UNIVERSITAT DE BARCELONA

Pseudopeptidic cages for biological applications

Lucía Tapia Pérez

UNIVERSITAT DE BARCELONA

DEPARTAMENT DE QUÍMICA ORGÀNICA

FACULTAT DE QUÍMICA
UNIVERSITAT DE BARCELONA

Programa de doctorat de Química Orgànica
Barcelona, 2021

Pseudopeptidic cages for biological applications

Memòria presentada per Lucía Tapia Pérez per optar al títol de doctor per la universitat de Barcelona.
Tesis realitzada al departament de Química Biològica del l'institut de Química Avançada de Catalunya (IQAC-CSIC)

Directors:
Dr Ignacio Alfonso Rodríguez
Dr. Jordi Solà Oller
Departament de Química Biològica
(IQAC-CSIC)

Tutor:
Dr. Ernesto Nicolás Galindo

Departament de Química Orgànica
Facultat de Química (UB)

Doctorand:

Lucía Tapia Pérez
Departament de Química Biològica
(IQAC-CSIC)

The journey of a thousand miles

beguins with one step

Lao Tzu

ACKNOWLEDGEMENTS

En primer lugar quiero agradecer a mis directores de tesis el Dr. Ignacio Alfonso y el Dr. Jordi Solà por brindarme la oportunidad de realizar la tesis bajo su dirección. Gracias por acompañarme, guiarme, animarme, motivarme, enseñarme y ayudarme durante estos cuatro años. Y sobre todo gracias por contagiarme la pasión por lo que hacéis.

A mis dos primeras compis de doctorado Anna y María, con vosotras al lado haría una y mil tesis más. A Dan, a Ciril y a cada una de las personas con las que he coincidido en el lab.309 que han conseguido que mire a la química con una sonrisa.

A Yolanda y Naiara por su inmensurable ayuda con la resonancia, al grupo de RUBAN por introducirme en el mundo de los cultivos celulares y a todas aquellas personas del IQAC que me han ayudado en todo lo que ha estado en su mano.

Al Dr. Roberto Quesada, por acogerme en su laboratorio y a todos los miembros de su grupo en la universidad de Burgos: no hay mejor manera de aprender técnicas nuevas.

To Dr. Miller for accepting me in his lab, and every single person I had the pleasure to meet at The miller Lab, sharing some time with you was an amazing experience.

Y sobre todo a mi familia. A mis padres por su apoyo incondicional y por animarme y creer en mi siempre. Y por supuesto a mi hermana, cualquier cosa parece más sencilla con alguien como tú al lado.

LIST OF ABBREVIATIONS

CF	Carboxyfluorescein
aa	amino acid
ACN	acetonitrile
ATP	adenosine triphosphate
Boc	*tert*-butyloxicarbonyl
BSA	bovine serum albumin protein
CB	Cucurbiturils
cmc	critical micellar concentration
COSY	Correlated spectroscopy
DCC	Diciclohexylcarbodiimide
DCM	Dichloromethane
DIPEA	N,N-diisopropylethylamine
DMEM	Dulbecco's Modified Eagle Medium
DMF	*N,N*-dimethylformamide
DMSO	Dimethyl sulfoxide
DOSY	Diffusion ordered spectroscopy
DPC	dodecylphosphorylcholine
EC_{50}	Half maximum effective concentration
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EGFR	epidermal growth factor receptor
ESI	Electrosprai ionization
FBS	fetal bovine serum
FDA	Food and drug administration
Fi	Fluorescence intensity
FMOC	Fluorenylmethoxycarbonyl
GPCR	G protein coupled receptors
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HMBC	heteronuclear multiple bond correlation
HoBt	1-hydroxybenxotriazole
HPLC	High performance liquid chromatography
HPTS	8-Hydroxypyrene-1,3,6-trisulfonic acid
HSQC	Heteronuclear single quantum correlation
IGFR	Insulin grown factor receptor
ISE	Ion selective electrode

K_{ex}	Exchange constant
K_{SV}	Stern-volmer constant
K_{SV}	Stern-Volmer constant
LUV	Large unilamelar vesicles
MeOH	Methanol
MS	Mass spectrometry
MT	Molecular tweezer
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NMR	Nuclear magnetic resonance
NRTK	non receptor tyrosin kinase
pH_e	Extracellular pH
pH_i	Intracellular pH
PIPES	piperazine-*N,N*'-bis(2-ethanesulfonic acid)
POPC	1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine
QSAR	Quantitative structure-activity relationship
RP-HPLC	Reverse phase HPLC.
SEC	Size exclusioin chromatography
SN_2	Bimilecular nucleophilic substitution
S-V	Stern-Volmer
TBACl	Tetrabutylamonium chloride
TES	Triethyl silane
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethanesulfonic acid
TK	tyrosin kinase
TLC	Thin laye chromatography
TOF	time of flight
TRIS	Tris(hydroxymethyl)aminomethane
trt	trityl
UPLC	ultra performance liquid chromatography
UV-vis	Ultraviolet visible
WASP	Wiskott-Aldrich sindrome protein
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength

INDEX

CHAPTER 1. GENERAL INTRODUCTION: SUPRAMOLECULAR CHEMISTRY

1.1 SUPRAMOLECULAR CHEMISTRY

Supramolecular chemistry is a field of chemistry that studies the interactions between molecules. It is focused on weak, non-covalent and reversible interactions which are mainly π- π, Van del Walls, hydrogen bonds, coulombic, dipole-dipole and hydrophobic interactions. Jean Marie Lehn, one of the fathers of this discipline, referred to supramolecular chemistry using the famous expression "chemistry beyond the molecule".

In biological systems there are thousands of examples of processes regulated by non-covalent interactions such as protein folding, protein-protein interactions, enzyme substrate recognition, cell membrane formation, replication, transcription, DNA folding, etc For years, supramolecular chemists have been focused on the design, preparation and study of small host-guest complexes, aiming to someday reach this complexity found in nature. The common approach to that is to gradually increase size and complexity of supramolecular systems to finally achieve the desired folding and functionality. Using Nature as source of inspiration, much progress has been achieved in the design of molecules with bio-like properties since 1987 when Lehn, Pedersen and Cram won the novel price in chemistry for their development and use of molecules with structure specific interactions of high selectivity.

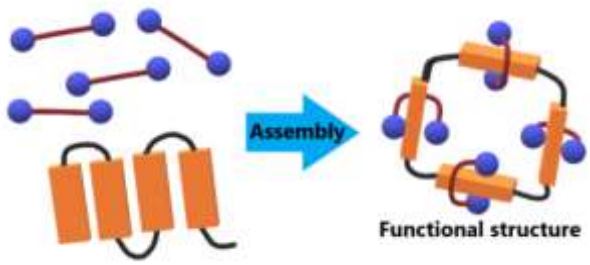


Figure 1.1. Cartoon representation of the self-sorting of different supramolecular elements to form a more complex and functional structure.

This evolution can be seen in the 2016 Nobel price to Sauvage, Stoddart and Feringa for their work on (supra)molecular machines.[1,2] Nowadays the synthesis of complex architectures with diverse functionalities (mimicking macroscopic biological systems) is possible and these new systems are being applied not only in biology but also in other fields such as pharmacology, catalysis and material chemistry.

Besides the direct use of supramolecular chemistry as a tool to control biological processes, it is also a suitable tool for monitoring them and understand how they function and to amend them when they do not work properly. Receptors used for this purpose are known as biosensors and have been identified as potential tools in medicine, agriculture industry or environmental monitoring.[3] .[4]

"A supramolecular element can be defined as an entity made of a molecular building block that folds or assembles through non covalent interactions into a well ordered structure."[5] Synthetic supramolecular systems were initially composed of simple building blocks but they have rapidly evolved to form complex 3D structures with specific functionality inaccessible through classic synthetic methods. Molecules or building blocks self-sorting processes can be active (formation of species out of the equilibrium) or passive (according to thermal equilibrium and association constants) which uses non covalent interactions. The major goal of supramolecular chemistry is to efficiently predict and tune the outcome of these processes.

Regardless of the difference in complexity between synthetic supramolecular assemblies and biomolecules (like enzymes or DNA), they share some important properties such as reversibility, topology or dynamics.

By combining both natural and synthetic supramolecular elements in a smart way, chemists are everyday closer to control and influence biological processes of high interest such as the photosynthesis or the function of enzymes. This will eventually lead to the design of new pharmacological or medicinal treatments based on supramolecular chemistry principles.

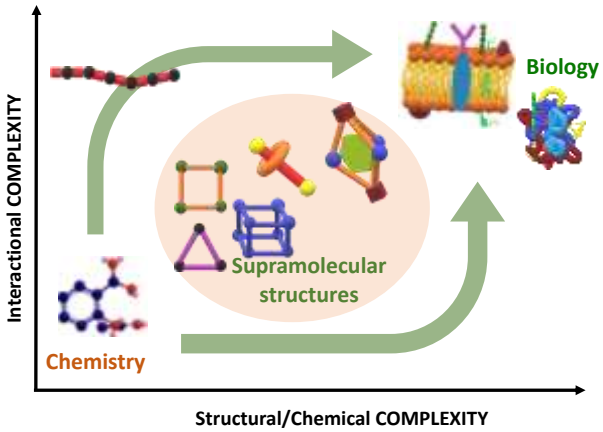


Figure 1.2. Schematic representation of the relation between interactional and structural complexity from simple molecules to complex biological structures.

While supramolecular chemistry is focused on molecular interactions, systems chemistry goes one step further aiming to study and exploit these interactions to design specific assemblies with emergent properties.[6] Systems chemistry is a field of chemistry focused on the development of complex molecular systems with properties above the sum of their components.

The design of macromolecular building blocks with specific functionality is challenging not only for their chemical complexity but because it also requires the development of analytical techniques for their characterization and the study of their functions. Moreover, getting supramolecular systems to work in water (the main component of biological mediums) is still a challenging goal. So far, a lot of progress has been done in this direction, from the design and preparation of micelles or vesicles to the self-assembly of some systems using water as the main solvent.

1.2 Dynamic combinatorial chemistry (DCC)

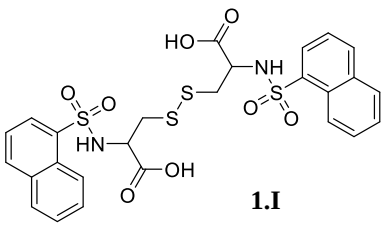
Dynamic chemistry is a promising method for the simultaneous formation of different bonds to prepare complex molecules. The principle is that self-organization processes are able to select the right molecule (either with kinetic or thermodynamic control) from a well-defined mixture of building blocks. However, when the desired products are formed, the purification of this mixture is often the main difficulty.[7] A Dynamic combinatorial Library (DCL) is generated by mixing several building blocks. In the absence of external stimuli the

binding blocks undergo reversible reactions until thermodynamic equilibrium is reached. However, the presence of selected guests that stabilize some components of the DCL changes the species distribution in the equilibrium leading to an amplification of ligands with higher affinity for the guest.[8] A DCC can be used to identify structures with unusual binding properties, potential ligands for relevant biological targets and also to provide simple routes for the synthesis of complex molecules.[8,9] A recent work reported by the supramolecular chemistry group managed to use live cells as templates for a DCL[10] broadening the potential of DCC and the use of DCLs in the fields of medicine, pharmaceuticals, smart materials and catalysis.

1.3 SUPRAMOLECULAR CHEMISTRY IN WATER

Self-assembly of complex structures in water is mainly driven by hydrophobic effects, but some other forces intervene like hydrogen bonding, electrostatic interactions and π-stacking. All these forces together with the role of proximity and specificity are used for the design and study of supramolecular systems. Furthermore, the presence of other salts or counter ions in the solution, as well as the temperature, might considerably affect the self-assembly allowing or preventing it to take place in the desired way. In biological systems water is also present and takes part in common interactions.[11] For instance in protein-ligand interactions water can affect both the binding pocket and the ligand. In synthetic systems water may also have an impact with solubility issues.

The use of self-sorting processes and DCC are also gaining importance in water based media. To date, several artificial receptors have been designed for the selective binding of a number of guests in aqueous media.[11] This is the case of the compounds reported by Milanesi et al (**1.I**) who by using DCC found a bidentate ligand that binds CaM (Calcium transducer calmodulin) protein with a K_d = 10 μM in aqueous media (Figure 1.3).[12]



1.I

Figure 1.3. Chemical structure of compound 1.I reported by Milanesi et al.

Of great interest for supramolecular chemist is anion recognition in water for its potential clinical applications. For instance Fujita, Luppert and Navarro among others used bowl-shape molecules for this purpose.[11] The different approaches for the design of ion hosts together with their many applications and the most relevant examples will be presented in chapter 2.

A supramolecular system designed for the study or control of biological processes has be water soluble, able to interact in biological media with strengths in the range of biomolecules (μM-pM) and contain a high selectivity for the guest biomolecules. Using selective hosts for amino acids or proteins able to efficiently work in aqueous media has arisen as a powerful tool for the study and control of many relevant biological processes. Details and examples of this strategy will be discussed in chapter 4.

Building more complex supramolecular structures able to be assembled in water has gained interest recently. Hence new strategies to get proper systems for aqueous solutions are constantly being explored.

1.4 SUPRAMOLECULAR CHEMISTRY IN BIOLOGICAL SYSTEMS

Many biological systems are a combination of building blocks to form larger structures that ultimately lead to a complex living organisms. All the major processes in life are somehow ruled by supramolecular interactions present in proteins, sugars, lipids and nucleotides.

Proteins

There are 20 different proteinogenic amino acids present in living organisms that combine to form proteins. Amino acids are covalently bound and their order determines the primary structure of the protein. This protein can then fold mainly forming α-helix or β-sheet subunits depending on the non-covalent interactions determined by the amino acid side chains, and this is the so called secondary structure of a protein. Upon folding, the secondary structure forms a 3D structure responsible for protein functionality. Quaternary structure is achieved when several proteins aggregate to generate a more complex functional structure like, for instance, an ion channel.

The first attempt to describe enzyme-substrate interactions was in 1895 when Fischer proposed the “lock and key hypothesis”. This hypothesis was further extended by Koshland who included the induced fit mechanism.[13]

Saccharides

Sugars are a group of carbohydrate molecules that often form 5 o 6 ring cyclic structures. These structures contain several chiral centres which are an additional complexity in the field of sugar chemistry.

When a saccharide is covalently attached to an amino-acid side chain of a protein it forms a glycopeptide or a glycoprotein. This binding is generally through threonine or asparagine. Glycoproteins affect in different ways to the proteins properties. They alter the folding, change the hydrophobicity and often increase the specificity in molecular recognition processes.[14]

Lipids

Lipids also interact to form complex structures through supramolecular interactions. The most well-known example is the formation of cell membranes,[15] phospholipidic bilayers responsible of cell compartmentalization and separation of cells from their environment. The fluidity of this structure allows the diffusion of small molecules through it and the presence of several biomolecules mainly proteins responsible for cell function. The great relevance of cell membranes in life has led to the use of artificial lipidic systems called vesicles (spherical lipid bilayers that separate their enclosed volume from the environment) mimic their properties for in vitro studies.[16]

Nucleic acids (RNA and DNA)

RNA and DNA are polymers composed of nucleosides (a monosaccharide and a nucleobase) joined together by a phosphate ester. They pack forming and helix due to the enforced twist in the phosphate ester backbone which exposes the nucleobases and allows/induced the formation of hydrogen bonds between the different bases in the α-helix. It is worth mentioning that DNA self-replication and transduction which are the key of live as we know it are processes governed by supramolecular interactions.

1.5 Applications in diagnosis and therapeutics

Supramolecular chemistry is becoming a promising tool both for diagnosis and therapeutic purposes.[17]

Scientific literature is now full of chemical sensors but very few have become a practical device.[4] In the field of imaging, magnetic resonance (MRI) has become one of the most powerful tools for diagnosis. It requires the use of contrast agents which are generally gadolinium complexes with strong non-covalent binding properties and whose function is to increase MRI signal.

To date, the greatest success of supramolecular chemistry in therapeutics is the monoclonal antibody therapy. It relays in the fact that antibodies interact with their targeted antigens using weak forces and these specific interactions can act by blocking receptors or altering biochemical pathways.

Supramolecular interactions are also the basis of molecules used in chelation therapy to correct imbalance of certain ions which bind the targeted metal and efficiently transport to the desired place or to excretion.

Macrocycles are currently being used for radiotherapy like the case of DOTA or as antibiotics mimicking natural ones like monensin or valiomicin by piercing the cell wall. The most famous example is the glycopeptide vancomycin, discovered by the company Eli Lilly and clinically used since 1959. Vancomycin is considered to be one of the last resources in hospitals against multi-resistant bacteria. Vancomycin binds the cell wall mucopeptide precursor in the terminal section of its sequence which is L-Lys-D-Alan-D-Ala.[18]

Although impressive progress is constantly achieved in controlling supramolecular interactions with biological systems, biomolecules form structures way easier than their synthetic competitors and closing this gaps is one of the most ambitious challenges of supramolecular chemistry. In fact, life is far from equilibrium and biological functionalities are only achieved by kinetically-driven processes which are constantly dissipating energy so that cells can function. [19] It is in far-from equilibrium systems where the most complex functions can be achieved: that is why supramolecular chemistry is currently evolving to kinetically controlled systems.[19]

1.6 SUPRAMOLECULAR STRUCTURES: CAGES

Macrocyclic structures have been widely studied in supramolecular chemistry for their potential Natural examples of macrocyclic molecules are for instance cyclodextrins, cyclic oligosaccharides glucose subunits α-1,4-linked with hydrophobic cavities (**1.V**). Typical

cyclodextrins have 6 (α-CD), 7 (β-CD) or 8 (γ-CD) glucose monomers in a ring creating a conic 3D structure. They are able to form host-guest complexes with hydrophobic molecules and have several uses in food, pharmaceutical and cosmetic industries. A relevant synthetic example of macrocycles are cucurbiturils (**1.VIII**), compounds with binding affinities for different types of molecules in water.[20] Examples of cucurbiturils and their applications are further explained in chapter 4.

Adding an extra level of complexity to macrocyclic structures we find molecular cage compounds which are macrobicycles with a closed 3D cavity designed for the encapsulation of guest molecules. The chemical preparation of cages has to face many synthetic goals but they are a very promising tool for many applications such as delivery, purification or as nano-reactors. Initially, the synthesis of these complex structures was through multiple steps routes but now there is a growing tendency to use DCC which is more efficient, faster and ideal for screening.

The first example of a synthetic cage is from the 60's when Eaton et al. prepared the first ever described cubane (**1.II**).[21] In 1967 Charles Peterson published the first papers on Crown Ethers (**1.III**)[22] which served as stimuli for Jean Marie Lehn, who started his research on cryptans.[23] Cryptans (**1.IV**) are probably the first example of cage-like molecules studied for the recognition of cations. Two examples are the [2.2.2]cryptand that has a very high selectivity for K^+ [24] and a binding constant of $10^{5.8}$ M^{-1} and the [2.2.1]cryptand whith a higher affinity for Na^+.[25] In 1984 Frits Vogtle's group succeeded in the synthesis of the first artificial shape persistent cage with a Fe^{3+} complexation capacity higher than EDTA (**1.VII**).[26] The same year Cram reported a cage compound that he called carcerand (**1.VI**).[27] In fact Cram and Co-workers were the first to use imine chemistry for the synthesis of a cage in high yields. In the 1990s, Fujita and Stang described the assembly of metallocycles and metallocages using metal-ligand coordination.[28] So far, the largest artificial cage ever prepared was reported by Fujita et al.[29] This cage was self-assembled in water and it is almost a spherical shell composed of 144 components (including metal ions in its structure).

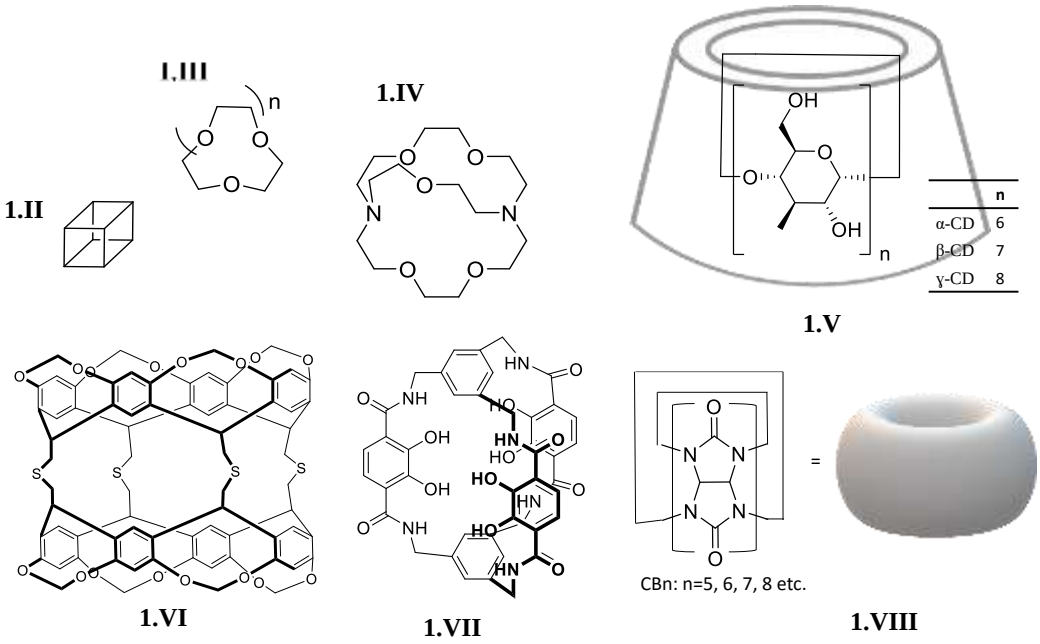


Figure 1.4. Chemical structure of compounds 1-5, Cyclodextrins and Cbn.

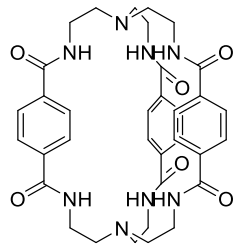
Although the majority of the currently described cage compounds contain metals, of special interest for their potential use in medicine are fully organic cages. In nature the main example of defined cavities is found in enzymes which control many biological processes. Enzymes have been an inspiration for chemists in the search and design of defined cage compounds for different purposes like the stabilization of reactive species, catalysis or sensing.

The main advantage of purely organic cages is that they are usually neutral and possess higher stability since they are formed by covalent bonds. There are mainly two approaches for the synthesis of shape persistent organic cage. The first is by using irreversible reactions, which render the compounds with high stability but usually in very low yields. The second one is using DCC which has emerged as the preferably tool to obtain cage-like compounds in reasonable yields from simple building blocks.

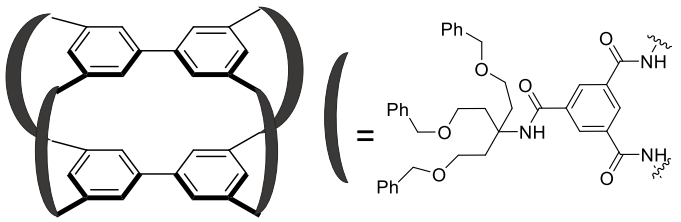
1.6.1 Cages synthetized by irreversible bond formation

There are few examples of cages formed by irreversible bonds with reasonable yields in literature. Some of these compounds are prepared through the formation of amide bond like the siderophore (**1.VI**) of Vogtle (13 % yield in the cyclization step).[26] Raymod's group

later reported the synthesis of a compound similar to the siderophore **1.VI** but using Fe^{3+} as a template (**1.IX**) increasing the yield to 70%.[30] There are also examples of cages synthetized without templates with reasonable yields. For instance, A. P. Davis' group used two consecutive [2+2] cyclizations tosynthetize cage **1.X** in acceptable yields at hight dilution (62% yield in the last cyclization step). Besides, Davis investigated those compounds as monosaccharide receptors. [31]



1.IX

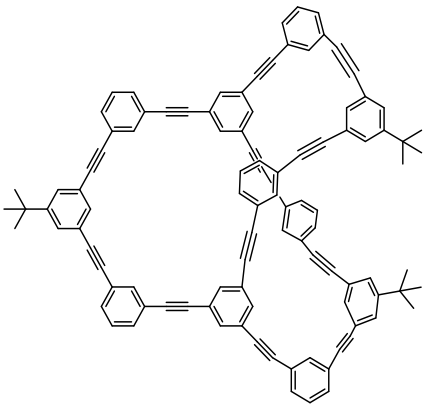


1.X

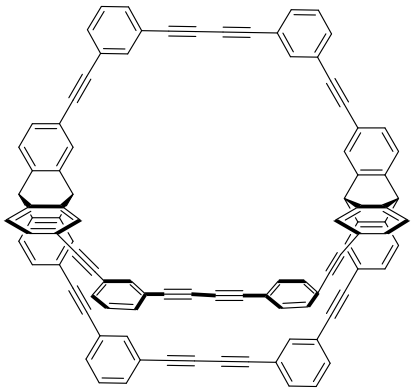
Figure 1.5. Chemical structure of compounds 1.IX and 1.X.

Nucleophilic aromatic substitution is another attractive method for the formation of covalent bonds for the synthesis of macrocycles,[32] with high yields like the ones described by Katz and c-workers[33] and allows the use of ion templates used for instance by I. Martí et al. for the synthesis of anion receptor cages.[34]

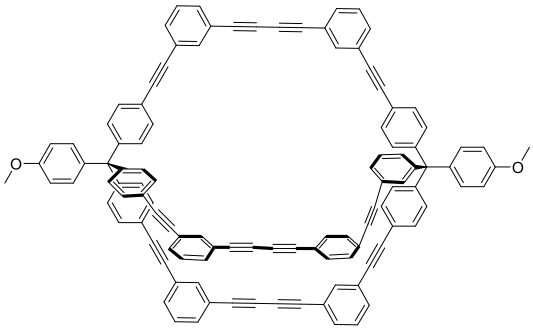
The first cage with a fully carbon framework was reported in 1977 by Wennerström.[35] More recent examples are compounds **1.XI** synthetized by Moore and co-workers using a Sonogahira-Hagihara approach (32% yield), and compounds **1.XII** and **1.XIII** reported by Chen et. Al. and Doonan et al. who used the Glaser-Eglington coupling reaction[31] (Figure 1.6).



1.XI



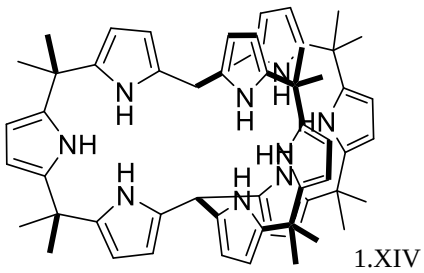
1.XII



1.XIII

Figure 1.6. Compounds 1.XI (synthetized in a multistep approach via Sonogashira Hagihara cross-coupling reactions) 1.XII and 1.XIII synthetized via Glaser-Eglington coupling.

One last example of cages synthetized via irreversible C-C bond formation can be the macrotricycle nanopyrrole 1.XIV from Sessler et al.[36] These compounds are attractive as host molecules because of the pyrrolic NH functionality. NH groups can serve as potential binding sites acting as H-bond donors or acceptors. In the case of molecules **1.XIV**, studied for its ion binding properties the NH group is part of an anion binding motif (fig 1. 7).



1.XIV

Figure 1.7. Chemical structure of compounds 1.XIV.

1.6.2 Cages synthetized via dynamic covalent bond formation

Regardless of the examples presented in the previous section, cage compounds synthetized with irreversible bonds are generally only accessible in very low yields.

The alternative strategy and the most popular nowadays is the synthesis of macrocycles via dynamic covalent chemistry (DCC). It involves making and breaking reversible bonds until the thermodynamically stable structures are formed. This strategy allows the preparation of complex structures from relatively simple building blocks in one pot and with relatively high yields.[37]

One of the most versatile reactions used in DCC is the imine bond formation. The formation of an imine from an aldehyde and an amine is a reversible reaction which when reaching the equilibrium forms the thermodynamically most stable products. This reaction is widely used in DCC and it is also used for the synthesis of large symmetric molecules starting from simple precursors. It is a well-known reaction in organic chemistry which involves the loss of a water molecule. It can be inter or intramolecular and many factors can influence the formation of the thermodynamic species such as concentration, pH or temperature. Steric and electronic factors can also affect the equilibrium, which once formed can be broken by adding water to the solution. The reversible imine exchange reaction can be frozen by reducing the imine bond to a secondary amine.[38]

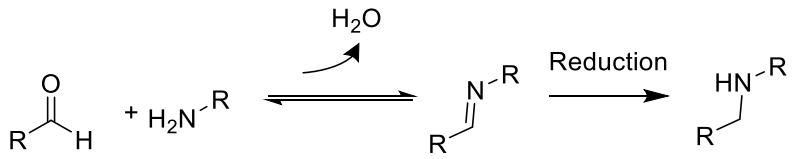


Figure 1.8. Imine reaction and reduction.

Examples of cages synthetized via imine reactions are for instance the tetrahedral cages 4+6 reported by Cooper and co-workers (1.XV). [39]

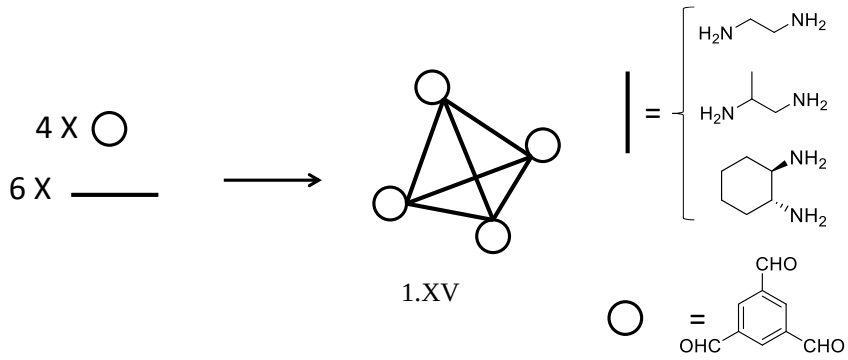


Figure 1.9. Synthesis of tetrahedral cages reported by Cooper.

Later the same precursors were successfully combined to form 12+8 structures. Moreover, it was observed that in certain conditions, this reaction leads to the formation of a catenane [40](Figure 1.10).

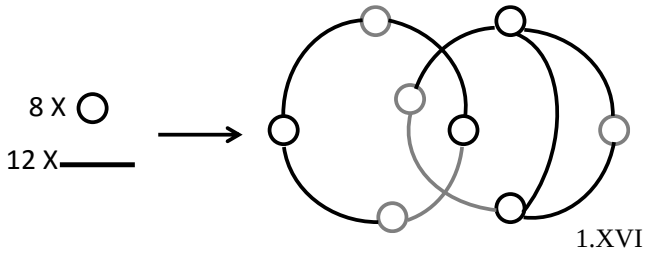


Figure 1.10. Synthesis of a catenane.

Also using the imine reaction for the formation of macrocycles, Gawronski's[41] group reported the formation of a 12+8 cage (**1X.XVII**).

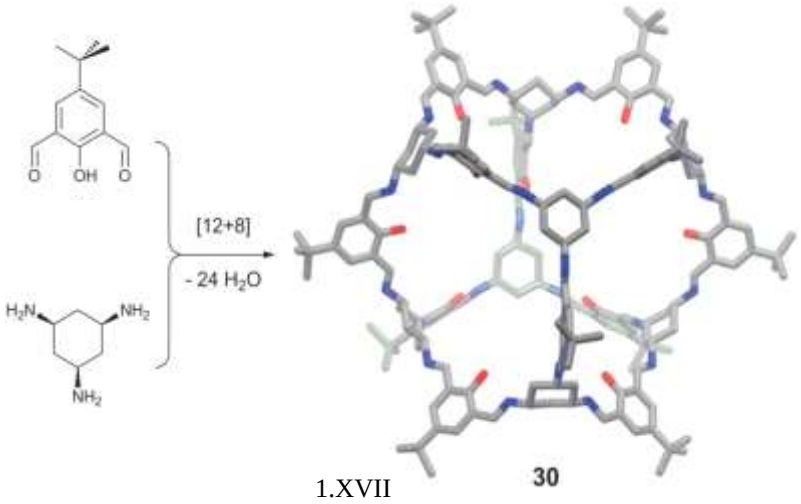


Figure 1.11. Synthesis of an octahedral [12+8] cage (figure adapted form referende[41]).

The use of templates for the synthesis of cages via DCC can be very useful to control the outcome of macrocyclization reactions to obtain the desired products in higher yields. This is seen for instance in the work of Wang and co-workers where they reported the preparation of several porphyrin-based cages via one-pot template directed imine condensation with different size and very high yields (89-quantitative, Figure 1.12). [42,43]

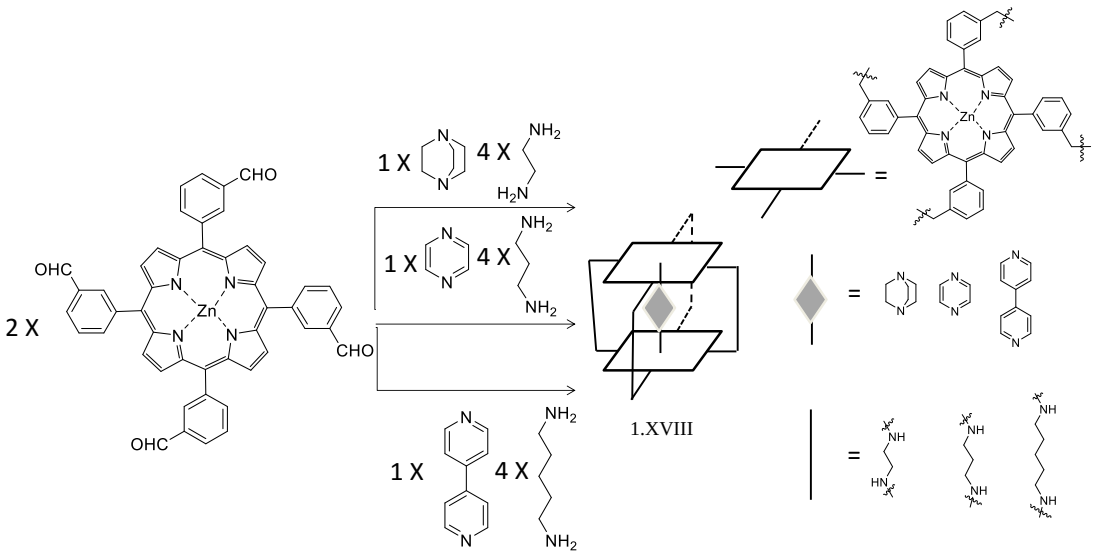
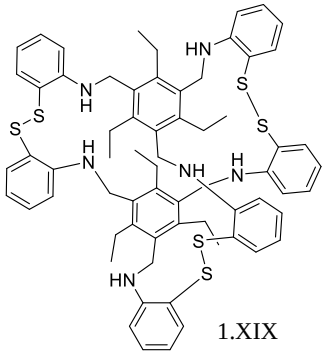


Figure 1.12. Synthesis of template directed cages by Wang.

Over the years, more complex and larger cages with specific properties are being prepared. Some of them even include interlocked systems.[31] For the preparation of larger cages with specific functionality, extended building blocks or multicomponent assemblies are required.

Apart from the imine reaction, other reversible reactions used for the synthesis of cages through DCC are hyadrazone condensation, boronic esters condensation, boronic acid condensation, alkyne methatesis,[44] alkene metathesis, disulphide exchange,[45] dynamic boronic ester[46] or phenyl/aldehyde condensation.[31]



1.XIX

Figure 1.13. Supramolecular cage formed by reversible disulfide bonds reported by Horng.[47]

The use of coordination metals for the synthesis of cage-like structures has also been widely explored by several authors like Raymond, Clever, etc.[48] For example Severin reported a large cage [49] containing Ruthenium ions (1.XX).

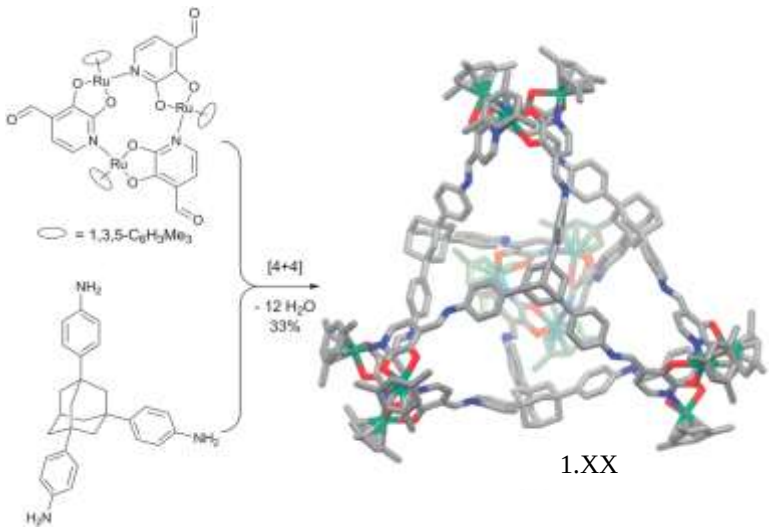


Figure 1.14. Synthesis of a [4+4] cage by connection of metallamacrocycles via imine bonds.Figure adapted from reference [49].

The understanding of the geometrical needs of the different building blocks would allow the synthesis of even more complex structures in a more efficient and predictable manner. Thus, usually, a small change in the shape or the size of a building block might lead to completely different cages. Moreover, the proper selection of the solvent has a huge influence in the resulting products of a chemical library generated by dynamic covalent bonds.[31]

1.6.3 Properties and applications of cages

Organic cages with a well-designed cavity (size, and shape) can be an ideal host for biologically relevant molecules and thus be interesting for many applications from research to medical treatments.[50] This PhD thesis puts the focus of the use of synthetic cages for the selective recognition of ions and biomolecules which are in fact a major research topic in supramolecular chemistry. A more detailed introduction regarding these topics can be found in chapters 2 and 4.

Cage like compounds are also used to stabilize reactive species that are protected from the environment and remain stable for longer periods of time when included inside the cage. Some other potential applications of these compounds are in the field of catalysis, material

science,[31] sensing of pollutants or even as templates for the synthesis of metal nanoparticles.[51]

A challenging limitation of synthetic cages is the size of the cavity because the synthetic complexity bursts when trying to make cages with bigger cavities to host biomolecules like proteins. The current solution to that is the use of multicomponent assembly. Fujita's group has successfully succeeded in the synthesis of protein metalorganic composites with protein binding affinity, whose potential applications are the stabilization of proteins, alteration of their function or the facilitation of protein crystallizations.[52]

1.7 Study of interactions in supramolecular chemistry: Determination of association constants

Interactions in supramolecular chemistry are often quantified using a titration (either by adding guest to a solution of the host or host to a guest solution) monitoring a change in a physical property sensitive to the interaction through NMR, fluorescence or any other analytical technique. The resulting data generally gives a lot of information about the interaction like the stoichiometry or the association constants.

To simplify the calculations, the concentration of one of the compounds of the titration is kept constant to only have one variable in the data. Changes in the monitored property can be plotted as a function of the concentration of the added species. Then, the appropriate mathematical model has to be selected for the fitting of the plotted isotherm.

Sometimes determining the stoichiometry of a system becomes a tricky stem in the analysis of the data. When we suspect the participation of complex stoichiometries, data is fitted to all the possible models and the results are carefully compared to accurately decide which model better explain the process. In some cases, titrations at different concentration ranges are necessary to better define the actual stoichiometry.

1.7.1 NMR titrations

Monitoring a titration with NMR spectroscopy allows to obtain quantitative info, shift changes, symmetry, host-guest interactions and stoichiometry. Although the classic disadvantage of this technique is the low sensitivity, modern equipment's are suitable for working at sub-millimolar concentrations with association constants in the range of up to $10^6\ M^{-1}$. The information derived from an NMR experiment is also affected by the exchange

rate of the monitored system which is sometimes a limiting factor for determining association constants with this technique.

1.7.2 Fluorescence titrations

Fluorescence is a highly sensitive technique (μM and sometimes nM), measurements can be made with very low concentrations and high association constants can be determined (up to $10^{9\text{-}10}$). The ideal situation is the one where fluorescence changes are the result of the complexation process only. However, different mechanisms affecting fluorescence must be considered when they are present. The two main types of quenching are dynamic (collisional) and static quenching (associated to supramolecular interactions) and these processes have to be differentiated form the supramolecular binding that is being studied.[53]

Dynamic collisional quenching occurs as a result of the excited state of the fluorophore deactivated by another molecule present in the solution called quencher. Collisional quenching can decrease fluorescence intensity as described by the Stern-Volmer equation, which is a plot of the F_0/F as a function of the concentration of the quencher. A typical 1:1 dynamic quenching fits to the Stern-Volmer equation in which the Stern-Volmer constant indicates the sensitivity of a fluorophore to a quencher.

$$\frac{F_0}{F} = 1 + K_{SV} \times [Q]$$

Stern-Volmer equation. $K_{S\text{-}V}$: Stern-Volmer quenching constant, [Q]: quencher concentration.

Sometimes the quenching observed is a combination of both static and dynamic quenching and this makes the analysis more challenging. Besides the formation of a complex with the fluorophore often alters the fluorescence intensity (or the shape of the emission spectrum) by modifying its environment, attenuating the incident light, etc.

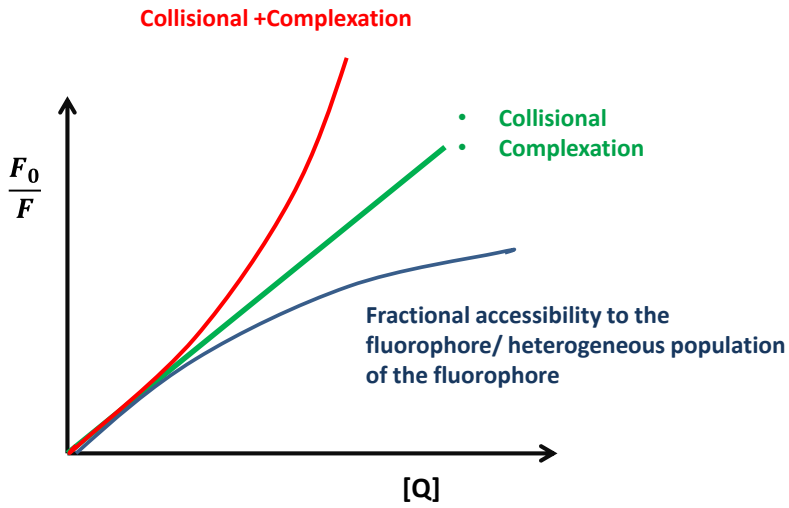


Figure 1.15. Typical deviation form linearity in a S-V plot.

Hence the most desirable situation in a fluorescence titration is the formation of a complex emitting at a different wavelength so that there is a new growing fluorescence band that can be used for the analysis. Data analysis and interpretation is also a crucial step in every titration. The most common strategy for data analysis nowadays is global analysis, possible with modern computers which adjust many observables at the same time without approximations or simplifications of the model.

The software used for the fitting of the different titrations conducted in this thesis were HypSpec[54] and HypNMR which are based on multicomponent analysis and are one of the most accurate methods available for the fitting of titration data.

1.7.3 Determination of association constants: *K*, log *K*, *β*, Log *β*, K_d K_{ass}.

Hypspec software determines the *β* (equilibrium constant for the formation of a complex in solution). Since the formation of complexes 1:1 and 1:2 occur via a step to step addition $K_{1:1}$ and $K_{1:2}$ are the equilibrium constants for each step (they indicate the extent of the formation of a specie without including information about the previous steps). In multiple steps equilibrium *β* is a cumulative constants includes that information.

$$H + G \leftrightarrows HG \quad K_{1:1} \quad \beta_1$$

$$HG + G \leftrightarrows HG_2 \quad K_{1:2}$$

$$H + 2G \leftrightarrows HG_2 \quad \beta_2$$

From the different log β determined with the fitting the $K_{1:1}$ and $K_{1:2}$ can be calculated which are what we consider the K_{ass} for each complex. The K_d is then directly determined as the inverse od the K_{ass}. For simple equilibrium where only a 1 : 1 complex is being formed log K = log β.

1.7.4 BC_{50} and BC_{50}^0

For hosts whose interactions with one guest can generate more than one adduct of different stoichiometry, all the complexes contribute to the binding ability of the host towards the ligand so none of the association constants itself will properly describe the host-guest affinity.

In order to properly quantify the affinities in those systems, especially if we want to compare them with other interaction pairs, there is a need to define a new parameter related to binding, which considers the quantity of each reagent contributing to the binding. The BC_{50} descriptor was defined[55] by S. Roelens and co-workers to fulfil this requirement. It is a parameter that allows a quantitative assessment of the affinity from the knowledge of the different association constants contributing to the binding. BC_{50} can be used for systems containing any number of complexes and to compare different binding systems fitted to different models under defined parity conditions. Those properties make BC_{50} a very useful parameter for comparative purposes.

BC50 is described as the total concentration of ligand ([L], mol/L) necessary for the binding of 50 % of receptor ([R]), thus a higher affinity implies a lower BC_{50}.

When the concertation of the analyte becomes negligible, BC_{50} becomes constant

$$\lim_{analyte \to 0} BC_{50} = cnt = BC_{50}^{0}$$

BC_{50}^0, often referred to as the intrinsic media binding concentrations, depends on all the equilibrium constants present in the system regardless of their concentration. In the most simple equilibrium situation where only the 1:1 complex is formed, BC_{50} is equal to the dissociation constant, K_d.

A software called "BC_{50} calculator" was developed by S. Roelens and co-workers. It calculates the BC_{50}° of any equilibrium system form the association constants of the different complexes involved in the interaction.[55]

BC_{50}° is a property that has been extensively used in this project to compare the affinity of different compounds for the guest when more than one equilibrium contributed to the interaction. In the different host-guest systems studied throughout this thesis "BC_{50} calculator" was used to calculate BC_{50}° form the association constants derived from the fittings of the titration data of the different systems.

1.8 INTRODUCTION TO THE SPECIFIC WORK OF THIS THESIS

To sum up, supramolecular chemistry is a potent approach to the study and control of biological systems, while synthetic organic cages are well-defined structures that can be ideal hosts for ions or biomolecules. Therefore, synthetic organic cages are useful tools for the study or control of biological systems.

The supramolecular chemistry group in IQAC-CSIC has a long experience in this field. For instance, the group reported the binding of Cu(II) ions to a family of open chain and macrocyclic pseudopeptides, and the ability of the obtained metal complexes to recognize different biologically important dicarboxylates, or even to transfer the Cu(II) ion to the biomolecules.[56]

Moreover, during the last years the group has been working in other pseudopeptidic receptors for different species, such as organic anions,[57] N-protected amino acids[58] and dipeptides.[59] Following an optimization process, we prepared pseudopeptidic cages of different sizes,[34,32] which have shown to bind challenging guests in very competitive media, including aqueous environments. For instance, small tripodal pseudopeptidic cages are able to selectively bind the chloride anion.in aqueous acetonitrile.[60] More interestingly, these molecules transport the chloride anion through lipid bilayers as models of cell membranes, with an interesting structure-activity relationship.[34]

On the other hand, the supramolecular chemistry group has also experience in the synthesis of larger pseudopeptidic cages using a multicomponent reductive amination reaction.[32] These cages recognize N-protected dipeptides in different media, showing a good selectivity toward dipeptides bearing an aromatic residue in the C-terminus.[61] One of the cages selectively interacts with the Ac-EY-OH dipeptide,[62] which is a model of the target sequence for tyrosine kinases. More recently, this interaction was proved to be very selective and can be studied in aqueous medium.[63] Considering the key role of kinases in many biological processes related to normal growth and metabolism, the molecular recognition properties of these cages allows envisioning potential applications of these hosts as modulators of kinase activity.[63]

The aim of this PhD thesis is to go one step further in the design, synthesis, characterization and study of pseudopeptidic cages for the recognition of both biological ions and specific peptide sequences.

The next 4 chapters of this thesis are divided as follows:

- Chapters 2 and 3 are dedicated to the synthesis and study of pseudopeptidic cages for chloride transport through cell membranes.
- Chapters 4 and 5 are dedicated to the synthesis and study of another family of cage-like compounds for Tyr-kinases inhibition purposes.

Each chapter contains a more detailed introduction on the specific topic followed by the design, synthesis, preparation and study of the corresponding cage compounds.

Bibliography

1 J. P. Sauvage, *Angew. Chemie - Int. Ed.*, 2017, **56**, 11080–11093.

2 B. L. Feringa, *Angew. Chemie - Int. Ed.*, 2017, **56**, 11060–11078.

3 J. Sun, M. Cai and J. J. Lavigne, *Supramol. Chem.*, , DOI:10.1002/9780470661345.smc177.

4 T. L. Mako, J. M. Racicot and M. Levine, *Chem. Rev.*, 2019, **119**, 322–477.

5 H. Schmidt and F. Würthner, *Angew. Chemie*, 2020, **59**, 2–12.

6 G. Ashkenasy, T. M. Hermans, S. Otto and A. F. Taylor, *Chem. Soc. Rev.*, 2017, **46**, 2543–2554.

7 F. B. L. Cougnon and J. K. M. Sanders, *Acc. Chem. Res.*, 2012, **45**, 2211–2221.

8 P. Frei, R. Hevey and B. Ernst, *Chem. - A Eur. J.*, 2019, **25**, 60–73.

9 J. Li, P. Nowak and S. Otto, *J. Am. Chem. Soc.*, 2013, **135**, 9222–9239.

10 D. Carbajo, Y. Pérez, J. Bujons and I. Alfonso, *Angew. Chemie*, 2020, **132**, 17355–17359.

11 G. V. Oshovsky, D. N. Reinhoudt and W. Verboom, *Angew. Chemie - Int. Ed.*, 2007, **46**, 2366–2393.

12 L. Milanesi, C. A. Hunter, S. E. Sedelnikova and J. P. Waltho, *Chem. – A Eur. J.*, 2006, **12**, 1081–1087.

13 J.-P. Behr, *The Lock-and-key Principle: The state of the art 100 years on.*, 2008.

14 M. Delbianco, P. Bharate, S. Varela-Aramburu and P. H. Seeberger, *Chem. Rev.*, 2016, **116**, 1693–1752.

15 G. T. Noble, R. J. Mart, K. P. Liem and S. J. Webb, in *Supramolecular Chemistry: from mMlecules to Nanomaterials*, 2012.

16 A. Barba-Bon, M. Nilam and A. Hennig, *ChemBioChem*, 2020, **21**, 886–910.

17 X. Ma and Y. Zhao, *Chem. Rev.*, 2015, **115**, 7794–7839.

18 D. H. Williams and B. Bardsley, *Angew. Chemie - Int. Ed.*, 1999, **38**, 1172–1193.

19 E. Mattia and S. Otto, *Nat. Nanotechnol.*, 2015, **10**, 111–119.

20 S. J. Barrow, S. Kasera, M. J. Rowland, J. Del Barrio and O. A. Scherman, *Chem. Rev.*, 2015, **115**, 12320–12406.

21 P. E. Eaton and T. W. Cole, *J. Am. Chem. Soc.*, 1964, **86**, 3157–3158.

22 C. J. Pedersen, *J. Am. Chem. Soc.*, 1967, **89**, 2495–2496.

23 J. M. Lehn, *Pure Appl. Chem.*, 1977, **49**, 857–870.

24 M. K. Chantooni and I. M. Kolthoff, *Cathet. Cardiovasc. Diagn.*, 1981, **7**, 7245–7247.

25 K. A. Arnold, J. C. Hernandez, C. Li, J. V Mallen, A. Nakano, O. F. Schall, J. E. Trafton, M.

Tsesarskaja, B. D. White and G. W. Gokel, *Supramol. Chem.*, 1995, **5**, 45–60.

26 B. W. Kiggen and F. Vogtle, *Angew. Chemie - Int. Ed.*, 1984, **23**, 714–715.

27 D. J. Cram, S. Karbach, Y. H. Kim, L. Baczynskyj and G. W. Kalleymeyn, *J. Am. Chem. Soc.*, 1985, **107**, 2575–2576.

28 R. Chakrabarty, P. S. Mukherjee and P. J. Stang, *Chem. Rev.*, 2011, **111**, 6810–6918.

29 D. Fujita, Y. Ueda, S. Sato, N. Mizuno, T. Kumasaka and M. Fujita, *Nature*, 2016, **540**, 563–566.

30 T. J. Mcmurry, S. J. Rodgers and K. N. Raymond, *J. Am. Chem. Soc.*, 1987, **109**, 3451–3453.

31 G. Zhang and M. Mastalerz, *Chem. Soc. Rev.*, 2014, **43**, 1934–1947.

32 A. Moure, S. V. Luis and I. Alfonso, *Chem. - A Eur. J.*, 2012, **18**, 5496–5500.

33 J. L. Katz, K. J. Selby and R. R. Conry, *Org. Lett.*, 2005, **7**, 3505–3507.

34 I. Martí, J. Rubio, M. Bolte, M. I. Burguete, C. Vicent, R. Quesada, I. Alfonso and S. V. Luis, *Chem. - A Eur. J.*, 2012, **18**, 16728–16741.

35 H. E. Högberg, B. Thulin and O. Wennerström, *Tetrahedron Lett.*, 1977, **18**, 931–934.

36 C. Bucher, R. S. Zimmerman, V. Lynch and J. L. Sessler, *J. Am. Chem. Soc.*, 2001, **123**, 9716–9717.

37 M. Mastalerz, *Angew. Chemie - Int. Ed.*, 2010, **49**, 5042–5053.

38 M. E. Belowich and J. F. Stoddart, *Chem. Soc. Rev.*, 2012, 41, 2003–2024.

39 T. Tozawa, J. T. A. Jones, S. I. Swamy, S. Jiang, D. J. Adams, S. Shakespeare, R. Clowes, D. Bradshaw, T. Hasell, S. Y. Chong, C. Tang, S. Thompson, J. Parker, A. Trewin, J. Bacsa, A. M. Z. Slawin, A. Steiner and A. I. Cooper, *Nat. Mater.*, 2009, **8**, 973–978.

40 T. Hasell, X. Wu, J. T. A. Jones, J. Bacsa, A. Steiner, T. Mitra, A. Trewin, D. J. Adams and A. I. Cooper, *Nat. Chem.*, 2010, **2**, 750–755.

41 P. Skowronek, B. Warzajtis, U. Rychlewska and J. Gawroński, *Chem. Commun.*, 2013, **49**, 2524–2526.

42 H. Ding, X. Meng, X. Cui, Y. Yang, T. Zhou, C. Wang, M. Zeller and C. Wang, *Chem. Commun.*, 2014, **50**, 11162–11164.

43 H. Ding, X. Wu, M. Zeller, Y. Xie and C. Wang, *J. Org. Chem.*, 2015, **80**, 9360–9364.

44 K. Konishi, F. Xu and Y. Murakami, *Chem. Lett.*, 2006, **35**, 476–477.

45 H. Ding, R. Chen and C. Wang, *Dyn. Covalent Chem.*, 2017, 165–205.

46 N. Nishimura, K. Yoza and K. Kobayashi, *J. Am. Chem. Soc.*, 2010, **132**, 777–790.

47 Y.-C. Horng, T.-L. Lin, C.-Y. Tu, T.-J. Sung, C.-C. Hsieh, C.-H. Hu, H. M. Lee and T.-S. Kuo, *European J. Org. Chem.*, 2009, **2009**, 1511–1514.

48 D. Yang, J. Zhao, X. J. Yang and B. Wu, *Org. Chem. Front.*, 2018, **5**, 662–690.

49 A. Granzhan, C. Schouwey, T. Riis-Johannessen, R. Scopelliti and K. Severin, *J. Am. Chem. Soc.*, 2011, **133**, 7106–7115.

50 A. Pöthig and A. Casini, *Theranostics*, 2019, **9**, 3150–3169.

51 K. Acharyya and P. S. Mukherjee, *Angew. Chemie*, 2019, **131**, 8732–8745.

52 D. Fujita, *Pure Appl. Chem.*, 2014, **86**, 3–11.

53 J. R. Lakowicz, *Principles of fluorescence spectroscopy, 3rd Edition, Joseph R. Lakowicz, editor*, 2006.

54 Http://www.hyperquad.co.uk/HypSpec2014.htm,

55 A. Vacca, O. Francesconi and S. Roelens, *Chem. Rec.*, 2012, **12**, 544–566.

56 E. Faggi, R. Gavara, M. Bolte, L. Fajarí, L. Juliá, L. Rodríguez and I. Alfonso, *Dalt. Trans.*, 2015, **44**, 12700–12710.

57 I. Alfonso, M. Bolte, M. Bru, M. I. Burguete, S. V. Luis and J. Rubio, *J. Am. Chem. Soc.*, 2008, **130**, 6137–6144.

58 I. Alfonso, M. I. Burguete, F. Galindo, S. V. Luis and L. Vigara, *J. Org. Chem.*, 2009, **74**, 6130–6142.

59 I. Alfonso, M. Bolte, M. Bru, M. I. Burguete, S. V Luis and C. Vicent, *Org. Biomol. Chem.*, 2010, **8**, 1329–1339.

60 I. Martí, M. Bolte, M. I. Burguete, C. Vicent, I. Alfonso and S. V. Luis, *Chemistry*, 2014, **20**, 7458–7464.

61 E. Faggi, A. Moure, M. Bolte, C. Vicent, S. V. Luis and I. Alfonso, *J. Org. Chem.*, 2014, **79**, 4590–4601.

62 E. Faggi, C. Vicent, S. V. Luis and I. Alfonso, *Org. Biomol. Chem.*, 2015, **13**, 11721–11731.

63 E. Faggi, Y. Pérez, S. V. Luis and I. Alfonso, *Chem. Commun.*, 2016, **52**, 8142–8145.

GENERAL OBJECTIVES

Supramolecular chemistry has important challenges when applied to chemical biology projects. On one hand, water is the natural solvent in life and the design of artificial receptors for the efficient recognition of molecules and ions in this medium is specially challenging. Moreover, selectivity is an additional issue when considering biological media, which are intrinsically complex with many interactional networks operating in competition. With this problematic in mind, the general objectives of the present PhD Thesis are to delve into the study of *ad hoc* synthetized pseudopeptidic cages for two different biologically relevant targets:

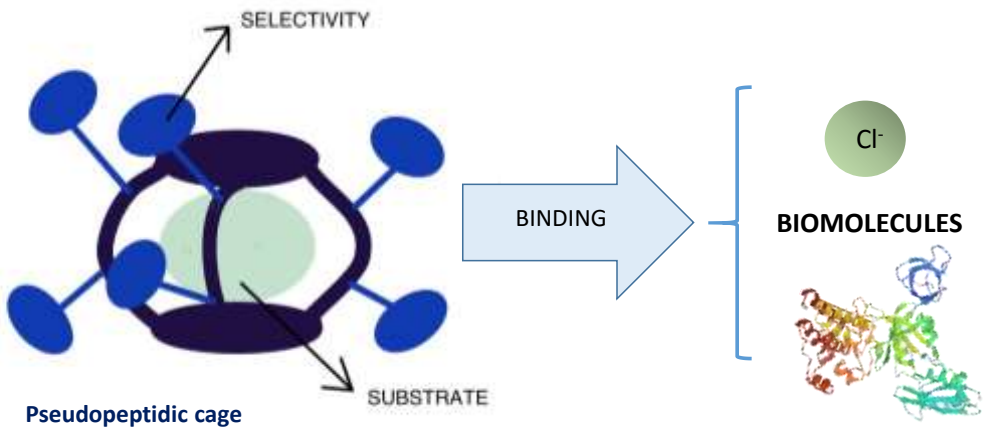


Figure 1. Pseudopeptidic cages cartoon

- *For the selective recognition of chloride anions in biological systems*.

Previous work in our group with small pseudopeptidic cages was used as the starting point for this research. In the present work we aim to introduce new cage-like compound with chloride transport capacity through cell membranes. Then main purpose is to get active compounds able to induce death to cancer cells and to correlate their physicochemical properties (lipophilicity, Cl binding, Cl transport in vesicles and pK_a) with their biological activity in cancer cell cultures mimicking a tumour-like environment. We aim to modulate the corresponding physicochemical and biological properties of the cages by slightly tuning the structures of the side chains coming from the amino acid precursors.

- *For the selective recognition of tyrosine residues within peptides in biological media.*

Prior research in our group regarding the synthesis of pseudopeptidic cages for the recognition of Tyr–containing peptides was also a key starting point for this part of the project. To expand this research our objectives are: 1) the synthesis and characterization of new pseudopeptidic cages decorated with different amino acid side chains, and 2) the study of their binding to biologically relevant Tyr-containing peptides. The challenging objective is to use these cages as tyrosine kinase inhibitors by substrate encapsulation and to study this property with different substrates and purified kinases in vitro. Finally we aim to correlate the binding capacity of each cage towards the different substrates with the kinase inhibition in a biomimetic environment.

The specific objectives regarding the different parts of the Thesis are detailed in the introduction of each chapter. Chapter 2 and 3 are dedicated to Cl^- transporters and chapters 4 and 5 to tyrosine binding and kinase inhibition.

CHAPTER 2 : PSEUDOPEPTIDIC CAGES FOR Cl^- TRANSPORT

2.1 BACKGROUND

Anion receptor chemistry started in the late 60s with the research from Shriver and Bialls[1] and Park and Simmons.[2] However it was not until the 80's that it bloomed thanks to the progresses of Graf and Lehn in their research with cryptanes[3,4,5] and Schmidtchen with macrocycles as anion receptors.[6,7] In the past decades, this research field has moved from simple neutral receptors in organic solvents to complex structures with strong binding properties in water increasing the potential application fields of anion receptors for sensing, pharmacology or waste treatment.[8]

The firsts interaction forces explored for anion binding were anion-π interactions[9] but since the blossom of the topic other interaction forces, mainly hydrogen bonding, are being explored too.[10,11]

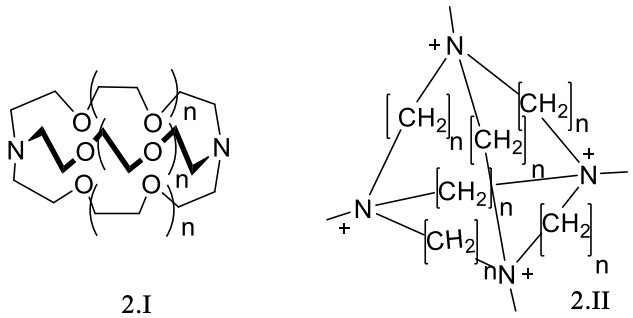
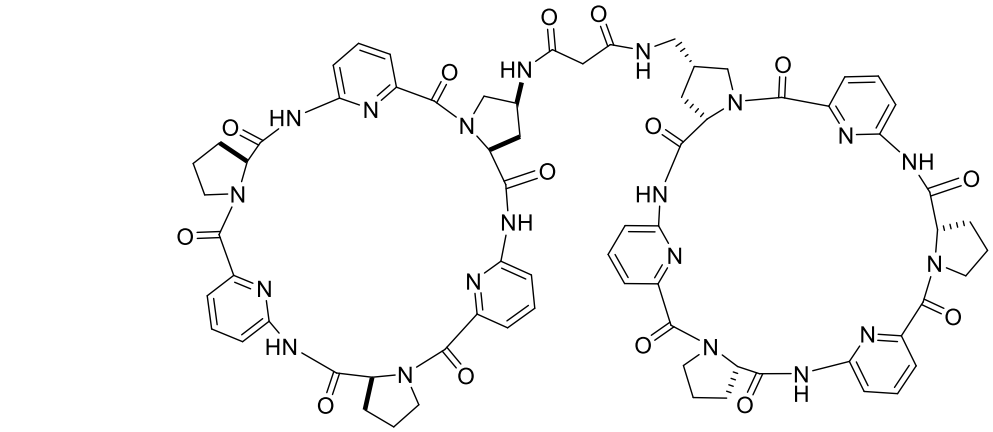
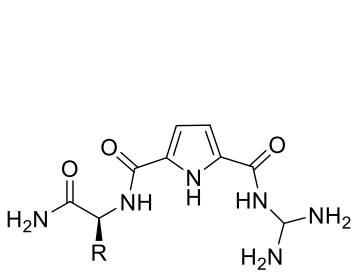


Figure 2.1. General structure of Cryptanes (2.I) and macrocycles described by Shcmidtchen (2.II) as anion receptors

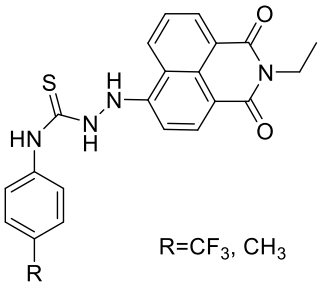
Kubik's compound able to encapsulate sulphate ions,[12] Schmuck's carboxylate receptor able to recognize carboxylate residues from biomolecules (2.III),[13] Gunnlaugsson's thiourea with colorimetric anion sensing properties in water (2.IV)[14] or Martinez-Mañez's cyanide sensor (2.V)[15] are some of the many examples of anion hosts present in bibliography.



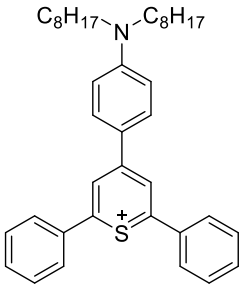
2.III



2.IV



2.V



2.VI

Figure 2.2. Structure of anion receptors 2.III-VI.

2.1.1 Chloride Ion (Cl^-) and chloride channelopathies

Chloride is the most abundant ion in living organisms where it is present forming different electrochemical gradients. However, cell membranes are not permeable to this ion so a more complex mechanism is necessary to transport chloride through them. The incredible progress in the development of techniques for in vivo sensing during the past twenty years has allowed the identification of many proteins involved in Cl^- regulation processes in living organisms. Transmembrane ion channels are the systems responsible for this regulation that affects different physiological functions like muscle contraction, cell volume, neuronal impulses, osmosis, cytoplasmatic calcium concentration or pH. They are membrane proteins that contain a pore and provoke the fast movement of ions across cell membranes,that is why they are also known as chloride pumps.

Alterations in the genes encoding those proteins have been associated with some diseases known as channelopathies.[16] In the past 25 years, many diseases have been associated with

the incorrect functioning of ion pumps like epilepsy, Startle disease, deafness, blindness, neurodegeneration, osteoporosis, lung infection, cystic fibrosis and cancer.[16,17] Hence, these proteins have become potential targets for the treatment of the mentioned diseases.

One of the most widely studied channelopathies is Cystic fibrosis, a multisystem disease whose main characteristic is the abnormal transport of salt and water across epithelia. It is a hereditary disease and patients who suffer from that, lack the CFTR ion channel, responsible of Cl^-/H^+ transport through cell membranes. This has taken scientists to look for alternative pathways to transport chloride trough cell membranes to at least partially compensate the CFTR deficiency.[16] More recently, chloride imbalance has also been associated with cancer and the overexpression of chloride pumps has been found to be a common trend in many solid tumours.[18]

Alternatively to targeting the malfunctioning proteins, the use of artificial transmembrane chloride carriers has emerged as a promising strategy with many potential clinical applications for the treatment of diseases characterized by the dysregulation of ion transport processes.

2.1.2 Ionophores

Synthetic ion carriers, also known as ionophores, are organic compounds that can effectively mediate anion transport through phospholipid bilayers without metabolic energy supply. To do so, the carrier needs to first recognize the ion (generally through electrostatic, hydrogen bond or anion-π interactions), make a supramolecular complex that will diffuse across the membrane and release the ion to the other side. Furthermore, these compounds need to be able to repeat the process and continue to transport ions.

There are several natural cation selective ionophores which have indeed served as an inspiration for the design of artificial cationophores like for instance polycyclic ethers or cyclodextrins used as antibiotics.[19] Meanwhile anionophores are much less common and they haven't reach the clinic yet.

In the cellular level, electrolytes imbalance, particularly chloride imbalance between intra and extracellular environments often leads to apoptosis.[20] This feature can be used as a strategy for the rational design of chemotherapeutical drugs acting either breaking up cell homeostasis or by restoring the normal homeostasis of cells. [19,21]

Promising results have already been published with anion transporters capable of transporting chloride though cell membranes replacing the function of a missing ion channel.[22] Cells exposed to these chloride transporters often get their pH gradient altered thus ending up in apoptosis. [23,24] This trend makes anion transporters suitable for potential uses as chemotherapeutic agents for cancer treatments.[25,26]

2.1.3 Chloride transporters

The biological relevance of Cl^- transport across cell membranes turned it into the perfect target for ion transporters. So far, great interest has been placed in the use of small molecules for chloride transport since they are more "drug-like". Many of the described Cl^- transporters are based on ureas,[27,28] thioureas,[29,30] sulphonamides,[31,32] squaramides,[33,34] pyrroles,[35] imdols,[36,37] imidazoles/benzimidazoles[38,39,40] and triaminopyrimidine[41] (**Figure 2.3**) and they employ ionic interactions, hydrogen bonds, salt bridges, halogen-bonds, anion-π and Lewis acid anion interactions for the formation of the supramolecular complex with Cl^-.

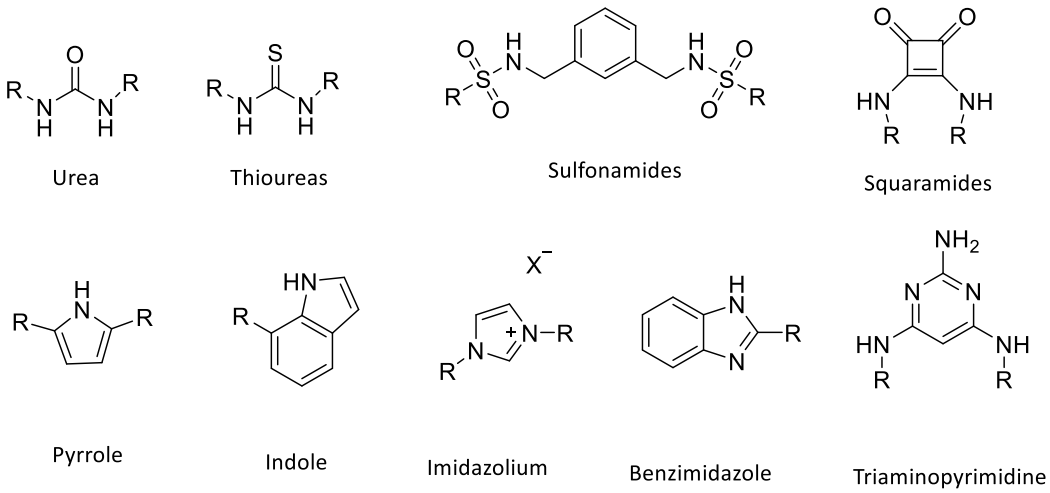
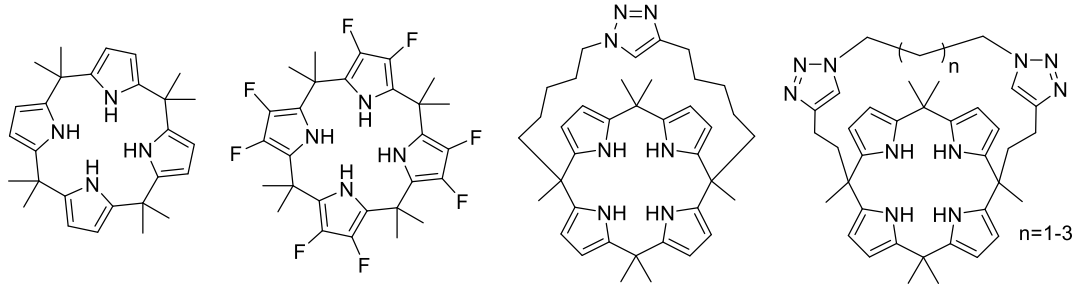


Figure 2.3. Typical structures used for the construction of Cl^- transporters.

Cl^- transport by synthetic molecules generally occur via a passive mechanism where the charge balance across the membrane remains unaltered. Those mechanisms are cotransport (or symport, process in which two ions are transported across the membrane together) and antiport (where two anions are transported in opposite directions across the membrane).

A widely studied family of compounds for this purpose are calix[4]pyrroles since they offer many possibilities in the development of transmembrane ion transporters via different mechanisms. The pyrrolic N-H provides the capacity to bind anions and the aromatic cup to bind larger polarizables cations.

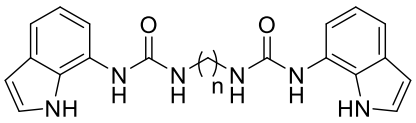
Examples of calix[4]pyrroles with chloride binding properties are presented in Figure 2.4. 2.VII acts as a CsCl contransporter thus being electroneutral.[42] Fluorination of the pyrrole rings enhances the NH acidity increasing the Cl^- binding and the transmembrane transport capacity of the molecules, that is why 2.VIII is an efficient nitrate/chloride and bicarbonate/chloride exchanger thus capable of uniport across a membrane.[43] Gale and co-workers also studied strapped calix[4]pyrroles (2.IX-X in Figure 2.4) which showed higher Cl^-/NO_3^- exchange in the presence of caesium salt. This is a clear example of the effect of both co-transport and anion exchange to the global transport activity of this family of compounds.[44]



2.VII 2.VIII 2.IX 2.X

Figure 2.4. Calix[4]pyrroles 2. VII-X with Cl^- transport properties.

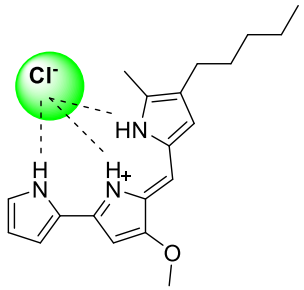
Gale et al. also reported bis-indolyurea derivatives, compounds with two anion binding units linked by an alkyl chain of different length, for anion transport (Figure 2.5). Moreover, through molecular dynamic simulations they demonstrated that the lipid bilayer is not disrupted by the internalization of these compounds.[45]



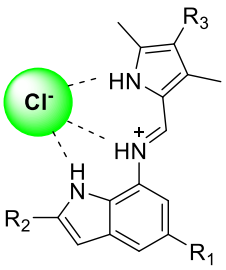
2.XI

Figure 2.5. Bis-indolyurea derivatives 2.XI.

Aside from synthetic transporters, there are natural compounds with potent Cl^- transport capacities. The most active known one is prodigiosin, a red pigment produced by several bacteria strains and a H^+/Cl^- symport cotransporters which transports Cl^- across lipid vesicles *via* an anion exchange (or antiport) mechanism.[46] Besides, prodigiosin can induce apoptosis in cancer cells without important side effects to healthy cells.[47] Using this natural product as inspiration, Gale's group designed a family of potent HCl transporters that he named perenosis which also function as HCl cotransporters and present anti-cancer activity (Figure 2.6).[37,48]



2.XII



2.XIII

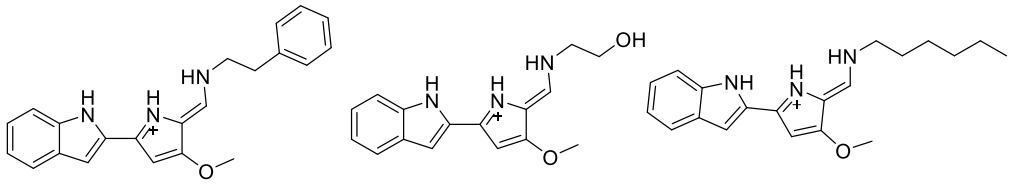
Figure 2.6. Structure Prodigiosin 2.XII and perenosin 2.XIII HCl complexes.

2.1.4 Biological effect of Cl^- transporters

Chloride concentration difference between the inside (4-60 mM) and outside /(120 mM) of a cell make Cl^- transport possible using anion transporters along Cl gradient.

Proper balance of ions in cells is essential for the normal functioning of cells.[49] Cl^- transporters will not only alter Cl^- imbalance but also the membrane potential and pH gradient which play a crucial role in cell proliferation, migration and chemoresistance (see section 2.1.7).

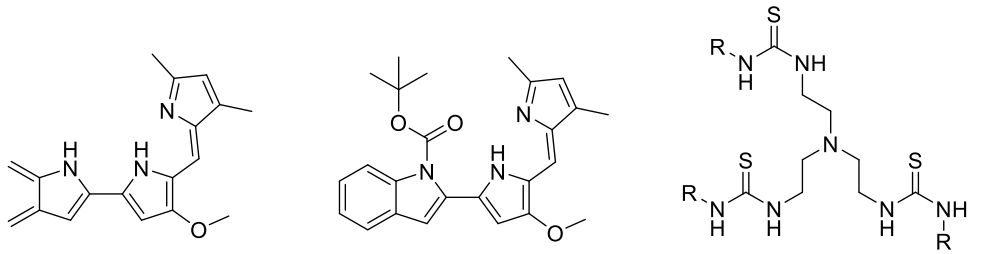
Quesada and co-workers studied the capacity of active anionophores inspired in tambjamines (**Figure 2.7**) to induce cytosol activity and hyperpolarization of the plasma cell membrane. In the same work they demonstrated how this double effect can lead to an effective elimination of cancer stem cells (a tumour cell subpopulation responsible of tumour initiation).[50]



2.XIV 2.XV 2.XVI

Figure 2.7. Structure of tambjamine 2.XIV-XVI derivatives reported by Quesada and co-workers.[50]

They also studied the biological effect of both natural ion transporters (pordigiosin derivatives) and synthetic ones (tren-based thioureas) (Figure 2.8). In that work they reported that active transporters both natural and synthetic triggered apoptosis in several cancer cell lines and have similar properties at least regarding chloride transport and biological effects. [51,52]



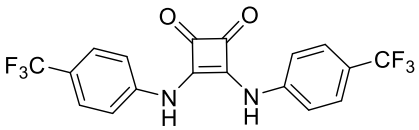
2.XVII 2.XVIII

R= Ph, *p*-CNAr, *n*-pentyl, *tert*-pentyl

2.XIX

Figure 2.8. Prodigiosin derivatives 2.XVII-XVIII and tren based thioureas 2.XIX reported as Cl^- transporters with biological activity.

Almost at the same time, Shin and Gale studied the effect of ion transporters based in squaramide moiety in cells and they found a correlation between the ion transport observed in model vesicles and cell death. Indeed they found and demonstrated that one of their compounds (**Figure 2.9**) disrupts autophagy.[33]



2.XX

Figure 2.9. Squaramide-base compound 2.XX studied by Shin and Gale.

The correlation between transport capacity of anionophores and their cytotoxicity has been attributed to their capacity to disrupt chloride balance in cells, lysosomal pH and cell membrane potentials.[24,23] Hence antitumoural activity of anionophores seems to be directly linked to their transport activity. [33,23,53]

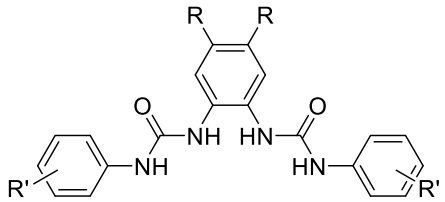
2.1.5 Optimization of Cl^- transport

The many examples present in the bibliography show that the development of efficient anion transporters has gained much attention and effort. As a result, a variety of strategies have been used to optimize transport activity thus improving the biological effect. The two main tuneable properties to improve anionophore activity are anion binding affinity and lipophilicity.

2.1.5.1 Anion binding affinity

As it was previously mentioned, the first step for Cl^- transport is the selective recognition of the ion by the host. To achieve the desired selective recognition, researchers relay on anion-μ, hydrogen bonding and halogen bonding interactions. Therefore, the anion binding strength can be enhanced by the introduction of electron withdrawing substituents which improve the acidity of NH and CH hydrogen bond donors.

In 2013, Moure et al reported a family of o-phenyldiamine based bisureas (**Figure 2.10**) aiming to study the effects of different electron withdrawing substituents in different positions of the molecule. They found that best transporter was the compound with R1=F, R'=H and that the position of R' significantly affects the Cl^- transport activity. Besides, some of these compounds demonstrated some anti-tumour activity against different cancer cell lines.[54] The effect of different electron withdrawing substituents in a family of compounds has also been studied by other groups like Yu and co-workers, who prepared a family of bis-benzimidazolyl-based anionophores also with different electro withdrawing substituents in different positions. [55]



R=H, F, Cl
R'=*p*-CN, *p*-CF_3, *p*-NO_2, *m*-NO_2, *o*-NO_2.

2.XXI

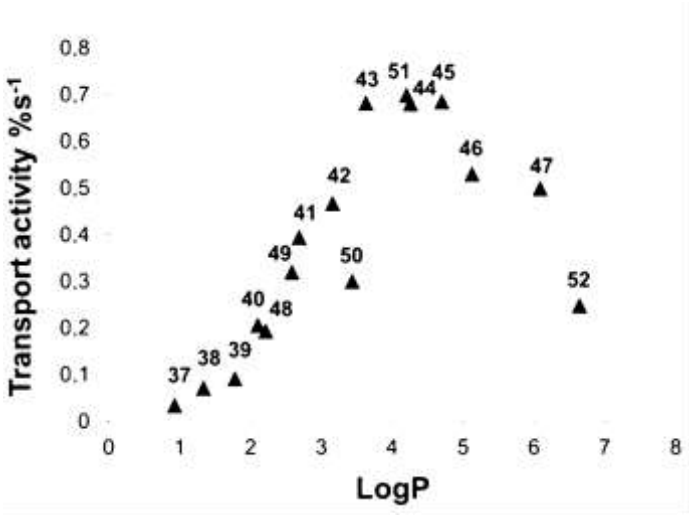
Figure 2.10. Chemical structure of bis-ureas 2.XXI described and studied by Moure et al.[54]

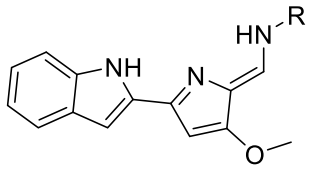
It is however important to highlight that anion transporters are considered to follow "Goldilocks principle": best transporters are not the ones with the highest neither those with the lowest affinity but those with the appropriate intermediate affinity values. Hence, for each family of compounds different strategies might be necessary to achieve the desired Cl^- binding strength.

2.1.5.2 Lipophilicity

A key parameter in the design of anionophores is their lipophilicity since these compounds have to shuttle back and forth through them transporting ions with a good turnover.

Several studies have tried to link lipophilicity to transport. Quesada and co-workers studied this correlation with tambjamines.[56] A representation of the initial transport rate against the log *P* showed a parabolic relationship with optimal values of log *P* around 4.2 (**Figure 2.11**). They hypothesized that compounds with lower log *P* were to hydrophilic to interact with lipidic membranes and those with higher log *P* had poor deliverability from the aqueous phase. Moreover they present lower mobility in the membrane since they mainly remain in the most lipophilic region avoiding the contact with the aqueous phase.[35]

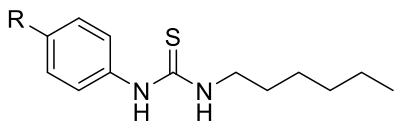




2.XXII

Figure 2.11. Representation of the transport activity measured as initial chloride efflux (%s^{-1}) vs calculated average log P for different tambjamine derivatives 2.XXII. [56] (Image from ref. [56])

Similarly, Gale and co-workers studied this effect in a series of 1-hexyl-3-phenylthioureas with different substituents in 4-position of the aromatic ring to have different lipophilicities (Figure 2.12). They performed a QSAR analysis aiming to determine the effect of each molecular parameter in the performance of a compound as anion transporter. Finally, they concluded that in this set of compounds, an increase in lipophilicity correlates with an increase in anion transport ability. Besides, together with Spooner, this set of compounds was enlarged to have a wider variety of log *P* values and they demonstrated that the optimal log *P* for transport in this series of compounds was the same, around 5.[57]



2.XXIII

R= Br, CF_3, Cl, CN, $COCF_3$, COMe, COOMe, F, H, I, NO_2, O(CO)Me, OCF_3, OEt, OMe, SMe, SO_2Me, CH_3, $(CH2)_nCH_3$ (n=1-4).

Figure 2.12. 1-Hexyl-3-phenylthioureas.

However, so far it seems that the relation between log *P* and transport is strongly dependent on the family of compounds, so a general rule to predict the best log *P* value for Cl^- transmembrane transporters is not available yet.

A typical strategy to improve biological activity and lipophilicity of small molecules is to introduce fluorine residues to them. In the particular case of Cl^- transporters, it has been observed that the introduction of a fluorine or a trifluoromethyl group to an anionophore can improve both anion transport and anti-tumour activity. [58,59]

This strategy has been used by several authors. For instance Gale and co-workers synthetized a series of mono-ureas and mono-thioureas some of them containing a trifluoromethyl group and they showed that fluorinated compounds have higher transport capacity in phospholipid vesicles. Besides, the best transporters from this series induced apoptosis in human cancer cell lines.[36] Gale also used this strategy with tren- based carriers (Figure 2.13), which upon fluorination became better transporters that their non-fluorinated analoges.[58] Therefore, the same effect was tested and found with thioureas and indole ureas. [36] This enhancement in the transport was associated to the increase in the lipophilicity of the fluorinated analogues. In addition, fluorinated compounds showed higher effect in cells deacidifying acid organelles and inducing cell death. [60,58]

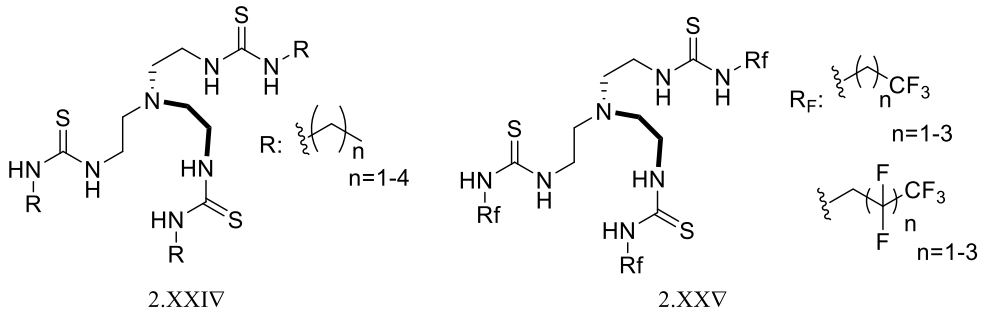


Figure 2.13 Gale's tren based fluorinated anion carriers. [58]

One must not forget, however, that solubility is also a key parameter closely linked to lipophilicity in the design of potential drugs, so the right balance between lipophilicity and solubility needs to be found for ion transporters to work in living systems.

The development of new Cl^- transporters is still a hot topic in supramolecular chemistry. To date, some other structures like anthracene derivatives, cyclic peptides, cyanurates etc. have also been used for the development of novel active Cl^- transporters.[25]

2.1.6 Approaches to get selective transporters

Synthetic ion transporters are not directly affected by efflux pumps present in tumour cells hence they have important advantages vs other chemotherapeutic agents such as strong

inhibition towards a wide range of cancer cells lines and the fact that they do not induce drug resistance. Even so, safety of ion transporters is still a challenge that requires the search for strategies to address their selectivity. Such strategies can be for instance the use of light controllable activity or the pH-dependent activity among others. [44]

The main challenge in drug design is always selectivity. Regardless of the demonstrated effectivity of ionophores inducing cell death, selectivity for cancer cells over healthy cells is still an issue to address. To get some selectivity in Cl^- transporters chemists have considered to link the anion transport capacity to specific conditions such as photosensitivity.[61,62] A very promising strategy in this direction is to link transport to pH since pH gradient is inverted in cancer cells specially those forming solid tumours (see next section for details and examples).

2.1.7 pH imbalance in cancer cells

Cells use glucose to yield energy by forming mainly ATP which when consumed ends as CO_2 and protons. This metabolic process is called glycolysis.

In the 20s Otto Warburg noticed that cancer cells were able to make glycolysis even in the presence of oxygen producing lactate. This was later named the Warburg effect and is considered to be a hallmark in cancer.[63] This upregulation of glycolysis creates an alkaline pH_i (intracellular) and an acidic pH_e (in the extracellular environment). pH gradient in a normal cell is pH_i (intracellular) around 7.2, lower than the extracellular pHe of 7.4, whereas in cancer cells pH_i is 7.4 or above and pH_e can drop to 7-6.7. These inverse pH conditions is one of the adaptive characteristics of tumours. It is achieved and maintained due to changes in the expression and activity of ion pumps responsible for the facilitated proton efflux, oxygen depletion and an increased activity of the carbonic anhydrase. Furthermore, it is ideal for the initiation of metastasis (principal cause of death in cancer patients). Various tumour properties are sensitive to small pH variations such as cell progression, adaption, migration, or metabolism among others. Moreover, many biological processes are susceptible to pH changes such as metabolism, protein expression, protein-ligand binding or ion channels activity which in tumour cells have been detected to create an electrolyte imbalance to favour the tumour progression.

Cancer cells use mitochondrial pathways and glycolysis to control their metabolism and their increased pH also limits apoptosis, which occurs in acidic intracellular conditions (with a pH ≤7.2) promoting cell survival. On the other hand, the acidic extracellular pH diminishes

the growth of healthy cells giving cancer cells more space and nutrients for their proliferation, evades the immune system, and alters some of the chemotherapeutic drugs reducing their efficiency. [63]

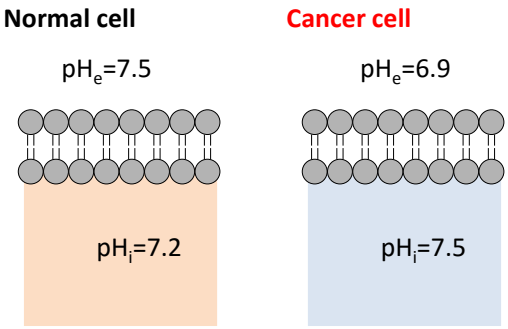


Figure 2.14. Schematic representation of the pH gradient in a normal cell and a cancer cell.

Additionally, reverse pH gradient has an effect in many drugs that have at least one protonation point. Those drugs have a different distribution profile between the intracellular and the extracellular environment. Besides, pH gradients also affect drug activity if they have a pK_a near the neutral pH. This drug distribution might be affected not only by the passive diffusion but also through the modulation of active efflux since generally weak bases are easier effluxes.[64]

This pH gradient can also have positive implications for the future design of cancer treatments. For instance, many pH-sensitive delivery systems have been developed to harness this characteristic.[65]

2.1.7.1 pH sensitive chemotherapeutic drugs

The main action mechanism of chemotherapy is the damage of DNA to induce apoptosis. Regardless of the constant progresses in the design of new cancer treatments, the main challenge is still selectivity to avoid the unpleasant side effects derived from chemotherapeutic treatments. Nevertheless, resistance to chemotherapy is also an issue to address.[66,63]

Reverse pH gradient is common in all types of cancer cells. This trademark has been considered an opportunity for the design of specific drugs with pH sensitivity so that they have a much higher effect in cancer tissue than in healthy one. This can be achieved for instance by using compounds that have a different protonation state in the acidic tumour microenvironment than in healthy cells environment.

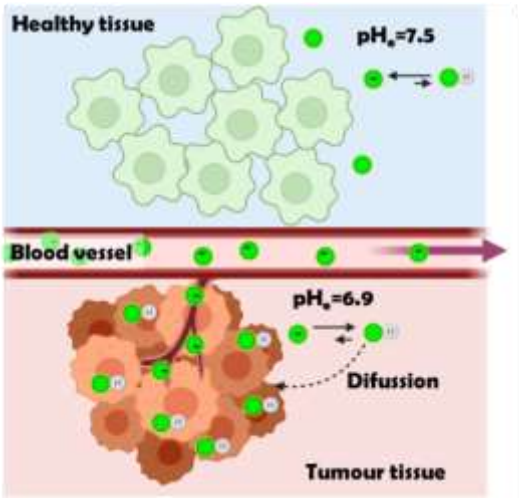


Figure 2.15. Schematic representation of pH sensitive chemotherapeutic drugs action mode (represented using Biorender).

Fukamachi et al[65] screened a library of almost 300 compounds against Hela cells at pH 7.5 and at pH 6.6 and they found that two compounds, lovastatin and cantharidin were better inhibitors of cell growth at acidic pH. Other statins (simbastatin and atorvastin) showed the same properties while some other compounds like manumycin A, Iomicin or doxorubicin showed no difference in activity at different pH.

An alternative approach is the use of compounds that selectively decrease the intracellular pH of tumour cells to equilibrate it with the extracellular pH with a disruption of the ion balance leading to the cell death. An example proposed by Dayanjali et al[67] consisted in the selective transfer of nano pores to cancer cells to dysregulated the cation balance inducing cell death.

It has also been considered to target molecules involved in metabolic pathways that work preferably under acidic conditions since those molecules are less active in non-cancer cells in a more basic microenvironment.[63]

Chemotherapy with acid-sensitive drugs might improve the selectivity of the actual cancer treatments, reduce the important side effects of currently used chemotherapeutic drugs and it might even avoid chemoresistance.[66] So far, few compounds have been found with this property but it is very likely that new pH dependent drugs will be developed aiming to exploit the potential of using pH as the selectivity factor for cancer treatments.

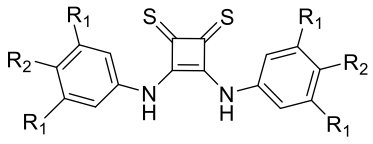
2.1.7.2 pH sensitive Cl transporters

A more recent research line combining the potential of ionophores with pH sensitive properties has emerged as an alternative approach for the treatment of solid tumours. The ideal situation would be an ionophore with high transport capacity at slightly acidic pH, enough to induce cell death, and no transport at physiological pH addressing both selectivity and chemo-resistance issues.

One of the first examples of a pH switchable Cl^- transporter through phospholipidic membranes was reported by Gale and co-workers and consists of a family of thiosquaramides with different substituents in the aromatic rings (**Figure 2.16**). They studied the different properties of these compounds including the binding with chloride, the pKa and the transport through POPC vesicle and they found that compound **1** with R_1=R_2=H and R_1=H, R_2=CF_3 whose pKa are 7.3 and 5.3 respectively, transport Cl^- more efficiently at pH 4 (with an EC_{50} down to 0.001% transporter:lipid molar ratio) than at pH 7.3.[68] (**Figure 2.16**; 2.XXVII). [69]

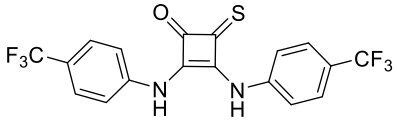
Similar to the previous example, A. Joliffe reported an oxothiosquaramide (2.XXVII) with two CF_3 substituents that binds Cl^- and displays a pH-sensitive chloride transport ability across phospholipidic bilayers at around pH 7 (Figure 2.16). It is in fact a Cl^-/nitrate exchanger that can act also as a HCl symport in the presence of a pH gradient acting as a ON/OFF pH switcher.[70]

A more recent compound with pH switchable transport activity was developed by Arundhati and co-workers who used a bis(sulphonamide) (2.XXVVIII) also with two CF_3 groups (Figure 2.16). [71]

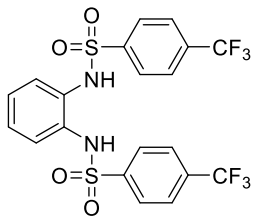


R_1=H, R_2= H
R_1=H, R_2= CF_3
R_1=CF_3, R_2= H
R_1=tBu, R_2= H

2.XXVI



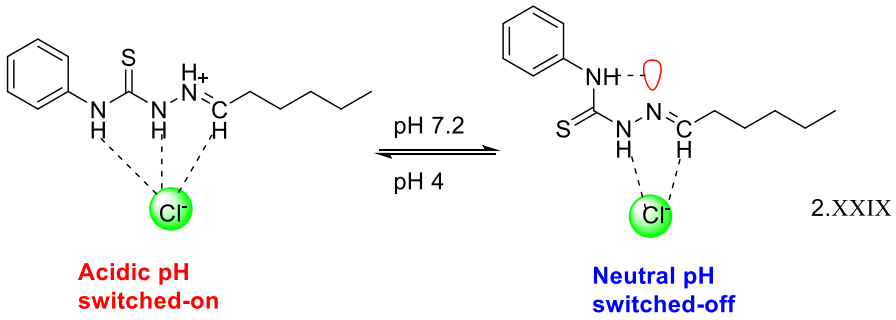
2.XXVII



2.XXVIII

Figure 2.16. Thiosquaramides reported by Gale(2.XXVI), Joliffe (2.XXXVII) and Arundhati (2.XXVIII).

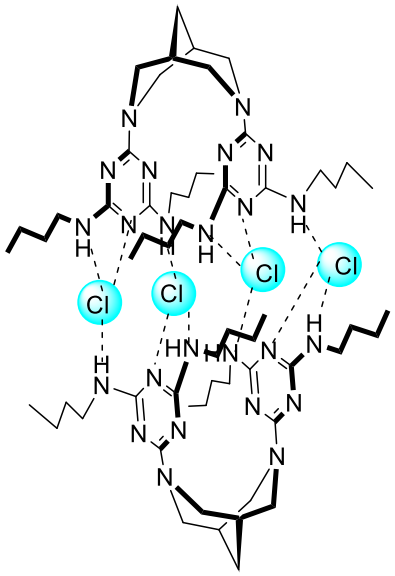
Gale's group have also provided another remarkable example of a pH switchable transporter. Phenylthiosemicarbazones are compounds that can be protonated at acidic pH and function as a highly active H^+/Cl cotransporters. This is the case of the compounds form Figure 2.17; their calculated EC_{50} in the chloride/nitrate assay at pH 4.0 was 0.0074 mol%, 640-times lower than that observed at pH of 7.2. Hence, this system is an example of a non-electrogenic transporter (a transporter that does not facilitate a net flow of charge as since both a H^+ and Cl^- are simultaneously transported) that displaying an interesting pH-switching behaviour between neutral and acidic conditions.[72]



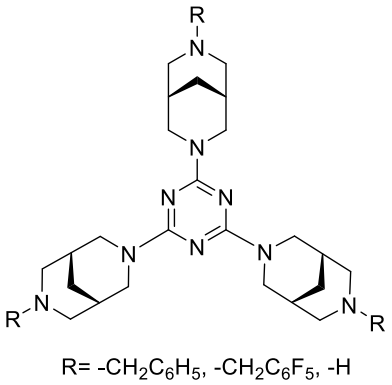
2.XXIX

Figure 2.17. Urea-thiosemicarbazone ON/OFF chloride transport mechanism.

A more complex system consisting on a dimeric Bis(melamine) substituted bispidine reported by Talukdar and co-workers also acts as a H^+/ Cl^- cotransporter through cell membranes (Figure 2.18). The high dependency of the transport activity vs pH was associated to the protonation state of the monomer at different pHs. At low pH the great degree of protonation of the aromatic amines induces a high repulsion between these two groups making both the chloride complexation and dimerization less favoured. An increase in the pH reduces the protonation in these groups preventing the repulsion and enhancing the dimerization and also the Cl^- complexation and transport through membranes.[73] The same group described tripodal semi-cages (Figure 2.18, right) based on trizma core that also displays pH-sensitive Cl^- transport activity across LUVs (large unilamelar vesicles).[74]



2.XXX



2.XXXI

Figure 2.18. bis(melamine)- substituted bispidine 2.XXX and tripodal receptors 2.XXXI reported by Talukdar.

2.1.8 Synthetic ion channels

The other interesting application of ion transmembrane transporters is to mimic natural ion channels.

Elsa and co-workers reported a family of highly efficient chloride transporters with structural similarity to prodigiosin, and studied them in living cells, where these compounds showed a permeabilization of the cell membrane to halides close to the one achieved by natural CFTR channels without compromising cell viability.[52] These compounds also showed a higher efficiency with slightly acidic extracellular pH probably due to the pK_a of these molecules (which is between 6.2 and 6.4).

The CFTR channel transports chloride and bicarbonate in the apical membrane of the epithelium. Over 2000 mutations of the gene encoding this channel have been found that can compromise its function, some of which are directly relate to CF. The use of anionophores to treat these pathologies has the main advantage that they can be used no matter what mutation is affecting the CFTR function. [52]

Gale's group, reported a new group of H^+/Cl^- cotransporters that he named peronosins and were capable of triggering cell death acting as strictly HCl cotransporters.[37]

2.2 OBJECTIVES

2.2.1 Background in the supramolecular chemistry group

The experience of the supramolecular chemistry group working with macrocycles especially pseudopeptidic macrocycles, lead them to in 2012 manage to synthetize tripodal pseudopeptidic cages using an SN_2 reaction.

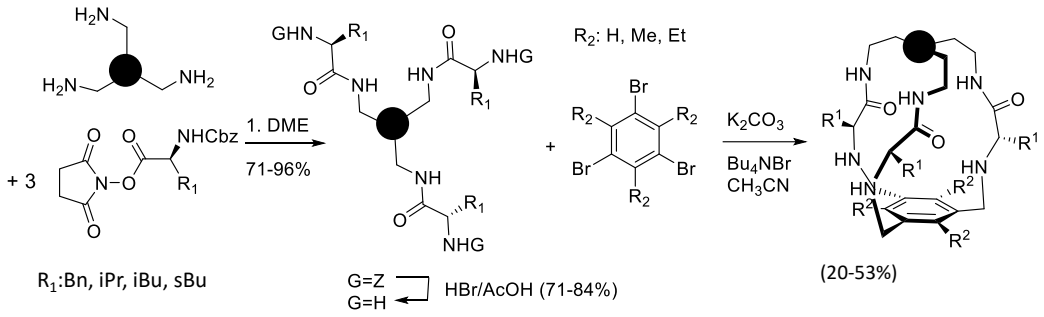


Figure 2.19). They tried different electrophilic bromides for the reaction that was conducted in acetonitrile at reflux with a phase transfer catalyst. [75,76]

Derivatives with a tren- scaffold were obtained in very good yields considering that in each step there is a triple reaction, whereas other scaffolds containing an aromatic ring gave very poor yields and complex mixtures hard to isolate. Besides tren based cages displayed high affinity for Cl^- ions.

The cages obtained with a tren-scaffold were crystalized in their tetra protonated form showing a preference for 4 chloride per unit. Amongst the different cages prepared, those containing phenylalanine residues in position R_1 and Me or Et in R_2 showed the formation of a tight inclusion complex with chloride that settles in the inner cavity of the cage. The stabilization of the anion in the cavity is achieved through H-bonding with the 4 ammonium protons and probably helped by anion-π interactions. There are clear differences in the binding model when the substituent in position 2 is Me, Et or H.[76]

The interaction of these cages with chloride was studied through NMR and gas phase MS (table 2.1) and using POPC vesicles it was possible to measure the transport capacity of these molecules through phospholipidic membranes using an ISE (ion selective electrode).[76]

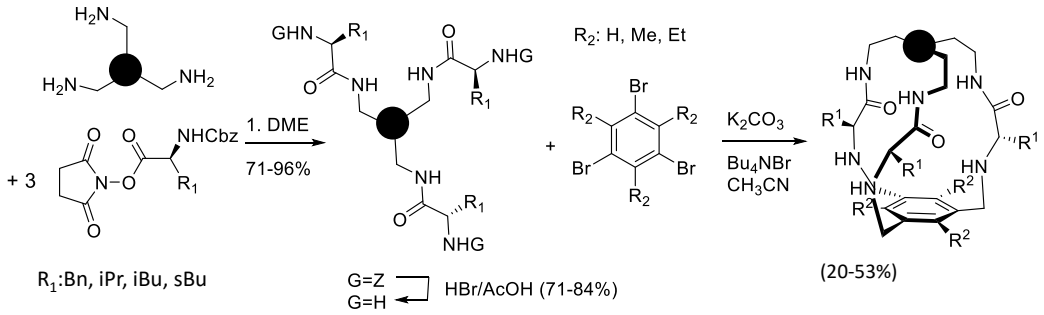


Figure 2.19. Synthetic scheme of the tripodal cages described by Alfonso et al.

Further studies through MS and NMR were successful proving the selectivity of these cages for Cl^- vs other ions like fluoride or bromide (**Figure 2.20**).[77]

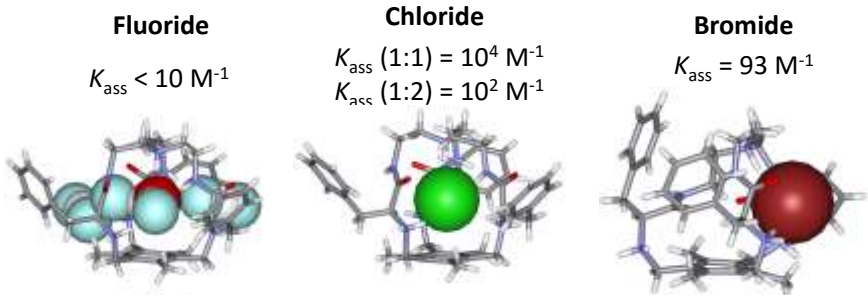


Figure 2.20. X Ray of cage 5e crystalized with fluoride, chloride and bromide.

Binding constants of tripodal cages from reference 18 with Cl^- and $BC_{50}{}^{078}$

Table 2.1. Binding constants of tripodal cages from reference [77] with Cl^- and $BC_{50}{}^0$.

Cage (ref [76])	R^1 (Aaa)	R^2	log β[a] (cage : Cl^-)	$BC_{50}{}^0$(μM)
1y	Bn (Phe)	H	2.35 ± 0.03 (1:1)[b]	4467 ± 309
1a	Bn (Phe)	Me	4.01 ± 0.08 (1:1)[b]	96 ± 17
			6.2 ± 0.1 (1:2)[c]	
1x	Bn (Phe)	Et	3.37 ± 0.09 (1:1)[b]	417 ± 83
			5.1 ± 0.1 (1:2)[c]	
1u	iPr (Val)	Et	3.68 ± 0.07 (1:1)[b]	202 ± 31
			5.86 ± 0.09 (1:2)[c]	
1v	iBu (Leu)	Et	4.35 ± 0.11 (1:1)[b]	44 ± 11
			6.9 ± 0.1 (1:2)[c]	
			7.9 ± 0.3 (1:3)[d]	

1w	sBu (Ile)	Et	3.87 ± 0.08 (1:1)[b] 5.9 ± 0.1 (1:2)[c]	133 ± 24

From the binding experiments and the transport assays conducted in phospholipidic vesicles it was concluded that among the substituents tested, best residues in position R1 are aromatic residues whereas in position 2 higher transport rates are achieved with Et and Me residues.

Moreover, these compounds were proven to display some Cl^- transport activity through POPC vesicles (Figure 2.21).

From the 3D plot of the different properties measured form these cages, a relation between lipophilicity and transport was extracted. The most active Cl^- transporter was the most lipophilic cage whereas a reduction in lipophilicity leads to a reduction of Cl^- transport. Meanwhile, a clear correlation between binding and transport properties form those molecules could not be stablished (Figure 2.21).

From this study, a great interest for the potential biological applications of these cages arouse since they are very stable compounds with a biological resemblance as they are mainly build up from amino acids and with different protonation states at physiological pH.

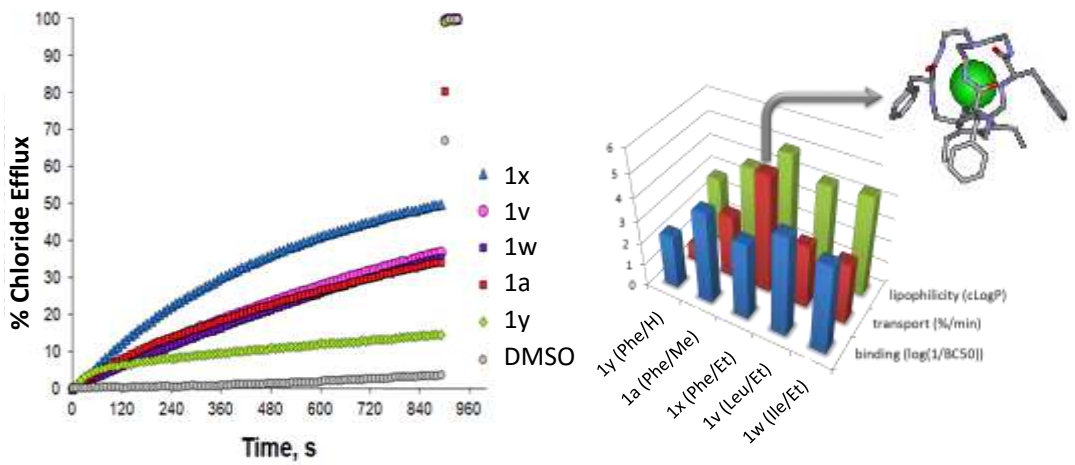


Figure 2.21. Chloride efflux promoted by 25 µM (5% mol carrier: lipid) of compounds from table 2.1 in unilamelar POPC vesicles loaded with 489 mM NaCl buffered at pH 7.0 with 10 mM phosphate dispersed in 489 mM $NaNO_3$ buffered at pH 7.0 (left). 3D plot of the lipophilicity, transport and binding of compounds from table 2.1 (right).

2.2.2 Objectives

Aiming to continue with this promising research line, to get more active anionophores and to get a better understanding of the parameters affecting both transport and biological activities we decided to exploit the structure of pseudopeptidic tripodal cages that had shown a very selective chloride binding when protonated,[79,76] and good abilities to transport this anion through artificial lipid bilayers.[76]

The following objectives regarding this project were proposed:

- Synthesis and characterization of small pseudopeptidic cages able to act as anionophores in aqueous media. Considering the general structure of the small tripodal cages able to selectively bind chloride anion (Figure 2.22),[79] we propose to prepare new derivatives by introducing variations in the side chain R_1 to modulate the chlorine binding ability and the hydrophobicity.
- Binding studies with the corresponding ion-host pairs: The binding constants for the pairs formed between different hosts and the ions will be determined using the most appropriate experimental technique (usually NMR).
- Transport studies of the ion-host pairs through lipidic bilayers carried out in collaboration with experts in the field outside the IQAC-CSIC (Prof. Roberto Quesada from Universidad de Burgos).
- Biological activity of the pseudopeptidic ionophores: The obtained molecules in the previous tasks will be assayed for their biological activity as potential anticancer drugs.

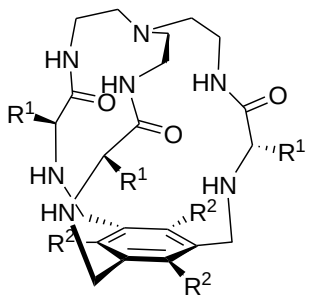
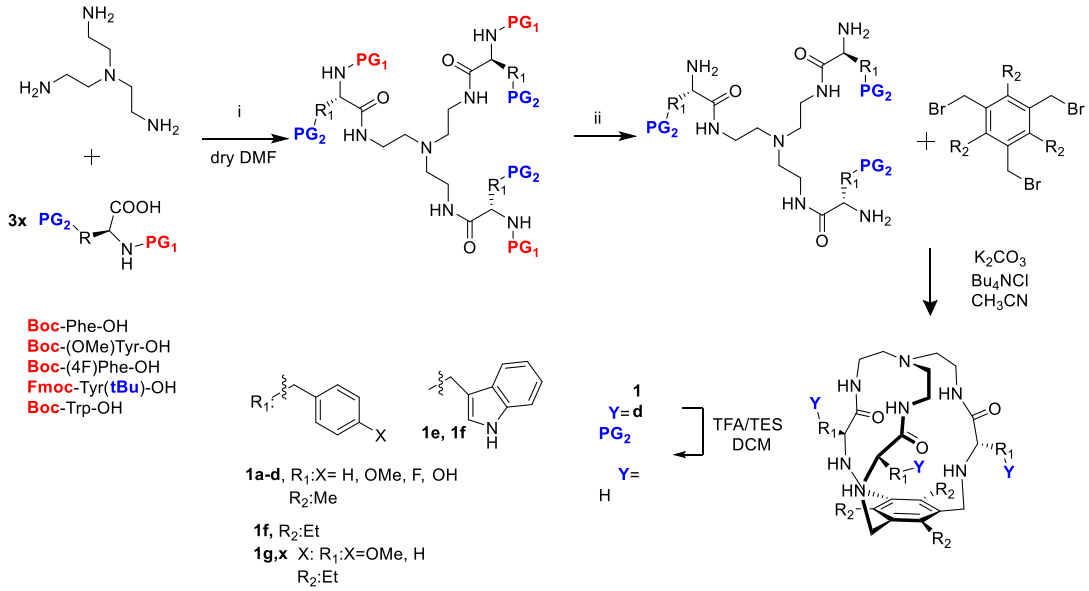


Figure 2.22. Chemical structure of the small pseudopeptidic cages proposed.

2.3 SYNTHESIS OF PSEUDOPEPTIDIC CAGES

In order to fine-tune the physicochemical properties and the potential interactions with the cell membrane, we considered different aromatic side chains deriving from naturally occurring Phe, Tyr and Trp, as well as the non-natural O-Me-Tyr and 4-F-Phe with a methyl group in R_2 (compounds **1a-e** in scheme 2.1). Moreover, compounds with an ethyl group in position R_2 were also studied (compounds **1f**, **1g** and **1x** in scheme 2.1).

The pseudopeptidic cages were prepared following a modification of the previously reported methodology (**Scheme 2.1**). Compound **1d** required the use of two orthogonal protecting groups one for the amine (Fmoc) and one for the alcohol of the tyrosine (*t*Bu) thus their synthesis requires one additional step for the deprotection of the amino acids side chains. Experimental details for the synthesis of each cage are provided in the experimental section. All the cages were accurately characterized by spectroscopic techniques including NMR, MS and HPLC. For two of them, the corresponding crystal structures of the tetra-HCl salts were resolved by X-ray diffraction.



Scheme 2.1. General synthetic pathway for the synthesis of the pseudopeptides cages 1a-g,x: (i) EDC, HoBt DIPEA in DMF(1a-c) or DCC, HoBt, DIPEA in DMF (1d) or HBTU, DIPEA in dry DMF (1e); (ii)) TFA/TES in CH_2Cl_2 (1a-c, 1e, 1f, 1x) or DEA in CH_2Cl_2 (1d).

2.4 STRUCTURAL CHARACTERIZATION OF THE COMPLEXES IN THE SOLID STATE

Crystals suitable for X-ray diffraction studies of compounds **1b** and **1d** were obtained by low evaporation of a methanolic solution of the corresponding compounds with an excess of concentrated aqueous HCl. Their corresponding crystal structures of the tetra-HCl salts were resolved by X-ray diffraction.

In these two examples (Figure 2.23) one of the chloride counter ions is tightly bound inside the cavity of the cage through strong H-bonding interactions, as we had previously observed for related hosts.[76,79] These solid-state structures supported the ability of the protonated systems to efficiently encapsulate a chloride anion.

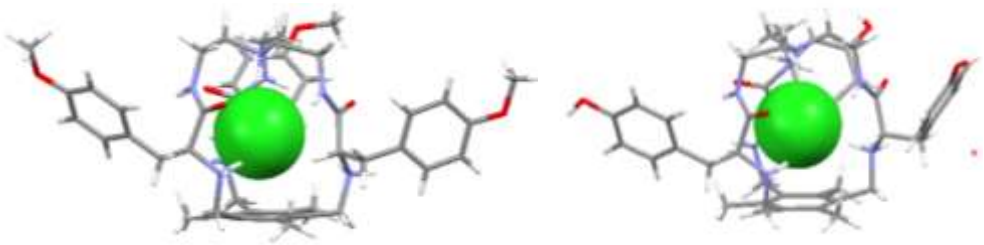


Figure 2.23 1b·4HCl and 1d·4HCl salt (X-ray diffraction) with the encapsulated chloride in CPK. Additional chloride counter ions, solvent molecules and non-polar hydrogen atoms have been omitted for clarity.

2.5 PROTONATION STUDIES OF 1A IN CHLOROFORM

Two stock solutions of **1a** (free base) 5 mM in $CDCl_3$, one of them containing 1 eq. of TBACl, were titrated with TFA 0.653M. ^{1}H NMR spectra were recorded after each addition of TFA. Tetrakistrimethylsililsilane 1 mM was used as the reference (Figure 2.24). The same titration was conducted in the presence of 3 eq. of TBACl (Figure 2.26).

2D NMR spectra of the cage stock solution and of the samples with excess of TFA were recorded to properly assign each signal to the corresponding proton.

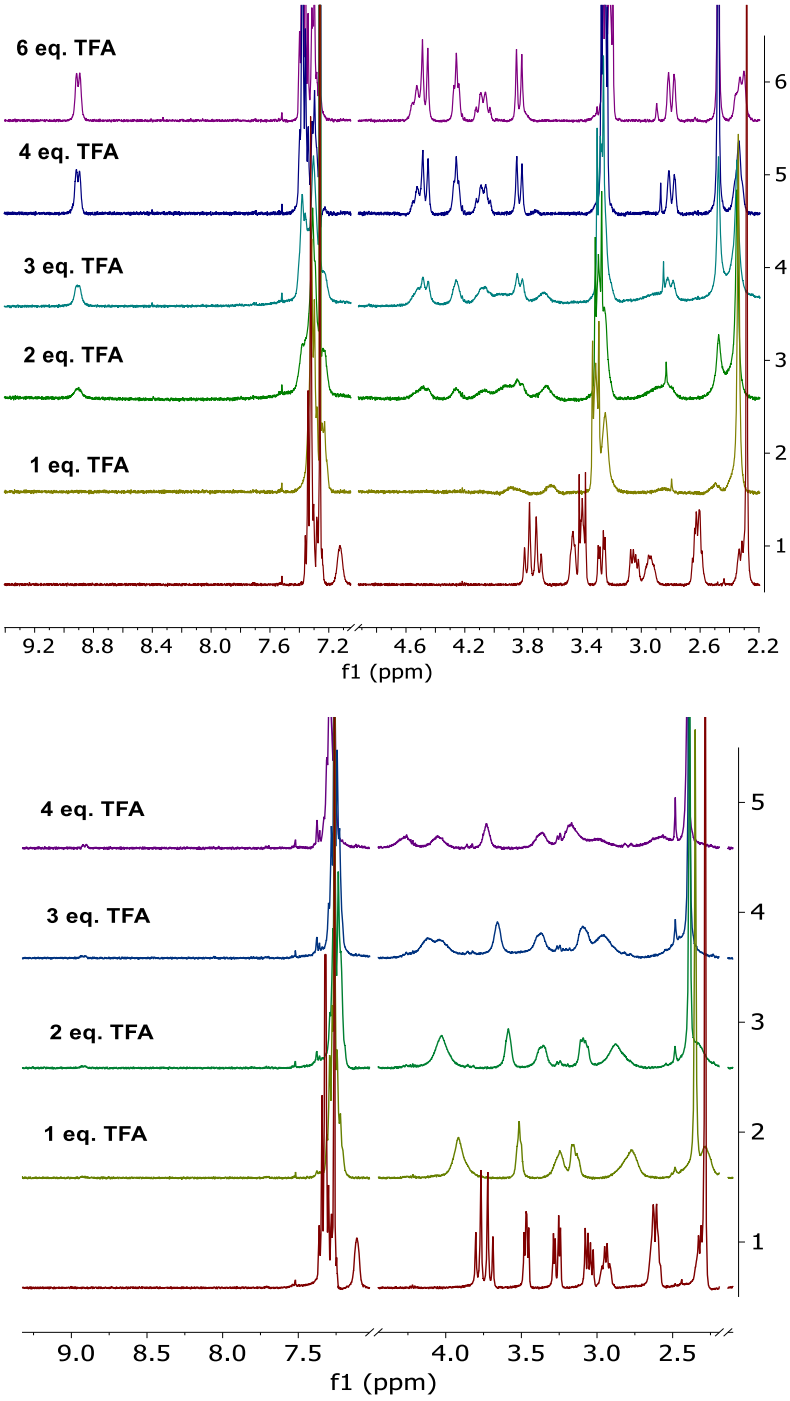


Figure 2.24. Stacked ^{1}H NMR spectra for the titration of 1a in the presence of 1 eq of TBACl (up) and without TBACl (down). After 4 eq. of TFA 1a in the absence of Cl^- precipitates.

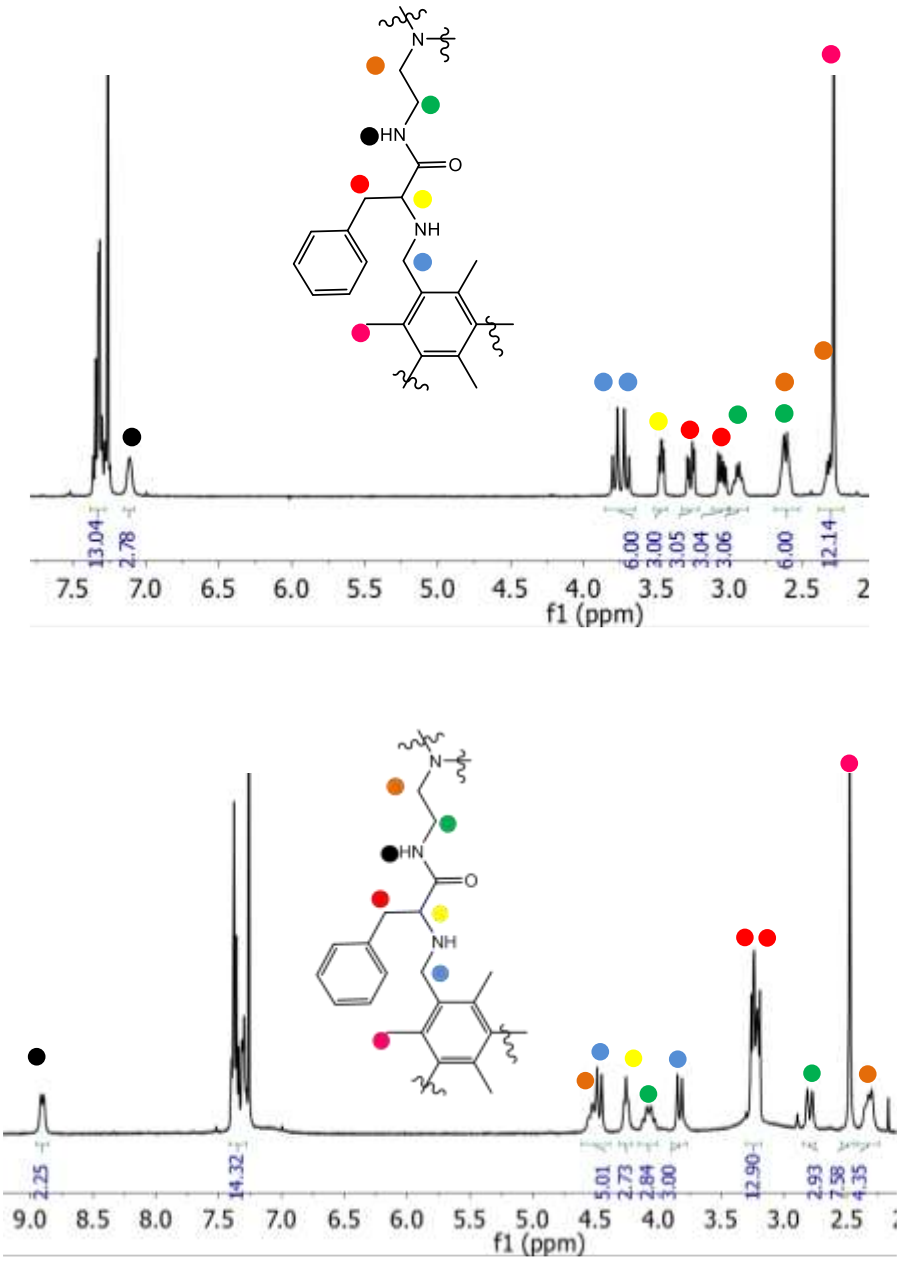


Figure 2.25. ^{1}H NMR of the cage (free base) with tetrakistrimethylsilylsilane (1mM) in $CDCl_3$ (up) and H NMR of the fully protonated cage with 1 eq. of TBACl in $CDCl_3$ (down) with all the signals assigned.

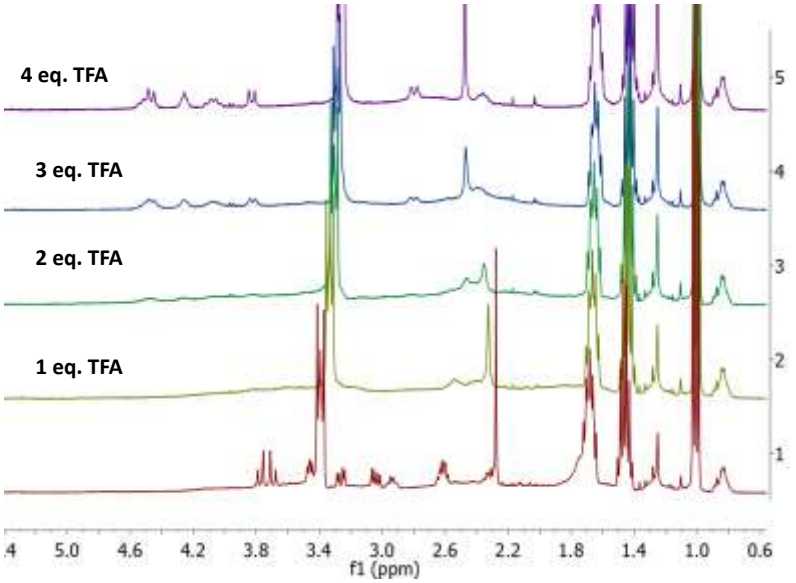


Figure 2.26. Stacked ^{1}H NMR spectra for the titration of 1a in the presence of 3 eq. of TBACl.

The different signal changes in the titration of **1a** with TFA in the presence and absence of 1 eq. of Bu_4NCl points out the interaction between the protonated cage and Cl^-. The complexation of the protonated cage with Cl^- renders a more rigid structure yielding narrow and well defined ^{1}H-NMR signals. This rigidity can be associated with the presence of hydrogen bonds between the protonated amines and the amide with the encapsulated chloride. Therefore, signals more affected by the protonation are those from H close to the protonable amines and amide.

The titration of **1a** containing more equivalents of Cl^- (Figure 2.26) gives the same changes as in the titration with only 1 eq. of Cl. This proofs that only the complexation with one Cl is affecting the rigidity of the structure of the protonated cage. The strong binding pocket being the cage cavity is in agreement with these observations as well as the fact that the association constant 1:1 (between one cage and one Cl) is two orders of magnitude higher than the K_{ass} of the formation of the complex 1:2.[76]

2.6 Chloride binding studies by NMR spectroscopy

We also studied the binding properties of cages **1b-g,** by NMR titration of their tetrakis(trifluoroacetate) salt forms with tetrabuthylamoniumchloride (TBACl) in 5% aqueous CD_3CN at 298 K since this mixture generally allowed a good solubility during the titration experiment and rendered reasonably sharp and well-defined ^{1}H NMR spectra.

Besides, under these conditions, the amide proton is detectable during the titration experiments. The binding with **1a** and **1x** had already been reported.[76] In all the cases the signals more affected by the interaction with chloride were those of the amide NH, the HB/HB' and the HE, all of them close to the binding pocket thus more affected by the interactions with an encapsulated anion. Therefore, the variations of those signals were used for fitting the data using HypNMR 2008 version 4.0.71 software.

$$Cage + Cl \underset{\leftarrow}{\rightarrow} CageCl \quad K_{1:1} \quad \beta_1$$

$$CageCl + Cl \underset{\leftarrow}{\rightarrow} CageCl_2 \quad K_{1:2}$$

$$Cage + 2Cl \underset{\leftarrow}{\rightarrow} CageCl_2 \quad \beta_2$$

The titration experiments and the fitting data for compound **1b** are plotted in Figure 2.28-Figure 2.29. The same analysis was conducted for the other compounds (see experimental section for details) to determine the association constants between chloride and compounds **1a-g,** (Table 2.2). In the titration of **1d** with TBACl partial precipitation in the NMR tube occurs which precludes the accurate fitting of the titration data. Data from **1f** titration was not successfully fitted probably due to the formation of more than 1:1 and 1:2 complexes which complicates the fitting with HypNMR.

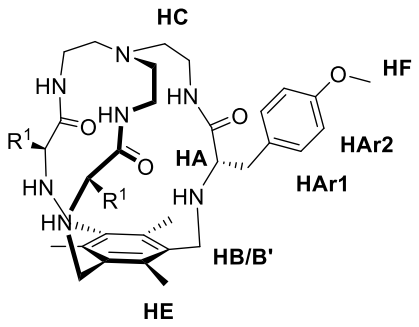


Figure 2.27. Structure of **1b**.

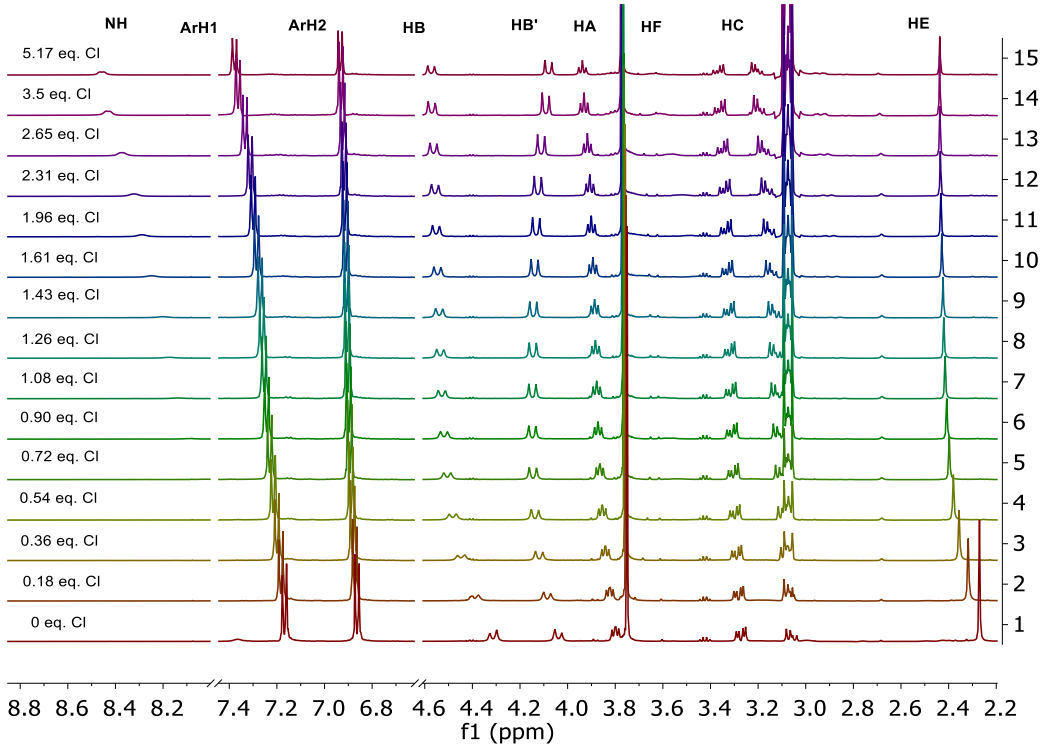


Figure 2.28 1b Stacked ^{1}H NMR spectra for the titration of **1b**.

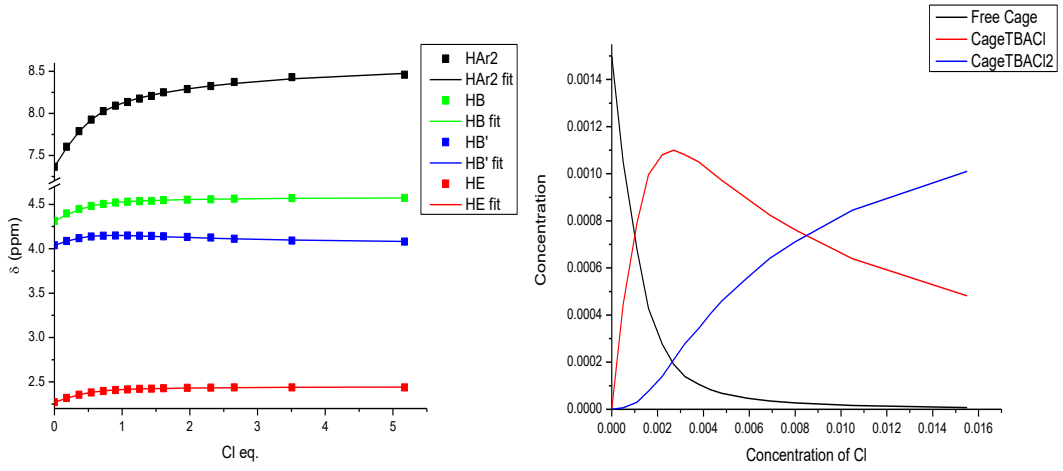


Figure 2.29 Plot of the experimental (symbols) and fitting (lines) data in 1b titration (left) Species distribution as a function of the chloride concentration in 1b titration (right).

Table 2.2 Chloride affinity ($logK_{1:1}$, $logK_{1:2}$, BC_{50}°) for the cages 1a-g,x.

cage	$logK_{1:1}$ (M^{-1})[a]	$logK_{1:2}$(M^{-1})[a]	BC_{50}° (μM)[b]
1a	4.01(8)	2.2(2)	96±17
1b	3.7(1)	2.2	198±44
1c	3.82(3)	2.2(1)	148±10
1d	n.d.	n.d.	n.d.
1e	3.40(3)	2.1(1)	380±24
1f	n.d.	n.d.	n.d.
1g	3.9(1)	2.57*	112*
1x	3.37(9)	1.8(1)	417±83

[a] By NMR titration, standard deviation on the last significant figure in parenthesis. [b] As defined in ref.[78] * Excessive error in β.

Interestingly, we observed their ability to bind two chloride anions with very different binding constants for the corresponding 1 : 1(~10^4 M^{-1}) and 1 : 2 complexes (~10^2 M^{-1}). This result is consistent with the encapsulation of an anion along with a weaker interaction with a second chloride, probably by electrostatic interactions outside the cage. For comparison purposes the chloride affinity was expressed in terms of the BC_{50}^0 parameter obtained by using the BC_{50} calculator version 2.37.1 program. Similar chloride affinities (BC_{50}°)[78] were found for the different cages prepared regardless of the substituents in R_1 and R_2 positions, reflecting a conserved binding pocket for all the receptors.

2.7 LIPOPHILICITY

The goal in the design on anion transporters is to get compounds within the appropriate range of lipophilicity that allow the compounds to interact with lipophilic membranes but also to interact with the aqueous phase to exchange Cl^- from one side to the other of the membrane. To achieve that, drugs have to be soluble both in water and lipids. Thus lipophilicity and solubility are two major properties responsible for drug localization and mobility within lipid membranes.

2.7.1 Theoretical calculations of the log *P*

Log *P* (1-octanol-water partition coefficient) is one of the major and most used parameters to estimate lipophilicity. Computational chemists have developed several methods for the theoretical prediction of log *P* of molecules.

Lipophilicity of compounds **1a-g, x** was theoretically calculated with the VCClab[80] software that uses different methods such as ALOGPS, AC, LogP, ALOGP, MLOGP, XLOPGP2 or XLOGP3, and we used consensus Log *P* expressed as the mean value of the different calculations given by the software for both the neutral and the tetraprotonated form of each cage (

Table 2.3). It is important to indicate that this software determines the lipophilicity using a statistical ensemble of associative neutral networks trained on a map set which is composed by different families of compounds but they are generally planar structures, which is quite far from the 3D structure of the cage-like compounds studied in the present work.

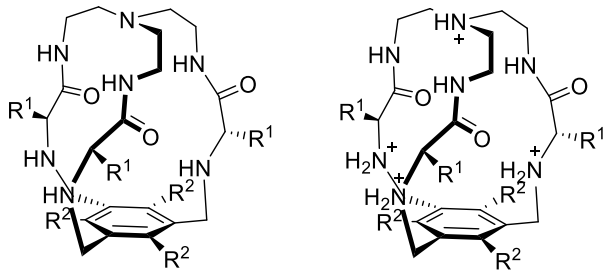


Figure 2.30 General structures of the neutral and tetraprotonated cage.

Table 2.3 calculated Log P values and retention times (minutes). Log P1 correspond to the non-protonated cage and Log P2 corresponds to the tetraprotonated form of the cage.

Cage	*LogP1*	*LogP2*	*Average LogP1 and LogP2*	*Retention time (min)*
1a	*3.56*	*0.26*	*1.91*	*8.02*
1b	*3.63*	*0.1*	*1.865*	*8.43*
1c	*4.31*	*0.77*	*2.54*	*9.44*
1d	*2.81*	*-0.69*	*1.06*	*5.07*
1e	*4.04*	*0.7*	*2.37*	*8.15*
1f	*5.15*	*1.56*	*3.35*	*9.95*
1g	*4.75*	*1.20*	*2.97*	*9.46*
1x	*4.59*	*1.40*	*2.99*	*9.64*

2.7.2 Analytical determination

To accurately determine the log *P* of a molecule we should dissolve it in a water-octanol mixture and analyse which proportion of the drug goes to each phase with quantitative analytical methods.

When the exact value of the log *P* is not necessary and we just need to compare the lipophilicity of different compounds of a series, HPLC analysis can be used. The retention time of a molecule in a reverse-phase HPLC column is directly related to its lipophilicity.

In order to establish a proper comparison between the lipophilicity of compounds 1**a-g,x** the retention time in an HPLC analysis was determined (see experimental section for the full HPLC chromatograms). Using the same gradient and column, compounds with a lower retention time are less lipophilic so they can be sorted from more to less lipophilic (

Table **2.3**). **Figure 2.31** shows the good correlation between the experimental data (retention time) and the theoretical calculations of the lipophilicity in this family of compounds both considering the neutral (Log *P1*) and the tetraprotonated form (Log *P2*) of the cages. The introduction of more hydrophobic substituents either in position R_2 or R_1 can be used to modulate the lipophilicity of the cage.

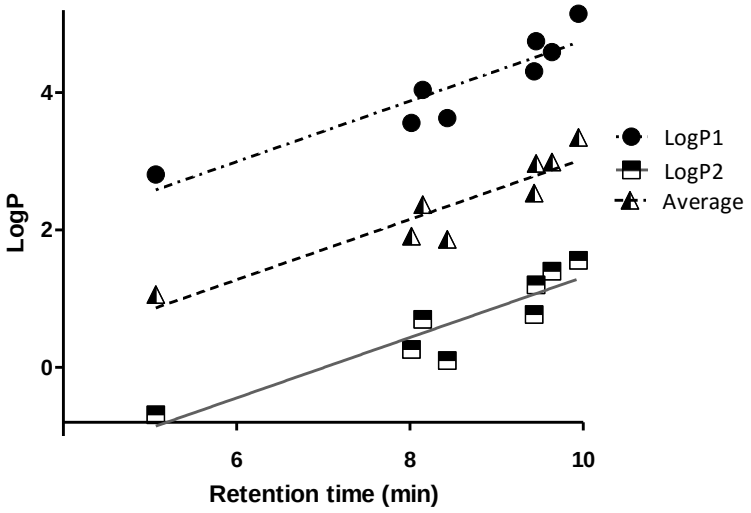
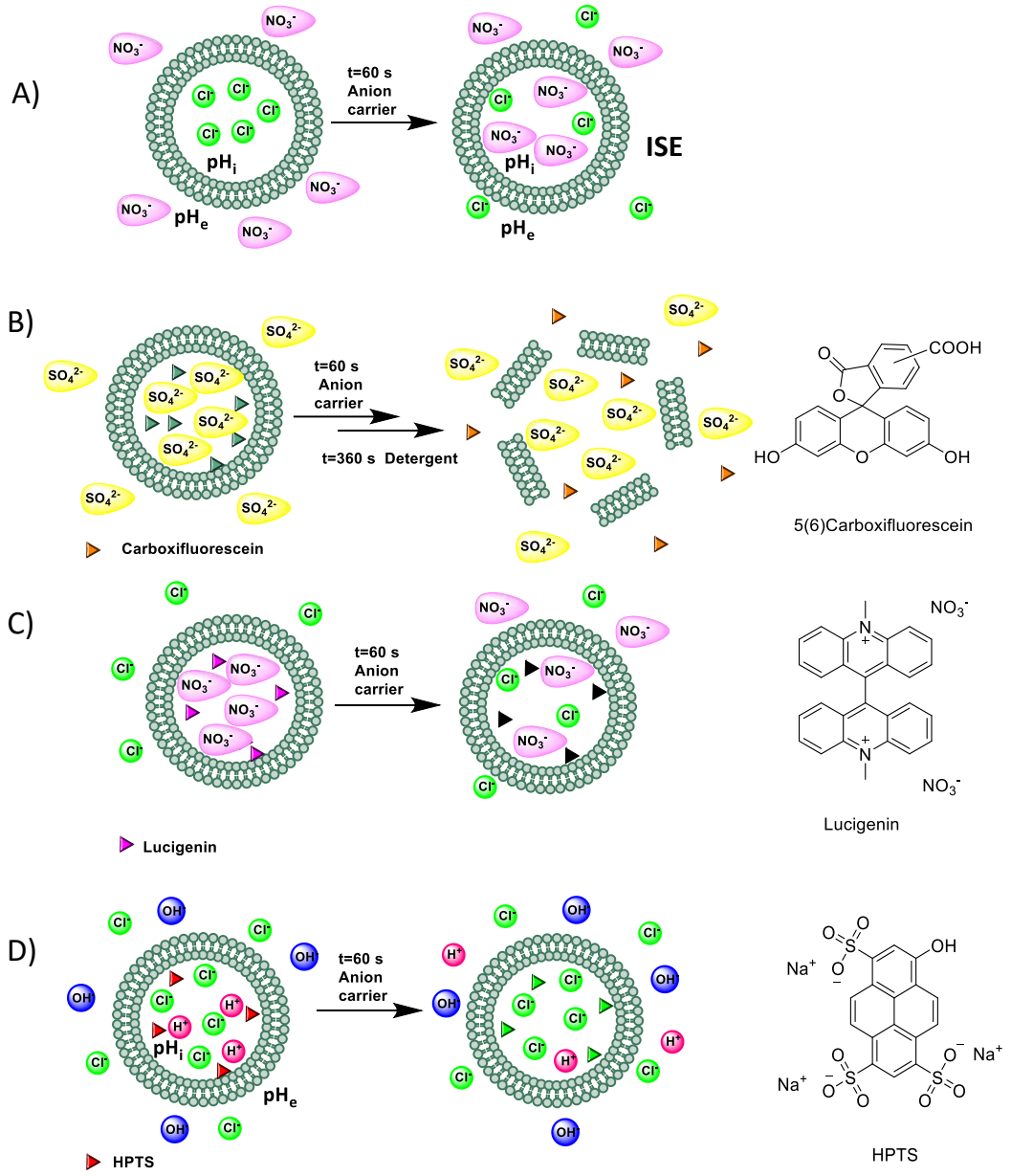


Figure 2.31. Plot of the calculated log P1/log P2/Average vs the retention time (min).

2.8 CHLORIDE TRANSPORT EXPERIMENTS

The ionophoric activity of compounds **1a**-**g, x** was explored in model phospholipid bilayers by using POPC liposomes.[51,81–83] To do so, both potentiometric (ISE) and luminescent (carboxyfluorescein, lucigenin and HPTS based) techniques were used (Scheme 2.2).

Standard literature procedures were followed to prepare the vesicles used in the transport studies.[84] The intravesicular and extravesicluar solutions were buffered to different pH values (6.2, 6.5, 7.2 or 7.5) depending on the experiment with NaH_2PO_4. The concentration of this salt varies depending on the assay. The ionic strength of both solutions (intra- and extravesicular) was always the same, in order to avoid damages of the vesicle membranes.



Scheme 2.2. Schematic representations of: A) ISE, B) Carboxyfluorescein, C) Lucigenin D) HPTS transport assays in POPC vesicles and chemical structure of fluorescent probes used in the transport experiments with vesicles.

First, the ability of these compounds to facilitate chloride efflux from chloride-loaded vesicles was monitored using a chloride selective electrode (ISE).[85] This assay is focused on determining the chloride concentration using a chloride selective electrode (ion selective

electrode, ISE, Scheme 2.2, A). The liposome suspension was placed in an isotonic, chloride-free medium and the studied cages were added as aliquots of stock solutions in DMSO. Percent chloride efflux for each compound was monitored over time, since release of all encapsulated chloride by addition of a detergent allowed the normalization of the chloride leakage. Blank traces correspond to the addition of the same volume of DMSO. Initial ISE experiments were conducted at physiological pH (7.2) both inside and outside the vesicles (Figure 2.32). Na^+ counterion was exchanged by K^+ (Figure 2.33) in the ISE experiment and showed the low influence of the counterion in the Cl^- transport capacity of the best transporters from this family.

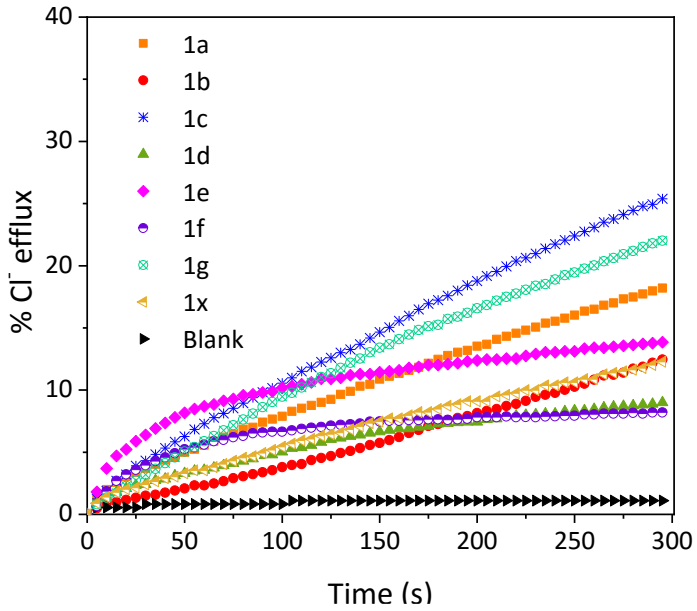


Figure 2.32. Anion transport activity of **1a-g,x** (50 µM =10 mol%) ISE experiment. 0.5 mM POPC. Vesicles loaded with 489 mM NaCl (phosphate buffer 5 mM, pH 7.2) where suspended in $NaNO_3$ 489 mM (phosphate buffer 5 mM, pH 7.2). At t = 300 s a pulse of detergent was added in order to release all chloride anions.

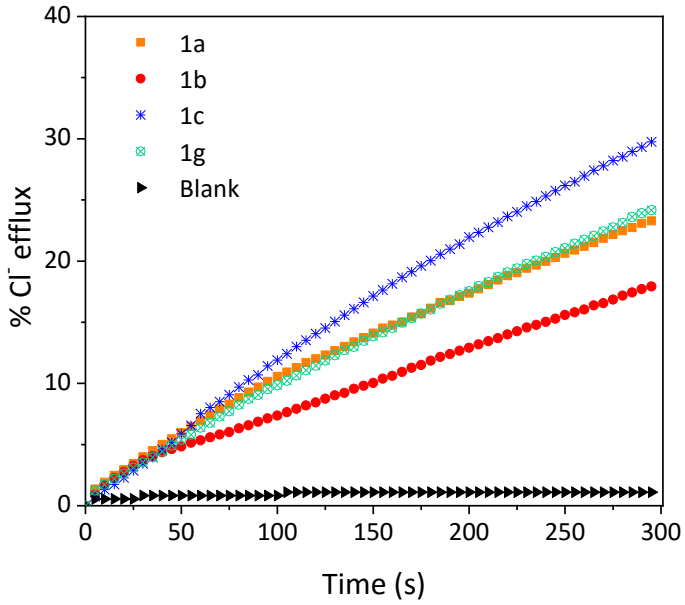


Figure 2.33. Anion transport activity of **1a-c,g** (50 µM =10 mol%) ISE experiment. 0.5 mM POPC. Vesicles loaded with 489 mM KCl (phosphate buffer 5 mM, pH 7.2) where suspended in $NaNO_3$ 489 mM (phosphate buffer 5 mM, pH 7.2). At t = 300 s a pulse of detergent was added in order to release all chloride anions.

Unspecific detergent effects induced by these compounds were ruled out by performing control experiments in vesicles loaded with carboxyfluorescein (CF),[86] a dye whose fluorescence is self-quenched at high concentrations. In Figure 2.34 it is shown that cages induce very low leakage of carboxyfluorescein demonstrating the stability of the vesicles in the presence of these compounds. The only exceptions are compounds **1e** and **1f** (tryptophan derivatives) which have a higher CF leakage effect so we associate the initial Cl^- concertation detected in the ISE experiment with this compounds to unspecific interactions with the vesicle and no direct Cl^- transport. The fluorescence emission of these compounds in the measuring conditions was measured to discard possible interferences with the assay and none of the studied cages showed detectable emission.

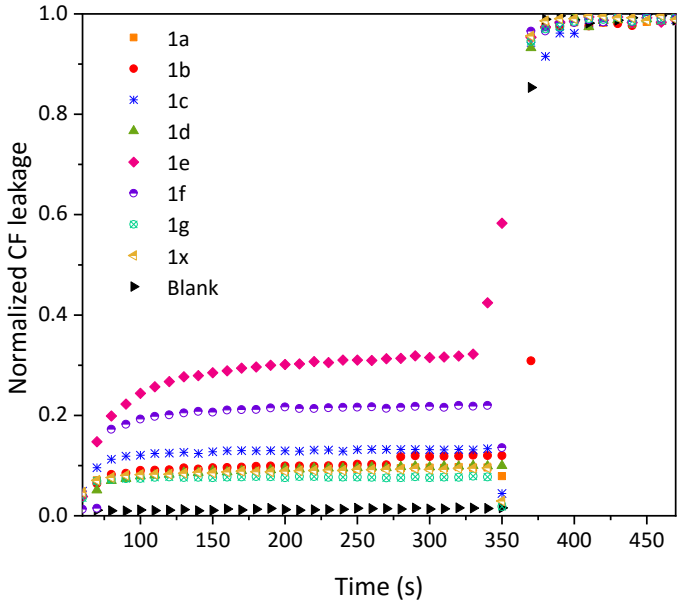


Figure 2.34. Carboxyfluorescein leakage upon addition of **1a-g,x** to POPC vesicles, 0.05 mM. Vesicles contained NaCl (451 mM, buffered with NaH_2PO_4 20 mM to pH 7.2, I.S. 500 mM and 50 mM CF) were suspended in Na_2SO_4 (150 mM, buffered with NaH_2PO_4 20 mM to pH 7.2, I.S. 500 mM). At t = 60 s addition of the anion carrier (10 mol% carrier to lipid). At t = 360 s addition of 20 µL of detergent. blank (12.5 µL DMSO). Each trace represents the average of at least three different trials, done with at least two different batches of vesicles.

Aiming to study the effect of pH in Cl^- transport capacity, different internal and external pH conditions were tested (**Figure 2.35**). The conditions for these ISE experiments were chosen to be comparable to those expected under physiological conditions in terms of pH values and gradients. The cages showed moderate to good chloride transport abilities, being the efficiency dependent on the nature of the side chains. Thus, in general, the most efficient transporter was **1c**, derived from 4-F-Phe, that is the most hydrophobic of the series containing Me in R_2. The specific behaviour in the presence of a pH gradient is also very noticeable (**Figure 2.39**). The use of acidic pH_i and neutral pH_e produced an increase of the chloride transport promoted by **1c** (X = F, R_2=Et), which was much lower with **1b** (X = OMe, R_2=Et) and negligible for **1a** (X = H, R_2=Et) and **1g** (X = OMe, R_2=Et) (compare orange and blue traces in **Figure 2.39**). The use of both acidic pH_i and pH_e increased the chloride transport rates of the four cages, though to a different extent (red traces in Figure 2.39). Thus, the fastest rate and highest effect of the acid medium was observed with the fluorinated cage **1c**.

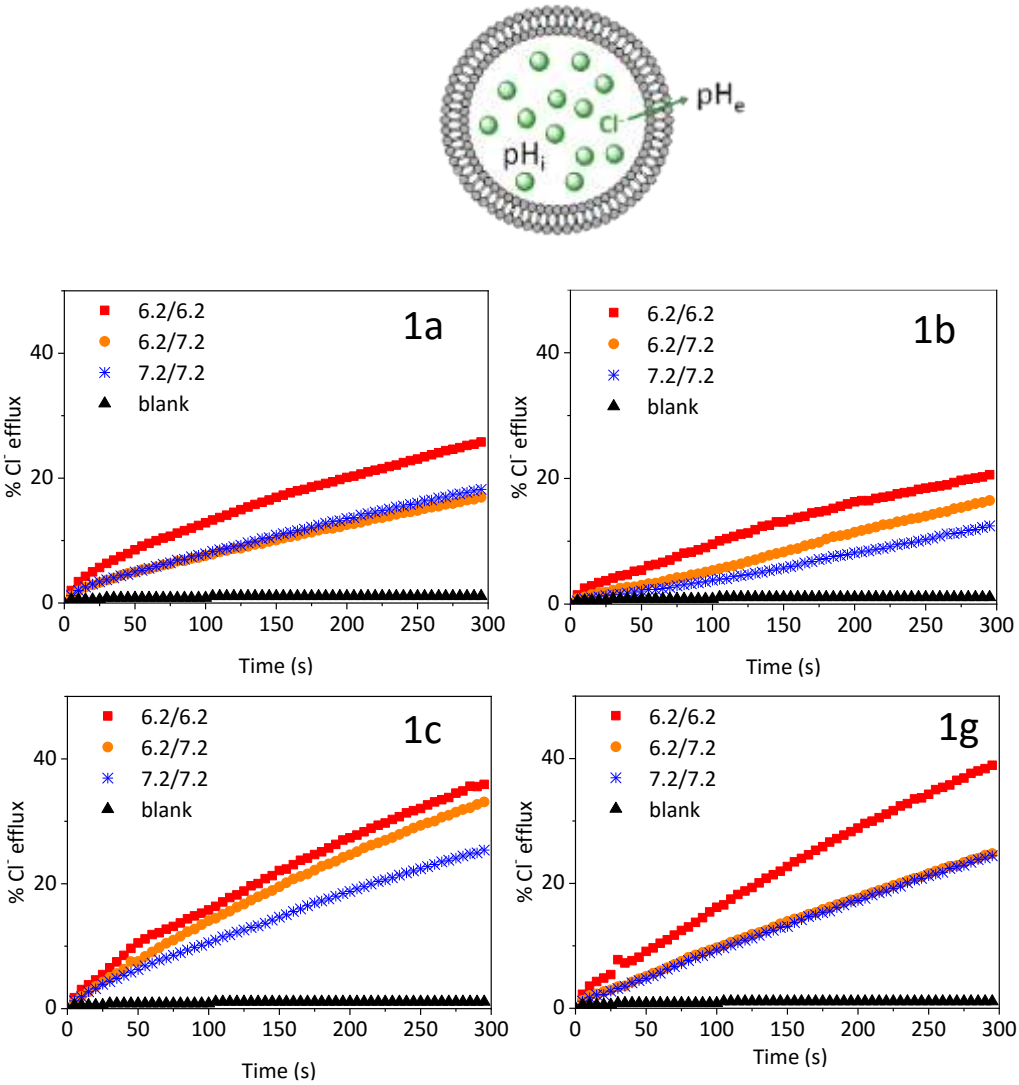


Figure 2.35. ISE experiments to measure the chloride efflux promoted by **1a-c, g** cages at different internal and external pH. Each plot shows the average of three experiments.

The highest transport capacity of **1c** was also confirmed by monitoring chloride influx using lucigenin, this dye has a fluorescence that is quenched in the presence of halides.[83] When measuring the Cl^- influx with the lucigenin based assay at low Cl^- concentrations in the extravesicular solution almost no transport was detected (Figure 2.81, experimental section). By switching the Cl^- concentration to 102.2 mM the most active compounds already detected by ISE experiments (**1c, 1b, 1g, 1a** and **1x**) showed higher transport rates with this technique too. The results obtained were comparable with those observed with

the ISE. **1c** is the most active Cl^- transporter followed by **1b**, **1g**, **1a** and **1x** (**Figure 2.36**)**.** The rest of the cages showed no transport activity.

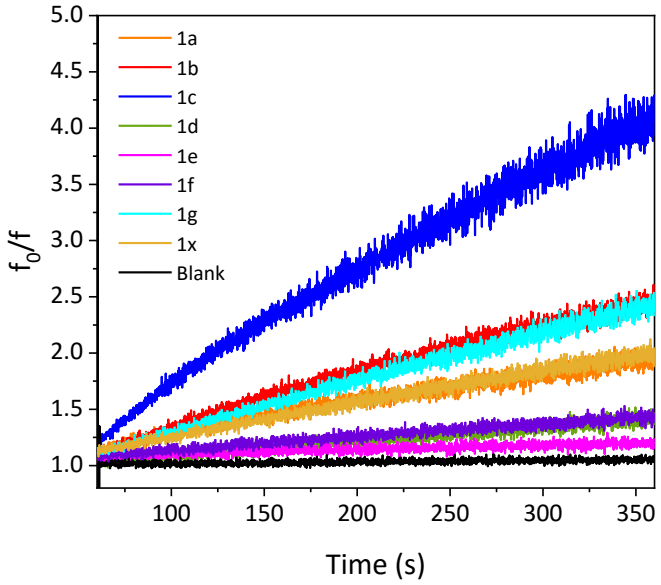


Figure 2.36. f_0/f normalization of lucigenin fluorescence emission upon addition of cages **1a-g,x** to POPC : Cholesterol (7 : 3) vesicles, 0.25 mM POPC. Vesicles contained $NaNO_3$ (102.2 mM $NaNO_3$, I.S. 150 mM, NaH_2PO_4 20 mM, pH 7.2; lucigenin 3 mM) were suspended in NaCl (102.2 mM NaCl, I.S. 150 mM, NaH_2PO_4 20 mM, pH 7.2). At t = 60 s the anion carrier was added (10 mol% carrier to lipid; 0.025 mM). Blank (12.5 µL MeOH). Each trace represents the average of at least three different trials, done with at least two different batches of vesicles.

Additional experiments with vesicles loaded with a fluorescent probe sensitive to pH (HPTS)[82,81] proved that these compounds efficiently dissipate pH gradients through lipidic bilayers (**Figure 2.37**). The protonated and deprotonated form of HPTS present different absorption bands while both emit at the same wavelength. Using a calibration curve it is possible to determine the vesicles internal pH using the fluorescence intensity relationship at 510 nm when the sample is excited at both 460 and 403 nm. HPTS experiments were conducted both using $NaNO_3$ and NaCl buffers and it was proven that the presence of Cl^- in the buffer enhances the capacity of these compounds to dissipate pH gradients. Cages complexating Cl^- have a different protonation profile and a higher rigidity that can be affecting the mobility of the complex trough the lipidic bilayer (section 2.5).

The relative efficiency was found well correlated to the relative activity as chloride transporters measured using ISE assays. Overall, the results suggest a H^+/Cl^- symport mechanism for the **1a-g,x** cages that is facilitated by acidic media and strongly dependent on the nature of the host side chain.

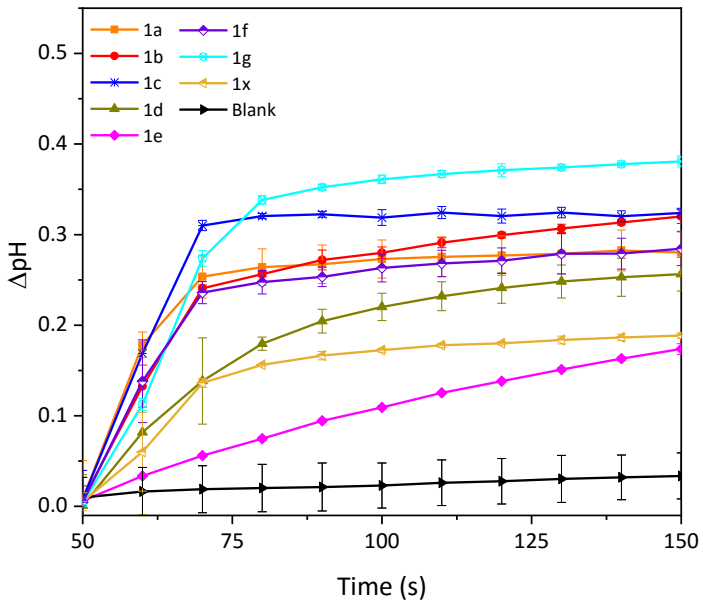


Figure 2.37. Schematic representation of HPTS transport assays with NaCl buffer. Variation of pH upon addition of compounds **1a-g**,**x** (10 mol%) to POPC : cholesterol (7: 3) vesicles, 0.5 mM POPC. Vesicles contained NaCl (126.25 mM NaCl, 10 mM buffer phosphate pH 6.5, I.S. 150 mM and HPTS 10 µM) were suspended in NaCl (126.25 mM NaCl, 10 mM buffer phosphate pH 7.5 and I.S. 150 mM). At t = 60 s addition of the anion carrier. (blank (12.5 µL DMSO)). Each trace represents the average of at least three different trials from at least two different batches of vesicles.

2.9 NMR EXPERIMENTS IN MICELLES

In order to obtain a more precise picture of the process at the molecular level, we carried out NMR studies of **1c** in an aqueous-lipid phase.[87,88] To that, deuterated-DPC micelles were suspended in D_2O and a solution of **1c** in DMSO-d_6 was added. Different conditions of the external aqueous phase were studied ranging from neutral to acidic pH, and also in the absence and presence of salt (Figure 2.38). The DOSY NMR experiments performed in all the samples showed self-diffusion rates compatible with the total incorporation of the cage within the micelles. Thus, **1c** showed an apparent hydrodynamic radius (r_H) of 2.9 nm, in a

very good agreement with the reported size of DPC micelles,[89] and much larger than the r_H of the cage when DOSY experiments were performed with **1c** in DMSO-d_6 in the absence of micelles (0.66 nm alone and 0.76 nm with protonated **1c** and TBACl). The observed changes of the NMR spectra of **1c** within the micelles upon variations in the composition of the external aqueous solution were also very illustrative. The proton NMR of the sample at pH 7.1 showed the presence of two set of signals for some of the protons, which were especially important for those of the aromatic side chains (Figure 3). Since there is always an unavoidable concentration of chloride in these samples (~2-3 mM for pH adjustment), we assigned these signals to the free and chloride-bound cage complexes in slow exchange in the chemical shift NMR timescale. The addition of external NaCl induced minor changes (line broadening) in the NMR spectra, as expected by the weak chloride binding of the non-protonated cage. Interestingly, when the pH of the bulk aqueous solution was lowered, the addition of NaCl produced a dramatic increase of the signals for the chloride-bound cage, demonstrating the ability of **1c** to strongly bind HCl within the lipid phase. Fortunately, the observation of split signals for the free and bound species allowed us to estimate the chloride exchange rate by EXSY experiments.[90] The sample in acidic medium showed a much faster chloride exchange (k_{ex} = 2.7 s^{-1} at 303 K) than the one at neutral pH (k_{ex} = 0.2 s^{-1} at 303 K). The difference in the exchange rate was observed in several signals of the cages (ESI). Accordingly, we concluded that the external aqueous acid medium produced a much stronger chloride binding and a much faster chloride exchange with a cage that remains in the lipid phase. To the best of our knowledge, this is the first direct experimental measurement of this type of processes occurring in an aqueous-lipid interface. Moreover, the fact that the aromatic signals of **1c** were strongly affected by the process also suggested the active participation of the fluorinated side chain, most likely through CH-anion or anion-π transient interactions.[91]

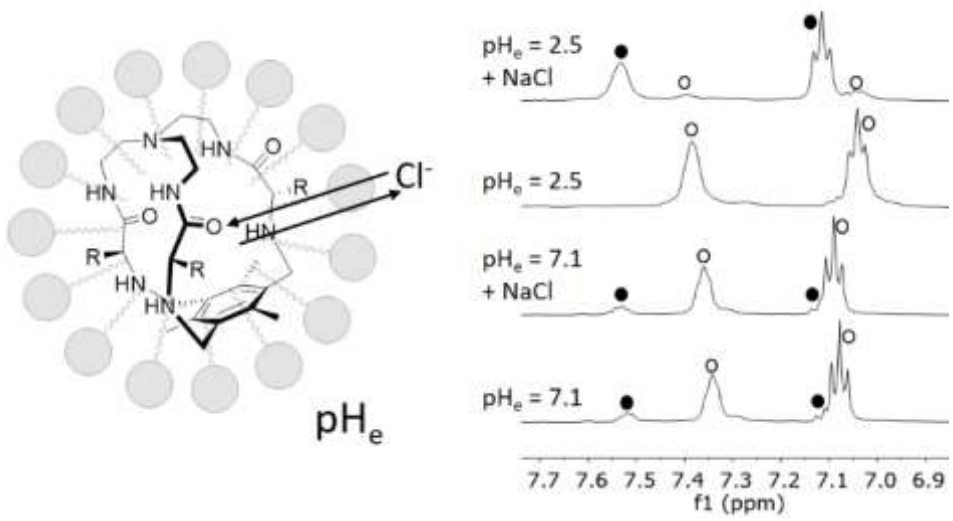


Figure 2.38. NMR experiments (500 MHz, 303 K) performed with **1c** (0.4 mM) in deuterated DPC micelles (25 mM, ~one molecule of 1c per DPC micelle) suspended in D_2O at different pH values and upon addition of 150 mM NaCl. The aromatic region showed signals for the free (○) and the chloride-bound (●) cage at slow exchange in the chemical shift NMR timescale.

2.10 BIOLOGICAL ASSAYS

Lung cancer is a common malignant solid tumour responsible of around 18 % of the total deaths by cancer world-wide.

The biological activity of the cages in living cells was tested in a human lung adenocarcinoma cell line (A549) as a proof of concept. Thus, the cytotoxicity of all the studied cages was determined using the MTT assay[92] (Figure 2.39).

2.10.1 Cytotoxicity of Compounds 1a-g,x (MTT 24H)

Cells were seeded 24 hours prior to treatment in 96-well plate (100 µL of a suspension $2.5 \cdot 10^5$ cell/mL). The following day, culture media from the wells was replaced by 100 µL of fresh media at each pH containing the desired concentration of the compound to be tested. All the compounds were previously dissolved in DMSO at a concentration of 20 mM. The final concentration of DMSO used in the corresponding wells did not exceed 1% (v/v). This concentration does not affect cell viability. Negative control cultures received the same concentration of solvent alone.

Cells were incubated for 24h in the presence of compounds **1a-g,x** at 37°C in a humidified atmosphere with 5% CO_2. At the end of incubation, the culture media was removed and 100 μL of MTT solution (5mg/mL diluted with plain culture media 1 : 5) was added to each well and incubated for 4 hours. Afterwards, MTT solution was discarded. The purple formazan crystal formed at the bottom of the wells was dissolved with 100 μL of DMSO and shaken for 30 minutes at room temperature. The absorbance at 570 nm was read on a spectrophotometer plate reader.

2.10.2 Cytotoxicity at different extracellular pH.

In order to test our hypothesis, we performed the experiments at different pH_e values, namely using the standard conditions (DMEM medium, pH_e 7.5) and with the presence of PIPES buffer to fix a slightly more acidic external medium (pH_e 7.2 and 6.2).[93] From the cages having a Me in R_2, compounds **1a-c** displayed cytotoxic activities with interesting pH dependence, whereas **1d, 1e** and **1f** were found essentially nontoxic. The Phe derivative rendered IC_{50} values unaffected by pH_e, while the O-Me-Tyr counterpart showed a slightly higher activity as the pH_e was lower. The performance of **1c** bearing the 4-F-Phe derivative was especially interesting, since this receptor systematically increased the cytotoxicity as the pH_e was decreased (Figure 2.39). Most remarkably, **1c** showed to be ~five-fold more cytotoxic at acidic pH_e than under the conventional conditions. This allows a relatively wide concentration window (50-150 μM) for which **1c** would be safe for cells surrounded by weakly basic pH (healthy cells conditions), but cytotoxic for cells in an acidic microenvironment (such as for solid tumours).

Derivatives with Et group in R_2 displayed a similar pH behaviour to their analogues with Me in R_2 but with lower cytotoxicity. Compound **1x** showed some cytotoxicity with a very low pH dependency in the 7.6-6.2 pH range whereas compounds **1f** and g were essentially non cytotoxic at 200 μM. The increased lipophilicity of the cages with an Et group in R_2 might be also affecting these cages solubility and bioavailability thus reducing their biological activity.

The trends observed for the biological activity at different pH_e values correlated well with the trends observed for the transport assays in model liposomes. The only compound whose cytotoxicity does not correlate with its transport capacity is **1g** which displays Cl^- transport through POPC vesicles and yet didn't induce cytotoxicity but this had already been associated to the high lipophilicity of the cage. These results are in good agreement

with a key role of the ionophoric activity of compounds **1a-c** in the observed biological activity.

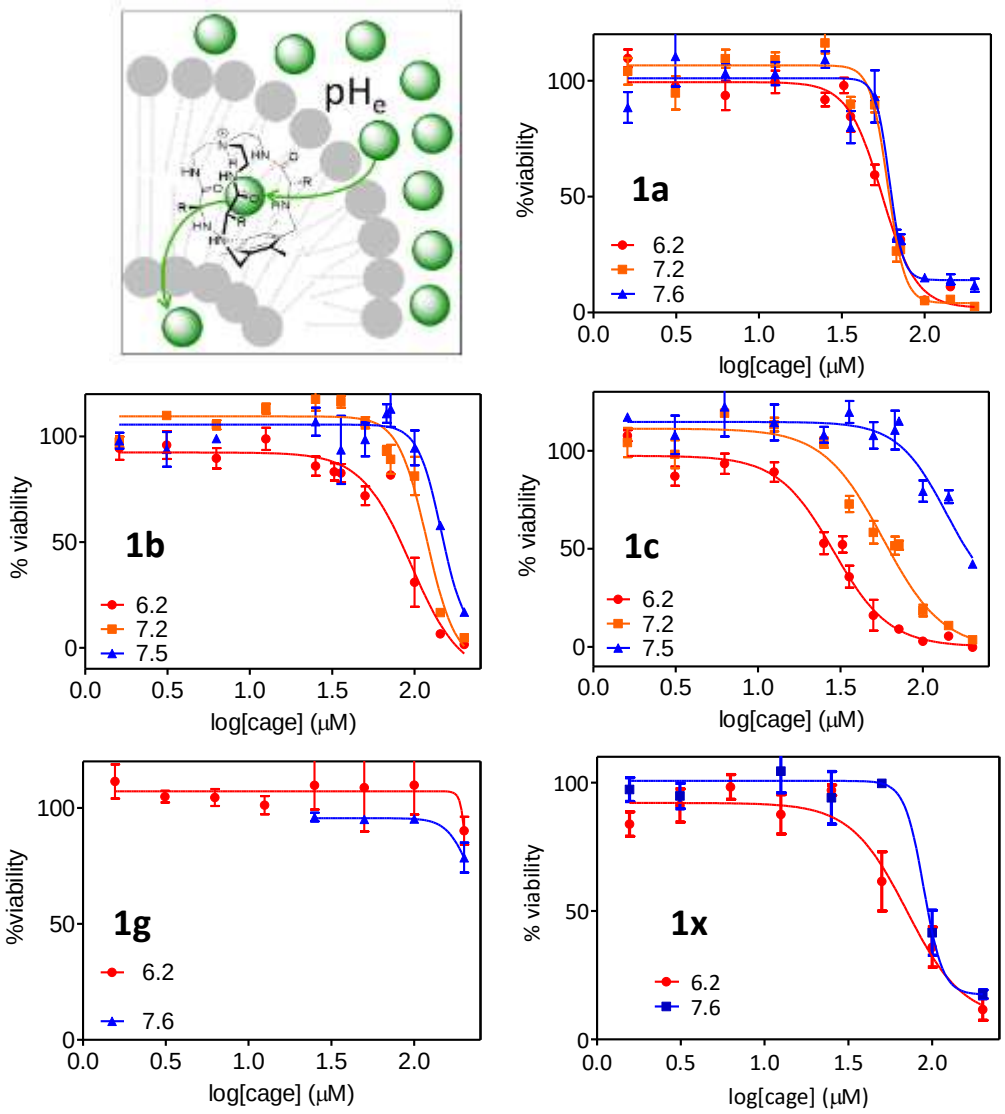


Figure 2.39. Schematic representation of the cytotoxicity experiments (A549 cell line) and plots of the percent cell viability versus the concentration of the cages **1a-c, 1g, 1x** performed at different pH_e values (referenced to their corresponding blanks, see ESI). The results correspond to the average of two independent experiments performed in triplicates each (error bars show standard deviation).

2.10.3 Cytotoxicity of 1c at different concentrations of Cl^-

To test the implication of Cl^- transport in the cell death mechanism we carried MTT assays with **1c** at a reduced concentration of chloride and slightly acidic pH (Figure 2.40). The lowest Cl^- concentration used was 14 mM which is similar to the Cl^- concentration inside human cells.[94] These assays were conducted In DPBS with different concentrations of Cl^-, prepared by replacing NaCl with sodium gluconate. pH was adjusted before each experiment to 6.4 by adding NaOH or HCl. The rest of the conditions in this MTT assay were unaltered.

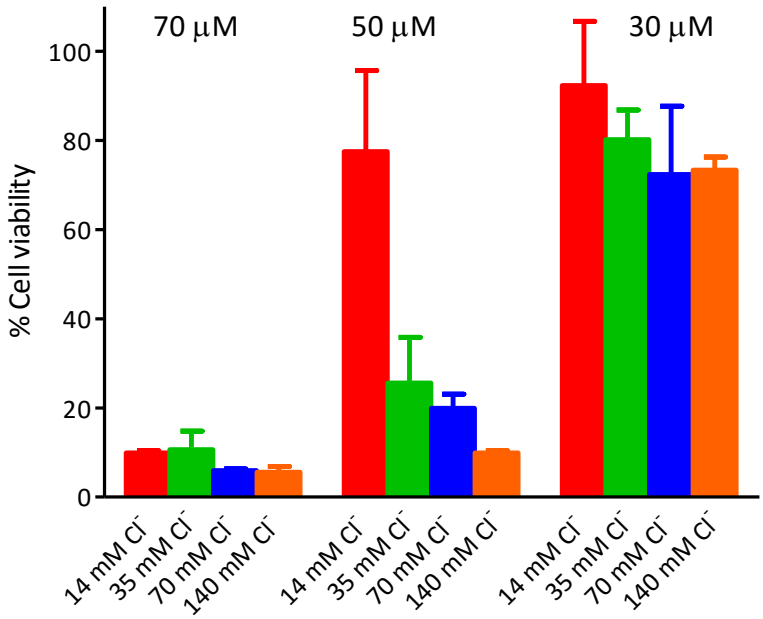


Figure 2.40. Cytotoxicity of compound 1c (70, 50 and 30 μM) vs. A549 cells determined by the MTT assay after 24h of incubation in DPBS containing different concentrations of Cl^- (14, 35, 70, 140 μM) at pH 6.4.

The reduction of the Cl^- concentration in the cell culture buffer decreased the cytotoxicity of compound **1c** (Figure 2.40). This was especially significant with 50 μM of **1c**. Overall, these results are in good agreement with a key role of the ionophoric activity of compounds **1a–c** in the observed biological activity.

2.10.4 Summary table

Table 2. 8. Chloride affinity (log$K_{1:1}$, log$K_{1:2}$, BC_{50}°), hydrophobicity (calculated log*P*, HPLC retention time), chloride efflux rates in model liposomes (%Cl·s^{-1}, at different pH_i / pH_e) and MTT cytotoxicity in A549 cell lines (IC_{50} in μM at different external pH) for the cages **1a-g,x**.

					Chloride efflux rate (%Cl·s^{-1}), pH_i / pH_e			Cytotoxicity to A549, IC_{50} (μM)		
cage	**log$K_{1:1}$ log$K_{1:2}$**	**BC_{50}° (μM)[b]**	**log*P*[c]**	**t_R (min)[d]**	**7.2 / 7.2**	**6.2 / 7.2**	**6.2 / 6.2**	**pH_e 7.5**	**pH_e 7.1**	**pH_e 6.2**
1a	4.01(8) 2.2(2)	96±17	3.9	8.02	0.092	0.091	0.159	60±4	60±3	55±5
1b	3.7(1) 2.2	198±44	3.63	8.43	0.039	0.055	0.078	145±35	118±16	95±25
1c	3.82(3) 2.2(1)	148±10	4.31	9.44	0.121	0.151	0.198	166±35	58±10	29±4
1d	n.d.[e]	n.d.[e]	2.81	5.07	0.059	0.1	0.16	>200	>200	>200
1e	3.40(3) 2.1(1)	380±24	4.04	8.15	0.14	0.05	0.128	>200	>200	~170
1f	n.d.[f]	n.d.[f]	5.15	9.46	0.09	0.16*	0.14	>200	>200	>200
1g	3.92(1) 6.4[g]	112±30	4.75	9.95	0.098	0.093	0.17	>200	>200	>200
1x	3.37(9) 1.8(1)	417±83	4.59	9.64	0.06	0.06	0.18*	88±17	88±22	70±38

[a] By NMR titration, standard deviation on the last significant figure in parenthesis. [b] As defined in ref.[78]. [c] Calculated using VCCLab software with the cages as free amines. [d] Retention time in HPLC C18 reverse phase (see ESI for details). [e] Partial precipitation precluded the accurate fitting of the titration data. [f] Data could not be fitted to the 1 : 1/1 : 2 interaction model proposed. [g] Excessive error in the K. * abrupt change in the first 60 seconds followed by a flat curve.

2.11 CONCLUSIONS

- We have successfully synthetized small pseudopeptidic cages able to efficiently pump HCl through artificial phospholipidic bilayers that simulate the aqueous-lipid interfaces of cell membranes.
- The specific behaviour of these HCl transporters strongly depend on the nature of the amino acid side chains which is responsible for the different lipophilicity, HCl transport capacity and cytotoxicity properties between cages. Since the Cl^- binding site of the cage is preserved between cages **1a-g,x** these physicochemical differences were associated with the possible non-covalent interactions between the side chains and the lipid-aqueous interfaces.
- A fluorinated Phe derivative (**1c**) showed a remarkably increased HCl transport ability in model POPC vesicles and a higher cytotoxicity to A549 cancer cells in the presence of pH gradients resembling those characteristic of tumour microenvironments (acidic pH_e and slightly basic pH_l).
- The results presented in this chapter pave the way to further design new pH-dependent anionophores for applications in cancer chemotherapy, with the aim to improve their selectivity for the target.

Most of the work presented in this chapter was successfully published in Angewandte Chemie.

Tapia, L.; Pérez, Y.; Bolte, M.; Casas, J.; Solà, J.; Quesada, R.; Alfonso, I. PH-Dependent Chloride Transport by Pseudopeptidic Cages for the Selective Killing of Cancer Cells in Acidic Microenvironments. *Angew. Chemie - Int. Ed.* **2019**, *58* (36), 12465–12468.

2.12 EXPERIMENTAL SECTION

2.12.1 General

Reagents and solvents were purchased from commercial suppliers (Adrich, Fluka or Merck) and were used without further purification. Compounds **1a** and **1x** were synthetized as previously described.[76] All the compounds prepared were fully characterized by the complete spectroscopic (NMR, ESI-MS) and analytical data. Preparative reverse phase purifications were performed on an Isolera Biotage instrument (KP-C18-HS, CH_3CN and water with 0.1% TFA). Analytical RP-HPLC was performed with a Hewlett Packard Series 1100 (UV detector 1315A) modular system using a reverse-phase Kromasil 100 C8 (15 x 0.46 cm, 5 µm) column. CH_3CN-H_2O mixtures containing 0.1% TFA at 1 mL/min were used as mobile phase and monitoring wavelengths were set at 220, 254 and 280 nm.

2.12.2 Instrumental techniques

2.12.2.1 NMR spectroscopy

The NMR experiments were carried out at 25˚C on a VNMRS-400 NMR spectrometer (Agilent Technologies 400 MHz for ^{1}H and 100 MHz for ^{13}C) for characterization and a Bruker Avance-III 500 MHz spectrometer equipped with a z-axis pulsed field gradient triple resonance (^{1}H, ^{13}C, ^{15}N) TCI cryoprobe (500 MHz for 1H and 125 MHz for 13C) for titrations. Chemical shifts are reported in ppm using tetrakis-trimethylsilylsilane as a reference. Data were processed with the software program MNova (Mestrelab Research).

2.12.2.2 ESI mass spectrometry

High resolution mass spectra (HRMS) were performed on Acquity UPLC System and a LCT PremierTM XE Benchtop orthogonal acceleration time-of-flight (oa-TOF) (Waters Corporation, Milford, MA) equipped with an electrospray ionization source. All sample solution (in the 1 x 10-4 to 1 x 10-6 M range) were prepared in methanol.

2.12.2.3 X-RAY diffraction

Data for all structures were collected on a STOE IPDS II two-circle diffractometer with a Genix Microfocus tube with mirror optics using MoKα radiation (λ = 0.71073 Å). The data were scaled using the frame scaling procedure in the X-AREA program system.[102] The structures were solved by direct methods using the program SHELXS [95,96] and refined against F2 with full-matrix least-squares techniques using the program SHELXL[97]

2.12.2.4 NMR spectroscopy in micelles

All data were acquired using a Bruker Avance-III 500 MHz spectrometer equipped with a z-axis pulsed field gradient triple resonance (^{1}H, ^{13}C, ^{15}N) TCI cryoprobe. For all NMR experiments with cages in presence of DPC-d_{38} micelles, we worked at 25 mM monomer concentration, well above cmc (1.36 mM in PBS). Due to the limited solubility of pseudopeptidic host in aqueous solution, the samples in DPC/D_2O solution have a final concentration of 0.4 mM cage (from a weighted stock solution in DMSO-d_6) with a ~1 : 1 cage : DPC micelles ratio (using known aggregation number ~ 55-60).

2.12.2.5 Fluorescence

Fluorescence emission was recorded using a HITACHI F-7000 Fluorescence spectrophotometer equipped with stirrer. Two different types of cuvettes were used depending on the assay: standard 10 mm quartz glass cells from Hellma Analytics High Precision Cell made of Quartz SUPRASIL or disposables MAPM-F10-100 labbox cuvettes. All measurements were performed at 25 °C unless specified.

2.12.2.6 Chloride Selective Electrode measurements

Chloride concentration in transport experiments was recorded using a Combination Chloride Electrode HI4107 Hanna Instruments and a Chloride Selective Electrode 96 52 C Crison.

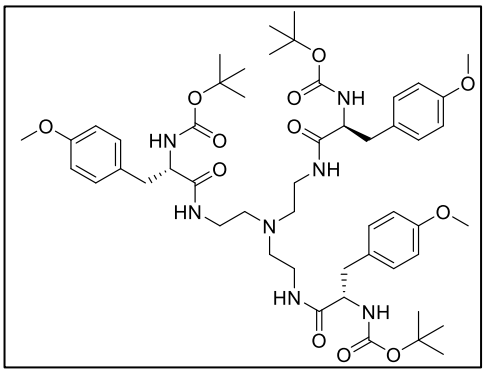
2.12.2.7 pH measurements

Were made using a CRISON pHmeter 50 14 T.

2.12.3 Synthesis of compounds 1b-g

***2.12.3.1** Synthesis of 1a*

2b: tri-tert-butyl((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(3-(4-methoxyphenyl)-1-oxopropane-1,2-diyl))tricarbamate



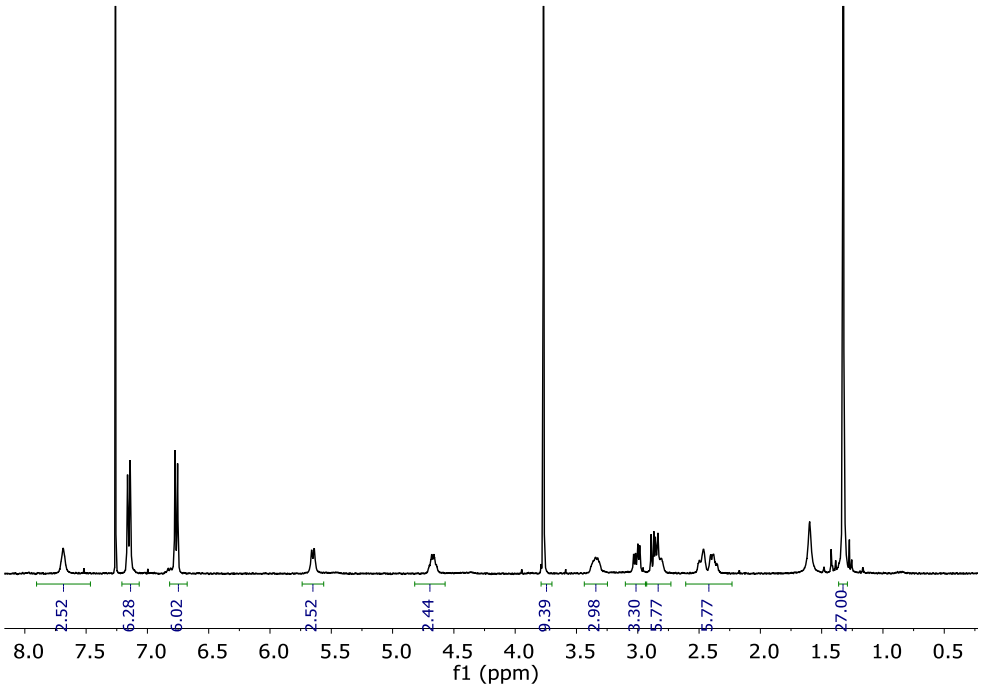
Boc-Tyr(Me)-OH (375 mg, 1.27 mmol) was dissolved in dry DMF (2.5 mL). N-(3 Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC·HCl O.394 mg, 1.53 mmol), 1-Hydroxybenzotriazole hydrate (HOBt, 0.207 mg, 1.53 mmol), N,N-Diisopropylethylamine (DIPEA, 0.8 mL, 4.6 mmol) and tris(2-

aminoethyl)amine (0.06 mL, 0.38 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using DCM : MeOH 95 : 5 to give 0.298 mg of **2b** (0.307 mmol, 80% yield).

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm) = 7.69 (s, 3H), 7.15 (A subsystem from AB, J_{AB}=8.3 Hz, 6H), 6.76 (B subsystem from AB, J_{AB}=8.4 Hz, 6H), 5.65 (d, J=8.8 Hz, 3H), 4.67 (X subsystem from ABX, J_{AX}=6.3, J_{BX}=8.6 Hz, 3H) 3.77 (s, 9H), 3.35 (m, 3H), 3.01 (A subsystem from ABX, J_{AB}=13.8, J_{AX}=6.3 Hz, 3H), 2.86 (dd, B subsystem from ABX, J_{AB}=13.8, J_{BX}=8.6 Hz, 3H) 2.84 (m, 3H), 2.48 (m, 3H), 2.39 (m, 3H), 1.32 (s, 27H).

^{13}C NMR (101 MHz, $CDCl_3$): δ(ppm)=189.3, 181.1, 172.9, 158.2, 158.2, 130.3, 113.6, 55.6, 55.1, 54.9, 38.7, 38.3, 28.3. (Data obtained from HSQC and HMBC spectra).

HRMS (ESI-TOF) m/z [**2b** + H]$^+$ Calc.: 978.5546, found: 978.6500.



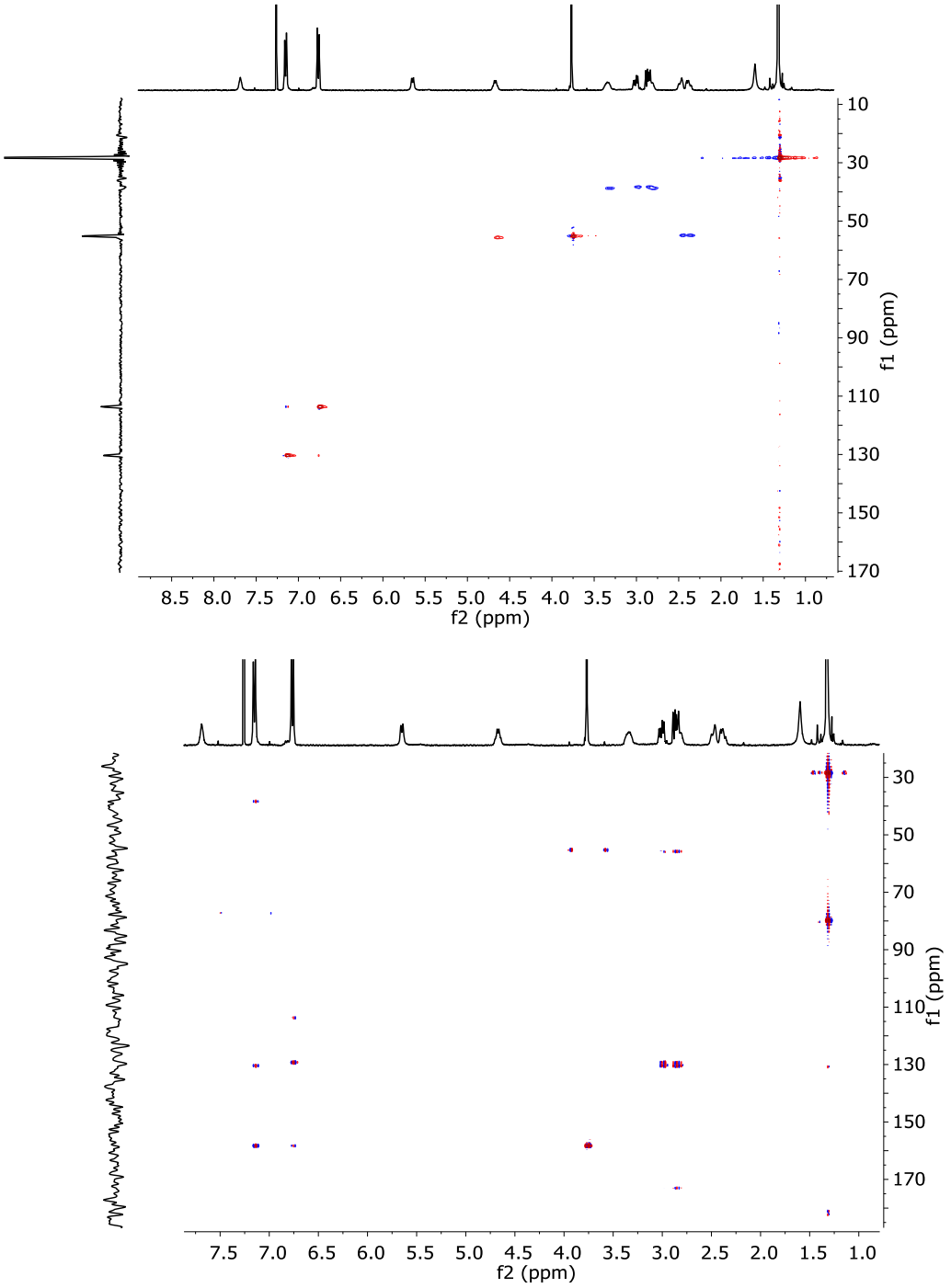


Figure 2.41. ^{1}H-NMR (400 MHz, CD_3OD), HSQC and HMBC spectra of 2b

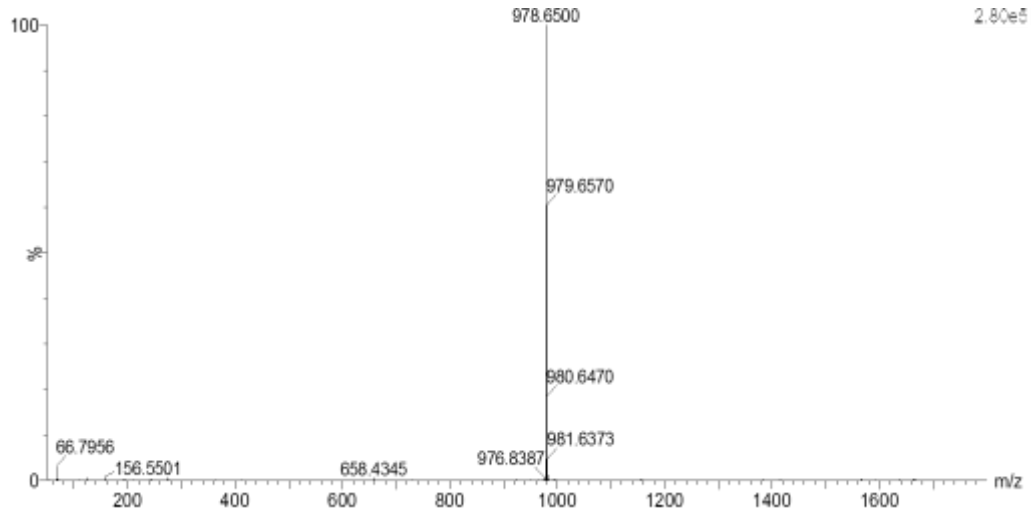
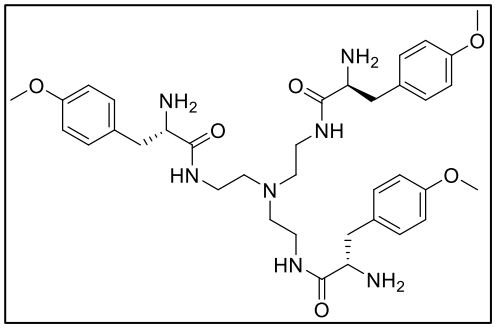


Figure 2.42. HRMS (ESI+) experimental spectrum of 2b.

3b: (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(4 methoxyphenyl) propanamide)



2b (153 mg, 0.156 mmol) was dissolved in DCM (1.5 ml). Triethylsilane (0.35 mL, 2.34 mmol) and trifluoroacetic acid (1 mL) were then added. The solution was stirred at room temperature during 3 hours and then solvents were evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3b**·4TFA as a white solid (166 mg, 0.145 mmol, 93% yield). The resulting solid was dissolved in NaOH 1M and extracted with ethyl acetate to eliminate the TFA and obtain the free amine.

^{1}H NMR (400 MHz, CD_3OD): δ(ppm)=7.18 (d, *J*=8.7 Hz, 6H), 6.91 (d, *J*=8.7 Hz, 6H), 4.03 (X subsystem from ABX, J_{AB}=7.3, J_{BX}=7.4 Hz, 3H), 3.79 (s, 9H), 3.29 (m, 6H), 3.12 (B subsystem from ABX, J_{AB}=14.1, J_{BX}=7.4 Hz, 3H), 3.00(A subsystem from ABX, J_{AB}=14.1, J_{AX}=7.3 Hz, 3H), 2.65 (m, 6H).

^{13}C NMR (101 MHz, CD_3OD): δ(ppm)= 197.9, 187.4, 158.7, 154.1, 142.5, 83.4, 82.9, 81.3, 64.7, 63.5 .

HRMS (ESI-TOF) m/z [**3b** + H]$^+$ Calc.: 678.3974, found: 678.4558.

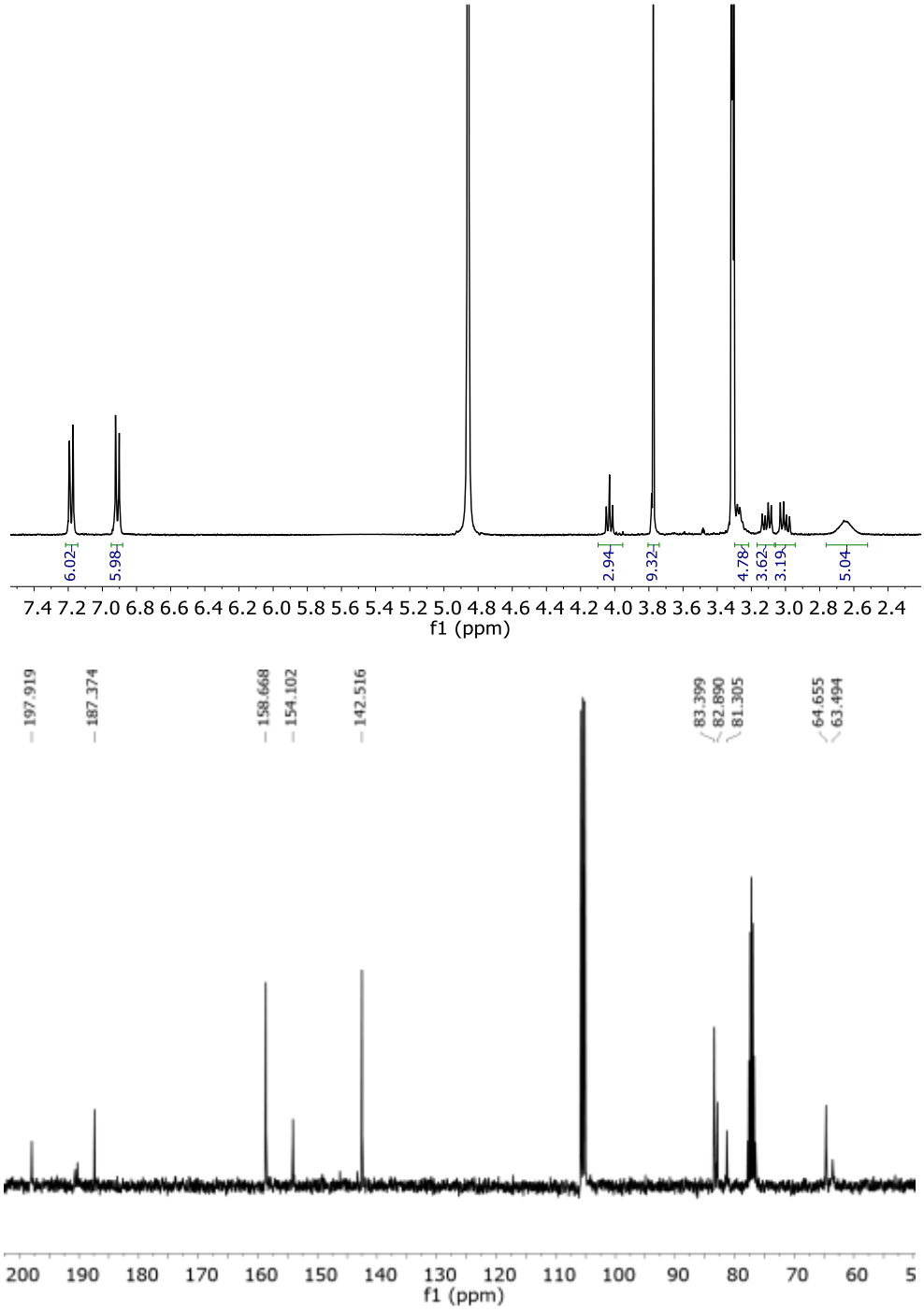


Figure 2.43. ^{1}H-NMR (400 MHz, CD_3OD) spectra of 3b·4TFA. and ^{13}C-NMR (101 MHz, CD_3Cl) spectra of **3b**.

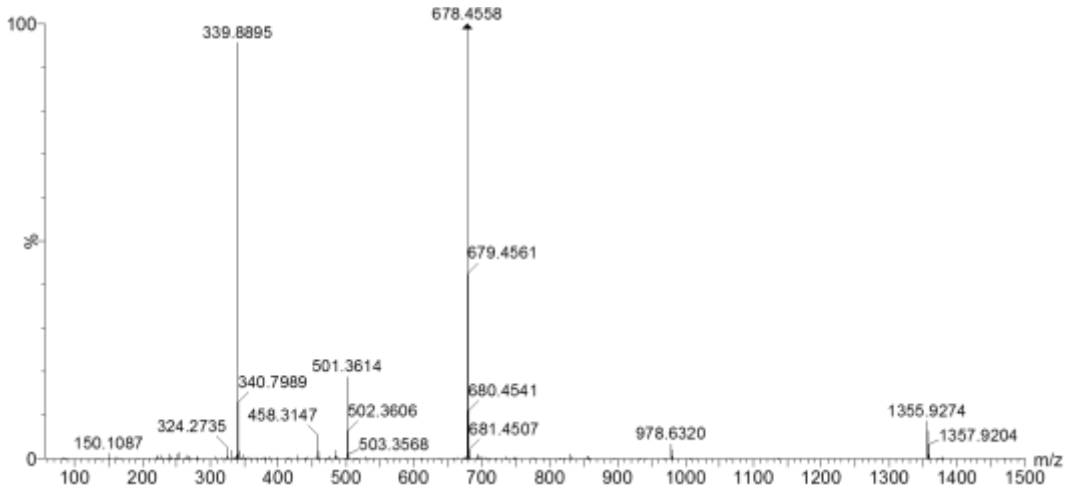
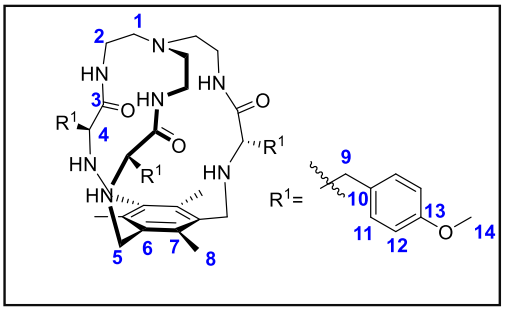


Figure 2.44. HRMS (ESI+) experimental spectrum of 3b·4TFA.

1b

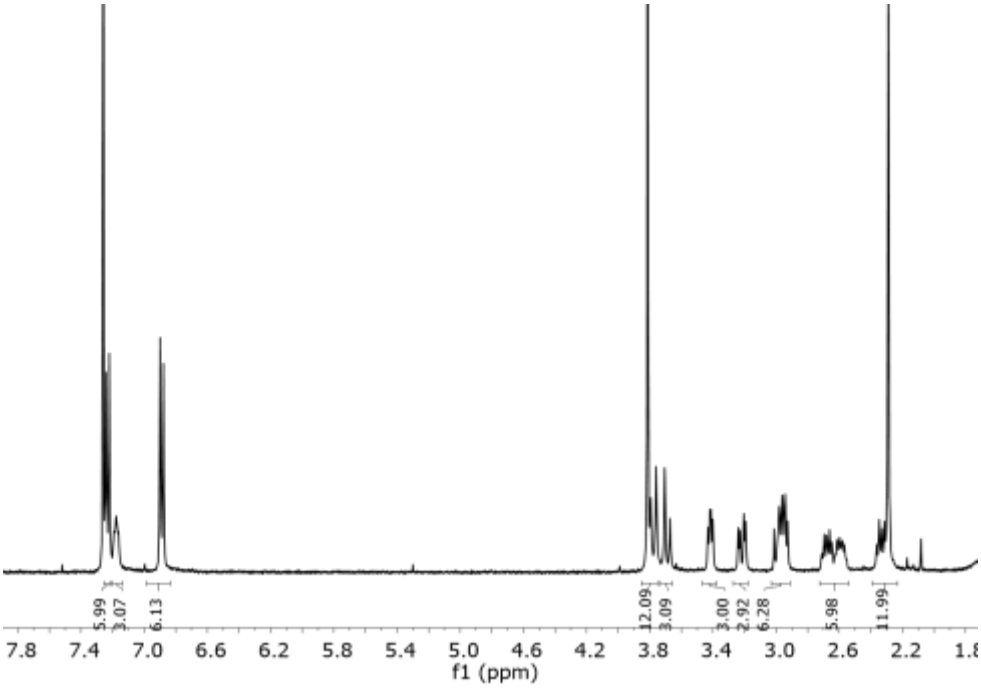


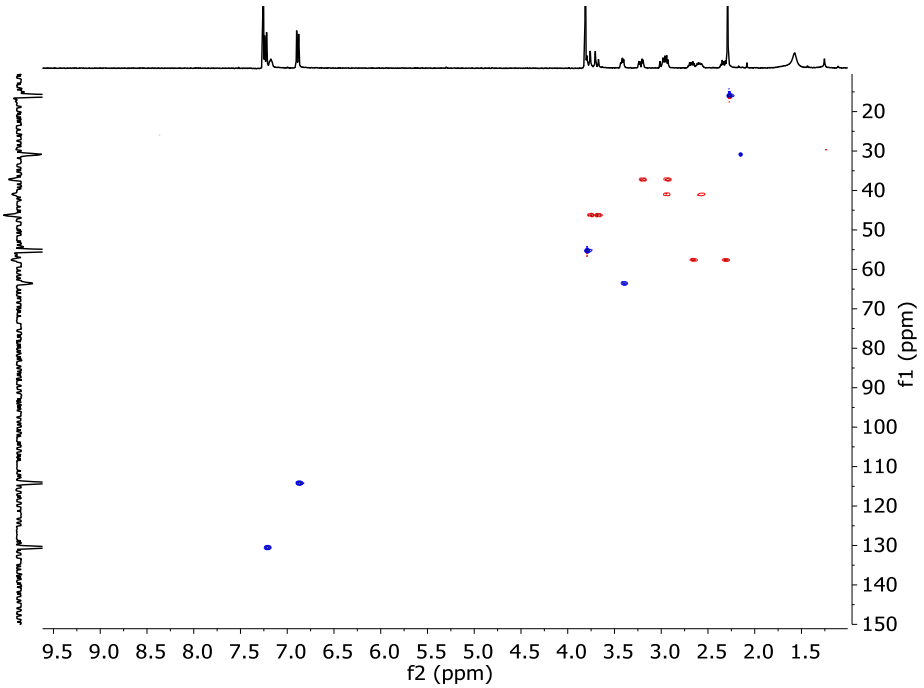
3b (143.2 mg, 0.126 mmol) was dissolved in dry acetonitrile (50 mL). Bu_4Cl (34 mg, 0.076 mmol) and K_2CO_3 (349 mg, 3 mmol) were added to the solution. 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (50 mg, 0.126 mmol) was dissolved in acetonitrile (10 mL) and the solution was added to the reaction containing **3b**. Reaction mixture refluxed for 16 hours. Solvent was evaporated and the resulting crude was purified by flash chromatography The residue was purified by flash chromatography using DCM:MeOH 95:5 as eluent to give **1b (**43 mg, 50 mmol, 40% yield).

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)=7.23 (d, *J*=8.6 Hz, 6H, H**12**), 7.18 (m, 3H, NH), 6.89 (d, *J*=8.6 Hz, 6H, H**11**), 3.81 (s, 9H, H**14**), 3.73 (AB_q, δ_A=3.77 , δ_B=3.69, J_{AB}=13.6 Hz, 6H, H**5**), 3.42 (X subsystem from ABX, J_{AX}=7.5, J_{BX}=4.6 Hz, 3H, H**4**), 3.22 (B subsystem from ABX, J_{AB}=4.0, J_{BX}=4.6 Hz, 3H, H**9**), 2.97 (m, 6H, H**2** and A subsystem from ABX H**9** overlapped), 2.65 (m, 6H, H**1** H**2**), 2.34 (m, 3H, H**1**), 2.29 (s, 9H, H**14**).

^{13}C NMR (100 MHz, $CDCl_3$): δ(ppm)=174.2 (C**3**), 158.5 (C**7**), 135.1 (C**6**), 130.7 (C**12**), 130.5 (C**10**),129.6 (C**13**), 114.3 (C**11**), 63.5 (C**4**), 57.6 (C**1**), 55.3 (C**14**), 46.3 (C**5**), 41.0 (C**2**), 37.1 (C**9**), 15.3 (C**8**). (Data obtained from HSQC and HMBC spectra)

HRMS (ESI-TOF) m/z [**1b** + H]$^+$ Calc.: 834.4913, found: 834.5082.





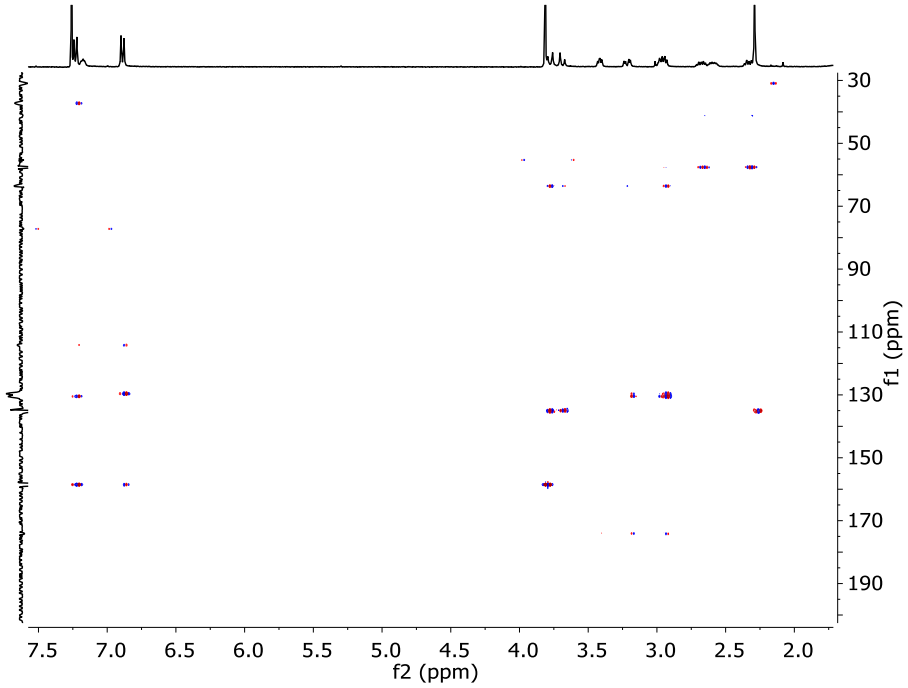


Figure 2.45. ^{1}H-NMR (400 MHz, CD_3OD), HSQC and HMBC spectra of 1b.

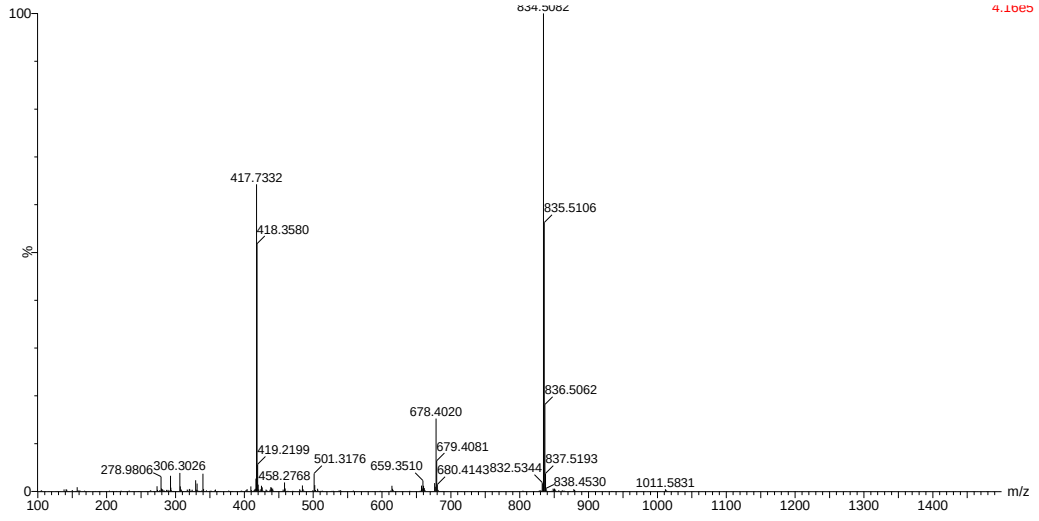
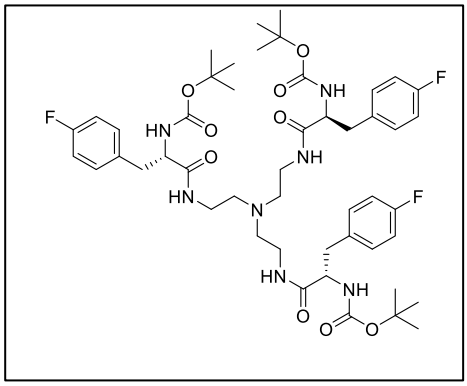


Figure 2.46. HRMS (ESI+) experimental spectrum of 1b.

2.12.3.2 Synthesis of 1c

2c: tri-tert-butyl ((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(3-(4-fluorophe nyl)-1-oxopropane-1,2-diyl))tricarbamate



Boc-Phe(4-F)-OH (299 mg, 1.057 mmol) was dissolved in dry DMF (2.5 mL). *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC·HCl O.254 mg, 1.28 mmol) and 1-Hydroxybenzotriazole hydrate (HOBt, 0.196 mg, 1.28 mmol), *N*,*N*-diisopropylethylamine (DIPEA, 0.67 mL, 3.84 mmol) and tris(2-aminoethyl)amine (0.05 mL, 0.32 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using 95:5 DCM:MeOH to give 0.254 mg of **2c** (0.239 mmol, 85% yield).

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)=7.72 (s, 3H), 7.20 (dd, *J*=8.5, 5.5 Hz, 6H), 6.90 (t, *J*=8.7 Hz, 2H), 5.71 (d, *J*=9.0 Hz, 3H), 4.70 (X subsystem from ABX J_{AX}=13.7, J_{BX}=8.8 Hz, 3H), 3.35 (m, 3H), 3.02 (A subsystem from ABX, J_{AX}=13.7, J_{AB}=6.4 Hz, 3H), 2.87 (B subsystem from ABX J_{BX}=8.8 Hz J_{AB}=6.4 Hz, 3H), 2.79 (m, 3H), 2.43 (m, 6H), 1.32 (s, 27H).

^{13}C NMR (101 MHz, CD_3Cl): δ(ppm)=173.0, 156.4, 131.1, 131.0, 115.2, 115.0, 80.1, 55.8, 55.1, 39.0, 38.7, 28.4.

HRMS (ESI-TOF) m/z [**2c** + H]$^+$ Calc.: 942.4947, found: 942.4931.

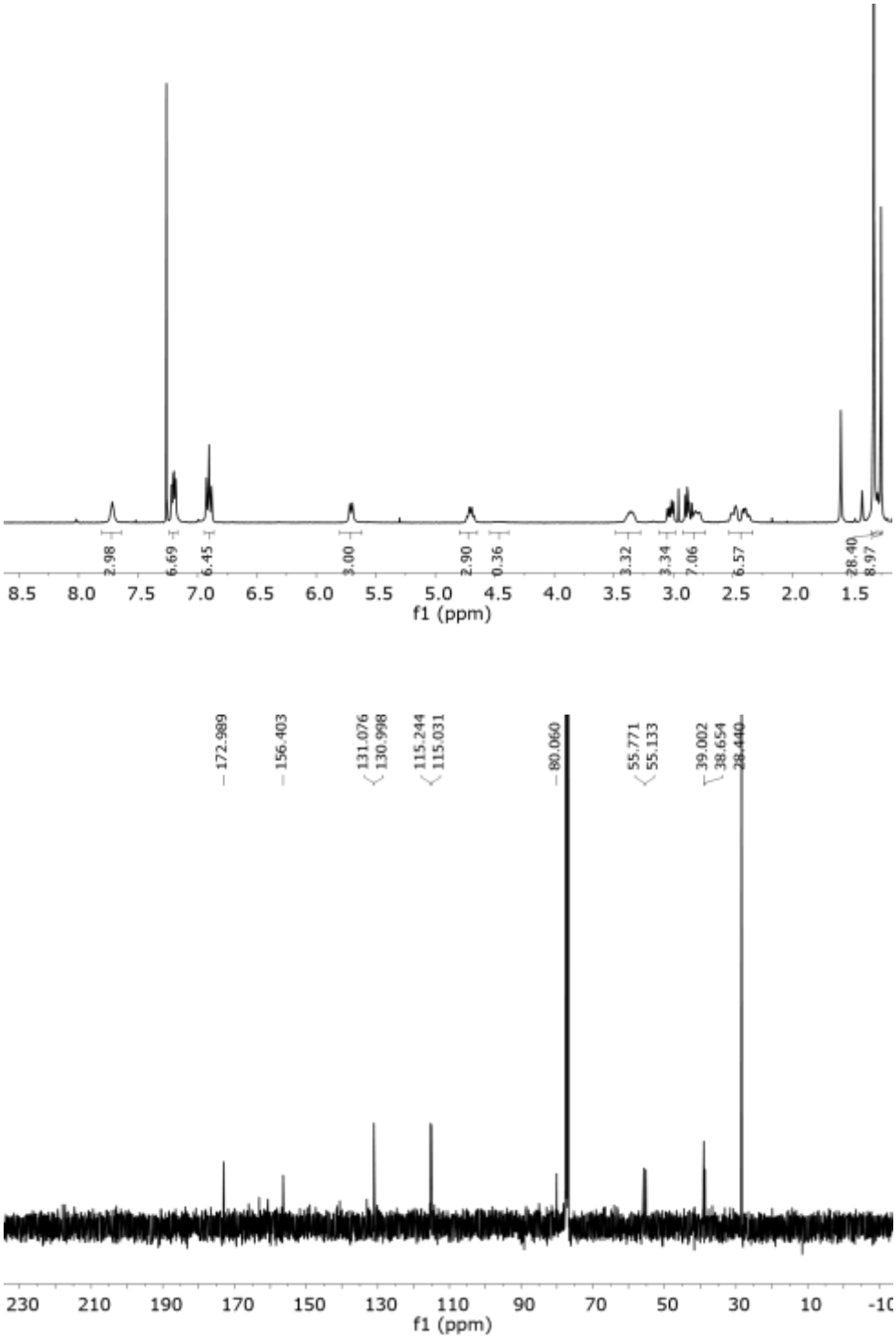


Figure 2.47. ^{1}H-NMR (400 MHz, CD_3Cl) and ^{13}C-NMR (101 MHz, CD_3Cl) spectra of **2c.**

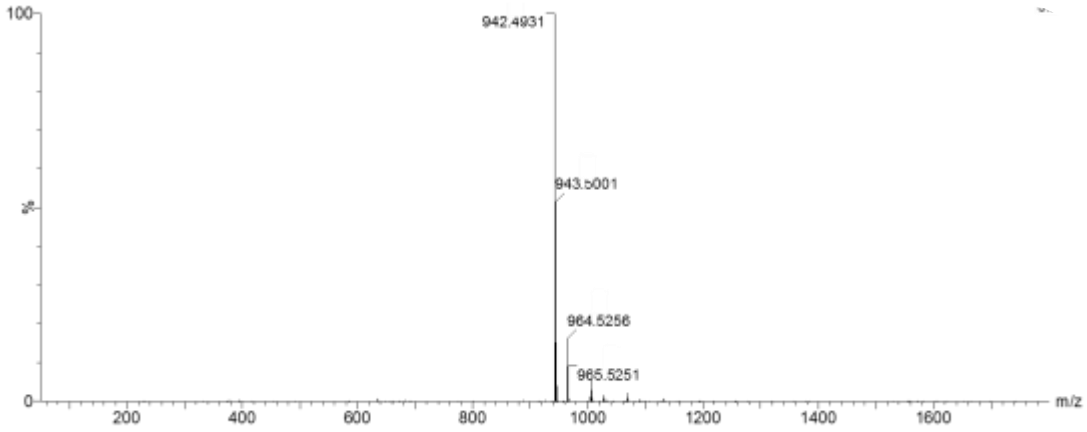
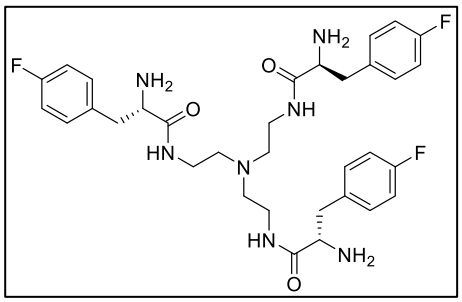


Figure 2.48. HRMS (ESI+) experimental spectrum of 2c.

3c: (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(4-fluorophenyl)propanamide)



2c (173 mg, 0.183 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.414 mL, 2.756 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3c·4TFA** as a white solid (166 mg, 0.145 mmol, 93% yield).

^{1}H NMR (400 MHz, CD_3OD): δ(ppm)=7.33 (dd, *J*=8.5, 5.2 Hz, 6H), 7.09 (t, *J*=8.7 Hz, 6H), 4.13 (X subsystem from ABX, J_{AX}=7.2, J_{BX}=7,6 Hz, 3H), 3.33 (m, 6H), 3.16 (A subsystem from ABX , J_{AX}=7.2, J_{AB}=14 Hz, 3H), 3.07 (B subsystem from ABX, J_{BX}=7.6, J_{AB}=14 Hz, 3H) 2.59 (m, 6H).

^{13}C NMR (101 MHz, CD_3OD): δ(ppm)=168.5, 131.0, 130.1, 115.5, 115.3, 54.3, 52.6, 36.4.

HRMS (ESI-TOF) m/z [**3c**+ H]$^{+}$ Calc.: 642.3374, found: 642.3666.

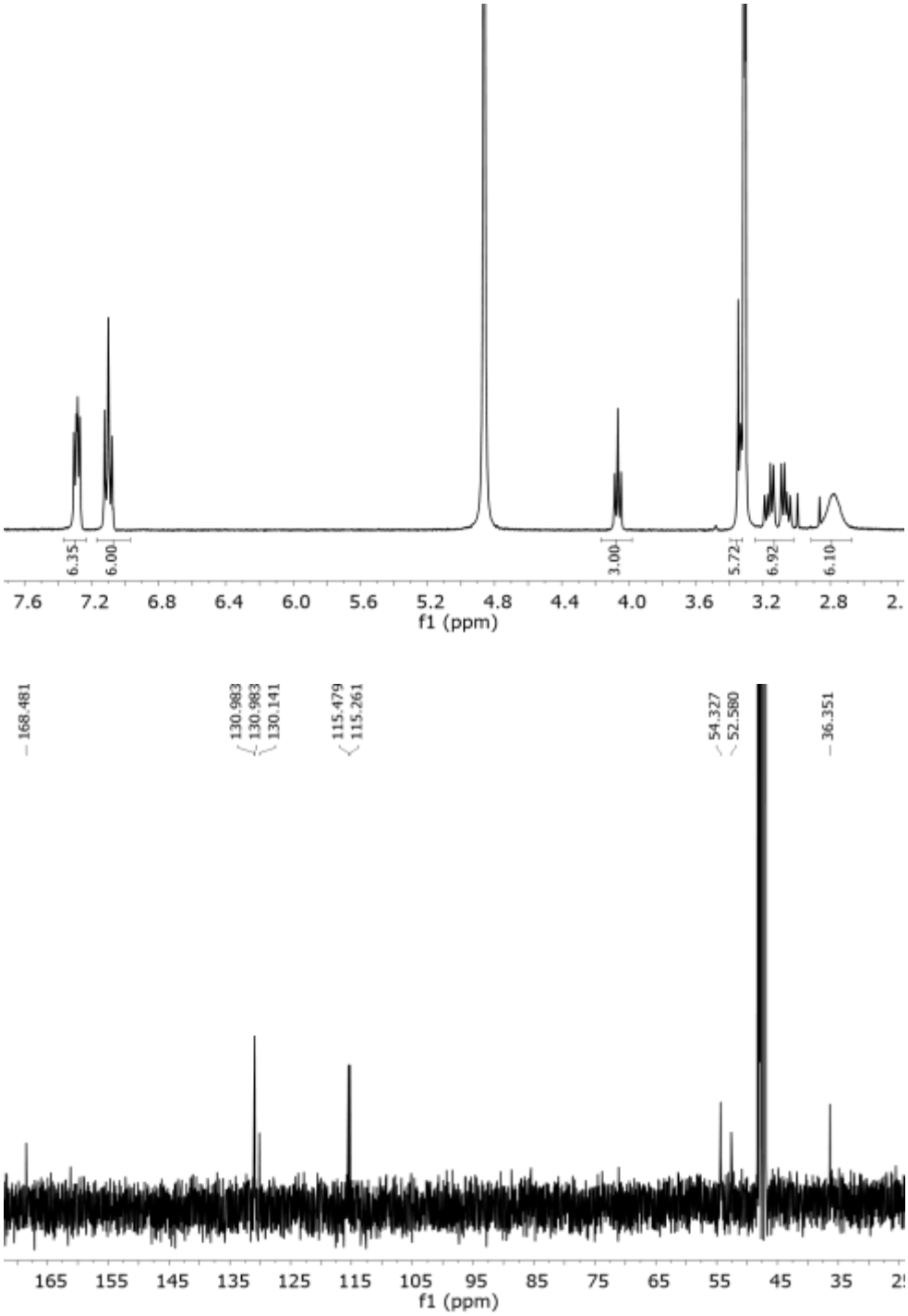


Figure 2.49. ^{1}H-NMR (400 MHz, methanol-d_3) and ^{13}C-NMR (101 MHz, methanol-d_3) spectra of **3c.**

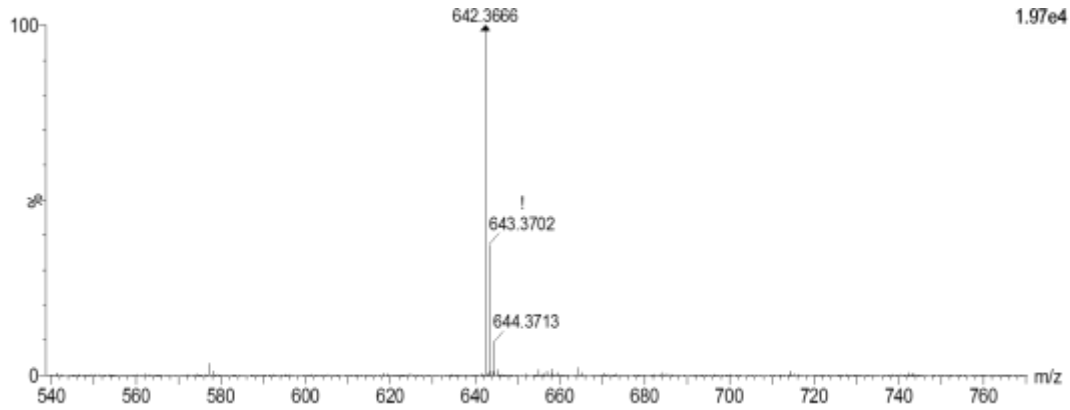
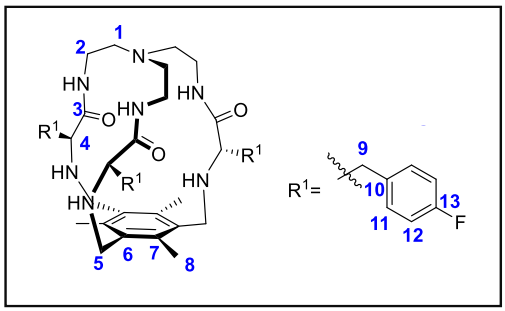


Figure 2.50. HRMS (ESI+) experimental spectrum of **3c.**

1c



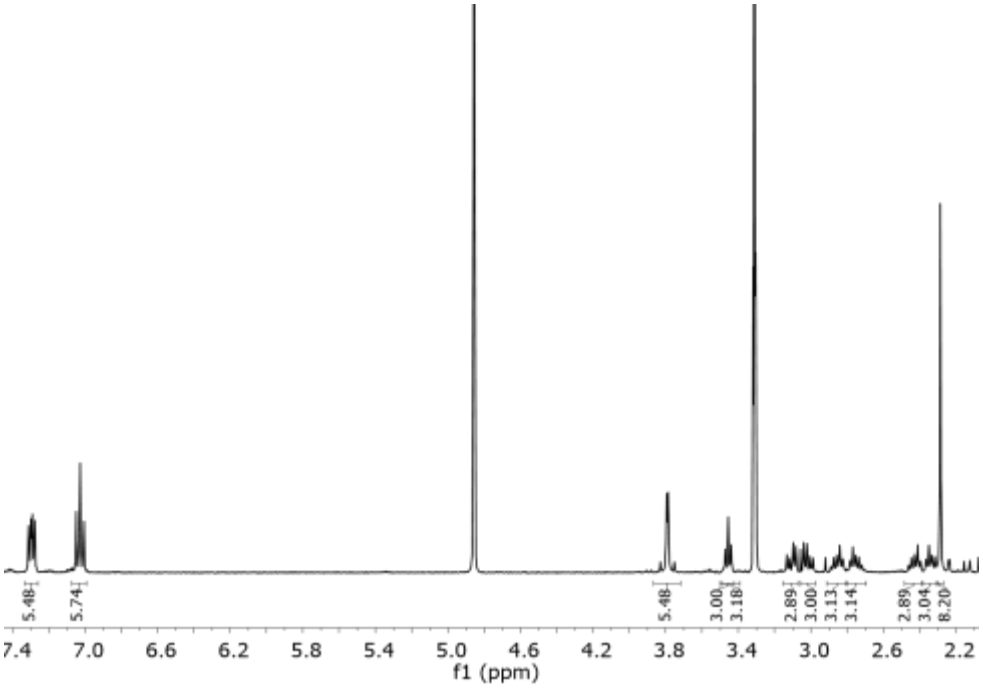
3c (173 mg, 0.158 mmol) was dissolved in ACN (40 mL). Tetrabuthylamonium chloride (22 mg, 0.079 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (62 mg, 0.158mmol) and potassium carbonate (435 mg, 3.15 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM:MeOH 95:5 as eluent to give **1c** (47mg, 59 mmol, 40% yield).

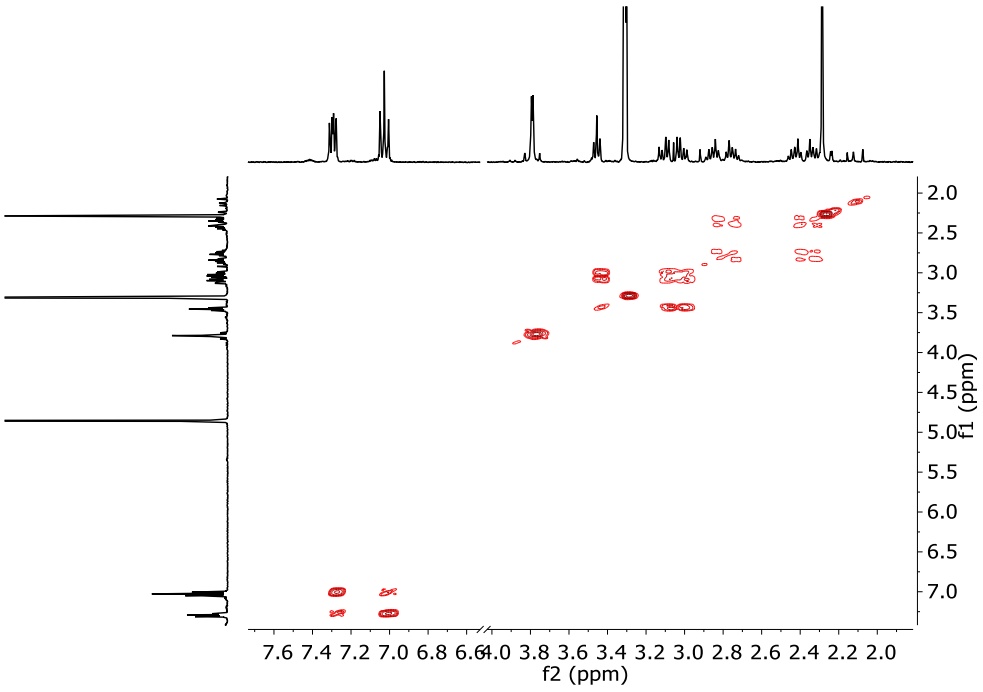
^{1}H NMR (400 MHz, CD_3OD): δ(ppm)=7.29 (m, 6H H**11**), 7.04 (m, 6H, H**12**), 3.79 (AB_q, δ_A=3.80, δ_B=3.78, J_{AB}=14 Hz, 6H, H**5**), 3.46 (X from ABX subsystem, J_{AX}=6.3, J_{BX}=6.6 Hz, 3H, H**4**), 3.11 (A subsystem from ABX, J_{AX}=6.3, J_{AB}=3.7 Hz, 3H, H**9**), 3.02 (B subsystem, from ABX J_{BX}=6.6, J_{AB}=3.7 Hz, 3H, H**9**), 2.86 (m, 3H, H**2**), 2.75 (m, 3H, H**2**), 2.43 (m, 3H, H**1**), 2.33 (m, 3H, H**1**), 2.29 (s, 9H, H**8**).

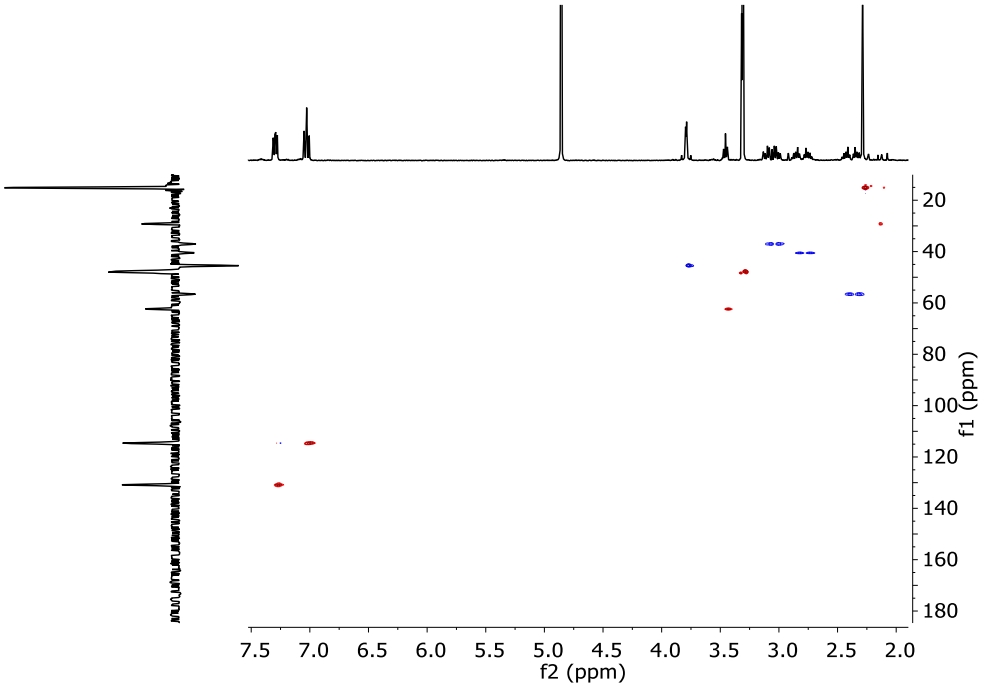
^{19}F NMR (376 MHz, CD_3OD): δ(ppm)= -118.64 (m).

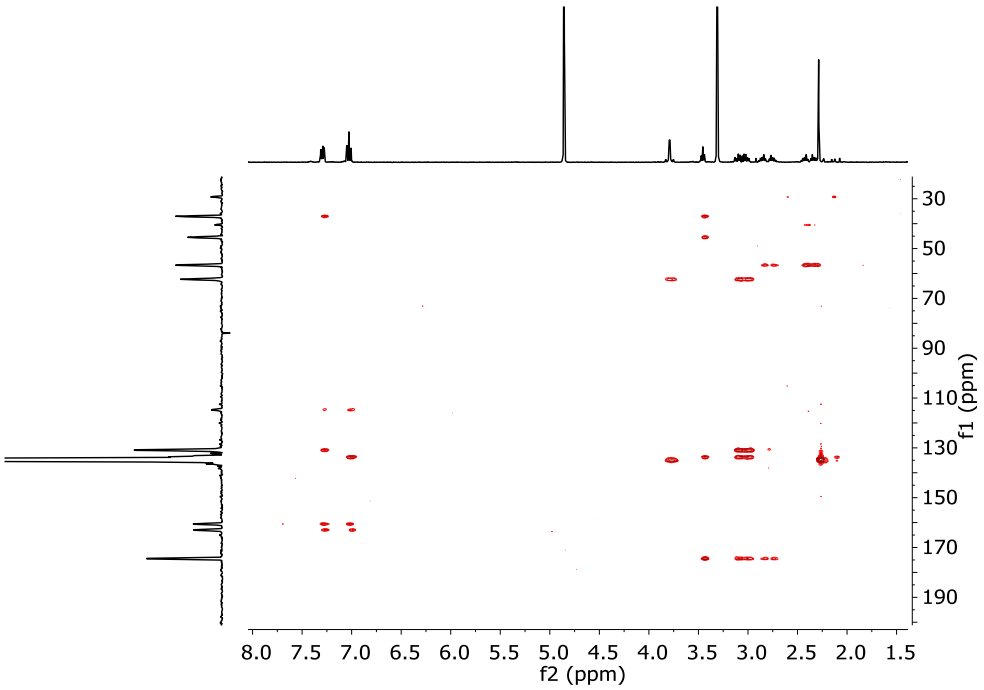
^{13}C NMR (101 MHz, CD_3OD): δ(ppm)=174.5 (C**3**), 163.0 and 160,6 (C**10**, C**13**), 134.6 (C**8**),133.5 (C**6**), 130.9 (C**11**), 114.6 (C**12**), 62.4 (C**4**), 56.6 (C**1**),45.5 (C**5**), 40.5 (C**2**), 37.1 (C**9**), 15.2 (C**8**).

HRMS (ESI-TOF) m/z [**1c**+ H]$^+$ Calc.: 798.4313, found: 798.4435.


5.48
5.74
5.48
3.00
3.18
2.89
3.00
3.13
3.14
2.89
3.04
8.20
7.4 7.0 6.6 6.2 5.8 5.4 5.0 4.6 4.2 3.8 3.4 3.0 2.6 2.2
f1 (ppm)


2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f1 (ppm)
7.6 7.4 7.2 7.0 6.8 6.6 4.0 3.8 3.6 3.4 3.2 3.0 2.8 2.6 2.4 2.2 2.0
f2 (ppm)


20
40
60
80
100
120
140
160
180
f1 (ppm)
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0
f2 (ppm)


30
50
70
90
110
130
150
170
190
f1 (ppm)
8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
f2 (ppm)

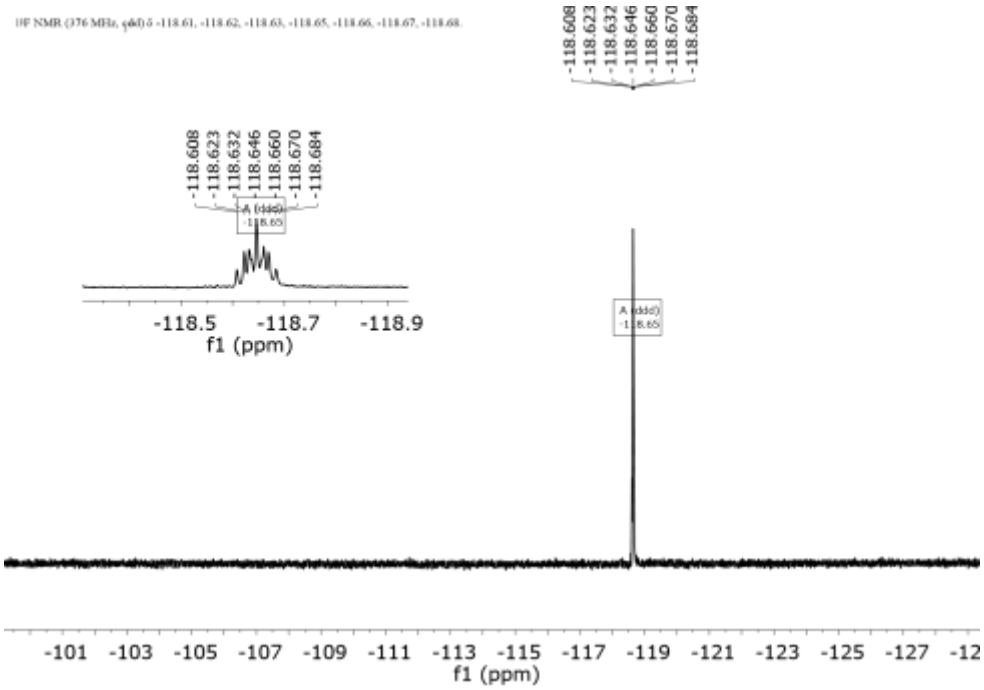


Figure 2.51. ^{1}H-NMR (400 MHz, CD_3OD), COSY, HSQC, HMBC and ^{19}F-NMR spectra of 1c.

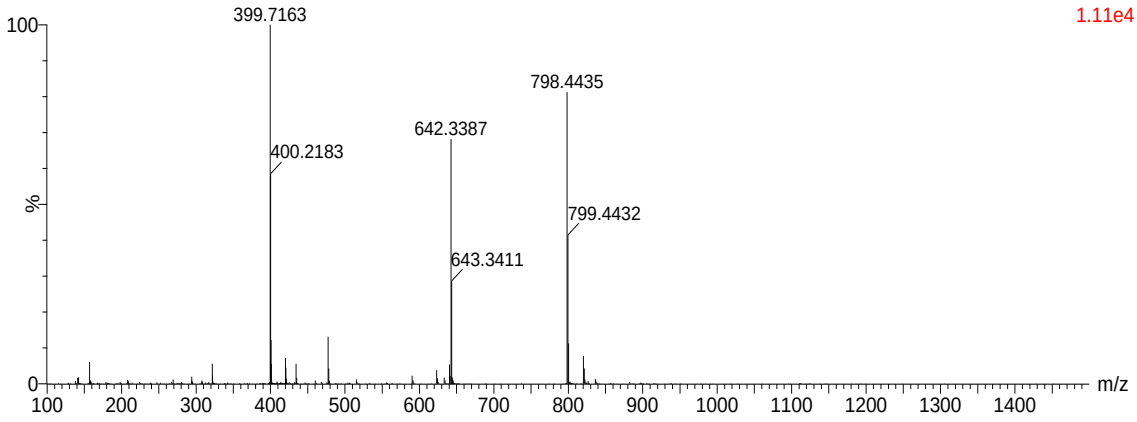
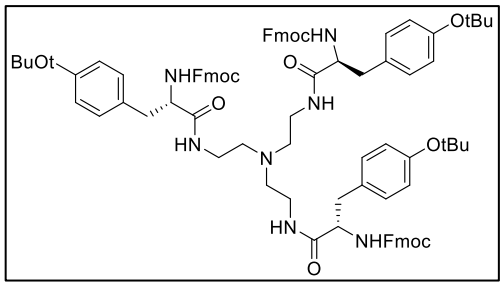


Figure 2.52. HRMS (ESI+) experimental spectrum of 1c.

2.12.3.3 Synthesis of 1d

2d: tris((9H-fluoren-9-yl)methyl) ((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azane-diyl))tris (3-(4-(tertbutoxy) phenyl)-1-oxopropane-1,2-diyl))tricarbamate

Fmoc-L-Tyr(tBu)-OH (767 mg, 1.67 mmol) was dissolved in dry DMF (10 mL). *N*,*N*'-Dicyclohexylcarbodiimide hydrochloride (DCC·HCl O.39 mg, 2.02 mmol) and 1-hydroxybenzotriazole hydrate (HOBt, 273 mg, 2.02 mmol), dissolved in DMF (5 mL) were

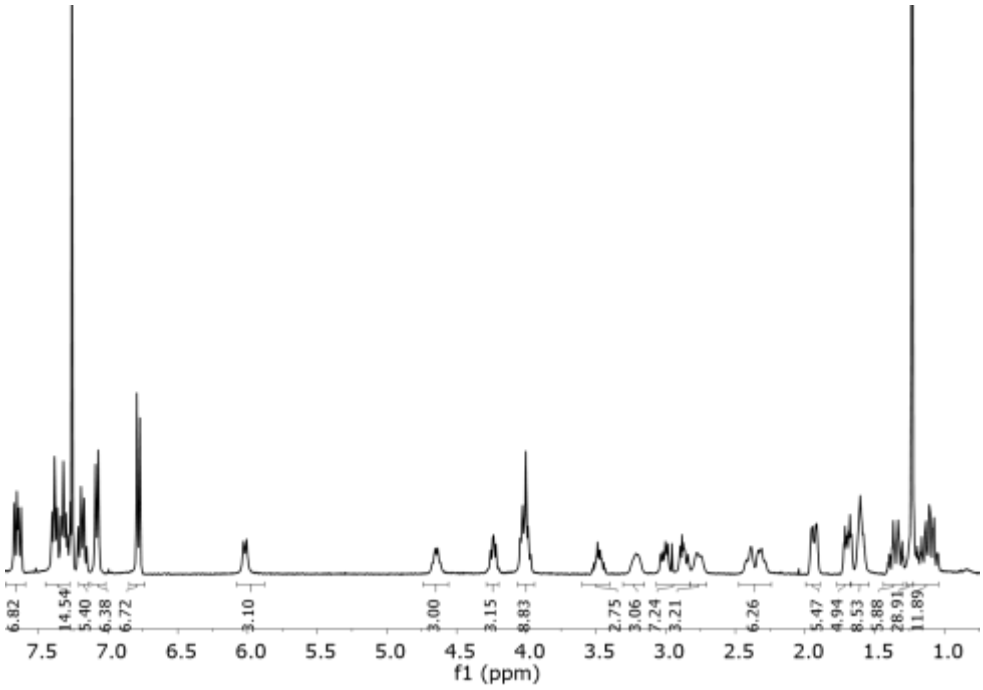


added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was then diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using DCM:MeOH 97:3 to give 614 mg of **2a** (0.417 mmol, 81% yield).

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)=7.65 (dd, *J*=12.0, 7.5 Hz, 6H), 7.38 (t, *J* = 7.4 Hz, 6H), 7.32 (t, *J*=7.4 Hz, 6H), 7.18 (q, *J*=8.0 Hz, 6H), 7.08 (d, *J*=8.1 Hz, 6H), 6.78 (d, *J*=8.3 Hz, 6H), 6.02 (d, *J*=8.7 Hz, 3H), 4.65(m, 3H), 4.24 (m, 3H), 4.015 (m, 9H), 3.48 (m, 3H), 3.20 (m, 3H), 3.01 (m, 3H), 2.87 (m, 3H), 2.77 (m, 3H), 2.36 (m, 6H), 1.94 (m, 3H), 1.70 (m, 6H), 1.69 (m, 12H), 1.36 (m, 6H), 1.23 (s, 27H), 1.10 (m, 12H).

^{13}C NMR (101 MHz, CD_3Cl): δ(ppm)=172.6, 156.7, 154.2, 143.8, 141.2,131.8, 129.9, 127.8, 127.2, 125.3, 124.213, 120.0, 78.3, 67.5, 56.3, 54.5, 46.9, 38.8, 28.9.

HRMS (ESI-TOF) m/z [**2d**+ H]$^+$ Calc.: 1471.9190, found: 1471.7548.



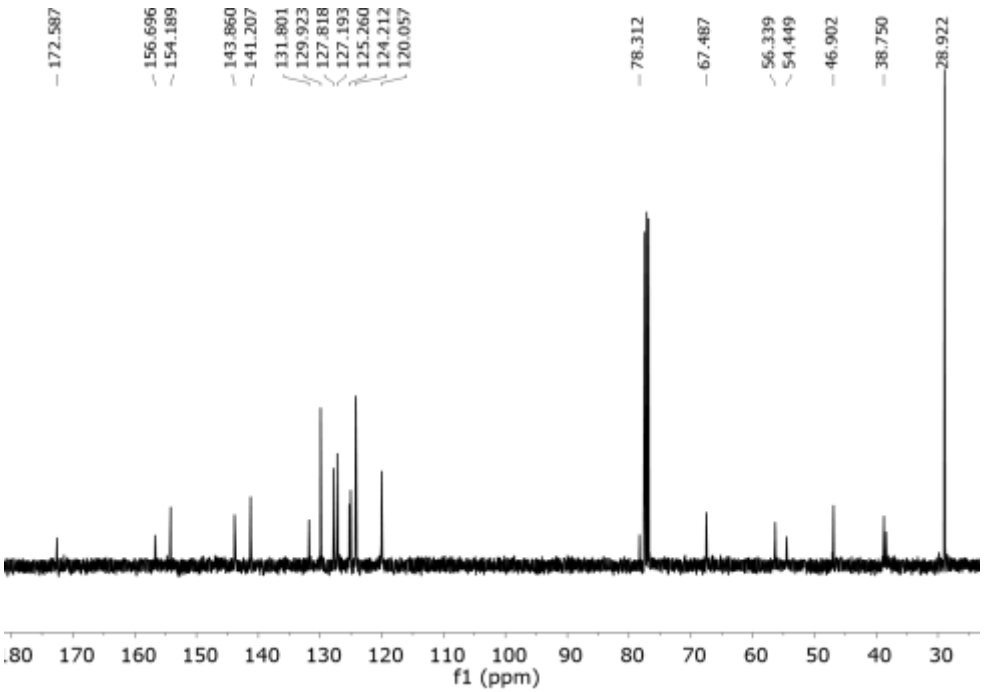


Figure 2.53. ^{1}H-NMR (400 MHz, CD_3Cl) and ^{13}C-NMR (101 MHz, CD_3Cl) spectra of **2d**.

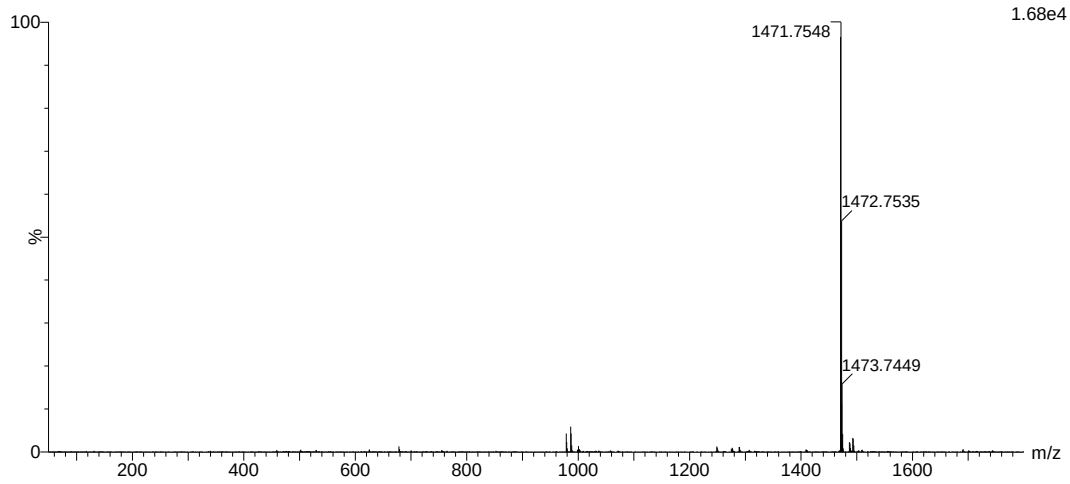
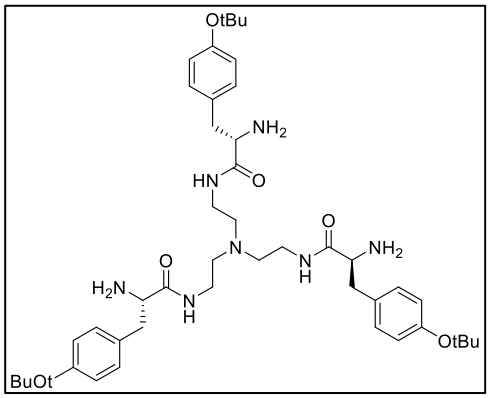


Figure 2.54. HRMS (ESI+) experimental spectrum of **2d.**

3d: (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(4-(tert-butoxy) phenyl) propanamide)

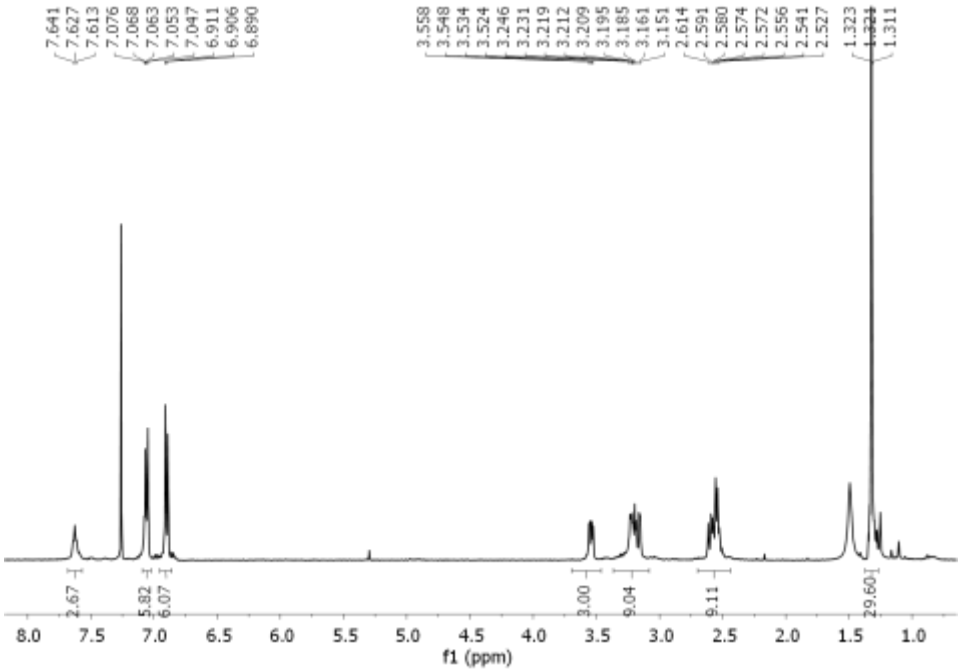


2d (500 mg, 0.340 mmol) was dissolved in DCM/diethylamine (1 mL, 1:1). The solution was stirred at room temperature for 4 hours when the complete conversion of **2d** was observed by TLC. After that, the solvent was evaporated and the solid was washed several times with hexane giving **3d** as a white solid (250 mg, 0.310 mmol, 91% yield)

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)=7.63 (t, *J*=5.6 Hz, 3H), 7.12 (m, 6H), 6.90 (m, 6H), 3.54 (X subsystem from ABX, J_{AX}=9.3, J_{BX}=4.0 Hz, 3H), 3.22 (m, 6H), 3.17 (B subsystem from ABX, J_{BX}=4.0, J_{AB}=3.7 Hz, 3H), 2.75–2.49 (A subsystem from ABX, J_{AX}=9.3, J_{AB}=3.7 Hz, 3H overlapped with m, 6H), 1.32 (s, 27H).

^{13}C NMR (101 MHz, CD_3Cl): δ(ppm)=173.8, 154.3, 132.2, 129.9, 124.4, 78.5, 56.5, 54.3, 49.2, 40.3, 29.0.

HRMS (ESI-TOF) m/z [**3d** + H]$^{+}$ Calc.: 804.5382, found: 804.5430.



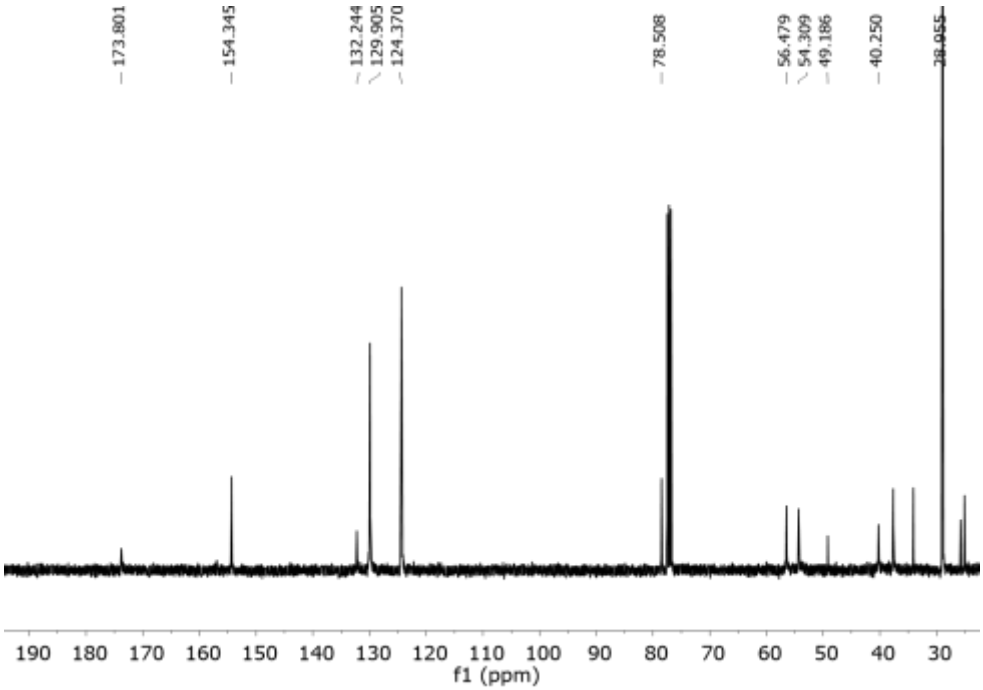


Figure 2.55. ^{1}H-NMR (400 MHz, CD_3Cl) and ^{13}C-NMR (101 MHz, CD_3Cl) spectra of **3d.**

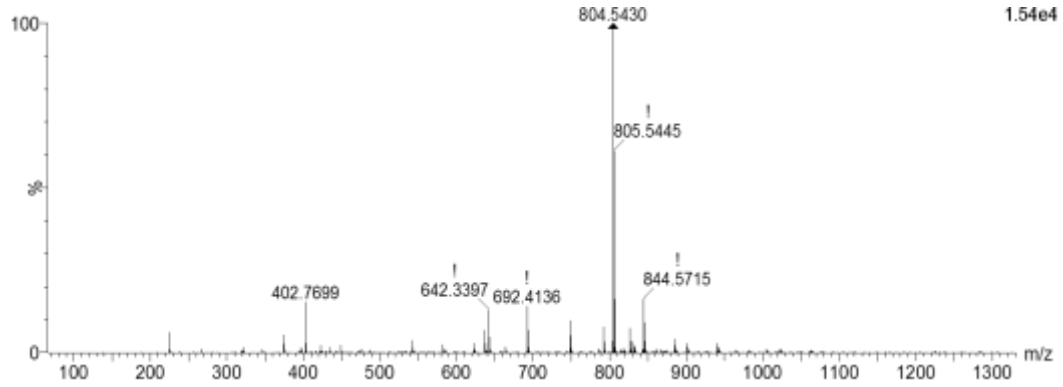
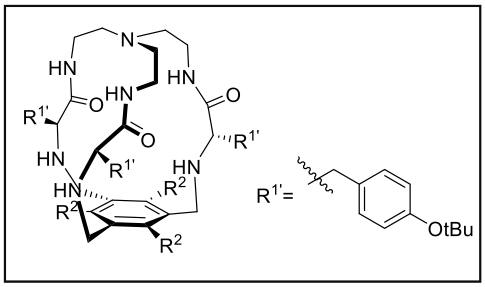


Figure 2.56. HRMS (ESI+) experimental spectrum of **3d**.

4d

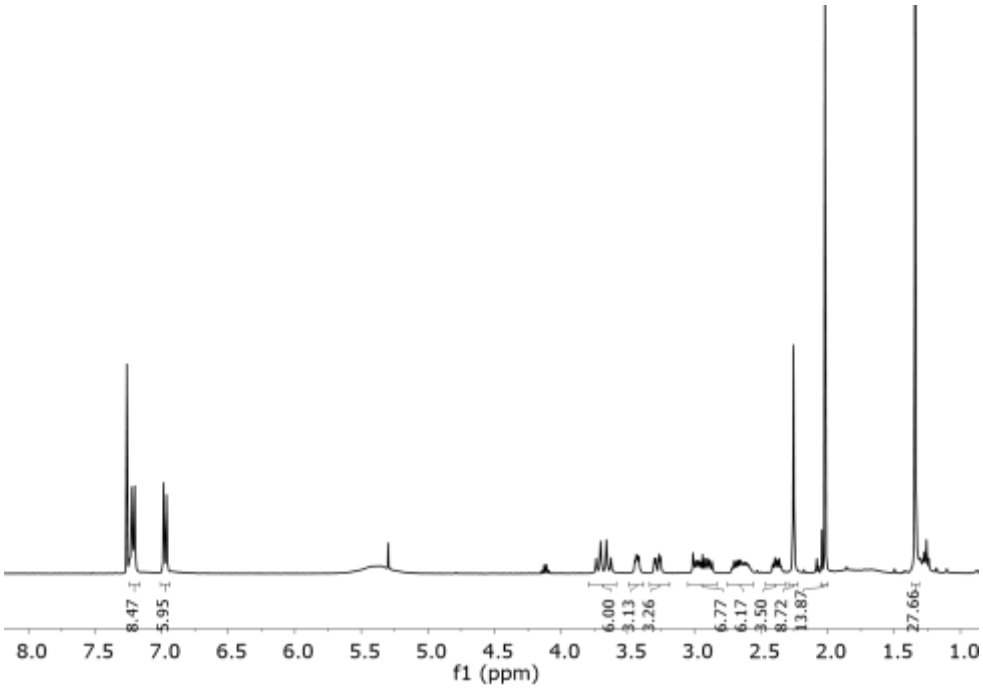


3d (230 mg, 0.286 mmol) was dissolved in ACN (50 mL). Tetrabuthylamonium chloride (40 mg, 0.143 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (116 mg, 0.286mmol) and potassium carbonate (790 mg, 5.72 mmol) were added to the solution. The reaction mixture was refluxed for 16 hours under nitrogen atmosphere. After cooling down to room temperature, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM:MeOH 97:3 as eluent to give **4d** (90mg, 93.7 mmol, 33% yield).

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)=7.21 (d, *J*=8.5 Hz, 6H), 6.97 (d, *J*=8.4 Hz, 6H), 3.68 (AB_q, δ_A=3.72, δ_B=3.65, J_{AB}=13.6 Hz, 6H), 3.43 (X subsystem from ABX, J_{AX}=4.4, J_{BX}=8.0 Hz, 3H), 3.28 (A subsystem from ABX, J_{AX}=4.4, J_{AB}=3.9 Hz, 3H), 2.94 (B subsystem from ABX overlapped with 3H, 6H), 2.65 (m, 6H), 2.38 (m, 3H), 2.26 (s, 9H), 1.34 (s, 27H).

^{13}C NMR (101 MHz, $CDCl_3$): δ(ppm)=174.3, 154.4, 135.2, 132.7, 130.0, 124.6, 78.6, 63.7, 57.9, 46.6, 41.3, 37.9, 29.9, 29.0, 16.2.

HRMS (ESI-TOF) m/z [**4d** + H]$^+$ Calc.: 960.6321, found: 960.7116.



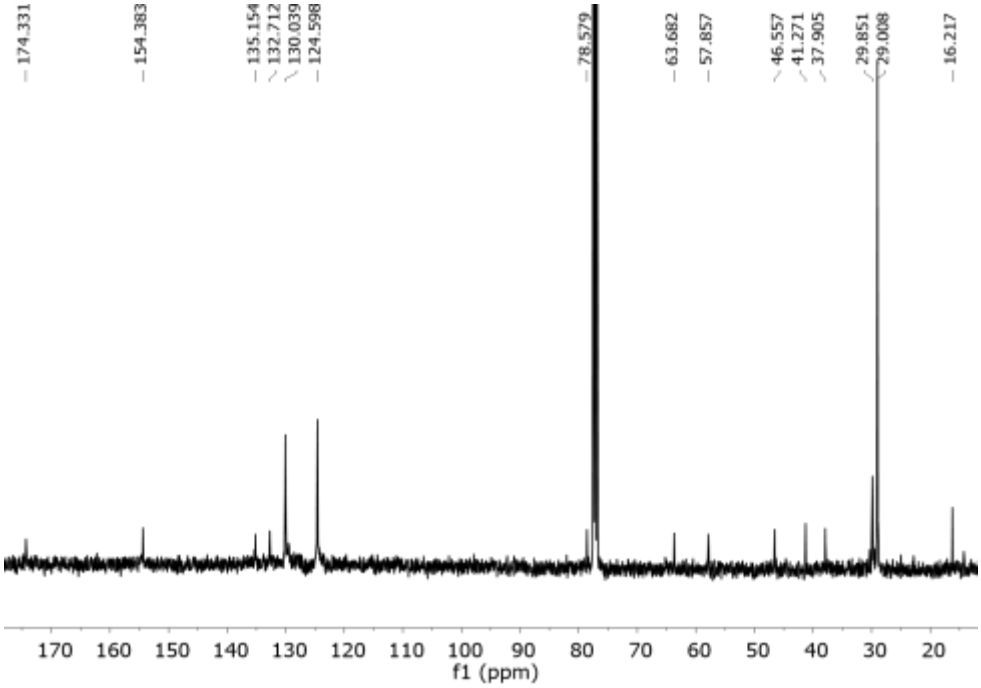


Figure 2.57.^{1}H-NMR (400 MHz, CD_3Cl) and ^{13}C-NMR (101 MHz, CD_3Cl) spectra of **4d**.

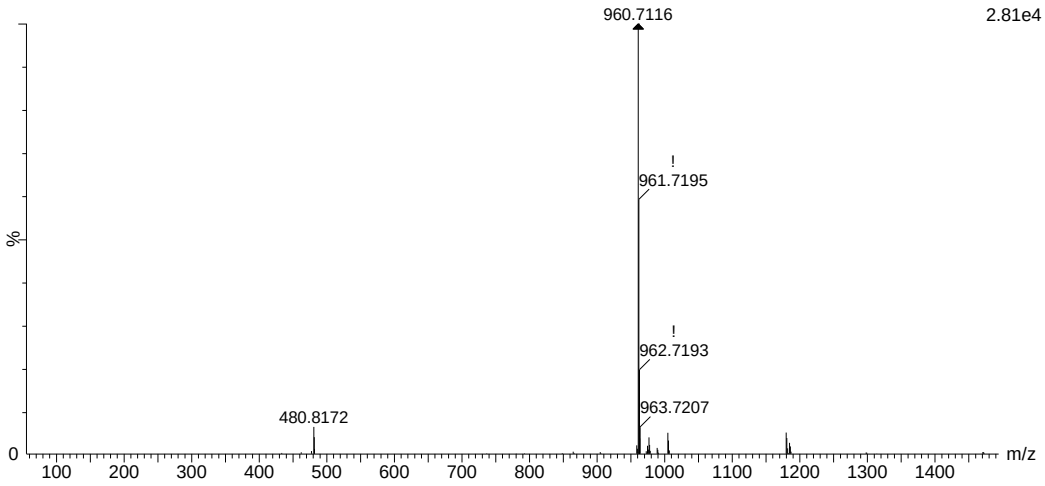
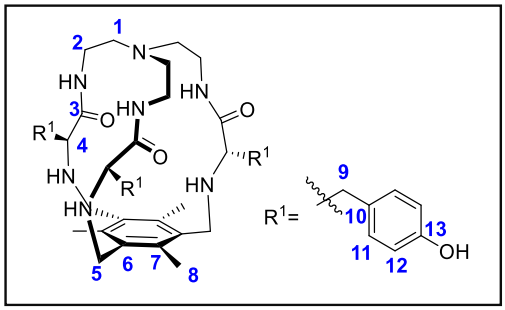


Figure 2.58. HRMS (ESI+) experimental spectrum of **4d**.

1d

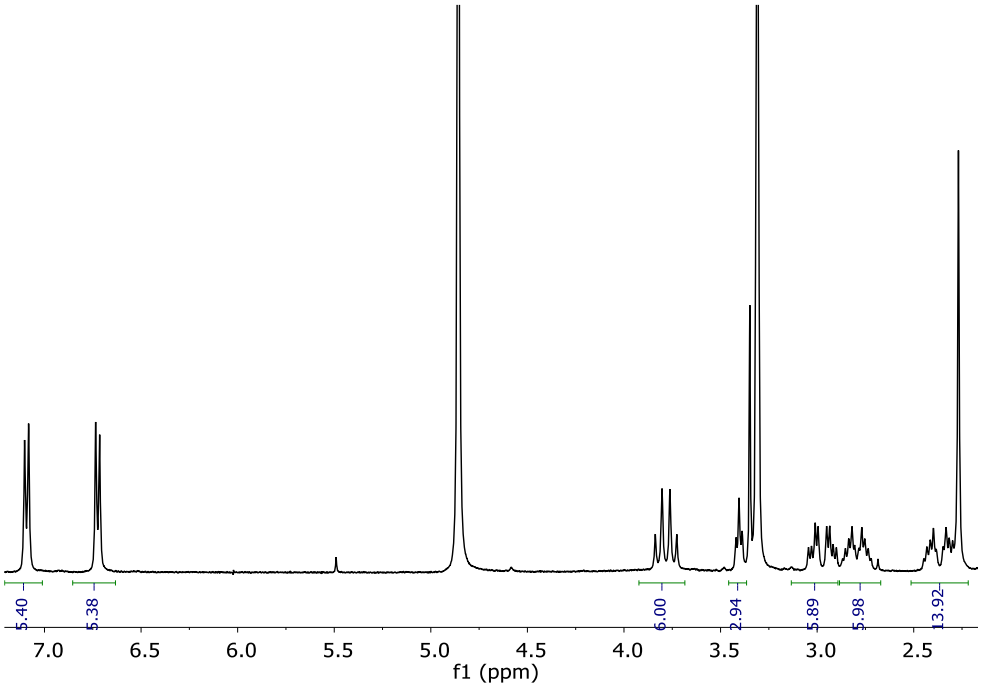


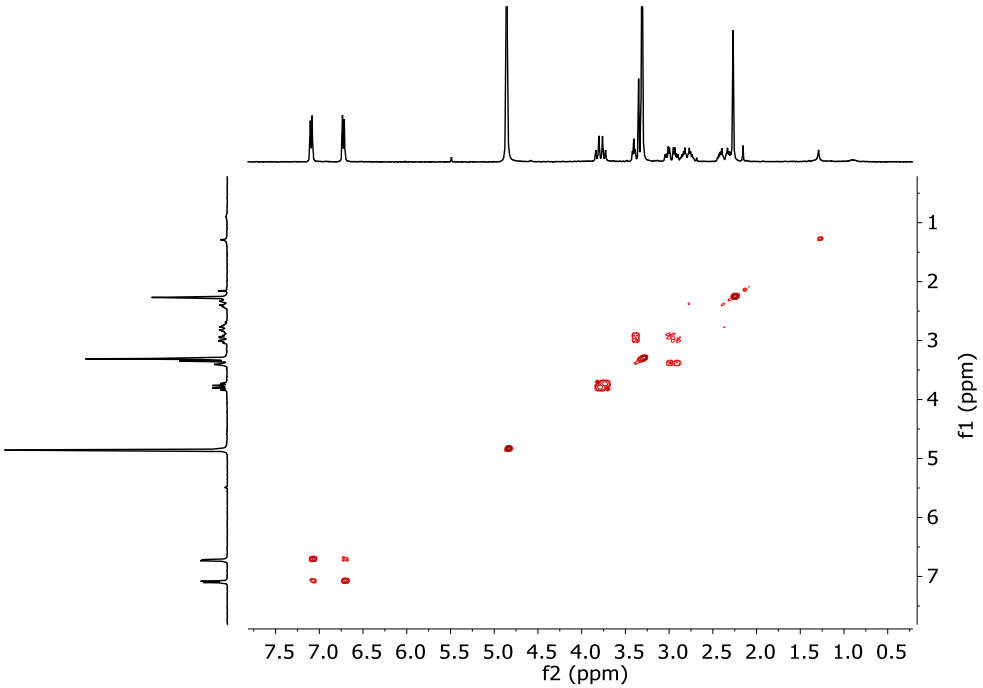
4d (90 mg, 93.7 mmol) was dissolved in DCM (1.5 ml). Triethylsilylsilane (0.15 mL, 1.1 mmol) and trifluoroacetic acid (1 mL) were added. The solution was stirred at room temperature during 3 hours and then solvents were evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and purified through reverse phase flash chromatography affording **1d·4TFA** as a white solid (108 mg, 86 mmol, 91 % yield). 1d·4FTA was dissolved in methanol and amberlite IRA-95 resin was added while stirring until neutral to basic pH was reached. Finally, amberlite was filtered off and the solvent was evaporated to obtain **1d** (43 mg, 34 mmol, 77 % overall yield)**.**

^{1}H NMR (400 MHz, CD_3OD): δ(ppm)=7.09 (d, *J*= 8.2 Hz, 2H, H**11**), 6.72 (d, *J*= 8.1 Hz, 2H, H**12**), 3.78 (AB_q, δ_A=3.82, δ_B=3.74, J_{AB}=14 Hz, 6H, H**5**), 3.40 (X subsystem from ABX, J_{AX}=6.2, J_{BX}=6.6 Hz, 3H, H**4**), 3.02 (A subsystem from ABX, J_{AX}=6.2, J_{AB}=3.7 Hz, 3H, H**9**) 2.93 (B subsystem from ABX, J_{BX}=6.6 J_{AB}=3.7 Hz, 3H, H**9**), 2.0 (m, 6H, H**2**), 2.37 (m, 6H, H**1**), 2.27 (s, 9H, H**8**).

^{13}C NMR (100 MHz, CD_3OD): δ(ppm)=174.7 (C**3**), 155.2 (C**13**), 135.9 (C**7**), 130.2 (C**11**), 128.2 (C**6**), 114.9 (C**12**), 62.5 (C**4**), 56.7 (C**1**), 45.6 (C**5**), 40.7 (C**2**), 37.1 (C**9**), 15.3 (C**8**).

HRMS (ESI-TOF) m/z [**1d** + H]$^+$ Calc.: 792.443, found: 792.4522


5.40
5.38
6.00
2.94
5.89
5.98
13.92
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
f1 (ppm)


1
2
3
4
5
6
7
f1 (ppm)
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f2 (ppm)

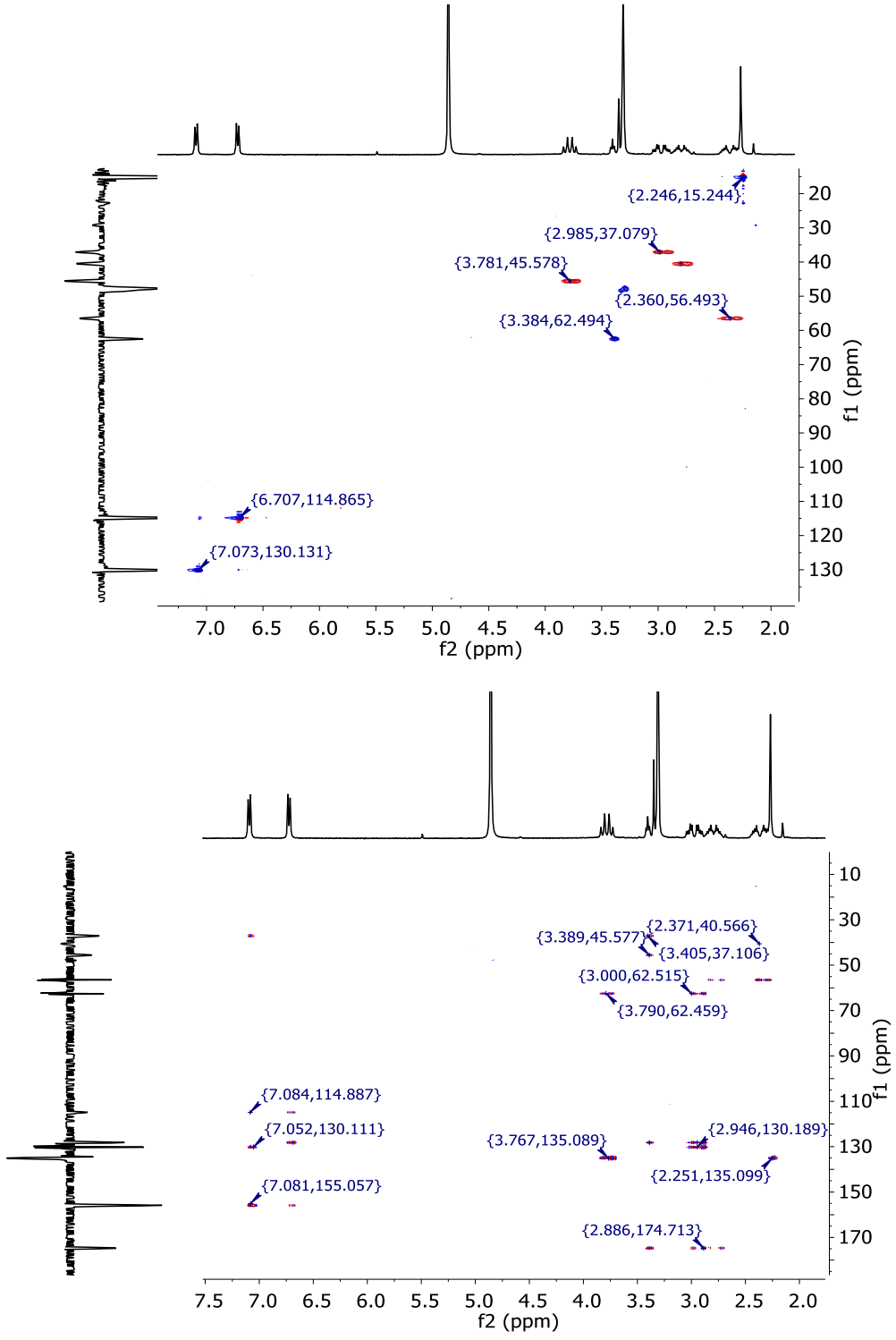


Figure 2.59. ^{1}H-NMR (400 MHz, CD_3OD), COSY, HSQC and HMBC spectra of **1d**.

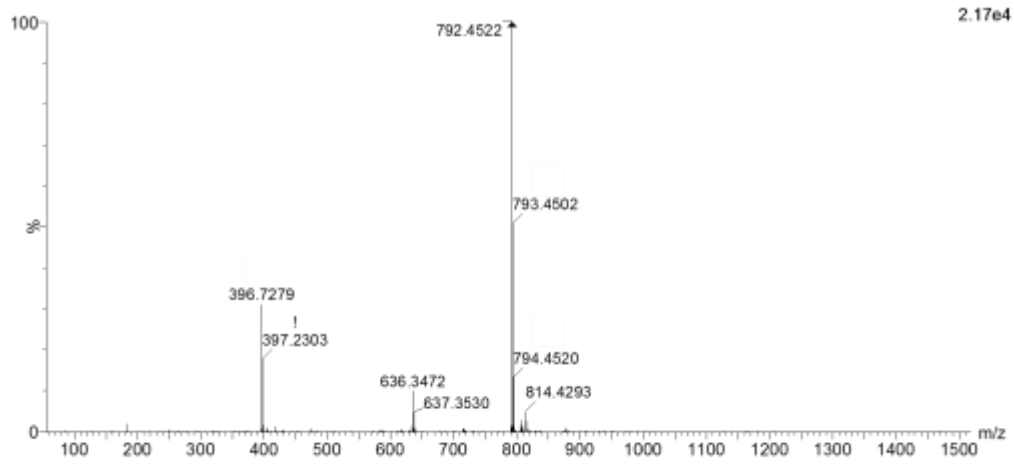
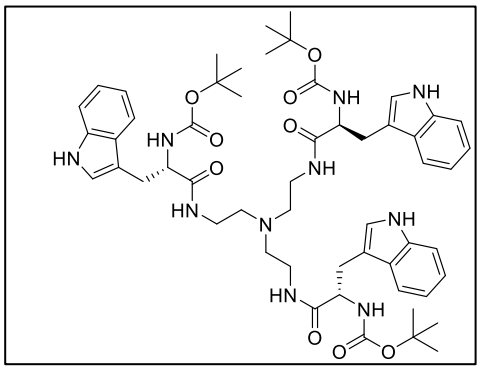


Figure 2.60. HRMS (ESI+) experimental spectrum of **1d.**

2.12.3.4 Synthesis of 1e

2e: tri-tert-butyl ((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(3-(1H-indol-3-yl)-1-oxopropane-1,2-diyl))tricarba-mate



Boc-Trp-OH (300 mg, 0.99 mmol) was dissolved in dry DCM (5 mL) and DMF (0.4 mL). (2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU O.375 mg, 0.989 mmol) *N*,*N*-diisopropylethylamine (DIPEA, 0.09 mL, 0.99 mmol) and tris(2-aminoethyl)amine (0.041 mL, 0.30 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, after no more conversion of the starting material was observed by TLC. The mixture was diluted with more DCM and washed with water (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using DCM:MeOH 95:5 to give 0.276 mg of **2e** (0.275 mmol, 92% yield).

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)=8.78 (s, 3H), 7.54 (d, *J*=7.9 Hz, 3H), 7.27 (d, *J*=8.1 Hz, 3H), 7.10 (t, *J*=7.5 Hz, 3H), 7.01 (t, *J*=7.5 Hz, 6H), 6.97 (s, 3H), 6.41 (s, 3H), 5.54 (s, 3H), 4.39 (X subsystem from ABX J_{AX}=8, J_{BX}=6 Hz, 3H), 3.30–3.03 (m, A and B subsystems from ABX, J_{AX}=8.0, J_{BX}=5.7, J_{AB}=4.4 Hz, 6H), 2.74 (m, 6H), 2.43 (m, 6H), 1.41 (s, 27H).

^{13}C NMR (100 MHz, CD_3Cl): δ(ppm)=172.5, 156.0, 136.2, 127.7, 123,7, 121.9, 119.4, 118.9, 111.4, 80.3, 55.9, 53.5, 53.0, 28.9, 28.4.

HRMS (ESI-TOF) m/z [**2e** + H]$^+$ Calc: 1005.5310, found: 1005.6048.

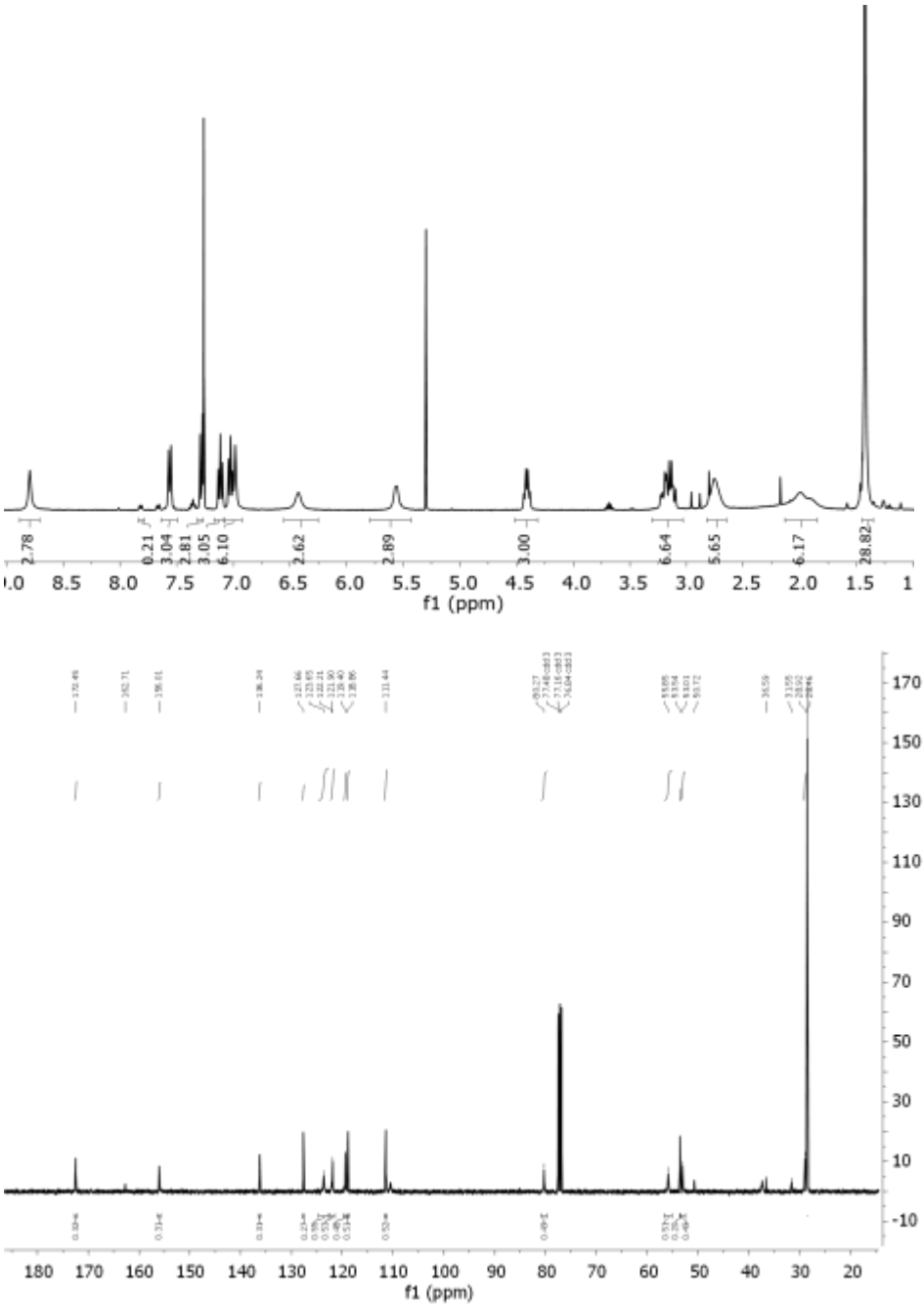


Figure 2.61. ^{1}H-NMR (400 MHz, CD_3Cl), and ^{13}C-NMR (101 MHz, CD_3Cl) spectra of **2e**

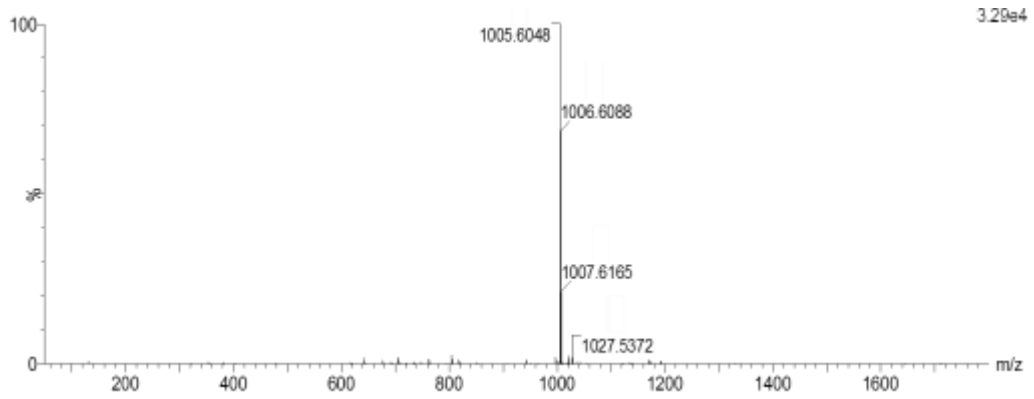
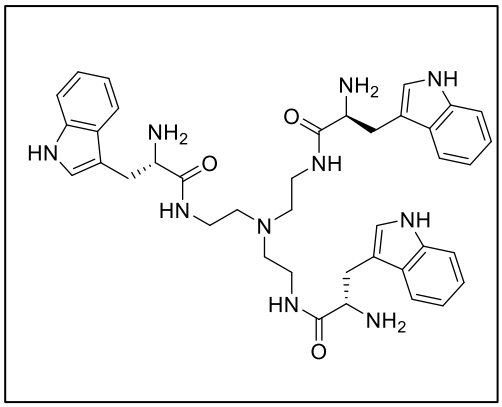


Figure 2.62. HRMS (ESI+) experimental spectrum of **2e.**

3e: (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(1H-indol-3-yl)propane-mide)



2e (250 mg, 0.249 mmol) was dissolved in DCM (1.5 ml). Triethylsilylsilane (0.44 mL, 3.73 mmol) and trifluoroacetic acid (1 mL) were added. The solution was stirred at room temperature during 3 hours and then solvents were evaporated under an air current affording a yellow oil. It was washed several times with diethyl ether and dried affording **3e·4TFA** as a white solid (253 mg, 0.299 mmol, 88% yield).

^{1}H NMR (400 MHz, CD_3OD): δ(ppm)=7.59 (d, *J*=7.8 Hz, 3H), 7.39 (d, *J*=8.1 Hz, 3H), 7.21 (s, 3H), 7.14 (t, *J*=7.5 Hz, 3H), 7.06 (t, *J*=7.4, 3H), 4.10 (X subsystem from ABX, J_{AX}=7.4, J_{BX}=7.4 Hz, 3H), 3.41–3.33 (A and B subsystems from ABX, J_{AX}=7.4, J_{BX}=7.4 J_{AB}=4.5 Hz overlapped with 6 additional H, 12H), 3.35 (m, 6H), 2.73 (m, 6H).

^{13}C NMR (101 MHz, CD_3OD): δ(ppm)=199.3, 166.3, 156.5, 153.8, 151.1, 148.5, 147.3, 140.9, 136.2, 83.4, 81.5, 63.9, 56.

HRMS (ESI-TOF) m/z [**3e** + H]$^{+}$ Calc.: 705.3911, found: 705.4070.

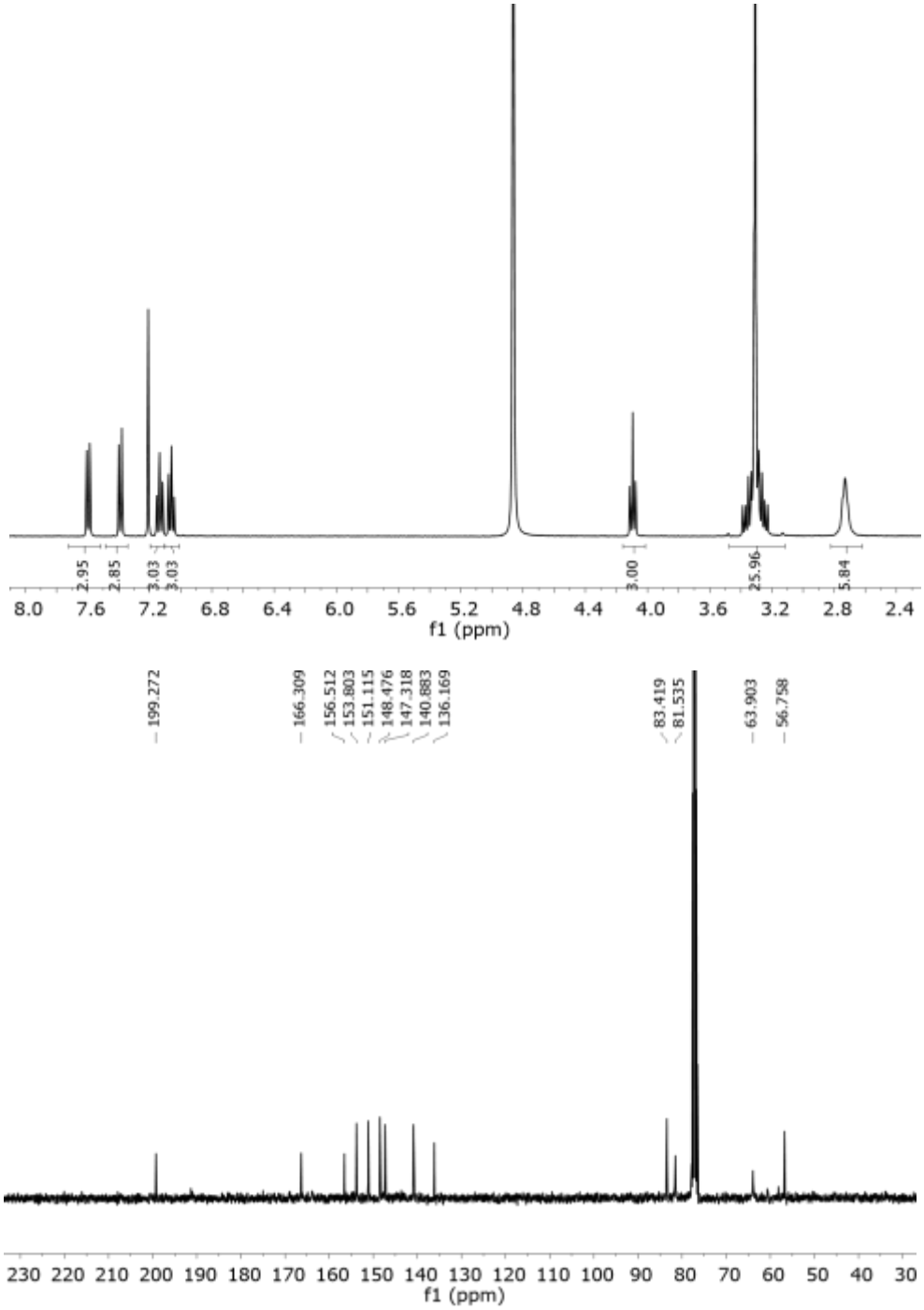


Figure 2.63. ^{1}H-NMR (400 MHz CD_3OD) and ^{13}C-NMR (101 MHz, CD_3Cl) spectra of **3e**.

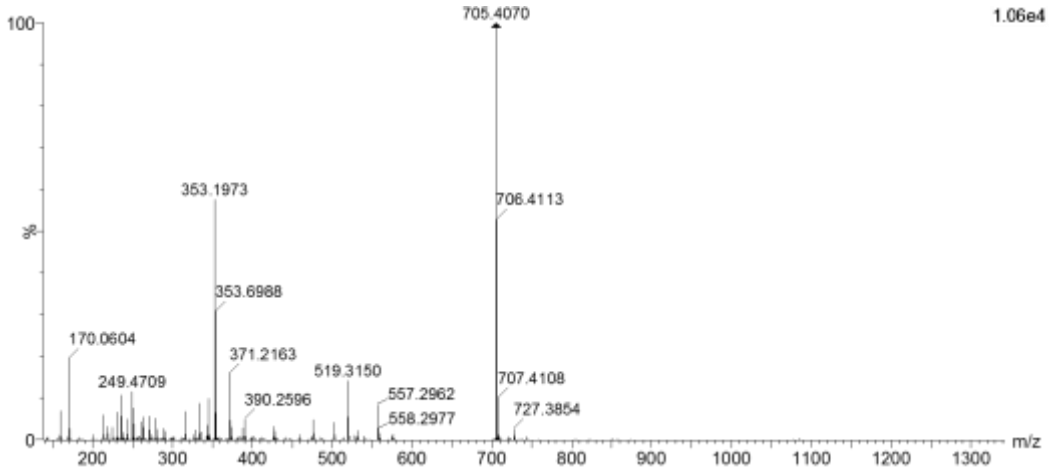
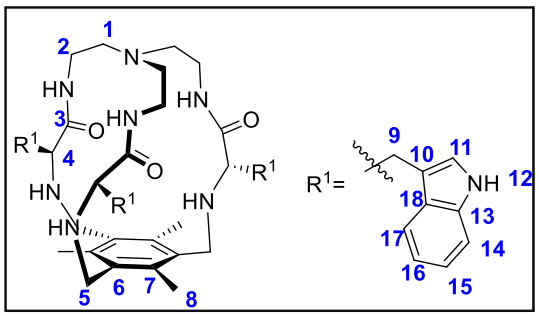


Figure 2.64. HRMS (ESI+) experimental spectrum of **3e.**

1e

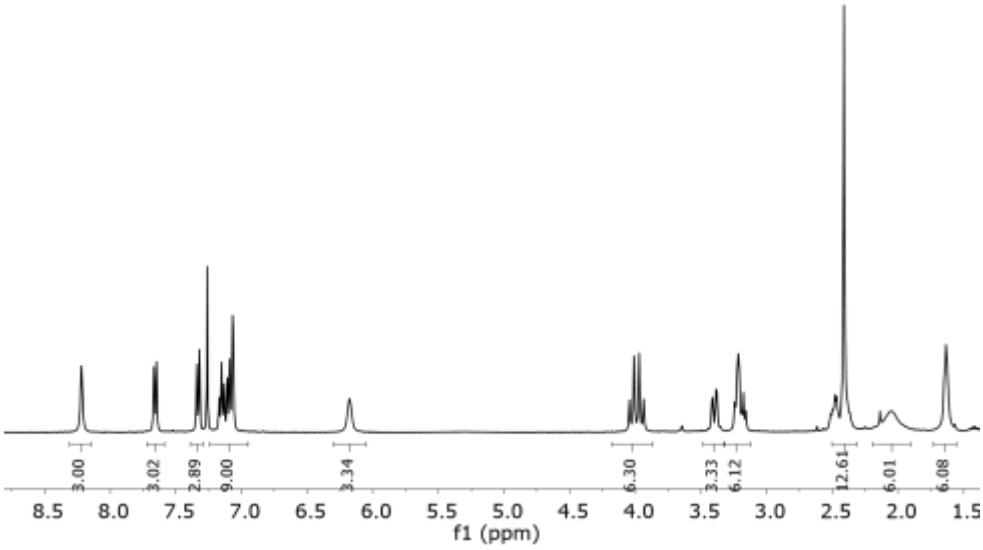


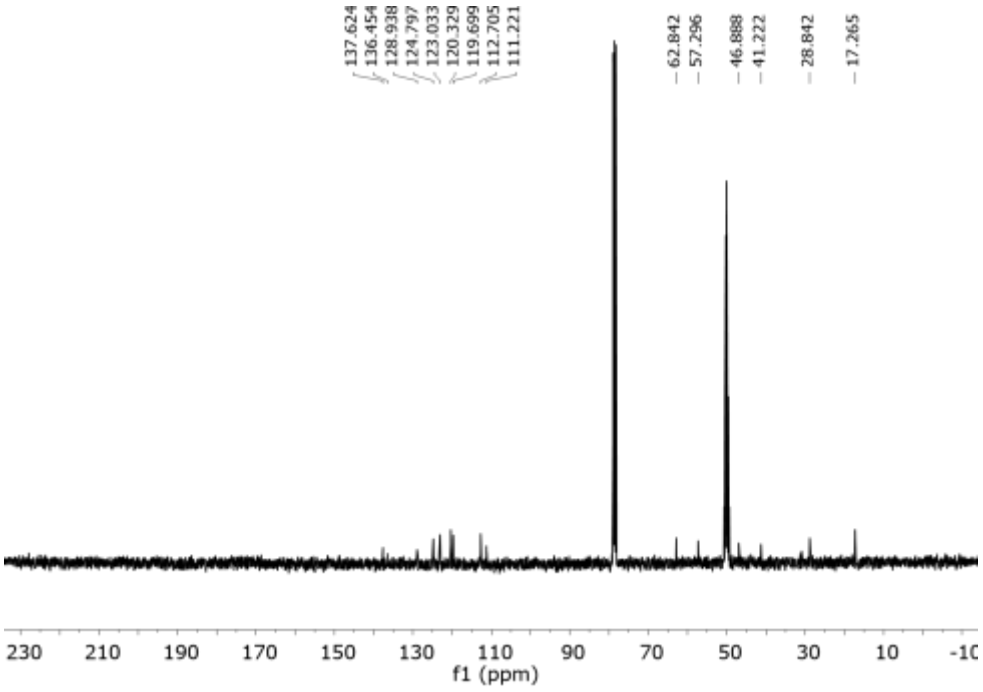
3e (150 mg, 0.213 mmol) was dissolved in dry acetonitrile (40mL). Bu_4Cl (29 mg, 0.106 mmol) and K_2CO_3 (588 mg, 4.25 mmol) were then added. 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (93.8 mg, 0.212 mmol) was dissolved in acetonitrile (10 mL) and the solution was added to the **3e** solution. The reaction mixture refluxed for 16 hours. Solvent was evaporated and the resulting crude was purified by flash chromatography using DCM:MeOH 95:5 as eluent to give **1e** as a white solid (64 mg, 57.7 mmol, 33% yield).

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)=8.22 (s, 3H, H**12**), 7.66 (d, *J*=7.8 Hz, 3H, H**17**), 7.33 (d, *J*=8.0 Hz, 3H, H**14**), 7.26 (s, 3H, H**11**), 7.10 (m, 9H, H**15**, H**16**), 6.18 (s, 3H, NH_{amide}), 3.99 (AB_q, δ_A=4.02, δ_B=3.94, J_{AB} = 14 Hz, 6H, H**5**), 3.40 (m, 3H, H**4**), 3.21 (m, 6H, H**9**), 2.46 (m, 3H, H**2**), 2.41 (s, 9H, H**8**), 2.06 (m, 6H, H**2** and NH_{amine}), 1.63 (m, 6H, H**1**).

^{13}C NMR (101 MHz, CD_3OD): δ(ppm)=176.0 (C**3**), 140.0 (C**7**), 136.6 (C**10**), 135.5 (C**6**), 127.9 (C**7**), 123.8 (C**15/16**), 122.0 (C**15/16**), 119.3 (C**11**), 118.7 (C**17**), 111.7 (C**14**), 110.2 (C**13**), 61.8 (C**4**), 56.3 (C**1**), 45.9 (C**5**), 40.2 (C**2**), 27.8 (C**9**), 16.3 (C**8**).

HRMS (ESI-TOF) m/z [**1e** + H]$^+$ Calc.: 861.485, found: 861.5199 .


3.00
3.02
2.89
9.00
3.34
6.30
3.33
6.12
12.61
6.01
6.08
8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
f1 (ppm)


137.624
136.454
128.938
124.797
123.033
120.329
119.699
112.705
111.221
62.842
57.296
46.888
41.222
28.842
17.265
230 210 190 170 150 130 110 90 70 50 30 10 -10
f1 (ppm)

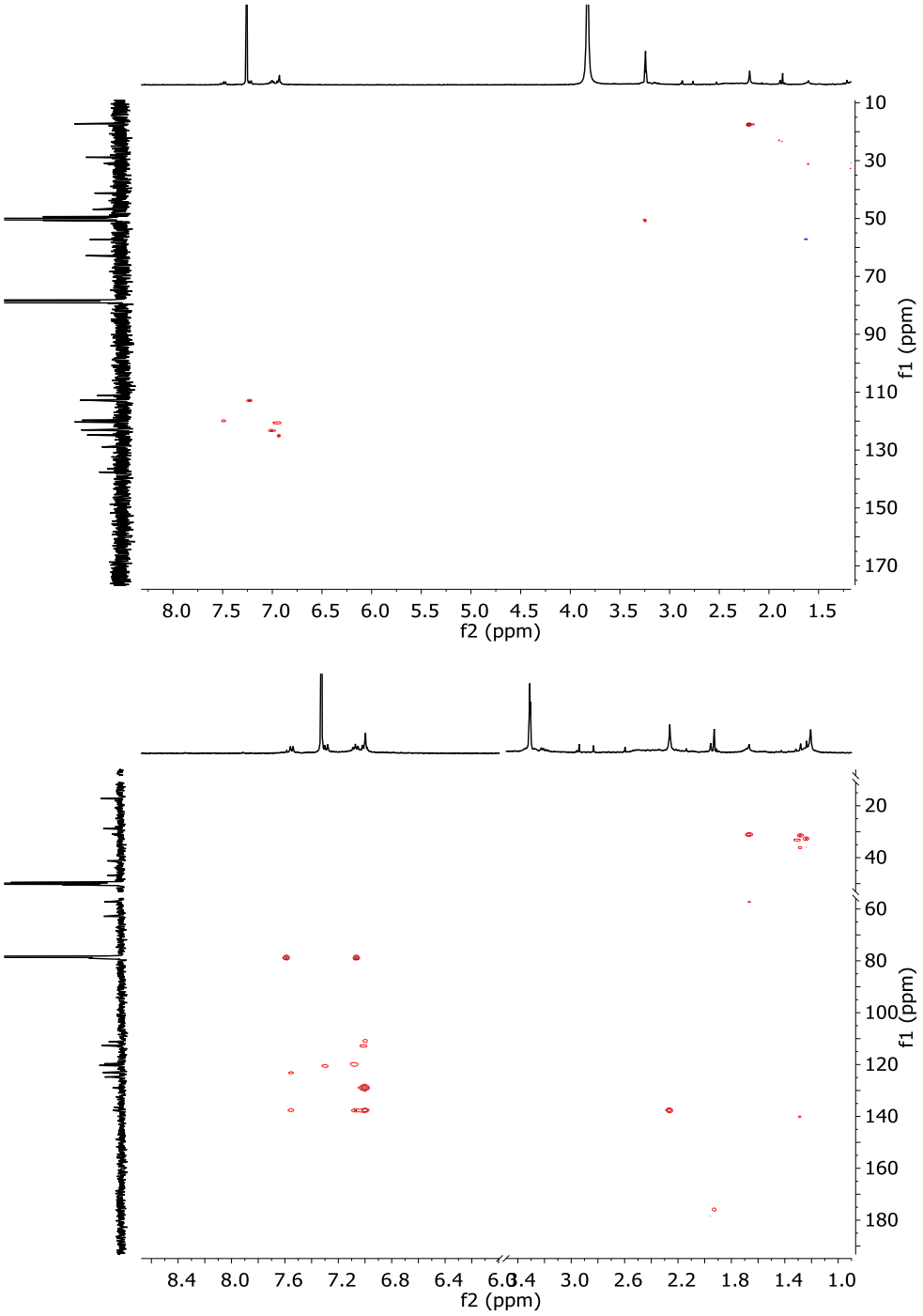


Figure 2.65. ^{1}H-NMR (400 MHz $CDCl_3$) ^{13}C-NMR (101 MHz, CD_3Cl/CD_3OD(6:4)), HSQC and HMBC spectra of **1e.**

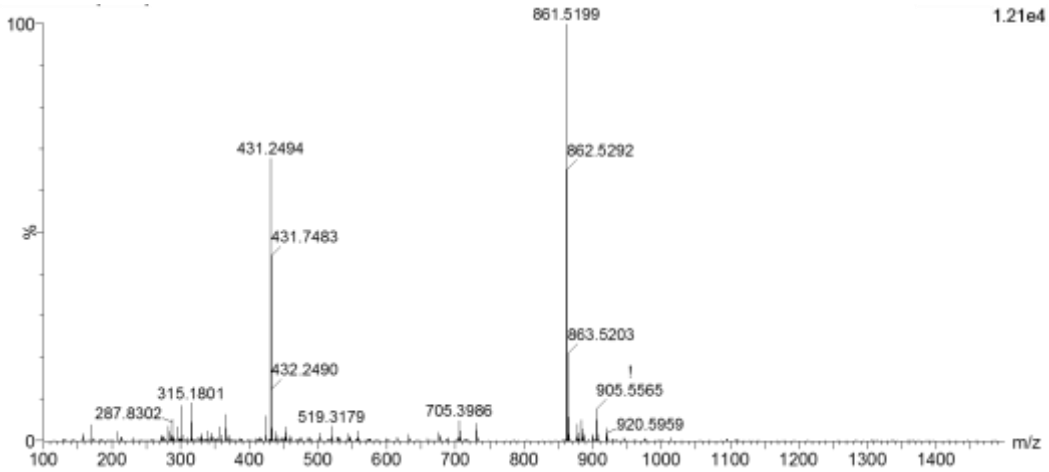
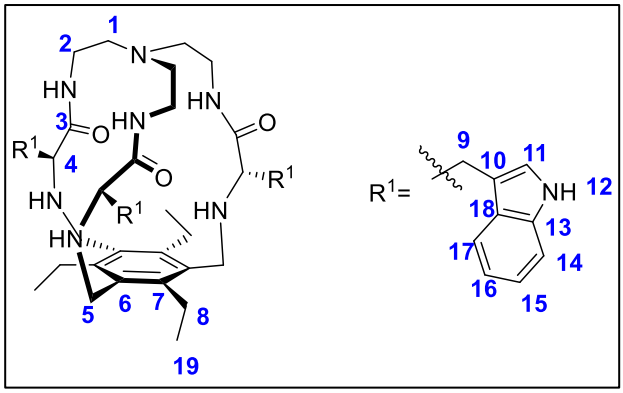


Figure 2.66. HRMS (ESI+) experimental spectrum of **1e**.

2.12.3.5 Synthesis of 1f

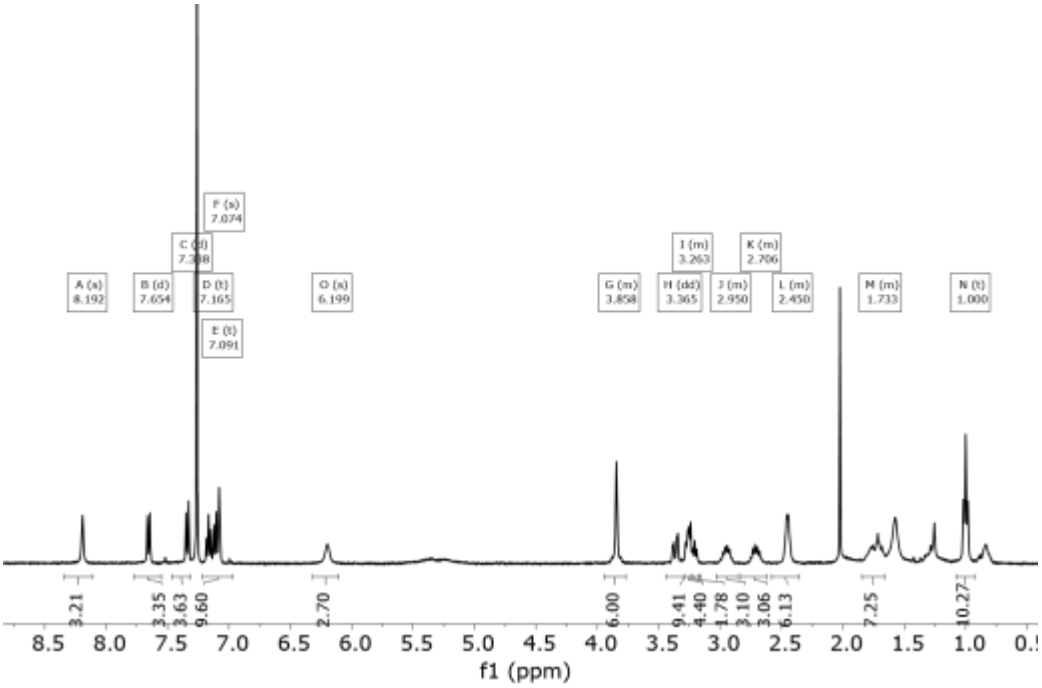


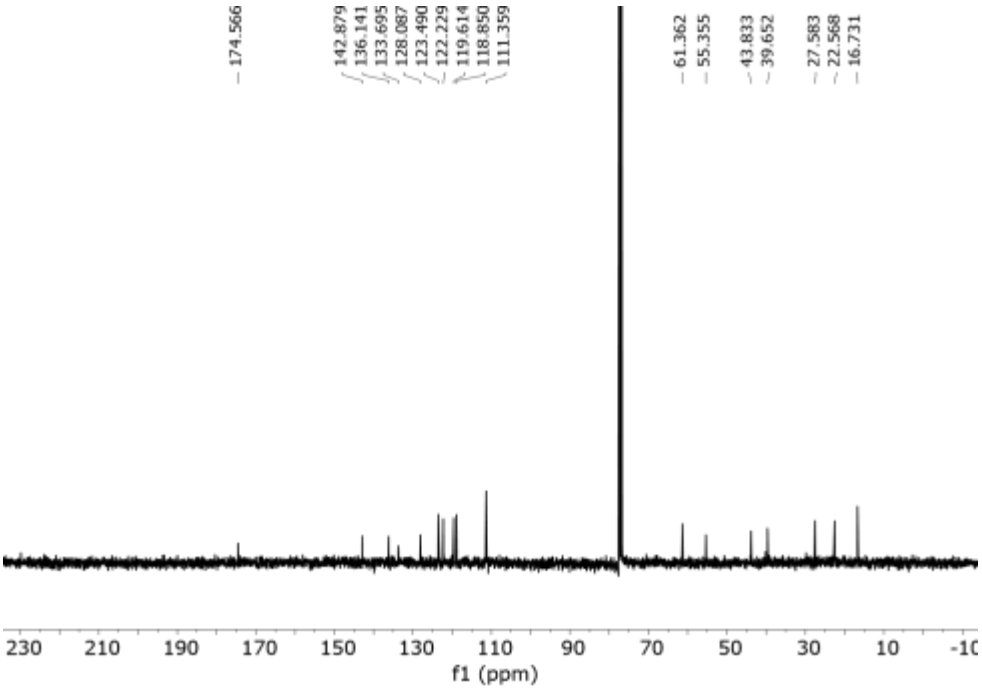
3e (80mg, 0.069 mmol) was dissolved in dry acetonitrile (20mL). Bu_4Cl (29 mg, 0.106 mmol) and K_2CO_3 (588 mg, 4.25 mmol) were then added. 1,3,5-tris(bromomethyl)-2,4,6-triethyl benzene (30 mg, 0.68 mmol) was dissolved in acetonitrile (5 mL) and the solution was added to the **3e** solution. The reaction mixture refluxed for 16 hours. Solvent was evaporated and the resulting crude was purified by flash chromatography using DCM:MeOH 95:5 as eluent to give **1f** as a white solid (21.6 mg, 25.1 mmol, 36% yield).

HRMS (ESI-TOF) m/z [**1f** + H]$^+$ Calc.: 903.5392, found: 903.5500 .

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)= δ 8.19 (s, 3H, H**12**), 7.65 (d, *J* = 7.8 Hz, 3H, H**17**), 7.34 (d, *J* = 8.0 Hz, 3H, H**14**), 7.16 (t, *J* = 7.5 Hz, 3H, H**15**), 7.09 (t, *J* = 9.0 Hz, 3H, H**16**), 7.07 (s, 3H, H**11**), 6.20 (s, 3H, NH_{amide}), 3.84 (AB_q, δ_A=3.33, δ_B=3.86, J_{AB} = 14 Hz, 6H, H**5**), 3.19-3.39 (m, 9H,H**9**, H**4**), 2.95 (m, 3H, H**8**), 2.71 (m, 3H, H**8**), 2.45 (m, 6H, H**2**), 1.81 – 1.71 (m, 2H, H**1**), 1.00 (t, *J* = 7.4 Hz, 6H, H**19**).

^{13}C NMR (101 MHz, CD_3OD): δ(ppm)= 174.57 (C**3**), 142.88 (C**7**), 136.14 (C**10**) , 133.69 (C**6**), 128.09 (C**17**), 123.49 and 122.23 (C**15,16**), 119.61 and 118.85 (C**11,17**), 111.36 (C**14 +13**), 61.36(C**4**), 55.36(C**1**), 43.83(C**5**), 39.65(C**2**), 27.58(C**9**), 22.57(C**8**), 16.73(C**19**).





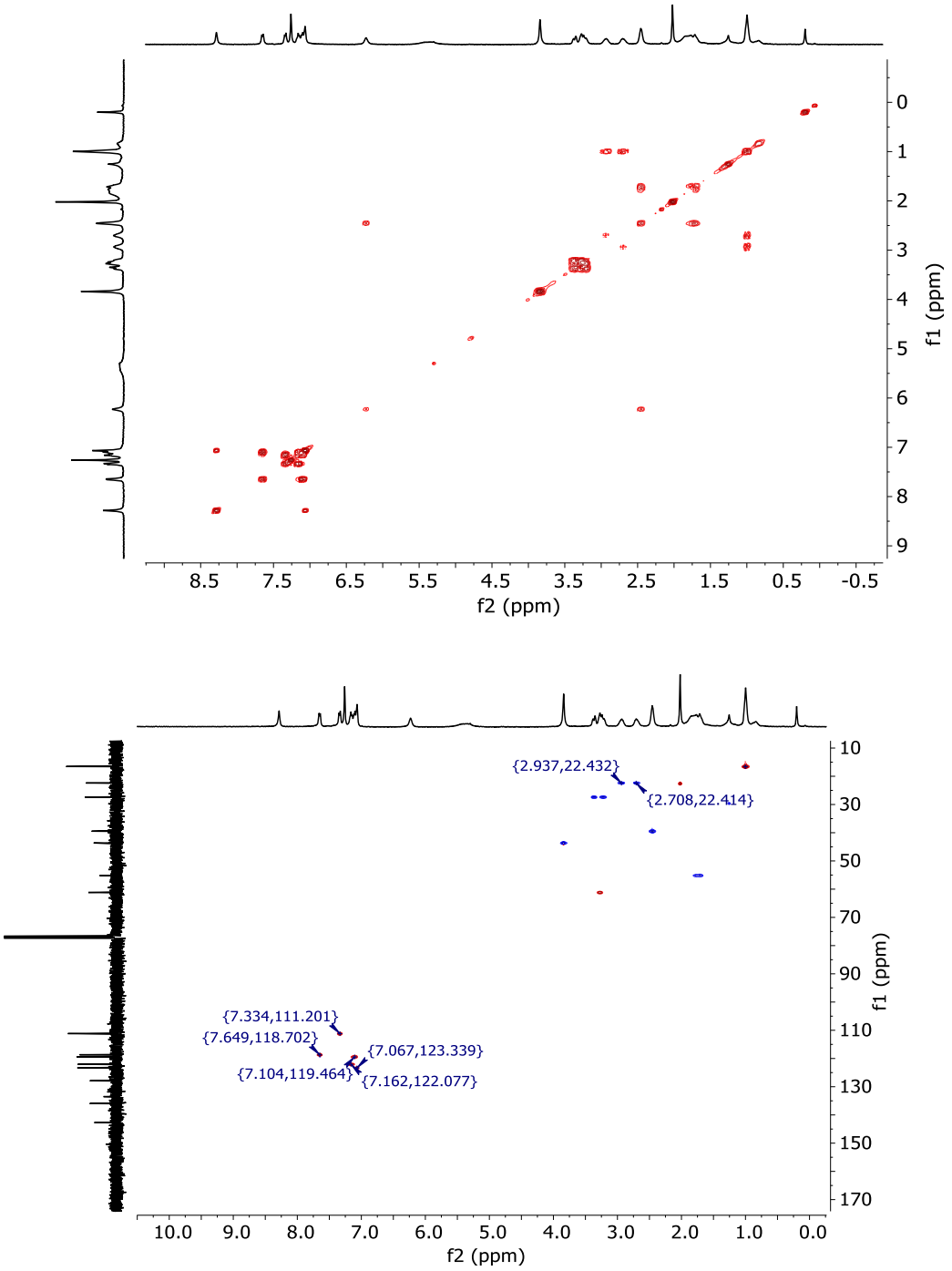


Figure 2.67. ^{1}H-NMR (400 MHz $CDCl_3$) ^{13}C-NMR (101 MHz, CD_3Cl), COSY and HSQC spectra of **1f.**

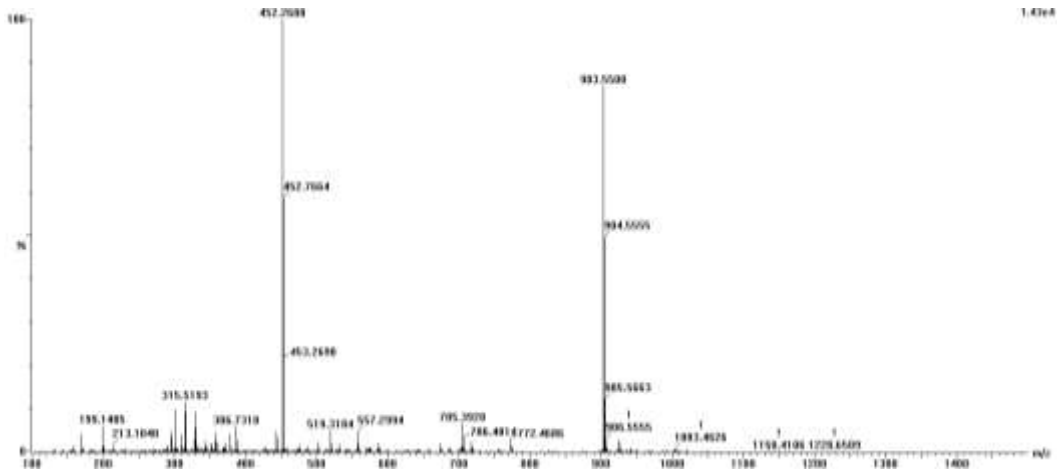
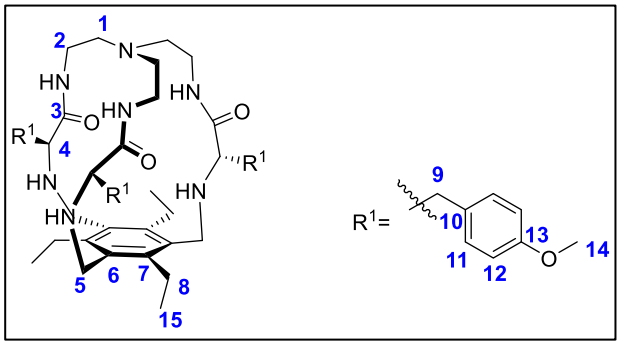


Figure 2.68. HRMS (ESI+) experimental spectrum of **1f**.

2.12.3.6 Synthesis of 1g

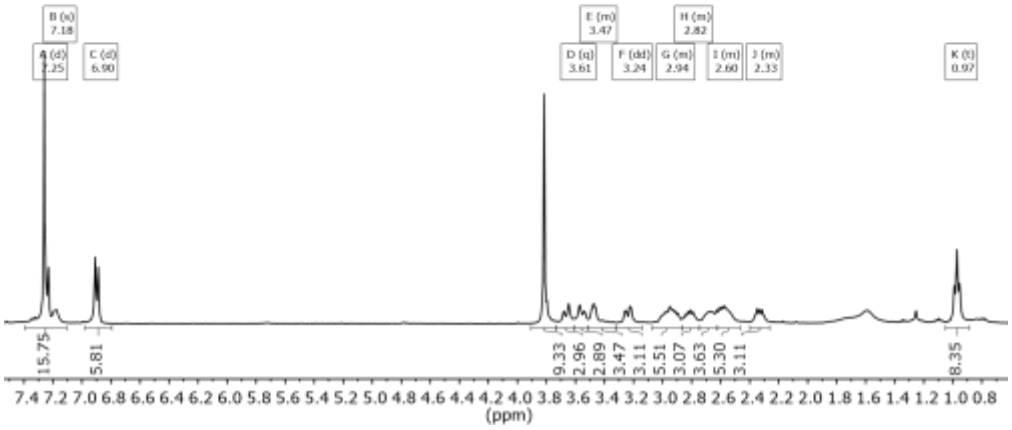


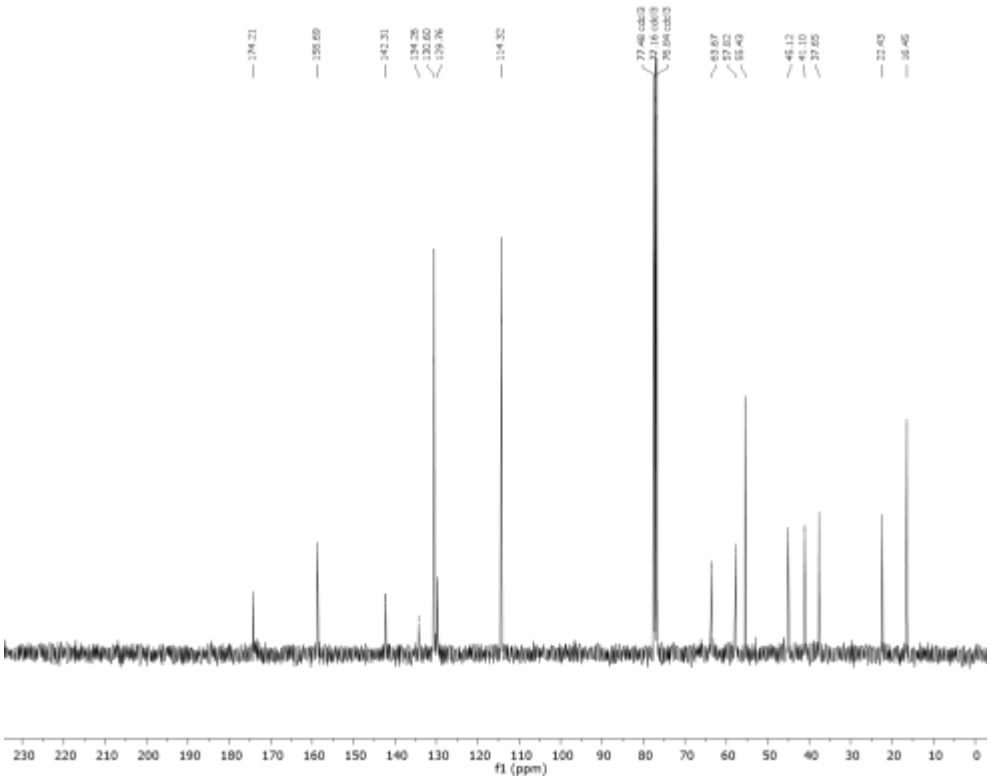
3b (80 mg, 0.706 mmol) was dissolved in dry acetonitrile (20 mL). Bu_4Cl (10 mg, 0.076 mmol) and K_2CO_3 (195 mg, 3 mmol) were added to the solution. 1,3,5-tris(bromomethyl)-2,4,6-triethyl-benzene (28 mg, 0.07 mmol) was dissolved in acetonitrile (10 mL) and the solution was added to the reaction containing **3b**. Reaction mixture refluxed for 16 hours. Solvent was evaporated and the resulting crude was purified by flash chromatography The residue was purified by flash chromatography using DCM:MeOH 95:5 as eluent to give **1g** (23 mg, 26 mmol, 37% yield).

^{1}H NMR (400 MHz, CD_3Cl): δ(ppm)= δ 7.246 (d, *J* = 8.5, Hz, 6H, H**11**), 7.181 (s, 3H, NH_{amide}), 6.898 (d, *J* = 8.5 Hz, 6H, H**12**), 3.816 (s, 9H, H**14**), 3.609 (AB_q, δ_A=3.66, δ_B=3.55, J_{AB} = 13.6 Hz, 6H, H**5**), 3.47 (m, 3H, H**4**), 3.235 (dd, *J* = 16.2, 2.9 Hz, 3H, H**9**), 2.94 (m, 6H, H**9'+2**), 2.82 (m, 3H, H**8**), 2.60 (m, 9H, H**8'+1+2**), 2.33 (m, 3H, H**1'**), 0.972 (t, *J* = 7.4 Hz, 9H, H**15**).

^{13}C NMR (101 MHz, $CDCl_3$) δ 174.21 (C**3**), 158.69 (C**13**), 142.31 (C**7**), 134.26 (C**6**), 130.60 (C**11**), 129.76 (C**10**), 114.32 (C**12**), 63.67 (C**4**), 57.82 (C**1**), 55.43 (C**14**), 45.12 (C**5**), 41.10 (C**2**), 37.65 (C**9**), 22.43 (C**8**), 16.45 (C**15**).

HRMS (ESI-TOF) m/z [**1g** + H]$^+$ Calc.: 876.5382, found: 876.5382.


B (s) 7.18
A (d) 7.25
C (d) 6.90
E (m) 3.47
H (m) 2.82
D (q) 3.61
F (dd) 3.24
G (m) 2.94
I (m) 2.60
J (m) 2.33
K (t) 0.97
15.75
5.81
9.33
2.96
2.89
3.47
3.11
5.51
3.07
3.63
5.30
3.11
8.35
7.4 7.2 7.0 6.8 6.6 6.4 6.2 6.0 5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4 3.2 3.0 2.8 2.6 2.4 2.2 2.0 1.8 1.6 1.4 1.2 1.0 0.8
(ppm)


174.21
159.89
142.31
134.25
132.60
129.76
114.52
77.46
77.16
76.84
63.67
57.82
55.43
46.12
41.10
37.65
22.43
16.46
230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

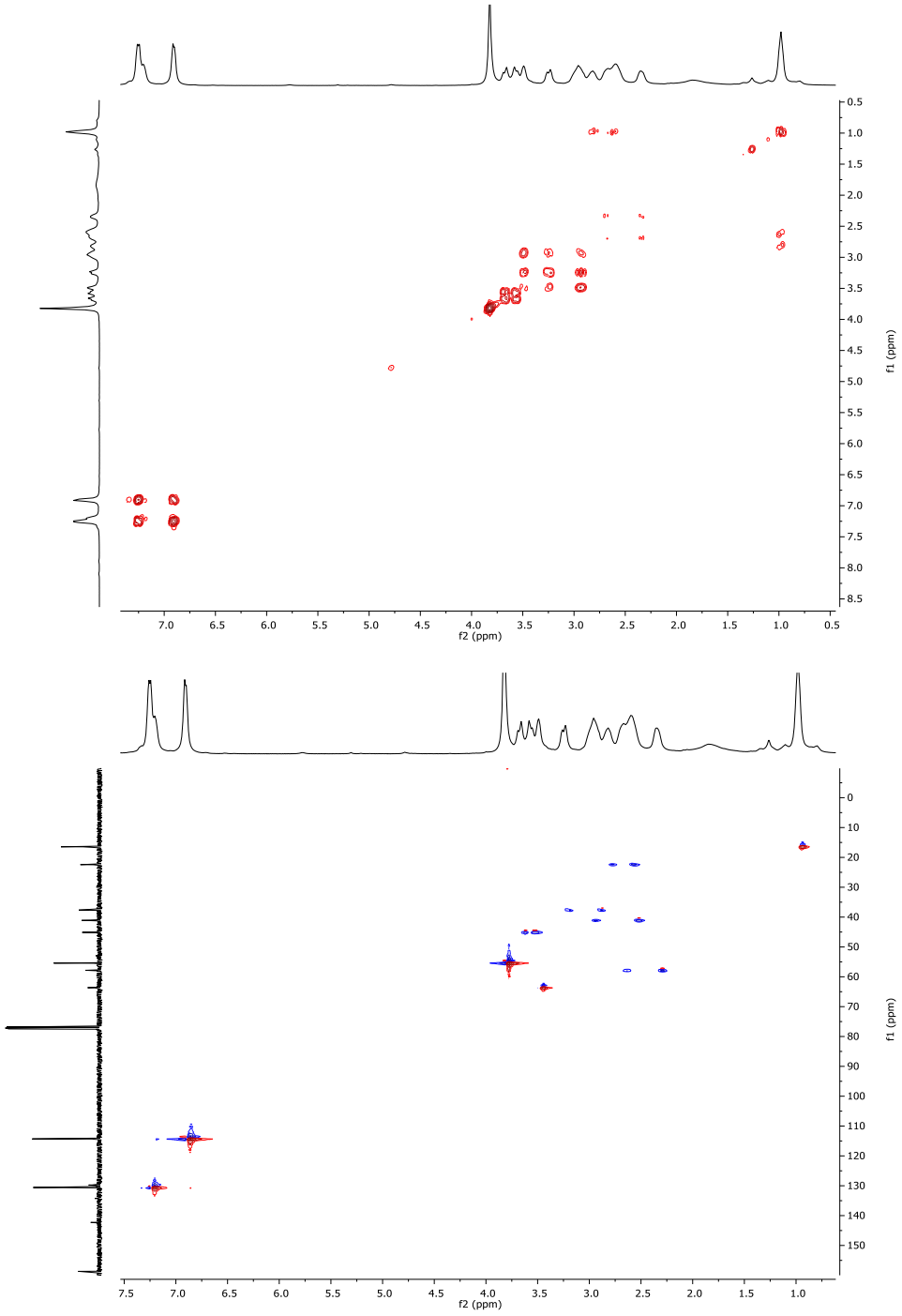


Figure 2.69.^{1}H-NMR (400 MHz $CDCl_3$) ^{13}C-NMR (101 MHz, CD_3Cl), COSY and HSQC spectra of **1g**.

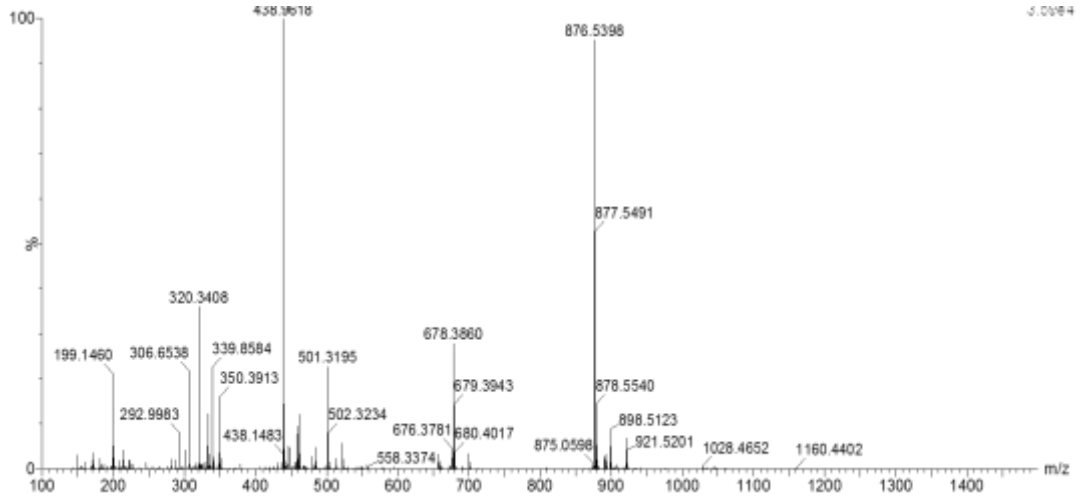


Figure 2.70. HRMS (ESI+) experimental spectrum of **1g.**

2.12.4 X-RAY crystal analysis.

Crystals were obtained by low evaporation of a methanolic solution of the corresponding compounds with an excess of concentrated aqueous HCl.

In **1b**, the methylene groups bonded to N1A are disordered over two sites with a site occupation factor of 0.576(11) for the major occupied sites. The absolute structure could be determined, Flack-x-parameter 0.05(4).

In **1d**, the methylene groups bonded to N1B, N1C, N1D, N1B, N1B, and the water O atom O56 are disordered over two sites with a site occupation factor of 0.627(16), 0.652(15), 0.652(16), 0.516(15), 0.689(16), 0.52(4), respectively, for the major occupied sites.

The displacement parameters of the atoms C17Z, C37Z and C57Z were restrained to an isotropic behaviour.

The H atoms of the solvent water molecules could not be reliably located and were therefore omitted from the refinement. The absolute structure could be determined, Flack-x-parameter -0.01(3).

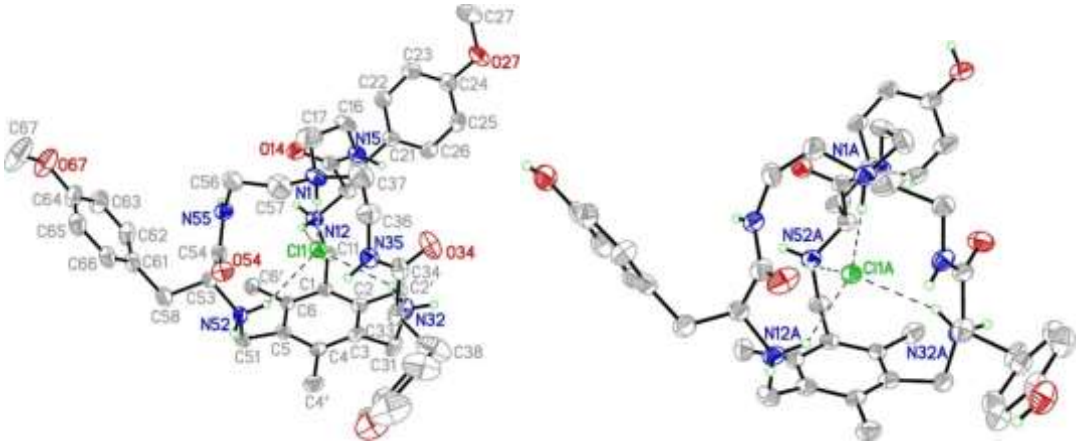


Figure 2.71 Perspective view of **1b** with the atoms labels. Displacement parameters are drawn at the 50% probability level. Additional chloride counterions, solvent molecules and non-polar hydrogen atoms have been omitted for clarity (left). Perspective view of 1d with the most important atoms labels. Displacement parameters are drawn at the 30% probability level. Additional chloride counterions, solvent molecules and non-polar hydrogen atoms have been omitted for clarity (right).

2.12.5 Titration of compounds 1b-g with tetrabuthylamonium chloride: Kass

The titrations were performed with the cage receptor as the fully protonated molecules, using trifluoroacetic acid (TFA). The corresponding tetra-TFA salts were prepared by dissolving each compound in methanol and adding an excess of trifluoroacetic acid, followed by the solvent evaporation and drying in vacuum.

Stock solutions of the cage were prepared by weighting the corresponding amount of the receptor and reaching a final concentration between 1 and 2 mM. The solvent used was 95:5 CD_3CN:H_2O, since this mixture generally allowed a good solubility during the titration experiment and rendered reasonably sharp and well-defined ^{1}H NMR spectra. Besides, under these conditions, the amide proton is detectable during the titration experiments. Additionally, a stock solution of the titrant containing 0.1 M TBACl was prepared by dissolving the salt in the stock solution of the cage. Thus, for each experiment, the solution of the titrant will be 0.1 M in TBACl and 0.001-0.002 M in the receptor therefore maintaining the concentration of the cage constant during the titration experiment. The stock solution of the cage was introduced in a NMR tube and the ^{1}H NMR spectrum (500 MHz, 303 K) was acquired, then small volumes of the stock solution of the titrant were added and the ^{1}H NMR spectrum recorded after each addition.

Different signals changed upon addition of chloride anion, and their variations were fitted to the simplest model using HypNMR 2008 version 4.0.71 software.[97,98] For a suitable comparison of the systems, we calculated the $BC50^0$ parameter, using the BC50 calculator version 2.37.1 program.[78]

Following, we show the stacked plot of the NMR spectra for the titration experiments, the corresponding data set introduced (experimental) and obtained (fit) during the fitting process, the output values (both logβ and $BC50^0$) for the binding for every supramolecular complex, the plot of the experimental (symbols) and the fitted (lines) values of the chemical shifts, and the plot of the simulated species distribution obtained.

2.12.5.1 Titration of 1b

Plots of the titration spectra and data fitting of this compound are included in the main text.

Table 2.4. Data set of 1b titration:

[Cl^-]	Cl^- eq.	HAr2	HAr2 fit	HB	HB fit	HB'	HB' fit	HE	HE fit
0.0000	0.0000	7.365	7.365	4.313	4.313	4.039	4.039	2.27	2.27
0.0005	0.1815	7.603	7.5911	4.396	4.3828	4.086	4.0847	2.318	2.316
0.0011	0.3623	7.789	7.7828	4.447	4.4403	4.12	4.1199	2.356	2.355
0.0016	0.5425	7.926	7.9256	4.482	4.4808	4.137	4.1409	2.38	2.382
0.0022	0.7220	8.025	8.0236	4.505	4.5059	4.146	4.1495	2.397	2.399
0.0027	0.9009	8.092	8.0919	4.519	4.5213	4.148	4.1507	2.408	2.409
0.0032	1.0791	8.133	8.1427	4.528	4.5313	4.148	4.1485	2.415	2.415
0.0038	1.2567	8.175	8.183	4.535	4.5383	4.146	4.1447	2.42	2.420
0.0043	1.4337	8.206	8.2162	4.539	4.5435	4.144	4.1404	2.423	2.423
0.0048	1.6100	8.25	8.2444	4.547	4.5476	4.139	4.1359	2.429	2.426
0.0059	1.9608	8.289	8.2904	4.553	4.5536	4.132	4.1273	2.432	2.430
0.0069	2.3091	8.323	8.3267	4.557	4.558	4.125	4.1195	2.435	2.432
0.0080	2.6549	8.372	8.3562	4.564	4.5613	4.111	4.1127	2.437	2.434
0.0105	3.5088	8.43	8.4108	4.571	4.5669	4.092	4.0991	2.438	2.438
0.0155	5.1724	8.458	8.4756	4.573	4.5731	4.081	4.0818	2.438	2.442

Results of the HypNMR fitting:

- Logß 1= 3.69 ± 0.1 (1:1)

- Log ß 2= 5.89 (1:2) excessive relative error on beta= 37%
- BC_{50}: 198 ±44µM

2.12.5.2 Titration of 1c

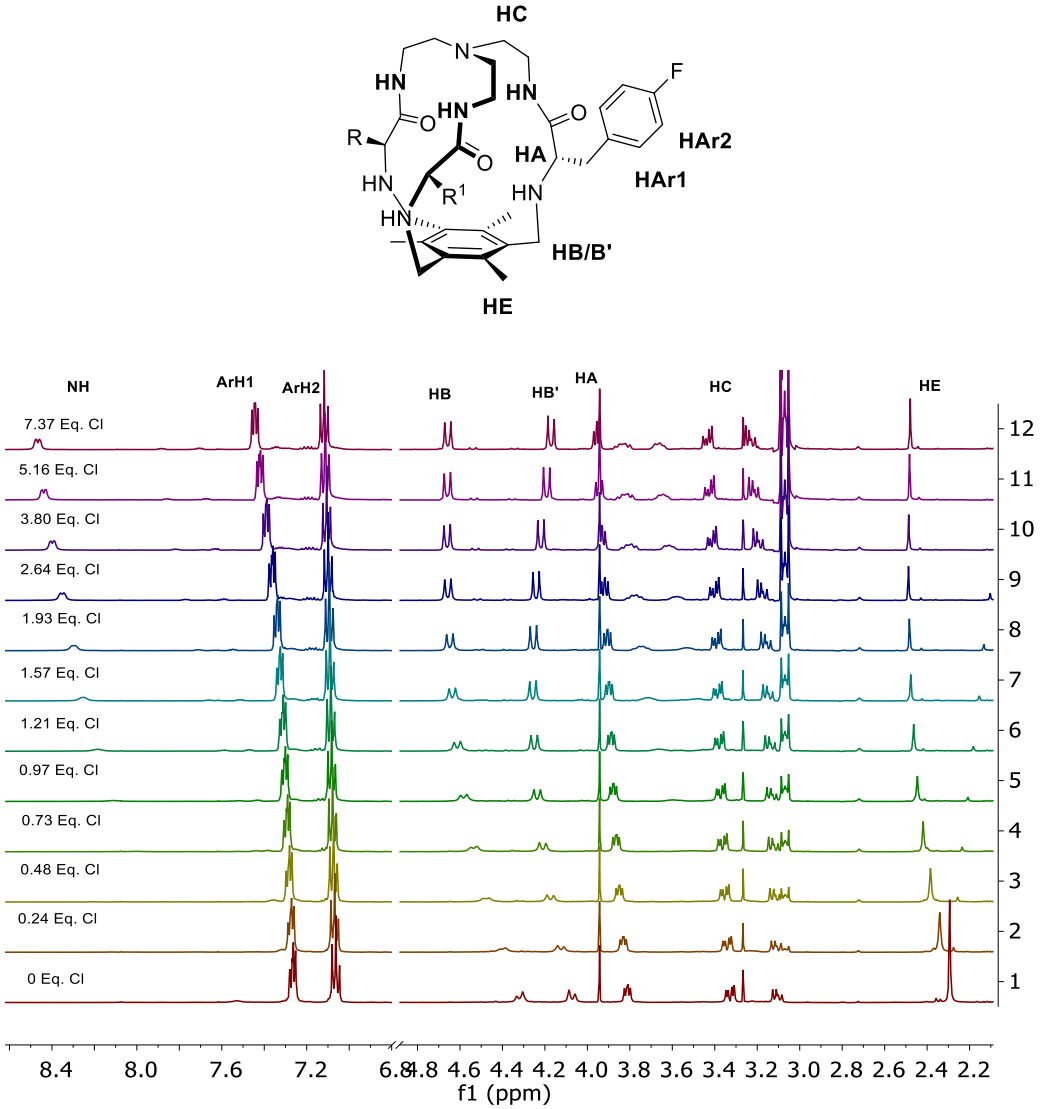


Figure 2.72 1c Stacked ^{1}H NMR spectra for the titration of **1c**

Table 2.5 Data set of **1c** titration:

[Cl^-]	Cl^- eq.	NH	NH fit	HB	HB fit	HB'	HB' fit	HE	HE fit
0.000	0.000	7.528	7.528	4.318	4.318	4.072	4.072	2.294	2.294
0.000	0.244	7.700	7.700	4.399	4.398	4.125	4.124	2.341	2.341

0.001	0.488	7.850	7.859	4.473	4.472	4.174	4.172	2.384	2.383
0.001	0.730	8.000	7.997	4.535	4.534	4.21	4.211	2.418	2.419
0.002	0.972	8.108	8.106	4.582	4.582	4.237	4.238	2.446	2.447
0.002	1.212	8.184	8.183	4.612	4.613	4.25	4.252	2.463	2.464
0.003	1.571	8.253	8.253	4.636	4.637	4.257	4.257	2.477	2.477
0.004	1.928	8.297	8.297	4.648	4.648	4.254	4.254	2.483	2.483
0.005	2.636	8.348	8.349	4.657	4.657	4.241	4.240	2.488	2.486
0.007	3.797	8.397	8.399	4.660	4.660	4.219	4.216	2.487	2.486
0.010	5.161	8.439	8.434	4.659	4.659	4.192	4.195	2.483	2.483
0.014	7.365	8.467	8.470	4.658	4.658	4.171	4.171	2.479	2.480

Results of the HypNMR fitting:

- Log *ß* 1 = 3.82 ± 0.028 (1 : 1)
- Log *ß* 2 = 5.99 ± 0.058 (1 : 2)
- BC_{50} = 148 ± 10 µM

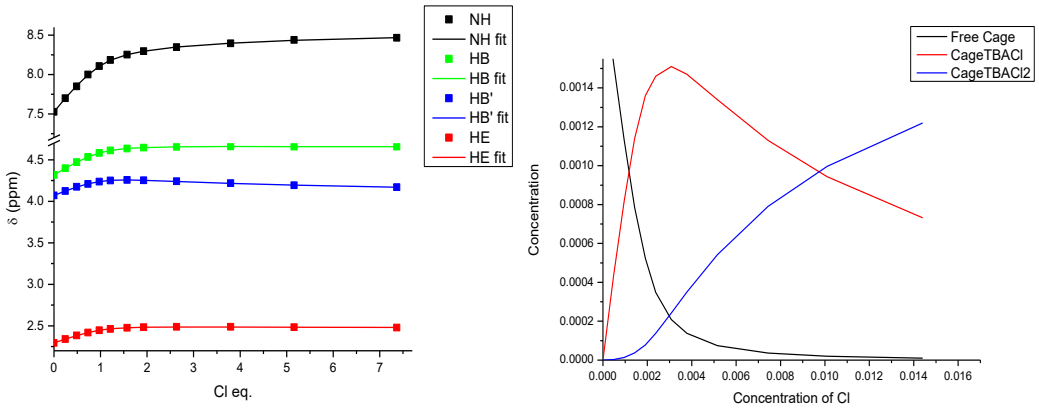


Figure 2.73. Plot of the experimental (symbols) and fitting (lines) data of **1c** titration Figure S45. Species distribution as a function of the chloride concentration of **1c** titration.

2.12.5.3 Titration of 1d

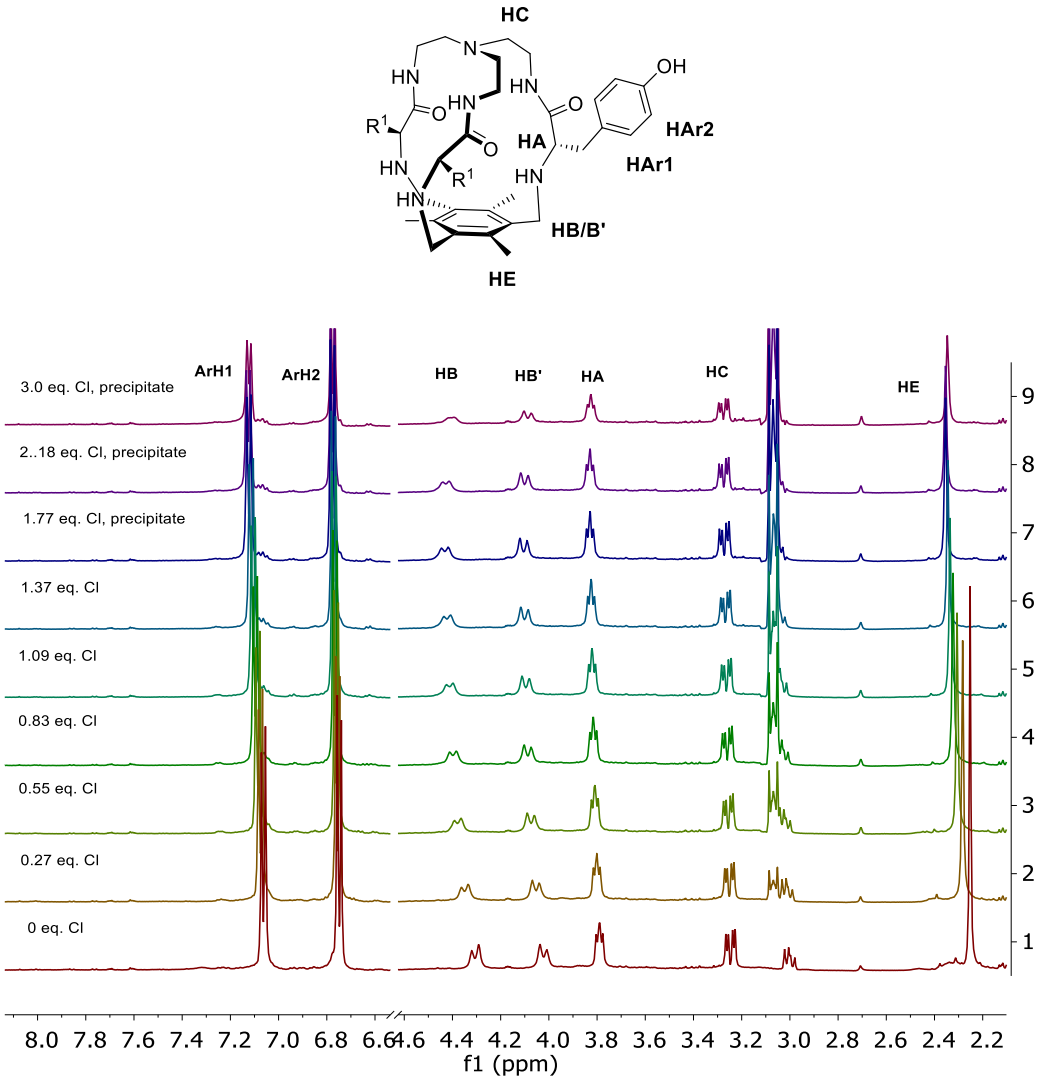


Figure 2.74. 1d Stacked ^{1}H NMR spectra for the titration of **1d.**

Tyrosine cage **1d** starts to precipitate when we add more than one equivalent of TBACl, so association constants in this case are harder to determine. However, it was observed that first and second association constants are within the same order of magnitude, data that we obtain if we fix one value and optimize by fitting the other one. Nuclei used for the adjustment were ArH1. HB/B' and HE.

Table 2.6 Data set of 1d titration:

[Cl^-]	Cl^- eq.	HAr1	HAr1 fit	HB	HB fit	HB'	HB'fit	HE	HE fit
0.0000	0.0000	7.0630	7.0630	4.3050	4.3050	4.0230	4.0230	2.2520	2.2522
0.0005	0.2766	7.0760	7.0755	4.3490	4.3490	4.0540	4.0544	2.2830	2.2834
0.0010	0.5521	7.0860	7.0863	4.3780	4.3783	4.0750	4.0744	2.3060	2.3057
0.0015	0.8266	7.0960	7.0961	4.4000	4.3990	4.0880	4.0880	2.3230	2.3227
0.0020	1.0999	7.1050	7.1050	4.4120	4.4132	4.0960	4.0965	2.3350	2.3356
0.0025	1.3721	7.1130	7.1129	4.4220	4.4215	4.1010	4.1007	2.3450	2.3448

Results of the HypNMR fitting:

- Log *ß* 1 = 3.74 ± 0.04 (1 : 1)
- Log *ß* 2 = 7.45 ± 0.5 (1:2)
- BC_{50}: 115 ± 34 µM

Species distribution as a function of the chloride concentration. In this case the species distribution is not accurate due to precipitation events of the 1:2 complex that is displacing the equilibrium towards the complexation of the cage.

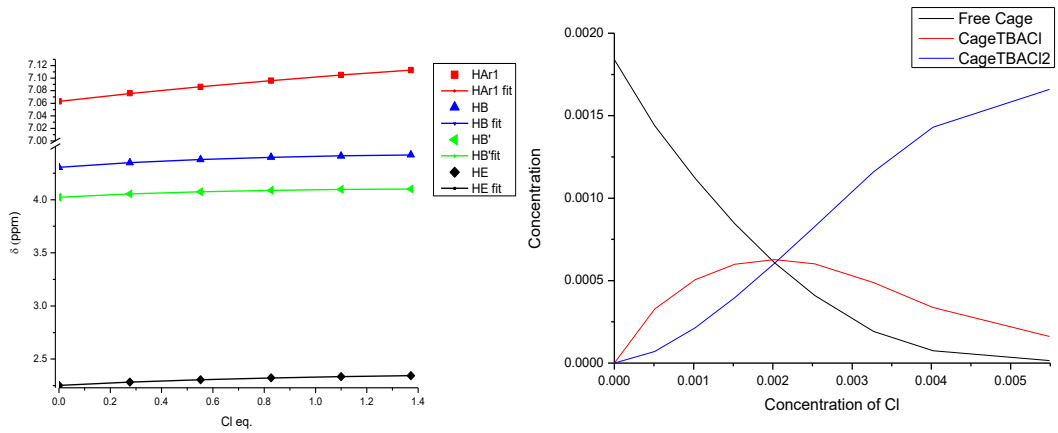


Figure 2.75. Plot of the experimental (symbols) and fitting (lines) data of 1d titration Figura S48. Species distribution as a function of the chloride concentration in 1d titration.

2.12.5.4 Titration of 1e

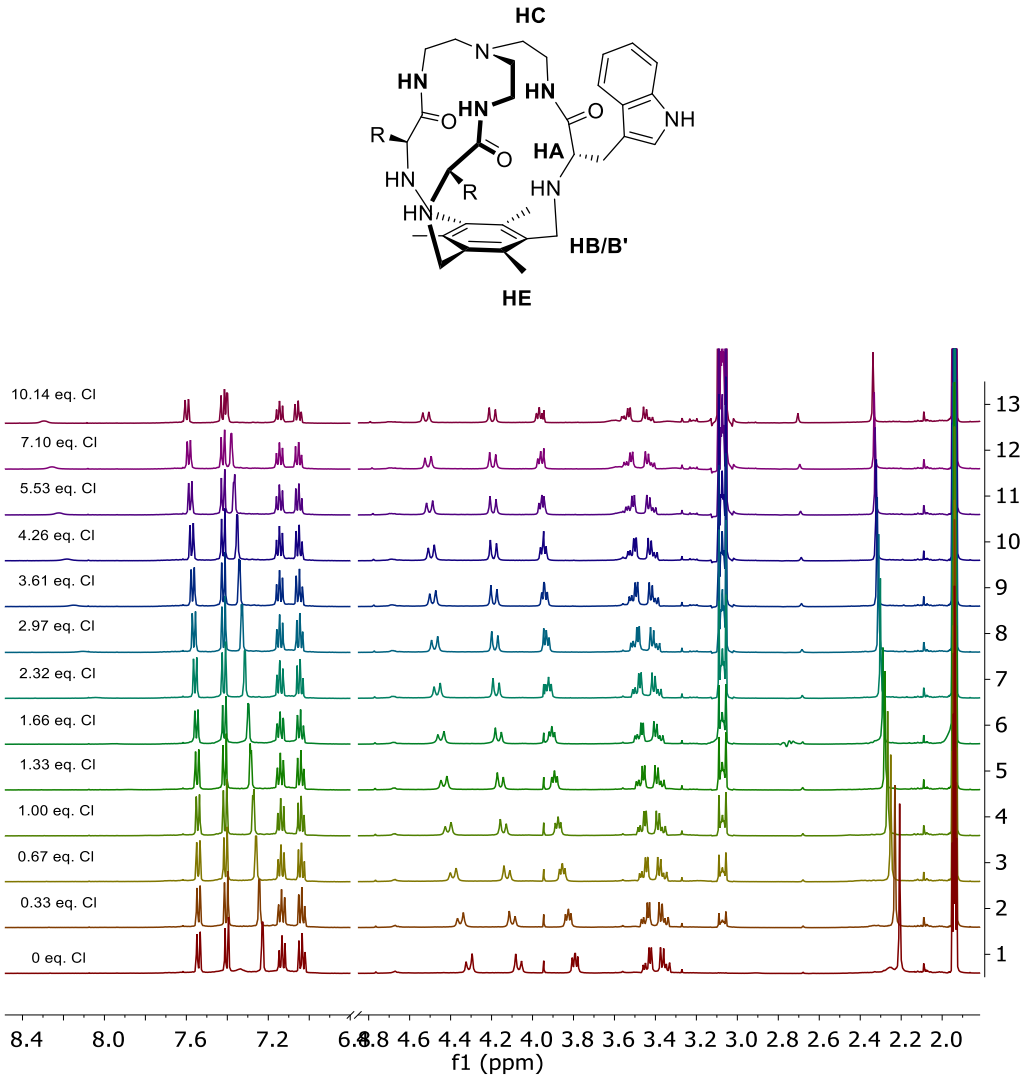


Figure 2.76. 4e Stacked ^{1}H NMR spectra for the titration of **1e.**

Table 2.7 Data set of **1e** titration

[Cl$^-$]	Cl$^-$ eq	NHa	NHa fit	NHb	NHb fit	HB	HB fit	HB'	HB'fit	HA	HA fit
0.00000	0.00	9.67	9.67	7.34	7.34	4.31	4.31	4.07	4.07	3.79	3.79
0.00028	0.33	9.69	9.69		7.52	4.35	4.35	4.10	4.10	3.83	3.83
0.00056	0.67	9.70	9.70		7.67	4.39	4.38	4.13	4.12	3.86	3.85
0.00084	1.00	9.72	9.72	7.79	7.79	4.41	4.41	4.14	4.14	3.88	3.88
0.00112	1.33	9.73	9.73	7.88	7.88	4.43	4.43	4.16	4.16	3.89	3.89

0.00140	1.66	9.74	9.74	7.96	7.95	4.45	4.45	4.17	4.17	3.91	3.91
0.00195	2.32	9.76	9.76	8.04	8.05	4.47	4.47	4.18	4.18	3.92	3.92
0.00250	2.97	9.77	9.77	8.11	8.11	4.48	4.48	4.18	4.19	3.93	3.93
0.00304	3.62	9.79	9.78	8.15	8.15	4.49	4.49	4.19	4.19	3.94	3.94
0.00358	4.26	9.80	9.80	8.18	8.18	4.49	4.49	4.19	4.19	3.95	3.95
0.00465	5.53	9.81	9.81	8.22	8.22	4.50	4.50	4.19	4.19	3.95	3.95
0.00596	7.10	9.83	9.83	8.25	8.25	4.51	4.51	4.19	4.19	3.96	3.96
0.00852	10.2	9.85	9.85	8.29	8.29	4.52	4.52	4.20	4.19	3.96	3.96

Results of the HypNMR fitting:

- Log *ß* 1= 3.40 ± 0.03 (1:1)
- Log *ß* 2 = 5.53 ± 0.12 (1:2)
- BC_{50} = 380 ± 24 µM

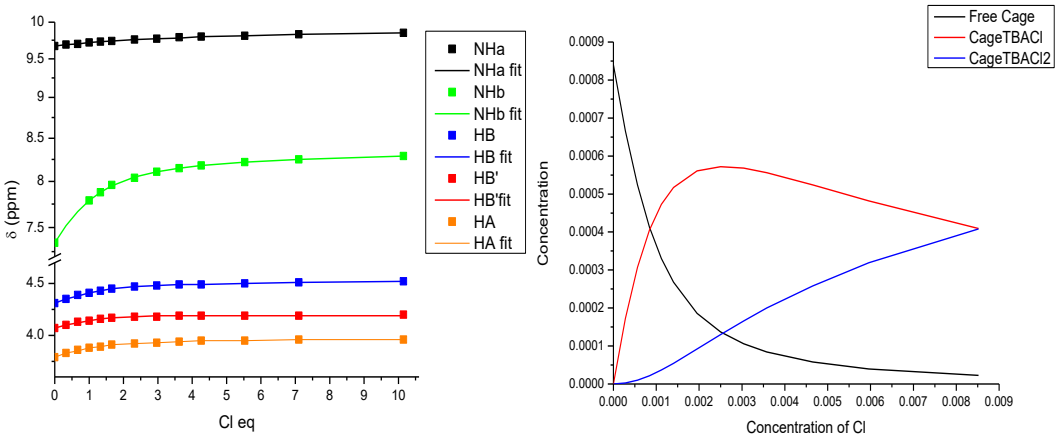


Figure 2.77. Plot of the experimental (symbols) and fitting (lines) data in **1e** titration. Species distribution as a function of the chloride concentration in **1e** titration.

2.12.5.5 Titration of 1f

Data from this titration was not successfully fitted to the 1:1 and 1:2 interaction model used for all the other compounds.

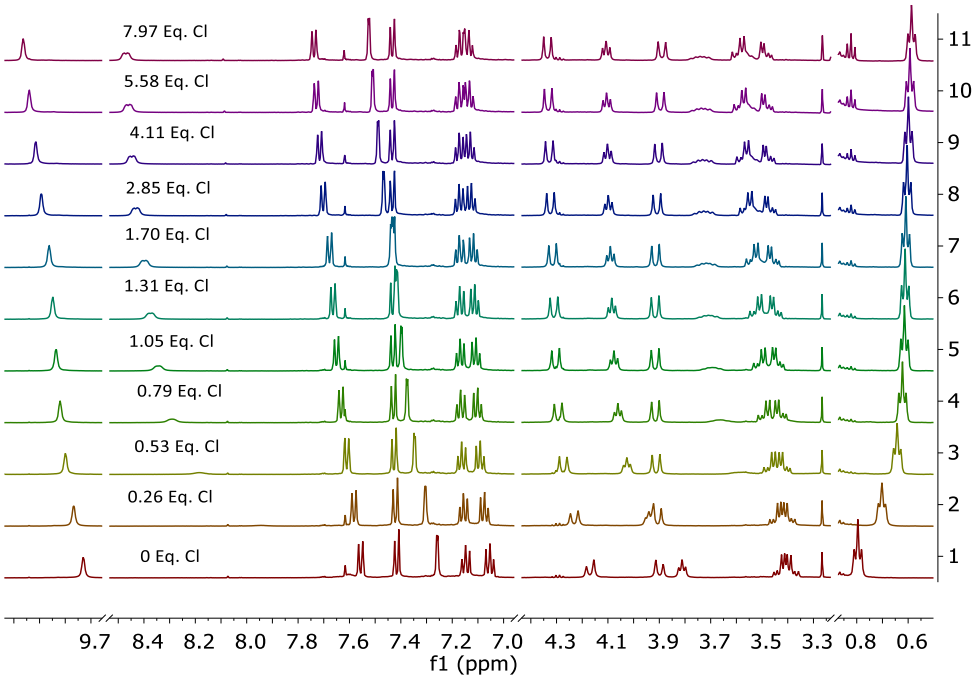


Figure. 2.78. Stacked ^{1}H NMR spectra for the titration of **1f.**

2.12.5.6 Titration of 1g

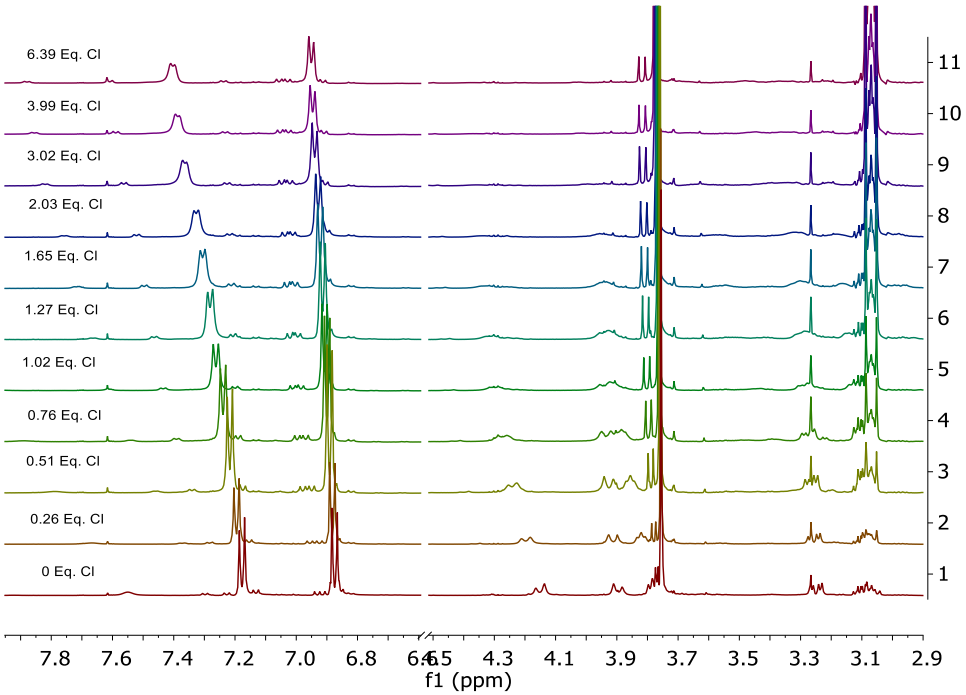


Figure 2.79. Stacked ^{1}H NMR spectra for the titration of **1g**.

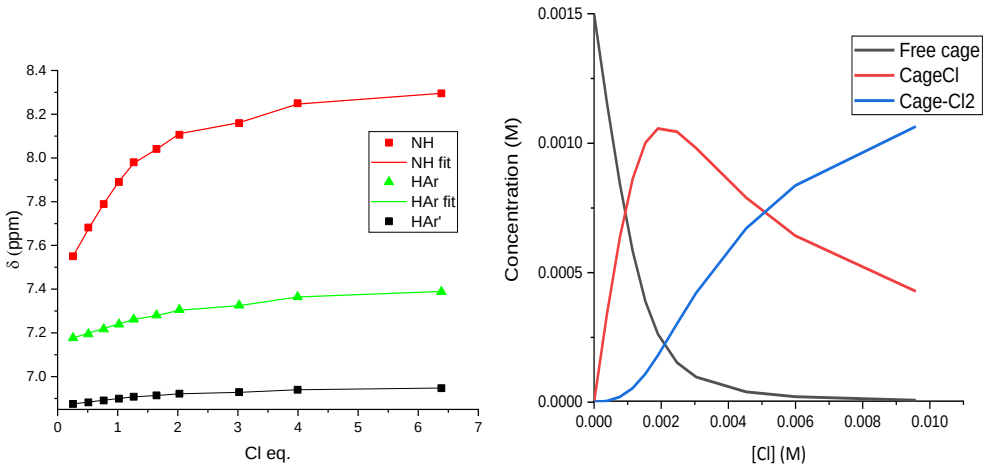


Table 2.8. Data set of **1g** titration

[Cl⁻]	Cl⁻ eq.	H (amide)	H(amide) fit	HAr1	HAr1 fit	HAr2	HAr2Fit
0.000	0.000	7.550	7.550	7.177	7.177	6.875	6.875
0.000	0.256	7.550	7.550	7.177	7.177	6.875	6.875
0.001	0.511	7.682	7.674	7.195	7.199	6.882	6.884
0.001	0.767	7.789	7.790	7.217	7.220	6.891	6.892
0.002	1.020	7.890	7.893	7.239	7.241	6.899	6.900
0.002	1.267	7.980	7.976	7.262	7.261	6.908	6.907
0.002	1.647	8.041	8.040	7.281	7.279	6.914	6.914
0.003	2.027	8.106	8.111	7.306	7.304	6.922	6.922
0.005	3.020	8.159	8.161	7.326	7.325	6.929	6.928
0.006	3.993	8.250	8.246	7.364	7.364	6.940	6.940
0.010	6.387	8.295	8.296	7.388	7.390	6.947	6.948

Results of the HypNMR fitting:

- Log *ß* 1 = 3.92 ± 0.13 (1:1)
- Log *ß* 2 = 6.47 (Excessive relative error)
- BC_{50} = 112 ± 30 µM

2.12.6 Log *P* calculation and Retention time determination

Log *P* values for compounds were calculated using VCCLab software. This software allows the calculation of Log *P* using different methods such as ALOGPS, AC LogP, ALOGP, MLOGP, MiLOGP, XLOPGP2 or XLOGP3, and we used consensus Log *P* as the average of these values.

Neutral (for logP1) and tetraprotonated (for LogP2) forms of the pseudopeptidic tripodal small cages were considered. Reverse phase HPLC (equipment: Agilent technologies 1100 series, column: X-terra MS C18 4.6*150 mm (5 µM)) was also used for measuring the relative lipophilicity of these compounds, since the retention time of each molecule on the reverse phase column is related to its lipophilicity.

All the products were dissolved in MeOH at concentration of 1 mM and eluted using a linear gradient of 20 % to 100 % of acetonitrile in water (using 0.1% of TFA) in 20 minutes.

2.12.7 Vesicles preparation

Two different kinds of vesicles were prepared depending on the assay in which they were going to be used. One of them were made only with POPC (1-palmitoyl-2- oleoyl-sn-glycero-3-phosphocholine), and the other were composed by POPC and cholesterol in a 7:3 ratio, respectively. Cholesterol is needed in some of the assays to increase the stability of the vesicles and obtain robust results. POPC (20 mg/mL = 26.32 mM) and cholesterol (26.32 mM) stock solutions were prepared in chloroform and were kept in the freezer. Below is detailed the vesicle preparation process when 2 mL of POPC solution are used.

For the POPC vesicles preparation, 2 mL of POPC stock solution were added to a 5 mL round bottom flask. In the case of POPC:cholesterol (7:3), a mixture of 2 mL of POPC and 0.6 mL of the cholesterol stock solutions were added to a 5 mL round bottom flask. The solvent was evaporated using a rotary evaporator (20° C) and dry overnight under high vacuum. On the next morning, the obtained lipid film was rehydrated with 1 mL of the required internal solution corresponding to each assay, followed by careful vortexing. The obtained lipid suspension, composed of multilamellar vesicles with different sizes, was subjected to seven freezer and thaw cycles (introducing the flask in a dewar with liquid nitrogen followed by the immersion into warm water). In this point, the lipid solution was made of unilamellar vesicles with different sizes. In order to standardize the size, the suspension was extruded through 200 nm polycarbonate nucleopore membrane using a LiposoFast Basic extruder (Avestin, Inc.).

Once the vesicles were prepared, the next step was removing the non-encapsulated rehydrating solution. There are two different manners of proceeding. One of them consists in placing the lipid suspension in a dialysis membrane and dialysed it against the external required solution (2 times x 500 mL). The other one lies in carrying out a Size Exclusion Chromatography (SEC) on a Sephadex G-50 column. Finally, the vesicles suspension was

place in a 10 mL volumetric flask, to obtain a known-concentration lipid solution (5.26 mM for 2 mL of POPC).

$$2\text{mL} \cdot \frac{20\text{ mg POPC}}{1\text{mL}} \cdot \frac{1\text{ mmol POPC}}{760\text{ mg POPC}} \cdot \frac{1}{10\text{ mL}} = 5.26 \cdot 10^{-3}\text{M} = 5.26\text{ mM POPC}$$

For example, if the POPC concentration in each assay is 0.5 mM in a total volume of 5 mL, each assay contains 475 µL of the 5.26 mM solution and 4525 µL of extravesicular solution:

Cc · Vc= Cd · Vd, 5.26mM · Vc = 0.5 Mm · 5000 µL Vc= 475 µL

In a vesicle assay, the transporter concentration could be expressed in molarity or as a function of the POPC amount. For example, in a 0.5 mM lipid assay, 5 µM is equivalent to 1 mol% carrier to POPC. In all the assays present in this thesis the transport concentration is expressed in mol% carrier to POPC.

$$\frac{0.005\text{ mM carrier}}{0.5\text{ mM POPC}} \cdot 100 = 1\text{ mol\% carrier to POPC}$$

2.12.8 Potentiometric transport assays: Chloride selective electrode assays

POPC vesicles were prepared as described in the section 2.12.7. The non-encapsulated rehydrating solution was removed by dialysis with the external solution (2 times x 500 mL). The intra- and extravesicular composition are outlined below. All ISE assays were carried out using a total volume of 5000 µL. The experimental time of most of the experiments was 300 s. Unilamellar vesicles were suspended in the external solution to give a final lipid concentration of 0.5 mM POPC. At t = 60 s, a DMSO solution of the anion carrier was added to the sample. This pulse was never higher than 25 µL in order to avoid influence of the solvent. The chloride efflux out of the vesicles was monitored over time, using the combination chloride electrode. At t = 300 s or 840 (depending on the experiment), a 20 µL pulse of Triton X-10% (v/v) (detergent) was added in order to lyse the vesicles and release all chloride anions. This total value of chloride anions was considered as 100 % and the rest of dates were calibrated as a function of it. Data was expressed as percentage of Cl^- efflux over time. The intravesicular and extravesicular solutions for each assay were:

NO_3^-/Cl^- exchange at pH 7.2 both in and out of the vesicles:

- Intravesicular solution: 489 mM NaCl, I.S. 500 mM, NaH_2PO_4 5 mM, pH 7.2
- Extravesicular solution: 489 mM $NaNO_3$, I.S. 500 mM, NaH_2PO_4 5 mM, pH 7.2

NO_3^-/Cl^- exchange at pH 7.2 outside and 6.2 inside the vesicles

- Intravesicular solution: 489 mM NaCl, I.S. 500 mM, NaH2PO4 5 mM, pH 6.2
- Extravesicular solution: 489 mM NaNO3, I.S. 500 mM, NaH2PO4 5 mM, pH 7.2

NO_3^-/Cl^- exchange at pH 6.2 outside and 6.2 inside the vesicles

- Intravesicular solution: 489 mM NaCl, I.S. 500 mM, NaH2PO4 5 mM, pH 6.2
- Extravesicular solution: 489 mM NaNO3, I.S. 500 mM, NaH2PO4 5 mM, pH 6.2

NO_3^-/Cl^- exchange with K^+ as counterion

- Intravesicular solution: 489 mM KCl, I.S. 500 mM, NaH_2PO_4 5 mM, pH 7.2
- Extravesicular solution: 489 mM KNO3, I.S. 500 mM, NaH_2PO_4 5 mM, pH 7.2

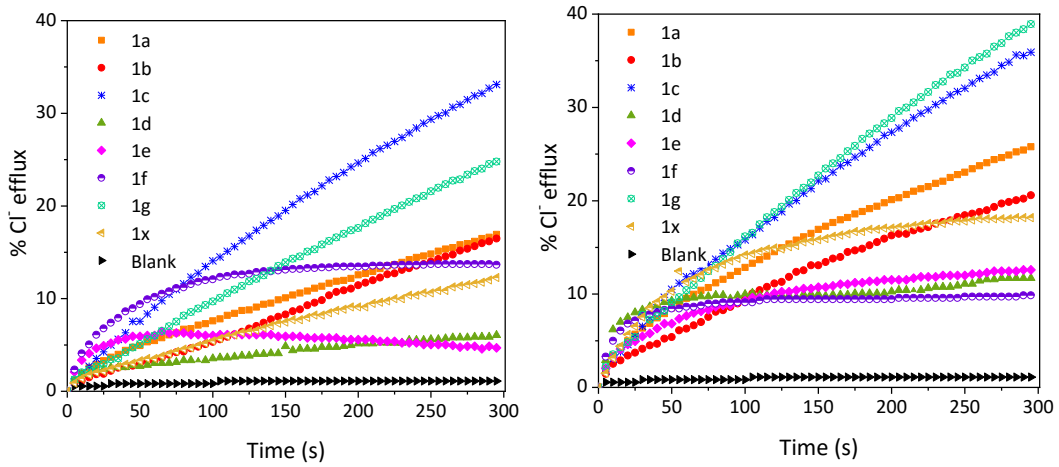


Figure 2.80. Anion transport activity of **1a-g, x** (50 µM =10 mol%) ISE experiment. 0.5 mM POPC.Left: Intravesicular pH 6.2, extravesicular pH 7.2; Right: Intravesicular pH 6.2, extravesicular pH 6.2.

2.12.9 Fluorescence based transport assays

These assays are based on fluorophores whose fluorescence varies as a function of changes in different variables such as concentration (carboxyfluorescein), pH (HPTS) or the presence of some ions (lucigenin is selectively quenched by chloride). POPC and POPC : cholesterol (7:3) vesicles were prepared as previously described. The non-encapsulated rehydrating solution was removed by SEC, using as mobile phase the external buffer. The intra- and extravesicular compositions are outlined below. The total volume in the cuvettes was 2500

µL. Unless other specification, the anion carrier was added at t = 60 s and fluorescence changes were monitored over time.

2.12.9.1 Lucigenin based assays

Lucigenin is a bisacridinium salt whose fluorescence emission is quenched by halides. (K_{sv}, Cl^- = 390 M^{1}).[99,100] Based on these properties, lucigenin could be employed to monitor the chloride transport inside vesicles at 503 nm after exciting at 372 nm. Data collected was represented as initial fluorescence divided by fluorescence vs time (f_0/f). In this assays, cage compounds have to be dissolved in methanol instead of DMSO since Lucigenin emission is quenched in the presence of DMSO.

POPC:cholesterol (7:3) vesicles were loaded with $NaNO_3$ (102.2 mM $NaNO_3$, I.S. 150 mM, NaH_2PO_4 20 mM, pH 7.2; lucigenin 3 mM). The non-encapsulated solution was removed by SEC, using as mobile phase the nitrate solution. The obtained vesicles were suspended over $NaNO_3$ (102.2 mM. $NaNO_3$, I.S. 150 mM, NaH_2PO_4 20 mM, pH 7.2). At the beginning of the experiment, a pulse of NaCl was added in order to obtain a final concentration of 10 mM in the sample. After one minute the compound was added and the fluorescence decrease was monitored over time during five more minutes. Data obtained was normalized as f_0/f and plotted against time.

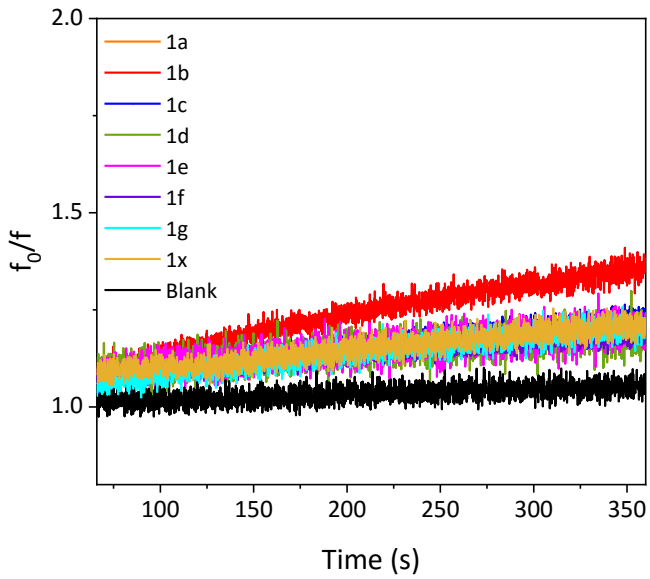


Figure 2.81. I_0/I normalization of lucigenin fluorescence emission upon addition of cages **1a-g,x** to POPC:Cholesterol (7:3) vesicles, 0.5 mM POPC. Vesicles contained $NaNO_3$ (102.2

mM $NaNO_3$, I.S. 150 mM, NaH_2PO_4 20 mM, pH 7.2; lucigenin 3 mM) were suspended in $NaNO_3$ (102.2 mM $NaNO_3$, I.S. 150 mM, NaH_2PO_4 20 mM, pH 7.2). At t = 0 s was added a pulse of NaCl to obtain a final concentration of 10 mM. At t = 60 s the anion carrier was added (10 mol% carrier to lipid; 0.05 mM). Blank(12.5 µL MeOH). Each trace represents the average of at least three different trials, done with at least two different batches of vesicles

The experimental conditions were optimized. Vesicles concertation was reduced to 0.25 mM. They were loaded with $NaNO_3$ (102.2 mM $NaNO_3$, I.S. 150 mM, NaH_2PO_4 20 mM, pH 7.2; lucigenin 3 mM) and suspended over NaCl (102.2 mM. NaCl, I.S. 150 mM, NaH_2PO_4 20 mM, pH 7.2) 0.25 mM POPC. After one minute the cage was added and the fluorescence decrease was monitored over time during five more minutes. Data obtained was normalized as f_0/f and plotted against time. λexc = 372 nm; λem = 503 nm.

2.12.9.2 Carboxyfluorescein based assays

The experimental conditions used in this assay were:

- POPC vesicles 0.05 mM
- Intravesicular solution: 451 mM NaCl, I.S. 500 mM, NaH_2PO_4 20 mM, pH 7.2; CF 50 mM.
- Extravesicular solution: 150 mM Na_2SO_4, I.S. 500 mM, NaH_2PO_4 20 mM, pH 7.2
- t = 0 s vesicles + buffer
- t = 60 s anion carrier addition in DMSO solution
- t = 360 s detergent. The fluorescence was recorded 2 more minutes
- Total time = 480 s
- λ_{exc} = 490 nm; λ_{em} = 520 nm

Data processing: The following equation is applied to the obtained data.

$$CF\ Leakage = \left[\frac{I_t - I_0}{I_{max} - I_0}\right]_{\lambda=520\ nm}$$

I_t → fluorescence intensity at time t.

I_0 → fluorescence intensity at time 0 s.

I_{max} → maximum fluorescence intensity observed after addition of detergent.

2.12.9.3 HPTS based assays

The experimental conditions used in each of the HPTS assays conducted were:

HPTS based assay with $NaNO_3$:

- POPC:cholesterol (7:3) vesicles 0.5 mM in POPC
- Intravesicular solution: 126.2 mM NaNO3, I.S. 150 mM, NaH2PO4 10 mM, pH 6.5; HPTS 10 µM
- Extravesicular solution: 126.2 mM NaNO3, I.S. 150 mM, NaH2PO4 10 mM, pH 7.5
- t = 0 s vesicles + buffer
- t = 60 s anion carrier addition in DMSO solution.
- Total time 360 s
- λexc = 403 nm and 460 nm; λem = 510 nm

HPTS based assay with NaCl:

- POPC:cholesterol (7:3) vesicles 0.5 mM in POPC
- Intravesicular solution: 126.2 mM NaCl, I.S. 150 mM, NaH_2PO_4 10 mM, pH 6.5; HPTS 10 µM
- Extravesicular solution: 126.2 mM NaCl, I.S. 150 mM, NaH_2PO_4 10 mM, pH 7.5
- t = 0 s vesicles + buffer
- t = 60 s anion carrier addition in DMSO solution.
- Total time 360 s
- λ_{exc} = 403 nm and 460 nm; λ_{em} = 510 nm

The relationship between the fluorescence intensity at 510 nm when the sample is excited at both 460 and 403 nm allows to determine the internal pH at each time:

$$\frac{Fi\ em\ 510\ nm\ (\lambda ex\ 460\ nm)}{Fi\ em\ 510\ nm\ (\lambda ex\ 403\ nm)}$$

A calibration curve for determining the actual pH value as function of the ratiometric emission of HPTS was carried out using 15 nM HPTS in the phosphate buffer solutions used in the measurements (with $NaNO_3$ and with NaCl) and adding consecutive aliquots of NaOH 0.5 M to the solution to raise the pH from 5.5 to 9.5. After each addition the fluorescence

rate was read. The representation of the emission ration (I_{406}/I_{403}) vs the pH was adjusted to a S-Logistic model.[101]

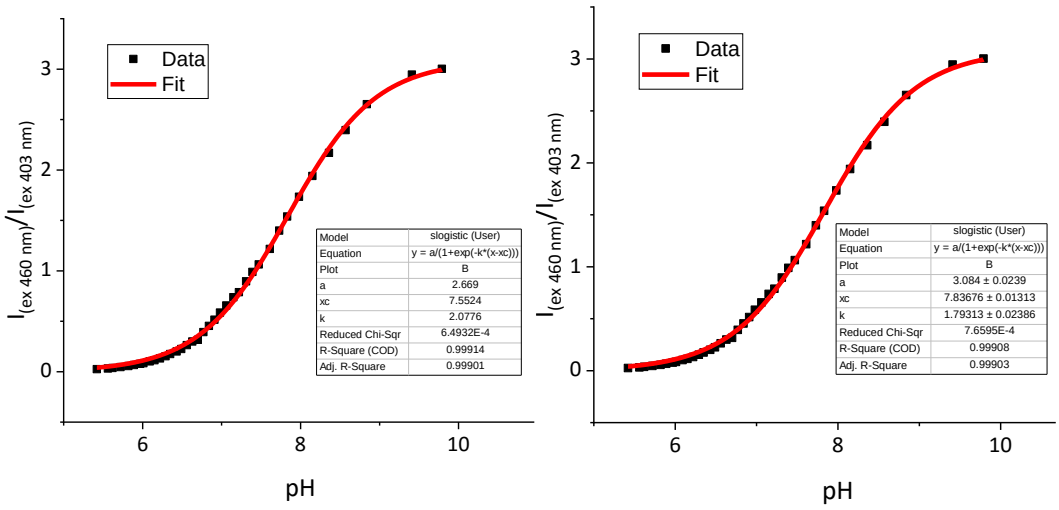


Figure 2.82. HPTS calibration s-logistic1 model (Left: 126.25 mM $NaNO_3$, 10 mM buffer phosphate; Right: 126.25 mM NaCl, 10 mM buffer phosphate).

The conversion of fluorescence data into pH was done with the corresponding S-Logistic1 model obtained from the calibration curve at the experimental condition.

2.12.10 Cell cultures

Human lung adenocarcinoma cells, A549 were maintained in Dulbecco's modified eagle's medium (DMEM; Gibco Thermo Fisher Scientific, USA) with carbonate buffer. Cells were kept in the logarithmic growth phase by routine passage every 2-3 days using 0.025% trypsin-EDTA treatment.

pH 7.2 and 6.2 media were prepared replacing the carbonate by PIPES buffer (10 mM) and adjusting the pH with HCl or NaOH. pH 7.6 media (with carbonate buffer) is the same used for cell growth.

2.12.11 Cytotoxicity vs cancer cells

The antiproliferative activity of compounds **1a-g,x** was evaluated using the MTT assay method. Cells were seeded 24 hours prior to treatment in 96-well plate (100 µL of a suspension $2.5 \cdot 10^5$ cell/mL). The following day, culture media from the wells was replaced by 100 µL of fresh media at each pH containing the desired concentration of the compound to be tested. All the compounds were previously dissolved in DMSO at a concentration of

20 mM. The final concentration of DMSO used in the corresponding wells did not exceeded 1% (v/v). This concentration does not affect cell viability. Negative control cultures received the same concentration of solvent alone.

Cells were incubated for 24h in the presence of compounds **1a-1g,x** at 37°C in a humidified atmosphere with 5% CO_2. At the end of incubation, culture media was removed and 100 µL of MTT solution (5mg/mL diluted with plain culture media 1 : 5) was added to each well and incubated for 4 hours. Afterwards, MTT solution was discarded. The purple formazan crystal formed at the bottom of the wells was dissolved with 100 µL of DMSO and stirred for 30 minutes at room temperature. The absorbance at 570 nm was read on a spectrophotometer plate reader. The proportion of surviving cells was calculated as Absorbance of treated sample/ Absorbance of control x 100. Dose-response curves were constructed to obtain IC_{50} values using Origin dose-response fitting function. All experimental data were derived from at least 3 independent experiments.

Control cells were grown in the same conditions (same cell culture media, buffer and pH) as treated cells to ensure that the difference in viability is only associated to the corresponding cages and not the buffer or the pH of the medium.

In order to ensure the suitability of PIPES buffer in the experimental conditions (10 mM, 24h incubation) pH of the media with PIPES at both pH 7.2 and 6.2 was measured before incubation, and after 24 hours incubation with A549 cancer cells in the presence of the cages. The corresponding pH changes in all the cases were lower than 0.12 pH units so we concluded that this buffer concentration was enough to maintain the extracellular pH constant during the experiment.

Statistical analysis: GraphPad Prism v5.0 software (GraphPad Software Inc., La Jolla, USA) was used for statistical analysis. For all experiments, the obtained results of the triplicates were represented as means with standard deviation (SD).

2.12.12 Cytotoxicity of 1c at different concentrations of Cl^-.

The anti-proliferative activity of compound **1c** in DPBS with different concentrations of Cl^- in the buffer was carried out as follows.

Cells were seeded 24 hours prior to treatment in 96-well plate (100 µL of a suspension $2.5 \cdot 10^5$ cell/mL). The following day, culture media from the wells was replaced by 100 µL of fresh DPBS with the desired concentration of Cl^- and previously adjusted to pH 6.4

containing different concentrations of compound 1c (from the stock solution 20 mM in DMSO). Negative control cultures received the same concentration of solvent alone. The rest of the assay was carried out as described for the previous MTT assays.

DPBS solutions with different concentrations of Cl^- were prepared by replacing NaCl for sodium gluconate. pH was adjusted before each experiment to 6.4 by adding NaOH or NaCl.

2.13 References

1 D. F. Shriver and M. J. Biallas, *J. Am. Chem. Soc.*, 1967, **89**, 1078–1081.

2 C. H. Park and H. E. Simmons, *J. Am. Chem. Soc.*, 1968, **90**, 2431–2432.

3 E. Graf and J. M. Lehn, *J. Am. Chem. Soc.*, 1975, **97**, 5022–5024.

4 E. Graf and J. M. Lehn, *J. Am. Chem. Soc.*, 1976, **98**, 6403–6405.

5 J. M. Lehn, *Pure Appl. Chem.*, 1977, **49**, 857–870.

6 F. P. Schmidtchen, *Angew. Chemie Int. Ed. English*, 1977, **16**, 720–721.

7 F. P. Schmidtchena and G. Mullerb, 1984, 1115–1116.

8 P. A. Gale, *Chem. Commun.*, 2011, **47**, 82–86.

9 B. L. Schottel, H. T. Chifotides and K. R. Dunbar, *Chem. Soc. Rev.*, 2008, **37**, 68–83.

10 K. Kavallieratos, C. M. Bertao and R. H. Crabtree, *J. Org. Chem.*, 1999, **64**, 1675–1683.

11 V. Amendola, L. Fabbrizzi and L. Mosca, *Chem. Soc. Rev.*, 2010, **39**, 3889–3915.

12 S. Kubik, R. Kirchner, D. Nolting and J. Seidel, *J. Am. Chem. Soc.*, 2002, **124**, 12752–12760.

13 C. Schmuck, *Chem. - A Eur. J.*, 2000, **6**, 709–718.

14 T. Gunnlaugsson, P. E. Kruger, P. Jensen, J. Tierney, H. D. P. Ali and G. M. Hussey, *J. Org. Chem.*, 2005, **70**, 10875–10878.

15 T. Ábalos, S. Royo, R. Martínez-Máñez, F. Sancenón, J. Soto, A. M. Costero, S. Gil and M. Parra, *New J. Chem.*, 2009, **33**, 1641–1645.

16 R. Planells-Cases and T. J. Jentsch, *Biochim. Biophys. Acta - Mol. Basis Dis.*, 2009, **1792**, 173–189.

17 S. Hong, M. Bi, L. Wang, Z. Kang, L. Ling and C. Zhao, *Oncol. Rep.*, 2015, **33**, 507–514.

18 K. O. Alfarouk, S. B. M. Ahmed, A. Ahmed, R. L. Elliott, M. E. Ibrahim, H. S. Ali, C. C. Wales, I. Nourwali, A. N. Aljarbou, A. H. H. Bashir, S. T. S. Alhoufie and S. S. Alqahtani, 2020, 1–22.

19 I. Alfonso and R. Quesada, *Chem. Sci.*, 2013, **4**, 3009–3019.

20 S. P. Yu, L. M. T. Canzoniero and D. W. Choi, *Curr. Opin. Cell Biol.*, 2001, **13**, 405–411.

21 X. Yu, X. Hong, Q. Mao and W. Chen, *Eur. J. Med. Chem.*, 2019, **184**, 111782.

22 S. P. Zheng, L. B. Huang, Z. Sun and M. Barboiu, *Angew. Chemie - Int. Ed.*, 2021, **60**, 566–597.

23 S. Ko, S. K. Kim, A. Share, V. M. Lynch, J. Park, W. Namkung, W. Van Rossom, N. Busschaert, P. A. Gale and J. L. Sessler, *Nat. Chem.*, 2014, **6**, 885–892.

24 N. Busschaert, S. H. Park, K. H. Baek, Y. P. Choi, J. Park, E. N. W. Howe, J. R. Hiscock, L. E. Karagiannidis, I. Marques, V. Félix, W. Namkung, J. L. Sessler, P. A. Gale and I. Shin, *Nat. Chem.*, 2017, **9**, 667–675.

25 J. T. Davis, P. A. Gale and R. Quesada, *Chem. Soc. Rev.*, 2020, **49**, 6056–6086.

26 E. Oosterwijk and R. J. Gillies, *Philos. Trans. R. Soc. B Biol. Sci.*, 2014, **369**, 1–7.

27 L. A. Jowett, E. N. W. Howe, X. Wu, N. Busschaert and P. A. Gale, *Chem. – A Eur. J.*, 2018, **24**, 10475–10487.

28 C. M. Dias, H. Li, H. Valkenier, L. E. Karagiannidis, P. A. Gale, D. N. Sheppard and A. P. Davis, *Org. Biomol. Chem.*, 2018, **16**, 1083–1087.

29 N. Akhtar, A. Saha, V. Kumar, N. Pradhan, S. Panda, S. Morla, S. Kumar and D. Manna, *ACS Appl. Mater. Interfaces*, 2018, **10**, 33803–33813.

30 S. N. Berry, V. Soto-Cerrato, E. N. W. Howe, H. J. Clarke, I. Mistry, A. Tavassoli, Y. T. Chang, R. Pérez-Tomás and P. A. Gale, *Chem. Sci.*, 2016, **7**, 5069–5077.

31 A. Roy, O. Biswas and P. Talukdar, *Chem. Commun.*, 2017, **53**, 3122–3125.

32 T. Saha, M. S. Hossain, D. Saha, M. Lahiri and P. Talukdar, *J. Am. Chem. Soc.*, 2016, **138**, 7558–7567.

33 N. Busschaert, S. H. Park, K. H. Baek, Y. P. Choi, J. Park, E. N. W. Howe, J. R. Hiscock, L. E. Karagiannidis, I. Marques, V. Félix, W. Namkung, J. L. Sessler, P. A. Gale and I. Shin, *Nat. Chem.*, 2017, **9**, 667–675.

34 X. Bao, X. Wu, S. N. Berry, E. N. W. Howe, Y.-T. Chang and P. A. Gale, *Chem. Commun.*, 2018, **54**, 1363–1366.

35 A. M. Rodilla, L. Korrodi-Gregório, E. Hernando, P. Manuel-Manresa, R. Quesada, R. Pérez-Tomás and V. Soto-Cerrato, *Biochem. Pharmacol.*, 2017, **126**, 23–33.

36 S. J. Moore, M. Wenzel, M. E. Light, R. Morley, S. J. Bradberry, P. Gómez-Iglesias, V. Soto-Cerrato, R. Pérez-Tomás and P. A. Gale, *Chem. Sci.*, 2012, **3**, 2501–2509.

37 L. A. Jowett, E. N. W. Howe, V. Soto-Cerrato, W. Van Rossom, R. Pérez-Tomás and P. A. Gale, *Sci. Rep.*, 2017, **7**, 1–11.

38 C. Montoya, R. Cervantes and J. Tiburcio, *Tetrahedron Lett.*, 2015, **56**, 6177–6182.

39 L. González-Mendoza, B. Altava, M. I. Burguete, J. Escorihuela, E. Hernando, S. V Luis, R. Quesada and C. Vicent, *RSC Adv.*, 2015, **5**, 34415–34423.

40 C. R. Elie, G. David and A. R. Schmitzer, *J. Med. Chem.*, 2015, **58**, 2358–2366.

41 P. Motloch, A. Guerreiro, C. Q. Azeredo, G. J. L. Bernardes, C. A. Hunter and I. Kocsis, *Org. Biomol. Chem.*, 2019, **17**, 5633–5638.

42 C. C. Tong, R. Quesada, J. L. Sessler and P. A. Gale, *Chem. Commun.*, 2008, 6321–6323.

43 P. A. Gale, C. C. Tong, C. J. E. Haynes, O. Adeosun, D. E. Gross, E. Karnas, E. M. Sedenberg, R. Quesada and J. L. Sessler, *J. Am. Chem. Soc.*, 2010, **132**, 3240–3241.

44 P. A. Gale, J. T. Davis and R. Quesada, *Chem. Soc. Rev.*, 2017, 46, 2497–2519.

45 C. J. E. Haynes, S. J. Moore, J. R. Hiscock, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2012, **3**, 1436–1444.

46 J. L. Seganish and J. T. Davis, *Chem. Commun.*, 2005, 5781–5783.

47 B. Montaner and R. Pérez-Tomás, *Life Sci.*, 2001, **68**, 2025–2036.

48 W. Van Rossom, D. J. Asby, A. Tavassoli and P. A. Gale, *Org. Biomol. Chem.*, 2016, **14**, 2645–2650.

49 J. R. Casey, S. Grinstein and J. Orlowski, , DOI:10.1038/nrm2820.

50 V. Soto-Cerrato, P. Manuel-Manresa, E. Hernando, S. Calabuig-Fariñas, A. Martínez-Romero, V. Fernández-Dueñas, K. Sahlholm, T. Knöpfel, M. García-Valverde, A. M. Rodilla, E. Jantus-Lewintre, R. Farràs, F. Ciruela, R. Pérez-Tomás and R. Quesada, *J. Am. Chem. Soc.*, 2015, **137**, 15892–15898.

51 X. Wu, E. N. W. Howe and P. A. Gale, *Acc. Chem. Res.*, 2018, **51**, 1870–1879.

52 E. Hernando, V. Capurro, C. Cossu, M. Fiore, M. García-Valverde, V. Soto-Cerrato, R. Pérez-Tomás, O. Moran, O. Zegarra-Moran and R. Quesada, *Sci. Rep.*, 2018, **8**, 1–10.

53 J. L. Sessler, L. R. Eller, W.-S. Cho, S. Nicolaou, A. Aguilar, J. T. Lee, V. M. Lynch and D. J. Magda, *Angew. Chemie Int. Ed.*, 2005, **44**, 5989–5992.

54 S. J. Moore, C. J. E. Haynes, J. González, J. L. Sutton, S. J. Brooks, M. E. Light, J. Herniman, G. J. Langley, V. Soto-Cerrato, R. Pérez-Tomás, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2013, **4**, 103–117.

55 X.-H. Yu, C.-C. Peng, X.-X. Sun and W.-H. Chen, *Eur. J. Med. Chem.*, 2018, **152**, 115–125.

56 V. Saggiomo, S. Otto, I. Marques, V. Félix, T. Torroba and R. Quesada, *Chem. Commun.*, 2012, **48**, 5274–5276.

57 N. Busschaert, M. Wenzel, C. J. E. Haynes, J. R. Hiscock, I. L. Kirby, L. E. Karagiannidis, S. J. Moore, N. J. Wells, J. Herniman, G. J. Langley, P. N. Horton, M. E. Light, I. Marques, P. J. Costa, V. Félix, J. G. Frey and P. A. Gale, *Chem. Sci.*, 2013, **4**, 3036–3045.

58 M. J. Spooner, H. Li, I. Marques, P. M. R. Costa, X. Wu, E. N. W. Howe, N. Busschaert, S. J. Moore, M. E. Light, D. N. Sheppard, V. Félix and P. A. Gale, *Chem. Sci.*, 2019, **10**, 1976–1985.

59 X. Yu, X. Hong and W. Chen, *Org. Biomol. Chem.*, 2019, **17**, 1558–1571.

60 N. Busschaert, M. Wenzel, M. E. Light, P. Iglesias-hern, P. Ricardo and P. A. Gale, 2011, 14136–14148.

61 Y. R. Choi, G. C. Kim, H.-G. Jeon, J. Park, W. Namkung and K.-S. Jeong, *Chem. Commun.*, 2014, **50**, 15305–15308.

62 S. B. Salunke, J. A. Malla and P. Talukdar, *Angew. Chemie Int. Ed.*, 2019, **58**, 5354–5358.

63 H. Kobayashi, *Cancers (Basel).*, , DOI:10.3390/cancers9040036.

64 B. A. Webb, M. Chimenti, M. P. Jacobson and D. L. Barber, *Nat. Rev. Cancer*, 2011, **11**, 671–677.

65 T. Fukamachi, Y. Chiba, X. Wang, H. Saito and M. Tagawa, *Cancer Lett.*, 2010, **297**, 182–189.

66 D. A. Senthebane, A. Rowe, N. E. Thomford, H. Shipanga, D. Munro, M. A. M. Al Mazeedi, H.

A. M. Almazyadi, K. Kallmeyer, C. Dandara, M. S. Pepper, M. I. Parker and K. Dzobo, *Int. J. Mol. Sci.*, , DOI:10.3390/ijms18071586.

67 D. Wijesinghe, M. C. M. Arachchige, A. Lu, Y. K. Reshetnyak and O. A. Andreev, *Sci. Rep.*, 2013, **3**, 2–8.

68 N. Busschaert, R. B. P. Elmes, D. D. Czech, X. Wu, I. L. Kirby, E. M. Peck, K. D. Hendzel, S. K. Shaw, B. Chan, B. D. Smith, K. A. Jolliffe and P. A. Gale, *Chem. Sci.*, 2014, **5**, 3617–3626.

69 N. Busschaert, I. L. Kirby, S. Young, S. J. Coles, P. N. Horton, M. E. Light and P. A. Gale, *Angew. Chemie Int. Ed.*, 2012, **51**, 4426–4430.

70 R. B. P. Elmes, N. Busschaert, D. D. Czech, P. A. Gale and K. A. Jolliffe, *Chem. Commun.*, 2015, **51**, 10107–10110.

71 A. Roy, O. Biswas and P. Talukdar, *Chem. Commun.*, 2017, **53**, 3122–3125.

72 E. N. W. Howe, N. Busschaert, X. Wu, S. N. Berry, J. Ho, M. E. Light, D. D. Czech, H. A. Klein, J. A. Kitchen and P. A. Gale, *J. Am. Chem. Soc.*, 2016, **138**, 8301–8308.

73 S. V. Shinde and P. Talukdar, *Angew. Chemie - Int. Ed.*, 2017, **56**, 4238–4242.

74 A. Roy, D. Saha, P. S. Mandal, A. Mukherjee and P. Talukdar, *Chem. - A Eur. J.*, 2017, **23**, 1241–1247.

75 A. Moure, S. V. Luis and I. Alfonso, *Chem. - A Eur. J.*, 2012, **18**, 5496–5500.

76 I. Martí, J. Rubio, M. Bolte, M. I. Burguete, C. Vicent, R. Quesada, I. Alfonso and S. V. Luis, *Chem. - A Eur. J.*, 2012, **18**, 16728–16741.

77 I. Martí, M. Bolte, M. I. Burguete, C. Vicent, I. Alfonso and S. V. Luis, *Chem. - A Eur. J.*, 2014, **20**, 7458–7464.

78 A. Vacca, O. Francesconi and S. Roelens, *Chem. Rec.*, 2012, **12**, 544–566.

79 I. Martí, M. Bolte, M. I. Burguete, C. Vicent, I. Alfonso and S. V. Luis, *Chemistry*, 2014, **20**, 7458–7464.

80 http://www.vcclab.org/lab/alogps/.

81 P. Xin, H. Kong, Y. Sun, L. Zhao, H. Fang, H. Zhu, T. Jiang, J. Guo, Q. Zhang, W. Dong and C. P. Chen, *Angew. Chemie - Int. Ed.*, 2019, **58**, 2779–2784.

82 P. Xin, H. Kong, Y. Sun, L. Zhao, H. Fang, H. Zhu, T. Jiang, J. Guo, Q. Zhang, W. Dong and C. Chen, *Angew. Chemie*, 2019, **131**, 2805–2810.

83 B. A. McNally, A. V. Koulov, B. D. Smith, J. B. Joos and A. P. Davis, *Chem. Commun.*, 2005, 1087–1089.

84 R. C. MacDonald, R. I. MacDonald, B. P. Menco, K. Takeshita, N. K. Subbarao and L. R. Hu, *Biochim. Biophys. Acta*, 1991, **1061**, 297–303.

85 L. A. Jowett and P. A. Gale, *Supramol. Chem.*, 2019, **31**, 297–312.

86 P. H. Schlesinger, N. K. Djedovič, R. Ferdani, J. Pajewska, R. Pajewski and G. W. Gokel, *Chem. Commun.*, 2003, **3**, 308–309.

87 E. Brunetti, A. Inthasot, F. Keymeulen, O. Reinaud, I. Jabin and K. Bartik, *Org. Biomol. Chem.*, 2015, **13**, 2931–2938.

88 D. A. Kallick, M. R. Tessmer, C. R. Watts and C. Y. Li, *J. Magn. Reson. B*, 1995, **109**, 60–65.

89 E. Pambou, J. Crewe, M. Yaseen, F. N. Padia, S. Rogers, D. Wang, H. Xu and J. R. Lu, *Langmuir*, 2015, **31**, 9781–9789.

90 C. L. Perrin and T. J. Dwyer, *Chem. Rev.*, 1990, **90**, 935–967.

91 A. Vargas Jentzsch, D. Emery, J. Mareda, P. Metrangolo, G. Resnati and S. Matile, *Angew. Chemie - Int. Ed.*, 2011, **50**, 11675–11678.

92 J. C. Stockert, R. W. Horobin, L. L. Colombo and A. Blázquez-Castro, *Acta Histochem.*, 2018, **120**, 159–167.

93 T. Fukamachi, X. Wang, Y. Mochizuki, C. Maruyama, H. Saito and H. Kobayashi, *Int. Immunopharmacol.*, 2013, **17**, 148—153.

94 et al. odish H, Berk A, Zipursky SL, *Molecular cell biology. 4th edition.*, 2000.

95 G. M. Sheldrick, *Acta Crystallogr. Sect. A*, 2008, **64**, 112–122.

96 A. L. Spek, *Acta Crystallogr. Sect. D*, 2009, **65**, 148–155.

97 C. Frassineti, S. Ghelli, P. Gans, A. Sabatini, M. S. Moruzzi and A. Vacca, *Anal. Biochem.*, 1995, **231**, 374–382.

98 C. Frassineti, L. Alderighi, P. Gans, A. Sabatini, A. Vacca and S. Ghelli, *Anal. Bioanal. Chem.*, 2003, **376**, 1041—1052.

99 J. Biwersi, B. Tulk and A. S. Verkman, *Anal. Biochem.*, 1994, **219**, 139–143.

100 C. Huber, K. Fähnrich, C. Krause and T. Werner, *J. Photochem. Photobiol. A Chem.*, 1999, **128**, 111–120.

101 N. Busschaert, P. A. Gale, C. J. E. Haynes, M. E. Light, S. J. Moore, C. C. Tong, J. T. Davis and W. A. Harrell, Jr., *Chem. Commun.*, 2010, **46**, 6252–6254.

102 Stoe & Cie, *X-AREA*. Diffractometer control program system. Stoe & Cie, Darmstadt, Germany, **2002**.

CHAPTER 3: FLUORINATED PSEUDOPEPTIDIC CAGES FOR Cl^- TRANSPORT

In chapter 2 we reported the possibility to use a small pseudopeptidic cage-like receptor to recognize chloride with a pH-modulated activity and enhanced cytotoxicity vs human lung adenocarcinoma cell cultures in acid media.

Among the different compounds studied, we found very good results with the cage containing a fluorine atom in the R_1 aromatic substituent (compound **1c**) and a methyl group in R_2. In order to improve the activity and selectivity, and to understand the physicochemical properties of those molecules leading to this particular behaviour, we decided to prepare a new family of cage-like compounds containing different fluorinated aromatic substituents in R_1 (different amount of fluorine atoms and in different positions of the aromatic ring) maintaining a methyl group in position R_2. We chose Me substituent for R_2 because lipophilicity is slightly lower with a methyl and solubility improves compared to using an ethyl. Moreover, the cyclization reaction proceeds with higher yields. To do so, we selected different commercially available (N-fluoro)-phenylalanine amino acids ranging from the mono fluorinated to the pentafluorinanted phenylalanine to get compounds **1h-o** from figure 3.1. The better performance of aromatic substituents in R_1 had also been observed previously[1,2] and remarked in chapter 2, that is why only phenylalanine derivatives were used.

3.1 RESULTS AND DISCUSSION

3.1.1 Synthesis of new fluorinated cages

This new family of pseudopeptidic cages containing fluorine in different positions of the phenyl ring were prepared following the previously described methodology for compound **1c** (chapter 2) which involves the amide coupling of the amino-protected aa with the TREN, followed by the deprotection of the amines and finally the cyclization of the open intermediate with 2,4,6-Tris-(bromomethyl)benzene (figure 3.1). For compounds **1m** and **1l** an additional step to protect the aa was required since the Boc-protected starting material was not commercially available. Experimental details for each compound (including yield and characterization) are provided in the experimental section. Table 3.1 summarizes the global yield for the synthesis of these cages. In all the cases the limiting step is the macrocyclization reaction but it still proceeds with reasonable yields to furnish the different fluorinated cages **1h-o**.

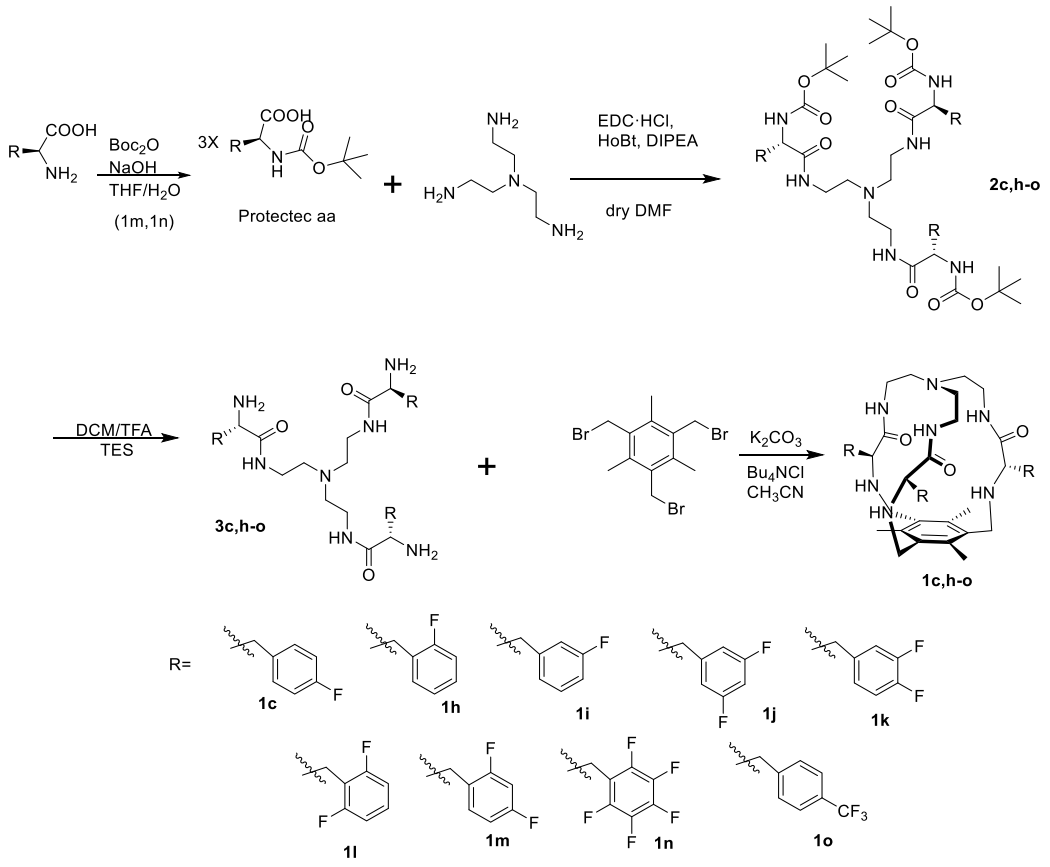


Figure 3.1. Synthetic scheme for the preparation of compounds **1c, 1h-o.**

Table 3.1 Overall yield of the synthesis of compounds **1h-1o** (calculated from the protected aa).

Cage	**1h**	**1i**	**1j**	**1k**	**1l**	**1m**	**1n**	**1o**
Yield %	25	37	43	36	34	33	33	20

3.1.2 X-Ray diffraction crystal strature analysis

Compounds **1c** (chapter 2) and **1h** were successfully crystallized as their tetra-HCl salt and their structures resolved by X-RAY diffraction (Figure 3.2).

The crystals obtained show the inclusion of a chloride molecule inside the inner cavity of the cage and three additional chloride ions per cage molecule externally bound. The main chloride binding site of **1h** is the same cavity observed in the cage-like hosts from chapter 2. This result was expected since the core of the cage and the binding cavity have not been

modified. Furthermore, this encourages us to believe that the chloride complexation for the other fluorinated cages also takes place mainly (with higher affinity) in the inner cavity of the cage although no suitable crystals for X-Ray analysis have been obtained for the rest of compounds.

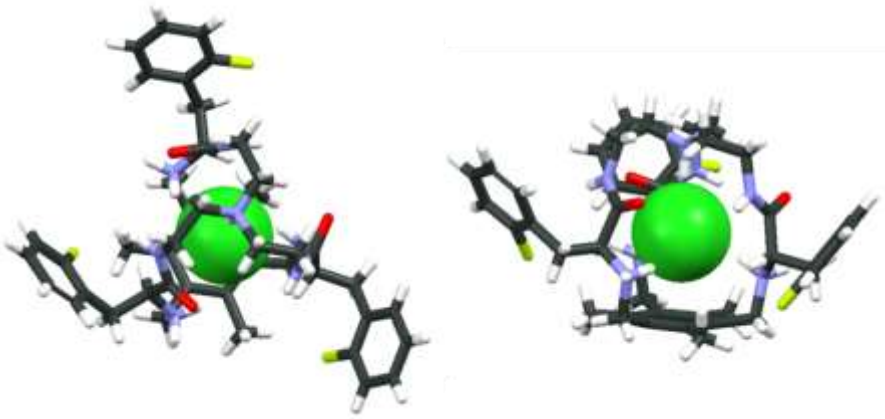


Figure 3.2. Solid state X-Ray structure of 1h·4HCl salt (X-ray diffraction) with the encapsulated chloride in CPK. Additional chloride counterions and solvent molecules have been omitted for clarity.

3.1.3 Chloride binding studies by NMR spectroscopy

The interaction of compounds **1h, 1i, 1m** and **1l** with the chloride ion was studied by ^{1}H NMR titration of the tetrakis(trifluoroacetate) salt of each compound with tetrabuthylamonium chloride (TBACl) in 5 % aqueous CD_3CN at 298 K. In the three cases the signals more affected by the interaction with chloride were those of the amide NH, the aromatic CH_2, the CH of the chiral centre and the CH_3 of the benzene ring. These signals were used for the fitting of the titrations with TBACl. As previously reported for compound **1c**, the newly synthetized cage compounds show very similar chloride binding properties between each other (table 3.2). They possess the ability to bind a Cl^- presumably inside their cavity and another chloride with lower strength somewhere outside the cavity. The binding constant corresponding to the 1 : 1 complex formation (~10^4 M^{-1}) is around two orders of magnitude higher than the one for the 1 : 2 complex (~10^2 M^{-1}) which in the case of **1h, 1m** and **1l** is not even detected.

For comparison purposes the chloride affinity was expressed in terms of the BC_{50}^{0} parameter (table 3.2). For the four compounds studied (**1h, 1i, 1m, 1l**) 1:1 constant is within the high μM range.

The affinity of the rest of the derivatives (**1j, 1k, 1n, 1o**) towards chloride was not directly measured through this method for mainly two reasons. The inner binding pocket is conserved between all the cages studied so the same range of affinities for the formation of the 1:1 complex was expected for the rest of compounds in their tetraprotonated form. Secondly, we have seen that with small to no variations in the binding constant measured through NMR with the mentioned conditions, the activity of the cages varies substantially so we inferred that the different biological activity is not directly linked to the affinity measurements for this family of cages. Besides, these compounds have no cytotoxic activity as it will be explained in section 3.1.7. The protonation state of the cages will also be very different in the membrane compared to the conditions of the NMR measurements, so this accounts for our reservations when linking Cl^- binding affinity with transport or biological activity.

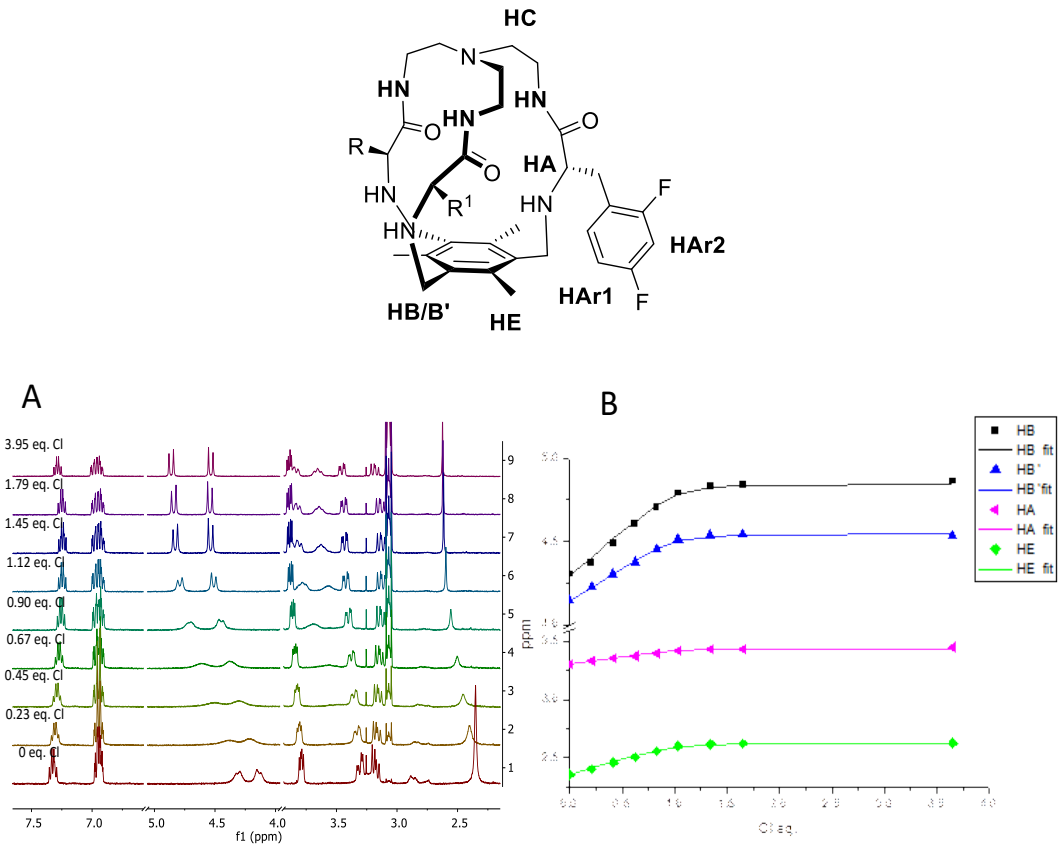


Figure 3.3. A) Stacked ^{1}H NMR spectra NMR (500 MHz, 5% aqueous CD_3CN at 303 K for the titration of **1m** (2mM as the TFA salt) with TBACl. B) Results of HypNMR for the fitting of

1m titration. Symbols represent the experimental chemical shifts; lines correspond to the calculated ones for the best fitting.

Table 3.2. $LogK_{1:1}$, $LogK_{1:2}$ and $BC_{50}{}^{0}$ for compounds **1c, h-l** with chloride determined through NMR titration.

Cage	R1	$LogK_{1:1}$ (M^{-1})	$LogK_{1:2}$ (M^{-1})	$BC_{50}{}^{0}$(μM)
1c	4-F	3.82 ± 0.03	2.2±0.1	148 ± 5
1h	2-F	3.75 ± 0.09	….	173 ± 4
1i	3-F	4.4 ± 0.1	6.5±0.2	36 ± 8
1l	2,6F	4.3 ± 0.2	…	50 ± 15
1m	2,4F	4.3 ± 0.15	….	50 ± 15

3.1.4 Lipophilicity

Lipophilicity of compounds **1h-o** was theoretically calculated with the VCClab software for the neutral form of each cage (table 3.2). Data introduction to the software was performed through the SMILES nomenclature. This software determines the lipophilicity using an statistical ensemble of associative neutral networks trained on a map set that is composed by different families of compounds but they are generally planar structures, which is quite far away from the 3D structure of the cage-like compounds studied in the present work. This is probably the reason why all the compounds with the same number of fluorine atoms resulted in the same theoretical log *P*. In order to stablish a proper comparison between the lipophilicity of compounds **1h-o,** their retention times in an HPLC analysis were determined (see experimental section for the full HLC chromatograms). Using the same gradient and column, compounds with a lower retention time are less lipophilic so they can be sorted from more to less lipophilic.

Compounds **1n** and **1o** are very insoluble in water and are the most lipophilic of the series. Their calculated log *P* are 5.9 and 5.51 respectively and the HPLC analysis also shows that the retention time of these two compounds is higher than the others of the family in more than two minutes: 13.07 and 12.44 min respectively. The poor solubility and high lipophilicity of **1n** and **1o** make them unappropriated for applications in aqueous media including biological assays.

Through HPLC analysis, we detected important differences in the lipophilicity of compounds with the same amount of F in different positions. For instance, compounds **1h** and **1l** which are the less lipophilic ones of the series have very similar lipophilicity (according to their retention time) although **1h** has 3 F-atoms compared to **1e** which has 6

F-atoms. The next ones in growing lipophilicity are compounds **1c** and **1l** also with different amount of F but sharing a fluorine in position 4 of the aromatic ring. The rest of compounds with two F substituents in the phenylalanine ring (**1k, 1j**) were surprisingly found to be more lipophilic than **1m** and **1l.**

The differences in lipophilicity between different compounds that arise only by changing the position of the fluorine in the aromatic ring are an indication that not only the fluorine but its position in the molecule is somehow affecting the 3D structure, its lipophilicity and its interaction properties with aqueous and organic solvents.

Table 3.3. Calculated Log *P* values and retention time (min). LogP corresponds to the non-protonated cage. All the products were dissolved in MeOH at concentration of 1 mM and eluted using a linear gradient of 20 % to 80 % of CH_3CN in H_2O (using 0.1% of TFA) in 20 minutes.in a reverse-phase Kromasil 100 C8 (15 x 0.46 cm, 5 µm) column with a 1 mL/min flow.

Compound	R	LogP neutral (calc)	HPLC retention time (min)
1c	4F-Phe	4.00	9.435
1h	2F-Phe	4.00	8.017
1i	3F-Phe	4.00	9.218
1j	(3,5-diF)Phe	4.38	10.640
1k	(3,4-diF)Phe	4.39	10.334
1l	(2,6-diF)Phe	4.37	8.28
1m	(2,4-diF)Phe	4.38	9.53
1n	Penta-F-Phe	5.51	13.070
1o	4-CF_3-Phe	5.90	12.440

3.1.5 Determination of the pKa of cages 1a 1c, 1h, 1i.

Log *P* and pK_a are two very important parameters in pharmacokinetics, especially in poorly soluble compounds. Besides, they are somehow related, since the ionization state of a molecule with several protonation states will be determined by its pK_a, and this will strongly influence its lipophilicity at a given pH. Drugs in their ionized states are more water soluble but less permeable to membranes (they are less hydrophobic) than their neutral form. Pseudopeptidic cages **1a-o** can have up to 5 different protonation states since they have 4 protonable amines. The protonation state of these molecules at the pH values with

physiological relevance might be strongly affecting all their properties: Cl^- recognition, lipophilicity, Cl^- transports, cell internalization and cytotoxicity. Hence we considered designing an experiment to determine the pK_a of these cages.

The most reliable technique for pK_a determinations is potentiometry but the poor water solubility of these compounds at basic pH made it impossible for us to use this technique.[3,4]

Alternately, we proposed to use fluorescence for this purpose. To do so, the fluorescence emission spectra of a solution (100-200 µM) of the cage in water: methanol (60 : 40) at different pH values was recorded. The excitation wavelength used was the maximum absorbance wavelength for phenylalanine (270 nm). We observed a change in the emission upon the change in the pH and this data was successfully fitted with HypSpec to the different acid-base equilibrium possible for the system:

$L + H \leftrightarrows LH$	pK_a 1
$LH + H \leftrightarrows LH_2$	pK_a 2
$LH_2 + H \leftrightarrows LH_3$	pK_a 3
$LH_3 + H \leftrightarrows LH_4$	pK_a 4

We used a mixture of water and methanol because the low solubility of these cages at basic pH made it impossible to use a 100 % aqueous system. There are some bibliographic examples of this technique to estimate the pK_a of organic compounds with poor solubility.[3–7] The drawback of this methodology is that the fitted pK_a might not be the exact one but pK_a between the 4 cages measured can well be compared since they were all determined under the same conditions.

We determined the pK_as for compounds **1h, 1i** and **1c** which correspond to the cages with a fluorine in the positions 2,3 and 4 of the aromatic ring of R, and compound **1a** which correspond to the phenylalanine cage without any fluorine to better understand the effect of both the presence and the position of F substituents. Compounds with more than one fluorine were not successfully fitted to estimate pK_a values because precipitation starts to occur at basic pH.

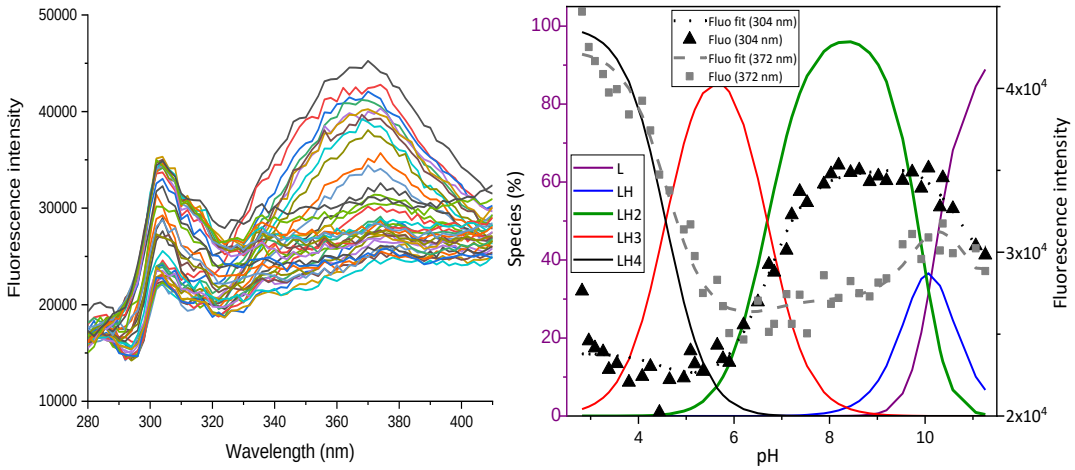


Figure 3.4. Fluorescence spectra of **1a** (Phe) at different pH, arrow indicates the increase of the fluorescence when pH increases (left). Data fitting with HypSpec including the species distribution at different pH (right).

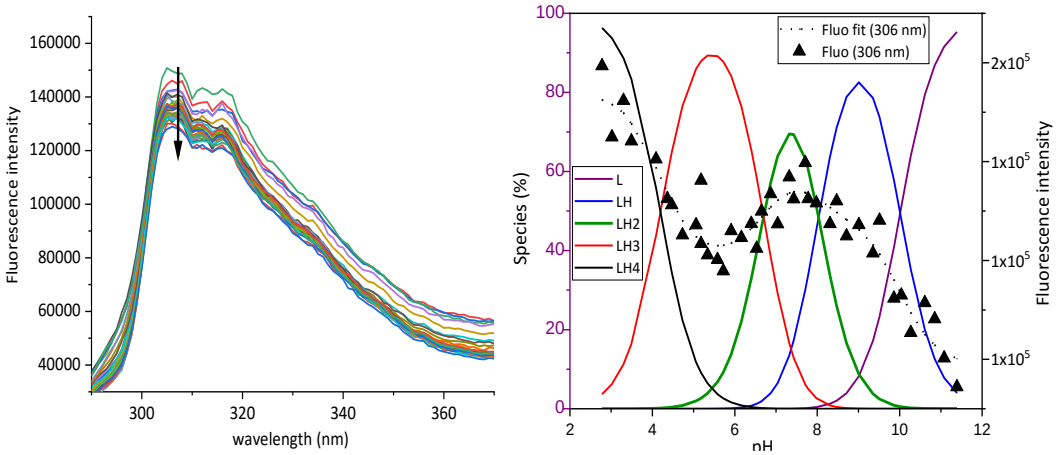


Figure 3.5. Fluorescence spectra of **1c** (4-F-Phe) at different pH, arrow indicates the increase of the fluorescence when pH increases (left). Data fitting with HypSpec including the species distribution at different pH (right).

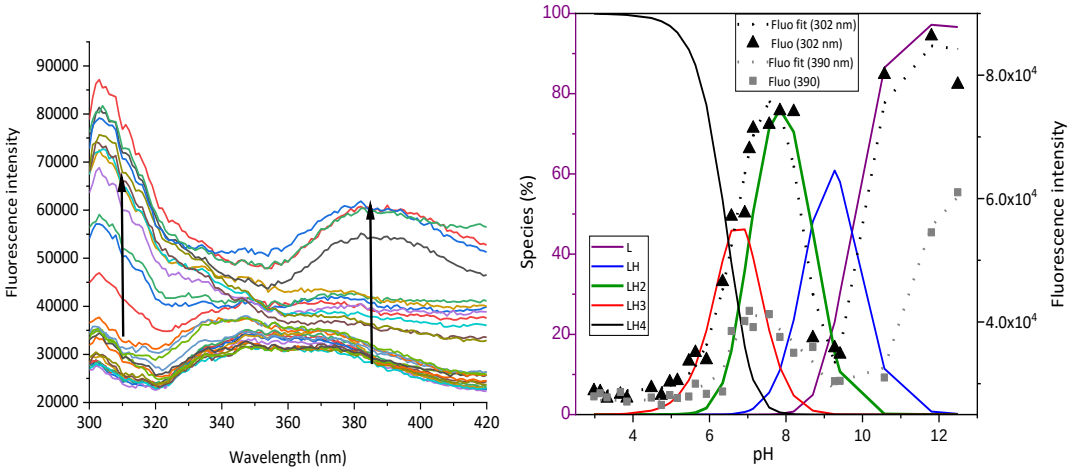


Figure 3.6. Fluorescence spectra of **1h** (2-F-Phe) at different pH, arrow indicates the increase of the fluorescence when pH increases (left). Data fitting with HypSpec including the species distribution at different pH (right).

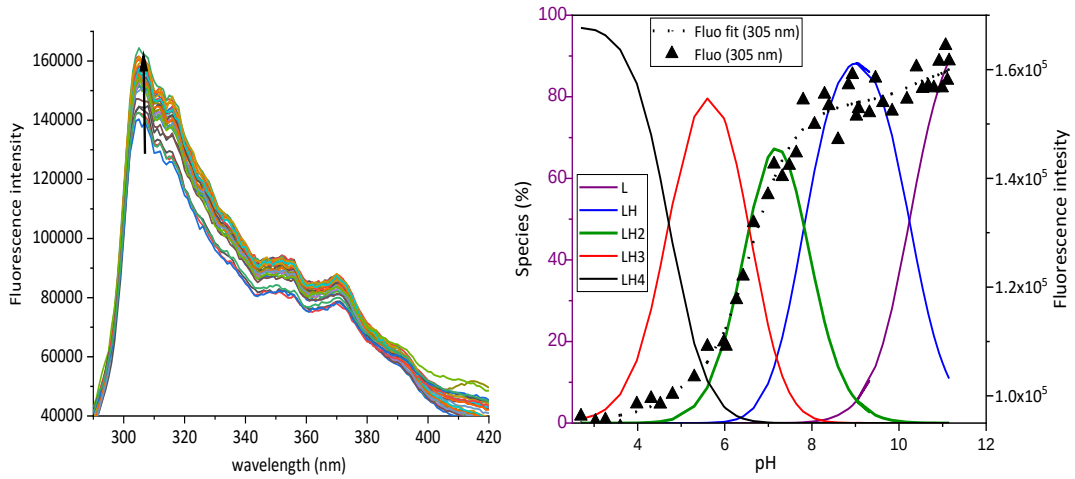


Figure 3.7. Fluorescence spectra of **1i** (3-F-Phe) at different pH, arrow indicates the increase of the fluorescence when pH increases (left). Data fitting with HypSpec including the species distribution at different pH (right).

Table 3.4. pK_a of compounds **1h, 1i, 1c** and **1a.**

R_1	Cage	pKa1	pKa2	pKa3	pKa4	Rt
2F-Phe	**1h**	9.69	8.67	7.03	6.48	8.01
3F-Phe	**1i**	10.23	7.83	6.55	4.72	9.21
4F-Phe	**1c**	10.01	8.03	6.7	4.21	9.40

Phe	**1a**	10.13	10.12	6.69	4.57	8.02

Compounds **1h** and **1c** have very similar pK_a values, hence their species distribution at the pH range 6.2-7.5 is very similar. At pH 6.2 almost half of the cage is triprotonated and half diprotonated. On the other hand, compound **1h** has a very different protonation profile. It is a more basic compound, pK_a2 and pK_a3 of **1h** are lower than those for the other two compounds. This means that the cage is more protonated at both pH 6.2 and 7.5 than compounds **1h** and **1c.** When comparing fluorinated cages with **1a** we see a clear effect of the presence of fluorine in pK_a2 which is 1,5 - 2 units higher than in fluorinated cages whereas pK_a1, pK_a3 and pK_a4 are similar to **1i** and **1c**. As a result **1a** is more basic than their fluorinated analogues which makes the cage comparatively more protonated at basic pH. The differences in the pK_a are clearly reflected in the species distribution.

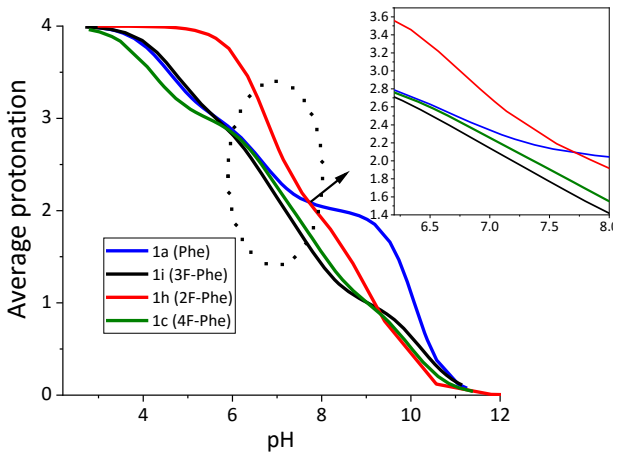


Figure 3.8. Average protonation of compounds **1a**, **1h**, **1i** and **1c** at different pH.

Through the study of the pK_a of these compounds we have seen that the presence of fluorine and its position in the aromatic ring have a strong influence in the acidity/basicity of the amines of the core in these cages which will deeply determine their physicochemical and biological properties.

3.1.6 Chloride transport in artificial vesicles

The ability of these compounds to act as anion transmembrane carriers was tested with the HPTS fluorescence assay (Figure 3.10) using POPC vesicles. The conditions used for the HPTS assay were modified from the ones reported in chapter 2. These measurements were

conducted by the group of Prof. Roberto Quesada at Burgos University and they optimized
the experimental conditions to improve the reproducibility of the assay.

The calibration curve for the new HPTS assay conditions was fitted to an S-logistic function with Origin software (Figure 3.9). The fitting equation was then used in the assays to convert fluorescence values into pH.

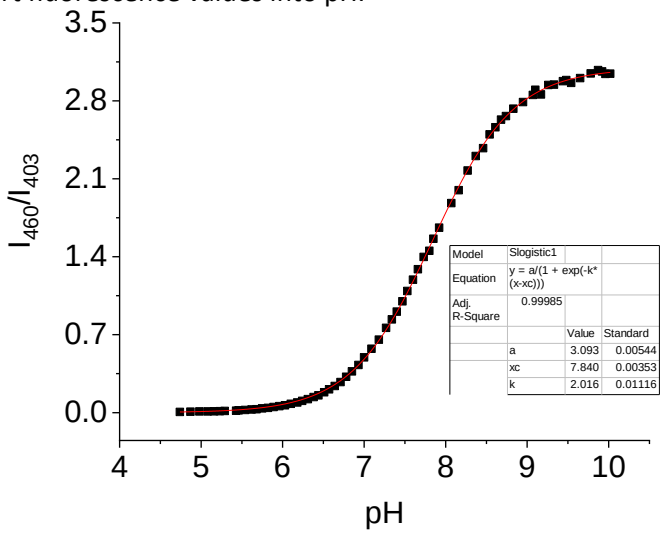


Figure 3.9. HPTS calibration s-logistic1 model (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM phosphate buffer).

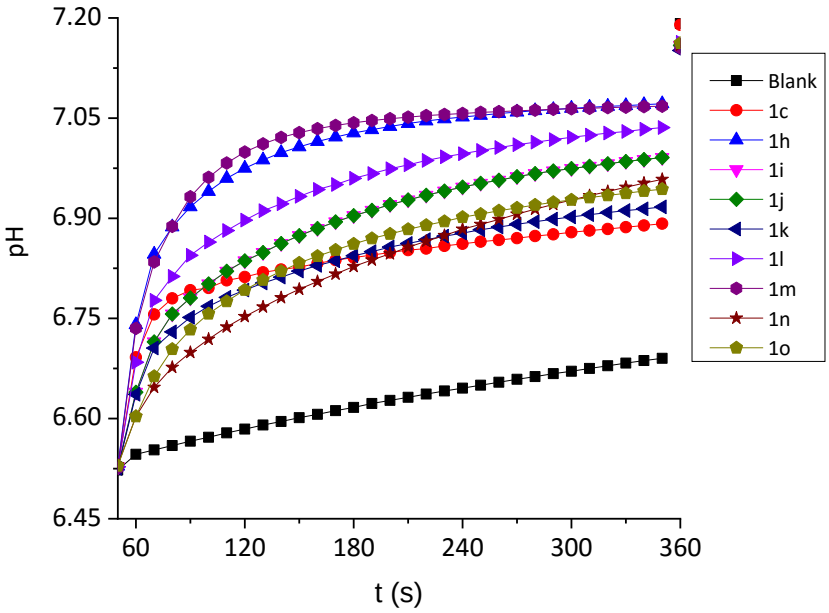


Figure 3.10. Variation of pH upon addition of compounds **1c,1h-o** (10 mol %) to POPC : cholesterol (7 : 3) vesicles, 0.5 mM POPC. Vesicles contained NaCl (10 mM NaCl, 123.9 mM $NaNO_3$, 10 mM buffer phosphate pH 6.5, I.S. 150 mM and HPTS 1 mM) were suspended in the same buffer. At t = 60 s addition of the anion carrier (blank: 12.5 µL DMSO). Each trace represents the average of at least three different trials from at least two different batches of vesicles.

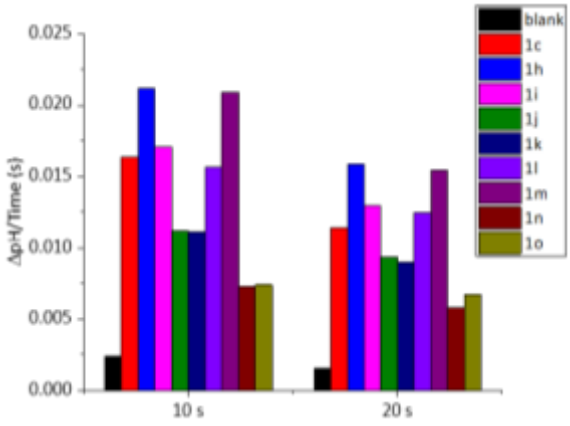


Figure 3.11. Initial rate of the pH variation induced by compounds **1c, h-o** in POPC vesicles measured with the HPTS assay in the first 10 and 20 seconds.

The trend observed in the HCl transport assays was: 1h (2-F) ≥ 1m (2,4-diF) > 1i (3-F) ≥ 1c (4-F) ≥ 1l (2,6-diF) > 1j (3,5-diF) ≈ 1k (3,4-diF) > 1o (4-CF3) ≥ 1n (penta-F).

Compounds containing one or two F-atoms in the aromatic ring with one F-atom in position 2 are better transporters than their analogues with fluorine in any other position. Compounds 1j (3.5-diF), 1k (3.4-diF), 1n (pentaF) and 1o (4-CF_3)are the weakest transporters probably due to their high lipophilicity that prevents them to interact with the aqueous phase.

For the different compounds with fluorine we found a correlation between the transport capacity determined with the HPTS assay and the lipophilicity measured by the reverse phase HPLC retention time. An increase of the lipophilicity reduces the transport activity (Figure 3.12).

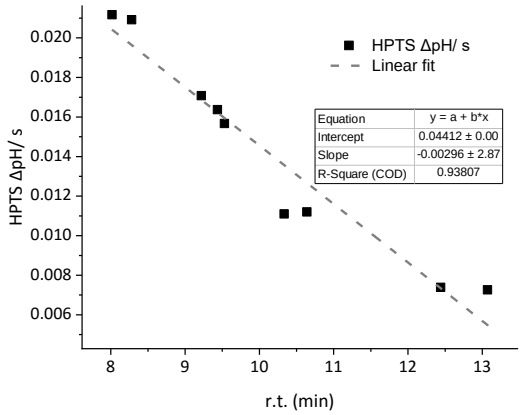


Figure 3.12. Plot of the transport capacity of compounds 1c-o vs their experimental lipophilicity (retention time). Red line corresponds to the linear fitting of the data.

The carboxy-fluorescein (CF) assay with compounds **1c,h-o** shows a fluorescence enhancement only when the detergent is added (fig. 3.13). This indicates that vesicles are stable in the presence of these compounds and that they do not form large pores in the vesicles or act as detergents. This result further supports the ability of these compounds to discharge pH gradients as a result of the ion transport activities.

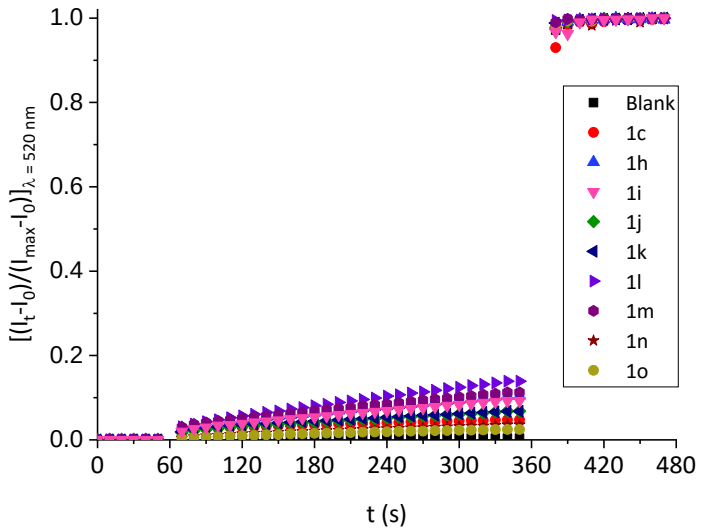
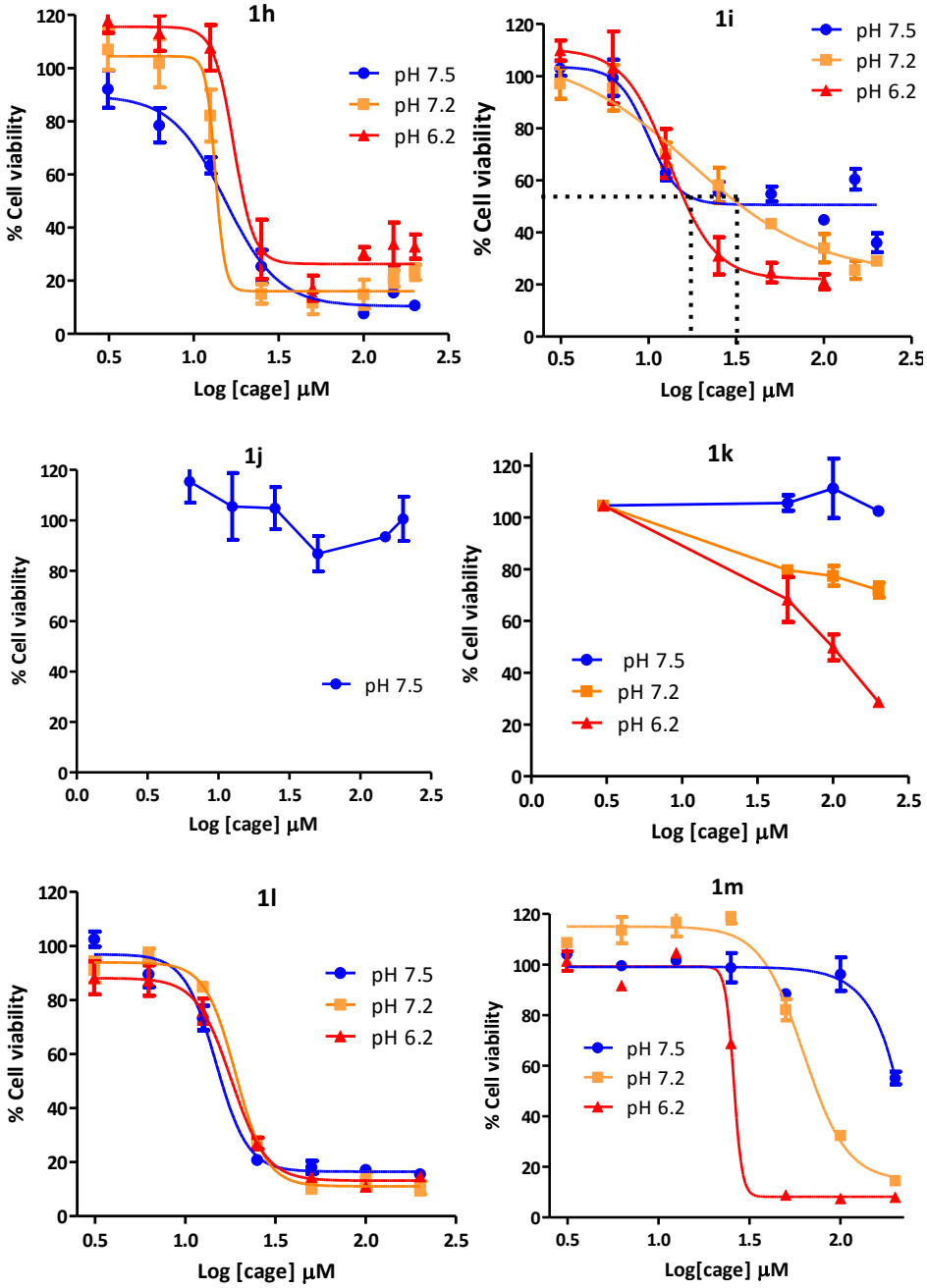


Figure 3.13. Carboxy-fluorescein leakage observed upon addition of **1c, h-o** compounds (10% mol carrier to lipid concentration) or the blank (DMSO) to POPC vesicles (0.05 mM). Vesicles, loaded with a $NaNO_3$ and NaCl buffered aqueous solution (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM NaH_2PO_4, 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a $NaNO_3$ and NaCl buffered aqueous solution (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM NaH_2PO_4, I.S. 150 mM, pH 7.2). At t = 60 s the compound (or the blank, DMSO, 1.25 µL) was added, and at t = 360 s a detergent (Triton-X, 10 % dispersion in water, 20 mL) was added. Each trace corresponds to the average of six different trials performed with three batches of vesicles.

3.1.7 Cellular assays

Cytotoxicity of the new derivatives towards A549 cells at pH 6.2, 7.2 and 7.5 was studied as described for compounds **1a-g** in chapter 2 (24 hour incubation time and MTT assay). Dose-response curves were constructed to obtain IC_{50} values using Graphpad dose-response fitting function for each trace.


1h
% Cell viability
Log [cage] µM
pH 7.5
pH 7.2
pH 6.2
1i
% Cell viability
Log [cage] µM
pH 7.5
pH 7.2
pH 6.2
1j
% Cell viability
Log [cage] µM
pH 7.5
1k
% Cell viability
Log [cage] µM
pH 7.5
pH 7.2
pH 6.2
1l
% Cell viability
Log [cage] µM
pH 7.5
pH 7.2
pH 6.2
1m
% Cell viability
Log[cage] µM
pH 7.5
pH 7.2
pH 6.2

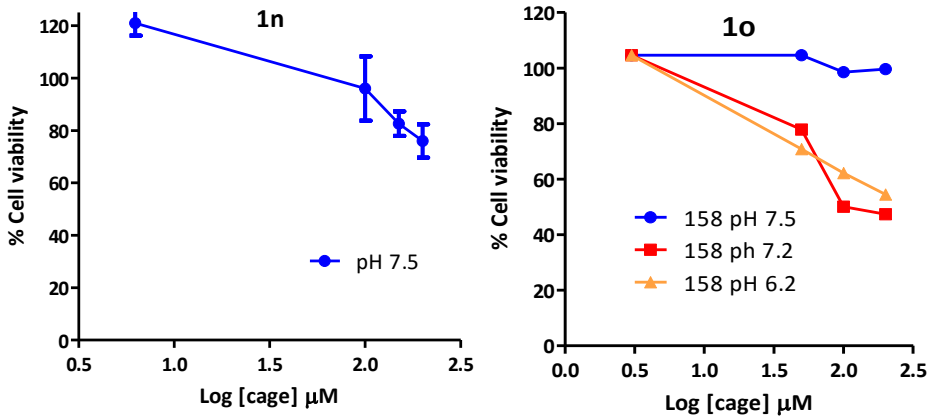


Figure 3.14. Cytotoxicity of compounds **1h-o** in A549 cells at pH 7.5,7.2 and 6.2 after 24 hour exposure measured with MTT and dose response calculated curves. Each trace corresponds to an average between at least three independent experiments.

Table 3.5. IC_{50} of compounds **1c,1h-o** at pH 7.5,7.2 and 6.2 after 24 hour exposure measured with MTT.

pH\Cage	**1c** 4-F	**1h** 2-F	**1i** 3-F	**1j** 3,5-F	**1k** 3,4-F	**1l** 2,6-F	**1m** 2,4-F	**1n** F_5	**1o** CF_3
6.2	29±4	14±2	13±3	n.d.	~100	17±2	25±8	n.d.	≥200
7.2	58±10	17±3	~25	n.d.	>200	18±2	62±10	n.d.	≥200
7.5	166±35	13±2	~25	>200	>200	14±2	>200	>200	>200

Compound **1c** described in chapter 2 (4F-Phe substituent in R_1) was used as the reference compound for its interesting biological properties. When comparing cages with one single fluorine in each R substituent, we observe that compound **1h** is the most cytotoxic and its activity does not change in the studied pH range compared to **1c** which is the least cytotoxic at neutral pH but its toxicity is highly increased at pH 6.2 almost matching **1h.**

Compound **1i** has a behaviour between **1h** and **1c**. It is not as cytotoxic as **1h** and the cytotoxicity is less pH dependent than in **1c**. Nevertheless, solubility issues might be preventing cytotoxicity to reach 100 % levels even at high concentration of **1i** at neutral pH, which has been found to be more lipophilic through the HPLC analysis. This trend is present in all the cages with a F-atom in position 3. Proper mathematical fitting of the data to estimate the IC_{50} values cannot be conducted with the data obtained but graphical estimation of it was possible by simply finding the concentration of the cage at a 50 % of

survival. At the 3 different pH the IC_{50} is similar, only slightly lower at pH 6.2. However the graph shape is considerably different depending on the pH. The MTT assays are maybe mediated by the lower aqueous solubility and higher hydrophobicity of this compound.

Compound **1l** has similar cytotoxicity to **1h** at the different pH's tested whereas for compound **1m** cytotoxicity resembles the one of **1c**. **1m** is even more selective for acidic pH than **1c**. Apparently, the introduction of a fluorine in the position 2 of the phenyl substituent enhances the cytotoxicity and the fluorine in position 4 is the responsible of the pH dependency. Very remarkably, the effects are additive, which makes compound **1m**, combining the substitution at both positions, specially interesting. It must be noted that a high toxicity change is observed in the range 6.2-7.5, conditions that can resemble those surrounding either a solid tumour or a healthy tissue, respectively.

Whereas for Compounds **1k, 1j, 1n** and **1o**, their higher lipophilicity suggests that the linear increase of the cytotoxicity especially at low pH is an effect of the solubility. Those compounds have a very low water solubility that can increase at acidic pH since they get protonated. None of these compounds have potentially interesting cytotoxic properties (low to non-detected cytotoxicity towards A549 cells).

With the given results we thought that the cytotoxicity shift for compounds **1h** and **1m** might be at a higher pH, since the pK_a measurements showed an increase of basicity produced by the 3-F-substitution. Accordingly, the IC_{50} of these two compounds was also measured at pH 8.0 and 8.5. The results show exactly what we predicted. At pH above 8.0 the cytotoxicity of compounds **1h** and **1m** was drastically reduced (Figure 3.15). The apparent increase in the cell survival can be associated with mitochondrial stress prior to cell death.

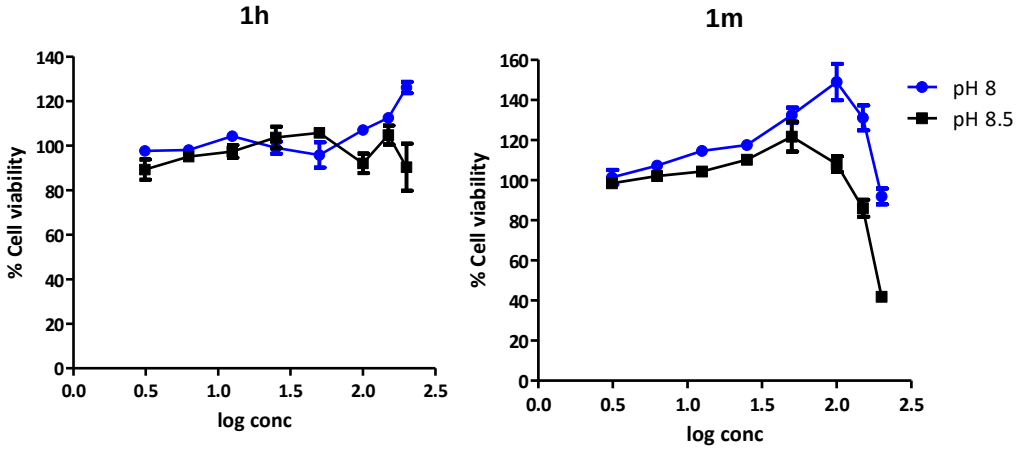


Figure 3.15. Cytotoxicity of compounds **1h** and **1m** towards A459 cells at pH 8.0 and 8.5 after 24 hours exposure measured with MTT. Each trace corresponds to an average between at least three independent experiments.

Cytotoxicity in alginate spheroids

We assessed the cytotoxicity of compounds **1c** and **1h** in a 3D cell culture model using alginate spheroids loaded with A549 cells.[8,9] The incubation time of the A549 in alginate spheroids with the cage compounds was 48 h since almost no cytotoxicity was detected at 24h probably due to diffusion events through the alginate gel. The cytotoxicity of compounds **1h** to A549 in spheroids was equivalent to the cytotoxicity measured in a 2D culture. For compounds **1c**, a higher cytotoxicity at neutral pH was also detected in spheroids compared to the 2D cell culture (Figure 3.16 and table 3.6). The lower pH differences between the 2D and 3D models were associated to the longer time exposures and to the possible interactions of the cage with alginate detected through fluorescence spectroscopy. Besides, some buffering effects of the alginate gel micro-environment network could be playing a role at slightly acidic pH. Overall we have successfully recreated the cytotoxicity and the pH dependence of these two compounds in a simple 2D culture into a more complex three-dimensional system, which is more similar to a living tissue. These preliminary results strengthen the potential of the use of pH sensitive anionophores as anticancer agents.

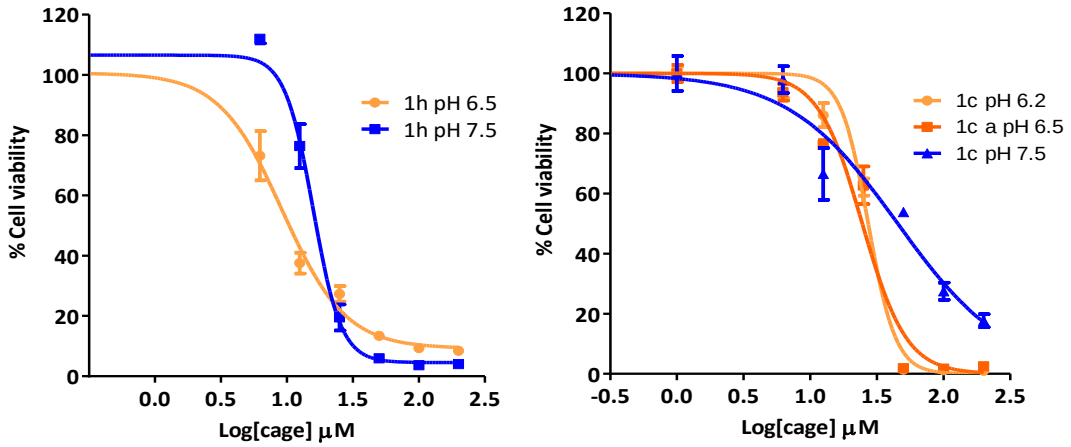


Figure 3.16. Cytotoxicity of compounds **1h** (2F-Phe) and **1c** (4F-Phe) at different pH in A549 cells present in alginate beads after 48 hour exposure measured with resazurin and dose response calculated curves. Each trace corresponds to an average between at least three independent experiments.

Table 3.6. IC_{50} of compounds **1h** (2F-Phe) and **1c** (4F-Phe) to A549 cells in alginate beads at pH 7.5, 6.5 and 6.2 after 48 hour exposure measured with resazurin.

pH\Cage	1c	1h
6.2	27±4	-
6.5	25±3	9±2
7.5	45±10	15±2

3.1.8 Molecular modelling of 1h

Molecular modelling studies of compound **1h** (2F-Phe) were conducted in collaboration with Dr. Jordi Bujons. The modelling of the structure of monoprotonated **1h** (protonated in the tertiary amine yielding a totally symmetric structure) predict the possible formation of an hydrogen bond between the fluoride and the secondary amine of the nearby arm of the cage (Figure 3.17). Moreover, this interaction will favour the interaction between H3 of the aromatic ring and the oxygen of the amide of the nearby branch of the cage forming a weak non-conventional C-H···O H-bond.[10,11] These two interactions are only possible in the cage with fluorine in position 2, and they can explain the different acidic behaviour and lower lipophilicity of cages **1h** and **1m**. . A hydrogen bond with fluorine in a protonated amine can have an stabilizing effect of the positive charge thus reducing the acidity of the amine. Hence the protonation state of the molecule at neutral pH is higher than their analogues without this F-H interactions.

The models obtained can also explain the different basicity between cage **1a** (Phe) and cages containing fluorine in different positions. The first pK_a being of the same order in cage **1a** (Phe) and cages with F substituents agrees with the first protonation to occur in the tertiary amine, the less affected by the side chains since they are very separated in their 3D structure. The second protonation is strongly affected by the presence of a F which can destabilize the charge in the corresponding ammonium due to an electron-withdrawing inductive effect of F, hence making it more acidic as we observed when comparing **1a** with all the fluorinated cages.

Moreover, in cage **1h** the intramolecular H-F bonds can be either one two or three of their R substituents. The overlap of the 3D structure of **1h** with 1, 2 or 3 F-H hydrogen bonds is represented in fFigure 3.18. The presence of these interactions yields a more compact structure with less exposure of the lipophilic side chains (R_1) that can be responsible of the lower lipophilicity of cages **1h** and **1m**.

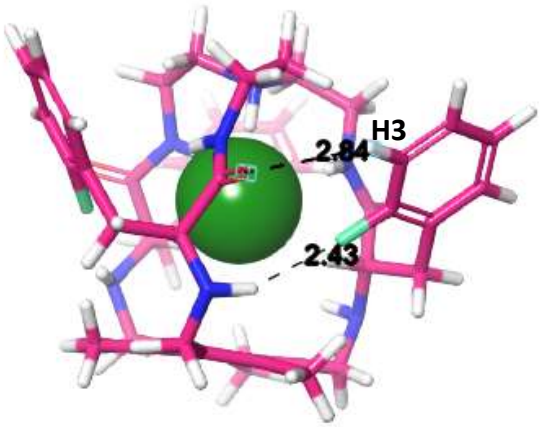


Figure 3.17. Proposed structure of cage **1h** (2F-Phe) with a hydrogen bond between an amine and the F of the R substituent.

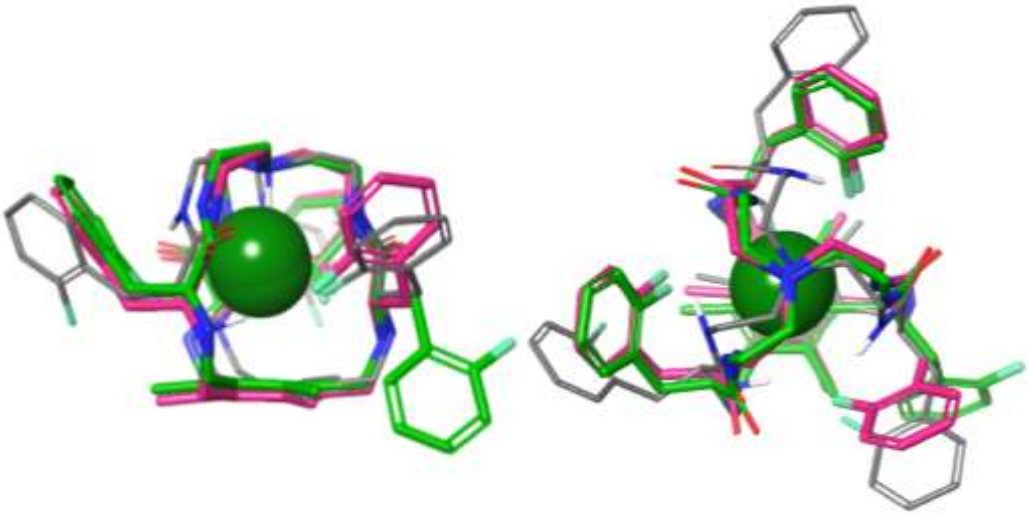


Figure 3.18. Overlay of the computational model of monoportonated **1h** forcing the formation of 1 (grey) 2 (green) or 3 (pink) hydrogen bonds between Fluor and the amine.

3.2 DISCUSSION

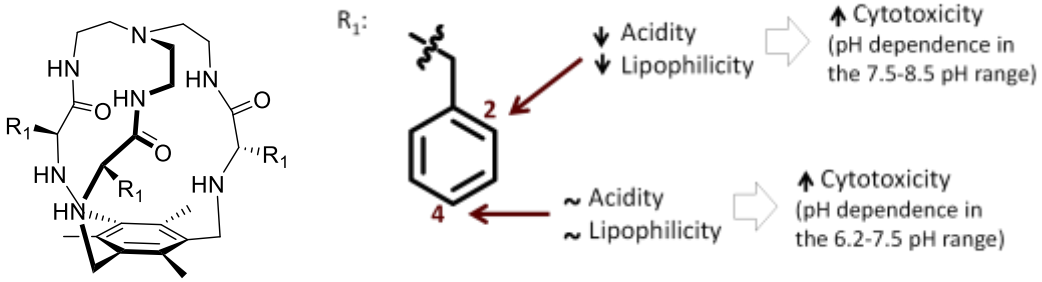


Figure 3.19. Schematic representation of the most relevant results from chapter 3.

A new family of cage-like compounds containing different fluorine in the R_1 position was successfully synthetized and studied. They have interesting properties which vary depending on the R_1 substituent. From this new family the most interesting compounds are **1h** and **1l** with higher cytotoxicity than the previously reported **1c** but with pH dependency in a more basic range (7.5-8.5) and compounds **1m** with higher pH dependent cytotoxicity than **1c** in the range of 6.2-7.5.

Several physicochemical properties of compounds **1h-o** have been studied aiming to establish a connection between the physico-chemical properties of these cages and their cytotoxicity. In particular we have been deeply focused on finding a structural explanation for the different pH dependent cytotoxicity behaviour of cages containing F at different positions of the aromatic substituent R_1. Experimental data indicates that differences in the association constants of the different cages with Cl^- are not relevant enough to be responsible of their different biological behaviours, moreover they are determined in conditions that differ substantially from the conditions encountered in a biological media (solvent and protonation state). Lipophilicity has proven to be a key parameter in this family of compounds. The presence of 2 or more F-atoms in the R substituent (meaning that the cage will have 6 or more F) increases the lipophilicity of the cage. This was associated with the lack of cytotoxicity of compounds **1j**, **1k**, **1n** (all of them containing at least one F-atom in position 3) and **1o**. These cages are too lipophilic to interact with the aqueous phase once they are incorporated into the lipid bilayer, hence they are unable to transport HCl or to further penetrate inside the cell to induce cytotoxicity. Moreover compounds **1n** and **1o** have very low water solubility at physiological pH which complicates the biological assays. A good correlation between lipophilicity and transport was found for compounds **1c,h-o**. This is in agreement with the observations previously made by other researchers.[12,13] For compounds **1c, h-o** we clearly observe that the increase in the lipophilicity reduces the transport capacity in POPC vesicles determined with the HPTS assay. From results of chapter 1 with compounds **1a-g** we would expect the inflection of the lipophilicity vs transport curve to be around a retention time of 8 min since for instance compound **1a** is already worse than **1b**, and **1e** and compound **1d** showed no transport activity. Best transporters of this series are compounds with one fluorine (**1h**, **1i** and **1c**) and compounds **1l** and **1m** which have two fluorine atoms but one or both are in position 2.

Table 3.6. Average protonation state of compounds **1h**, **1i** and **1c** at different pH.

	pH 6,2	**pH 7,2**	**pH 7,6**	**pH 8**
1h	3.64	2.52	2.18	1.88
1i	2.59	2.02	1.65	
1c	2.77	2.21	1.74	
1a	2.78	2.24	2.14	

The pH modulated cytotoxicity of compounds **1h**, **1i**, **1c**, **1l** and **1m** can be explained with the different pK_a of these molecules. pK_a measurements indicate that cage **1h** is more basic than their analogues **1i** and **1c**. If we determine the average protonation of cages **1h**, **1i** and **1c** at pH 6.2, 7.2, 7.6 and 8.0 (table 3.6) we see important differences. The average protonation states of **1h** at pH 7.6 is 0.5 units higher than **1i** and **1c** which might account for the different cytotoxicity at this pH. There seems to be a threshold at the average protonation degree of 2.0-2-2. Cages start to be cytotoxic when their average protonation reaches this value. This is in agreement with the pH responsive observed for compounds **1c** between pH 6.2 and 7.5 and for compounds **1h** between pH 7.5 and 8.5. The same behaviour would be expected for compounds **1l** and **1m** although their pK_a were not successfully determined due to their lower solubility at basic pH. One would expect the pK_a of **1l** to be very similar to **1h** and **1c** very similar to **1m** since they share the F substituents in the relevant positions 2 and 4 and their cytotoxicity is equivalent.

Finally a structural proposal has been made to explain the particular properties of compounds with fluorine in position 2. The close special disposition of fluorine and the amine of the core in the cage allows the formation of hydrogen bonds between the fluorine and the amine. These presumably weak interactions can occur in the 3 substituents of the molecule (C3 symmetry) making them statistically more relevant. These interactions will then be responsible of a lower size of the molecule (property observed in NMR diffusion experiments), the lower lipophilicity experimentally determined (the more compact state of the hydrophobic side chains reduces their lipophilic interactions) and their higher basicity.

3.3 CONCLUSIONS

- A new family of pseudopeptidic cages containing fluorine atoms in the extrenal side chains has been successfully synthetized and studied.
- Cages **1h** (derived from 2-F-Phe) and **1l** (from 2,6-diF-Phe) are the most cytotoxic at slightly acidic and neutral pH and their cytotoxicity decreases when pH is raised to 8.
- Cages **1c** (from 4-F-Phe) and **1m** (from 2,4-diF-Phe) are cytotoxic at slightly acidic pH (6.2) and their cytotoxicity decreases at physiological pH (7.5).
- The increase of the basicity of cage **1h** with fluorine at position 2 can explain both the lipophilicity and the cytotoxicity differences between cages.
- Cages with higher cytotoxicity also present higher transport activity in artificial POPC vesicles.
- Cages with 2 or more fluorine atoms (except with one or the two F at position 2) are very lipophilic hence they are bad transporters and non-cytotoxic.

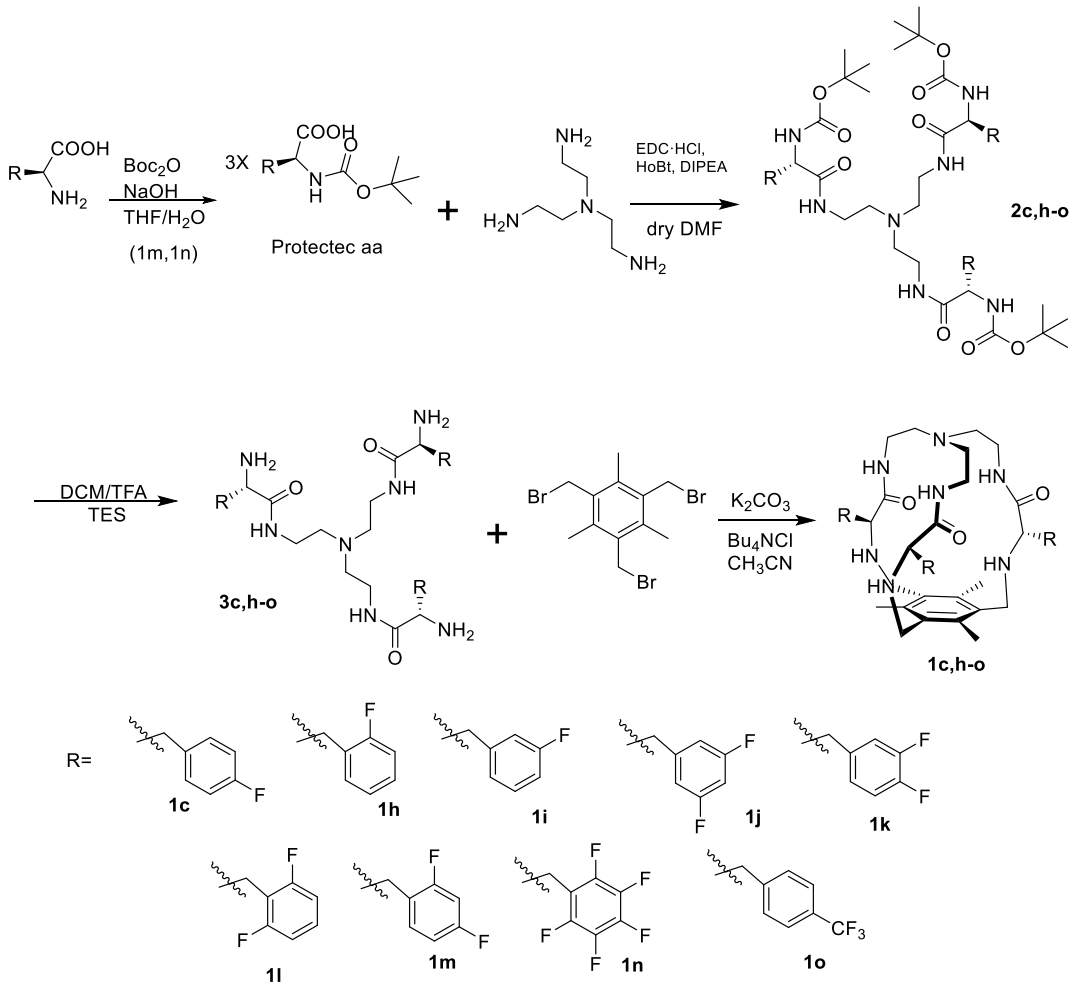
3.4 EXPERIMENTAL SECTION

3.4.1 Materials and methods

Materials and methods used in the different sections of this chapters were the ones reported in the experimental section of chapter 1 unless otherwise specified.

Synthesis of compounds 1h-o

The synthetic procedure followed is described in Scheme 3.1. Compounds **1m** and **1n** required an additional step (protection of the amine residue of the amino acid) whereas for the rest of the compounds the synthesis was started with the N-Boc protected amino acid.



Scheme 3.1. Synthetic scheme of compounds **1c, 1h-o.**

3.4.2 Synthesis of 1h

2h. tri-tert-butyl (((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(3-(2-fluorophenyl)-1-oxopro-pane-1,2-diyl))tricarbamate

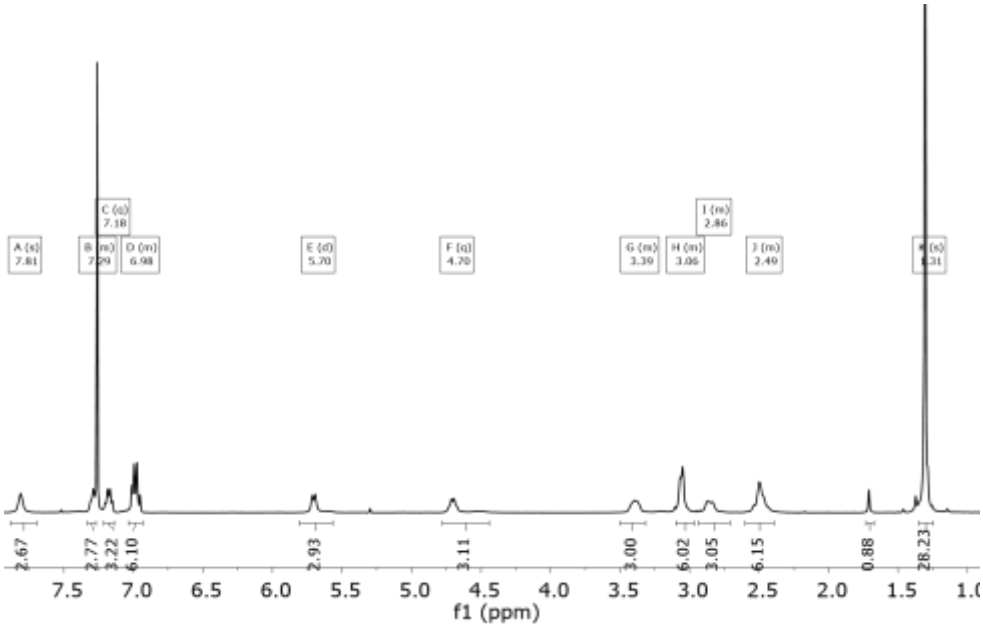
Boc-Phe(2-F)-OH (419 mg, 1.48 mmol) was dissolved in dry DMF (2.5 mL). *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC·HCl O.344 mg, 1.79 mmol) and 1-Hydroxybenzotriazole hydrate (HOBt, 0.274 mg, 1.79 mmol), *N*,*N*-diisopropylethylamine (DIPEA, 0.93 mL, 5.38 mmol) and tris(2-aminoethyl)amine (0.07 mL, 0.44 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5 % w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using 95:5 DCM : MeOH to give 0.330 mg of **2h** (0.351 mmol, 78 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81 (s, 3H), 7.29 (m, 3H), 7.18 (m, 3H), 6.98 (m, 6H), 5.70 (d, *J* = 8.8 Hz, 3H), 4.703 (m, 3H), 3.39 (m, 3H), 3.06 (m,6H), 2.86 (m, 3H), 2.49 (m, 6H), 1.306 (s, 27H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 172.8, 162.8, 160.4, 156.2, 132.0, 128.3, 123.9, 115.4, 115.1 , 79.9, 55.2, 54.8, 39.0, 32.3, 28.4.

^{19}F NMR (376 MHz, CD_3OD) δ -117.23.

HRMS (ESI-TOF) m/z [**2h** + H]$^{+}$ Calc: 942.4947, found: 942.5418



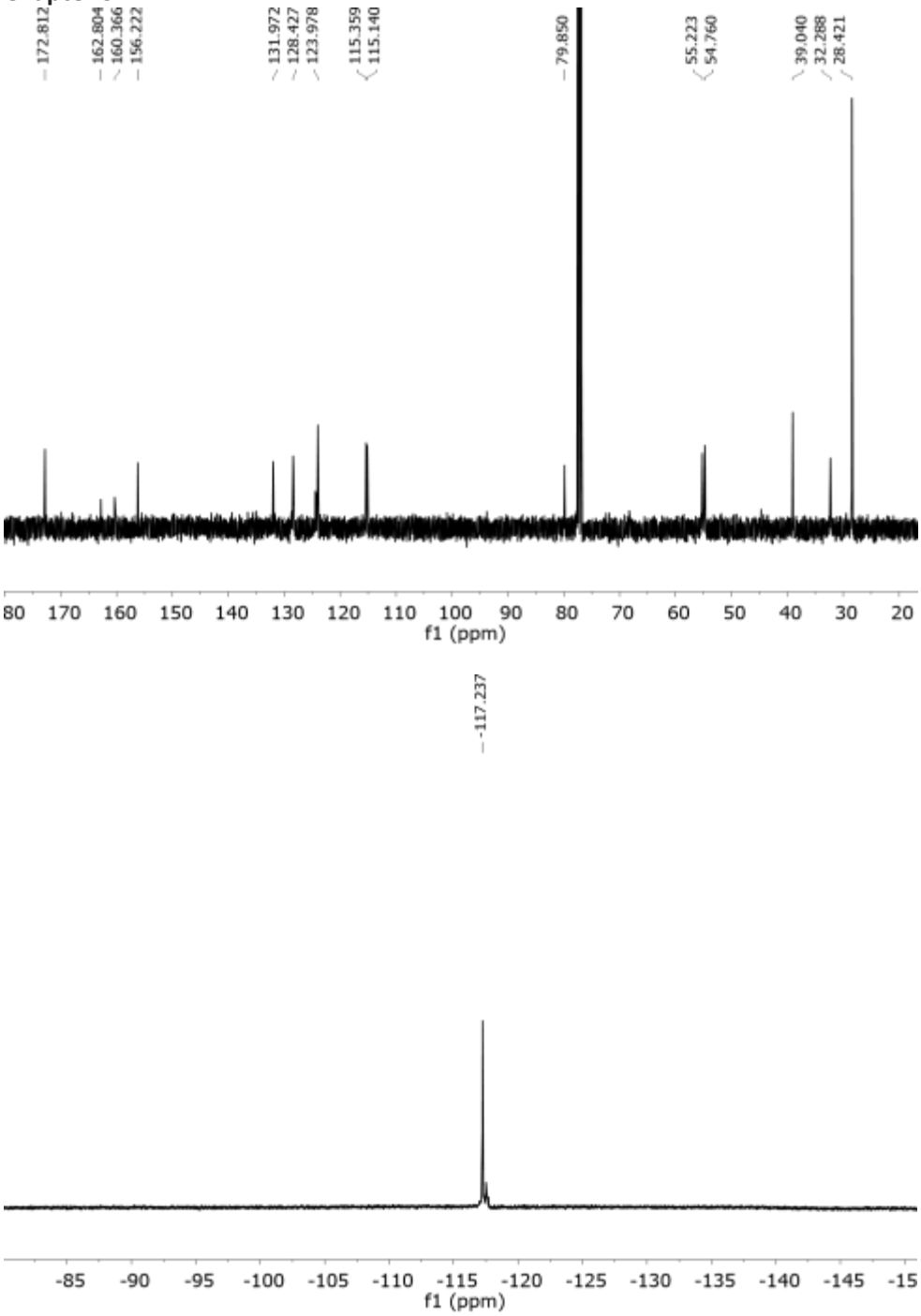


Figure 3.20. ^{1}H-NMR (400 MHz, CD_3Cl), ^{13}C-NMR (101 MHz, CD_3Cl) and ^{19}F NMR (376 MHz, CD_3OD) spectra of **2h**.

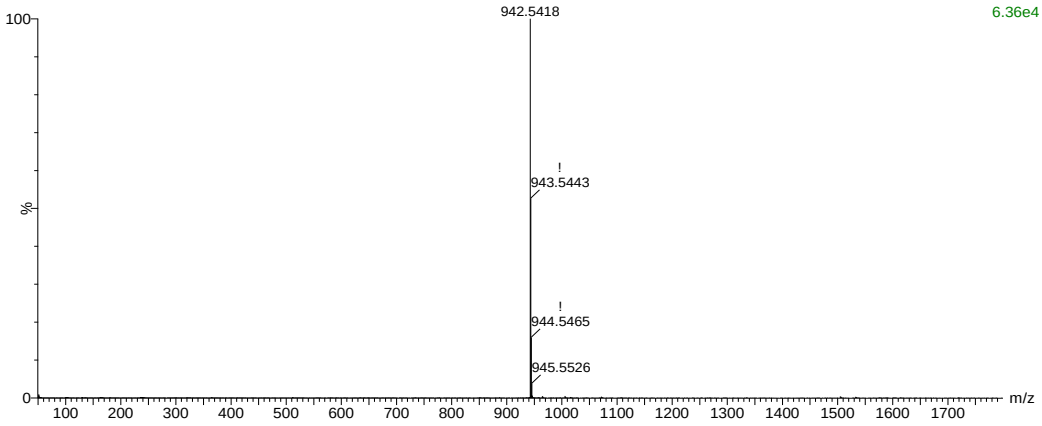


Figure 3.21. HRMS (ESI+) experimental spectrum of **2h**.

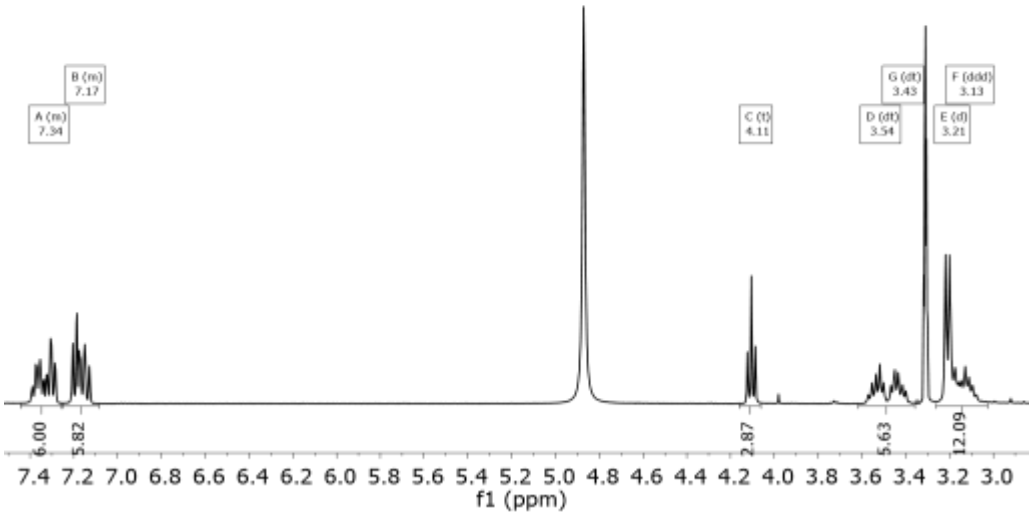
3h (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(2-fluorophenyl) propan amide)

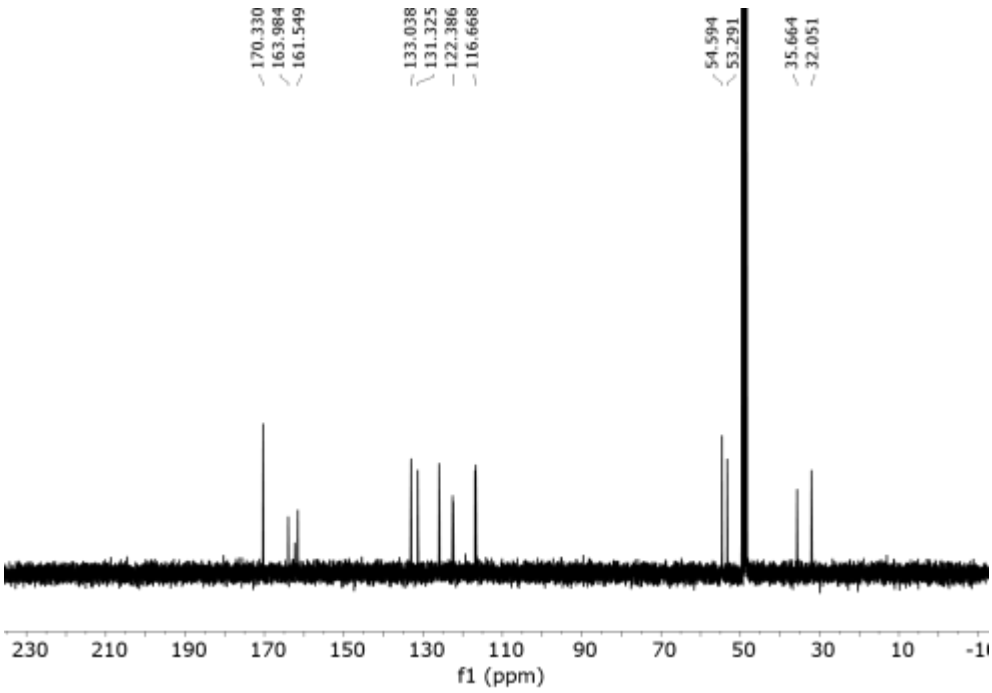
2h (280 mg, 0.287 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.65 mL, 4.31 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3h·4TFA** as a white solid (279 mg, 0.250 mmol, 89 % yield).

^{1}H NMR (400 MHz, CD_3OD) δ 7.34 (m, 6H), 7.17 (m, 6H), 4.11 (X subsystem form ABX, J_{AX} = 7.0, J_{BX} = 6.5 Hz, 3H), 3.54 (A subsystem form ABX, J_{AX} = 7, J_{AB} = 14 Hz, 3H), 3.43 (B subsystem form ABX, J_{BX} = 6.5, J_{AB} = 14 Hz, 3H), 3.21 (d, J = 7.3 Hz, 6H), 3.13 (m, 6H).

^{13}C NMR (101 MHz, CD_3OD) δ 170.3, 162.8, 133.0, 131.3, 125.9, 122.5, 116.8, 54.6, 51.5, 35.7, 32.1.
^{19}F NMR (376 MHz, CD_3OD) δ -118.73.

HRMS (ESI-TOF) m/z [**3h** + H]$^+$ Calc: 642.3374, found: 642.3651


A (m)
7.34
B (m)
7.17
C (t)
4.11
D (dt)
3.54
G (dt)
3.43
E (d)
3.21
F (ddd)
3.13
6.00
5.82
2.87
5.63
12.09
7.4 7.2 7.0 6.8 6.6 6.4 6.2 6.0 5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4 3.2 3.0
f1 (ppm)


170.330
163.984
161.549
133.038
131.325
122.386
116.668
54.594
53.291
35.664
32.051
230 210 190 170 150 130 110 90 70 50 30 10
f1 (ppm)

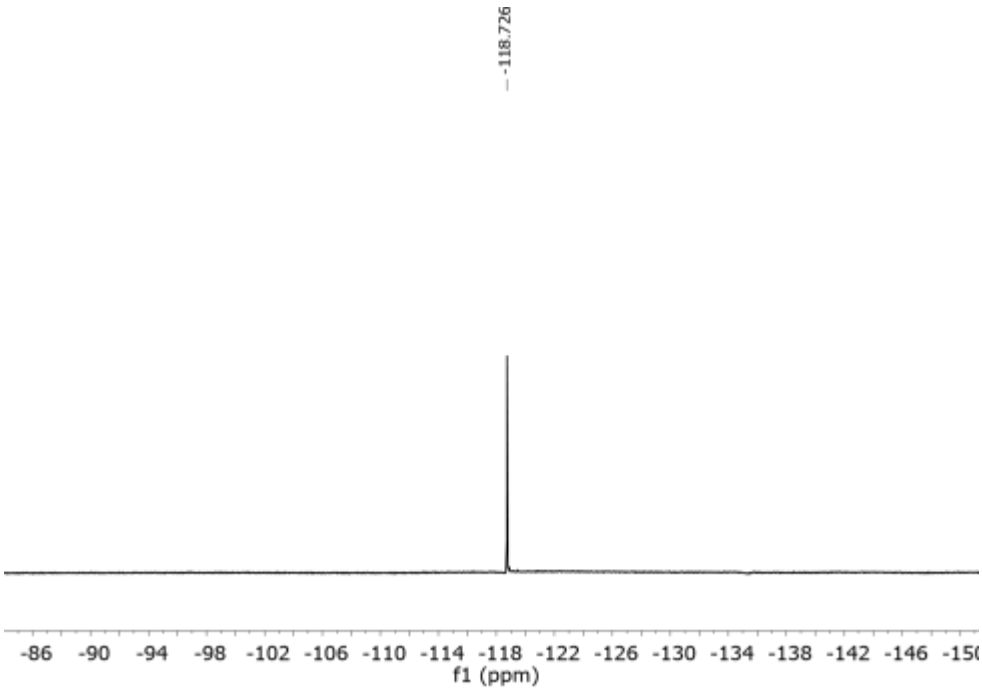


Figure 3.22. ^{1}H-NMR (400 MHz, CD_3OD), ^{13}C-NMR (101 MHz, CD_3OD) and ^{19}F NMR (376 MHz, CD3OD) spectra of **3h.**

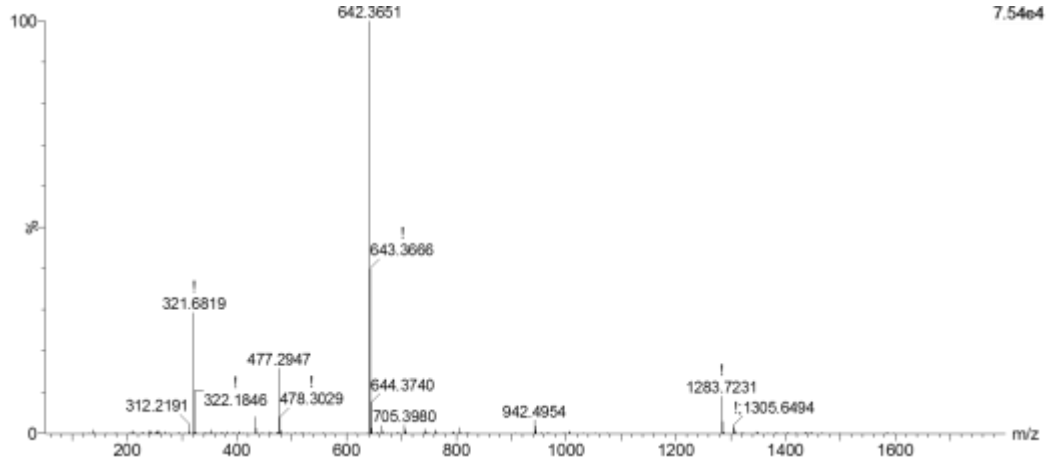
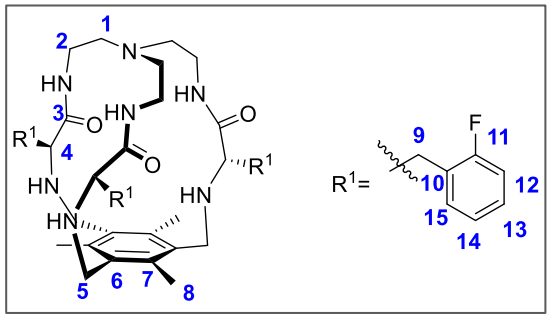


Figure 3.23. HRMS (ESI+) experimental spectrum of **3h.**

1h

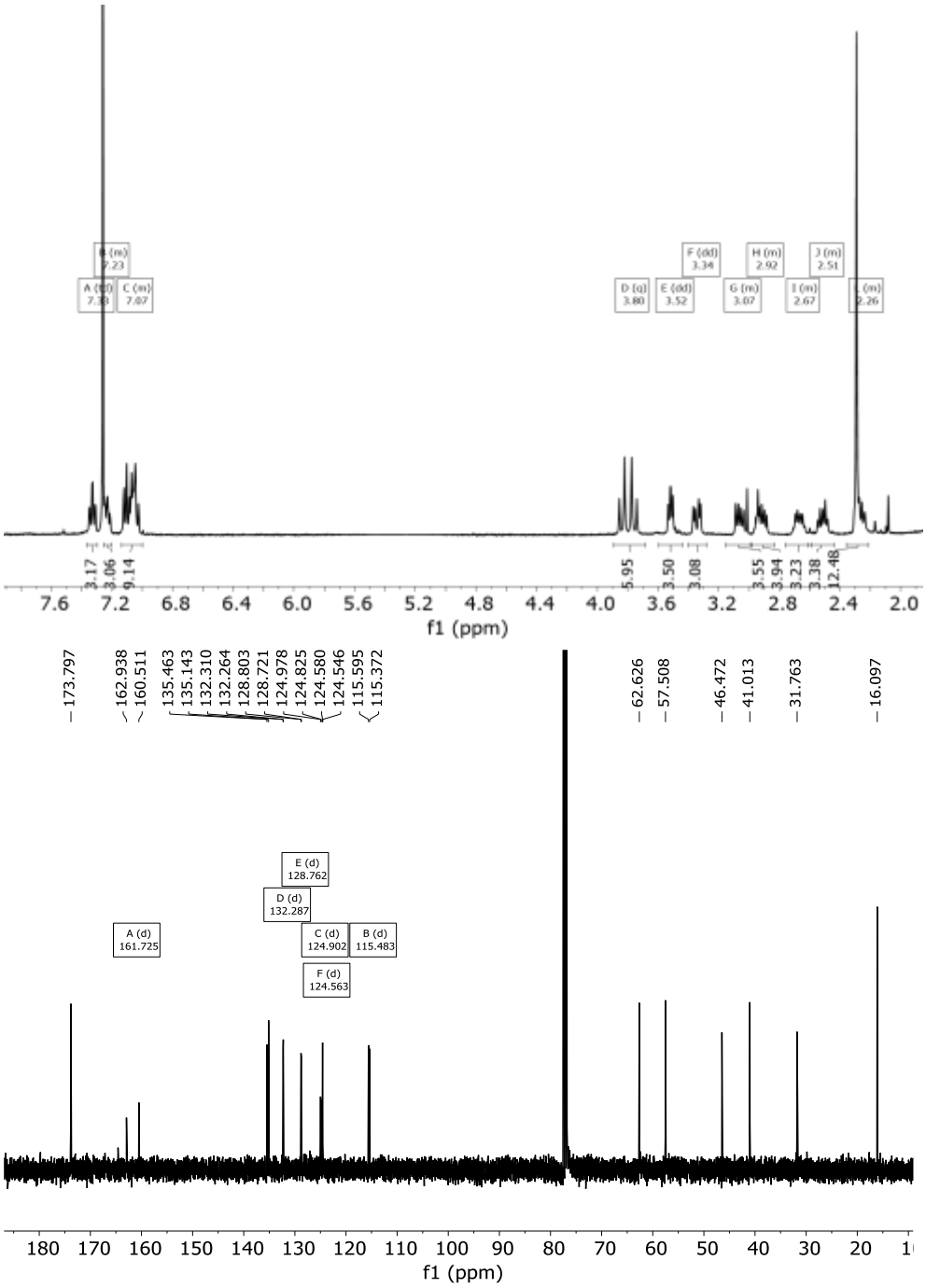


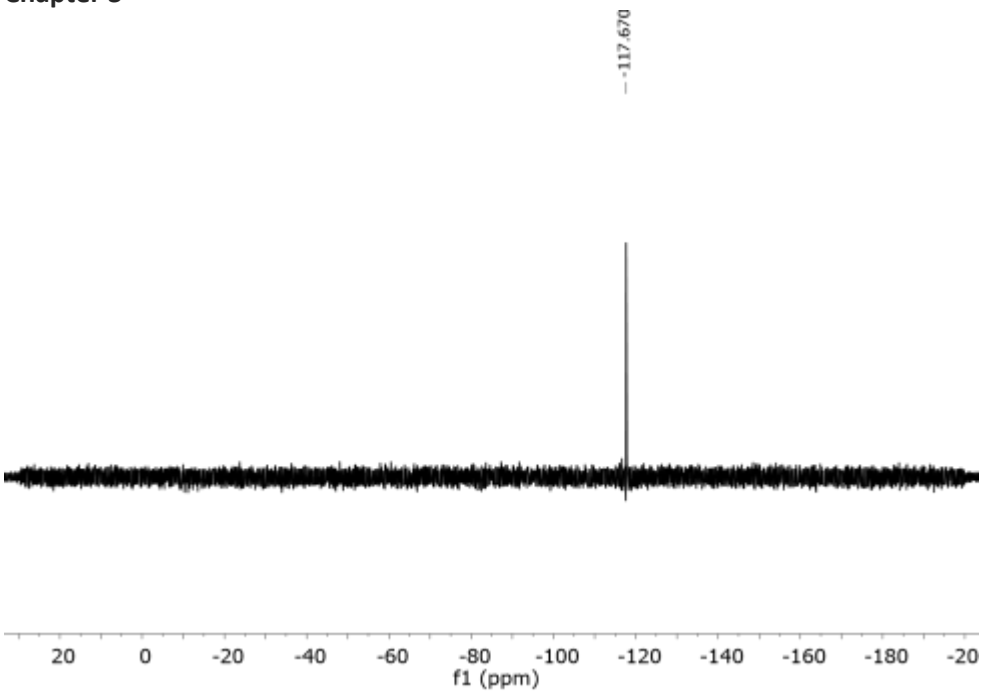
3h (160 mg, 0.146 mmol) was dissolved in ACN (40 mL). Tetrabuthylamonium chloride (20 mg, 0.073 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (58 mg, 0.146 mmol) and potassium carbonate (402 mg, 2.91 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM:MeOH 95 : 5 as eluent to give **1h** as a white solid (40 mg, 0.05 mmol, 35 % yield).

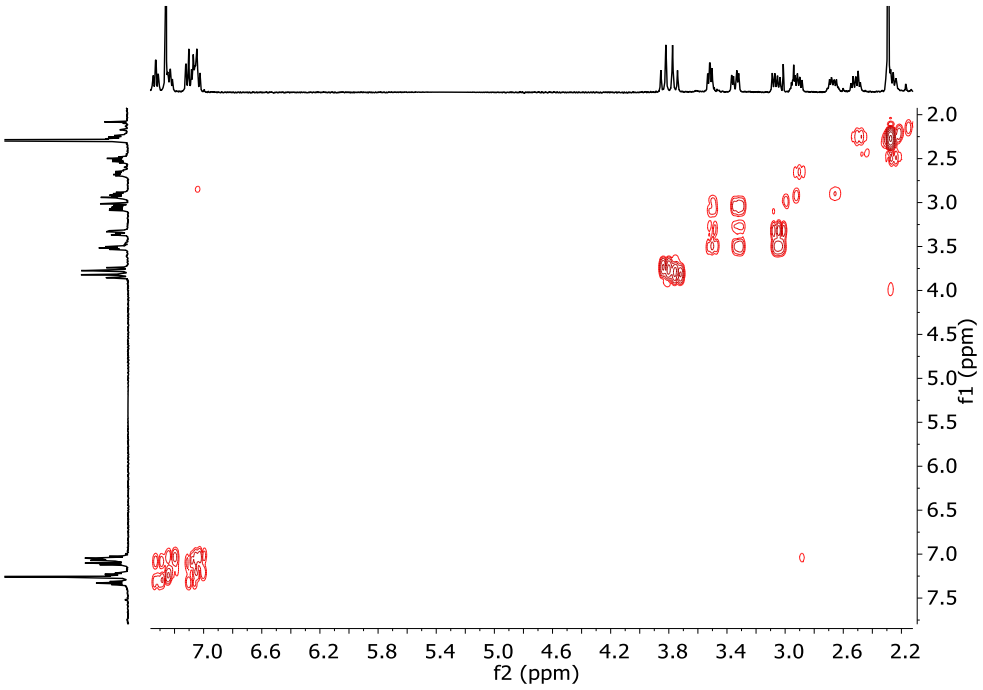
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33 (td, J = 7.6, 1.8 Hz, 3H (H**12/13**)), 7.22(m, 3H, (H**12/13**)), 7.07 (m, 6H, (H**14,15**)), 3.80 (AB_q, δ_A = 3.81, δ_B = 3.74, J_{AB} = 13.6 Hz, 6H, (H**5**)), 3.52 (X subsystem from ABX, J_{AX} = 5.1, J_{BX} = 7.1 Hz, 3H (H**4**)), 3.34 (A subsystem from ABX, J_{AX} = 5.1, J_{AB} = 14.0 Hz, 3H, (H**9**)), 3.06 (B subsystem from ABX, J_{BX} = 7.1, J_{AB} = 14.0 Hz, 3H, (H**9**)), 2.92 (m, 3H, (H**2**)), 2.67 (m, 3H, (H**2**)), 2.53 (m, 3H, (H**1**)), 2.26 (s, 9H, (H**8**) overlapped with m, 3H, (H**1**)).

^{13}C NMR (101 MHz, $CDCl_3$) δ 173.8 (C**3**), 161.7 (d, $^{1}J_{C\text{-}F}$ = 244.2 Hz, C**11**), 135.5 (C**6/7**), 135.1 (C**6/7**), 132.3 (d, $^{n}J_{C\text{-}F}$ = 4.7 Hz, C**13**), 128.8 (d, $^{n}J_{C\text{-}F}$ = 8.2 Hz, C**14/15**), 125.0 (d, $^{n}J_{C\text{-}F}$ = 15.4 Hz, C**10**), 124.6 (d, $^{n}J_{C\text{-}F}$ = 3.5 Hz, C**14/15**), 115.4 (d, $^{n}J_{C\text{-}F}$ = 22.4 Hz, C**12**), 62.6 (C**4**), 57.5 (C**1**), 46.47 (C**5**), 41.0 (C**2**), 31.8 (C**9**), 16.1 (C**8**).^{19}F-{^{1}H} NMR (376 MHz, $CDCl_3$) δ -117.67.

HRMS (ESI-TOF) m/z [**1h**+ H]$^{+}$ Calc: 798.4313, found: 798.4215


B (m) 7.23
A (td) 7.33
C (m) 7.07
F (dd) 3.34
H (m) 2.92
J (m) 2.51
D (q) 3.80
E (dd) 3.52
G (m) 3.07
I (m) 2.67
(m) 2.26
3.17
3.06
9.14
5.95
3.50
3.08
3.55
3.94
3.23
3.38
12.48
f1 (ppm)
173.797
162.938
160.511
135.463
135.143
132.310
132.264
128.803
128.721
124.978
124.825
124.580
124.546
115.595
115.372
62.626
57.508
46.472
41.013
31.763
16.097
E (d) 128.762
D (d) 132.287
A (d) 161.725
C (d) 124.902
B (d) 115.483
F (d) 124.563
f1 (ppm)


-117.670
20
0
-20
-40
-60
-80
-100
-120
-140
-160
-180
-20
f1 (ppm)


2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
f1 (ppm)
7.0
6.6
6.2
5.8
5.4
5.0
4.6
4.2
3.8
3.4
3.0
2.6
2.2
f2 (ppm)

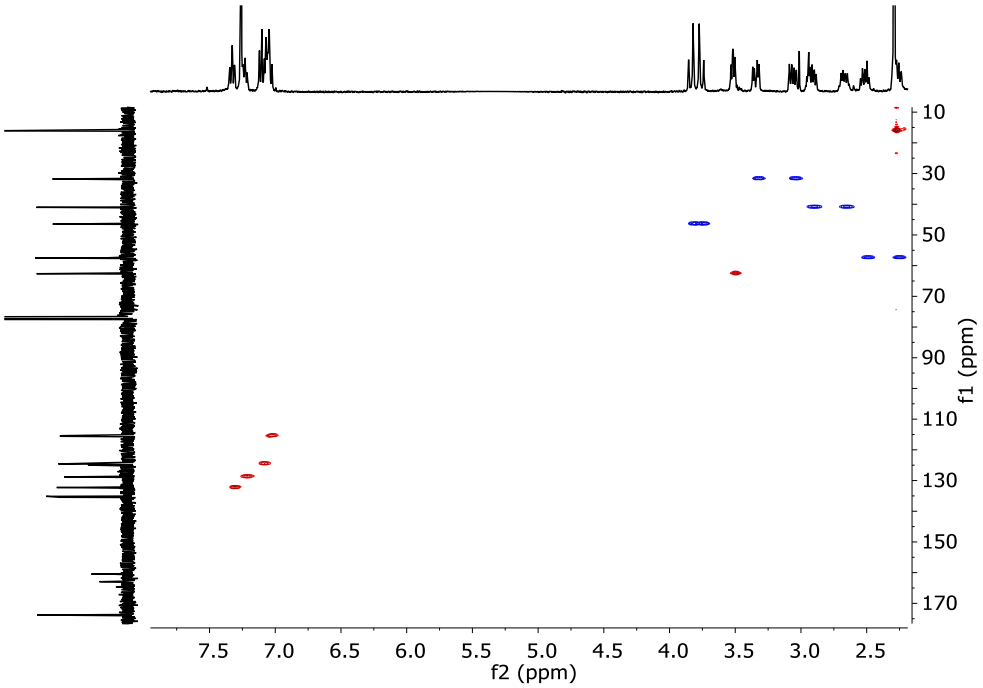


Figure 3.24. ^{1}H-NMR (400 MHz, CD_3OD), ^{13}C-NMR (101 MHz, CD_3OD), ^{19}F NMR (376 MHz, $CDCl_3$), COSY and HSQC spectra of **1h.**

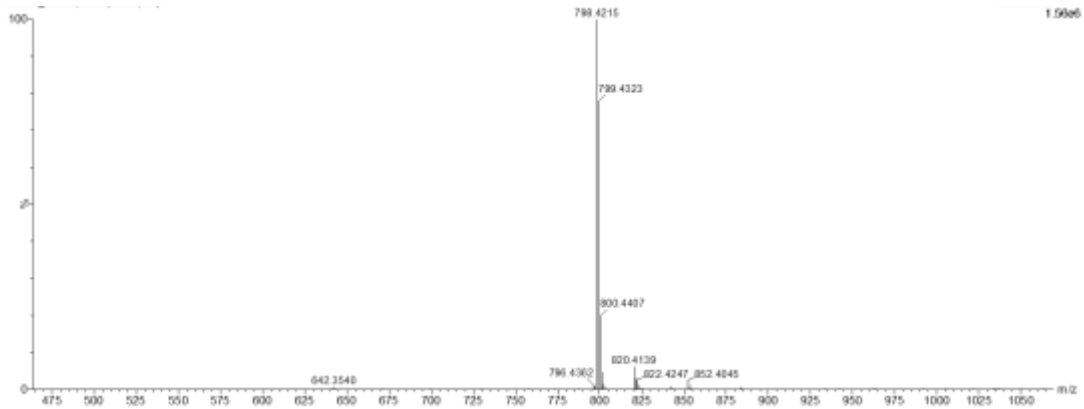


Figure 3.25. HRMS (ESI+) experimental spectrum of **1h.**

3.4.3 Synthesis of 1i

2i. tri-tert-butyl (((nitrilotris (ethane-2,1-diyl))tris(azanediyl))tris(3-(3-fluorophenyl)-1-oxopro-pane-1,2-diyl))tricarbamate

Boc-Phe(3-F)-OH (500 mg, 1.76 mmol) was dissolved in dry DMF (3 mL). *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl O.384 mg, 2 mmol)

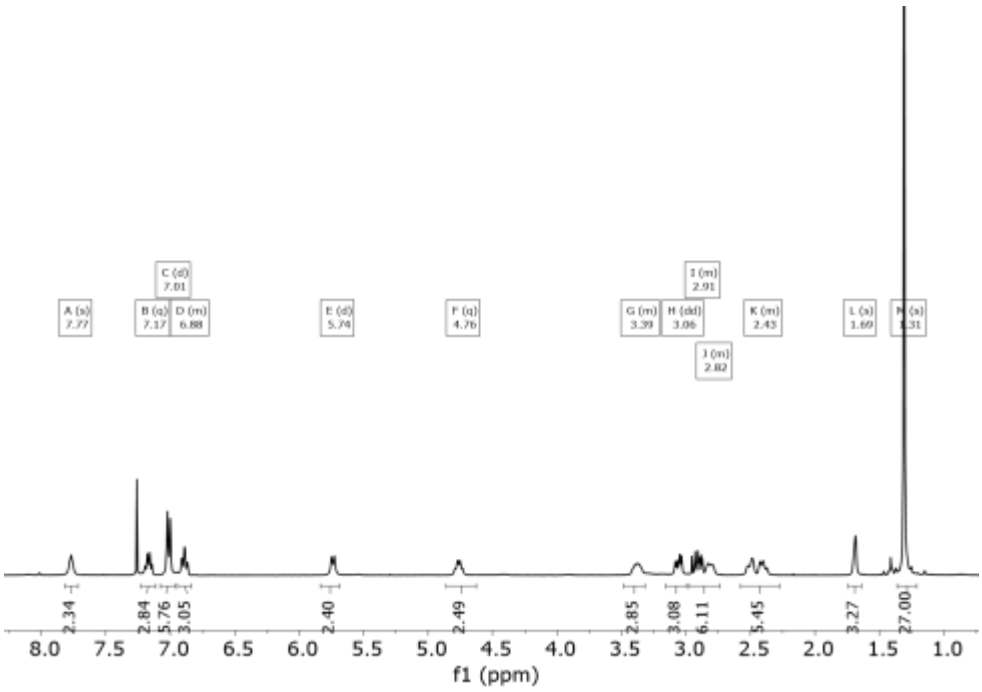
and 1-Hydroxybenzotriazole hydrate (HOBt, 0.306 mg, 2 mmol), *N*,*N*-diisopropylethylamine (DIPEA, 1 mL, 5.78 mmol) and tris(2-aminoethyl)amine (0.08 mL, 0.56 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using 95:5 DCM:MeOH to give 0.412 mg of **2i** (0.437 mmol, 90 % yield).

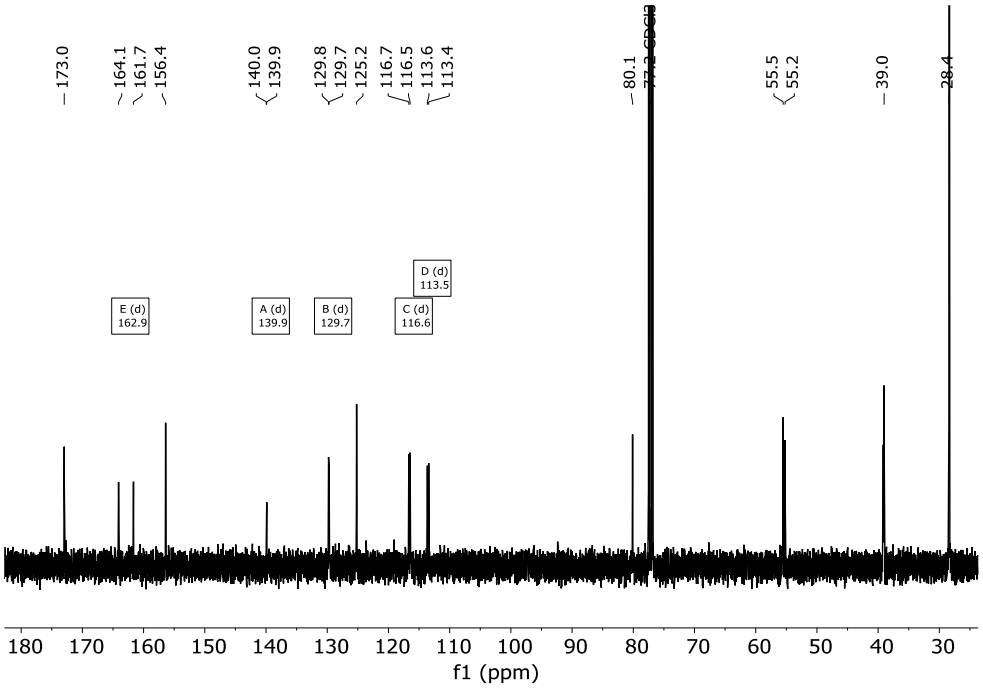
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (s, 3H), 7.17 (q, *J* = 7.3 Hz, 3H), 7.01 (d, *J* = 8.1 Hz, 6H), 6.95 – 6.79 (m, 3H), 5.74 (d, *J* = 8.9 Hz, 3H), 4.76 (q, *J* = 8.2 Hz, 3H), 3.39 (m, 3H), 3.06 (dd, *J* = 13.6, 6.1 Hz, 3H), 2.91 (m, 3H), 2.82 (m, 3H), 2.43 (m, 6H), 1.69 (s, 3H), 1.31 (s, 27H).

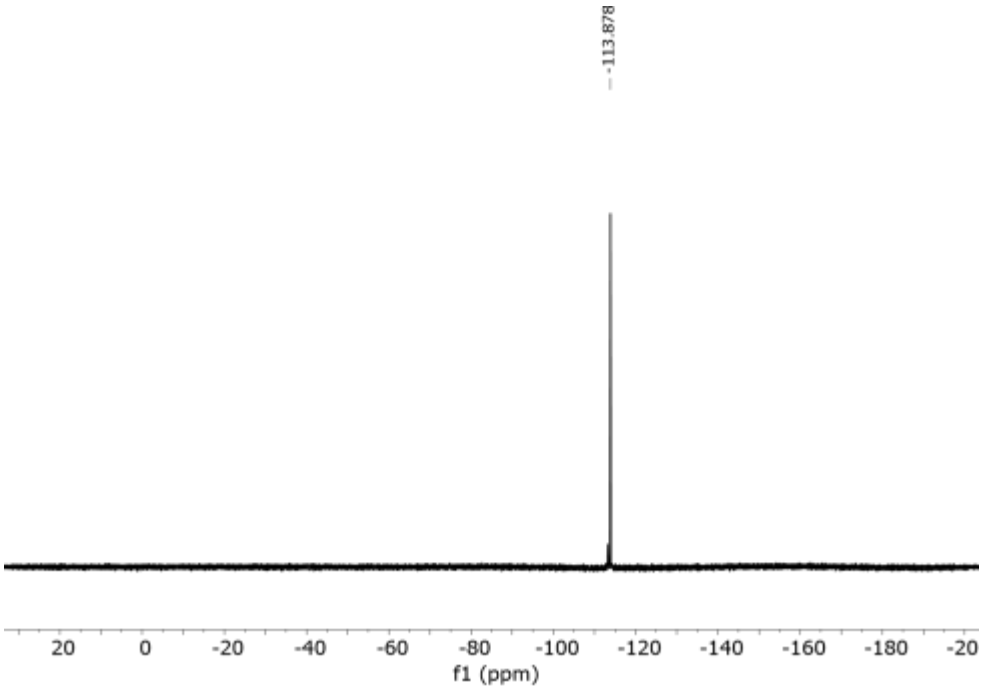
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.0, 162.9 (d, $^{1}J_{C\text{-}F}$ = 245.1 Hz), 156.4, 139.9 ((d, $^{n}J_{C\text{-}F}$ = 7.4 Hz), 129.7 (d, $^{n}J_{C\text{-}F}$ = 8.6 Hz), 125.2, 116.6 (d, $^{n}J_{C\text{-}F}$ = 21.4 Hz), 113.5(d, $^{n}J_{C\text{-}F}$ = 20.7 Hz) 80.1, 55.5, 39.1, 39.0, 28.4.

^{19}F-{^{1}H} NMR (376 MHz, $CDCl_3$) δ -113.88.

HRMS (ESI-TOF) m/z [**2i** + H]$^{+}$ Calc: 942.4947, found: 942.5024




173.0
164.1
161.7
156.4
140.0
139.9
129.8
129.7
125.2
116.7
116.5
113.6
113.4
80.1
55.5
55.2
39.0
28.4
E (d) 162.9
A (d) 139.9
B (d) 129.7
C (d) 116.6
D (d) 113.5
180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30
f1 (ppm)


-113.878
20 0 -20 -40 -60 -80 -100 -120 -140 -160 -180 -20
f1 (ppm)

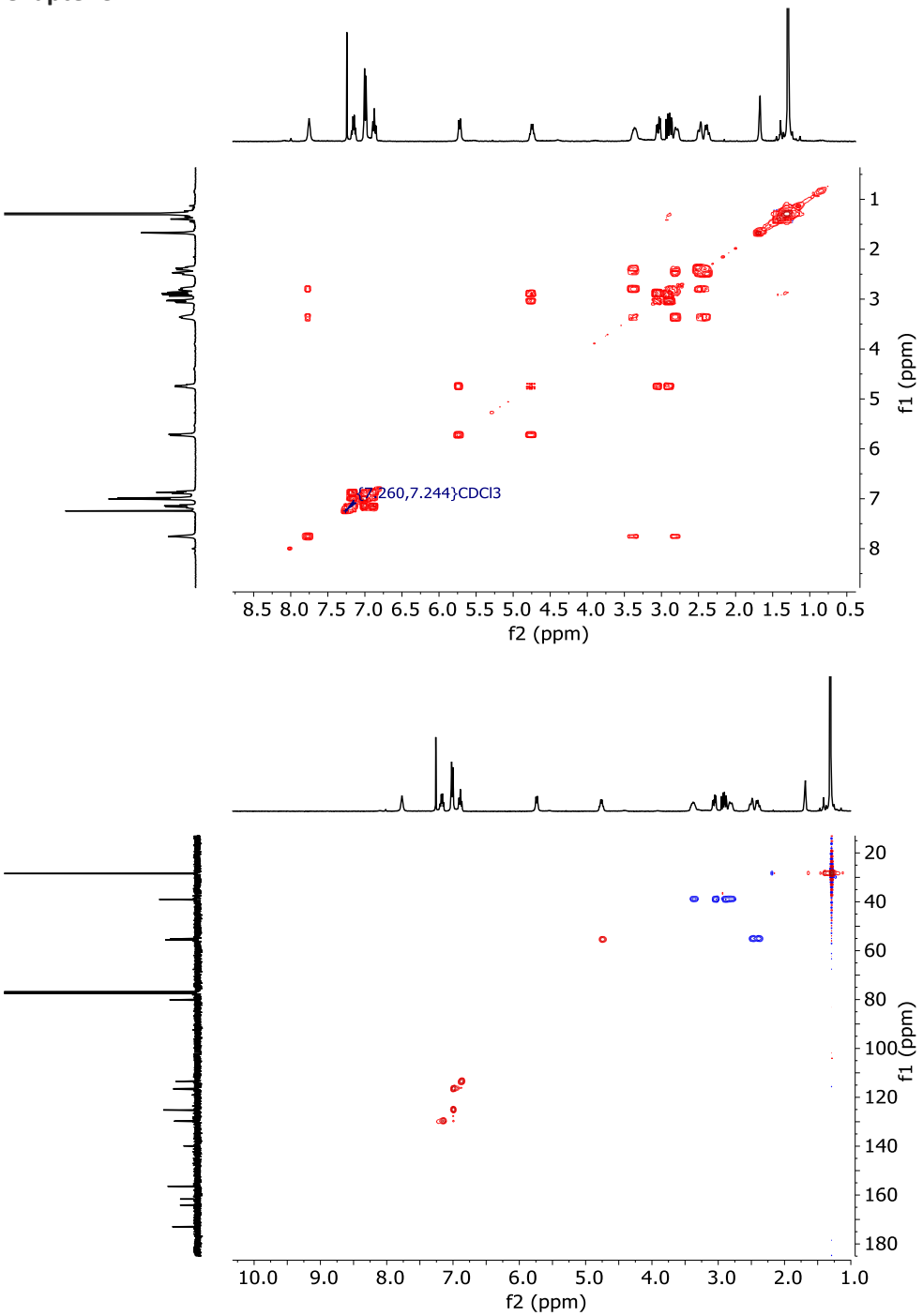


Figure 3.26. ^{1}H-NMR (400 MHz, CD_3OD), ^{13}C-NMR (101 MHz, CD_3OD), ^{19}F NMR (376 MHz, $CDCl_3$), COSY and HSQC spectra of **2i.**

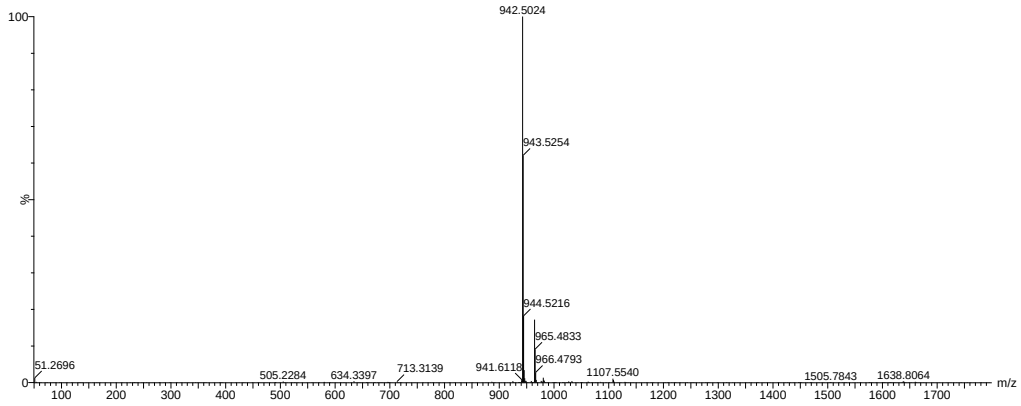


Figure 3.27. HRMS (ESI+) experimental spectrum of **2i.**

3i (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(3-fluorophenyl) propan amide)

2i (400 mg, 0.425 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.65 mL, 4.31 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3i·4TFA** as a white solid. The solid was dissolved in water NaOH 0.1 M and extracted with DCM. Organic layers were dryedo affording 3i as a white solid (255 mg, 0.397 mmol, 93% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.65 (m, 3H), 7.23 (m, 3H), 6.93 (m, 9H), 3.60 (dd, *J* = 9.1, 4.3 Hz, 3H), 3.21 (m, 9H), 2.69 (dd, *J* = 13.7, 9.1 Hz, 3H), 2.54(m, 6H).

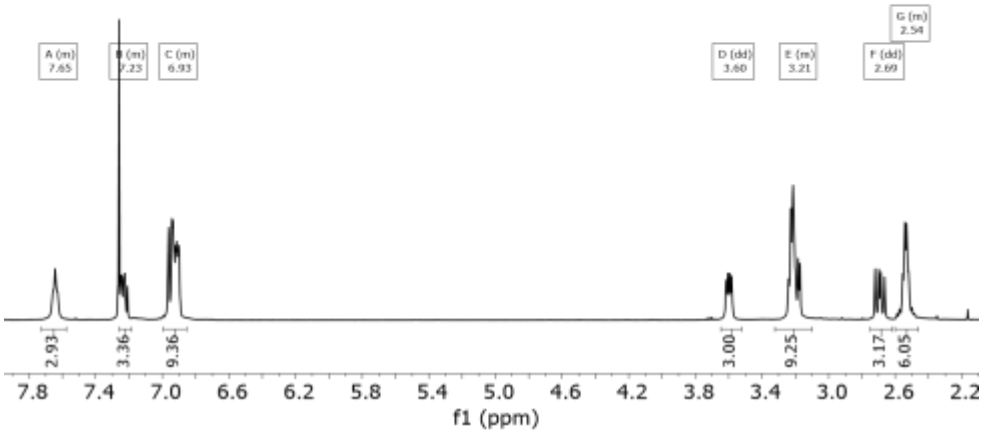
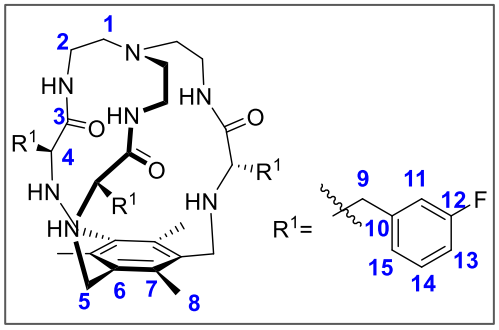


Figure 3.28. ^{1}H-NMR (400 MHz, CD_3OD) spectra of **3i.**

1i

3i (200 mg, 0.311 mmol) was dissolved in ACN (40 mL). Tetrabuthylamonium chloride (30 mg, 0.1.095 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (116 mg, 0.622 mmol) and potassium carbonate (500 mg, 3.62 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM:MeOH 97:3 as eluent to give **1i** as a white solid (90 mg, 0.112 mmol, 38 % yield).

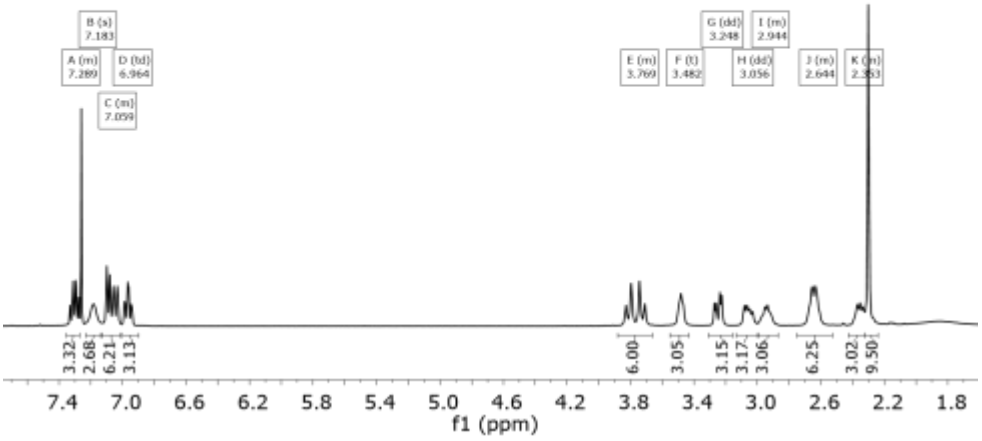


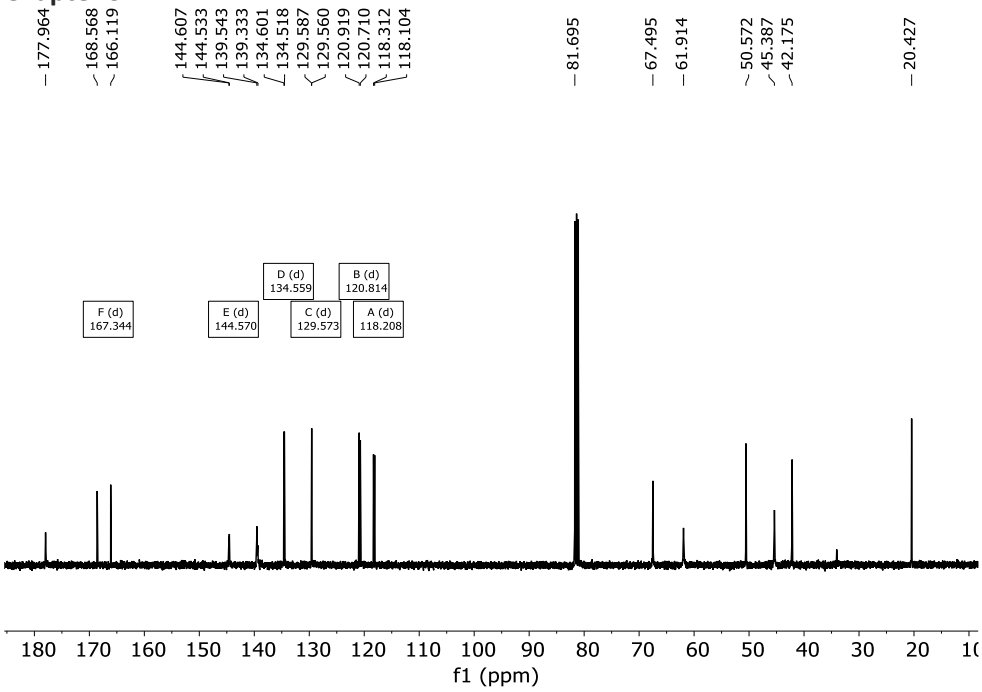
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 (m, 3H, H**ar**), 7.18 (s, 3H (NH_{amide})), 7.06 (m, 6H, (H**ar**)), 6.96 (M, 3H, (H**ar**)), 3.77 ((AB_q, δ_A = 3.838, δ_B=3.757, J_{AB} = 13.7 Hz 6H, (H**5**)), 3.48 (X subsystem from ABX, J_{AX} = 5.1, J_{BX} = 7.1 Hz, 3H (H**4**)), 3.25 (A subsystem from ABX, J_{AX} = 5.1, J_{AB} = 14.0 Hz, 3H, (H**9**)), 3.06 (B subsystem from ABX, J_{BX} = 7.1, J_{AB} = 14.0 Hz, 3H, (H**9**)), 2.94 (m, 3H, (H**2**)), 2.64 (m, 6H, (H**2+1**)), 2.35 (m, 3H, (H**1**)), 2.30 (s, 9H, (H**8**)).

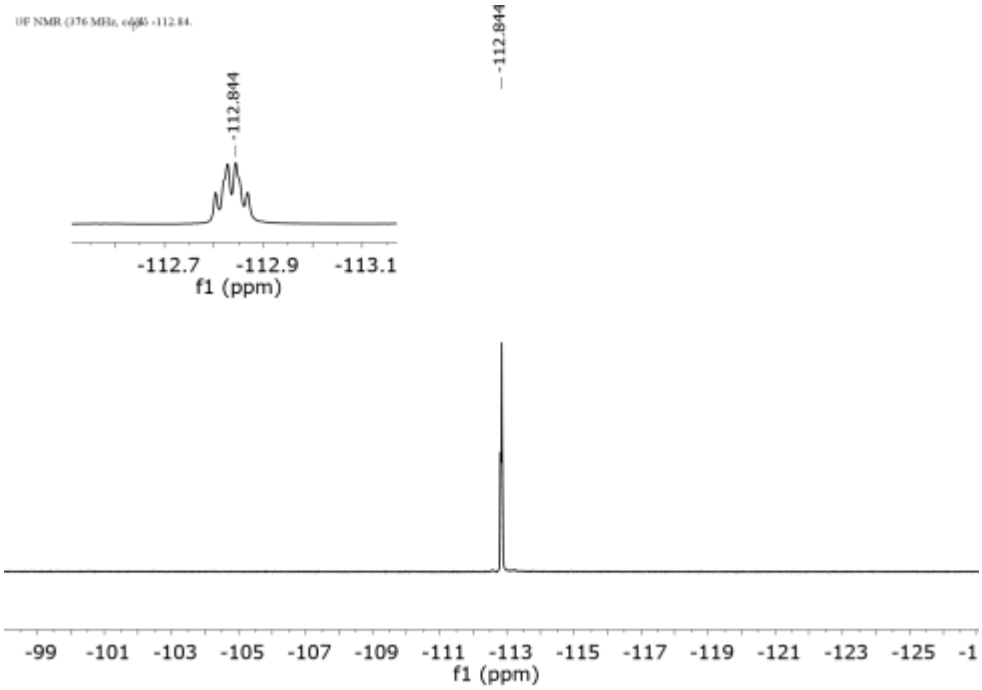
^{13}C NMR (101 MHz, $CDCl_3$) δ 178.0 (C**3**), 167.3 (d, $^{1}J_{C\text{-}F}$ = 246.4 Hz, C**10**), 144.6 (d, $^{n}J_{C\text{-}F}$ = 7.4 Hz C**6/7/12**), 139.5 (C**6/7/12**), 139.3 (C**6/7/12**) 134.6 (d, $^{n}J_{C\text{-}F}$ = 8.3 Hz, C**14/15/11/13**), 129.6 (d, $^{n}J_{C\text{-}F}$ = 2.7 Hz, C**14/15/11/13**), 120.8 (d, $^{n}J_{C\text{-}F}$ = 21.0 Hz, C**14/15/11/13**), 118.1 (d, $^{n}J_{C\text{-}F}$ = 21.0 Hz, C**14/15/11/13**), 67.5 (C**4**), 61.9 (C**1**), 50.6 (C**5**), 45.4 (C**2**), 42.2 (C**9**), 20.4 (C**8**).

^{19}F NMR (376 MHz, $CDCl_3$) δ -112.8 (m, due to H-F coupling).

HRMS (ESI-TOF) m/z [**1i** + H]$^{+}$ Calc: 798.4313, found: 798.4662




177.964
168.568
166.119
144.607
144.533
139.543
139.333
134.601
134.518
129.587
129.560
120.919
120.710
118.312
118.104
81.695
67.495
61.914
50.572
45.387
42.175
20.427
F (d) 167.344
E (d) 144.570
D (d) 134.559
C (d) 129.573
B (d) 120.814
A (d) 118.208
180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20
f1 (ppm)


-112.844
-112.844
-112.7 -112.9 -113.1
f1 (ppm)
-99 -101 -103 -105 -107 -109 -111 -113 -115 -117 -119 -121 -123 -125
f1 (ppm)

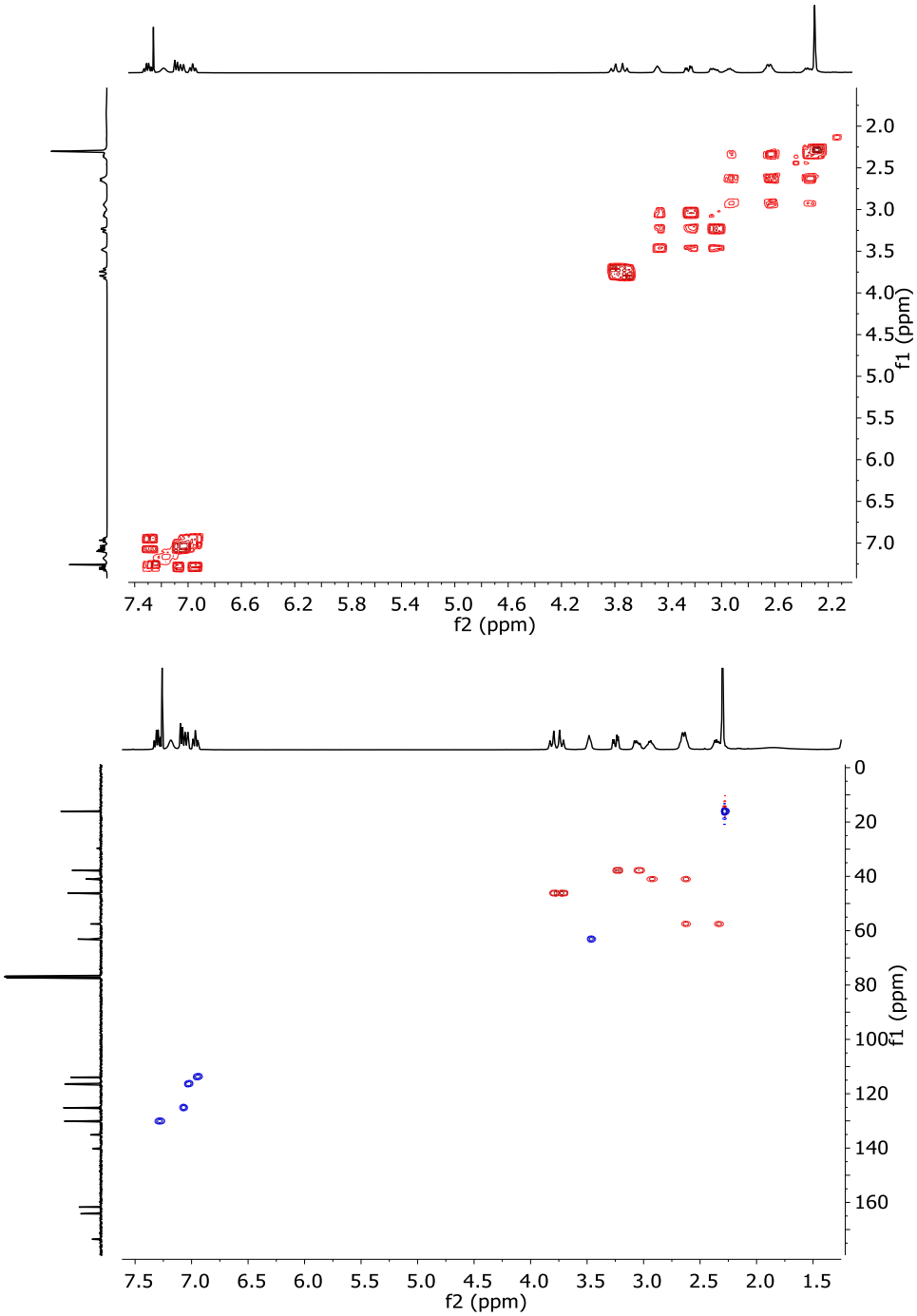


Figure 3.29. ^{1}H-NMR (400 MHz, CD_3OD), ^{13}C-NMR (101 MHz, CD_3OD), ^{19}F NMR (376 MHz, $CDCl_3$), COSY and HSQC spectra of **1i.**

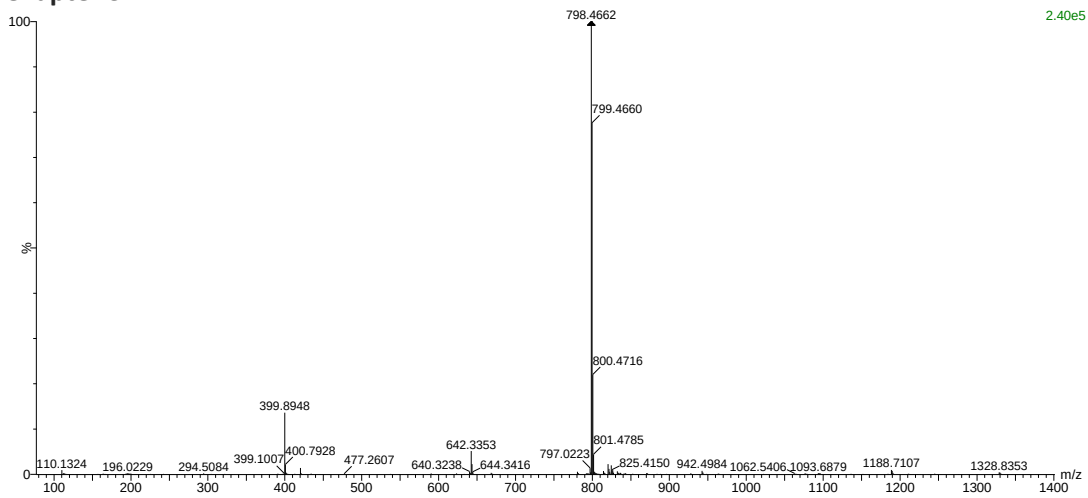


Figure 3.30. HRMS (ESI+) experimental spectrum of **1h.**

3.4.4 Synthesis of 1j

2j. tri-tert-butyl (((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(3-(3,5-difluorophenyl)-1-oxo-propane-1,2-diyl))tricarbamate

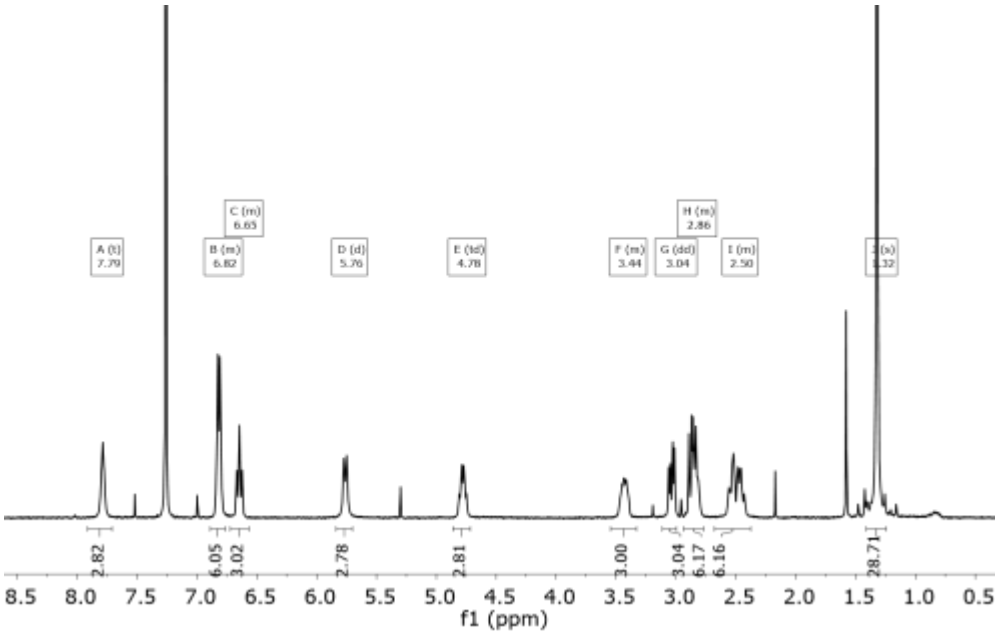
Boc-Phe(3,5-F)-OH (254 mg, 0.84 mmol) was dissolved in dry DMF (3 mL). *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC·HCl O.196 mg, 1.02 mmol) and 1-Hydroxybenzotriazole hydrate (HOBt, 0.157 mg, 1.024 mmol), *N*,*N*-diisopropylethylamine (DIPEA, 0.53 mL, 53.07 mmol) and tris(2-aminoethyl)amine (0.04 mL, 0.25 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5 % w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using 95:5 DCM : MeOH to give 0.220 mg of **2j** (0.256 mmol, 87 % yield).

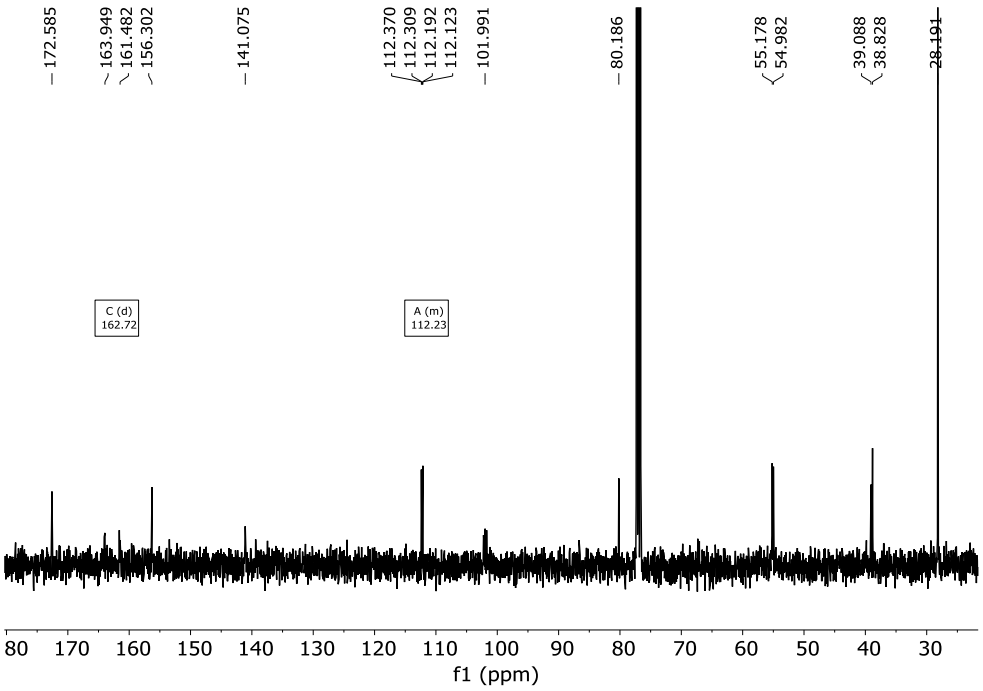
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (t, J = 5.1 Hz, 3H), 6.82 (m, 6H), 6.65 (m, 3H), 5.76 (d, J = 9.1 Hz, 3H), 4.78 (td, J = 9.2, 5.8 Hz, 3H), 3.44 (m, 3H), 3.04 (dd, *J* = 13.6, 5.9 Hz, 3H), 2.86 (m, 6H), 2.55 (m, 6H), 1.32 (s, 27H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 172.6, 162.7. (d, $^{n}J_{C\text{-}F}$ = 248.Hz*)*, 156.3, 141.1, 112.4(m), 101.9 (m), 80.2, 55.2, 55.0, 39.1, 38.8, 28.2.

HRMS (ESI-TOF) m/z [**2j** + H]$^{+}$ Calc: 996.4664, found: 996.5342

^{19}F –{^{1}H}-NMR (376 MHz, $CDCl_3$) δ -110.74.


A (t)
7.79
B (m)
6.82
C (m)
6.65
D (d)
5.76
E (td)
4.78
F (m)
3.44
G (dd)
3.04
H (m)
2.86
I (m)
2.50
2.82
6.05
3.02
2.78
2.81
3.00
3.04
6.17
6.16
28.71
f1 (ppm)


172.585
163.949
161.482
156.302
141.075
112.370
112.309
112.192
112.123
101.991
80.186
55.178
54.982
39.088
38.828
28.191
C (d)
162.72
A (m)
112.23
f1 (ppm)

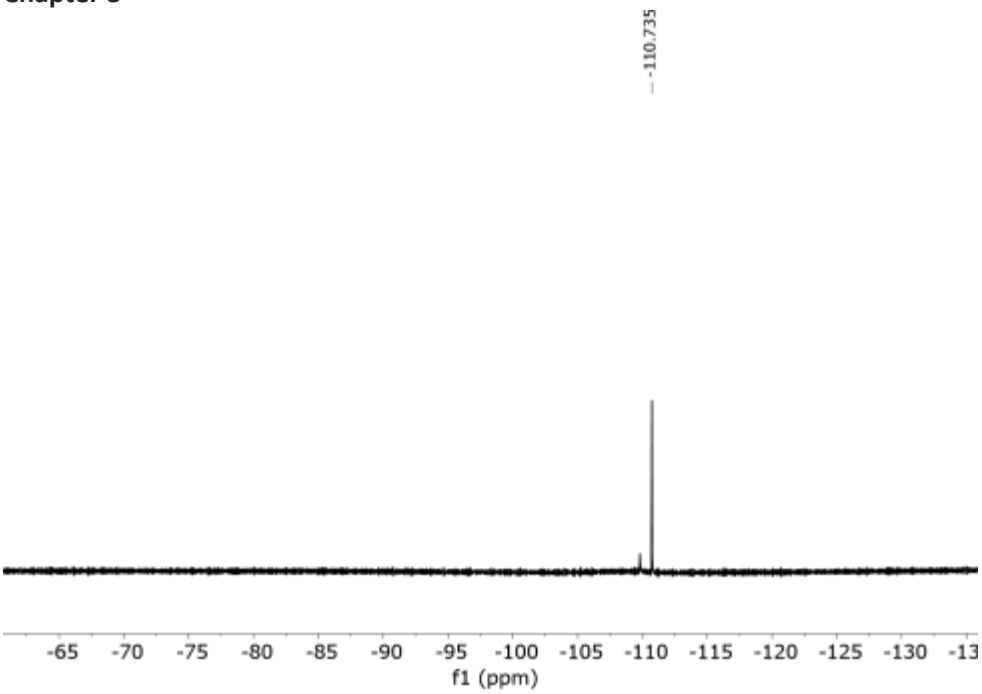


Figure 3.31. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, $CDCl_3$) and ^{19}F NMR (376 MHz, $CDCl_3$) of **2j.**

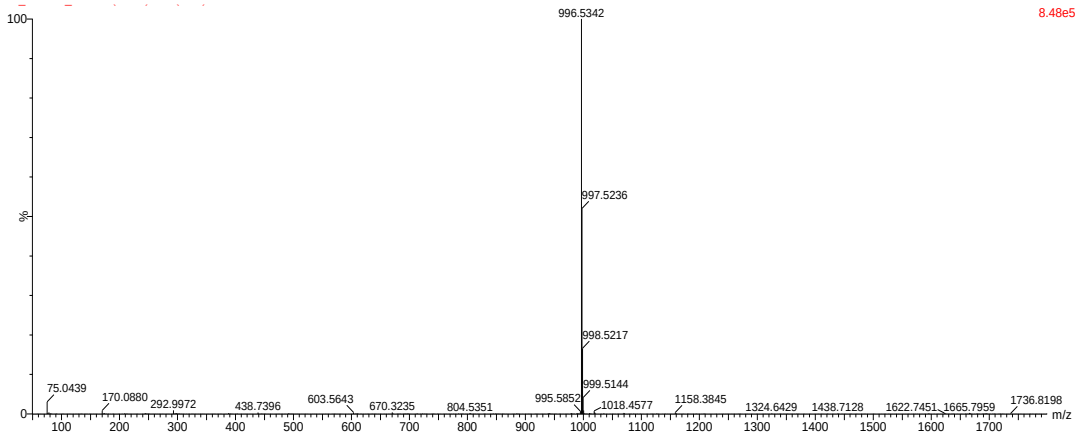


Figure 3.32. HRMS (ESI+) experimental spectrum of **2j.**

3j.N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(3,5-difluorophenyl)propanamide)

2j(220 mg, 0.225 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The

residue was washed several times with diethyl ether and dried affording **3j·4TFA** as a white solid (279 mg, 0.250 mmol, 91% yield).

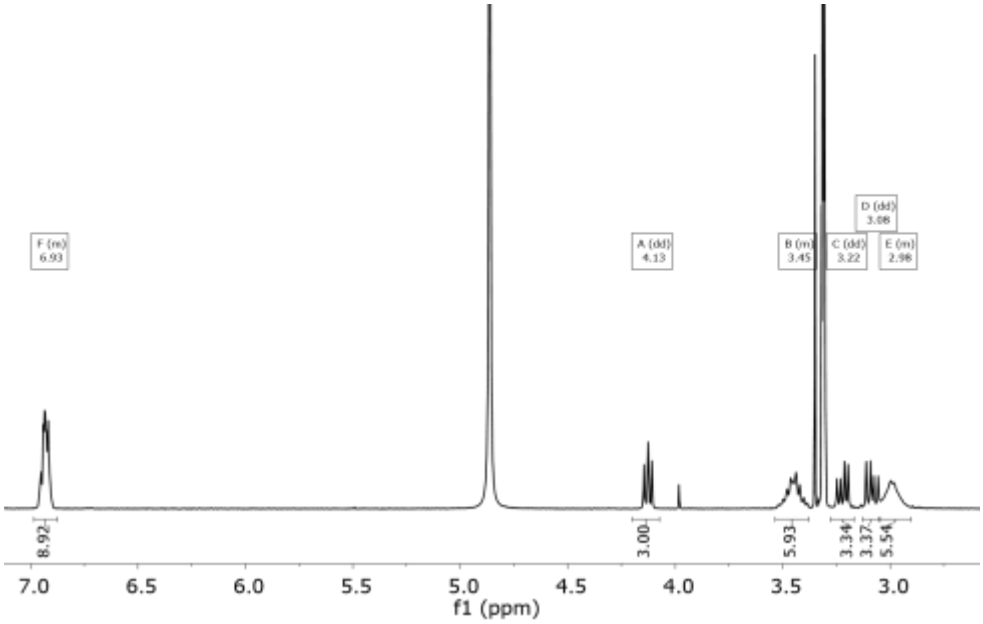
TFA was removed by dissolving the solid in NaOH 1M and extract several times with CH_2Cl_2.

1H NMR (400 MHz, CD_3OD) δ 6.93 (m, 9H), 4.13 (m, 3H), 3.45 (m, 6H), 3.22 (dd, *J* = 14.0, 6.7 Hz, 3H), 3.08 (dd, *J* = 14.0, 8.1 Hz, 3H), 2.98 (m, 6H).

^{13}C NMR (101 MHz, $CDCl_3$) δ168.5, 163.3 (d, $^1J_{C\text{-}F}$ = 248.3 Hz), 138.5 (m), 112.1 (m), 102.8 (m), 54.0, 52.5, 36.6, 35.4.

^{19}F NMR–{^1H}- (376 MHz, $CDCl_3$) δ -110.74.

HRMS (ESI-TOF) m/z [**3j** + H]$^+$ Calc: 696.3091, found: 696.2816



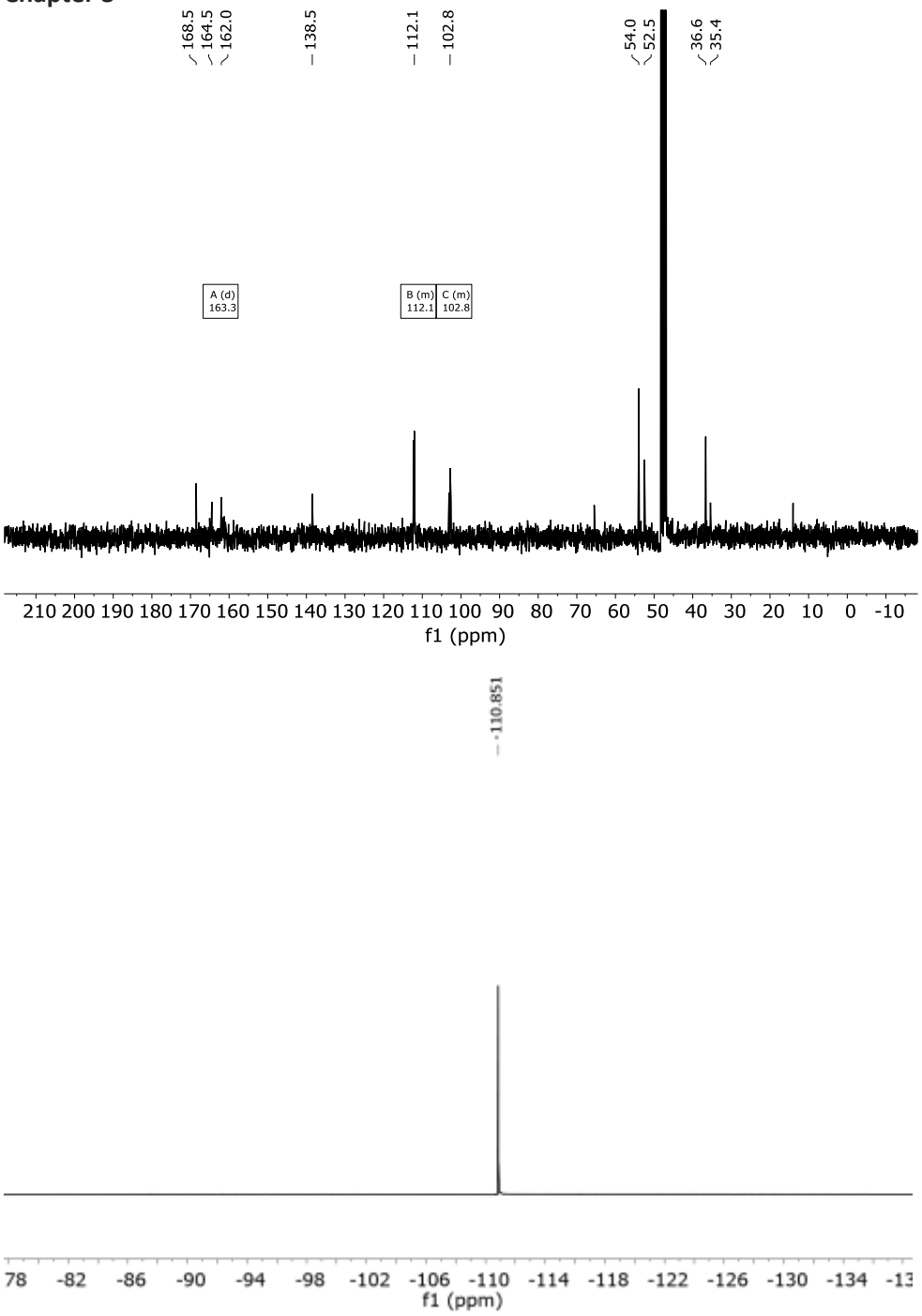


Figure 3.33. ^{1}H-NMR (400 MHz, CD_3OD), ^{13}C-NMR (101 MHz, CD_3OD) and ^{19}F NMR (376 MHz, CD_3OD) of **3j.**

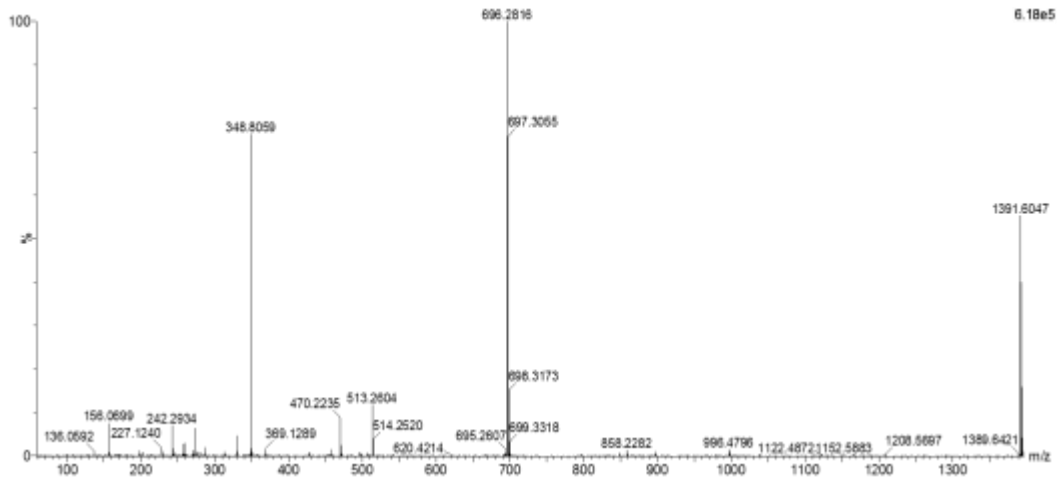
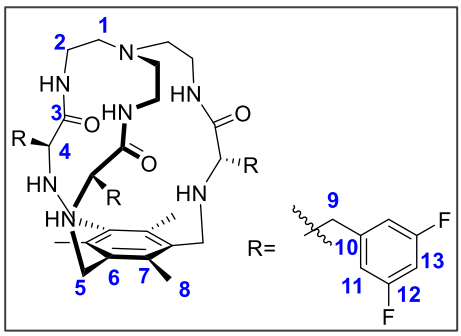


Figure 3.34. HRMS (ESI+) experimental spectrum of **3j.**

1j.



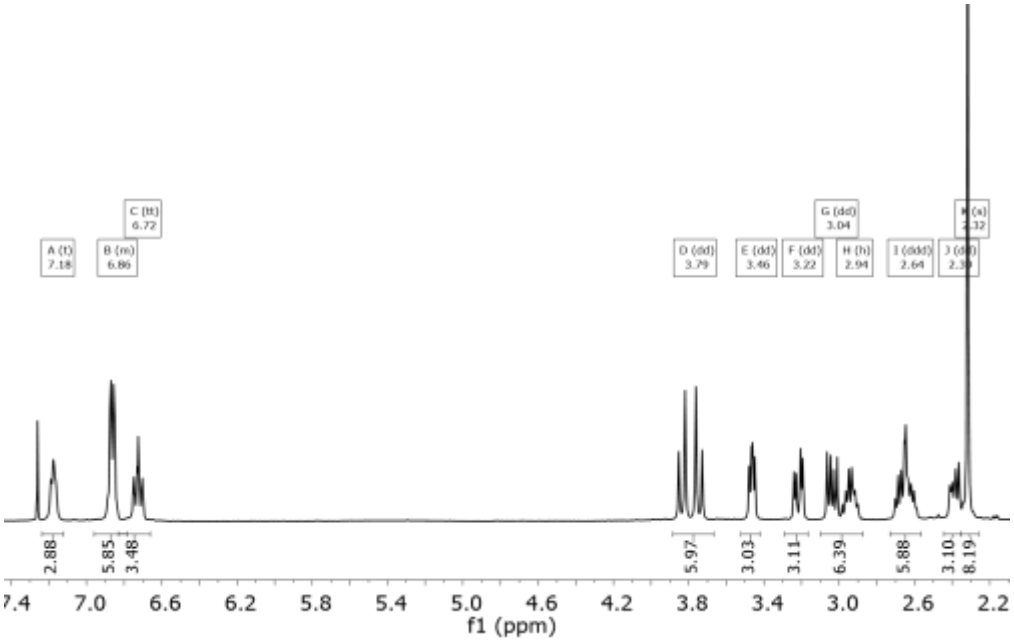
3j (120 mg, 0.121 mmol) was dissolved in ACN (40 mL). Tetrabuthylamonium chloride (24 mg, 0.087 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (68.8 mg, 0.171 mmol) and potassium carbonate (477 mg, 3.45 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM:MeOH 97:3 as eluent to give **1j** as a white solid (80 mg, 0.09 mmol, 55 % yield).

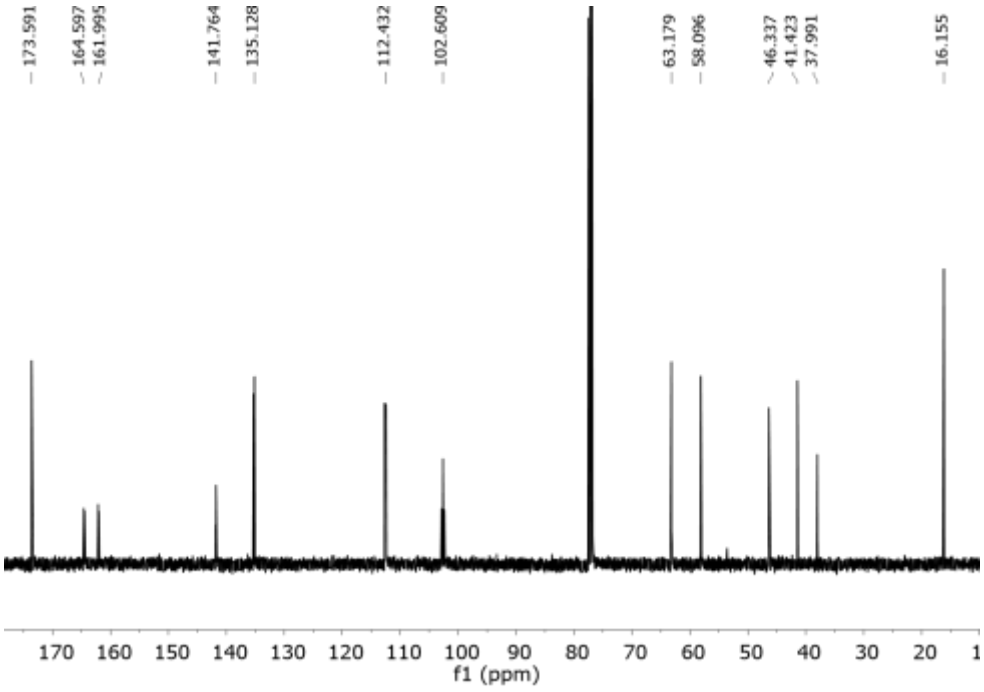
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.175 (t, J = 5.4 Hz, 3H, NH_{amide}), 6.86 (m, 6H, H**11**), 6.725 (tt_{ap}, J = 8.9, 2.4 Hz, 3H, H**13**), 3.790 (AB_q, δ_A = 3.84, δ_B = 3.74, J_{AB} = 13.4 Hz, 6H, H**5**), 3.464 (X subsystem from ABX, J_{AX} = 7.3, J_{BX} = 4.8 Hz, 3H, H**4**), 3.217 (B subsystem from ABX, J_{AB} = 13.9, J_{BX} = 4.8 Hz, 3H, H**9**), 3.037 (dd, A subsystem form ABX, J_{AX} = 7.3, J_{AB} = 13.9 Hz, 3H, H**9**), 2.941 (m, 3H, H**2**), 2.645 (m, 6H, H**2+1**), 2.389 (m, 3H, H**1**), 2.317 (s, 9H, H**8**) .

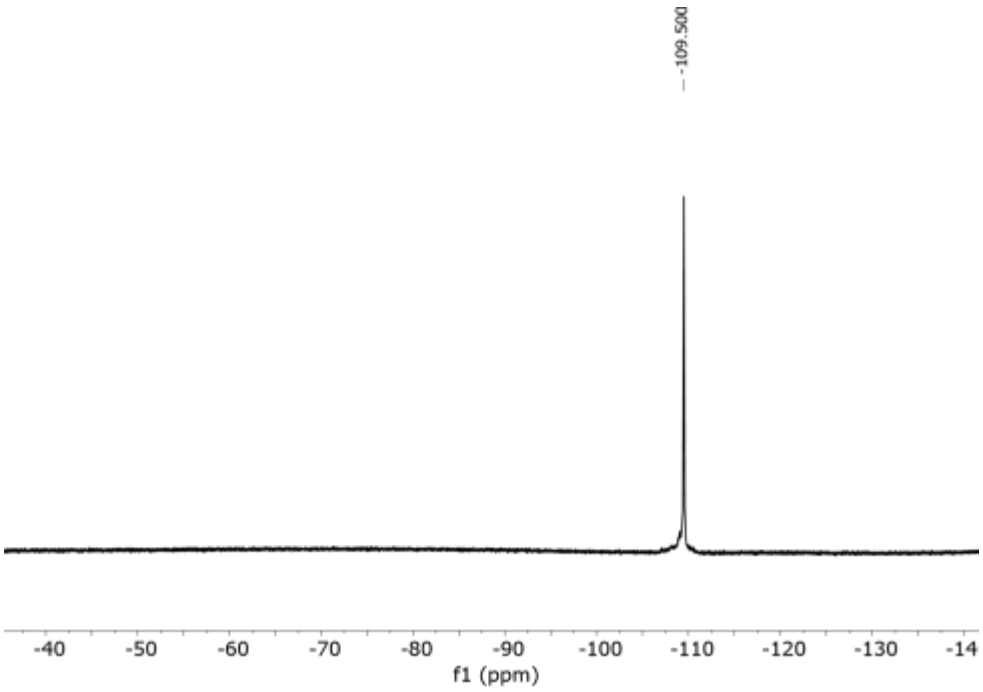
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.6(C**3**), 164.6(C**11/12**), 162.0(C**11/12**), 141.8(C**7**), 135.1(C**6**), 112.4(C**11**), 102.6(C**13**), 63.2(C**4**), 58.1(C**1**), 46.3(C**5**), 41.4(C**2**), 38.0(C**9**), 16.2(C**8**).

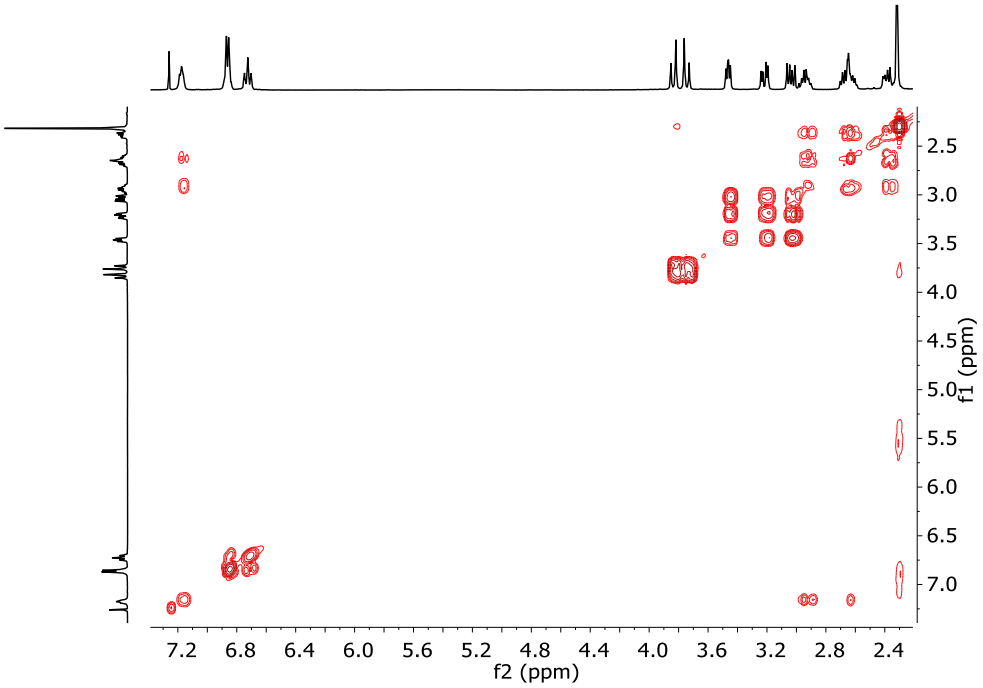
^{19}F NMR-{^{1}H}- (376 MHz, $CDCl_3$) δ -109.50.

HRMS (ESI-TOF) m/z [**1c**+ H]$^+$ Calc: 852.4030, found: 852.4062






-109.500
-40
-50
-60
-70
-80
-90
-100
-110
-120
-130
-14
f1 (ppm)


2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f1 (ppm)
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
f2 (ppm)

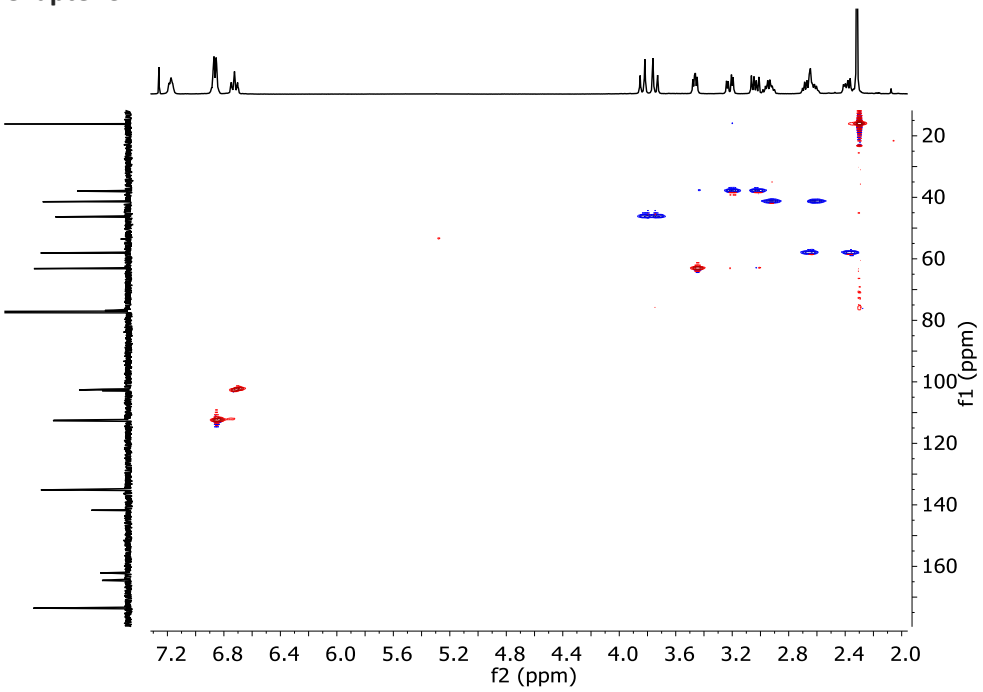


Figure 3.35. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, $CDCl_3$), ^{19}F NMR (376 MHz, $CDCl_3$), COSY, and HSQC spectra of **1j.**

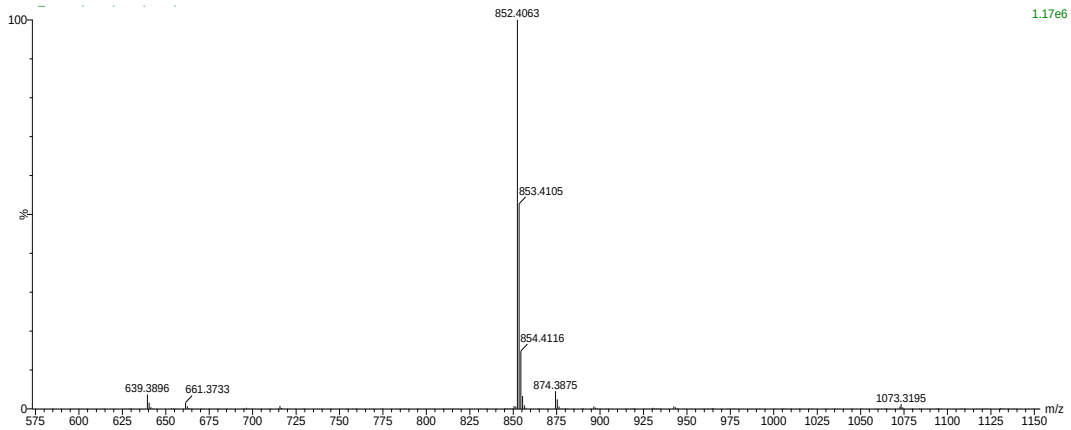


Figure 3.36. HRMS (ESI+) experimental spectrum of **1j.**

3.4.5 Synthesis of 1k

2k. tri-tert-butyl ((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(3-(3,4-difluoro-phenyl)-1-oxopropane-1,2-diyl))tricarbamate

Boc-Phe(3,4-F)-OH (0.445 mg, 1.48 mmol) was dissolved in dry DMF (4 mL). *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC·HCl O.343 mg, 1.793

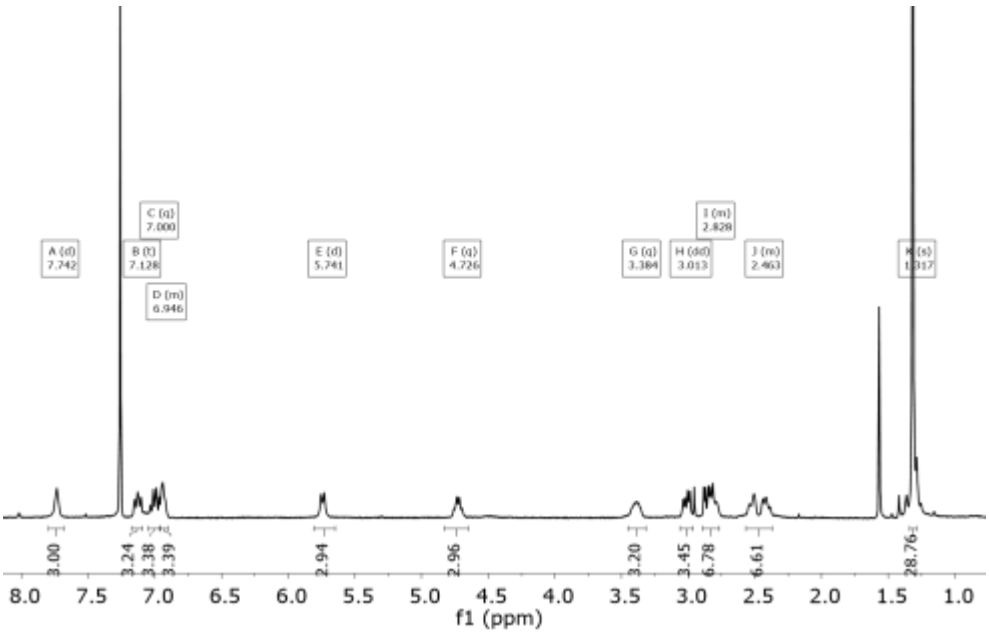
mmol) and 1-Hydroxybenzotriazole hydrate (HOBt, 0.274 mg, 1.79 mmol), *N*,*N*-diisopropylethylamine (DIPEA, 0.93 mL, 5.37 mmol) and tris(2-aminoethyl)amine (0.07 mL, 0.448 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash c chromatography using 97:3 DCM:MeOH to give 0.320 mg of **2k** (0.321 mmol, 72% yield).

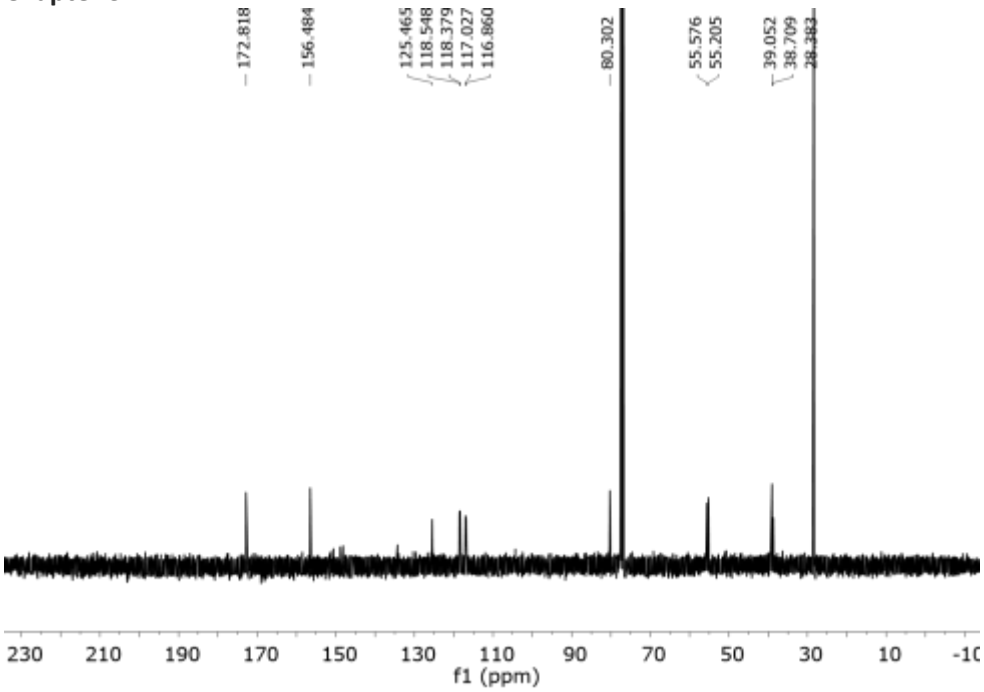
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.74 (d, *J* = 5.2 Hz, 3H), 7.13 (t, *J* = 9.6 Hz, 3H), 7.00 (q, *J* = 8.9 Hz, 3H), 6.94 (m, 3H), 5.74 (d, J = 9.1 Hz, 3H), 4.73 (q, J = 8.4 Hz, 3H), 3.38 (q, J = 7.5, 6.6 Hz, 3H), 3.01 (dd, J = 13.7, 6.2 Hz, 3H), 2.83 (m, 6H), 2.46 (m, 6H), 1.32 (s, 27H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 172.8, 156.5, 125.5, 118.6, 118.4, 117.0, 116.9, 80.3, 55.6, 55.2, 39.1, 38.7, 28.4.

^{19}F NMR (376 MHz, $CDCl_3$) δ -112.55/-112.78 (d, *J* = 21.1 Hz), -115.08/-117.44 ((d, *J* = 21.1 Hz). Each F gives two signals of different intensity due to the presence of rotamers.

HRMS (ESI-TOF) m/z [**2k**+ H]$^+$ Calc: 996.4664; found: 996.5171





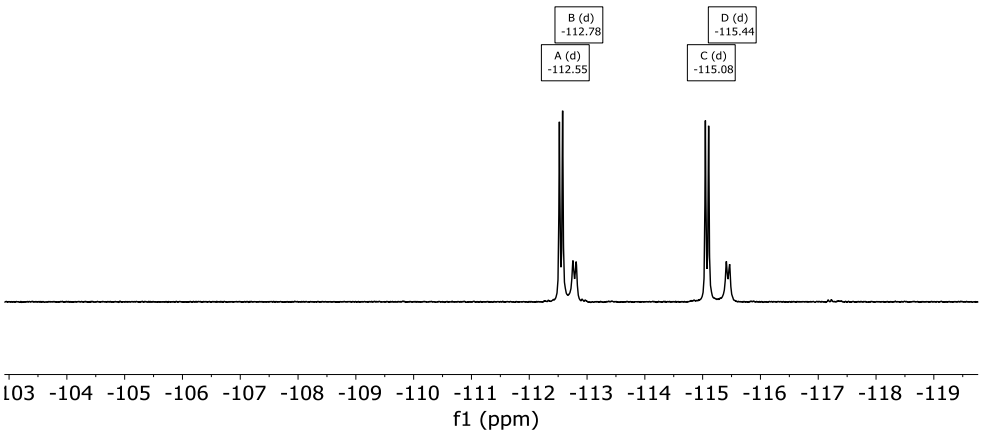


Figure 3.37. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, $CDCl_3$) and ^{19}F NMR (376 MHz, $CDCl_3$) of **2k.**

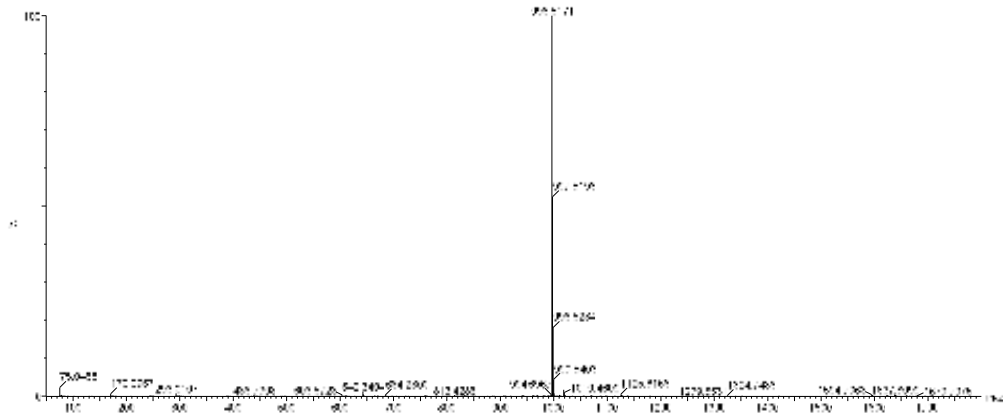


Figure 3.38. HRMS (ESI+) experimental spectrum of **2k.**

3k. (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(3,4-difluorophenyl) pro-panamide)

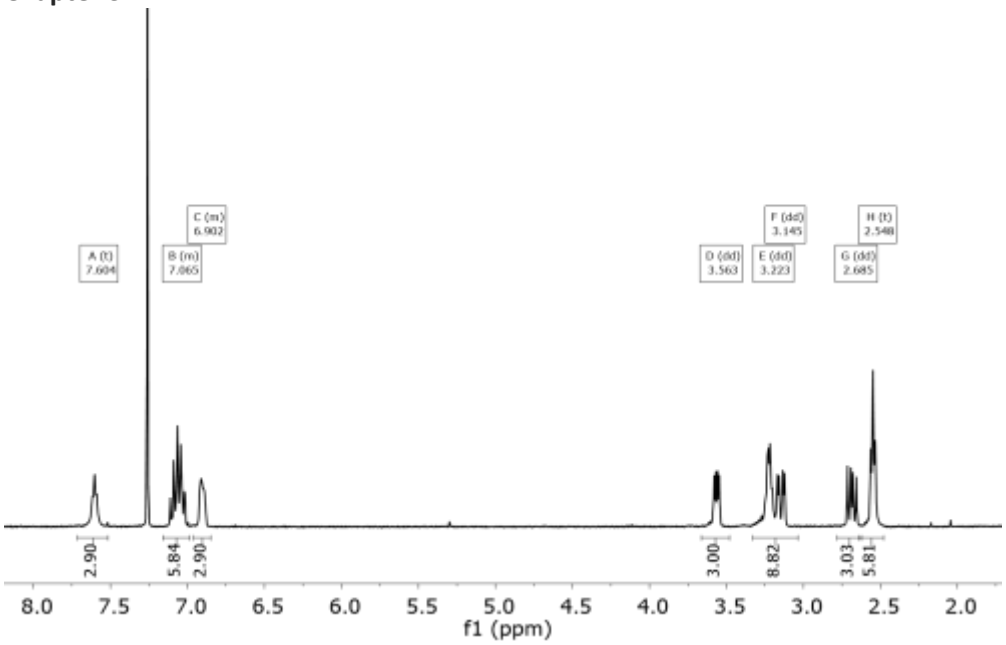
2k (250 mg, 0.251 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3k·4TFA** as a white solid (360 mg, 0.35 mmol, 90% yield). TFA was removed by dissolving the solid in NaOH 1M and extract several times with CH_2Cl_2.

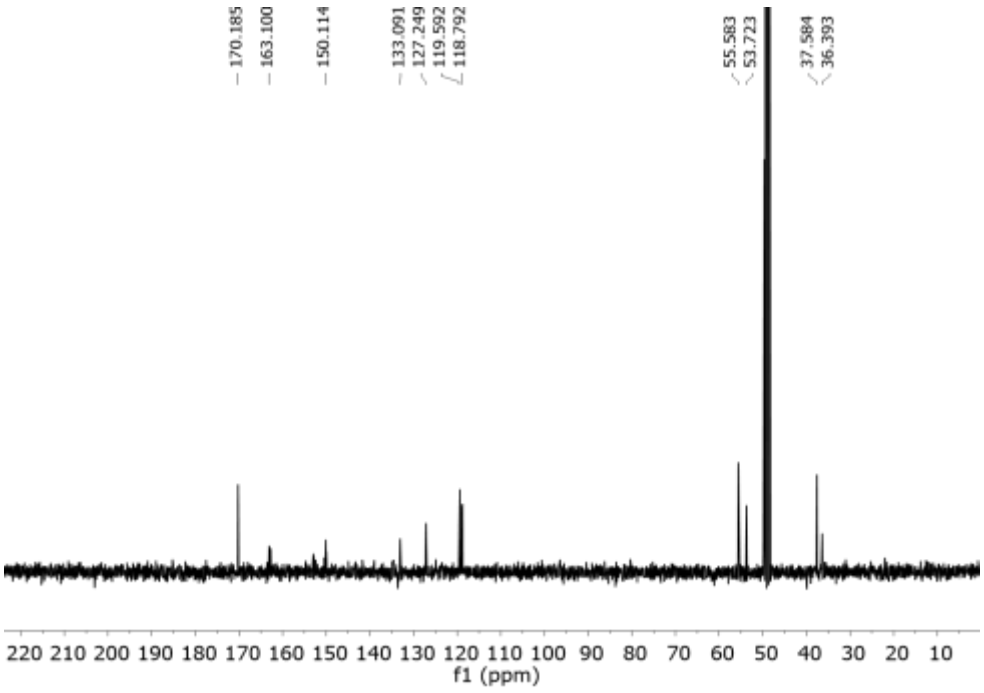
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.604 m, 3H), 7.07 (m, 6H), 6.90 (m, 3H), 3.56 (X subsystem from ABX, J_{AX} = 8.9, J_{BX} = 4.3 Hz, 3H), 3.22 (A subsystem from ABX, J_{AX} = 8.9, J_{AB} = 13.8 Hz, 3H), 3.15 (B subsystem from ABX, J_{BX} = 4.3, J_{AB} = 13.8 Hz, 3H), 2.69 (dd_{ap}, J = 13.8, 8.9 Hz, 3H), 2.55 (t_{ap}, J = 5.7 Hz, 6H).

^{13}C NMR (101 MHz, CD_3OD) δ 170.2, 163.1, 150.1, 133.1, 127.3, 119.6, 118.8, 55.6, 53.7, 37.6, 36.4.

^{19}F NMR (376 MHz, $CDCl_3$) δ -137.48, -140.57.

HRMS (ESI-TOF) m/z [**3k** + H]$^+$ Calc: 696.3091, found: 696.3204


A (t)
7.604
B (m)
7.065
C (m)
6.902
D (dd)
3.563
E (dd)
3.223
F (dd)
3.145
G (dd)
2.685
H (t)
2.548
2.90
5.84
2.90
3.00
8.82
3.03
5.81
8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0
f1 (ppm)


170.185
163.100
150.114
133.091
127.249
119.592
118.792
55.583
53.723
37.584
36.393
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10
f1 (ppm)

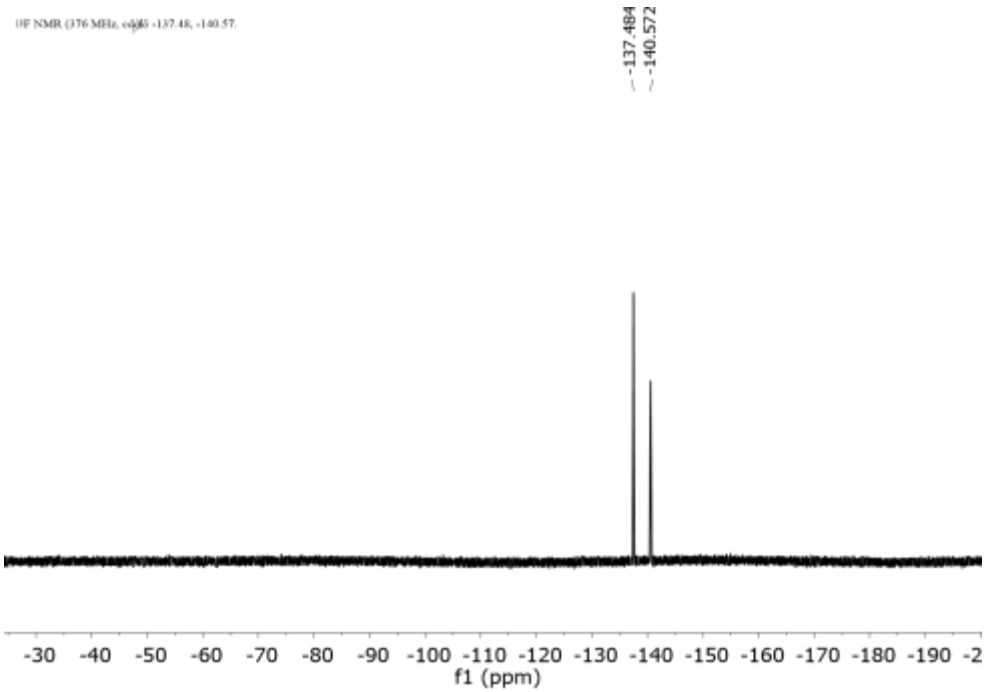


Figure 3.39. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, CD_3OD) and ^{19}F NMR (376 MHz, $CDCl_3$) of **3k.**

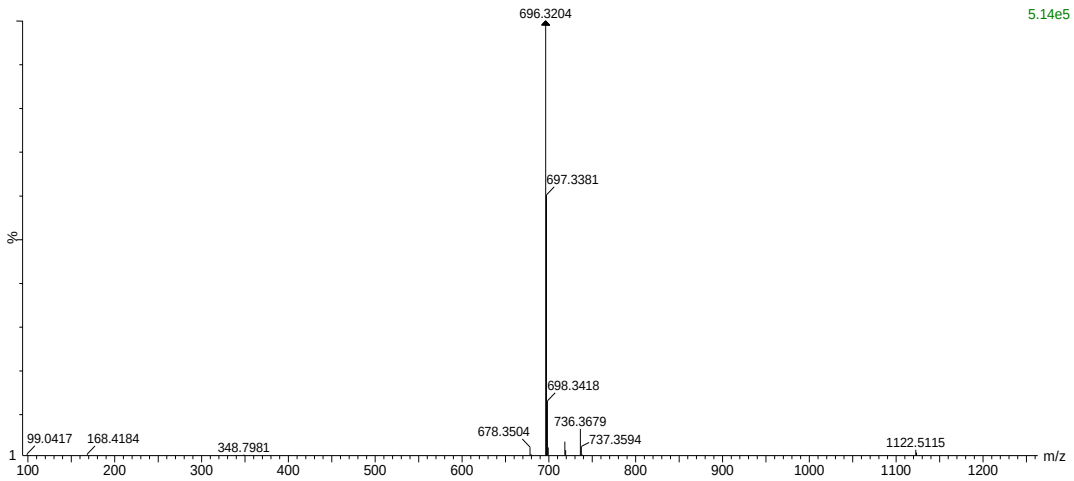
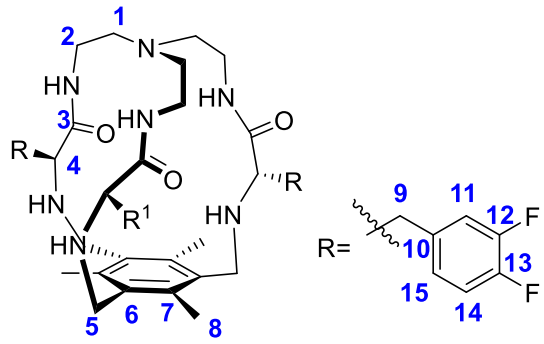


Figure 3.40. HRMS (ESI+) experimental spectrum of **3k.**

1k.

3k (112 mg, 0.113 mmol) was dissolved in ACN (40 mL). Tetrabuthylamonium chloride (22.4 mg, 0.080 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (64.3 mg, 0.160 mmol) and potassium carbonate (445 mg, 3.22 mmol) were added over the solution. The reaction



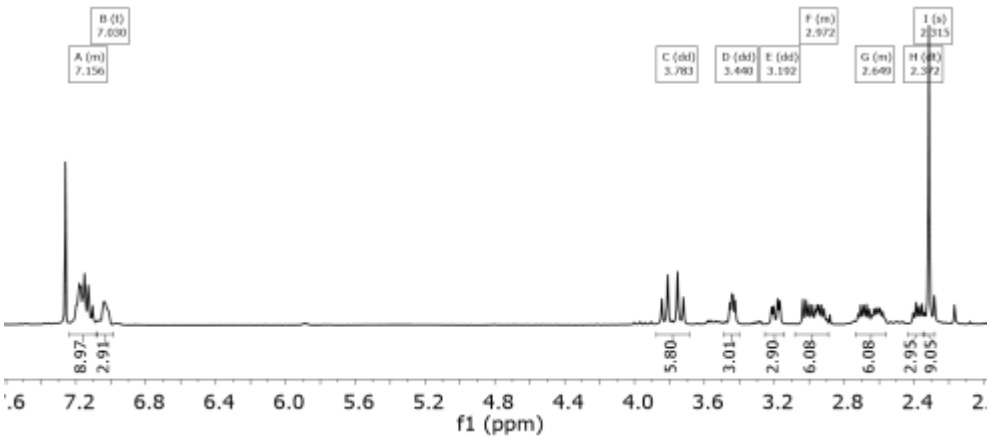
mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM : MeOH 97 : 3 as eluent to give **1k** as a white solid (75 mg, 0.09 mmol, 55 % yield).

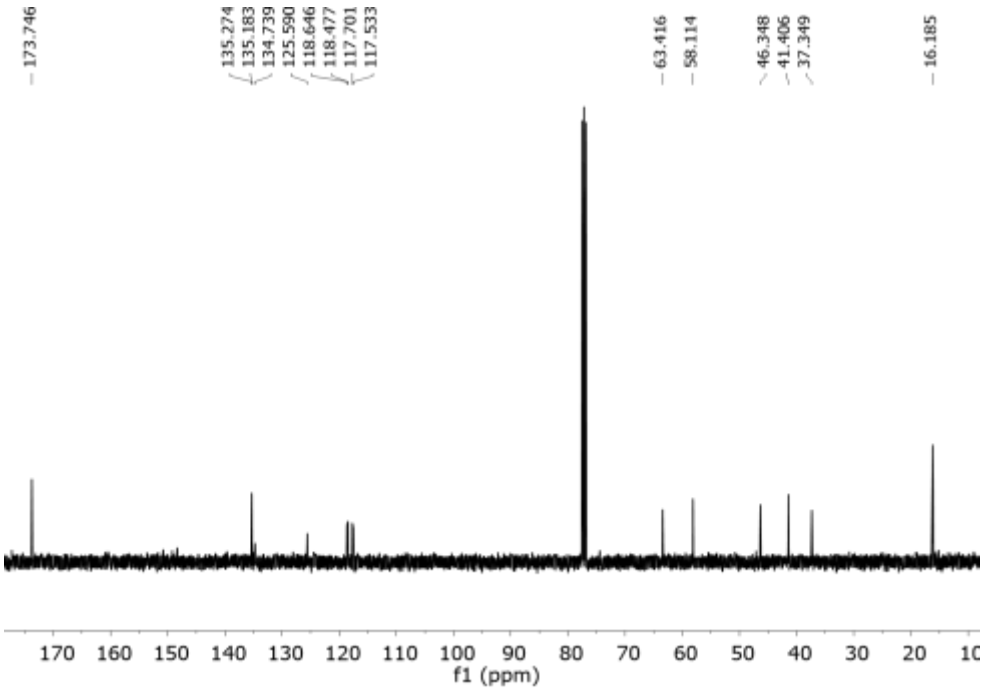
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.15 (m, 9H, (H**11,14,15**), 7.03 (m, 3H, (NH_{amide}), 3.78 ((AB_q, δ_A=3.81, δ_B=3.72, J_{AB}= 13.4 Hz, 6H, H**5**)), 3.44 ((X subsystem from ABX, J_{AX} = 4.7, J_{BX} = 7.2 Hz, 3H (H**4**)), 3.19 (A subsystem from ABX, J_{AX}=4.7, J_{AB} = 14.0 Hz, 3H, (H**9**)), 2.97 (B subsystem from ABX, J_{BX}=7.2, J_{AB} = 14.0 Hz, 3H, (H**9**) + m, 3H (H**2**)), 2.65 (m, 6H, (H**2+1**)), 2.37 (m, 3H, (H**1**)), 2.32 (s, 9H, (H**8**)).

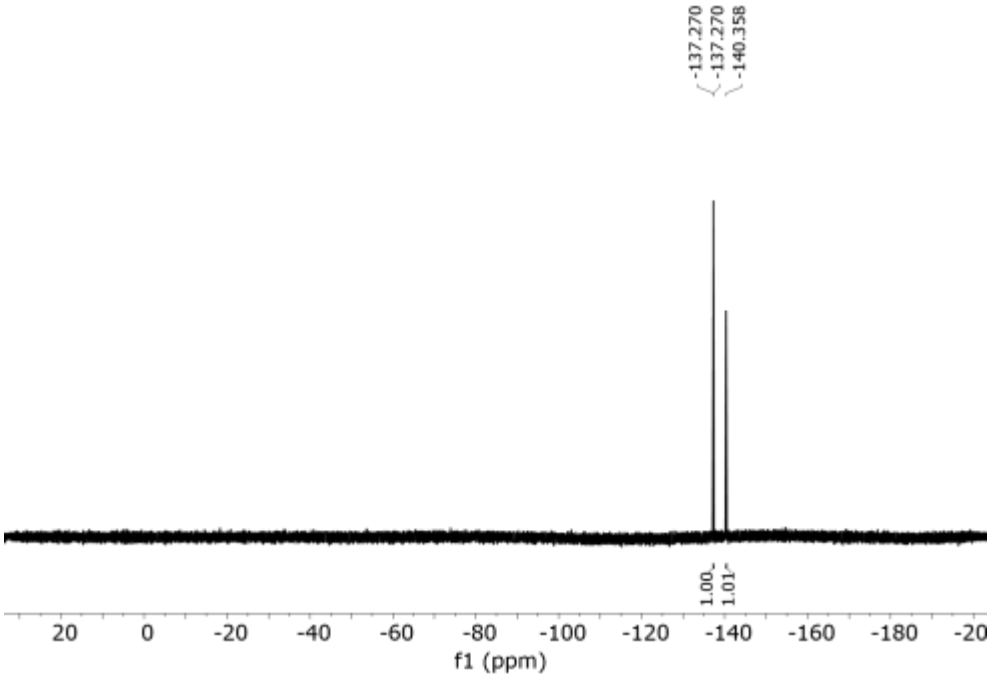
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.8 (C**3**), 135.3(C**6/7**), 135.2 (C**6/7**), 134.8 (C**14**), [125.6, 118.7, 118.5, 117.7, 117.5, (C**10/11/12/13/15**)], 63.4 (C**4**), 58.1 (C**1**), 46.4(C**5**), 41.4 (C**2**), 37.4 (C**9**), 16.2 (C**8**).

^{19}F NMR (376 MHz, $CDCl_3$) δ -137.27, -140.36.

HRMS (ESI-TOF) m/2z [**1k**+ 2H]$^{+}$/2 Calc: 852,9469, found: 852.4105




173.746
135.274
135.183
134.739
125.590
118.646
118.477
117.701
117.533
63.416
58.114
46.348
41.406
37.349
16.185
170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10
f1 (ppm)


-137.270
-137.270
-140.358
1.00
1.01
20 0 -20 -40 -60 -80 -100 -120 -140 -160 -180 -20
f1 (ppm)

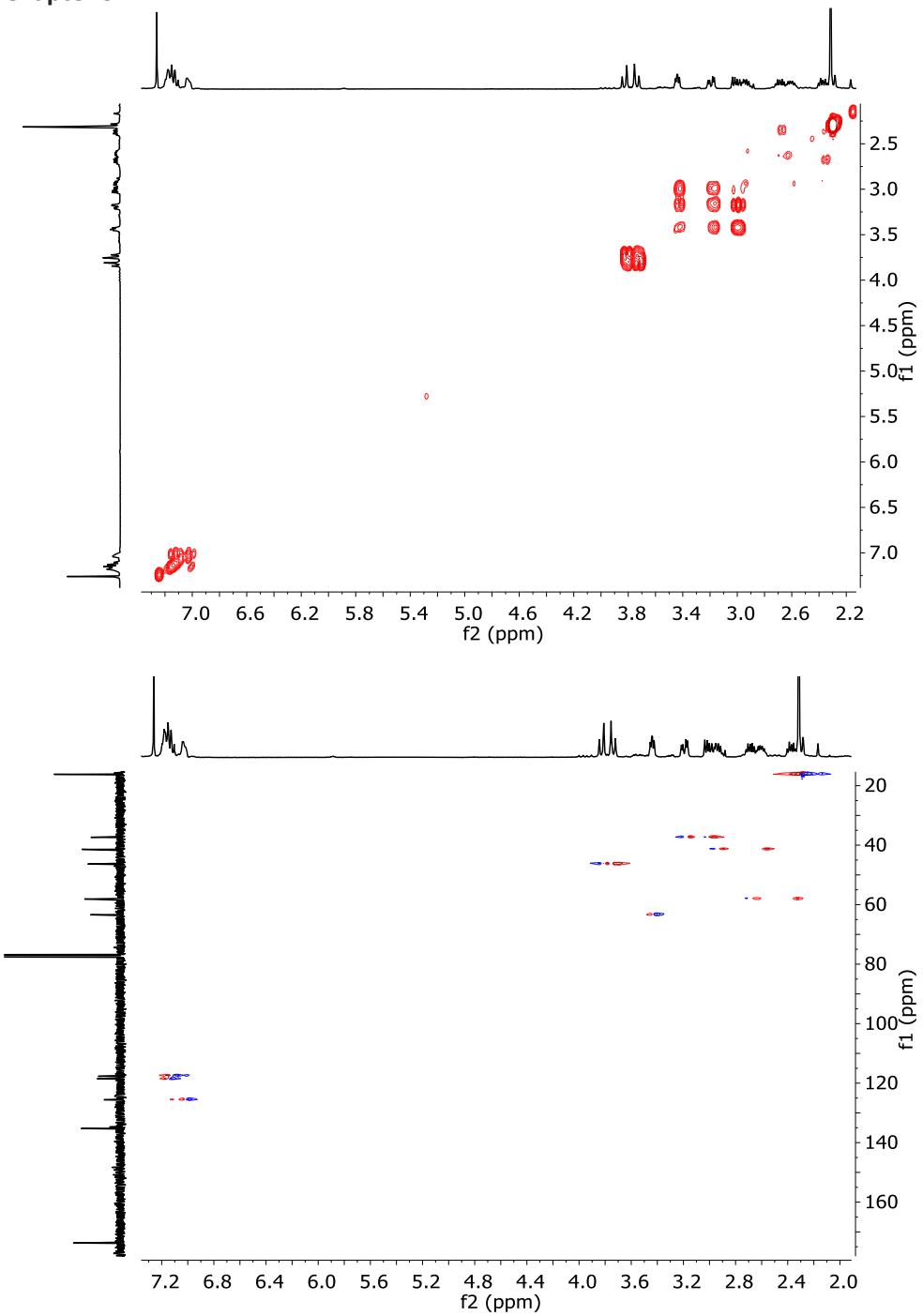


Figure 3.41. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, $CDCl_3$), ^{19}F NMR (376 MHz, $CDCl_3$), COSY and HSQC of **1k.**

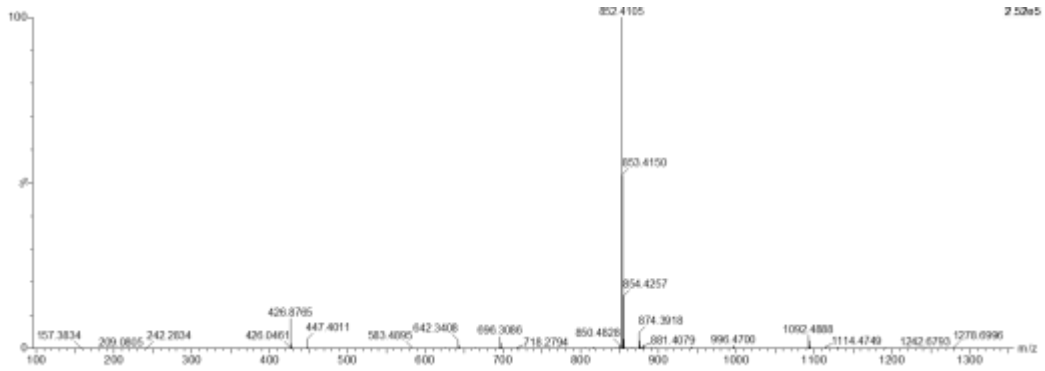


Figure 3.42. HRMS (ESI+) experimental spectrum of **1k.**

3.4.6 Synthesis of 1l

Boc (2,6F-Phe)-OH

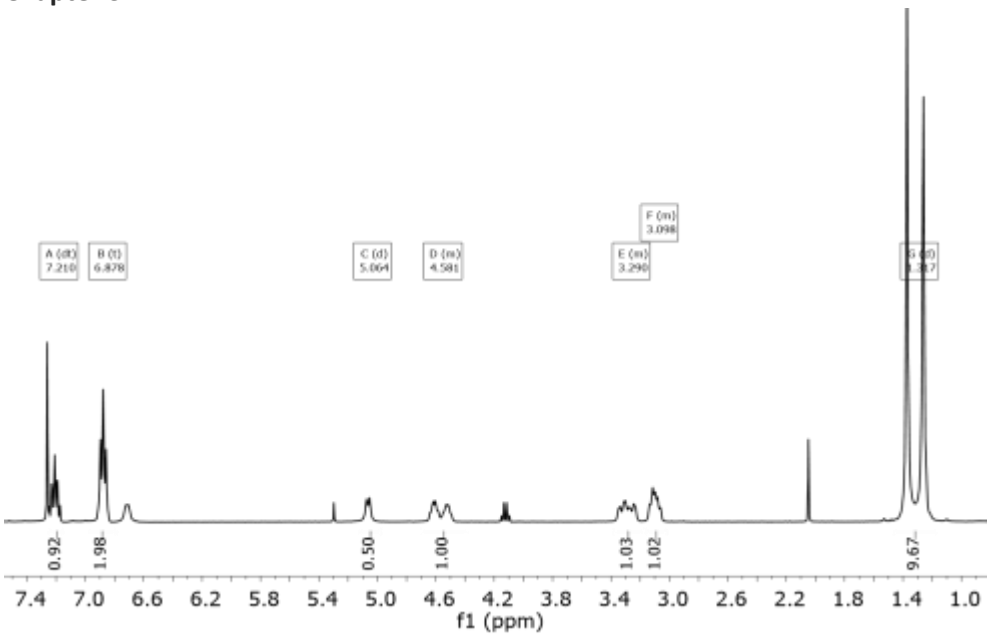
To a solution 0f 2,6-difluoro-phenylalanine *HCl (0.5g, 2.1 mmol) in THF/H_2O 1:1 (8.5 mL) was added Di-tertButyl dicarbonate (0.5 g, 2.3 mmol) and sodium hydroxide (0.278 g, 6.95 mmol). The reaction was stirred at room temperature for 16 hours. The THF was removed in vacuum and DCM was added to the reaction flask. 10% HCl was added drop wise to this solution while stirring vigorously until the precipitate ceased (~PH=4). The reaction mixture was extracted with CH_2Cl_2 (3x150mL). The organic layer was then dried over anhydrous sodium sulphate (Na_2SO_4), filtered and concentrated in vacuum to afford **Boc (2,6F-Phe)-OH** as a white solid (0.6058 g, 2.01 mmol, 96 % yield).

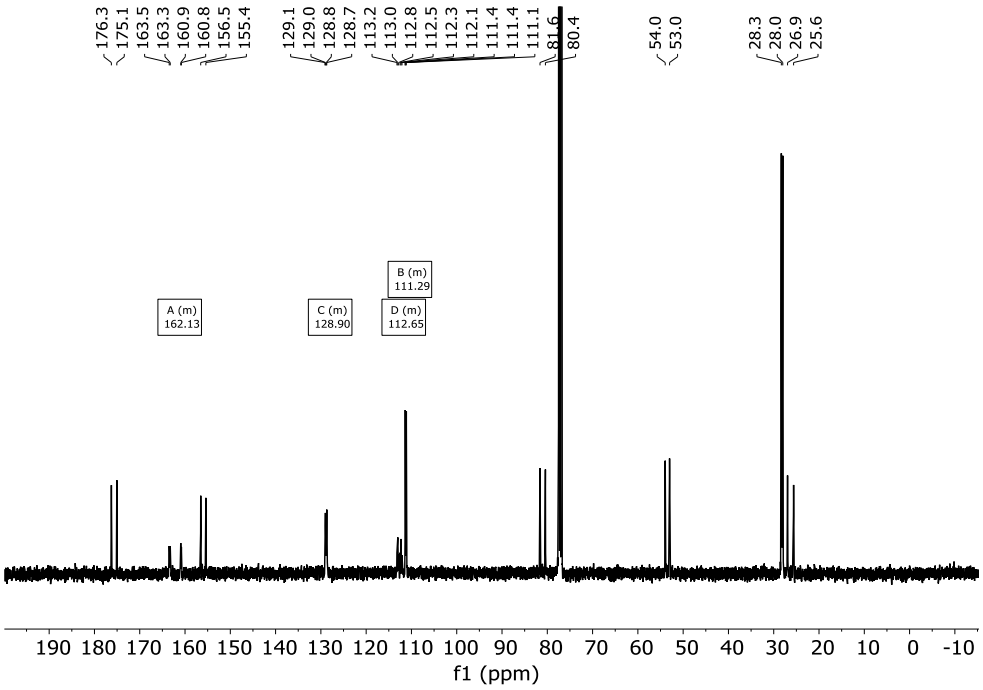
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.21 (m, 3H), 6.88 (t_{ap}, *J* = 7.7 Hz, 3H), 5.06 (d, *J* = 8.5 Hz, 3H), 4.68 (m, 3H), 3.29 (m, 3H), 3.10 (m, 3H), 1.32 (m, 9H)

^{13}C NMR (101 MHz, $CDCl_3$) (most of the C give two signals due to the presence of rotámers in the solution) δ 176.3/175.1, 163.4/160.9 (m, F-C_{Arr} coupling),156.5/155.4, 128.9 (m, F-C_{Arr} coupling), 112.7 (m, F-C_{Arr} coupling), 111.4/111.1 (m, F-C_{Arr} coupling), 81.6/80.4, 50.4/53.0, 28.3/28.0, 29.9/25.6.

^{19}F NMR (376 MHz, $CDCl_3$) δ -114.18 (dt, *J* = 61.2, 6.8 Hz). Spectra multiplicity due to H coupling.

HRMS (ESI-TOF) m/z [Boc (2,6F-Phe)-OH - H]$^+$ Calc: 300.1053, found: 300.1022


A (dt) 7.210
B (t) 6.878
C (d) 5.064
D (m) 4.581
E (m) 3.290
F (m) 3.098
0.92
1.98
0.50
1.00
1.03
1.02
9.67
7.4 7.0 6.6 6.2 5.8 5.4 5.0 4.6 4.2 3.8 3.4 3.0 2.6 2.2 1.8 1.4 1.0
f1 (ppm)


176.3
175.1
163.5
163.3
160.9
160.8
156.5
155.4
129.1
129.0
128.8
128.7
113.2
113.0
112.8
112.5
112.3
112.1
111.4
111.4
111.1
81.6
80.4
54.0
53.0
28.3
28.0
26.9
25.6
A (m) 162.13
C (m) 128.90
B (m) 111.29
D (m) 112.65
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

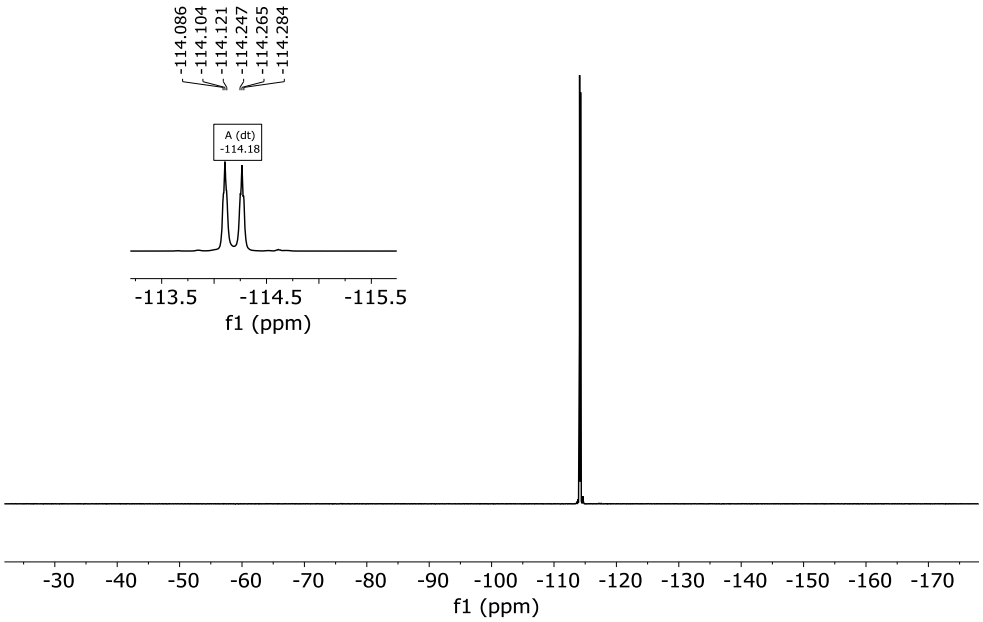


Figure 3.43. ^{1}H-NMR (400 MHz, CD_3Cl), ^{13}C-NMR (101 MHz, CD_3Cl) and ^{19}F NMR (376 MHz, $CDCl_3$) spectra of **Boc (2,6F-Phe)-OH.**

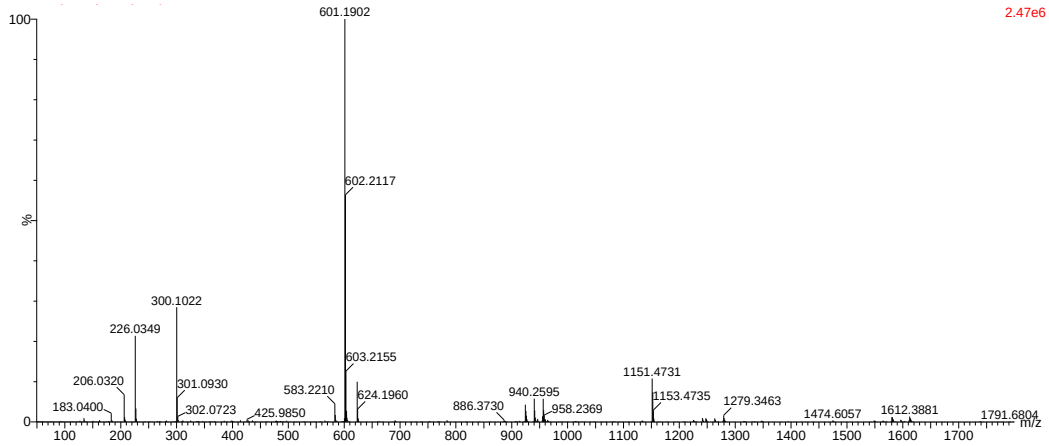


Figure 3.44. HRMS (ESI+) experimental spectrum of **Boc (2,6F-Phe)-OH.**

2l. tri-tert-butyl ((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(3-(2,6-difluoro-phenyl)-1-oxopropane-1,2-diyl))tricarbamate

The protected amino acid (0.509mg, 1.69 mmol) was dissolved in dry DMF (3 mL). In a separate flask, EDC·HCl (0.393 mg, 2.05 mmol) and HoBt (0.313 mg, 2.05 mmol) were

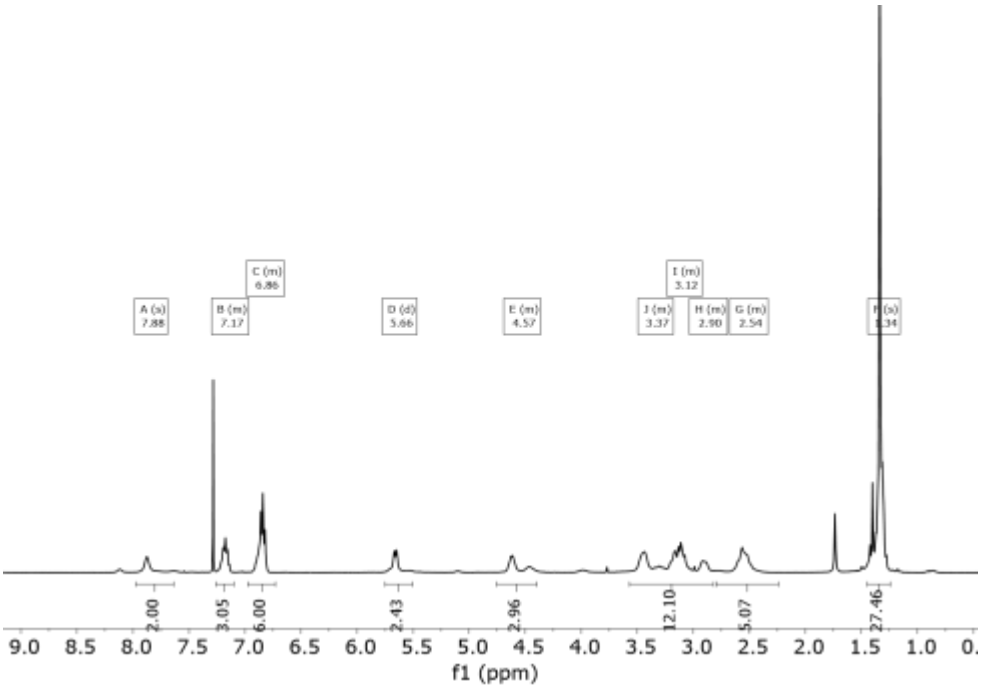
dissolved in dry DMF and added to the amino acid solution and diisopropylethylamine (DIPEA, 1 mL, 6.15 mmol) was also added to the solution. The reaction mixture was stirred under nitrogen at room temperature of 10 minutes to allow the activation of the acid. Then Tris(2-aminoethyl)amine (0.08 mL, 0.512 mmol) was added. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5 % w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using 97:3 DCM:MeOH to give 0.350 mg of **2l** (0.351 mmol, 69% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.88 (s, 3H), 7.17 (m, 3H), 6.86 (m, 6H), 5.66 (d, *J* = 8.8 Hz, 3H), 4.57 (m, 3H), 3.37 (m, 3H), 3.12 (m, 6H), 2.90 (m, 3H), 2.54 (m, 6H), 1.34 (s, 27H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 172.2, 160.7, 155.9, 128.2, 111.1, 110.8, 79.7, 55.5, 54.0, 39.2, 25.9.

^{19}F NMR (376 MHz, $CDCl_3$) δ -113.86.

HRMS (ESI-TOF) m/z [**2l** - H]$^{+}$ Calc: 996.3091 found: 996.4661



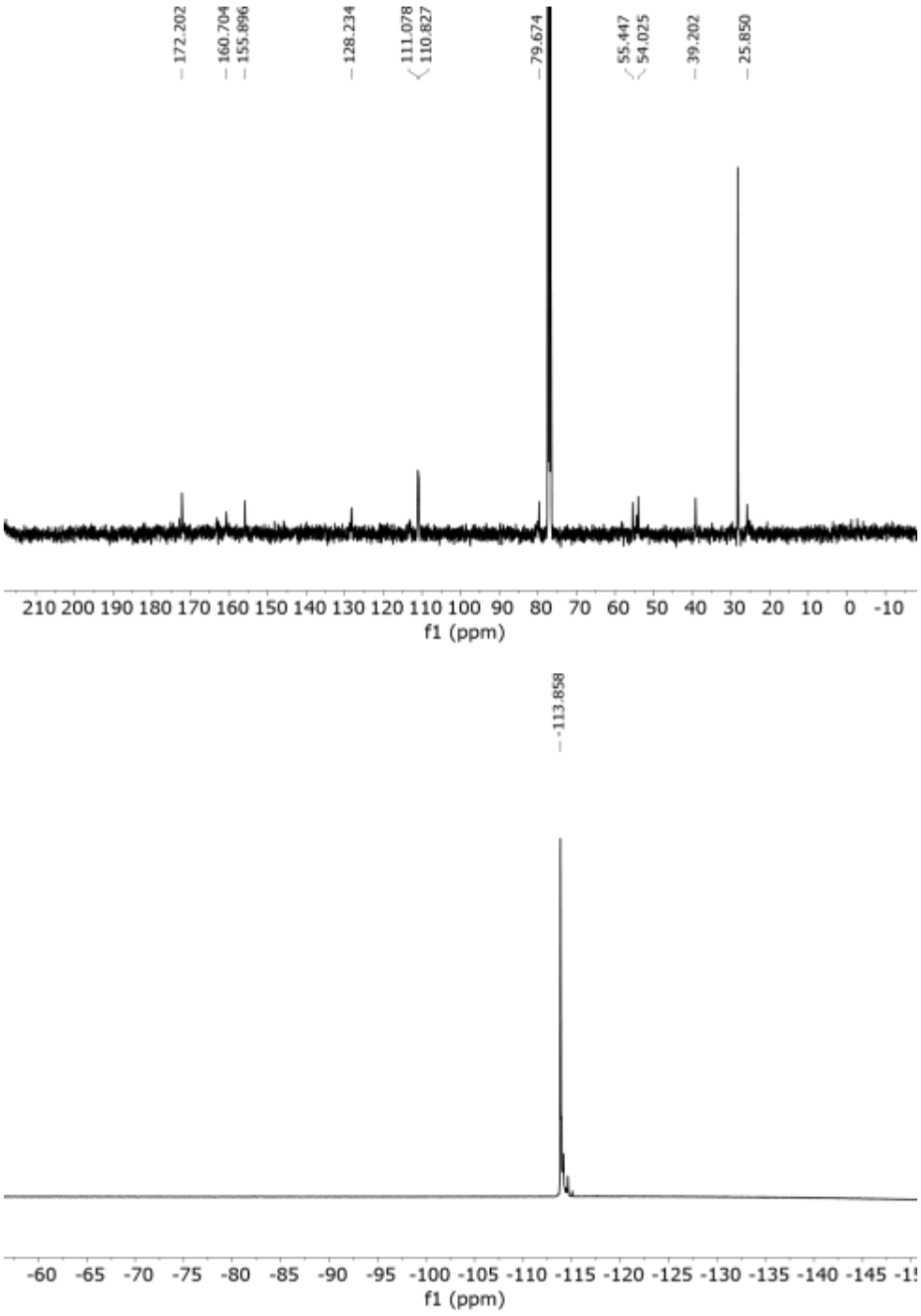


Figure 3.45. ^{1}H-NMR (400 MHz, CD_3Cl), ^{13}C-NMR (101 MHz, CD_3Cl) and ^{19}F NMR (376 MHz, $CDCl_3$) spectra of **2l.**

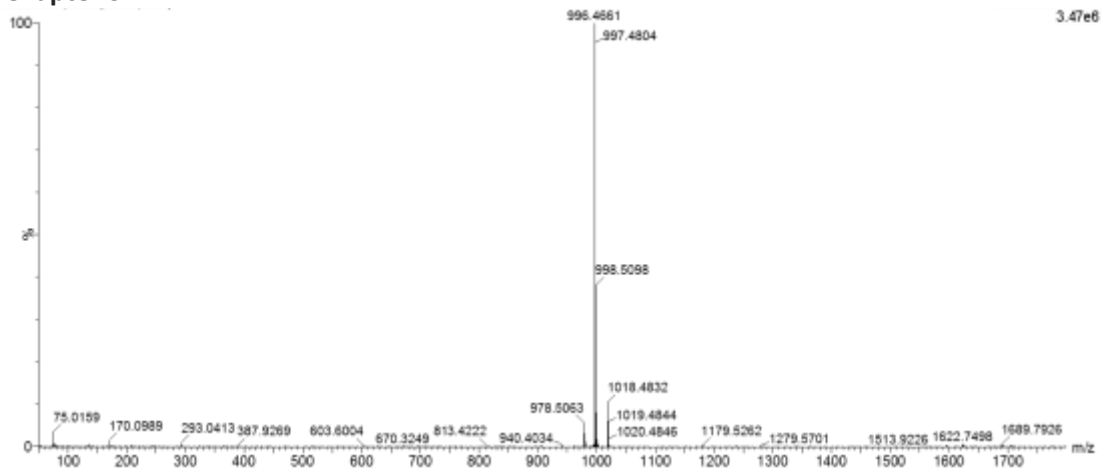


Figure 3.46. HRMS (ESI+) experimental spectrum of **2l.**

3l. (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(2,6-difluorophenyl) pro-panamide)

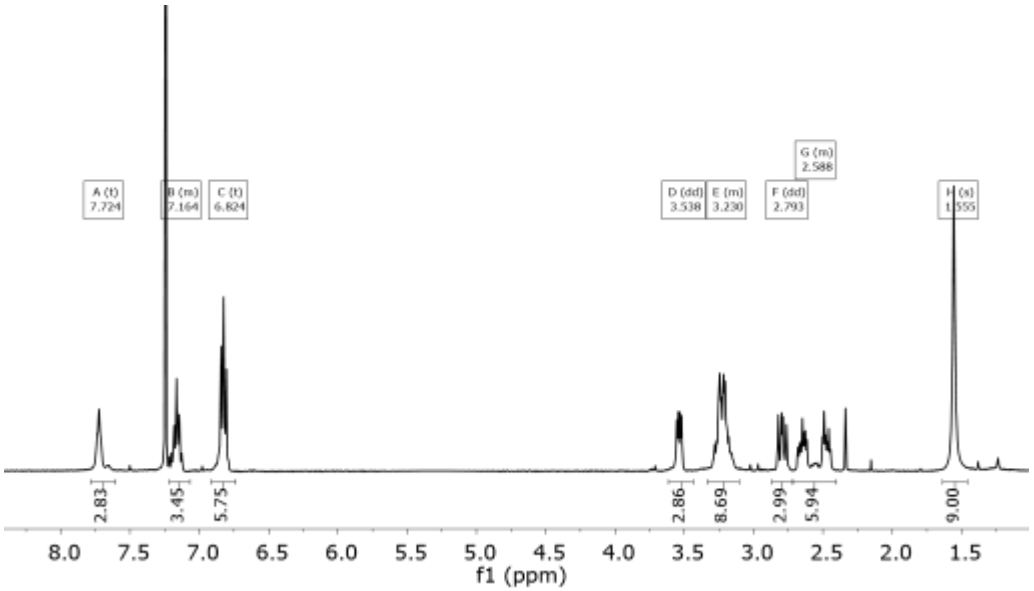
2l(300 mg, 0.301 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3l·4TFA** as a white solid (279 mg, 0.250 mmol, 91% yield). TFA was removed by dissolving the solid in NaOH 1M and extract several times with CH_2Cl_2.

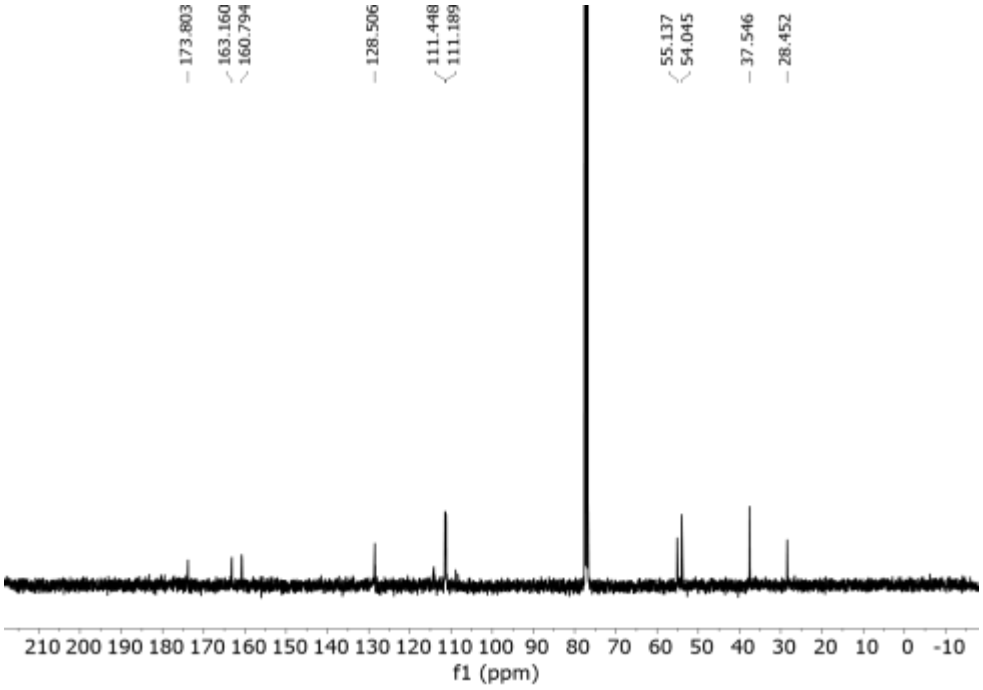
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.72 (t, *J* = 5.5 Hz, 3H), 7.16 (m, 3H), 6.82 (t, *J* = 7.7 Hz, 6H), 3.538 ((X subsystem from ABX, J_{AX}=4.5, J_{BX}=9.6 Hz, 3H,), 3.32 (m, 6H + A subsystem from ABX, J_{AX}=4.5, J_{AB}=13.9 Hz, 3H), 2.793 (B subsystem from ABX, J_{BX}=9.5 J_{AB}=13.9 Hz, 3H), 2.59 (m, 6H), 1.555 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 173.8, 163.2, 160.8, 128.5, 111.4, 111.2, 55.1, 54.1, 37.6, 28.5.

^{19}F NMR (376 MHz, $CDCl_3$) δ -114.32. Spectra multiplicity due to H coupling.

HRMS (ESI-TOF) m/z [**3l** - H]$^+$ Calc: 696.3091 found: 696.3162


A (t)
7.724
B (m)
7.164
C (t)
6.824
D (dd)
3.538
E (m)
3.230
F (dd)
2.793
G (m)
2.588
H (s)
1.555
2.83
3.45
5.75
2.86
8.69
2.99
5.94
9.00
8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
f1 (ppm)


173.803
163.160
160.794
128.506
111.448
111.189
55.137
54.045
37.546
28.452
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

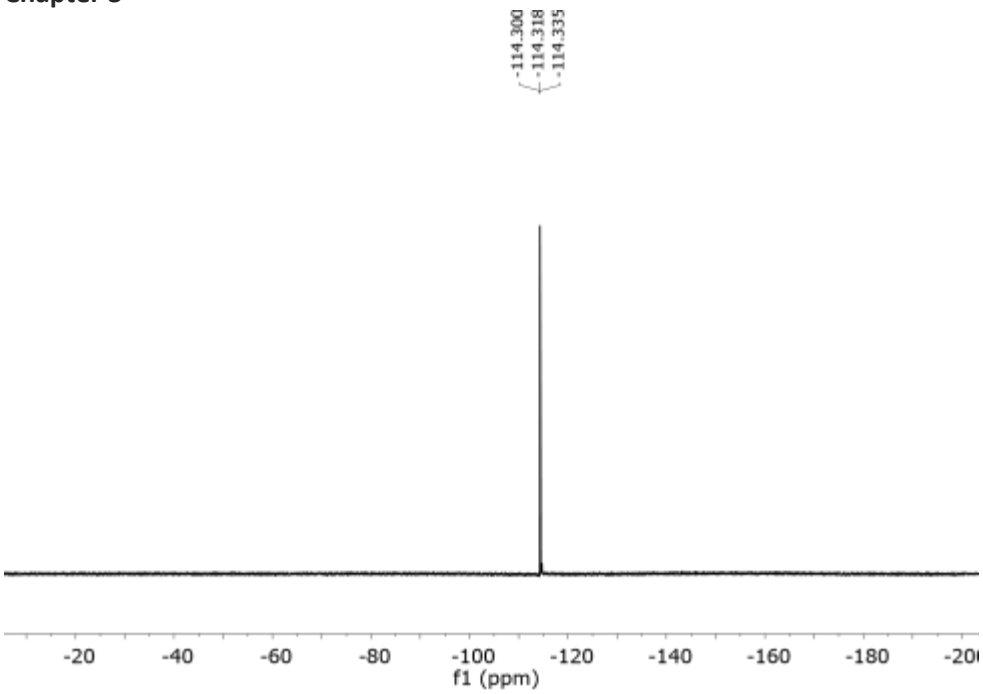


Figure 3.47. ^{1}H-NMR (400 MHz, CD_3Cl), ^{13}C-NMR (101 MHz, CD_3Cl) and ^{19}F NMR (376 MHz, $CDCl_3$) spectra of **3l.**

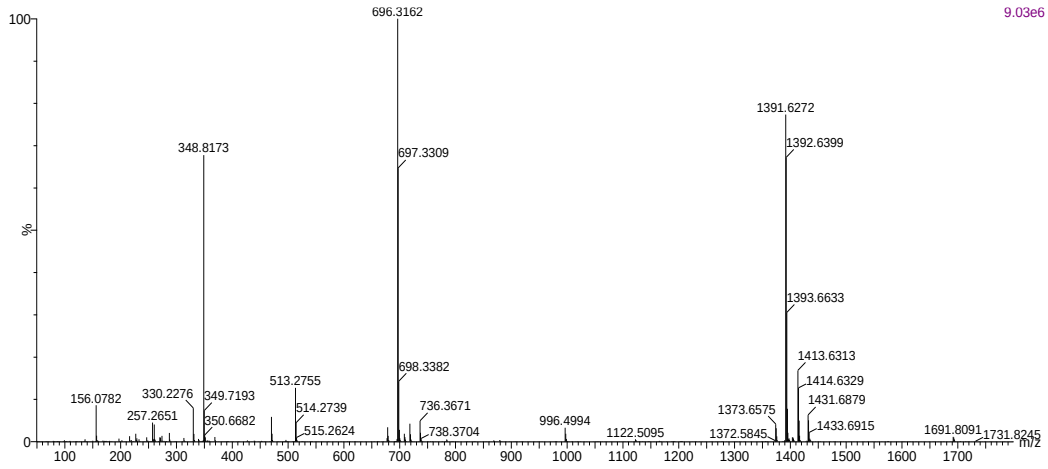
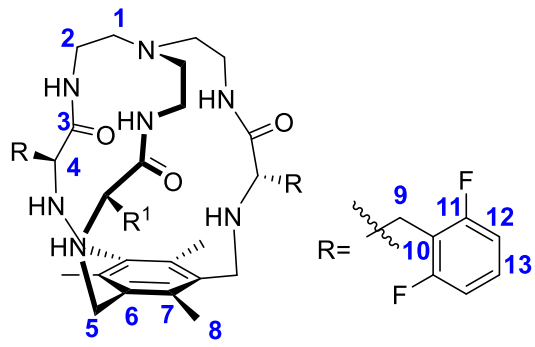


Figure 3.48. HRMS (ESI+) experimental spectrum of **3l.**

1l

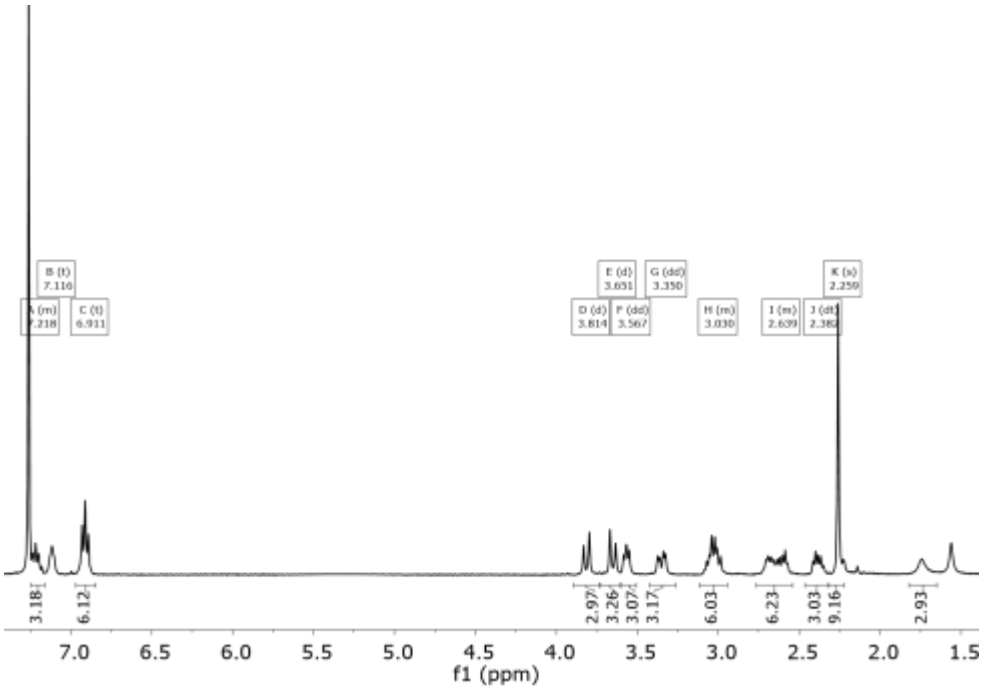
3l (150 mg, 0.150 mmol) was dissolved in ACN (50 mL). Tetrabuthylamonium chloride (30 mg, 0.097 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenze ne (86 mg, 0.216 mmol) and potassium carbonate (597 mg, 4.32 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM : MeOH 97 : 3 as eluent to give **1l** as a white solid (97 mg, 0.073 mmol, 53 % yield).



^{13}C NMR (101 MHz, $CDCl_3$) δ 163.2 (C**3**), 160.7 (C**11**), 135.2 (C**6+7**), 128.7 (C**13**), 111.4 (C**12**), 111.2 (C**10**), 62,1 (C**4+1**) 46.4 (C**5**), 40.2 (C**2**), 26.2 (C**9**), 16.0 (C**8**)

^{19}F-{^{1}H} NMR (376 MHz, $CDCl_3$) δ -113.63

HRMS (ESI-TOF) m/z [**1l** - H]$^+$ Calc: 852.4030; found: 852.4113



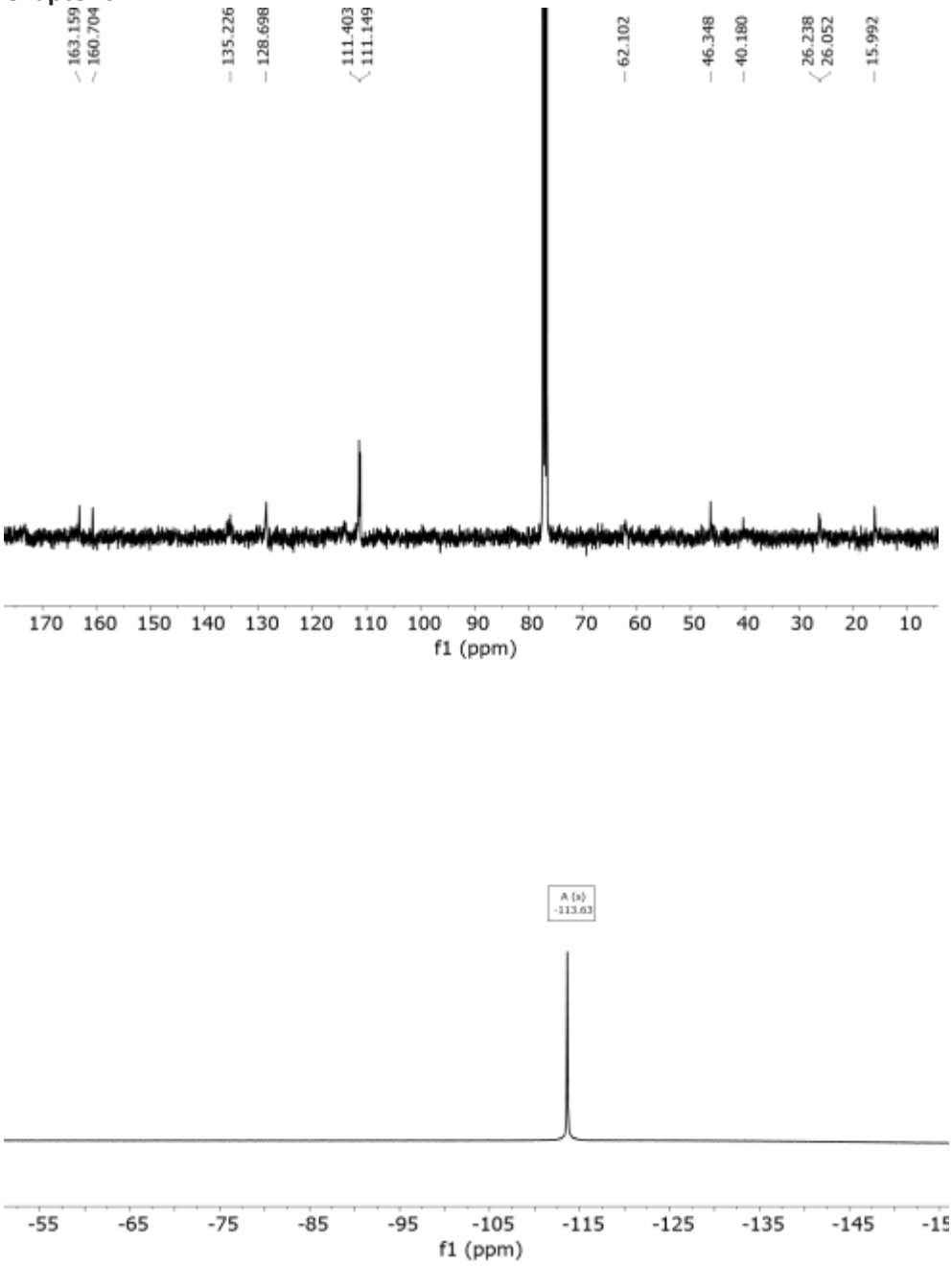


Figure 3.49. ^{1}H-NMR (400 MHz, CD_3Cl), ^{13}C-NMR (101 MHz, CD_3Cl) and ^{19}F-NMR (376 MHz, $CDCl_3$) spectra of **1l.**

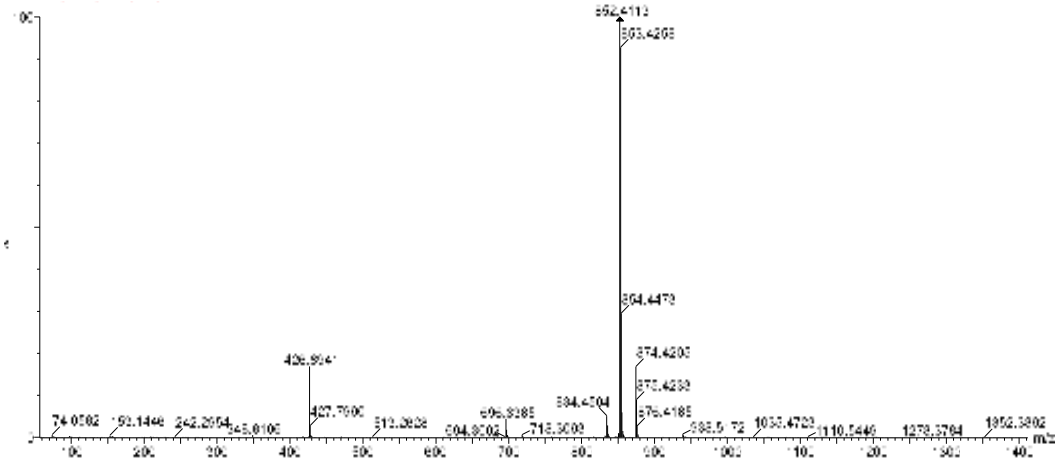


Figure 3.50. HRMS (ESI+) experimental spectrum of **1l.**

3.4.7 Synthesis of 1m

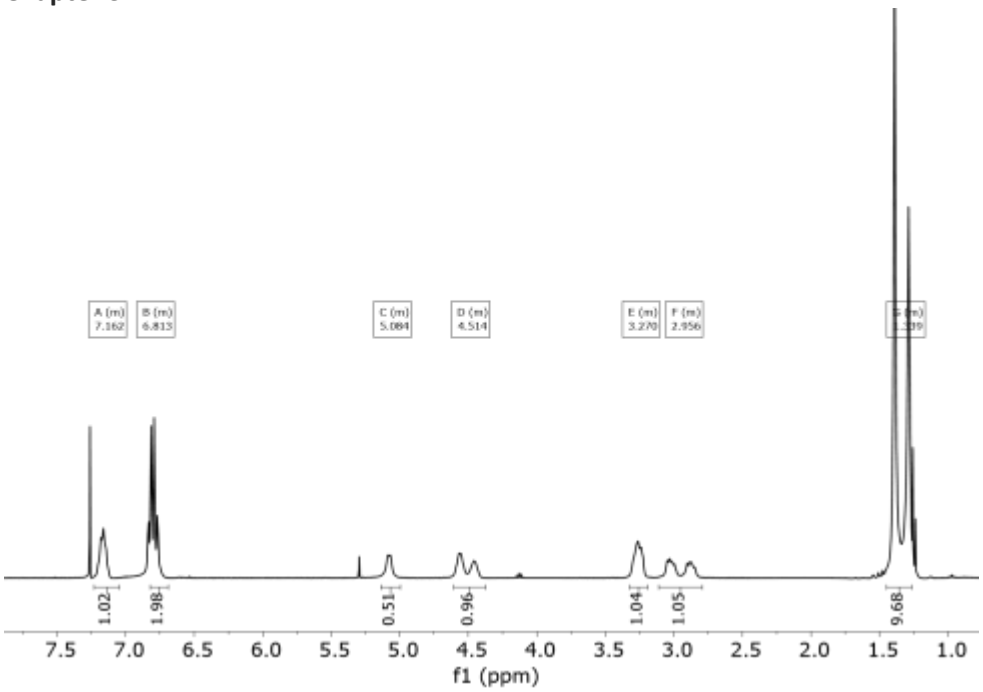
Boc-(2,4F)Phe-OH

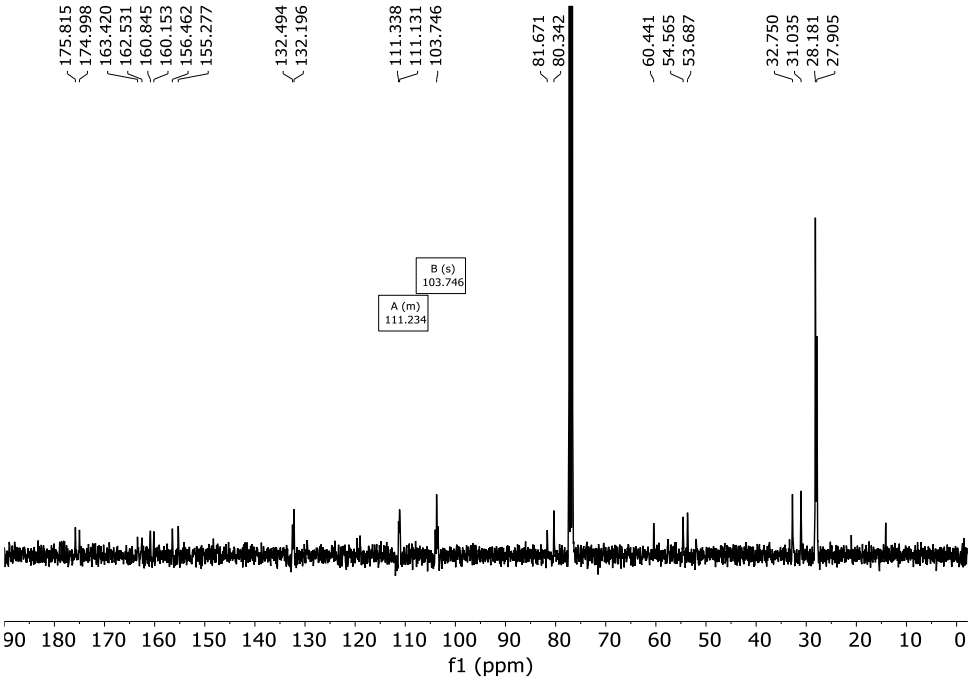
To a solution of 2,4-difluoro-phenylalanine (0.5g, 2.1 mmol) in THF/H_2O 1:1 (8.5 mL) was added Di-tertButyl dicarbonate (0.5 g, 2.3 mmol) and sodium hydroxide (0.184 g, 4.62 mmol). The reaction was stirred at room temperature for 16 hours. The THF was removed in vacuum and DCM was added to the reaction flask. 10% HCl was added drop wise to this solution while stirring vigorously until the precipitate ceased (~PH=4). The reaction mixture was extracted with CH_2Cl_2 (3x150mL). The organic layer was then dried over anhydrous sodium sulphate (Na_2SO_4), filtered and concentrated in vacuum to afford **Boc-(2,4F)Phe-OH** as a white solid (0.606 g, 2.03 mmol, 97% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.16 (m, 1H), 6.81 (m, 2H), 5.08 (m, 1H), 4.57-4.46 (m, 1H, rotamers), 3.27 (m, 1H), 3.03-2.88 (m, 1H, rotamers), 1.39-1.29(m, 9H, rotamers).

^{13}C NMR (101 MHz, $CDCl_3$) δ 175.8/175.0, 163.4/162.5, 160.8/160.15, 156.5/155.3, 132.5/132.2, 111.3 (m, C-F coupling), 103.8 (m, C-F coupling), 81.7/80.3, 60.4, 54.6/53.7, 32.8/31.1, 28.2/27.9. Most of the C give two different signals due to the presence of rotamers.

^{19}F NMR (376 MHz, $CDCl_3$) δ -111.31/-112.87 (d, *J* = 7.4 Hz, 3F), -111.78/-113.06 (d, *J* = 7.1 Hz, 3F). The two F have two signals associated to rotamers.
HRMS (ESI-TOF) m/z [Boc-(2,4F)Phe-OH - H]$^+$ Calc: 300.2824, found: 300.1204


A (m)
7.162
B (m)
6.813
C (m)
5.084
D (m)
4.514
E (m)
3.270
F (m)
2.956
1.02
1.98
0.51
0.96
1.04
1.05
9.68
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
f1 (ppm)


175.815
174.998
163.420
162.531
160.845
160.153
156.462
155.277
132.494
132.196
111.338
111.131
103.746
81.671
80.342
60.441
54.565
53.687
32.750
31.035
28.181
27.905
B (s)
103.746
A (m)
111.234
90 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

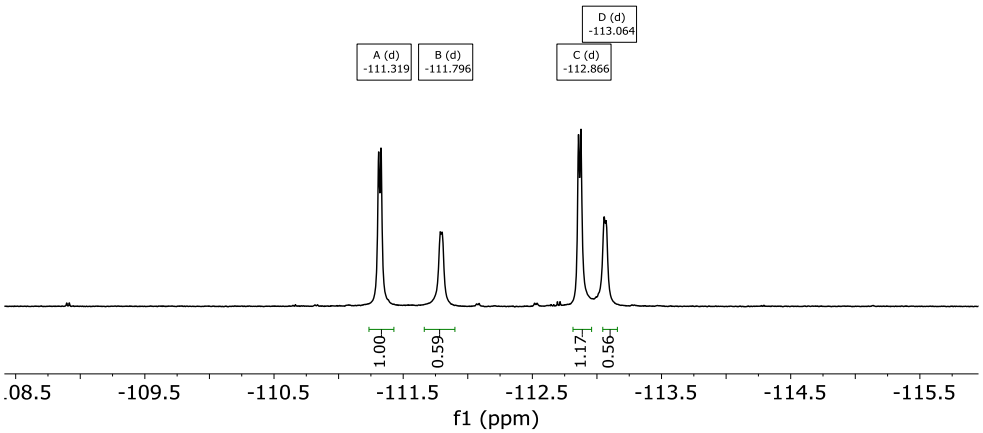


Figure 3.51. ^{1}H-NMR (400 MHz, CD_3OD), ^{13}C-NMR (101 MHz, CD_3OD) and ^{19}F NMR (376 MHz, $CDCl_3$) of **Boc-(2,4F)Phe-OH.**

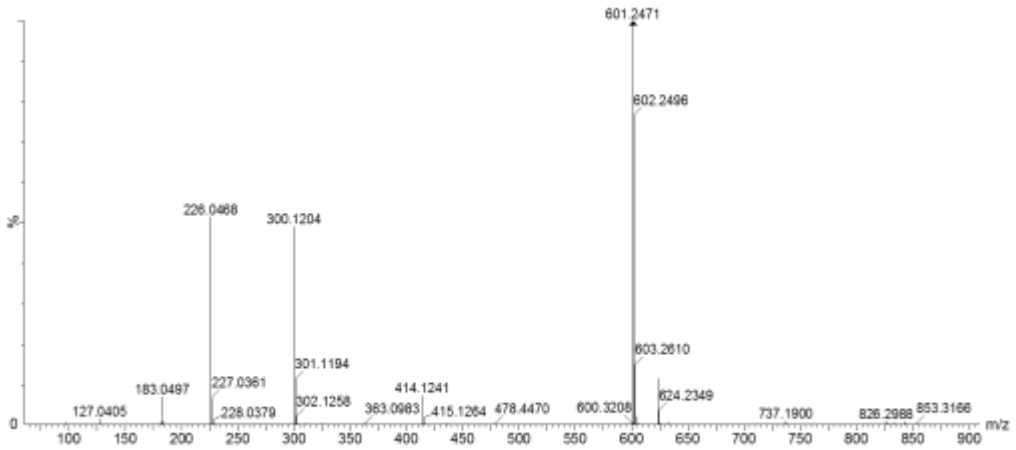


Figure 3.52. HRMS (ESI+) experimental spectrum of **Boc-(2,4F)Phe-OH.**

2m tri-tert-butyl ((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(3-(2,4-difluorophenyl)-1-oxopropane-1,2-diyl))tricarbamate

Boc-(2,4F)Phe-OH (0.50 mg, 1.68 mmols) was dissolved in dry DMF (3 mL). In a separate flasc, EDC·HCl (0.393 mg, 2.05 mmols) and HoBt (0.313 mg, 2.05 mmols) were dissolved in

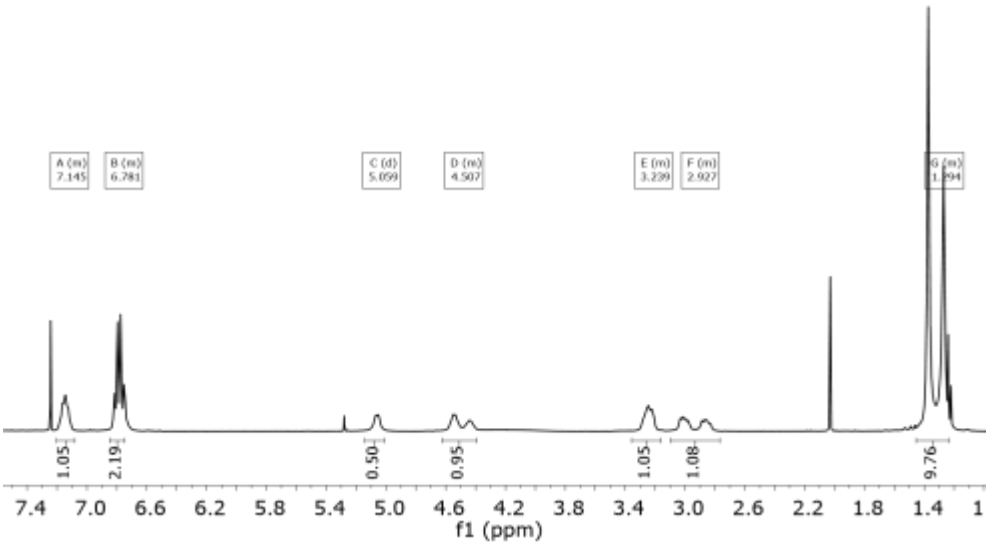
dry DMF and added to the amino-acid solution and diisopropylethylamine (DIPEA, 1 mL, 6.15 mmols) was also added to the solution. The reaction mixture was stirred under nitrogen at room temperature fo 10 minutes to allow the activation of the acid. Then Tris(2-aminoethyl)amine (0.08 mL, 0.512 mmols) was added. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using 97:3 DCM:MeOH to give 0.435 mg of **2m** (0.436 mmol,84% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.15 (m, 3H), 6.78 (m, 6H), 5.06 (d, *J* = 8.0 Hz, 3H), 4.51 (m, 3H), 3.24 (m, 3H), 2.93 (m, 3H), 1.35 (m, 27H (rotamers)).

^{13}C NMR (101 MHz, $CDCl_3$) δ 175.8, 163.4, 160.2, 155.3, 132.2, 111.1, 103.7, 80.3, 60.4, 53.6, 31.0, 28.2.

^{19}F NMR (376 MHz, $CDCl_3$) δ -111.380, -112.876

HRMS (ESI-TOF) m/z [**2m** + H]$^{+}$ Calc: 996,4664, found: 996,4399



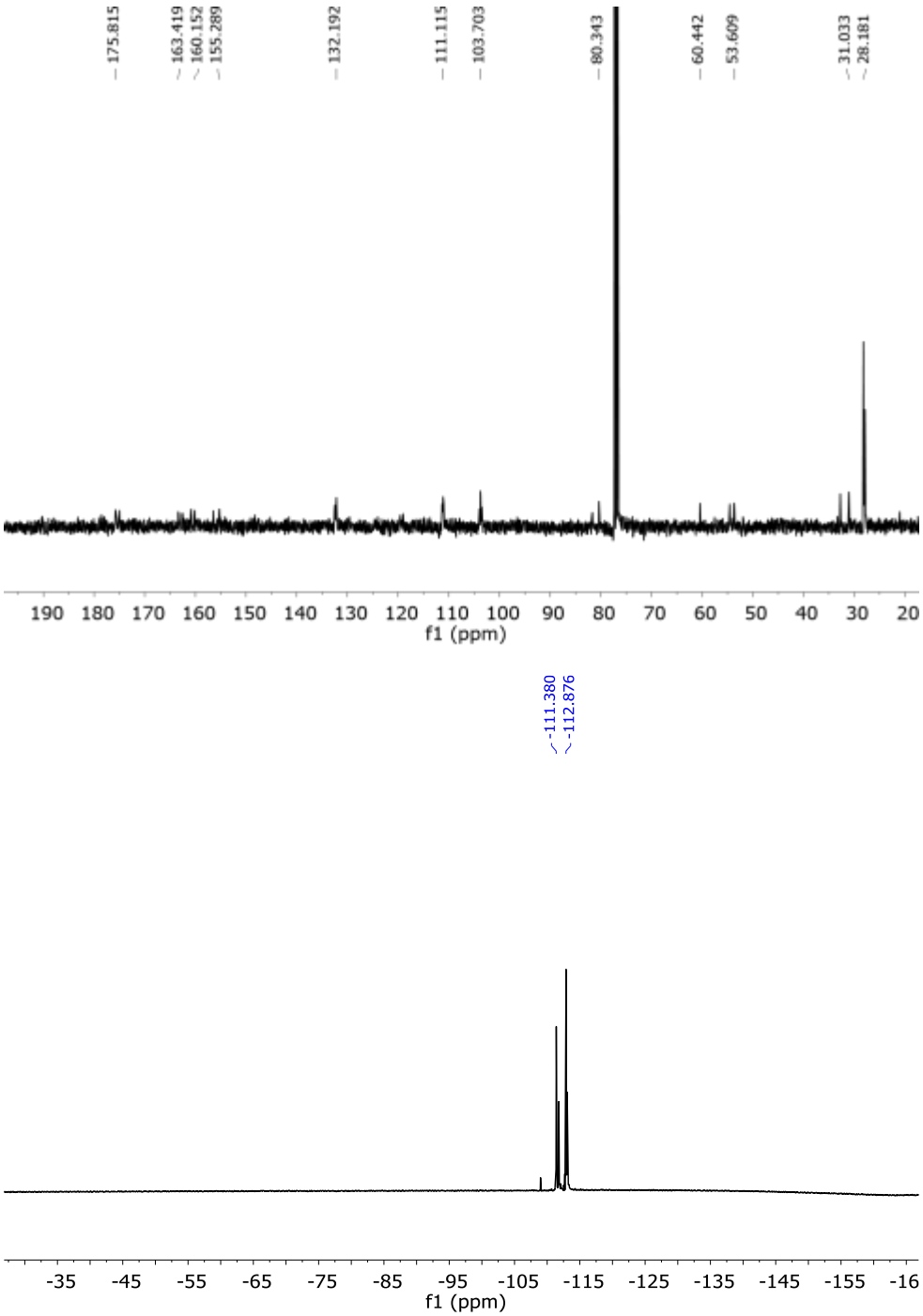


Figure 3.53. ^{1}H-NMR (400 MHz, CD_3Cl) and ^{13}C-NMR (101 MHz, CD_3Cl) and ^{19}F-NMR spectra of **2m.**

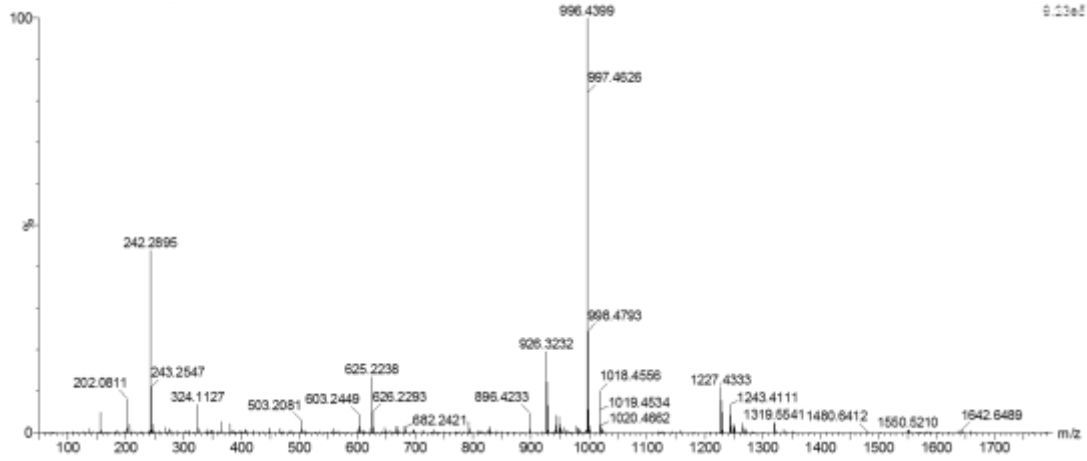


Figure 3.54. HRMS (ESI+) experimental spectrum **of 2m.**

3n. (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(2,4-difluorophenyl) propa namide)

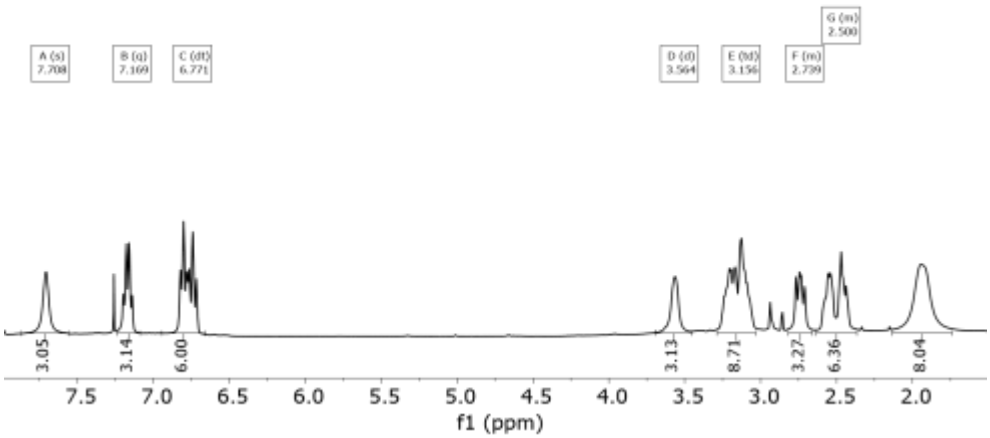
2m (220 mg, 0.250 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3m·4TFA** as a white solid (279 mg, 0.245 mmol, 95% yield). TFA was removed by dissolving the solid in NaOH 1M and extract several times with CH_2Cl_2.

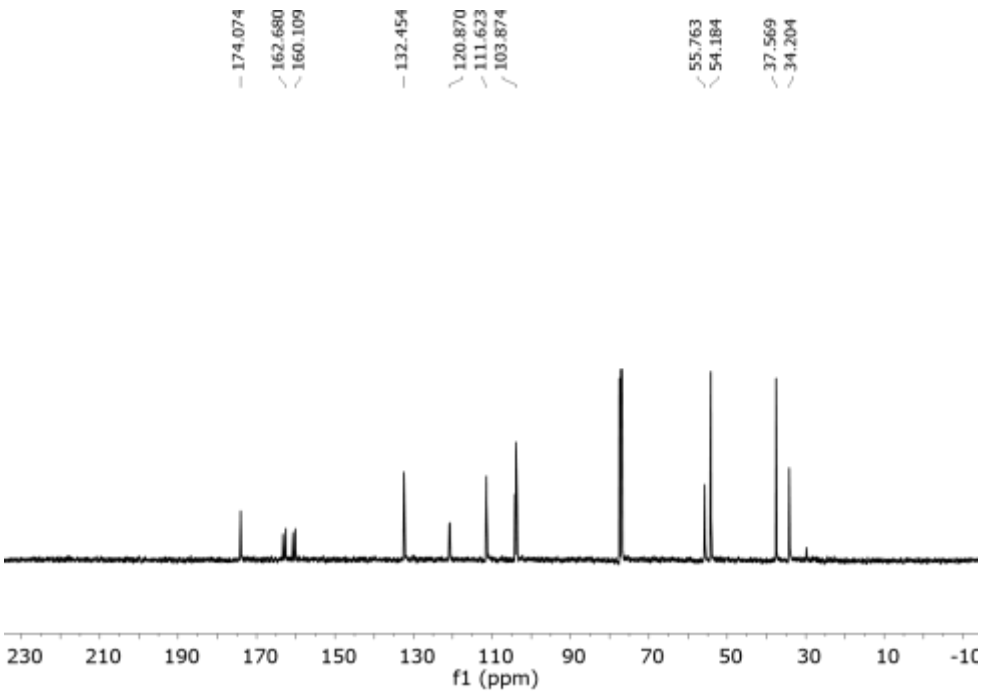
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.71 (s, 3H), 7.17 (tt, *J* = 8.0 ,9.0 Hz, 3H), 6.77 (dt, *J* = 24.8, 9.0 Hz, 6H), 3.54 (m, 3H), 3.16 (m, 9H), 2.74 (m, 3H), 2.5 (m, 6H), 1.95 (m, 9H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 174.1, 162.7, 160.1, 132.5, 120.9, 111.6, 103.9, 55.8, 54.8, 37.6, 34.2.

^{19}F NMR (376 MHz, $CDCl_3$) δ -113.67, -111.85

HRMS (ESI-TOF) m/z [**3m**+ H]$^+$ Calc: 696,3091, found: 696.3078


A (s)
7.708
B (q)
7.169
C (dt)
6.771
D (d)
3.564
E (bd)
3.156
F (m)
2.739
G (m)
2.500
3.05
3.14
6.00
3.13
8.71
3.27
6.36
8.04
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
f1 (ppm)


174.074
162.680
160.109
132.454
120.870
111.623
103.874
55.763
54.184
37.569
34.204
230
210
190
170
150
130
110
90
70
50
30
10
-10
f1 (ppm)

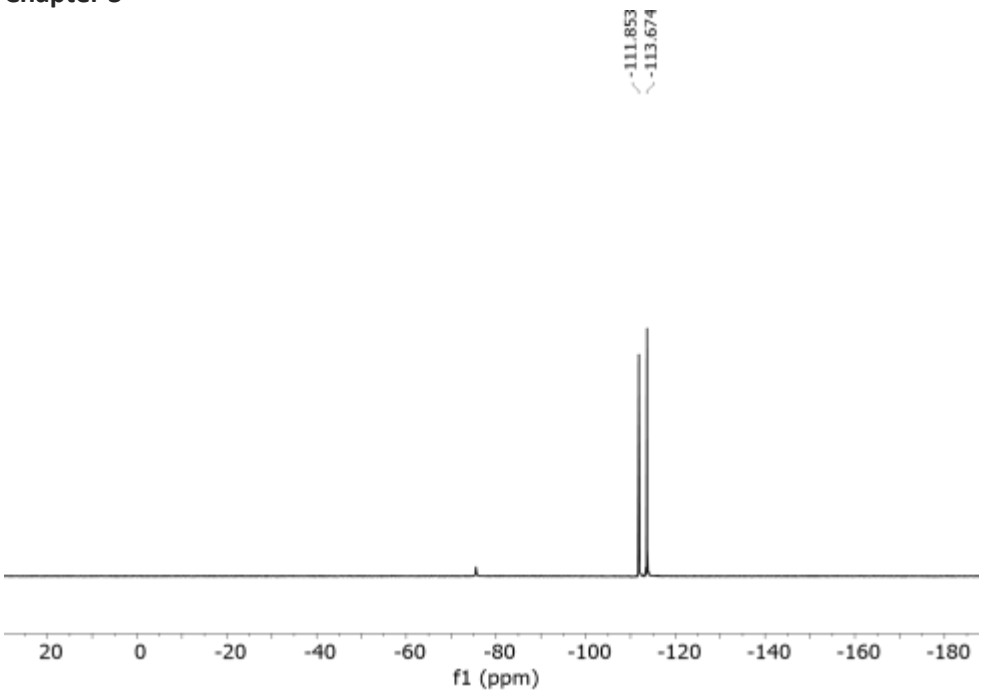


Figure 3.55. ^{1}H-NMR (400 MHz, CD_3Cl) and ^{13}C-NMR (101 MHz, CD_3Cl) and ^{19}F-NMR spectra of **3m.**

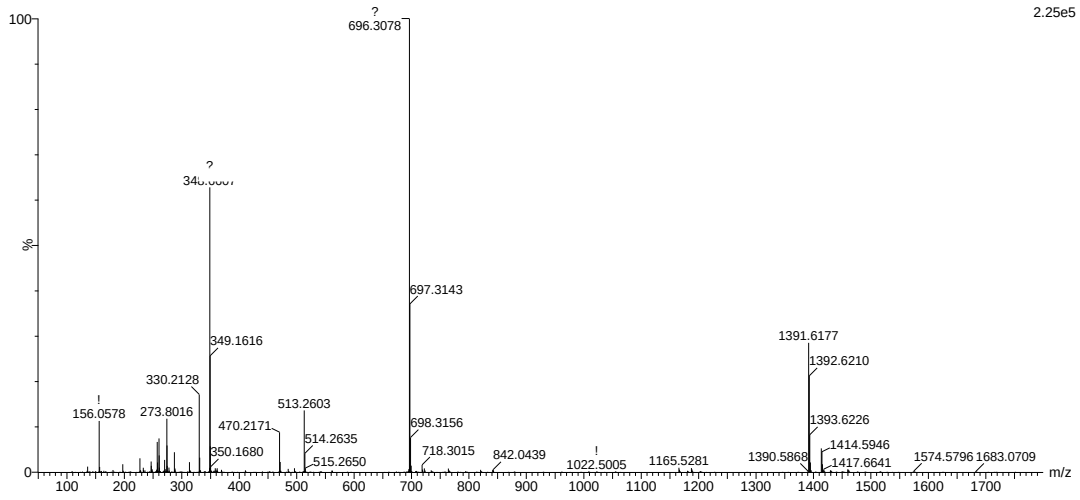
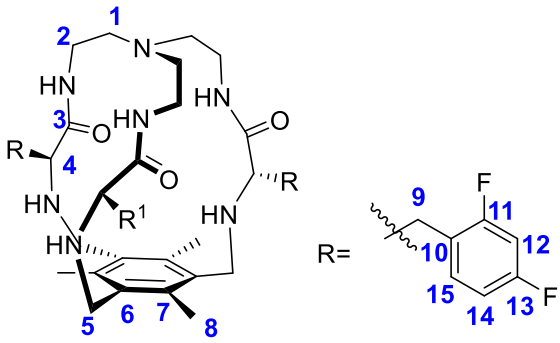


Figure 3.56. HRMS (ESI+) experimental spectrum of **3m.**

1m



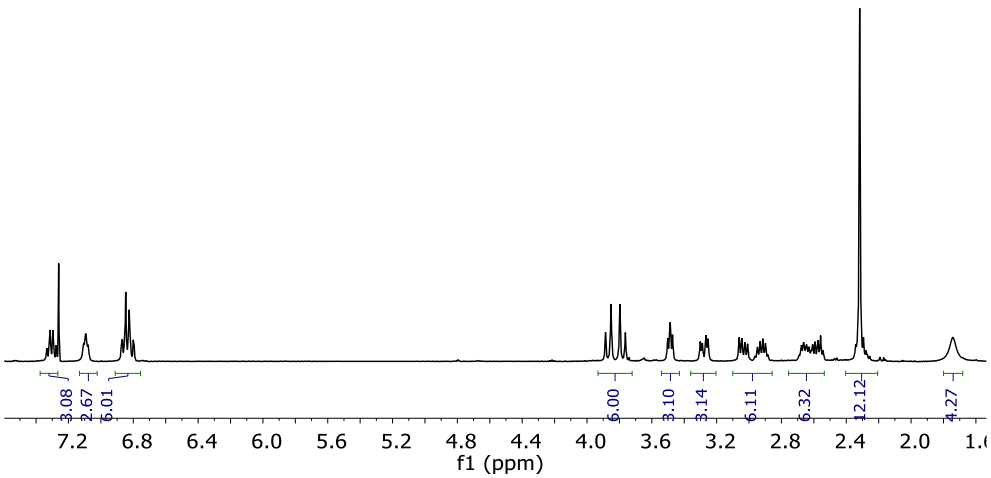
3m (150 mg, 0.21 mmol) was dissolved in ACN (60 mL). Tetrabuthylamonium chloride (30 mg, 0.10 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (86.1 mg, 0.2158 mmol) and potassium carbonate (596 mg, 4.31 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM : MeOH 97 : 3 as eluent to give 1m as a white solid (80 mg, 0.094 mmol, 43 % yield).

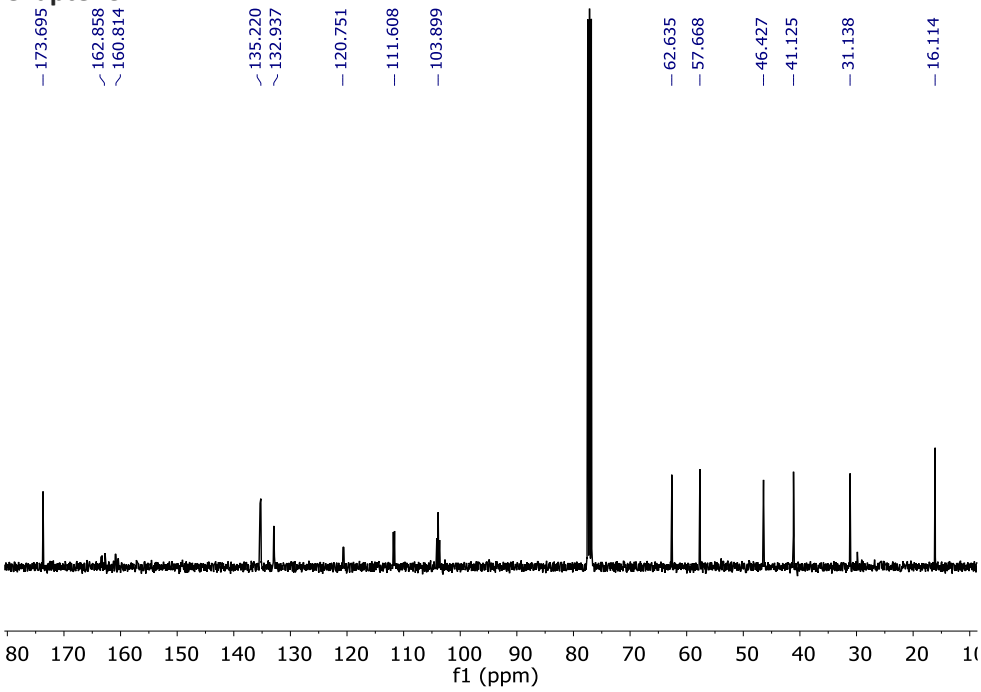
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 (m, 3H, H**15**), 7.1 (t, *J* = 5.8 Hz, 3H, NH_{amide}), 6.83 (m, 6H, H**12**+H**14**), 3.81 (AB_q, δ_A = 3.87, δ_B = 3.78, J_{AB} = 13.6 Hz, 6H, H**5**), 3.49 (X subsystem from ABX, J_{AX} = 5.2, J_{BX} = 6.8 Hz, 3H, H**4**), 3.28 (A subsystem form ABX, J_{AX} = 5.2, J_{AB} = 14.1 Hz, 3H, H**9**), 3.04 (B subsystem from ABX, J_{AB} = 14.1, J_{BX} = 6.8 Hz, 3H, H**9**), 2.93 (m, 3H, H**2**), 2.67 (m, 3H, H**2**), 2.56 (m, 3H, H**1**), 2.31 (s + m, 4H, H**1** + H**8**). 1.74 (s, 3H, NH_{amine}).

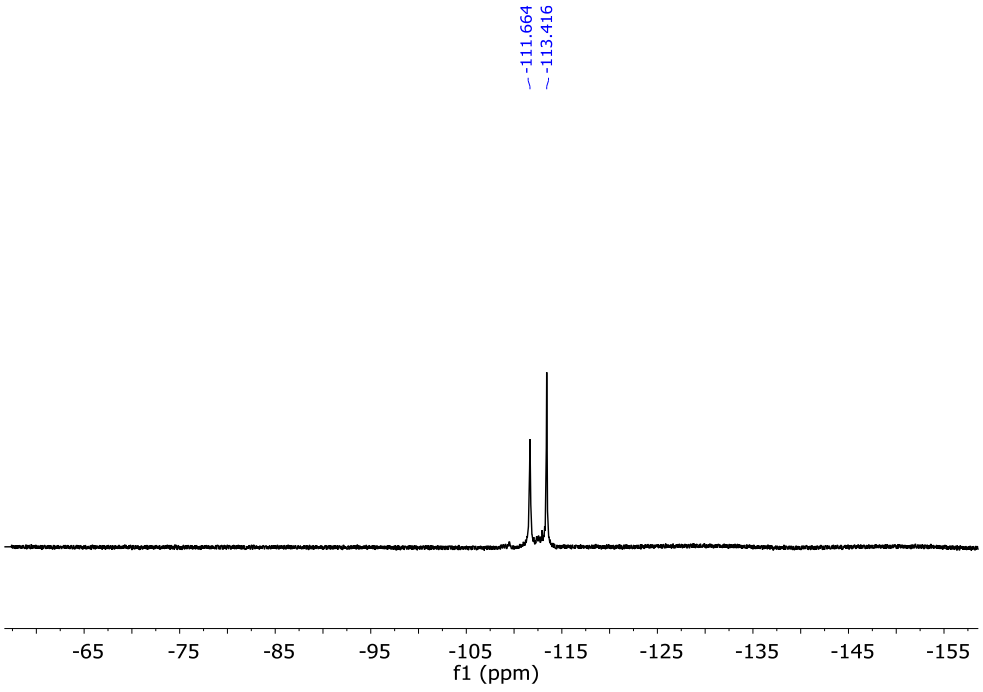
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.7 (C**3**), 162.9 and 160.8 (C**11**+C**13**), 135.2 (C**6**+C**7**), 132.9 (C**15**), 120.8 (C**10**), 111.6 (C**14**), 103.9 (C**12**), 62.6 (C**4**), 57.7 (C**1**), 46.4 (C**5**), 41.1 (C**2**), 31.1 (C**9**), 16.1 (C**8**).

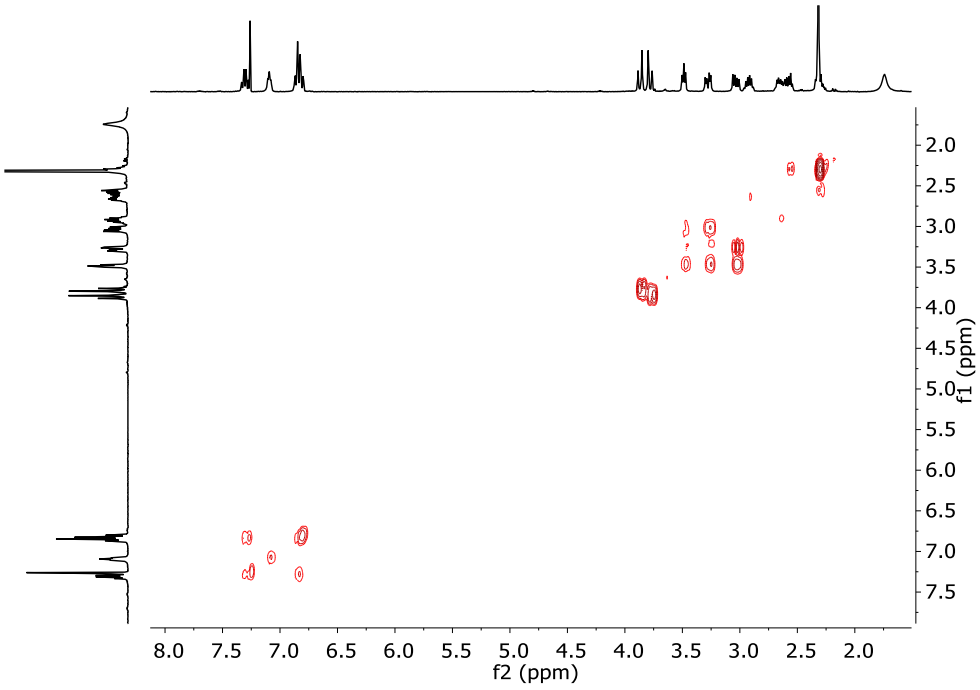
^{19}F NMR (376 MHz, $CDCl_3$) δ -111.664, -113.416.

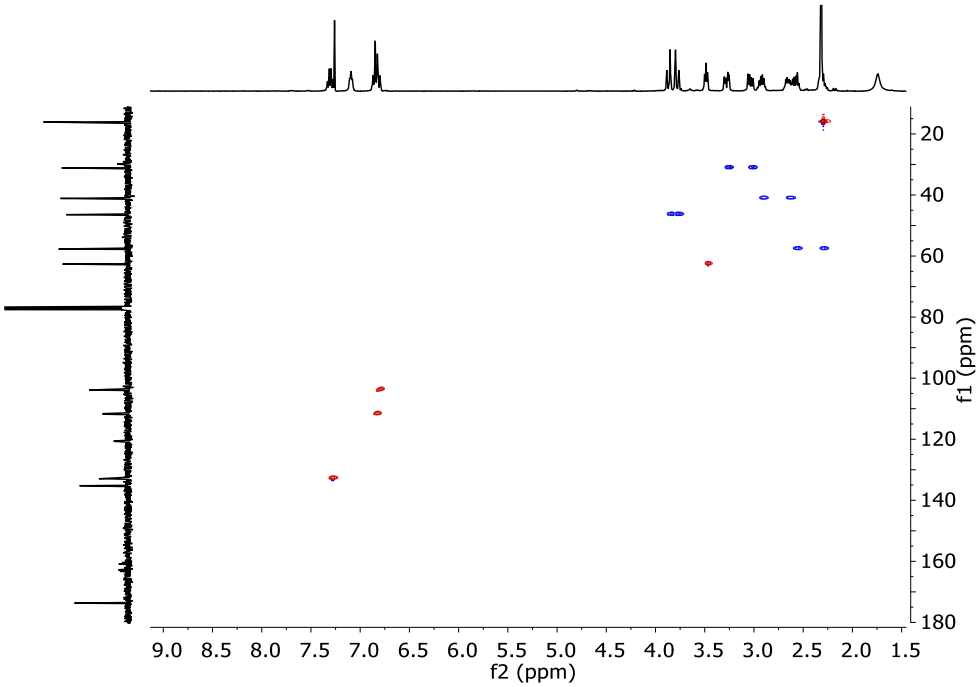
HRMS (ESI-TOF) m/z [**1m**+ H]$^{+}$ Calc: 852.4030, found: 852.3644




173.695
162.858
160.814
135.220
132.937
120.751
111.608
103.899
62.635
57.668
46.427
41.125
31.138
16.114
80 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10
f1 (ppm)


-111.664
-113.416
-65 -75 -85 -95 -105 -115 -125 -135 -145 -155
f1 (ppm)


2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
f1 (ppm)
8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0
f2 (ppm)


20
40
60
80
100
120
140
160
180
f1 (ppm)
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
f2 (ppm)

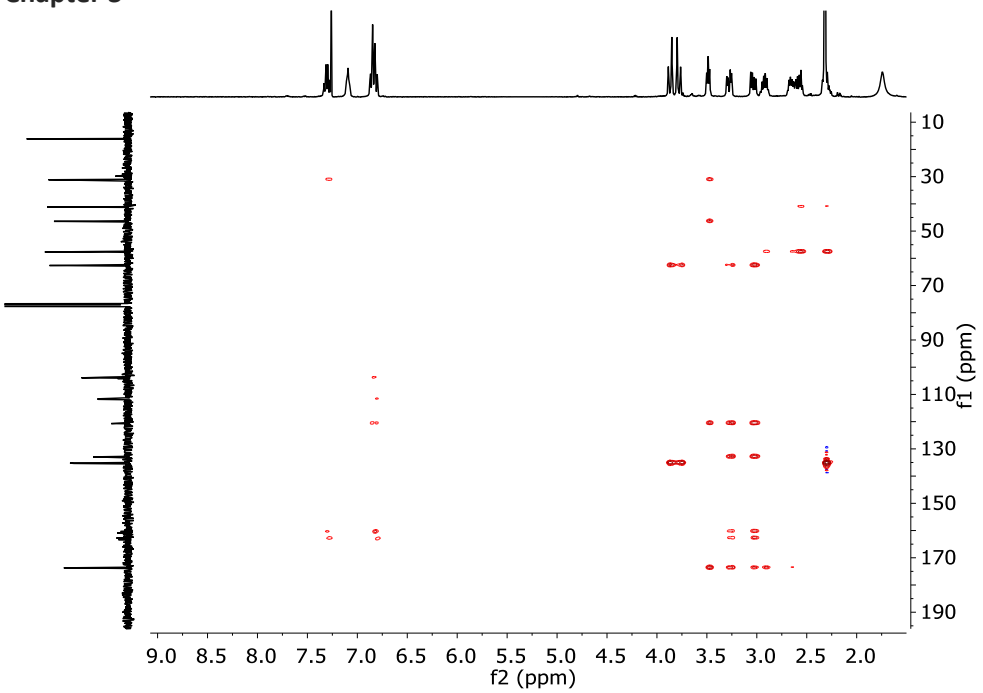


Figure 3.57. ^{1}H-NMR (400 MHz, CD_3Cl), ^{13}C-NMR (101 MHz, CD_3Cl) and ^{19}F-NMR (376 MHz, $CDCl_3$) spectra of **1m.**

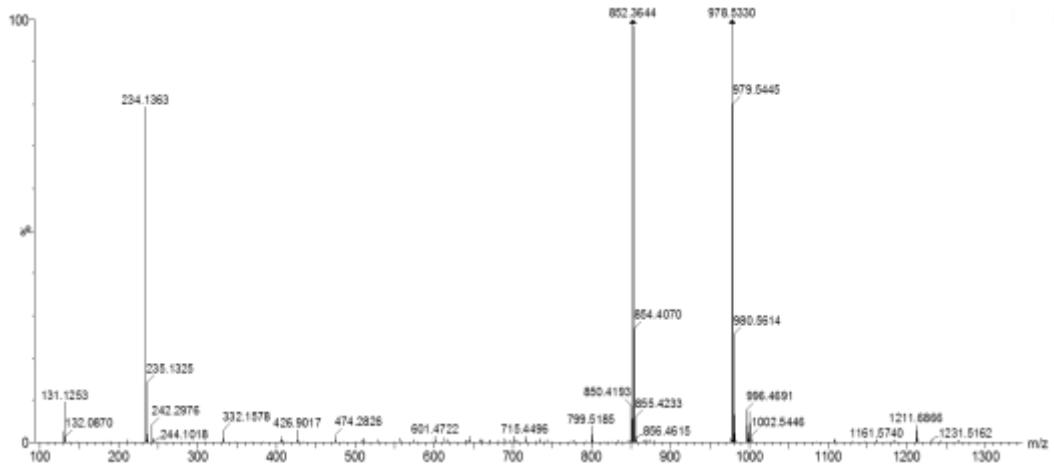


Figure 3.58. HRMS (ESI+) experimental spectrum of **1m.**

3.4.8 Synthesis of 1n

2n. tri-tert-butyl ((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(1-oxo-3-(per-fluorophenyl)propane-1,2-diyl))tricarbamate

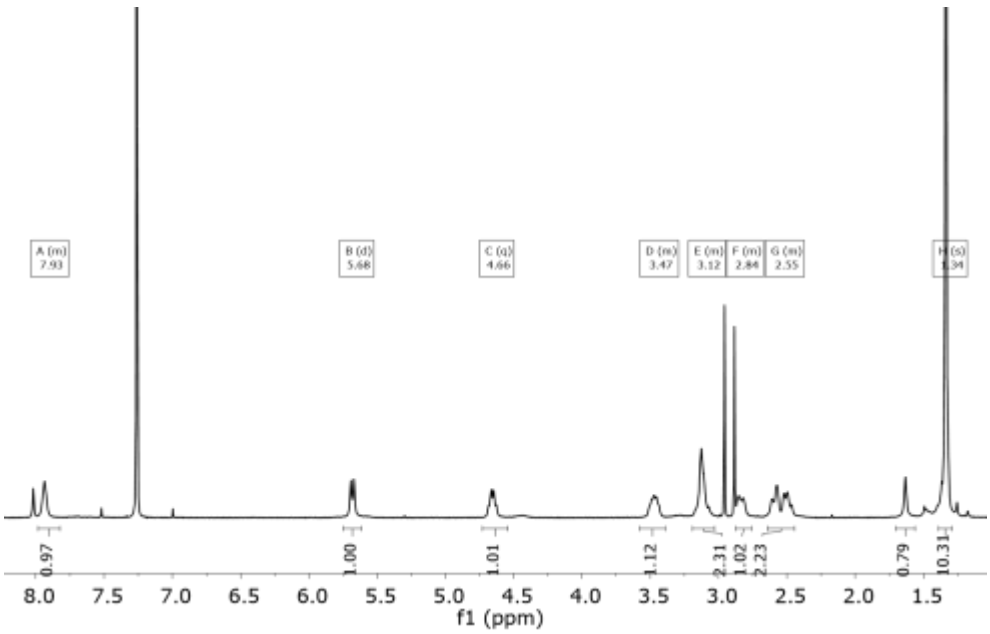
Boc-(Penta-F)Phe-OH (405 mg, 1.14 mmol) was dissolved in dry DMF (4 mL). N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC·HCl O.265 mg, 1.38 mmol) and 1-Hydroxybenzotriazole hydrate (HOBt, 0.21 mg, 1.38 mmol), N,N-diisopropylethylamine (DIPEA, 0.72 mL, 4.15 mmol) and tris(2-aminoethyl)amine (0.051 mL, 0.32 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using 98:2 DCM:MeOH to give 0.390 mg of **2n** (0.330 mmol, 92% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.93 (m, 3H), 5.68 (d, J = 9.1 Hz, 3H), 4.66 (m, 3H), 3.47 (m, 3H), 3.12 (m, 6H), 2.84 (m, 3H), 2.55 (m, 6H), 1.34 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.83, 156.29, 80.53, 55.75, 53.65, 39.41, 28.23, 26.18.

^{19}F NMR (376 MHz, $CDCl_3$) δ -142.18 (dd, *J* = 22.0, 7.7 Hz, 6F),, -156.36 (t, *J* = 20.8 Hz, 3F) , -162.94 (td, *J* = 21.7, 7.6 Hz, 6F)

HRMS (ESI-TOF) m/z [**2n**+ H]$^{+}$ Calc: 1158,3816, found: 1158.4625



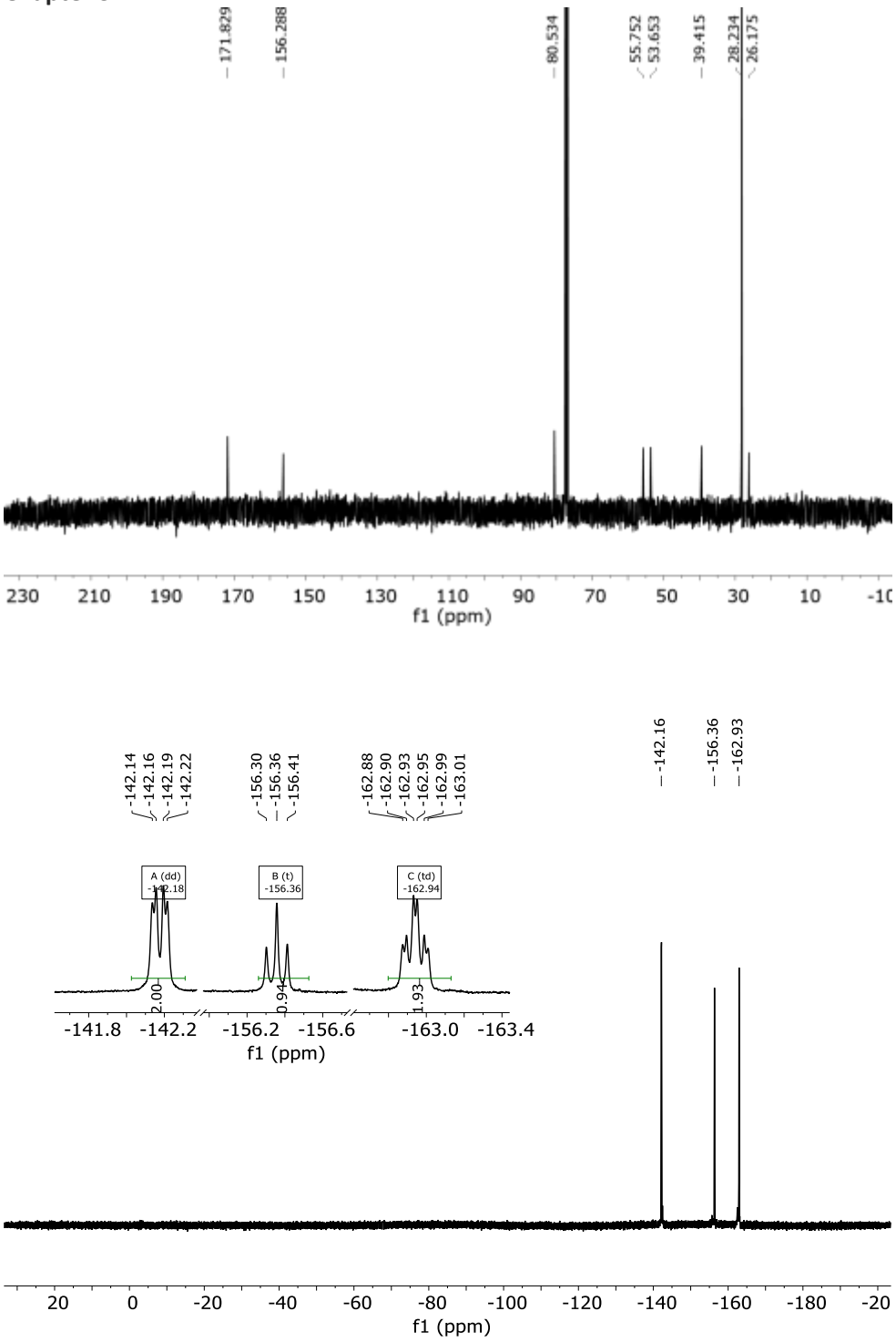


Figure 3.59. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, $CDCl_3$) and ^{19}F NMR (376 MHz, $CDCl_3$) of **2n.**

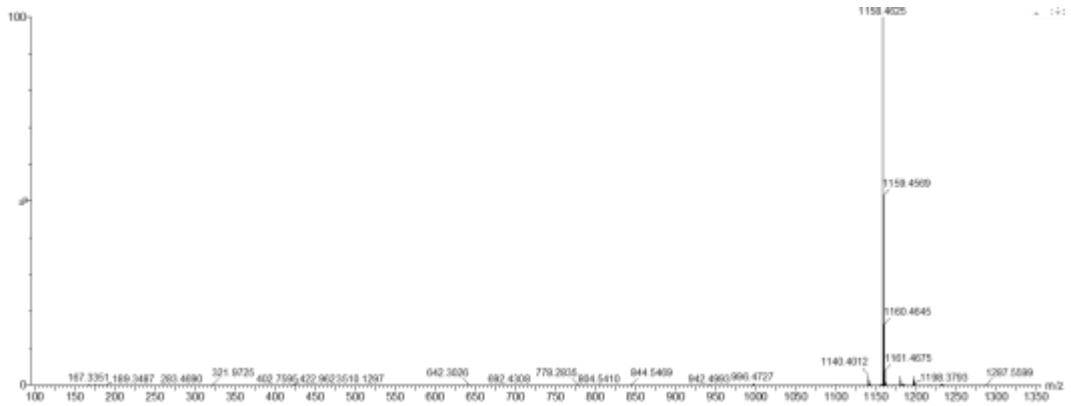


Figure 3.60. HRMS (ESI+) experimental spectrum of **2n.**

3n. (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(perfluorophenyl) propanemide)

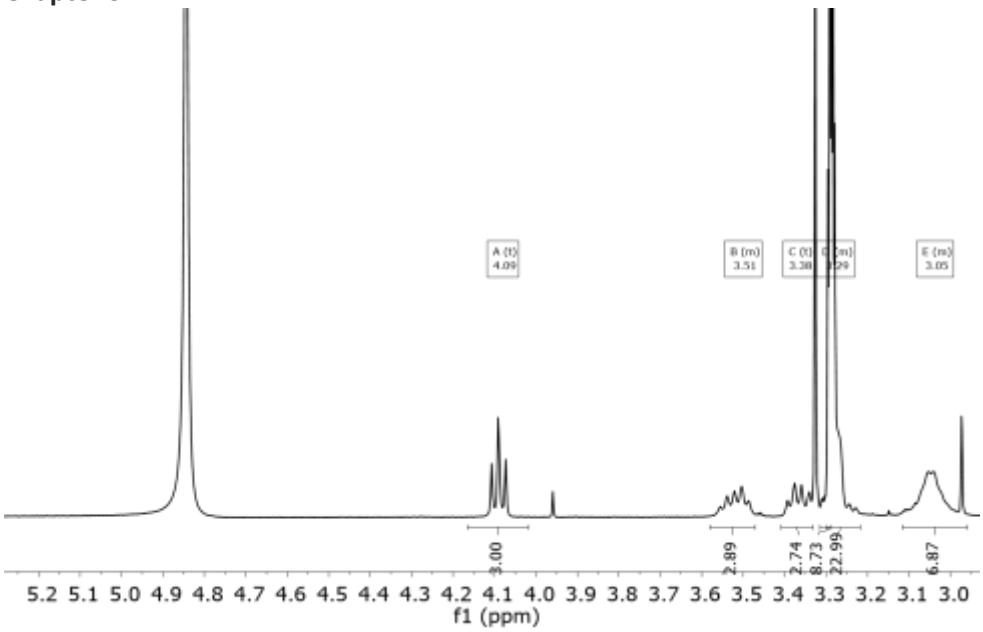
2n (385 mg, 0.31 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.75 mL, 5.05 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording 3n·4TFA as a white solid (395 mg, 0.28 mmol, 90% yield). TFA was removed by dissolving the solid in NaOH 1M and extract several times with CH_2Cl_2.

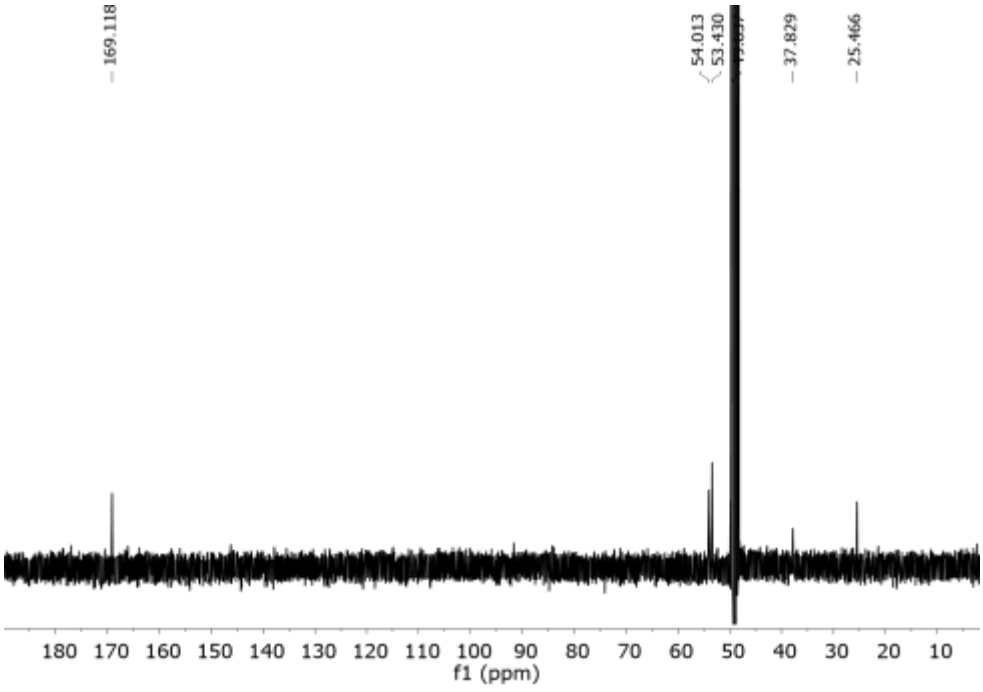
^{1}H NMR (400 MHz, CD_3OD) δ 4.09 (t, J = 6.8 Hz, 3H), 3.51 (m, 3H), 3.38 (t, *J* = 6.8 Hz, 3H), 3.263 (m, 6H), 3.05 (m, 6H).

^{13}C NMR (101 MHz, CD_3OD) δ 169.1, 54.0, 53.4, 49.6, 37.8, 25.5(C_{ar} are not seen in this spectra).

^{19}F –{^{1}H}-NMR (376 MHz, $CDCl_3$) δ -143.311 (dd, *J* = 20.9, 7.0 Hz, 6F)., -157.377 (t, *J* = 19.9 Hz, 3F)., -164.649 (td, *J* = 21.1, 7.4 Hz, 6F).

HRMS (ESI-TOF) m/z [**3n**+ H]$^+$ Calc: 858,2234, found: 858,2258


A (t)
4.09
B (m)
3.51
E (m)
3.05
3.00
2.89
2.74
8.73
22.99
6.87
5.2 5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2 3.1 3.0
f1 (ppm)


169.118
54.013
53.430
37.829
25.466
180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10
f1 (ppm)

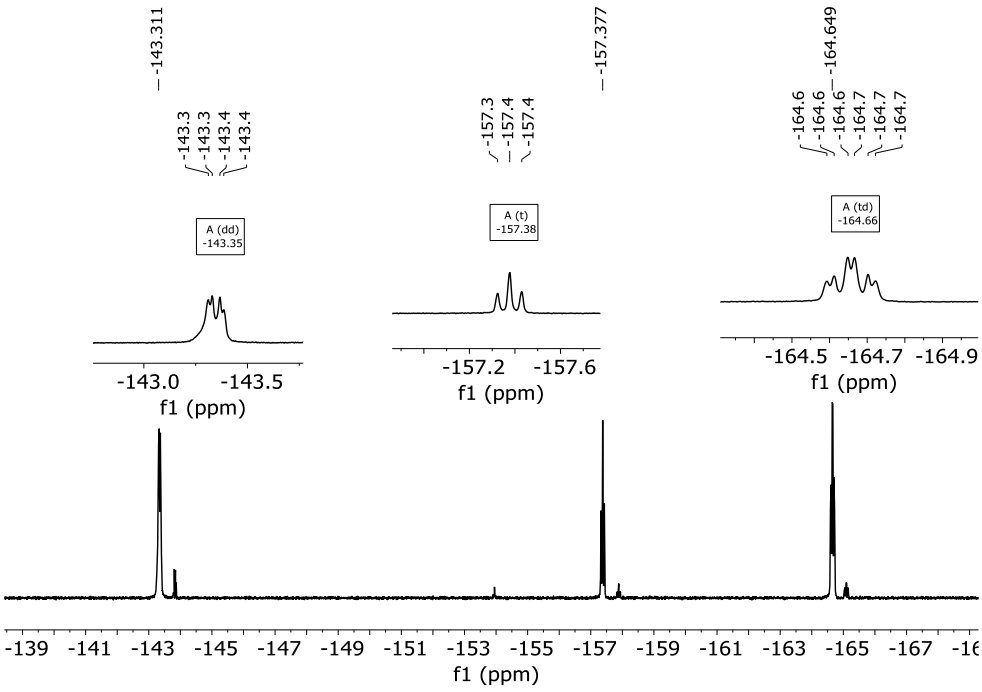


Figure 3.61. ^{1}H-NMR (400 MHz, CD_3OD), ^{13}C-NMR (101 MHz, CD_3OD) and ^{19}F NMR (376 MHz, CD_3OD) of **3n.**

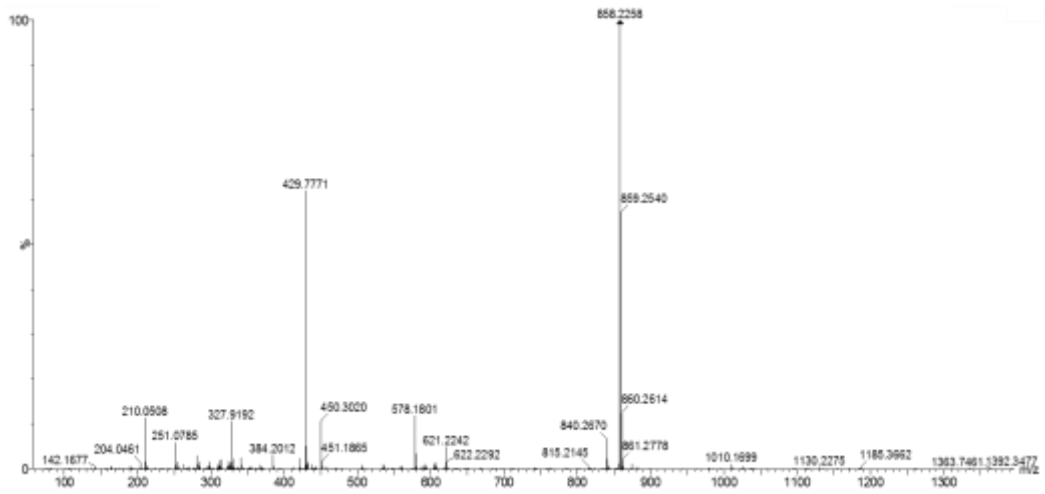
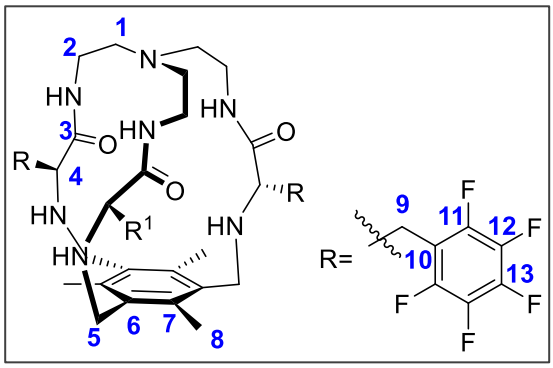


Figure 3.62. HRMS (ESI+) experimental spectrum of **3n.**

1n

3n (205 mg, 0.156 mmol) was dissolved in ACN (60 mL). Tetrabuthylamonium chloride (21 mg, 0.078 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (62.4 mg, 0.156 mmol) and potassium carbonate (432 mg, 3.12 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM:MeOH 97:3 as eluent to give **1n** as a white solid (40 mg, 0.039 mmol, 40% yield).

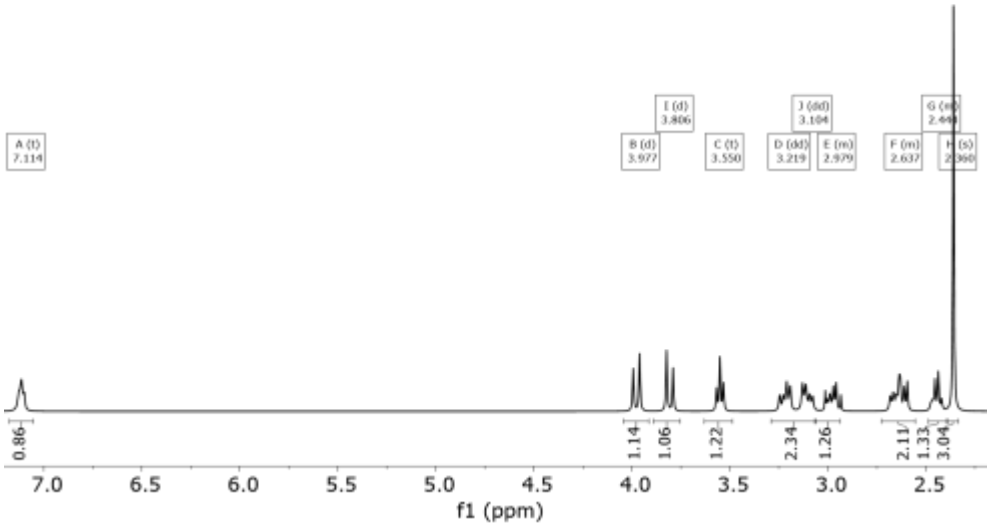


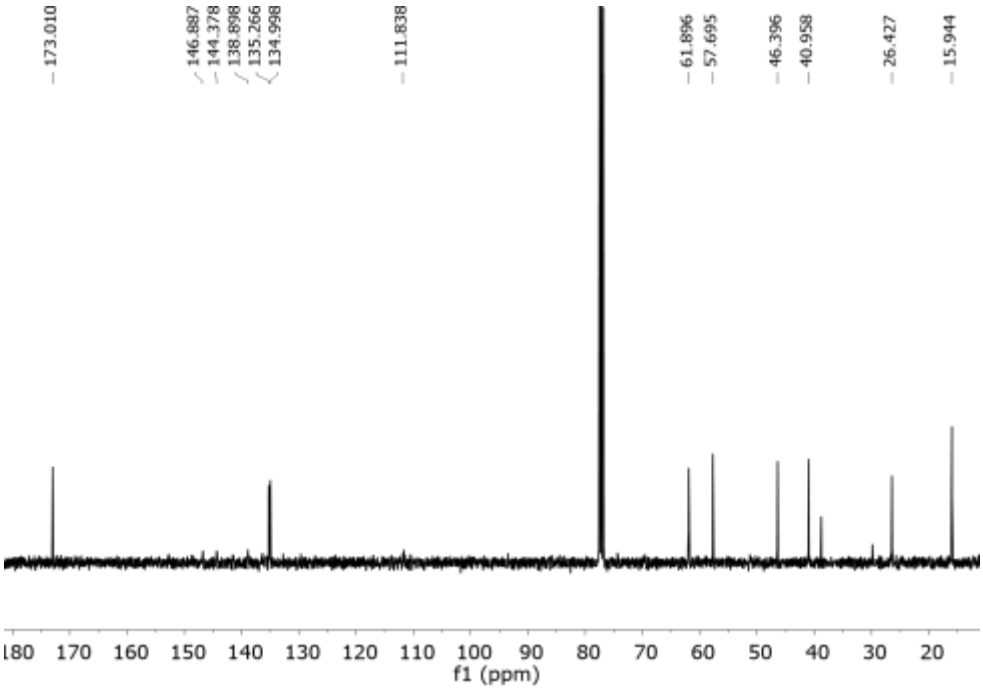
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.11 (m, 3H, NH_{amide}), 3.89 ((AB_q, δ_A=3.98, δ_B=3.81, J_{AB} = 13.1 Hz, 6H, H**5**)), 3.550 (X subsystem from ABX, J_{AX}=7.3, J_{BX}=6.5 Hz, 3H, H**4**), 3.21 (A subsystem from ABX, J_{AX}=7.3 J_{AB}=14 Hz, 3H, H**9**), 3.10 (B subsystem from ABX, J_{BX}=6.5, J_{AB}=14 Hz, 3H, H**9**), 2.98(m, 3H, H**2**), 2.64 (m, 6H, H**2'+1**), 2.44 (m, 3H, H**1'**), 2.36 (s, 9H, H**8**).

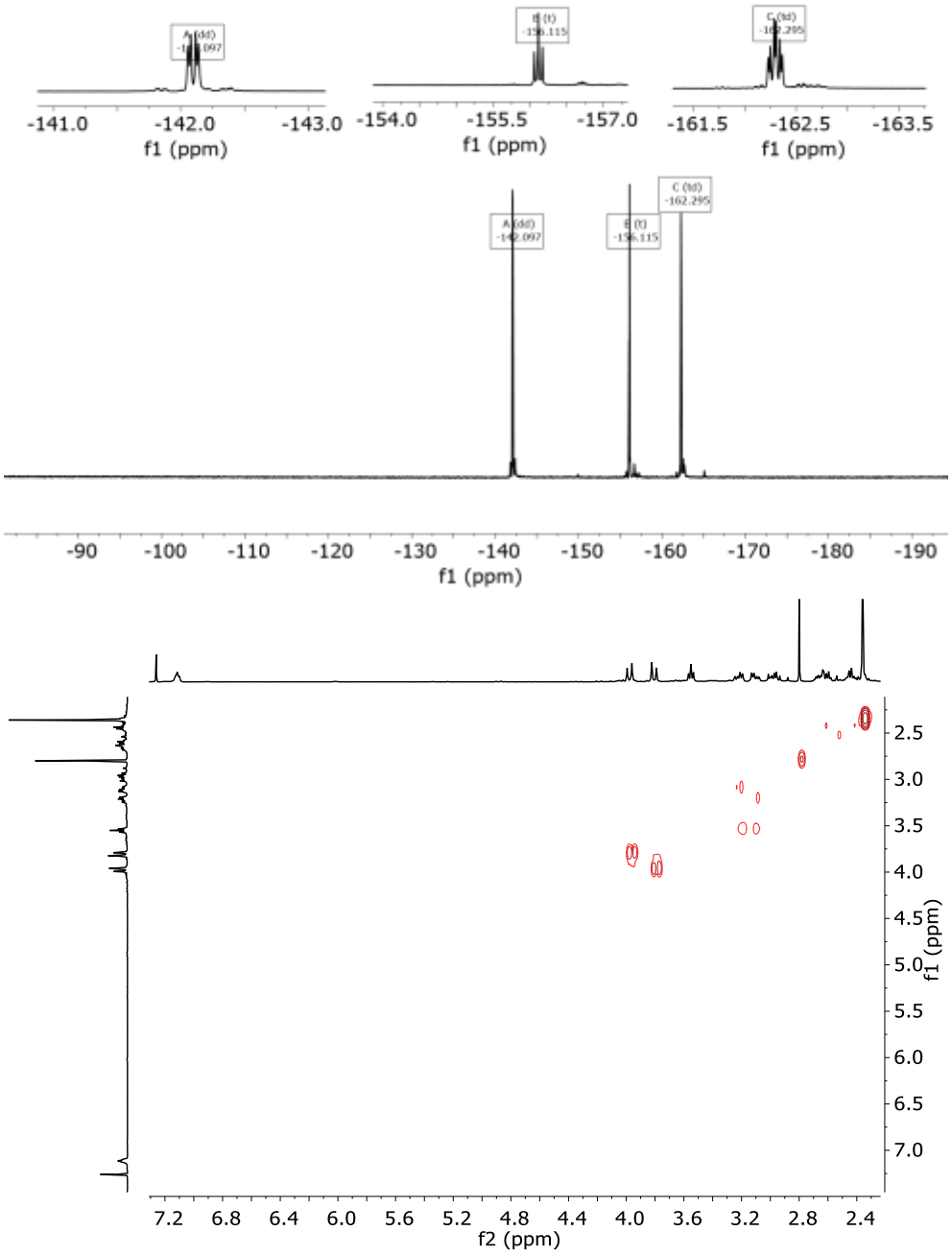
^{13}C NMR (101 MHz, $CDCl_3$) δ 173.0 (C**3**), 146.9 (m, C**11/12**), 144.4 (m, C**11/12**), 138.9 (m, C**13**), 135.3 (C**6**), 135.0 (C**7**), 111.8 (m, C**10**), 61.9 (C**4**), 57.7 (C**1**), 46.4 (C**5**), 41.0 (C**2**), 26.4 (C**9**), 15.9 (C**8**).

^{19}F –{^{1}H}-(NMR 376 MHz, $CDCl_3$) δ -142.10 (dd, *J* = 22.7, 8.2 Hz), -156.12 (t, *J* = 20.9 Hz), -162.29 (td, *J* = 21.9, 8.2 Hz).

HRMS (ESI-TOF) m/z [**3n**+ H]$^{+}$ Calc: 1014.2008, found: 1014.3086


A (t)
7.114
B (d)
3.977
I (d)
3.806
C (t)
3.550
D (dd)
3.219
J (dd)
3.104
E (m)
2.979
F (m)
2.637
G (m)
2.449
H (s)
2.360
0.86
1.14
1.06
1.22
2.34
1.26
2.11
1.33
3.04
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
f1 (ppm)


173.010
146.887
144.378
138.898
135.266
134.998
111.838
61.896
57.695
46.396
40.958
26.427
15.944
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
f1 (ppm)


-141.0
-142.0
-143.0
f1 (ppm)
-154.0
-155.5
-157.0
f1 (ppm)
-161.5
-162.5
-163.5
f1 (ppm)
-90
-100
-110
-120
-130
-140
-150
-160
-170
-180
-190
f1 (ppm)
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f1 (ppm)
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
f2 (ppm)

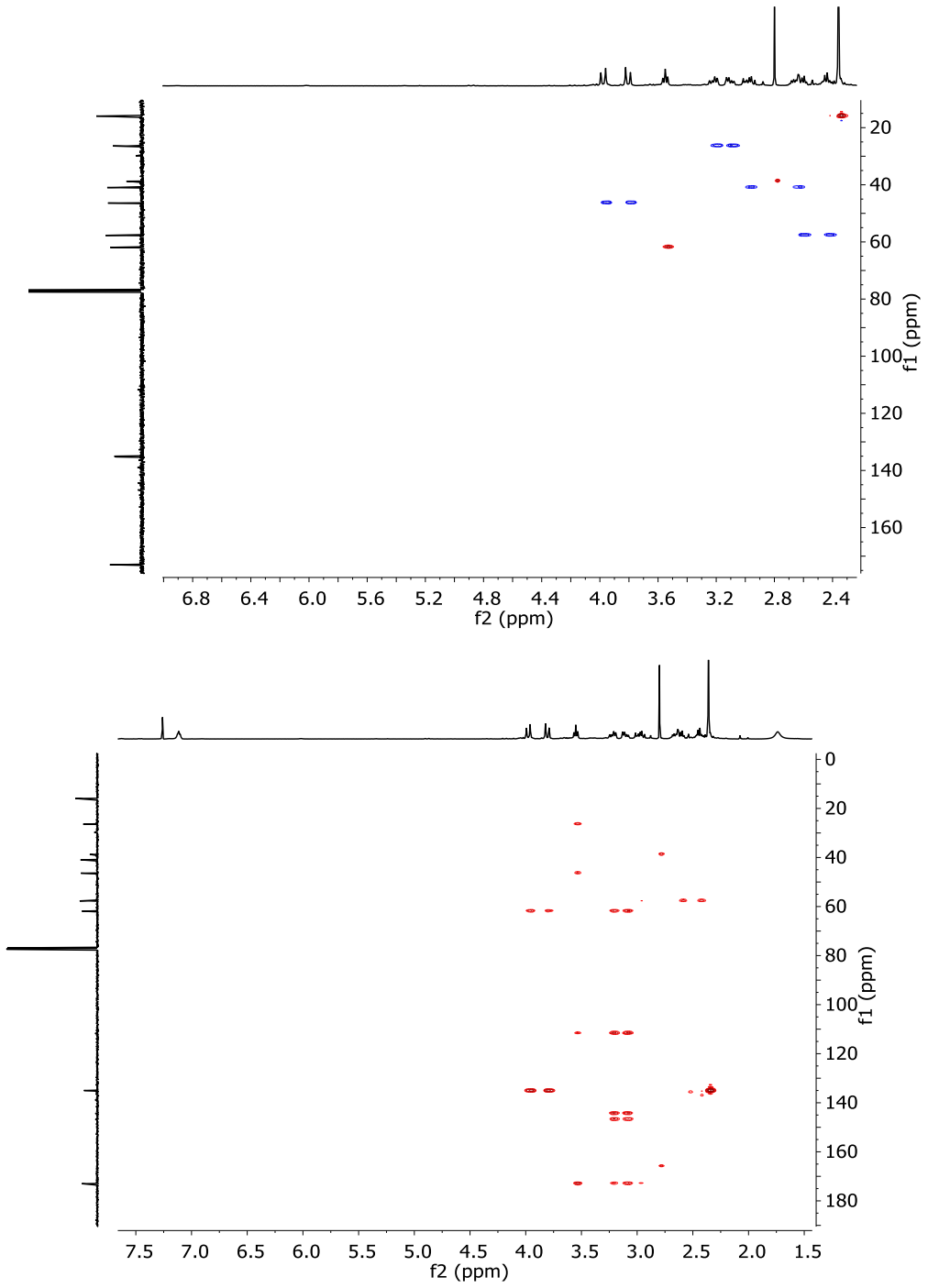


Figure 3.63. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, $CDCl_3$), ^{19}F NMR (376 MHz, $CDCl_3$) COSY, HSQC, HMBC of **1n.**

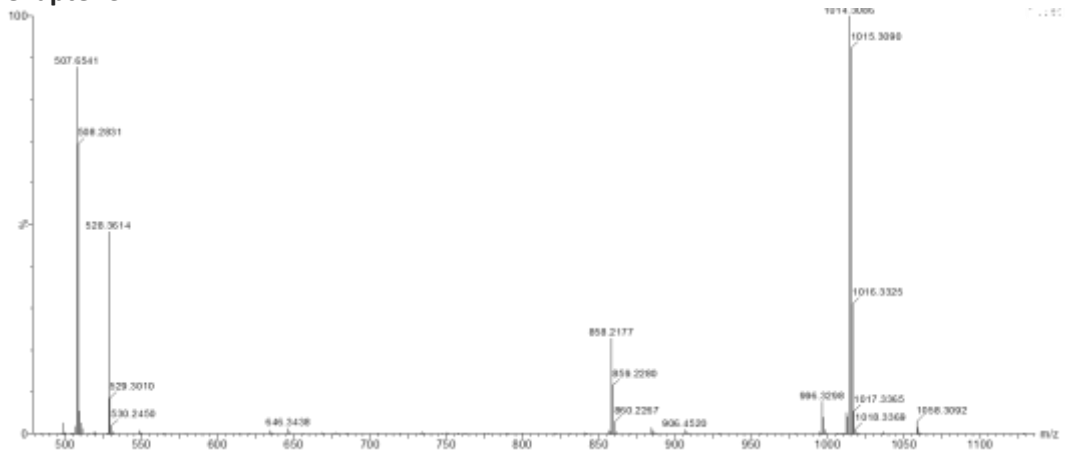


Figure 3.64. HRMS (ESI+) experimental spectrum of **1n.**

3.4.9 Synthesis of 1o

2o. tri-tert-butyl ((2S,2'S,2''S)-((nitrilotris(ethane-2,1-diyl))tris(azanediyl))tris(1-oxo-3-(4-(trifluoromethyl)phenyl)propane-1,2-diyl))tricarbamate

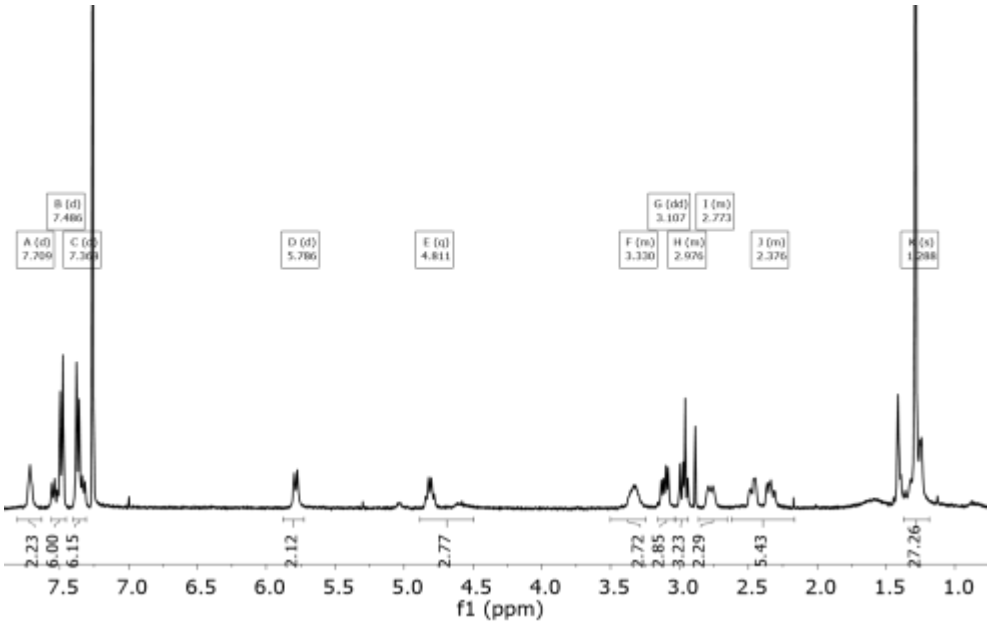
Boc-Phe(3-CF_3)-OH (493 mg, 1.48 mmol) was dissolved in dry DMF (4 mL). *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC·HCl O.343 mg, 1.79 mmol) and 1-Hydroxybenzotriazole hydrate (HOBt, 0.274 mg, 1.793 mmol), *N*,*N*-diisopropylethylamine (DIPEA, 0.93 mL, 5.37 mmol) and tris(2-aminoethyl)amine (0.07 mL, 0.448 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using 98:2 DCM:MeOH to give 0.216 mg of **2o** (0.197 mmol, 44% yield).

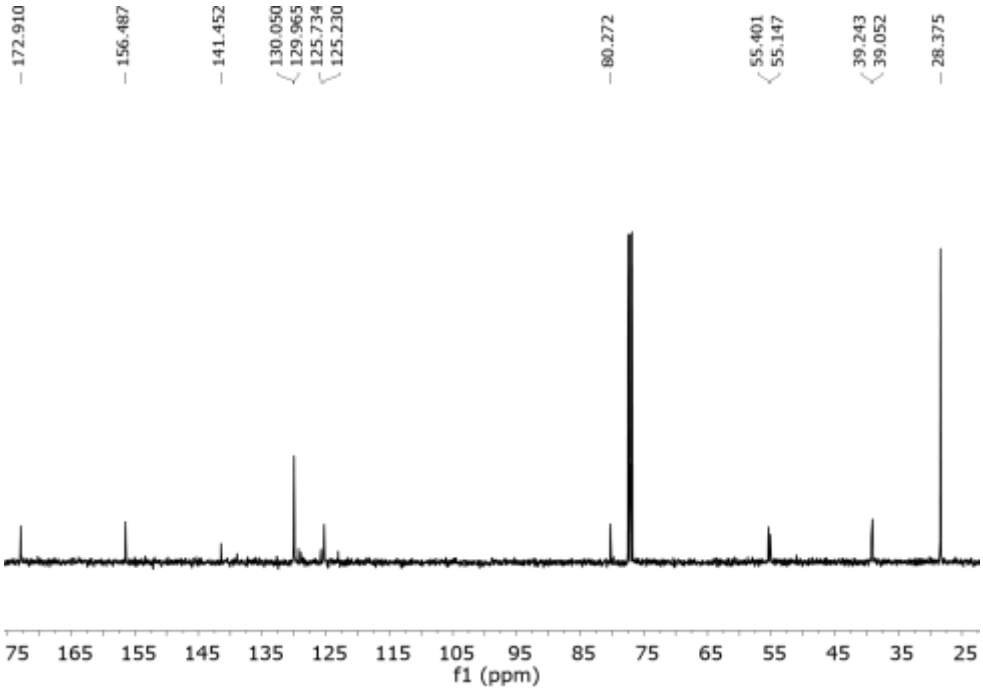
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.711 (d, *J* = 5.2 Hz, 3H), 7.49 (d, *J* = 7.9 Hz, 6H), 7.37 (d, *J* = 7.9 Hz, 6H), 5.79 (d, *J* = 9.1 Hz, 3H), 4.81 (q_{ap}, *J* = 8.3 Hz, 3H), 3.33 (m, 3H), 3.11 (dd, *J* = 13.5, 6.6 Hz, 3H), 2.98 (m, 3H), 2.77 (m, 3H), 2.38 (m, 6H), 1.29 (s, 27H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 172.9, 156.5, 141.5, 130.1, 130.0, 125.7, 125.2, 80.3, 55.4, 55.2, 39.2, 39.1, 28.4.

^{19}F –{^{1}H}- NMR (376 MHz, CD_3OD) δ -62.545.

HRMS (ESI-TOF) m/z [**2o** + H]$^+$ Calc: 1092.4851, found: 1092.4929


A (d) 7.709
B (d) 7.486
C (dd) 7.369
D (d) 5.786
E (q) 4.811
F (m) 3.330
G (dd) 3.107
H (m) 2.976
I (m) 2.773
J (m) 2.376
K (s) 1.288
2.23
6.00
6.15
2.12
2.77
2.72
2.85
3.23
2.29
5.43
27.26
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
f1 (ppm)


172.910
156.487
141.452
130.050
129.965
125.734
125.230
80.272
55.401
55.147
39.243
39.052
28.375
75 165 155 145 135 125 115 105 95 85 75 65 55 45 35 25
f1 (ppm)

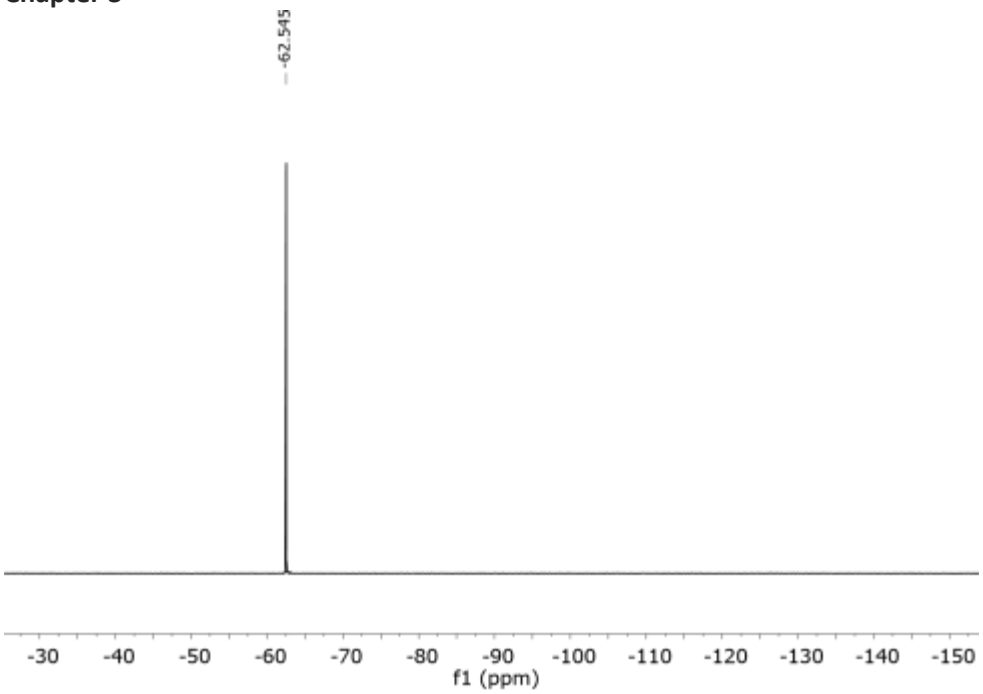


Figure 3.65. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, $CDCl_3$) and ^{19}F NMR (376 MHz, CD_3OD) spectra of **2o.**

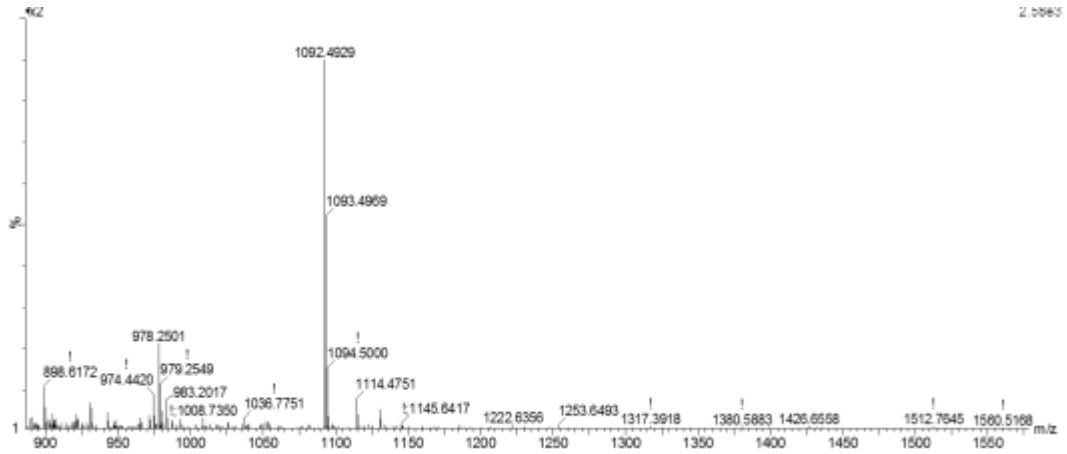


Figure 3.66. HRMS (ESI+) experimental spectrum of **2o.**

3o. (2S,2'S,2''S)-N,N',N''-(nitrilotris(ethane-2,1-diyl))tris(2-amino-3-(4-(trifluoromethyl) phe-nyl)propanamide)

2o (200 mg, 0.183 mmol) was dissolved in DCM (1mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The

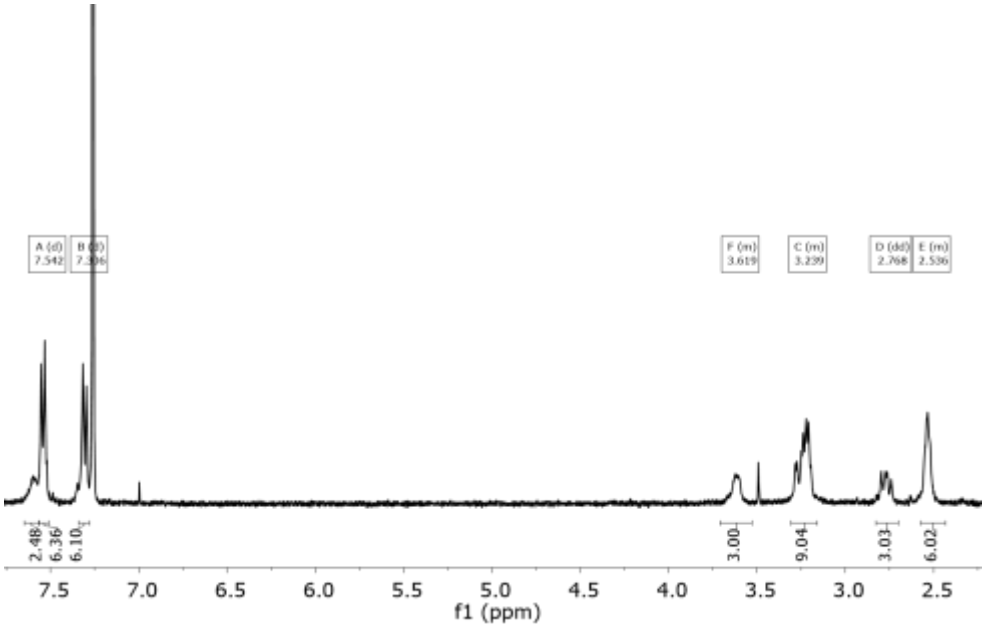
residue was washed several times with diethyl ether and dried affording 3o·4TFA as a white solid. TFA was removed by dissolving the solid in NaOH 1M and extract several times with CH_2Cl_2, dried and washed with hexane. Final product **3o** was isolated as a white solid (120 mg, 0.155 mmol, 85% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.54 (d, J = 7.9 Hz, 6H), 7.31 (d, *J* = 8.0 Hz, 6H), 3.62 (m, 3H), 3.24 (m, 9H), 2.77 (dd, *J* = 13.6, 9.1 Hz, 3H), 2.54 (m, 6H).

^{13}C NMR (101 MHz, CD_3OD) δ 174.0, 142.1, 129.6, 129.6, 125.5, 125.5, 56.3, 54.5, 41.0, 37.6.

^{19}F NMR (376 MHz, $CDCl_3$) δ -62.442.

HRMS (ESI-TOF) m/z [**3o** + H]$^+$ Calc: 792,3278, found: 792.3239



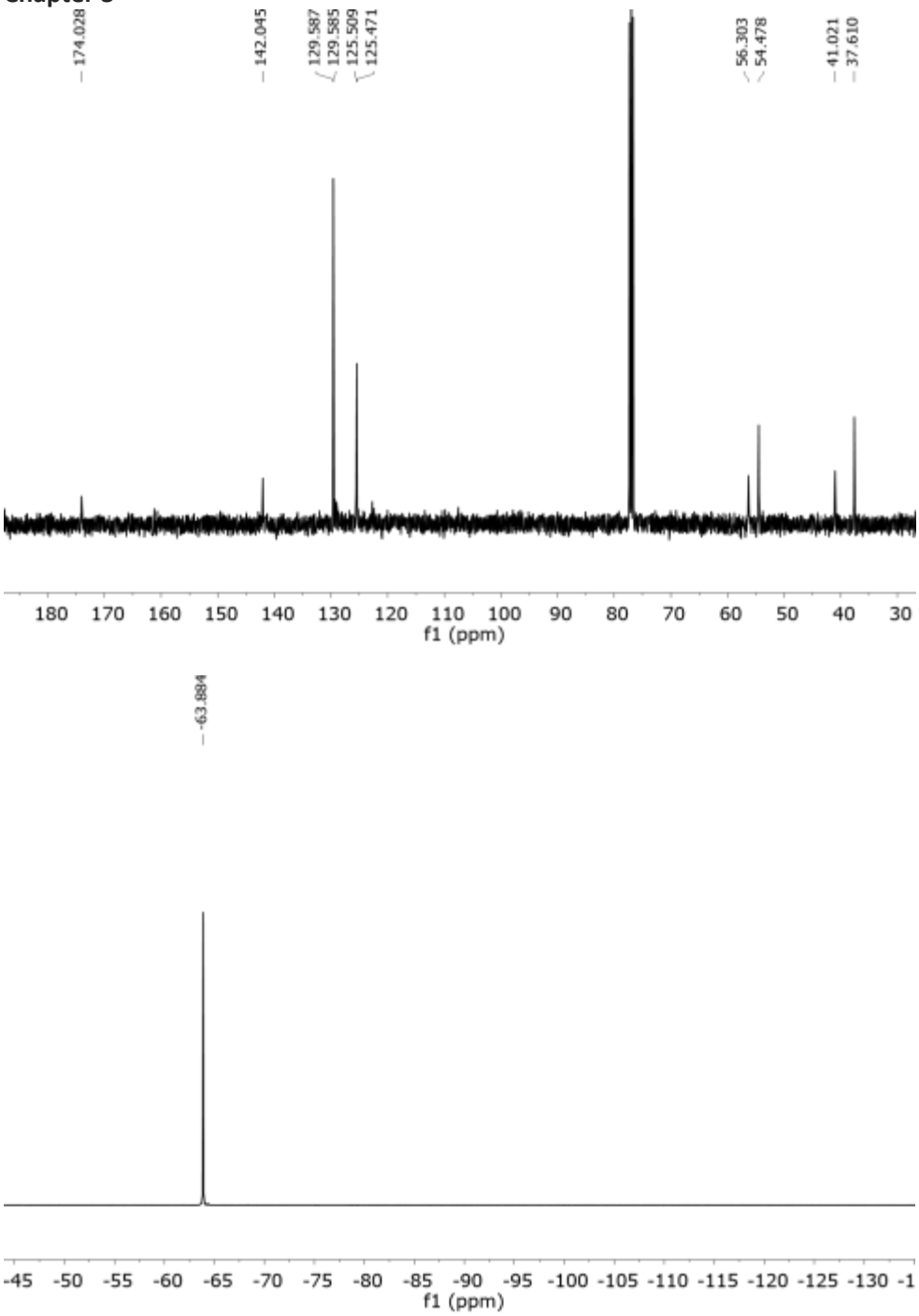


Figure 3.67. ^{1}H-NMR (400 MHz, $CDCl_3$), ^{13}C-NMR (101 MHz, $CDCl_3$) and ^{19}F NMR (376 MHz, CD_3OD) spectra of **3o.**

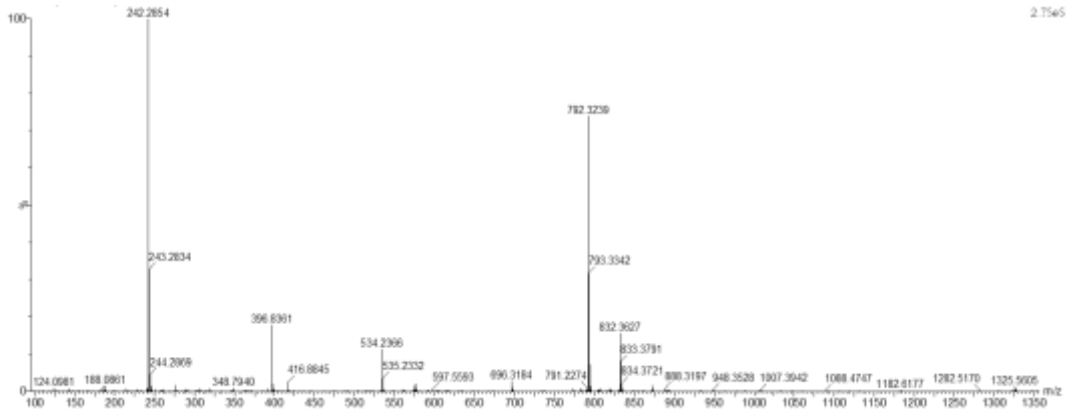
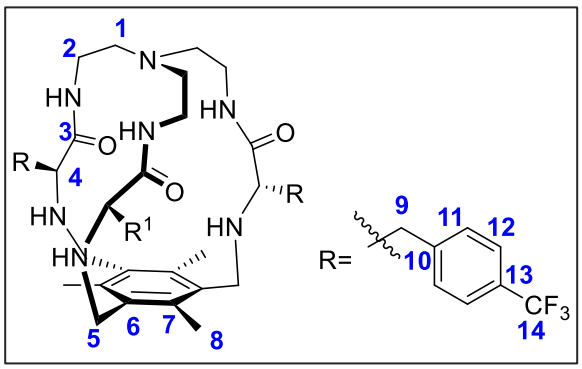


Figure 3.68. HRMS (ESI+) experimental spectrum of **3o.**

1o



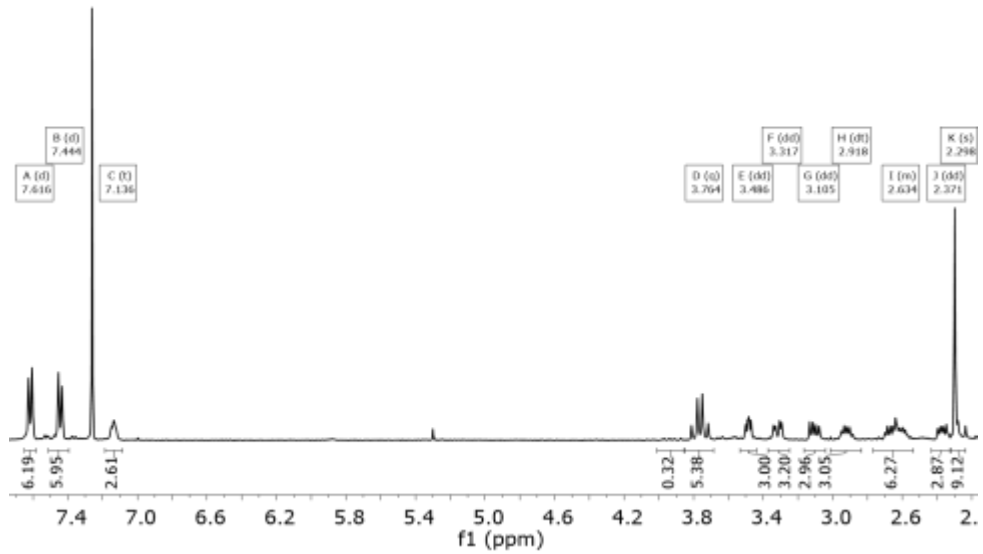
3o (120 mg, 0.151 mmol) was dissolved in ACN (40 mL). Tetrabuthylamonium chloride (21 mg, 0.076 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (60.5 mg, 0.151 mmol) and potassium carbonate (419 mg, 3.03 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM : MeOH 98 : 2 as eluent to give 1o as a white solid (70 mg, 0.0738 mmol, 50% yield).

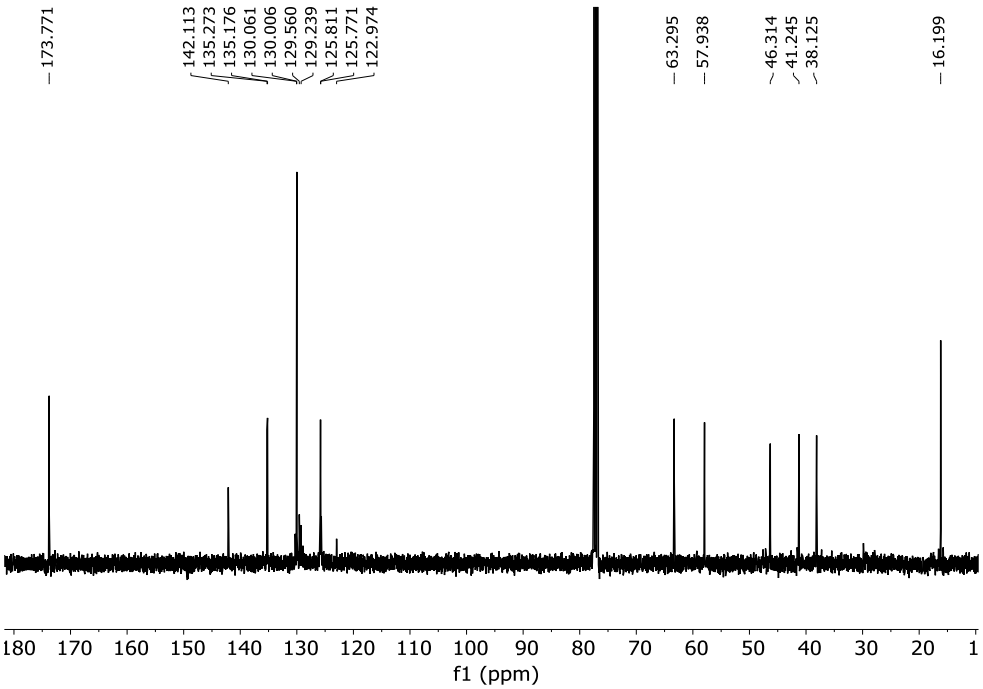
^{1}H NMR (400 MHz, CD_3Cl) δ 7.62 (d, J = 7.9 Hz, 6H (H**12**)), 7.44 (d, *J* = 8.0 Hz, 6H (H**11**)), 7.14 (t, *J* = 5.4 Hz, 3H (NH_{amide})), 3.76 (AB_q, δ_A = 3.79, δ_B = 3.73, J_{AB} = 13.5 Hz, 6H, (H**5**)), 3.49 (X subsystem from ABX, J_{AX} = 4.8, J_{BX} = 7.4 Hz, 3H (H**4**)), 3.32 (A subsystem from ABX, J_{AX} = 7.4, J_{AB} = 13.8 Hz, 3H, (H**9**)), 3.11 (B subsystem from ABX, J_{BX} = 7.4, J_{AB}=13.5 Hz, 3H, (H**9**)), 2.92 (m, 3H, (H**2**)), 2.63 (m, 6H, (H**2**) overlapped with (H**1**)), 2.37 (m, 3H, (H**1**)), 2.30 (s, 9H (H**8**)).

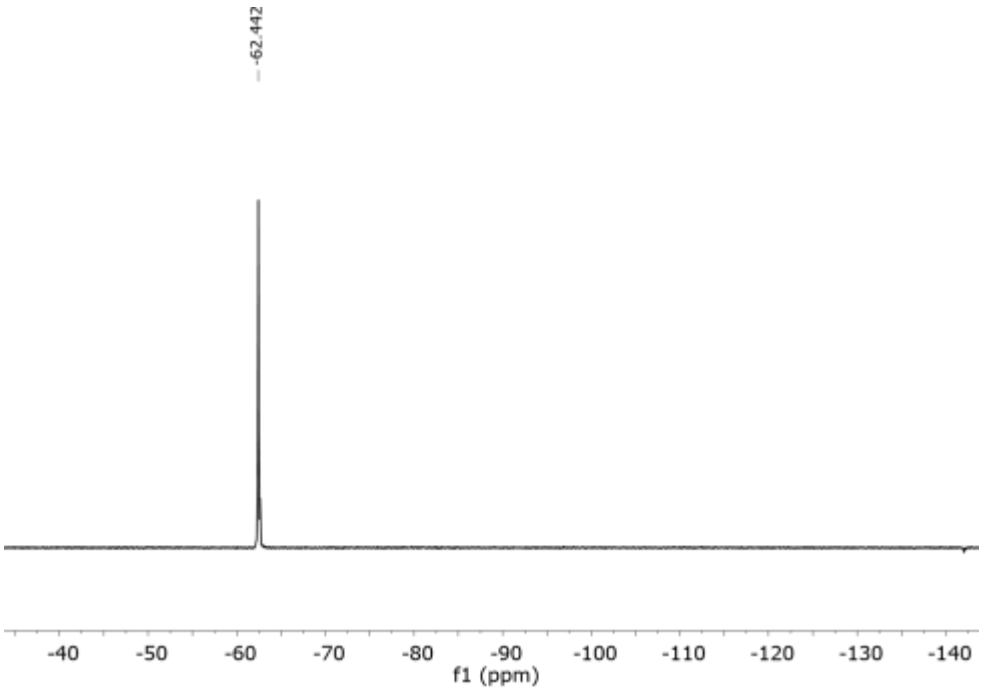
^{13}C NMR (101 MHz, CD_3Cl) δ 173.8 (C**3**), 142.1 (C**10**), 135.3 (m, C-F coupling C**14**), 135.2 (C**7**), 130.1 (C**11**), 130.0 (C**12**), 129.2 (C**4**) 125.8 (m, C-F coupling **13**), 125.7 ((C**6**), 63.3 (C**9**), 57.9 (C**1**), 46.3 (C**5**), 41.2 (C**2**), 38.13 (C**9**), 16.2 (C**8**).

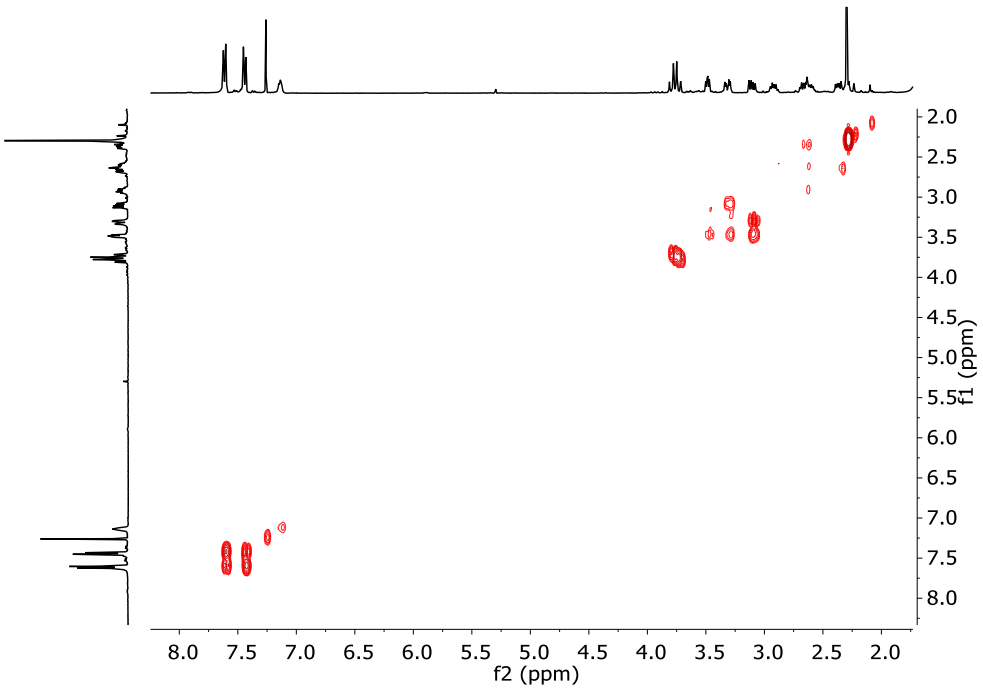
^{19}F –{^{1}H}-NMR (376 MHz, CD_3OD) δ -62.442.

HRMS (ESI-TOF) m/z [**1o** + H]$^{+}$ Calc: 948.4217, found: 948.5418


A (d) 7.616
B (d) 7.444
C (t) 7.136
D (q) 3.764
E (dd) 3.486
F (dd) 3.317
G (dd) 3.105
H (dt) 2.918
I (m) 2.634
J (dd) 2.371
K (s) 2.298
6.19
5.95
2.61
0.32
5.38
3.00
3.20
2.96
3.05
6.27
2.87
9.12
7.4 7.0 6.6 6.2 5.8 5.4 5.0 4.6 4.2 3.8 3.4 3.0 2.6 2.
f1 (ppm)


173.771
142.113
135.273
135.176
130.061
130.006
129.560
129.239
125.811
125.771
122.974
63.295
57.938
46.314
41.245
38.125
16.199
180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 1
f1 (ppm)


-62.442
-40
-50
-60
-70
-80
-90
-100
-110
-120
-130
-140
f1 (ppm)


2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
f1 (ppm)
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
f2 (ppm)

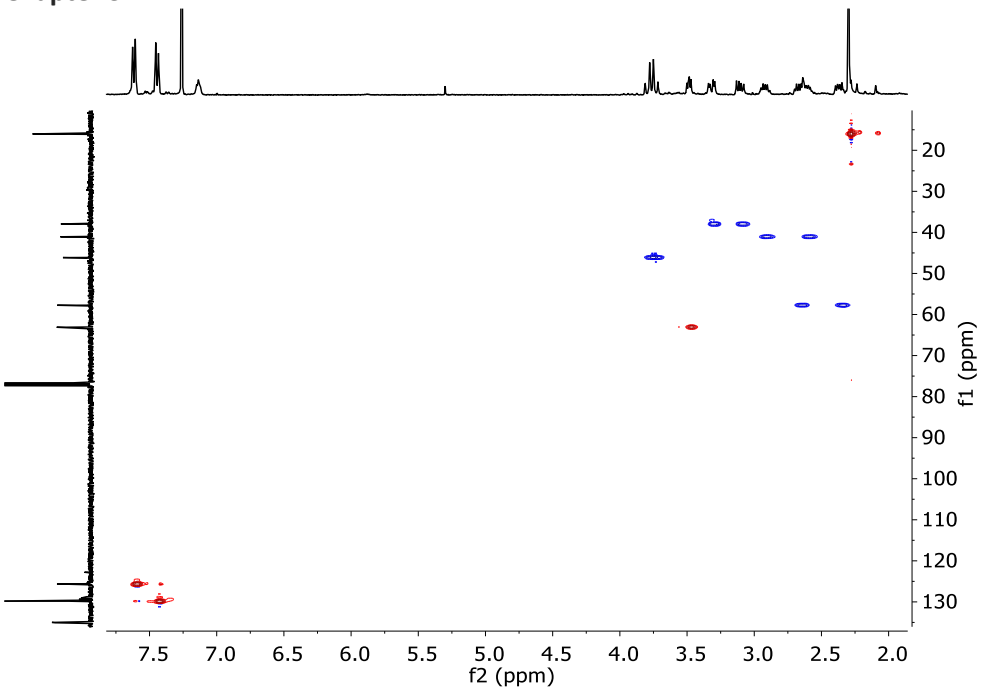


Figure 3.69. ^{1}H-NMR (400 MHz, CD_3Cl), ^{13}C-NMR (101 MHz, CD_3Cl), ^{19}F NMR (376 MHz, CD_3OD), COSY and HSQC spectra of **1o.**

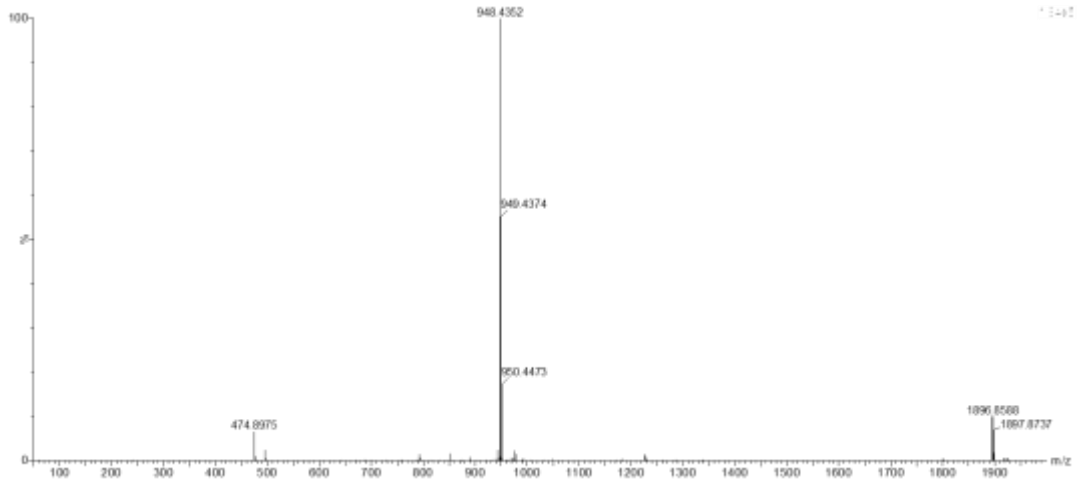


Figure 3.70. HRMS (ESI+) experimental spectrum of **1o.**

3.4.10 X-ray crystal analysis

Crystals suitable for X-ray diffraction studies of compound 1h were obtained by slowly evaporating a solution of **1h** fully protonated with TFA in acetonitrile with 5% H_2O and an excess of TBACl.

Table 3.7 Crystal data and structure refinement for **1h**

Identification code	n32	
Empirical formula	C45 H72 Cl4 F3 N7 O10	
Formula weight	1069.89	
Temperature	173(2) K	
Wavelength	0.71073 Å	
Crystal system	Orthorhombic	
Space group	P 21 21 21	
Unit cell dimensions	a = 13.9421(4) Å	a= 90°.
	b = 17.0957(4) Å	b= 90°.
	c = 21.6955(7) Å	g = 90°.
Volume	5171.1(3) $Å^3$	
Z	4	
Density (calculated)	1.374 Mg/m^3	
Absorption coefficient	0.301 mm^{-1}	
F(000)	2264	
Crystal size	0.160 x 0.150 x 0.120 mm^3	
Theta range for data collection	3.293 to 25.651°.	
Index ranges	-16<=h<=16, -18<=k<=20, -26<=l<=26	
Reflections collected	47047	
Independent reflections	9693 [R(int) = 0.0488]	
Completeness to theta = 10.000°	95.2 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	1.000 and 0.818	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	9693 / 6 / 721	
Goodness-of-fit on F^2	1.099	
Final R indices [I>2sigma(I)]	R1 = 0.0483, wR2 = 0.1050	
R indices (all data)	R1 = 0.0568, wR2 = 0.1088	
Absolute structure parameter	0.00(3)	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.649 and -0.236 $e.Å^{-3}$	

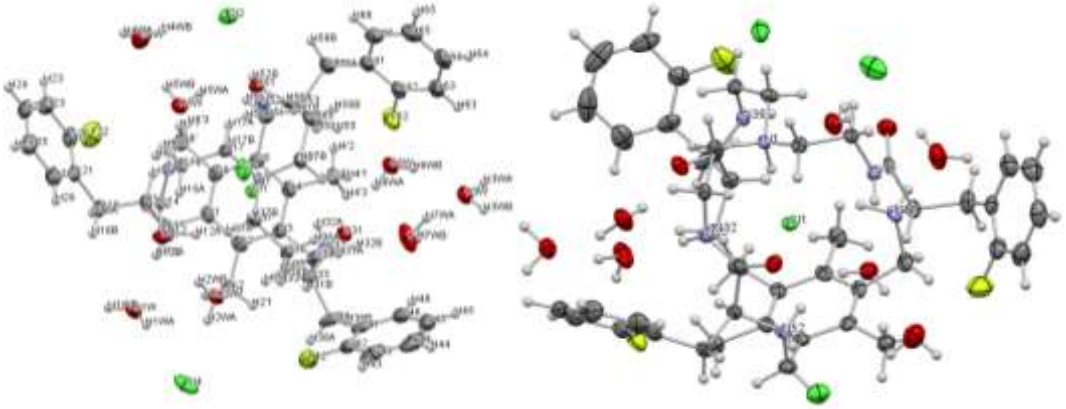
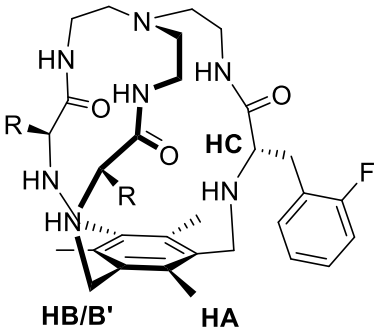


Figure 3.71. Perspective view of 1b with the atoms labels. Displacement parameters are drawn at the 50 % probability level. Additional chloride counterions, solvent molecules and non-polar hydrogen atoms have been omitted for clarity (left). Perspective view of 1d with the most important atoms labels.

3.4.11 Titration of compounds 1h, 1m and 1l with TBACl: K_{ass}

3.4.11.1 1h



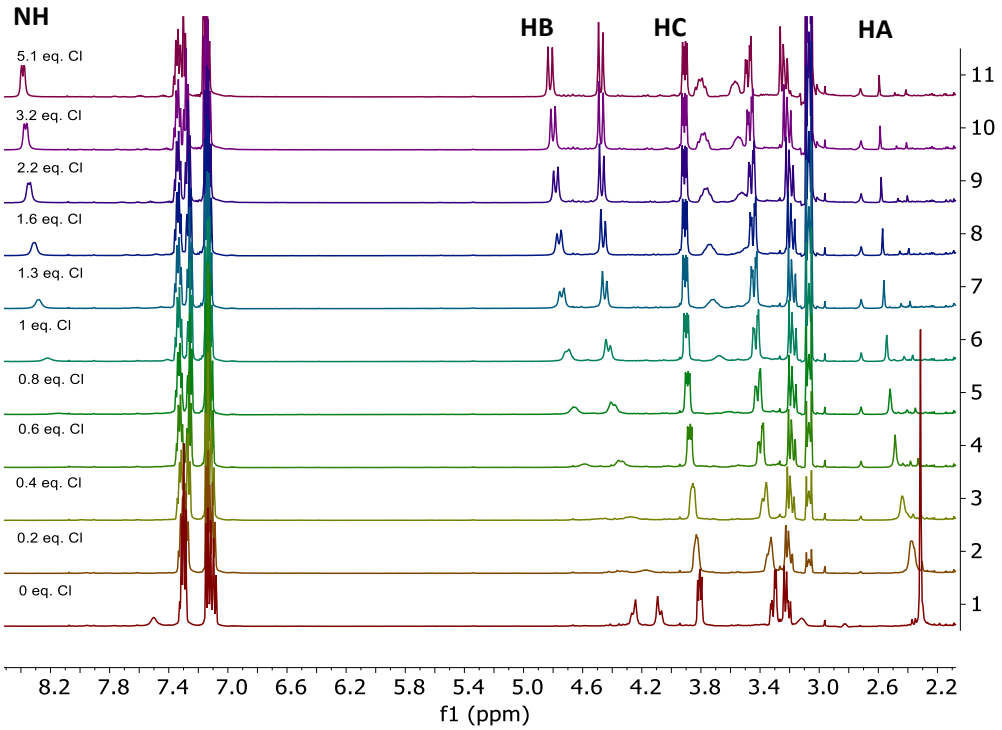


Figure 3.72. Stacked ^{1}H NMR spectra for the titration of **1h.**

Table 3.8 data set of **1h** titration:

[Cl] (M)	NH	NH fit	HB	HB fit	HC	HC fit	HA	HA fit
0.0005	7.696	7.696	4.335	4.374	3.339	3.343	2.368	2.374
0.0010	7.970	7.871	4.461	4.482	3.368	3.373	2.437	2.429
0.0016	8.046	8.030	4.588	4.581	3.396	3.401	2.468	2.479
0.0021	8.154	8.152	4.661	4.656	3.415	3.422	2.520	2.517
0.0026	8.222	8.234	4.705	4.707	3.428	3.436	2.543	2.543
0.0033	8.281	8.299	4.742	4.748	3.441	3.447	2.562	2.563
0.0041	8.311	8.330	4.760	4.767	3.448	3.452	2.571	2.573
0.0056	8.352	8.358	4.782	4.784	3.458	3.457	2.581	2.581
0.0081	8.370	8.375	4.800	4.794	3.468	3.460	2.588	2.587

Results of the HypNMR fitting: Logß1= 3.76 ± 0.094 (1 : 1), BC_{50} = 173 ± 4 µM

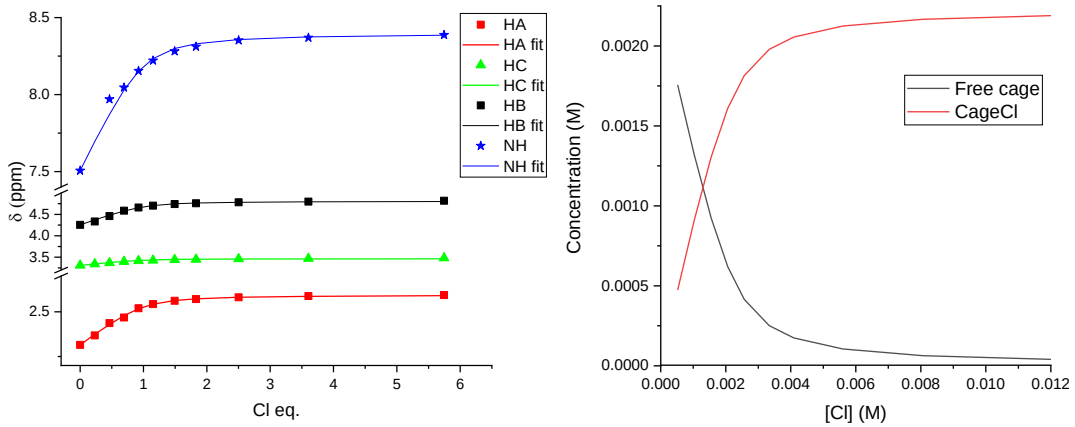
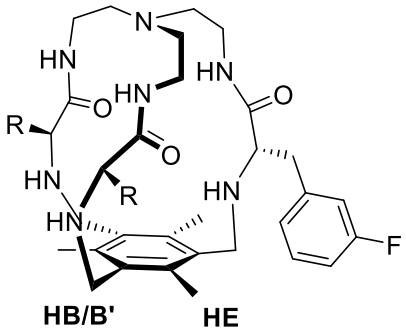


Figure 3.73. Plot of the experimental (symbols) and fitting (lines) data of 1h titration (lelft). Species distribution as a function of the chloride concentration of 1h titration (right).

3.4.11.2 1i



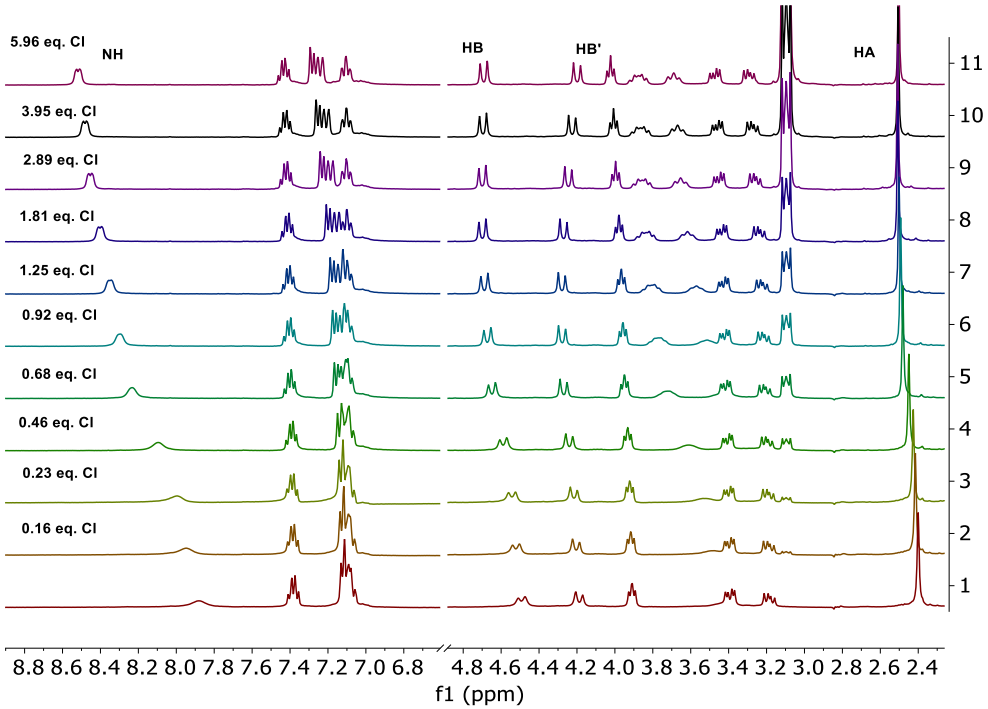


Figure 3.74. Stacked ^{1}H NMR spectra for the titration of **1i.**

Table 3.9: data set of **1i** titration

[Cl] M	NH	NH fit	HB	HB fit	HB'	HB' fit	HA	HA fit
0.0002	7.948	7.938	4.522	4.517	4.205	4.2012	2.415	2.413
0.0004	7.997	7.9943	4.543	4.5418	4.217	4.2142	2.426	2.4259
0.0008	8.095	8.1042	4.589	4.5899	4.241	4.2393	2.45	2.4509
0.0012	8.233	8.2076	4.648	4.6347	4.271	4.2622	2.481	2.4743
0.0016	8.296	8.2938	4.672	4.671	4.279	4.2794	2.493	2.4931
0.0022	8.348	8.3642	4.688	4.696	4.28	4.2855	2.502	2.5061
0.0031	8.402	8.407	4.698	4.7019	4.271	4.2737	2.508	2.5092
0.0050	8.453	8.4507	4.699	4.6992	4.246	4.2469	2.508	2.5077
0.0068	8.481	8.4796	4.695	4.6955	4.225	4.2258	2.506	2.5057

Results of the HypNMR fitting: Logβ1 = 4.4 ± 0.1 (1:1), Log β2 = 6.5 ± 0.2 (1 : 2)**,** BC_{50} = 36.8 ± 8 µM

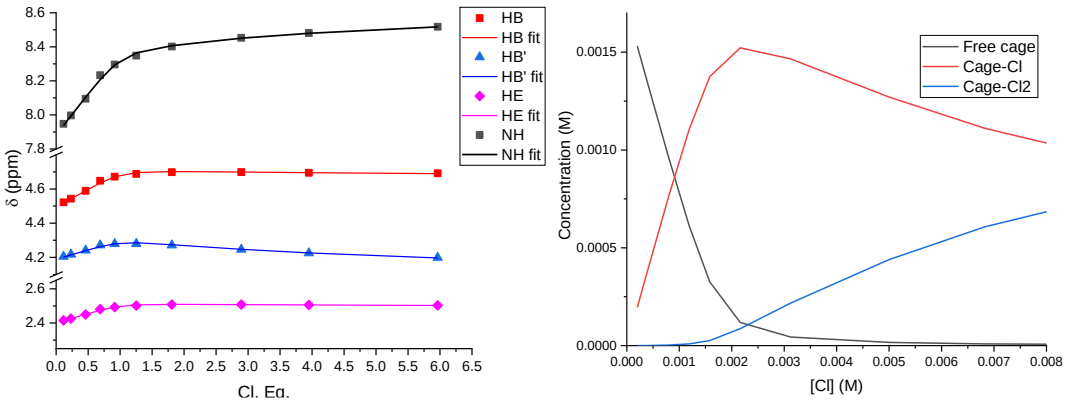
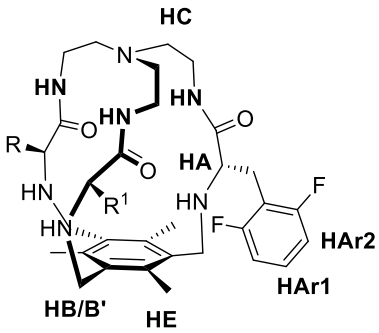


Figure 3.75. Plot of the experimental (symbols) and fitting (lines) data of 1i titration (lelft). Species distribution as a function of the chloride concentration of 1i titration (right).

3.4.11.3 1l



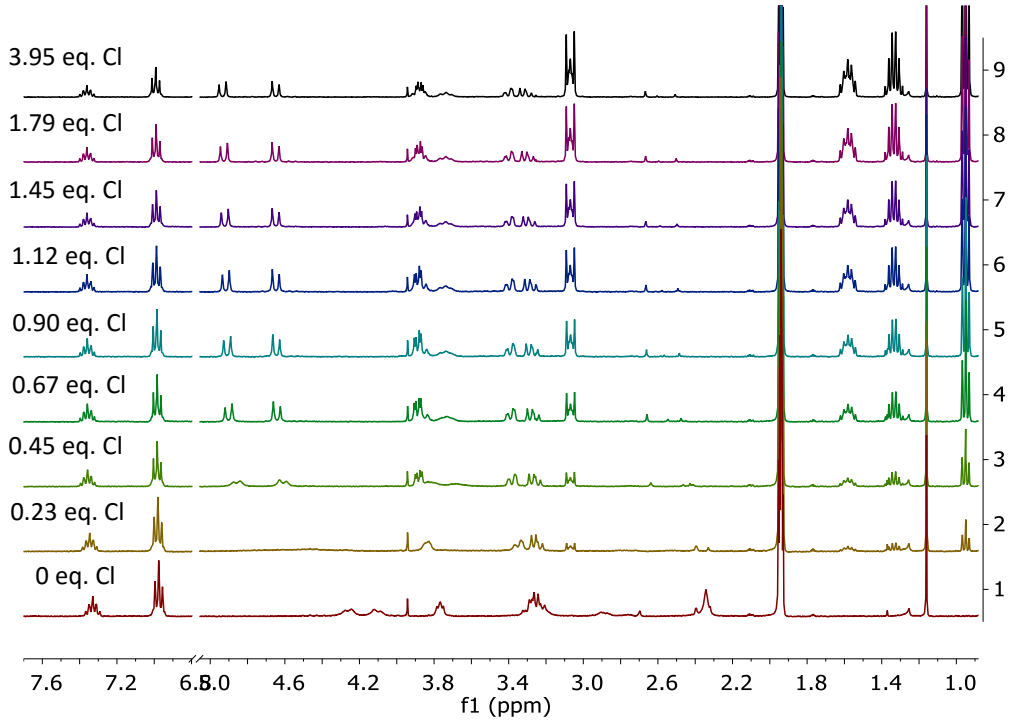


Figure 3.76. Stacked ^{1}H NMR spectra for the titration of **1l.**

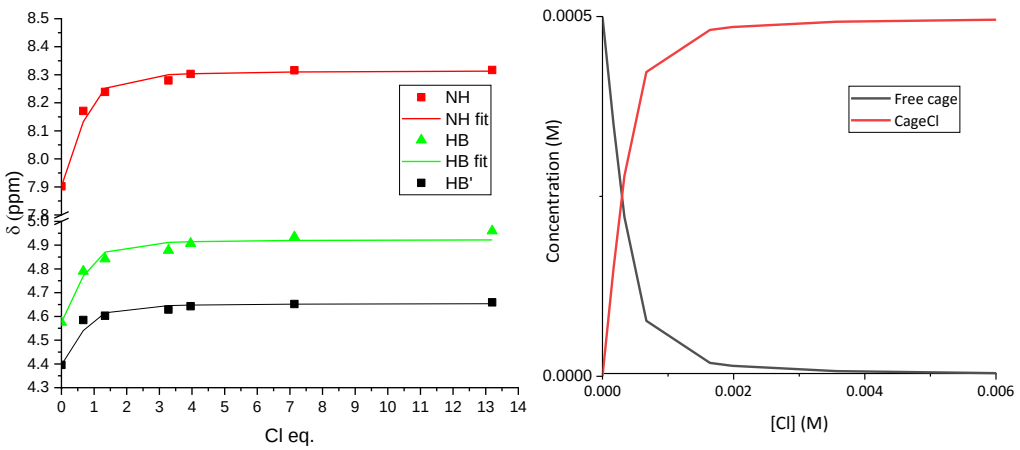


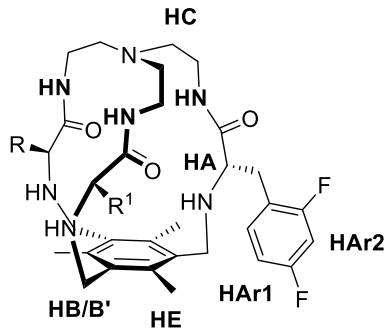
Figure 3.77. Plot of the experimental (symbols) and fitting (lines) data of **1l** titration (lelft). Species distribution as a function of the chloride concentration of **1l** titration (right).

[Cl]	Cl eq.	HA	HA fit	HC	HC fit	HB	HB fit
0	0	7.902	7.902	4.575	4.575	4.396	4.396
0.000349	0.67	8.171	8.1326	4.79	4.7698	4.585	4.5404
0.000687	1.3378	8.239	8.252	4.842	4.8707	4.603	4.6153

0.00137	3.28	8.28	8.3004	4.878	4.9116	4.629	4.6456
0.001879	3.96	8.303	8.3038	4.907	4.9145	4.643	4.6477
0.00256	7.14	8.316	8.3099	4.935	4.9197	4.652	4.6515
0.00306	13.2	8.317	8.3129	4.96	4.9221	4.659	4.6534

Results of HypNMR fitting: Logß = 4.3 ± 0.13; BC_{50} = 50 ± 15 µM

3.4.11.4 1m



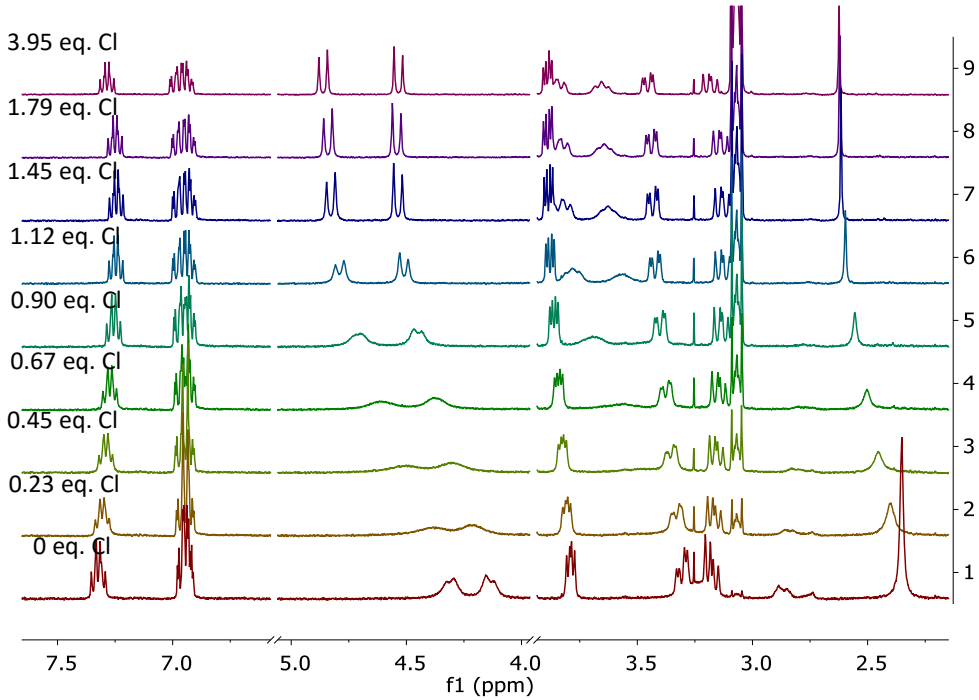


Figure 3.78. Stacked ^{1}H NMR spectra for the titration of **1m.**

Table . Data set of **1m** titration:

[Cl]	Cl eq.	HB	HB fit	HB'	HB'fit	HA	HA fit	HE	HE fit

0.000	0.000	4.303	4.2903	4.149	4.1382	3.306	3.3061	2.352	2.352
0.000	0.209	4.373	4.4039	4.226	4.222	3.329	3.3334	2.401	2.408
0.001	0.417	4.489	4.515	4.304	4.3041	3.354	3.3601	2.457	2.462
0.001	0.624	4.603	4.6215	4.375	4.3827	3.376	3.3857	2.503	2.514
0.001	0.830	4.705	4.7154	4.45	4.452	3.401	3.4083	2.555	2.560
0.002	1.036	4.79	4.7807	4.511	4.5002	3.423	3.424	2.597	2.591
0.002	1.344	4.828	4.8207	4.537	4.5297	3.433	3.4336	2.615	2.611
0.002	1.650	4.84	4.8332	4.542	4.539	3.439	3.4367	2.621	2.617
0.005	3.652	4.861	4.8471	4.535	4.5492	3.454	3.44	2.624	2.624

Results of HypNMR fitting: Logß = 4.3 ± 0.13; BC_{50} = 50 ± 15 µM

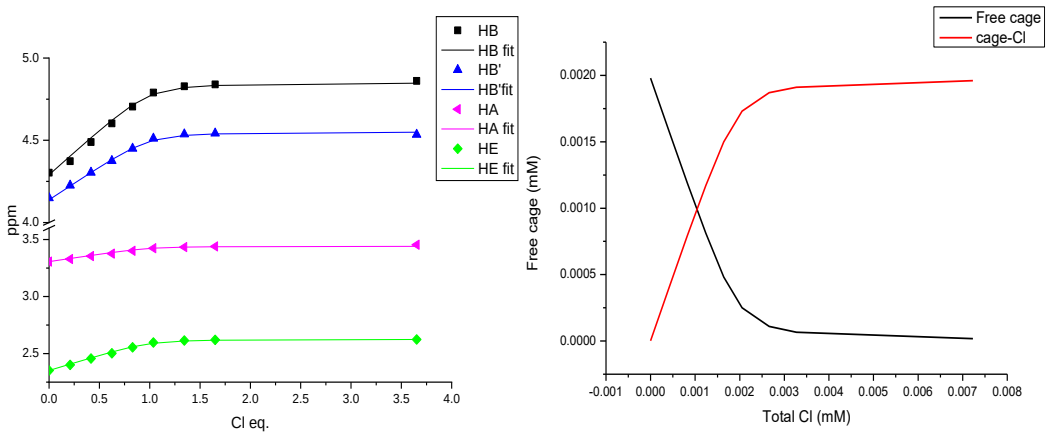


Figure 3.79. Plot of the experimental (symbols) and fitting (lines) data of 1m titration (left) Species distribution as a function of the chloride concentration of 1m titration (right).

3.4.12 LOG P calculation and Retention time determination

The HPLC program used to determine the retention time (r.t.) consists of a linear gradient of 20 % to 80 % of acetonitrile in water (using 0.1 % of trifluoroacetic acid) in 20 minutes.

Table 3.9: Calculated Log P values and retention times (minutes) of compounds 1c**,h-o.**

Cage	R	Log P neutral (calc)	r.t. (min)
1c	4F-Phe	4.00	9.44
1h	2F-Phe	4.00	8.02
1i	3F-Phe	4.00	9.22
1j	(3,5-F)Phe	4.38	10.64
1k	(3,4-F)Phe	4.39	10.33
1l	(2,6-F)Phe	4.38	8.28
1m	(2,4-F)Phe	4.37	9.53
1n	(PentaF)Phe	5.51	14.22
1o	(4-CF_3)Phe	5.90	12.44

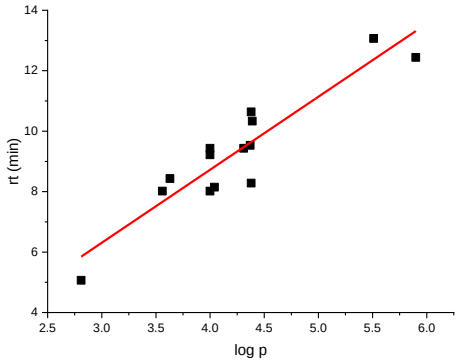


Figure 3.80. Plot of the Rt VS lop P (calculated for the neutral form) for compounds **1a-o.**

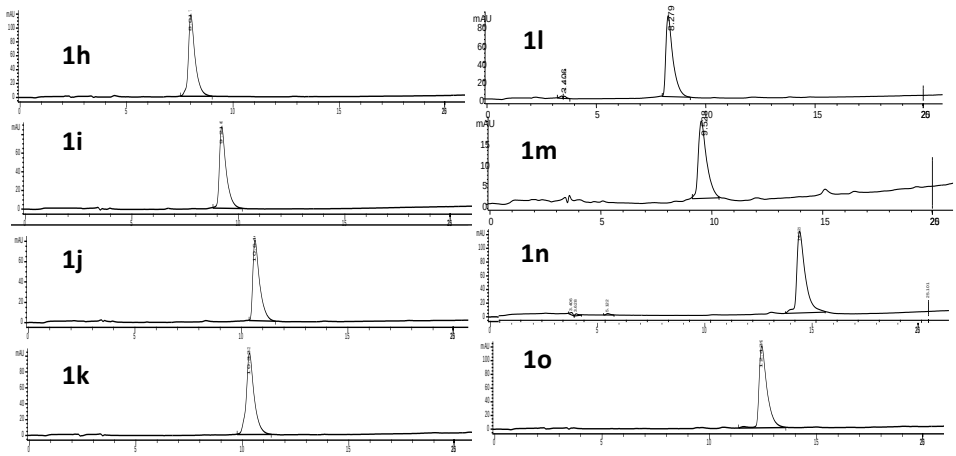


Figure 3.81. HPLC chromatogram of compounds **1h-o.**

3.4.13 Vesicles transport assays

Phospholipid vesicles were prepared following the protocol described in chapter 1.

3.4.13.1 HPTS-based assay

Calibration curve: First of all, a calibration curve matching I_{460}/I_{403}, the relationship between the emission intensity at 460 nm and that at 403 nm (corresponding to the excitation wavelengths of the dye's deprotonated and protonated forms, respectively) of an HPTS aqueous solution (15 nM), prepared with a $NaNO_3$/NaCl aqueous solution (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM NaH_2PO_4, I.S. 150 mM), and the pH was built. In order to do it, aliquots of a sodium hydroxide aqueous solution (0.5 M), prepared with a $NaNO_3$/NaCl aqueous solution (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM NaH_2PO_4, I.S. 150 mM), were successively added to the HPTS solution, and after each addition the ratio of emission intensities at 460 and 403 nm and the pH value of the solution were recorded. Data were fitted to an S-logistic model, which provided an $R^2 > 0.999$.

HPTS-based assay: POPC:cholesterol vesicles were loaded with a $NaNO_3$/NaCl aqueous solution (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM NaH_2PO_4, 1 mM HPTS, I.S. 150 mM, pH 6.5) and treated according to the procedure described in *chapter 2.* The assays were conducted in 1-cm disposable cells, the final POPC concentration in the cuvette being 0.5 mM and the final volume 2.5 mL. Just before starting the measurements the required volume of the vesicles stock solution was suspended in the outer solution (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM NaH_2PO_4, I.S. 150 mM, pH 7.5). At t = 60 s an aliquot of a solution of the compound in DMSO (or the blank, DMSO, 12.5 µL) was added, and the ratio of emission intensities at 460 and 403 nm recorded for five more minutes. At t = 360 s a detergent (Triton-X, 10% dispersion in water, 20 µL) was added, to lyse the vesicles and balance the pH.

3.4.13.2 Carboxy-fluorescein-based assay

POPC vesicles were loaded with a $NaNO_3$/NaCl aqueous solution (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM NaH_2PO_4, 50 mM CF, I.S. 150 mM, pH 7.2). The experiments were performed in 1-cm disposable cells, the final POPC concentration in the cuvette being 0.05 mM and the total volume 2.5 mL. Just before starting the measurements the required volume of the vesicles stock solution was suspended in the outer solution (123.9 mM $NaNO_3$, 10 mM NaCl, 10 mM NaH_2PO_4, I.S. 150 mM, pH 7.2). At t = 60 s the compound was added and the emission changes were recorded during five minutes. At t = 360 s a pulse of a detergent

(Triton-X, 10% dispersion in water, 20 μL) was added to lyse the vesicles and release all the entrapped carboxy-fluorescein. This value was regarded as 100% release and used to normalise the data.

3.4.14 pKa determination

A solution 100-200 μM of the corresponding cage in MeOH/H_2O (40:60, 11 mL) containing 150 mM of $KClO_4$ (to maintain the ionic strength constant through the assay) was adjusted to different pH's (3-11) using small and controlled volumes of NaOH or $HClO_4$. The fluorescence emission of the solution was measured at room temperature for each pH using λ_{ex} = 270 nm (maximum absorbance of the phenylalanine) and λ_{em} = 290-420 mn. There is a fluorescence increase upon basification of the solution associated with changes in the protonation of the cage.

Titration data was fitted using HypSpec software to the protonation model for these systems:

- $L + H \leftrightarrows LH$ $\quad \log \beta_1$
- $LH + H \leftrightarrows LH_2$ $\quad \log \beta_2$
- $LH_2 + H \leftrightarrows LH_3$ $\quad \log \beta_3$
- $LH_3 + H \leftrightarrows LH_4$ $\quad \log \beta_4$

From the results of the fitting the different pK_a were calculated as decrived in the introduction section.

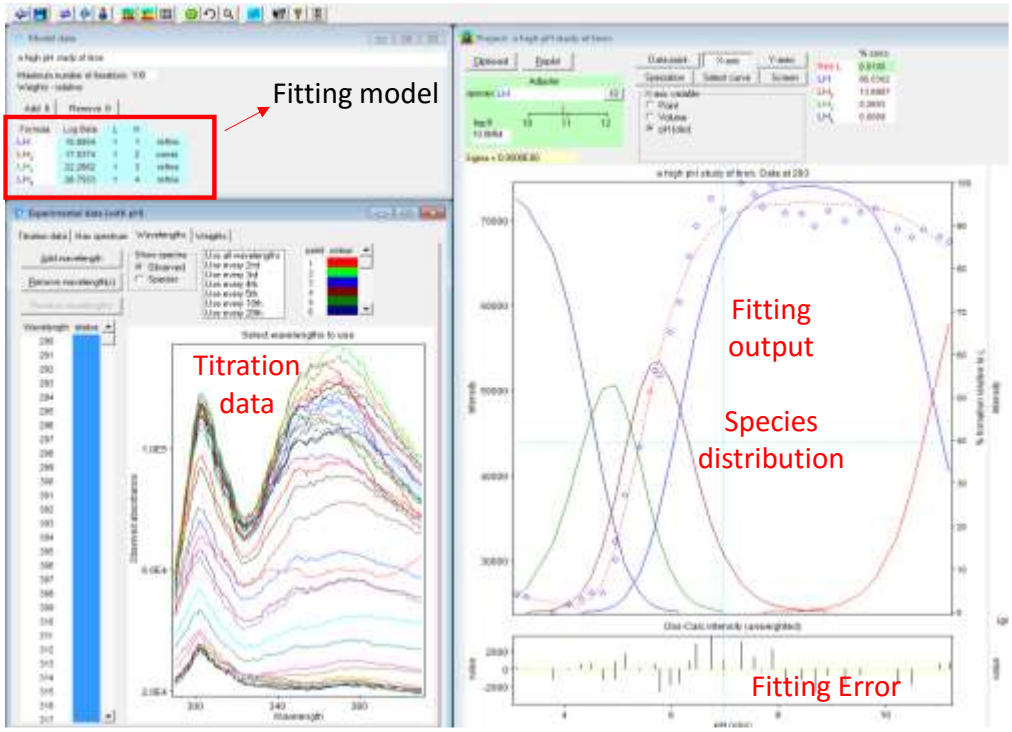


Figure 3.82. Example of the Output of a fitting with hypSpec to the different protonation states.

3.4.15 Computational methods

All molecular modelling was performed with the package Schrödinger Suite 2021,[14] through its graphical interface Maestro.[14] The software Macromodel,[15] with its default force field OPLS4[16] and GB/SA water solvation conditions,[17] was used for MM (Molecular Mechanics) energy minimization. The software Jaguar was used for DFT (Density Functional Theory) optimizations using the default B3LYP-D3 (B3LYP[18,19] plus Grimme dispersion corrections[20,21]) functional and the 6-31G** basis set, with PBF (standard Poisson–Boltzmann continuum solvation model implemented in Jaguar[22,23]) implicit water solvation.

Compound T-2FPhe-TMB was modelled with its axial nitrogen protonated and with a chloride anion bound into the cage, resulting in a neutral complex. A conformational analysis was performed using Macromodel's Mixed Torsional/Low Mode Sampling method to explore the flexibility of the 2-fluorobenzyl side chains. In this way and after removal of symmetry related conformations, 82 distinct conformers with energies within 20 kcal/mol from the lowest energy minimum were generated, which were then submitted to further DFT optimization.

The results of these DFT calculations show that the three lowest energy conformers determined correspond to conformers with 2, 3 and 1 intramolecular F-H bonds, established between the 2-fluoro aromatic substituents of one arm of the cage and the NH amino groups of the contiguous arm, with relative water solution energies of 0, 0.74 and 3.05 kcal/mol (Figure 1). Therefore, despite it cannot be discarded that a more exhaustive conformational search could render some other lower energy conformers, these results suggest that conformers with such intramolecular interactions could be relatively abundant in solution. Furthermore, it could be reasoned that they could be even more stabilized depending on the protonation state of the amino groups or in apolar environments such as lipid membranes. These conformers, in particular the one with three F-H bonds, have a more compact structure than other more extended conformations, i.e. those where the F-H interactions are absent, and it is reasonable to argue that they can be responsible for the particular properties (chromatographic mobility, chloride transport,…) that distinguish this compound from others that have a fluor in other positions of the aromatic ring, or that have other type of substituents, and that cannot adopt such compact conformations.

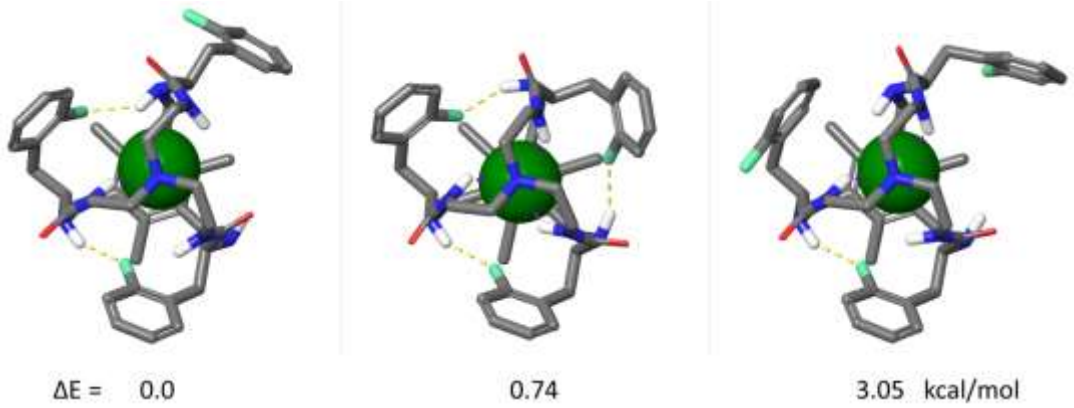


Figure 3.83. Lowest energy conformers found for compound **1h** and their relative solution energies calculated at B3LYP-D3/6-31G** level.

3.4.16 Cell culture studies

2D. Cytotoxicity assays were conducted as described in chapter 1. Media at pH 8 and 8.5 were prepared using 10 mM HEPES buffers adjusted to the desired pH.

3D models with Ca-alginate spheroids. A549 cells incubated in a T-75 flask were suspended in DMEM and mixed in a 1 : 1 proportion with a Sodium Alginate 2.4 % solution in water. Cells suspension was prepared to get beads with 20000-40000 cells each. The mixture was taken up with a syringe coupled to a 24G needle. Then it was drop by drop poured into a

102 mM $CaCl_2$ solution with magnetic stirring forming the Ca-alginate spheroids. Finally, the alginate beads containing cells were washed with a NaCl 0.9% and transferred into DMEM.

A 96 well plate was used for the cytotoxicity assay. In each well a spheroid was suspended in 100 µL of DMEM buffer. After 24 hours, the buffer was replaced by the desired buffer (pH 6.2, 6.5 or 7.5) containing different concentrations of compounds **1c** or **1h**. Beads were incubated at 37˚C with the cages for 48 hours. Then 20 µL of CellTiter-Blue were added to each well. The plate was incubated at 37˚C for 3 hour and absorbance of each well was measured using a plate reader λ = 570 nm.

3.5 References

1 I. Martí, M. Bolte, M. I. Burguete, C. Vicent, I. Alfonso and S. V. Luis, *Chemistry*, 2014, **20**, 7458–7464.

2 I. Martí, J. Rubio, M. Bolte, M. I. Burguete, C. Vicent, R. Quesada, I. Alfonso and S. V. Luis, *Chem. - A Eur. J.*, 2012, **18**, 16728–16741.

3 J. L. Wilcox and P. C. Bevilacqua, *J. Am. Chem. Soc.*, 2013, **135**, 7390–7393.

4 M. F. Hossain, C. Obi, A. Shrestha and M. F. Khan, *Mod. Chem. Appl.*, 2014, **02**, 1–7.

5 S. K. Dubey, G. Singhvi, A. Tyagi, H. Agarwal and K. V. Krishna, *J. Appl. Pharm. Sci.*, 2017, **7**, 155–158.

6 T. Gunnlaugsson, P. E. Kruger, P. Jensen, J. Tierney, H. D. P. Ali and G. M. Hussey, *J. Org. Chem.*, 2005, **70**, 10875–10878.

7 R. Ruiz, C. Ràfols, M. Rosés and E. Bosch, *J. Pharm. Sci.*, 2003, **92**, 1473–1481.

8 A. Khurana and C. Godugu, eds. B. H. A. Rehm and M. F. Moradali, Springer Singapore, Singapore, 2018, pp. 157–183.

9 T. Andersen, P. Auk-Emblem and M. Dornish, *Microarrays (Basel, Switzerland)*, 2015, **4**, 133–161.

10 S. Scheiner, eds. C. E. Dykstra, G. Frenking, K. S. Kim and G. E. B. T.-T. and A. of C. C. Scuseria, Elsevier, Amsterdam, 2005, pp. 831–857.

11 Y. Itoh, Y. Nakashima, S. Tsukamoto, T. Kurohara, M. Suzuki, Y. Sakae, M. Oda, Y. Okamoto and T. Suzuki, *Sci. Rep.*, 2019, **9**, 1–5.

12 V. Saggiomo, S. Otto, I. Marques, V. Félix, T. Torroba and R. Quesada, *Chem. Commun.*, 2012, **48**, 5274–5276.

13 N. Busschaert, M. Wenzel, C. J. E. Haynes, J. R. Hiscock, I. L. Kirby, L. E. Karagiannidis, S. J. Moore, N. J. Wells, J. Herniman, G. J. Langley, P. N. Horton, M. E. Light, I. Marques, P. J. Costa, V. Félix, J. G. Frey and P. A. Gale, *Chem. Sci.*, 2013, **4**, 3036–3045.

14 Schrödinger Release 2019-1: Maestro, Schrödinger, LLC: New York, NY, 2019.

15 Schrödinger Release 2019-1: Macromodel, Schrödinger, LLC: New York, NY, 2019.

16 C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner and E. D. Harder, *J. Chem. Theory Comput.*, , DOI:10.1021/acs.jctc.1c00302.

17 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.

18 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372–1377.

19 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

20 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

21 L. Goerigk and S. Grimme, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6670–6688.

22 B. Marten, K. Kim, C. Cortis, R. A. Friesner, R. B. Murphy, M. N. Ringnalda, D. Sitkoff and B. Honig, *J. Phys. Chem.*, 1996, **100**, 11775–11788.

23 D. J. Tannor, B. Marten, R. Murphy, R. A. Friesner, D. Sitkoff, A. Nicholls, B. Honig, M. Ringnalda and W. A. Goddard, *J. Am. Chem. Soc.*, 1994, **116**, 11875–11882.

CHAPTER 4. PSEUDOPEPTIDIC CAGES FOR TYROSINE ENCAPSULATION

4.1 INTRODUCTION

Complexation of bio-molecules has many potential biomedical applications. For instance, sensing, enzyme monitoring or drug design. Besides, it can provide insights into significant interaction mechanisms.[14]

The complexation of natural products using artificial receptors is a major topic in supramolecular chemistry. Non covalent interactions with biomolecules (amino acids, biopolymers, enzymes or nucleic acids), which are the fundaments of many pharmacological activities, can be studied using supramolecular model complexes. Also in the biochemistry field, supramolecular complexes are being used for the protection, transport or targeting of bioactive molecules. Encapsulation in a suitable host renders for instance selectivity in the protection of natural drugs from metabolism or even the manipulation of the guest redox properties.[14]

Although this chapter is mainly focused on the use of synthetic receptors, it is worth mentioning that natural structures and their derivatives can be also very efficient hosts of special interest in view of their quality for the recognition of chiral compounds. Some well-known natural antibiotics also use the supramolecular host-guest complexation. For instance, some ionophores (chapter 2) bind ions and transport them to disrupt the transmembrane ion concentration. Host-guest action mechanism is also found in glycopeptide antibiotics which bind to polypeptide intermediates and prevent the transpeptidation thus inhibiting the biosynthesis of the bacteria cell wall.[14]

Supramolecular complexation can act both increasing and decreasing the bioavailability of the desired compounds. The literature is filled with examples of supramolecular hosts for biomolecules and peptides, nucleotides, nucleosides and many other biomolecules like carbohydrates or hormones.[14] Previous research in the group of Dr. Ignacio Alfonso had been centered in the design of pseudopeptidic receptors for amino acids and short peptides in aqueous buffers because of their potential use in biological systems.[15,1]

This introductory section is focused on the described hosts for amino acids and peptides, the type of structures used for this purpose and their applications. The few examples of hosts able to selectively bind amino acids in aqueous systems encouraged us to continue working in this research line, developing new compounds and studying their properties.

4.1.1 Supramolecular hosts for amino acids and peptides

Most kinases function by interacting with other enzymes through protein-protein interactions, and this is just one example of the relevance of proteins activity for cell regulation. Consequently, the interest for controlling these interactions has grown in chemistry, biology and pharmacology fields. The control of protein-protein interactions often demands targeting a long peptide or a specific region of a protein, which are in fact very challenging targets for small molecules. Besides, protein-protein interactions are complex processes mainly driven by electrostatic and hydrophobic interactions but in which shape and configuration are the features that render selectivity.

Peptides recognition in biological media has been attracting considerable interest due to its potential applications in biology or medicine.[16] Synthetic host molecules have emerged as promising strategy to modulate protein-protein interactions thus acting as drugs. There are several examples of synthetic molecules with peptide recognition properties in the literature with structures like crown ethers, cucurbiturils, calixarenes and molecular tweezers [17] In all the reported cases, the driven forces of the interaction are non-covalent interactions like the hydrophobic effect, electrostatic interactions or hydrogen bonding.

Peptide recognition in biological media (water based) is still a growing field because of the complexity of the many interactions present in biomolecules like proteins. Water is a highly coordinating solvent which stablish strong electrostatic interactions with molecules in solution contributing to the high energetic cost of desolvation, hence ligands have to compete with water fort the guest. All those interactions must be considered for the design of synthetic hosts for peptides. For instance the hydrophobicity, which is one of the major driving forces of protein-protein interactions, has to be enough to allow the interaction and avoid the presence of water in the recognition motif, but not too high to prevent precipitation.[18]

Although compounds with specific protein interactions are still just a few, there are several examples of different structures with amino acid selectivity. Those can be the precedent for the future development of more complex structures targeting proteins through interactions between the compound and the specific amino acids in the peptidic sequence.

The progress in the development of new analysis techniques like NMR, ITC, fluorescence, crystallography and modelling allow to deeply study the interactions and sometimes predict how a modification will affect the binding. In fact, more complex systems require

more fancy analytical tools for their study so the progress in supramolecular chemistry goes hand in hand with the improvement of analytical methods.

Receptors that selectively bind specific aa residues or short peptides in biological media are interesting targets for chemists. However, examples of amino acids hosts that work in water are still limited. The interactions of both the host and the guest with water molecules through hydrogen bonding complicate the specific host-guest interactions that in this case need to be strengthened through hydrophobic, ion pair or cation-π interactions.

The huge advances in the recognition of amino acids and short peptide sequences present an opportunity to use synthetic supramolecular systems to modulate protein-protein interactions thus acting as drugs. So far, the most relevant structures described for this purpose are cucurbiturils, molecular tweezers, calixarenes and metal complexes. Despite the structural differences, they all have a very organized structure with an accessible hydrophobic cavity and are susceptible of chemical modifications.

4.1.1.1 Cucurbiturils

Cucurbiturils (CBs) are formed with 6-8 glycoluril monomers and have been used for the recognition of protein regions.[19] They have a cavity ringed with carbonyl oxygens facing inwards so they are accessible to interact with potential guests acting as hydrogen bond donors. The binding sites of CB are amphiphilic and they preferably bind cationic and hydrophobic elements. They are water soluble, easy to synthetize and have interesting binding properties in water. Cucurbiturils are made by an acidic catalysed reaction of glucoluryl with formaldehyde (**Figure 4.1**). Moreover, they are easily modified leading to the formation of their acylic analogues which are versatile and tuneable structures.

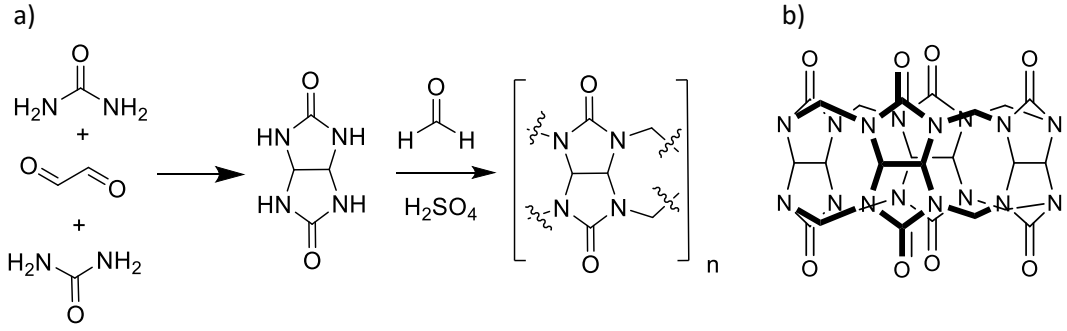


Figure 4.1. a) Cucurbiturils general synthesis. b) Structure of a CB4.

There are several examples of CBs studied as receptors. CB[7]s have a high binding affinity for aromatic residues. They have also been reported as a receptor to monitor enzymatic

reactions when using acridine orange as the indicators.[20] CB[6] complexation towards different peptides sequences was studied by Schollmeyer who came to the conclusion that CB[6] does not form inclusion complexes with dipeptides due to the polarity if the amide bond. Hence they form exclusion complexes.[21]

CBs are also used for peptide or aa sensing. CB[8] binds aromatic aa with a log *K* of 3.3 for Tyr, 3.7 for Phe and 4.6 for Trp. There are CBs also described which binds di- and tripeptides.[22] For instance, Urbach and co-workers described a CB[8] (Q8) that acts as a receptor for the tripeptide Tyr-Leu-Ala in aqueous media with nanomolar affinity.[23]

4.1.1.2 Molecular tweezers

A molecular tweezer (MT) is a host molecule with an open cavity formed by rigid curved frameworks. These structures have been used to lessen the activity of proteins. Schrader reported a molecular tweezer able to bind cationic peptides with affinities in the micromolar range.[24] Together with Klaner, Schrader also developed a molecular tweezer that binds lysine and arginine side chains of some short peptides.[25]

A successful example is the CLR01 molecular tweezer which embraces the Lys214 residue of a protein called 14-3-3σ and inhibits its binding with interaction partners as the proteins C-RafpS259 and ExoS (Figure 4.2). CLR01 has been used to inhibit the self-assembly of amyloid and is now moving to clinical trials. [26,27]

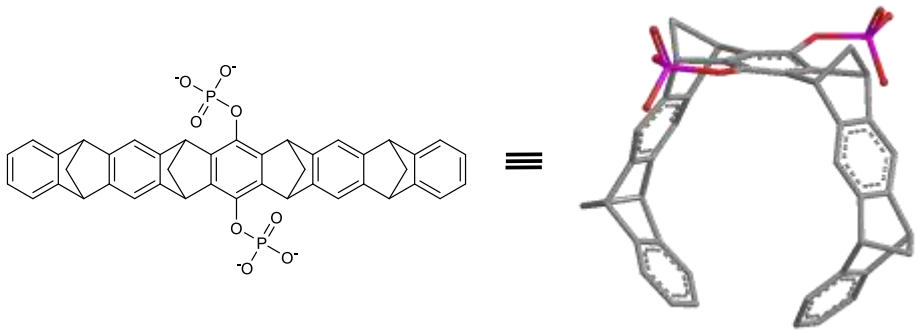


Figure 4.2. Chemical structure and 3D structure of the molecular tweezer CLR01.

4.1.1.3 Calixarenes

Calixarenes are one of the most widely studied organic hosts in supramolecular chemistry. Since the 1990's there has been a boom in the interests of calixarenes and many compounds of this type have been reported for interesting applications.

Calixarenes are compounds formed through a condensation reaction between a phenol and an aldehyde. They can be of different sizes and with different modifications in the upper and lower rims of the cavitand. This versatility of the calixarenes allows them to target a variety of biologically relevant targets by means of supramolecular interactions with their modified rims. For instance, the introduction of polar moieties to the calixarenes allows the establishment of new electrostatic interactions and hydrogen bonds. Using different modifications and adding metals there are examples of water soluble calixarenes able to recognise free cysteine, histidine, aspartic and glutamic acids.[28] *p*-Sulfonatocalix[n]arenes can encapsulate basic amino acids (Lys and Arg) either free or as part of a di- or tripeptidic sequence. The driving forces of the interaction are electrostatic and can be modulated by hydrophobic-aromatic interactions (Figure 4.3 a).[29] Zhang and co-workers reported a fully soluble *p*-sulfonatocalix[4]arene that forms a 1 : 1 complex with L-Tyrosine in water (Figure 4.3 d).[30] Another relevant work with the same *p*-sulfonatocalix[4]arene (Figure 4.3 d) lead by P. Crowley reported the crystal structure of these calixarene with a protein. *p*-sulfonatocalix[4]arene binds lysine rich cytochrome C at three different sites and acts as a mediator of protein-protein interactions.[31] Following the same strategy Crowley also reported the crystal structure of lysozyme complex with *p*-sulfonatocalix[4]arene. [32]

Hioki and co-workers developed a library of peptidocalix[4]arene and studied their interactions with different amino acids an peptides in aqueous media and found that some of their compounds were selective hosts for guest peptides.[33] They also came to the conclusion that electrostatic interactions were crucial for the binding in water.

Kalchenko et al. reported a phosphonato-calix[n]-arene that can host amino acids by hydrophobic interactions between the lateral side chain of the aa and the calixarene cavity and carboxylate-calix[n]arenes are another example of compounds with amino acid binding affinity (Figure 4.3 b).[28]

Besides their properties as amino acid binders, by smartly tuning the calixarene structures they can become hosts for bigger biomolecules like proteins. For instance, the group of Coleman demonstrated the complexation of some calixarenes to BSA.[34] Bezouska et al successfully studied a calixarene that acts an antagonist of the CD69 receptor providing a complete protection of it to further biological modifications leading to apoptosis (Figure 4.3 c).[35] Hamilton and co-workers first described a calix[4]arene that targets proteins covering a part of their surface and inhibits their protein-protein interactions.[36,37,38] Using

this strategy they described a calix[4]arene that forms a 1 : 1 complex with cytochrome C inhibiting Cyt-C Fe^{3+} reduction by ascorbate.[36] In this same direction, calix[4]arene with guanidinium groups in the upper ring described by Mendoza and co-workers were reported as inhibitors of the Shaker potassium channel.[39] One last example to illustrate the potential of calixarenes as enzymes modulators was the Calix[4]arene containing methylenebisphosphonic acid fragments reported by Kalchenko and co-workers which displays a strong inhibition of the calf intestine alkaline phosphatase by coordinating the metal cations present in the active site of this enzyme.[40]

Calixarenes have also been studied as hosts for pharmacologically active compounds and proposed as an attractive strategy in the design of new therapeutic formulations.[28] Besides, their biological activity has been explored against viruses, fungi and bacteria with promising results in all the cases (Figure 4.3e).[28]

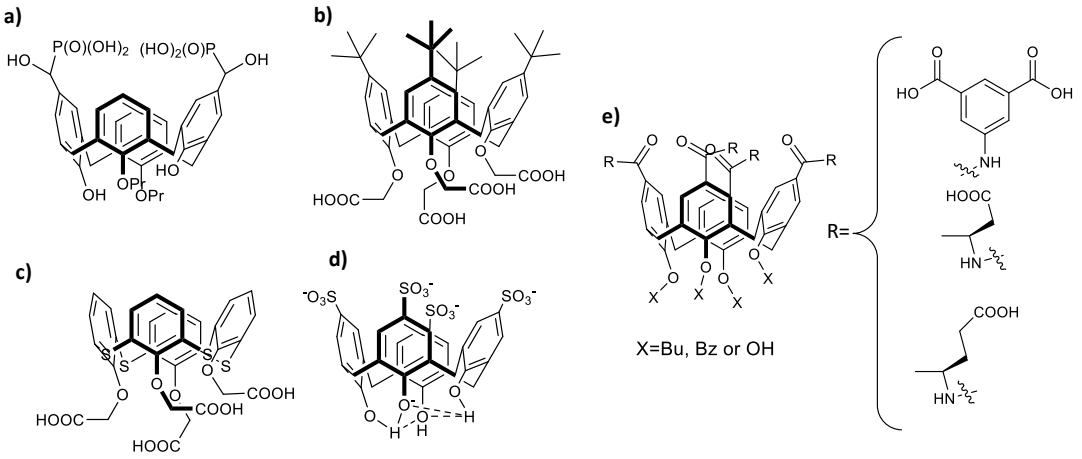


Figure 4.3. Chemical structure of the described calixarenes. a) Calix[4]arene bis-hydroxy methylphosphonic acid. b) tert-Butyl tetracarboxymethoxy calix[4]arene c) Carboxylatocalix-[4]arene with affinity for CD69 d) PSCX4 from Zhang with affinity for L-tyrosine e) Calix[4]arene carboxylic acid derivatives presenting anti-HIV and anti-HCV activities.

4.1.1.4 Metal complexes

Metals can also be used for the synthesis of synthetic receptors both for aa and peptides. For instance, Severin and co-workers reported a Ruthenium complex for detection of histidine and methionine containing peptides.[41] Fujita and co-workers used a Zinc complex

for the encapsulation of oligopeptides[42] and they showed that an increase of the ionic strength of the solution enhanced the hydrophobic interactions. More examples of metal complexes for peptide recognition can be found in the work from the groups of Wilson[43] and Nitschke,[44] work among others.[45,46,47]

4.1.1.5 Other structures

Despite the previously described structures are the most relevant ones developed for the recognition of amino acids and peptides, other types of molecules have also been described with peptide binding properties. For instance, the cyclophane reported by Dougherty and co-workers which strongly binds Arg in aqueous media by cation-π interactions.[48] Schmuck introduced the guanidiniocarbony purrole group as an efficient oxoanion binding site and used it in a dicationic hosts containing Ser and Lys described as a receptor for Ala containing dipeptide carboxylates in aqueous media (Figure 4.4 a).[49] One last example is the first synthetic receptor for RGD (Arg-Gly-Asp) peptide in water reported by Schrader designed as a combination of an arginine host (Figure 4.4 b)[50] with a linker containing primary ammonium residues able interact with the aspartic residues.[51]

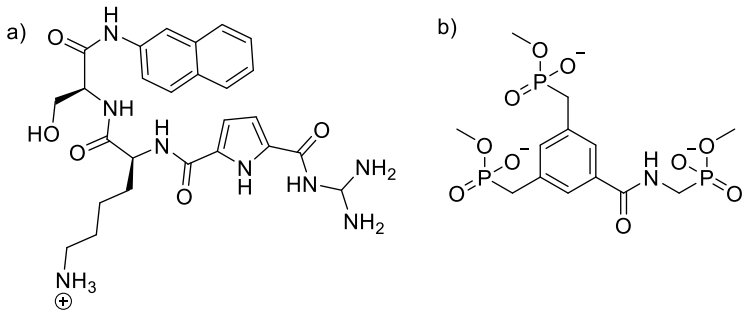


Figure 4.4. Structure a) Ala receptor b) Arginine receptor

One of the research lines in the supramolecular chemistry group at IQAC is focused on the development of pseudopeptidic cage-like macrocyclic compound for the recognition of Tyr residues as it will be explained in the next section.

4.1.1.6 Targeting proteins

When looking into the bibliography, a few examples of compounds that interact with bio-relevant proteins can be found. Huc and co-workers described a foldamer that binds carbonic anhydrase (enzyme responsible of CO_2 hydration) with nanomolar affinity.[52] Sulphonamides can also be used to inhibit the activity of this enzyme.[53,54]

4.1.1.7 Applications in sensing

Another application of interest of peptide-host supramolecular structures is the design of optical sensors. Small molecule sensors can provide selectivity for the recognition site. In this direction, Liu and co-workers reported a supramolecular tandem (cavitand) capable of discriminating between phospho-tyrosine and phosphate linked to other structures. To do so, they synthetized a cavitand capable of binding peptide motifs and the sensing was by fluorescence displacement processes. Besides they studied the effect of metal ions on the modulation of the selective sensing.[55] Given its huge biological importance, many peptide sensors have been developed.[56]

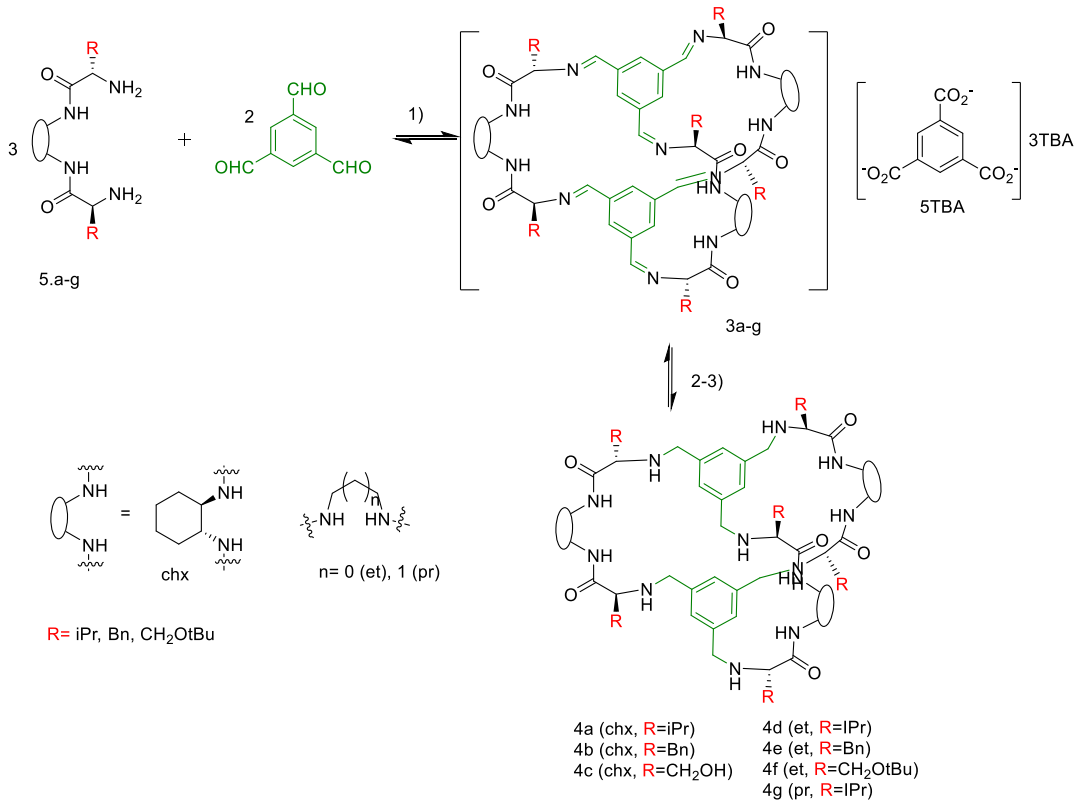
4.1.2 Background in the supramolecular chemistry group

Previous research in the supramolecular chemistry group at IQAC was focused on the efficient obtaining of macrocyclic structures with supramolecular properties of interest for biological applications.

Inspired by different works on macrobicyclic peptides like vancomycin, we envisioned that the introduction of amino acid derived chemical moieties (pseudopeptides) to macrocycles could increase the functional possibilities leading to future applications in biological systems.[57]

As it was previously mentioned in the general introduction, the main drawback in the preparation of molecular cages is to find a synthetic strategy that leads to the desired product in high yield and allowing structural diversity. In 2012, More et al[57] described the efficient synthesis of macrobicyclic pseudopeptides with two strategies, using organic anion templates and driving the formation of the product by controling the configuration of the building blocks. They described the synthesis of a family of pseudopeptidic macrobicycles with relatively high yields (they obtained enough quantity for different supramolecular studies) using a 3+2 imine condensation to form the thermodynamically most stable product, followed by the in situ reduction of the imine group (Scheme 4.1.). The methodology used for the formation of the hexaimine intermediate was DCC thus the cage formed was the most stable in the experimental conditions used. It is worth mentioning that this one pot reaction involves the formation of 6 covalent bonds in two steps which means a total of 12 reactions. This reaction can be carried out using a template but the final products are also obtained without template if the mixture is allowed to reach the equilibrium for enough time (24 h).

The critical step is the formation of the hexaimine intermediate by dynamic covalent chemistry. This successful process was followed by NMR which surprisingly indicated the formation of only one cyclic compound pointing out the good selectivity or this reaction. The variety of moieties (ciclohexane-1,2-diyl, ethylenyl, propilenyl) and side chains (*i*Pr, Bn, CH_2OH) all with good yields (even in the case of the CH_2OH side chain that requires and additional synthetic step to deprotect the alcohol) increases the potential of this dynamic covalent system for the preparation of highly different cage-like pseudopeptidic structures with different chemical properties (polarity, size, flexibility) for numerous applications.



Scheme 4.2. Synthesis of pseudopeptidic cages. 1) Equilibrium at RT either in the absence or the presence of 5TBA. 2)BH_3·py 3) HCl (scheme from reference [57])

Further efforts were made aiming to use these cages for the recognition of short peptides tuning the side chains to modulate the affinity for different substrates. Different cages prepared following the synthesis described in Scheme 4.3 with different moieties and side chains were tested against N-protected dipeptides.[6] The interactions were measured through different techniques (NMR, MS and in some cases fluorescence) which allowed the determination of the association constant for each cage-receptor pair. Very interesting conclusions were extracted from this work. It was demonstrated that cyclohexane gives a rigidity to the moiety that is beneficial for the recognition of some dipeptides. Moreover, the side chains (R substituents) of the cage also play an important role in the recognition. For instance, Ser and Val enhance the interaction whereas Phe residues reduce it. This deep interaction analysis marked a clear direction to further explore the possibilities of these structures. Cage-peptide interactions were higher for the dipeptides containing aromatic rings (Phe and Tyr). The best guest was Ac-EY-OH due to the additional H-bond properties of the electron-rich aromatic ring. Besides, interactions with sequences containing Tyr

residues can be measured through fluorescence titrations experiments, which is a methodological advantage.

Cages with the cyclohexane-1,2-diyl scaffold were the ones further developed and still under study. Since these compounds are chiral, selectivity for dipeptides bearing different configuration (L,D) were performed with two cages (CySer and CyThr) with the dipeptide (Ac-YE-OH). Synthetic pseudopeptidic cages were proven to be better binders for dipeptides containing naturally occurring amino acids (L,L configuration) being the configuration of the two amino acids (Y and E) relevant for the binding. A binding model was proposed for the interaction using all the data collected from the different interaction studies (NMR, MS and molecular modelling analysis) based on the type of interactions expected for this complex (electrostatic contacts, hydrogen bonding, π–π, hydrophobic and steric interactions, Figure 4.5).

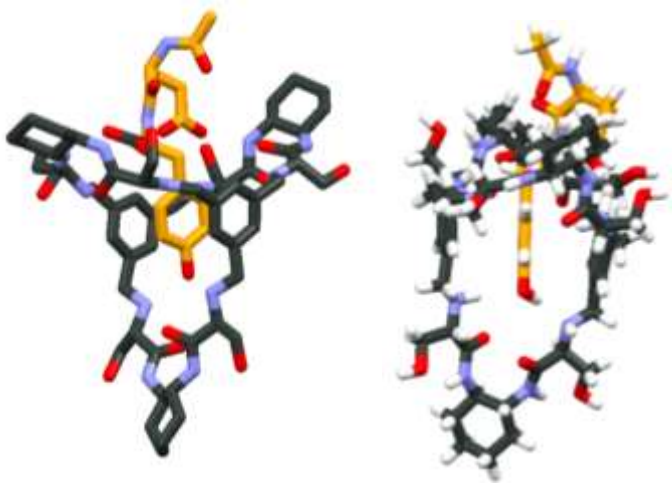


Figure 4.5. Binding model for the recognition of the dipeptide Ac-L-Glu-L-Tyr-OH with CySer.

Given the promising results, new cage-like compounds were synthetized with Lys and Orn residues. . In 2016 E. Faggi *et al*. reported a deep study of these interactions between cages and both polypeptides containing several tyrosine residues short Tyr containing peptide. They showed better affinity for Tyr residues in short peptides and also in a polypeptide ($PolyE_4Y$). This is relevant because the EYE epitope is a known substrate for protein tyrosine kinases (PTKs). The affinity observed of both CyOrn and CyLys for this sequence was better than the one of the other cages (CySer, CyThr) due to the positive charge of the Lys and Orn residues that allows electrostatic interaction between glutamic residues of the

polypeptide with the cage. CyLys and CyOrn were proven to inhibit PTK activity by protecting the Tyr from phosphorylation using a commercial assay kit. The kinase inhibition was in agreement with the K_d found for CyLys and CyOrn and showed that CyLys is the most efficient cage for kinase inhibition in this system.[1]

This chapter is focused on widening this family of pseudopeptidic cage-like compounds and the study of their interactions with different polypeptides through different techniques (fluorescence, NMR and MS. Spec).

4.2 OBJECTIVES

The main objective is the synthesis and study of pseudopeptidic receptors for the molecular recognition of specific peptide sequences with biological relevance. This goal has been divided in the following points:

a. Synthesis and characterization of large pseudopeptidic cages with different side chains able to interact with specific peptide sequences. The supramolecular chemistry group has already described the synthesis and the binding properties of a family of pseudopeptidic cages that are able to selectively recognise specific peptides sequences. Here our aim is to prepare a family of different large cages with different spacers and side chains to improve their performance.
b. Assess the binding properties of the prepared pseudopeptidic cages toward different peptide sequences. NMR and fluorescence spectroscopy techniques will be used for this purpose.

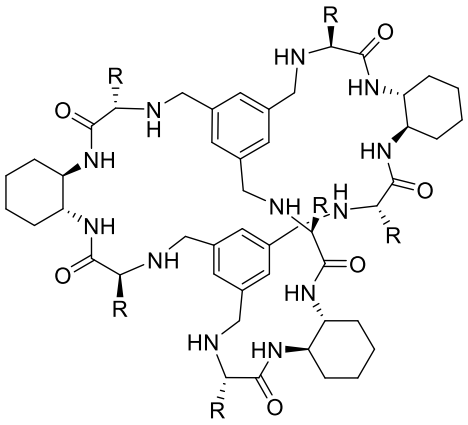


Figure 4.6. Structure of the proposed large pseudopeptidic cages.

4.3 SELECTION OF THE NEW CAGE-COMPOUNDS

In the background section of this chapter we described the pseudopeptidic cages studied so far by the supramolecular chemistry group. They include cages with uncharged residues derived from valine and alanine and positively charged residues at physiological pH derived from lysine and ornithine (with a pK_a around 10).

To have a more complete family of these compounds we decided to synthetize cages with histidine residues (one protonation point with a pK_a close to the physiological pH) and aspartic and glutamic residues, which are expected to be mostly deprotonated at physiological pH since the pK_a of their side chains is around 4.

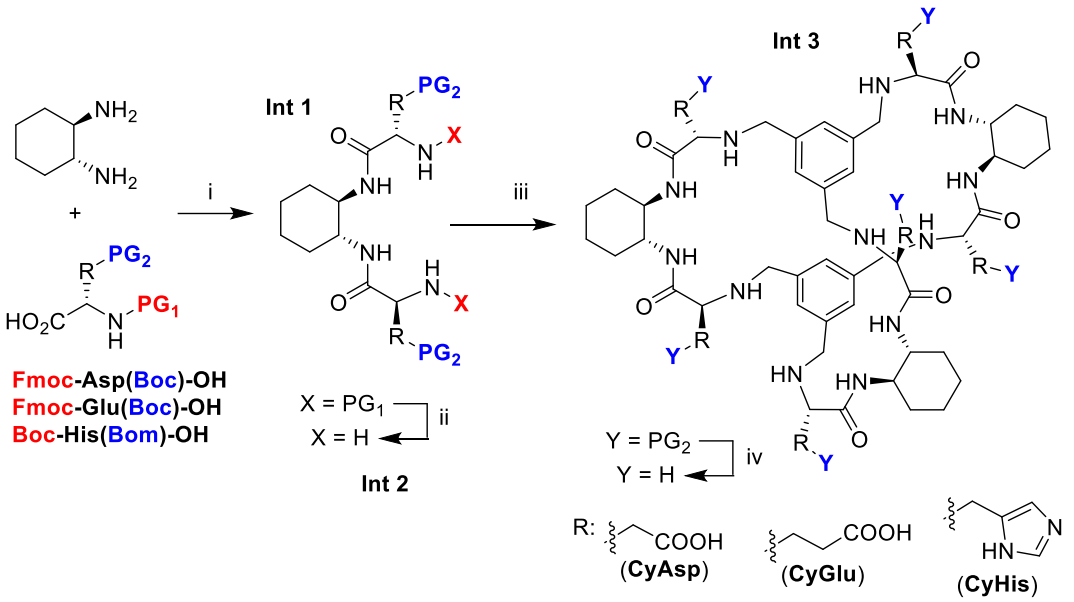
Our initial hypothesis was that histidine cage may induce stronger interactions with tyrosines in solutions due to its intermediate protonation state at pH 7.5. On the contrary, aspartic and glutamic were chosen because given the studied interaction of CyLys and CyOrn with polyE_4Y and short peptides containing tyrosine and glutamic residues, we hypothesized that aspartic and glutamic cages would interact better with tyrosines surrounded by positively charged residues. Besides, as in the case of CyLys and CyOrn, aspartic and glutamic only differ from each other in a CH_2 but this minimal difference might have important influence on the interaction properties of the macrocycles proposed worth studying.

4.4 SYNTHETIC PROCEDURE

Using the reported synthesis of **CyOrn** and **CyLys** as a reference,[1] the synthesis of **CyHis, CyAsp** and **CyGlu** was accomplished following the general procedure depicted in Scheme 4.4. The key macrobicyclization is a [3+2] reductive amination reaction, where the structural preorganization of the bis(amidoamine) precursor governs the process.[58] The rest of the steps in the synthetic scheme correspond to amide coupling and orthogonal deprotection reactions as in conventional solution phase peptide synthesis. **CyOrn** and **CyLys** were synthetized following the previously described protocol[1] whereas **CyAsp**, **CyGlu** and **CyHis** are new receptors, for which synthetic and characterization details are given in experimental section. Final products purification was carried out through reverse phase chromatography using H_2O/ACN/0.1 % TFA. TFA in the salts of the cages obtained was then exchanged by Cl^-. The concentrations of each purified cage was determined by NMR since they are hydroscopic salts.

For each cage, the selection of the protecting groups of the NH_2 and the amino acid side chain were crucial. The side chain protecting group must allow the 3+2 cyclization to occur and be orthogonal to the NH_2 protecting group. The selection of the protecting groups for each cage are detailed next.

Regarding the overall yield for the synthesis of these cages, it is worth mentioning that in each synthetic step there are six chemical reactions occurring simultaneously including in the 3+2 cyclization which is the most critical step which takes place with reasonably high yields. Hence the synthesis of these compounds is considered to proceed in overall good yields.



Scheme 4.5. General synthetic pathway for the synthesis of the pseudopeptides cages: (i) HBTU, DIPEA in DMF; (ii) Piperidine in DMF (Asp and Glu) or TFA/TES in CH_2Cl_2 (His); (iii) benzene-1,3,5-tricarbaldehyde in MeOH (12 h at RT) and then $NaBH_4$; (iv) TFA/TES in CH_2Cl_2 (Asp) or H_3PO_4/TES in CH_2Cl_2 (Glu) or TFA/TFMSA/p-cresol (His).

4.4.1 CyHis

CyHis was synthetized following the route described in Scheme 4.5. Firsts attempts to obtain this cage were conducted using the Fmoc-His(Trt)-OH but the 3 + 2 cyclization reaction rendered very poor yields and mixtures hard to separate so this strategy was discharged. This was probably caused by the steric hindrance of the trityl protecting group.

Fortunately, using Boc to protect the primary amine and Bom to protect the histidine, CyHis was obtained with an overall 25% yield.

Benzyloxymethyl group (Bom) is a protecting group of specific interest for the protection of histidine residues that can be removed by HF, TFMSA or hydrogenolysis and is completely stable to bases and nucleophiles.[59] For this synthesis the removal of the Bom group was performed using TFMSA.

The final product was purified trough reverse phase chromatography and obtained as a TFA salt. Since TFA is not a usual counter-ion in biological systems, it was exchanged by Cl^- for the different interaction assays. To do so, **CyHis**·TFA was dissolved in methanol containing HCl and evaporated several times. The elimination of TFA was confirmed by ^{19}F-NMR. Quantitative analysis of the cage was done by ^{1}H-NMR to determine which percentage of the solid was **CyHis** since in these salts obtained from lyophilization there are always counter-ions and high amounts of water.

4.4.2 CyAsp

The starting amino acid for the synthesis was Fmoc-Asp(O*t*Bu)-OH since Fmoc and *t*Bu are completely orthogonal protecting groups. The final product was obtained with an overall yield of 27 % as a TFA salt. TFA counter ions were exchanged by Cl^- as described for **CyHis**.

4.4.3 CyGlu

The starting amino acid selected for the synthesis was Fmoc-Glu(O*t*Bu)-NH_2 to use the same synthetic route as for **CyAsp**. Unfortunately the deprotection of the carboxylic acid residues (last synthetic step) turned out to be the trickiest step towards **CyGlu**. CyGlu(*t*Bu) was then isolated trough column chromatography before *t*Bu removal. The 1st deprotection attempt using TFA/DCM leads to the cyclization of the side chain of glutamic detected by UPLC-MS and Maldi-Tof mass spectrometry (Figure 4.7). After trying different methodologies changing reaction times, temperature and different acids, it was found that reaction with phosphoric acid at room temperature for 1 hour yields **CyGlu** and avoids the cyclization of the carboxylic residues.

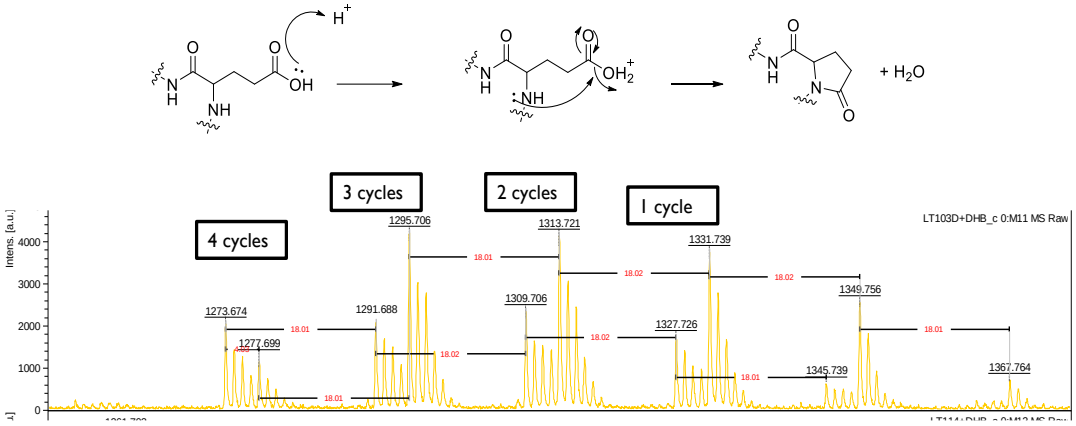


Figure 4.7. Mechanism for the cyclization reaction of the glutamic residue of CyGlu in the presence of an acid and Maldi-Tof spectra of the mixture containing CyGlu with zero to 4 residues in the cyclic form.

In this case TFA was not exchanged for Cl^- because even in mild conditions some cyclization stars to occur so this compound was quantified by NMR in its TFA salt form. The quantitative analysis showed that the resulting solid contains 49 % of CyGlu mixed with TFA and water molecules.

4.4.4 CyOrn and CyLys

Synthesis was conducted as described in the previous paper of the group.[60] The purified cages obtained as TFA salts were converted into their corresponding HCl salts and the concentration of cage in the resulting solids were also determined by quantitative ^{1}H-NMR.

4.5 Interaction studies with Tyr containing polypeptides

4.5.1 Peptide selection to study the interaction between pseudopeptidic cages and Tyrosine: poly (Glu, Tyr) 4:1, Poly (Lys, Tyr) 4:1, Poly (Glu, Lys, Tyr) 6:3:1

The different pseudopeptidic cages were designed to be hosts for tyrosine. To test the affinity of these new compounds towards tyrosine we selected 3 copolymers Poly (Glu, Tyr) 4:1, Poly (Lys, Tyr) 4:1, Poly (Glu, Lys, Tyr) 6:3:1 and the trimer Ac-EYE-NH_2 as Tyr containing peptides models and studied their interaction with the cages.

Poly (Glu, Tyr) 4:1 is a random copolymer which contains L-glutamic acid and L-tyrosine in 4:1 ratio and it is generally isolated as a sodium salt.

The activity of PTK can be assayed by analysing the autophosphorylation of the transphosphorylation of artificial substrates (see chapter 5). In the beginning of the studies with kinases, the biological substrate for each kinase was unknown so they used artificial substrates to asses kinase activity.[61] An important type of substrates where copolymers because they are prepared in a relatively easy manner. PolyE_4Y is one of the artificial substrates that is still being used nowadays.[62]

PolyE_4Y is a well-known substrate for EGFR (epidermal growth receptor factor) but it is also a good standard substrate for c-Src (two well-known and biologically relevant kinases).[61] In fact, it has been proven to be substrate for almost all the kinases so it could be used to assess their activity or as a positive control.[63] The good affinity of this polypeptide for kinases is associated to the presence of glutamic acid residues. Since this is a well-known substrate, we also decided to study the interactions with the short peptide Ac-EYE-NH_2. This is the simplification of the biding motif of the polypeptide and is useful to proof the interaction mechanism proposed between the cage and tyrosine.

Poly (Lys, Tyr) 4:1 is also a random copolymer which contains L-lysine and L-tyrosine in 4:1 ratio and it is generally isolated as a bromide salt. It has lower affinity in general for all the kinases since it contains lysine instead of glutamic which implies a global positive charge of the polypeptide which are detrimental to substrate specificity. It can be used as a negative control since it is expected to have lower affinity for all the kinases.

Poly (Glu, Lys, Tyr) 6:3:1 (PolyE_6K_3Y) is also a random copolymer, not commonly used in kinase assays (not many examples in the literature) but it's a commercially available peptide sold as sodium salt and it is in between of the two previously described polypeptides in terms of charge. The availability of this peptide and the presence of both lysine and glutamic residues make it an interesting substrate to study with the new cage-like compounds prepared.

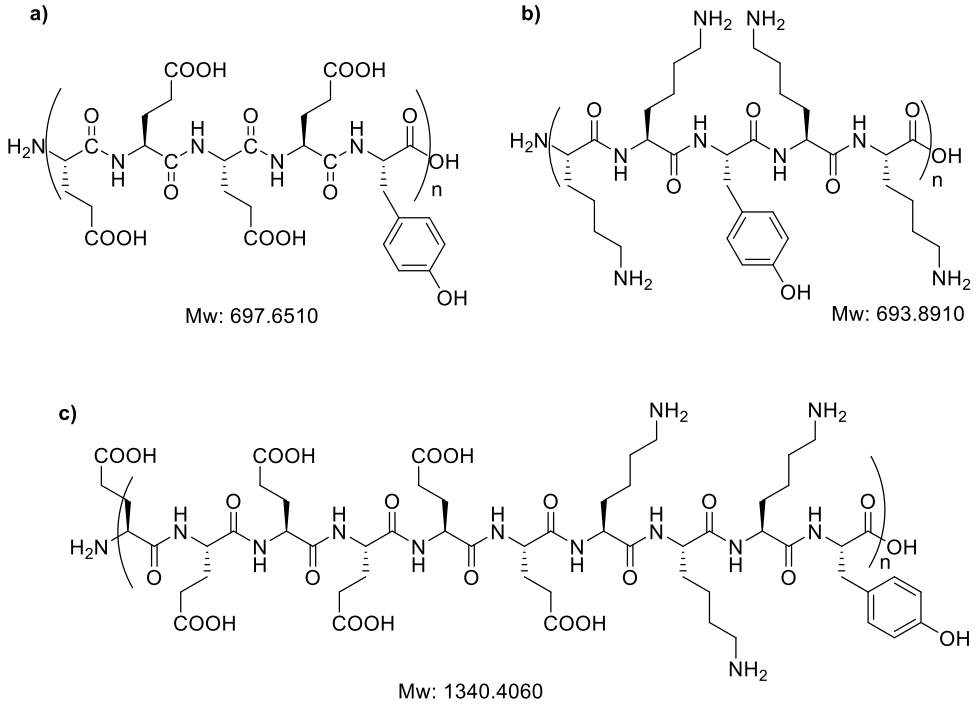


Figure 4.8. Chemical structure of the copolymers a) PolyE_4Y, b) PolyK_4Y and c) PolyE_6K_3Y.

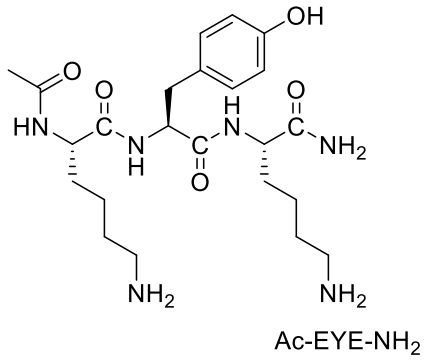


Figure 4.9. Chemical structure of Ac-EYE-NH_2

4.6 INTERACTION STUDIES:

The interaction of these pseudopeptidic cages with the selected copolymers was studied mainly with three different techniques: Fluorescence, NMR and MS in the gas phase.

The association constants between the different host-guest pairs were determined using fluorescence titrations and the data was carefully analysed for each host-guest pair.

4.6.1 Determination of the association constants using fluorescence titrations

The determination of the association constants between two species is essential to study supramolecular interactions. In the general introduction it was already explained how this can be achieved using techniques like NMR and fluorescence spectroscopy. Besides, it was also mentioned the software HypSpec and HypNMR for the fitting of the titration data to the selected interaction model.

The main technique used for the study of supramolecular interactions between the selected polypeptides and the pseudopeptidic cages was fluorescence spectroscopy. Different situations can be encountered when trying to fit the titration data to an appropriate interaction model. To ensure the accuracy of the fittings the Stern-Vomer plots and two different fitting methods were used when necessary.

4.6.1.1 Stern-Volmer (S-V) analysis

The fluorescence decrease of a fluorophore due to collisional quenching in the absence of other types of quenching can be detected by representing F_0/F vs the concentration of titrant agent. If the resulting plot can be adjusted to a linear equation the main contribution to this fluorescence decrease at the represented wavelength can be to either mainly binding or collisional quenching but not both simultaneously (section 1.7). However, at relatively high concentrations of the substrate (µM range) fluorescence decrease is often associated at least partially to collisional quenching. If during the titration of a guest with a host the only spectroscopic change detected is the fluorescence decrease of a band associated to the guest, the S-V plot allows to determine if this quenching is purely collisional or also due to host-guest interactions.

4.6.1.2 Fluorescence data fitting

The two methods used for data fitting in this chapter were the use of equation 4.1 and the use of HypSpec software. For most of the titrations, the K_{diss} was determined with the two different methods getting a very similar result in both cases. However, as it will be explained next, the most accurate and with less error fittings were obtained with HypSpec software.

The fitting with Eq. 4.1 requires the data from an emission band that appears after the addition of the host to the guest. [64] The equation is derived from the different species concentrations in a 1 : 1 equilibrium system. The use of eq. 4.1 is only possible for the simplest interaction model: a + b ⇆ ab and it can only be used at one wavelength of the

spectra which are their main limitations. That is why we also considered an alternative and more accurate method to determine K_{diss}.

$$\Delta F = F - F0 = \left(\frac{\Delta Fmax}{[P]}\right) \times \frac{[P] + [C] + K_D - \sqrt{([P] + [C] + K_D)^2 - 4[P][C]}}{2}$$

Equation 4.1. [P] and [C] are the total concentrations of peptide and cage, respectively. F is the fluorescence of the system with a given [C]. *F0* is the fluorescence of the system when C = 0. *ΔF*max and K_D are the fitting parameters.

The software **HypSpec** was the alternative fitting method used for the fitting of the titration data to a proposed interaction model for each cage-peptide pair. HypSpec is based on multicomponent analysis, one of the most accurate methods available for the fitting of titration data.[65,66] Hence, it allows the global fitting of the whole emission band (or the range desired) for each titration point to the interaction model selected thus reducing the error in the fitting.

4.6.1.3 Fluorescence titrations

Titrations of each copolymers with each cage were conducted. In every case, the concentrations of the compounds were adjusted to get a proper fitting (enough points with sufficient concentration of the complex formed) and avoid a precipitation events. All the Tyr containing polypeptides have a characteristic fluorescence emission band centred at 302 nm corresponding to the tyrosine fluorescence emission.

In this section when we quote peptide concentration we are referring to the concentration of tyrosine in the peptide solution. For this work, all the tyrosines present in each polydisperse co-polymer were considered equivalent thus equally available for the formation of a 1 : 1 complex with each of them. For the preparation of the peptide stock, the molecular weight of the smallest peptide subunit containing one tyrosine was used but the exact concentration was determined through absorbance.

Throughout the titration with the different cages, two main phenomena were detected in the fluorescence spectra:

- The fluorescence of the tyrosine at 302 nm decreases or increases in the presence of the cage depending on the cage-peptide pair and the concentrations used for the titration experiment.

- A new fluorescence emission band at higher wavelengths appears as a result of the formation of the cage-tyrosine excimer. The intensity of this band and its shape is strongly dependent on the cage compounds.

Tyrosine fluorescence decrease was at least partly associated with collisional quenching since this phenomena was detected to be more intense at higher peptide concentrations. For the different titrations, the representation of the S-V plot was used to interpret this observation. Deviations from the linearity indicate that some other events are affecting the Tyr emission, for instance the formation of a complex with the cage.

In the titrations where a new band appears, the fitting of the data was centred in this new band since it can only be associated to the formation of a complex regardless of the fluorescence chances in the Tyr fluorescence band.

Titrations where the Tyr fluorescence band decrease is the only observed change are considered individually to determine whether if complexation is taking place and how to measure or study it.

The previous reported data for the K_d values towards the different peptides with **CyLys** and **CyOrn** had been calculated with equation 4.1. Since the determination of association constants is more accurate using HypSpect, the fitting of these titrations was repeated with this alternative method. It was proven in all the cases that the K_d obtained is very similar to the reported one with lower error, which reinforces the value of the available data. Additionally, for comparative purposes we considered necessary to use the same methods for all the fittings.

4.6.2 Titrations with CyHis

This section contains a detailed explanation of the titrations of the different polypeptides with **CyHis** and the data analysis carried out in each case. **CyHis** was selected as an example to explain the diverse analysis required for the determination of the K_d between the cages and the polypeptides. All the other cages were studied following the same steps (details of each titration are in the experimental section).

4.6.2.1 Titration of PolyE$_4$Y by CyHis

A $2.0 \cdot 10^{-4}$ M solution of PolyE$_4$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a $1 \cdot 10^{-3}$ M solution of **CyHis**.

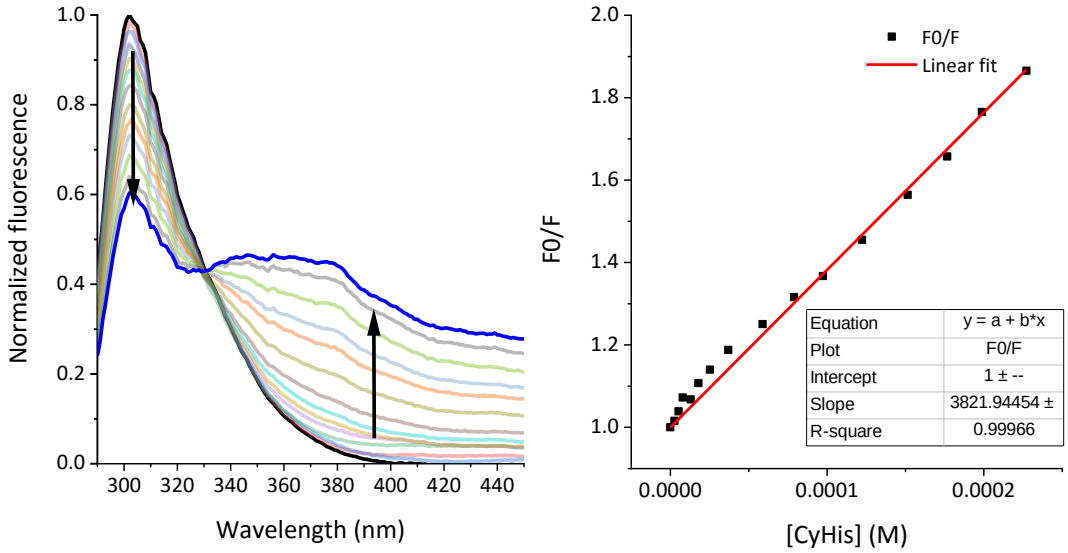


Figure 4.10. Normalized emission spectra of PolyE$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyHis** ([CyHis] = 0-0.5 · 10^{-3} M). [PolyE$_4$Y] = 2 · 10^{-4} M. λexc = 276 nm (left). Stern-Volmer plot of the titration of PolyE$_4$Y by **CyHis** at 302 nm and linear Fit of the data (right).

The slope of the S-V plot linear regression corresponds to the K_{sv}. In this case it is K_{sv} = 3821 ± 60 M^{-1}. Log (K_{sv}) = 3.58, 1/K_{sv} = 261 µM.

The firsts points of the titration indicate a small deviation of the S-V plot from the linearity but the overall titration allows a good fitting of the data to a linear equation. From this analysis we can conclude that the variation of the Tyr emission through the titration was mainly due to dynamic quenching and that the specific binding does not contribute significantly to this band change.

The new band that appears in the titration (380 nm) was associated to the complex formation thus the determination of the K_d was mainly done using the data form this new band. The fitting with eq. 4.1 was done using the maximum emission wavelength of the excimer at 380 nm (Figure 4.11) and the K_d obtained from this fitting was: 61 ± 35 µM.

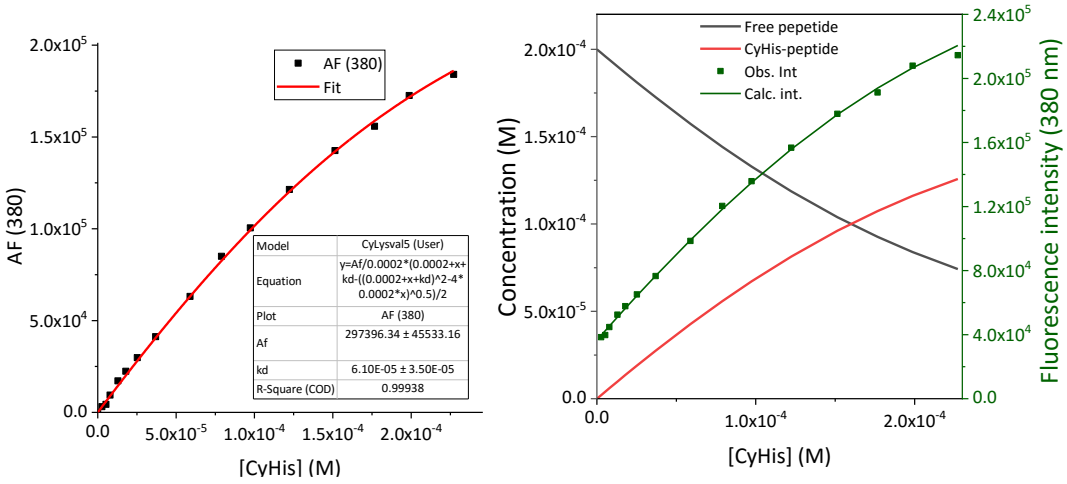


Figure 4.11. Fitting of the titration of PolyE$_4$Y with **CyHis** at pH 7.3 using eq.4.1 and fluorescence at 380 nm (left). Species distribution and fluorescence intensity at 380 nm (observed and calculated) for the titration of PolyE$_4$Y with **CyHis** at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec (right).

This fitting was also conducted with HypSpec using the whole spectra for each point of the titration.

HypSpec result:

$$\text{CyHis + Peptide} \leftrightarrows \text{CyHis -peptide Log } \beta = 4.22 \pm 0.04,\ K_d = 60 \pm 6\ \mu\text{M}.$$

The increase of a band with its maximum at 380 nm is an evidence of the formation of an excimer between **CyHis** and the peptide. The data from the titration can be fitted either with Eq.4.1 and Origin (using the data at 380 nm since it's the maximum emission of the new species formed) or using HypSpec. The K_d obtained with the two fitting strategies gives the same result but the fitting with HypSpec is more accurate since it uses the whole fluorescence band without making any simplification and reducing the error of the K_d. Because of that, K_d determined through HypSpec is the one used for the results analysis and discussion of each cage-peptide pair.

4.6.2.2 Titration of AC-EYE-NH$_2$ by CyHis

A $2.0 \cdot 10^{-4}$ M solution of Ac-EYE-NH$_2$ in buffered water (50 mM TRIS pH = 7.3) was titrated with a $1.0 \cdot 10^{-3}$ M solution of **CyHis** (Figure 4.12).

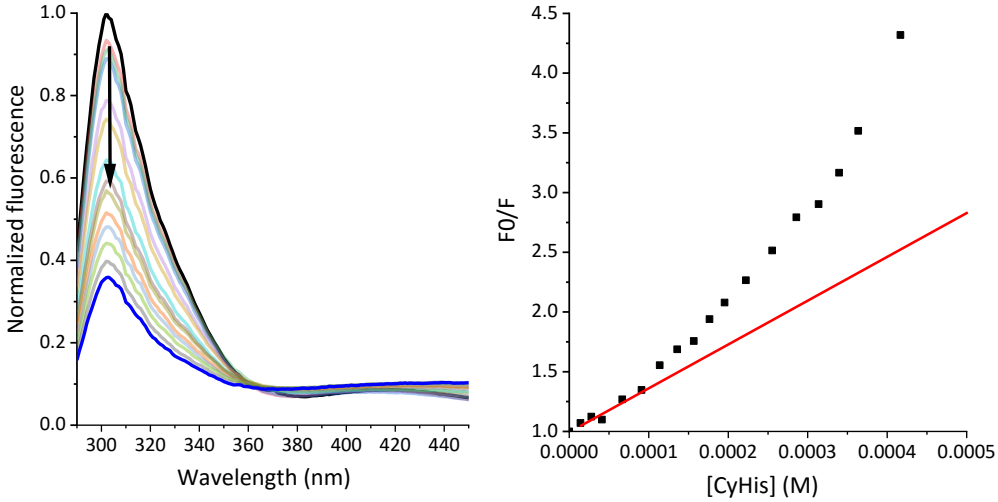


Figure 4.12. Normalized emission spectra of Ac-EYE-NH_2 in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyHis** ([CyHis]= 0-0.5 · 10^{-3} M). [Ac-EYE-NH_2] = 2 · 10^{-4} M. λexc =2 76 nm (left). Stern-Volmer plot of the titration of Ac-EYE-NH_2 with **CyHis** at 302 nm and linear fit of the data (solid red line) (right). Dashed red line represents the trend that the data should follow if it really had a linear response.

The absence of a fluorescence band corresponding to the excimer complicates the analysis of the interaction and the determination of binding constants in this case.

The S-V plot has a clear up-deviation from linearity typical in systems where complexation occurs (Figure 4.12). So the Tyr fluorescence variation is a result of both collisional quenching and complexation simultaneously. In this case since complexation is clearly affecting this band we can use HypSpec for the fitting of the data even there is no a new band corresponding to the formation of a complex (figure 4.12). We believe that the emission of the Tyr is much higher than that of the complex thus we are not able to see the emission band of the excimer in this concentration range.

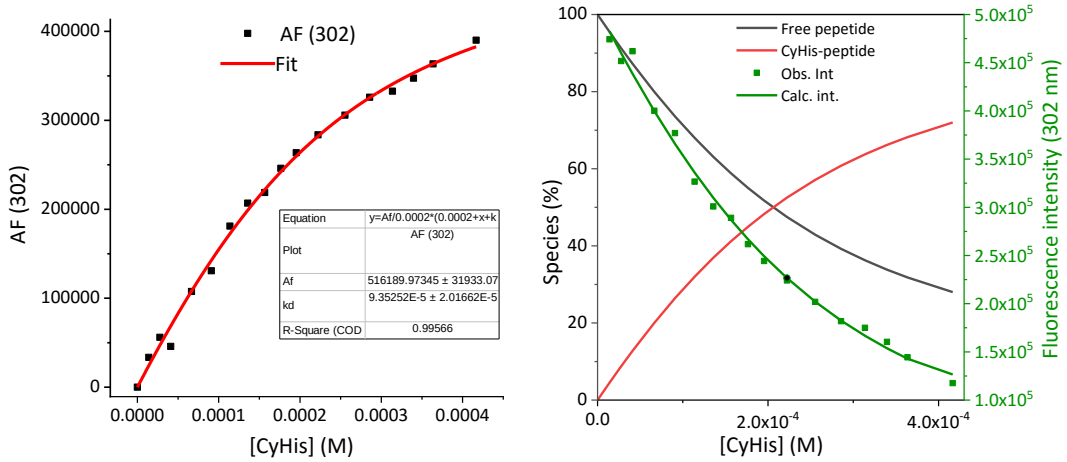


Figure 4.13. Fitting of the titration of Ac-EYE-NH_2 with **CyHis** at pH 7.3 using equation 4.1 and fluorescence at 302 nm. The K_d determined with the fitting was 93 ± 20 µM (left). species distribution and fluorescence intensity at 302 nm (observed and calculated) for the titration of Ac-EYE-NH_2 with CyHis at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec.

HypSpec result:

$$\text{CyHis} + \text{Peptide} \leftrightarrows \text{CyHis-peptide} \quad \text{Log } \beta = 4.023 \pm 0.002, \; K\text{d} = 94 \pm 4 \; \mu\text{M}.$$

The same titration was conducted with a more diluted peptide solution to better study this interaction and ensure the validity of the deductions applied for the fitting in the absence of an excimer band.

A $2.0 \cdot 10^{-5}$ M solution of Ac-EYE-NH_2 in buffered water (50 mM TRIS pH = 7.3) was titrated with a $1 \cdot 10^{-3}$ M solution of **CyHis** (HCl salt) (Figure 4.14).

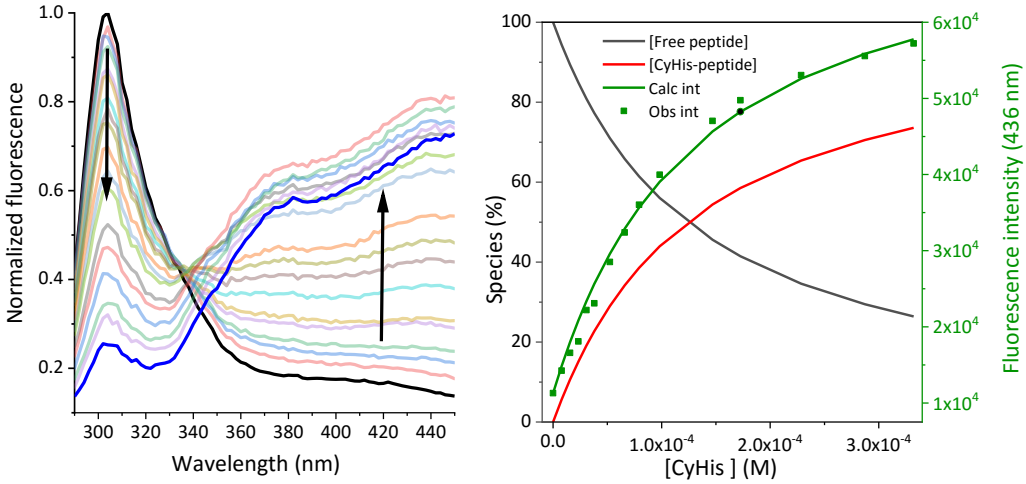


Figure 4.14. Normalized emission spectra of Ac-EYE-NH_2 in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyHis** ([CyHis] = 0-1 · 10^{-3} M). [PolyK_4Y] = 2 · 10^{-5} M. λexc = 276 nm (left). Species distribution and fluorescence intensity at 436 nm (observed and calculated) for the titration of Ac-EYE-NH_2 with **CyHis** at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec (right).

The S-V plot in this case at 301 nm (figure 4.15) can be properly adjusted to a linear equation which indicates that the fluorescence quenching of the Tyr band is mainly due to dynamic quenching events (K_{sv} = 2818 M^{-1}, log K_{sv} = 3.44). However, the band corresponding to the formation of an excimer is well appreciated in these conditions. The fitting of the titration data to a 1 : 1 model was conducted with HypSpec.

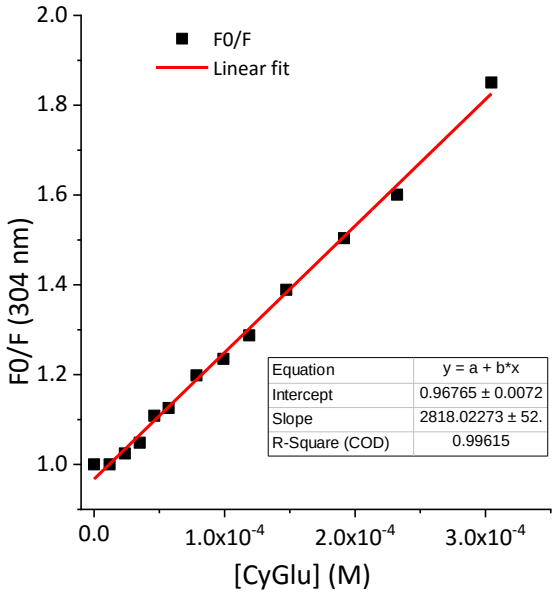


Figure 4.15. Stern-Volmer plot of the titration of Ac-EYE-NH_2 (20 µM) by **CyHis** at 302 nm and linear Fit of the data.

HypSpec result:

$$\text{CyHis} + \text{Peptide} \leftrightarrows \text{CyHis-peptide} \quad \text{Log } \beta : 3.94 \pm 0.01, K_d = 114 \pm 10 \text{ µM}.$$

This titration with a more diluted solution of the peptide proofs that a complex is being formed and that the K_d for this complex can be determined with both 200 and 20 µM peptide solutions since very similar results are obtained in both fittings.

4.6.2.3 Titration of polyK₄Y by CyHis

A $2.0 \cdot 10^{-4}$ M solution of PolyK_4Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a $1 \cdot 10^{-3}$ M solution of **CyHis** in its HCl salt.

The addition of **CyHis** to polyK_4Y only affects the fluorescence spectra in the Tyr emission band by reducing it. The representation of the S-V plot for this band gives a non-linear relation between [CyHis] and F_0/F associated to the formation of a complex between the Tyr and CyHis. HypSpec was used to fit this data to a 1 : 1 complex model (Figure 4.16).

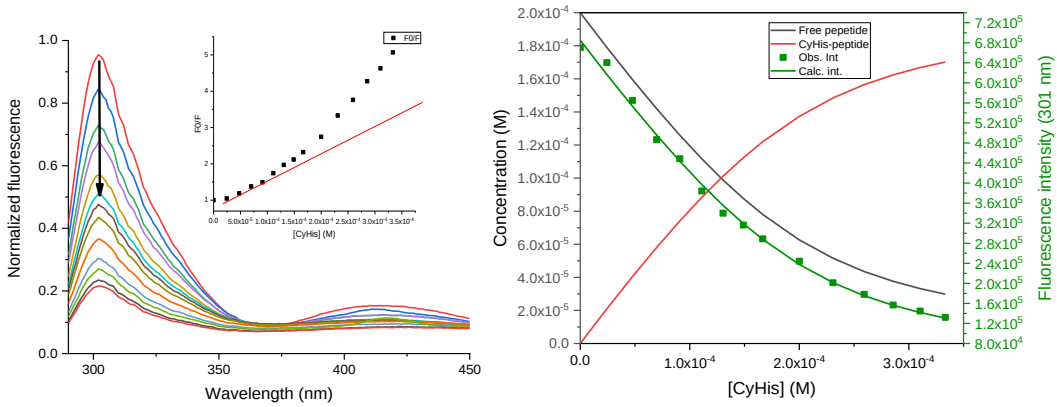


Figure 4.16. Normalized emission spectra of PolyK$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyHis** ([CyHis] = 0-0.4 · 10^{-3} M) and S-V plot at 301 nm (red line corresponds to a linear reponse). [PolyK$_4$Y] = 2 · 10^{-4} M. λexc = 276 nm (left). Species distribution and fluorescence intensity at 301 nm (observed and calculated) for the titration of PolyK$_4$Y with **CyHis** at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec (right).

HypSpec result:

$$\text{CyHis} + \text{Peptide} \leftrightarrows \text{CyHis-peptide} \quad \text{Log } \beta = 4.43 \pm 0.02,\ K_d = 28 \pm 2\ \mu\text{M}.$$

The same titration was conducted with a 20 µM solution of PolyK$_4$Y (Figure 4.17). In this titration we observed both the decrease of the tyrosine emission band and the increase of a new fluorescence band that we associate to the CyHis-Tyr complex. Titration data was fitted to a 1 : 1 model with HypSpec and the K_d obtained was 93 ± 2 µM.

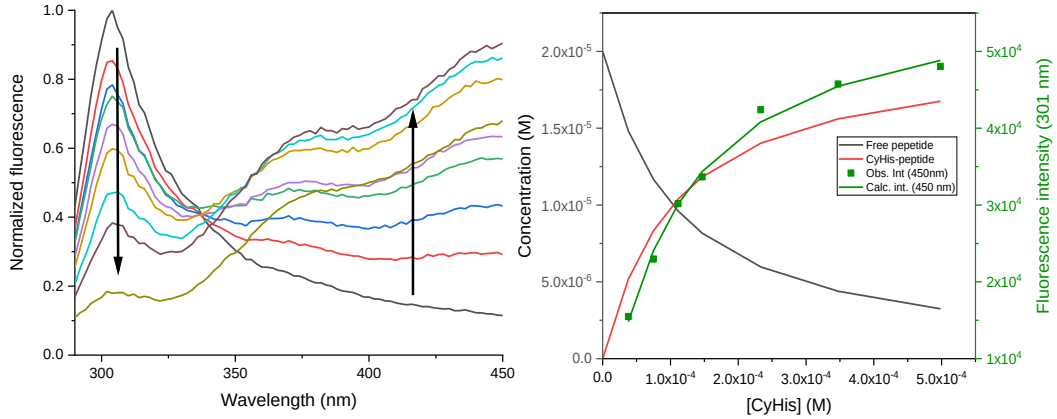


Figure 4.17. Normalized emission spectra of PolyK$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyHis** ([CyHis] = 0-0.5 · 10^{-3} M). [PolyK$_4$Y] = 2 · 10^{-5} M. λexc = 276 nm (left). Species distribution and fluorescence intensity at 450 nm (observed and calculated) for the titration of PolyK$_4$Y with **CyHis** at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec (right).

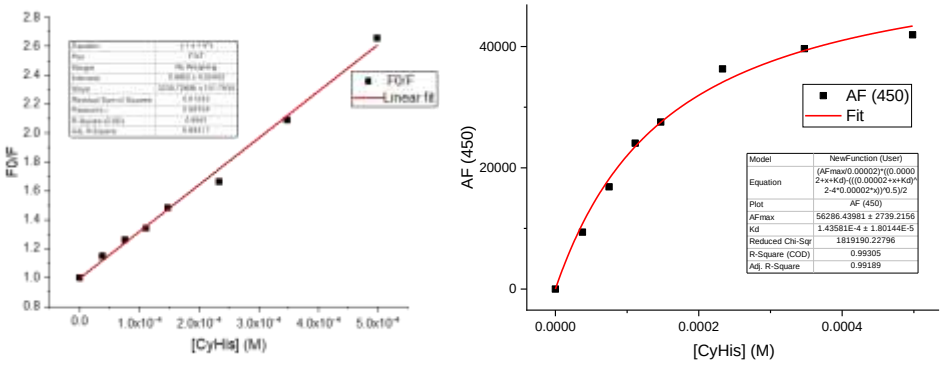


Figure 4.18. S-V plot of the titration of PolyK$_4$Y (20 µM) by CyHis at 302 nm and linear Fit of the data. K_{sv} derived from the linear fitting is 3231 ± 101 M^{-1} (left) Fitting of the titration of PolyK$_4$Y (20 µM) with CyHis at pH 7.3 using equation 4.1 and fluorescence at 450 nm (right) The Kd determined with the fitting was **143 ± 20 µM.**

HypSpec result:

$$\text{CyHis} + \text{Peptide} \leftrightarrows \text{CyHis-peptide} \quad \text{Log } \beta = 4.08 \pm 0.008,\ K_d = 93 \pm 2\ \mu\text{M}.$$

S-V plot of the band at 302 nm fits to a linear equation and the data can also be fitted with equation 4.1 but with a higher error associated to the fitting (figure 4.18).

Two conclusions were extracted from the two titrations of PolyK$_4$Y at 20 and 200 µM.

- Even in the absence of a band corresponding to the formation of the excimer, data can be fitted to a 1 : 1 model when the S-V plot is clearly deviated from linearity.
- The K_d obtained in the two titrations is slightly different but within the same order of magnitude. We kept the value obtained from the second titration since the fitting when a band grows is more accurate because there are probably less events affecting the band apart from the complexation. However, the K_d obtained in the first titration would also be useful for comparative purposes.

4.6.2.4 Titration of polyE$_6$K$_3$Y by CyHis

A 2.0 · 10^{-5} M solution of polyE$_6$K$_3$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 0.85 · 10^{-3} M solution of **CyHis** in its HCl salt. In this titration we also observe a new fluorescence band growing with a similar profile to the previous ones as a result of the formation of a complex. Data was fitted with HypSpec giving a K_d of 194 ± 3 µM.

S-V plot gives a non-linear relation and the fitting with eq. 4.1 gives result of the K_d slightly higher than the one obtained with HypSpec (figure 4.20).

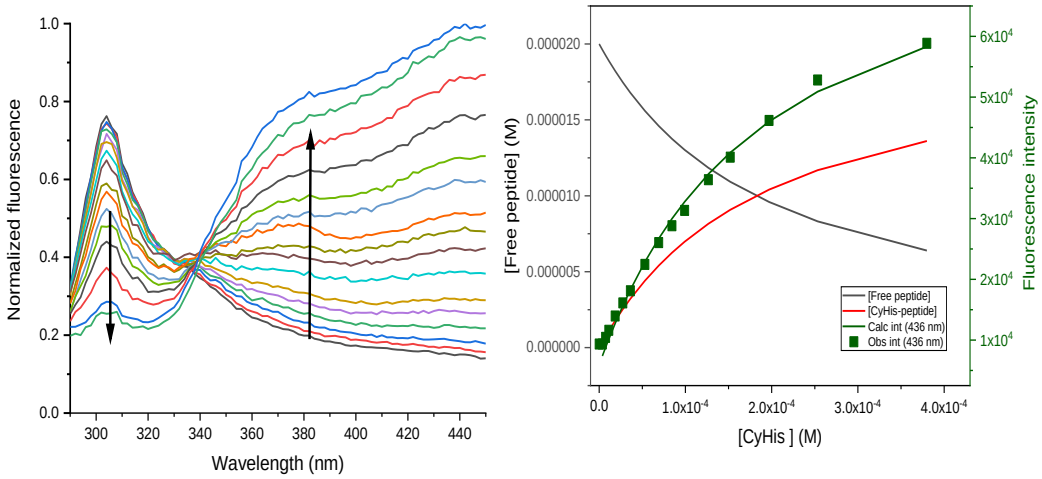


Figure 4.19. Normalized emission spectra of PolyE$_6$K$_3$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyHis** ([CyHis] = 0-0.4 · 10^{-3} M). [PolyE$_6$K$_3$Y] = 2 · 10^{-5} M. λexc = 276 nm (left). Species distribution and fluorescence intensity at 436 nm (observed and calculated) for the titration of PolyE$_6$K$_3$Y with CyHis at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec (right).

HypSpec result:

CyHis + Peptide ⇆ CyHis-peptide Log β = 3.710 ± 0.005, K_d = 194 ± 3 µM.

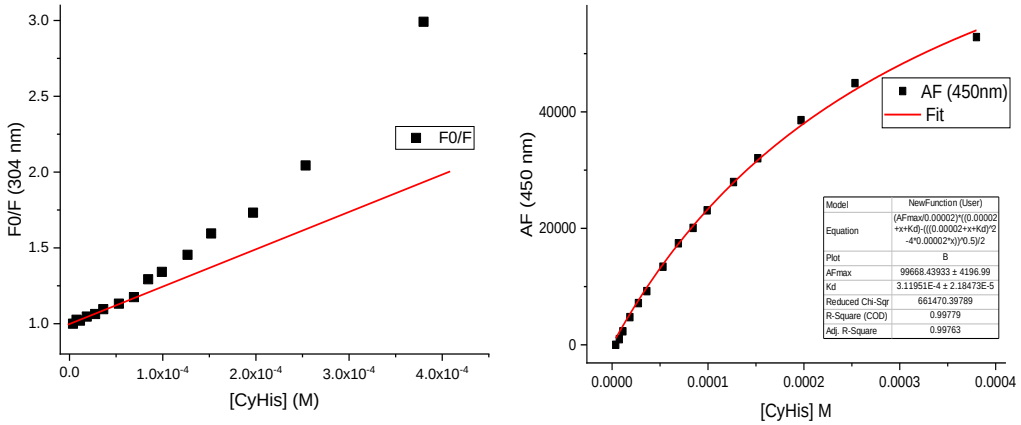


Figure 4.20. S-V plot of the titration of $PolyE_6K_3Y$ (20 µM) by **CyHis** at 304 nm. Red line would be the trend if the response were linear (left). Fitting of the titration of $PolyE_6K_3Y$ (20 µM) with **CyHis** at pH 7.3 using equation 1 and fluorescence at 450 nm. The K_d determined with the fitting was **311 ± 22 µM** (right).

4.7 RESULTS AND DISCUSSION

4.7.1 Determination of the K_d using fluorescence

The same strategy described for the titrations, fitting and analysis described for **CyHis** was used for the other pseudopeptidic cages with the selected polypeptides. Table 4.1 summarizes the K_d for the different cage-peptides determined through fluorescence spectroscopy titrations.

Table 4.1. Dissociation constants (K_d, μM) for the complexes formed by polypeptides and cages obtained by fluorescence excitation at 276 nm (50 mM TRIS, pH 7.3).

	$PolyE_4Y$	AC-EYE-NH_2	$PolyK_4Y$	$PolyE_6K_3Y$
CyOrn	920 ± 24	740 ± 26	No fit.	114.5 ± 2
CyLys	450± 20	398 ± 16	S-V linear ≥ 825	27.5 ± 0.6
CyHis	61 ± 6	114 ± 10	93 ± 2	194 ± 3
CyAsp	S-V linear ≥ 1225	2055 ± 42	398 ± 10	S-V linear ≥ 461
CyGlu	1737	4365 ± 190	400 ± 10	S-V linear ≥ 525

Throughout the fluorescence titrations of $polyE_4Y$ (200 μM) with the different cages, we observed that **CyHis** (entry 3, table 4.1) shows the strongest binding towards Tyrosine in this polypeptide with a K_d of 61 μM followed by **CyLys** and **CyOrn** (K_d was determined with HypSpec fitting the titration data from previous work).[1] The K_d of **CyAsp** with $polyE_4Y$ cannot be accurately determined in the titration conditions used since in this case there is only a fluorescence decrease of the Tyr fluorescence band upon the addition of the cage and the S-V plot of this band gives a linear response which indicates that this change in emission can be mainly associated to dynamic quenching (Figure 4.21). However, the inverse of the K_{sv} can be considered the lower limit of the K_d which in this case is 1225 μM. A similar situation was found for **CyGlu**. The change in fluorescence associated to the excimer was very small and the fitting to a 1 : 1 complex was not possible (there might be some other events affecting the fluorescence like aggregation or strong electrostatic interactions/repulsions). In this case we conducted the titration at a lower concentration

of polyE$_4$Y (20 µM) and successfully fitted the collected data to the 1 : 1 interaction model with the polypeptide and the K_d derived from the fitting was found to be 1737 µM.

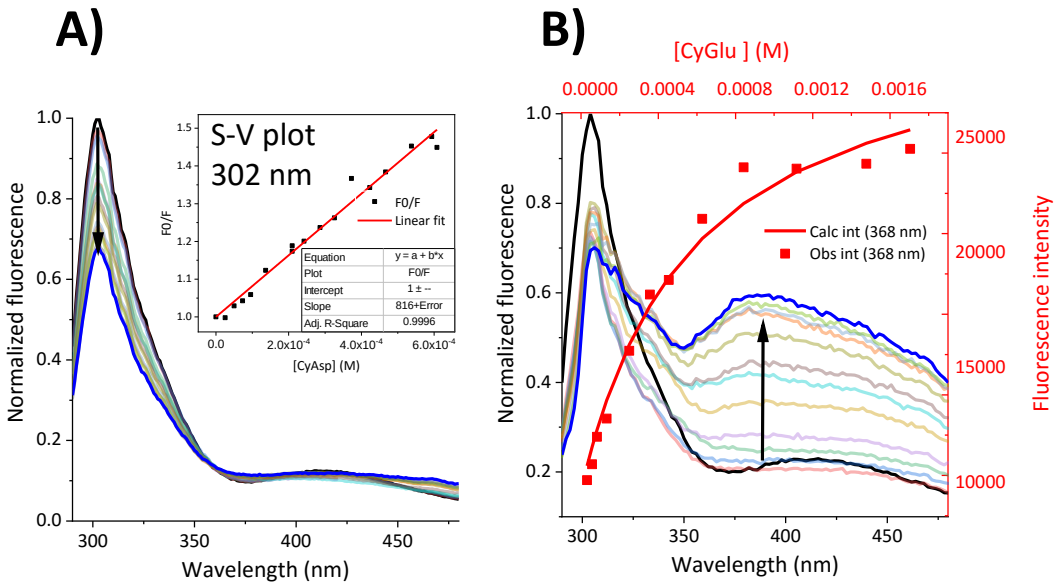


Figure 4.21. Normalized emission spectra of polyE$_4$Y in buffered water (50 mM TRIS pH = 7.3; λexc = 276 nm, in absence and in presence of different amounts of: A) **CyAsp** ([CyAsp] = 0-1 · 10^{-3} M), [polyE$_4$Y] = 2 · 10^{-4} M and S-V plot at 302 nm. B) **CyGlu** ([CyGlu] = 0-2 · 10^{-3} M), [polyK$_4$Y] = 2 · 10^{-5} M.

The interactions between polyE$_4$Y and the cages were associated with the electrostatic interactions between the glutamic residues of the polypeptide and the R substituents of the cages. **CyLys** and **CyOrn** have a stronger binding due to the positive charge of their R substituents which can better interact with the glutamic residues next to the tyrosine whereas the carboxylic residues of **CyAsp** and **CyGlu** might be producing a repulsion effect reducing the affinity of those cages for the tyrosine. The strong interaction with **CyHis** was associated with the pK_a of histidine being close to the physiological pH which allow this cage to be in different protonation states thus enhancing the interaction with tyrosine by modulating its protonation state to its environment.

When we studied the interactions with the short peptide AC-EYE-NH$_2$ (binding motif of polyE$_4$Y) we found a very similar trend as for the polypeptide. The association constants were also in the same range with **CyHis** being the strongest binder with a K_d = 114 µM, followed by **CyLys** and **CyOrn**. **CyAsp** and **CyGlu** showed a weaker binding with a K_d of 2055

and 4266 µM respectively. The weak interaction in this case lead to a less reliable fitting since in both cases **CyAsp** and **CyGlu** less than 60% of complexation is reached during the titration (Figure 4.22). Anyway the results are in agreement with the results obtained using polyE$_4$Y. The previously reported results which indicate that polyE$_4$Y can be used as a model polypeptide for the EYE motif in an aqueous buffer at physiological pH for this family of compounds were corroborated with the newly synthetized cages. For the different cages studied, the K_d with AC-EYE-NH$_2$ and with PolyE$_4$Y are in the same range and the relative strength between the cages is the same.

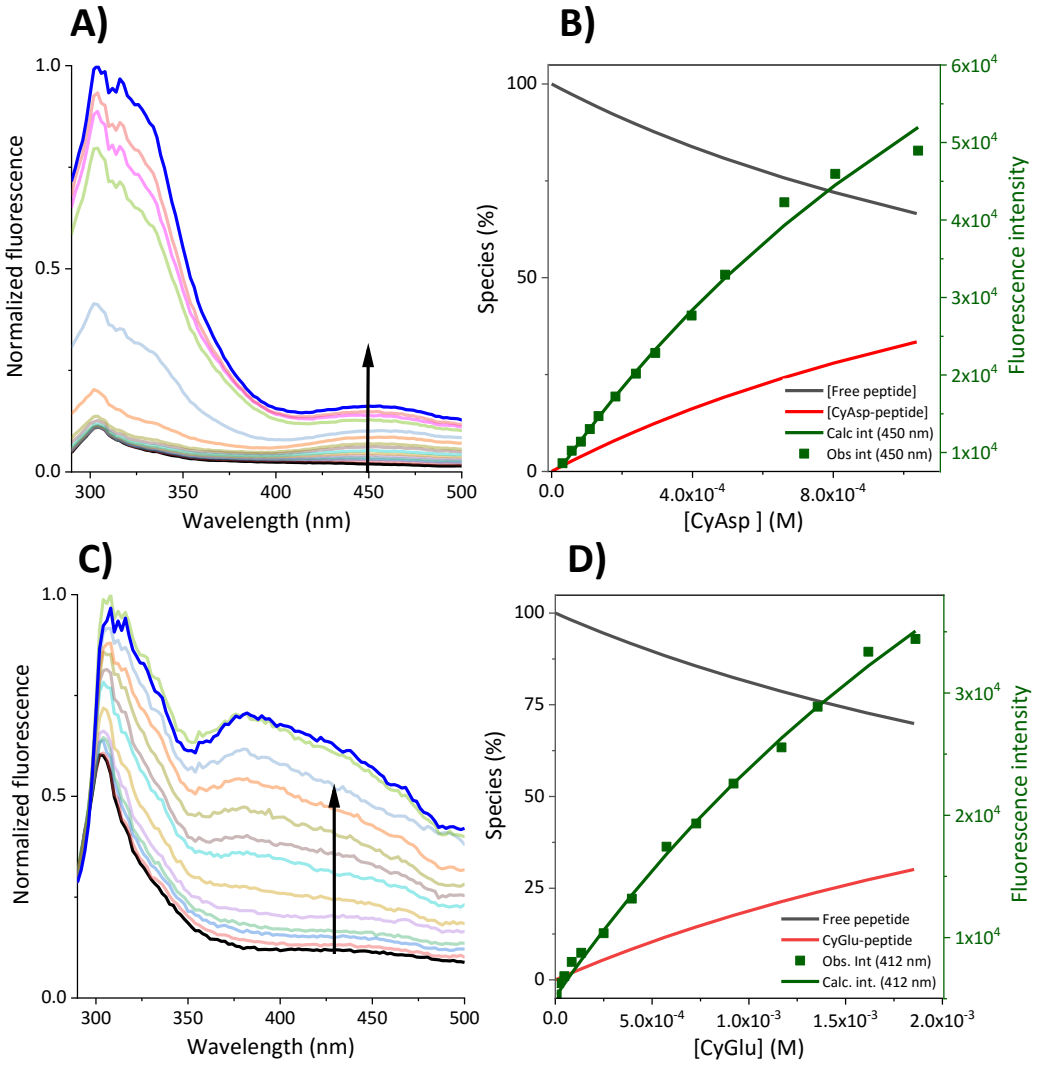


Figure 4.22. Normalized emission spectra of AC-EYE-NH$_2$ in buffered water (50 mM TRIS pH = 7.3; λexc = 276 nm, in absence and in presence of different amounts of: A) **CyAsp** ([CyAsp] = 0-1 · 10^{-3} M), [AC-EYE-NH$_2$] = 1 · 10^{-5} M; B) **CyGlu** ([CyGlu] = 0-2 · 10^{-3} M), [AC-EYE-NH$_2$] =

$2 \cdot 10^{-5}$ M. Species distribution and fluorescence intensity (observed and calculated) for the titration of AC-EYE-NH_2 with: C) **CyAsp** D) **CyGlu** obtained from a fitting model for the formation of complex 1 : 1 using HypSpec**.**

The results obtained from the titrations with the positively charged polypeptide polyK_4Y (table 4.1, 5th column) were in agreement with the interaction model proposed and the affinity dependence on the electrostatic interactions between the side chains of the cages and the amino acids nearby the tyrosine in the peptides. The interactions between polyK_4Y and **CyOrn/CyLys** were so weak that they could not be determined through fluorescence titrations regardless of the cage/polypeptide concentrations used. In the titration with **CyLys** only the emission band at 302 nm (emission of the tyrosine) decreased upon the addition of the cage and the S-V plot at this wavelength fits to a linear regression indicating that the major contribution to this fluorescence change is dynamic quenching. Hence, the interaction between **CyLys** and the Tyr of this peptide is too weak to be determined (Figure 4.23). The interaction between polyK_4Y and **CyOrn** was measured at two different concentrations of the peptide. Using 200 μM polyK_4Y data collected shows a non-linear S-V plot for the Tyr band and an isosbestic point at 355 nm. A particularity in the titration of polyK_4Y 20 μM with **CyOrn** during the titration was observed: in the first addition the Tyr emission band decreases but after the 3th point the emission in the whole registered range increases with the **CyOrn** concentration (Figure 4.24). A band centred at 375 nm seems to be appearing that could be associated with the formation of the excimer but the global fluorescence increase hides it. The data collected from this titration cannot be fitted to a simple 1 : 1 interaction model. The fluorescence changes observed might be associated to non-specific interactions such as electrostatic repulsions that affect the electronic transitions of the fluorophores.

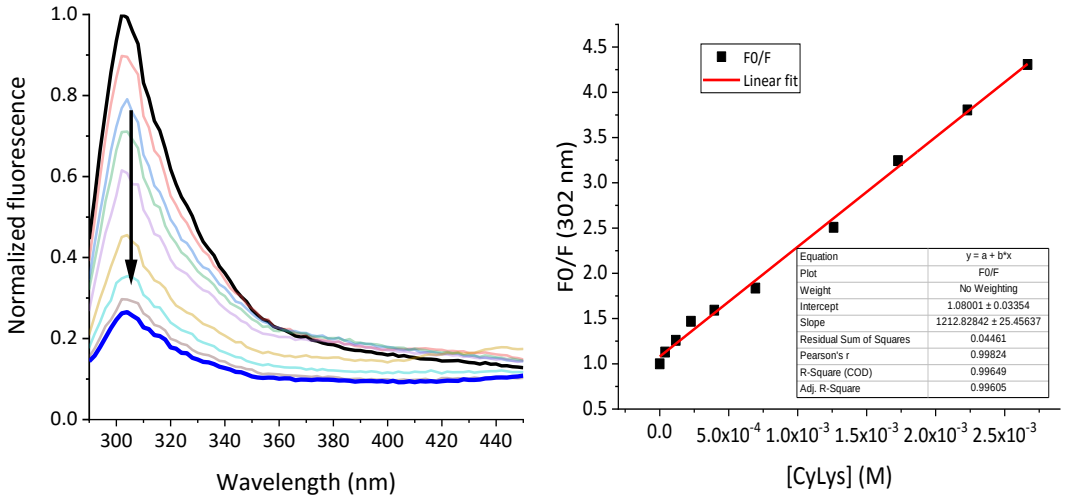


Figure 4.23. Normalized emission spectra of polyK$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyLys** ([CyLys] = 0-2.3 · 10^{-3} M). [polyK$_4$Y] = 2 · 10^{-5} M. λexc = 276 nm (left) and S-V plot of the data, $K_{S\text{-}V}$ = 1212 ± 25 M^{-1} (right).

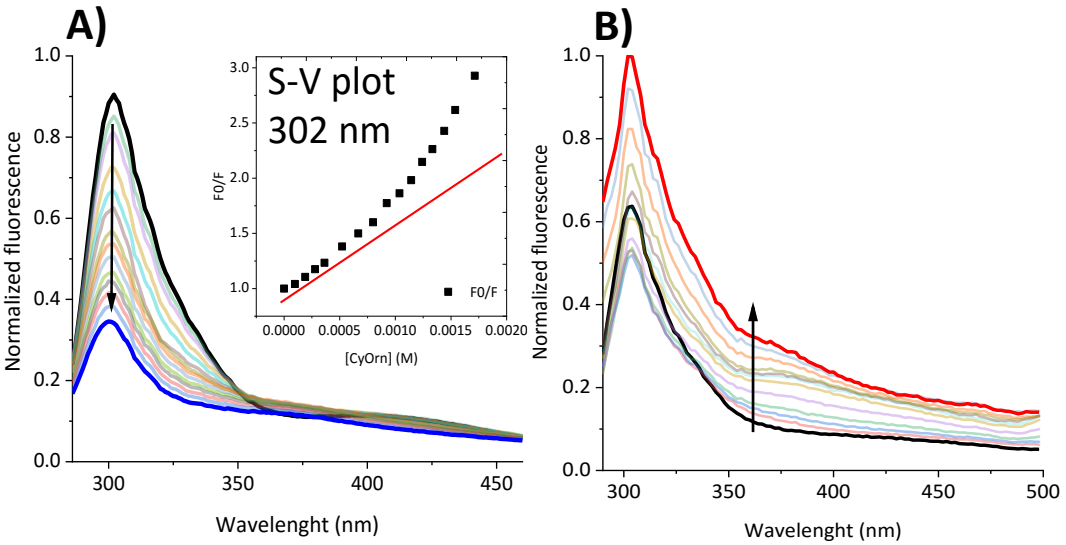


Figure 4.24. Normalized emission spectra of polyK$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyOrn** ([CyOrn] = 0-2 · 10^{-3} M) . λexc = 276 nm. A) [polyK$_4$Y] = 2 · 10^{-4}M and S-V plot at 302 nm. B) [polyK$_4$Y] = 2 · 10^{-5} M. λexc = 276 nm.

On the contrary, **CyAsp** and **CyGlu** form a complex with tyrosine in this polypeptide with a lower K_d (398 and 400 µM respectively). This is one order of magnitude lower than the K_d with the glutamic polypeptide where a stronger binding is observed thanks to the attractive electrostatic interactions between the amino acid residues from the cage and the peptide (Figure 4.25). Not surprisingly, **CyHis** has a very similar affinity for the two copolymers (polyK$_4$Y and polyE$_4$Y) and is the cage with the strongest binding of the series. This versatility was again associated to the This versatility was associated to the versatility of the imidazole group in molecular interactions because of its unique molecular structure.[67] It can act as both hydrogen bond donor (polar hydrogen atom) and acceptor (basic nitrogen). Histidine interactions are fully influenced by pK_a (~6.5) that allows it to be partially protonated depending on its microenvironment. As a consequence histidine can participate in cation-π interactions with negatively charged amino acids (acting as the π-donor)[68] or with aromatic amino acids [69] when protonated thus acting as an organic cation. The non-protonated form of the imidazole ring can also make π-π stacking interactions.[70,71]

Imidazole group can be considered an amphoter in water acting as both an acid or a base (either a Brönsted or by partial protonation), its pK_a and protonation state can switch depending on its microenvironment which will determine the interactions formed. In fact histidine interactions are very relevant in protein structure, interactions and function[72–77]

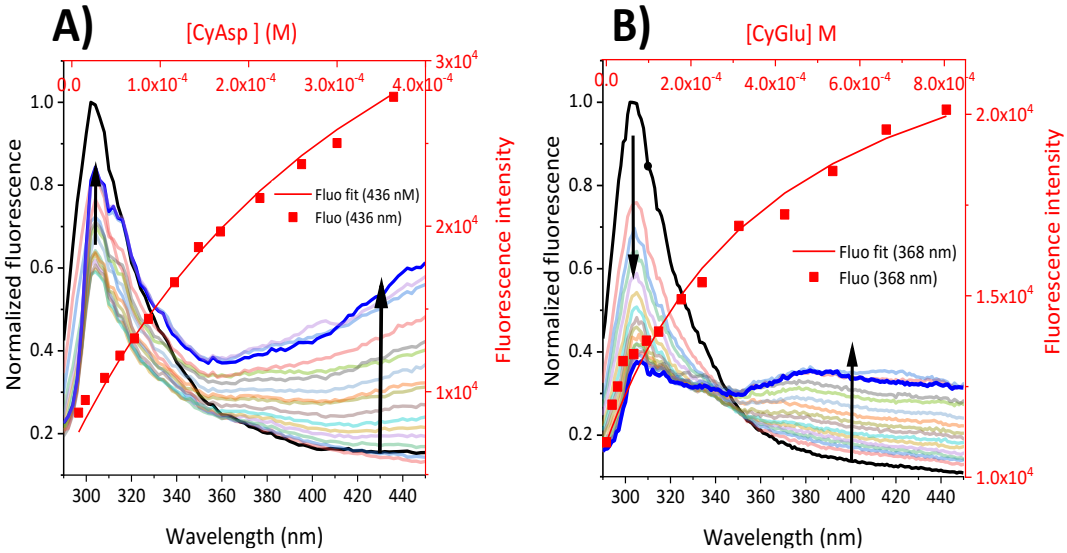


Figure 4.25. Normalized emission spectra of PolyK$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of A) **CyAsp** ([CyAsp] = 0-0.6 · 10^{-3} M) B) **CyGlu** ([CyGlu] = 0-0.8 · 10^{-3} M) (left axis). [PolyK$_4$Y] = 2 · 10^{-5} M. λexc = 276 nm and fitting of the data to a complex 1 : 1 using HypSpec (right axis).

Finally, we studied the binding with the copolymer polyE_3K_6Y (table 3.1, column 6). The results for the fluorescence titrations indicate that **CyLys/CyOrn** have the ability to encapsulate tyrosine in this polypeptide even with higher affinity than with polyE_4Y since the K_{diss} determined was 27 and 114 μM, respectively. Once again, the affinity of the tyrosine for **CyLys** is higher than for **CyOrn** even though the difference is a simple CH_2 in the side chain of the cage. This repeated trend found for **CyLys/CyOrn** indicates that the affinity for tyrosine can be modulated with very small structural changes. CyLys is slightly more basid than CyOrn, therefore a more positive cage on average at equivalent pH values which can be responsible of the affinity differences of these cages for the different copolymers studied.On the other hand, **CyHis** shows a lower affinity for this copolymer, in this case the binding detected was weaker than with **CyLys** and **CyOrn** although be in the same order of magnitude. In the titrations of this copolymer with **CyAsp** and **CyOrn** there was no sign of a new fluorescence band associated to the excimer formation, the only changes in the spectra upon the addition of the polypeptide were the decrease of the tyrosine emission which in both cases gives a linear S-V plot. With this data we could not determine the K_d but if the interaction with the tyrosine occurs its lower limit has to be 465 μM for **CyAsp** and 525 μM for **CyGlu**.

The case of Poly E_3K_6Y has two particularities that can be influencing the binding studies and the fitting. First, we assumed the equivalency between all the Tyr residues in the polypeptide but in this case because of the polypeptide heterogeneity it can be expected to have many non-equivalent Tyr surrounded by glutamic acid, lysine or both. Secondly, the presence of both glutamic and lysine residues allows the presence of positive-negative interactions within the polymer which can result in the polymer collapse both intre- and intramolecularly thus keeping some Tyr inaccessible to the cage.. On average data indicates that cages have a similar behaviour towards this peptide and polyE_4Y because the average charge of it is negative although it contains lysine.

From the fitting with HypSpec software, the fluorescence emission of the Tyr-Cage complex in each fitting can be extracted as a simulated spectra for the 100% complex species. The similarity of the predicted spectra for the Tyr-**CyHis** complex in the different polymers (Figure 4.26, titrated at the same concentrations) indicates that the chemical environment of the Tyr in the copolymers is equally affected by the binding of **CyHis** and this can only be possible if the binding of the cage to the Tyr is by the formation of a very similar inclusion

complex with the three copolymers used. This reinforces the interaction mechanism proposed for the Tyrosine with the newly prepared cages.

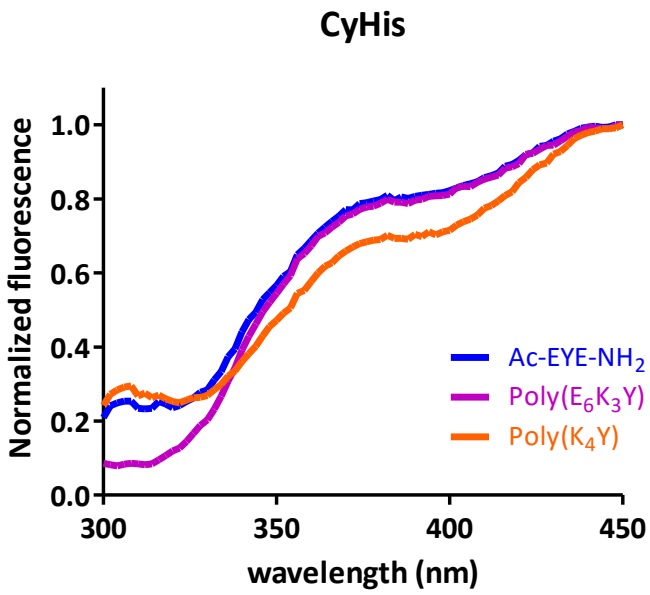


Figure 4.26. Normalized fluorescence emission of the complexes of CyHis with Ac-EYE-NH_2, PolyE_6K_3Y and PolyK_4Y predicted with HypSpec.

4.7.2 ESI-MS Interaction studies

In collaboration with Dr. Cristian Vicent and Prof. Santiago V. Luis from Universitat Jaume I in Castellon we conducted several experiments to get more information about the binding of the pseudopeptidic cages and Tyrosine containing peptides. This is an on-going project so further details will be reported in a scientific publication.

The development of the soft ionization techniques (ESI and MALDI) allows the use of Mass spectrometry (MS) not only for the identification of the molecular weight but also for a more detailed study of compounds properties.[64] The softness sensitivity and specificity of ESI-MS allows its use in supramolecular chemistry to study complexes with low sample and time consumption. Properties like stoichiometry, binding affinities or stability of host-guest complexes in the gas phase can be studied.[83] Collision induced dissociation (CID) is the most used dissociation technique and it allows to follow the ion dissociation as a function of the internal energy supplied thus comparing the stability of different complexes in the gas phase. [84]

Several attempts were conducted aiming to detect pokyE$_4$Y with ESI-MS but it was not possible to identify the mass corresponding to the polypeptide neither in water or using formic acid or ammonia buffers. This issue was associated with the high molecular weight of the polypeptide and its length dispersity. The lack of results using this big polypeptide lead us to decide that further experiments should be conducted with the short peptide Ac-EYE-NH_2.

ESI-MS and CID of CyHis, CyOrn and CyLys with Ac-EYE-NH_2

ESI of **CyHis** 1 · 10^{-5} mM was measured in the presence of increasing concentrations of Ac-EYE-NH_2 (0.5, 1, 4 eq) using adjusted conditions to detect the supramolecular adduct.

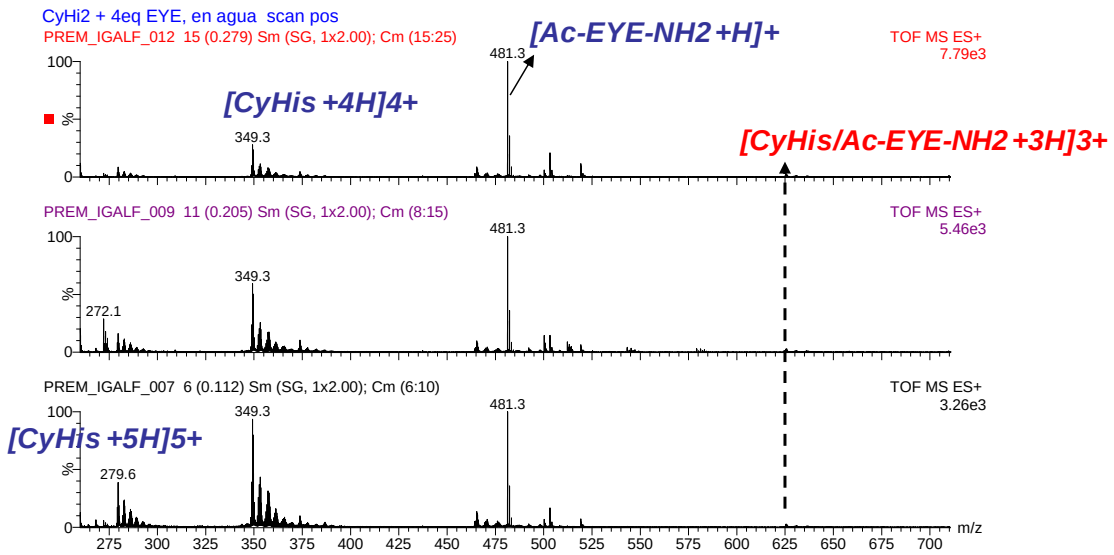


Figure 4.27. ESI(+) spectra of *[CyHis] = 1 · 10-5 M* after adding 0.5 (bottom), 1 (medium) and 4 (top) equivalents of *Ac-EYE-NH_2*.

By recording the MS spectra of **CyHis** in the presence of different concentrations of guest (**¡Error! No se encuentra el origen de la referencia.** we could observe the following:

- **CyHis** with 0.5 eq. of Ac-EYE-NH_2 is present with 3 different protonation states: $[CyHis+5H]^{5+}$ (279.6), $[CyHis+4H]^{4+}$ (349.3) and $[CyHis+3H]^{5+}$ (465.7, very low intensity). It is also seen the 1 : 1 adduct of $[CyHis:Ac\text{-}EYE\text{-}NH_2 + 3H]^{3+}$ (625.8, highlighted in red)
- At higher concentrations of guest the signals corresponding to the host shift to lower charge states (due to the guest basicity). In the spectra with 4 equivalents the signal at 279.6 is very low (almost absent) compared the one at 349.3.

In this case trying to obtain constants form the relative intensities is not possible because we cannot assume that the size of both the host and the guest are similar. Besides there are ions with three positive charges which are intrinsically unstable so there is probably in-source fragmentation.

Spectra of samples 1 : 1 host-guest were also recorded using a basic (pH 9.7) and an acidic (pH 5) buffers of NH_4Ac (figure 4.28)

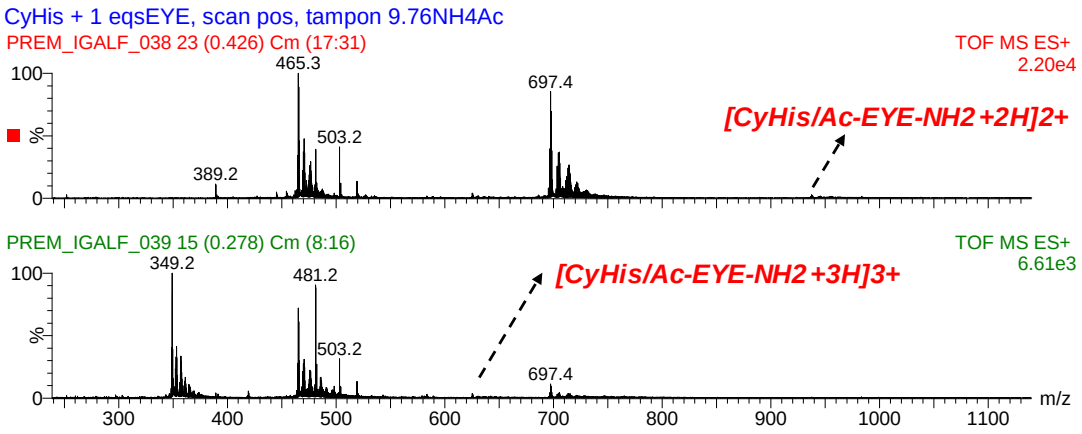


Figure 4.28. ESI(+)-M of Ac-EYE-NH_2 with **CyHis** (10^{-5}) M at pH 5 (top) and 9.7 (bottom) in water.

Spectra at acidic conditions is very similar to the one registered at neutral pH whereas at pH 9.7 there is a clear shift of the peaks to lower charge compounds. In fact at basic pH **CyHis** is detected in the form of $[CyHis + 3H]^{3+}$ (465) and $[CyHis + 2H]^{2+}$ (697) and the 1 : 1 complex is found as $[CyHis+G + 3H]^{3+}$ (625) and $[CyHis+G+2H]^{2+}$ (937).

We examined the fragmentation reaction of the complex $[CyHis + G + 3H]^{3+}$. After the mass selection, the protonated complex was subject to collision-induced dissociation (CID) experiments (figure 4.29).

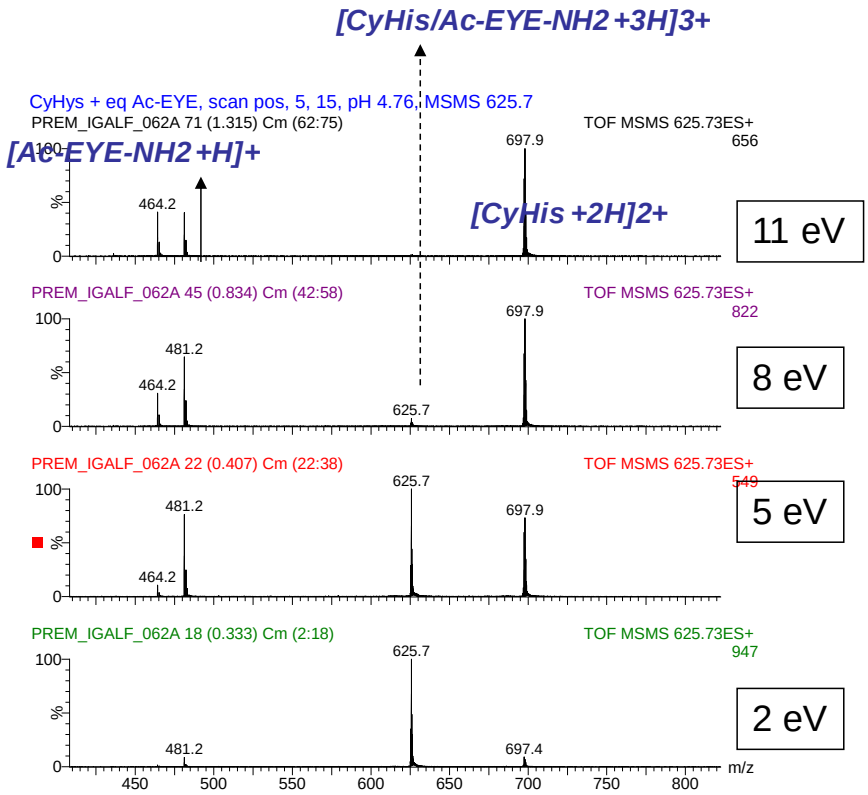


Figure 4.29. CID of $[CyHis + G + 3H]^{3+}$ at growing CE.

The fragmentation pathway observed goes through the adduct separation with charge separation:

$$[CyHis + G + 3H]^{3+} \rightarrow [CyHis + 2H]^{2+} + [G+H]^{+}$$

The CE value determined from this CID experiment is 4.5 eV.

The same measurements were conducted with **CyLys** and **CyOrn** focusing our attention in the 1 : 1 stoichiometry experiments.

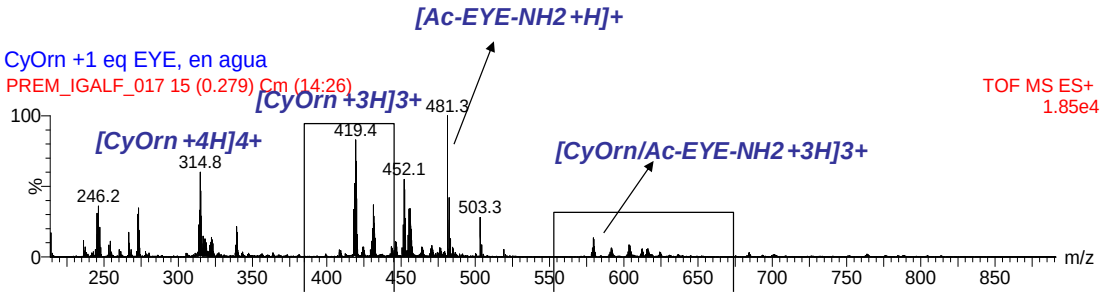


Figure 4.30. ESI(+) spectra of Ac-EYE-NH_2 and **CyOrn** 1 : 1 dissolutions in water. [CyOrn] = 1 x 10^{-5} M.

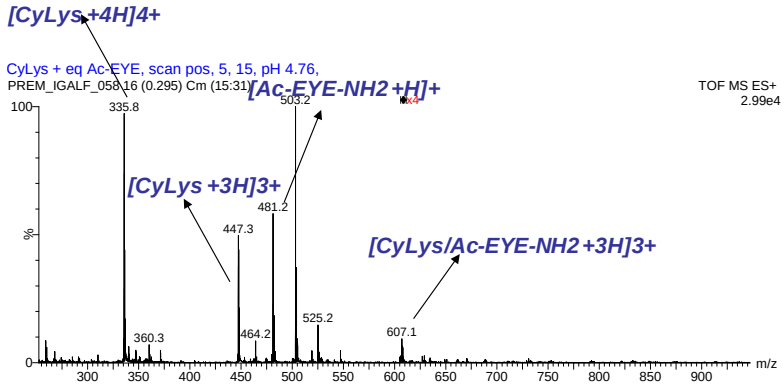


Figure 4.31. ESI(+) spectra of Ac-EYE-NH_2 and **CyLys** 1 : 1 dissolutions in water. [CyLys] = 1 x 10^{-5} M.

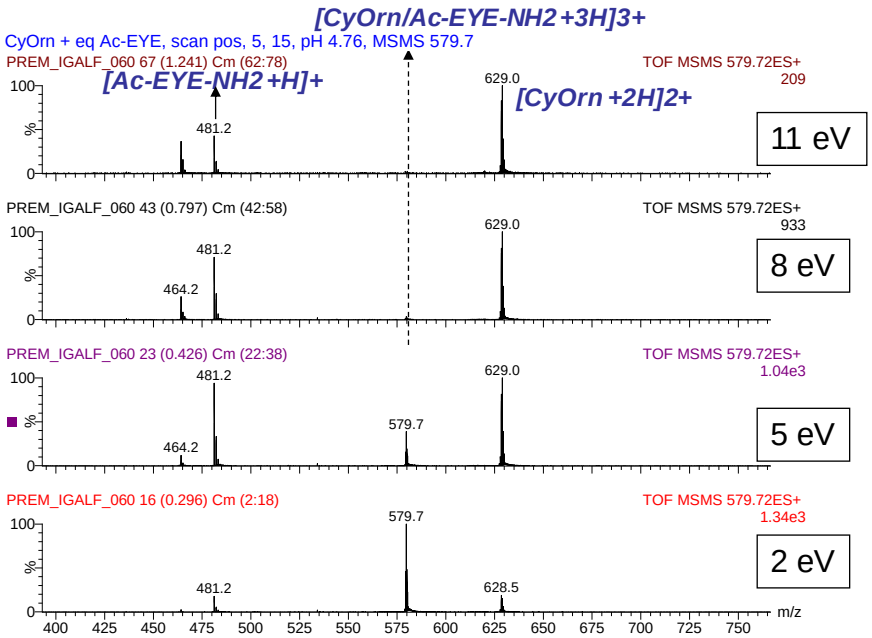


Figure 4.32. CID of $[CyOrn + G + 3H]^{3+}$ at growing CE.

The stoichiometry for both **CyLys** and **CyOrn** is again 1 : 1 And the two adducts have one single fragmentation pathway:

$$[Cage + G + 3H]^{3+} \rightarrow [Cage + 2H]^{2+} + [G + H]^{+}$$

The CE values for **CyOrn** and **CyLys** adducts were 3.1 and 3.2 eV respectively.

Comparing the CID data of **CyHis**, **CyLys** and **CyOrn** we can see that in the gas phase, **CyHis** adduct with Ac-EYE-NH_2 is more stable than **CyLys** followed by **CyOrn** which is in agreement with the results from fluorescence experiments. Finally we analysed the interaction of **CyAsp** with Ac-EYE-NH_2 using the same methodology.

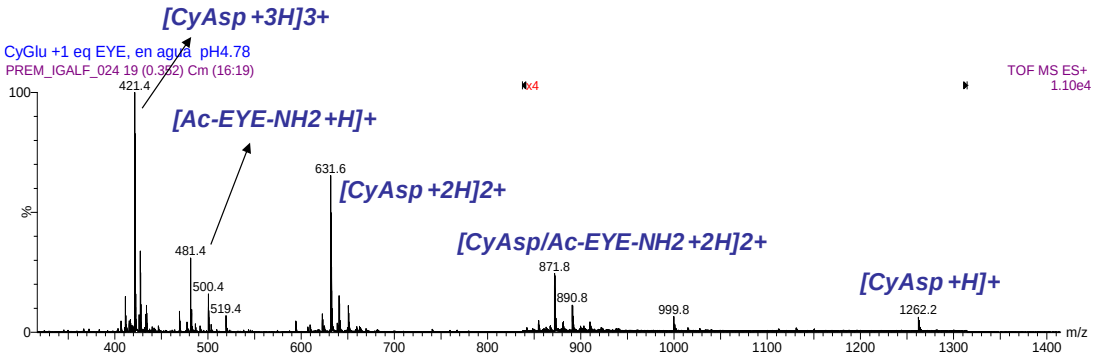


Figure 4.33.-ESI(+) spectra of Ac-EYE-NH_2 and CyAsp 1 : 1 dissolutions in water. [CyAsp] = 1 x 10^{-5} M

In the ESI(+) spectra of Ac-EYE-NH_2 with **CyAsp** (**¡Error! No se encuentra el origen de la referencia.**) **CyAsp** is detected in three protonation states $(CyAsp + H)^+$, $(CyAsp + 2H)^{2+}$, $(CyAsp + 3H)^{3+}$. The adduct of CyAsp with Ac-EYE-NH_2 has a 1 : 1 stoichiometry and the protonation state of the cage forming the adduct is 2+ $[CyAsp/Ac\text{-}EYE\text{-}NH_2 + 2H]^{2+}$.

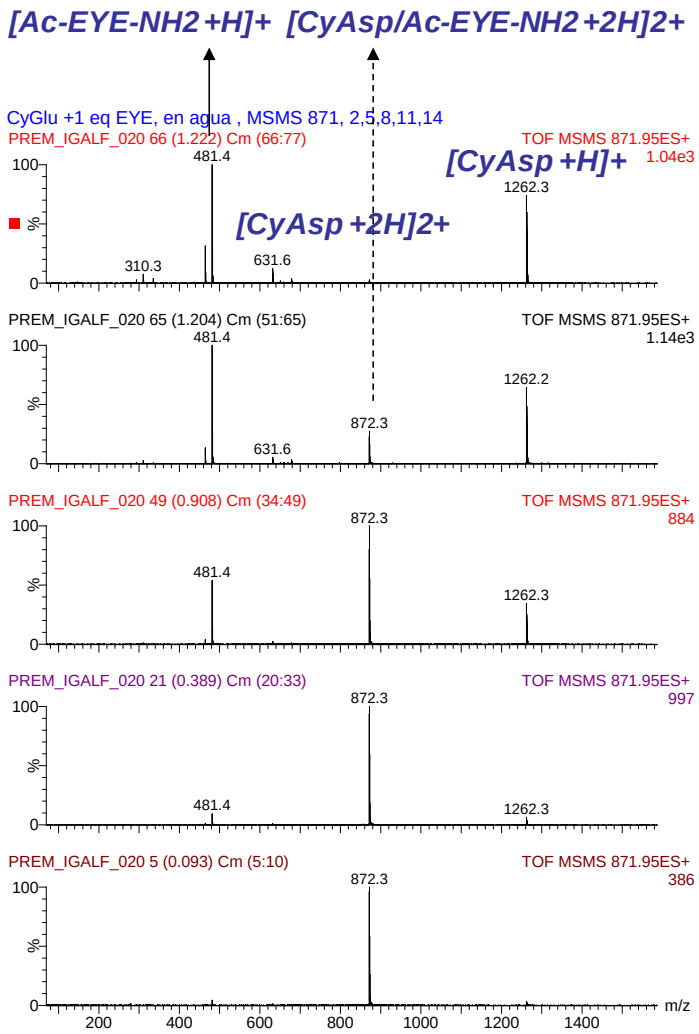


Figure 4.34. CID of [CyAsp +Ac-EYE-NH_2+3H]$^{3+}$ at growing CE.

With CID experiments we detected two main fragmentation pathways of the adduct:

[CyAsp + Ac-EYE-NH_2 + 3H]$^{2+}$ → [CyAsp + 2H]$^{2+}$ + [Ac-EYE-NH_2] *(eq. 1)*

[CyAsp + Ac-EYE-NH_2 +3H]$^{2+}$ → [CyAsp + H]$^{+}$ +[Ac-EYE-NH_2 +H]$^{+}$ (eq. 2)

The predominant fragmentation pathway in this case was the second one which includes both fragmentation and charge separation.

CyAsp was detected in lower protonation states compared to **CyLys, CyOrn CyHis** because it lacks basic lateral groups. Moreover in the adduct with Ac-EYE-NH_2 **CyAsp** is present with a 2+ charge compared to the 3+ charge of **CyLys, CyOrn** and **CyHis** in the adduct.

4.7.3 Molecular modeling

Molecular modeling of cages **CyOrn**, **CyGlu** and **CyHis** with the tripeptide Ac-EYE-NH_2 was conducted. The different models confirm the inclusion of the tyrosine inside the designed binding pocket.

A structural proposal for the complexes of cages **CyOrn**, **CyGlu** and **CyHis** with Ac-EYE-NH_2 is shown in figures 4.35-37 obtained after a Monte Carlo conformational search using Macromodel and the OPLS3e force-field in implicit water. The proposed structures are In agreement with the relative affinity between different cages and the peptides determined through fluorescence and mass spectroscopy. The model proposed for the interaction with **CyHis** shows a perfect inclusion of the tyrosine inside the cavity of the cage inducing small conformation changes to adapt the guest. The case of **CyOrn** also shows a perfect fit of the tyrosine inside the cage inducing slightly more conformational changes in the cage affected by the attractive interactions between the positive charges of the cage and the negative residues of the ligand. Finally the model proposed with the negative cage **CyGlu** indicates the enlargement of the cage cavity thus inducing higher conformational changes which also have to stabilize the repulsion charges between the side chains of the cage and the ligand. The proposed models are in agreement with the stronger interactions between **CyHis** and cages with positive residues and the EYE peptide.

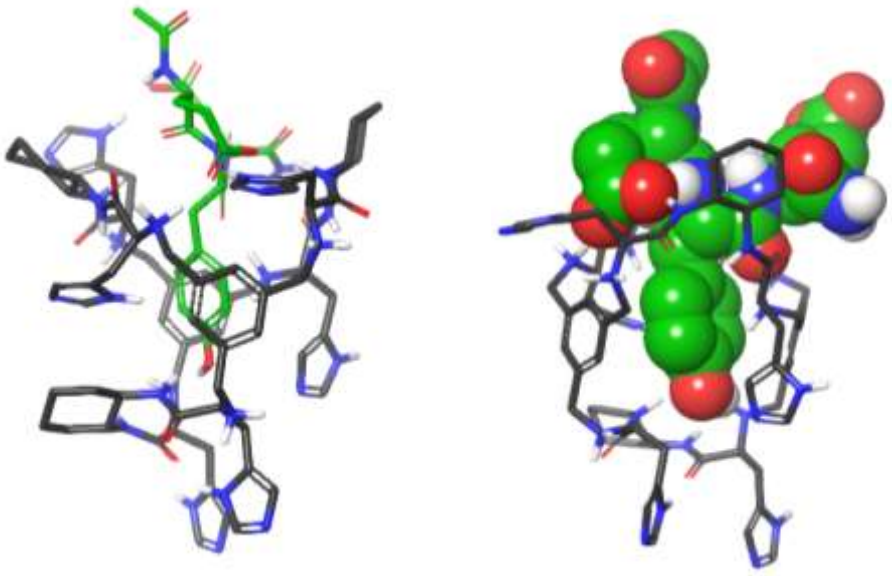


Figure 4.35. Proposed model for the interaction between Ac-EYE-NH_2 and **CyHis**

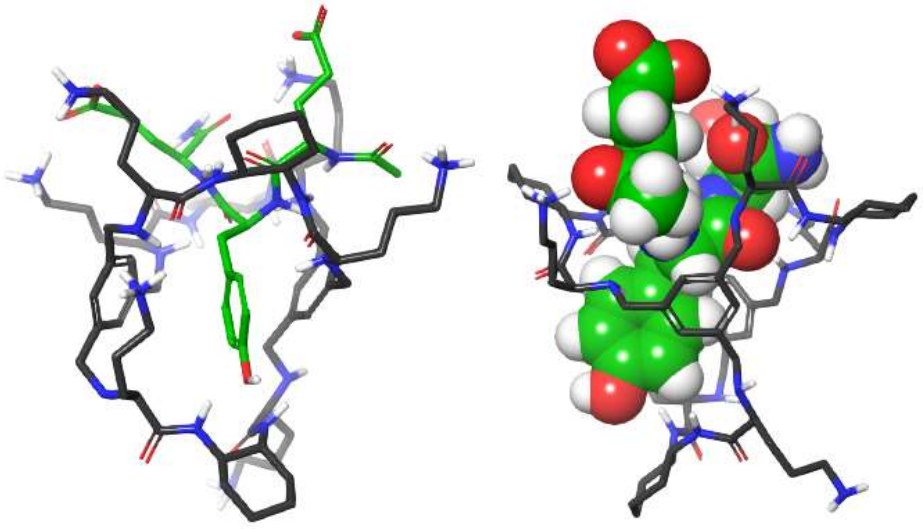


Figure 4.36. Proposed model for the interaction between Ac-EYE-NH_2 and **CyOrn**

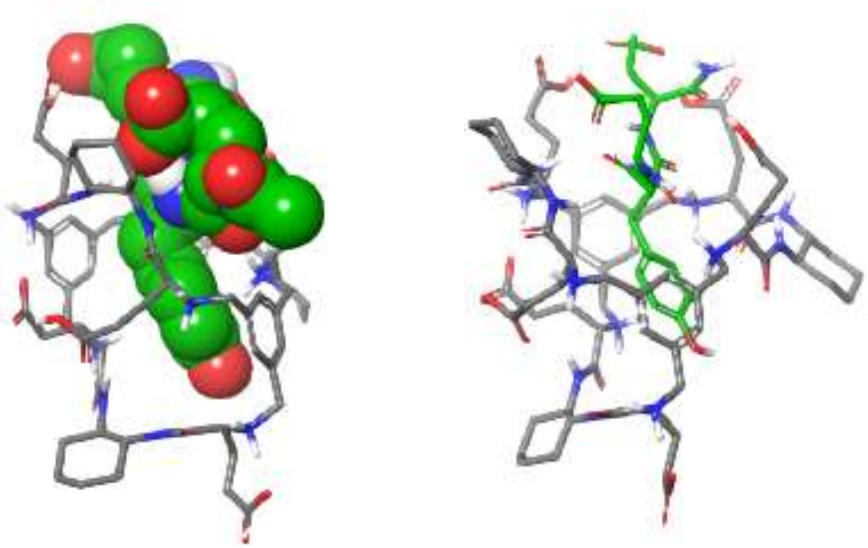


Figure 4.37. Proposed model for the interaction between Ac-EYE-NH_2 and **CyGlu.**

These cage-polypeptide systems have also been studied by NMR through relaxation and DOSY experiments. A very good correlation between the interaction tendencies found via fluorescence titrations and NMR experiments was found. However, since this part of the project was conducted by another researcher from our group we rather not include more details which will be reported in a future publication (currently under preparation).

4.8 CONCLUSIONS

- The pseudopeptidic cages family previously described by the supramolecular chemistry group has been successfully enlarged adding new cages derived from histidine, aspartic and glutamic acids by adapting the previously described synthetic procedure.[1] Therefore cages with neutral, positively and negatively charged residues in the external residue with Tyr binding properties in aqueous buffered solutions have been successfully synthetized.
- Our results from the interaction studies between the cages and Tyr containing polypeptides demonstrate that the binding is modulated by electrostatic interactions between the side chains of the amino acids around the tyrosine in the polypeptide and those decorating the external side of the cage. These interactions boost the Tyr binding inside the cavity when attractive electrostatic interactions are established, while work against it when electrostatic repulsions occur.
- The interaction mechanism between the cages and the tyrosine was proven to be the tyrosine inclusion within cage cavity, independently on the polypeptide substrate.
- These results show the potential of pseudopeptidic cages towards the selective encapsulation of tyrosine residues present in proteins by carefully modulating the secondary interactions between the cage and the protein side chains. The potential applications of these compounds range from the fields of sensing, diagnosis or research to even therapy, among others

4.9 EXPERIMENTAL SECTION

Materials

Reagents and solvents were purchased from commercial suppliers (Adrich, Fluka or Merck) and were used without further purification. Compounds **CySer, CyThr, CyLys** and **CyOrn** were synthetized as previously described.[1,6] All the compounds prepared were fully characterized by the complete spectroscopic (NMR, ESI-MS) and analytical data. Preparative reverse phase purifications were performed on an Isolera Biotage instrument (KP-C18-HS, CH_3CN and water with 0.1% TFA

NMR spectroscopy

The NMR experiments were carried out at 25˚C on a VNMRS-400 NMR spectrometer (Agilent Technologies 400 MHz for ^{1}H and 100 MHz for ^{13}C) Longer relaxation delays (5s) were used for the ^{1}H NMR spectra of the final products in order to get signals with the right integration. Chemical shifts (δ) are quoted in parts per million (ppm) and referenced against the solvent residual peak

ESI-MS

High resolution mass spectra (HRMS) were performed on Acquity UPLC System and a LCT PremierTM XE Benchtop orthogonal acceleration time-of-flight (oa-TOF) (Waters Corporation, Milford, MA) equipped with an electrospray ionization source. All sample solutions (in the 1×10^{-4} to 1×10^{-6} M range) were prepared in methanol.

HPLC analysis

Analytical RP-HPLC was performed with a Hewlett Packard Series 1100 (UV detector 1315 A) modular system using a reverse-phase Kromasil 100 C8 (15 x 0.46 cm, 5 µm) column. CH_3CN-H_2O mixtures containing 0.1% TFA at 1 mL/min were used as mobile phase and monitoring wavelengths were set at 220, 254 and 280 nm.

pH measurements were made using a CRISON pHmeter 50 14 T.

Fluorescence titrations

Fluorescence emission spectra were acquired on a SpectraMax M5 instrument using 10 mm path length cuvettes, excitation bandwidth: 9 nm, emission bandwidth: 15 nm, light source: Xenon flash lamp (1 joule/flash), emission read every 1 nm. All the fluorescence

experiments were performed at 20°C and specific measuring details and fitting procedures are given in the corresponding section.

The different polypeptide-cage titrations were all conducted in a 700 µL fluorescence cuvette following the following protocol: A solution of the peptide was prepared in buffered water (50mM TRIS, pH 7.5). 300 µL of the peptide solution was titrated with a solution of the cage (1-4 mM) in buffered water (50 mM TRIS, pH 7.3,) containing the titrated peptide in the same concentration to maintain the peptide concentration constant throughout the whole titration. The peptide concentration was adjusted for each titration to the concentration that prevents precipitation events and allows to get a larger number of meaningful experimental points for the fitting. The excitation wavelength was λ_{ex}: 276 nm and the emission window recorded was adjusted for each peptide to acquire the whole emission bad for the excimer λ_{em} 290-500/550 nm. HypSpec[65] was the software used for the fitting of the titration data to a proposed interaction model for each cage-peptide pair.

HypSpec fitting procedure

Titration data was introduced to HypSpec and fitted to a 1 : 1 model using either the whole spectra recorded or just the band associated to the formation of the excimer.

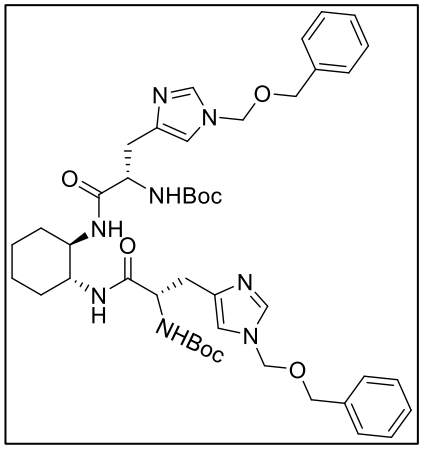
Absorbance measurements were recorded on a SpectraMax M5 spectrophotometer.

Molecular modeling

All molecular simulations were carried out with the package Schrödinger Suite 2019,[78] through its graphical interface Maestro.[79] The program Macromodel,[80] with its default force field OPLS3[81] and GB/SA water solvation conditions,[82] was used for energy minimization.

4.9.1 Synthetic methodology

CyHis_int1



Boc-L-His(bom)-OH (2.5 mg,6.6 mmol) was dissolved in dry DMF (5 mL). (1*R*,2*R*)-*trans*-1,2-cyclohexanediamine (0.59 mg,3.17 mmol) dissolved in DMF (7 mL), HBTU (2.52 mg,6.65 mmol) and *N*,*N*-diisopropylethylamine (DIPEA,7.3 mL,42 mmol) were added. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. Then water was added and the product was extracted with DCM. Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using DCM : MeOH (100-95:5) to give 2.15 g of CyHis_int1 as a white solid (3.0 mmol, 80% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.455 (m, 2H), 7.329 (m, 10H), 6.881 (s, 2H), 6.532 (s, 2H), 5.597 (s, 2H), 5.273 (ABq,J_{ab}= 11 Hz, 4H), 4.494 (s, 4H), 4.316 (X subsystem from ABX, J_{ax} = 8, J_{bx} = 14 Hz, 2H), 3.480 (m, 2H), 3.163 (B subsystem from ABX, J_{aB} = 15.4, J_{bx} = 14 Hz, 2H), 2.946 (A subsystem from ABX, J_{ax} = 8, J_{ab} = 15.4 Hz, 2H), 1.858 (m, 2H), 1.471 (m, 2H), 1.395 (s, 18H), 1.228 (m, 2H), 1.079 (m, 2H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 171.39, 155.85, 137.75, 135.90, 128.76, 128.44, 128.16, 127.92, 80.49, 73.39, 70.11, 53.80, 31.87, 28.33, 26.40, 24.47.

HRMS (ESI-TOF) m/z [CyHis_int1 + H]$^+$ Calc :829.4607, found: 829.4377.

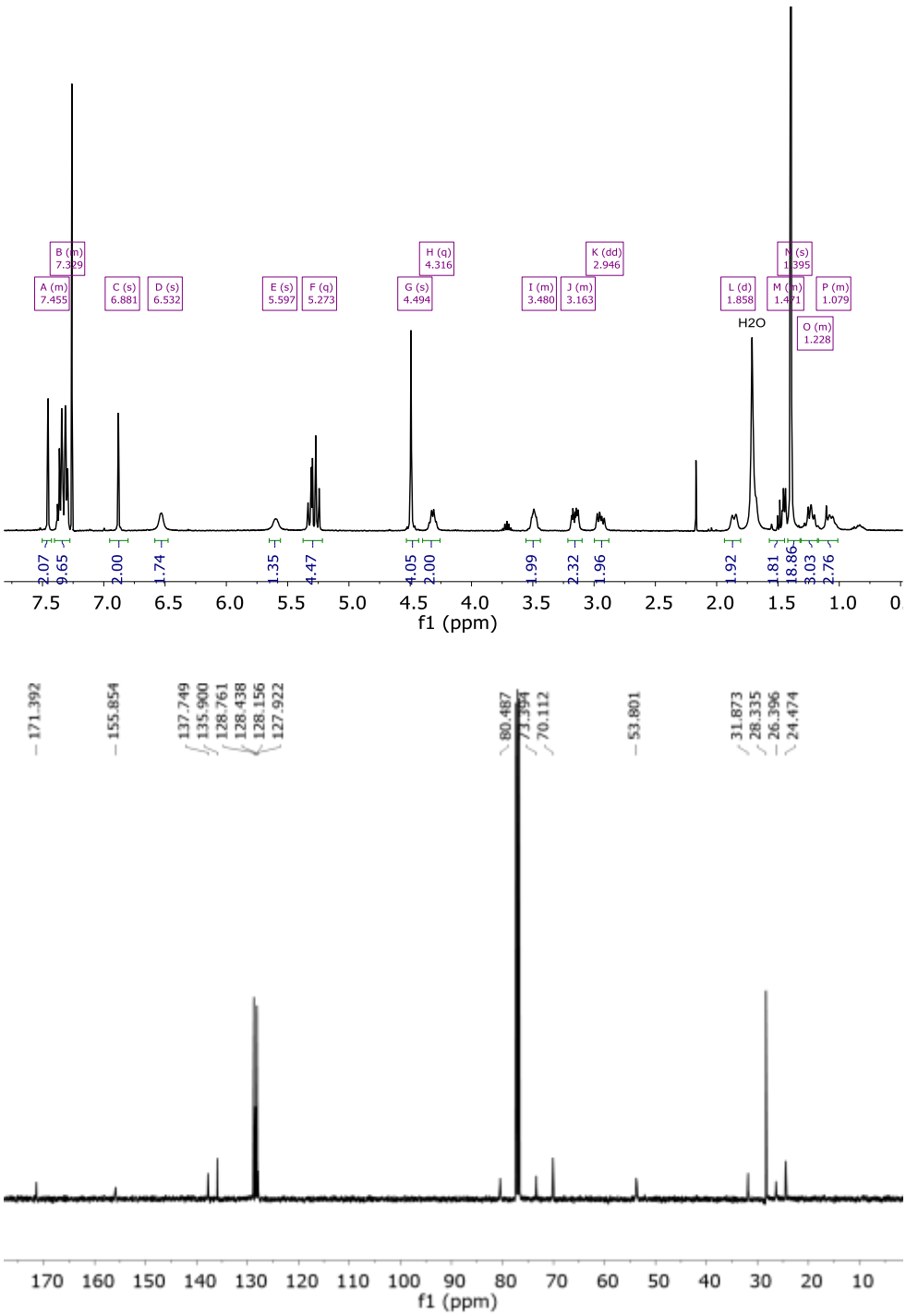


Figure 4.38. ^{1}H-NMR (400 MHz, $CDCl_3$) and ^{13}C NMR (101 MHz, $CDCl_3$) of **CyHis_int1.**

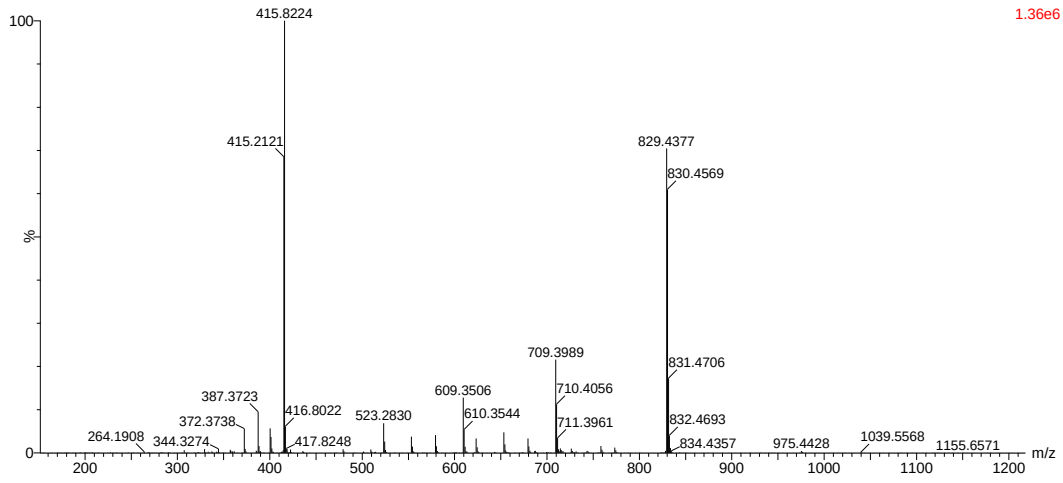
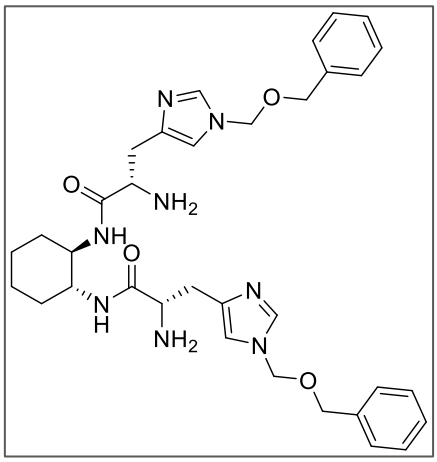


Figure 4.39. HRMS (ESI+) experimental spectrum of **CyHis_int1.**

CyHis_int2



CyHis_int1 (2g, 2.4 mmol) was dissolved in DCM/TFA 1 : 1 (4 mL) with triethylsilane (0.2 mL). The reaction mixture was stirred at room temperature for 3 hours. The solvent was evaporated and the resulting solid was dissolved in NaOH (1 M), extracted with $CHCl_3$/IPA and dried to give **CyHis_int2** as a white solid (1.35 g, 90% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.470 (s, 2H), 7.331 (m, 8H), 7.232 (m, 2H), 6.893 (s, 2H), 5.257 (AB_q, J_{ab}=11.1 Hz, 4H), 4.422 (s, 4H), 3.600 (m, 2H), 3.517 (X subsystem from ABX, J_{ax} = 8.8, J_{bx} = 4.4 Hz, 2H), 3.192 (B subsystem from ABX, J_{ab} = 15.3, J_{bx} = 4.4 Hz, 2H), 2.770 (A subsystem from ABX, J_{ax} = 8.8, J_{ab} = 15.3 Hz, 2H), 1.982 (m, 2H), 1.743 (m, 22H), 1.249 (m, 2H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 173.99, 138.29, 136.01, 128.87, 128.69, 128.31, 127.93, 112.27, 73.02, 69.87, 54.61, 53.18, 29.51, 24.61.

HRMS (ESI-TOF) m/z [**CyHis_int2** + H]$^+$ Calc :629.3558, found: 629.3506.

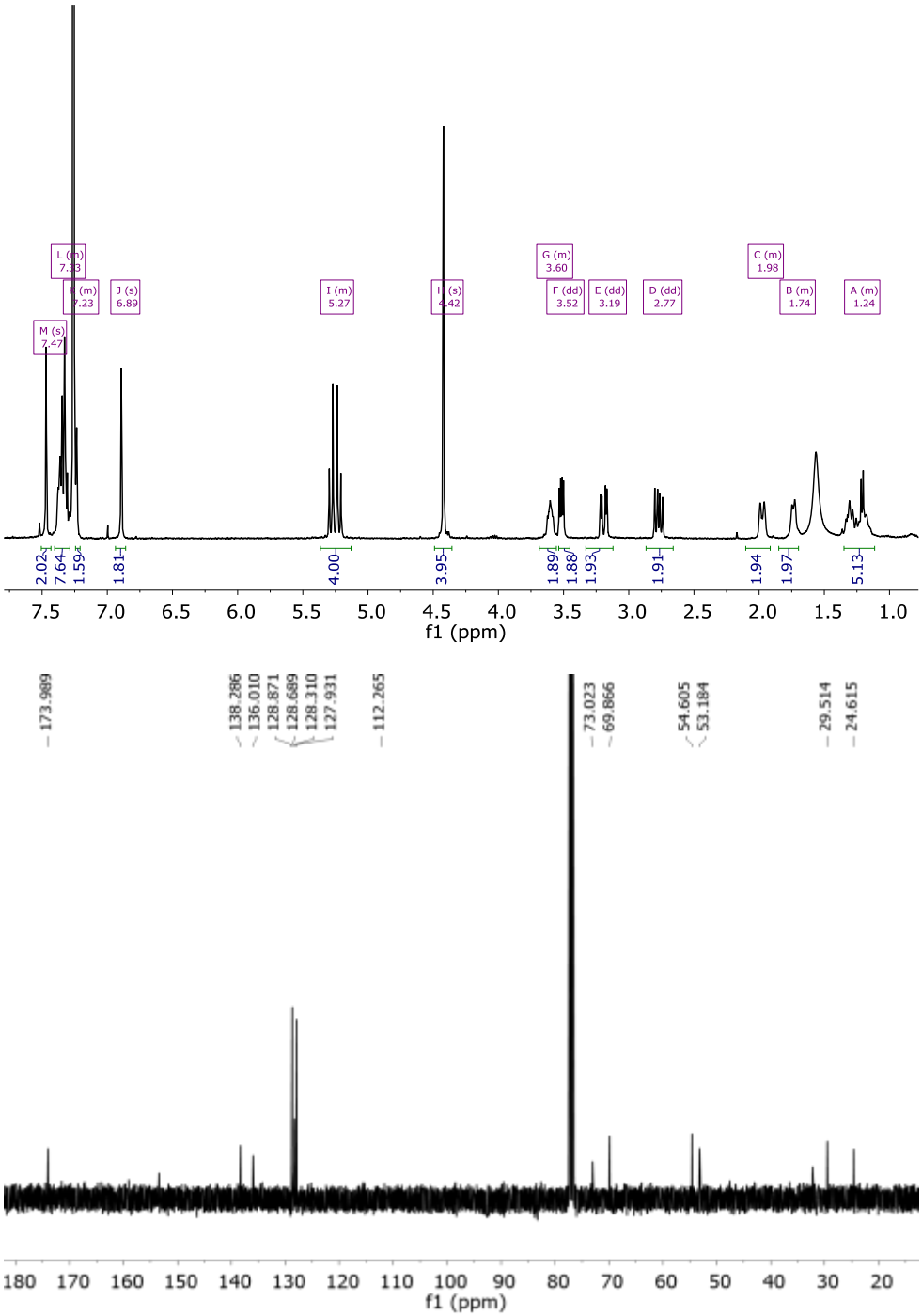


Figure 4.40. ^{1}H-NMR (400 MHz, $CDCl_3$) and ^{13}C NMR (101 MHz, $CDCl_3$) of **CyHis_int2.**

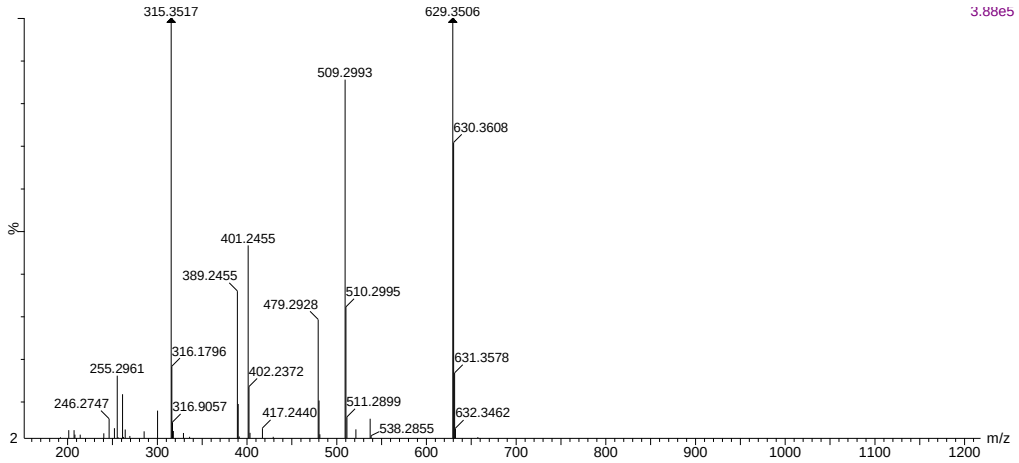
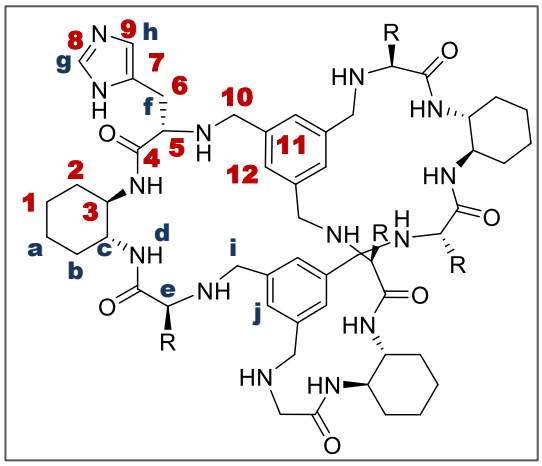


Figure 4.41. HRMS (ESI+) experimental spectrum of **CyHis_int2**

CyHis

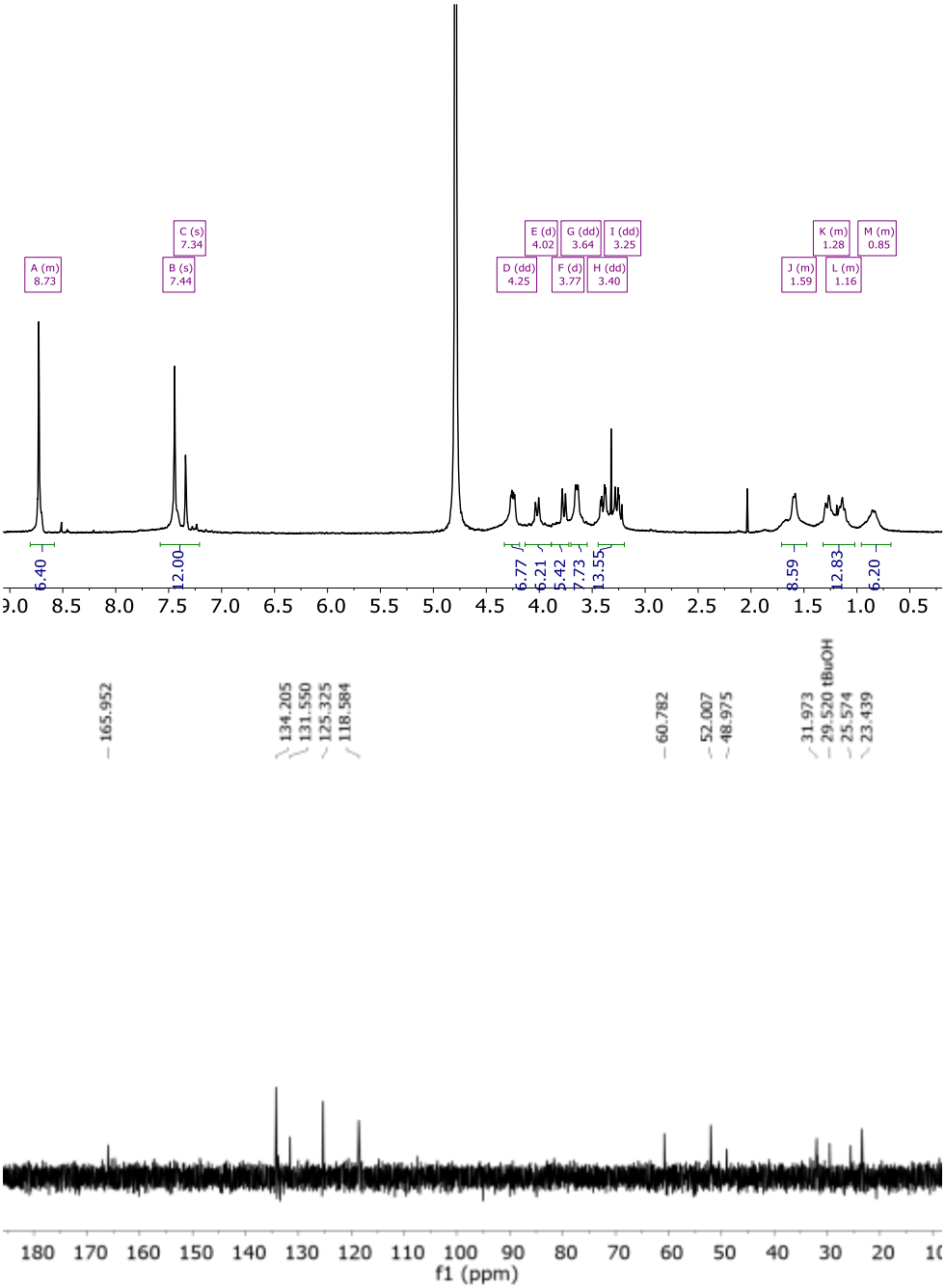


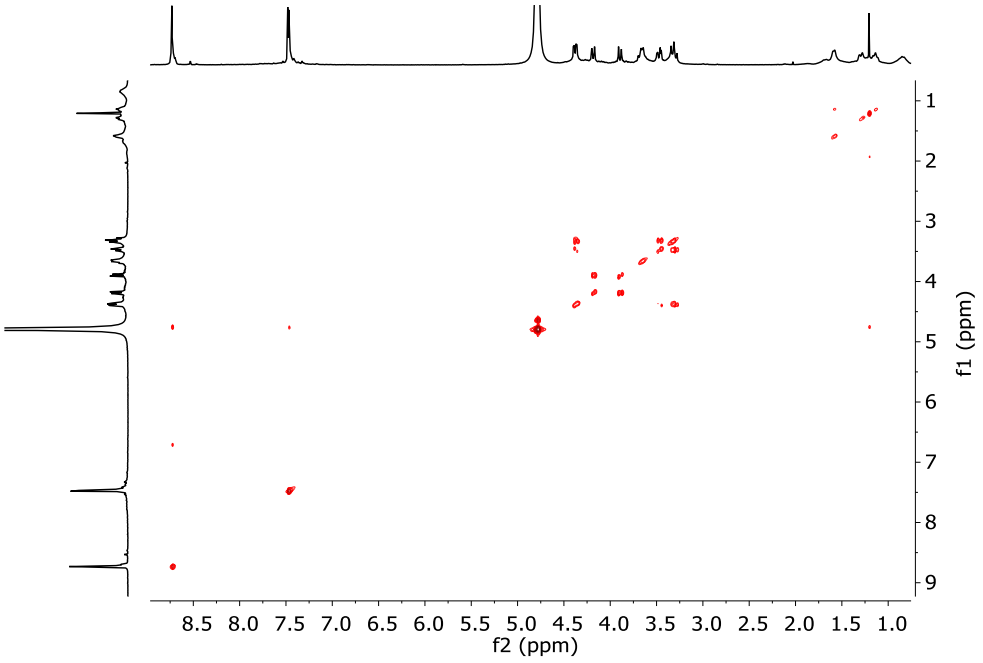
CyHis_int2 (0.4g, 0.64 mmol) was dissolved in anhydrous MeOH (25 mL). Benzene-1,3,5-tricarbaldehyde (0.069g, 0.42 mmol) was added. The solution was stirred at room temperature for 24 hours. Then $NaBH_4$ (0.160 g, 4.24 mmol) was added and the reaction mixture was stirred for 24 hours more. After reaction was completed, NH_4Cl (aq) was added to neutral pH and the solid was filtered off dried and dissolved in TFA/TFMSA/P-cresol (2.7/0.3/0.3 mL). The reaction crude was stirred for 3.5 hours at room temperature. Then TFA was evaporated under a N_2 current and the precipitate washed several times with ether. The resulting solid was purified using reverse phase chromatography to give **CyHis** in its TFA salt as a white solid (180 mg, 35% yield).

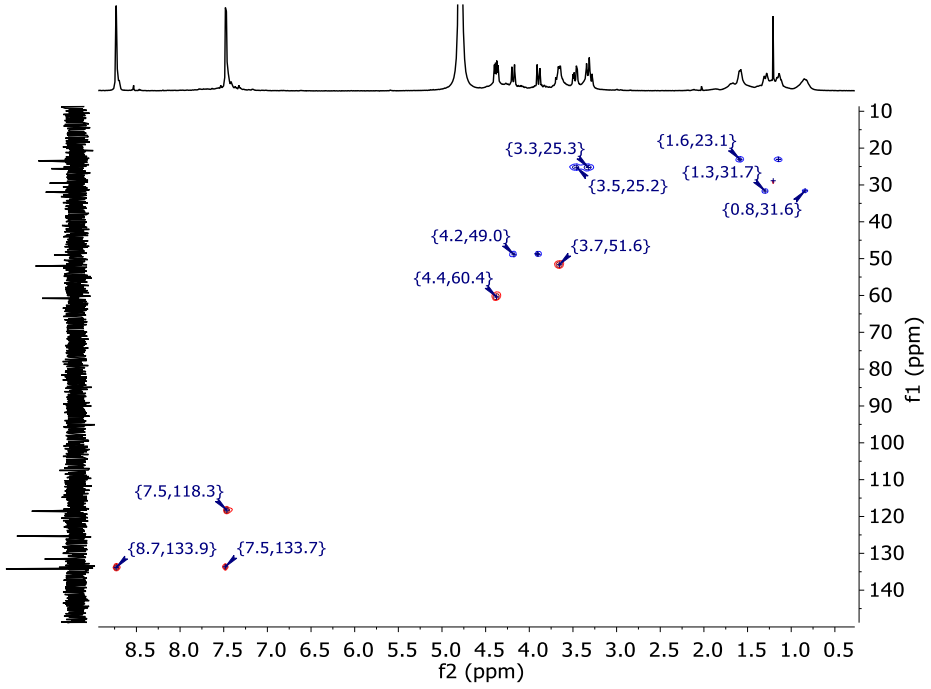
^{1}H NMR (400 MHz, D_2O) δ 8.726 (m, 6H, (**g**)), 7.443 (s, 6H, (**h**)), 7.340 (s, 6H, (**j**)), 4.251 (X subsystem from ABX, J_{bx} = 11, J_{ax}= 4.9 Hz, 6H, (**e**)), 4.037-3.753 (q_{ab}, J = 11.8 Hz, 12H, (**i**)), 3.645 (m,6H, (**c**)), 3.396 (A subsystem from ABX, J_{ab} = 14.6, J_{ax} = 4.9 Hz, 6H, (**f**)), 3.252 (B subsystem from ABX, J_{ab} = 14.6, J_{bx} = 11.0 Hz, 6H, (**f**)), 1.588 (m, 6H, (**a**)), 1.283 (m, 6H, (**b**)), 1.142 (m, 6H, (**a**)), 0.847 (m, 6H, (**b**)).

^{13}C NMR (101 MHz, D_2O) δ 165.95(**4**), 134.21(**8+12**), 131.55(**11**), 125.32(**7**),118.58(**9**), 60.78(**5**), 52.01(**3**), 48.97(**10**), 31.97(**2**), 25.57(**6**), 23.44(**1**).

HRMS (ESI-TOF) m/z [**CyHis** + H]$^+$ Calc 1393.7810, found: 1393.7982.




f1 (ppm)
f2 (ppm)


{3.3,25.3}
{3.5,25.2}
{1.6,23.1}
{1.3,31.7}
{0.8,31.6}
{4.2,49.0}
{3.7,51.6}
{4.4,60.4}
{7.5,118.3}
{8.7,133.9}
{7.5,133.7}
f1 (ppm)
f2 (ppm)

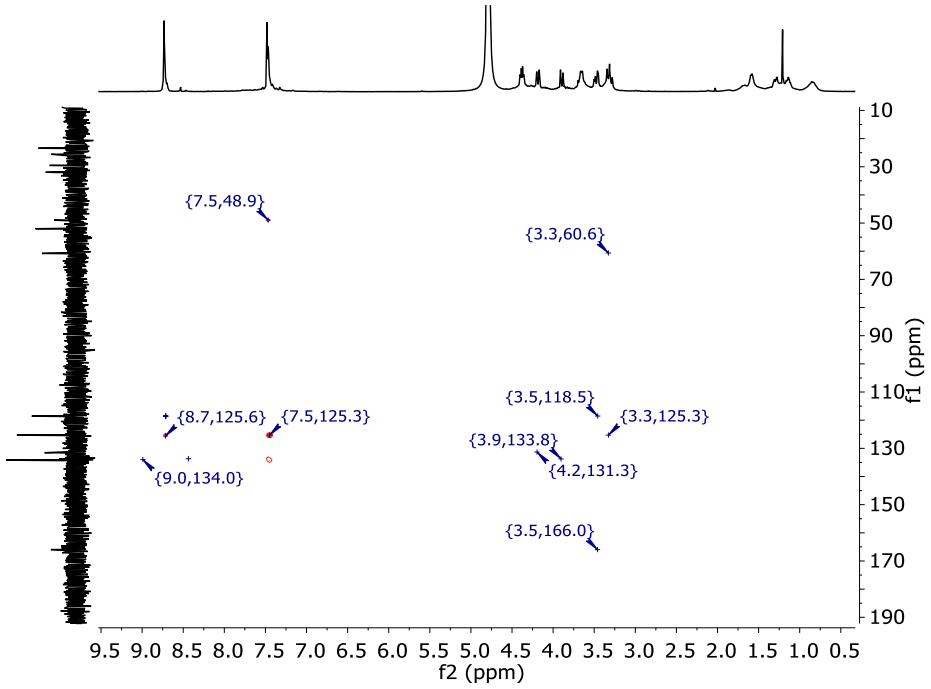


Figure 4.42. ^{1}H-NMR (400 MHz, D_2O), COSY, ^{13}C NMR (101 MHz, , D_2O), HSQC and HMBC spectra of **CyHis.**

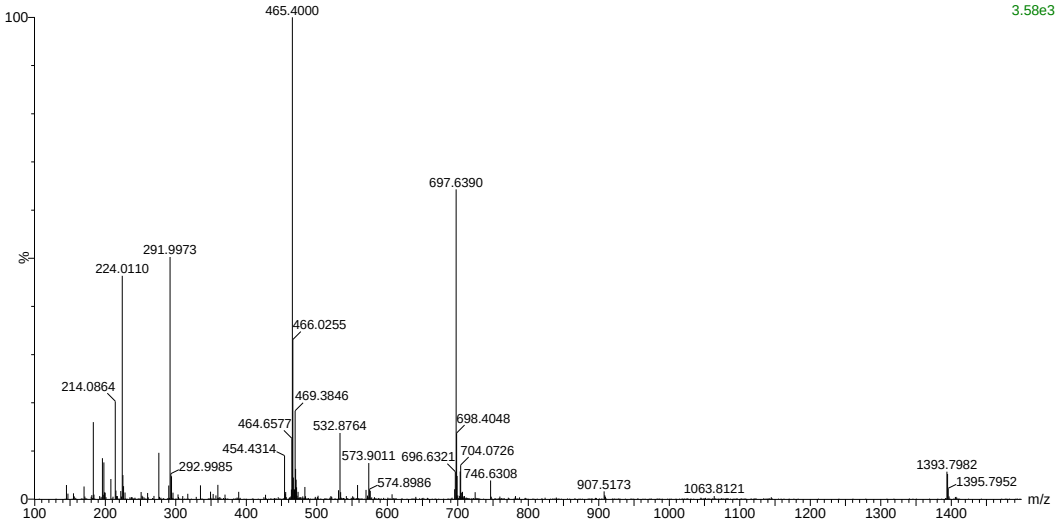


Figure 4.43. HRMS (ESI+) experimental spectrum of **CyHis.**

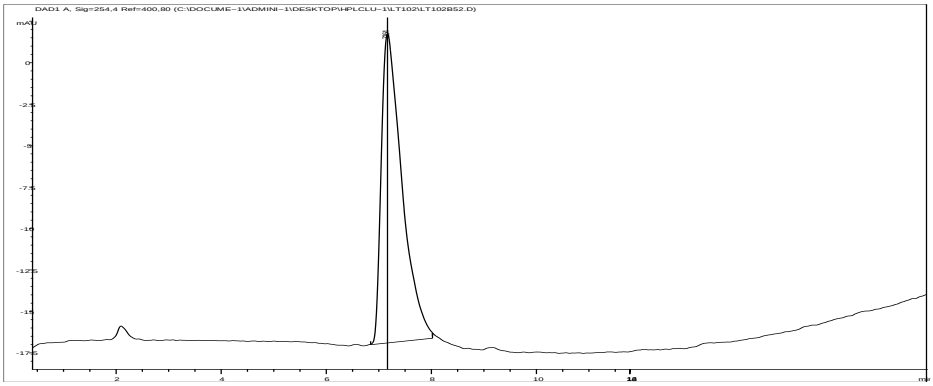
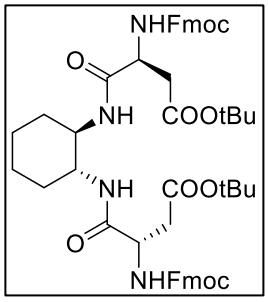


Figure 4.44. HPLC chromatogram of compound **CyHis.**

CyAsp_int1



Fmoc-L-Asp(*t*Bu)-OH (2 g, 4.86 mmol), (1*R*,2*R*)-*trans*-1,2-cyclohexanediamine·2HCl (0.45 g, 2.43 mmol), HBTU (1.84 g, 4.86 mmol) and DIPEA (1.65 mL, 9.49 mmol) were dissolved in dry DMF (9 mL). The reaction mixture was stirred at room temperature under nitrogen atmosphere for 3 hours (fully conversion of the starting materials was observed by TLC). Then 25 mL of water were added to the reaction mixture. The white precipitate formed was filtered and washed several times with water, dissolved in DCM and dried over magnesium sulfate. Finally solvent was evaporated to dryness. Product **CyAsp_int1** was obtained as a white solid (2.1 g, 2.33 mmol, 95% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.22 (m, 4H), 7.586 (d, *J* = 7.5 Hz, 4H), 7.364 (q, *J* = 7.2 Hz, 4H), 6.710 (m, 2H), 6.079 (d, *J* = 9.0 Hz, 2H), 4.604 (m, 2H), 4.362 (m, 4H), 4.198 (m, 2H), 3.687 (s, 2H), 2.896 (dd, *J* = 16.8, 5.8 Hz, 2H), 2.780 (dd, *J* = 17.4, 5.7 Hz, 2H), 2.006 (d, *J* = 11.1 Hz, 2H), 1.765 (m, 2H), 1.407 (s, 18H), 1.298 (m, 4H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.99, 170.79, 156.38, 143.67, 141.21, 127.66, 127.03, 125.13, 119.91, 81.53, 67.42, 53.86, 51.27, 46.97, 37.32, 32.13, 28.05, 24.57.

HRMS (ESI-TOF) m/z [**CyAsp_int1** + H]$^{+}$ Calc: 901.4309, found:901.4305.

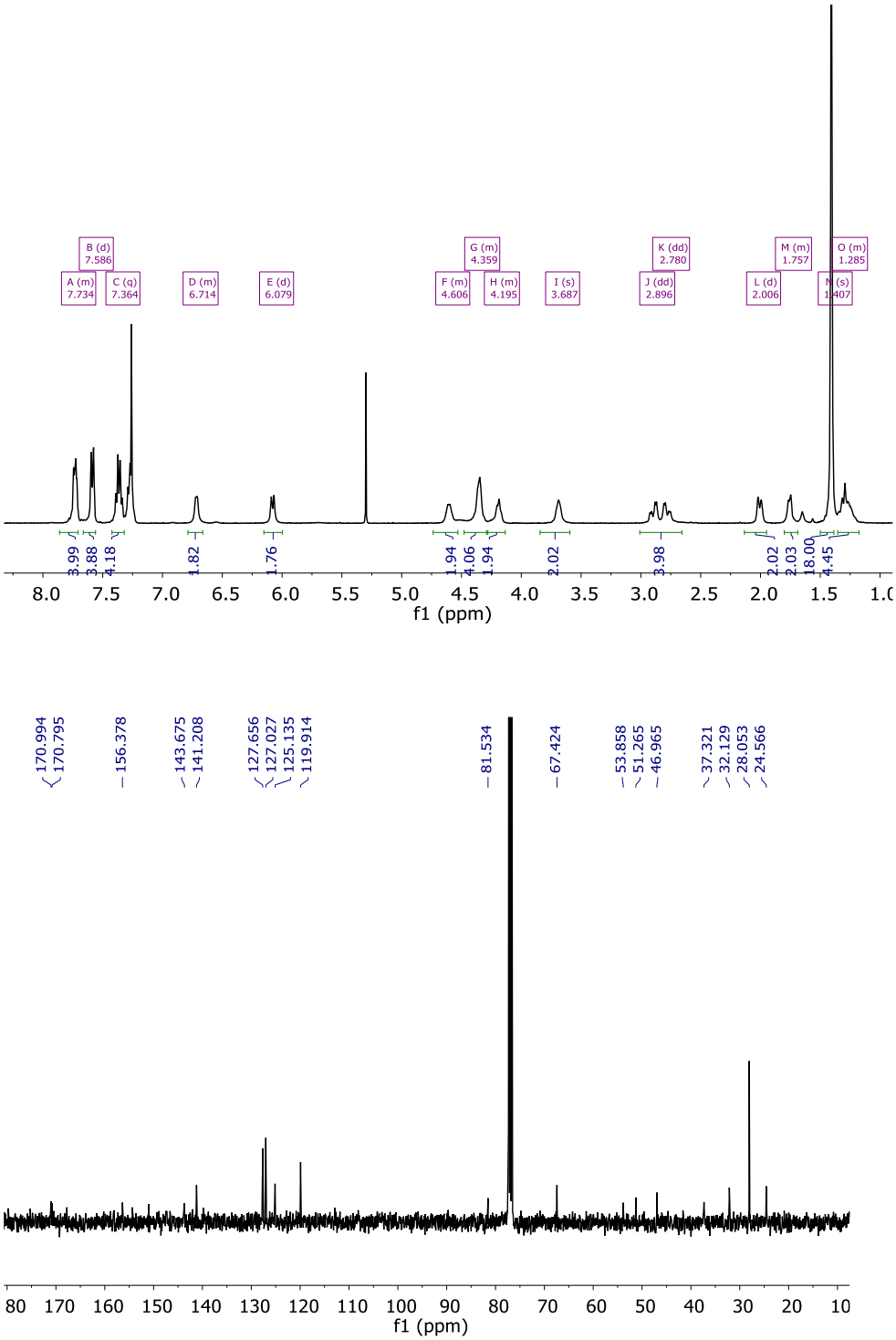


Figure 4.45. ^{1}H-NMR (400 MHz, $CDCl_3$), COSY, ^{13}C NMR (101 MHz, $CDCl_3$) of **CyAsp_int1.**

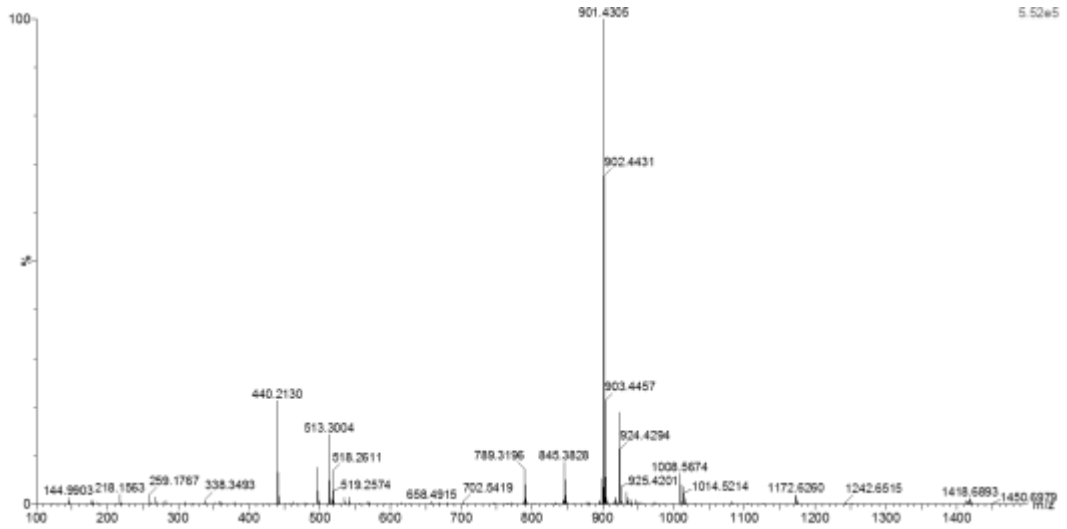
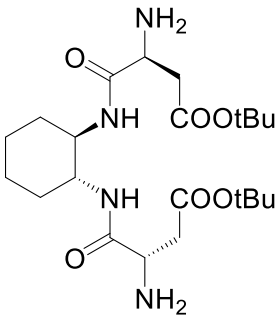


Figure 4.46. HRMS (ESI+) experimental spectrum of **CyAsp_int1.**

CyAsp_Int2



CyAsp_int1 (2.1 g, 2.33 mmol) was dissolved in 15 mL of a solution of piperidine 20% in DMF. The solution was stirred at room temperature for 6 hours. Then water (15 mL) was added to the mixture. The white precipitate formed was filtered off. Product was extracted from the aqueous phase with DCM (3x15 mL). Organic layer was dried over Mg_2SO_4 and evaporated to dryness. **CyAsp_int2** was obtained as a white solid (0.8 g, 1.75 mmol, 82% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.434 (d, *J* = 8.0 Hz, 2H(NH amide), 3.592 (m, 2H), 3.513 (X subsystem form ABX, J_{ax} = 8.1, J_{bx} = 3.9 Hz, 2H), 2.756 (B subsystem from ABX, J_{ab} = 8.1, J_{bx} = 3.9Hz, 2H), 2.505 (A subsystem from ABX, J_{ab} = 8.1, J_{ax} = 8.1 Hz, 2H), 1.937 (m, 2H), 1.678 (m, 2H), 1.239 (m, 4H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 173.87, 171.53, 81.03, 53.11, 52.12, 40.65, 32.61, 28.30, 24.88.

HRMS (ESI-TOF) m/z [**CyAsp_int2** + H]$^+$ Calc: 457.2942, found: 457.2963

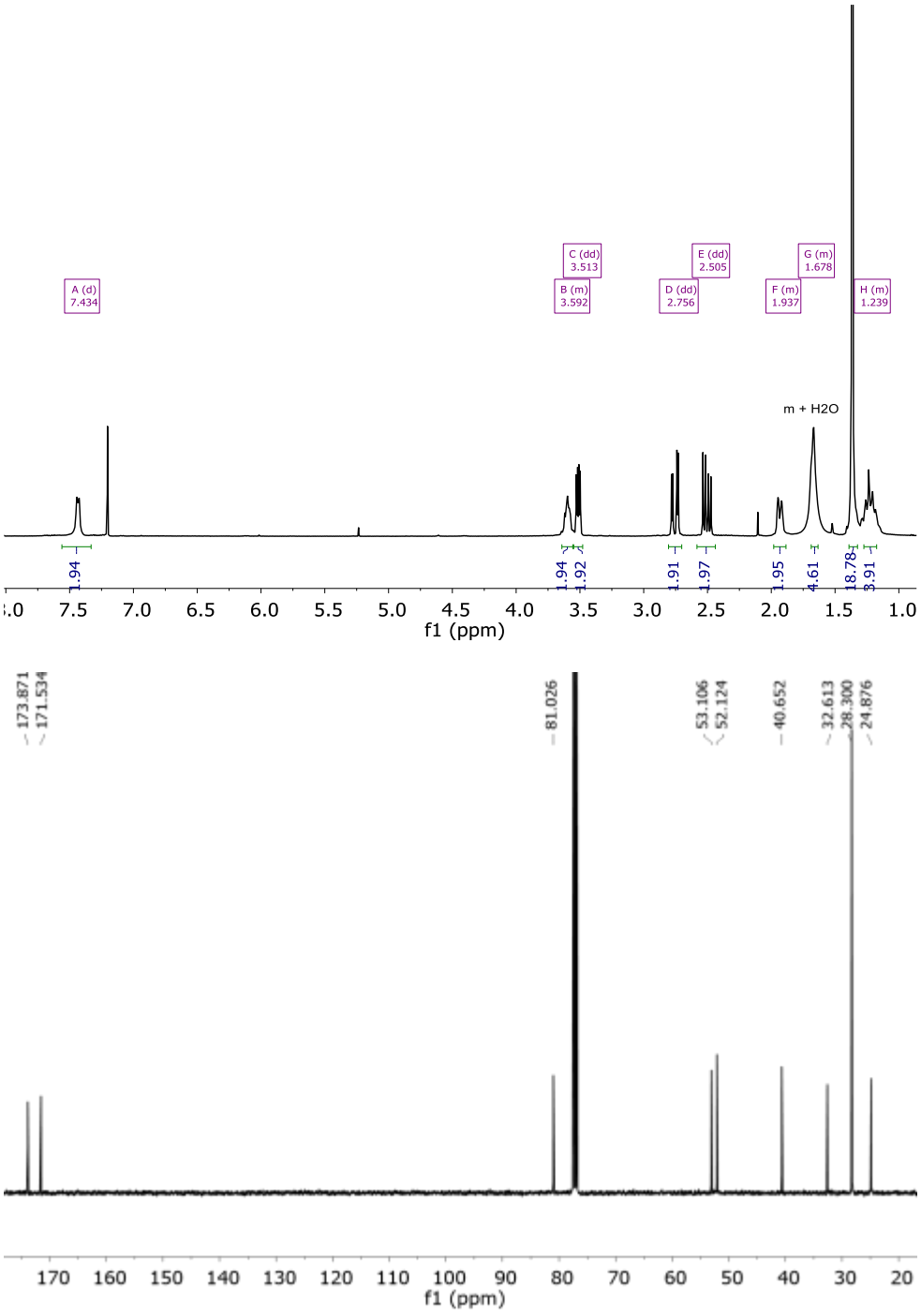


Figure 4.47. ^{1}H-NMR (400 MHz, $CDCl_3$), and^{13}C NMR (101 MHz, $CDCl_3$) of **CyAsp_int2.**

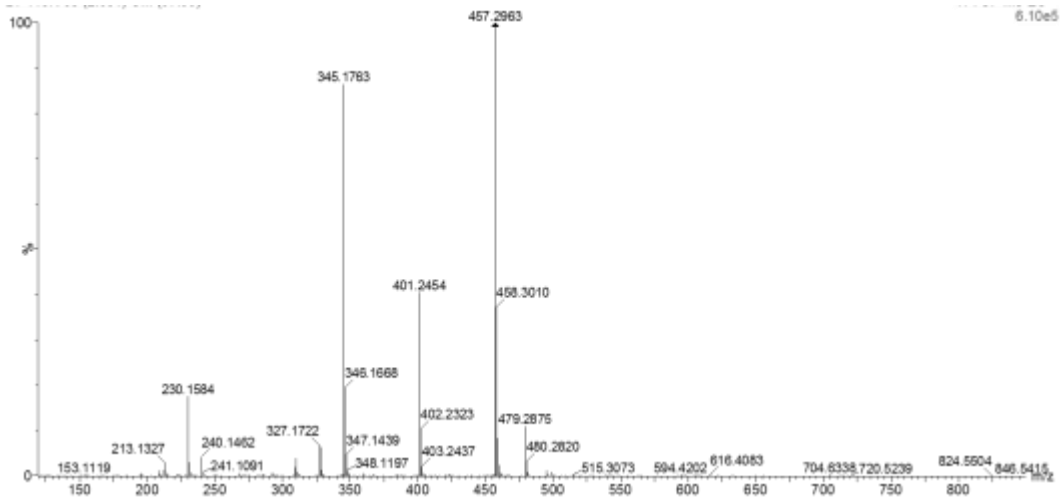
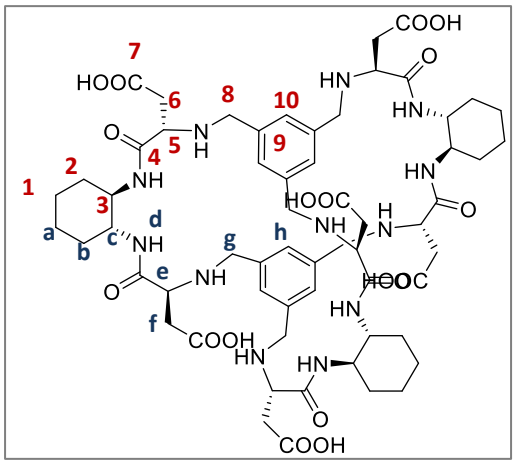


Figure 4.48. HRMS (ESI+) experimental spectrum of **CyAsp_int2.**

CyAsp

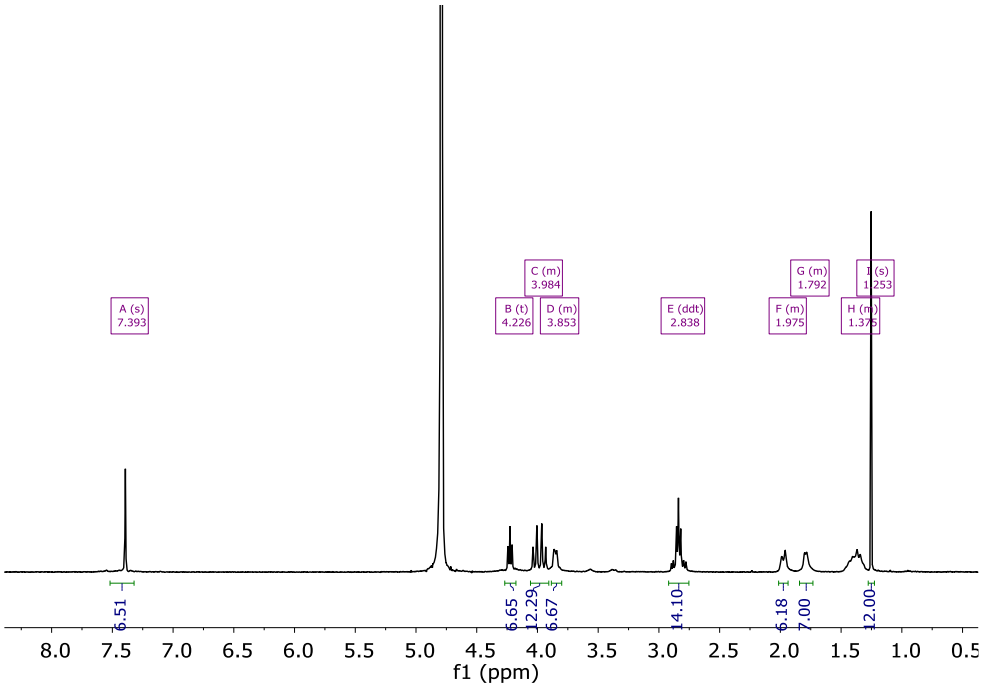


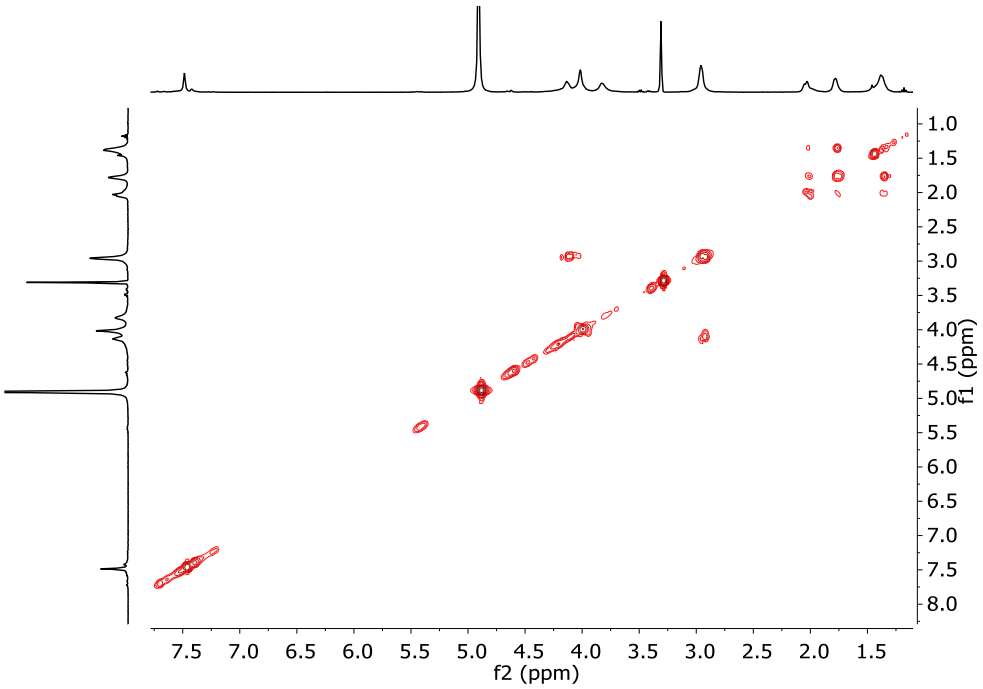
CyAsp_int2 (0.3g, 0.657 mmol) was dissolved in anhydrous MeOH (25 mL). Benzene-1,3,5-tricarbaldehyde (0.071g, 0.42 mmol) was added. The solution was stirred at room temperature for 24 hours. Then $NaBH_4$ (0.116 g, 4.38 mmol) was added and the reaction mixture was stirred for 24 hours more. After reaction was completed, NH_4Cl (aq) was added to neutral pH and the product was extracted with DCM. The resulting crude was purified by column chromatography (DCM:MeOH 95:5). The white solid obtained was dissolved in DCM/TFA 1 : 1 (2 mL) with 0.2 mL if triethylsilane (TES). The solution was stirred for 3 hours. After, TFA was evaporated and the solid was washed several times with diethyl ether. Final product was purified by reverse phase chromatography. **CyAsp** was obtained as a TFA salt (white solid, 35% yield)

^{1}H NMR (400 MHz, CD_3OD) δ 7.399 (s, 6H, (**h**)), 4.226 (t, 6H, (**e**)), 3.984 (m, 12H, (**g**)), 3.853 (m, 6H, (**c**)), 2.838 (m, 6H, (**f**)), 1.975 (m, *3*H, (**b**)), 1.783 (m, 3H, (**a**)), 1.379 (m, 6H, (**b+a**)).

^{13}C NMR (101 MHz, CD_3OD) δ 173.28(**7**), 169.11(**4**), 134.84 (**9**), 132.06(**10**), 58.16 (**5**), 54.17(**3**), 50.58(**8**), 35.98(**6**), 33.00(**2**), 25.64(**1**).

HRMS (ESI-TOF) m/z [**CyAsp**+ H]$^+$ Calc: 1261.6099, found: 1261.6128.





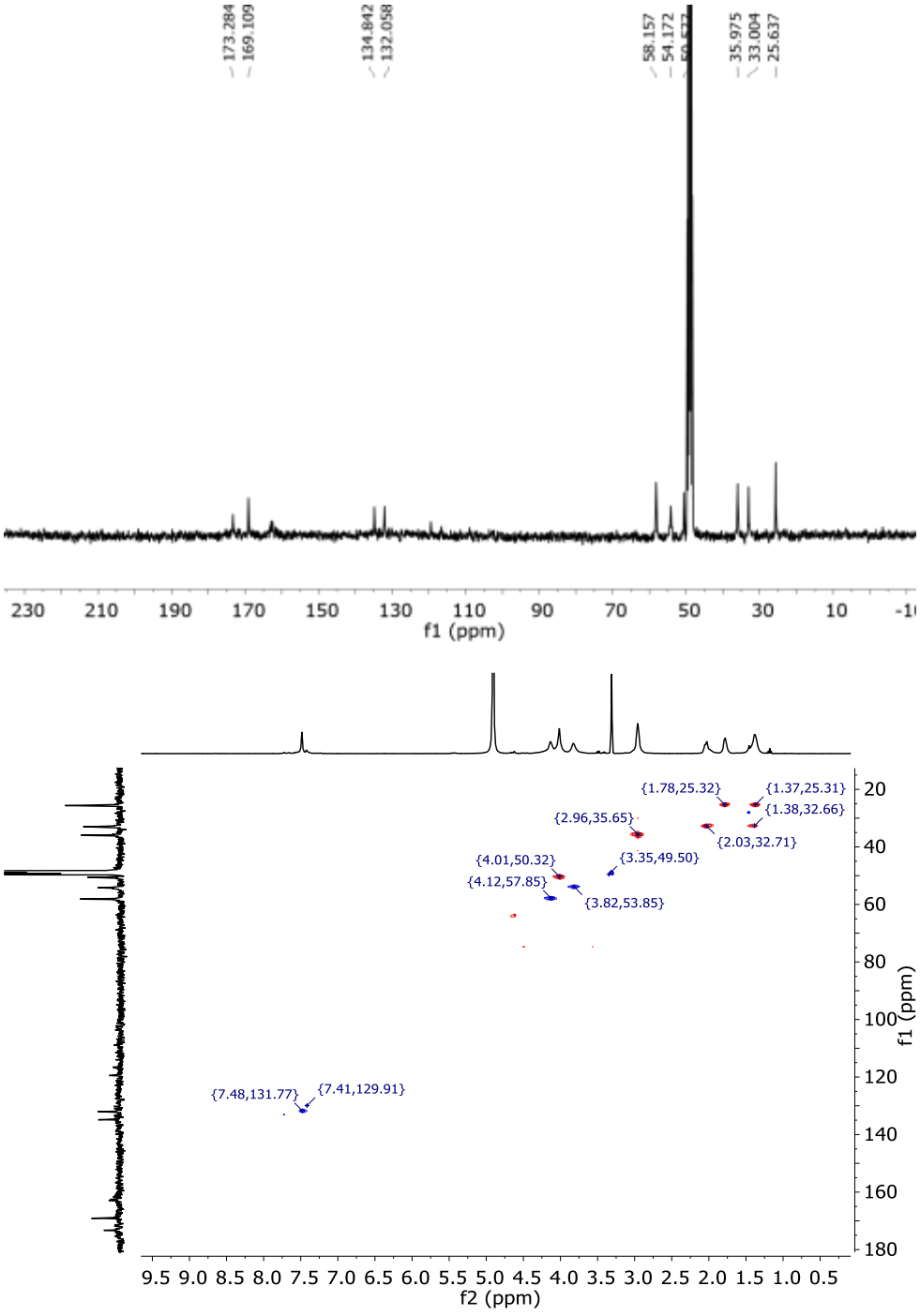


Figure 4.49. ^{1}H-NMR (400 MHz, D_2O), COSY, ^{13}C NMR (101 MHz, CD_3OD) and HSQC of **CyAsp**

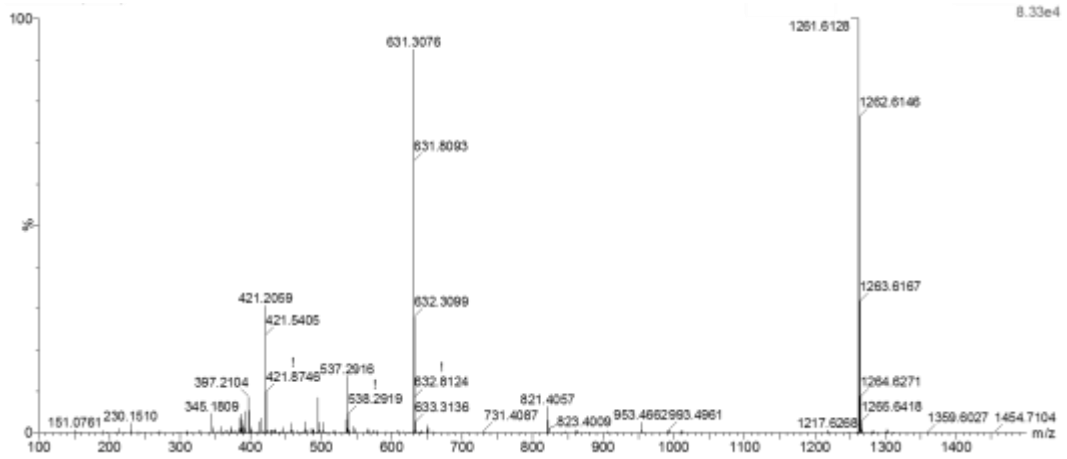


Figure 4.50. HRMS (ESI+) experimental spectrum of **CyAsp.**

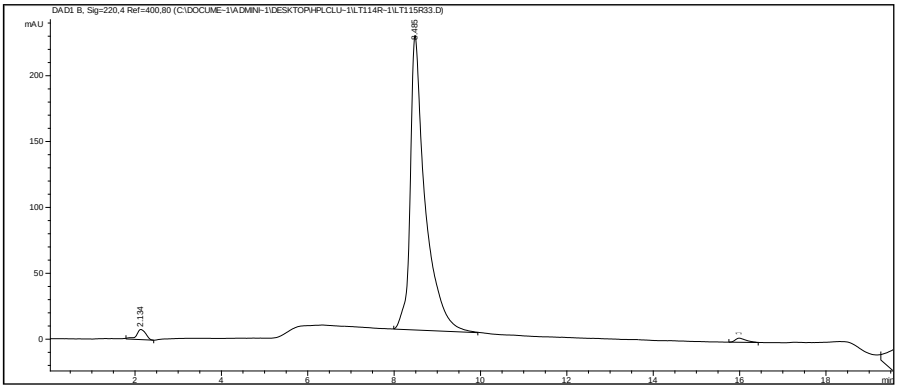
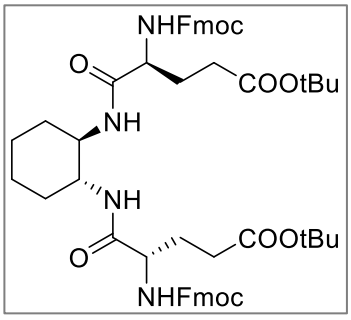


Figure 4.51. HPLC chromatogram of compound **CyAsp.**

CyGlu_ int1



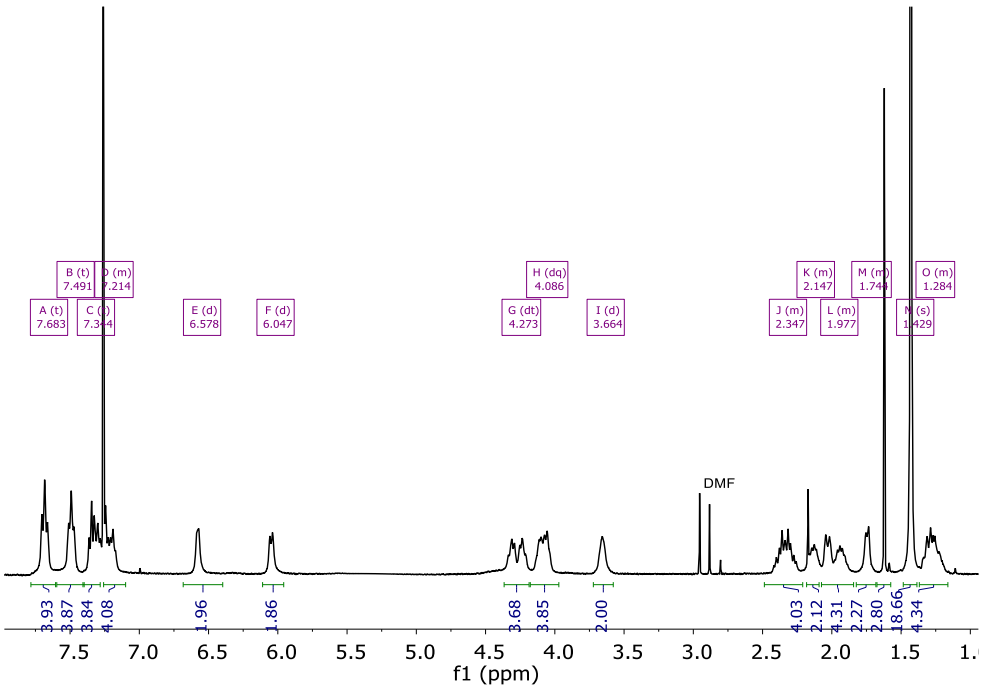
Fmoc-L-Glu(*t*Bu)-OH (2 g, 4.5 mmol), (1*R*,2*R*)-*trans*-1,2-cyclohexanediamine·2HCl (0.40 g, 2.14 mmol), HBTU (1.71 g, 4.50 mmol) and DIPEA (3.2 mL, 18 mmol) were dissolved in dry DMF (8 mL). The reaction mixture was stirred at room temperature under nitrogen atmosphere for 3 hours (fully conversion of the starting materials was observed by TLC). Then 25 mL of water were added to the reaction mixture. The white precipitate formed was filtered and washed several times with water, dissolved in DCM and dried over magnesium sulphate. Solvent was evaporated to dryness and

purified through a silica gel column (0-5% MeOH in DCM). Product **CyGlu_int1** was obtained as a white solid (1.97 g, 2.33 mmol, 92% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.683 (t, *J* = 8.4 Hz, 4H), 7.491 (t, *J* = 8.3 Hz, 4H), 7.344 (t, *J* = 7.6 Hz, 4H), 7.214 (m, 4H), 6.578 (d, *J* = 6.0 Hz, 2H), 6.047 (d, *J* = 8.2 Hz, 2H), 4.273 (m, 4H), 4.086 (m, *4*H), 3.664 (m, *2*H), 2.347 (m, 4H), 2.147 (m, 2H), 1.977 (m, 2H), 1.744 (m, 2H), 1.429 (s, 18H), 1.284 (m, 4H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 193.99, 174.55, 172.29, 144.44, 134.72, 134.14, 129.10, 124.35, 120.33, 80.10, 68.78, 63.04, 62.59, 46.34, 38.69, 31.23, 28.12, 27.88, 24.46.

HRMS (ESI-TOF) m/z [**CyGlu_int1**+ H]$^{+}$ Calc: 929.4695, found: 929.5586.



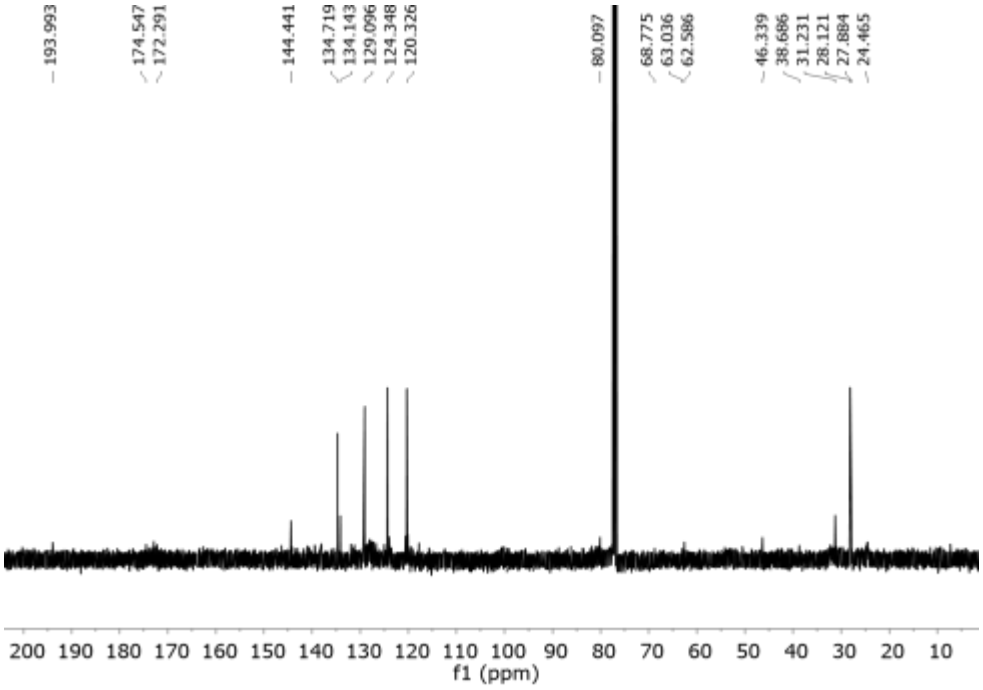


Figure 4.52. ^{1}H-NMR (400 MHz, $CDCl_3$) and ^{13}C NMR (101 MHz, $CDCl_3$) of **CyGlu_int1.**

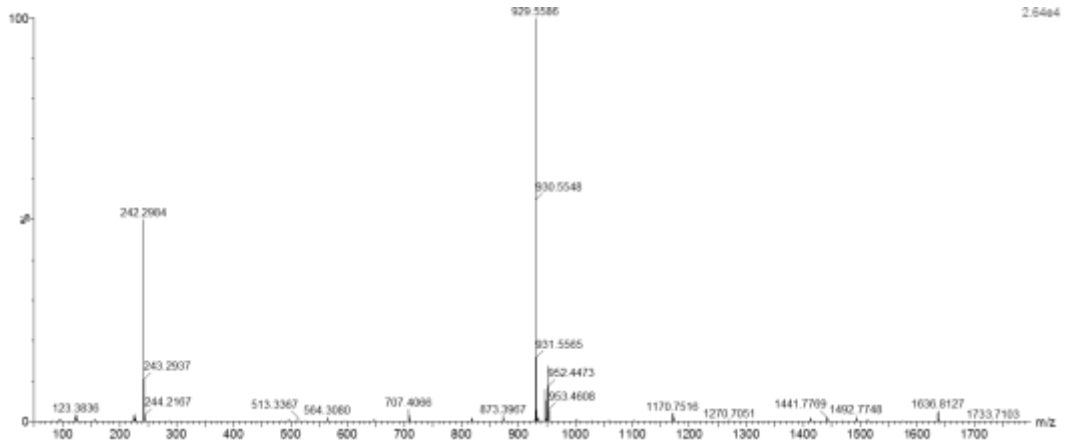
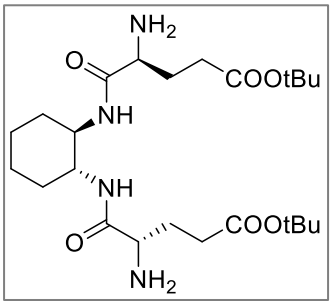


Figure 4.53. HRMS (ESI+) experimental spectrum of **CyGlu_int1.**

CyGlu_int2

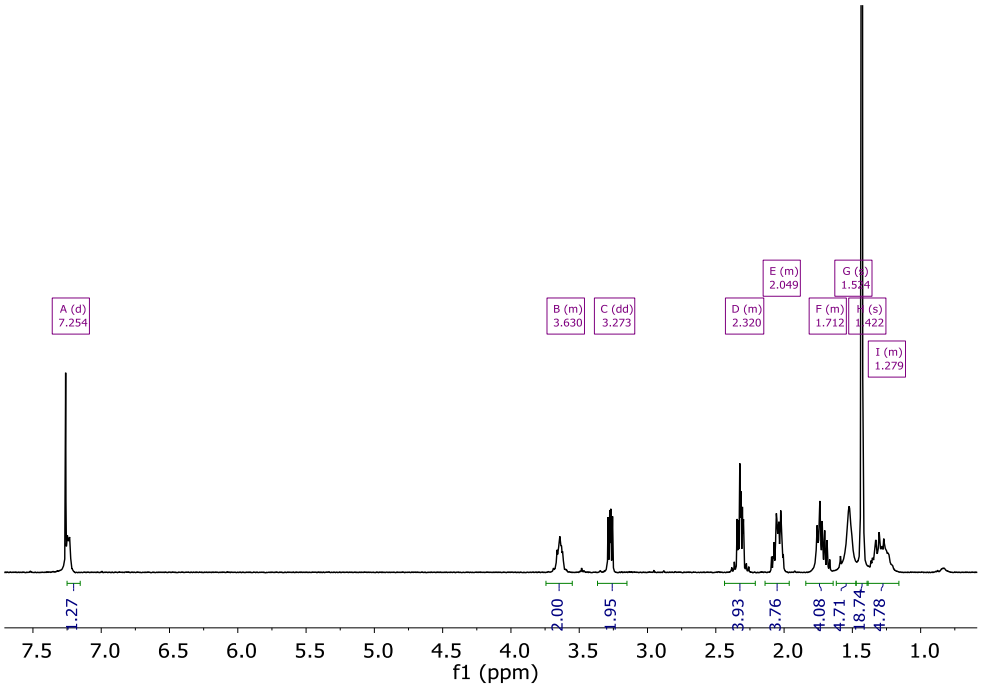


3a (1.97. g, 2.12 mmol) was dissolved in 15 mL of a solution of piperidine 20% in DMF. The solution was stirred at room temperature for 6 hours. Then water (15 mL) was added to the mixture. The white precipitate formed was filtered off. Product was extracted from the aqueous phase with dichloromethane (3x15 mL). Organic layer was dried over magnesium sulfate and evaporated to dryness. **3b** was obtained as a white solid (0.82g, 1.75 mmol, 80% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.254 (d, *J* = 5.1 Hz, 2H), 3.630 (m, 3H), 3.273 (dd, *J* = 8.0, 5.1 Hz, 2H), 2.320 (m, 4H), 2.049 (m, 4H), 1.712 (m, 4), 1.524 (s, 2H), 1.422 (s, 18H), 1.279 (m, 4H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 175.18, 172.64, 80.62, 54.71, 53.27, 32.48, 32.19, 30.81, 28.18, 24.79.

HRMS (ESI-TOF) m/z [**CyGlu_int2**+ H]$^{+}$ Calc: 485.3334, found: 485.335.



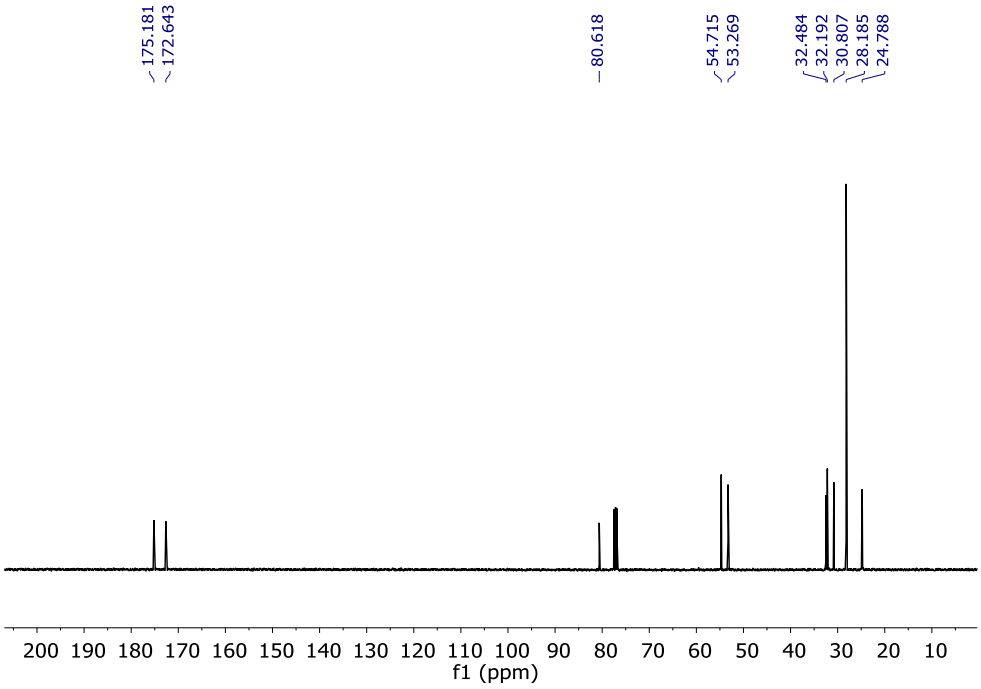


Figure 4.54. ^{1}H-NMR (400 MHz, $CDCl_3$) and ^{13}C NMR (101 MHz, $CDCl_3$) of **CyGlu_int2.**

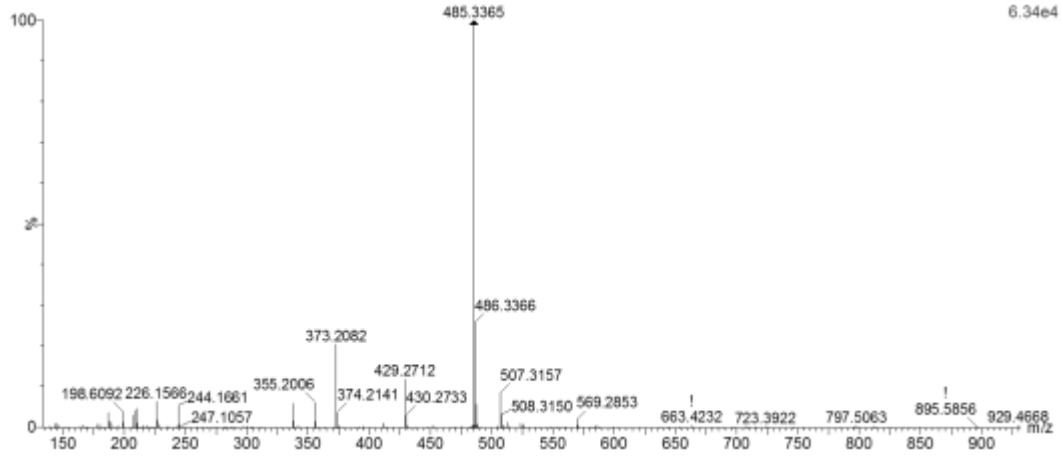
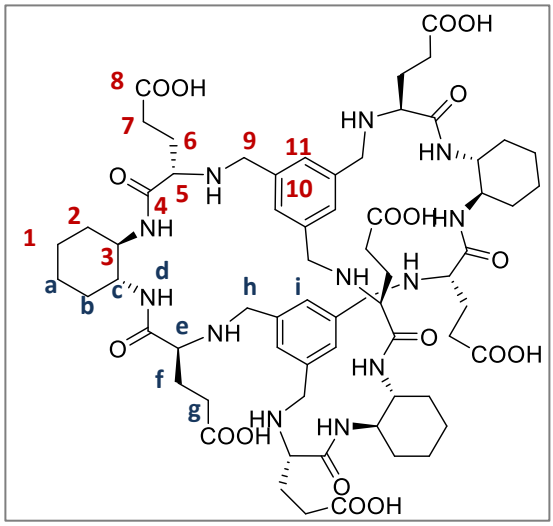


Figure 4.55. HRMS (ESI+) experimental spectrum of **CyGlu_int2.**

CyGlu



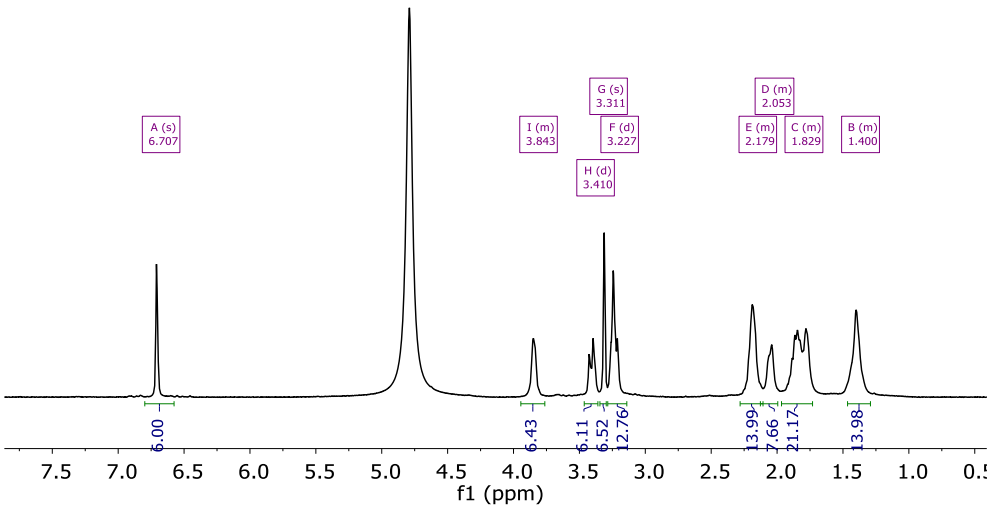
CyGlu_int2 (0.65g, 1.34 mmol) was dissolved in anhydrous MeOH (15 mL). 1,3,5-benzenecarbaldehyde (144 g, 0.89 mmol) was added dissolved in 10 mL of MeOH. The solution was stirred at room temperature for 24 hours. Then $NaBH_4$ (0.338 g, 8.9 mmol) was added and the reaction mixture was stirred for 24 hours more. After reaction was completed, NH_4Cl (aq) was added to neutral pH and the product was extracted with DCM. **CyGlu_int3** was purified by column chromatography (0-5% MeOH in DCM). **CyGlu_int3** was obtained as a white solid (550 mg, 0.32 mmol, 75%yield).

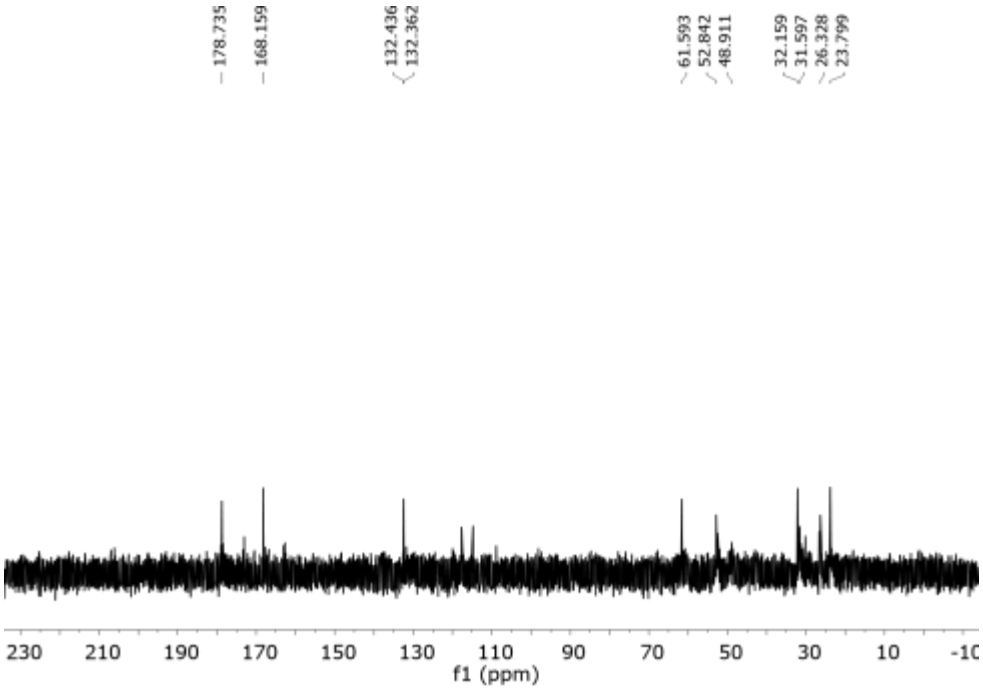
CyGlu_int3 (200 mg, 118 mmol) was dissolved in DCM (2 mL). Triethylsilane (0.2 mL) and H_3PO_4 (2 ml) were added. The reaction mixture was stirred at room temperature for 1 hour. Then K_2CO_3 (conc) was added until neutral pH. The product was extracted with chloroform/IPA (3:1) dried over $MgSO_4$, evaporated and further purified by reverse phase chromatography giving **CyGlu** as a TFA salt white solid (100 mg, 0.074 mmol, 40% yield).

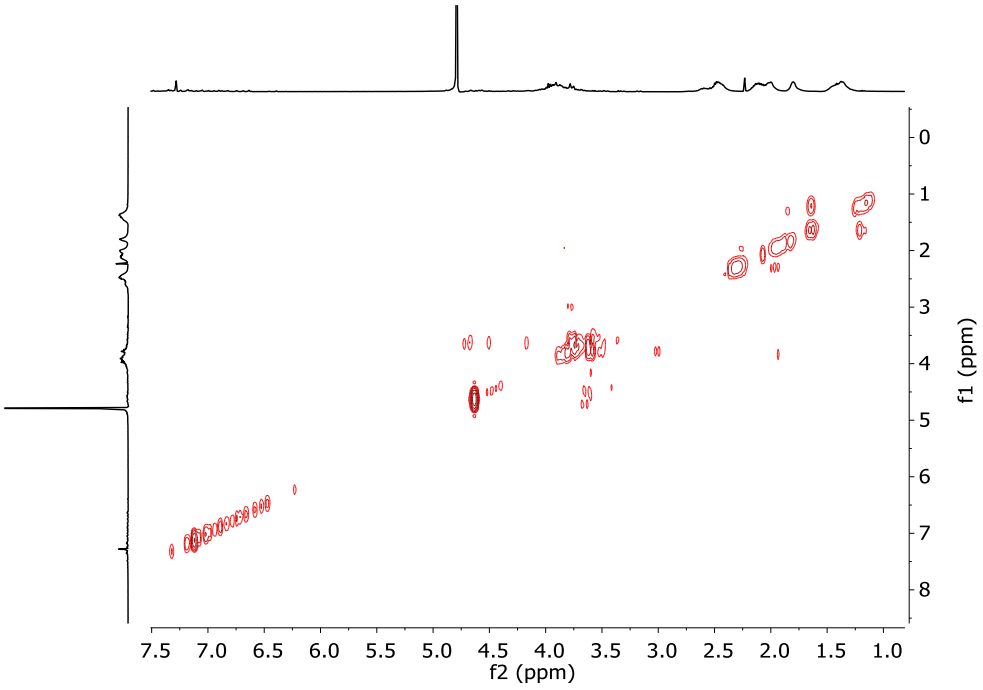
^{1}H NMR (400 MHz, CD_3OD) δ 6.707 (s, 6H, (**h**)), 3.843 (m, 6H, (**c**)), 3.410 (A subsystem from AB_q d, J_{ab} = 12.2 Hz, 6H, (**h**)), 3.311 (m, 6H, (**e**)), 3.227 (B subsystem from AB_q J = 12.2 Hz, 6H and another m, 6H, (**h**)), 2.179 (m, 12H, (**g**)), 2.053 (m, 6H, (**f**)), 1.829 (m, 21H(**f+b+a**)), 1.400 (m, 12H, (**b+a**)).

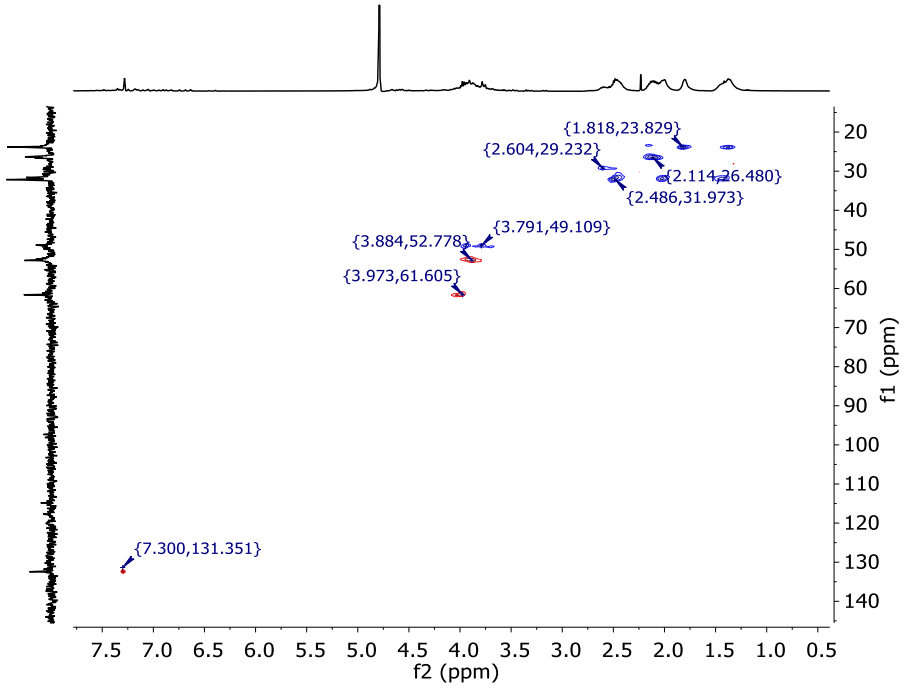
^{13}C NMR (101 MHz, D_2O) δ 178.74(**8**), 168.16(**4**), 132.44(**10**), 132.36(**11**), 61.59(**3**), 52.84(**5**), 48.91(**9**), 32.16(**2/7**), 31.60(**2/7**), 26.33(6), 23.80(**1**).

HRMS (ESI-TOF) m/z [**CyGlu**+ H]$^+$ Calc: 1345.6966, found: 1345.6829


A (s)
6.707
I (m)
3.843
G (s)
3.311
F (d)
3.227
H (d)
3.410
D (m)
2.053
E (m)
2.179
C (m)
1.829
B (m)
1.400
6.00
6.43
6.11
6.52
12.76
13.99
7.66
21.17
13.98
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f1 (ppm)


178.735
168.159
132.436
132.362
61.593
52.842
48.911
32.159
31.597
26.328
23.799
230 210 190 170 150 130 110 90 70 50 30 10 -10
f1 (ppm)


0
1
2
3
4
5
6
7
8
f1 (ppm)
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
f2 (ppm)


{1.818,23.829}
{2.604,29.232}
{2.114,26.480}
{2.486,31.973}
{3.791,49.109}
{3.884,52.778}
{3.973,61.605}
{7.300,131.351}
20
30
40
50
60
70
80
90
100
110
120
130
140
f1 (ppm)
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f2 (ppm)

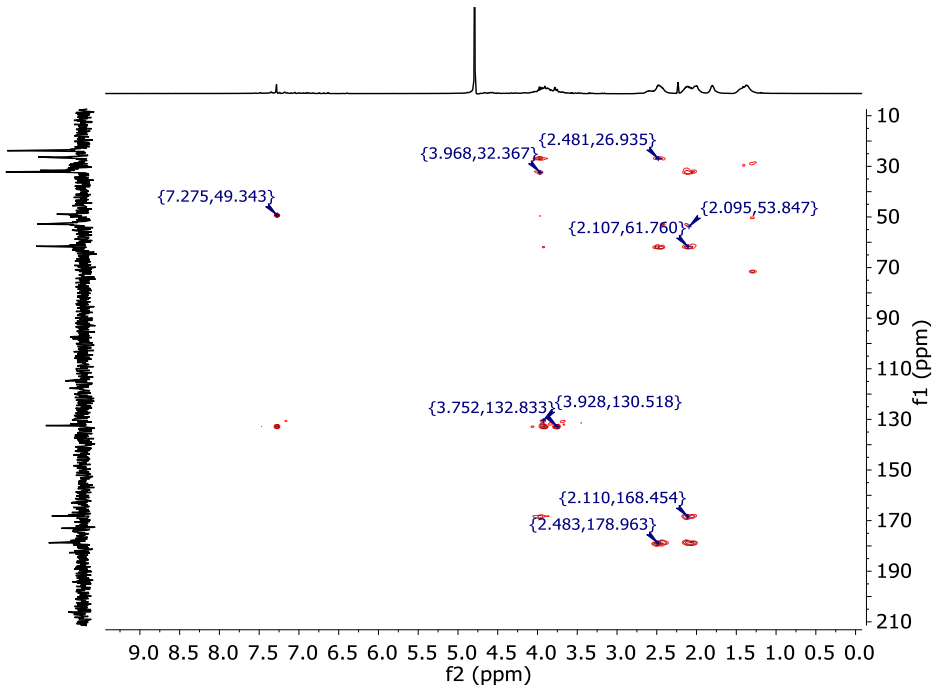


Figure 4.56. ^{1}H-NMR (400 MHz, D_2O), ^{13}C NMR (101 MHz, D_2O), COSY HSQC and HMBC of **CyGlu**

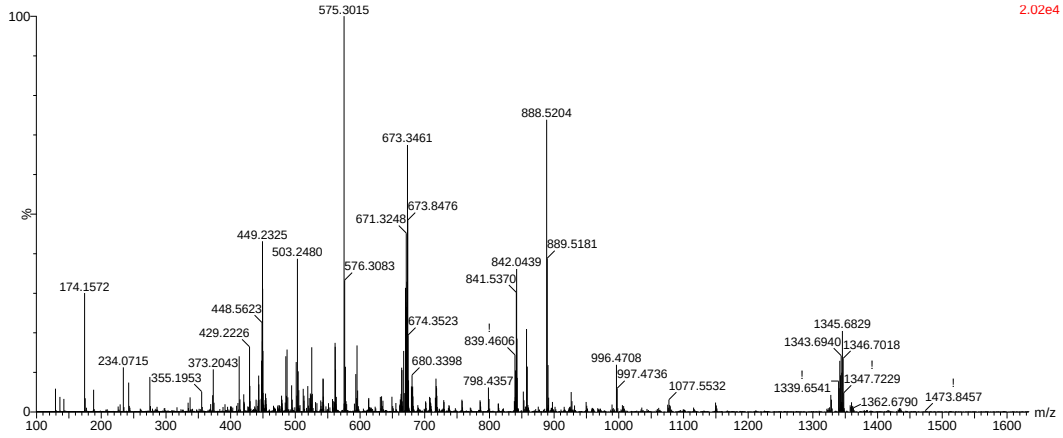


Figure 4.57. HRMS (ESI+) experimental spectrum **of CyGlu.**

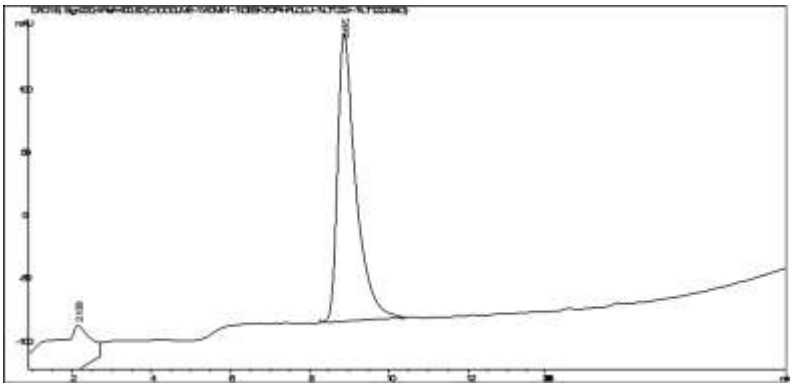


Figure 4.58. HPLC spectra of **CyGlu.**

4.10 FLUORESCENCE TITRATIONS

4.10.1 Titration of PolyE$_4$Y by CyAsp

A 2.0 · 10^{-4} M solution of PolyE$_4$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 1 · 10^{-3} M solution of **CyAsp** in its HCl salt.

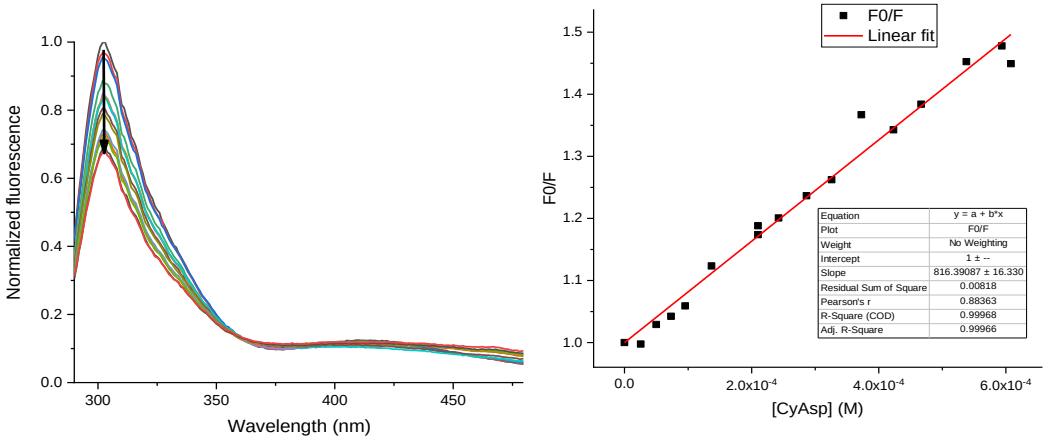


Figure 4.59. A) Normalized emission spectra of PolyE$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyAsp** ([CyAsp] = 0-0.8 · 10^{-3} M). [PolyE$_4$Y] = 2 · 10^{-4} M. λexc = 276 nm. B) Stern-Volmer plot of the titration of PolyE$_4$Y by **CyAsp** at 302 nm and linear Fit of the data.

The addition of **CyAsp** to the buffered PolyE$_4$Y solution provokes a quenching of the Tyr fluorescence band of this polypeptide. to The Stern-Volmer plot at 300 nm shows a linear relation vs the cage concentration with a K_{sv} = 816 M^{-1}. The major contribution to the

fluorescence quenching is in this case collisional quenching. Moreover, no band associated to the excimer formation appears. The linearity does not necessarily mean that there is no interaction but that this interaction is lower than the K_{sv}. The K_d for this interaction cannot be properly determined in the conditions assayed but we can conclude that it is necessarily higher than 1225 µM (inverse of the K_{sv}).

4.10.2 Titration of AC-EYE-NH_2 by CyAsp

A $2.0 \cdot 10^{-4}$ M solution of AC-EYE-NH_2 in buffered water (50 mM TRIS pH = 7.3) was titrated with a $1.4 \cdot 10^{-3}$ M solution of CyAsp in its HCl salt.

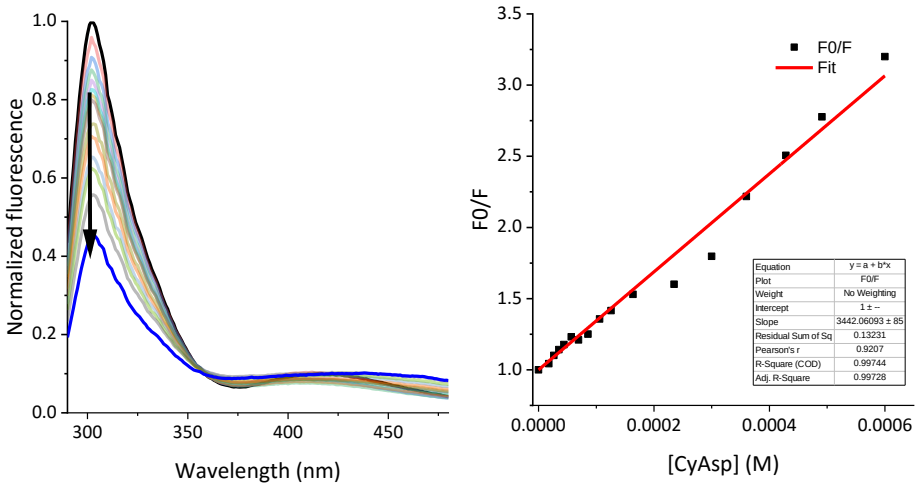


Figure 4.60. Normalized emission spectra of AC-EYE-NH_2 in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyAsp** ([CyAsp] = 0-0.8 · 10^{-3} M). [AC-EYE-NH_2] = 2 · 10^{-4} M. λexc = 276 nm. Stern-Volmer plot of the titration of AC-EYE-NH_2 by **CyAsp** at 302 nm and linear Fit of the data.

K_{sv} = 3442 ± 85 M^{-1}, $K_d \geq 1/K_{sv}$ = 290 µM (lower limit for the K_d constant).

Once again the linearity of the S-V plot and the absence of a new band corresponding to the formation of the complex indicates that the binding is lower than the K_{sv}. No accurate value for the K_d can be extracted from de titration data collected with any of the data fitting procedures described hence the same titration was conducted lowering the peptide concentrations.

A $.0 \cdot 10^{-5}$ M solution of AC-EYE-NH_2 in buffered water (50 mM TRIS pH = 7.3) was titrated with a $2 \cdot 10^{-3}$ M solution of **CyAsp** (Figure 4.22, Figure 4.61).

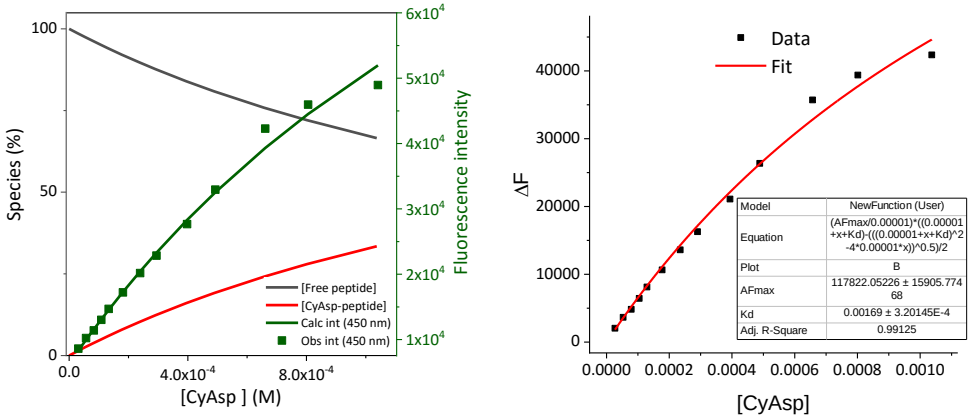


Figure 4.61. Species distribution and fluorescence intensity at 450 nm (observed and calculated) for the titration of AC-EYE-NH_2 (10 µM) with **CyAsp** at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec (left) and origin (right).

At this new conditions the fitting of the data to a 1 : 1 interaction model was possible both using HypSpec and Eq. 3.1.

- HypSpec results: CyAsp + Peptide ⇆ CyAsp-peptide Log β = 2.69 ± 0.01 K_d = 2120 ± 78 µM.
- Fitting with eq. 3.1: K_d = 1690 ± 320 µM.

4.10.3 Titration of polyK$_4$Y by CyAsp

A 2.0 · 10^{-4} M solution of PolyK$_4$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 4 · 10^{-3} M solution of **CyAsp** in its HCl salt. The addition of **CyAsp** to PolyK$_4$Y is quenching the fluorescence of the Tyrosine band. In figure 4.48 we can observe an isosbestic point probably caused by a weak interaction between the tyrosine and **CyAsp**. The very low fluorescence change in the region of the isosbestic point together with the little deviation of the S-V plot from linearity points out that once again the specific interaction is lower than the K_{sv} (K_{sv} = 4048 M^{-1}). So $K_d \geq$ 247 µM.

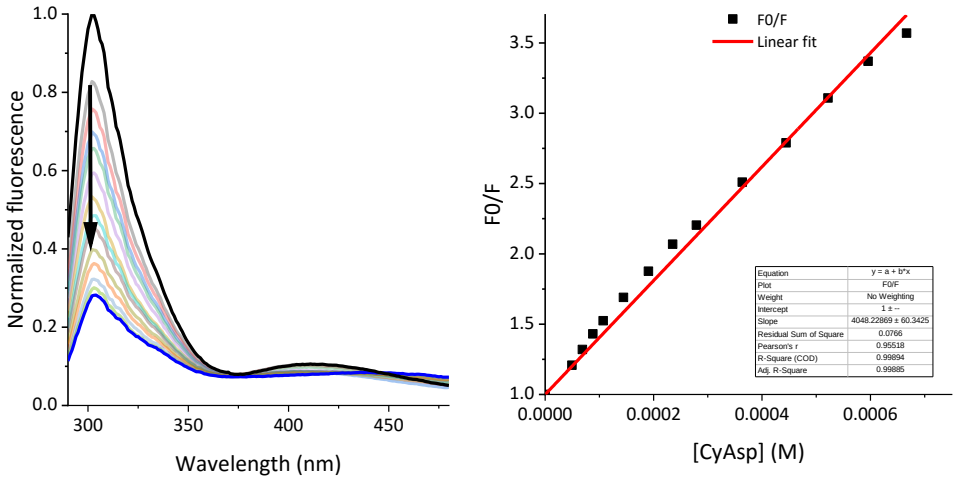


Figure 4.62. Normalized emission spectra of PolyK$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyAsp** ([CyAsp] = 0-0.8 · 10^{-3} M). [PolyK$_4$Y] = 2 · 10^{-4} M. λexc = 276 nm. S-V plot of the titration of PolyK$_4$Y by **CyAsp** at 302 nm and linear Fit of the data.

The same titration was conducted with a 2.0 · 10^{-5} M solution of PolyK$_4$Y.

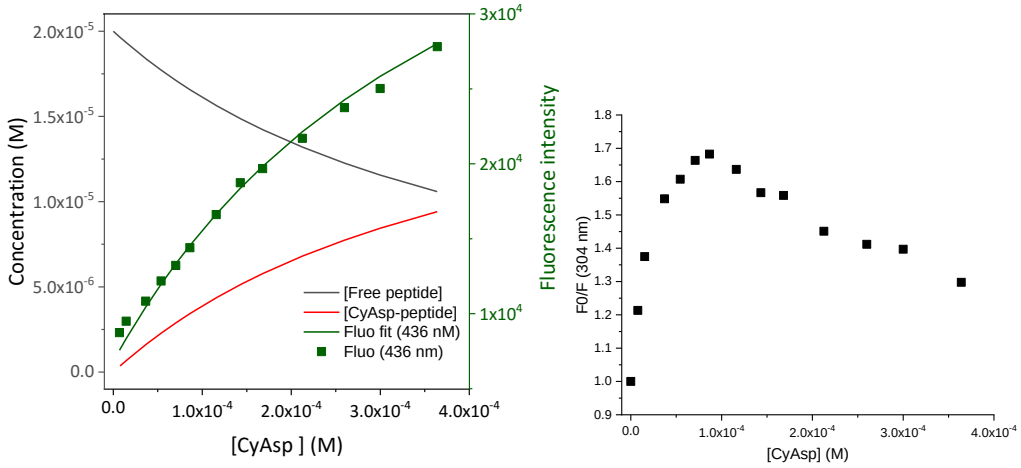


Figure 4.63. Normalized emission spectra of PolyK$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyAsp** ([CyAsp] = 0-0.6 ·10^{-3} M). [PolyK$_4$Y] = 2 · 10^{-5} M. λexc = 276 nm (left). Species distribution and fluorescence intensity at 305 nm (observed and calculated) for the titration of PolyK$_4$Y with **CyAsp** at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec (left). S-V plot of the titrations (right).

The presence of a growing new fluorescence band that increases upon the addition of CyAsp is an indication of the binding between CyAsp and the peptide. The S-V plot representation at 304 nm shows no linearity.

The fitting with HypSpec was carried out using the wavelength range from 350 to 450 nm and HypSpect results were:

CyAsp + Peptide ⇆ CyAsp-peptide Log β = 3.40 ± 0.01 K_d = 398 ± 10 µM

From this second titration we determined the K_d which is indeed as we predicted higher than the inverse of the K_{sv}.

The fitting with eq.4.1 was not possible in this case, data could not be fitted properly with that method.

4.10.4 Titration of polyE$_6$K$_3$Y by CyAsp

A 2.0 · 10^{-4} M solution of Poly E_6K_3Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 1 · 10^{-3} M solution of CyAsp in its HCl salt.

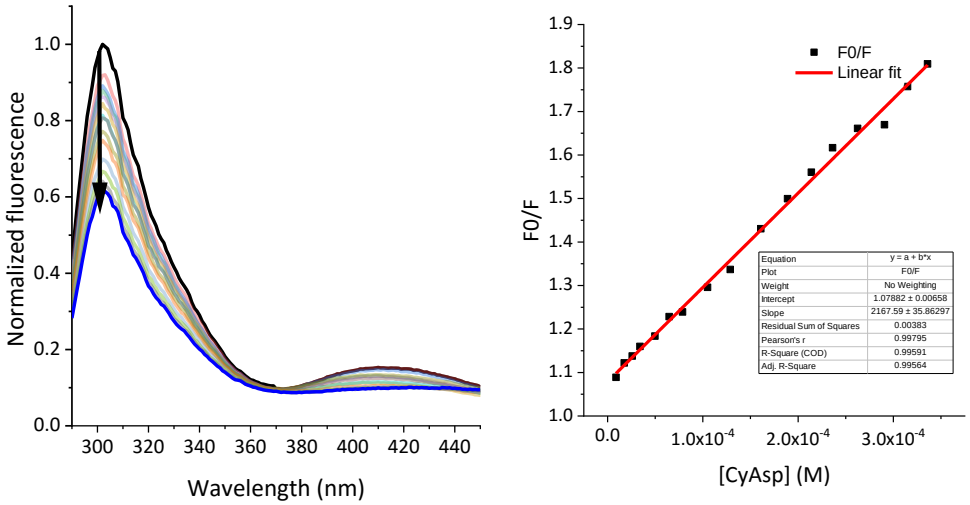


Figure 4.64. A) Normalized emission spectra of PolyE_6K_3Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyAsp** ([CyAsp] = 0-0.5 · 10^{-3} M). [E_6K_3Y] = 2 · 10^{-4} M. λexc = 276 nm. B) S-V plot of the titration of PolyE_6K_3Y by **CyAsp** at 302 nm and linear Fit of the data.

The S-V plot at the tyrosine emission band has a linear shape so this fluorescence change can be attributed to collisional quenching and no band associated to the complex formation

was detected in this case. The binding of **CyAsp** with PolyE$_6$K$_3$Y if happening lower than the K_{sv} (K_{sv} = 2167 M^{-1}). $K_d \geq$ 461 µM.

4.10.5 Titration of PolyE$_4$Y by CyGlu

A 2.0 · 10^{-4} M solution of polyE$_4$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 4 · 10^{-3} M solution of CyGlu. The effect of adding CyGlu are a quenching of the Tyrisine fluorescence mainly produced by collisional quenching (Linear S-V plot, supporting) with an isosbestic point at 350 nm linked to a complexation process.

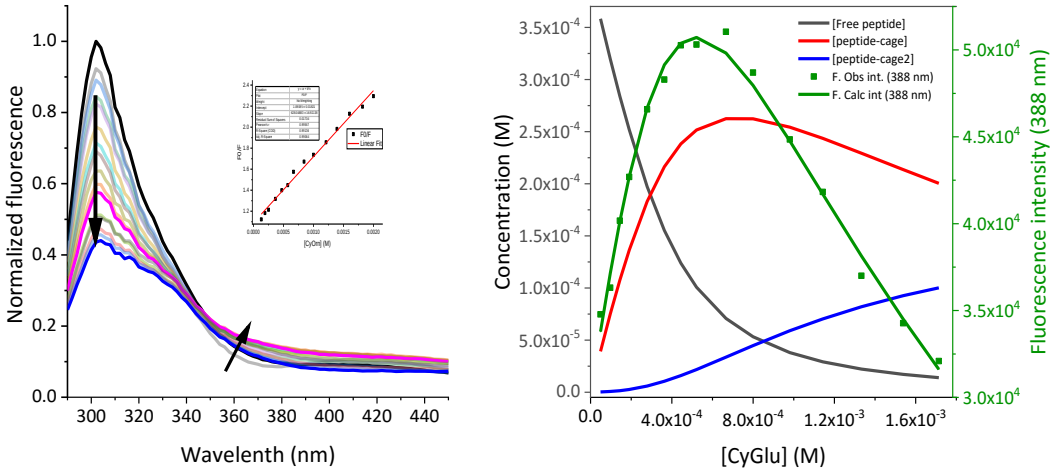


Figure 4.65. Normalized emission spectra of polyE$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of CyGlu ([CyGlu] = 0-2 · 10^{-3} M). [polyE$_4$Y] = 2 · 10^{-4} M. λexc = 276 nm and S-V plot at 302 nm, K_{sv} = 628 ± 17 M^{-1} (left). Species distribution and fluorescence intensity at 388 nm (observed and calculated) for the titration of PolyE$_4$Y vs CyGlu with a fitting model including the formation of the complexes 1 : 1 and 1 : 2 (right).

The plot of Fluorescence vs concentration at 388 nm with a big fluorescence increase followed by a big fluorescence decrease upon the addition of CyGlu might be a consequence of aggregation events at high concentrations of both peptide and CyGlu.

The same titration was conducted with a diluted solution of the peptide ([polyE$_4$Y] = 2 · 10^{-5} M) and data was fitted with HypSpec. In the diluted conditions the fluorescence of the

tyrosine is almost unaffected by the presence of CyGlu and the band corresponding to the complex formation is detected centred at 380 nm.

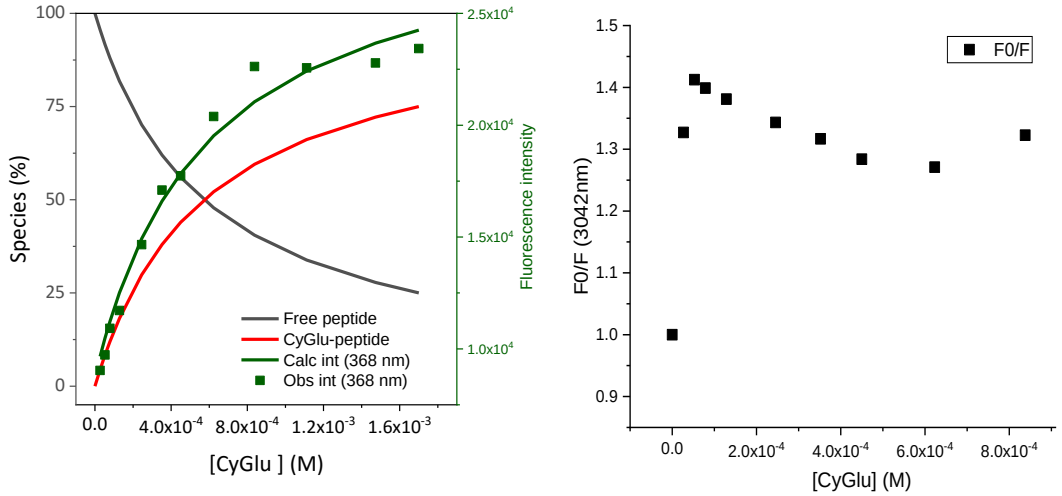


Figure 4.66. Species distribution and fluorescence intensity at 368 nm (observed and calculated) for the titration of PolyE$_4$Y with **CyGlu** at pH 7.3 with a fitting model for the formation of 1 : 1 complex using HypSpec (right). S-V plot at 302 nm and linear Fit of the data. K_{sv} can not be determined.

HypSpect results:

CyGlu + Peptide ⇆ CyGlu-peptide Log β = 3.25 ± 0.008, K_d = 560 ± 6 µM

4.10.6 Titration of AC-EYE-NH$_2$ by CyGlu

A 2.0 · 10^{-5} M solution of of Ac-EYE-NH$_2$ in buffered water (50 mM TRIS pH = 7.3) was titrated with a 2.67 · 10^{-3} M solution of **CyGlu.**

- HypSpec results: CyGlu + Peptide ⇆ CyGlu-peptide Log β = 2.37 ± 0.01 K_d = 4365 ± 190 µM.
- Fitting with eq. 3.1: K_d = 4100 ± 722 µM.

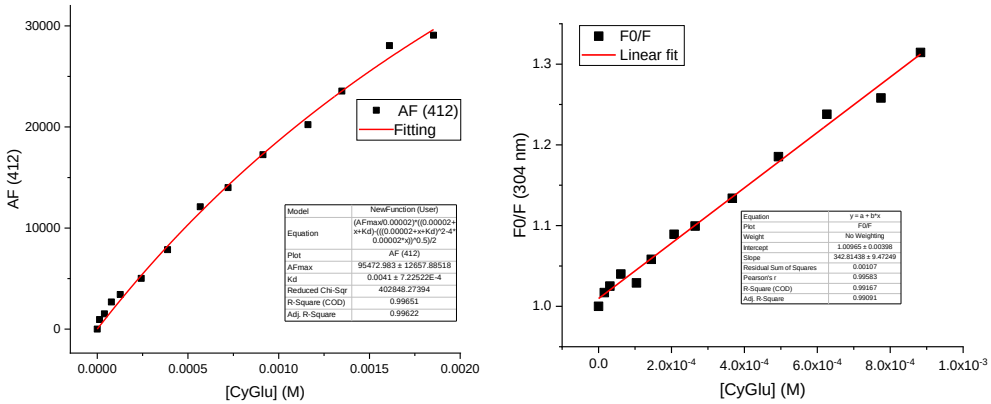


Figure 4.67. Fitting of the titration of Ac-EYE-NH_2 with **CyGlu** at pH 7.3 using eq. 3.1 and emission at 412 nm. S-V plot at 302 nm and linear Fit of the data. . K_{sv} = 342±9 M^{-1}.

4.10.7 Titration of polyK$_4$Y by CyGlu

A 2.0 · 10^{-5} M solution of polyK$_4$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 1.4 · 10^{-3} M solution of CyGlu. Data was fitted with HypSpect to determine the K_d of the complex formation.

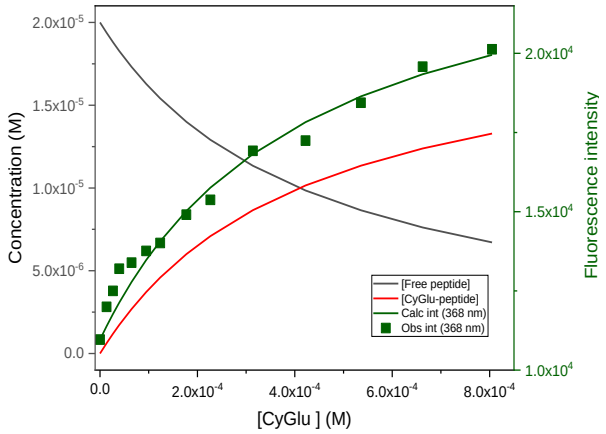


Figure 4.68 Normalized emission spectra of polyK$_4$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyGlu** ([CyGlu] = 0-1 · 10^{-3} M). [polyK$_4$Y] = 2 · 10^{-5} M. λexc = 276 nm. Species distribution and fluorescence intensity at 388 nm (observed and calculated) for the titration of polyK$_4$Y vs **CyGlu** with a fitting model for the formation of 1 : 1 complex using HypSpec.

HypSpec results:

CyGlu + Peptide ⇆ CyGlu-peptide Log β = 3.398 ± 0.009 K_d = 400 ± 10 µM.

The S-V plot of the Tyr bad for this titration is not linear and the K_d was also determined using eq 4.1 giving the same value (figure 4.69).

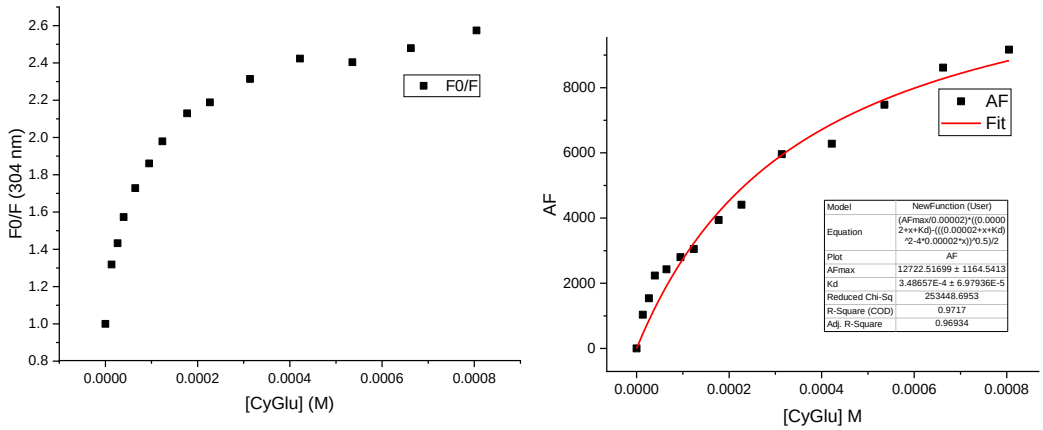


Figure 4.69. S-V plot of the titration of PolyK$_4$Y (20 µM) by **CyGlu** at 302 nm (left). Fitting of the data with equation 4.1 at λ = 368 nm (right). The K_d determined with the fitting was 348 ± 70 µM.

4.10.8 Titration of polyE$_6$K$_3$Y by CyGlu

A 2.0 · 10^{-4} M solution of polyE$_6$K$_3$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 2 · 10^{-3} M solution of **CyGlu**.

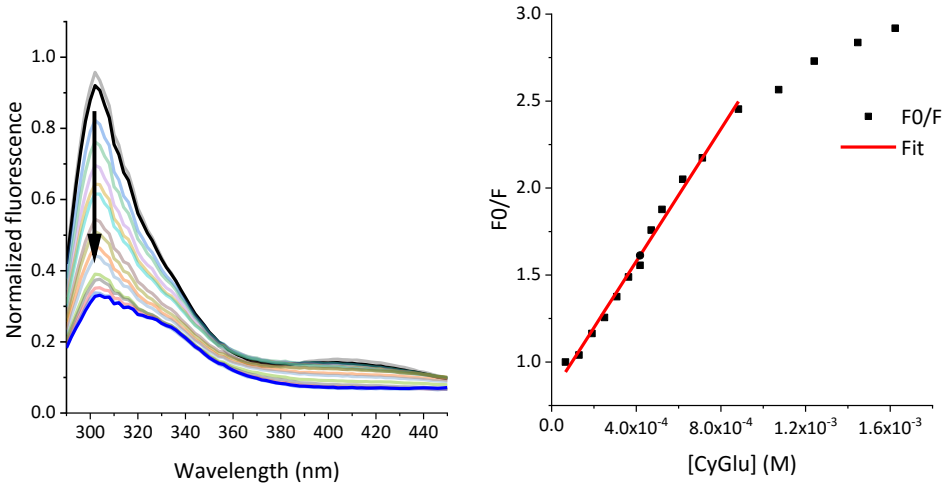


Figure 4.70. Normalized emission spectra of polyE$_6$K$_3$Y in buffered water (50 mM TRIS pH=7.3; in absence and in presence of different amounts of **CyGlu** ([CyGlu]= 0-1.7 · 10^{-3} M).

[polyE$_6$K$_3$Y] = 2 · 10^{-4} M. λexc = 276 nm. B) S-V plot of the titration of PolyE$_6$K$_3$Y by **CyGlu** at 302 nm and linear Fit of the data. K_{sv} = 1902 ± 56 M^{-1}

S-V plot renders a linear response for the fluorescence intensity up to a [CyGlu] = 1 mM. The last points suffer a deviation but since it is at high concentrations this can be a result of some aggregation events that thus reduce the availability of the fluorophore. There is no band corresponding to the excimer formation in this case either so no binding detected. If complexation occurs it must have a K_{ass} lower than K_{sv} therefore $K_d \geq 526$ µM.

4.10.9 Titration of polyK$_4$Y by CyLys

A 2.0 · 10^{-5} M solution of polyK$_4$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 3 · 10^{-3} M solution of **CyLys**. Graphs of the titration are in the results section.

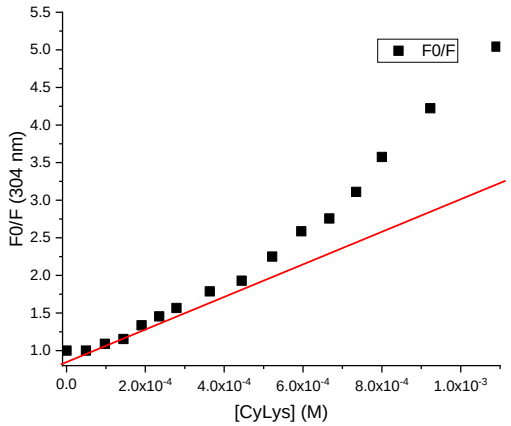


Figure 4.71. S-V plot of the titration of PolyK$_4$Y (200 µM) by **CyLys** at 304 nm. Red line would be the trend if the response were linear.

4.10.10 Titration of polyE$_6$K$_3$Y by CyLys

A 2.0 · 10^{-5} M solution of polyE$_6$K$_3$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 1.2 · 10^{-3} M solution of **CyLys.** Data was fitted with HypSpec to determine the K_d.

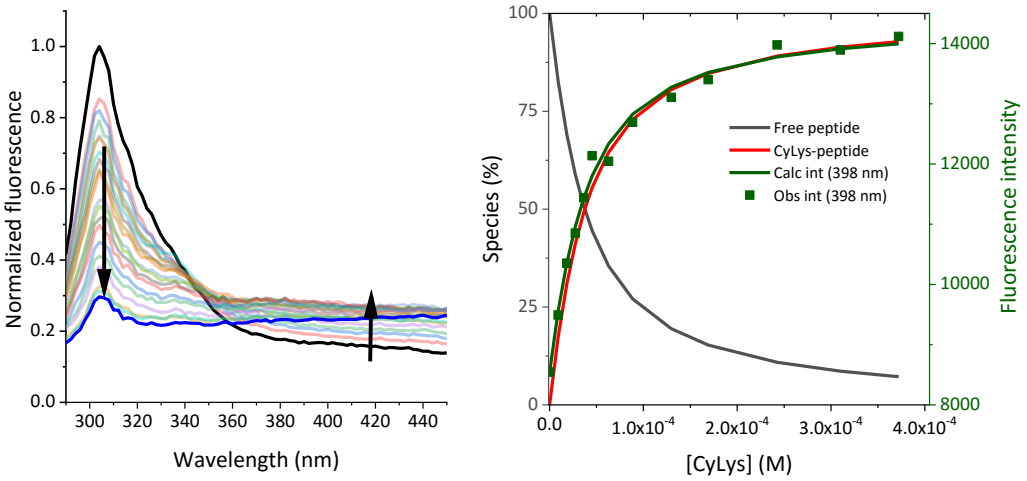


Figure 4.72. Normalized emission spectra of polyE$_6$K$_3$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyLys** ([CyLys] = 0 - 0.6 · 10^{-3} M). [polyE$_6$K$_3$Y] = 2 · 10^{-5} M. λexc = 276 nm (left). Figure: species distribution and fluorescence intensity at 302 nm (observed and calculated) for the titration of PolyE$_6$K$_3$Y vs **CyLys** with a fitting model for the formation of a 1 : 1 complex.

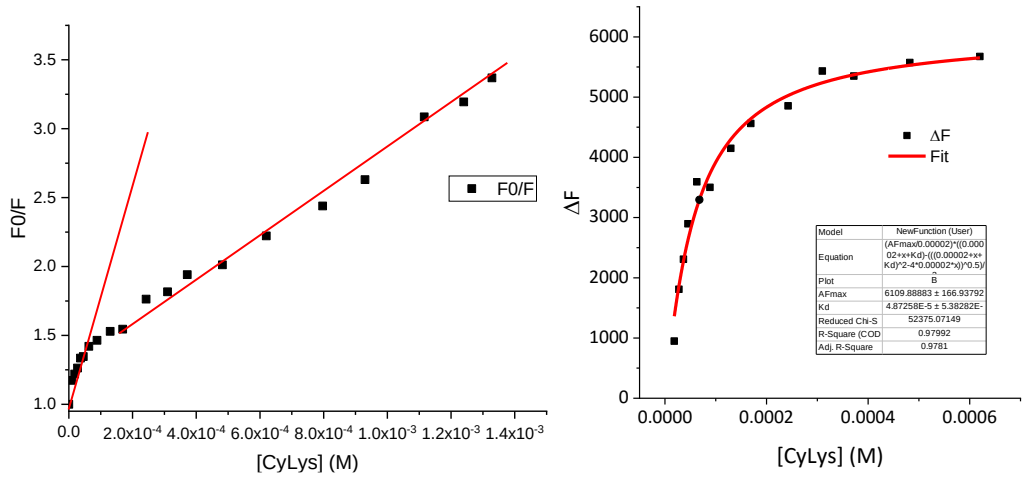


Figure 4.73. Stern-Volmer plot of the titration of polyE$_6$K$_3$Y (20 µM) by CyLys at 302 nm. Red line would be the trend if the response were linear. Up to 0.2 mM of CyLys the response is linear and data could be fitted to a linear equation getting a K_{sv} = 1502 M^{-1}. Fitting of the titration of polyE$_6$K$_3$Y with CyLys at pH 7.3 using eq. 3.1 and emission at 398 nm.

- HypSpec results: CyLys + Peptide ⇆ CyLys-peptide Log β = 4.56 ± 0.01 K_d = 27.5± 0.6 µM.

- Fitting with eq. 3.1: 01 K_d = 48 ± 5 µM.

4.10.11 Titration of polyK$_4$Y by CyOrn

A 2.0 · 10^{-4} M solution of polyK$_4$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 2 · 10^{-3} M solution of **CyOrn.**

Through this titration a new band corresponding to the formation of the complex was not detected, however there is an isosbestic point at 360 nm which can be associated to some weak interaction. Moreover the S-V plot at 302 nm (Tyr emission band) is deviated from the linearity which indicated that not just dynamic quenching is producing the fluorescence decrease upon the addition of the cage so a fitting with HypSpect was possible for this data. The same titration was conducted with a more diluted solution of the peptide:

A 2.0 · 10^{-5} M solution of polyK$_4$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 3.2 · 10^{-3} M solution of **CyOrn.**

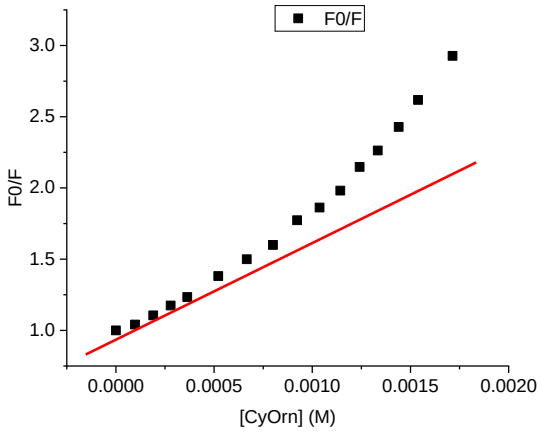


Figure 4.74. Stern-Volmer plot of the titration of PolyK$_4$Y by **CyOrn** at 302 nm. Red line would be the trend if the response were linear.

The results of this titration cannot be fitted to the simple 1 : 1 model not with HypSpec neither with eq.4.1. Spectra are plotted in the results section.

4.10.12 Titration of polyE$_6$K$_3$Y by CyOrn

A 2.0 · 10^{-5} M solution of polyE$_6$K$_3$Y in buffered water (50 mM TRIS pH = 7.3) was titrated with a 1.8 · 10^{-3} M solution of **CyOrn.** The presence of a new fluorescence band centered at 380 nm was linked to the complex formation. Data was successfully fitted to a 1 : 1 model using HypSpec. Fitting with eq.4.1 was also possible with a very similar result.

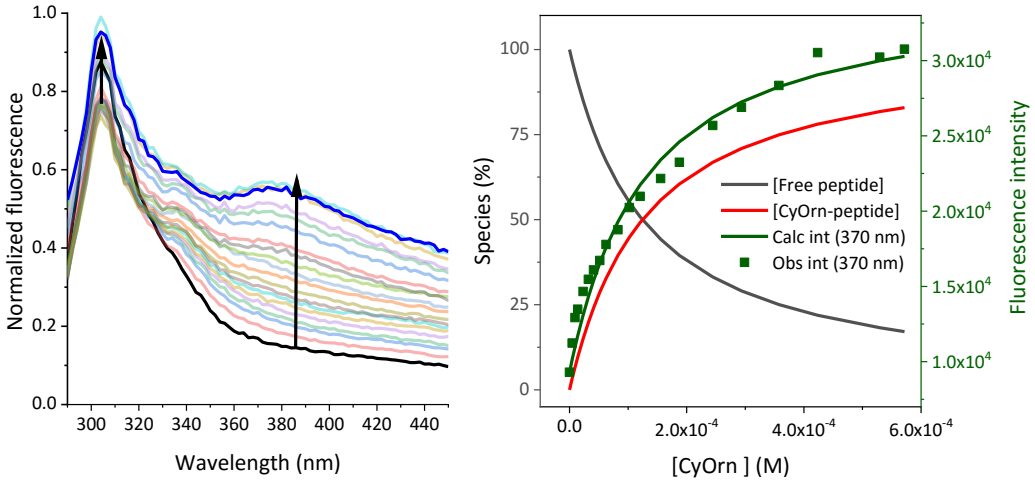


Figure 4.75. A) Normalized emission spectra of polyE$_6$K$_3$Y in buffered water (50 mM TRIS pH = 7.3; in absence and in presence of different amounts of **CyOrn** ([CyOrn] = 0-0.7 · 10^{-3} M). [poly polyE$_6$K$_3$Y] = 2 · 10^{-5} M. λexc = 276 nm. B) Species distribution and fluorescence intensity at 370 nm (observed and calculated) for the titration of polyE$_6$K$_3$Y with CyOrn at pH 7.3 with a fitting model for the formation of complex 1 : 1 using HypSpec.

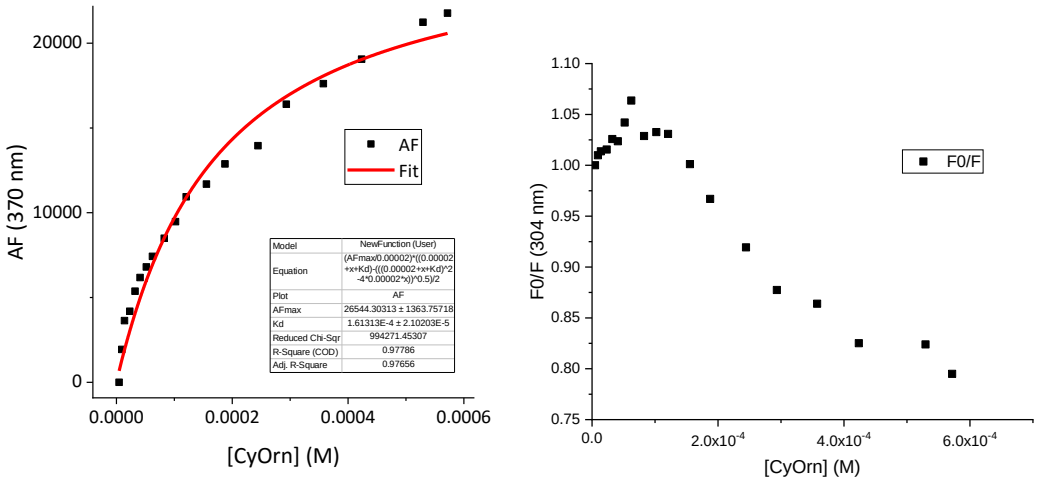


Figure 4.76. Fitting of the titration of polyE$_6$K$_3$Y (20 µM) with **CyOrn** at pH 7.3 using eq. 3.1 and emission at 370 nm. The Kd determined with the fitting was **161 ± 20 µM.** S-V plot at 302 nm.

- HypSpec results: CyOrn + Peptide ⇆ CyOrn-peptide Log β = 3.94 ± 0.007 K_d = 114 ± 2 µM.
- Fitting with eq. 4.1: K_d = 161 ± 21 µM.

4.11 BIBLIOGRAPHY

1 E. Faggi, Y. Pérez, S. V. Luis and I. Alfonso, *Chem. Commun.*, 2016, **52**, 8142–8145.

2 I. Martí, J. Rubio, M. Bolte, M. I. Burguete, C. Vicent, R. Quesada, I. Alfonso and S. V. Luis, *Chem. - A Eur. J.*, 2012, **18**, 16728–16741.

3 I. Martí, M. Bolte, M. I. Burguete, C. Vicent, I. Alfonso and S. V. Luis, *Chem. - A Eur. J.*, 2014, **20**, 7458–7464.

4 L. Tapia, Y. Pérez, M. Bolte, J. Casas, J. Solà, R. Quesada and I. Alfonso, *Angew. Chemie - Int. Ed.*, 2019, **58**, 12465–12468.

5 E. Faggi, S. V Luis and I. Alfonso, *Curr. Med. Chem.*, 2019, **26**, 4065–4097.

6 E. Faggi, A. Moure, M. Bolte, C. Vicent, S. V. Luis and I. Alfonso, *J. Org. Chem.*, 2014, **79**, 4590–4601.

7 E. Faggi, R. Gavara, M. Bolte, L. Fajarí, L. Juliá, L. Rodríguez and I. Alfonso, *Dalt. Trans.*, 2015, **44**, 12700–12710.

8 G. S. Martin, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 467–475.

9 H. Krishnan, W. T. Miller and G. S. Goldberg, *Genes Cancer*, 2012, **3**, 426–435.

10 R. Roskoski, *Pharmacol. Res.*, 2015, **100**, 1–23.

11 V. Jha, M. Macchia, T. Tuccinardi and G. Poli, *Cancers* , 2020, 12.

12 S. Martellucci, L. Clementi, S. Sabetta, V. Mattei, L. Botta and A. Angelucci, *Cancers* , 2020, 12.

13 A. Simatou, G. Simatos, M. Goulielmaki, D. Spandidos A., S. Baliou and V. Zoumpourlis, *Mol Clin Oncol*, 2020, **13**, 21.

14 H. J. Schneider, P. Agrawal and A. K. Yatsimirsky, *Chem. Soc. Rev.*, 2013, **42**, 6777–6800.

15 I. Alfonso, M. Bolte, M. Bru, M. I. Burguete, S. V Luis and C. Vicent, *Org. Biomol. Chem.*, 2010, **8**, 1329–1339.

16 G. Yu and X. Chen, *Theranostics*, 2019, **9**, 3041–3074.

17 S. Van Dun, C. Ottmann, L. G. Milroy and L. Brunsveld, *J. Am. Chem. Soc.*, 2017, **139**, 13960–13968.

18 G. V. Oshovsky, D. N. Reinhoudt and W. Verboom, *Angew. Chemie - Int. Ed.*, 2007, **46**, 2366–2393.

19 S. J. Barrow, S. Kasera, M. J. Rowland, J. Del Barrio and O. A. Scherman, *Chem. Rev.*, 2015, **115**, 12320–12406.

20 G. Ghale, V. Ramalingam, A. R. Urbach and W. M. Nau, *J. Am. Chem. Soc.*, 2011, **133**, 7528–7535.

21 H. J. Buschmann, L. Mutihac, R. C. Mutihac and E. Schollmeyer, *Thermochim. Acta*, 2005, **430**, 79–82.

22 K. I. Assaf and W. M. Nau, *Chem. Soc. Rev.*, 2015, **44**, 394–418.

23 L. C. Smith, D. G. Leach, B. E. Blaylock, O. A. Ali and A. R. Urbach, *J. Am. Chem. Soc.*, 2015, **137**, 3663–3669.

24 T. Schrader, G. Bitan and F. G. Klärner, *Chem. Commun.*, 2016, **52**, 11318–11334.

25 M. Fokkens, T. Schrader and F.-G. Klärner, *J. Am. Chem. Soc.*, 2005, **127**, 14415–14421.

26 A. Attar, C. Ripoli, E. Riccardi, P. Maiti, D. D. Li Puma, T. Liu, J. Hayes, M. R. Jones, K. Lichti-Kaiser, F. Yang, G. D. Gale, C. Tseng, M. Tan, C.-W. Xie, J. L. Straudinger, F.-G. Klärner, T. Schrader, S. A. Frautschy, C. Grassi and G. Bitan, *Brain*, 2012, **135**, 3735–3748.

27 S. Prabhudesai, S. Sinha, A. Attar, A. Kotagiri, A. G. Fitzmaurice, R. Lakshmanan, M. I. Ivanova, J. A. Loo, F.-G. Klärner, T. Schrader, M. Stahl, G. Bitan and J. M. Bronstein, *Neurotherapeutics*, 2012, **9**, 464–476.

28 F. Perret and A. W. Coleman, *Chem. Commun.*, 2011, **47**, 7303–7319.

29 N. Douteau-Guével, F. Perret, A. W. Coleman, J. P. Morela and N. Morel-Desrosiers, *J. Chem. Soc. Perkin Trans. 2*, 2002, 524–532.

30 G. Zhang, Y. Li and X. Zhao, *J Incl Phenom Macrocycl Chem*, 2012, **72**, 473–479.

31 R. E. McGovern, H. Fernandes, A. R. Khan, N. P. Power and P. B. Crowley, *Nat. Chem.*, 2012, **4**, 527–533.

32 R. E. McGovern, A. A. McCarthy and P. B. Crowley, *Chem. Commun.*, 2014, **50**, 10412–10415.

33 M. Kubo, E. Nashimoto, T. Tokiyo, Y. Morisaki, M. Kodama and H. Hioki, *Tetrahedron Lett.*, 2006, **47**, 1927–1931.

34 A. W. Coleman, F. Perret, A. Moussa, M. Dupin, Y. Guo and H. Perron, ed. T. Schrader, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 31–88.

35 K. Bezouska, R. Snajdrová, K. Krenek, M. Vancurová, A. Kádek, D. Adámek, P. Lhoták, D. Kavan, K. Hofbauerová, P. Man, P. Bojarová and V. Kren, *Bioorg. Med. Chem.*, 2010, **18**, 1434–1440.

36 Y. Hamuro, M. C. Calama, H. S. Park and A. D. Hamilton, *Angew. Chemie Int. Ed. English*, 1997, **36**, 2680–2683.

37 Y. Wei, G. L. Mc Lendon, A. D. Hamilton, M. A. Case, C. B. Purring, Q. Lin, H. S. Park, C. S. Leec and T. Yua, *Chem. Commun.*, 2001, **1**, 1580–1581.

38 H. S. Park, Q. Lin and A. D. Hamilton, *J. Am. Chem. Soc.*, 1999, **121**, 8–13.

39 V. Martos, S. C. Bell, E. Santos, E. Y. Isacoff, D. Trauner and J. De Mendoza, *PNAS*, 2009, **106**, 10482–10486.

40 A. Vovk, V. Kalchenko, O. Muzychka, V. Tanchuk, I. Muravyova, A. Shivanyuk, S. Cherenok and V. Kukhar, *Phosphorus. Sulfur. Silicon Relat. Elem.*, 2008, **183**, 625–626.

41 A. Buryak and K. Severin, *Angew. Chemie Int. Ed.*, 2004, **43**, 4771–4774.

42 S. Tashiro, M. Kobayashi and M. Fujita, *J. Am. Chem. Soc.*, 2006, **128**, 9280–9281.

43 S. H. Hewitt, M. H. Filby, E. Hayes, L. T. Kuhn, A. P. Kalverda, M. E. Webb and A. J. Wilson, *ChemBioChem*, 2016, **18**, 151.

44 J. Mosquera, B. Szyszko, S. K. Y. Ho and J. R. Nitschke, *Nat. Commun.*, 2017, **8**, 6–11.

45 S. Sun, M. A. Fazal, B. C. Roy and S. Mallik, *Org. Lett.*, 2000, **2**, 911–914.

46 A. F. Martins, D. M. Dias, J. F. Morfin, S. Lacerda, D. V. Laurents, É. Tõth and C. F. G. C. Geraldes, *Chem. - A Eur. J.*, 2015, **21**, 5413–5422.

47 B. E. Collins and E. V. Anslyn, *Chem. - A Eur. J.*, 2007, **13**, 4700–4708.

48 S. M. Ngola, P. C. Kearney, S. Mecozzi, K. Russell and D. A. Dougherty, *J. Am. Chem. Soc.*, 1999, **121**, 1192–1201.

49 C. Schmuck and L. Hernandez-Folgado, *Org. Biomol. Chem.*, 2007, **5**, 2390–2394.

50 S. Rensing, M. Arendt, A. Springer, T. Grawe and T. Schrader, *J. Org. Chem.*, 2001, **66**, 5814–5821.

51 S. Rensing and T. Schrader, *Org. Lett.*, 2002, **4**, 2161–2164.

52 M. Jewginski, L. Fischer, C. Colombo, I. Huc and C. D. Mackereth, *ChemBioChem*, 2016, **17**, 727–736.

53 Y. Akbaba, E. Bastem, F. Topal, İ. Gülçin, A. Maraş and S. Göksu, *Arch. Pharm. (Weinheim).*, 2014, **347**, 950–957.

54 S. Göksu, A. Naderi, Y. Akbaba, P. Kalın, A. Akıncıoğlu, İ. Gülçin, S. Durdagi and R. E. Salmas, *Bioorg. Chem.*, 2014, **56**, 75–82.

55 Y. Liu, J. Lee, L. Perez, A. D. Gill, R. J. Hooley and W. Zhong, , DOI:10.1021/jacs.8b08693.

56 T. L. Mako, J. M. Racicot and M. Levine, *Chem. Rev.*, 2019, **119**, 322–477.

57 A. Moure, S. V. Luis and I. Alfonso, *Chem. - A Eur. J.*, 2012, **18**, 5496–5500.

58 A. Moure, S. V. Luis and I. Alfonso, *Chem. Eur. J.*, 2012, **18**, 5496–5500.

59 A. Isidro-Llobet, M. Alvarez and F. Albericio, *Chem. Rev.*, 2009, **109**, 2455–2504.

60 E. Faggi, Y. Pérez, S. V. Luis and I. Alfonso, *Chem. Commun.*, 2016, **52**, 8142–8145.

61 J. S. M. Raymond J.A. Budde *, Nihal U. Obeyesekere, Shi Ke, *Biochim. Biophys. Acta*, 1995, **1248**, 50–56.

62 Y. Yang, L. H. Guo, N. Qu, M. Y. Wei, L. X. Zhao and B. Wan, *Biosens. Bioelectron.*, 2011, **28**, 284–290.

63 J. Blouin, P. Roby, M. Arcand, L. Beaudet and F. Lipari, *Curr. Chem. Genomics*, 2011, **5**, 115–121.

64 C. Schalley, *Analytical Methods in Supramolecular Chemistry*, 2007.

65 P. Gans, A. Sabatini and A. Vacca, *Talanta*, 1996, **43**, 1739–1753.

66 Http://www.hyperquad.co.uk/HypSpec2014.htm,

http://www.hyperquad.co.uk/HypSpec2014.htm.

67 A. Mrozek, J. Karolak-Wojciechowska and K. Kieć-Kononowicz, *J. Mol. Struct.*, 2003, **655**, 397–403.

68 S. K. Burley and G. A. Petsko, *FEBS Lett.*, 1986, **203**, 139–143.

69 S. M. Liao, Q. S. Du, J. Z. Meng, Z. W. Pang and R. B. Huang, *Chem. Cent. J.*, 2013, **7**, 1.

70 D. C. Bas, D. M. Rogers and J. H. Jensen, *Proteins*, 2008, **73**, 765–783.

71 H. Li, A. D. Robertson and J. H. Jensen, *Proteins Struct. Funct. Bioinforma.*, 2005, **61**, 704–721.

72 R. Loewenthal, J. Sancho and A. R. Fersht, *J. Mol. Biol.*, 1992, **224**, 759–770.

73 C. D. M. Churchill and S. D. Wetmore, *J. Phys. Chem. B*, 2009, **113**, 16046–16058.

74 J. Heyda, P. E. Mason and P. Jungwirth, *J. Phys. Chem. B*, 2010, **114**, 8744–8749.

75 K. Haghani, K. Khajeh, H. Naderi-Manesh and B. Ranjbar, *Int. J. Biol. Macromol.*, 2012, **50**, 1040–1047.

76 E. DASHTBAN MOGHADAM, N. AKBARI and K. KHAJEH, *Mol. Biol. Res. Commun.*, 2014, **3**, 119.

77 L. Zhang, J.-C. Wang, L. Hou, P.-R. Cao, L. Wu, Q.-S. Zhang, H.-Y. Yang, Y. Zang, J.-P. Ding and J. Li, *Sci. Rep.*, 2015, **5**, 10115.

78 Schrödingher, 2019.

79 Schrödinger Release 2019-1: Maestro, Schrödinger, LLC: New York, NY, 2019.

80 Schrödinger Release 2019-1: Macromodel, Schrödinger, LLC: New York, NY, 2019.

81 E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, *J. Chem. Theory Comput.*, 2016, **12**, 281–296.

82 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.

83 J. S. Brodbelt, *Int. J. Mass Spectrom.*, 2000, **200**, 57–69.

84 E. Kalenius and P. Vainiotalo, *Spectrosc. Eur.*, 2008, **20**, 13–16.

CHAPTER 5 MODULATION OF THE SRC KINASE ACTIVITY BY SELECTIVE SUBSTRATE RECOGNITION WITH PSEUDOPEPTIDIC CAGES

5.1 INTRODUCTION: PROTEIN TYROSINE KINASES

Kinases are enzymes that control nearly every aspect of cell function. They catalyse the reaction between ATP and a substrate (generally a protein) to give ADP and the phosphorylated substrate. A divalent ion is necessary for this reaction which in nature is Mg^{2+} since it is the most abundant dication in biological systems.

$$\text{Mg–ATP} + \text{Substrate–OH} \underset{\text{Phosphatase}}{\overset{\text{PTK}}{\rightleftharpoons}} \text{Substrate–}OPO_3^{2-} + M\text{g–ADP} + H^+$$

Protein phosphorylation can increase or decrease enzymatic activity. In fact, kinases control most of the biological processes such as metabolism, transcription, cell division, movement, apoptosis, immune response and even some functions of the nervous system.

Protein phosphatases are enzymes that catalyse the dephosphorylation of proteins so they make the whole phosphorylation/dephosphorylation process reversible.[1] The importance of protein phosphorylation has prompted researchers to develop several methodologies to measure kinase activity both in vitro and in vivo. Until now, the "gold standard" method is based on the transference of radioactive labelled [γ-^{32}ATP] to a substrate and use phosphocellulose paper P81 (anion exchange) to hold the substrate and determine the phosphorylation rate using a scintillation counter that measures the radioactivity of a sample.[2,3]

Animal Kinases are divided in two main groups depending on the residue that they phosphorylate. The first group are serine/threonine kinases which phosphorylate the hydroxyl group of serine or threonine residues. The second group, are tyrosine kinases (PTKs) which phosphorylate the tyrosine hydroxyl group. There are 90 PTKs 58 are receptor PTKs and the other 32 are non-receptor PTKs.[4,5]

Tyrosine kinases (TK) play a key role in cell growth, division, migration and survival signalling pathways.[6] Moreover, TK are also closely associated with several oncogenic processes which makes them relevant targets for the development of anticancer drugs.[7]

Each kinase has generally two conformations, the active and the inactive one. The activation of a kinase is the result of conformational change to render a catalytically competent structure. The kinase then toggles between the active and inactive conformation carrying out its catalytical function.

5.1.1 RTKs

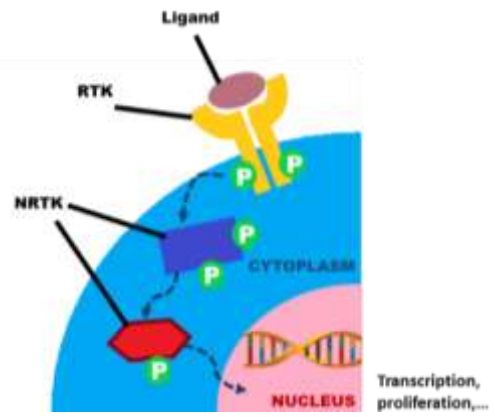
Receptor tyrosine kinases (RTKs) are transmembrane glycoproteins which are activated by the binding of their cognate ligand to the extracellular subunit. Then they transduce the corresponding signal to the intracellular media through the phosphorylation of the Tyr unit of a specific receptors (which can either be another protein or themselves) triggering a downstream signalling cascade. They possess an extracellular portion that allows the selective binging of the ligand, a transmembrane helix connecting the intra and extracellular media and a cytoplasmatic part that contains the kinase activity. The regulation of RTK is through different mechanisms but the most frequent ones are auto-phosphorylation and dimerization.

RTKs are known to activate several signalling pathways controlling proliferation, differentiation, migration and metabolic changes. IGFR (insulin growth receptor factor) is an examples of RTKs.

5.1.2 NRTKs

Non receptor tyrosine kinases are generally cytoplasmatic enzymes that are activated by RTKs and other receptors like GPCR or signalling from the immune system. Examples of NRTKs are the SRC family, JAKS or ABL kinases. They do not have the receptor feature but possess domains able to mediate interactions like protein-protein, protein-lipid or protein-DNA.

The activation of NRTKs is through the phosphorylation of their Tyr in the activation loop by another kinase, generally a RTKs previously activated through auto-phosphorylation.



Since the characterization of the crystal structure of myoglobin in 1958,[8] over 400 crystal structures of proteins have been resolved and uploaded to the Protein Data Bank. The human body has about 500 PTKs identified and so far the structure of over 50 of them has been determined. The structural properties of kinases (when available) are the most valuable information for the discovery of a new lead compound. The crystal structure of a protein provides a molecular framework for the elucidation of the

mechanism of action of different kinases and a starting point for the design of molecules able to regulate the activity of these enzymes, especially those involved in complex disease like cancer or neurodegenerative disorders. Even so, the catalytic domain of different kinases share a conserved structure challenging the development of small molecules with selective pharmacological properties.[6]

5.1.3 Small molecules as kinase inhibitors

Science needs to face several challenges to efficiently target kinases for the selective treatment of the mentioned diseases. There has to be a validation of the target kinase, they have to face drug resistance issues, address selectivity and spread its use to non-oncological areas where kinases are also an appropriate target. Even so, there are several small molecules approved by the FDA[4] and many others in clinical trials[9] that act as PTK inhibitors, most of them directed to neoplastic diseases.

The main strategy to inhibit PKs has been to target the ATP binding site and several of the approved drugs act this way. Nevertheless, the ATP binding site has a highly conserved structure which is common in several kinases[10] and this makes selectivity the main disadvantage of targeting ATP biding pockets.[11] A studied alternative is to design inhibitors targeting less conservative surrounding pockets which show more differences between kinases and allows to achieve selectivity.[6] Although most of the reported inhibitors target the inactive conformation of the kinase, there are some others that target the active conformation since it is thought to be less tolerant to develop resistance mutations. Some other inhibitors that do not target the ATP binding site are designed to target either the active site or non-catalytic domains which appears to be a promising strategy to improve selectivity since they have a less conserved structure.[10] In any case, all the PTKs modulation strategies studied so far directly target the kinase.

The uses of kinase inhibitors in oncology can be divided in those who target transcription and those focused in regulating the immune response, two major physiological processes controlled by kinases.

5.1.3.1 Kinase inhibitors targeting transcription

Extraordinary levels of transcription detected in many tumours are required to maintain their proliferation rates. Kinases associated to transcription are essential machinery to mediate oncogenic transformation, therefore they are potential targets for the inhibition

of tumour growth. Different inhibitors targeting kinases involved in transcription processes are already in clinical trials or approved for their use in oncology. [9]

5.1.3.2 Kinase inhibitors targeting immune response

One of the hallmarks in cancer is the immune system, that is why some researchers have focused their efforts in the design of drug molecules that stimulate the immune system for the recognition of malignant cells. Several kinase families and their activity have been linked to immune system regulation pathways. Compounds targeting those kinases have already been design to induce anticancer immunity and some of them are currently in clinical trials.[9]

So far the main application of kinase inhibitors is focused in oncology, but targeting kinases is also being explored successfully for the treatment of other pathologies like inflammatory diseases, degenerative diseases and even infectious diseases.

The big challenges in the design of new kinase inhibitors remain the selectivity and the rapid development of chemo-resistance. Hopefully, with the aid of new preclinical validation targets, computation technology and biochemical experiments, this field has a great potential to rapidly grow and bring improved kinase inhibitors to clinical applications.

5.1.4 SRC, IGFIR and WASP

As it will be noted in the experimental section of this chapter, the kinases catalytic domains used for the kinase assays were from SRC ACK1 and IGF1R kinases. The kinases used were selected in terms of availability and previous knowledge of their behaviour towards several substrates.

This section attempts to give a very general overview of the used kinases, their structure and functionality in human cells.

5.1.4.1 SRC kinases

Among the most widely studied NRTKs there is the SRC family which comprises 9 members of kinases: c-SRC, YES, FYN, LYN, LCK, HCK, FGR, BIK and YRK, all of them with similar structure.[12] SRC kinase have a key role in mediating signal transduction pathways via interactions with other proteins and enzymes. They interact with several enzymes from the plasma membrane stablishing a bidirectional flow of information with this region.[12]

SRC are controlled by integrin receptors, GPCRs and some other types or receptors and they participate in cell migration and mobility, cell survival, proliferation and gene

expression. Besides they have also been found to participate in bone formation. An aberrant activation of this family of kinases is responsible for tumour progression in many cancer types where they have been demonstrated to control apoptosis, cell proliferation, adhesion, migration, cell invasion angiogenesis and metabolism.[13]

SRC kinases have been known for many years and yet a lot of research is still being conducted on these kinases. All the kinases in the SRC family share a common structure which is composed of 7 parts: Unique domain, SH4, SH3, SH2, SH1 (Catalytic domain), SH2-SH3 linker, and the C-terminal regulatory region. The transition from the inactive to the active state is accompanied by conformational changes.[12]

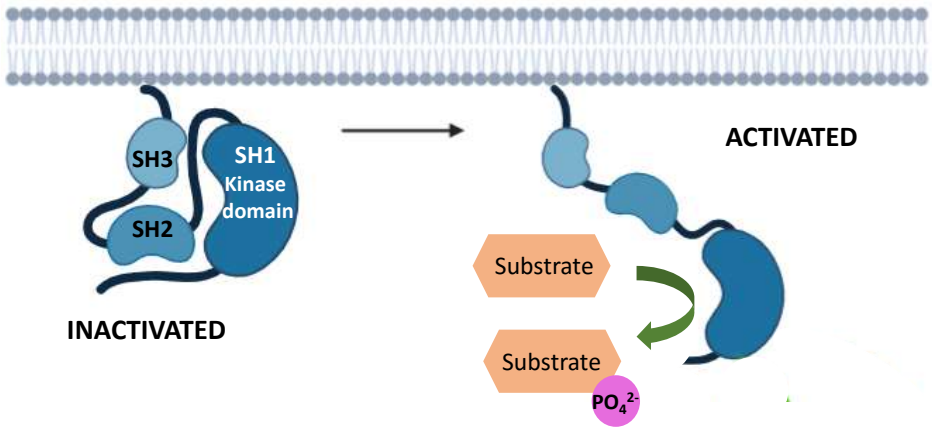


Figure 5.1. SRC kinase schematic representation in its inactivated and activated form.

SRC kinases are becoming more and more important in medicine because of its implication in many cancer processes and they are becoming a relevant target for the design of anticancer drugs.[14] For instance, an over-activation or overexpression of c-SRC- has been found in numerous human tumours like lung, breast, pancreatic, colorectal and prostatic ones. Despite major focus has been catered in the implications of kinases with cancer, kinase dysregulation is also linked to immunological, inflammatory, degenerative, metabolic, cardiovascular and infectious diseases. Despite PTKs being the subject of much systematic investigation, the complexity of the many signalling pathways in which they are involved are making the transition from biological studies to the clinic a hard and time-consuming work.

5.1.4.2 IGFR

Insulin growth factor receptor 1 (IGFR1) together with insulin, regulates the physiology of the organism as a whole system but also at a cellular level. However, its dysregulation plays a key role in the transforming ability of several oncogenes.[17] It has been demonstrated to stimulate tumour proliferation and its genetic modification to reduce IGFR signalling in animal models has shown a reduction of tumour growth.[17] This has contributed to the study of therapies based on the regulation of IGFR1 but little success has been achieved in the clinical trials with the dozens of compounds tested.

IGFR1 is a trans-membrane receptor tyrosine kinase that belongs to the insulin receptor family. It is formed by an extracellular ligand binding domain, a trans-membrane region and the intracellular part which includes the catalytic domain. The activation of the IGF1R occurs when the polypeptide (insulin, IGF1 ir IGF2) binds the ligand binding domain and triggers a conformation change in the receptor.

IGF1R together with the insulin receptor act as part of a signalling cascade by phosphorylating members of the insulin receptor substrate family of proteins. This leads to the activation of other enzymes and triggers their functions which are the inhibition of gluconeogenesis and the activation of glycogen storage in the liver or the stimulation of apoptosis and the inhibition of proliferation in epithelial cells. IGFR are widely spread in both neoplasic and normal tissues with low differences but there is evidence that variations between the insulin secretion rates in different individuals influences cancer prognosis since insulin can increase cancer cells proliferation.[18]

The main IGFR production site is the liver and the growth hormone is the principal stimulus for IGF1. Insulin is on the other hand produced by pancreatic β-cells and spread through circulation.

The first clinical application of insulin was realized in the twenties (for the control of diabetes) but it has been in the past 30 years that the relevance of IGF1R has substantially increased as it was proposed as a target for human cancers.

Several strategies including small molecules for receptor inhibition, anti-receptor antibodies (designed to spare the insulin interfering with ligand binding to the IGFR1) and anti-ligand antibodies have been proposed to target IGFR (Figure 5.2), some of which even reached clinical trials for their anti-neoplasic activity in vivo.[17]

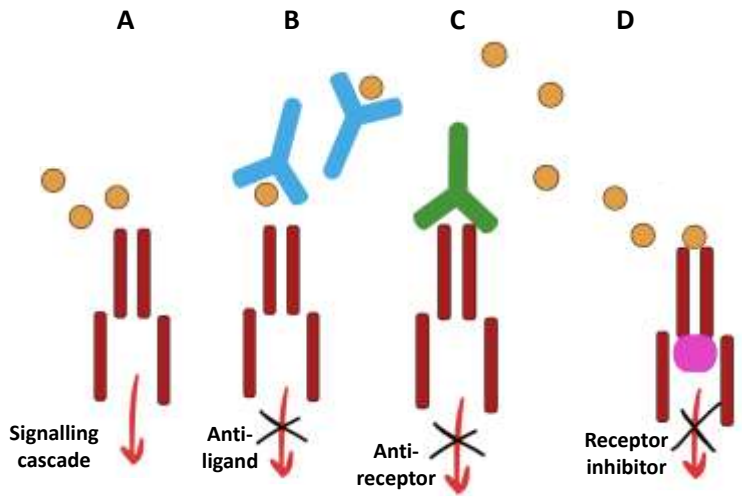


Figure 5.2. Anti-ligand, anti-receptor and receptor tyrosine kinase inhibition strategies. (A) Binding of insulin to IGFR Inducing the corresponding signalling cascade. (B) Ligand target strategy: attempt to reduce the ligand concentration with specific ligand antibodies. (C) Antireceptor strategy: use of IGF1R antibody to block the binding of tyrosine. (D) Receptor inhibition strategy: a compound that binds the IGF1R blocking its function even in the presence of insulin. (image adapted from reference [17]).

Although some phase-II clinical trials showed activity with low toxicity they also showed a lack of efficacy in phase-III trials. These results have arisen the debate about the convenience of investigating IGFR as a target. In any case, more research is necessary to clarify the role of IGF1R and the molecular downstream to better design inhibitors with clinical applications.

WASP peptide

The Wiskott Aldrich syndrome protein (WASP) is a Cdc42 effector that plays an important role in the formation of new filaments[15] and is one of the main substrates of ACK1 kinase. Wasp plays a significant role in integrating and transforming signals from critical receptors on the cell surface to actin remodelling. Actin cytoskeleton is a complex network of actin filaments constantly being polymerized/depolymerized in response to a variety of external signals. What´s more, the regulation of the actin cytoskeletons is crucial for transduction, cell shape changes, mobility and polarity in response to the appropriate stimuli.[16]

The Wasp fragment containing the tyrosine whose phosphorylation is related to the activation of Wasp has the following aa sequence: KVIYDFIEKKG.

5.2 OBJECTIVES

As an alternative to directly inhibit the kinase to modulate its catalytic activity, we propose to target the specific substrate that triggers the kinase function that needs to be controlled (Figure 5.3.A).

This strategy was previously presented by the supramolecular chemistry group in a previous contribution.[19] We reported on the use of certain pseudopeptidic macrobicycles (**CyLys** and **CyOrn** from chapter 4) able to inhibit the TK-promoted phosphorylation, as shown by a TK commercial kit that employs the non-natural universal polyE_4Y substrate. Later on, cationic pillarenes were also used in a very similar fashion.[20] Although these preliminary results demonstrated the proof of concept, the important issue of substrate selectivity still remains unaddressed. Only very recently, Zhao and co-workers reported the use of molecularly imprinted nanoparticles able to inhibit serine kinases by substrate binding, though the method was less effective in Tyr protection.[21]

The previous successful results and the relevance of this almost unexplored research strategy lead us to carefully assay the TK inhibition activity of macrocyclic compounds reported in chapter 4 towards different substrate-kinase systems.

To do so we stablished a collaboration with the Miller Lab in Stony Brook University, where I had the opportunity to carry out a three-months research stay, during which I was able to perform several in-vitro kinase inhibition experiments.

The specific objectives of this chapter were:

- To study the TK inhibition capacity of **CyLys, CyOrn, CySer, CyThr, CyAsp, CyGlu** and **CyHis** towards different kinase-substrate systems.
- To study the interactions between the cages and the substrates by fluorescence and NMR.
- To conduct several control experiments to corroborate the action mechanism proposed for the TK inhibition.
- To experimentally find a correlation between the chemical interactions measured through spectroscopic techniques and the kinase inhibition potency of the different cages towards several substrates.

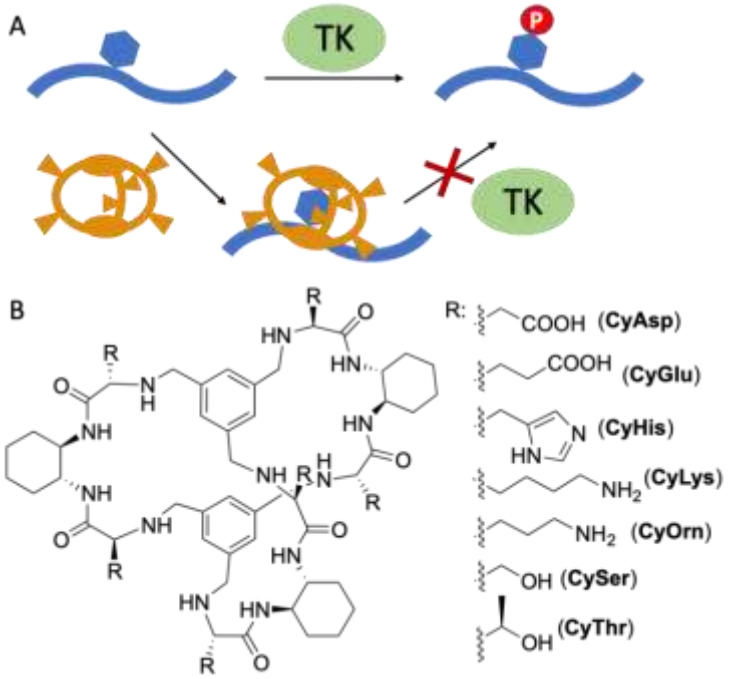


Figure 5.3. (A) Schematic representation of the proposed mechanism for the substrate-recognition TK modulation. (B) Chemical structures of the pseudopeptidic cages used in this work.

5.3 RESULTS AND DISCUSSION

5.3.1 Kinase inhibition assays

The **CySer/CyThr**[22] and **CyLys/CyOrn**[19] cages used in this Thesis were prepared as previously described by the supramolecular chemistry group, whereas **CyAsp**, **CyGlu** and **CyHis** were prepared as described in chapter 4.

As the benchmark TK, we chose Src for its wide substrate acceptance and biological relevance in cell physiology and cancer.[5,23–27] In order to study the effect of CyAsp, **CyGlu, CyHis**, **CyLys**, **CyOrn**, **CySer** and **CyThr** on Src kinase activity, we used the phosphocellulose paper binding assay (see experimental section for details).[28] Initially we tested polyE_4Y as the reference nonspecific substrate, which showed that all the cages inhibit the kinase activity of Src to some degree (Figure 5.5A). However, we observed significant differences between the receptors, with the Asp, Lys and Orn cages being the most active ones. The activity observed with **CyLys** and **CyOrn** is in agreement with our previous results using a TK commercial kit,[19] and can be explained by the electrostatic attraction between their cationic side chains and the negatively charged polypeptide sequence that mediate substrate recognition. The performance of **CyAsp** was counterintuitive following the same rationale, and must be due to other factors (see below).

For a deeper study of the substrate selectivity, we tested a series of synthetic peptides: the Src peptide substrate identified by Cantley and co-workers (Src-PS),[29,30] a sequence derived from the Wiskott-Aldrich syndrome protein (WASP)[15,31] and the Ile5Val mutant of the hormone Angiotensin II (V5Ang-II)[32] (Figure 5.5B). They differ in several structural aspects, such as their length, hydrophobicity and charge distribution along the sequence. Furthermore, the theoretical prediction[33–35] of their corresponding secondary structures also rendered important differences in the most probable conformations and in the spatial distribution of the polar/non-polar residues. These predictions were carried out using PEPFOLD 3.5. The most populated cluster for each peptide is represented in Figure 5.4 while the rest are represented in the experimental section. In these representations, the colour code in the backbone is: red (acid), blue (basic) and green (hydrophobic) residues. A Gly6Ala variant of Src-PS (A6Src-PS) was also considered to check the effect of a single substitution close to the phosphorylation site with minimal structural perturbation.

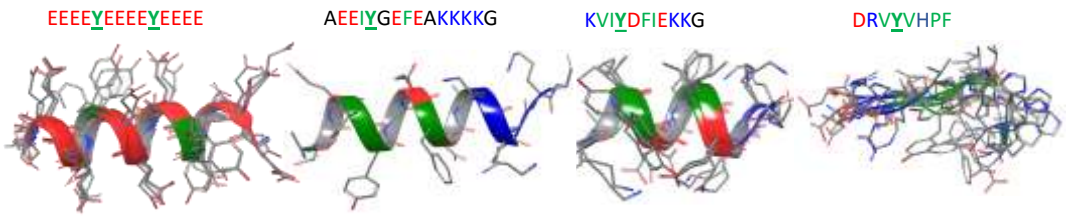


Figure 5.4. Sequences and overlay of the most probable secondary structures for peptides E_4YM_4, Src-PS, WASP and V5Ang-II (left to right).

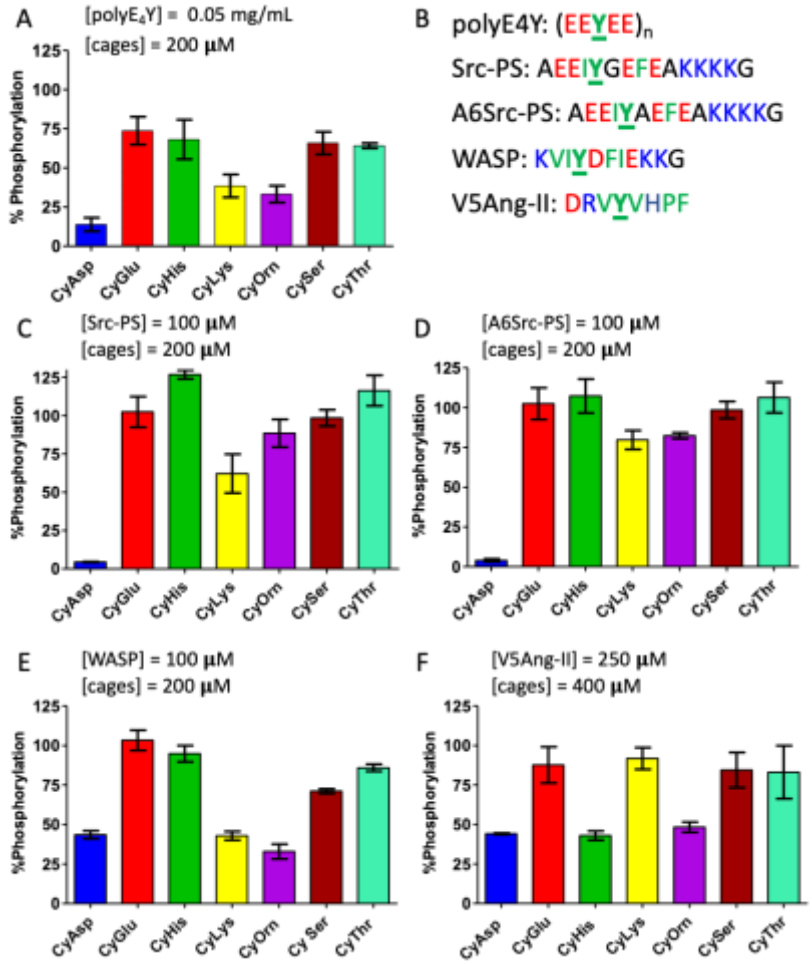


Figure 5.5. (A) Effect of the cages on the Src-promoted Tyr-phosphorylation of the polyE$_4$Y substrate. (B) Peptide sequences used in this work. (C-F) Effect of the cages on the Src-promoted Tyr-phosphorylation of the substrates: (C) Src-PS, (D) A6Src-PS, (E) WASP and (F) V5Ang-II. In all the cases, the plots show the percent phosphorylation relative to the corresponding control reactions in the absence of a cage.

The protection of the Tyr residues from the Src-promoted phosphorylation depends on both the cage structure and the peptide sequence (Figure 5.5C-F). As a general trend, **CyAsp** is the most efficient phosphorylation inhibitor, leading to almost complete inhibition in the case of Src-PS and A6Src-PS (Figure 5.5C,D, blue bar) and to > 50 % inhibition with the other peptide substrates (Figure 5.5E,F). The cages made from Lys, Orn and His also show interesting activities with marked differences between substrates. The comparison between **CyLys** and **CyOrn** is especially noteworthy, as the relative inhibitory activities vary for the different peptides (compare yellow and purple bars in Figure 5.5C-D). Thus, **CyLys** inhibits **Src-PS** phosphorylation more efficiently than **CyOrn**, while for Va5Ang-II the trend is significantly reversed **(WASP** lies in between). Even more remarkably, **CyHis** displays a noticeable activity only for the **V5Ang-II** substrate (Figure 5.5 F, green bar), being comparable with the best inhibitors in this case, **CyAsp** and **CyOrn** (blue and purple bars, respectively). Overall, the observed differences suggest that the inhibitory activity occurs by the selective encapsulation of the Tyr side chain of the peptide substrates, which depends on the accessory interactions between the side chains of the different cages and the amino acid side chains surrounding the Tyr residue.

5.3.2 Interactions measurements by fluorescence

Since the cage-substrate interaction is a key factor in the inhibitory activity, we decided to study the supramolecular structures for selected cages by exploiting the changes in fluorescence emission of the Tyr side chain upon inclusion inside the cage cavity[19] (Table 5.1).

HypSpec[36,37] software was used to fit the fluorescence titration data to every proposed interaction model. This software performs the global fitting of the whole emission band (or a selected range) for each titration point, to satisfy the interaction model in each case. This model can include several association constants between 1, 2 or more components (**¡Error! No se encuentra el origen de la referencia.**). In these particular cases the emission fluorescence spectra of the peptides were first measured in a wide concentration range and, for those showing non - linear emission vs. concentration plot, a simple dimerization model was used to extract K_{dim}. Following, the titration of sample with a fixed concentration of the peptides with increasing amount of the cages was carried out. For the peptides prone to dimerize, K_{dim} was included in the fitting process of the cage-peptide titration as a constant value, and the corresponding cage-peptide stability constant $\beta(AB_n)$ was obtained

from the fitting. As a default, the simplest 1 : 1 binding mode was tested and, if necessary, more complex 1 : 2 binding mode was used (**CyHis**). In the specific case of **CyAsp**, the formation of a Mg(II) complex was considered. To that, the **CyAsp**-Mg(II) interaction was measured by ^{1}H NMR titration and the corresponding K_{Mg} was obtained by fitting the data using HypNMR[38] (Figure 5.51). Accordingly, K_{Mg} was included as a constant value in the fitting of the fluorescence emission titrations of all the peptides with **CyAsp**. Despite this method performs a global fitting of the whole emission band, Figure 5.6, Figure 5.8 and Figure 5.15 only show the fitting of a single wavelength for simplicity.

The first important observation is the non-linear concentration dependence of the fluorescence emission of the peptides in solution. Fitting the self-dilution titration data rendered a relatively strong dimerization constant for the three studied peptides (Table 5.1, entry 1, Figure 5.29Figure 5.32), which implies their effective self-assembly at μM concentrations, such as those used in the phosphorylation assays. These dimerization processes were taken into account in all the equilibria shown next.

Regarding the cage-peptide binding, the fluorescence titrations show important differences between the systems. Thus, the fitting of the fluorescence titration data with **CyAsp** (entry 2) requires considering the formation of [**CyAsp**-Mg-peptide] ternary complexes (Figure 5.6). The **CyAsp**-Mg interaction was independently confirmed by ^{1}H NMR titration experiments (Figure 5.51Figure 5.52). Besides, the corresponding fluorescence titration experiments in the absence of Mg salt only showed dynamic quenching and a weaker binding (Figure 5.43), confirming the key role of the Mg(II) ion in the peptide-cage interaction. For a fair comparison, the corresponding equilibrium constants for the binding of the [**CyAsp**-Mg] species to the peptides are included in entry 4 of Table 4.1. In all the cases this interaction is very strong ($K_1^{Mg} \approx$ 2-4 x 10^4 M^{-1}), explaining the efficient inhibition by **CyAsp** of the Src-catalysed phosphorylation reactions, which were performed in a relatively high concentration of Mg ions.

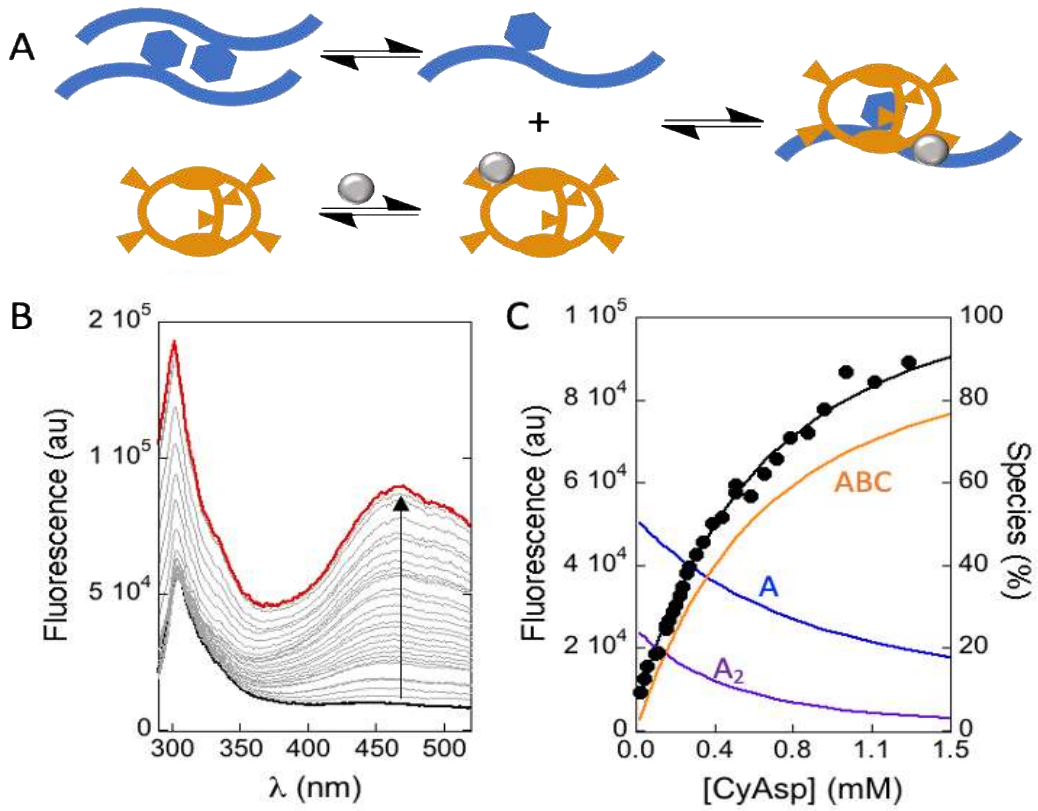


Figure 5.6. (A) Binding mode of peptides to the CyAsp-Mg species. The peptide is shown in blue, the cage is orange and the Mg(II) ion is depicted as a silver sphere (B) Fluorescence emission spectra of Src-PS peptide (black, 10 μM, 30 mM Tris buffer at pH 7.5, 20 mM $MgCl_2$) upon increasing concentrations of CyAsp (grey to red). (C) Fluorescence at 467 nm (experimental values as black symbols, fitting line) and species concentrations: free Src-PS (blue), dimer (purple) and [CyAsp-Mg-Src-PS] complex (orange).

Table 5.1. Logarithm of the equilibrium constants (Logβ, Log$K$) for the complexes formed by peptides (A), cages (B) and Mg(II) ion (C), obtained by fluorescence titrations upon excitation at 276 nm (30 mM Tris buffer at pH 7.5 and 20 mM $MgCl_2$). Values in parentheses correspond to the standard deviation on the last significant figure and the corresponding reaction equation defining each equilibrium is included in each case.

entry	cage	reaction	property	Src-PS	WASP	V5Ang-II
1	Peptide[a]	2A ⇆ A_2	Logβ(dimer)	4.97(3)	5.20(3)	3.56(2)
2	CyAsp	A+B+C⇆ABC	Logβ(ABC)	5.24(3)	5.640(3)	5.30(1)
3	CyAsp	B+C⇆BC	LogK_{Mg}[b]	0.95(3)[b]	0.95(3)[b]	0.95(3)[b]
4	CyAsp	A+BC⇆ABC	LogK_1^{Mg}	4.29	4.69	4.35
5	CyGlu	A+B⇆AB	Logβ(AB) = LogK_1	3.258(6)	3.306(3)	2.844(6)
6	CyHis	A+2B⇆AB_2	Logβ(AB_2)	6.45(1)	6.61(2)	6.53(1)
7	CyHis	A+B⇆AB	Logβ(AB) = LogK_1	3.697(6)	3.89(2)	3.44(1)
8	CyHis	AB+B⇆AB_2	LogK_2	2.75	2.72	3.09
9	CyHis	n.a.	$\alpha = 4K_2/K_1$[c]	0.45	0.27	1.79
10	CyLys	A+B⇆AB	Logβ(AB) = LogK_1	3.82(3)	3.447(6)	2.80(1)
11	CyOrn	A+B⇆AB	Logβ(AB) = LogK_1	3.489(4)	3.537(6)	3.419(3)

[a]Corresponding to the dimerization of the peptide. [b]Obtained by ^{1}H NMR titration experiments. [c]Cooperativity parameter as defined in ref.[39]

Rather surprisingly, the binding mode with **CyGlu** was found to be simpler (entry 5). In this case, the Mg ion does not directly intervene in the formation of the supramolecular complexes, which was additionally confirmed by control titration experiments at low Mg concentration (compare Figure 5.44 and Figure 5.45). The different Mg coordination trends of Asp/Glu amino acids explains the behaviour observed for **CyAsp** and **CyGlu**.[40,41] Thus, Mg(II) could form a 6-membered ring chelate with the amino and carboxylate groups in the case of **CyAsp**, which is not possible with **CyGlu**. Remarkably, **CyGlu** rendered much lower cage-peptide binding constants (at least one order of magnitude, compare entries 4 and 5 in Table 5.1), which is in very good agreement with its poorer phosphorylation inhibition abilities.

The binding data obtained with **CyLys** and **CyOrn** are especially significant (entries 10 and 11 in Table 5.1). Thus, **CyLys** binds to **Src-PS** more strongly than **CyOrn**, with the corresponding inhibition showing the same trend (**CyLys** > **CyOrn**, Figure 5.5C). However, the binding of **WASP** peptide to **CyOrn** is slightly stronger than to **CyLys**,

with the same order in the inhibition experiments (**CyOrn** > **CyLys**, Figure 5.5E). Finally, the recognition of V5Ang-II peptide is fivefold more efficient with **CyOrn** than with **CyLys**, which is clearly reflected in the phosphorylation experiments, where **CyOrn** is a more potent inhibitor (**Figure 5.5**F). Thus, for these two structurally similar cages, **CyLys** and **CyOrn**, the binding and the inhibitory efficiencies are strongly correlated. The higher basicity of **CyLys** renders a more positively charged host than **CyOrn** at neutral pH, on average.[19] On the other hand, the number of negatively charged residues in the studied substrates varies in the series: Src-PS > WASP > V5Ang-II. Thus, the attractive secondary electrostatic interactions favour the complexes with **CyLys** also in the same trend, giving a reasonable explanation to both the binding and inhibition results.

The case of **CyHis** is more difficult to rationalize, because all the tested peptides lead to complexes with 1 : 1 and 1 : 2 peptide : **CyHis** stoichiometry (entry 6). Cages of this type can recognize other aromatic residues, although the Phe binding is usually weaker.[23] Our titration data shows that **CyHis** is less selective for Tyr than the other receptors, binding to both aromatic residues (Tyr and Phe) present in the three peptidic substrates. This equilibrium scheme requires a more careful analysis: Table 1 shows the corresponding stepwise constants for the individual binding events (entries 7 and 8). Moreover, for the 1 : 1 supramolecular species, there are two possible microspecies (A and B in Figure 5.8) leading to encapsulation of either Tyr or Phe residues. This is very relevant, since the Phe-bound species would be less efficient in the Tyr protection. In such complex processes, the evaluation of the cooperativity parameter[39] is highly convenient (α in entry 9). Thus, the recognition of Src-PS and WASP by **CyHis** shows the usual negative cooperativity ($\alpha < 1$) meaning that the first binding event hinders the second one. This can be reasonably explained by electrostatic and steric repulsions between consecutive binding of two **CyHis** to close residues in the peptide sequence. Thus the peptide with closer Tyr/Phe residues (WASP) also shows lower α value, reflecting this hindering effect. However, for V5Ang-II-CyHis, an uncommon positive cooperativity ($\alpha > 1$) implies that the first **CyHis** binding favours the formation of the 1 : 2 complex. This remarkable difference could account for the different inhibitory abilities in the phosphorylation assays. In most cases (i.e. Src-PS and WASP), **CyHis** binding to Phe sequesters the cage and reduces Tyr protection, whereas **CyHis** inhibits the phosphorylation of the V5Ang-II

substrate either through a Tyr encapsulation in the 1 : 1 complex or by a cooperative Tyr encapsulation following the initial Phe binding.

The next section contains experimental results designed to better understand CyHis behabiour.

5.3.3 Structural studes of CyHis with V5Ang-II

In order to better understand the singular behaviour of **CyHis** towards **V5Ang-II**, we prepared samples that maximize the 1 : 1 and 1 : 2 V5Ang-II : **CyHis** supramolecular complexes to compare their ^{1}H NMR spectra with that of the peptide alone. To properly determine the concentrations of V5Ang-II : **CyHis** in the samples for the NMR experiments, we performed a simulation of species using HySS2009[42] at different conditions and selected the two that fitted out requirements (Figure 5.7):

The chemical shift perturbations of the peptide ^{1}H NMR signals (Figure 5.8C, Figure 5.9) show upfield shifts of the Asp1, Tyr4 and His6 signals in the 1 : 1 complex (grey bars in Figure 5.8C), while the aromatic protons of Phe8 are shielded mainly in the 1 : 2 complex (black bars in Figure 5.8B). These results strongly suggest the Tyr4 residue as the site for the first binding event. The spatial proximity between Tyr4 and His6 in Ang-II-type peptides has been reported by several authors, affecting the properties of the Tyr side chain.[43] The Tyr inclusion in the 1 : 1 complex with **CyHis** perturbs the His6 signals (Figure 5.8C), suggesting an interaction between the cage and peptide imidazole rings,[44] which further stabilize the supramolecular complex. All these observations are in agreement with the particularly efficient **CyHis** inhibition of V5Ang-II phosphorylation.

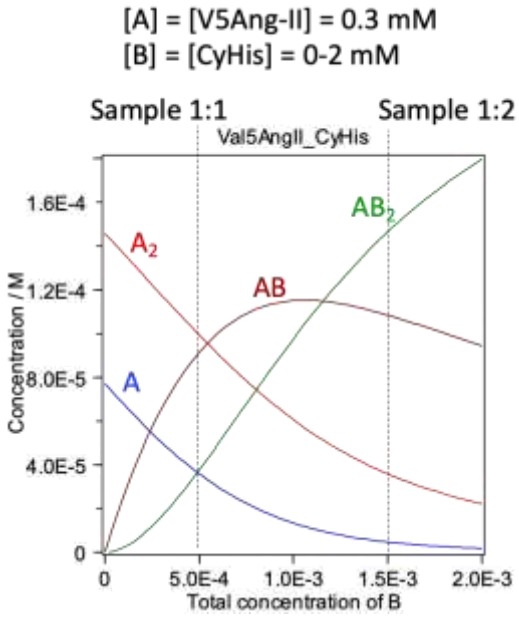


Figure 5.7. Simulation of the species distribution in equilibrium for the V5Ang-II-CyHis system. The prepared samples for the corresponding 1:1 and 1:2 complexes are shown in dashed lines. Samples containing a higher concentration of **CyHis** led to partial precipitation and thus showed unsuitable for NMR analysis.

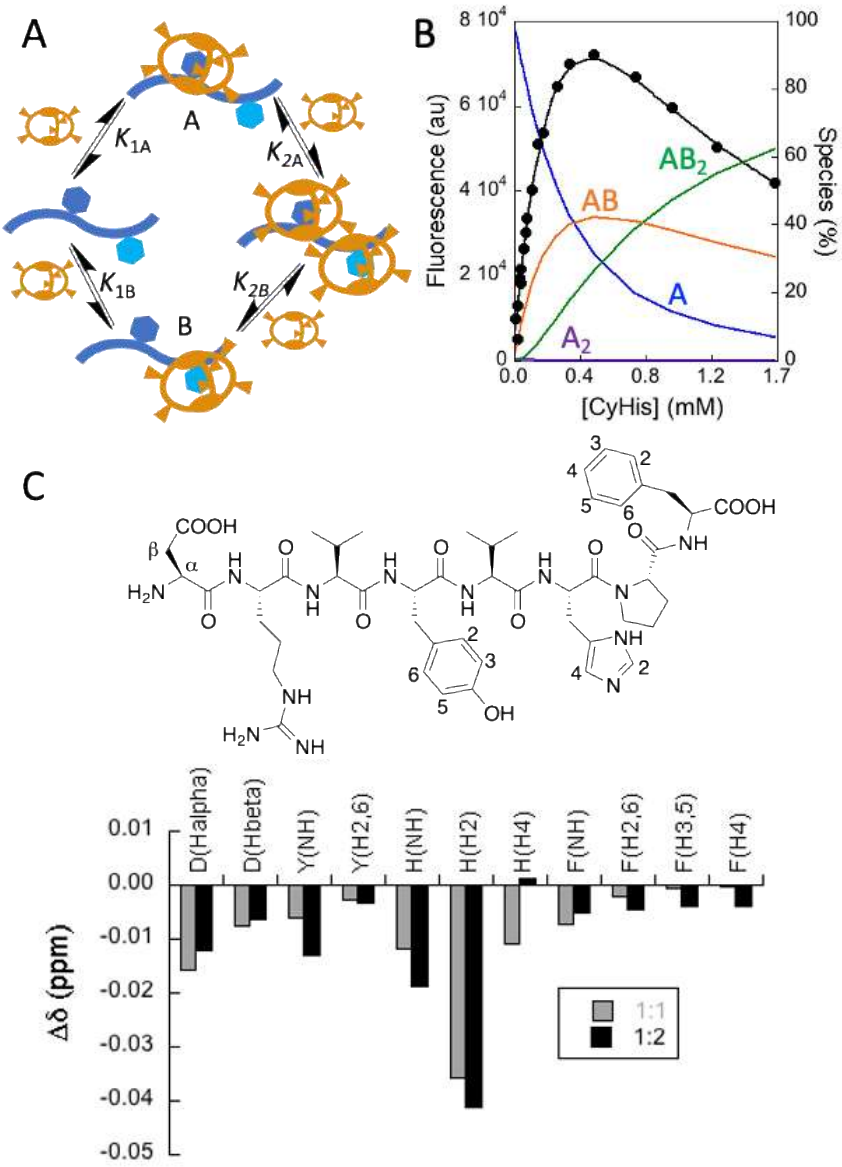


Figure 5.8. (A) Schematic representation of the binding mode for **CyHis** receptor (orange), implicating 1 : 1 and 1 : 2 supramolecular complexes The peptide is shown in blue. Dimerization of the peptide is omitted for simplicity. (B) Fluorescence (442 nm) titration of V5Ang-II (black, 1 μM, 30 mM Tris buffer at pH 7.5, 20 mM $MgCl_2$) upon addition of **CyHis**: observed fluorescence (black symbols), fitting (line); equilibrium species: free V5Ang-II (blue), dimer (purple), 1 : 1 (orange) and 1 : 2 (green) complexes. (C) Plot of the chemical shift perturbation of the ^{1}H NMR signals of V5Ang-II upon binding to **CyHis** at the conditions for the main formation of the 1 : 1 (grey) and 1 : 2 (black) peptide : cage complexes.

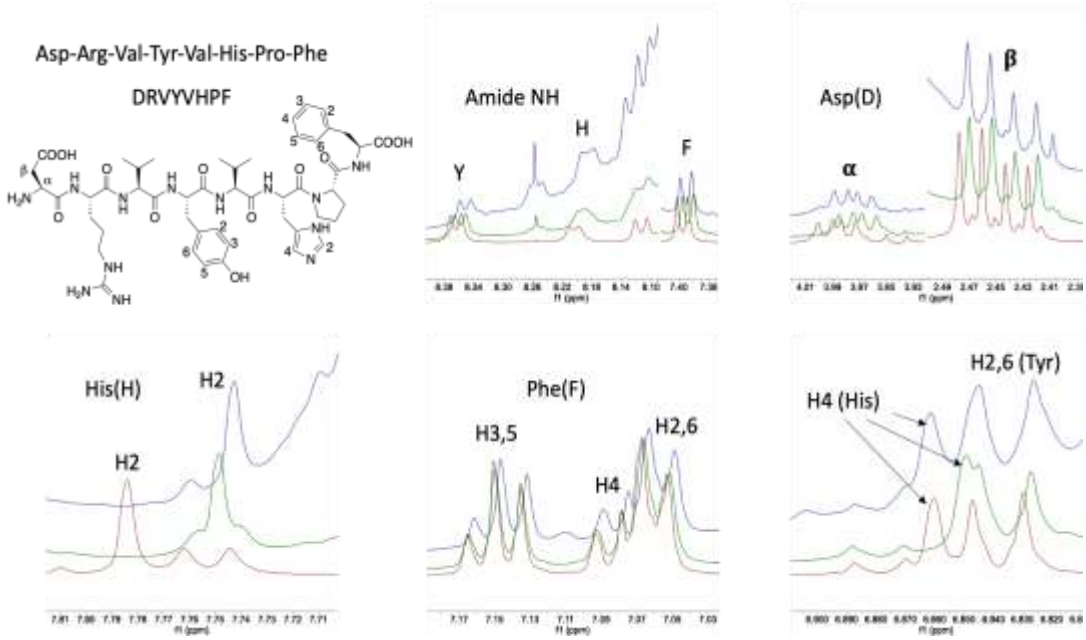


Figure 5.9. Selected regions of the overlayed ^{1}H NMR spectra (90% H_2O/10% D_2O, 30 mM HEPES-d18 with 5 mM $MgCl_2$ buffer, 288 K) for the V5Ang-II peptide (maroon), the 1:1 sample (green) and the 1:2 sample (blue). Assigned protons that are perturbed by the formation of the complex are indicated in each figure.

A structural proposal for the V5Ang-II-CyHis complex is shown in Figure 5.10 (also see figures A4.36-37) obtained after a Monte Carlo conformational search using Macromodel and the OPLS3e force-field in implicit water. In agreement with the changes observed in the fluorescence emission spectra, the Tyr4 aromatic side chain (green CPK in Figure 4D) fits inside the cage cavity, favoring the interaction of the surrounding amino acids with the receptor. In this minimum, the *i*Pr groups of Val3 and Val5 lay on top of two cyclohexane rings of **CyHis**, possibly establishing hydrophobic contacts. On the other hand, since Tyr4 and His6 are correlated in Ang-II-type peptides,[43] the presence of Tyr4 within the host cavity causes His6 to approach to one of the imidazole rings of CyHis (highlighted CPK in Figure 5.10) in an edge-to-face disposition that explains the shielding observed in the His6 ^{1}H NMR signals, especially for H2 (Figure 4C). The complex is additionally stabilized by six host-guest H-bonds, mainly implicating Tyr4 from the guest and the imidazole rings from the host. For instance, two cage His residues bind the backbone amide carbonyls of Val3 and Tyr4 from V5Ang-II. This model illustrates how the Tyr4 inclusion within the cage cavity

is enforced by additional interactions between side chains, thus explaining why **CyHis** is particularly efficient in the inhibition of V5Ang-II phosphorylation.

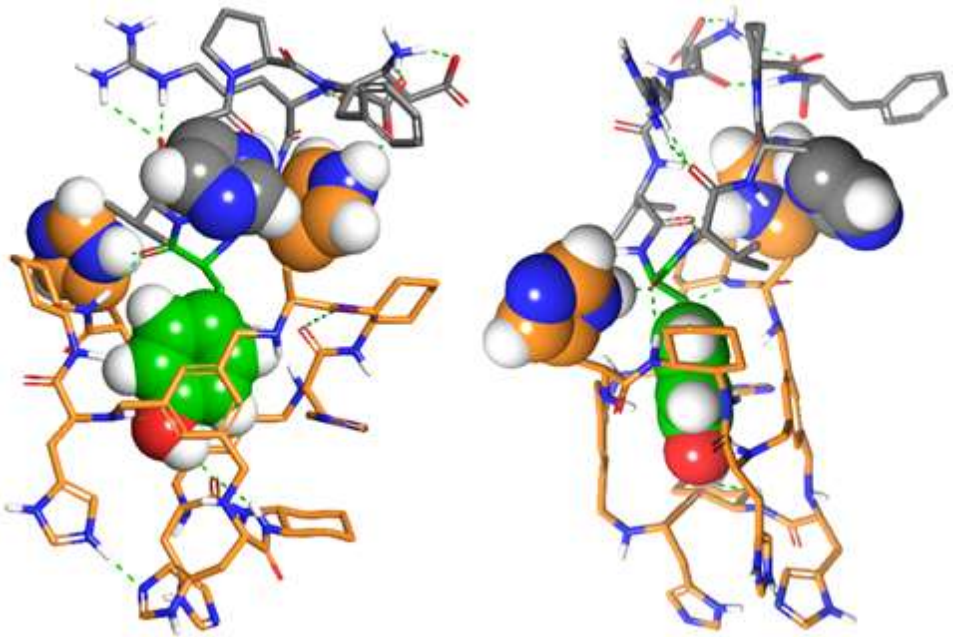


Figure 5.10. Two views of the proposed [CyHis-V5Ang-II] supramolecular complex (Macro Model OPLS3e minimized structure). The cage C-atoms are shown in orange, and the Tyr4 V5Ang-II C-atoms are shown in green, most non-polar H-atoms are omitted for clarity and selected aromatic rings are displayed as CPK. H-bonds are represented as green dashed lines.

5.3.4 Src inhibition vs K_{ass}

For the peptide Src-PS, the plot of the kinase inhibition of Src in the presence of the different cages (same concentration of peptide (100 µM) and cage (200 µM)) vs the binding constants (Cage-peptide) in a logarithmic scale renders a linear response (**Figure 5.11**). This result accounts for the good correlation between the inhibition and the binding reported. This linear trend is less clear for the other peptides mainly because the affinity of the Src towards them is lower hence the Src concentration in the inhibition experiments was higher. The affinity of the Src for the peptide is in these cases more similar to the affinity of the cage for the peptide leading to a competition for the substrate that deviates the Inhibition/K_{ass} trend form linearity.

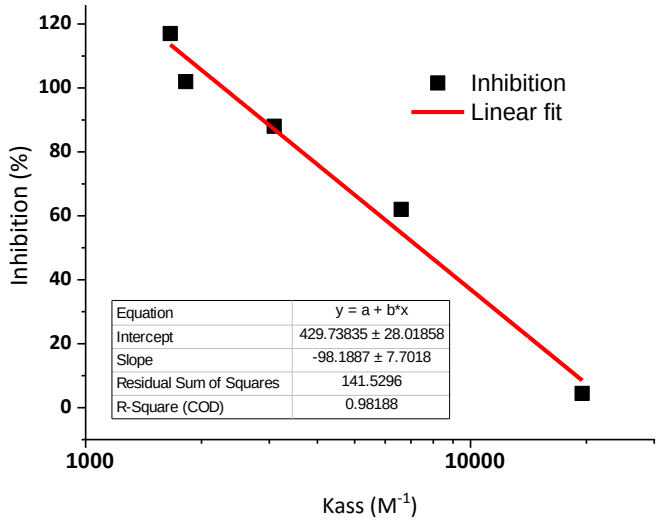


Figure 5.11. Inhibition of Src-PS phosphorylation by Src kinase vs the binding (K_{ass}) of the different cages towards this peptide.

Overall, although the direct comparison between substrates must be done carefully due to the intrinsic dimerization properties of the sequences and distinct Src affinities, our cage-peptide binding data successfully explain the observed trends in the phosphorylation assays.

5.3.5 Kinase inhibition assays in the presence of BSA

With the aim of confirming the proposed inhibition mechanism, we performed additional phosphorylation experiments. First, we considered the potential competition of a protein displaying several solvent-exposed Tyr residues. We examined the effect of bovine serum albumin (BSA), which is commonly included in phosphorylation assays to prevent non-specific hydrophobic interactions. BSA contains 20 Tyr residues, of which at least 12 are located at the protein surface or in an accessible pocket (PDB ID: 4F5S). Selected kinase inhibition assays were conducted in the presence of BSA (5 mg/mL). Under these conditions, we found that the presence of BSA reduces the inhibitory ability of the cages. Different concentrations of the corresponding cages were tested in order to observe this effect. The inhibition is lower even at higher concentrations of the cage compound but the trend between the different cages is generally preserved (Figure 5.12- Figure 5.13), suggesting a competitive effect of the Tyr residues of BSA through partial complexation of the cages. The corresponding BSA-cage interaction was also confirmed by additional NMR experiments (Figure 5.21). Nevertheless, the observation of inhibition activities even

in the presence of competing BSA implies a stronger binding of the cages to the target peptides.

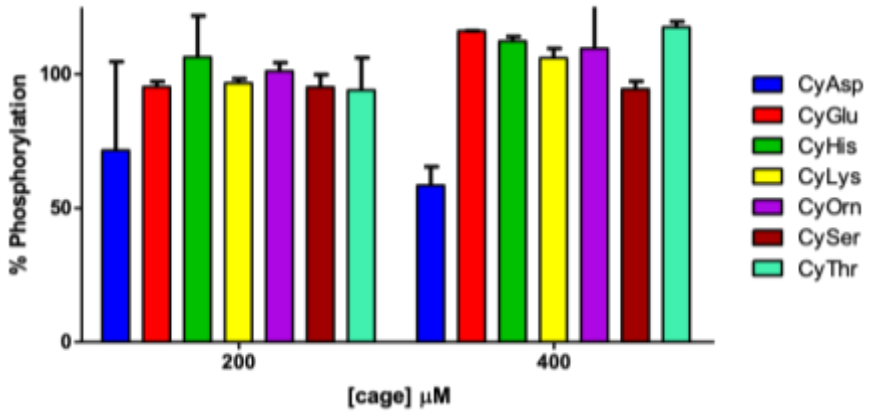


Figure 5.12. Effect of the cages on the Src-promoted Tyr-phosphorylation of Src-PS (100 μM) peptide in the presence of BSA.

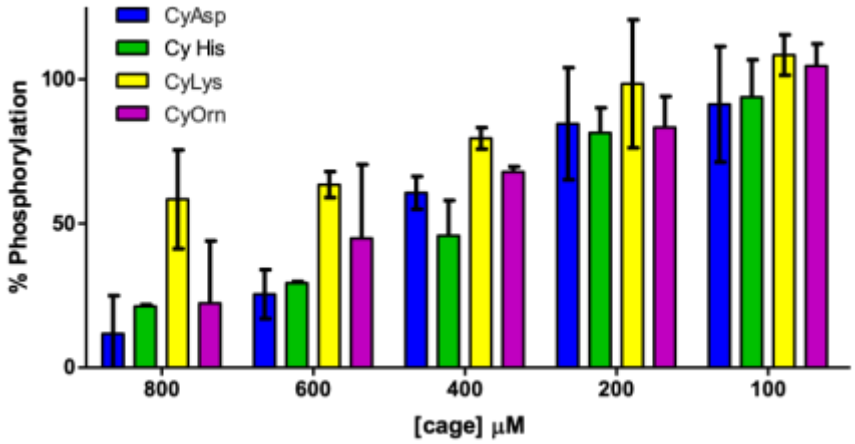


Figure 5.13. Effect of the cages on the Src-promoted Tyr-phosphorylation of WASP (100 μM) peptide in the presence of BSA.

5.3.6 Src activity modulation at different substrate concentrations

We also built a dose-response curve for the **CyAsp**/**Src-PS** pair, at three different concentrations of the peptide substrate (Figure 5.14). Remarkably, the calculated IC_{50} of **CyAsp** is proportional to the concentration of **Src-PS** at a constant concentration of Src, which means that the cage inhibits the phosphorylation through the specific

supramolecular peptide recognition. These additional results further support the inhibition mechanism depicted in Figure 5.3.

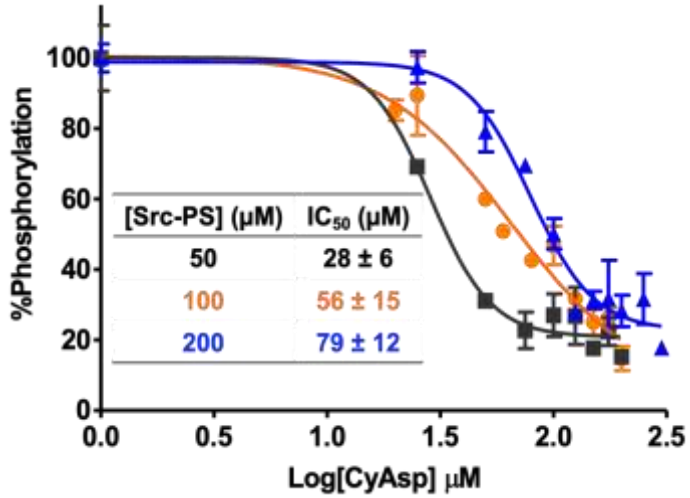


Figure 5.14. Dose-response plot for the CyAsp inhibition of the Src-mediated phosphorylation of different concentrations of Src-PS substrate.

5.3.7 Kinase assays in a different TK-Peptide system

In order to validate the generality of the approach, we tested our strategy in a completely different TK-system: the insulin-like growth factor 1 receptor (IGF1R) and the KKEEEE**Y**MMMMG peptide substrate (**E**4YM4).[45] Several of the cages inhibited the Tyr phosphorylation, with **CyAsp** and **CyHis** being the most potent ones in this case (Figure 5.15). These activities can be also explained with the corresponding supramolecular complexes. Thus, also in this case, a very stable [**CyAsp**-Mg-E4YM4] species was confirmed by fluorescence titration (Figure 5.15B) rendering a Logβ(ABC) = 5.406(3) and LogK_1 = 4.46. On the other hand, the fluorescence titration of E4YM4 with **CyHis** can be successfully fitted to a simpler 1 : 1 binding mode (Logβ(AB) = 3.784(8)), in agreement with the absence of a second aromatic amino acid in the substrate. This strong binding also explains the ability of **CyHis** to inhibit the IGF1R-promoted phosphorylation of E4YM4.

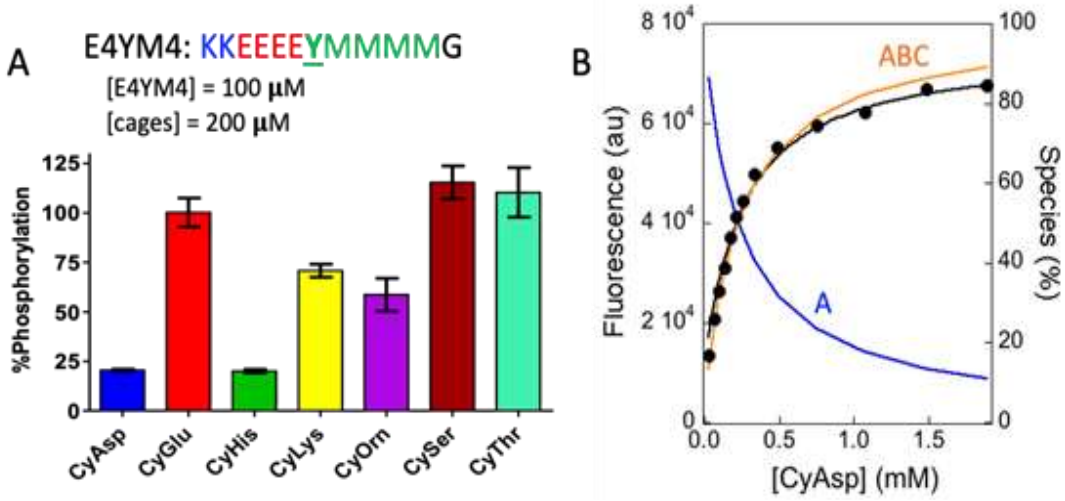


Figure 5.15 (A) Effect of the cages on the IGF1R-promoted Tyr-phosphorylation of the E4YM4 substrate (the plots show the percent phosphorylation relative to the corresponding control reactions in the absence of a cage). (B) Fluorescence emission (465 nm) of E4YM4 (black, 10 µM, 30 mM Tris buffer at pH 7.5, 20 mM $MgCl_2$) upon increasing concentrations of CyAsp and species distribution throughout the titrations.

5.4 CONCLUSIONS

- Our results demonstrate that the supramolecular binding of Tyr residues in peptides represents an efficient method to protect the substrates from the TK-mediated phosphorylation.
- The inhibitory activities of the pseudopeptidic cages reported can be rationalized considering the cage-peptide complexes, supporting the proposed mechanism of the Tyr inclusion inside the cage cavity.
- Thus, the binding constants and the inhibitory activities are modulated by the secondary interactions established between the side chains of the peptide substrates and the cages, as demonstrated through advanced NMR analysis, which complement the binding of the Tyr within the cage cavity.
- This approach paves the way towards the selective modulation of an individual kinase-stimulated signalling pathway, without interfering with other functions of the kinase, potentially leading to the development of improved tools for research, diagnosis, or therapy in biomedicine.

5.5 EXPERIMENTAL SECTION

5.5.1 Materials

Reagents and solvents were purchased from commercial suppliers (Adrich, Fluka, Iris Biotech or Merck) and were used without further purification. Compounds **CySer, CyThr, CyLys** and **CyOrn** were synthetized as previously described.[19,46] Details for the synthesis of **CyAsp**, **CyGlu** and **CyHis** are provided in chapter 3. The different polypeptides studied were designed and purchased in Inmuno-Dynamics Inc. and purified by preparative HPLC before use.

5.5.2 NMR spectroscopy

The NMR experiments were carried out at 25˚C on Bruker Avance-III 500 MHz spectrometer equipped with a z-axis pulsed field gradient triple resonance (^{1}H, ^{13}C, ^{15}N) TCI cryoprobe (500 MHz for ^{1}H and 125 MHz for ^{13}C). Chemical shifts are reported in ppm using. NMR spectra were acquired and processed using Bruker TopSpin 3.5 and MNova11 software (Mestrelab Research), respectively.

V5Ang-II:CyHis binding experiments: All the samples for this set of experiments were prepared in 90% H_2O/10% D_2O 30 mM HEPES-d18 with 5 mM $MgCl_2$ buffer, pH 7.0 and were acquired at 288K. The experiments for the characterization of the peptide were carried out using a 0.3 mM V5Ang-II reference sample. The samples to maximize the 1:1 and 1:2 V5Ang-II:CyHis supramolecular complexes were composed of 0.3:0.5 mM and 0.3:1.5 mM of peptide:cage, respectively. The pulse programs used for spectra acquisition were zggpw5 (1D ^{1}H with water suppression using watergate W5 pulse sequence) and roesygpph19 (2D ^{1}H-^{1}H ROESY, mix = 70 ms) from Bruker pulse sequence library, and clmlevgpphw5 (2D ^{1}H-^{1}H TOCSY, mix = 75 ms) as reported in the literature.[47]

BSA:CyLys binding experiments: For this set of experiments, a reference sample of 1 mM CyLys and a 4:1 CyLys:BSA (1 mM CyLys:250 µM BSA) sample were prepared in 90% H_2O/10% D_2O 30 mM HEPES-d18 5 mM MgCl2 pH=7,4 buffer. All the experiments were acquired at 298K. The pulse programs used for spectra acquisition were zggpw5 (1D ^{1}H with water suppression using watergate W5 pulse sequence), t1rho_esgp2d (pseudo 2D ^{1}H $T_{1\rho}$-filtered, two spectra acquired with 10 and 300 ms filter) cmpg_esgp2d (pseudo 2D ^{1}H T2-filtered, two spectra acquired with 10 and 300 ms filter) and stebpgp1s191d (1D ^{1}H diffusion-filtered, two spectra acquired with 5% and 95% gradient strength, Δ=120 ms, δ=2 ms) from Bruker pulse sequence library.

Table 5.2. Proton resonance assignment of Val5AngII peptide in 30 mM HEPES-d18 with 5 mM $MgCl_2$ buffer, pH 7.0, 288 K. Two rotamers of the Pro tertiary amide bond were observed in slow exchange in the proton chemical shift NMR timescale. We tentatively assigned these isomers as *trans* (84%) and *cis* (16%) rotamers of the prolinamide bond. Chemical shifts are reported in ppm. NO, not observed.

Residue		δ (*trans*)	δ (*cis*)
Asp1	NH	NO	
Arg2	NH	NO	NO
	*α*H	4.15	3.74
	*β*H1	1.55	1.55
	*β*H2	1.55	1.79
	*γ*Hs	1.34	1.37
		2.97	3.19
Val3	NH	8.11	
	*α*H	3.89	
	*β*Hs	1.79	
	*γ*H1	0.68	
	*γ*H2	0.73	
Tyr4	NH	8.37	8.36
	*α*H	4.33	4.42
	*β*H1	2.66	2.60
	*β*H2	2.70	2.74
	2,6H	6.84	6.88
	3,5H	6.53	6.57
Val5	NH	7.75	7.86
	*α*H	3.79	3.98
	*β*Hs	1.66	1.81
	γH1	0.63	0.69
	*γ*H2	0.63	0.71

Residue		δ (*trans*)	δ (*cis*)
His6	NH	8.20	8.09
	*α*H	4.51	4.15
	*β*H1	2.80	2.73
	*β*H2	2.86	2.77
	*2*H	7.79	7.60
	*4*H	6.86	6.77
Pro7	*α*H	4.17	
	*β*H1	1.68	
	*β*H2	1.97	
	*γ*Hs	1.77	
	*δ*H1	3.24	
	*δ*H2	3.58	
Phe8	NH	7.40	7.82
	*α*H	4.25	4.14
	*β*H1	2.82	2.90
	*β*H2	2.95	2.95
	2,6H	7.06	6.98
	3,5H	7.15	7.08
	4H	7.08	7.08

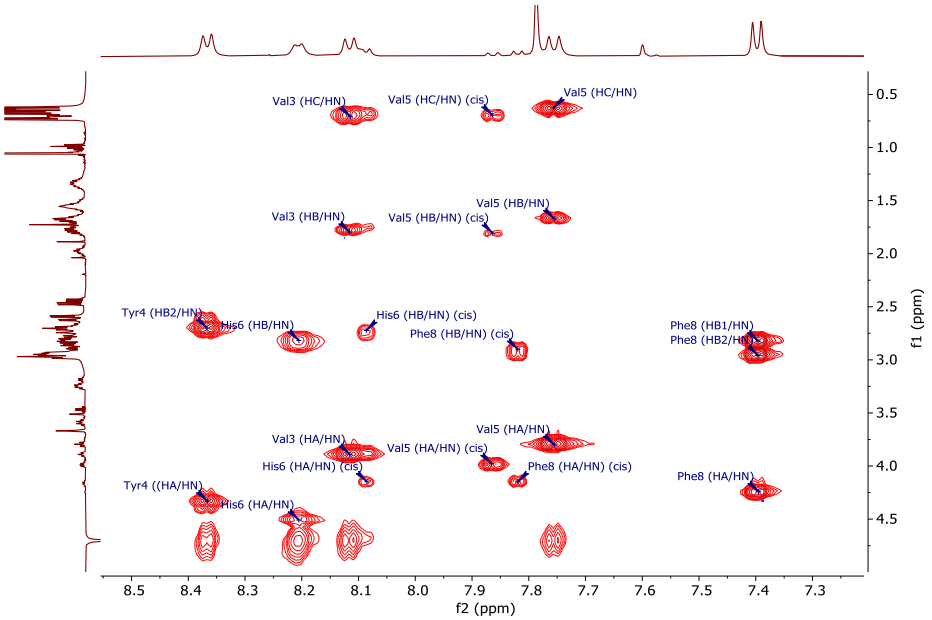


Figure 5.16. Amide region of the 500 MHz TOCSY spectrum of Val5AngII in 90% H_2O/10% D_2O, 30 mM HEPES-d18 with 5 mM $MgCl_2$ buffer at 288K. All the crosspeaks shown correspond to the main *trans* conformer except when stated otherwise.

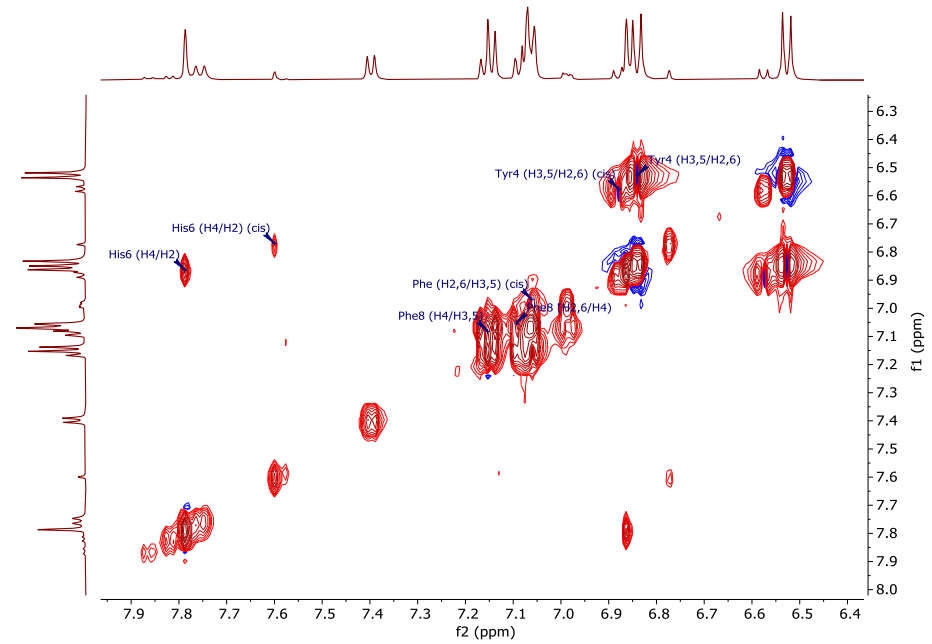


Figure 5.17. Aromatic ring region of the 500 MHz TOCSY spectrum of Val5AngII in 90% H_2O/10% D_2O 30 mM HEPES-d18 with 5 mM $MgCl_2$ buffer at 288K. All the crosspeaks shown correspond to the main *trans* conformer except when stated otherwise.

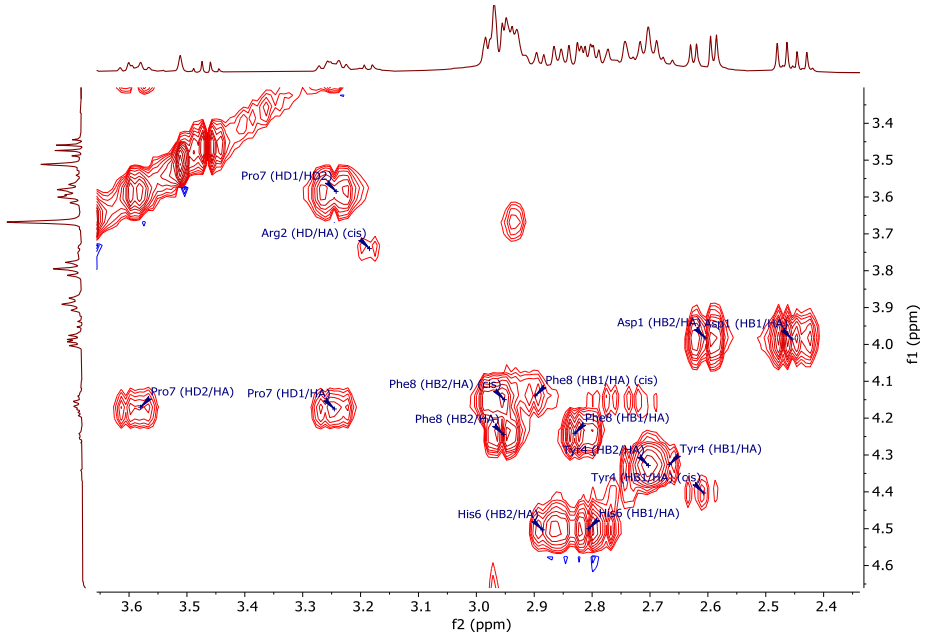


Figure 5.18. 3.7-2.3 ppm region of the 500 MHz TOCSY spectrum of Val5AngII in 90% H_2O/10% D_2O 30 mM HEPES-d18 with 5 mM $MgCl_2$ buffer at 288K. All the crosspeaks shown correspond to the main *trans* conformer except when stated otherwise.

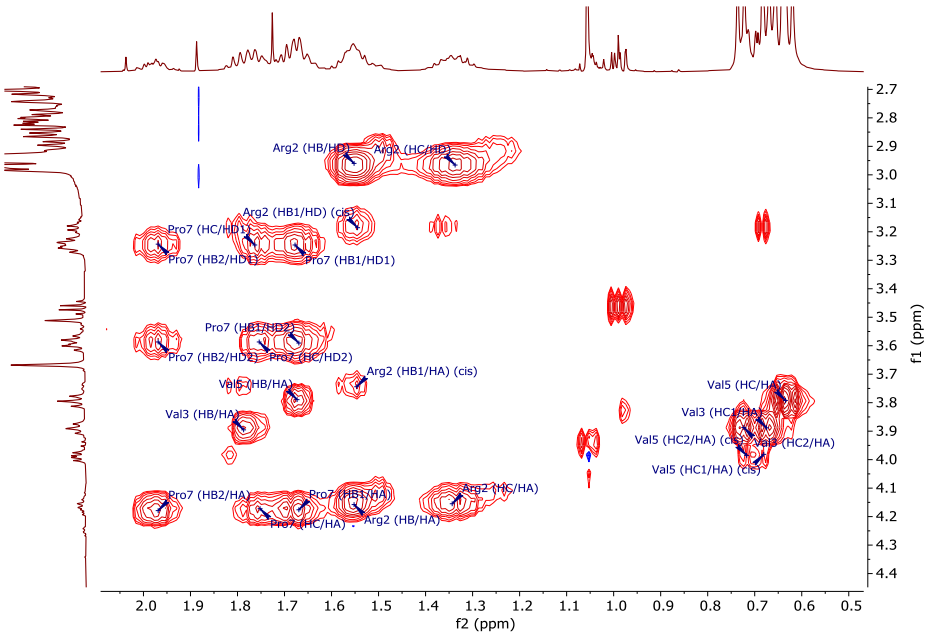


Figure 5.19. 2.2-0.5 ppm region of the 500 MHz TOCSY spectrum of Val5AngII in 90% H_2O/10% D_2O 30 mM HEPES-d18 with 5 mM $MgCl_2$ buffer at 288K. All the crosspeaks shown correspond to the main *trans* conformer except when stated otherwise.

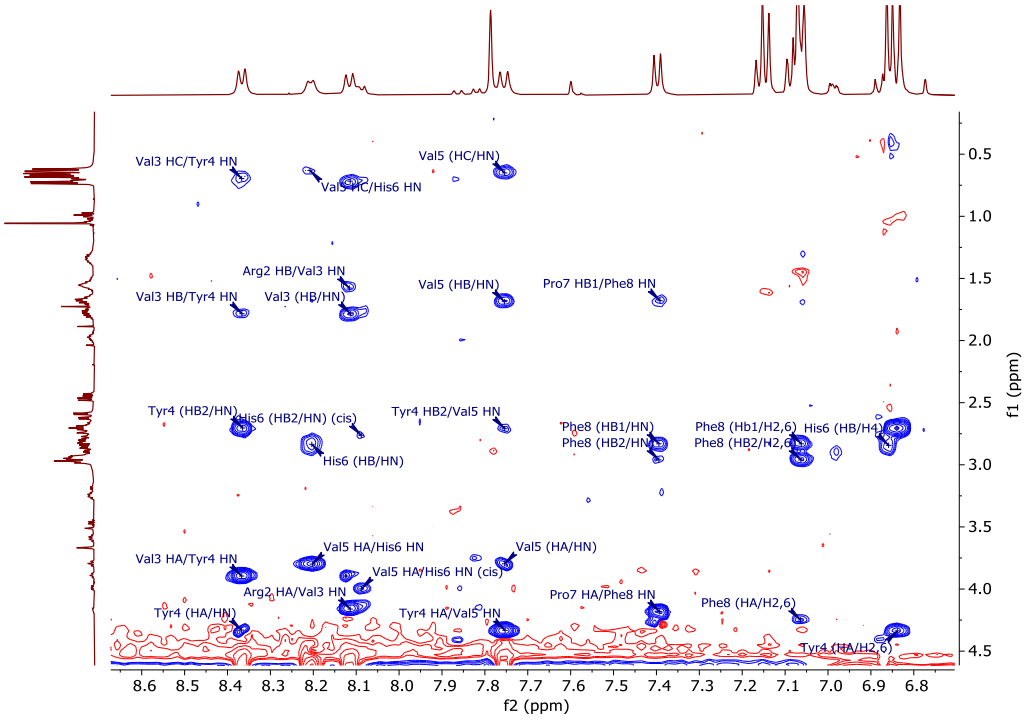


Figure 5.20. Amide region of the 500 MHz ROESY spectrum of Val5AngII in 90% H_2O/10% D_2O, 30 mM HEPES-d18 with 5 mM $MgCl_2$ buffer at 288K. All the crosspeaks shown correspond to the main *trans* conformer except when stated otherwise.

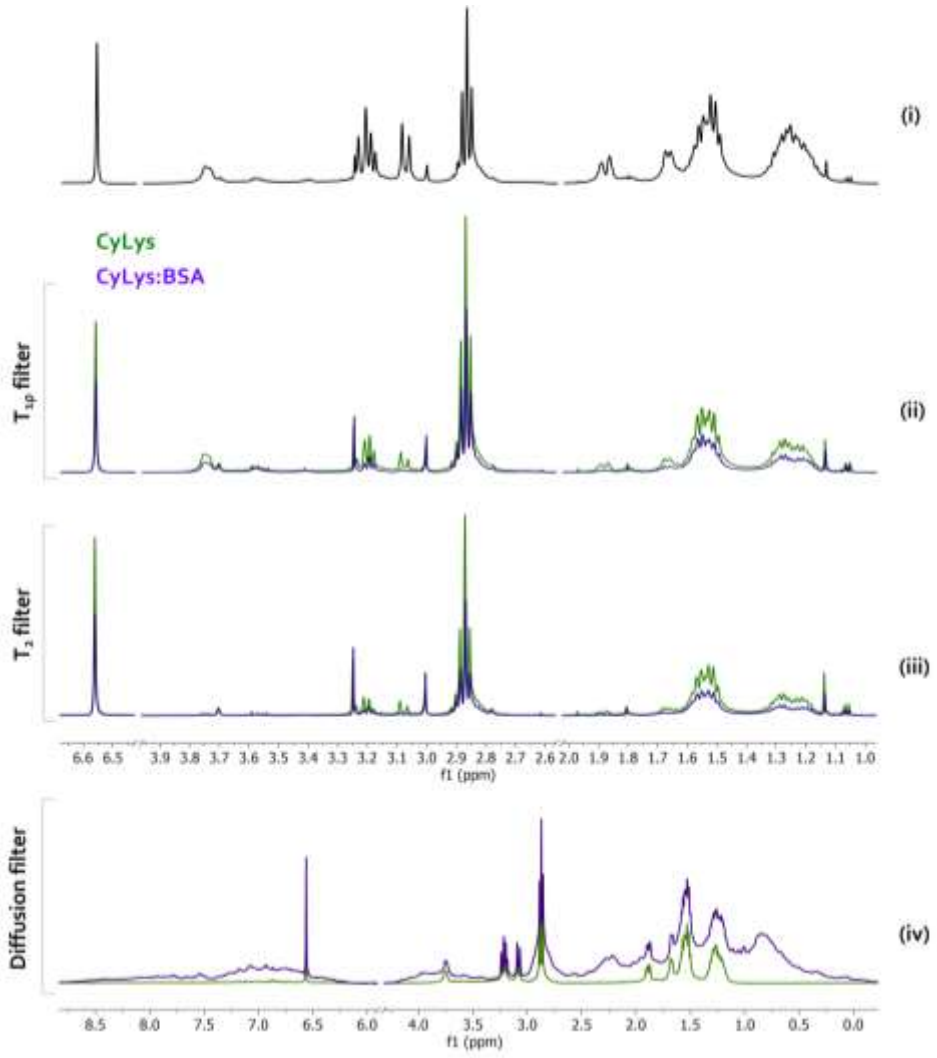


Figure 5.21. NMR binding experiments showing the interaction of CyLys with BSA. For these experiments, a reference sample of 1 mM CyLys and a 4:1 CyLys:BSA (1 mM CyLys:250 µM BSA) sample were prepared in 30 mM HEPES-d18 5 mM $MgCl_2$ pH=7.4 buffer. All the spectra were acquired at 298K in a Bruker Avance™ III HD spectrometer operating at 500 MHz (^{1}H resonance frequency) and equipped with a 5mm TCI (^{1}H/^{13}C/^{15}N) cryoprobe. From top to bottom: (i) ^{1}H reference spectrum of CyLys (ii) $T_{1\rho}$-filtered spectra (300 ms) of CyLys (green) and CyLys:BSA (purple). (iii) T_2-filtered spectra (300 ms) of CyLys (green) and CyLys:BSA (purple). The implementation of a $T_{1\rho}$ or T_2 relaxation filter results in attenuation of the NMR signals of species showing slow tumbling in solution, like large molecules. When small molecules bind to large biomolecules like BSA, their signals decrease compared to the unbound molecules. (iv) Diffusion-filtered spectra (95% gradient strength) of CyLys

(green) and CyLys:BSA (purple). Opposite to the relaxation filter experiments, the implementation of a diffusion filter results in attenuation of the NMR signals of species with small diffusion coefficients, like small molecules. When small molecules bind to large biomolecules like BSA, their signals increase compared to the unbound molecules. All the experiments (ii-iv) showed an apparent larger size of the CyLys compound in the presence of BSA, implying the effective binding of the cage to the protein under the experimental conditions.

5.5.3 Kinase activity assays

Src Kinase catalytic domain was expressed and purified from *E. coli* as described[45] and was a kind gift from Dr. Markus Seeliger (Stony Brook University). The IGF1R kinase domain was purified from *Spodoptera frugiperda* Sf9 cells using a recombinant baculovirus as described.[48] Radioactivity measurements were conducted in a HIDEX-300 SL. Scintillation counter. For assaying PTK (protein tyrosine kinase) activities, we measured the phosphorylation of the different polypeptides using the acid precipitation onto filter paper assay.[28,3] Phosphorylation reactions were conducted in Eppendorf tubes with a total volume of 25 µL per reaction. Each reaction mixture contains TRIS buffer (30 mM, pH 7.5), 20 mM $MgCl_2$, 400 µM ATP, peptide (**Src-PS**, **WASP** or **Val5-AngII** at the desired concentration), the cage (at different concentrations depending on the assay), ^{32}P-ATP (0.1 µL, 10 mci/mL) and SRC (at the optimal concentration for each substrate). Each reaction mixture was incubated at 30˚C for 5-10 minutes. At the end of the reaction time, 5 µL of trichloroacetic acid were added to the reaction Eppendorf and the solution was centrifuged. Then 35 µL of the supernatant were spotted into a 2x2 cm square Whatman filter paper which was subsequently washed in 5% H_3PO_4 3 times for 10 minutes each time followed by a final wash with acetone. Finally, each paper's radioactivity was determined by scintillation counting.

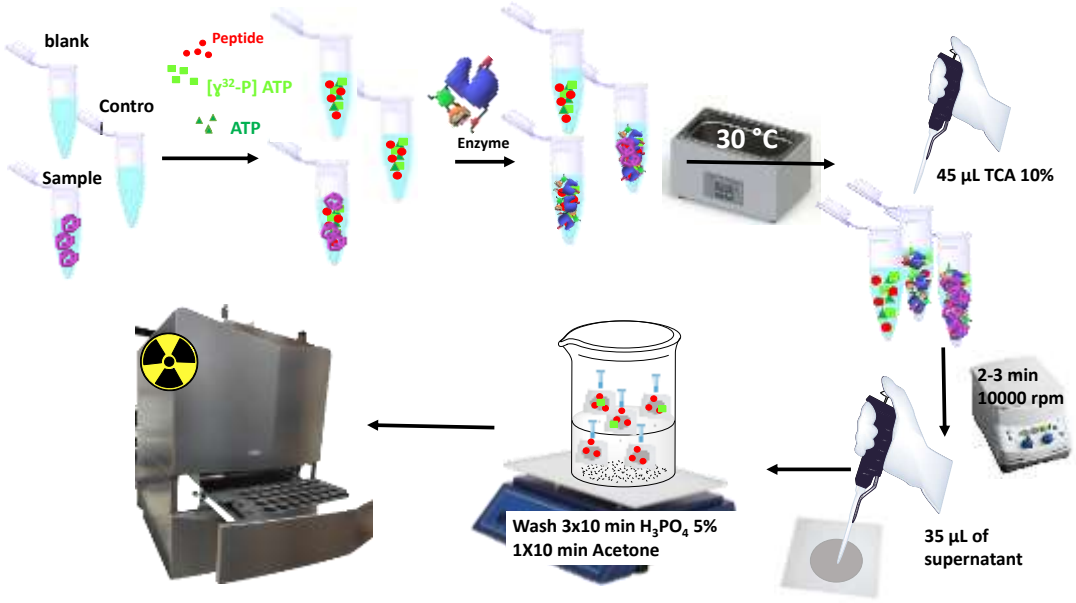


Figure 5.22. Graphical representation of the phosphocellulose paper assay.

The assay conditions (peptide concentrations, enzyme concentrations and incubation time) for each peptide were accurately optimized to ensure that the selected ones fall within the linear phase of the enzymatic reaction. Conditions:

- **PolyE_4Y**: 10 min incubation, 50 nm Src and 0.1 mg/ml of the polypeptide
- **Src-PS** and **A6Src-PS**: 7 min incubation, 10 nm Src and 100µM of the polypeptide
- **WASP**: 10 min incubation, 50 nm Src and 100µM polypeptide
- **V5Ang-II**: 15 min incubation, 200 nm Src and 250µM polypeptide.

Results were expressed as % of phosphorylation compared to the control experiments (assay without the cage compound) so controls were considered to be the 100%. Each plot is a representation of the average of at least 3 repetitions and error bars correspond to the standard deviation.

Kinase inhibition assays with the system **E_4YM_4**: The kinase inhibition activity of the cage compounds was tested with a different Kinase/peptide system as a control experiment for the action mechanism proposed. The system selected was the peptide E_4YM_4: (whose sequence is KKEEEE**Y**MMMMG) with the IGFIR kinase. The assay was conducted with 0.64 µM IGFIR and 100 µM polypeptide, and the incubation time was set to 10 minutes.

5.5.4 PEPFOLD 3.5 simulations

PEP-FOLD 3.5 (https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/) is a free online peptide structure simulation based on the concept of structural alphabet.[33–35] This software was used to predict the most probable structure of the different peptides studied using the amino acid sequence of each peptide as the input data. The obtained conformers were classified by conformational similarity in different clusters. The conformers within each cluster were aligned with Maestro package. The most populated cluster for each peptide is represented in Figure 2 of the main manuscript while the rest are represented in the ANEX. In these representations, the colour code in the backbone is: red (acid), blue (basic) and green (hydrophobic) residues.

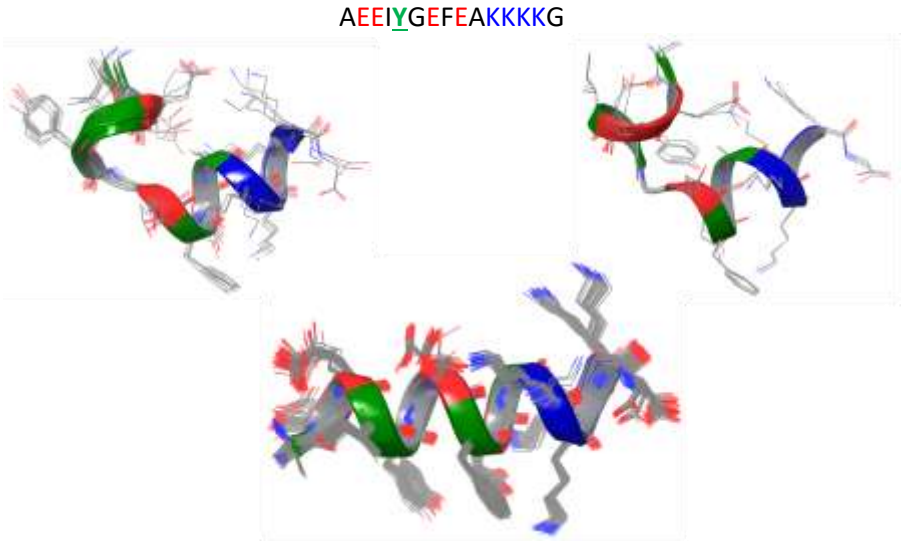


Figure 5.23. Superposition of the different conformers predicted for Src-PS peptide, in three different clusters. PEPFOLD-3.5 predicts an α-helix with some flexibility on the Gly residue at position 5. The helix structure shows a polar/non-polar distribution of the residues that could account for a coiled-coil dimerization of the helices as observed by self-dilution fluorescence experiments.

AEEIYAEFEAKKKKG

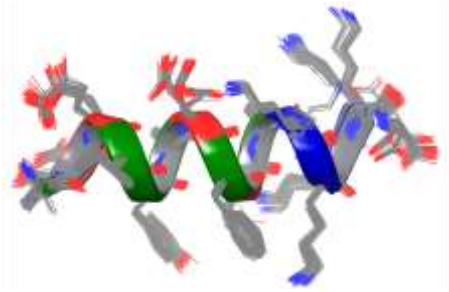


Figure 5.24. Superposition of the different conformers predicted for A5Src-PS peptide, in just one cluster. PEPFOLD-3.5 predicts an α-helix, where the Gly5Ala mutation reduces the helix flexibility, in agreement with the effect of this mutation reported in the literature.[49]

KVIYDFIEKKG

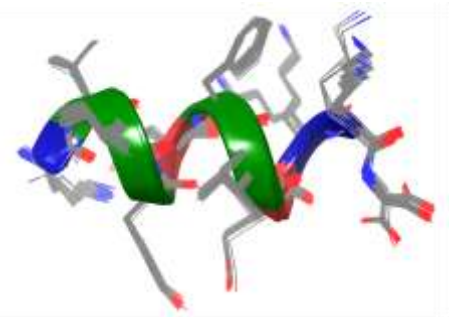


Figure 5.25. Superposition of the different conformers predicted for WASP peptide, in just one cluster. PEPFOLD-3.5 predicts an α-helix. The helix structure shows a polar/non-polar distribution of the residues that could account for a coiled-coil dimerization of the helices as observed by self-dilution fluorescence experiments.

DRVYVHPF

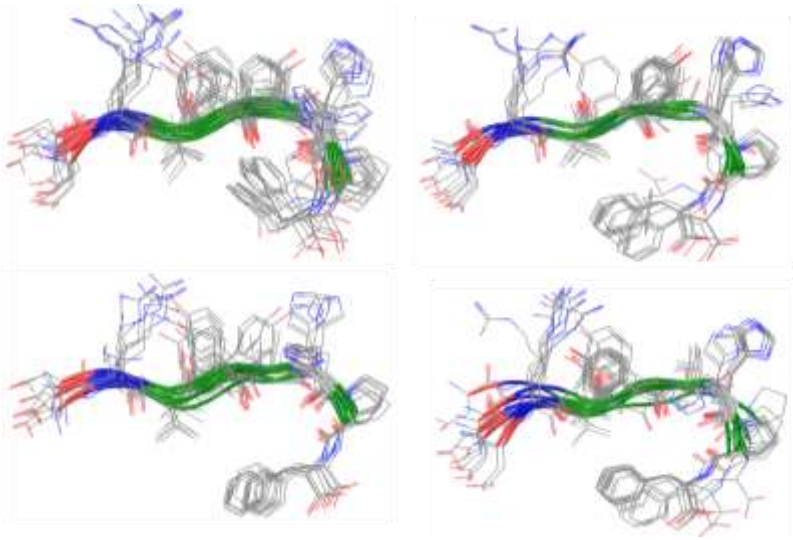


Figure 5.26. Superposition of the different conformers predicted for V5Ang-II peptide, showing the four most populated clusters (additional less populated clusters were located by the software). PEPFOLD-3.5 predicts a flexible conformation (random coil) with an incipient turn induced by the Pro in position 7. The overall conclusion is that this peptide is less structured than the others here studied. The hydrophobicity of the sequence (green) and the cluster formed by the three aromatic amino acids in close proximity for most of the conformations could account for the observed dimerization by fluorescence self-dilution experiments.

KKEEEEYMMMMG

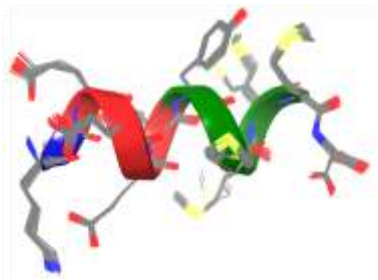


Figure 5.27. Superposition of the different conformers predicted for E4YM4 peptide, in just one cluster. PEPFOLD-3.5 predicts an α-helix as the main conformation. In this case, the polar/non-polar residues are not distributed in the typical amphiphilic faces leading to coiled-coils, explaining the absence of self-assembling properties of the peptide from the fluorescence self-dilution experiments.

5.5.5 Fluorescence spectroscopy

Fluorescence emission spectra were acquired on a SpectraMax M5 instrument using 10 mm path length cuvettes, excitation bandwidth: 9 nm, emission bandwidth: 15 nm, light source: Xenon flash lamp (1 joule/flash), emission read every 1 nm. All the fluorescence experiments were performed at 20°C and specific measuring details and fitting procedures are given in the corresponding section for each titration example. The different peptide-cage titrations were all conducted in a 700 µL fluorescence cuvette following the following protocol: A solution of the peptide (100, 10 or 1 µM) was prepared in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$). 300 µL of the peptide solution was titrated with a solution of the cage (1-3 mM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) containing the titrated peptide in the same concentration (100, 10 or 1 µM) to maintain the peptide concentration constant throughout the whole titration. The peptide concentration was adjusted for each titration to the concentration that prevents precipitation events and allows to get a larger number of meaningful experimental points for the fitting. The excitation wavelength was λ_{ex}: 276 nm and the emission window recorded was adjusted for each peptide to acquire the whole emission bad for the excimer λ_{em} 290-500/550 nm. HypSpec[36] was the software used for the fitting of the titration data to a proposed interaction model for each cage-peptide pair.

Data fitting-

HypSpec was the software used for the fitting of the titration data to a proposed interaction model for each cage-peptide pair. HypSpec is based on multicomponent analysis, one of the most accurate methods available for the fitting of titration data.[5,6]

HypSpec allows the global fitting of the whole emission band (or the desired range) for each titration point to the interaction model selected, which can include several association constants between 1, 2 or more components. Additionally, the software permits to fix one of the constants to fit the others (especially useful if one of the interactions has been previously studied with another method or in a previous experiment).

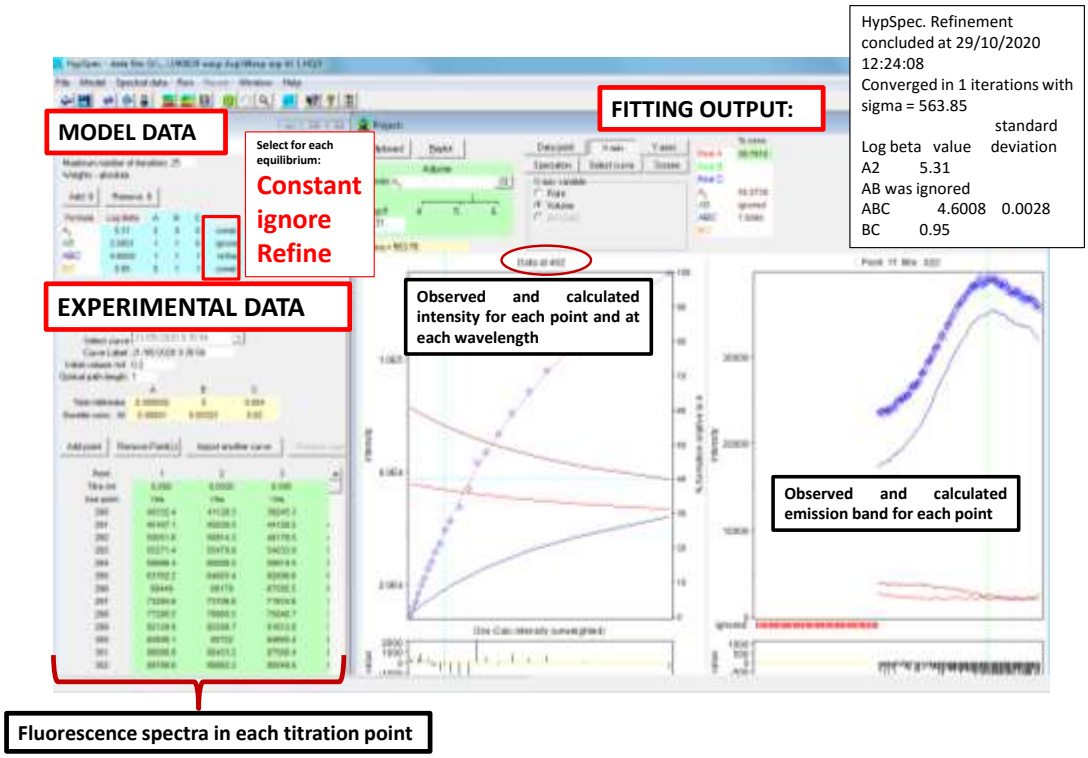


Figure 5.28. HypSpec output screen with additional comments on the most relevant sections.

Peptides aggregation

Fluorescence self-dilution experiments were conducted to determine the aggregation of the peptides. The four peptides studied (Src-PS, WASP, Val5-AngII and E_4YM_4) were assayed using the same protocol.

The fluorescence of the peptide was measured at growing concentrations in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) using λ_{ex}: 276 nm and measuring its emission from 290 to 360 nm at 20˚C. The fitting of the data to the aggregation model proposed for each peptide using HypSpec I detailed next.

Self-dilution titration of Src-Ps peptide

Src-Ps fluorescence emission was recorded at different concentrations (0.5-60 µM) in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) using λex = 276 nm.

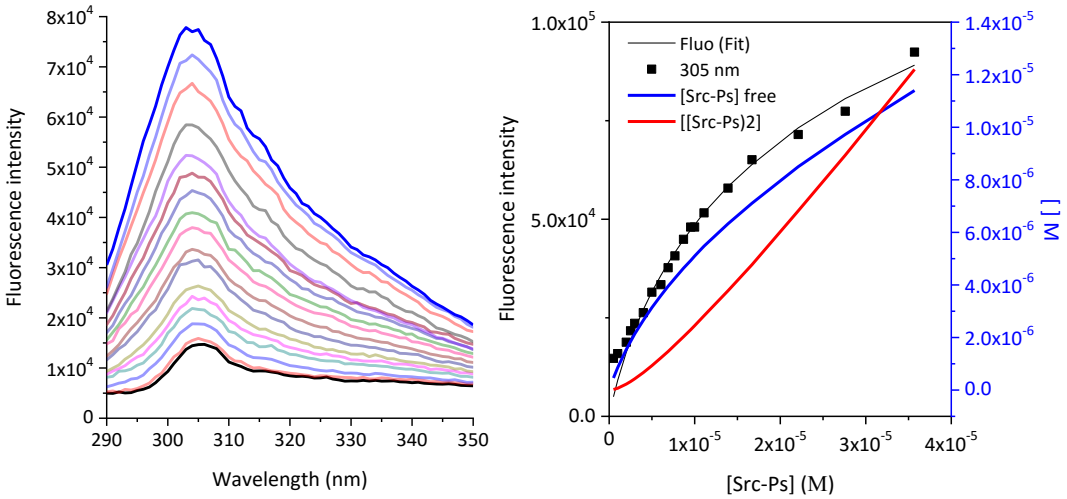


Figure 5.29. Emission spectra of Src-Ps at different concentrations (0.5-60 µM) in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) (left). Fitting of the self-dilution titration emission of Src-PS using HypSpec (left axis) and species distribution throughout the titration (right axis) (right).

The plot of the fluorescence emission of Src-PS in the λ_{max} = 305 nm (Figure 5.29) vs the peptide concentrations renders a non-linear response which suggests the formation of aggregates. The emission data was fitted to a dimerization model using HypSpec software.

Dimerization model for Src-PS and HypSpec result of this fitting: A + A ⇆ A2 log β = 4.97 ± 0.03

Self-dilution titration of WASP peptide

WASP fluorescence emission was recorded at different concentrations (0.5-120 µM) in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) using λ_{ex}: 276 nm.

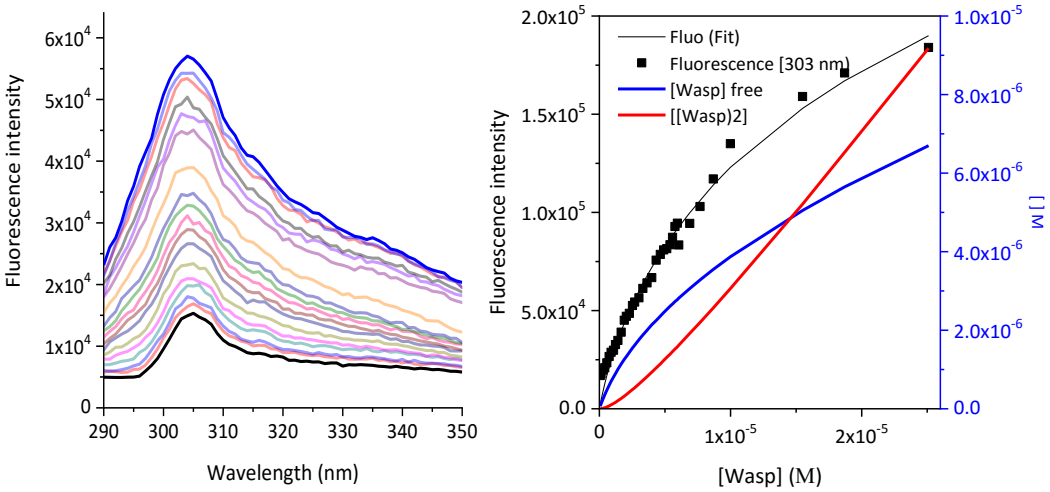


Figure 5.30. Emission spectra of WASP at different concentrations (0.5-120 µM) in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) (left). Fitting of the autodilution titration emission of WASP using HypSpec (left axis) and species distribution throughout the titration (right axis) (right).

The plot of the fluorescence emission at the λ_{max}: 303 nm (Figure 5.30) vs WASP concentrations renders a non-linear response which again suggests the formation of aggregates. The emission data was fitted to a dimerization model using HypSpec. The model and the dimerization constant derived from the fitting were:

$A + A \leftrightarrows A_2 \quad \log \beta = 5.20 \pm 0.03$

Data collected from the self-dilution experiment overlaps with the fitting to the dimerization model for WASP peptide. The dimerization constant for WASP peptide is high enough to expect the presence of dimer even at very low peptide concentrations.

Self-dilution titration of V5Ang-II

Val5-angiotensin II fluorescence emission was recorded at different concentrations (0.5-100 µM) in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) using λ_{ex}: 276 nm.

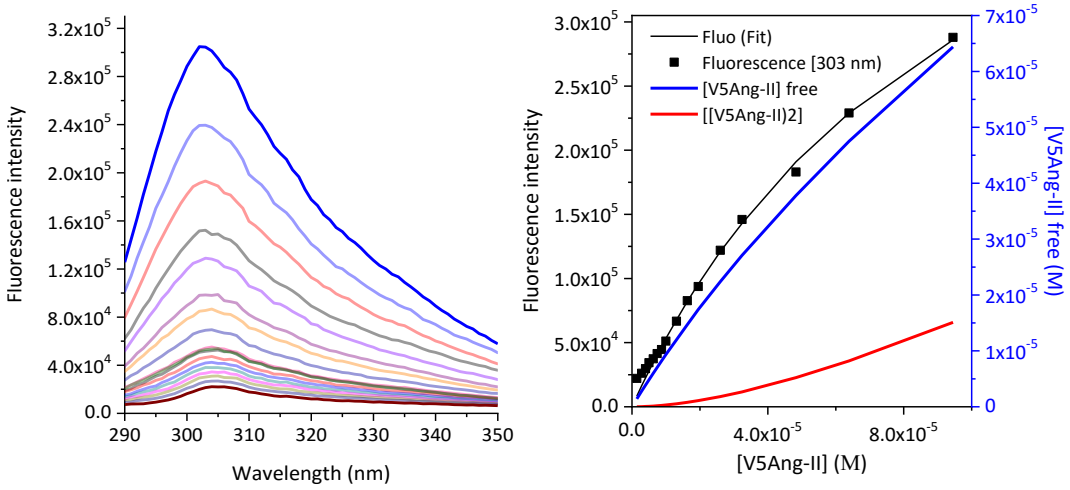


Figure 5.31. Emission spectra of Val5-Angiotensin II at different concentrations (0.5-100 µM) in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) (left). Fitting of the autodilution titration emission of Val5-Angiotensin II using HypSpec (left axis) and species distribution throughout the titration (right axis) (right).

The plot of the fluorescence emission of V5Ang-II at λ_{max}: 303 nm (Figure 5.31) vs V5Ang-II concentrations results in a nearly linear relation but when studied carefully it also indicates the formation of some type of aggregates detectable at high concentrations of the peptide. The emission data at 303 nm was fitted to a dimerization model using HypSpec software:

$A + A \leftrightarrows A_2 \quad \log \beta = 3.56 \pm 0.02$

The good overlap between the experimental data and the fitting points confirms the validity of the dimerization model at least within the concentrations range studied for V5Ang-II.

Self-dilution titration of E_4YM_4

E_4YM_4 fluorescence emission was recorded at different concentrations (0.5-100 µM) in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) using λ_{ex}: 276 nm.

The plot of the fluorescence emission of E_4YM_4 in the λ_{max}: 305 nm (figure S36) vs E_4YM_4 concentrations shows a linear response which indicates that no aggregation is happening within the studied concentrations range.

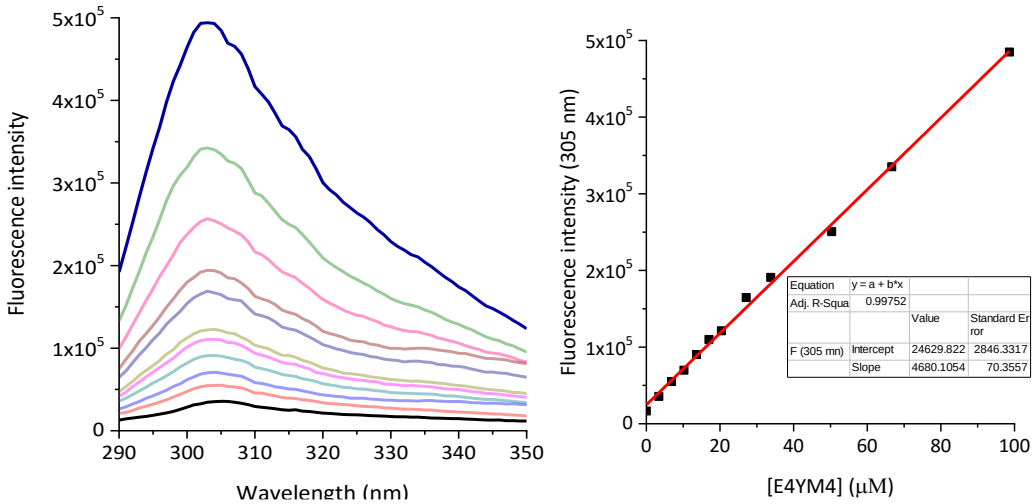


Figure 5.32. Emission spectra of E_4YM_4 at different concentrations (0.5-100 μM) in buffered water (30mM TRIS, pH 7.5, 20 mM $MgCl_2$) (left). Fitting of the self-dilution titration emission of E_4YM_4 to a linear equation using Origin V8. (right).

Titration of Src-PS Peptide by CyGlu

A 5 μM solution of Src-Ps in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 2 mM solution of **CyGlu** that contains Src-PS ([Src-Ps = 5 μM). using λ_{ex} = 276 nm.

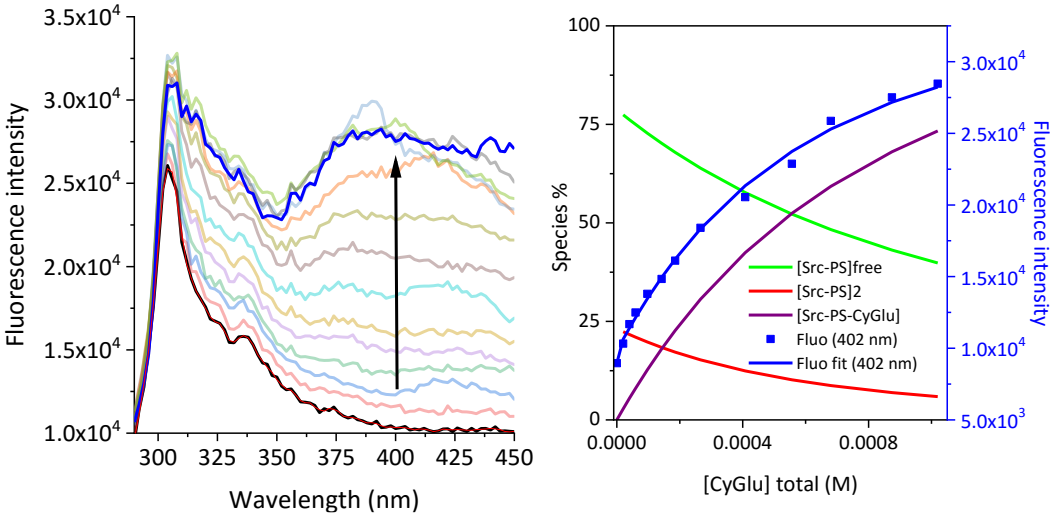


Figure 5.33. Emission spectra of Src-Ps (5 μM) in buffered water (30 mM TRIS, pH 7 .5, 20 mM $MgCl_2$) in the presence of different concentrations of CyGlu ([CyGlu = 0-1.2 mM) using λ_{ex} = 276 nm (left). Fitting of the titration of Src-Ps (5 μM) with CyGlu at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model proposed for the fitting of the different equilibrium present were:

$$A + A \leftrightarrows A_2 \log \beta = 4.97 \text{ (constant)}$$
$$A+B \leftrightarrows AB \log = 3.258 \pm 0.006 \text{ (refine)}$$

Where A corresponds to Src-Ps and B to **CyGlu.**

Titration of Src-PS Peptide by CyHis

A 1 µM solution of Src-Ps in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 2.3 mM solution of **CyHis** that contains Src-PS ([Src-Ps = 1 µM). using λ_{ex} = 276 nm. In this case the concentration of the peptide was reduced up to 1 µM to avoid aggregation events and the formation of precipitates at high concentrations of the cage.

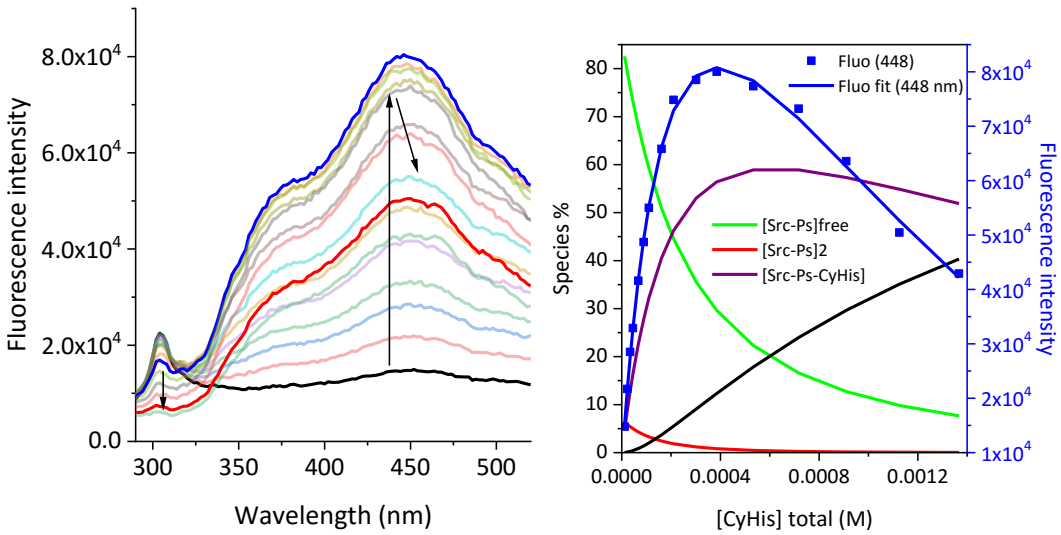


Figure 5.34. Emission spectra of Src-Ps (1 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyHis ([CyHis = 0-1.4 mM) using λ_{ex}: 276 nm (left). Fitting of the titration of Src-Ps (1 µM) with CyHis at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model used for the fitting of the log β for each equilibrium was:

$$A + A \leftrightarrows A_2 \log \beta = 4.97 \text{ (constant)}$$
$$A+B \leftrightarrows AB \log \beta = 3.697 \pm 0.006 \text{ (refine)}$$
$$A + 2B \leftrightarrows AB_2 \log \beta = 6.45 \pm 0.012 \text{ (refine)}$$

Where A corresponds to Src-Ps, B to **CyHis.** The dimerization of Src-PS had already been studied through fluorescence (section: Self-dilution titration of Src-Ps peptide) in the same

conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The best model for the data fitting includes the formation of both complexes AB and AB_2. The log β for the formation of these two complexes were successfully determined with HypSpec with the model proposed and a low deviation between the experimental data and the data derived from the model.

Titration of Src-Ps by CyLys

A 10 µM solution of Src-Ps in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 4 mM solution of CyLys that contains Src-PS ([Src-Ps= 10 µM). using λ_{ex}: 276 nm.

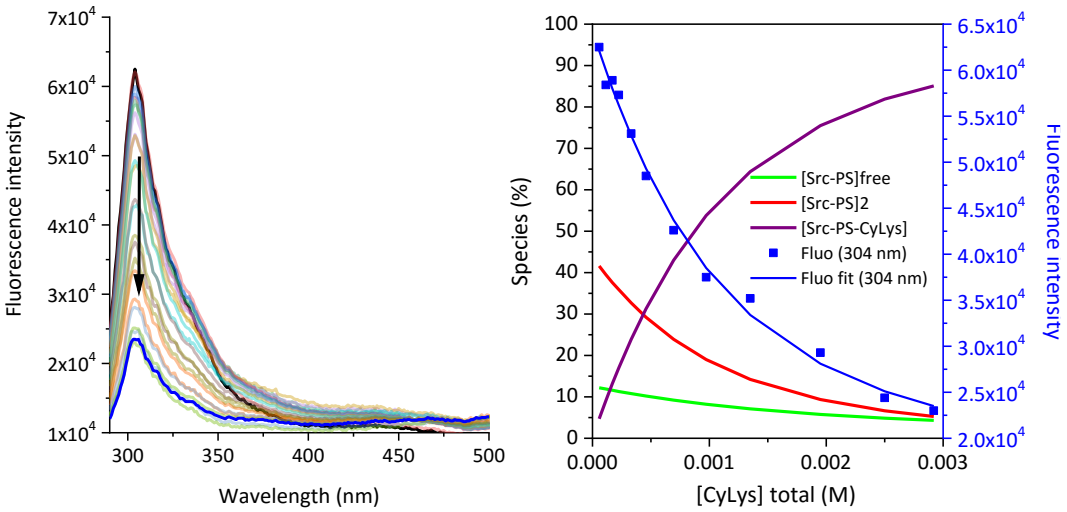


Figure 5.35. Emission spectra of Src-Ps (10 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyLys ([CyLys = 0-2.8 mM) using λ_{ex}: 276 nm (left). Fitting of the titration of Src-PS (10 µM) with CyLys at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model used for the fitting and the log β obtained were:

$$A + A \leftrightarrows A_2 \quad \log \beta = 4.97 \text{ (constant)}$$
$$A+B \leftrightarrows AB \quad \log \beta = 3.82 \pm 0.03 \text{ (refine)}$$

Where A corresponds to Src-Ps, B to **CyLys**. The dimerization of Src-PS had already been studied through fluorescence (section: self-dilution titration of Src-Ps peptide) in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The association constant for the formation of the complex AB was

successfully determined with HypSpec with the model proposed and a low deviation between the experimental data and the data derived from the model (Figure 5.35).

Titration of Src-Ps by CyOrn

A 1 µM solution of Src-Ps in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 4 mM solution of CyOrn that contains Src-PS ([Src-Ps = 1 µM). using λ_{ex} = 276 nm.

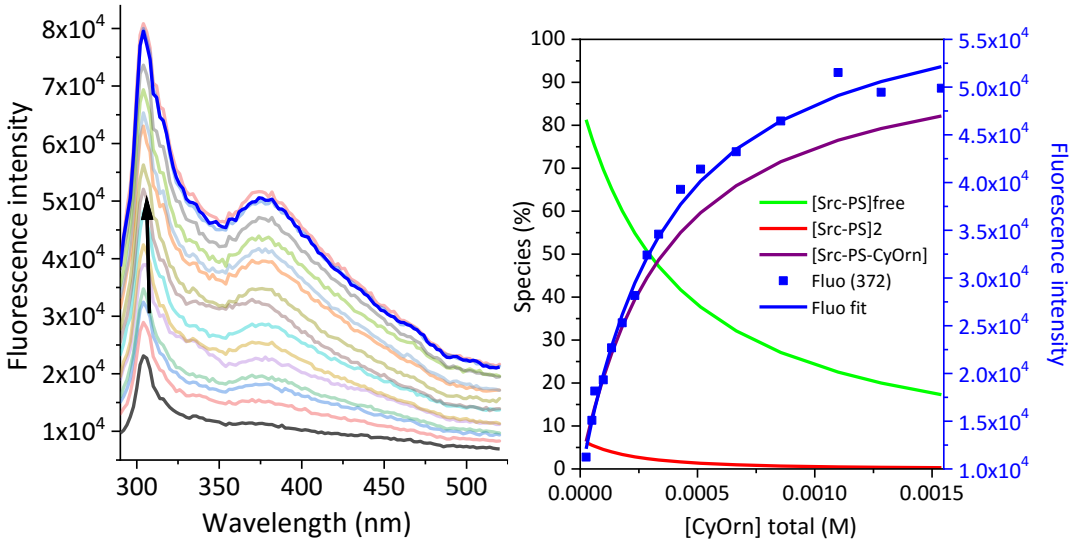


Figure 5.36. Emission spectra of Src-Ps (1 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyOrn ([CyOrn = 0-2.8 mM) using λ_{ex}: 276 nm (left). Fitting of the titration of Src-PS (1 µM) with CyOrn at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β = 4.97 (constant)
$A + B \leftrightarrows AB$ log β = 3.489 ± 0.004 (refine)

Where A corresponds to Src-Ps, B to **CyOrn**. The dimerization of Src-PS had already been studied through fluorescence (section: Self-dilution titration of Src-Ps peptide) in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The association constant for the formation of the complex AB was successfully determined with HypSpec.

Titration of WASP peptide by CyAsp

A 1 µM solution of WASP in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 2.4 mM solution of CyAsp that contains WASP ([WASP] = 1 µM). using λ_{ex} = 276 nm.

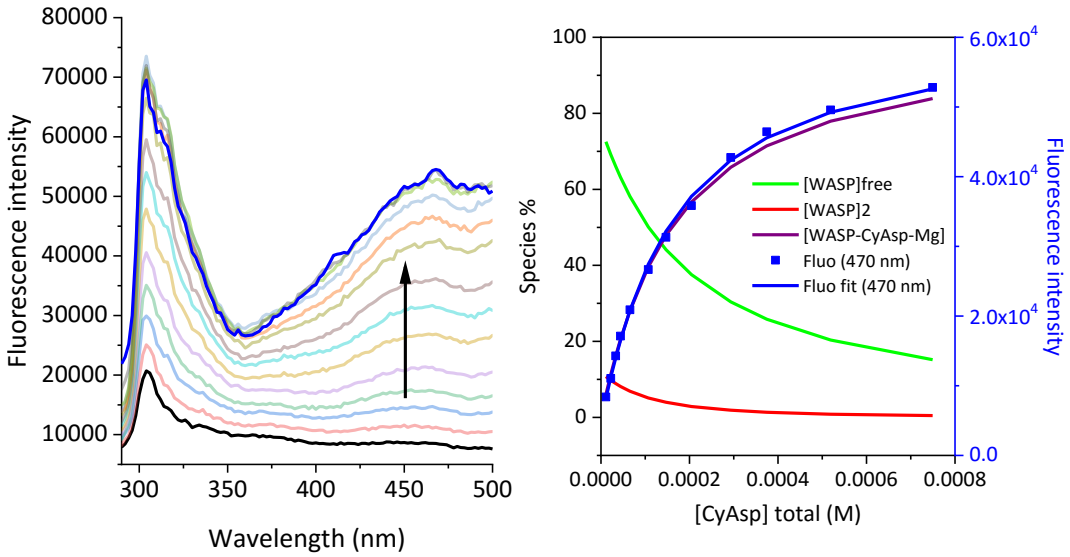


Figure 5.37. Emission spectra of WASP (1 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyAsp ([CyAsp = 0-1.4 mM) using λ_{ex}: 276 nm (left). Fitting of the titration of WASP (1 µM) with CyAsp at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β: 5.201 (fix)
$B + C \leftrightarrows BC$ log β: 0.95 (fix)
$A + B + C \leftrightarrows ABC$ log β: 5.640 ± 0.003

Where A corresponds to WASP, B to **CyAsp** and C to Mg^{2+} The dimerization of WASP had already been studied through fluorescence (section: Self-dilution titration of WASP peptide) in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The association between **CyAsp** and Mg^{2+} was also introduced to the model as a constant. The proposed interaction model for the pair WASP/**CyAsp** consists of the mentioned equilibria plus the formation of the trimer CyAsp-Mg^{2+}-WASP which is the only constant left to be fitted with HypSpec (Figure 5.37). The good overlap between the experimental data and the data from the fitting model validates the equilibrium model proposed for this system.

Titration of Wasp Peptide by CyGlu

A 10 µM solution of Wasp in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 1.66 mM solution of CyGlu that contains Wasp ([Wasp = 10 µM). using λ_{ex}: 276 nm.

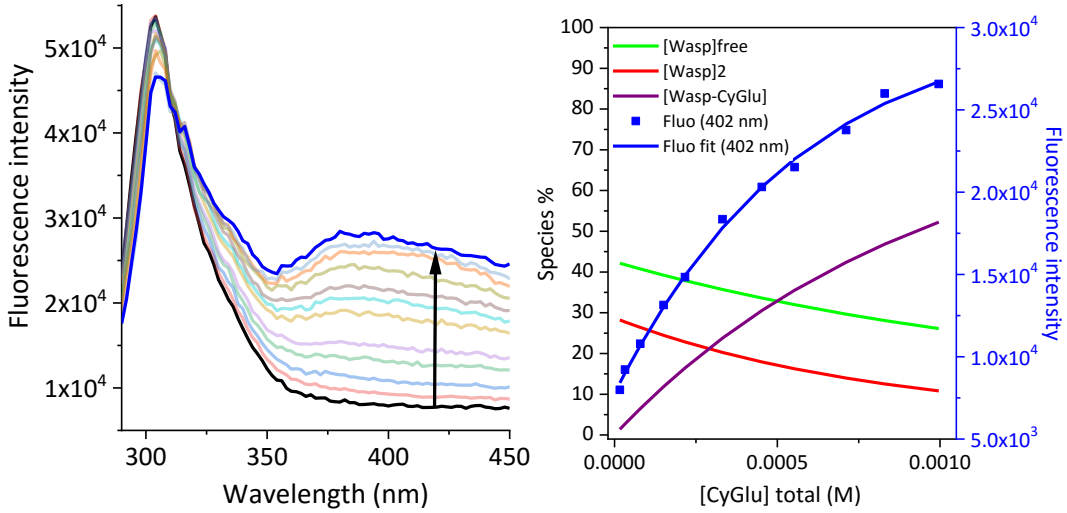


Figure 5.38. Emission spectra of Wasp (1 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyGlu ([CyGlu = 0-1.2 mM) using λ_{ex}: 276 nm. (left) Fitting of the titration of Wasp (10 µM) with CyGlu at pH 7.5 (right axis) and species distribution throughout the titration (left axis). (right).

The equilibrium model proposed for the fitting of the different equilibrium present was:

$A + A \leftrightarrows A_2$ log β = 5.201 (constant)

$A+B \leftrightarrows AB$ log β = 3.3055 ± 0.0033 (refine)

Where A corresponds to WASP and to **CyGlu.**

Titration of WASP peptide by CyHis

A 10 µM solution of WASP in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 2.4 mM solution of CyHis that contains WASP ([WASP] = 10 µM). using λ_{ex} = 276 nm.

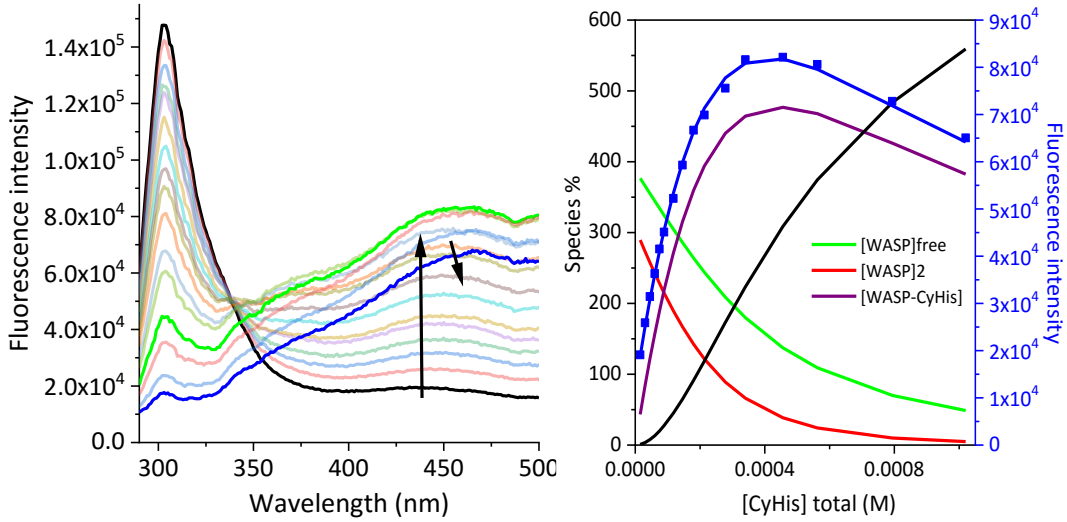


Figure 5.39. Emission spectra of WASP (10 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyHis ([CyHis = 0-1.1 mM) using λ_{ex}: 276 nm. Fitting of the titration of WASP (10 µM) with CyHis at pH 7.5 (right axis) and species distribution throughout the titration (left axis).

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β = 5.201 (constant)
$A+B \leftrightarrows AB$ log β = 3.89 ± 0.02 (refine)
$A+2B \leftrightarrows AB_2$ log β = 6.61 ± 0.02 (refine)

Where A corresponds to WASP and B to **CyHis.** The dimerization of WASP had already been studied through fluorescence (section: Self-dilution titration of WASP) in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The best model for the data fitting includes the formation of both complexes AB and AB2. The log β for the formation of these two complexes were successfully determined with HypSpec with the model proposed and a low deviation between the experimental data and the data derived from the model (Figure 5.39).

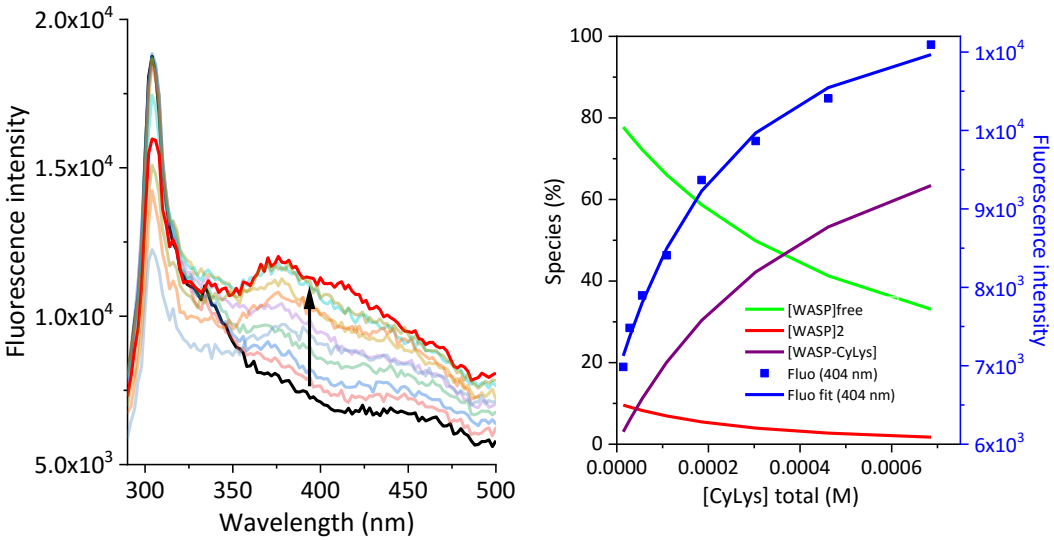


Figure 5.40. Emission spectra of WASP (11 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyLys ([CyLys=0-2.2 mM) using λ_{ex}: 276 nm.(left) Fitting of the titration of WASP (1 µM) with CyLys at pH 7.5 (right axis) and species distribution throughout the titration (left axis).(right).

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β = 5.201 (constant)
$A+B \leftrightarrows AB$ log β = 3.447 ± 0.006 (refine)

Where A corresponds to WASP, B to **CyLys**. The dimerization of WASP had already been studied through fluorescence (section: Self-dilution titration of WASP peptide) in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The association constant for the formation of the complex AB was successfully determined with HypSpec.

Titration of WASP peptide by CyOrn

A 1 µM solution of WASP in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 4 mM solution of **CyOrn** that contains WASP ([WASP] = 1 µM). using λ_{ex} = 276 nm.

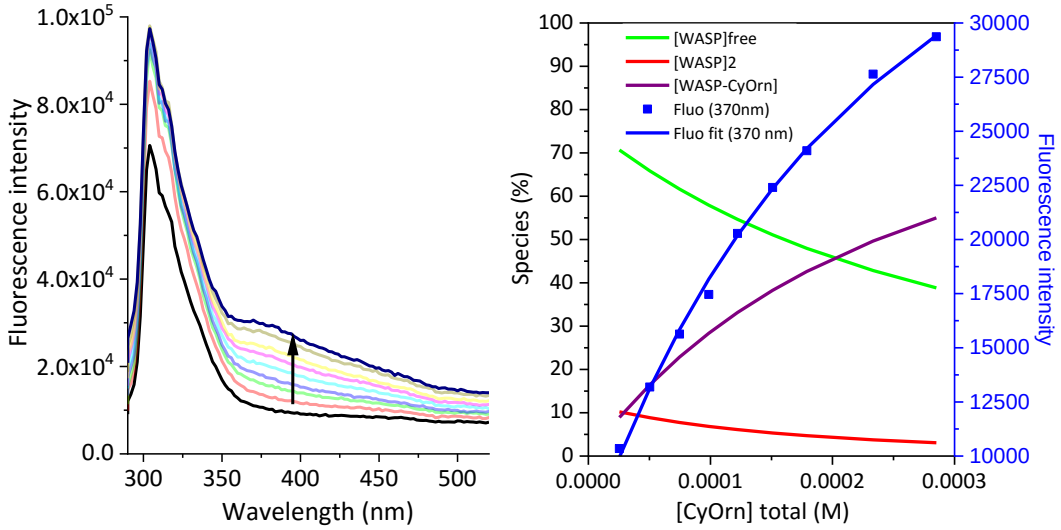


Figure 5.41. Emission spectra of WASP (1 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyOrn ([CyOrn=0-3 mM) using λ_{ex}: 276 nm. (left) Fitting of the titration of WASP (1 µM) with CyOrn at pH 7.5 (right axis) and species distribution throughout the titration (left axis).(right)

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β = 5.201 (constant)
$A+B \leftrightarrows AB$ log β = 3.537 ± 0.006 (refine)

Where A corresponds to WASP and B to **CyOrn**. The dimerization of WASP had already been studied through fluorescence (section: Self-dilution titration of WASP peptide) in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The association constant for the formation of the complex AB was successfully determined with HypSpec.

Titration of V5Ang-II by CyAsp

A 100 µM solution of V5Ang-II in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 3.2 mM solution of **CyAsp** that contains V5Ang-II ([V5Ang-II] = 100 µM). λ_{ex} = 276 nm.

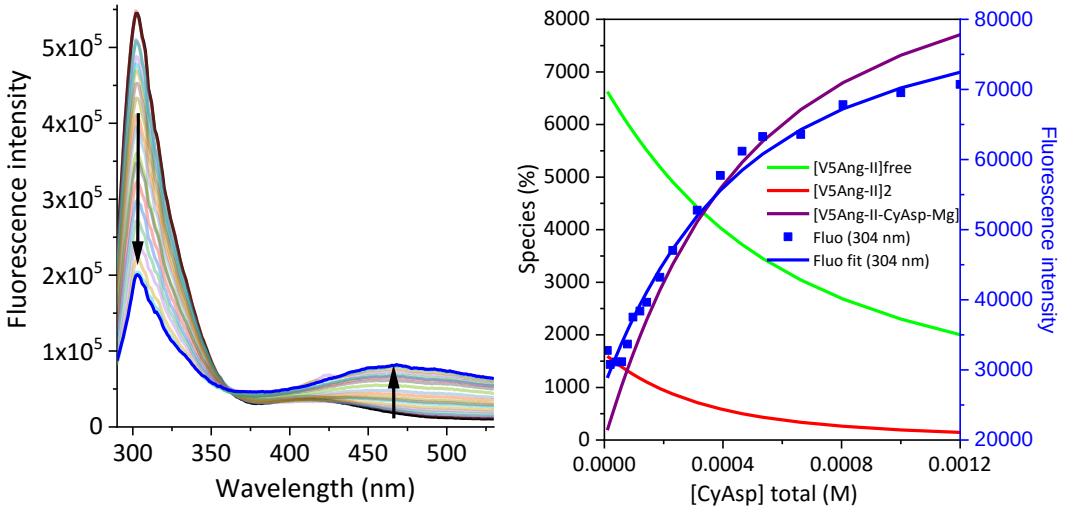


Figure 5.42. Emission spectra of V5Ang-II (100 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyAsp ([CyAsp = 0-1.3 mM) using λ_{ex}: 276 nm (left). Fitting of the titration of V5Ang-II (100 µM) with CyAsp at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β = 3.56 (constant)
$B + C \leftrightarrows BC$ log β = 0.95 (constant)
$A + B + C \leftrightarrows ABC$ log β = 5.30 ± 0.01 (refine)

Where A corresponds to V5Ang-II, B to **CyAsp** and C to Mg^{2+} The dimerization of V5Ang-II had already been studied through fluorescence (section: Self-dilution titration of V5Ang-II peptide) in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The association between **CyAsp** and Mg^{2+} was also introduced to the model as a constant. The proposed interaction model for the pair V5Ang-II/**CyAsp** consists of the mentioned equilibria plus the formation of the trimer **CyAsp**-Mg^{2+}-V5Ang-II which is the only constant left to be fitted with HypSpec (Figure 5.42). The good overlap between the experimental data and the data from the fitting model validates the equilibrium model proposed for this system.

Fluorescence titration of V5Ang-II by CyAsp in the absence of $MgCl_2$

A 100 µM solution of V5Ang-II in buffered water (30 mM TRIS, pH 7.5) was titrated with a 3.3 mM solution of CyAsp that contains V5Ang-II ([V5Ang-II]= 100 µM). using λ_{ex}: 276 nm.

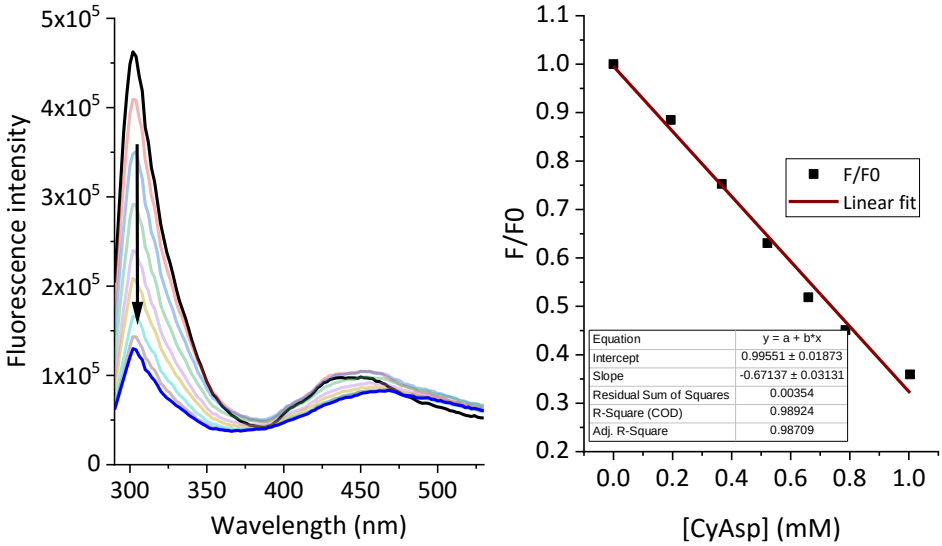


Figure 5.43. Emission spectra of V5Ang-II (100 µM) in buffered water (30 mM TRIS, pH 7.5) in the presence of different concentrations of CyAsp ([CyAsp=0-1.2 mM) using λ_{ex}: 276 nm (left). S-V plot of the titration of V5Ang-II (100 µM) by CyAsp at 302 nm (right).

Titration of V5Ang-II Peptide by CyGlu (20 mM $MgCl_2$)

A 10 µM solution of V5Ang-II in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 2.2 mM solution of **CyGlu** that contains V5Ang-II ([V5Ang-II = 10 µM). using λ_{ex} = 276 nm.

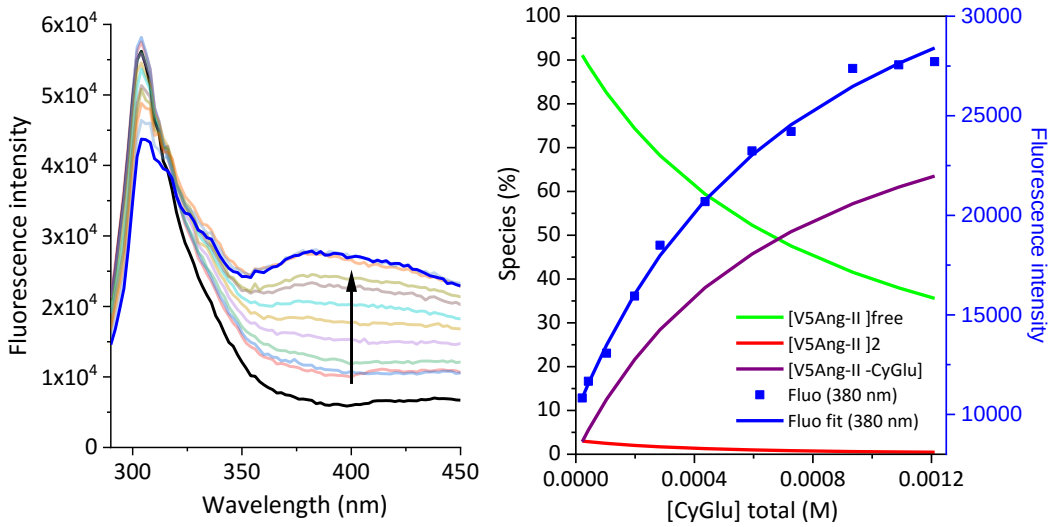


Figure 5.44. Emission spectra of V5Ang-II (10 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyGlu ([CyGlu=0-1.2 mM) using λ_{ex}: 276 nm (left). Fitting of the titration of V5Ang-II 10 µM) with CyGlu at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model proposed for the fitting and the log β of the different equilibrium present were:

$A + A \leftrightarrows A_2$ log $\beta = 3.56$ (constant)

$A+B \leftrightarrows AB$ log $\beta = 2.844 \pm 0.006$ (refine)

Where A correspond to V5Ang-II and B to **CyGlu.**

Titration of V5Ang-II Peptide by CyGlu (0.2 mM $MgCl_2$)

A 10 µM solution of V5Ang-II in buffered water (30 mM TRIS, pH 7.5, 0.2 mM $MgCl_2$) was titrated with a 2.0 mM solution of **CyGlu** that contains V5Ang-II ([V5Ang-II = 10 µM). using λ_{ex} = 276 nm.

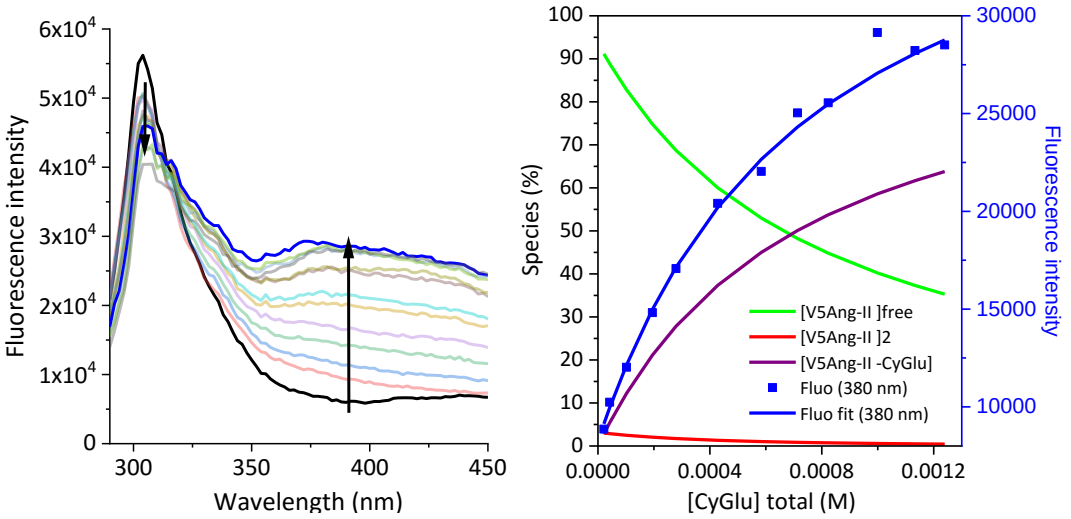


Figure 5.45. Emission spectra of V5Ang-II (10 µM) in buffered water (30 mM TRIS, pH 7.5, 0.2 mM $MgCl_2$) in the presence of different concentrations of CyGlu ([CyGlu = 0-1.2 mM) using λ_{ex} = 276 nm (left). Fitting of the titration of V5Ang-II 10 µM) with CyGlu at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model proposed for the fitting and the log β of the different equilibrium present were:

$A + A \leftrightarrows A_2$ log β = 3.56 (constant)

$A+B \leftrightarrows AB$ log β = 2.848 ± 0.007 (refine)

Where A corresponds to V5Ang-II and B to **CyGlu.**

Titration of V5Ang-II by CyHis

A 1 µM solution of V5Ang-II in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 3.2 mM solution of **CyHis** that contains V5Ang-II ([V5Ang-II] = 1 µM). using λ_{ex} = 276 nm.

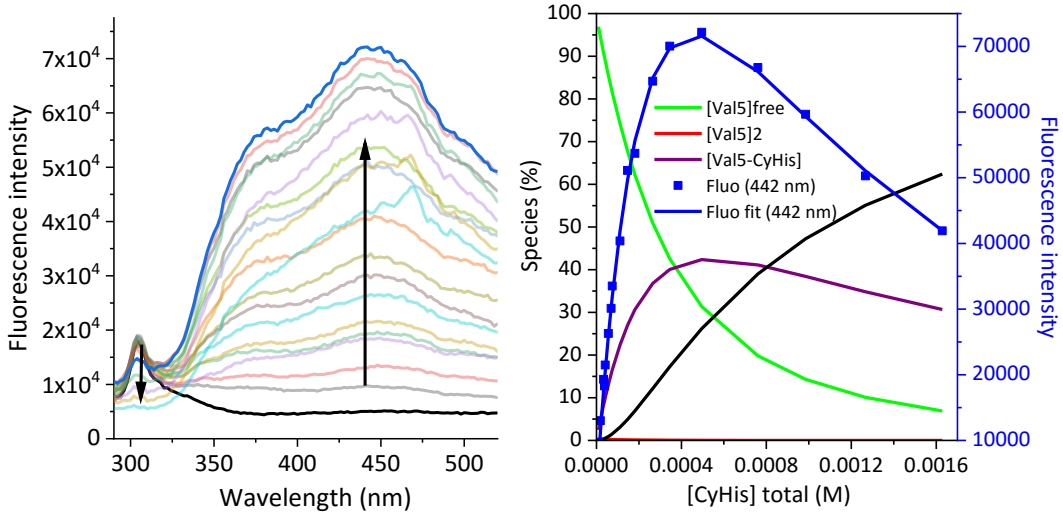


Figure 5.46. Emission spectra of V5Ang-II (1 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyHis ([CyHis=0-1.6 mM) using λ_{ex}: 276 nm. (left) Fitting of the titration of V5Ang-II (1 µM) with CyHis at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β = 3.56 (constant)

$A+B \leftrightarrows AB$ log β = 3.44 ± 0.013 (refine)

$A+2B \leftrightarrows AB_2$ log β = 6.53 ± 0.014 (refine)

Where A corresponds to V5Ang-II, B to **CyHis**. The dimerization of V5Ang-II had already been studied in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The best model for the data fitting includes the formation of both complexes AB and AB_2. The log β for the formation of these two complexes were successfully determined with HypSpec with the model proposed and a low deviation between the experimental data and the data derived from the model. In Figure 5.46 we can see the importance of the formation of the complex AB2 which coexists with the AB complex in in most of the concentrations range studied.

Titration of V5Ang-II by CyLys

A 10 µM solution of V5Ang-II in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 3.2 mM solution of **CyLys** that contains V5Ang-II ([V5Ang-II]= 10 µM). λ_{ex} = 276 nm.

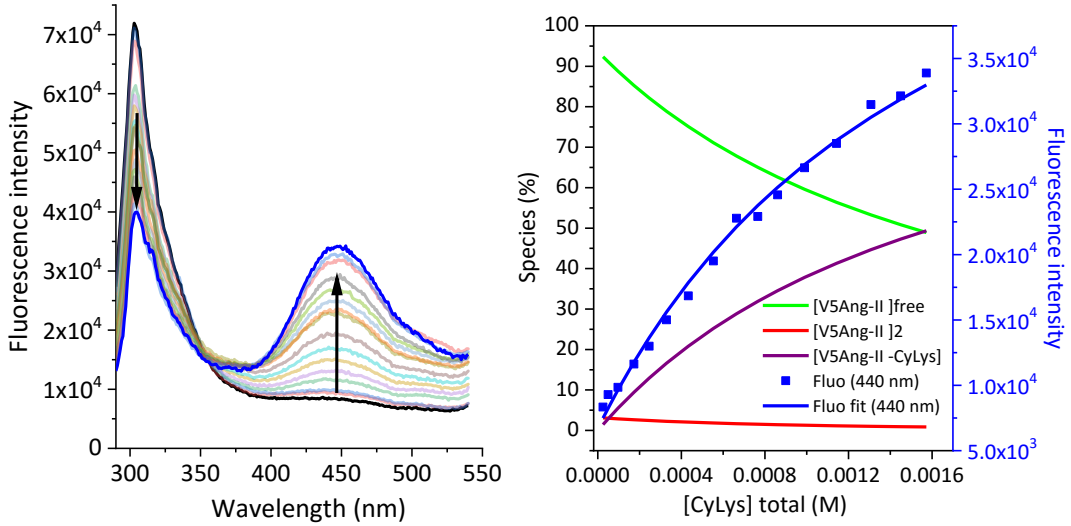


Figure 5.47. Emission spectra of V5Ang-II (10 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyLys ([CyLys=0-1.6 mM) using λ_{ex}: 276 nm (right). Fitting of the titration of V5Ang-II (10 µM) with CyLys at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (left).

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β = 3.56 (constant)

$A + B \leftrightarrows AB$ log β = 2.80 ± 0.01 (refine)

Where A corresponds to V5Ang-II, B to **CyLys**. The dimerization of V5Ang-II had already been studied in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The association constant for the formation of the complex AB was successfully determined with HypSpec.

Titration of V5Ang-II by CyOrn

A 10 µM solution of V5Ang-II in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 3.2 mM solution of **CyOrn** that contains V5Ang-II ([V5Ang-II] = 10 µM). λ_{ex} = 276 nm.

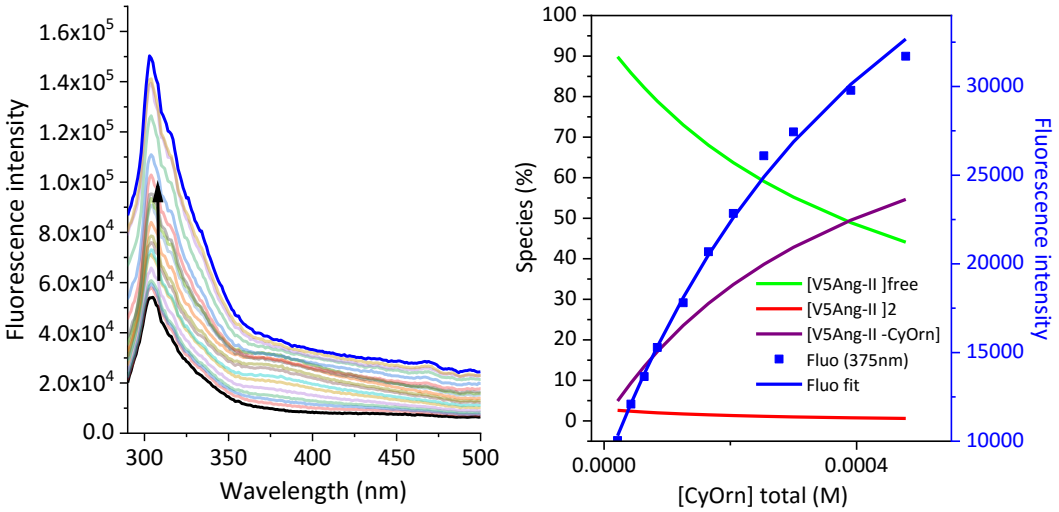


Figure 5.48. Emission spectra of V5Ang-II (10 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyOrn ([CyOrn=0-0.5 mM) using λ_{ex}: 276 nm (left). Fitting of the titration of V5Ang-II (10 µM) with CyOrn at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

The equilibrium model used for the fitting and the log β obtained were:

$A + A \leftrightarrows A_2$ log β = 3.56 (constant)

$A + B \leftrightarrows AB$ log β = 3.419 ± 0.003 (refine)

Where A corresponds to V5Ang-II, B to **CyOrn**. The dimerization of V5Ang-II had already been studied in the same conditions of this titration, so the dimerization constant of the peptide was introduced in this model as constant. The association constant for the formation of the complex AB was successfully determined with HypSpec.

Titration of E_4YM_4 by CyAsp

A 10 µM solution of E_4YM_4 in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 2 mM solution of **CyAsp** that contains E_4YM_4. ([E_4YM_4] = 10 µM). λ_{ex} = 276 nm.

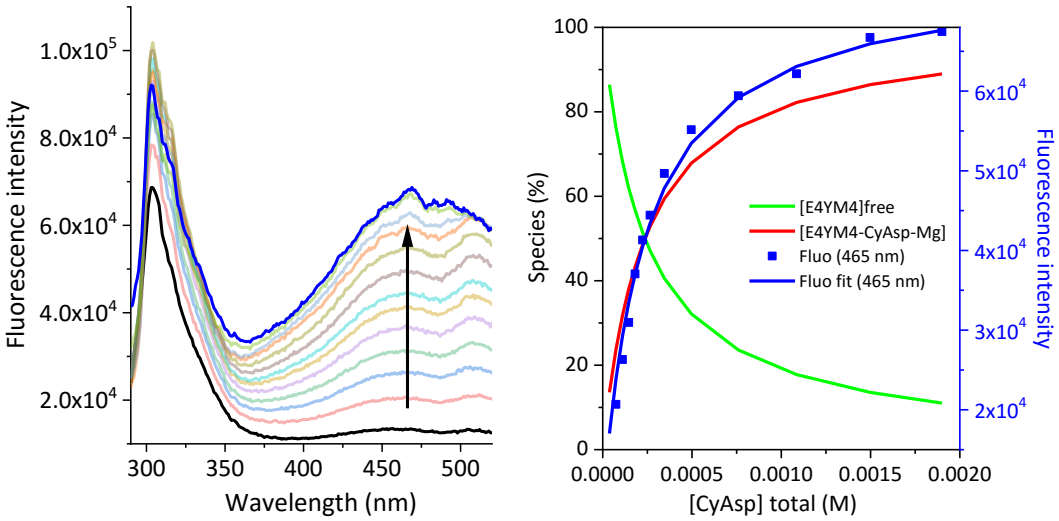


Figure 5.49. Emission spectra of E_4YM_4 (10 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyAsp ([CyAsp = 0-1.8 mM) λ_{ex} = 276 nm.

The equilibrium model used for the fitting and the log β obtained were:

B + C ⇆ BC log β = 0.95 (constant)

A + B + C⇆ AB log β = 5.406 ± 0.003 (refine)

Where A is E_4YM_4, B is **CyAsp** and C is Mg^{2+}.

Titration of E_4YM_4 by CyHis

A 10 µM solution of E_4YM_4 in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) was titrated with a 2 mM solution of **CyHis** that contains E_4YM_4. ([E_4YM_4] = 10 µM). λ_{ex} = 276 nm.

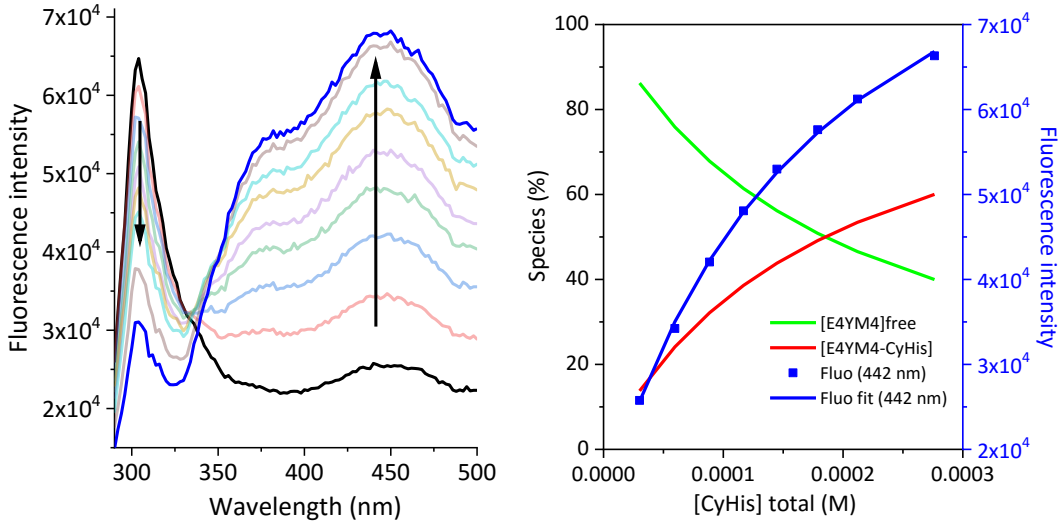


Figure 5.50. Emission spectra of E_4YM_4 (10 µM) in buffered water (30 mM TRIS, pH 7.5, 20 mM $MgCl_2$) in the presence of different concentrations of CyHis ([CyHis=0-0.3 mM) using λ_{ex}: 276 nm (left). Fitting of the titration of E_4YM_4 (10 µM) with CyHis at pH 7.5 (right axis) and species distribution throughout the titration (left axis) (right).

Since E_4YM_4 doesn't form aggregates in the working conditions, the model used for the data fitting in this case was the simplest interaction model:

$$A+B \leftrightarrows AB \quad \log \beta = 3.784 \pm 0.008 \text{ (refine)}$$

Where A corresponds to E_4YM_4 and B to **CyHis.**

5.5.6 NMR titration of CyAsp with $MgCl_2$ and binding constant determination

The titration was performed in D_2O with deuterated HEPES 15 mM at pH 7.3. Stock solution was prepared by weighing the corresponding amount of CyAsp to obtain a 1 mM solution. Additionally, a stock solution of the titrant containing 2 M $MgCl_2$ was prepared by dissolving the salt in the cage stock solution. Hence the concertation of the cage will remain constant during the titration. The stock solution of the cage was introduced into the NMR tube and the ^1H NMR spectrum (400 MHz, 298 K) was acquired, then small volumes of the stock solution of the titrant were added and the ^1H NMR spectrum recorded after each addition.

Different proton signals changed upon the addition of Mg^{2+}. The variation of ArH_1, H_A and H_B were fitted to the simplest model using HypNMR 2008 version 4.0.71 software[7,8]. The following figures (A4.29-30) show the stacked plot of the NMR spectra for this titration experiment, the data set introduced (experimental) and obtained (fit) during the fitting

process, the output value for the log ß, the plot of the experimental (symbols) and the fitted (lines) values of the chemical shifts and the plot of the simulated species distribution obtained. The log ß for the interaction between Mg and **CyAsp** determined was 0.95 which corresponds to a K_{ass} of 8.91 M^{-1}. The interaction between **CyAsp** and Mg is quite weak but since the concentration of Mg is 20 mM in all the kinase assays it will play an important role.

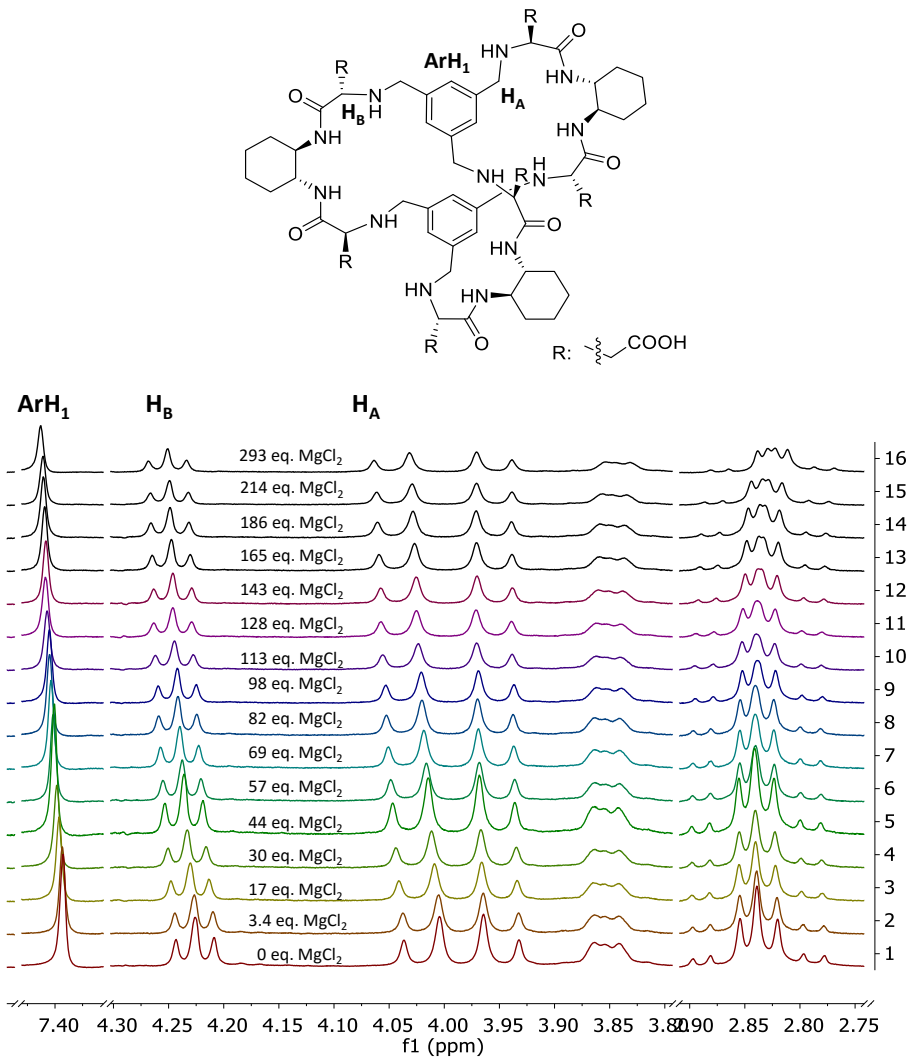


Figure 5.51. Stacked ^{1}H NMR spectra for the titration of CyAsp.

Table 5.3. Data set of CyAsp titration**:**

[Mg]	Mg eq.	ArH	ArH fit	HB	HB fit	HA	HA fit

0.003	3.413	7.393	7.394	4.227	4.227	4.038	4.038
0.017	16.953	7.396	7.397	4.230	4.231	4.041	4.041
0.030	30.315	7.398	7.399	4.233	4.234	4.044	4.045
0.042	43.502	7.401	7.401	4.236	4.236	4.049	4.047
0.055	56.518	7.402	7.402	4.238	4.238	4.049	4.049
0.068	69.367	7.404	7.404	4.240	4.240	4.051	4.051
0.080	82.051	7.405	7.405	4.243	4.242	4.052	4.053
0.095	97.680	7.406	7.406	4.243	4.243	4.053	4.054
0.110	113.063	7.407	7.407	4.245	4.244	4.056	4.056
0.125	128.205	7.409	7.408	4.246	4.246	4.057	4.057
0.140	143.113	7.408	7.408	4.246	4.247	4.058	4.058
0.161	165.046	7.409	7.409	4.248	4.248	4.059	4.059
0.182	186.480	7.411	7.410	4.249	4.249	4.061	4.060
0.209	214.313	7.411	7.411	4.249	4.250	4.062	4.062
0.286	293.040	7.412	7.413	4.254	4.253	4.064	4.064

Results of the HypNMR fitting: Log β = 0.95 ± 0.02

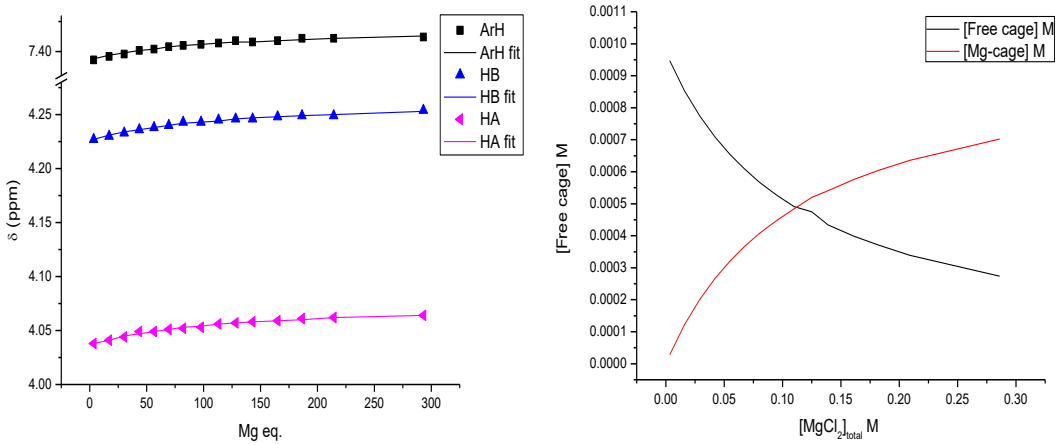


Figure 5.52. Plot of the experimental (symbols) and fitting (lines) data in CyAsp titration (left). Species distribution as a function of the magnesium chloride concentration in CyAsp titration (right).

5.5.7 Molecular modeling

All molecular simulations were carried out with the package Schrödinger Suite 2019,[50] through its graphical interface Maestro.[51] The program Macromodel,[52] with its default force field OPLS3[47] and GB/SA water solvation conditions,[53] was used for energy minimization. Different geometries were generated as starting points for Monte Carlo conformational

searches as implemented in the software, leading to the generation of more than 10000 local minima (1000 structures per each conformational search). The obtained minima were ordered in energy and those within an energy gap of 10 kJ·mol^{-1} were analyzed for each conformational search. The minimum shown in the manuscript corresponds to that of the lowest energy among the minima thus located.

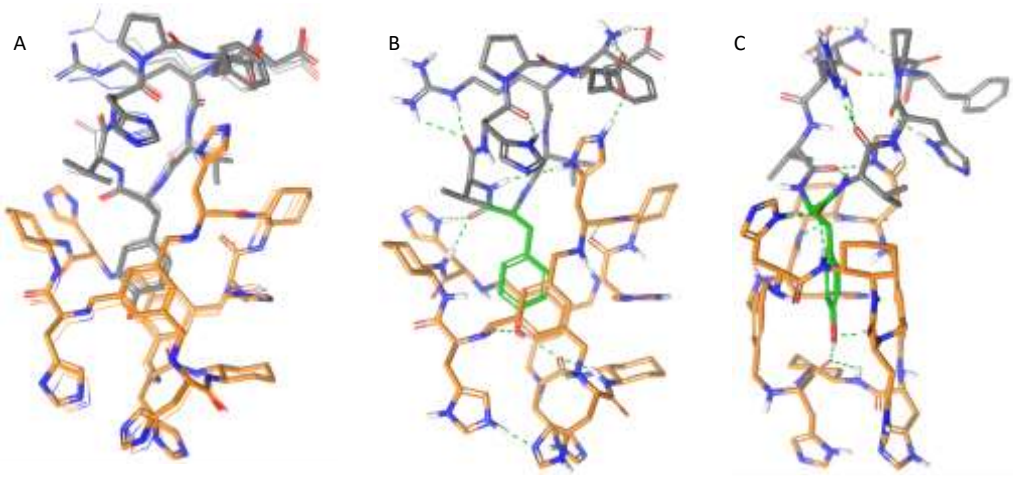


Figure 5.53. (A) Wire representation of overlaid energetically accessible minima (within 10 kJ·mol^{-1}) of the [CyHis-V5Ang-II] complex, with the global minimum as stick plot. C-atoms from the cage are shown in orange and all H-atoms are omitted for clarity. This representation shows some conformational flexibility in the C and N termini of the peptide substrate. (B,C) Front (B) and side (C) views of the stick representation of the global minimum with the Tyr residue of V5Ang-II represented with green C-atoms. Non-polar H-atoms are omitted and H-bonds are shown as green dashed lines. These two representations illustrate the inclusion of Tyr4 within the cage cavity and the establishing of different H-bonding interactions within the supramolecular complex (host-guest, host-host and guest-guest).

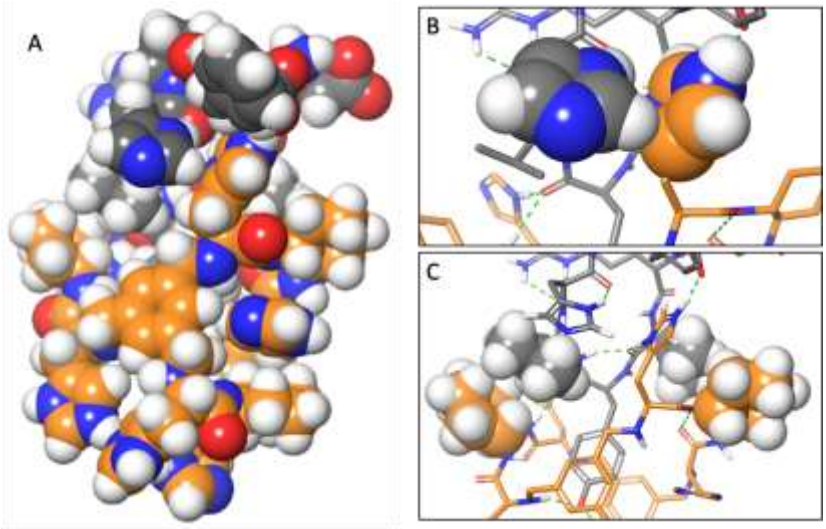


Figure 5.54. (A) Space-filling (CPK) representation of the global minimum for the [CyHis-V5Ang II] complex. C-atoms from the cage are shown in orange. The CPK model illustrates how the OH from the Tyr4 residue of V5Ang-II is fully protected inside the host cavity. (B,C) Details of the edge-to-face imidazole-imidazole (B) and cyclohexane-^{i}Pr (C) host-guest interactions.

5.6 References

1 A. Alonso, J. Sasin, N. Bottini, I. Friedberg, I. Friedberg, A. Osterman, A. Godzik, T. Hunter, J. Dixon and T. Mustelin, *Cell*, 2004, **117**, 699–711.

2 J. J. Witt and R. Roskoski, *Anal. Biochem.*, 1975, **66**, 253–258.

3 C. J. Hastie, H. J. McLauchlan and P. Cohen, *Nat. Protoc.*, 2006, **1**, 968–971.

4 S. R. Hubbard and J. H. Till, *Annu. Rev. Biochem.*, 2000, **69**, 373--398.

5 R. Roskoski, *Pharmacol. Res.*, 2015, 100, 1–23.

6 R. Roskoski, *Pharmacol. Res.*, 2015, **100**, 1–23.

7 C. Pottier, M. Fresnais, M. Gilon, G. Jérusalem, R. Longuespée and N. E. Sounni, *Cancers (Basel).*, , DOI:10.3390/cancers12030731.

8 M. M. Bluhm, G. Bodo, H. M. Dintzis, J. C. Kendrew and M. F. Perutz, *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.*, 1958, **246**, 369–389.

9 F. M. Ferguson and N. S. Gray, *Nat. Rev. Drug Discov.*, 2018, **17**, 353–376.

10 M. E. M. Noble, J. A. Endicott and L. N. Johnson, *Science (80-.).*, 2004, **303**, 1800–1805.

11 H. Krishnan, W. T. Miller, F. J. Blanco and G. S. Goldberg, *Drug Discov. Today*, 2019, **24**, 241–249.

12 R. Roskoski, *Pharmacol. Res.*, 2015, **94**, 9–25.

13 S. J. Parsons and J. T. Parsons, *Oncogene*, 2004, **23**, 7906–7909.

14 L. C. Kim, L. Song and E. B. Haura, *Nat. Rev. Clin. Oncol.*, 2009, **6**, 587–595.

15 N. Yokoyama, J. Lougheed and W. T. Miller, *J. Biol. Chem.*, 2005, **280**, 42219–42226.

16 X. Sun, Y. Wei, P. P. Lee, B. Ren and C. Liu, *Cell. Immunol.*, 2019, **341**, 103919.

17 M. Pollak, *Nat. Rev. Cancer*, 2008, **8**, 915–928.

18 M. Pollak, *Nat. Rev. Cancer*, 2012, 12, 159–169.

19 E. Faggi, Y. Pérez, S. V. Luis and I. Alfonso, *Chem. Commun.*, 2016, **52**, 8142–8145.

20 C. P. Li, Y. X. Lu, C. T. Zi, Y. T. Zhao, H. Zhao and Y. P. Zhang, *Int. J. Mol. Sci.*, 2020, **21**, 1–15.

21 X. Li, T. M. Palhano Zanela, E. S. Underbakke and Y. Zhao, *J. Am. Chem. Soc.*, 2021, **143**, 639–643.

22 E. Faggi, C. Vicent, S. V. Luis and I. Alfonso, *Org. Biomol. Chem.*, 2015, **13**, 11721–11731.

23 G. S. Martin, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 467–475.

24 H. Krishnan, W. T. Miller and G. S. Goldberg, *Genes Cancer*, 2012, **3**, 426–435.

25 V. Jha, M. Macchia, T. Tuccinardi and G. Poli, *Cancers* , 2020, 12.

26 S. Martellucci, L. Clementi, S. Sabetta, V. Mattei, L. Botta and A. Angelucci, *Cancers* , 2020,

12.

27 A. Simatou, G. Simatos, M. Goulielmaki, D. Spandidos A., S. Baliou and V. Zoumpourlis, *Mol Clin Oncol*, 2020, **13**, 21.

28 J. E. B. T.-M. in E. Casnellie, in *Protein Phosphorylation Part A: Protein Kinases: Assays, Purification, Antibodies, Functional Analysis, Cloning, and Expression*, Academic Press, 1991, vol. 200, pp. 115–120.

29 Z. Songyang, K. L. Carraway, M. J. Eck, S. C. Harrison, R. A. Feldman, M. Mohammadi, J. Schlessinger, S. R. Hubbard, D. P. Smith, C. Eng, M. J. Lorenzo, B. A. J. Ponder, B. J. Mayer and L. C. Cantley, *Nature*, 1995, **373**, 536–539.

30 N. Yokoyama and W. T. Miller, *FEBS Lett.*, 1999, **456**, 403–408.

31 N. Yokoyama and W. T. Miller, *Methods Enzymol.*, 2006, **406**, 250–260.

32 T. W. Wong and A. R. Goldberg, *J. Biol. Chem.*, 1983, **258**, 1022–1025.

33 A. Lamiable, P. Thévenet, J. Rey, M. Vavrusa, P. Derreumaux and P. Tufféry, *Nucleic Acids Res.*, 2016, **44**, W449-54.

34 P. Thévenet, Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux and P. Tufféry, *Nucleic Acids Res.*, 2012, **40**, W288-93.

35 Y. Shen, J. Maupetit, P. Derreumaux and P. Tufféry, *J. Chem. Theory Comput.*, 2014, **10**, 4745–4758.

36 P. Gans, A. Sabatini and A. Vacca, *Talanta*, 1996, **43**, 1739–1753.

37 Http://www.hyperquad.co.uk/HypSpec2014.htm, http://www.hyperquad.co.uk/HypSpec2014.htm.

38 C. Frassineti, S. Ghelli, P. Gans, A. Sabatini, M. S. Moruzzi and A. Vacca, *Anal. Biochem.*, 1995, **231**, 374–382.

39 P. Thordarson, *Chem. Soc. Rev.*, 2011, **40**, 1305–1323.

40 H. Schmidbaur, H. G. Classen and J. Helbig, *Angew. Chemie Int. Ed. English*, 1990, **29**, 1090–1103.

41 H. Schmidbaur, H. G. Classen and J. Helbig, *Angew. Chemie*, 1990, **102**, 1122–1136.

42 L. Alderighi, P. Gans, A. Ienco, D. Peters, A. Sabatini and A. Vacca, *Coord. Chem. Rev.*, 1999, **184**, 311–318.

43 N. Zhou, G. J. Moore and H. J. Vogel, *J. Protein Chem.*, 1991, **10**, 333–343.

44 S. Favelyukis, J. H. Till, S. R. Hubbard and W. T. Miller, *Nat. Struct. Biol.*, 2001, **8**, 1058–1063.

45 P. Pellicena, D. S. King, A. M. Falick, J. Kuriyan, M. A. Seeliger, M. Young and M. N. Henderson, *Protein Sci.*, 2005, **14**, 3135–3139.

46 E. Faggi, A. Moure, M. Bolte, C. Vicent, S. V. Luis and I. Alfonso, *J. Org. Chem.*, 2014, **79**, 4590–4601.

47 E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, *J. Chem. Theory Comput.*, 2016, **12**, 281–296.

48 M. Z. Cabail, S. Li, E. Lemmon, M. E. Bowen, S. R. Hubbard and W. T. Miller, *Nat. Commun.*, , DOI:10.1038/ncomms7406.

49 L. Serrano, J.-L. Neira, J. Sancho and A. R. Fersht, *Nature*, 1992, **356**, 453–455.

50 Schrödinger Release 2019-1, Schrödinger, LLC: New York, NY, 2019.

51 Schrödinger Release 2019-1: Maestro, Schrödinger, LLC: New York, NY, 2019.

52 Schrödinger Release 2019-1: Macromodel, Schrödinger, LLC: New York, NY, 2019.

53 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.

GENERAL CONCLUSIONS

Through a deep analysis of the results obtained from the research conducted in this Thesis, we can conclude that:

Small pseudopeptidic cages

1. Small pseudopeptidic cages can be used for the selective killing of cancer cells using the typical reverse-pH present in tumour microenvironments as the selectivity feature. The transport capacity studied in POPC vesicles of these compounds can be well correlated with their cytotoxicity at different pH. This is in agreement with the transport and cell killing mechanisms proposed that involve the transport of HCl through cell membranes disrupting cell homeostasis and leading cell death.
2. Moreover, the introduction of F-atoms in different positions of the side chains of these small cages allows the tuning of their pK_a and effective lipophilicity, which have been associated with their pH-dependent HCl transport and cytotoxicity. The introduction of an F-atom substituent reduces the basicity of the cage changing its protonation state at physiologically relevant pH thus tuning their transport capacity and cytotoxicity to be pH-dependent within the pH range that differentiates a normal cell form a cancer microenvironment. Through the systematic study of different fluorinated cages, we have shown that subtle modifications in the cage at positions far from the chloride binding site have a strong effect in their properties and their cytotoxicity. This characteristic can become a potent tool to easily modulate the biological activity of chloride anionophores without modifying the binding pocket and their affinity for the targeted ion.

Large pseudopeptidic cages

3. Large flexible pseudopeptidic cages with a conserved pocket can be used for the selective encapsulation of Tyr residues from random copolymers. The amino acids used for the synthesis of the cage determine the strength of the interaction with the tyrosine through secondary interactions with the substrate amino acids sequence. Electrostatic interactions play a very important role in determining the affinity of the cage for a specific tyrosine. Thus, those cages containing positively

charged amino acid residues interact better with tyrosine surrounded by negative charges at physiological pH. The cage made from His is especially interesting, since the amphoteric nature of the imidazole ring promotes the efficient binding of peptides with the Tyr surrounded either by basic or acidic amino acids.

4. The supramolecular recognition of Tyr residues in polypeptides constitutes a powerful method to prevent the TK-mediated phosphorylation of the substrates. The secondary interactions between the peptide side-chains and the cage modulate the binding and the inhibitory activity complementing the inclusion of the tyrosine within the cage cavity. The reported kinase inhibition approach paves the way towards the selective modulation of an individual kinase signalling pathway, without interfering with other functions of the same kinase potentially leading to the development of better tools for research, diagnosis or therapy in biomedicine.

The results from the present work represent an important contribution to broaden the potential of pseudopeptidic cages for different applications in biological systems. Regarding the use of cages as chloride transporters, we have successfully presented an alternative selectivity mechanism for cancer cells using pH-dependent Cl^- transporters whose activity is closely related to their pK_a. As for the second part of the project we have demonstrated the possibility of selective inhibition of tyrosine kinases by targeting their corresponding substrates. Overall we have shown the efficient modulation of the supramolecular properties of cage-like receptors by changing the nature of the precursor amino acid while keeping intact the core defining the inner binding site for each substrate (chloride or tyrosine, respectively). Thus, the obtained results underscore the success of the pseudopeptidic design. The fine tuning of the supramolecular properties is translated into potent and selective biological activities. Moreover, we have shown this concept in two very different chemical biology systems, leading to two new applications of pseudopeptidic cages with a huge potential for the future development of clinical tools based on the selective encapsulation of biomolecules and ions.