification.

^{1}H NMR (Acetone-d_6, 400 MHz): δ ppm = 8.58 (s, 4H, N***H***), 7.17 (m, 4H, Ar***H***), 6.72 (m, 4H, Ar***H***), 6.60 (m, 4H, Ar***H***), 6.55 (m, 4H, Ar***H***), 6.07 (d, 8H, Ar$\boldsymbol{H}_{\beta Pr}$, J= 2.66 Hz), 2.42 (m, 8H, -C***H₂***CH_2COOH), 2.26 (m, 8H, -C***H₂***COOH), 1.39 (m, 8H, -CC**H₂**CH_2-). **ATR-FTIR**: 3490 cm^{-1} N-H; 2921 cm^{-1} C-H (aliphatic); 1703 cm^{-1} C=O; 1429 cm^{-1} C-O-C; 1245 cm^{-1} C-N; 1043 cm^{-1} C-O. **TGA** (air, 25°C - 900 °C, flow rate: 20 °C/min): 51.22 % w/w loss at 266 °C; 15.18 % w/w loss at 451 °C; 33.29 % w/w loss at 596 °C.

Synthesis of ferromagnetic nanoparticles (FeNPs)

Iron magnetite nanoparticles were synthesized by alkaline co-precipitation of Fe^{3+} and Fe^{2+}. In two separate vials, $FeCl_3 \cdot 6H_2O$ (0.73 g, 2.70 mmol) and $FeCl_2 \cdot 4H_2O$ (0.27 g, 1.36 mmol) were solubilized in 3 mL and 1.5 mL of water, respectively. In a flask, 2.2 mL of NH_3 (30% aqueous) were added to

50 mL of water and, keeping the system under manual stirring, the two solutions of $FeCl_3$ and $FeCl_2$, were added in quick succession. A dark coloured suspension was immediately observed and left to decantation at room temperature for 2 hours. The supernatant was filtered off and the solid residue washed twice with water. The obtained FeNPs were isolated from the solvent by magnetic decantation and dried under nitrogen flux. **ATR-FTIR**: 3000 cm^{-1} OH, 551 cm^{-1} Fe-O.

Functionalized FeNPs

In two different flasks, $FeCl_2 \cdot 4H_2O$ (21.7 mg, 109 μmol) and $FeCl_3 \cdot 6H_2O$ (60 mg, 224.95 μmol) were dissolved in 1.5 and 3 mL of water, respectively. Separately, the cavitand was dissolved in a previously prepared solution of NH_3(0.125 mL), some drops of methanol and H_2O (20 mL). To this solution, kept under manual stirring, the two Fe^{3+} and Fe^{2+} solutions were added sequentially. The formation of a dark precipitate was visible. After magnetic decantation, the supernatant was filtered off and the solid residue washed twice with water and dried under nitrogen flux.

3.5 Appendix

3.5.1 NOESY NMR characterization

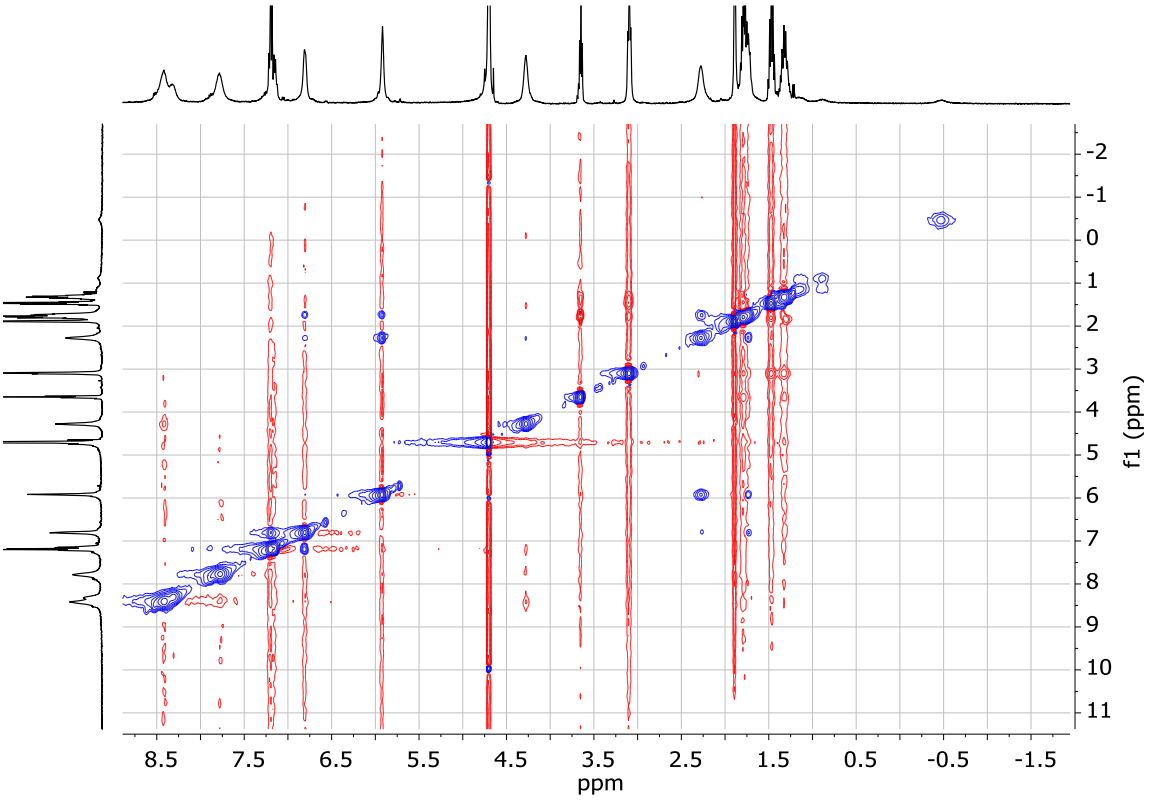


Figure 26. ^{1}H NOESY-NMR spectrum of **H1** + 5eq of Kac (D2O, 400MHz, 25°C).

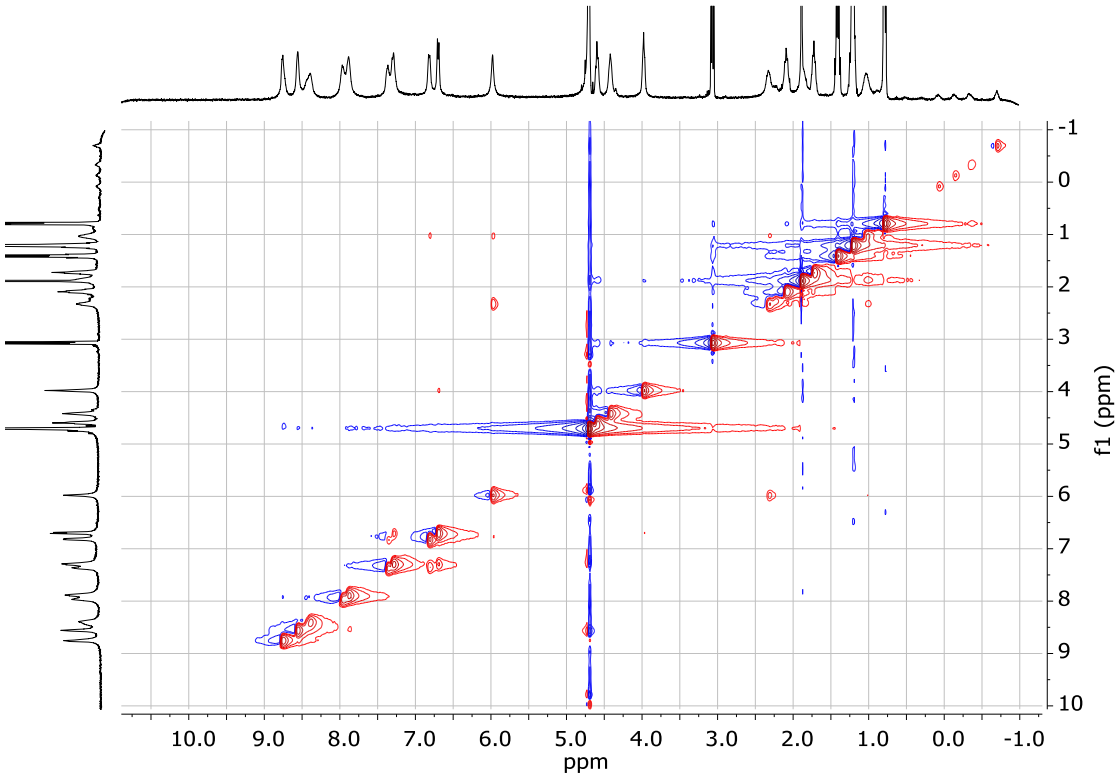


Figure 27. ^{1}H NOESY-NMR spectrum of **H2** + 4eq of PA (D2O, 400MHz, 25°C).

3.5.2 IR-ATR and TGA characterizations of H4 and H5

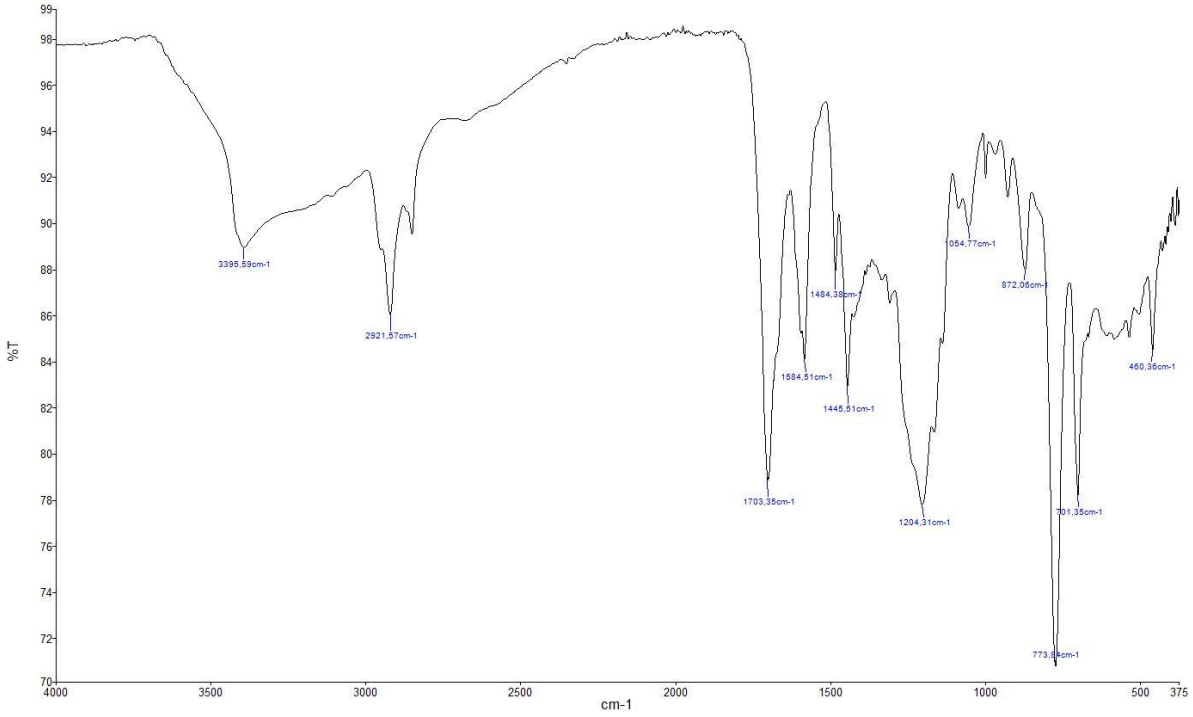


Figure 28. FT-ATR spectrum of **H4**

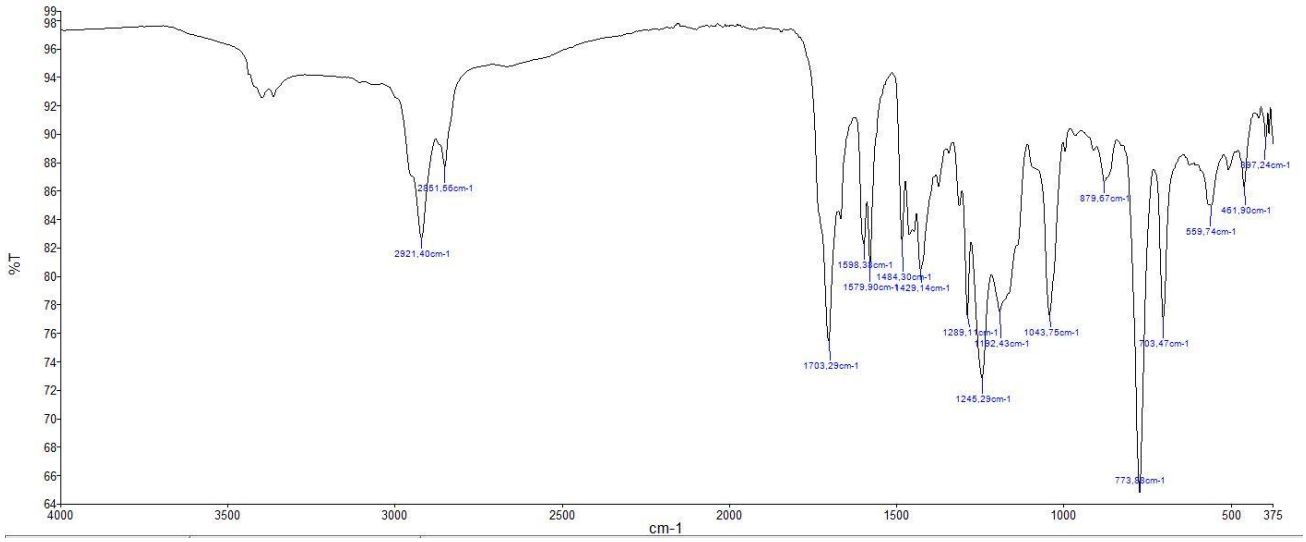


Figure 29. FT-ATR spectrum of **H5**

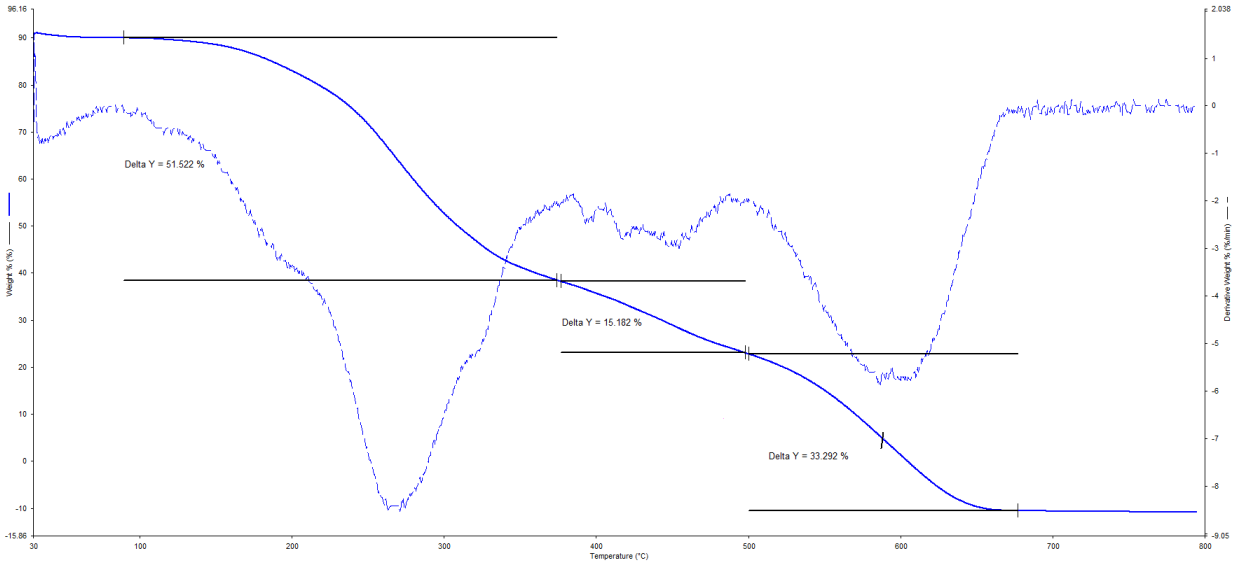


Figure 30. TGA thermogram of **H5**

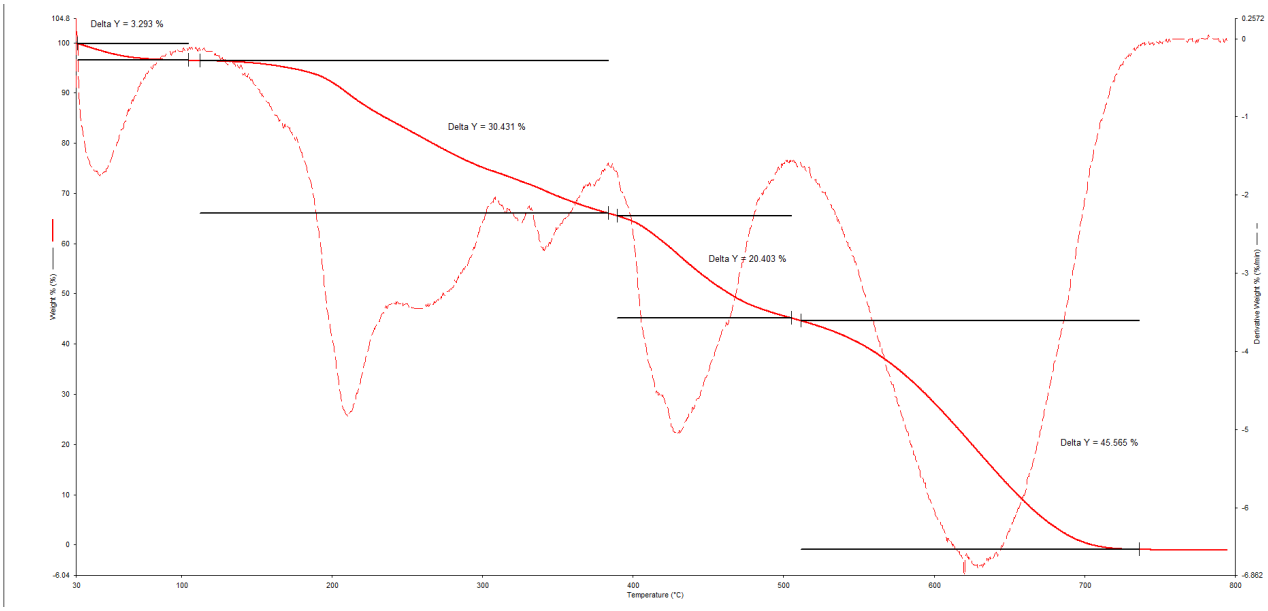


Figure 31. TGA thermogram of **H4**

3.5.3 TGA characterization of H5@FeNPs

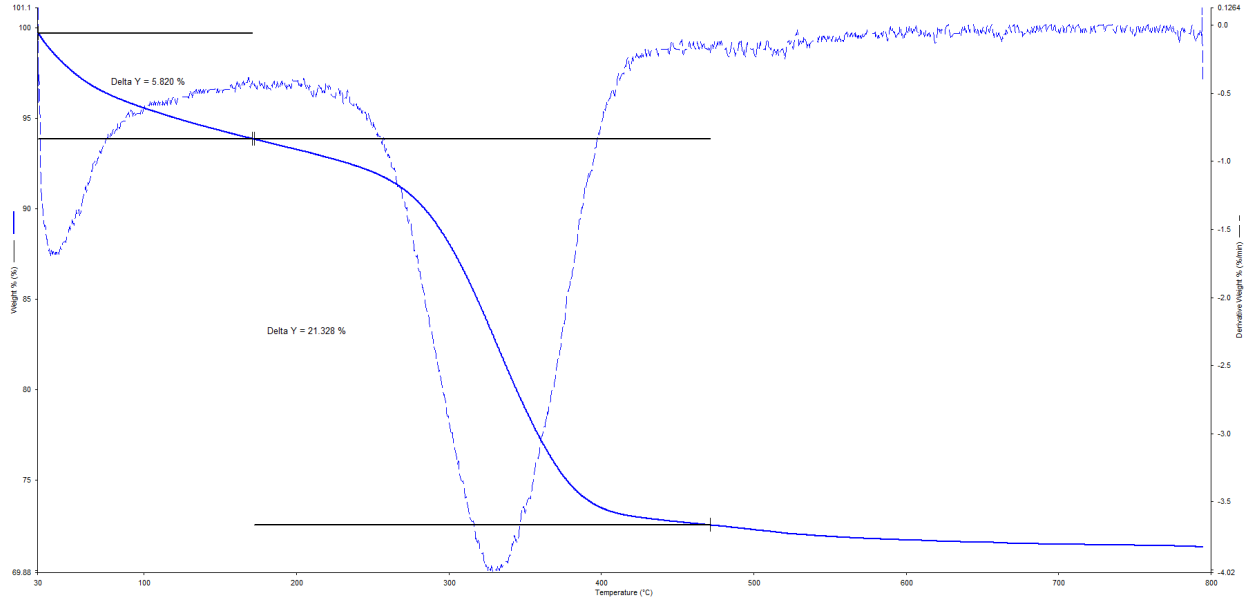


Figure 32. TGA thermogram of **H5@FeNPs**

3.6 References

1. Zhang, X., Wen, H. & Shi, X. Lysine methylation: beyond histones. *Acta Biochim. Biophys. Sin.* **44**, 14–27 (2012).
2. Levy, D. Lysine methylation signaling of non - histone proteins in the nucleus. *Cell. Mol. Life Sci.* **76**, 2873–2883 (2019).
3. Peleg, S., Feller, C., Ladurner, A. G. & Imhof, A. The Metabolic Impact on Histone Acetylation and Transcription in Ageing. *Trends Biochem. Sci.* **41**, 700–711 (2016).
4. Kalakonda, N. *et al.* Histone H4 lysine 20 monomethylation promotes transcriptional repression by L3MBTL1. *Oncogene* **27**, 4293–4304 (2008).
5. Liu, X. *et al.* The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* **451**, 10–14 (2008).
6. Zhang, X. *et al.* Catalytic Mechanism of Histone Acetyltransferase p300: From the Proton Transfer to Acetylation Reaction. *J. Phys. Chem. B* **118**, 8 (2019).
7. Jiang, J. *et al.* Investigation of the Acetylation Mechanism by GCN5 Histone Acetyltransferase. Plos one **7**, 36660 (2012).
8. Berndsen, C. E. *et al.* Catalytic Mechanism of a MYST Family Histone Acetyltransferase. *Biochemistry* **46(3)**, 623-629 (2007).
9. Delcuve, G. P., Khan, D. H. & Davie, J. R. Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. *Clinical Epigenetics* **4**, 1–13 (2012).
10. Seto, E. & Yoshida, M. Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harbor Perspectives Biology* **6(4)**, 018713 (2014).
11. Schiedel, M. & Conway, S. J. Small molecules as tools to study the chemical epigenetics of lysine acetylation. *Curr. Opin. Chem. Biol.* **45**, 166–178 (2018).
12. Katsyuba, E. & Auwerx, J. Modulating NAD^+ metabolism , from bench to bedside. *EMBO J.* **36**, 2670–2683 (2017).
13. Ford, E. *et al.* Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I *Genes Dev*. **20(9)**. 1075–1080 (2006).
14. Liszt, G., Ford, E., Kurtev, M. & Guarente, L. Mouse Sir2 Homolog SIRT6 is a nuclear ADP-ribosyltransferase. *J. Biol. Chem* **280(22)**, 21313–21320 (2005).
15. North, B. J. *et al.* The human Sir2 ortholog, SIRT2, is an NAD^+ - Dependent Tubulin Deacetylase. *Molecular Cell* **11**, 437–444 (2003).
16. Nakagawa, T., Lomb, D. J., Haigis, M. C. & Guarente, L. SIRT5 deacetylates carbamoyl phosphate synthetase 1 and regulates the urea cycle. *Cell* **137**, 560–570 (2009).
17. German, N. J. *et al.* SIRT4 coordinates the balance between lipid synthesis and catabolism by repressing malonyl CoA Decarboxylase. *Molecular Cell* **50(5)**, 686-698 (2012).
18. Michishita, E., Park, J. Y., Burneskis, J. M. & Barrett, J. C. Evolutionarily conserved and nonconserved cellular localizations and functions of human SIRT proteins. *Mol. Biol. Cell.* **16(10)**, 4623–4635 (2005).
19. Syntichaki, P., Topalidou, I. & Thireos, G. The Gcn5 bromodomain co-ordinates nucleosome

remodelling. *Nature* **404**, 414-417 (2000).

20. Cieniewicz, A. M. *et al.* The Bromodomain of Gcn5 Regulates Site Specificity of Lysine Acetylation on Histone H3. *Mol. Cell. Proteomics* **13**, 2896–2910 (2014).

21. Finnin, M. S. *et al.* Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature* **401**, 188–193 (1999).

22. Pan, B. *et al.* Epigallocatechin gallate reverses cTnI-low expression-induced age-related heart diastolic dysfunction through histone acetylation modification. *J. Cell. Mol. Med.* **21**, 2481–2490 (2017).

23. Kim, S. Y. *et al.* An alternative strategy for pan-acetyl-lysine antibody generation. *PLoS One* **11**, 1–16 (2016).

24. Escobar, L. & Ballester, P. Molecular recognition in water using macrocyclic synthetic receptors. *Chem. Rev.* **121**, 2445–2514 (2021).

25. Wu, Y. D., Wang, D. F. & Sessler, J. L. Conformational features and anion-binding properties of calix[4]pyrrole: A theoretical study. *J. Org. Chem.* **66**, 3739–3746 (2001).

26. Anzenbacher, P., Jursíková, K., Lynch, V. M., Gale, P. A. & Sessler, J. L. Calix[4]pyrroles containing deep cavities and fixed walls. Synthesis, structural studies, and anion binding properties of the isomeric products derived from the condensation of p-hydroxyacetophenone and pyrrole. *J. Am. Chem. Soc.* **121**, 11020–11021 (1999).

27. Díaz-Moscoso, A., Hernández-Alonso, D., Escobar, L., Arroyave, F. A. & Ballester, P. Stereoselective synthesis of lower and upper rim functionalized tetra-α isomers of calix[4]pyrroles. *Org. Lett.* **19**, 226–229 (2017).

28. Adriaenssens, L. & Ballester, P. Hydrogen bonded supramolecular capsules with functionalized interiors: The controlled orientation of included guests. *Chem. Soc. Rev.* **42**, 3261–3277 (2013).

29. Lee, J. *et al.* Ship in a breakable bottle: Fluoride-induced release of an organic molecule from a Pr(III)-linked molecular cage. *Chem. Commun.* **52**, 8514–8517 (2016).

30. Peñuelas-Haro, G. & Ballester, P. Efficient hydrogen bonding recognition in water using aryl-extended calix[4]pyrrole receptors. *Chem. Sci.* **10**, 2413–2423 (2019).

31. Escobar, L., Aragay, G. & Ballester, P. Super Aryl-Extended Calix[4]pyrroles: Synthesis, Binding Studies, and Attempts To Gain Water Solubility. *Chem. - A Eur. J.* **22**, 13682–13689 (2016).

32. Hernández-Alonso, D., Zankowski, S., Adriaenssens, L. & Ballester, P. Water-soluble aryl-extended calix[4]pyrroles with unperturbed aromatic cavities: synthesis and binding studies. *Org. Biomol. Chem.* **13**, 1022–1029 (2015).

33. Verdejo, B., Gil-Ramírez, G. & Ballester, P. Molecular recognition of pyridine N-oxides in water using calix[4]pyrrole receptors. *J. Am. Chem. Soc.* **131**, 3178–3179 (2009).

34. Li, Y. S. *et al.* Relative hydrophilicities of cis and trans formamides. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 19815–19820 (2019).

35. Escobar, L., Díaz-Moscoso, A. & Ballester, P. Conformational selectivity and high-affinity binding in the complexation of N-phenyl amides in water by a phenyl extended calix[4]pyrrole. *Chem. Sci.* **9**, 7186–7192 (2018).

36. Stewart, W.E. & Siddal, T.H. Nuclear magnetic resonance studies on amides. *Chem. Rev.* **70**, 517-551 (1970).

37. Sierra, A. F. *et al*. Optical Supramolecular Sensing of Creatinine. *J. Am. Chem. Soc.* **142**, 4276–4284 (2020).

38. Williams, D. B. G. & Lawton, M. Drying of organic solvents: Quantitative evaluation of the efficiency of several desiccants. *J. Org. Chem.* **75**, 8351–8354 (2010).

Chapter 4

Host-guest templated DNA duplex formation based on CB[8] heteroternary complexes

Acronyms and abbreviations

DCM	Dichloromethane
DMSO	Dimethylsulphoxyde
MeOH	Methanol
NaOH	Sodium hydroxide
KOH	Potassium hydroxide
THF	Tetrahydrofuran
THPTA	Tris (3-hydroxypropyltriazolylmethyl) amine
PEG	Polyethylene glycol
TEG	Triethylene glycol
DMAP	4-Dimethylaminopyridine
EDTA	Ethylenediaminetetraacetic acid
CB	Cucurbituril
MV	Methylviologen
BP-TMSD	Base-Pair Toehold-Mediated Strand Displacement
HG-TMSD	Host -Guest Toehold-Mediated Strand Displacement
DNA	Deoxyribonucleic acid
CuAAC	Copper-Catalysed Azide-Alkyne Cycloaddition
AAC	Azide-Alkyne Cycloaddition
ESI-MS	Electrospray Ionization - Mass Spectrometry
ESI-Orbitrap-MS	Electrospray Ionization - Orbitrap - Mass Spectrometry
TLC	Thin- Layer Chromatography
UV-Vis	Ultraviolet-visible spectroscopy

4.1 Introduction

4.1.1 Cucurbiturils

Cucurbit[n]urils (CB[n]s, n= 5-8, 10, 13-15) are synthetic macrocycles that have recently attracted great interest due to their peculiar properties, such as the ability to form stable host-guest complexes in water, their chemical inertness and non-toxicity. Cucurbit[n]urils are synthesised through an acid catalyzed condensation reaction between glycoluril and formaldehyde from which a mixture of oligomers is obtained and can be separated by exploiting their different solubility in aqueous solvents. The solubility, in fact, varies depending on the number of repetitive units forming the macrocycle.[1]

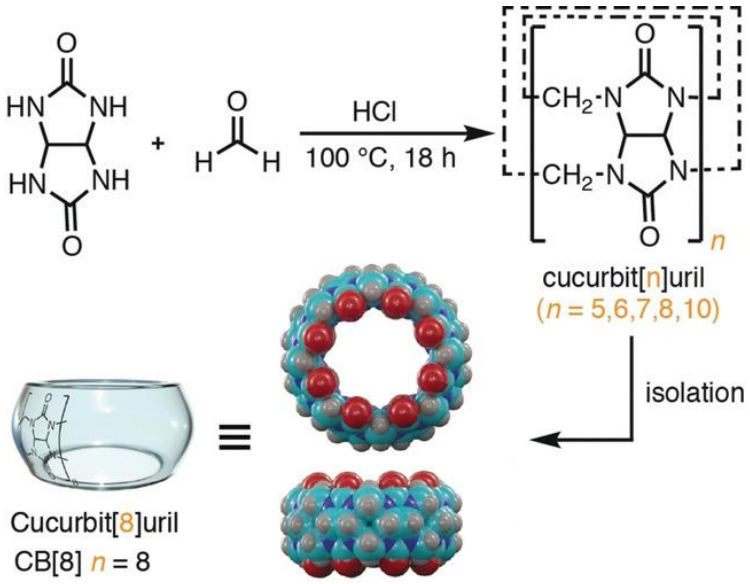


Figure 1. Synthesis of CB[n]s and isolation of CB[8].[2]

Structurally, cucurbit[n]urils are formed by *n* glycoluril units bound together *via* 2n methylene bridges, forming a toroidal macrocycle whose shape resembles that of a pumpkin. The size of the cavity varies with *n*, the depth remains constant at 9.1 Å, while the diameter increases as *n* increases (Table 1).

Table 1. Structural parameters for CB[n]s.[1]

	CB[5]	CB[6]	CB[7]	CB[8]	CB[10]
Portal diameter (Å)	2.4	3.9	5.4	6.9	9.5-10.6
Cavity diameter (Å)	4.4	5.8	7.3	8.8	11.3-12.4
Cavity volume (Å^3)	82	164	279	479	870
Outer diameter (Å)	13.1	14.4	16.0	17.5	20.0
Height (Å)	9.1	9.1	9.1	9.1	9.1

The use of CB[n]s in host-guest molecular recognition has the distinct advantage of forming complexes with high binding affinity in water. In fact, this kind of macromolecules behaves as hosts capable of assembling inclusion complexes with neutral or cationic organic guests through cation–dipole interactions, dispersion and hydrophobic forces.[3] The interior of the cavity is hydrophobic and none of the functionalities capable of acting as hydrogen bond acceptors or donors point inwards. The upper and the bottom rim (portals) present an excess of negative charge, making cucurbiturils an excellent receptor for cations (Figure 2).

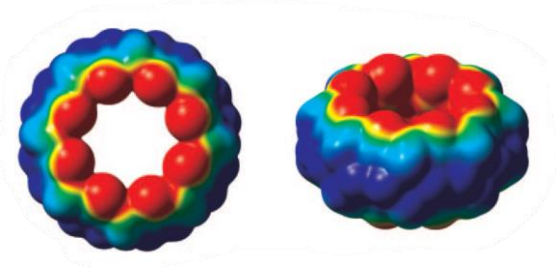


Figure 2. Electrostatic potential map of the CB8; excess of negative charge in the portal regions (*red*) and electron-poor equatorial region (*blue*).

The main driving force of the complex formation is the release of "high-energy" water molecules from the cavity of the host after the entrance of the guest. Specifically, the water confinement is unfavourable both enthalpically and entropically.[4] These water molecules experience a reduced number hydrogen bonds if compared to bulk solvent, so their liberation allows the formation of a higher number of interactions that are reflected in a relevant enthalpic contribution to the process.

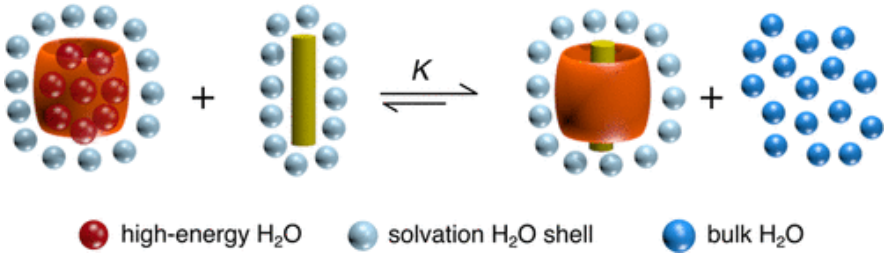


Figure 3. CB[n] host-guest complex driven by the release of high-energy water.[1]

Thanks to their properties in molecular recognition,[1] cucurbiturils have applications in multiple fields including drug delivery,[5] nanomaterials,[6] fluorescence sensing[7] and catalysis.[8]

4.1.2. Cucurbit[8]uril-based complexes

An important aspect that deserves to be considered is the shape complementarity between a guest and the cavity of cucurbituril. Depending on the nature and dimension of the guest and the size of the cavity, different types of complexations can be achieved. Complexes formed by smaller CB[n]s present a 1:1 stoichiometry while the larger ones, like CB[8] or CB[10], can host more than one guest.

As shown in Figure 4 (A), with a monocationic guest bringing a bulky hydrophobic part CB[8] tends to generate a binary complex 1:1 (*Type A*). Guests with two charges can establish cation-dipole interactions with both portals. Stoichiometry remains 1:1 because the repulsion between the charges does not allow the dimerization inside the host. For instance, in *Type B* complexes the butyl chain resides inside the cavity while in *Type C* the methyl viologen cation prefers to insert the aromatic part. A different kind of supramolecular complex results from the entrance of two monocationic guests with a proper size; *type D* homoternary 1:2 complexes are formed by two identical guests that enter the cavity in an antiparallel mode to maximize the host-guest interactions.[3]

Thus, a second type of ternary complexes can be observed: the inclusion of two guests of a different nature, resulting in a heteroternary complex with 1:1:1 stoichiometry (Figure 4, B). In this case the formation of a 1:1 binary complex with an electron-acceptor guest is initially observed, which facilitates subsequent inclusion of an electron-donor guest that otherwise would not have been complexed by the CB[8]. The driving force that governs the formation of this complex is both the instauration of charge-transfer and π-π stacking interactions between the two guests and the liberation of residual "high-energy" water molecules.[9] Indeed, it has been showed that both the hydrophobic and polar characteristics of the CB[8] cavity, due to the presence of a negative but uniform potential within, allow the stabilization of the charge-transfer complexes.

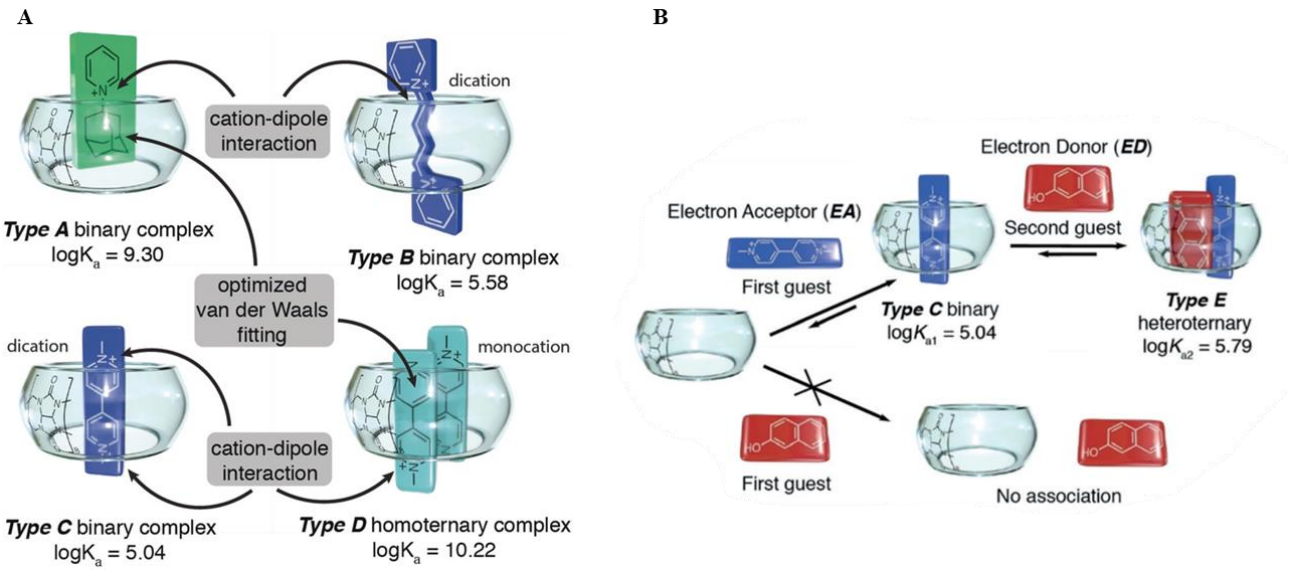


Figure 4. Expected type of CB[8]-guest complexes.[3]

In recent years, several studies have been reported exploting a wide variety of guests, such as the pair methylviologen and 2,6-dihydroxynaphtalene[10] or the pair 2,7-dimethyldiazapirene and catechol.[11] Several dicationic compounds that can be included in the CB[8] cavity are also reported in the literature, such as tetramethyl benzobis(imidazolium) salts,[12] salts of aryl-bisimidazolium,[13] aryl-viologen derivatives[5] and perylene bisdiimides.[14] With regard to homoternary complexes, however,

monocationic molecules capable of forming stable dimers once complexed are often used, such as phenylpyridinium derivatives,[15] bipyridinium salts[9] and anthracene derivatives.[16]

By exploiting the formation of a heteroternary complex it is possible to develop systems capable of acting as supramolecular fluorescent probes.[17] Examples of this are the afore mentioned complexes whose guests are 2,7-dimethyldiazapirene and catechol, or systems with pyrene and viologen derivatives (4,4'-bipyridinium salts). Viologen derivatives in particular are widely used in combination with cucurbiturils due to their synthetic versatility.[3] They are indeed able to give different complexation stoichiometries depending on the nature and number of pyridinium substituents. Furthermore, due to their structure, they are excellent electron acceptors, so they act as fluorescence quenchers.[18,19] In the supramolecular systems based on heteroternary complexes described above, in fact, the guests behaves usually as a fluorophore and a quencher. The simultaneous inclusion of the two compounds within the cavity of the cucurbituril induces a proximity effect that causes charge-transfer interactions to take place and to change the shape and/or the intensity of the fluorophore emission spectrum.

4.1.3 DNA derivatization and cucurbiturils as templating agents to induce molecular confinement

The possibility to program DNA sequences and derivatize them with a wide range of functional groups is boosting the spreading of DNA nanotechnology. The formation of a duplex between two complementary oligonucleotide strands through the establishment of specific interactions (Watson-Creek) is a fundamental tool of this field. In recent years host-guest interactions have been effectively applied to control this mechanism developing highly structured networks with a wide range of applications, such as stimuli responsive DNA switches or enzyme-free catalytic systems.[20,21] Cucurbiturils, thanks to their recognition properties and solubility in water can represent a powerful tool that can act as supramolecular input for the control of biological events.

Jayawickramarajah and *co-workers* reported various examples of DNA engineering based on CB[7].[22] They exploited a DNA strand-displacement mediated by a supramolecular input, namely the formation of a binary complex between CB[7] and adamantane derivatives. In contrast to conventional base-pair toehold-mediated strand displacement (BP-TMSD), which can result in unwanted structures or in off-target hybridization, the supramolecular approach is orthogonal to DNA base-pairing and can be easily integrated into oligonucleotides leading to DNA displacement. The authors investigated the effects of this host-guest toehold-mediated strand displacement (HG-TMSD) in different systems (Figure 5).

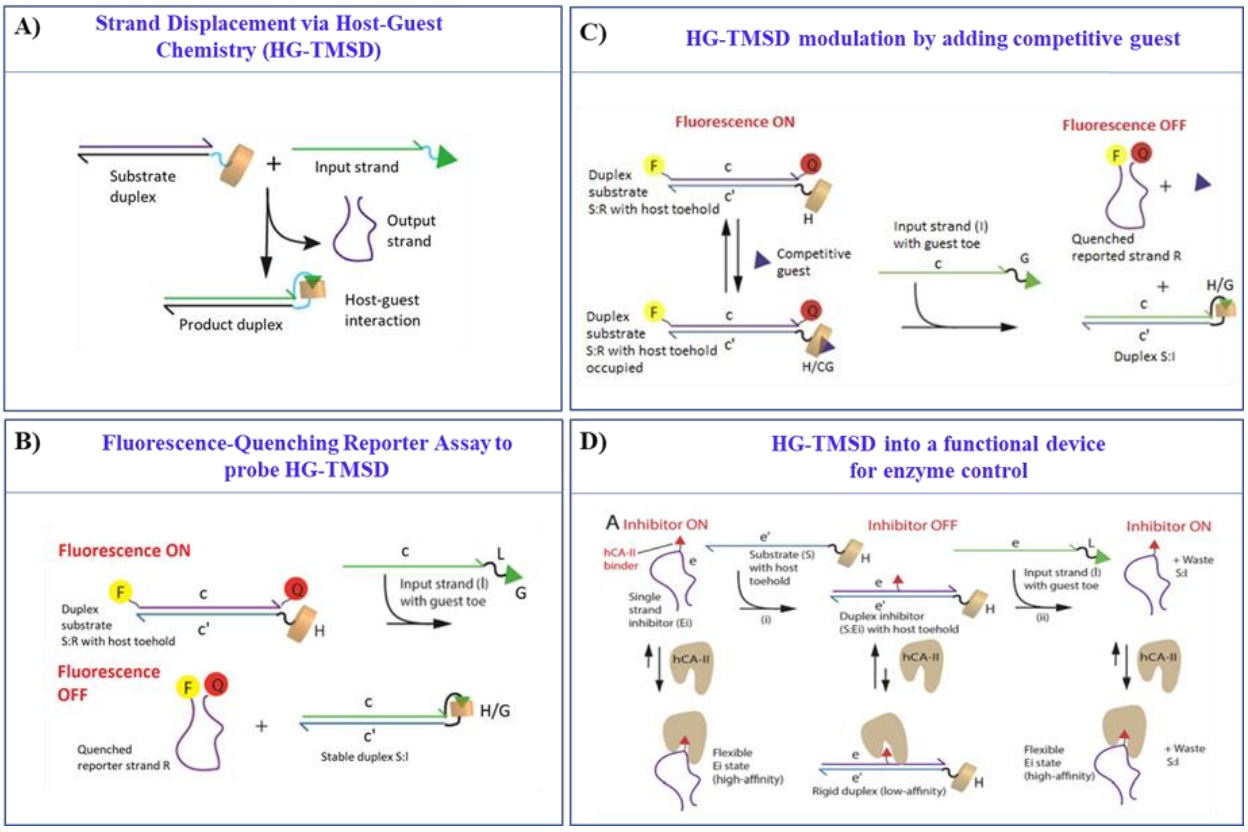


Figure 5. HG-TMSD in different approaches: **A)** Strand Displacement via Host-Guest chemistry (HG-TMSD); **B)** Fluorescence-Quenching Reporter Assay to probe HG-TMSD; **C)** HG-TMSD modulation by adding competitive guest; **D)** HG-TMSD into a functional device for enzyme control.[22]

Initially a strand derivatized with the bio-orthogonal CB[7] receptor was found to interact with a guest-linked input strand leading to the displacement of an output one (Figure 5, A). HG-TMSD was proved *via* fluorescence quenching assay; the CB[7]-functionalised strand interacted with a guest-toe input strand allowing the displacement of a quenched reporter strand (Figure 5, B). Moreover, the authors probe HG-TMSD investigating the effects of guest head-group to achieve a fine control over the rate of strand displacement, demonstrating the versatility of the supramolecular approach to strand displacement. Following the process with fluorescence, for a given input sequence the addition of secondary input molecules bearing a competitive guest such as phenylalanine, (trimethylsilyl)methyl ammonium or an adamantane derivative with high affinity to CB[7] can compete with toehold domain of the duplex and is capable to modulate HG-TMSD displacement process (Figure 5, C). Finally, to integrate this process into a functional device, HG-TMSD was incorporated into functional DNA-based systems. An ad hoc designed HG-TMSD based-DNA machine was able to control enzyme activity via toggling a DNA-sulphonamide conjugate (inhibitor of human carbonic anhydrase-II, hCA-II) from an active state ON, single-stranded, to a duplex OFF state (Figure 5, D).

Xu *et al.* exploited a ligand-invasion approach based on host-guest interactions for controllable DNA hybridization.[23] The invadent-ligand was constituted by a CB[7] receptor, while well-studied guest molecules were integrated as addition of artificial functional groups into the nucleobase sequence at

the N^4-position of cytosine or the N^6-position of adenine. The receptor, due to the competitive recognition of the guest groups, was able to break the base pairs through steric clash, disrupting the Watson-Crick interactions involved in the duplex and allowing the dissociation of the DNA duplex structure (Figure 6).

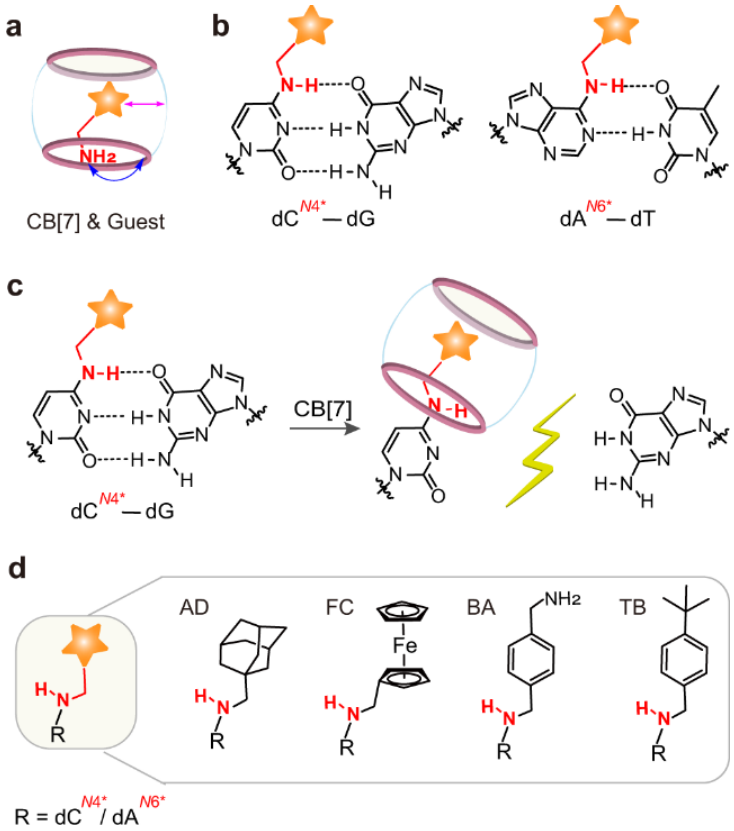


Figure 6. a) Representation of CB[7]@guest complex. **b)** Interactions between the nucleobases of the guest in the absence of CB[7]. **c)** CB[7] induces a ligand-invasion breaking the base-pair interactions resulting in a steric clush. **d)** Library of selected guest molecules.[23]

In general, the natural processes that concern protein-protein interactions and their biological networks involved thousands of reactants that are found dissolved in the same solution at concentrations too low (nM to μM) to allow random intermolecular interactions. Thus, the use of molecular templates that regulate the physical distances inducing a biomolecular confinement between molecules, scaffold that can orient functional units in space, organizing the orientation of binding sites of the receptor bringing them into proximity to initiate an effect, represents an interesting tool in biochemical processes.[24] For example, this natural colocalization process is critical in metabolism, transcriptional control, and signaling.[25] In this context, the derivatization of DNA strands in conjugation with affinity ligands involving in the interaction with multiple binding sites can allow the strands to be localized in a restricted volume, substantially increases the local concentration of the strands and promoting effective hybridization by enhancing their encounter rate. In this way, the

duplex formed is more stable than the intermolecular one resulting in a stable intramolecular hybridization (Figure 7).[26]

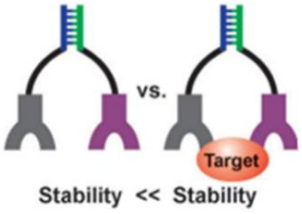


Figure 7. The binding of two properly conjugated and complementary strands to a target molecule results in a more stable intramolecular system.[26]

In most cases, the complementary DNA sequences are designed in such a way that the duplex formation cannot occur, or is unstable, in the absence of the target molecule. In fact, the DNA duplex is dependent on its length and sequence composition and moreover, the number of bases and base-stacking interaction should be considered in the design of complementary DNA sequences.[25] Thus, by properly designing DNA sequences and using biomolecule templates as substrates proximity-induced DNA assembly can be engineered to mimic molecular encountering in a confined space (Figure 8).

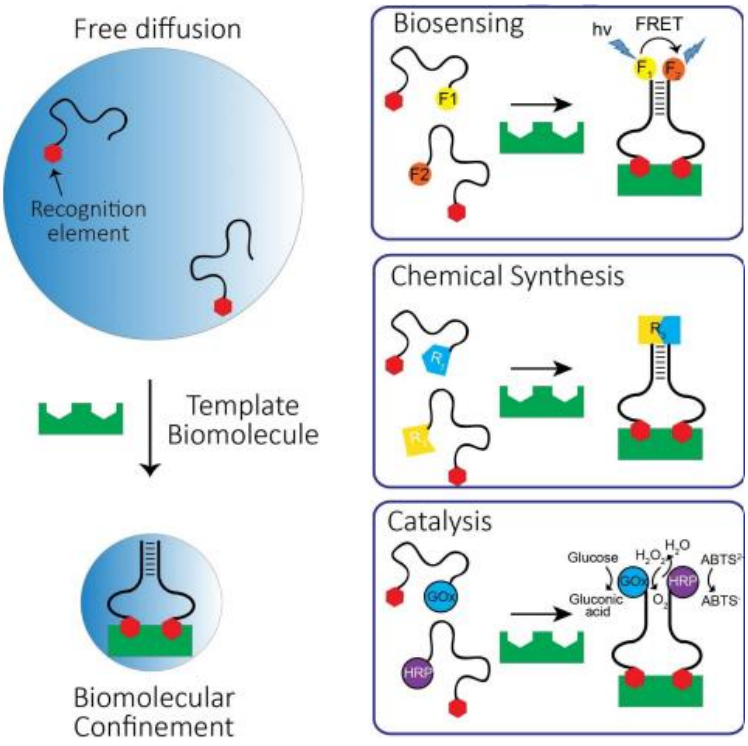


Figure 8. DNA-based system engineering through biomolecular confinement.[27]

In this chapter the design and synthetic approaches towards a host-guest templated DNA duplex formation based on the complexation properties of CB[8] are presented. The reported system, developed in collaboration with the research group of Prof. Alessandro Bertucci at University of Parma, is based on the formation of an heteroternary complex between CB[8] and two properly

derivatized oligonucleotides (Figure 9). In our design the duplex formation can be triggered both by host-guest complexation (supramolecular template) and by the introduction of a complementary DNA sequence (biological template).

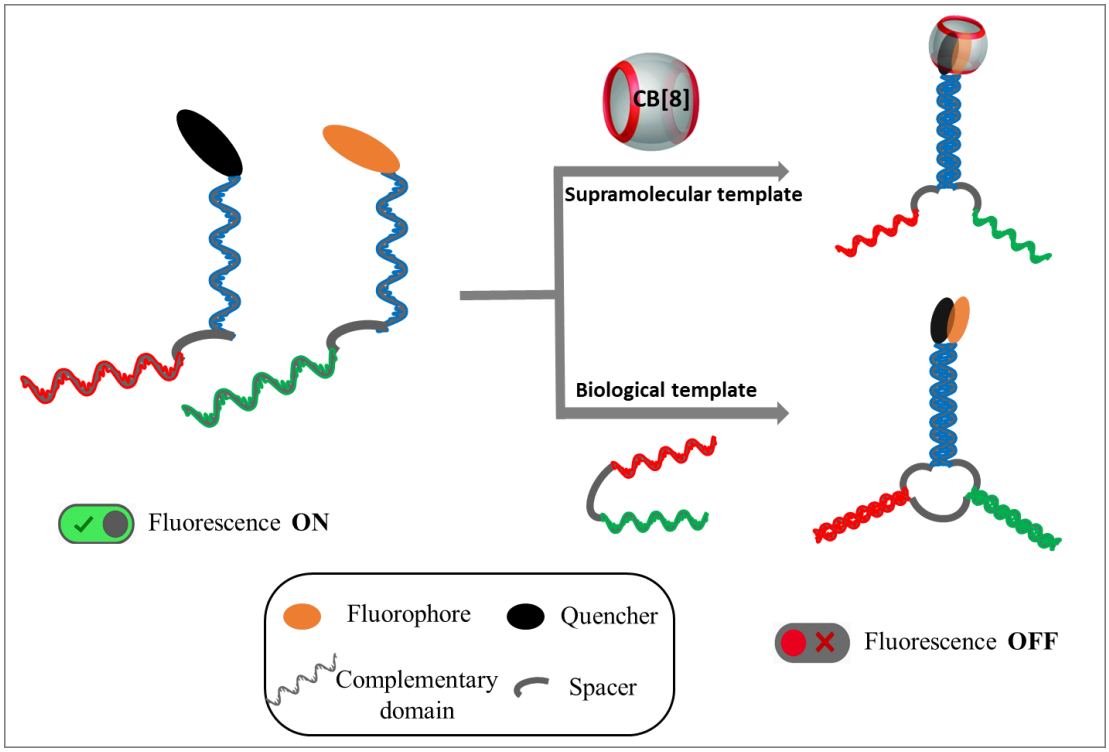


Figure 9. System design. Development of a supramolecular fluorescent probe alternative to a biological system

The use of CB[8] receptor as template represents an orthogonal alternative to biological systems and can provide the confinement input through the formation of a ternary complex bringing in proximity the two oligonucleotide strands. The detailed designed system is represented in Figure 10. The oligonucleotide strands employed are commercially available and are designed to be complementary, but with a sequence of bases such that, under low concentration conditions, duplex formation is possible only if mediated by a templating agent, which in this case can be either the CB[8] receptor or an antibody-mimic (Ab-mimic) DNA strand. The properly functionalization of the strands respectively with a fluorophore and a quencher units allow the dual possibility to exploit the recognition properties of CB[8] in water forming heteroternary complex and to follow duplex formation by fluorescence measurements. Thus, the entering of fluorophore and quencher units in the CB[8] cavity results in the confinement of the two strands forming the duplex and allows an optical response quenching the fluorescence in water solution.

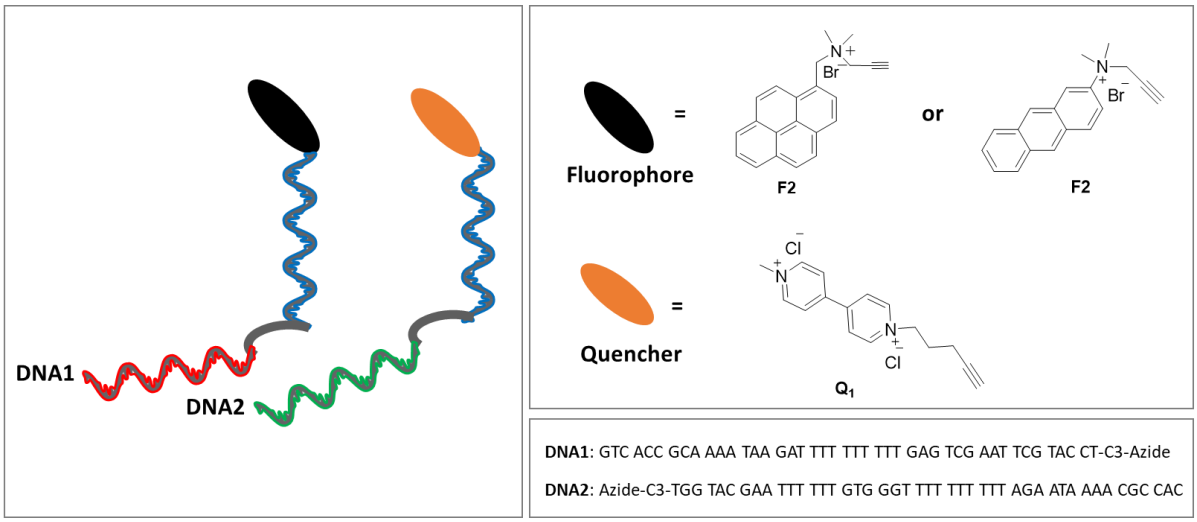


Figure 10. Detailed representation of the systems design

To conjugate the *ad hoc* designed fluorophores and quencher units with oligonucleotides a CuAAC (Copper-Catalysed Azide-Alkyne Cycloaddiction) click reaction with azide-functionalized commercial strands was selected (Figure 11). The choice of this reaction is driven by its effectiveness in the selective derivatization of biomolecules in the presence of wide variety of functionalities. In fact, "Click chemistry" represents an example of a series of chemical reactions that are orthogonal to functional groups found in complex biological systems allowing to the formation of a specific product. Moreover, azide and alkyne groups are stable in the presence of nucleophiles or electrophiles, and common to standard reaction conditions and show an intrinsic reactivity together with a quite selectivity and stability in water.[28] Moreover, the azide present a small size that allow its introduction in a biomolecular target with the minor structural perturbation.[29]

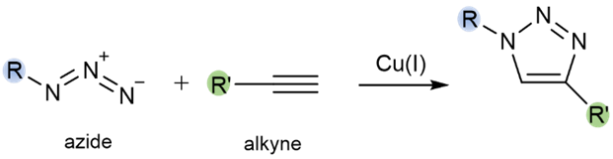


Figure 11. CuAAC (Copper-catalysed azide-alkyne cycloaddiction) click reaction.

Thus, in this chapter the possibility to derivatize through biorthogonal click reaction oligonucleotides with guests molecules suitable to be complexed by CB[8] was explored.

4.2 Results and Discussion

4.2.1 Synthesis of fluorophore and quencher guests

As synthetic targets we selected a fluorophore, namely pyrene derivative **F1**, and a viologen derivative **Q1** as quencher. Both the guests were functionalised with a short chain presenting a terminal alkyne unit necessary to be reacted in the subsequent CuAAC click reaction with azide-terminal oligonucleotide strands (Figure 12).[30]

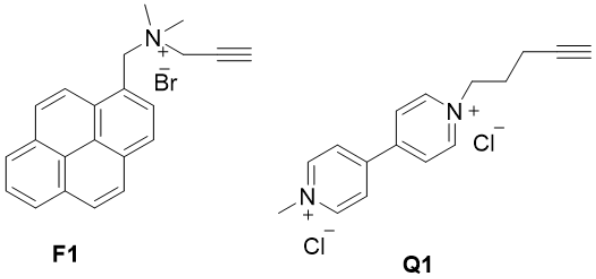
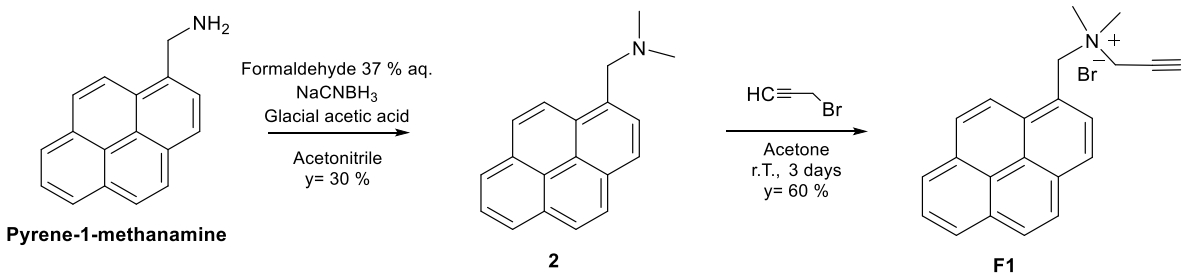


Figure 12. Fluorophore and quencher structures.

Synthesis of fluorophore F1

The following step involved the synthesis of the fluorophore **F1** (Scheme 1), obtained in two synthetic steps, adapting a procedure reported in literature.[31]



Scheme 1. Synthesis of fluorophore **F1**

The first step involved the reductive double amination of pyrene-1-methanamine, obtained by treating pyrene-1-methanamine hydrochloride with a NaOH solution, to obtain product **2** as a tertiary amine. The mechanism initially involves the acid catalysed formation of an imine intermediate between the amine function of pyrene and carbonyl group of formaldehyde in the presence of glacial acetic acid. Finally, imine was converted to methylated amine by reduction with sodium cyanoborohydride. The reaction was conducted in acetonitrile as solvent, for two hours at room temperature. Compound **2** was obtained after purification by flash chromatographic column with 30% of yield and characterized by ^{1}H NMR and ESI-MS.

The following step involved the quaternization of the amine by reacting compound **2** with propargyl bromide, in acetone at room temperature for three days, yielding product **F1** in 60% of yield. This reaction allowed both the formation of a positively charged quaternary amine, which is useful for improving molecule solubility in aqueous solvents and **F1** interaction with CB[8] portals, and the introduction of the terminal alkyne necessary for the subsequent functionalization of the fluorophore by click reaction with oligonucleotides. The molecule was characterized by ^{1}H NMR and ESI-MS and UV-Vis and fluorescence spectra were collected.

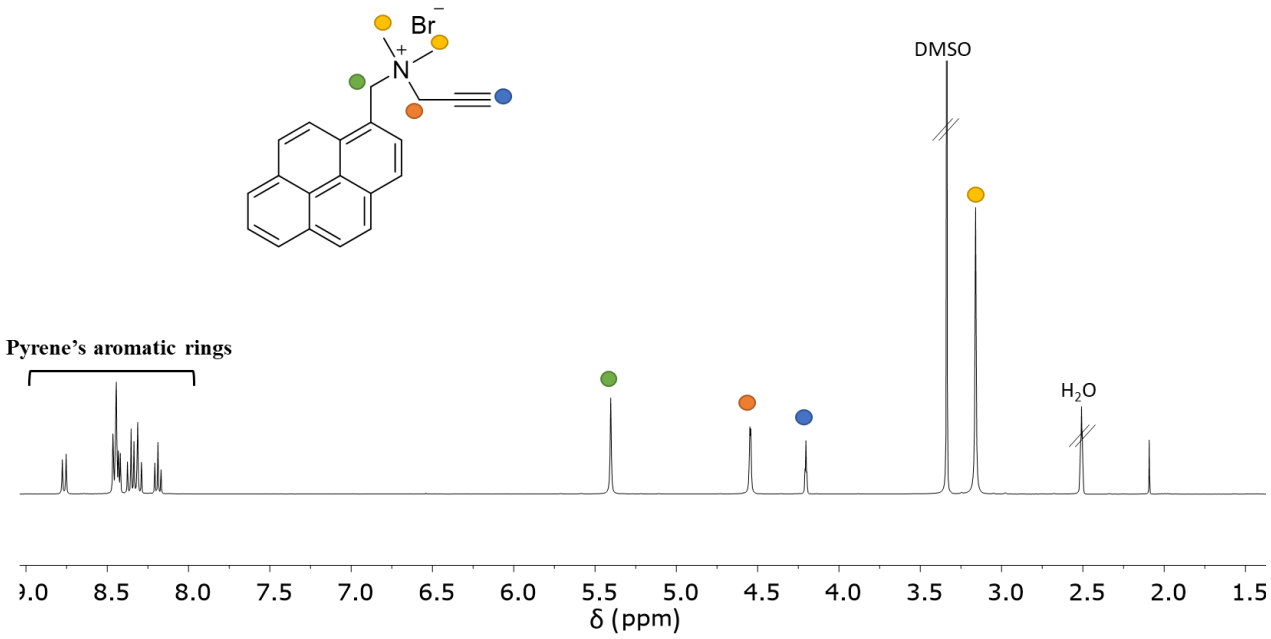
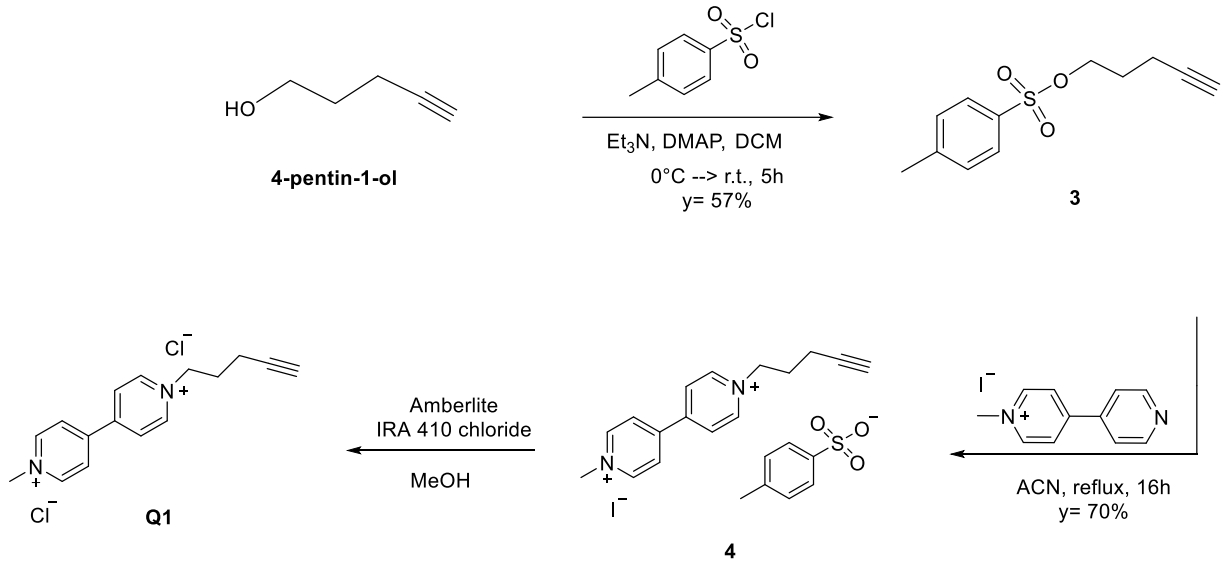


Figure 13. ^{1}H NMR spectrum of **F1** (DMSO-d_6, 600 MHz, 25°C).

In the ^{1}H NMR spectrum (Figure 13) the singlets at 4.15 and 4.48 ppm (blue and orange dots, respectively) are ascribable to the CH_2 in α position of the amine and the CH of propargyl group, respectively. In the aromatic zone, between 8.10 and 8.80 ppm, signals related to the aromatic rings of pyrene are present. In the aliphatic zone, at 3.11 ppm (yellow dot) the signal for amine methyl groups is observed.

Synthesis of quencher Q1

The synthesis of quencher **Q1** is reported in Scheme 2.



Scheme 2. Synthesis of **Q1**

The first synthetic step was the tosylation of 4-pentin-1-ol with tosyl chloride to introduce a good leaving group for the subsequent nucleophilic substitution reaction. The reaction was conducted in dichloromethane for 5 hours at room temperature, in the presence of triethylamine as base and 4-dimethylaminopyridine (DMAP) as scavenger. Product **3** was obtained in 57% yield and characterized by ^{1}H NMR and ESI-MS.

The following step was the nucleophilic substitution of the tosyl group of **3** with 1-methyl bipyridinium. The reaction, conducted in acetonitrile at 85 °C for 24 hours, allowed the formation of **4** in 70 % of yield.

Finally, the counterions of **4** were exchanged with two chlorides using Amberlite IRA 410 chloride exchange resin to improve the solubility in aqueous medium. **Q1** was purified by trituration in acetonitrile resulting in the isolation of a brown solid with 32% yield, which was characterized by fluorescence, ^{1}H NMR and ESI-MS.

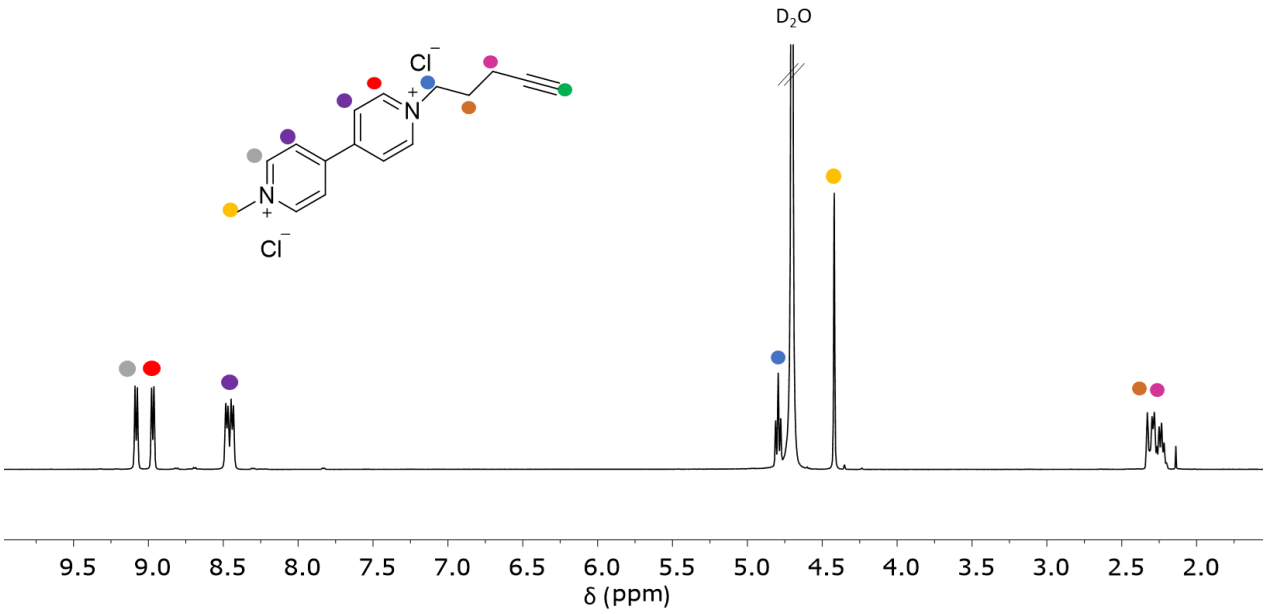


Figure 14. ^{1}H NMR of **Q1** (D_2O, 400 MHz, 25°C).

In the ^{1}H NMR spectrum shown in Figure 14, the peaks of the bipyridyl protons are visible in the aromatic zone between 8.35 and 9.28 ppm (purple, red and grey dots). In the region comprised between 2.31 and 2.38 ppm the alkyl chain signals are present (pink and orange dots), while the singlet related of the methyl group is present at 4.42 ppm (yellow dot). However, the signal of the alkyne proton (green dot) is not clearly discernible, being probably covered by the alkyl signals between 2.1 and 2.4 ppm, as it was confirmed by the proton integration resulting in 5 protons. This singlet, instead, is clearly visible at 2.85 ppm in NMR spectrum of **Q1** collected using deuterated dimethyl sulphoxide (DMSO-d_6) as solvent (Figure 15).

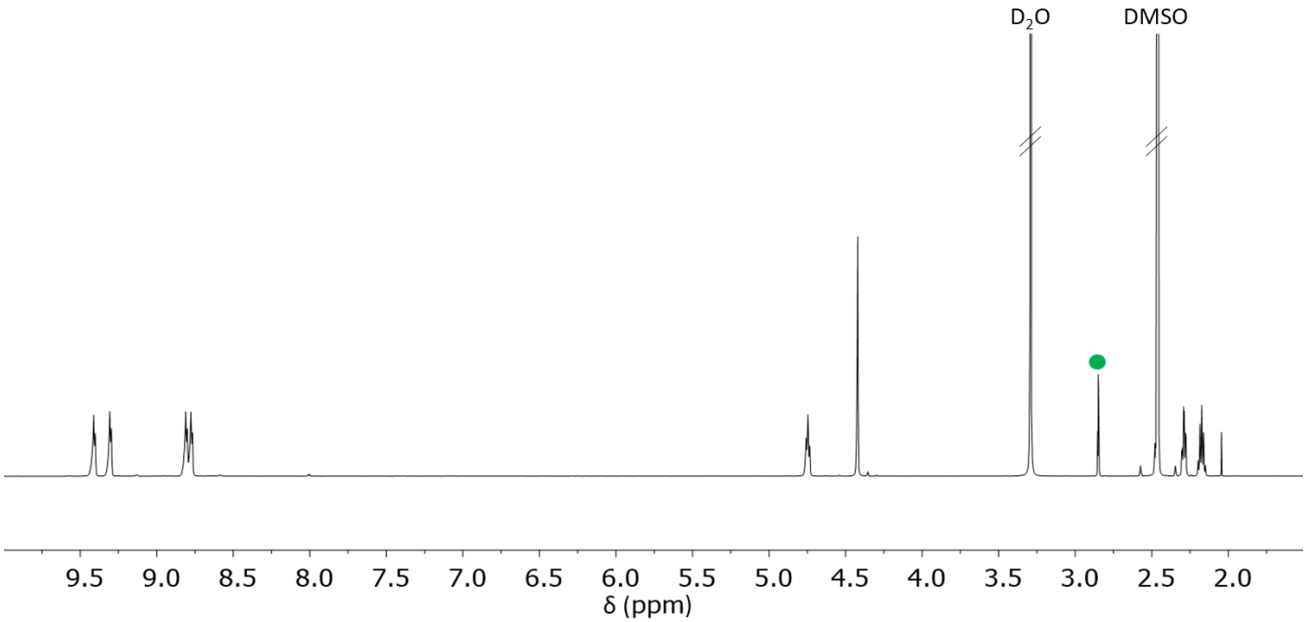
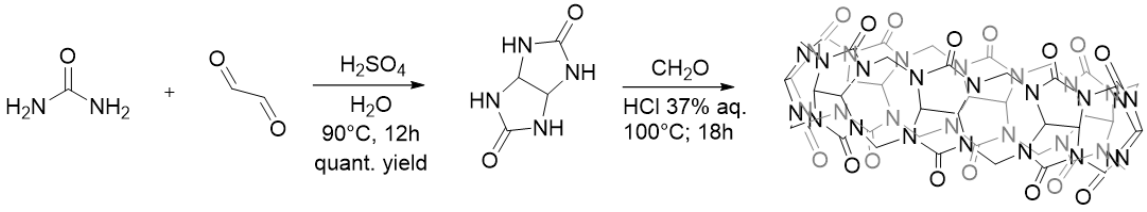


Figure 15. ^{1}H NMR of **Q1** (DMSO-d_6, 600 MHz, 25°C).

4.2.2 Synthesis of CB[8]

The CB[8] preparation is reported in the Scheme 3 above.



Scheme 3. CB[8] synthetic pathway

The first step involved the synthesis of the glycoluril monomer by acyl nucleophilic substitution of urea on glyoxal in acid environment. The reaction was conducted in water at 90°C for 12 hours in the presence of sulphuric acid. After 12 hours the reaction mixture was basified to pH 14, allowing the formation of a white solid suspension, which was filtered off to obtain **1** with quantitative yield. The product was characterized by ^{1}H NMR and ESI-MS.

In the next step CB[8] was synthesized following a literature procedure optimised by the research group.[2] The macrocyclization was obtained by reacting **1** with paraformaldehyde at 100°C in strongly acidic environment, using HCl (37% aq.) as solvent. Under these conditions, all isomers of CB[n] with n from 5 to 8 are present in the reaction mixture and can be separated by exploiting their different solubility in water. After cooling at room temperature CB[6] and CB[8] precipitated as white solids and were separated by filtration and then resuspended in a 1:1 solution of water and formic acid. By heating to 90°C only CB[6] passed in solution and CB[8] were isolated by filtration. The product was characterized by ^{1}H NMR.

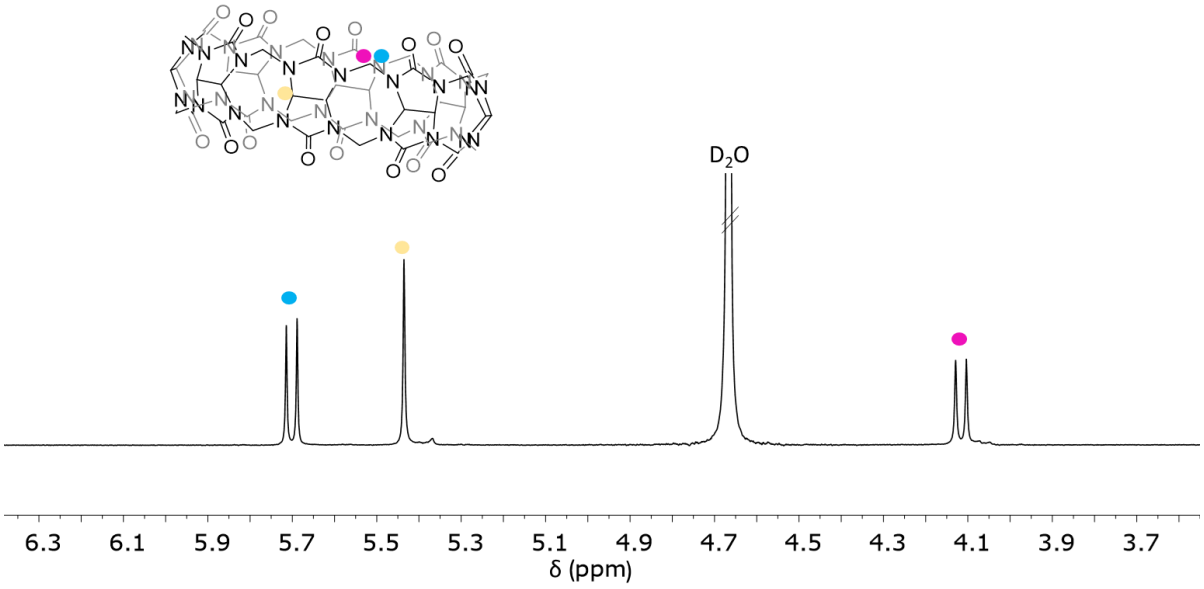


Figure 16. ^{1}H NMR spectrum of **CB[8]** (D_2O-DCl, 600 MHz, 25°C).

In ^{1}H NMR spectrum reported in Figure 16 the three characteristic CB[8] signals are visible. The two doublets at 4.38 and 5.81 ppm (pink and blue dots, respectively) are relate to the protons of the methylene bridge, which are split as they are not magnetically equivalent. The doublet highlighted in blue refers to the protons facing inwards the cavity, at lower fields than those facing outwards (pink dot) since they are affected by the electronic density of the carbonyl group. Finally, the singlet at 5.68 ppm (yellow dot) is related to the protons of the glycoluril units.

4.2.3. Supramolecular studies with F1 and Q1

In order to understand the complexation behaviour of the CB[8]/**F1**/**Q1** (1:1:1) system, a ^{1}H NMR titration was carried out.

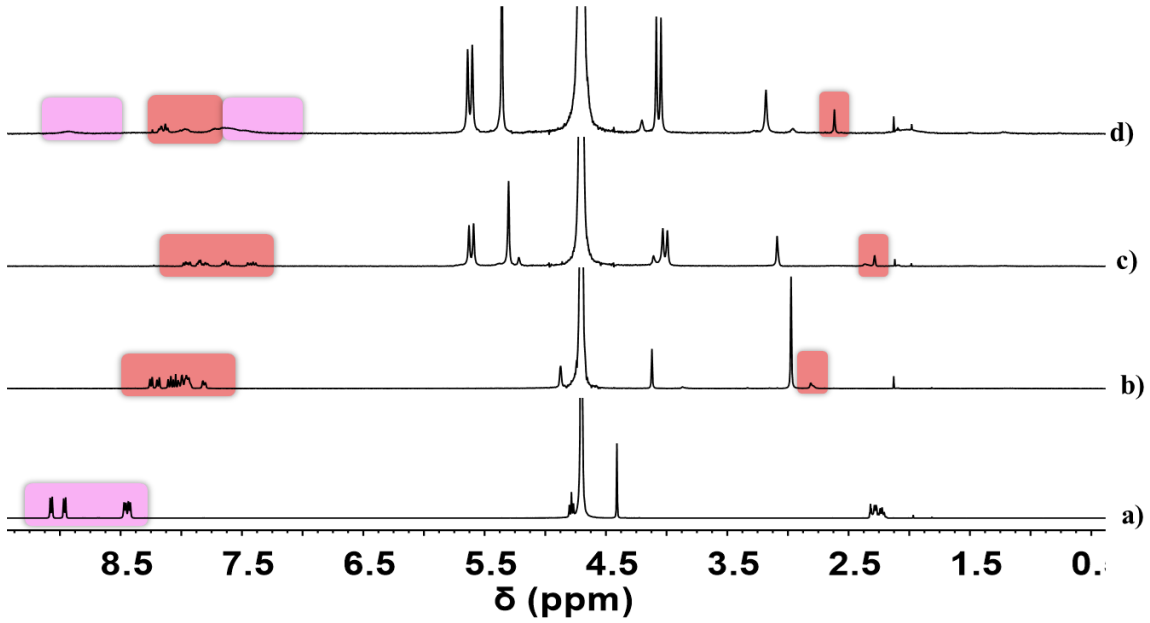


Figure 17. ^{1}H NMR spectra of **a)** Q1, **b)** F1, **c)** CB[8]+F1 1:1 and **d)** CB[8]+F1+Q1 1:1:1, in D_2O, 400 MHz, 25°C.

In Figure 17 the ^{1}H NMR spectrum resulting from an equimolar addition of CB[8] to a **F1** solution in deuterium oxide is reported and compared to the one of free **F1** (the shifted signals are highlighted with red circles). After the addition of CB[8], an overall shift of the aromatic signals of **F1** to lower fields was observed, as indication of the effective guest complexation in the cucurbituril cavity. The alkyne proton signal also presents a upfield shift of about 0.4 ppm. This can be ascribed to the folding of the alkyne chain of **F1**, which is then complexed in the CB[8] cavity not only with the aromatic part of the molecule but also with its aliphatic part (Figure 18).

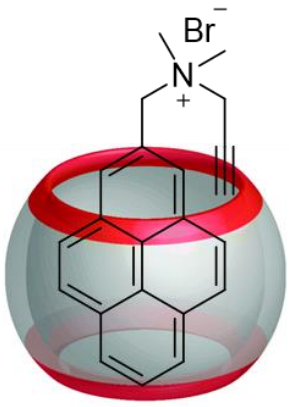


Figure 18. Folding of **F1** chain in the CB[8] cavity.

To the equimolar solution of CB[8] and **F1** was then added one equivalent of **Q1**. In the aromatic region of ^{1}H NMR spectrum (Figure 18, d), the shifts to the aromatic protons of **Q1** indicate that this guest is complexed by the cavity. However, at the same time the signals of **F1** protons tend to return to their initial position. The broadening of most signals is indicative of a fluxional situation, with the establishment of competition equilibria between **F1** and **Q1** for occupation of the CB[8] cavity. In contrast to the system reported in the literature,[18] the formation of the heteroternary complex could therefore not be observed unambiguously, probably due to the presence of the propargyl group, which folds and strengthens the binary complex between **F1** and CB[8].

These considerations are also supported by the results obtained with fluorescence titration (Figure 19). The initial addition of about 1 equivalent of **Q1** to a 1:1 solution of CB[8]/**F1** resulted in an increase of fluorescence intensity, as indication of the release of **F1** from the receptor cavity. After the subsequent addition of tens equivalents of **Q1** a partial decrease of fluorescence, probably caused by dynamic quenching, was observed.

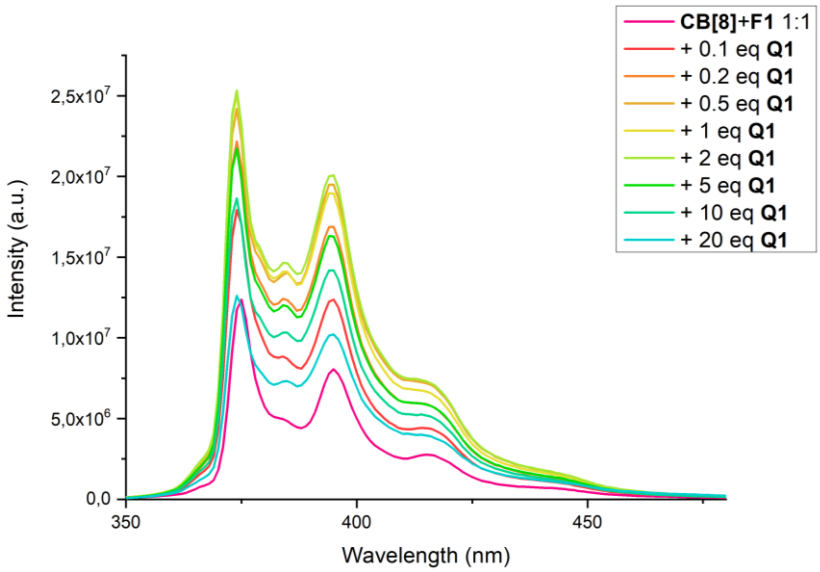


Figure 19. Fluorescence spectra (λ_{ecc}: 340 nm) of CB[8]+**F1** 1:1 titration with **Q1**.

However, in the target DNA-based system the terminal alkyne group of **F1** is no longer present as it has been reacted *via* click reaction with a DNA strand. In this case, the folding should no longer be possible and the formation of the heteroternary complex could became favoured. To investigate this hypothesis before proceeding with the synthetic demanding oligonucleotide functionalization, a model system based on the derivatization of **F1** and **Q1** with a TEG-N_3 chain was synthesized and complexation tests were carried out to verify formation of the heteroternary complex and the subsequent quenching of F1 fluorescence emission (Figure 20).

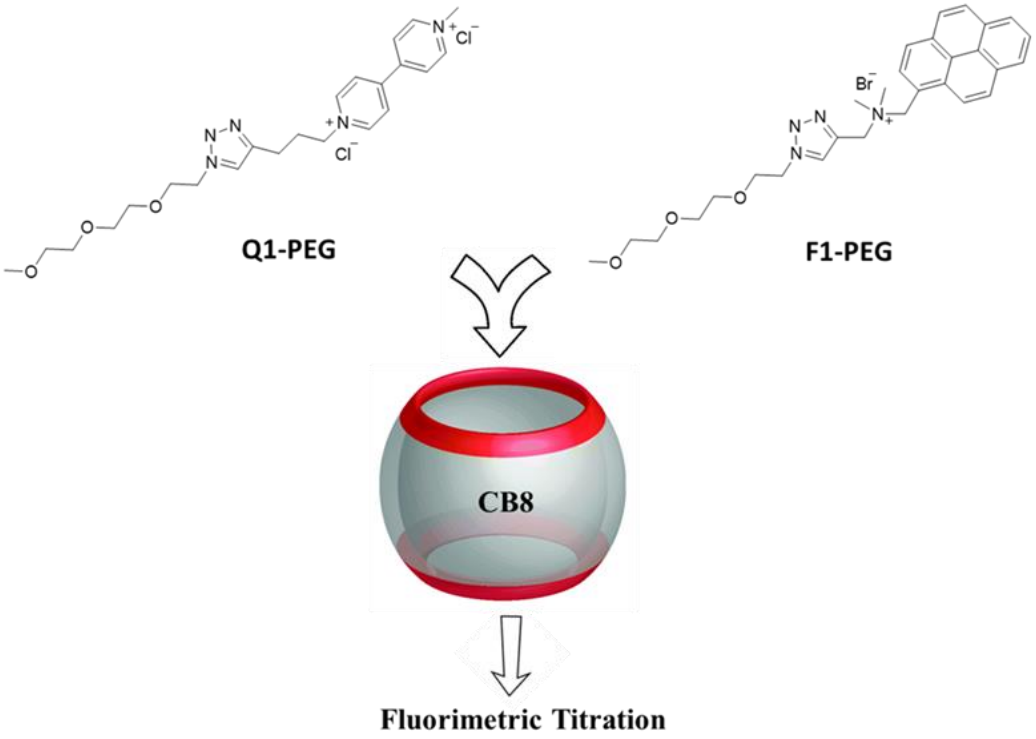
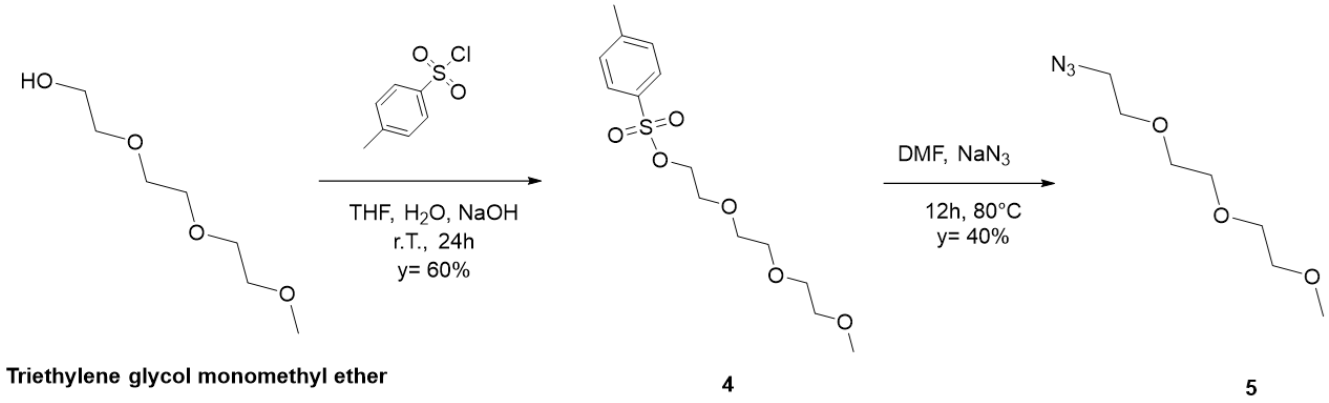


Figure 20. Representation of the second model system using TEG-functionalised **F1** and **Q1**.

4.2.4 Synthesis of TEG-N_3 chain

A triethylene glycol (TEG) chain with a terminal azide group was chosen for the derivatization of **F1** and **Q1** via click chemistry. The synthesis was conducted in two steps following a procedure reported in literature (Scheme 4).[32]



Scheme 4. Synthesis of **5**.

Initially, the terminal OH group of triethylene glycol monomethyl ether was replaced with a tosyl group. The reaction was conducted in a mixture of THF and H_2O, with NaOH as base, at room temperature for 24 hours. Product **4** was obtained in 60% yield and characterized by ^{1}H NMR (Figure 21a) and ESI-MS. Subsequently, the tosyl group was substituted by an azide *via* reaction of **4** with NaN_3 under anhydrous conditions in DMF at 80 °C for 12 hours. Product **5** was obtained in 40 % yield and characterized by ^{1}H NMR and ESI-MS.

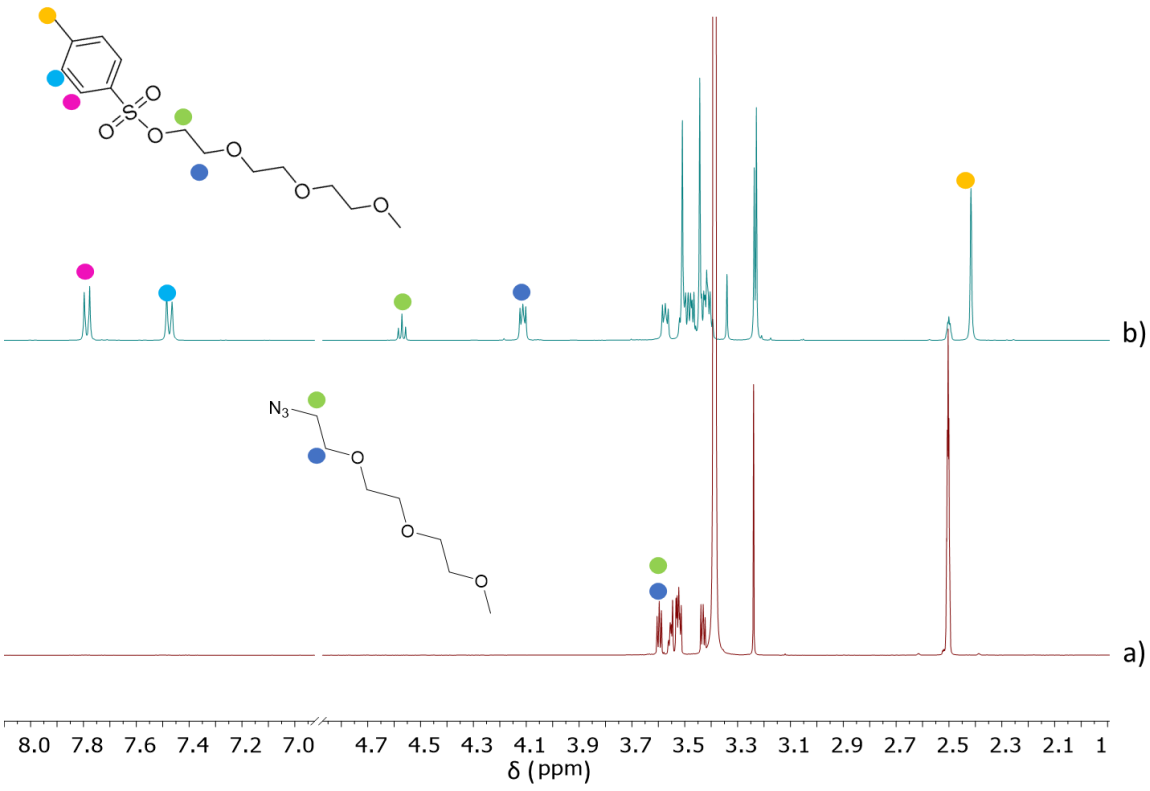


Figure 21. ^{1}H NMR spectrum of **4** a) and **5** b), in DMSO-d_6, 400MHz, 25 °C.

In the ^{1}H NMR spectrum (Figure 21b) peaks related to the tosyl group (blue, pink and yellow dots) are no longer present, as indication of the complete conversion of the reagent **4**. Moreover, the signals related to methylene groups in α and β positions of tosyl group (highlighted in green and blue dots respectively) can be observed at higher fields (from 4.57 ppm to 4.12 ppm).

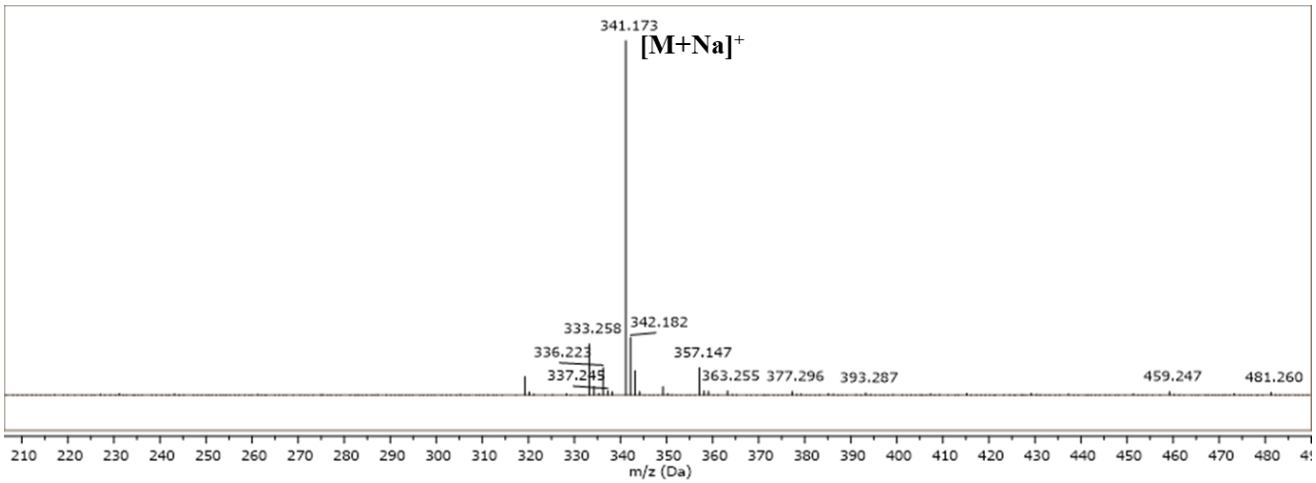
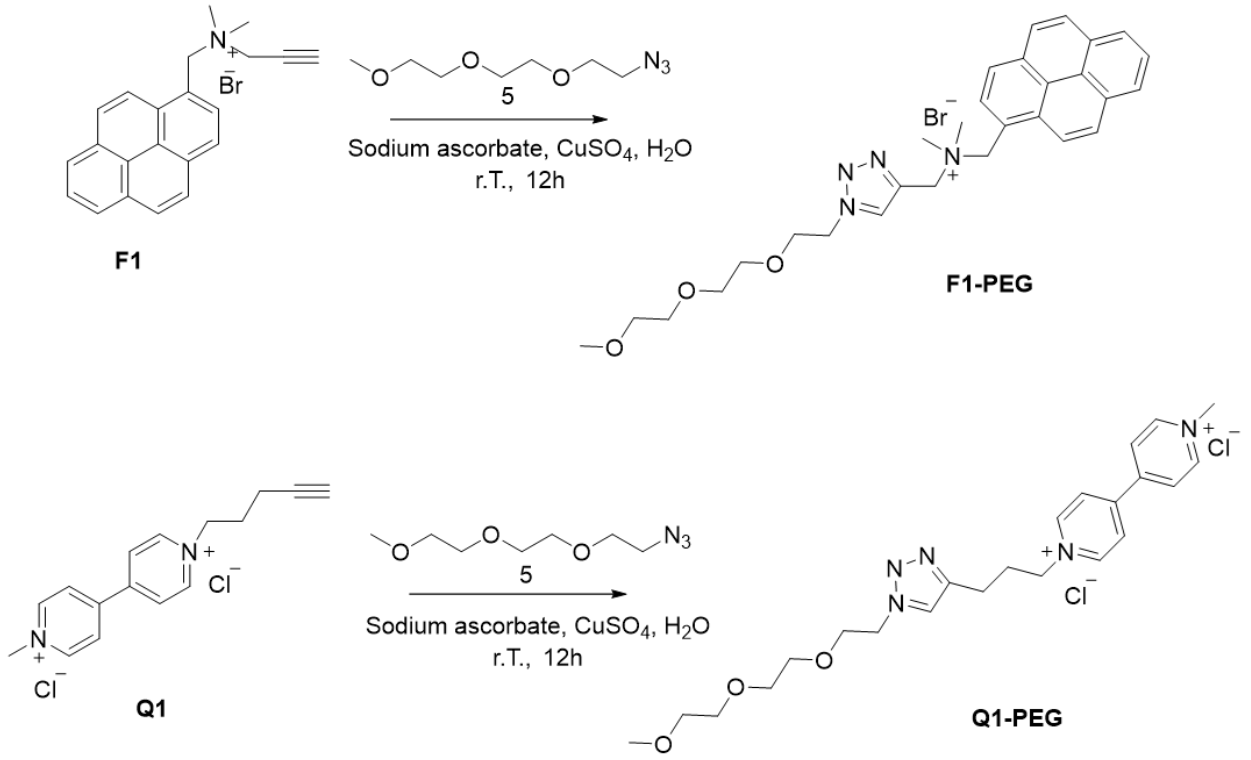


Figure 22. ESI-MS spectrum of **5**

The outcome of the reaction was also confirmed by ESI-MS analysis that shows the presence of the sodium adduct (Figure 22).

4.2.5 Click reaction of F1 and Q1 with PEG-N_3 chain

In the next step click reactions between alkyne groups of **F1** and **Q1** with the azide group of **5** were performed under the same reaction conditions (Scheme 7).



Scheme 7. Click reaction of **F1** and **Q1** with **5**

The click protocol followed was based on the use of copper (II) salt as source of the catalytic species in presence of sodium ascorbate as reducing agent to induce the *in situ* formation of Cu(I) catalyst. The reaction proceeded for 12 hours at room temperature.

F1-TEG was obtained, after recrystallization from a 1:1 mixture of dichloromethane/diethyl ether with a post-purification yield of 11% and fully characterized by ^{1}H NMR and ESI-MS.

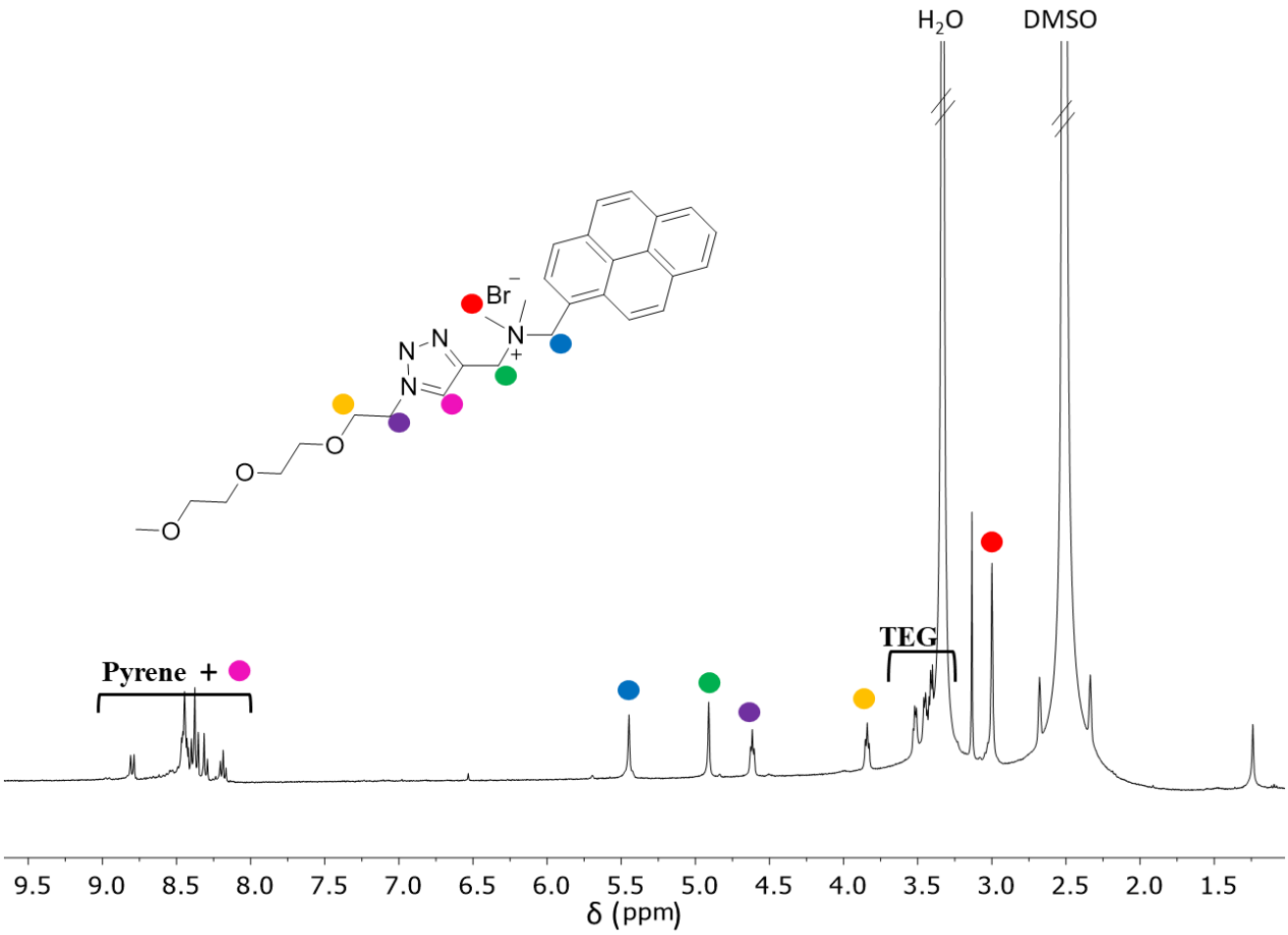


Figure 23. ^{1}H NMR spectrum of **F1-TEG** (DMSO-d_6, 400 MHz, 25 °C).

In the ^{1}H NMR spectrum reported in Figure 23 the diagnostic peak ascribable to the proton of the triazole ring is overlapped with the aromatic protons of pyrene, as confirmed by the integration of the multiplet from 3.56 to 3.33 ppm. Another diagnostic signal is the triplet present at 4.62 ppm (purple dot), relative to the CH_2 of the TEG chain in a to the triazole ring.

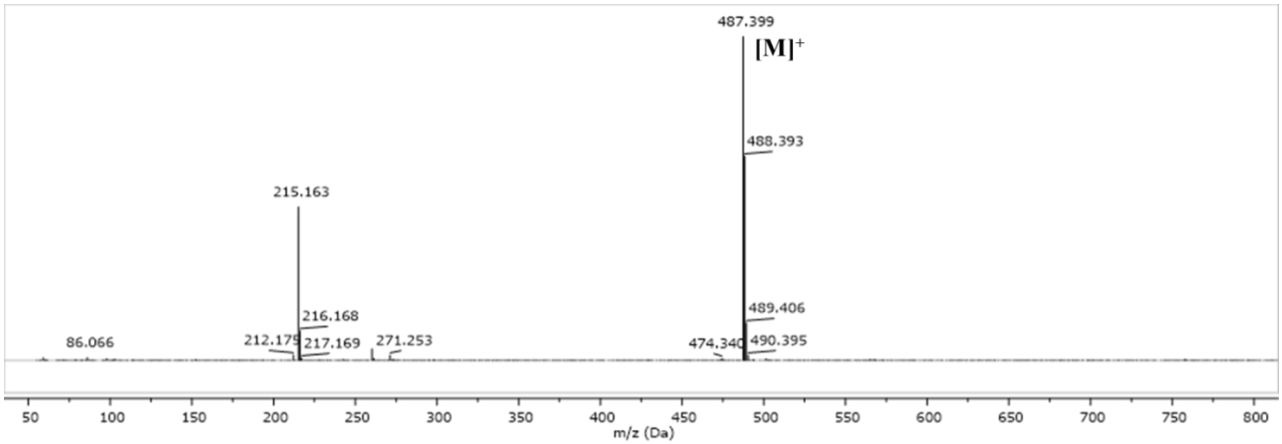


Figure 24. ESI-MS spectrum of **F1-TEG**.

The isolation of product **F1-TEG** was also confirmed by the presence of the molecular ion signal (487.4 m/z) in the ESI-MS spectrum (Figure 24).

As regard **Q1-TEG**, the purification, in this case, was carried out by several triturations of the crude with methanol. The product was obtained with a post-purification yield of 16%. Characterization of **Q1-TEG** was performed by ^{1}H NMR and ESI-MS.

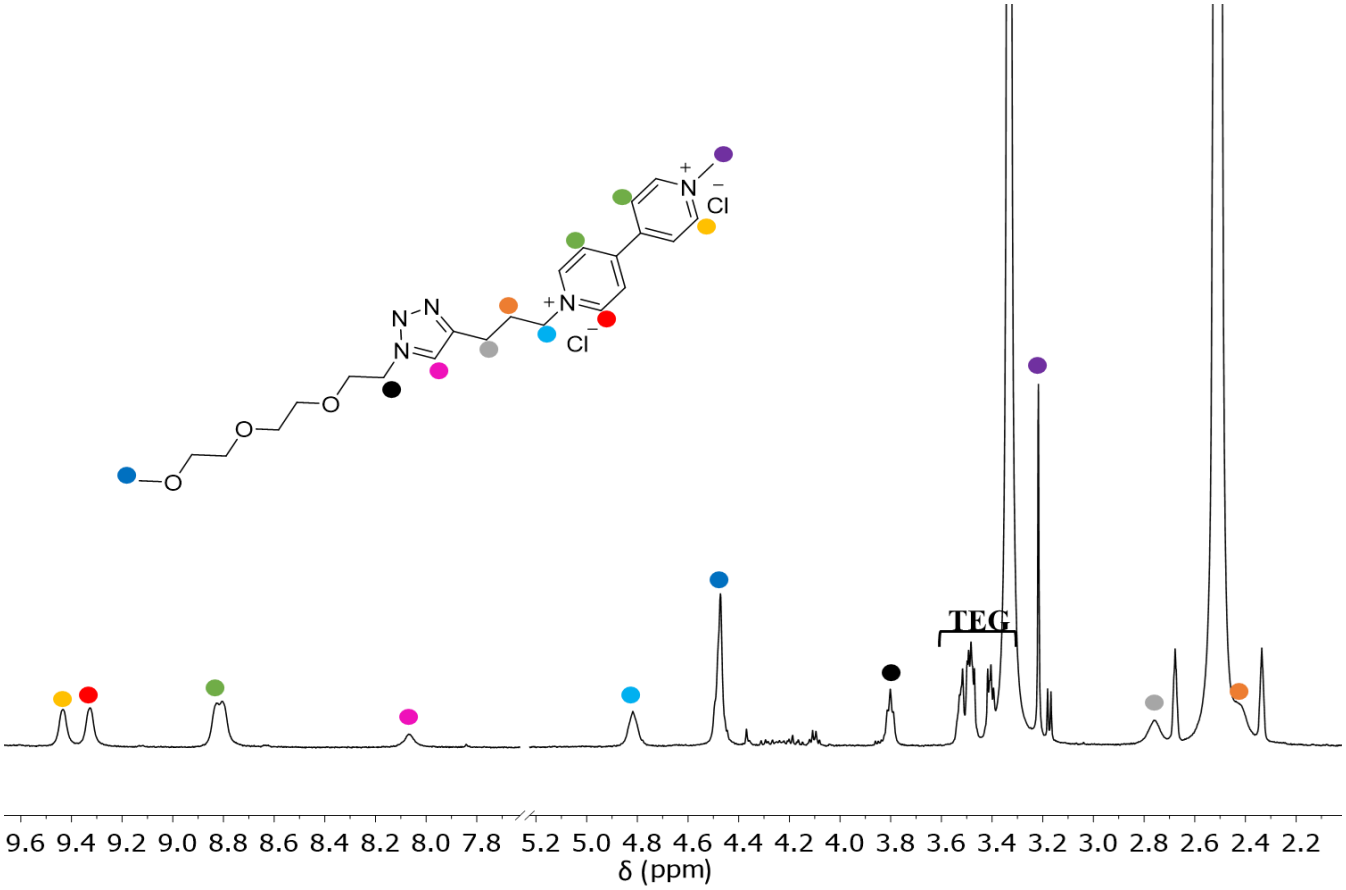


Figure 25. ^{1}H NMR spectrum of **Q1-TEG** (DMSO-d_6, 400 MHz, 25 °C).

In the ^{1}H NMR spectrum reported in Figure 25 the diagnostic peak of the triazole proton can be observed at 8.17 ppm (pink dot). After the click reaction the signal of the -CH_2- in a to the azide is now at 3.81 pm, as confirmation of the formation of the triazole ring.

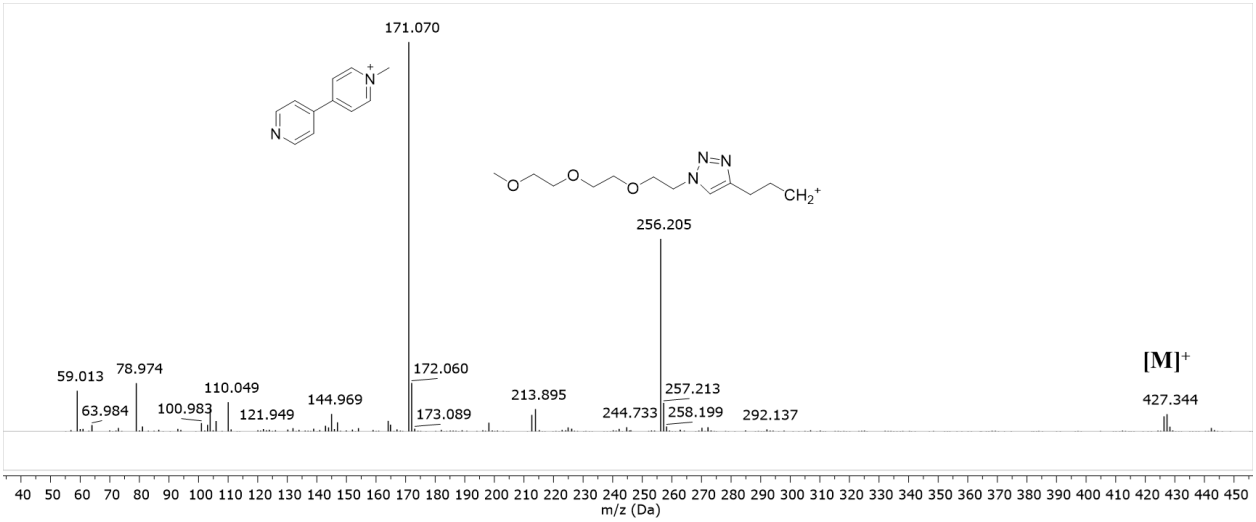


Figure 26. ESI-MS spectrum of **Q1-TEG**.

Moreover, the peak of the molecular ion is present at 427.3 m/z in the mass spectrum (Figure 26), although at a relatively low intensity. This can be explained by the tendency of viologen species to undergo fragmentation, as indicated by the two peaks at 171.1 and 256.2 m/z.

4.2.6 Fluorescence tests

The formation of a ternary complex between CB[8], TEG-functionalised **F1** and **Q1** units was directly tested by fluorescence analysis since the low solubility of **F1-TEG** element precluded an ^{1}H NMR titration.

Fluorescence analysis was performed starting from an equimolar solution of **CB[8]** and **F1-PEG** in MilliQ water and then titrated with aliquots of a solution of **Q1-PEG** in the same solvent. The recorded fluorescence spectra do not show a quenching, but after an initial decrease in fluorescence intensity with the addition of 0.5 equivalents of **Q1-PEG** a small drop of the emission intensity was observed (Figure 27). Overall, the intensity of the emission is not significantly altered by the additions of the quencher.

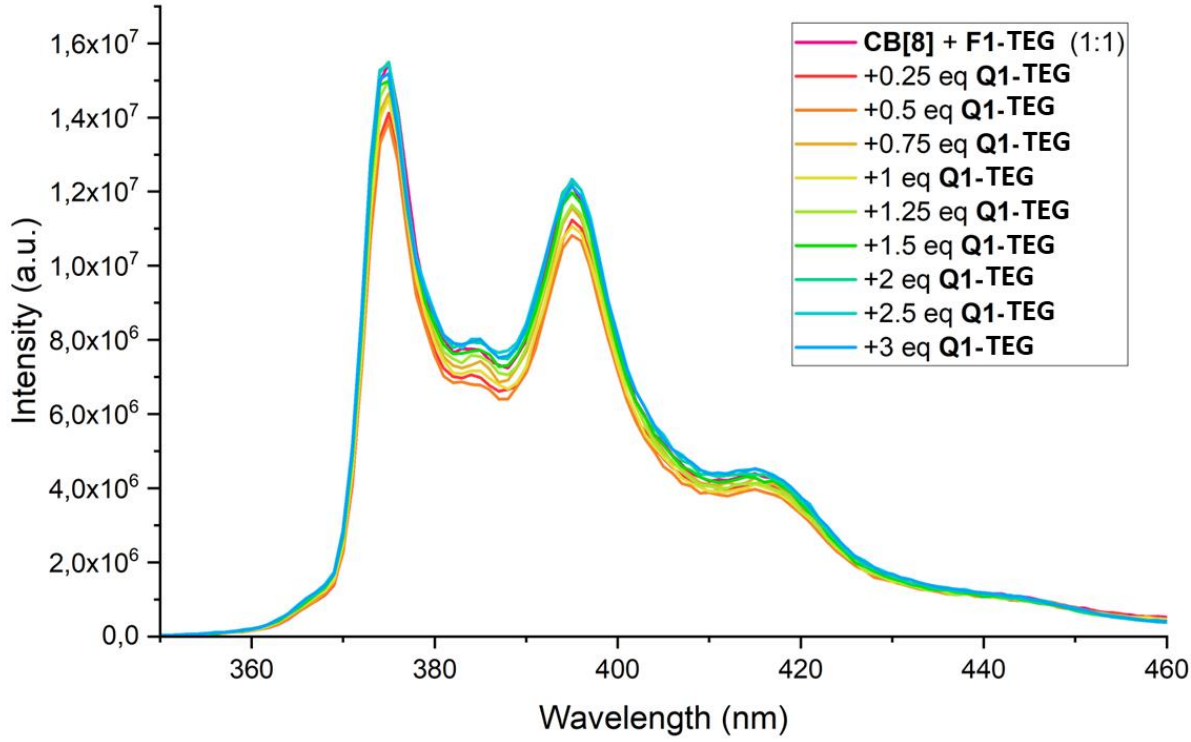


Figure 27. Fluorescence titration (λ_{ecc} 340 nm) of **CB[8]/F1-TEG** (1:1) with **Q1-TEG**

4.2.7 Supramolecular studies with F2 and Q1

Following these results, it was possible to conclude that bulkiness of the pyrene molecule does not allow the engulfment of another guest in the cavity of CB[8]. Therefore, it was decided to move to the less hindered anthracene-derivative fluorophore **F2**, even if it is well known in literature its tendency to form homoternary complexes in solution with a K_a value of 1.0-0.5 10^{12} M^{-1} (Figure 28).[16,33]

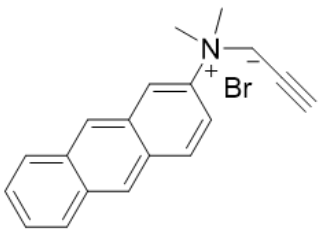


Figure 28. Fluorophore **F2**

Firstly, the complexation event was studied by ^{1}H NMR titration for the CB[8]/**F2** /**Q1** (1:1:1) system (Figure 29).

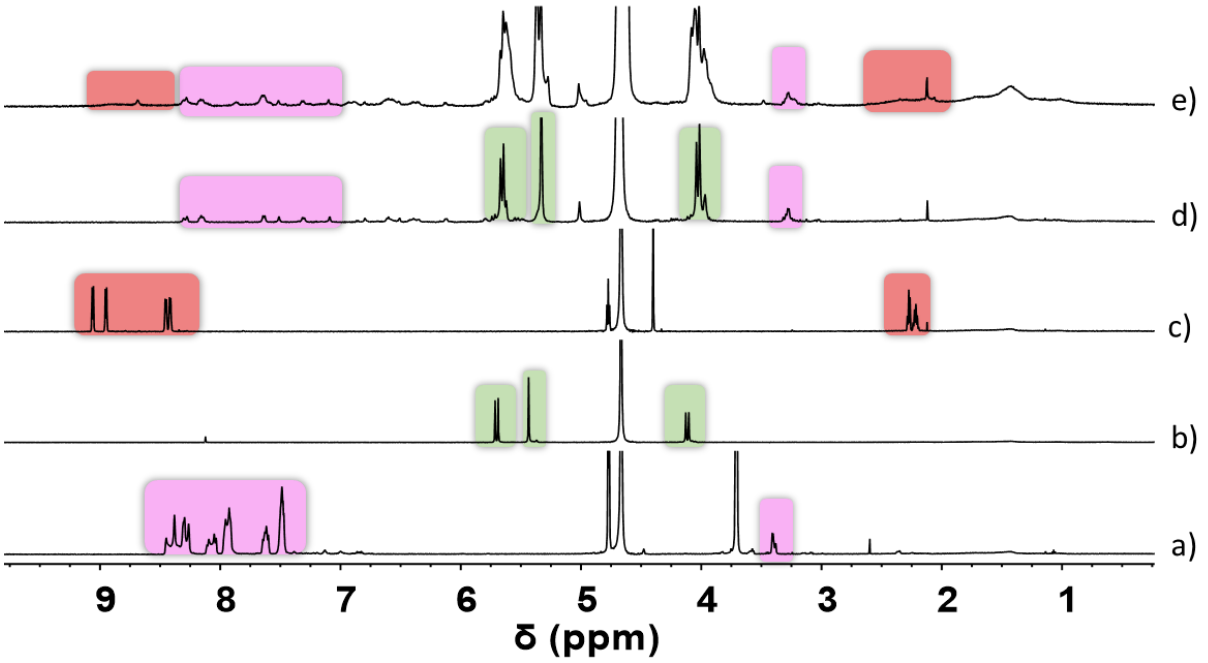


Figure 29. ^{1}H NMR spectra of **a)** F2, **b)** CB[8], **c)** Q1 and **d)** CB[8] + F2 1:1 and **d)** CB[8] + F2 + Q1 1:1:1 in D_2O, 600MHz, 25°C.

The ^{1}H NMR spectrum resulting from an equimolar addition of CB[8] to a **F2** solution in deuterium oxide is reported (Figure 29,d) and compared to the one of free **F2** and CB[8] (Figure 29, a and b). Because of the limit solubility of **F2** in water, it was not possible a precise stoichiometric control during the titration, however, after the addition of CB[8], a shift of the receptor signals (green circles) and an overall shift of the aromatic signals of **F2** to lower fields (pink circles) were observed, confirming the guest complexation in the cucurbituril cavity. Then, one equivalent of **Q1** was added to the equimolar solution of CB[8] and **F2**. In the aromatic region of ^{1}H NMR spectrum (Figure 29, e), the shifts and broadening of the aromatic protons of **Q1** indicate that the quencher is successfully complexed by the receptor. At the same time, the signals of **F2** protons do not show any shifts respect to the previous equimolar solution (Figure 29, d), confirming the simultaneous presence of both the guests inside the cucurbituril cavity forming the heteroternary complex, as desired.

Moreover, to verify the formation of the heteroternary complex in the presence of **Q1**, a fluorescence titration of an equimolar solution of CB[8] and **F2** with a solution of **Q1** in water was carried out.

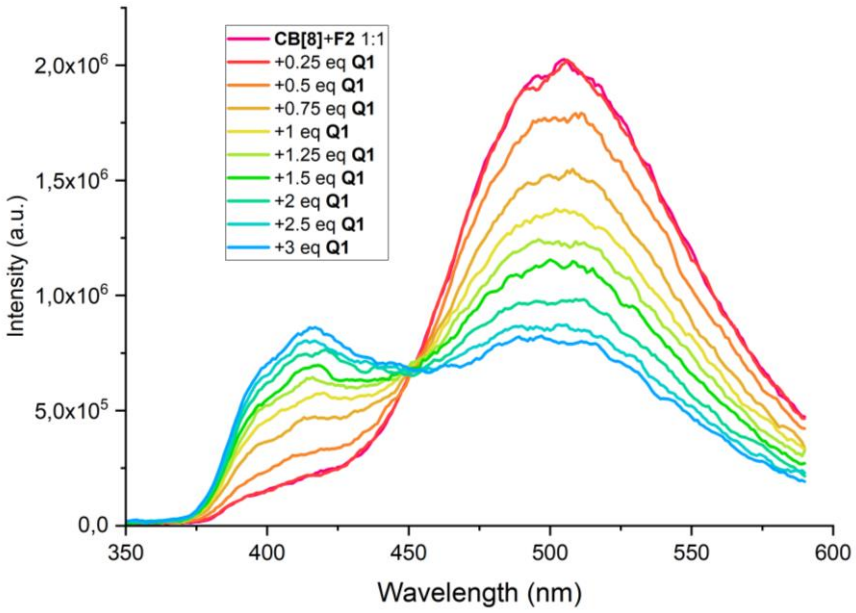


Figure 30. Fluorescence titration (λ_{ecc} 300 nm) of **CB[8]**/**F2** (1:1) with **Q1**.

In the fluorescence spectrum reported in Figure 30, a decrease of the fluorescence band at 505 nm, relative to the excimer, can be observed from the addition of 0.5 equivalents of **Q1**. At the same time, the fluorescence intensity of the 420 nm band increase progressively with the increasing of **Q1** equivalents. This behaviour indicated that the addition of quencher **Q1** to an equimolar solution of **CB[8]**:**F1** caused that, the equilibrium reported in Figure 31 was shifted towards the formation of the heteroternary complex. **Q1** enters the cavity of CB[8], dislocated one of the two **F2** molecules forming the heteroternary complex. This explains both the lowering of the band at 505 nm, due to the breaking of the excimer,[31] and the increase of the band at 420 nm, related to the formation of the ternary complex.

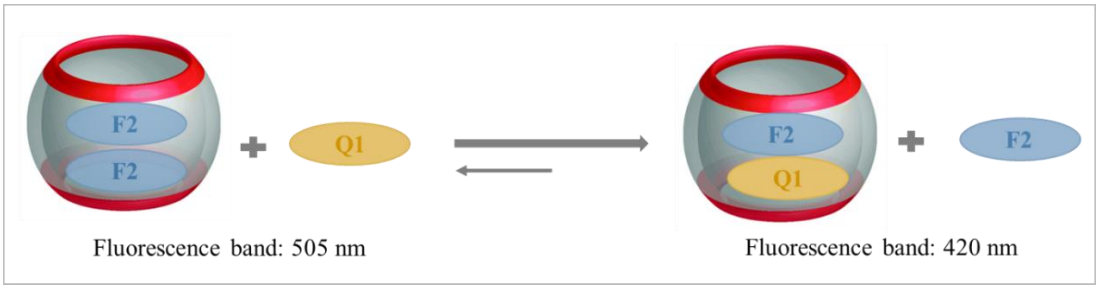
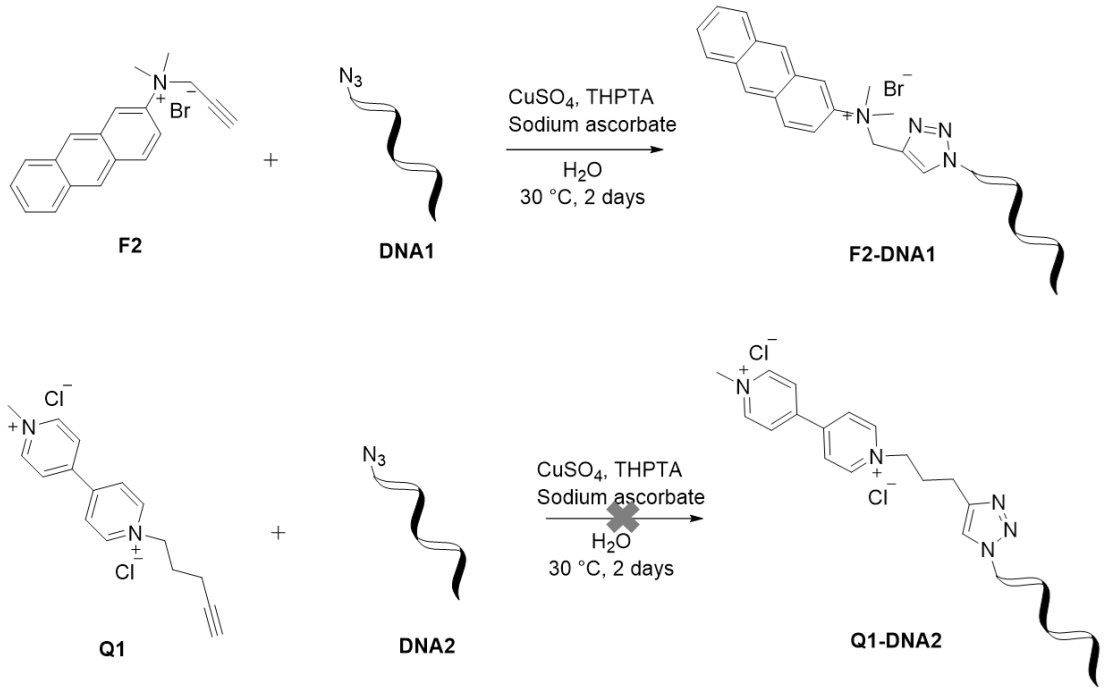


Figure 31. Representation of the equilibrium state, after **Q1** addition, in solution

This optical signal can be exploited for the desired sensing, and it was therefore decided to proceed with the functionalization of **F2** and **Q1** with the oligonucleotide strands necessary for the creation of the final recognition system.

4.2.8 Click reaction tests with oligonucleotide strands

Click reactions performed between oligonucleotides functionalized with a terminal azide group at the 3' or 5' end and **F1**/**Q1** are summarized in Scheme 8. The commercially available DNA strands were chosen taking inspiration from the recent publication of F. Ricci and co-workers,[34] and resulting to be complementary but with a sequence of bases such that, under low concentration conditions, duplex formation is possible only if mediated by a templating agent, which in this case can be either the CB[8] receptor.



Scheme 8. Click reaction of **F2** and **Q1** with oligonucleotide strands azide- terminal (**DNA1** and **DNA2**)

For both reactions involving **F2** and **Q1**, the same procedure was followed. The click protocol was based on the use of copper (II) salt as source of the catalytic species in presence of sodium ascorbate as reducing agent to induce the *in situ* formation of Cu(I) catalyst. THPTA, a polytriazole-based ligand often used in click reactions involving biological molecules, was added to stabilize the Cu(I) catalyst formed during the reaction, preventing oxidation or disproportionation reactions.[35,36] The reactions proceeded at 30 °C for 2 days, after which the solvent was evaporated, and the product was characterized using ESI-Orbitrap mass analysis.

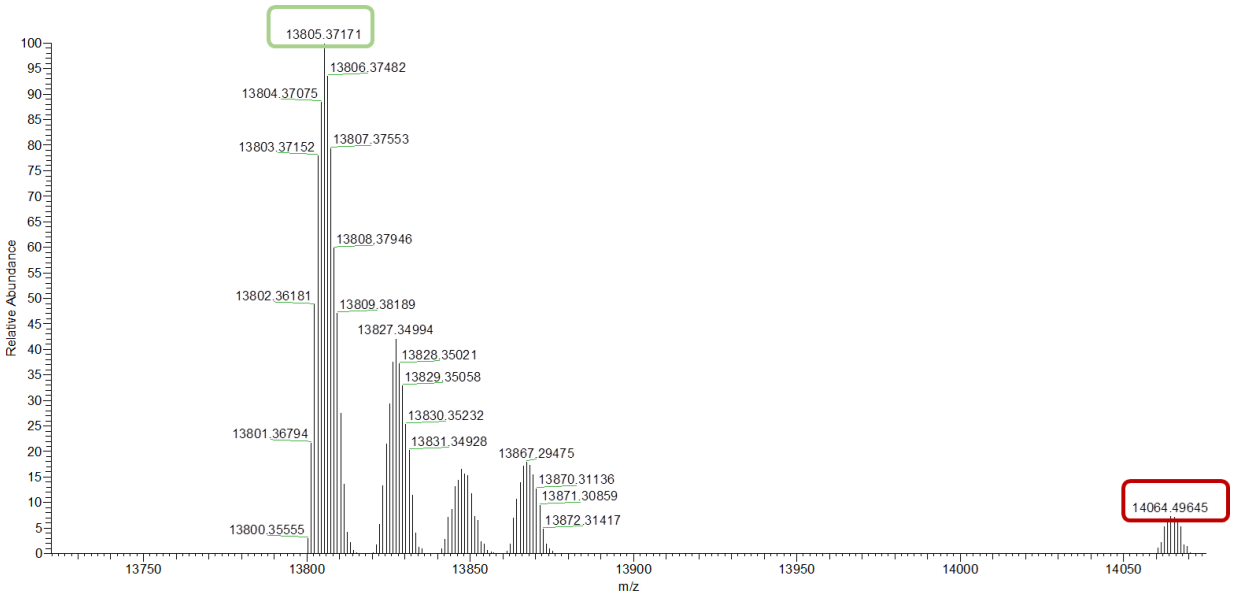


Figure 32. ESI -MS Orbitrap spectrum of **F2-DNA1**.

In the mass spectrum of the crude mixture of the first reaction (Figure 32), the peak of the molecular ion of **F2-DNA1** can be observed at 14064.49 m/z. However, the peak related to the unreacted DNA strand at 13805.37 m/z (highlighted in green) is also present with a intensity higher of one order of magnitude.

The ESI-MS Orbitrap spectrum of crude mixture of the second reaction, instead, the presence of signals related to **Q1-DNA2** was not detected and only the peak related to the react oligonucleotide was visible (Figure 33).

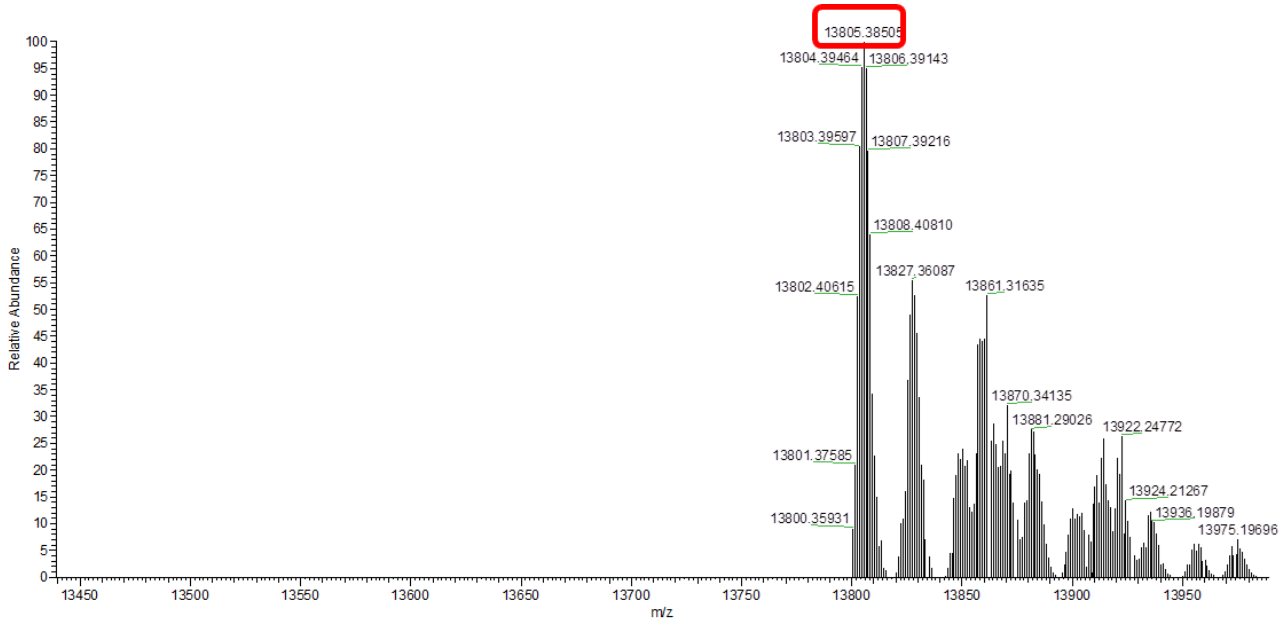


Figure 33. ESI -MS Orbitrap spectrum of Q1-DNA1. In red the peak related to the oligonucleotide reactant.

4.3 Conclusions

The anthracene-derivative fluorophore **F2** proved to be the best guest for the formation of the heteroternary complex. Fluorescence analysis demonstrated that the formation of **F2/Q1/CB[8]** complex results in a variation in the fluorescence profile of the system. Preliminary trials of click reaction carried out with commercial azide-functionalized oligonucleotide strands allowed the formation of the **F2-DNA** product. However, a future optimization of the reaction conditions is necessary to increase the final yield. Regarding **Q1**, the formation of the desired product was not observed. The law reactivity under CuAAC click conditions of methyl-viologen salts and derivatives and the dependence of the yield on the counterions was identified as the probable cause. In fact, in literature the difficulty to make methyl-viologen react by click reaction is well known.[37] This issue can be partially solved by exchanging the chloride counterion with PF_6^-. Respect to other ions, PF_6^- is not able to bridge two copper metal centres at the same time, a factor that significantly increases the conversion and reaction speed when compared to other ions such as chloride or iodide. However, the employment of PF_6^- would lower the solubility in water and adversely affect the outcome of the reaction.

As alternative approach, we can employ commercial and complementary oligonucleotide strands functionalized with N-terminal phenylalanine-glycine-glycine (FGG) peptide motif and CB[8] to control a dye displacement. In literature is known that homotetramerization of proteins can be induced exploiting the recognition abilities of CB[8] in binding simultaneously two FGG tripeptides in the cavity, with high association constant.[38,39,40] In this redesigned system, CB[8] can induced a proximity effect between the two complementary strands due to the homoternary complex formation hosting in its cavity two FGG-oligonucleotide strands and therefore increasing the local concentration allowing the dye displacement (Figure 34).

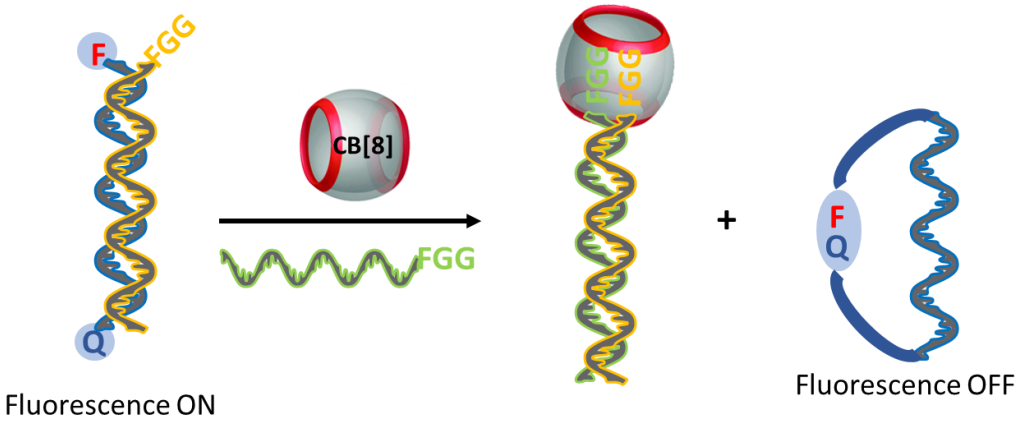


Figure 34. Future perspectives: supramolecular input to control a dye displacement.

4.4 Experimental section

4.4.1 Reagent, Chemicals and Instruments used

All reagents and solvent were purchased from certified commercial sources and used as received, without further purification. For the synthesis, all solvents were dried and distilled according to standard procedures known in the literature.[41]

Thin layer chromatography was performed using TLC Analytical Chromatography F254®, Merck.

Oligonucleotide strands were purchase from Metabion International AG (Germany) in lyophilized form, subsequently dissolved in MIlliQ water at a concentration of 100 μM and stored at -20 °C. The sequences used (5'-3') are shown in Table 2.

Table 2. Oligonucleotide sequences (5'-3'). In bold are highlighted the complementary portions.

Name	Sequence
DNA1	GTC ACC GCA AAA TAA GAT TTT TTT TTT GAG TCG **AAT TCG TAC C**T-C3-Azide
DNA2	Azide-C3-T**GG TAC GAA TT**T TTT GTG GGT TTT TTT TTT AGA ATA AAA CGC CAC

^{1}H and ^{31}P NMR spectra were recorded on Bruker 400 MHz AVANCE or Jeol 600 MHz using DMSO-d_6, D_2O, DCl and $CDCl_3$ as solvents. All chemical shifts (δ) were reported in parts for million (ppm) relative to proton resonances resulting from incomplete deuteration of NMR solvents. The abbreviations: s, d, t, m and sb indicated the spectrum peaks referred to: singlet, doublet, triplet, multiplet and broad singlet, respectively. The coupling constant (J) are expressed in Hz.

ESI-MS characterization experiments were performed on a Waters ACQUITY Ultra Performance LC HO6UPS-823M with ESI source ionization (electrospray ionization) in positive modality. High-resolution analyses were performed with a Thermo LTQ-Orbitrap.

UV-Vis absorption spectra were collected with Perkin Elmer Lambda 750 UV-VIS-NIR instrument. Fluorescence spectra were recorded using Fluoromax-3 instrument.

Synthesis of glycoluril (1)

In a two-necked flask, urea (30.01 g, 0.5 mol) and glyoxal (30 mL, 0.2 mol) were dissolved in 100 mL of water. The solution was left under magnetic stirring until it became homogeneous. Then, concentrated sulphuric acid (2.5 mL) was added drop by drop. The reaction was left under magnetic stirring, at 90 °C. After 12 hours, a 6.25 M of NaOH solution was added until a pH of 14 was reached. A suspension was formed which was filtered off and dried under vacuum to obtain glycoluril **1** as a beige solid (27.40 g), in quantitative yield. **^{1}H NMR** (400 MHz, DMSO-d_6, 25 °C): δ ppm = 7.17 (s, 4H, N***H***), 5.24 (s, 2H, C***H*****$_2$**).

Synthesis of CB[8]

Glycoluril **1** (3 g, 21.1 mmol) was dissolved in a 37 % aqueous HCl solution (4.6 mL) and paraformaldehyde (1.2 g, 21.1 mmol) was added under vigorous magnetic stirring. The solution was heated to 100 °C and kept reacting for 2 days under magnetic stirring. A withe precipitate was observed, which was filtered off and purified by suspending it in a 1:1 solution of water/formic acid (1 mL). The suspension was heated to 80 °C to dissolve the product and left at -4 °C for 3 days. The precipitate obtained was filtered off and dried under vacuum to obtain the desired product **CB[8]**. **^{1}H NMR** (600 MHz, D_2O-DCl, 25 °C): δ ppm = 5.17 (d, *J* = 15.4 Hz, 16H, ***H*****$_{in}$**), 5.09 (s, 16H, C***H***), 3.83 (d, *J* = 15.5 Hz, 16H, ***H*****$_{out}$**).

Synthesis of N,N-dimethyl-1-pyren-1-yl)methanamine (2)

Pyren-1-ylmethanamine (512 mg, 2.21 mmol) was dissolved in acetonitrile (40 mL) and 37% w/w formaldehyde (1 mL, 22.1 mmol) was added. The solution was brought to 0 °C and sodium cyanoborohydride (500 mg, 6.63 mmol) was slowly added. Glacial acetic acid (0.265 mL) was added drop by drop and the solution was restored to room temperature. After 2.5 hours, dichloromethane (50 mL) was added, and the organic phase was extracted with an aqueous solution of 1 M KOH, brine and water. The organic phase was dried with sodium sulphate and the solvent evaporated under reduced pressure. The crude was purified by flash chromatographic column (DCM/MeOH 95:5). Product **2** was obtained with 27% of yield and appeared as a yellow oil (152 mg). **^{1}H NMR** (400 MHz, $CDCl_3$, 25 °C): δ ppm = 8.49 (d, *J* = 9.3 Hz, 1H), 8.27 – 8.16 (m, 4H, Pyrene), 8.14- 8.00 (m, 4H, Pyrene), 4.25 (s, 2H, C***H*****$_2$**), 2.46 (s, 6H, C***H*****$_3$**). **ESI-MS:** m/z 260.18 $[M+H]^+$; 215.12 $[M-N(CH_3)_2]^+$.

Synthesis of N,N-dimethyl-N-(pyrene-1-ylmethyl)prop-2-in-1-ammonium bromide (F1)

Compound **2** (152 mg, 0.59 mmol) was dissolved in acetone (25 mL) and propargyl bromide (0.252 mL, 2.93 mmol) was added. The reaction was left under magnetic stirring at room temperature for 4 days. The precipitate formed was filtered off and washed with diethyl ether and dried under vacuum,

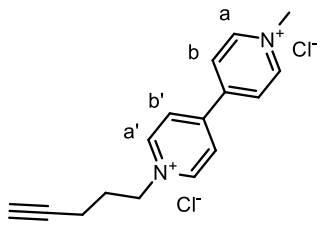
yielding product **F1** as a white solid (100 mg), in 57% of yield. **^{1}H NMR** (600 MHz, DMSO-d_6, 25 °C): δ ppm = 8.71 (d, *J* = 9.8 Hz, 1H), 8.43 – 8.36 (m, 4H), 8.34 - 8.24 (m, 3H), 8.14 (t, *J* = 7.6 Hz, 1H), 5.35 (s, 2H, ArC***H*$_2$**), 4.48 (d, *J* = 2.8 Hz, 2H, C***H*$_2$**CCH), 4.15 (t, J= 2.4 Hz, 1H, C***H***), 3.11 (s, 6H, C***H*$_3$**). **ESI-MS:** m/z 298.50 $[M]^+$; 214.46 $[M\text{-}C_5H_9N]^+$.

Synthesis of 4-penthyn-1-tosyl (3)

Under dry conditions, 4-penthyn-1-ol (0.110 mL, 1.19 mmol) was dissolved in 6 mL of dichloromethane and DMAP (14 mg, 0.12 mmol) and triethylamine (0.215 mL, 1.54 mmol) were added in small portions. The reaction was left under magnetic stirring at room temperature for 2 days. Then, the mixture was extracted with water and brine and the organic phase was dried with sodium sulphate and the solvent evaporated under reduced pressure. The resulting solid was dried under vacuum to obtain **3** (282 mg) in quantitative yield as a dark yellow oil. **^{1}H NMR** (400 MHz, DMSO-d_6, 25 °C): δ ppm = 7.86 – 7.75 (m, 2H, C***H***CHCCH$_3$), 7.55 – 7.43 (m, 2H), 4.08 (t, *J* = 6.2 Hz, 2H), 2.77 (t, *J* = 2.6 Hz, 1H), 2.43 (s, 3H), 2.18 (td, *J* = 7.1, 2.7 Hz, 2H), 1.74 (p, *J* = 6.6 Hz, 2H).

Synthesis of Q1

Compound **3** (282 mg, 1.18 mmol) and 1-methyl-4-bipyridinium iodide (352 mg, 1.18 mmol) were dissolved in acetonitrile (5 mL). The reaction was left under magnetic stirring at 85 °C for 16 hours The solvent was evaporated under reduced pressure. The solid obtained was dissolved in methanol and the counterion was exchange by passing the solution through a column packed with anion exchange resin (Amberlite IRA 410 chloride). The crude was triturated in acetonitrile and filtered off allowing product **Q1** (115 mg) in 32% of yield.



^{1}H NMR (400 MHz, D_2O, 25 °C): δ ppm = 9.41 (d, *J* = 6.7 Hz, 2H, **H$_a$**), 9.30 (d, *J* = 6.0 Hz, 2H, **H$_{a'}$**), 8.79 (dd, *J* = 20.6, 6.5 Hz, 4H,**H$_b$**+**H$_{b'}$**), 4.75 (t, *J* = 7.1 Hz, 2H, N$^+$C**H$_2$**), 4.42 (s, 3H, C**H$_3$**), 2.85 (t, *J* = 2.7 Hz, 1H, C**H**), 2.29 (td, *J* = 7.1, 2.6 Hz, 2H, C**H$_2$**CCH), 2.17 (q, *J* = 7.1 Hz, 2H, C**H$_2$**CH$_2$CCH). **ESI-MS:** m/z 237.18 $[M\text{-}H]^+$.

Synthesis of 4

Triethylene glycol monomethyl ether (2.43 mL, 15.22 mmol) and NaOH (3.28 g, 82.21 mmol) were dissolved in water (33 mL) and placed at 0 °C using an ice bath. A solution of *p*-toluenesulphonyl chloride (3.48 g, 18.27 mmol) in THF (10 mL) was slowly added *via* a dripping funnel. After 24 hours, the reaction mixture was extracted with dichloromethane. The organic phase was evaporated under reduced pressure, resulting in product **4** as a pale-yellow oil (3.10 g) in 65 % of yield. **^{1}H NMR** (400 MHz, DMSO-d_6, 25 °C): δ ppm = 7.80 (d, *J* = 8.3 Hz, 2H, C***H***CS), 7.49 (d, *J* = 8.1 Hz, 2H,

C***H***CCH_3), 4.15 – 4.10 (m, 2H, C***H₂***OS), 3.62 – 3.44 (m, 12H, C***H₂***), 3.24 (s, 3H, OC***H₃***), 2.43 (s, 3H, C***H₃***Ar). **ESI-MS:** m/z 341.18 $[M+Na]^+$.

Synthesis of 5

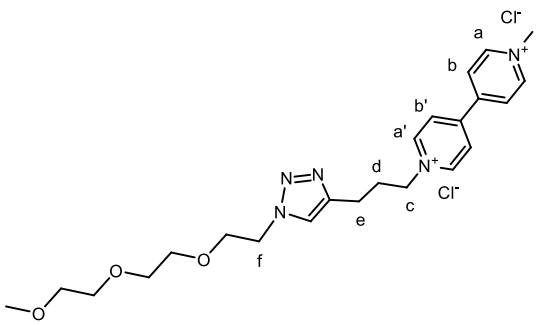
In dry conditions, **4** (1.2 g, 3.77 mmol) and NaN_3 (720 mg, 11.07 mmol) were dissolved in DMF (30 mL). The solution was left reacting at 80 °C under magnetic stirring. After 12 hours, the reaction was restored at room temperature, ethyl acetate was added, and a liquid-liquid extraction was performed with an aqueous solution of first HCl 1 M and then brine. The organic phase was dried, and the solvent evaporated under reduced pressure to afford product **5** as a yellow oil, in 40 % of yield. **^{1}H NMR** (600 MHz, $CDCl_3$, 25 °C): δ ppm = 3.68 – 3.63 (m, 8H), 3.56 – 3.53 (m, 2H), 3.40 – 3.36 (m, 5H). **ESI-MS:** m/z 212.11 $[M+Na]^+$.

Synthesis of F1-PEG

In a Schlenk, **F1** (55 mg, 0.14 mmol) and **5** (27.5 mg, 0.14 mmol) were dissolved in H_2O (10 mL). Then, in another Schlenk, $CuSO_4$ (54 mg, 0.22 mmol) and sodium ascorbate (43 mg, 0.22 mmol) were dissolved in water (6 mL). Both the solutions were subjected to three cycles of degassing by freeze-pump-thaw. The two solutions were combined in one Schlenk and the resultant was subjected to a further three degassing cycles. The reaction was left under magnetic stirring at room temperature. After 12 hours, the solution was extracted in $CHCl_3/H_2O$ and washed with EDTA, water and brine. The organic phase was dried and the solvent evaporated under reduced pressure, resulting in **F1-PEG** (9 mg) as a yellow ochre solid, in 11 % of yield. **^{1}H NMR** (400 MHz, DMSO-d_6, 25 °C): δ ppm = 8.80 (d, *J* = 9.3 Hz, 1H), 8.49 – 8.34 (m, 7H), 8.30 (d, *J* = 8.9 Hz, 1H), 8.19 (t, *J* = 7.7 Hz, 1H, ***H*** ***triazole***), 5.45 (s, 2H, ArC***H₂***), 4.91 (s, 2H, C***H₂***N^+), 4.62 (t, *J* = 5.0 Hz, 2H, C***H₂***NNN), 3.84 (t, *J* = 5.1 Hz, 2H, OC***H₂***CH_2N), 3.52 –3.38 (m, 8H), 3.00 (s, 6H, C***H₃***), 1.24 (s, 3H, OC***H₃***). **ESI-MS:** m/z 487.07 $[M]^+$.

Synthesis of Q1-PEG

In a first Schlenk, **Q1** (30 mg, 0.097 mmol) and **5** (18 mg, 0.097 mmol) were dissolved in water (5mL). Then, in a second Schlenk tube, $CuSO_4$ (36 mg, 0.145 mmol) and sodium ascorbate (28 mg, 0.145 mmol) were dissolved in water (3 mL). Both solutions were subjected to three cycles of degassing freeze-pump-thaw. The reaction was left under magnetic stirring at room temperature. After 12 hours, the solvent was evaporated. The resulting solid was dissolved in methanol and left overnight at - 4 °C. The precipitate was filtered off, obtaining **Q1-PEG** (8 mg) as a light brown solid, in 16 % of yield.



^{1}H NMR (400 MHz, DMSO-d_6, 25 °C): δ ppm = 9.43 (s, 2H, ***H*$_a$**), 9.33 (s, 2H, ***H*$_{a'}$**), 8.81 (d, *J* = 11.8 Hz, 4H, ***H*$_b$+*H*$_{b'}$**), 8.07 (s, 1H, ***H*$_{triazole}$**), 4.82 (s, 2H, ***H*$_c$**), 4.48 (s, 3H, C***H*$_3$**O), 3.80 (t, *J* = 5.0 Hz, 2H, ***H*$_f$**), 3.57 – 3.38 (m, 10H, C***H*$_2$** PEG), 3.22 (s, 3H, C***H*$_3$**N), 2.75 (br s, 2H, ***H*$_e$**), 2.43 (br s, 2H, ***H*$_d$**).

ESI-MS: m/z 427.34 [M]+; 256.21 [M-$C_{11}H_{11}N_2$]$^+$; 171.07 [M-$C_{12}H_{12}N_3O_3$]$^+$.

Synthesis of F2-DNA1 and Q1-DNA2

In an Eppendorf tube, 40 μL of a 100 μM solution of oligonucleotide in H_2O MilliQ and 50 μL of a 80 μM solution of reagent (**F2** or **Q1**) in H_2O MilliQ were inserted. Then, a solution containing tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), $CuSO_4$ and sodium ascorbate in a 1:1:1 ratio (10 μL, 333 μM) was added. The reaction was left at 30 °C for two days without stirring and the solvent was evaporated.

ESI-Orbitrap-MS for **F2-DNA1.** Theoretical: 14065 m/z; Experimental: 14064.49 m/z [M]$^-$.

4.5 References

1. Barrow, S. J., Kasera, S., Rowland, M. J., Del Barrio, J. & Scherman, O. A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.* **115**, 12320–12406 (2015).

2. Bardelang, D. *et al.* Cucurbit[n]urils (n = 5-8): A comprehensive solid state study. *Cryst. Growth. Des.* **11**, 5598–5614 (2011).

3. Pazos, E., Novo, P., Peinador, C., Kaifer, A. E. & García, M. D. Cucurbit[8]uril (CB[8])-Based Supramolecular Switches. *Angew. Chemie - Int. Ed.* **58**, 403–416 (2019).

4. Biedermann, F., Uzunova, V. D., Scherman, O. A., Nau, W. M. & De Simone, A. Release of high-energy water as an essential driving force for the high-affinity binding of cucurbit[n]urils. *J. Am. Chem. Soc.* **134**, 15318–15323 (2012).

5. Walker, S., Oun, R., McInnes, F. J. & Wheate, N. J. The potential of cucurbit[n]urils in drug delivery. *Isr. J. Chem.* **51**, 616–624 (2011).

6. Gürbüz, S., Idris, M. & Tuncel, D. Cucurbituril-based supramolecular engineered nanostructured materials. *Org. Biomol. Chem.* **13**, 330–347 (2015).

7. Ye, R., Cui, Q., Yao, C., Liu, R. & Li, L. Tunable fluorescence behaviors of a supramolecular system based on a fluorene derivative and cucurbit[8]uril and its application for ATP sensing. *Phys. Chem. Chem. Phys.* **19**, 31306–31315 (2017).

8. Pemberton, B. C. *et al.* Supramolecular photocatalysis: Insights into cucurbit[8]uril catalyzed photodimerization of 6-methylcoumarin. *Chem. Commun.* **47**, 6323–6325 (2011).

9. Ko, Y. H., Kim, E., Hwang, I. & Kim, K. Supramolecular assemblies built with host-stabilized charge-transfer interactions. *Chem. Commun.* 1305–1315 (2007).

10. Kim, H. J. *et al.* Selective inclusion of a hetero-guest pair in a molecular host: Formation of stable charge-transfer complexes in cucurbit[8]uril. *Angew. Chemie - Int. Ed.* **40**, 1526–1529 (2001).

11. Sindelar, V. *et al.* Supramolecular assembly of 2,7-dimethyldiazapyrenium and cucurbit[8]uril: A new fluorescent host for detection of catechol and dopamine. *Chem. - A Eur. J.* **11**, 7054–7059 (2005).

12. Biedermann, F. *et al.* Benzobis(imidazolium)-cucurbit[8]uril complexes for binding and sensing aromatic compounds in aqueous solution. *Chem. - A Eur. J.* **16**, 13716–13722 (2010).

13. Jiao, D., Biedermann, F., Tian, F. & Scherman, O. A. A systems approach to controlling supramolecular architecture and emergent solution properties via host-guest complexation in water. *J. Am. Chem. Soc.* **132**, 15734–15743 (2010).

14. Biedermann, F., Elmalem, E., Ghosh, I., Nau, W. M. & Scherman, O. A. Strongly fluorescent, switchable perylene bis(diimide) host-guest complexes with cucurbit[8]uril in water. *Angew. Chemie - Int. Ed.* **51**, 7739–7743 (2012).

15. Zhang, Y. *et al.* Encapsulation enhanced dimerization of a series of 4-aryl-N-methylpyridinium derivatives in water: New building blocks for self-assembly in aqueous

media. *Chem. - An Asian J.* **9**, 1530–1534 (2014).

16. Biedermann, F., Ross, I. & Scherman, O. A. Host-guest accelerated photodimerisation of anthracene-labeled macromolecules in water. *Polym. Chem.* **5**, 5375–5382 (2014).

17. Yang, X., Wang, R., Kermagoret, A. & Bardelang, D. Oligomeric Cucurbituril Complexes: from Peculiar Assemblies to Emerging Applications. *Angew. Chemie - Int. Ed.* **59**, 21280–21292 (2020).

18. Chen, H., Yang, H., Xu, W. & Tan, Y. A supramolecular switch based on three binding states of a pyrene derivate: A reversible three-state switch with only two stimuli. *RSC Adv.* **3**, 13311–13317 (2013).

19. Jiao, Y., Xu, J. F., Wang, Z. & Zhang, X. Visible-Light Photoinduced Electron Transfer Promoted by Cucurbit[8]uril-Enhanced Charge Transfer Interaction: Toward Improved Activity of Photocatalysis. *ACS Appl. Mater. Interfaces* **9**, 22635–22640 (2017).

20. He, D. *et al.* Reversible stimuli-responsive controlled release using mesoporous silica nanoparticles functionalized with a smart DNA molecule-gated switch. *J. Mater. Chem.* **22**, 14715–14721 (2012).

21. Shimron, S., Wang, F., Orbach, R. & Willner, I. Amplified detection of DNA through the enzyme-free autonomous assembly of hemin/G-quadruplex DNAzyme nanowires. *Anal. Chem.* **84**, 1042–1048 (2012).

22. Kankanamalage, D. V. D. W. *et al.* DNA Strand Displacement Driven by Host-Guest Interactions. *J. Am. Chem. Soc.* **144**, 16502–16511 (2022).

23. Xiao, L. *et al.* Controllable DNA hybridization by host–guest complexation-mediated ligand invasion. *Nat. Commun.* **13**, (2022).

24. Seitz, O. Templated chemistry for bioorganic synthesis and chemical biology. *J Pep Sci* **25**, 1–19 (2019).

25. Kuriyan, J. & Eisenberg, D. The origin of protein interactions and allostery in colocalization. *Nature* **450**, 983–990 (2007).

26. Le, X. C. *et al.* Assembling DNA through Affinity Binding to Achieve Ultrasensitive Protein Detection. *Angew. Chem. Int. Ed* **52**, 10698–10705 (2013).

27. Rossetti, M., Bertucci, A., Patiço, T., Baranda, L. & Porchetta, A. Programming DNA-Based Systems through Effective Molarity Enforced by Biomolecular Confinement. *Chem.Eur.J.* **26**, 9826-9834 (2020).

28. Finn, M. G. & Fokin, V. V. Click chemistry: Function follows form. *Chem. Soc. Rev.* **39**, 1231–1232 (2010).

29. Best, M. D. Click chemistry and bioorthogonal reactions: Unprecedented selectivity in the labeling of biological molecules. *Biochemistry* **48**, 6571–6584 (2009).

30. Pickens, C. J., Johnson, S. N., Pressnall, M. M., Leon, M. A. & Berkland, C. J. Practical Considerations, Challenges, and Limitations of Bioconjugation via Azide−Alkyne

Cycloaddition. *Bioconjug. Chem.* **29(3)**, 686-701 (2018)

31. Biedermann, F., Ross, I. & Scherman, O. A. Host-guest accelerated photodimerisation of anthracene-labeled macromolecules in water. *Polym. Chem.* **5**, 5375–5382 (2014).

32. Xu, X. *et al.* [2 × 2] metallo-supramolecular grids based on 4,6-bis((1H-1,2,3-triazol-4-yl)-pyridin-2-yl)-2-phenylpyrimidine ligands: From discrete [2 × 2] grid structures to star-shaped supramolecular polymeric architectures. *Dalt. Trans.* **50**, 8746–8751 (2021).

33. Das, A. *et al.* Dynamics of Anthracene Excimer Formation within a Water-Soluble Nanocavity at Room Temperature. *J. Am. Chem. Soc.* **143**, 2025–2036 (2021).

34. Baranda Pellejero, L. *et al.* Using antibodies to control DNA-templated chemical reactions. *Nat. Commun.* **11**, (2020).

35. Soriano Del Amo, D. *et al.* Biocompatible copper(I) catalysts for in vivo imaging of glycans. *J. Am. Chem. Soc.* **132**, 16893–16899 (2010).

36. Presolski, S. I., Hong, V., Cho, S. H. & Finn, M. G. Tailored ligand acceleration of the cu-catalyzed azide-alkyne cycloaddition reaction: Practical and mechanistic implications. *J. Am. Chem. Soc.* **132**, 14570–14576 (2010).

37. Janeček, E. R., Rauwald, U., Del Barrio, J., Cziferszky, M. & Scherman, O. A. A facile route to viologen functional macromolecules through azide-alkyne [3+2] cycloaddition. *Macromol. Rapid Commun.* **34**, 1547–1553 (2013).

38. Dang, D. T., Schill, J. & Brunsveld, L. Cucurbit[8]uril-mediated protein homotetramerization. *Chem. Sci.* **3**, 2679–2684 (2012).

39. Urbach, A. R. & Ramalingam, V. Molecular recognition of amino acids, peptides, and proteins by cucurbit[n]uril receptors. *Isr. J. Chem.* **51**, 664–678 (2011).

40. Heitmann, L. M., Taylor, A. B., Hart, P. J. & Urbach, A. R. Sequence-specific recognition and cooperative dimerization of N-terminal aromatic peptides in aqueous solution by a synthetic host. *J. Am. Chem. Soc.* **128**, 12574–12581 (2006).

41. Williams, D. B. G. & Lawton, M. Drying of organic solvents: Quantitative evaluation of the efficiency of several desiccants. *J. Org. Chem.* **75**, 8351–8354 (2010).

Chapter 5

Design of a fluorescent sensor for protein detection based on aptamer conjugated tetraphosphonate cavitands

Acronyms and abbreviation

DMSO	Dimethylsulphoxyde
MeOH	Methanol
THF	Tetrahydrofuran
THPTA	Tris (3-hydroxypropyltriazolylmethyl) amine
PEG	Polyethylene glycol
CB	Cucurbituril
MV	Methylviologen
DNA	Deoxyribonucleic acid
CuAAC	Copper-Catalysed Azide-Alkyne Cycloaddition
ESI-MS	Electrospray Ionization - Mass Spectrometry
ESI-Orbitrap-MS	Electrospray Ionization - Orbitrap - Mass Spectrometry
TLC	Thin- Layer Chromatography
UV-Vis	Ultraviolet-visible spectroscopy
ITC	Isothermal calorimetry

5.1 Introduction

Supramolecular sensing can be defined as a process that take advantage of the dynamic weak interactions (hydrogen bonds, cation-π, CH-π, π-π and Van der Waals) involved in the formation of host-guest assemblies as source of signal modulation when analytes are added. In general, a sensor is constituted by an active layer that interacts with the analyte and a transducer, which reports a detectable signal response. The signal generated is strictly dependent on chemical interactions involved and need to be detected by instrumentation able to transform the established interactions into an analytical detectable signal, such as pH, temperature or optical properties (Figure 1).[1]

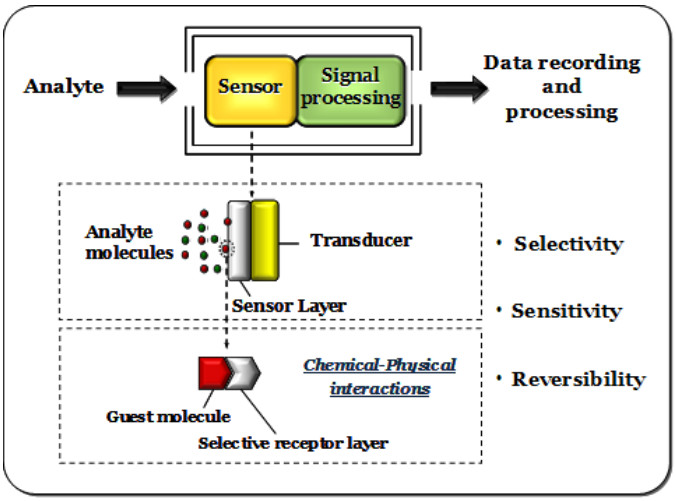


Figure 1. Supramolecular sensor

For biological sensing, in general, the processes based on optical responses are widely used even if most of the biological processes of interest are non-fluorescent. The possibility to obtain fluorescence-based sensors that exploit the use of small molecules or indicators represents an interesting tool to study biological events through a change in fluorescence.[2] This behaviour is usually reversible and can be exploited to follow dynamic processes over time. Fluorescent-based detection *via* supramolecular chemistry represent an interesting strategy since it takes advantage of the instauration of non-covalent interactions between host and guest molecules. However, this strategy requires a sufficiently strong interaction between host and guest to enable the detection of the fluorescence change upon the complexation. To date, many solutions have been reported exploiting different position of the fluorescent dye in the system, as illustrated in Figure 2. These sensing designs can either be based on the inclusion/expulsion of the dye-guest from the cavity of the receptor (Figure 2, a) or involve the direct attachment of the dye as a label to the guest or host molecule (Figure 2, b, c).[1]

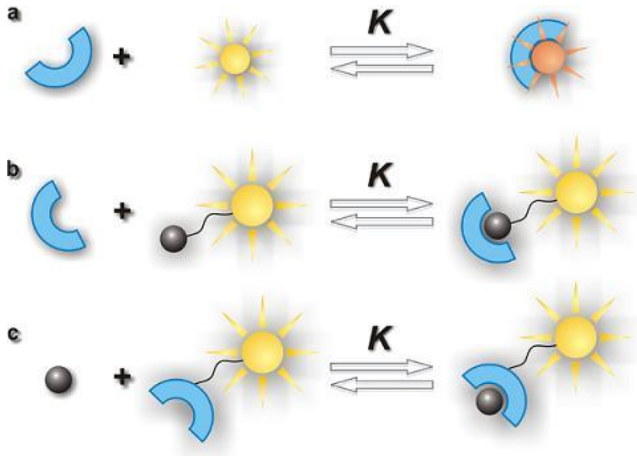


Figure 2. Fluorescence-based sensor systems.[3]

As widely illustrated in Chapter 2, in aqueous media tetraphosphonate cavitands can form host-guest complexes with K_a values in the order of 10^4 M^{-1} with specific guest molecules, such as methylammonium salts, exploiting their hydrophobic cavity and the establishment of simultaneous non-covalent interactions.[4] Taking advantage of these complexation properties, it is possible to design systems based on Tiiii cavitands acting as hosts and fluorescent ammonium-derivatives as guests.

In literature several examples in which these properties have been exploited to create a fluorescent probe for the detection of molecules of interest are reported. Dalcanale and co-workers exploited a pyrene derivative bearing a methylpyridinium group to demonstrate the fully reversible complexation of a methylammonium salt, monitoring the phenomena *via* fluorescence analysis in organic solution. It is well known that Tiiii preferentially binds methylammonium guests respect to the methylpyridinium ones, as with the latter no simultaneous hydrogen bonds can be established.[4] Thus, the addition of Tiiii to a methylpyridinium solution (1 x 10^{-5} M) in dichloromethane, led to a progressive increase of fluorescence intensity profile. While, after the addition of the methylammonium salt to the equimolar solution of Tiiii and methylpyridinium (1 x 10^{-5} M), the fluorescence emission profile decreased until reaching the typical value of the pyrene guest alone.[5]

In 2015, the same authors reported a study based on the design of a fluorescent probe for the detection and quantification, in water solution, of illicit drugs containing the N-methylated group and belonging to class of amphetamine-type-stimulants (ATS). In this case, the authors synthesized a Tiiii functionalized at the upper rim with two pyrene moieties. The two fluorophores were placed in distal position and oriented so that an optimal overlapping between the two units was possible and thus resulting in the formation of the excimer species, behaving as a transducer element. Because of the low solubility of the receptor in water medium, the cavitand was successfully loaded into water soluble organo-silica core nanoparticles (PluS NPs). The guest inclusion inside the receptor cavity was observed only for the bulkier methylammonium guests (e.g., MDMA) since they are able to bring

apart the pyrene units, leading to the perturbation of the excimer and thus resulting in a change in the fluorophore emission profile (Figure 3).[6]

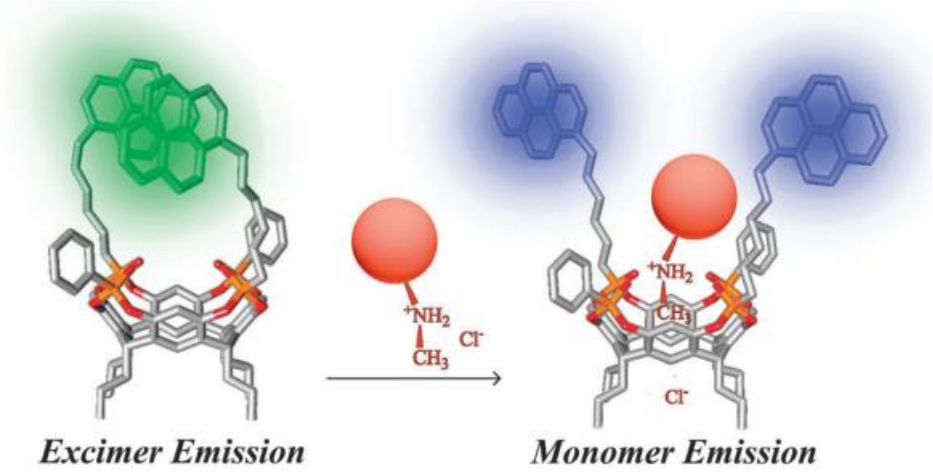


Figure 3. Fluorescence chemosensor based on Tiiii.[6]

More recently, Zhong and co-workers showed the possibility to detect oligonucleotide secondary structures exploiting the use of an array based on water-soluble receptors as hosts and cationic dyes that exhibit affinity for both their cavities and DNA G-quadruplexes.[7] The sensor provided a variation of fluorescence response in the presence of DNA strands resulting in both a classification and differentiation of G-quadruplex structures (Figure 4).

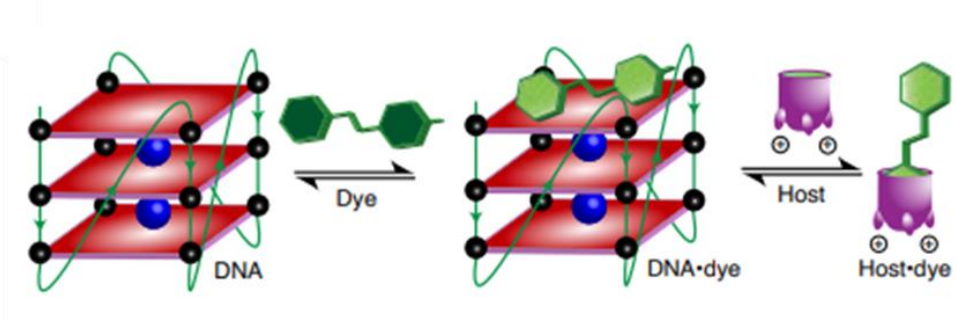
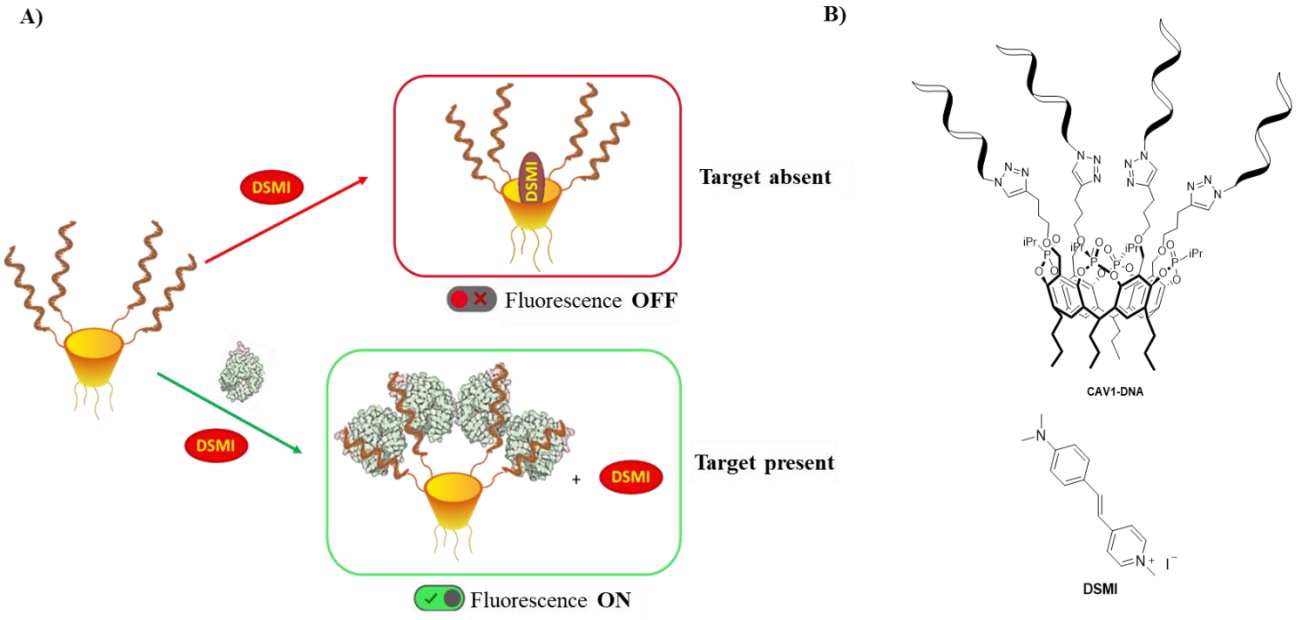


Figure 4. Illustration of the supramolecular approach based on an array of hosts and different guest dyes that show variable affinity for each other and with the oligonucleotide target allowing variable fluorescence responses.[7]

By exploiting the peculiarities in the recognition properties of the tetraphosphonate cavitands, in this chapter we designed, in collaboration with Dr. A. Bertucci (University of Parma), a system based on aptamer conjugated tetraphosphonate cavitands and fluorophores bearing a methylpyridinium group as guests. In particular, trans-4- [4- (dimetilammino)stiril]-1-metilpiridinio iodide (**DSMI**) was selected as dye guest. This fluorescent probe has been already applied by Hooley and *co-workers* for the recognition of citrulline modifications in cationic histone peptides, exploiting its complexation by water-soluble deep cavitands. They demonstrated that these host-guests pairs can distinguish the

citrulline modifications in peptides modifying their fluorescence responses.[8] Moreover, **DMSI** inclusion in the Tiiii cavity, due to the presence of the methylpyridinium group, was already been demonstrated in water (K_a in the order of 10^4 M^{-1}) and caused the quenching of the fluorescence of the dye.[7]

The general scheme of the sensing mechanism is represented in Scheme 1.



Scheme 1. A) Representation of the supramolecular switch induced by the cavitand complexation properties; **B)** Structures of the supramolecular receptor (**CAV1-DNA**) and fluorophore guest (**DSMI**).

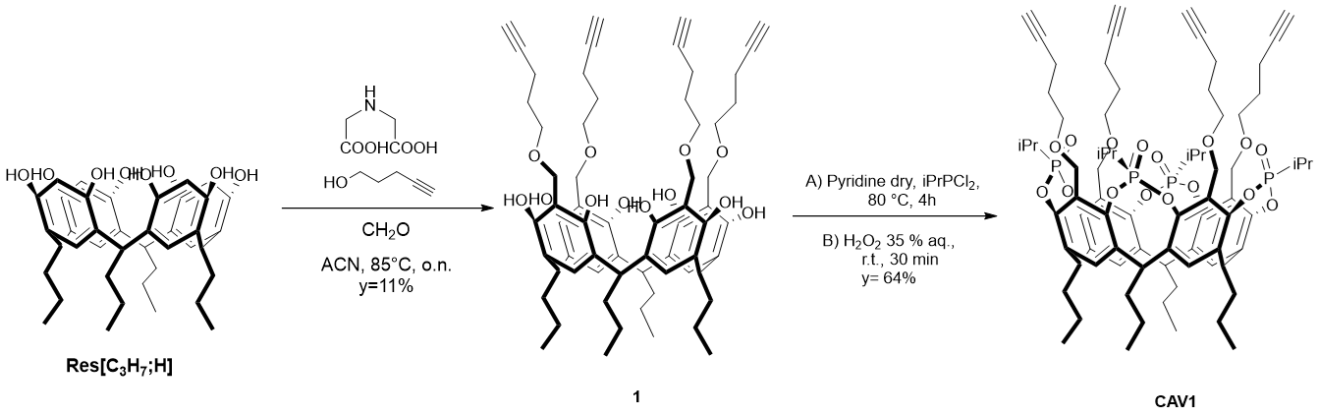
The DSMI is used as a fluorescent guest while the supramolecular receptor is a tetraphosphonate cavitand functionalized at the upper rim, in apical position, with four aptamer units. The aptamer, by definition, is a short, single-stranded DNA or RNA (ssDNA or ssRNA) molecules able to bind a specific target, in this case a protein, with high affinity. In this case, we selected an aptamer complementary to the protein "Vascular endothelial growth factor" (VGEF), which is a potent angiogenic mitogen commonly overexpressed in cancerous cells.[9] For its conjugation with tetraphosphonate cavitand CuAAC click reaction, already discussed in Chapter 4, was selected and *ad hoc* Tiiii functionalized at the upper rim with four alkyl chains bearing a terminal alkyne (**CAV1**) was synthesized. The proposed biosensor is based on a fluorescence turn-on switch and foresees two different scenarios depending on whether the target protein is present in the sample or not: if the target is absent, the formation of the host-guest complex between the cavity and the fluorophore is possible and the fluorophore is quenched; in contrast, when the aptamers bind the target protein, the formation of the cavitand-fluorophore complex is hampered since the macrocycle cavity is sterically occluded. In this case, the recognition of the protein by one or more of the oligomeric units located at the upper rim of the cavity leads to a cavity that is less available for the complexation. It is therefore expected

that the fluorescent guest will preferentially remain in solution while keeping its fluorescence profile almost unaltered.

5.2 Results and discussion

5.2.1 Synthesis of CAV1

Cavitand **CAV1** was synthesized in two synthetic steps from the propyl-footed resorcinarene, Res[C_3H_7;H] (Scheme 2).



Scheme 2. Synthesis of cavitand **CAV1**.

In the first synthetic step the functionalization of the resorcinarene Res[C_3H_7;H] in apical position with an alkyne-terminated chain was performed *via* Mannich reaction. The resorcinarene was reacted with formaldehyde and 4-pentin-1-ol, with iminodiacetic acid as catalyst, in acetonitrile at 85 °C, for 16 hours. After purification by column chromatography, compound **1** was obtained as white powder in 11 % yield and characterized by ^{1}H NMR and ESI-MS. The following step involved the bridging of the phenolic groups of **1** by nucleophilic substitution reaction with dichloroisopropylphosphine. The reaction was conducted using pyridine as both solvent and base at 80 °C for 4 hours to obtain a tetraphosphonito intermediate. This step allows the formation exclusively of the tetraphosphonito cavitand presenting the lone pair on the P(III) groups pointing inward the cavity. The following *in situ* oxidation to P(V) with hydrogen peroxide, which proceeded with retention of configuration at phosphorous center, led to the isolation of **CAV1** with all the four P=O groups pointing inward the cavity. **CAV1** was obtained as white solid in 64 % of yield and fully characterized by ^{1}H and ^{31}P NMR, ^{1}H-^{1}H COSY NMR and ESI-MS.

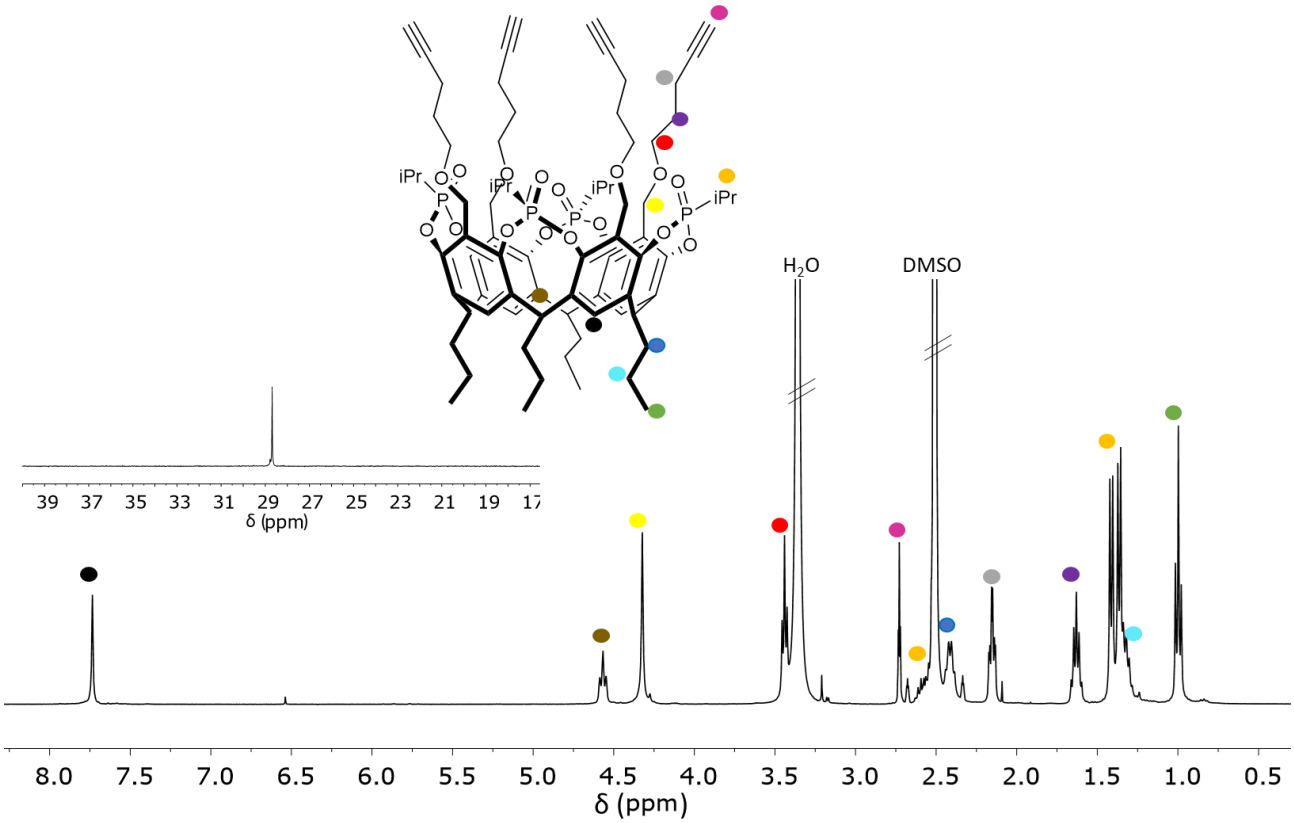


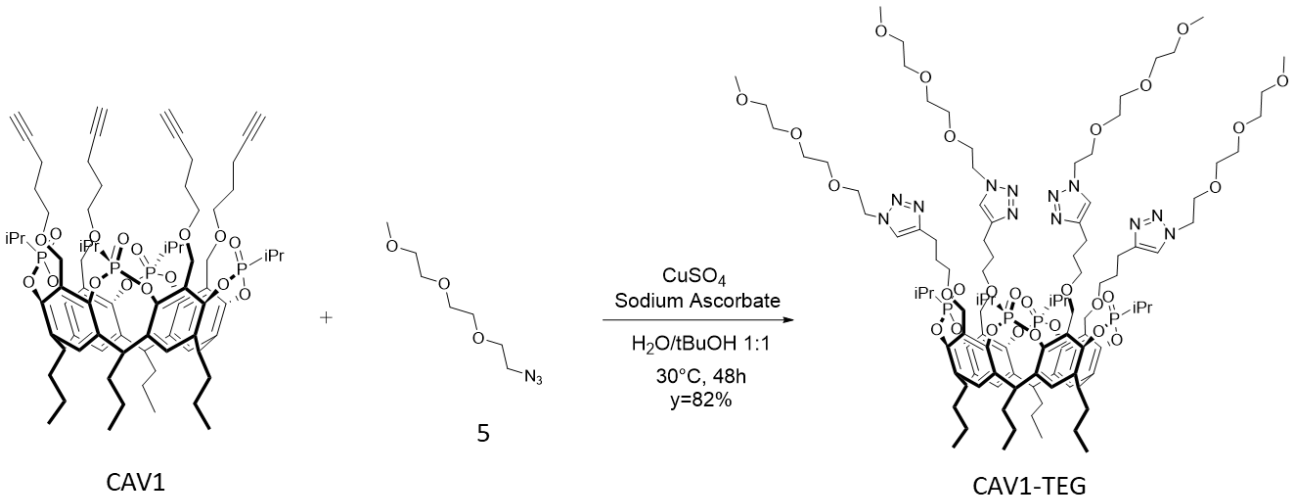
Figure 5. ^{1}H and ^{31}P NMR spectra of **CAV1**, in DMSO-d_6, 400 MHz, 25 °C.

In the proton spectrum, reported in Figure 5, the triplet relative to the methine bridge (brown dot) is visible at 4.57 ppm, while the singlet at 2.73 ppm is ascribable to the alkyne proton (pink dot). Diagnostics are the multiplets at 2.65 ppm (hided by the residual solvent peak) and the doublets at 1.39 ppm, which are related to the protons of the isopropyl group of the phosphonate bridging groups. In the ^{31}P NMR spectrum, only one peak is visible, as clear indication that the isolated product is only one out of the possible isomers, namely the cavitand with all the P=O groups facing inwards the cavity (Tiiii).

5.2.2 Synthesis of CAV1-TEG

Similarly, to the approach described in Chapter 4, before proceeding with the functionalization of the receptor with the oligonucleotide strands, we decided to test **CAV1** reactivity under CuAAC click conditions with four short TEG chains. These functionalities allowed us also to investigate the influence of the apical functionalization on the binding of fluorescent dye through NMR, ITC and fluorescence analysis.

The synthesis of **CAV1-TEG** was performed using the azide-functionalized TEG chain **5** (Scheme 3). The procedures reported in Chapter 4 were followed both for the general click protocol and the preparation of compound **5**.



Scheme 3. Synthesis of cavitand **CAV1-TEG**

As slight modification, a mixture of water and *tert*-butanol in 1:1.43 ratio was used as solvent, being able to solubilize both **CAV1** and the click reactants and compatible with the target DNA strands. The click protocol followed was based on the use of copper (II) salt as source of the catalytic species, in presence of sodium ascorbate as reducing agent to induce the *in situ* formation of Cu(I) catalyst. The reaction proceeded for 48 hours at 30 °C. **CAV1-TEG** was obtained, after purification with several trituration in dichloromethane, as an ochre solid in 82 % yield and fully characterized by ^{1}H and ^{31P} NMR and MALDI-TOF.

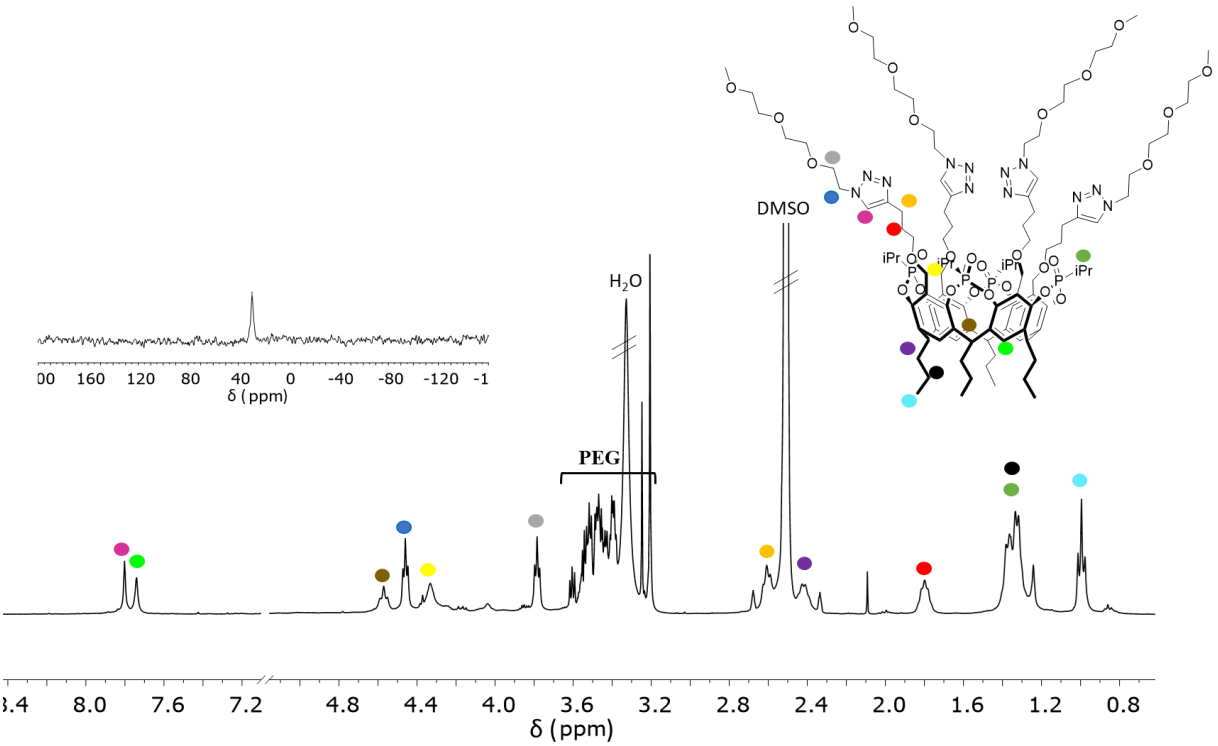


Figure 6. ^{1}H and ^{31}P NMR spectra of **CAV1-TEG**, in DMSO-d_6, 400 MHz, 25 °C.

In the ^{1}H NMR spectrum of **CAV1-TEG**, illustrated in Figure 6, the signals related to the chains of TEG are visible in the area between 3.1 and 3.58 ppm. The singlet present at 7.72 ppm, related to the triazole proton (*pink spot*), is diagnostic of the formation of the triazole ring. Other confirmations of the isolation of **CAV1-TEG** are the downfield shift (from 2.10 to 2.46 ppm) of the signal relative to the CH_2 in α to the triazole (orange spot) and the disappearance of the triplet at 2.73 ppm relative to the alkyne proton. Finally, the effective apical functionalization of all four TEG chains was confirmed by the presence of the signal at 2151.054 m/z in the MALDI-TOF, relative to the molecular ion spectrum (Figure 7).

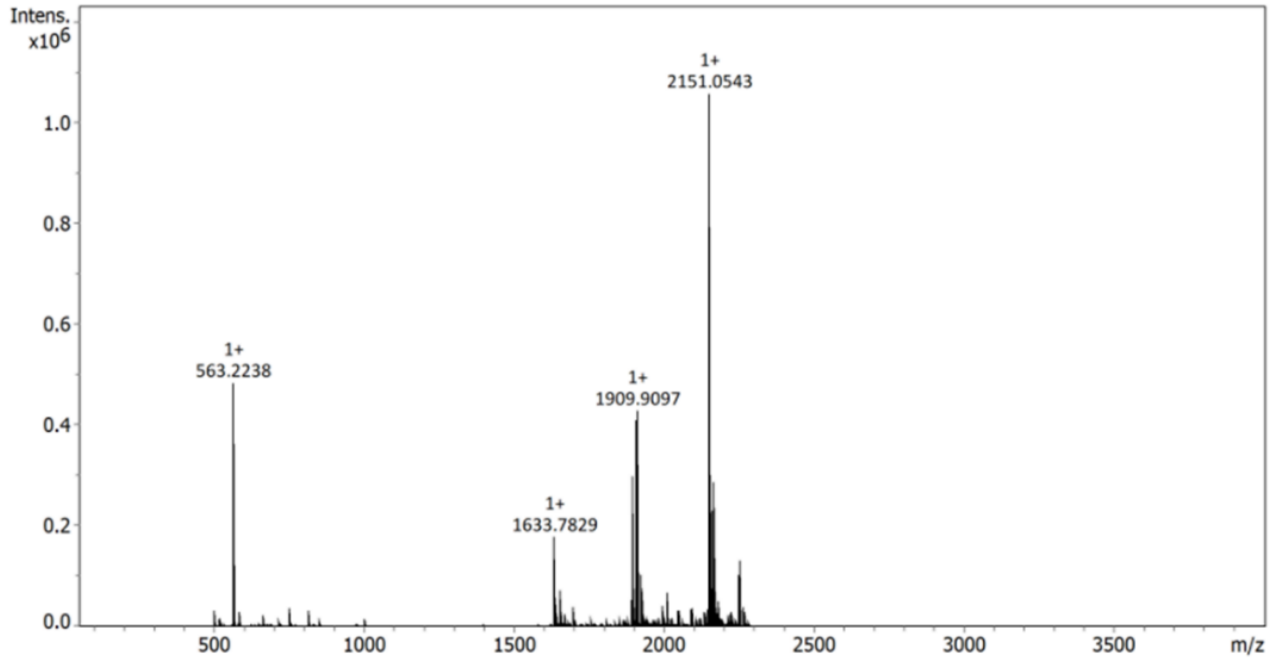


Figure 7. MALDI-TOF spectrum of **CAV1-TEG**.

5.2.3 Host-guest complexation studies

After the isolation and full characterization of **CAV1-TEG**, we studied its complexation of **CAV1-** with **DSMI** dye. Firstly, an NMR titration in deuterated water was performed.

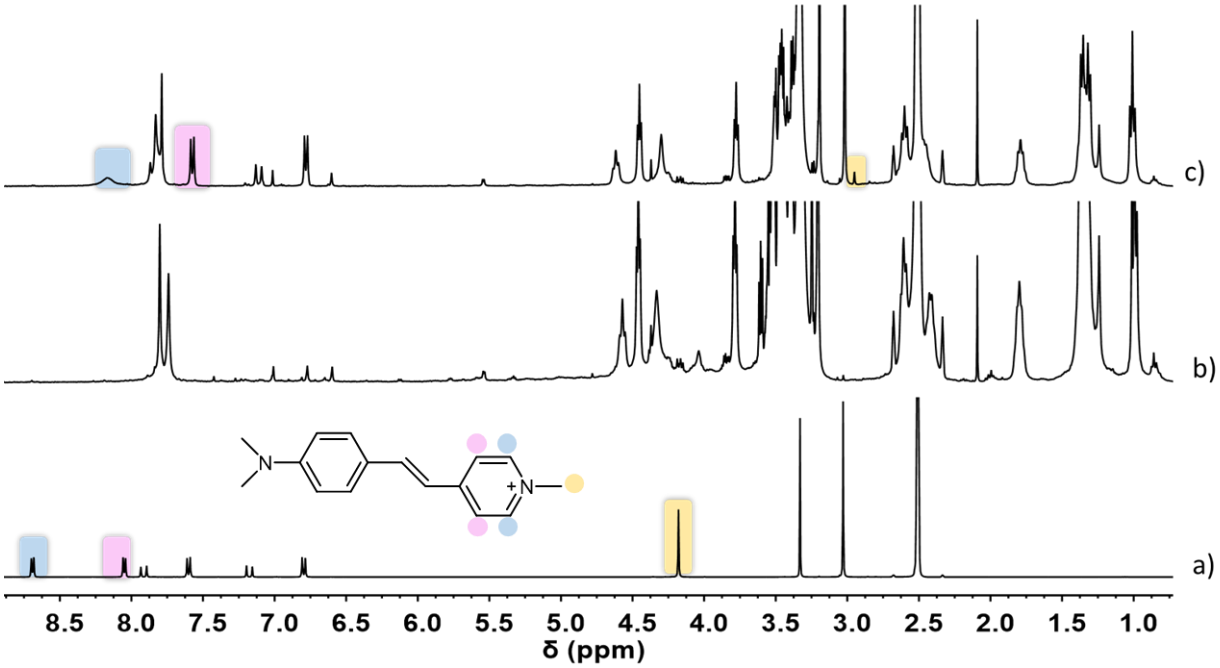


Figure 9. ^{1}H NMR spectra of a) **DMSI**, b) **CAV1-TEG**, c) **CAV1-TEG/DMSI** 1:1, in DMSO-d_6, 400 MHz, 25 °C.

In the ^{1}H NMR spectrum (Figure 9c) obtained by adding 1 equivalent of **DMSI** to a solution of **CAV1-TEG** in deuterated dimethylsulphoxide is reported in Figure 9c: the shift from 4.18 to 2.95 ppm of CH_3 of the methylpyridinium group of **DSMI** (yellow mark) is diagnostic of the inclusion in the Tiiii cavity. Moreover, the signal relating to the proton in position 2 of methylpyridinium (blue mark) shifts from 8.70 to 8.17 ppm and, the proton at position 3 of methylpyridinium (pink mark) moves shifts from 8.05 to 7.79 ppm.

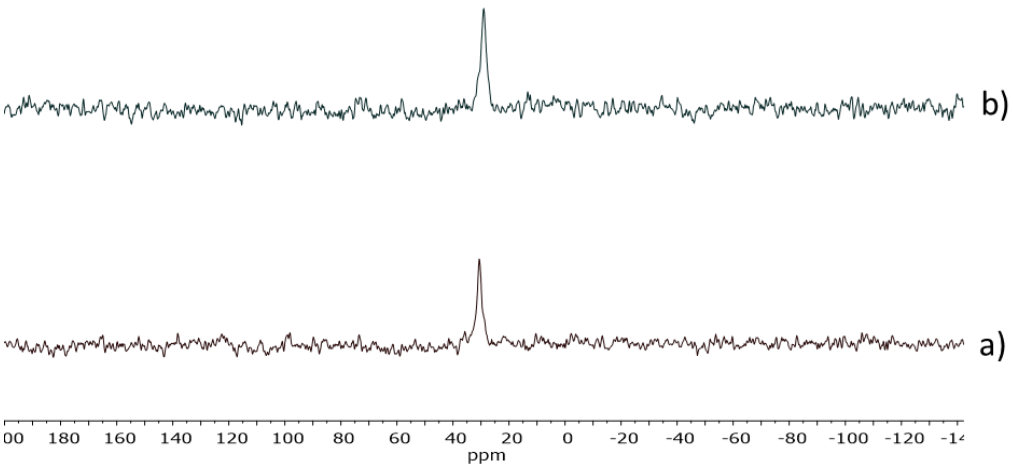


Figure 10. ^{31}P NMR spectra comparison of a) **CAV1-TEG/DMSI** 1:1 and b) **CAV1-TEG**

In addition, in ^{31}P NMR spectra, illustrated in Figure 10, it is possible to note the shift from 29.0 to 30.7 ppm leading to the conclusion that the guest complexation within the host cavity took place.

As further confirmation of the complex formation, a fluorescence titration was performed, in methanol as solvent, titrating a 10^{-5} M solution of DSMI with a 10^{-3} M solution of **CAV1-TEG**.

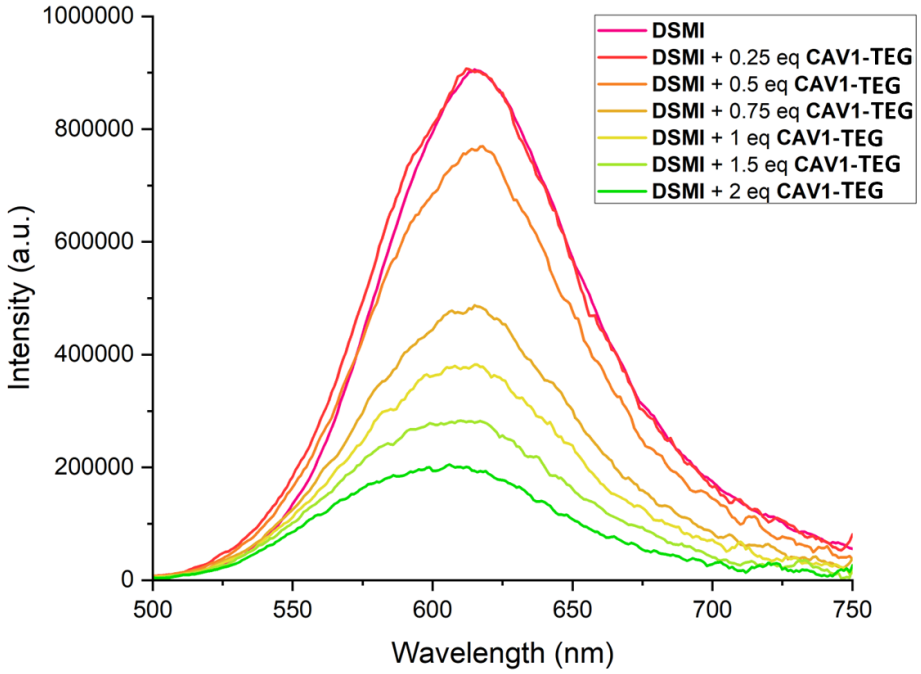


Figure 11. Fluorescence spectra (λ_{ecc} 485 nm) relative to the titration of **DSMI** with **CAV1-TEG**, in methanol.

In the fluorescence titration spectra (Figure 11), a quenching of the fluorescence is clearly visible even with the addition of 0.5 equivalents of **CAV1-TEG**, indicating the effective complexation of the guest by the cavitand despite the increased steric hindrance in the upper rim of the cavitand due to the TEG chains.

These results clearly indicate that apical functionalization with groups with limited bulkiness do not hamper the inclusion and the quenching of **DSMI**.

Finally, a calorimetric study was conducted on the host-guest system through ITC titration. The thermodynamic parameters obtained were compared to those of a model system involving **DSMI** with cavitand **CavPOEt**, that does not show a functionalization in apical position, to assess the effect of the functionalization at the upper rim on the complexation with **DSMI** guest (Figure 12).

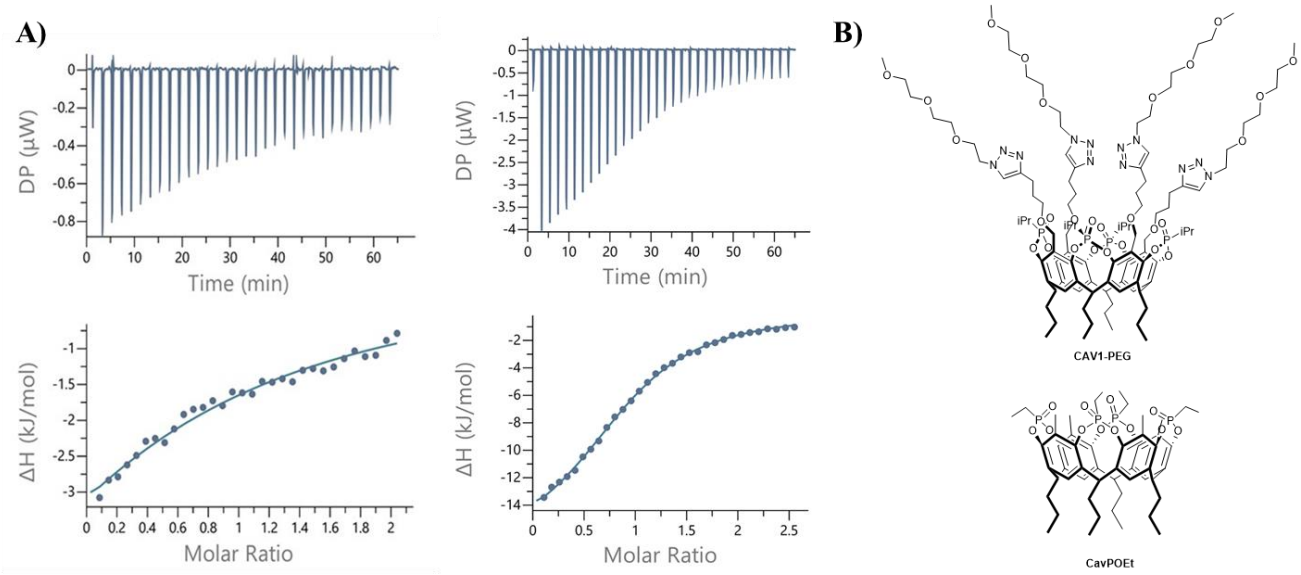


Figure 12. A) ITC measurement of **CAV1-TEG** with **DSMI** (*left*) and **CavPOEt** with **DSMI** (*right*). **B)** Structures of **CAV1-TEG** and **CavPOEt**.

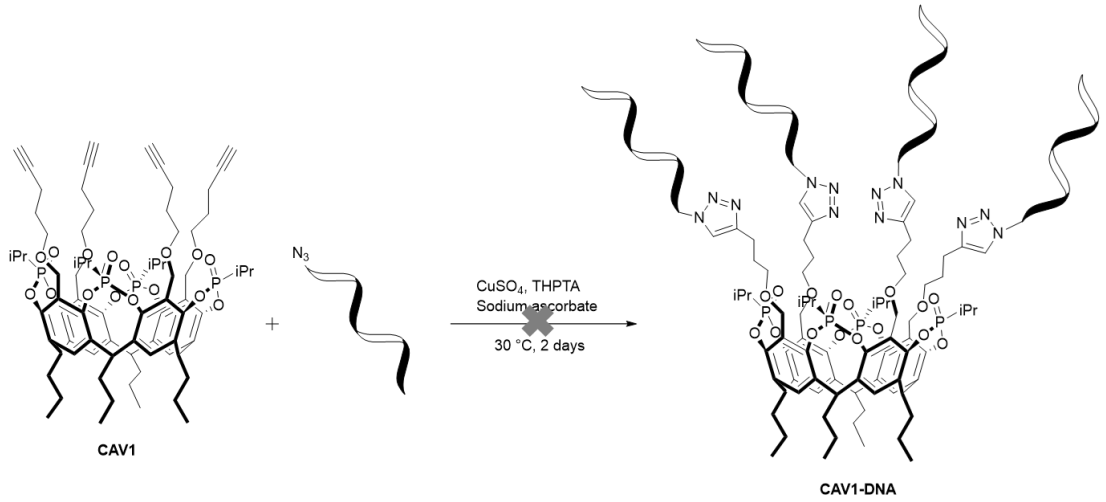
From the obtained ITC parameters, it is possible to see that in both cases the complexation was favoured at entropic and enthalpic levels (Table 1). The K_a values obtained are in the order of 10^4 M^{-1} for **CavPOEt** and 10^3 M^{-1} for **CAV1-TEG** cavitand. Therefore, it can be concluded that the increased steric hindrance at the upper rim of **CAV1-TEG** resulted in the loss of one order of magnitude in the association constant. This observation is promising for the target detection of VGEF proteins with aptamer-decorated cavitands as the increase of the steric hinderance around Tiiii cavity have affects the interaction with the fluorescence probe.

Table 1. Thermodynamic parameters from ITC performed in methanol at 25 °C.

Host	n	K_a (M^{-1})	ΔH (KJ/mol)	ΔG (KJ/mol)	-TΔS (KJ/mol)
CavPOEt	0.9(±0.05)	$2.22x10^4$ (± $0.09x10^4$)	-24.8 (± 0.7)	-16.3 (± 1.7)	-8.5 (± 1.6)
CAV1-PEG	0.82(±0.02)	$5.74x10^3$ (± $0.093x10^3$)	-5.1 (± 0.5)	-21.4 (± 0.07)	-16.4 (± 0.4)

5.2.4 Click reactions of CAV1 with oligonucleotide strands

Click reaction between **CAV1** and two different oligonucleotide strands were performed following the procedure described in Chapter 4 (Scheme 4).



Scheme 4. Synthesis of cavitand **CAV1-DNA.**

Since **CAV1** was not soluble in water, a mixture of organic solvent and water was used. The click protocol was based on the use of copper (II) salt as source of the catalytic species in presence of sodium ascorbate as reducing agent to induce the *in situ* formation of Cu(I) catalyst and THPTA ligand to stabilize the Cu(I) catalyst formed during the reaction. Several click reaction trials, exploiting the use of two aptamer units (DNA4 and DNA5) that differ in terms of sequences length, were performed by varying the equivalents of DNA or catalyst, the solvents ratio, and the type of organic solvent and are summarized in Table 2. The reaction mixtures were kept at 30 °C for 2 days, after which the solvent was evaporated, and the crude was characterized using ESI-LTQ-Orbitrap or MALDI-TOF mass analyses.

Table 2. Click reaction conditions tested

Product	DNA	DNA equivalents	Catalyst equivalents	solvent
CAV1-DNA4_1	DNA4	6	5	tBuOH/H_2O 1:1
CAV1-DNA4_2	DNA4	4	5	tBuOH/H_2O 1:1
CAV1-DNA5_1	DNA5	1	1.1	tBuOH/H_2O 1:1
CAV1-DNA5_2	DNA5	2	2.2	tBuOH/H_2O 1:1
CAV1-DNA5_3	DNA5	4	4.4	tBuOH/H_2O 1:1
CAV1-DNA5_4	DNA5	4	4.4	MeOH/H_2O 1:1
CAV1-DNA5_5	DNA5	4	4.4	MeOH/H_2O 99:1

With regard to the first tests (CAV1-DNA4_1 and CAV1-DNA4_2), the characterization of the crude was carried out through MALDI-TOF mass analysis but the product was not observed. It was therefore decided to test a shorter oligonucleotide (DNA5), so that the analysis could be performed with the ESI-LTQ-Orbitrap instrumentation present in the University of Parma facilities. Also in this

case, after several tests (CAV1-DNA5_1-5) the characterization by ESI-LTQ-Orbitrap did not allow to the identification of product formation.

5.3 Conclusions

CAV1 reporting four alkyne groups in apical position was successfully synthetized and fully characterized. Subsequently, it was functionalized with four TEG chains to obtain a model system with a dual purpose: to test its reactivity under CuAAC click conditions and to evaluate the effect of an increased hindrance at the upper rim in complexing properties of the resulting CAV1-PEG cavitand toward DSMI, mimicking the hindrance of the aptamers units. The host-guest model system was studied by ^{1}H NMR and fluorescence, proving the formation of the binary complex and by ITC analyses reporting a K_a value acceptable for the purpose, in the order of 10^3 M^{-1}. Thanks to these preliminary studies, **CAV1** proved to be suitable to be used in the final system based on aptamer-protein interaction. Thus, several click reaction tests were carried out with two aptamer units, different in terms of length, by varying different reaction parameters, but neither the final product nor partially reacted cavitand products were observed. Future perspectives will provide the click reaction performance with different reaction conditions such as the use of copper-free click AAC reported by Bertozzi and *co-workers*.[10]

5.4 Experimental section

All reagents and solvent were purchased from certified commercial sources and used as received, without further purification. For the synthesis, all solvents were dried and distilled according to standard procedures known in the literature.

Thin layer chromatography was performed using TLC Analytical Chromatography F254®, Merck.

Oligonucleotide strands were purchase from Metabion International AG (Germany) in lyophilized form, subsequently dissolved in MIlliQ water at a concentration of 100 μM and stored at -20 °C. The sequences used (5'-3') are shown in Table 3.

Table 3. Oligonucleotide sequences (5'-3'). In bold are highlighted the complementary portions.

Name	Sequence
DNA4	Azide-C3-CAA TTG GGC CCG TCC GTA TGG TGG GT
DNA5	Azide-C3-TGG TAC GAA TTC GAC

^{1}H and ^{31}P NMR spectra were recorded on Bruker 400MHz AVANCE or Jeol 600MHz using DMSO-d_6 and $CDCl_3$ as solvents. All chemical shifts (δ) were reported in parts for million (ppm) relative to proton resonances resulting from incomplete deuteration of NMR solvents. The abbreviations: s, d, t, m and sb indicated the spectrum peaks referred to: singlet, doublet, triplet, multiplet and broad singlet, respectively. The coupling constant (J) are expressed in Hz.

ESI-MS characterization experiments were performed on a Waters ACQUITY Ultra Performance LC HO6UPS-823M with ESI source ionization (electrospray ionization) in positive modality. High-resolution analyses were performed with a Thermo LTQ-Orbitrap.

UV-Vis absorption spectra were collected with Perkin Elmer Lambda 750 UV-VIS-NIR instrument. Fluorescence spectra were recorded using Fluoromax-3 instrument.

Synthesis of 1

Res[C_3H_7; H] (300 mg, 0.456 mmol), formaldehyde 37 % aq. (64 μL, 2.28 mmol), 4-pentin-1-ol (1.7 mL, 18.27 mmol) and iminodiacetic acid (30 mg, 0.228 mmol) were dissolved in acetonitrile (4.8 mL). The reaction mixture was heated at 85 °C under magnetic stirring. After 16 hours, the solvent was evaporated and the solid obtained was dissolved in ethyl acetate and extracted with water and brine. The organic phase was anhydrified with sodium sulphate and the solvent was evaporated under reduced pressure. The crude was purified by flash chromatographic column (silica gel, hexane/ethyl acetate 8:2) and subsequently triturated in water, obtaining **1** (50.4 mg) as a white solid in 11% yield. **^{1}H NMR** (400 MHz, $CDCl_3$, 25 °C), δ (ppm) = 8.68 (d, *J* = 11.8 Hz, 4H, O*H*), 7.19 (s, 4H, **H_{down}**), 4.83 (s, 8H, ArC**H_2**), 4.32 (t, *J* = 7.9 Hz, 4H, C**H_{bridge}**), 3.67 (t, *J* = 6.2 Hz, 8H, OC**H_2**CH_2), 2.30 (td, *J* = 7.0, 2.7 Hz, 8H, C**H_2**CCH), 2.19 (q, *J* = 7.6 Hz, 8H, CH_2C**H_2**CH_2), 1.99 (t, *J* = 2.7 Hz, 4H, CC**H**), 1.86 (p, *J* = 6.6 Hz, 8H, C**H_2**CH_2CH_3), 1.42 – 1.31 (m, 8H, C**H_2**CH_3), 1.00 (t, J = 7.4 Hz, 12H, C**H_3**).

Synthesis of CAV1

Compound **1** (50.4 mg, 0.086 mmol) was dissolved in pyridine (1.6 mL) under anhydrous conditions. The solution was brought to 0 °C and dichloroisopropylphosphine (26 μL, 0.38 mmol) was added. Then, the reaction was heated to 80 °C and left reacting under magnetic stirring and nitrogen flow. After 4 hours, the mixture was cooled at 0 °C and H_2O_2 (0.8 mL) was added. The reaction mixture was stirred for 1 hour and then poured into 100 mL cold water. The formation of a white solid was observed. The suspension was allowed to decant at 4 °C for two days, after which it was centrifugated and the supernatant removed. The white solid was finally washed with water obtaining **CAV1** (32.6 mg) as a white solid, in 64 % yield. **^{1}H NMR** (400 MHz, DMSO-d_6, 25° C), δ (ppm) = 7.73 (s, 4H, **H_{down}**), 4.57 (t, *J* = 7.6 Hz, 4H, C**H_{bridge}**), 4.32 (s, 8H, ArC**H_2**), 3.44 (t, *J* = 6.1 Hz, 8H, OC**H_2**CH_2), 2.73 (t, *J* = 2.6 Hz, 4H, CC**H**), 2.64 – 2.55 (m, 4H,PC**H**), 2.41 (q, *J* = 7.8 Hz, 8H, C**H_2**CH_2CH_3), 2.15 (td, *J* = 6.9, 2.6 Hz, 8H, C**H_2**CCH), 1.63 (p, *J* = 6.6 Hz, 8H, OCH_2C**H_2**), 1.39 (dd, *J* = 19.7, 7.1 Hz, 24H, PCH(C**H_3**)$_2$), 1.35 – 1.26 (m, 8H, C**H_2**CH_3), 1.00 (t, J= 7.3 Hz, 12H, CH_2C**H_3**). **^{31}P NMR** (162 MHz, DMSO-d_6, 25 °C), δ (ppm) = 28.7 (s, **P**O). **ESI-MS:** m/z 1416.05 $[M+Na]^+$.

Synthesis of CAV1-TEG

CAV1 (30 mg, 0.021 mmol) and **5** (24.4 mg, 0.129 mmol) were solubilized in H_2O/tert-butanol 2:5 mixture (7 mL). Separately, $CuSO_4$ (26.8 mg, 0.107 mmol) and sodium ascorbate (21.3 mg, 0.107 mmol) were solubilized in H_2O (3 mL). Both the solutions were degassed through three cycles of freeze-pump-thaw. Then, the two solutions were mixed together in the same Shlenk and the resulting mixture was degassed again with three cycles of freeze-pump-thaw. The solution was kept under magnetic stirring at 30 °C. After 2 days, a pale blue precipitate was observed. It was filtered off,

purified by several triturations in DCM and dried. **CAV1-PEG** (37.2 mg) was obtained pure as ochre solid in 82% of yield. **^{1}H NMR** (600 MHz, DMSO-d_6, 25 °C), δ (ppm) = 7.79 (s, 4H, $\boldsymbol{H_{down}}$), 7.72 (s, 4H, $\boldsymbol{H_{triazole}}$), 4.56 (t, *J* = 8.1 Hz, 4H, C$\boldsymbol{H_{bridge}}$), 4.45 (t, *J* = 5.2 Hz, 8H, NC$\boldsymbol{H_2}$), 4.34 (s, 8H, ArC$\boldsymbol{H_2}$), 3.77 (t, *J* = 5.3 Hz, 8H, $NCH_2C\boldsymbol{H_2}$), 3.64 – 3.34 (m, 40H, C$\boldsymbol{H_2}$ PEG + $OC\boldsymbol{H_2}CH_2CH_2$), 3.20 (s, 12H, OC$\boldsymbol{H_3}$), 2.65 – 2.56 (m, 8H, C$\boldsymbol{H_2}$CN), 2.44 – 2.35 (m, 8H, $C\boldsymbol{H_2}CH_2CH_3$), 1.79 (t, *J* = 7.4 Hz, 8H, $OCH_2C\boldsymbol{H_2}CH_2$), 1.46 – 1.27 (m, 32H, $CH(C\boldsymbol{H_3})_2$ + $C\boldsymbol{H_2}CH_3$), 0.98 (t, *J* = 7.7 Hz, 12H, $CH_2C\boldsymbol{H_3}$). **^{31}P NMR** (162 MHz, DMSO-d_6, 25 °C), δ (ppm) = 28.93 (s, *P*O). **MALDI-TOF**: Theoretical: 2150.3842 m/z; Experimental: 2151.0543 m/z $[M+H]^+$.

Protocol attempted for CAV1-DNA preparation

In an Eppendorf tube the DNA solution (40 μL, 100 μM) and **CAV1** (50 μL, 1,332 x 10^{-5}M) were solubilized in organic solvent/water (in according with the ratios reporting in Table 2). 10 μL (3,33 x 10^{-4}M) of a water solution containing $CuSO_4$, sodium ascorbate and THPTA was added. The mixture was kept reacting for 2 days, at 30 °C. Then, the solvent was evaporated, and the crude was directly characterized by MALDI-TOF or ESI-MS Orbitrap.

5.5 References

1. You, L., Zha, D. & Anslyn, E. V. Recent Advances in Supramolecular Analytical Chemistry Using Optical Sensing. *Chem. Rev.* **115**, 7840–7892 (2015).

2. Lemke, E. A. & Schultz, C. Principles for designing fluorescent sensors and reporters. *Nat. Chem. Biol.* **7**, 480–483 (2011).

3. Granadero, D., Bordello, J., Pérez-Alvite, M. J., Novo, M. & Al-Soufi, W. Host-guest complexation studied by fluorescence correlation spectroscopy: Adamantane-cyclodextrin inclusion. *Int. J. Mol. Sci.* **11**, 173–188 (2010).

4. Pinalli, R. & Dalcanale, E. Supramolecular sensing with phosphonate cavitands. *Acc. Chem. Res.* **46**, 399–411 (2013).

5. Biavardi, E. *et al.* Fully reversible guest exchange in tetraphosphonate cavitand complexes probed by fluorescence spectroscopy. *Chem. Commun.* 1638–1640 (2008).

6. Masseroni, D. *et al.* A fluorescent probe for ecstasy. *Chem. Commun.* **51**, 12799–12802 (2015).

7. Chen, J. *et al.* Selective discrimination and classification of G-quadruplex structures with a host–guest sensing array. *Nat. Chem.* **13**, 488–495 (2021).

8. Gill, A. D. *et al.* Sensing of citrulline modifications in histone peptides by deep cavitand hosts. *Chem. Commun.* **55**, 13259–13262 (2019).

9. Kaur, H. & Yung, L. Y. L. Probing high affinity sequences of DNA aptamer against VEGF 165. *PLoS One* **7**, 19–26 (2012).

10. Jewett, J. C. & Bertozzi, C. R. Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev.* **39**, 1272–1279 (2010).