UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN

"Scienza e Tecnologia dei Materiali"

CICLO XXXI

Molecular recognition, metal coordination and self-assembly with macrocyclic receptors and multidentate pyridine ligands.

Coordinatore:

Prof. Enrico Dalcanale

Tutore:

Prof. Chiara Massera

Co-tutore:

Prof. Enrico Dalcanale

Dottorando: Francesca Guagnini

Anni 2015/2018

To my family, who always supported me

To Andreas, Anjali and Jonathan,

who turned these 3 years into an adventure

Table of Contents

The purpose of this thesis

Supramolecular chemistry was defined for the first time by J. M. Lehn as "*the chemistry beyond molecules*". It studies molecular assemblies driven by non-covalent interactions where molecules assemble by spontaneous thermodynamically-driven processes. Being based on non-covalent interactions, supramolecular systems are reversible and highly adaptable. Changes can often be triggered by external stimuli, yielding controllable and tunable systems. Thus, tailored supramolecular systems can exploit a variety of functions, for example mimic complex biological processes. For these reasons, supramolecular chemistry is an extremely active and fast developing field that finds applications from life sciences to smart materials.

Molecular recognition, self-assembly and metal coordination are common strategies to achieve sophisticated supramolecular architectures. Hence, in this thesis, I report three case studies that involve the synthesis of supramolecular building blocks and the study of their assemblies. In the first chapter, the synthesis of multidentate pyridine ligands is shown. Such ligands are employed to construct assemblies in the solid state and to prepare coordination polymers. The second chapter comprises the synthesis and study of metal complexes obtained from the macrocyclic tetra-phosphonate cavitands. In the third chapter, three macrocyclic receptors are employed in the molecular recognition of a model protein. In this case, surface recognition is shown to drive protein assembly in the solid state.

CHAPTER 1

Supramolecular assembly of multidentate pyridine ligands *via* metal coordination, hydrogen bond and halogen bond*

*Part of this work has been carried out in collaboration with Prof. Lara Righi (University of Parma, Parma), Dr. Mauro Gemmi (Italian Institute of Technology, Pisa) and Dr. Arianna Lanza (Italian Institute of Technology, Pisa).

1. Pyridine based ligands: cages, assembly and molecular recognition

Self-assembly is the most used strategy to generate sophisticated architectures, from discrete cages to 3-D networks. Weak interactions and pre-organization are the two paradigms driving self-assembly of natural and synthetic molecules. Hence, ligands with functionalities suitable for both metal coordination and weak interactions are interesting building blocks for generating new materials. In this context, ligands bearing pyridines proved to be extremely useful. Pyridines can interact with metals but are also excellent hydrogen and halogen bond acceptors. Planarity is another crucial feature, that usually is achieved through aromatic conjugated systems. The presence of aromatic rings enables π-π stacking, C-H-π and C-F-π interactions.

Metal coordination driven self-assembly represents a powerful tool for the obtainment of self-assembled cages and capsules. Supramolecular cages generate confined environments, where guests can be loaded, and chemical reactions can be performed. Rigid planar ligands, characterized by a specific angle, have been preferred for the development of cages (Figure 1).[1–5] Their pre-organization facilitates the design and prediction of the final coordination geometry.[6]

James Crowley and co-workers employed a multidentate pyridine ligand (Figure 1.A) to develop a silver (I) six-member ring metallocycle (Figure 1.B).[2] The assembly was very promising to generate nanotubes and columnar architectures. Nevertheless, due to the formation of coordination polymers and to the instability of the product, the authors focused on Pd(II) coordination cages and metallocycles. Indeed in this field, Pd (II) and Pt (II)

M_2L_4 coordination cages have been widely studied by several groups (Figure 1.C).[1,3,4,7–11] Pd(II) cages showed good complexation abilities towards guests with the right size and electrostatic complementarity.[1] Furthermore, the complexation behaviour could be tuned playing with experimental conditions like solvent and counter ions.[4]

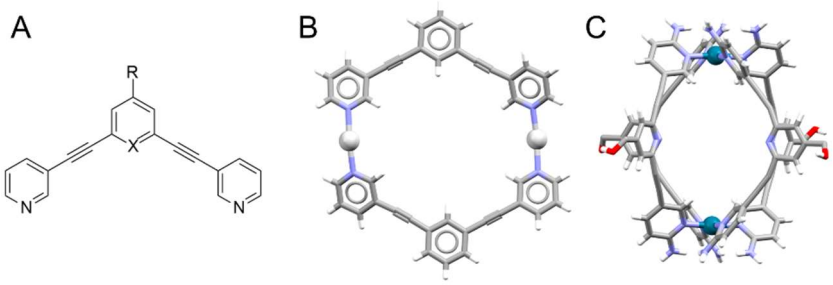


Figure 1. (A) Pyridine ligand employed in the preparation of cages and metallocycles. X = CH/N. (B) Metallocycle obtained from silver triflate and a pyridine ligand (CCDC number 795011). Ag^+ cations are represented as white spheres. In this case, the obtainment of a discrete structure over a coordination polymer was dictated by the choice of the counter-ion. (C) An example of $Pd_2L_4^{2+}$ cage (CCDC number 1439952). Pd^{2+} cations are represented as teal-blue spheres. Solvent molecules and counter-anions are omitted for clarity.

The behaviour of $Pd_2L_4^{4+}$ cages can be tuned and controlled by design and chemical stimuli. For example, kinetic stability can be enhanced by accurate choice of the substituents on the ligand.[7] Functionalization of the cages for further application was achieved via click chemistry[12] or amide bond formation.[10] Furthermore, interconversion of the cage assembly to metallocycle triggered by Cl^- anions was reported.[13] Thanks to their properties, $Pd_2L_4^{4+}$ cages are actively investigated as drug delivery

systems[8–10] and as catalysts.[11] Multidentate pyridine ligands have also been employed to generate cages and metallocycles with multiple cavities (Figure 2).[5,14]

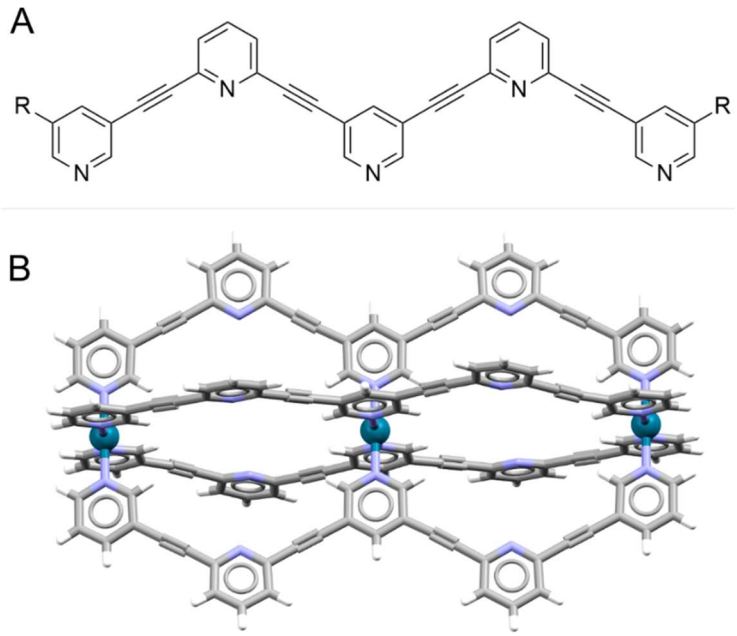


Figure 2. A multidentate pyridine ligand (A) was exploited to obtain a di-cavities palladium cage $Pd_3L_4^{6+}$ (B) (CCDC number 1517673). A tri-cavities cage $Pd_4L_8^{8+}$ was prepared in similar conditions and characterized in solution. Both cages were shown to bind to cisplatin.[14]

Metal coordination is not the only tool to exploit pyridine building blocks properties. Pyridine-based synthons can also be assembled *via* weak interactions or covalent bonds. The Shimizu group synthesized a pyridyl-bis-urea macrocycle that assembles into columnar structures.[15,16] Despite the lack of pores, the assembly could uptake guests like TFE, I_2, H_2 or CO_2. Furthermore, the columnar assemblies have been shown to absorb and

organize small guest molecules that interact via X-bond.[17] The same group reported the preparation of a wider bis-urea macrocycle, based on a pyridyl-acetylene scaffold (Figure 3.A and 3.B).[18] In this case, the columnar assembly (Figure 3.C) acted as a nanoreactor for the polymerization of isoprene.[18]

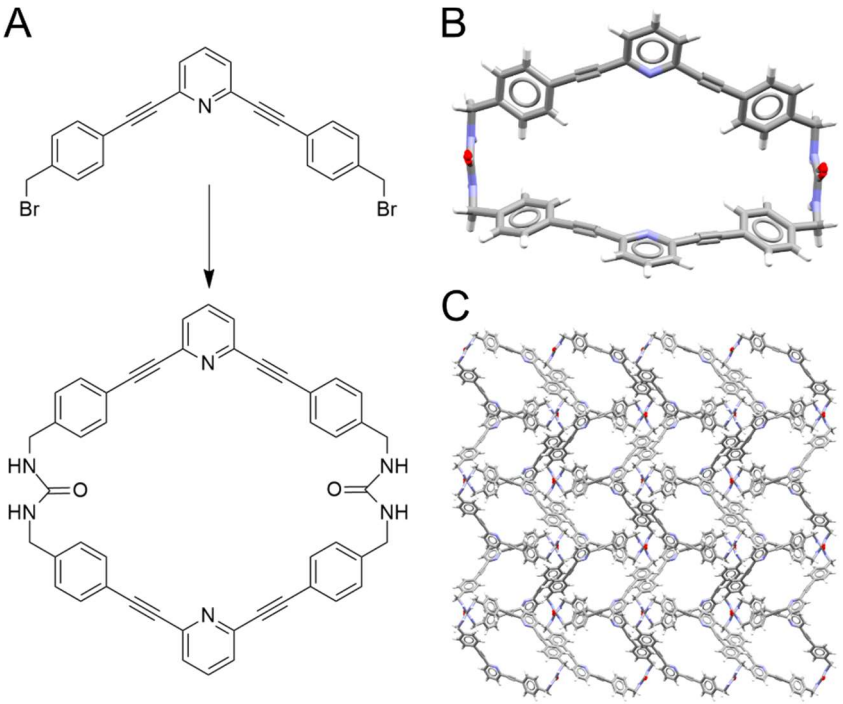


Figure 3. (A) The Shimizu group obtained a pyridyl-acetylene bis-urea macrocycle by coupling of a pyridyl-acetylene building block. The crystal structure of the macrocycle showed a bent conformation (B) that assembled into columnar architectures (C).

Pre-organization is a key feature also in guest binding. Johnson and co-workers synthesised a library of inherently fluorescent pyridine-based urea receptors (Figure 4.A).[19–24] Their receptors were successfully employed

in the sensing of Cl^- and other anions (Figure 4.B).[20,24] In their numerous reports, they suggest how water molecules and Cl^- anions interact in a similar fashion with pyridinic nitrogen.[19]

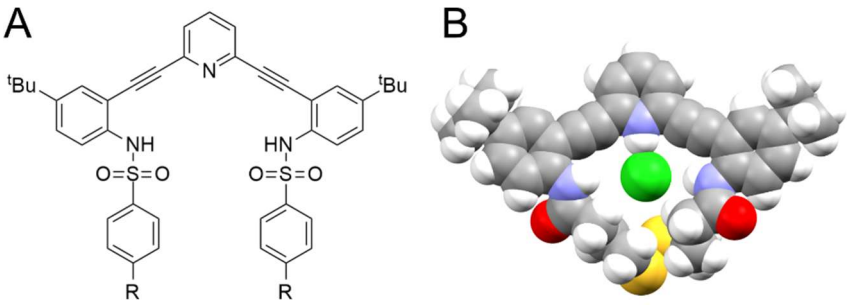


Figure 4. Inherently fluorescent pyridine-urea receptors (A) and, once protonated, the Cl^- complex (B). The molecule is represented as space fill. Chlorine anion is green.

Berryman and co-workers suggested a possible strategy to pre-organize pyridine-acetylene receptors via intramolecular hydrogen bond-halogen bond (HB-XB).[25] The resulting receptor (Figure 5) was more planar and showed a 9-fold increase in binding to halides (I^- and Br^-).[25]

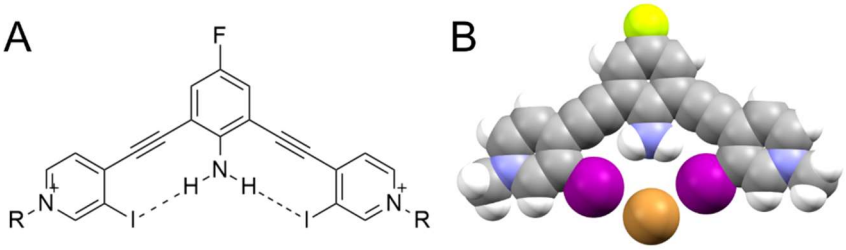


Figure 5. (A) 1,3-bis(4-ethynylpyridinium) receptors were pre-organized *via* B-XB and complex halides. (B) 1,3-bis(4-ethynylpyridinium) receptor binding to Br^-. The molecule is represented as space fill. Bromine anion is brown.

1.1. The design of new pyridine ligands

With these considerations in mind, we designed and synthesized a series of multidentate pyridine ligands (Figure 6). Our ligands are characterized by a pyridyl-acetylene scaffold comprising three aromatic rings connected by triple bonds. Functionalization of the rings (Figure 6) endows the ligands with different properties.

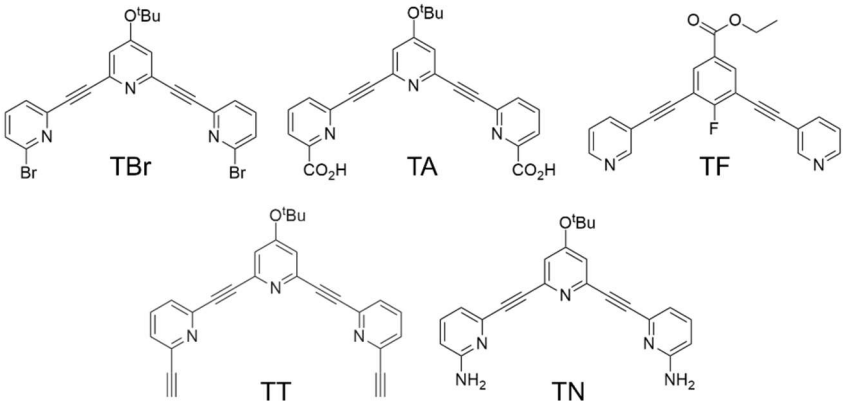


Figure 6. Pyridine ligands prepared for the present work.

These ligands can be combined *via* covalent or hydrogen bond to yield shape persistent macrocycles. **TA** and **TF** are also suitable for metal coordination either through the carboxylic moieties (**TA**) or through the exposed pyridine moieties (**TF**). Depending on the rotation of the pyridine rings around the triple bonds, the three-pyridine ligands can adopt three different conformations: W, S or C (Figure 7).

Functionalization of the ligands in para position on the central pyridine guarantees solubility in most common organic solvents.

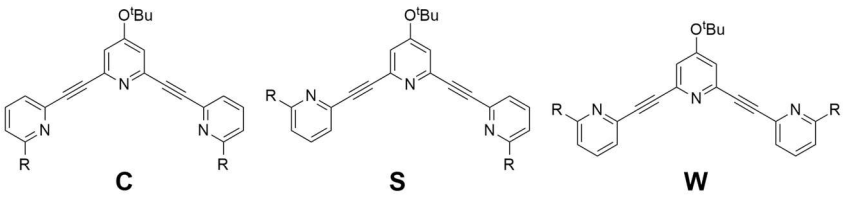
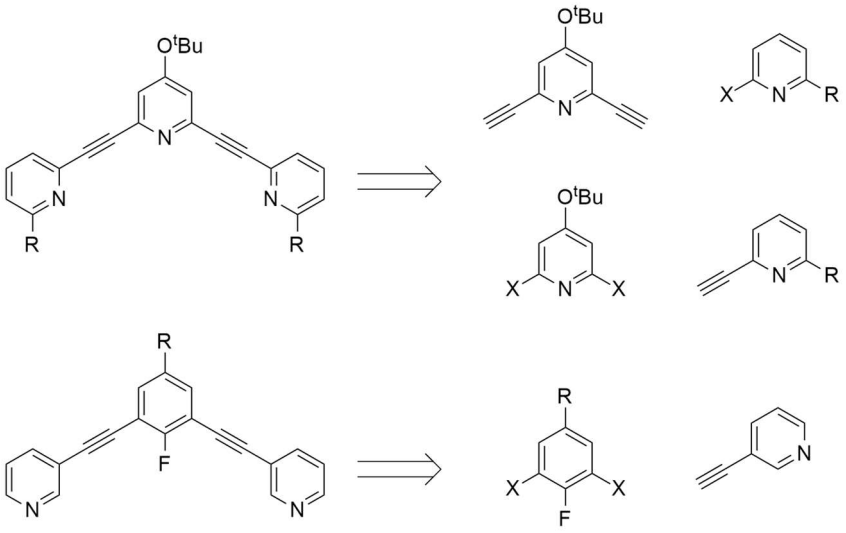


Figure 7. Different conformations that **TBr**, **TA**, **TT** and **TN** can adopt by rotation around the acetylene bonds.

We then started to pursue a systematic study on the solid state structures of these compounds through X-ray diffraction on single crystals, in order to understand their behaviour and potential applications (paragraphs 3-5).

2. Synthesis

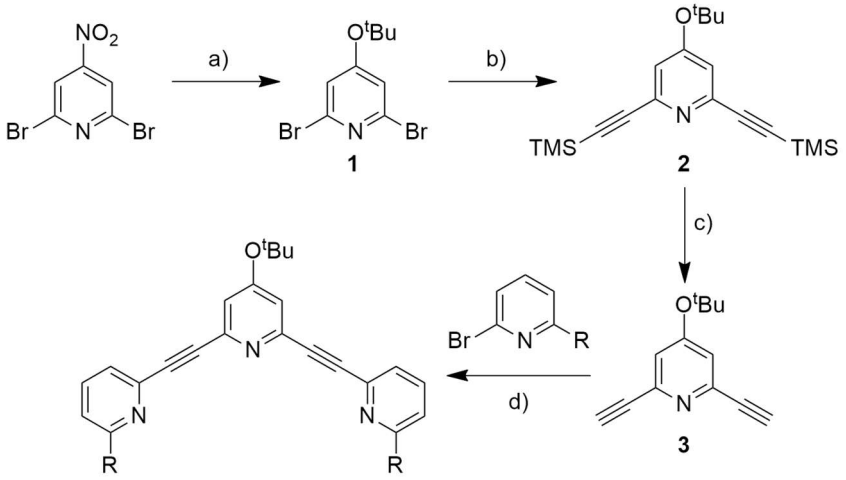
Pyridine ligands have been obtained following the reported retrosynthetic strategy (Scheme 1). The disconnection of the triple bonds led to pyridinic synthons (Scheme 1). Functionalization was achieved by changing the substituents on the starting pyridine.



Scheme 1. Retrosynthesis for **TBr**, **TA**, **TT** and **TN** (top panel) and for **TF** (bottom panel). The first disconnection led to synthons for Sonogashira coupling. The functional group for the coupling, halogen and triple bond, could be either on the axial or the central pyridine.

TA and **TBr** share a similar synthetic route (Scheme 2). The ligands were constructed starting from the central pyridine. The synthesis started from commercially available 2,6-dibromo-4-nitropyridine. The first step was the aromatic nucleophilic substitution of the nitro group with a *tert*-butoxy group.[26] The reaction showed high selectivity for the para position

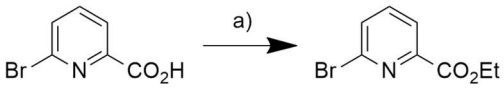
of the pyridine over the bromines in 2 and 6. The presence of only a small amount of di-substituted by-product was detected. The insertion of silyl-trimethyl protected triple bonds was achieved *via* Sonogashira coupling. The deprotection of the acetylene functionalities was performed in the presence of potassium fluoride and exploited the high affinity between Si and F^-.



Scheme 2. General synthesis of **TBr** and **TA** with R = Br, CO_2Et respectively. a) KO^tBu, THF, r.t, 5h, 69%; b) trimethylsilylacetylene, $Pd(PPh_3)_2Cl_2$, PPh_3, CuI, THF, NEt_3, r.t, 6h, 54%; c) KF, THF, MeOH, r.t, 1h, 75%; d) $Pd(PPh_3)_2Cl_2$, PPh_3, CuI, THF, NEt_3, 45-55°, 18h, 40-80%.

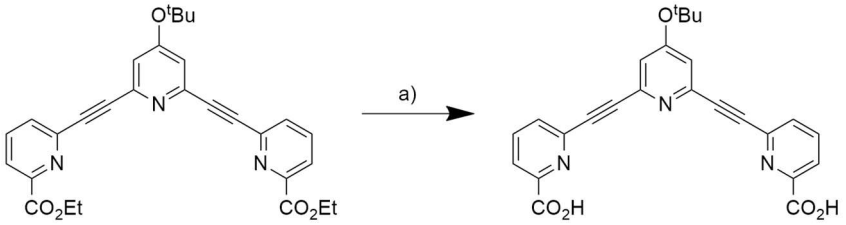
Intermediate **3** was reacted with either ethyl 6-bromopicolinate or 2,6-dibromopyridine in a Sonogashira coupling to yield **TA** and **TBr** respectively. The reaction to obtain **TBr** gave a high amount of impurities. Nevertheless, tetramers or in general higher oligomers could not be isolated. In the case of **TA**, protection of the carboxylic acid moieties, as ethyl esters, was necessary to perform the coupling. Hence, 6-

bromopicolinic acid was heated to reflux in ethanol in the presence of concentrated sulfuric acid (Scheme 3).



Scheme 3. Protection of the acid functionality as ethyl esters. a) EtOH, H_2SO_4, 80°, 24h, 85%.

TA was therefore obtained protected as ethyl ester. Deprotection was achieved *via* basic hydrolysis in the presence of lithium hydroxide (Scheme 4).

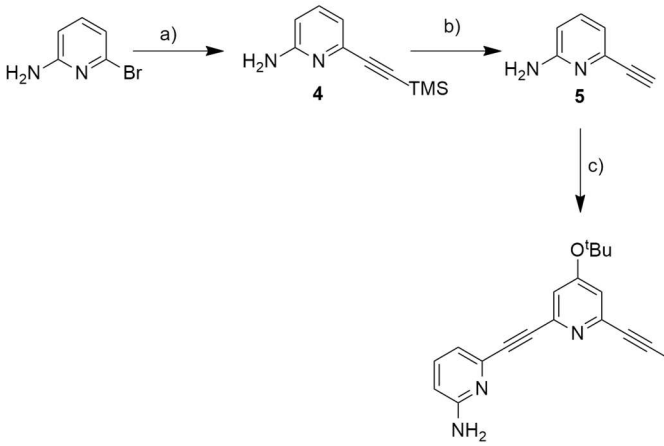


Scheme 4. Deprotection of **TA**. a) Li(OH)·H_2O, THF, H2O, r.t, 30', quant.

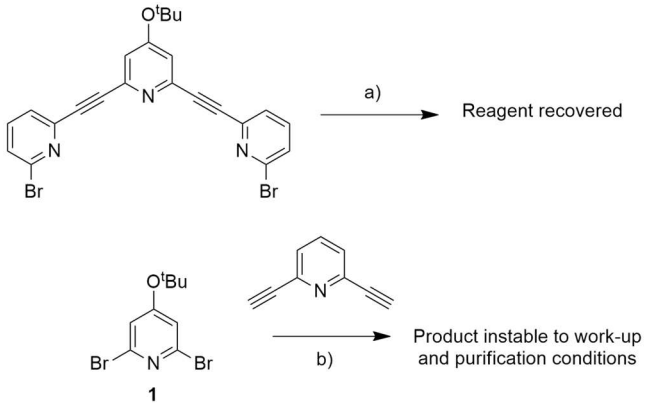
TN was obtained *via* a coupling in which the acetylene function was on the axial pyridine (Scheme 5). After several trials, this alternative synthetic route was chosen because it provided better overall yields. 2-amino-6-bromopyridine was reacted with trimethylsilylacetylene in a Sonogashira coupling to yield product **4**. Deprotection of triple bonds was achieved similarly to product **3**. Coupling between intermediate **5** and **1** gave **TN**.

The synthesis of **TT** was more challenging as it required more synthetic steps. Direct functionalization of **TBr** with trimethylsilylacetylene did not yield the desired product (Scheme 6).

Moreover, reaction of product **1** with 2,6-diethynylpyridine led to a crude product instable during work-up and purification conditions (Scheme 6).

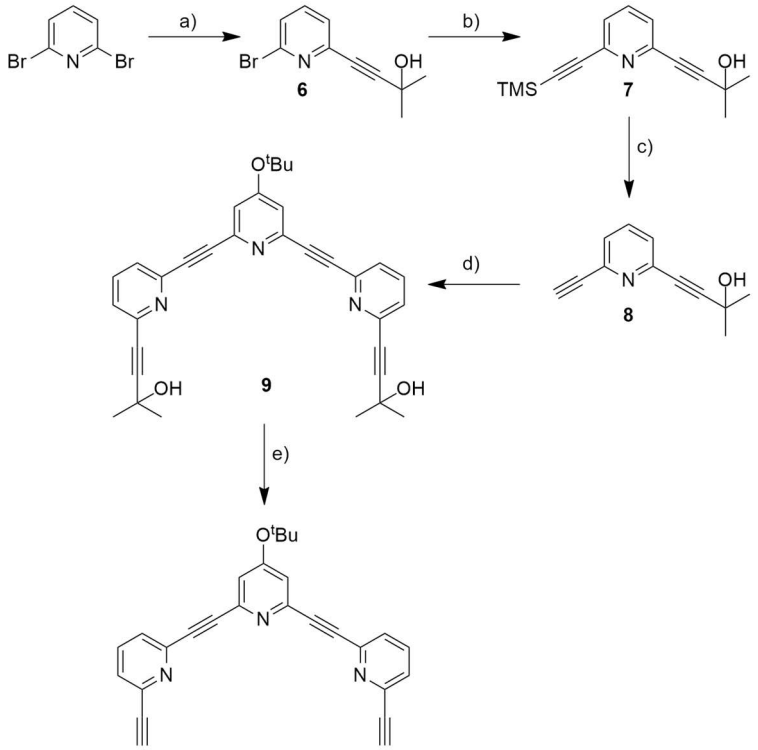


Scheme 5. Synthesis of **TN**. a) trimethylsilylacetylene, $Pd(PPh_3)_2Cl_2$, PPh_3, CuI, THF, NEt_3, 55°, 2h, 89%; b) KF, THF, MeOH, r.t, 2h, 46%; c) **1**, $Pd(PPh_3)_2Cl_2$, PPh_3, CuI, THF, NEt_3, 60°, 24h, 87%.



Scheme 6. Failed strategies in the synthesis of **TT**. a) trimethylsilylacetylene, $Pd(PPh_3)_2Cl_2$, PPh_3, CuI, NEt_3, r.t, 18h; b) $Pd(PPh_3)_2Cl_2$, PPh_3, CuI, NEt_3, 55°, 18h.

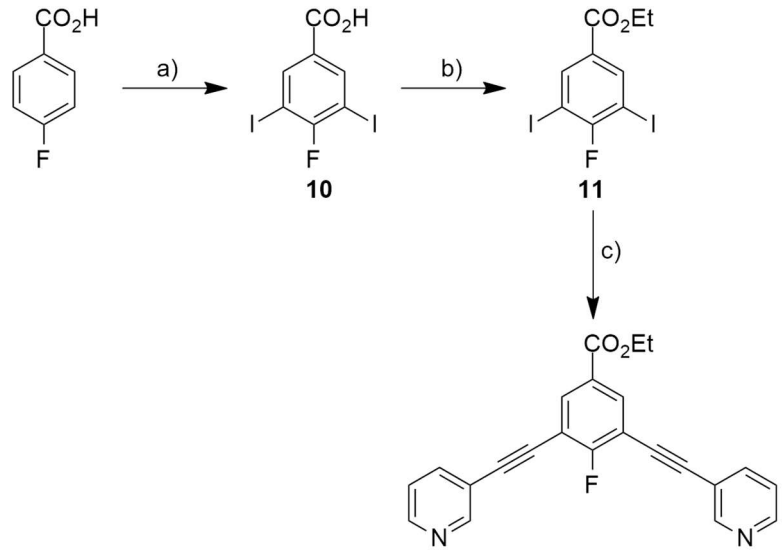
Orthogonal protection of acetylene moieties (Scheme 7) proved to be the successful strategy. 2,6-dibromopyridine was reacted firstly with 2-methylbut-3-yn-2-ol in stoichiometric amounts and then with trimethylsilylacetylene, yielding di-protected **7**. The two protecting groups are orthogonal and selective deprotection of only one triple bond was achieved in the presence of KF. Sonogashira coupling between **8** and **1** gave protected **TT**. Deprotection *via* basic hydrolysis was performed in the presence of sodium hydroxide.



Scheme 7. Synthesis of **TT**. a) 2-methylbut-3-yn-2-ol, $Pd(PPh_3)_2Cl_2$, PPh_3, CuI, NEt_3, r.t, 20h, 63%; b) trimethylsilylacetylene, $Pd(PPh_3)_2Cl_2$, PPh_3, CuI,

NEt_3, r.t, 20h, quant.; c) KF, THF, MeOH, r.t, 1h, 47%.; d) **1**, $Pd(PPh_3)_2Cl_2$, PPh_3, CuI, NEt_3, 55°, 16h, 34%; e) NaOH, toluene, 100°, 24 h, 80%.

TF was obtained in 3 steps from commercially available 4-fluorobenzoic acid (Scheme 8). Iodination of 4-fluorobenzoic acid in ortho to the fluorine was achieved in the presence of molecular iodine and chromium (III) oxide. Protection of the carboxylic acid as ester was performed under acid conditions, in toluene and ethanol. **TF** was obtained by Sonogashira reaction from intermediate **11** and 3-ethynylpyridine.



Scheme 8. Synthesis of **TF**. a) Cr_2O_3, I_2, H_2SO_4, r.t, 18h, quant.; b) EtOH, H_2SO_4, Toluene, reflux, 4h, quant.; c) 3-ethynylpyridine, $Pd(PPh_3)_4$, CuI, NEt_3, 80°, 24h, 14%.

3. Self-assembly of multidentate pyridine ligands: the role of solvent and substituents

Single crystals suitable for X-ray diffraction were obtained for **TBr**, **TN** and **TA**, for which it was thus possible to analyse in detail the solid-state assembly.

Crystals of **TBr** were grown by slow evaporation of different solvents, namely toluene, benzene, chloroform and methanol. The first three solvents were not included in the final structure, but they nevertheless influenced the polymorphic forms obtained, namely **TBrT**, **TBrB** and **TBrC**, respectively. Crystallization from methanol led to a solvated form, indicated as **TBrM**.

In the molecular structure of **TBrT** (Figure 8) the conformation of the pyridine rings around the triple bonds is C (Figure 7). The angles between the mean planes passing through the central ring (acetylene moieties included) and the two lateral pyridine fragments C8-C12/N2/Br1 and C15-C19/N3/Br2 are 55.85(6)° and 50.78(4)°, respectively. These lateral pyridine fragments are at 87.99(6)° one with respect to the other. In the crystal structure, the **TBr** molecules pack through a series of C-H···N and π-π stacking interactions (Figure 9). More precisely, each **TBr** acts both as donor and acceptor towards a symmetry-related molecule ii (ii = x-1, y, z), forming four interactions [C11-H11···N2, 3.328(5) Å and 170.78(2)°; C12-H12···N1, 3.534(4) Å and 149.70(4)°; C18-H18···N3, 3.324(2) Å and 167.74(8)°; C19-H19···N1, 3.508(4) Å and 155.92(8)°]. On the other side, π-π stacking is present between the pyridine rings C15-C19/N3/Br2 and C8-C12/N2/Br1 (in position 1+x, 1/2-y, -1/2 + z) (Figure 9). The main geometrical parameters are the two distances centroid···C8 of 3.526(5) Å and centroid···C15 of 3.529(6) Å. These sets of

interactions contribute to the overall packing of the molecule, which is very compact without accessible voids (Figure 10).

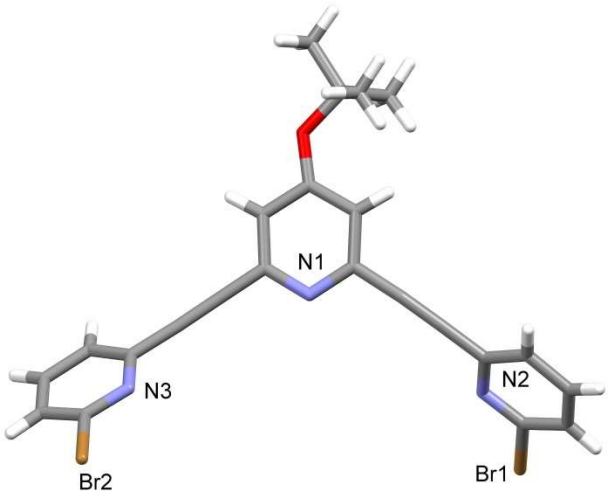


Figure 8. Molecular structure of **TBrT** with partial labelling scheme.

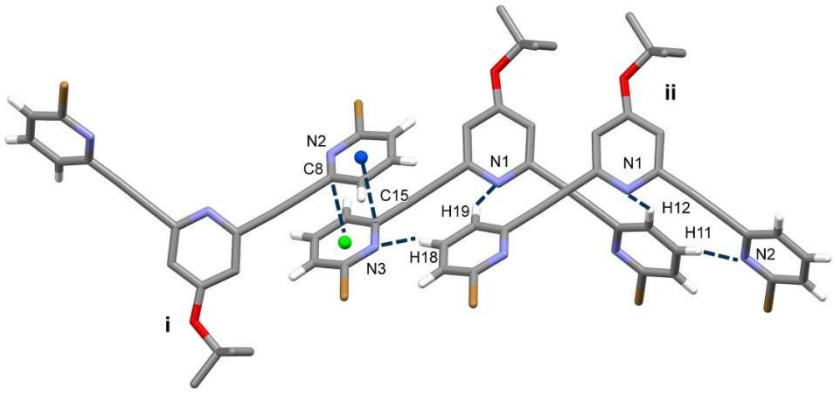


Figure 9. Relevant weak interactions in **TBrT**. Symmetry codes: i = 1+x, 1/2-y, -1/2 + z; ii = x-1, y, z.

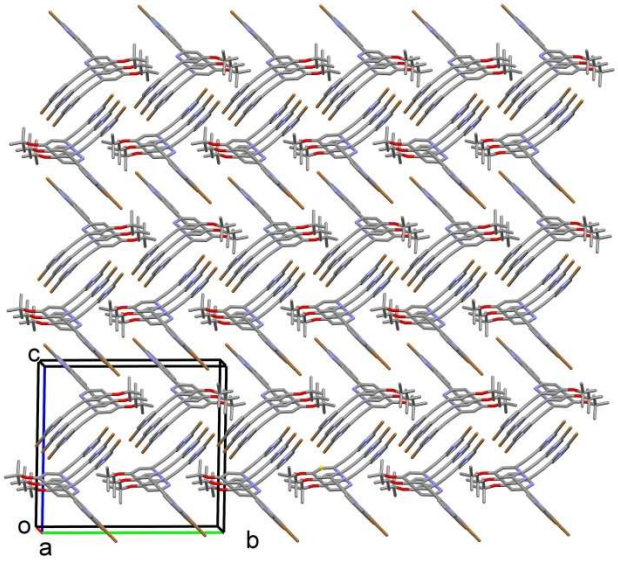


Figure 10. Packing in the crystal structure of **TBrT**. H atoms have been omitted for clarity.

The analysis of the crystal structure of **TBrT** highlighted the tendency of the pyridine rings to act as H-bond acceptors towards aromatic C-H groups (Figure 9). We therefore decided to use benzene as crystallization solvent (less sterically hindered than toluene), to probe its possible templating effect on the crystal packing of **TBr**. However, when crystals grown from benzene were analysed, no solvent appeared to be included. Nevertheless, a different polymorph of **TBr** was obtained from benzene (**TBrB**). The conformation of the **TBr** turned out to be of the S type (Figure 7, Figure 11) with the pyridine ring containing the nitrogen atom N3 pointing upwards with respect to N1 and N2.

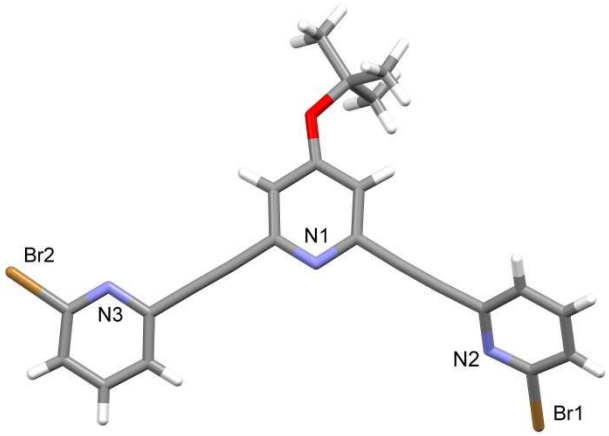


Figure 11. Molecular structure of **TBrB** with partial labelling scheme.

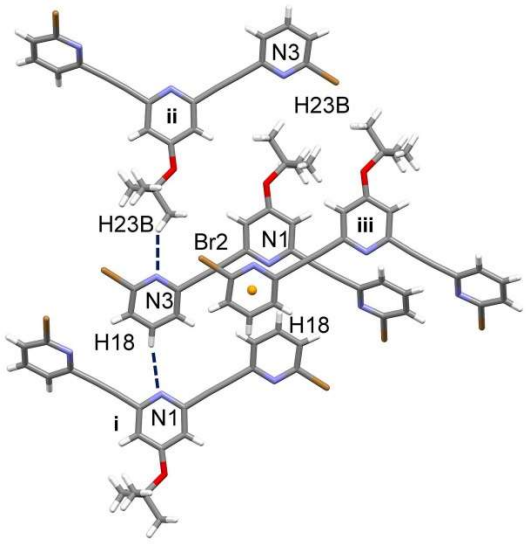


Figure 12. Relevant weak interactions in **TBrB**. Symmetry codes: i = 2-x, 1-y, 1-z; ii = 1-x, 1-y, -z; iii = x-1, y, z.

TBr molecules are more planar that in the previous case, as can be seen from the angles formed by the mean planes passing through the rings [43.11(7)°, 13.42(7)° and 33.16(8)°, following the scheme given for **TBrT**]. This different conformation obviously affects the crystal structure and the set of supramolecular interactions stabilizing it (Figure 12). In this case, each **TBr** molecule forms two sets of centrosymmetric C-H···N interactions with two other pyridine ligands in position *i* and *ii*, of the type C23-H23B···N3 [3.387(5) Å and 134.34(8)°] and C18-H18···N1 [3.492(4) Å and 157.58(7)°]. A short π-π interaction is also present, involving the π cloud of one triple bond and the C15-C19/N3/Br2 ring of a **TBr** molecule in position *iii* [centroid···C13, 3.696(3) Å]. Also in this case, the final packing is extremely compact without any accessible permanent porosity (Figure 13).

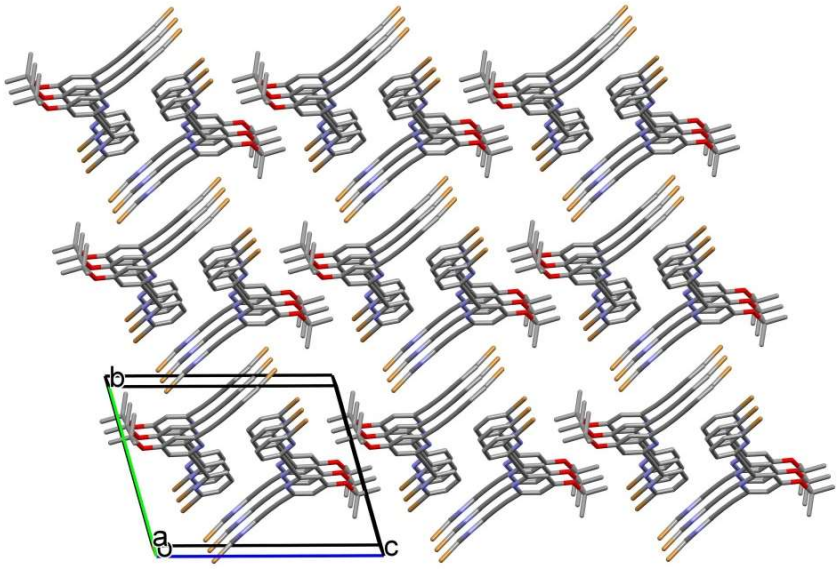


Figure 13. Crystal structure of **TBrB**. H atoms have been omitted for clarity.

A third batch of crystals was grown from chloroform, yielding again the ligand in a C conformation (**TBrC**, Figure 14). **TBrC** molecular structure is almost perfectly superimposable with **TBrT** structure, grown from toluene. However, in the chloroform case the molecule crystallizes in the triclinic space group *P-1* as opposed to the monoclinic *P21/c* of **TBrT**, leading to some minor differences in the crystal packing (Figure 15).

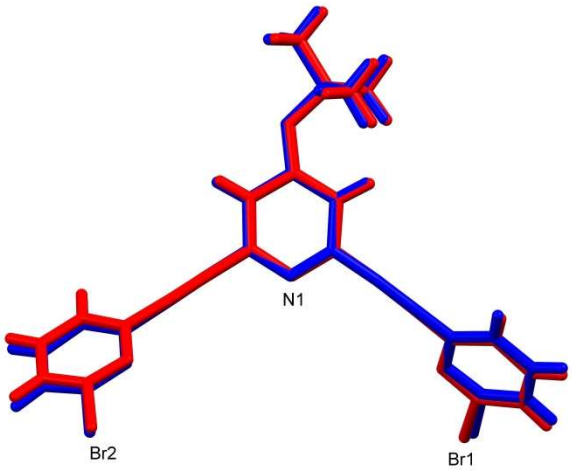


Figure 14. Comparison between the molecular structure of **TBrC** (blue) and **TBrT** (red).

In all the three polymorphic forms **TBrT**, **TBrB** and **TBrC**, the bromine atoms do not play an active role in the formation of supramolecular interactions. Their only influence on the packing is due to their steric hindrance.

In our crystallization experiments, the only solvent which was included in the final structure was methanol, which yielded the solvated form **TBrM** in a 1:1 ligand: solvent ratio (Figure 16). The conformation of the

ligand is S as in the case of the crystals grown from benzene (Figure 11 and 16).

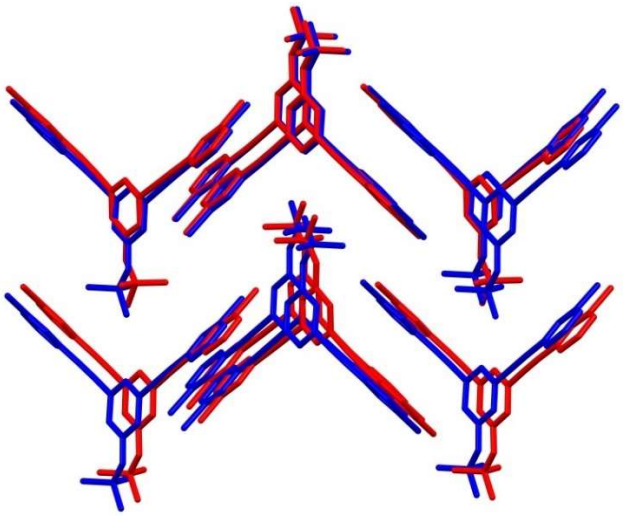


Figure 15. Superimposition of a portion of crystal structure of **TBrC** (blue) and **TBrT** (red).

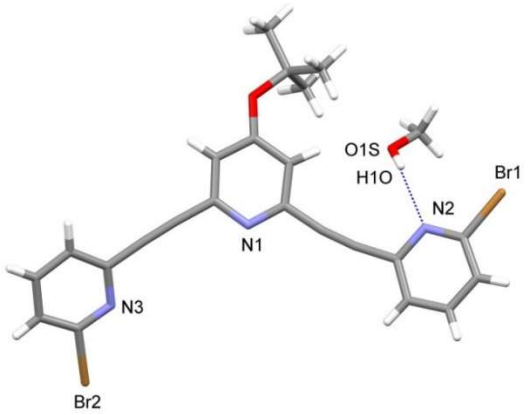


Figure 16. Molecular structure of **TBrM**, highlighting the H-bond (blue dashed line) between one pyridine ring and the methanol molecule.

The pyridine ring C8-C12/N2/Br1, which points upward with respect to the other two pyridine moieties, is held in place by a O1S-H1O···N2 H-bond [2.975(3) Å, 171.23(6)°]. Despite this interaction, methanol plays no other role in the final crystal packing, again dominated by π-π and C-H···N interactions (Figure 17).

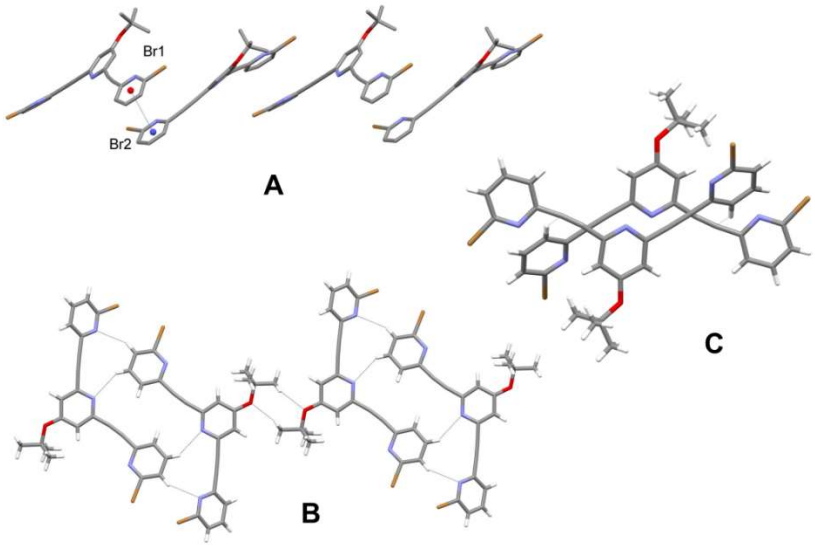


Figure 17. Main weak interactions contributing to the overall general packing of **TBrM**: (**A**) centroid···centroid interactions between pyridine rings; (**B**): C-H···N and C-H···O interactions; (**C**): C-H···π $_{\text{triple bond}}$ interactions.

In particular, Figure 17.A shows the weak π-π stacking involving the pyridine rings C8-C12/N1/Br1 and C15-C19/N2/Br2 [3.804(5) Å]. Furthermore, each **TBr** is involved in a supramolecular dimer held together by C-H···N interactions [C11-H11···N1, 3.434(2) Å and 144.03(2)°; C10-H10···N3, 3.506(4) Å and 152.66(7)°] (Figure 17.B). Supramolecular dimers assemble in tetramers (Figure 17.B) stabilized by two sets of weak centrosymmetric C-H···O hydrogen bonds [C21-H21B···O1, 3.619(4) Å and 158.01(4)°]; with this

layer parallel to the (101) plane. In addition, a supramolecular dimer is also formed by the interaction of a C-H group from the pyridine molecule C15-C19/N2/Br2 and the electronic π cloud of the triple bond [C19-H19···C7, 3.646(4) Å and 145.82(8)°] (Figure 17.C).

Crystals of **TN** were obtained by slow evaporation of a chloroform solution. In this case, solvents capable of forming H-bonds were excluded on purpose, to avoid any influence on the packing of the trimer. Due to the presence of $-NH_2$ substituents on the lateral pyridine rings, the main interaction directing the final crystal packing is H-bonding (Figure 18).

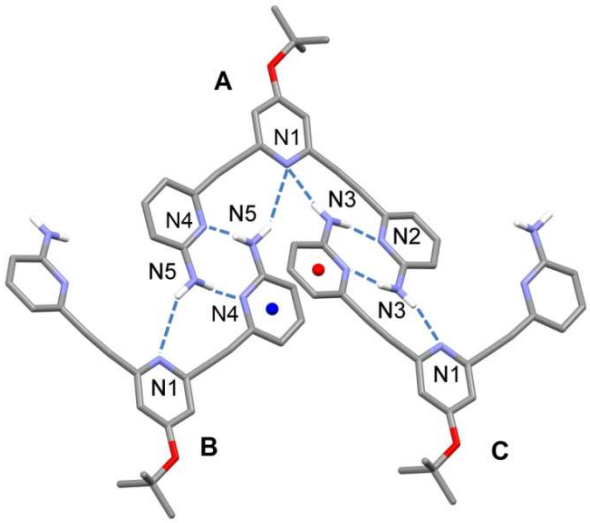


Figure 18. Molecular structure of **TN** showing the H-bonding network connecting **TN** molecules along the *c*-axis direction. The molecules **B** and **C** are in special positions (2-x, -y, -z and 2-x, -y, 1-z, respectively).

Each pyridine ligand acts both as donor and as acceptor of H-bonds with other two adjacent molecules, forming a chain along the c-axis

direction [N3-H3NB···N1, 3.301(3) Å and 152.08(8)°; N3-H3NA···N2, 3.085(2) Å and 173.39(7)°; N5-H5NA···N1, 3.474(4) Å and 148.64(5)°; N5-H5NB···N4, 3.123(4) Å and 171.89(7)°] (Figure 18). This structure is further stabilized by weak π-π stacking interactions involving the side pyridine rings (centroid···centroid distance of 3.751(9) Å). The overall crystal packing of **TN** consists of sets of H-bonded chains held together by van der Waals forces, leading to a very efficient occupancy of space with no accessible voids (Figure 19).

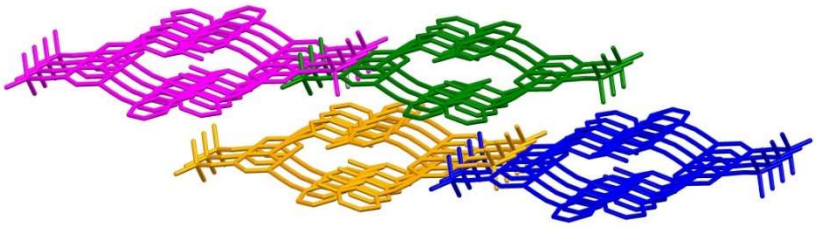


Figure 19. Packing of **TN** showing the reciprocal position of sets of H-bonded chains highlighted in green, magenta, blue and orange.

Crystals of **TA** were obtained by slow evaporation of an acetone solution, yielding the monohydrate form of the ligand [Figure 20; O2-H2O···O1W, 2.600(2) Å and 173.88(4)°]. The presence of water strongly interferes with the packing and could be possibly avoided by growing the crystals in less polar solvents. Nevertheless, attempts were hampered by the low solubility of **TA** in common apolar solvents.

We were aiming at assembling the trimer in a dimeric fashion, exploiting the H-bond ability of the carboxylic group, possibly creating a six-member ring. A dimeric species had indeed been observed via ESI-MS spectrometry. Nevertheless, in the crystal structure, carboxylic acids of

different **TA** ligands interact *via* H-bonding (Figure 21) and form supramolecular chains instead of discrete six-member rings.

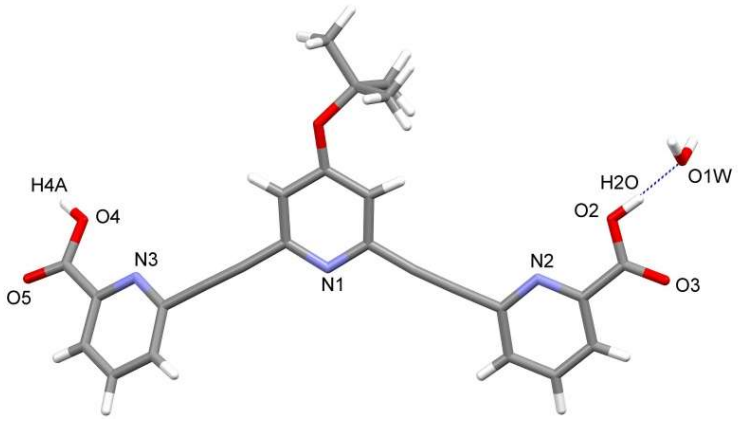


Figure 20. Molecular structure of the monohydrate form of **TA**, highlighting the H-bond (blue dashed line) between the pyridine ring and the water lattice molecule.

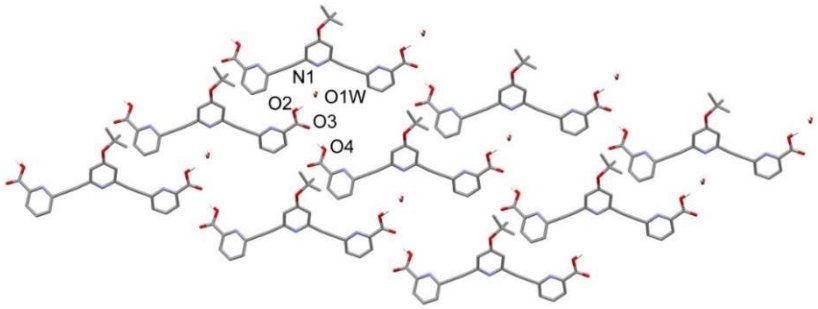


Figure 21. View of the layer parallel to (111) formed by water-mediated, H-bonded **TA** molecules. The atoms involved in H-bonds are labelled.

The oxygen atom O3 of one **TA** acts as hydrogen bond acceptor towards the O4-H4O fragment of an adjacent ligand, while the O2-H2O

group is then involved in a H-bond with water [O4-H4A···O3, 2.600(2) Å and 173.88(4)°], instead of interacting with the C=O5 moiety. These interactions are promoting supramolecular chains formation. Such chains pack in layers parallel to the plane (111) through van der Waals forces and H-bonds interactions involving the lattice water molecule and the pyridine rings of the ligands [O1W-H1W···N1, 2.968(3) Å and 149.70(6)°; O1W-H2W···N2, 2.926(5) Å and 165.57(4)°].

The systematic analysis of the crystal structures of the ligands showed a marked tendency of forming tightly packed networks, very often independently from the crystallization solvent employed. Hence, we introduced as templating agent 1,4-diiodotetrafluorobenzene (**DITFB**). **DITFB** has been widely used in crystal engineering for its ability of forming N···I halogen bonds.[27,28] The crystallization experiments were performed by either mixing equimolar amounts of the trimers with **DITFB** in chloroform, or by first grinding the powders together before recrystallization from a chloroform solution.

In the case of **TBr**, both methods yielded the same 1:1 co-crystal (Figure 22). The nitrogen atom N1 of the central pyridine ring is halogen bonding to the iodine atom I1 of **DITFB** [I1···N1, 3.009(3) Å] (Figure 22). **TBr** adopts a W conformation (Figure 7) which is stabilized by a series of Br···Br interactions of 3.574(6) Å. Two sets of chains are formed, oriented along different planes of the cell, that propagate at an angle of roughly 69° one with respect with the other. These chains are held together by **DITFB** molecules that bridge pairs of **TBr** in a zig-zag fashion through halogen bonds I1···N1 and I2···N3 [3.058(3) Å] (Figure 23). One halogen bond is involving the central pyridine, while the second is formed with one of the two lateral pyridine rings (Figure 23).

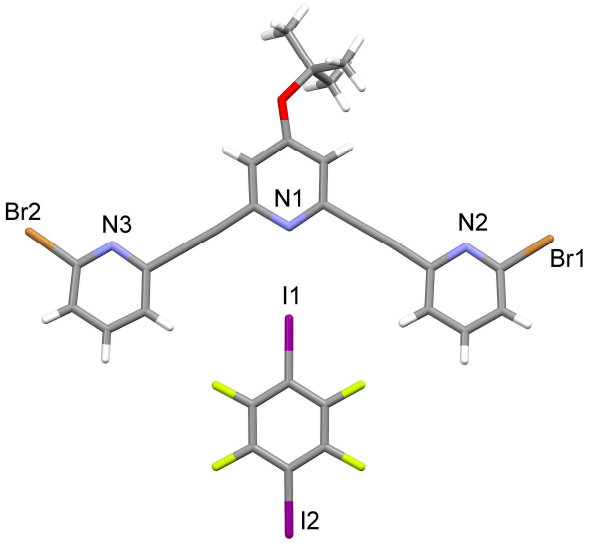


Figure 22. Molecular structure of the co-crystal formed by **TBr** and **DFTIB**.

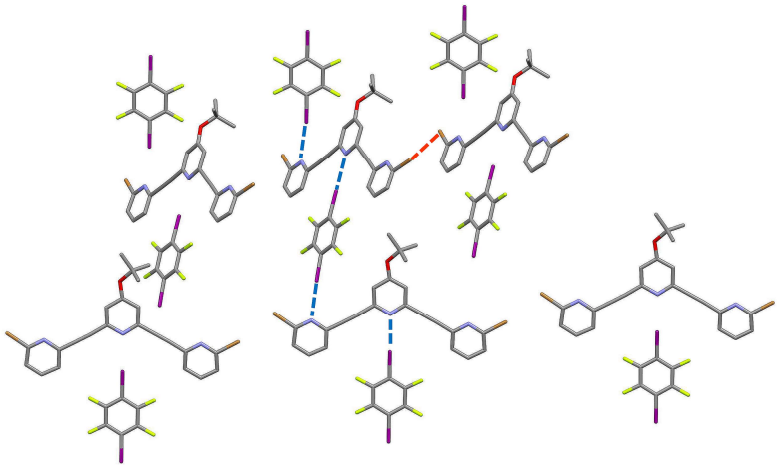


Figure 23. Packing of the co-crystal **TBr-DITFB**, showing two supramolecular chains with different orientations, held together by Br···Br interactions (orange dotted lines). The chains are connected through halogen bonds (blue dotted lines). H atoms have been omitted for clarity.

In addition, centrosymmetric supramolecular dimers are formed by mutual π-π stacking of the lateral pyridine moieties [Cg1···Cg2, 3.536(8) Å] and by the weak interaction of **DITFB** with the central ring of the trimer [Cg3···Cg4, 3.536(8) Å] (Figure 24).

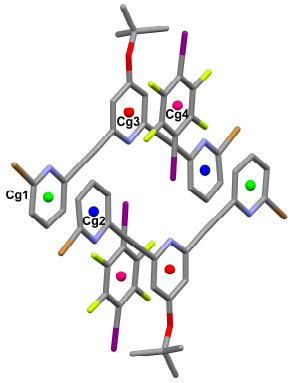


Figure 24. Centrosymmetric dimers formed by π-π stacking involving the aromatic rings of the trimer and of **DITFB**.

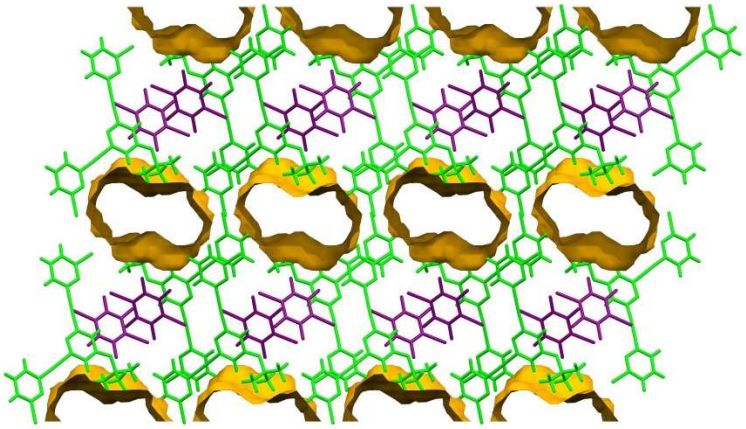


Figure 25. Crystal packing of **TBr-DITFB** viewed along the *b*-axis direction highlighting the voids as yellow cylinders.

The crystal packing is stabilized by the set of interactions described above and shows a moderate porosity. The overall packing has ≈22% of voids in the unit cell (calculated with the program *Mercury*). The resulting channels, parallel to the *b*-axis direction, are filled by loosely-bound chloroform molecules. The highly disordered solvent was treated with the program SQUEEZE (see experimental section).

The co-crystal between **TN** and **DITFB** was obtained only through grinding, since the crystals grown from solution turned out to be the two separate components. Most likely, harsher conditions are required for the formation of the co-crystals, due to the extensive network of H-bonds that strongly stabilizes the structure of **TN** (Figure 18).

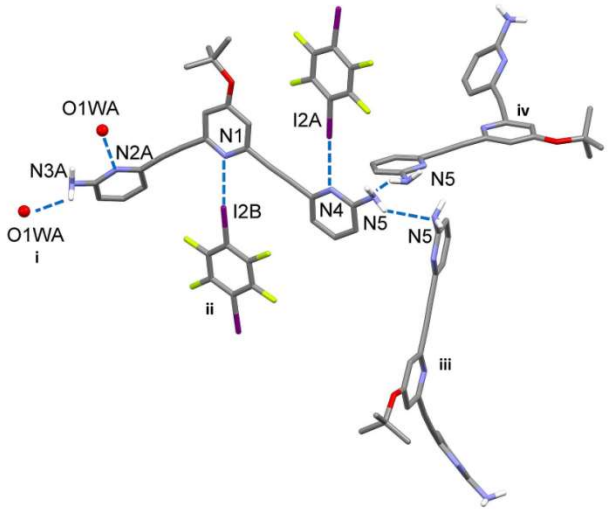


Figure 26. View of the trimer **TN** forming supramolecular interactions with symmetry-related neighbouring molecules. Symmetry codes: i = x-y, x, 2/3+z; ii = -x+y, 1-y, 1,66667+z; iii = -x+y, 1-x, -1/3+z; iv = 1-y, 1-x+y, 1/3+z. Only H atoms relevant for the interactions are shown.

The asymmetric unit of **TN-DITFB** consists of a **TN** molecule in W conformation, two **DFTIB** molecules and an adventitious water molecule. The pyridine ring containing N2 and N3 was found to be disordered over two positions, as it was the water lattice molecule O1W, for which it was not possible to find the H atoms in the difference Fourier map. Figure 26 shows all the relevant weak interactions that the trimer in general position forms with symmetry-related water, **TN** and **DITFB** molecules through hydrogen (from strong to very weak) and halogen bonds, respectively. [O1WA···N2A, 2.717(2) Å; O1WA···N3A, 2.930(2) Å; N5-H5N2···N5iii and N5iv-H5N2iv··· N5, 3.574(3) Å and 136.10(9)°; N1···I2B, 2.948(6) Å; N4···I2A, 2.966 Å]. Furthermore, the crystal packing shows the presence of channels along the *c*-axis direction clearly. Such channels are centred around the 6-fold axis of the hexagonal space group *P64* in which the co-crystal crystallizes. The channels are very small (3% of the unit cell volume calculated with the program *Mercury*), and in fact no chloroform is included in the structure.

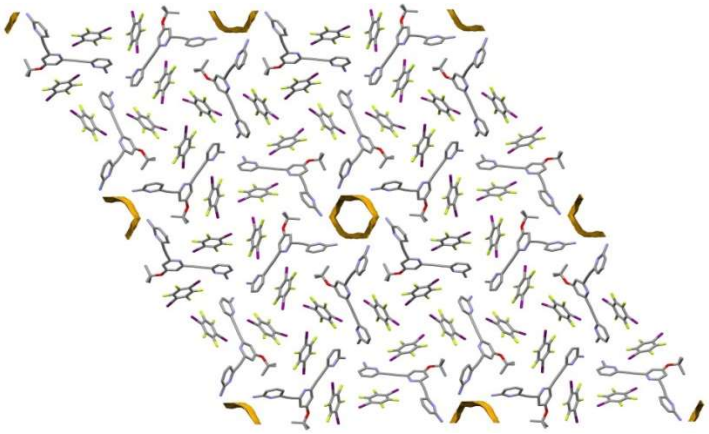


Figure 27. Packing of **TN-DITFB** viewed along the *c*-axis direction. Water molecules and H atoms have been omitted for clarity. Empty channels are represented as yellow cylinders.

Several attempts of growing co-crystals of **TA** and **DITFB** both from solution and from grinding always yielded crystals of the components alone.

We then analysed the behaviour of the trimer **TT**, in which the substituent on the lateral pyridine rings is a triple bond (Figure 28). The crystals were obtained directly from solution. At the time being, no results have been obtained from the grinding experiment. The asymmetric unit of the co-crystal comprises one **TT** molecule in a W conformation and two **DITFB** molecules. In this case, all the three pyridine rings are simultaneously involved in halogen bonds with **DITFB** molecules [I1A···N1, 2.951(2) Å; I2A···N2, 2.942(2) Å; I2B···N3, 3.029(4) Å]

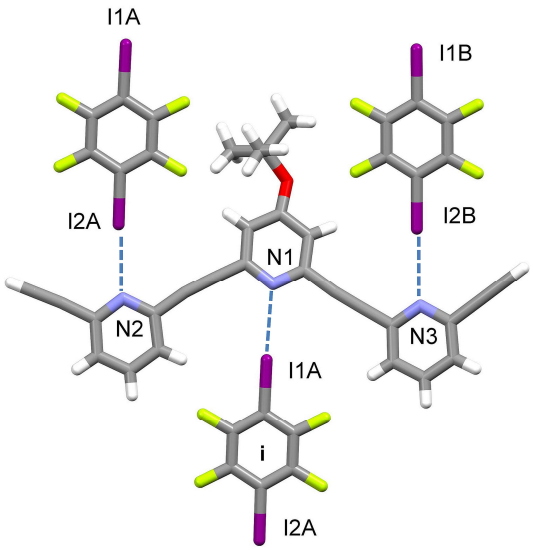


Figure 28. View of the halogen bonds formed in the crystal structure of **TT-DITFB**. The molecule labelled *i* is in special position (i = x, 1+y, z).

The resulting 1:3 assembly is roughly planar, and this planarity is maintained in the crystal structure, which consists of layers parallel to the (-101) plane, stabilized by a series of weak C-H···F interactions involving the acetylene (H14) and the *tert*-butyl (H27A) H atoms [Figure 29; F1B···H14, 2.62 Å; F1B···C14, 3.181(3) Å; C27-H27A···F4B, 3.576(4) Å, 155.67(8)°].

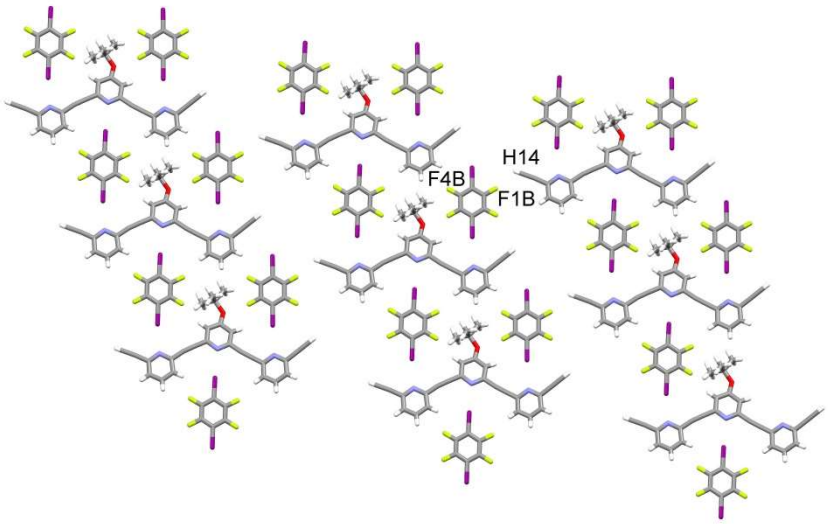


Figure 29. Packing of **TT-DITFB** showing layers parallel to the (-101) plane. Sets of atoms involved in the weak interactions have been labelled.

In this case, the packing does not show the presence of accessible voids, and no solvent is included. It is worth noticing that, in the absence of bulky Br substituents, and of H-bond active moieties such as the carboxylic or amino groups, the **DITFB** molecule can assemble around itself three different ligands by the interplay of N···I and C-H···F interactions (Figure 30). The almost perfect match of geometry and size between the trimer and **DITFB** confirms the potential of these ligands for molecular recognition purposes.

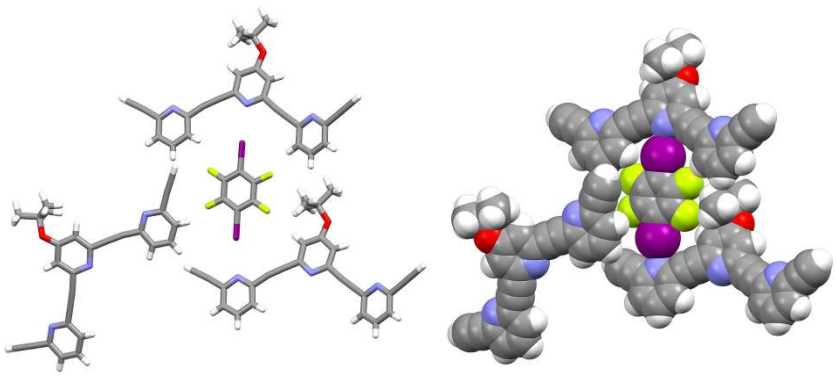


Figure 30. Stick (left) and space filling (right) assembly of **TT** molecules around a central **DITFB**.

4. Fluorine-pyridine ligand

The ligand **TF** was synthesized to study the effect of a fluorine atom in the trimer scaffold, which could provide an alternative site for the formation of C-H···F interactions. This could be very important in directing the final packing of the structures, as evidenced by the case of the **TT-DITFB** co-crystal, and by other examples found in the literature. A first explorative reaction to obtain a metal complex was successfully performed by stirring **TF** in the presence of $AgPF_6$ in acetone at room temperature for ten minutes.

Colourless crystals of the product were obtained *via* vapour diffusion of ethyl ether in acetonitrile. The crystals, which were analysed with single-crystals X-ray diffraction, turned out to be the discrete complex $[Ag(TF)_2]PF_6$ (Figure 31). The asymmetric unit comprises a silver(I) cation which coordinates two symmetry-related **TF** ligands (silver itself is on an inversion centre) through the nitrogen atom N2 of one of the available pyridine rings, while the other one remains uncoordinated. Due to symmetry, the coordination around the metal is perfectly linear, with Ag1-N2 distances of 2.166(9) Å. The molecule is overall planar, and the PF_6 counterion is located closed to the metal centre, with an Ag1···F2 contact of 2.869(9) Å.

In the crystal, the molecules are closely packed, yielding a non-porous solid, stabilized by weak, offset (due to the presence of the bulky anion) π-π stacking interactions involving the pyridine rings (Figure 32).

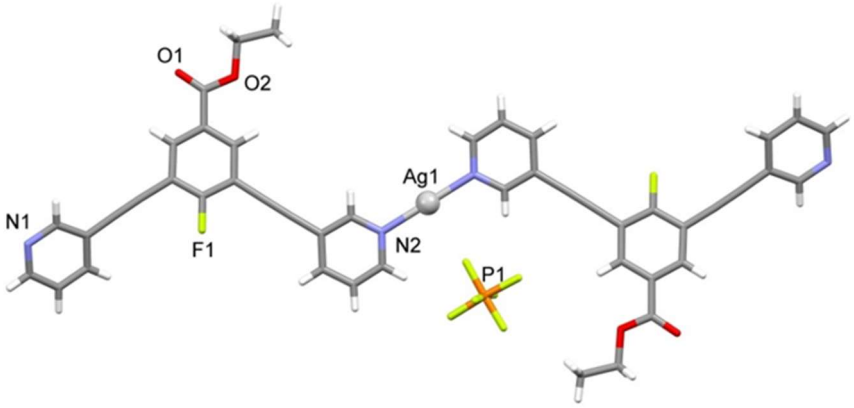


Figure 31. View of the complex [Ag(**TF**)$_2$]PF$_6$. Silver is represented as a grey sphere.

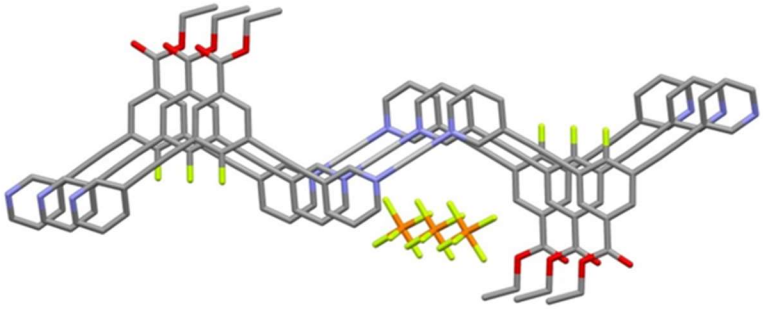


Figure 32. View of [Ag(**TF**)$_2$]PF$_6$ along the *b*-axis direction. H atoms have been omitted for clarity.

The complex forms also sets of supramolecular chains stabilized by couples of centrosymmetric weak C-H···N H-bonds [Figure 33; geometrical parameters: C5-H5···N1, 3.428(9) Å and 148.7(6)°]. The sets are roughly perpendicular one with respect to the other and interact through weak C-H···F interactions involving the ethoxy groups (Figure 34).

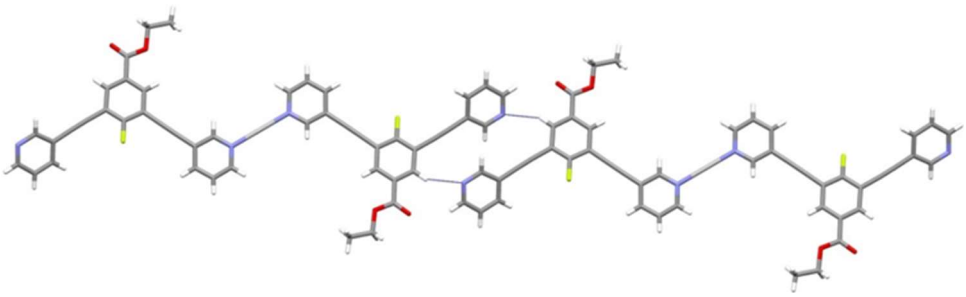


Figure 33. View of the supramolecular chains of [Ag(**TF**)$_2$]PF$_6$ running along the *a*-axis direction. H-bonds are shown as blue dotted lines.

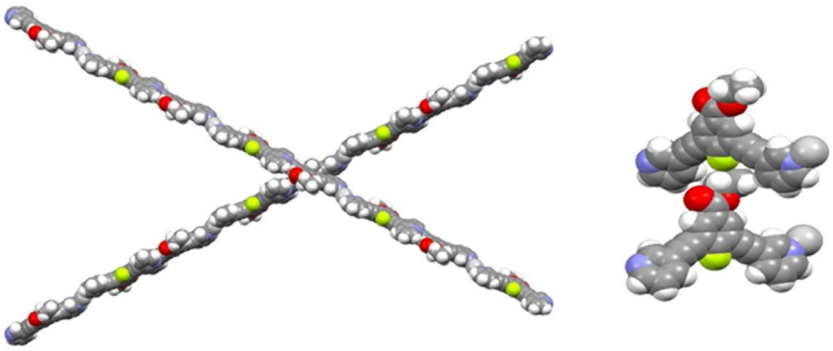


Figure 34. View along the *c*-axis direction of two sets of chains (left) interacting through C-H···F interactions (right).

Further studies are ongoing to obtain new complexes of **TF** with different metals, especially with the aim of obtaining a cyclic dimeric complex for sensing and molecular recognition applications.

5. Coordination polymers of zinc with pyridine-carboxylic ligands: mechanochemical and solvothermal synthesis

Coordination polymers (CPs) and metal organic frameworks (MOFs) are crystalline coordination compounds.[29,30] They comprise multidentate organic ligands that, by bridging metal cores, form chains or frameworks in 1, 2 or 3 dimensions. MOFs differ from CPs due to their high porosity.[30] CPs and MOFs are extremely versatile materials; properties can be tuned by an accurate choice of inorganic and organic components. For this reason, CPs and MOFs found application in several fields, like gas adsorption and storage, sensing and catalysis.[31,32] The choice of the organic ligands is the main factor influencing the properties of the resulting material, in terms of porosity, flexibility, functional groups and coordination geometry.[33] Functionalization of the material can lead to improved selectivity and enhanced recognition properties, which are crucial in applications like selective adsorption, sensing and catalysis.[34,35]

We decided to exploit the properties of the ligand **TA** to prepare functional CPs. Pyridines are in fact excellent H-bond and X-bond acceptors and can act as anchoring points for guests and substrates. **TA** can interact with metals via the two carboxylic groups, leaving the central pyridine free for further guest complexation (Figure 35). The pyridyl-acetylene scaffold facilitates π-π stacking and the obtainment of ordered architectures.

We employed two different synthetic routes to prepare zinc coordination polymers from pyridine-carboxylic ligands: solvothermal and mechanochemical synthesis. Solvothermal synthesis is the most common strategy employed for preparing CPs and MOFs, [33] exploiting temperature

and pressure generated in the reaction vessel. With this strategy, crystalline compounds are obtained directly in high quantities. The reaction is usually performed at high temperature, generally over 90°. In many systems, the reaction temperature influences the degree of crystallinity, yield and morphology of the product.[33] Thus, solvothermal synthesis is not compatible with organic ligands that are thermolabile or have low decomposition temperatures.

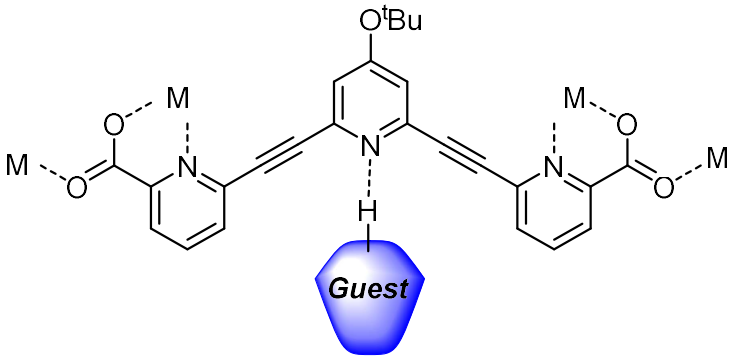


Figure 35. Binding of a **TA**-based coordination polymer to a generic guest through interaction with the central pyridine.

An emerging alternative is mechanochemistry.[36,37] Mechanochemical reactions exploit the energy gained from grinding the reagents together. This technique has lately attracted growing attention as a possible synthetic route to CPs and MOFs. Several versions exist, which differ for the grinding methodology (Figure 36).[36]

Mechanochemistry involves opposite reaction conditions with respect to solvothermal synthesis. Grinding has the advantages of employing low energies to obtain crystalline products in short reaction times (in the order of minutes) with quantitative yields[38]. Very low amounts of solvent are

required, making grinding a truly green technique.[38] In the mechanochemical formation of CPs or MOFs, liquids and counter-ions act as templating agents, facilitating the obtainment of porous structures.[39,40] Furthermore, the use of small quantities of solvent (Liquid Assisted Grinding, LAG) increases the mobility of the reagents and hence the reaction rates.[40] Recent studies have highlighted the advantages of using hydrate metallic salts with basic counter-ions.[33]

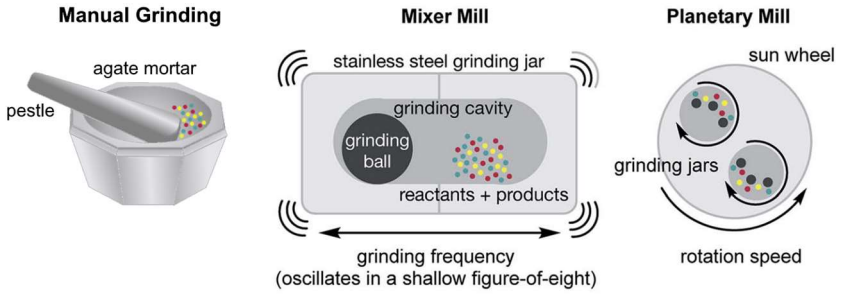


Figure 36. Grinding techniques.

The final characterization of the product is the main drawback in the synthesis of CPs and MOFs *via* grinding. CPs and MOFs are generally insoluble, making solution techniques (like NMR) unsuitable. Even if the products are crystalline powders, crystals are too small to obtain single crystal diffraction data. In these cases, powder diffraction is the main technique employed to obtain structural information. Nevertheless, the solution of crystal structures from powder data can be a challenging task, especially when in the presence of more than one phase. In this chapter, we obtained structural data *via* electron diffraction tomography, with the collaboration of Dr. M. Gemmi and his team (IIT, Pisa).

In this study, we employed two pyridine-carboxylic ligands: **TA** and the commercially available 2,6-pyridinedicarboxylic acid (**H_2PDC**). We

exploited **H_2PDC** to set reaction conditions and test our methodology. We performed solvothermal and mechanochemical synthesis with both **H_2PDC** and **TA** and we compared the results between the two synthetic routes and the two ligands. Zinc (II) has been chosen as metal centre because it forms complexes with a coordination number of 5 or 6 and can adopt different coordination geometries. This versatility aids the formation of reticulated structures. In addition, zinc complexes are generally stable to humidity, air and light, simplifying handling and storage.

Several solvothermal reactions have been performed exploiting **H_2PDC** and $Zn(NO_3)_2{\cdot}6H_2O$ (Table 1). Stoichiometric amounts have been varied and, in some cases, ancillary ligands (pillars) have been exploited. Firstly, DMF has been used as a solvent. Nevertheless, no crystalline material was obtained under these conditions.

Table 1. Solvothermal reactions of **H_2PDC** with $Zn(NO_3)_2{\cdot}6H_2O$. The third column reports stochiometric ratios between ligand (L), metal salt (M) and pillar (P).

ENTRY	PILLAR	L:M:P	SOLVENT	OUTCOME
1	/	1 : 1 : 0	DMF	Amorphous
2	4,4′-bipyridyl	1 : 1 : 0.5	DMF	Amorphous
3		1 : 2 : 0.5	DMF	Amorphous
4		1 : 1 : 0.5	EtOH + NH_3 aq.	Crystals
5	2,4,6-tri(4-pyridyl)-1,3,5-triazine	1:1:1	DMF	Amorphous
6		1:2:0.33	DMF	Amorphous
7		1:3:0.33	DMF	Amorphous

Only the synthetic strategy shown in entry 4, with ethanol and ammonia solution as solvents, gave needle-like transparent crystals with two different morphologies. The crystals were studied by single-crystal X-ray diffraction analysis and turned out to be two discrete complexes, from now on indicated as complex 1 and complex 2 (Figure 37). The pillar, 4-4'bipyridyl was not involved in zinc coordination, even if used as reagent.

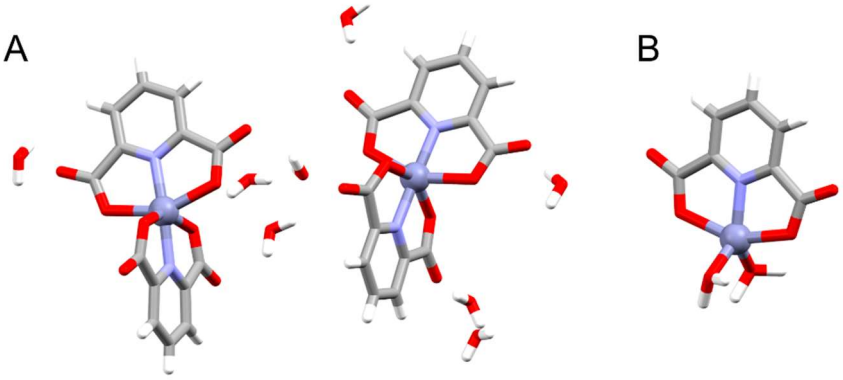


Figure 37. (A) Molecular structure of complex 1 and (B) complex 2. Zn^{2+} cations are represented as violet spheres.

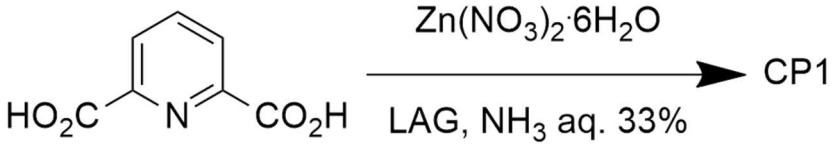
The asymmetric unit of complex 1 comprises two independent molecules of $[Zn(HPDC)_2]\cdot 4H_2O$. Each zinc cation coordinates two mono-deprotonated ligands **$HPDC^-$** and shows a distorted octahedral coordination geometry, with the two pyridinic nitrogens in axial positions and four carboxylate oxygens in equatorial positions (Figure 37.A). The Zn-O distances vary from 2.114(2) to 2.332(2) Å, while the Zn-N bond lengths are in the range 2.015(2) - 2.030(2) Å. The crystal structure is stabilized through a network of H-bonds involving the lattice water molecules.

Complex 2, $[Zn(PDC)(H_2O)_2]$, comprises a zinc cation coordinating one fully deprotonated **PDC^{2-}** ligand and two water molecules. The complex

geometry can be described as distorted trigonal bipyramid with two oxygen atoms from the ligand in axial positions [O2-Zn1-O4 152.37(8)°], while the pyridinic nitrogen and the oxygen atoms of the two water molecules occupy an equatorial position. In this case the two Zn-O bonds are of 2.289(2) and 2.138(2) Å, while the Zn-N distance is of 2.017(2) Å.

In the crystal, the complexes interact one with the other through a series of H-bonds involving the water molecules and the oxygen atoms of the carboxylic groups.

Since solvothermal reactions only led to discrete complexes, we decided to employ the LAG methodology. A 1D-coordination polymer (CP1) was obtained by manual grinding of **H_2PDC** and zinc nitrate in the presence of microliters of ammonia solution 33% (Scheme 9).



Scheme 9. Synthesis of CP1.

After washing of the crude with ethanol, PXRD analysis confirmed the crystallinity of the product (Figure 38).

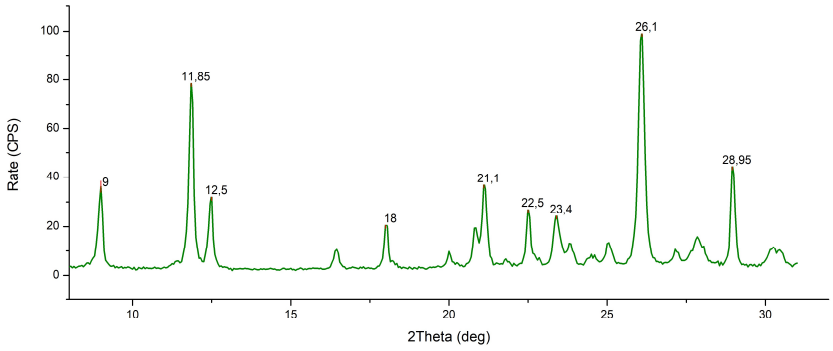


Figure 38. PXRD spectrum of CP1 after washing with ethanol.

Transparent crystals were obtained *via* recrystallization of the crude product in warm ethanol. The crystals, which were analysed with single-crystals X-ray diffraction, turned out to be a $\{[Zn(PDC)(H_2O)]0.5H_2O\}_n$ CP (Figure 39).

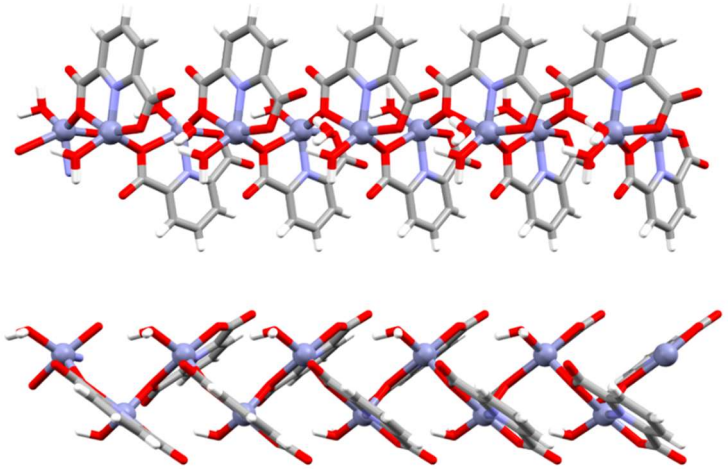


Figure 39. Structure of CP1 along the direction of the crystallographic axis *b*. The two panels are related by a 90° rotation. Zn^{2+} cations are

represented as violet spheres. Lattice water molecules have been omitted for clarity.

Similarly to complex 2, zinc shows a trigonal bipyramidal coordination geometry. Each Zn^{2+} coordinates one deprotonated **PDC^{2-}**, which acts as a tridentate chelating ligand through two oxygen atoms of the carboxylic groups and the pyridinic nitrogen atom, forming two 5-members condensed chelation rings. The fourth coordination position is occupied by an oxygen from an adjacent **PDC^{2-}** ligand, which bridges two metal centers, thus prompting polymerization along the *b* crystallographic axis. Zinc coordination sphere is completed by a water molecule. Distances and bond angles are similar to those found for complex 2. In the crystal structure, the polymeric chains are stabilized by a H-bond network, also involving crystallization water molecules. Comparison between experimental and calculated PXRD spectra (Figure 40) confirmed that the structure obtained from single crystal analysis coincides with the bulk phase obtained from the LAG reaction.

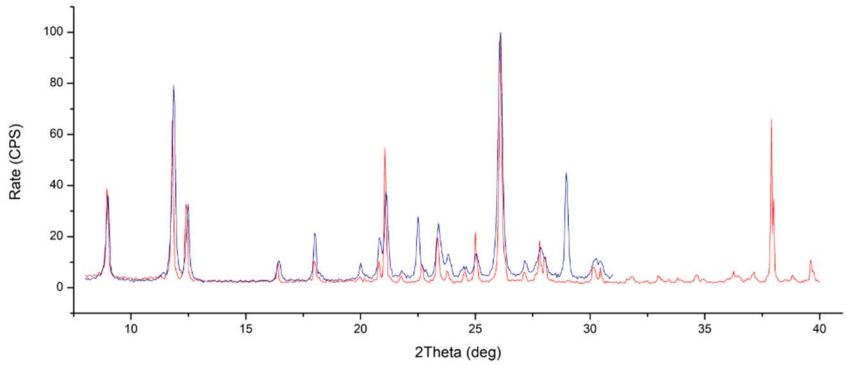


Figure 40. Comparison between the experimental PXRD spectrum of CP1 after the LAG reaction (blue) and predicted PXRD spectrum from CP1

crystals (red). The two spectra match well, confirming that CP1 is the bulk phase obtained after the LAG.

By monitoring the reaction *via* PXRD, we observed the formation of a stable intermediate after addition of half the amount of ammonia (Figure 41). The intermediate could be converted into CP1 either by further LAG or by recrystallization in ethanol and small quantities of ammonia.

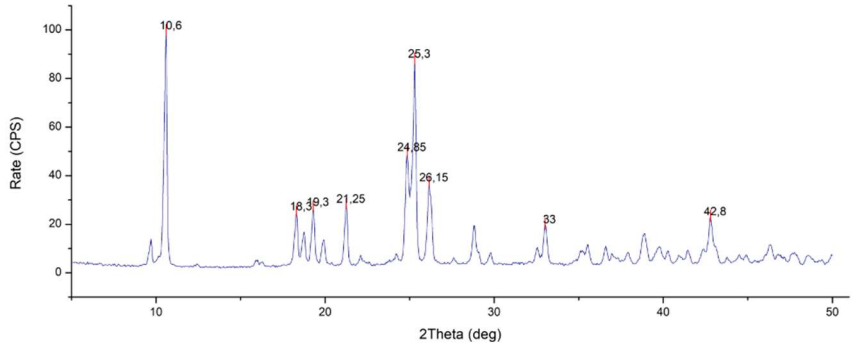


Figure 41. PXRD spectrum of the stable intermediate observed during the synthesis of CP1.

In CP1, the coordination around zinc was completed by a water molecule. To obtain 2D or 3D architectures, we thought of substituting this water molecule with a pillar ligand. Hence, LAG of **H_2PDC** and $Zn(NO_3)_2{\cdot}6H_2O$ was carried out in the presence of di-, tri- and tetra-dentate pyridinic pillars: 4,4'-bipyridyl (BiPy), 2,4,6-tri-(4-pyridyl)-1,3,5-triazine (TriPy) and 1,1'',1''',1''''''-(methanetetrayltetrakis(benzene-4,1-diyl))tetrakis(([4,4'-bipyridin]-1-ium)) tetrachloride (TetraPy) (Figure 42).

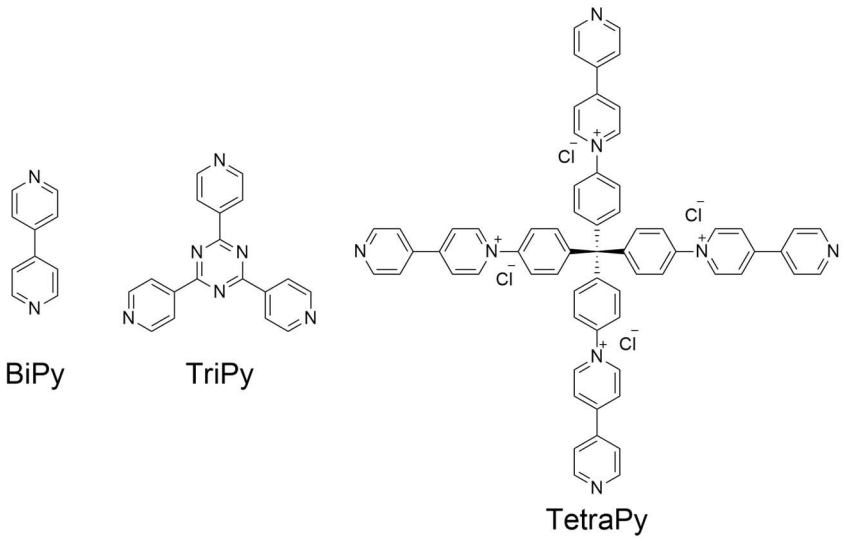
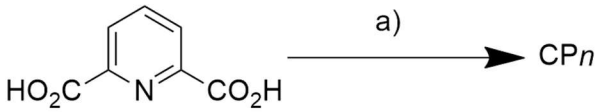


Figure 42. Molecular structures of BiPy, TriPy and TetraPy.

LAG reactions were performed with small quantities of aqueous ammonia 33% (Scheme 10).



Scheme 10. Synthesis of CPs from **H_2PDC** and pillars. a) BiPy/TriPy/TetraPy, NH_3 33%, LAG, 10', r.t.

The reaction of **H_2PDC** with BiPy gave a 2D coordination polymer (CP2). After work-up, the PXRD analysis of the crude showed the presence of a microcrystalline product, different from CP1 (Figure 43).

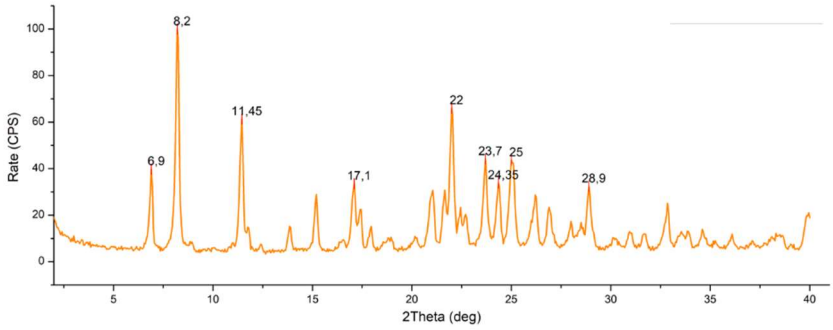


Figure 43. PXRD spectrum of CP2 after washing with ethanol.

CP2 was not soluble in most common organic solvent, thus recrystallization was impossible. The structure was solved through electron diffraction tomography, in collaboration with the group of Dr. Mauro Gemmi (IIT, Pisa). Two different crystal morphologies were observed at the electron microscope: square-shaped and rod-shaped (Figure 44).

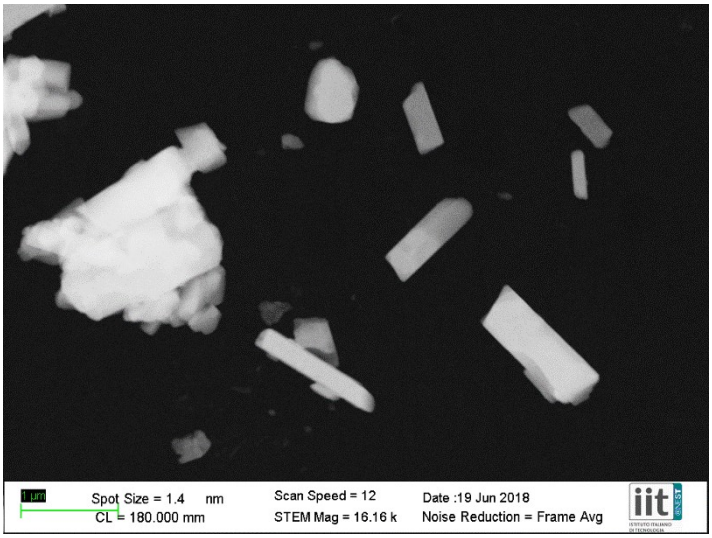


Figure 44. STEM figure of CP2 crystals.

Instability of the square-shaped crystals to the electron beam thwarted a complete dataset collection. Nevertheless, a dataset was collected for the rod-shaped crystals. The structure was solved in Jana and Vesta software and showed a 2D coordination polymer with formula $\{[Zn_2(PDC)_2(BiPy)]\}_n$ (Figures 45 and 46).

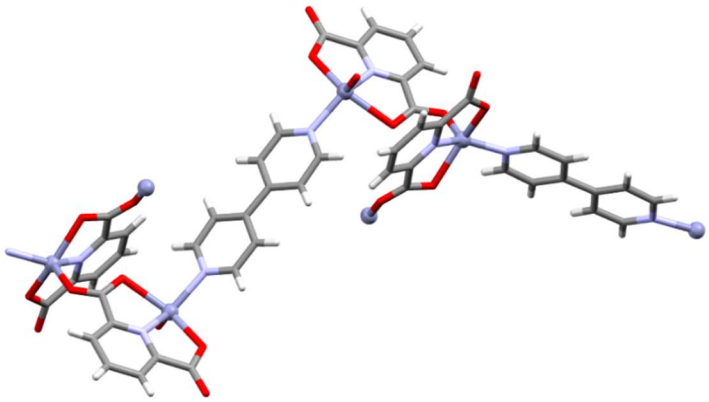


Figure 45. View of a portion of the coordination polymer CP2. Zn^{2+} is represented as a violet sphere.

As already observed for complex 2 and CP1, also in CP2 the ligand **H_2PDC** is fully deprotonated in the form **PDC^{2-}** and acts as a tridentate chelating ligand forming two 5-member condensed rings. The coordination geometry can be described as a distorted squared-based pyramid, with a BiPy pyridinic nitrogen occupying the fourth equatorial position, while the apical site is provided by an oxygen from a bridging carboxylate. The two independent zinc cations Zn1 and Zn2 are bridged by a carboxylic moieties, thus creating a zig-zag $[Zn_2(PDC)_2]_n$ polymer chain. On the other hand, BiPy pillars connect these $[Zn_2(PDC)_2]_2$ dimers along the unit cell diagonal (Figure 46), giving rise to a bi-dimensional structure.

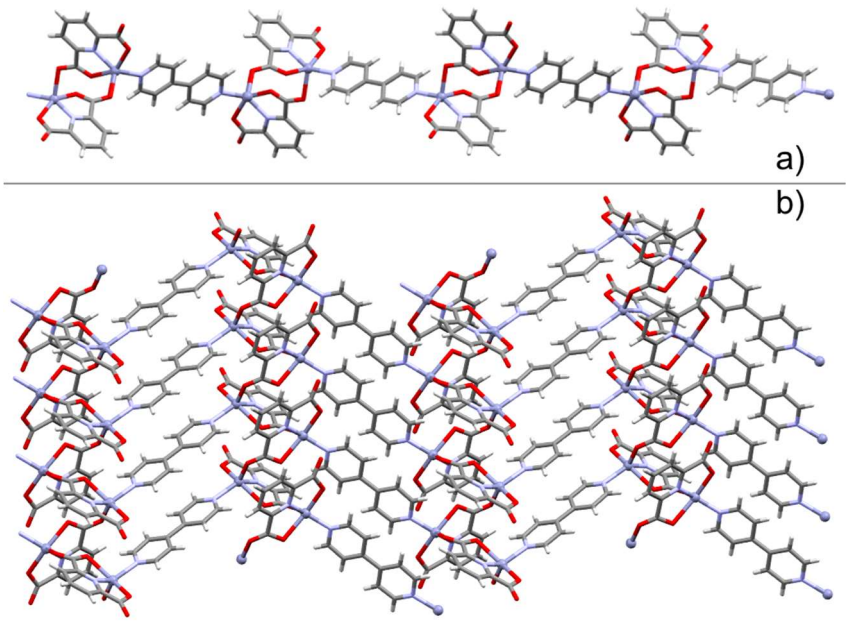


Figure 46. (A) CP2 structure along the direction of the crystallographic axis *a*. (B) packing of CP2. The bottom panel is obtained by a 90° rotation of the upper panel and an expansion of the $[Zn_2(PDC)_2]_n$ chains.

CP2 showed moderate porosity (≈ 21%), highlighted in Figure 47 as yellow channels. Most likely, these channels contain solvent molecules that could not be identified due to the low quality of the data.

Even if CP1 and CP2 present different coordination geometries, the two polymers show a similar concatenation of **PDC^{2-}** ligands (Figure 48). In CP1, the metallic cores are separated by an oxygen atom, while in CP2 by a whole O-C-O group. The fifth ligand completing the zinc coordination sphere is a water molecule in CP1 and a BiPy in CP2, which play two very similar roles in the overall coordination geometry.

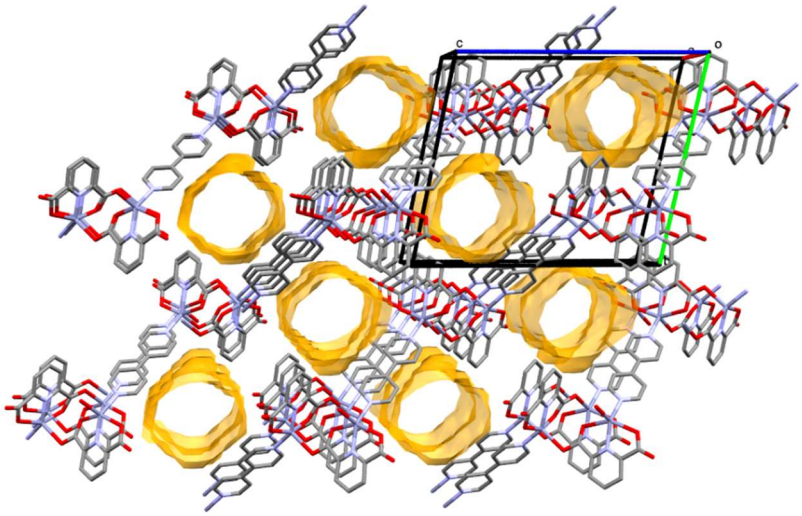


Figure 47. Packing in CP2. Voids are represented as yellow cylinders.

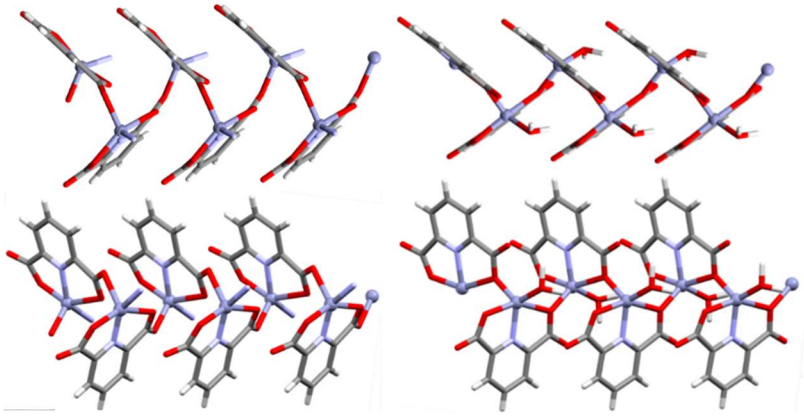


Figure 48. Comparison between the disposition of the ligand **PDC^{2-}** in CP2 (left) and CP1 (right).

The chain of **PDC^{2-}** ligands can be represented as a cylinder with Zn^{2+} laying on the outside surface (Figure 49). In this view, H_2O and BiPy are protruding from the surface with the same directionality. This graphical

analysis shows how the pillar BiPy takes the same position previously occupied by water and leads to a 2D polymer. Indeed, the insertion of a pillar increases the dimensionality from 1D to 2D and creates a partially porous network.

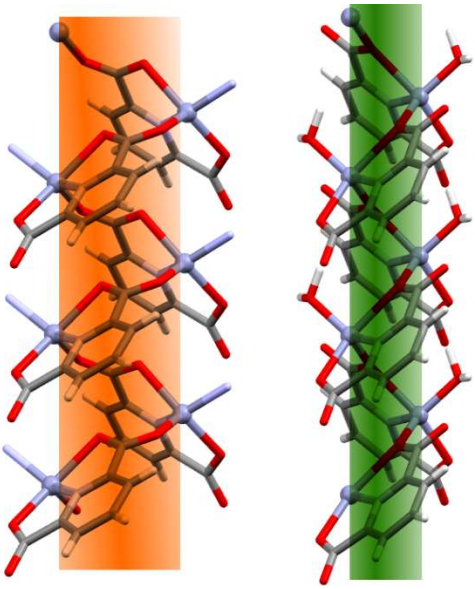


Figure 49. Graphical representation of the **PDC^{2-}** concatenation in CP2 (left) and CP1 (right). The one-dimensional chains are highlighted by an orange and green cylinder, respectively, with the Zn^{2+} cations (violet spheres) laying on the surface. BiPy and H_2O are protruding from the surface showing highly similar directionalities.

The promising results obtained with **H_2PDC** and BiPy led us to employ tri- and tetra-dentate pillars to obtain networks with different geometries and dimensionality.

The reaction through LAG of **H_2PDC** and zinc nitrate with the tri-and tetra-dentate pillars TriPy or TetraPy gave crystalline powders whose PXRD

spectra did not resemble any of the reagents (Figure 50 and 51). Structural characterization of this CPs through electron microscopy tomography is, at the time being, on-going.

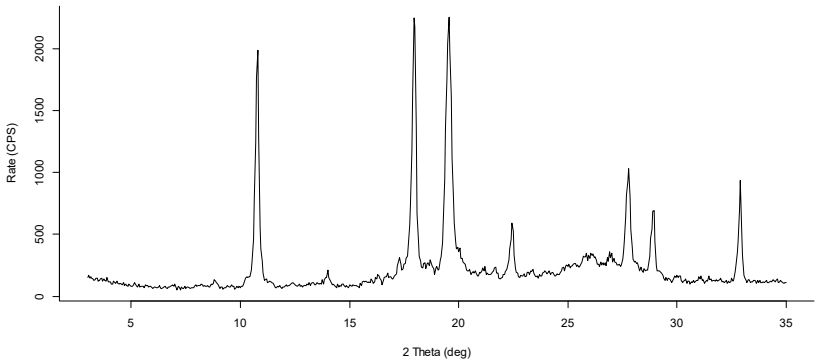


Figure 50. PXRD spectrum of the crude from the LAG reaction between **H_2PDC**, zinc nitrate and TriPy.

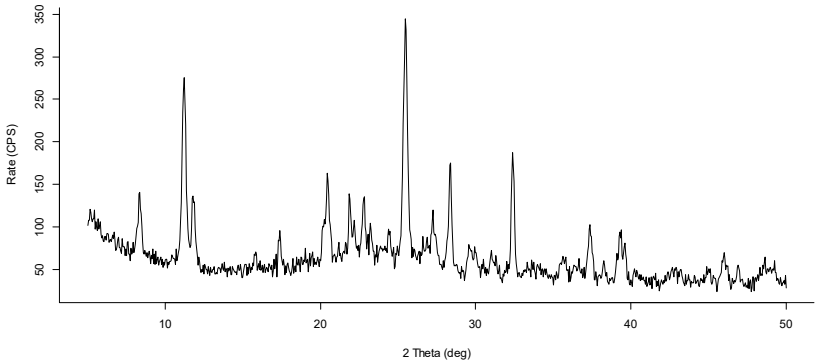
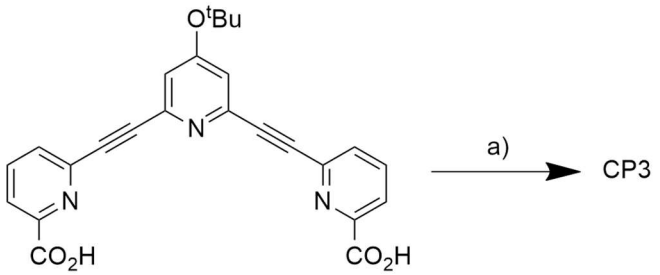


Figure 51. PXRD spectrum of the crude from the LAG reaction between **H_2PDC**, zinc nitrate and TetraPy.

After the results obtained for the commercially available ligand **H_2PDC**, we repeated the procedures with our synthetic ligand **TA**.

A 1D coordination polymer (CP3) was obtained from **TA** and $Zn(NO_3)_2{\cdot}6H_2O$ *via* solvothermal synthesis in DMF at 100° (Scheme 11).



Scheme 11. Synthesis of CP3. a) $Zn(NO_3)_2{\cdot}6H_2O$, DMF, 3 d, 100°.

After the reaction a black precipitate was obtained, most likely due to thermal decomposition of **TA** which is not stable over 100°. The precipitate was re-crystallized in hot DMSO, giving white, needle-shaped crystals. X-ray diffraction revealed the presence of a CP propagating along the direction of the crystallographic axis *b*, of formula $\{[Zn(TA)(H_2O)]{\cdot}H_2O{\cdot}MeOH{\cdot}DMF\}_n$ (Figure 52).

TA, which is in the W conformation, is completely deprotonated (Figure 52). Both its pyridine-carboxylate groups act as a bi-dentate chelating ligand *via* one carboxylic oxygen and the pyridinic nitrogen (Figure 52). On the other side, each zinc cation coordinates two pyridine-carboxylates belonging to two different, symmetry-related **TA** ligands. The coordination around the metal is completed by a water molecule, giving rise to an overall coordination geometry of a trigonal-based pyramid with two nitrogens in the axial positions and the three oxygens in the equatorial positions.

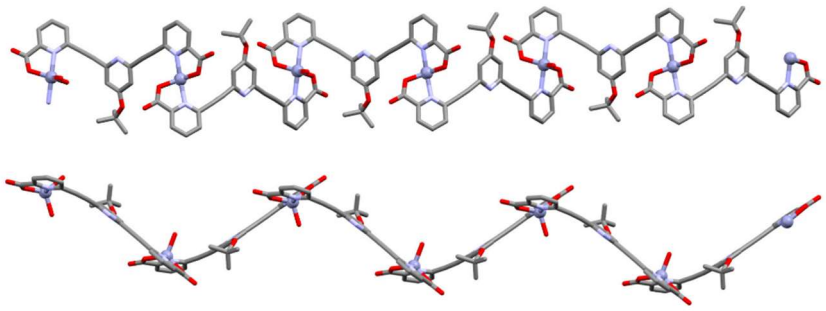


Figure 52. Structure of CP3 along the *b* crystallographic axis. The two panels are related by a 90° rotation. Zn^{2+} cations are represented as violet spheres. H atoms and solvent molecules have been omitted for clarity.

In the crystal lattice, the polymeric chains interact with each other via H-bonding and π-π stacking (Figure 53) leading to a packing with no porosity (Figure 53).

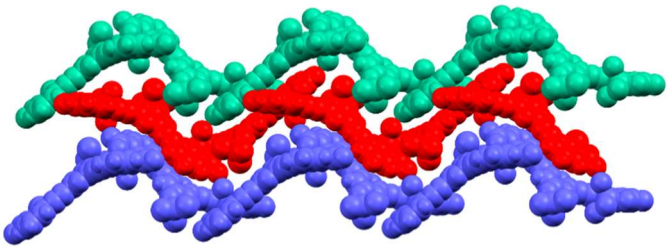


Figure 53. Packing in CP3 along the b crystallographic axis. Each polymeric chain is highlighted with a different colour. Solvent molecules have been omitted for clarity.

As for CP1, the water molecule on the zinc centre could have been substituted by a pillar ligand, increasing the dimensionality of the resulting product. To address this aim, we decided to employ only mechanochemical

synthesis as **TA** proved not stable enough under solvothermal conditions. Various LAG reactions have been performed employing **TA** and zinc nitrate, in ethanol and ammonia, with or without BiPy as a pillar. Unfortunately, the products of the reactions were amorphous, as showed by PXRD analysis (Figure 54). Low solubility of the products in most common organic solvents hampered recrystallization.

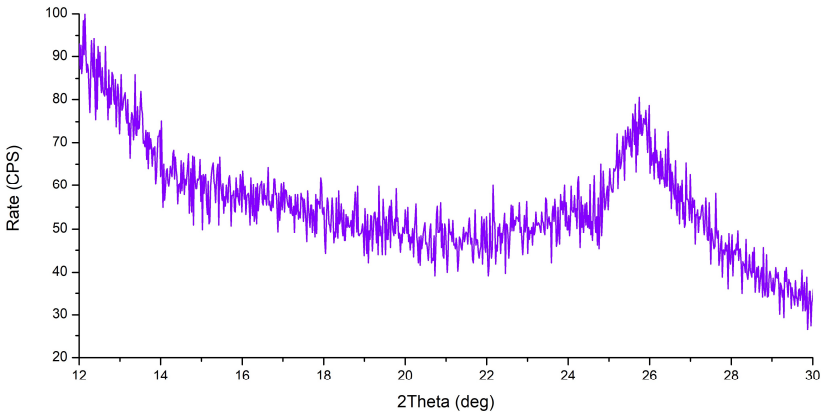


Figure 54. PXRD spectrum for the LAG reaction of **TA** with zinc nitrate and BiPy.

6. Conclusions and future perspective

In this chapter, the synthesis of five multidentate pyridine ligands is described. Such ligands possessed a V-shaped aromatic scaffold comprising a pyridyl-acetylene system. The properties of the ligand can be tuned by functionalization, which also influences its solid state assemblies. The main interactions driving the formation of the crystal structures are H-bonds C-H···π and π-π interactions, with a preference towards non-porous assemblies. Functional groups and solvent also appear to play a crucial role in the final packing. Furthermore, co-crystallization experiments with the halogen bond donor 1,4-diiodotetrafluorobenzene led to intricated architectures. The assembly, driven by X-bond, was again influenced by functional groups on the scaffold. Bromine substituents seem to promote porous assemblies, *via* Br-Br weak interactions.

In addition, functionalization with carboxylic-acid moieties enabled the use of such ligands for metal coordination. These pyridine-carboxylic ligands were exploited to develop coordination polymers (CPs) via mechanochemical synthesis. Thanks to the collaboration with Prof. Lara Righi (Department SCVSA, University of Parma), Dr. Mauro Gemmi (Italian Institute of Technology, Pisa) and his team, we characterized the materials with electron diffraction tomography. Electron diffraction is gaining increasing attention in crystallography as a technique to obtain structural information when single crystal XRD is not feasible. To our knowledge, mechanochemistry and electron diffraction have never been coupled to synthesize and characterize CPs. The combination of the two methodologies is highly effective.

7. Acknowledgments

We would like to acknowledge Dr. Stefano Canossa for **TBrT** structure data collection and for advice on crystal growth. Prof. Lara Righi (University of Parma), is thanked for powder diffraction data acquisition. Dr. Mauro Gemmi and Dr. Arianna Lanza (Italian Institute of Technology, Pisa) are thanked for electron diffraction data acquisition and elaboration. We acknowledge Danilo Marchetti for carrying out part of the LAG synthesis.

8. Experimental

Synthesis

2,6-dibromo-4-(*tert*-butoxy)pyridine (1): To a solution of 2,6-dibromo-4-nitropyridine (1 g, 3.5 mmol) in 15 mL of THF, potassium *tert*-butoxyde (2 g, 17.5 mmol) was added under argon at 0°. The flask was sealed, and the mixture was stirred a room temperature for 5 hours. The reaction was quenched with 20 mL of distilled water. A saturated solution of NH_4Cl was added dropwise until neutrality was reached. The aqueous phase was extracted with DCM and the organics were collected. The solvent was dried under reduced pressure and column chromatography (gradient from DCM/Hex 5:5 to DCM/Hex 7:3) afforded product **1** as a colourless oil (y: 69%).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.00 (s, 2H, Ar**H**); 1.48 (s, 9H, ArO^t**Bu**).

ESI-MS: m/z= 331.27 $[M+Na]^+$.

4-(*tert*-butoxy)-2,6-bis((trimethylsilyl)ethynyl)pyridine (2): A solution of **1** (500 mg, 1.625 mmol) in 7 mL of THF and 2 mL of NEt_3 was degassed *via* freeze-pump-thaw cycles. Under argon, PPh_3 (17 mg, 0.065 mmol), trimethylsilylacetylene (573 µL, 4.06 mmol), $Pd(PPh_3)_2Cl_2$ (45 mg, 0.065 mmol) and CuI (31 mg, 0.163 mmol) were added. The flask was sealed, and the mixture was stirred at room temperature. After 6 hours, volatiles were removed under reduced pressure. Column chromatography (Hex/AcOEt 9:1) afforded product **2** as a colourless oil (y: 54%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 6.98 (s, 1H, Ar**H**); 1.46 (s, 9H, ArO^t**Bu**); 0.23 (s, 9H, Si(C**H$_3$**)$_3$).

GC-MS: m/z= 343.3 $[M]^+$; 270.2 $[M-^tBu+H]^+$.

4-(*tert*-butoxy)-2,6-diethynylpyridine (3): To a solution of **2** (298 mg, 0.87 mmol) in 10 mL of THF and 10 mL of MeOH, KF (202 mg, 3.5 mmol) was added. The mixture was stirred at room temperature for 1 hour. The reaction was quenched with 20 mL of distilled water and the aqueous phase was extracted with DCM. The organics were collected, and the solvent was evaporated under reduced pressure. Column chromatography (gradient from Hex/AcOEt 7:3 to Hex/AcOEt 6:4) afforded product **3** as a brownish oil (y: 75%).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 6.94 (s, 2H, Ar**H**); 3.03 (s, 2H, ArCC**H**); 1.37 (s, 9H, ArO^t**Bu**).

ESI-MS: m/z = 200 $[M+H]^+$; 222 $[M+Na]^+$; 238 $[M+K]^+$.

TBr: A solution of **3** (77 mg, 0.39 mmol) and 2,6-dibromopyridine (366 mg, 1.55 mmol) in 3 mL of NEt_3 was degassed *via* freeze-pump-thaw cycles. Under argon, PPh_3 (16 mg, 0.062 mmol), $Pd(PPh_3)_2Cl_2$ (43 mg, 0.062 mmol) and CuI (29 mg, 0.15 mmol) were added. The flask was sealed, and the mixture was stirred at 65°. After 5 hours, volatiles were removed under reduced pressure. Column chromatography (gradient from DCM/MeOH 99:1 to DCM/MeOH 9:1) and preparative TCL (Hex/AcOEt 5:5) afforded **TBr** as a brownish solid (y: 25%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.55 (m, 6H, Ar**H**); 7.23 (s, 2H, Ar**H**); 1.59 (s, 9H, ArO^t**Bu**).

^{13}C-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 169; 143; 142; 141; 138; 128; 127; 120; 89; 86; 32; 30; 29; 23; 14.

ESI-MS: m/z = 512 $[M+H]^+$; 534 $[M+Na]^+$.

ethyl 6-bromopicolinate: A suspension of 6-bromopicolinic acid (500 mg, 2.5 mmol) in 15 mL of EtOH was heated until dissolution. 1 mL of H_2SO_4 98% was added dropwise and the mixture was stirred at 80° for 24 hours. The reaction was quenched with 1 mL of distilled water and neutralized with a saturated $NaHCO_3$ solution until pH 8 was reached. The aqueous phase was extracted with DCM. The organics were collected, and the solvent was removed under reduced pressure affording ethyl-6-bromopicolinate as a white solid (y: 85%).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.07 (s, 1H, Ar**H**); 7.66 (m, 2H, Ar**H**); 4.46 (q, J=7 Hz, 2H, C(O)OC**H_2**CH_3); 1.42 (t, J=7 Hz, 3H, C(O)OCH_2C**H_3**)

ESI-MS: m/z= 230.11 $[M+H]^+$; 252.11 $[M+Na]^+$; 270.13 $[M+K]^+$.

TA step 1: A solution of **3** (128 mg, 0.64 mmol) and ethyl 6-bromopicolinate (370 mg, 1.60 mmol) in 5 mL of THF and 2 mL of NEt_3 was degassed *via* freeze-pump-thaw cycles. Under argon, PPh_3 (7 mg, 0.026 mmol), $Pd(PPh_3)_2Cl_2$ (18 mg, 0.026 mmol) and CuI (12 mg, 0.064 mmol) were added. The flask was sealed, and the mixture was stirred at 55°. After 18 hours, volatiles were removed under reduced pressure. Column chromatography (DCM/AcOEt 8:2) and preparative TLC (DCM/AcOEt 7:3) afforded protected **TA** as a white solid (y: 80%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 8.08 (d, J=8, 2H, Ar**H**); 7.83 (m, 4H, Ar**H**); 7.28 (s, 2H, Ar**H**); 4.49 (q, J=7 Hz, 4H, C(O)OC**H_2**CH_3); 1.51 (s, 9H, ArO**^{t}Bu**); 1.44 (t, J=7 Hz, 6H, C(O)OCH_2C**H_3**)

ESI-MS: m/z = 498.24 $[M+H]^+$.

TA step 2: To as solution of protected **TA** (68 mg, 0.14 mmol) in 5 mL of THF and 3 mL of H_2O, $LiOH{\cdot}H_2O$ (23 mg, 0.55 mmol) was added. The mixture was stirred at room temperature for 30 minutes. The reaction was quenched

with 10 mL of distilled water, a solution of HCl (1N) was added dropwise until pH 7 was reached and a white precipitate was obtained. The precipitate was collected affording **TA** as a white solid (y: quant.).

^{1}H-NMR (acetone-d6, 300 MHz): δ (ppm) = 8.20 (m, 4H, Ar**H**); 7.98 (m, 2H, Ar**H**); 7.42 (s, 2H, Ar**H**); 1.58 (s, 9H, ArO**^{t}Bu**).

ESI-MS: m/z = 442.38 $[M+H^+]$; 464.38 $[M+Na]^+$; 480.29 $[M+K]^+$.

2-amino-6-((trimethylsilyl)ethynyl)pyridine (4): A solution of 2-amino-6-bromopyridine (1 g, 5.78 mmol) in 7 mL of THF and 2 mL of NEt_3 was degassed *via* freeze-pump-thaw cycles. Under argon, PPh_3 (60 mg, 0.23 mmol), trimethylsilylacetylene (978 µL, 6.90 mmol), $Pd(PPh_3)_2Cl_2$ (162 mg, 0.23 mmol) and CuI (110 mg, 0.58 mmol) were added. The flask was sealed, and the mixture was stirred at 55°. After 2 hours, volatiles were removed under reduced pressure. Column chromatography (DCM/MeOH 95:5) afforded product **4** as a white solid (y: 89%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.37 (t, J=9, 1H, Ar**H**); 6.87 (d, J=9, 1H, Ar**H**); 6.45 (d, J=9, 1H, Ar**H**); 4.57 (s, 2H, N**H**$_2$); 0.26 (s, 9H, Si(C**H$_3$**)$_3$).

GC-MS: m/z= 191.1 $[M+H]^+$; 175 $[M-NH_2]$; 145 $[M-NH_2\text{-}2Me]$.

2-amino-6-ethynylpyridine (5): To a solution of **4** (980 mg, 5.15 mmol) in 10 mL of THF and 10 mL of MeOH, KF (600 mg, 10.3 mmol) was added. The mixture was stirred at room temperature for 2 hours. The reaction was quenched with 20 mL of distilled water and the aqueous phase was extracted with DCM. The organics were collected, and the solvent was evaporated under reduced pressure. Column chromatography (gradient from DCM/MeOH 99:1 to DCM/MeOH 94:6) afforded product **5** as a brown solid (y: 46%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.36 (t, J=9, 1H, Ar**H**); 6.83 (d, J=9, 1H, Ar**H**); 6.47 (d, J=9, 1H, Ar**H**); 4.75 (s, 2H, N**H**$_2$); 3.04 (s, 1H, ArCC**H**).

GC-MS: m/z= 120 $[M+H]^+$; 104 $[M-NH_2]$.

TN: A solution of **1** (332 mg, 1.08 mmol) and **5** (318 mg, 2.69 mmol) in 5 mL of THF and 2 mL of NEt_3 was degassed *via* freeze-pump-thaw cycles. Under argon, PPh_3 (11 mg, 0.043 mmol), $Pd(PPh_3)_2Cl_2$ (30 mg, 0.043 mmol) and CuI (20 mg, 0.11 mmol) were added. The flask was sealed, and the mixture was stirred at 60°. After 24 hours, volatiles were removed under reduced pressure. Column chromatography (gradient from DCM/MeOH 98:2 to DCM/MeOH 94:6) afforded **TN** as a brownish solid (y: 87%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.47(t, J=9, 2H, Ar**H**); 7.21 (s, 2H, Ar**H**); 7.05 (d, J=9, 2H, Ar**H**); 6.55 (d, J=9, 2H, Ar**H**); 4.5 (N**H$_2$**); 1.51 (s, 9H, ArO$^{\mathbf{t}}$**Bu**).

ESI-MS: m/z = 384.34 $[M+H]^+$.

4-(6-bromopyridin-2-yl)-2-methylbut-3-yn-2-ol (6): A solution of 2,6-dibromopyridine (500 mg, 2.11 mmol) in 5 mL of NEt_3 was degassed *via* freeze-pump-thaw cycles. Under argon, $Pd(PPh_3)_4$ (12 mg, 0.010 mmol), 2-methyl-3-butynol (205 µL, 2.11 mmol) and CuI (16 mg, 0.084 mmol) were added. The flask was sealed, and the mixture was stirred at room temperature. After 20 hours, volatiles were removed under reduced pressure. Column chromatography (gradient from Hex/AcOEt 7:3 to Hex/AcOEt 5:5) afforded product **6** as a brown solid (y: 63%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.55 (m, 3H, Ar**H**); 1.65 (s, 6H, C(C**H$_3$**)$_2$OH).

GC-MS: m/z= 238 $[M+H]^+$.

2-methyl-4-(6-((trimethylsilyl)ethynyl)pyridin-2-yl)but-3-yn-2-ol (7): A solution of **6** (315 mg, 1.32 mmol) in 5 mL of NEt_3 was degassed *via* freeze-

pump-thaw cycles. Under argon, PPh_3 (14 mg, 0.053 mmol), trimethylsilylacetylene (570 µL, 3.95 mmol), $Pd(PPh_3)_2Cl_2$ (37 mg, 0.053 mmol) and CuI (25 mg, 0.13 mmol) were added. The flask was sealed, and the mixture was stirred at room temperature. After 16 hours, volatiles were removed under reduced pressure. Column chromatography (Hex/AcOEt 5:5) afforded product **7** as a white solid (y: quant.).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.65 (t, J = 8 Hz, 1H, Ar**H**); 7.39 (m, 2H, Ar**H**); 1.64 (s, 6H, C(C**H$_3$**)$_2$OH); 0.28 (s, 9H, Si(C**H$_3$**)$_3$).

GC-MS: m/z= 258 $[M+H]^+$.

4-(6-ethynylpyridin-2-yl)-2-methylbut-3-yn-2-ol (8): To a solution of **7** (399 mg, 1.55 mmol) in 2 mL of THF and 2 mL of MeOH, KF (90 mg, 1.55 mmol) was added. The mixture was stirred at room temperature for 1 hour. The reaction was quenched with 20 mL of distilled water and the aqueous phase was extracted with DCM. The organics were collected, and the solvent was evaporated under reduced pressure. Column chromatography (Hex/AcOEt 5:5) afforded product **8** as a white solid (y: 47%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.65 (m, 1H, Ar**H**); 7.42 (m, 2H, Ar**H**); 3.18 (s, 1H, ArCC**H**); 1.65 (s, 6H, (C**H$_3$**)$_2$OH).

GC-MS: m/z= 187 $[M+H]^+$.

TT step 1 (9): A solution of **1** (93 mg, 0.30 mmol) and **8** (134 mg, 0.72 mmol) in 6 mL of NEt_3 was degassed *via* freeze-pump-thaw cycles. Under argon, $Pd(PPh_3)_4$ (2 mg, 0.0017 mmol) and CuI (2 mg, 0.011 mmol) were added. The flask was sealed, and the mixture was stirred at 55°. After 16 hours, volatiles were removed under reduced pressure. Column chromatography (gradient from DCM/AcOEt 7:3 to DCM/AcOEt 5:5) afforded protected **TT** as a white solid (y: 34%).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.67 (m, 2H, Ar**H**); 7.58 (m, 2H, Ar**H**); 7.47 (m, 2H, Ar**H**); 7.21 (s, 2H, Ar**H**); 1.65 (s, 6H, (C**H$_3$**)$_2$OH); 1.51 (s, 9H, ArO$^{\mathbf{t}}$**Bu**).

ESI-MS: m/z = 518.35 $[M+H]^+$; 540.32 $[M+Na]^+$; 556.20 $[M+K]^+$.

TT step 2: To a solution of **9** (53 mg, 0.10 mmol) in 5 mL of toluene, NaOH (77 mg, 0.76 mmol) was added. The mixture was stirred at 100°for 24 hours. The reaction was quenched with 20 mL of distilled water and the aqueous phase was extracted with DCM. The organics were collected, and the solvent was evaporated under reduced pressure. Column chromatography (DCM/AcOEt 9:1) afforded **TT** as a white solid (y: 80%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.68 (m, 2H, Ar**H**); 7.60 (m, 2H, Ar**H**); 7.47 (m, 2H, Ar**H**); 7.22 (s, 2H, Ar**H**); 3.20 (s, 2H, ArCC**H**); 1.52 (s, 9H, ArO$^{\mathbf{t}}$**Bu**).

ESI-MS: m/z = 402.58 $[M+H]^+$; 425.00 $[M+Na]^+$.

4-fluoro-3,5-diiodobenzoic acid (10): To a suspension of I_2 (2.95 g, 9.99 mmol) in 20 mL of H_2SO_4, CrO_3 (1 g, 9.99 mmol) was added and the mixture turned green. The mixture was stirred at room temperature for 30 minutes in which a colour change from green to black was observed. 4-fluorobenzoic acid (1 g, 7.12 mmol) was added and the mixture was stirred at room temperature. After 24 hours, the reaction was poured in distilled H_2O at 0°C. The white precipitate was collected and dried under reduced pressure affording product **6** as an off-white solid (y: quant.).

^{1}H-NMR (DMSO-d_6, 400MHz): δ = 8.28 (d, J = 4 Hz, 2H, Ar**H**).

^{19}F-NMR (DMSO-d_6, 162 MHz): δ = 69.49 (s, 1F, Ar**F**).

ethyl 4-fluoro-3,5-diiodobenzoate (11): To a suspension of **10** (500 mg, 1.28 mmol) in 20 mL of toluene, H_2SO_4 (1 mL) and EtOH (10 mL) were added. The mixture was refluxed for 4 hours. After cooling at room temperature, the

reaction was quenched with distilled water. The aqueous phase was extracted in toluene and evaporation of the solvent under reduced pressure afforded product **11** as a white solid (y: quant.).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.40 (d, J = 8 Hz, 2H, Ar**H**); 4.39 (q, J = 8 Hz, 2H, C(O)OC**H_2**CH_3); 1.44 (t, J = 8 Hz, 3H, C(O)OCH_2C**H_3**).

^{19}F-NMR ($CDCl_3$, 162 MHz): δ (ppm) = 65.84 (s, 1F, Ar**F**).

^{13}C-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 141; 140; 62; 14.

GC-MS: m/z= 420 $[M+H]^+$; 392 [M-Et]; 375 [M-Et-F]; 347 [M-Et-I].

TF: A solution of **11** (80 mg, 0.9 mmol) and 3-ethynylpyridine (59 mg, 0.57 mmol) in 3 mL of THF and 3 mL of NEt_3 was degassed *via* freeze-pump-thaw cycles. Under argon, $Pd(PPh_3)_4$ (9 mg, 0.0076 mmol) and CuI (4 mg, 0.019 mmol) were added. The flask was sealed, and the mixture was stirred at 80°. After 24 hours, volatiles were removed under reduced pressure. Preparative TLC (DCM/MeOH 98:2) and recrystallization in acetonitrile afforded **TF** as a white solid (y: 14%).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.85 (s, 2H, Ar**H**); 8.63 (s, 2H, Ar**H**); 8.23 (d, J = 8 z, 2H, Ar**H**); 7.89 (m, 2H, Ar**H**); 7.35 (m, 2H, Ar**H**); 4.41 (q, J = 8 Hz, 2H, C(O)OC**H_2**CH_3); 1.44 (t, J = 8 Hz, 3H, C(O)OCH_2C**H_3**).

^{19}F-NMR ($CDCl_3$, 162 MHz): δ (ppm) = 99.67 (s, 1F, Ar**F**).

ESI-MS: m/z= 371.16 $[M+H]^+$; 393.16 $[M+Na]^+$.

Solvothermal Synthesis

Complex 1 and 2: $Zn(NO_3)\cdot 6H_2O$ (77 mg, 0.26 mmol), **H_2PDC** (43 mg, 0.26 mmol) and BiPy (20 mg, 0.13 mmol) were dissolved in 7 mL of EtOH and NH_3 33% (100 µL) was added. The tube was sealed with a Teflon cap and heated to 80°. After 5 days, needle-like crystals were obtained.

CP3: $Zn(NO_3)\cdot 6H_2O$ (13 mg, 0.045 mmol) and **TA** (20 mg, 0.045 mmol) were dissolved in 3 mL of DMF. The tube was sealed with a Teflon cap and heated to 100°. After 3 days, a black precipitate was obtained. The precipitate was recrystallized from hot DMSO, yielding white needle-like crystals.

Mechanochemical Synthesis

General LAG procedure: H_2PDC or **TA**, $Zn(NO_3)_2\cdot 6H_2O$ in equimolar amounts were transferred in an agate mortar. 0.5 equivalents of the chosen pillar (see main text) were added followed by 50 µL of NH_3 33%. The reagents were ground together manually until complete evaporation of the solvent. The crude was washed with EtOH to eliminate reagents traces.

X-ray Crystallography for Section 3: Supramolecular assembly of pyridine V-shaped ligands

The crystal structures of **TBrT**, **TBrB**, **TBrC**, **TBrM**, **TN**, **TA**, **TBr-DITFB**, **TN-DITFB** and **TT-DITFB** were determined by X-ray diffraction methods. Crystal data and experimental details for data collection and structure refinement are reported in Tables S1-5. Intensity data and cell parameters were recorded at 190(2) K on a Bruker Apex II diffractometer equipped with

a CCD area detector, using a MoKα radiation (= 0.71073 Å). The raw frame data were processed using SAINT and SADABS to yield the reflection data file.[41] **TBrC** was twinned and 4832 corrected reflections were written to the file twin4.hkl using the program TWINABS v2008/1.[41]The structures were solved by Direct Methods using the SIR97 program[42] and refined on F_o^2 by full-matrix least-squares procedures, using SHELXL-2014[43] in the WinGX suite v.2014.1.[44] All non-hydrogen atoms were refined with anisotropic atomic displacements, except when disorder was present. The hydrogen atoms were included in the refinement at idealized geometry (C-H 0.95-0.99 Å, O-H 0.82 Å) and refined "riding" on the corresponding parent atoms with U_{iso}(H) set to 1.2U_{eq}(C) and 1.5U_{eq}(O, C_{met}). When possible, the H atoms of water molecules were found in the difference Fourier map. In the case of **TBr-DITFB** the PLATON SQUEEZE procedure[45] was used to treat regions of diffuse solvent which could not be sensibly modelled in terms of atomic sites. Their contribution to the diffraction pattern was removed and modified F_o^2 written to a new HKL file. The number of electrons located were included in the formula, formula weight, calculated density, and F(000). The residual electron density of 362 electrons was assigned to 4 molecules of chloroform per unit cell. The weighting schemes used in the last cycle of refinement were $w = 1/ [\sigma^2F_o^2 + (0.0383P)^2]$, $w = 1/ [\sigma^2F_o^2 + (0.0476P)^2 + 1.8617P]$, $w = 1/ [\sigma^2F_o^2 + (0.0365P)^2 + 1.3951P]$, $w = 1/ [\sigma^2F_o^2 + (0.0269P)^2 + 4.8588P]$, $w = 1/ [\sigma^2F_o^2 + (0.0926P)^2]$, $w = 1/ [\sigma^2F_o^2 + (0.0687P)^2 + 0.2082P]$, $w = 1/ [\sigma^2F_o^2 + (0.0185P)^2]$, $w = 1/ [\sigma^2F_o^2 + (0.0293P)^2 + 13.3548P]$ and $w = 1/ [\sigma^2F_o^2 + (0.495P)^2 + 2.6128P]$, where $P = (F_o^2 + 2F_c^2)/3$, for **TBrT**, **TBrB**, **TBrC**, **TBrM**, **TN**, **TA**, **TBr-DITFB**, **TN-DITFB** and **TT-DITFB**, respectively.

Table S1. Crystallographic data for **TBrT** and **TBrB**.

	TBrT	**TBrB**
Formula	$C_{23}H_{17}N_3OBr_2$	$C_{23}H_{17}N_3OBr_2$
Formula weight	511.21	511.21
Crystal system	Monoclinic	Triclinic
Space group	*P21/c*	*P-1*
a/Å	6.0870(2)	6.0679(6)
b/Å	19.5378(7)	12.315(1)
c/Å	17.9833(6)	15.658(1)
α/°	-	103.775(4)
β/°	92.834(2)	98.920(4
γ/°	-	93.544(4)
V/Å^3	2136.1(1)	1116.7(2)
Z	4	2
D_c/g cm^{-3}	1.590	1.520
F(000)	1016	508
μ/mm^{-1}	3.814	3.648
$\theta_{min,max}$/°	2.496, 26.456	3.352, 26.644
Reflections collected	22204	18144
Independent reflections	4397 (*R*int = 0.0930)	4613 (*R*int = 0.0879)
Observed reflections	2877	2380
Data/restr./param.	4397 / 0 / 265	4613 / 0 / 265
S[a]	1.002	1.009
R [Fo>4σ(Fo)][b], w*R*2[b]	0.0445, 0.0835	0.0602, 0.1170
$\Delta\rho_{min,max}$/e Å^{-3}	0.620, -0.471	1.123, -0.857

[a]Goodness-of-fit S = [Σw(F_o^2-F_c^2)^2/ (n-p)]1/2, where n is the number of reflections and p the number of parameters. [b]R_1 = Σ‖F_o| - |F_c‖/Σ|F_o|, wR_2 = [Σ[w(F_o^2 F_c^2)^2]/Σ[w(F_o^2)^2]]$^{1/2}$.

Table S2. Crystallographic data for **TBrC** and **TBrM**.

	TBrC	**TBrM**
Formula	$C_{23}H_{17}N_3OBr_2$	$C_{24}H_{21}N_3O_2Br_2$
Formula weight	511.21	543.26
Crystal system	Triclinic	Monoclinic
Space group	*P-1*	*C2/c*
a/Å	6.1318(7)	22.700(4)
b/Å	9.8814(11)	10.680(2)
c/Å	18.050(2)	18.910(4)
α/°	93.461(2)	-
β/°	93.268(2)	97.607(4)
γ/°	90.630(2)	-
V/Å^3	1089.8(2)	4544(2)
Z	2	8
D_c/g cm^{-3}	1.558	1.588
F(000)	508	2176
μ/mm^{-1}	3.738	3.594
$\theta_{min,max}$/°	1.132, 27.207	1.810, 28.347
Reflections collected	4832	30987
Independent reflections	4832 (*R*int = 0.0472)	5651(*R*int = 0.0569)
Observed reflections	3510	4359
Data/restr./param.	4832 / 0 / 266	5651 / 0 / 288
S[a]	1.020	1.018
R [Fo>4σ(Fo)][b], w*R*2[b]	0.0466, 0.0914	0.0332, 0.0710
$\Delta\rho_{min,max}$/e Å^{-3}	0.859, -0.660	0.738, -0.613

[a]Goodness-of-fit S = [$\Sigma w(F_o^2-F_c^2)^2/(n-p)$]1/2, where n is the number of reflections and p the number of parameters. [b]$R_1 = \Sigma \| F_o | - | F_c \| / \Sigma | F_o |$, $wR_2 = [\Sigma[w(F_o^2 F_c^2)^2]/\Sigma[w(F_o^2)^2]]^{1/2}$.

Table S3. Crystallographic data for **TN** and **TA**.

	TN	**TA**
Formula	$C_{23}H_{21}N_5O$	$C_{25}H_{21}N_3O_6$
Formula weight	383.45	459.45
Crystal system	Triclinic	Triclinic
Space group	*P-1*	*P-1*
a/Å	8.840(1)	8.0663(3)
b/Å	11.231(2)	11.6644(4)
c/Å	11.328(2)	12.8449(4)
α/°	94.917(2)	99.567(1)
β/°	105.702(2)	99.312(1)
γ/°	107.879(3)	101.369(1)
V/Å^3	1012.7(3)	1144.46(7)
Z	2	2
D_c/g cm^{-3}	1.257	1.333
F(000)	404	480
μ/mm^{-1}	0.081	0.097
$\theta_{min,max}$/°	1.900, 30.855	1.641, 26.058
Reflections collected	16787	21133
Independent reflections	6280 (*R*int = 0.0301)	4509 (*R*int = 0.0395)
Observed reflections	3845	3975
Data/restr./param.	6280 / 0 / 346	4509 / 0 / 388
S[a]	1.028	1.081
R [Fo>4σ(Fo)][b], w*R*$_2$[b]	0.0577, 0.1489	0.0419, 0.1113
$\Delta\rho_{min,max}$/e Å^{-3}	0.275, -0.192	0.211, -0.289

[a]Goodness-of-fit S = $[\Sigma w(F_o^2-F_c^2)^2/(n-p)]1/2$, where n is the number of reflections and p the number of parameters. [b]$R_1 = \Sigma \| F_o | - | F_c \| / \Sigma | F_o |$, $wR_2 = [\Sigma[w(F_o^2 F_c^2)^2]/\Sigma[w(F_o^2)^2]]^{1/2}$.

Table S4. Crystallographic data for **TBr-DITFB** and **TN-DITFB**

	TBr-DITFB	**TN-DITFB**
Formula	$C_{30}H_{18}N_3OBr_2Cl_3F_4I_2$	$C_{35}H_{23}N_5O_2F_8I_4$
Formula weight	1032.44	1205.18
Crystal system	Monoclinic	Hexagonal
Space group	*P21/c*	*P64*
a/Å	18.064(4)	32.4780(5)
b/Å	8.097(2)	-
c/Å	25.527(6)	6.6219(1)
α/°	-	-
β/°	101.216(4)	-
γ/°	-	-
V/Å^3	3663(1)	6049.1(2)
Z	4	6
D_c/g cm^{-3}	1.872	1.985
F(000)	1960	3408
μ/mm^{-1}	4.166	3.165
$\theta_{min,max}$/°	1.149, 26.369	2.508, 25.089
Reflections collected	26620	55937
Independent reflections	7457 (*R*int = 0.1018)	7147 (*R*int = 0.0483)
Observed reflections	4240	6602
Data/restr./param.	7457 / 0 / 373	7147 / 1 / 486
S[a]	1.003	1.005
R [Fo>4σ(Fo)][b], wR_2[b]	0.0479, 0.0808	0.0296, 0.0656
$\Delta\rho_{min,max}$/e Å^{-3}	1.004, -0.862	0.777, -0.420
Absolute structure parameter	-	-0.012(9)

[a]Goodness-of-fit S = $[\Sigma w(F_o^2-F_c^2)^2/(n-p)]1/2$, where n is the number of reflections and p the number of parameters. [b]$R_1 = \Sigma \| F_o | - | F_c \| / \Sigma | F_o |$, $wR_2 = [\Sigma[w(F_o^2 F_c^2)^2]/\Sigma[w(F_o^2)^2]]^{1/2}$.

Table 6. Crystallographic data for **TT-DITFB**

	TT-DITFB
Formula	$C_{39}H_{19}N_3OF_8I_4$
Formula weight	1205.17
Crystal system	Monoclinic
Space group	*P21/c*
a/Å	15.191(2)
b/Å	13.537(2)
c/Å	19.814(3)
α/°	-
β/°	104.127(6)
γ/°	-
V/Å^3	3951(1)
Z	4
D_c/g cm^{-3}	2.026
F(000)	2264
μ/mm^{-1}	3.228
$\theta_{min,max}$/°	2.691, 27.481
Reflections collected	56420
Independent reflections	9040(*R*int = 0.0247)
Observed reflections	8084
Data/restr./param.	9040 / 0 / 499
S[a]	1.004
R [Fo>4σ(Fo)][b], w*R*2[b]	0.0253, 0.0730
$\Delta\rho_{min,max}$/e Å^{-3}	2.537, -2.260

[a]Goodness-of-fit S = [Σw(F_o^2-F_c^2)^2/ (n-p)]1/2, where n is the number of reflections and p the number of parameters. [b]*R*1 = Σ‖F_o| - |F_c‖/Σ|F_o|, *w*R2 = [Σ[w(F_o^2 F_c^2)^2]/Σ[w(F_o^2)^2]]$^{1/2}$.

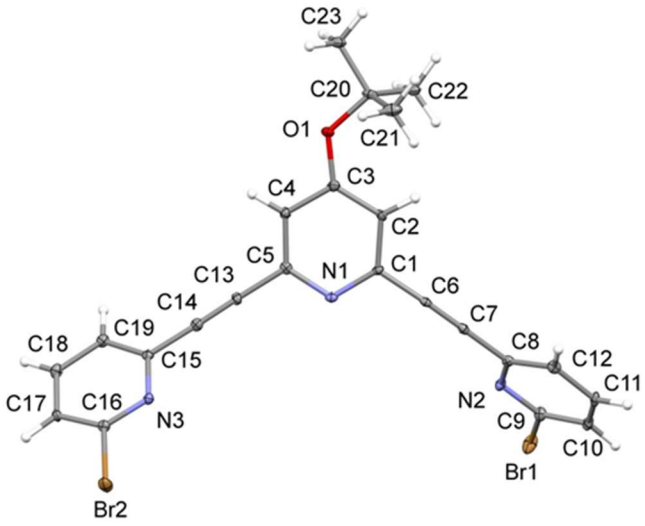


Figure S1. Ortep view of **TBrT** with displacement ellipsoids drawn at the 20% probability level.

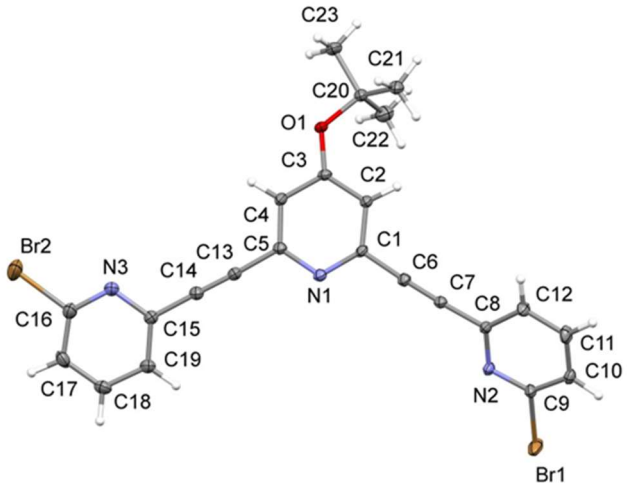


Figure S2. Ortep view of **TBrB** with displacement ellipsoids drawn at the 20% probability level.

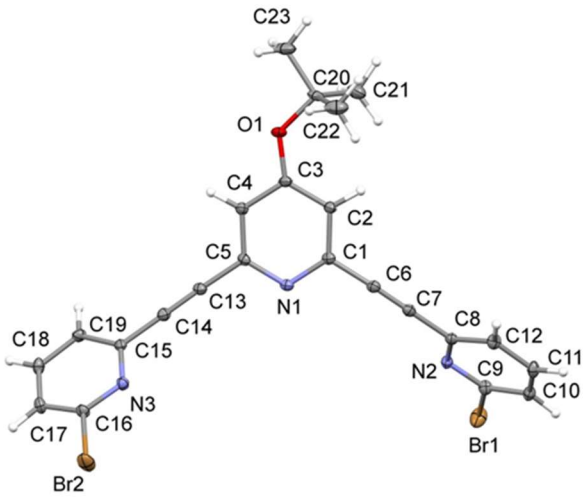


Figure S3. Ortep view of **TBrC** with displacement ellipsoids drawn at the 20% probability level.

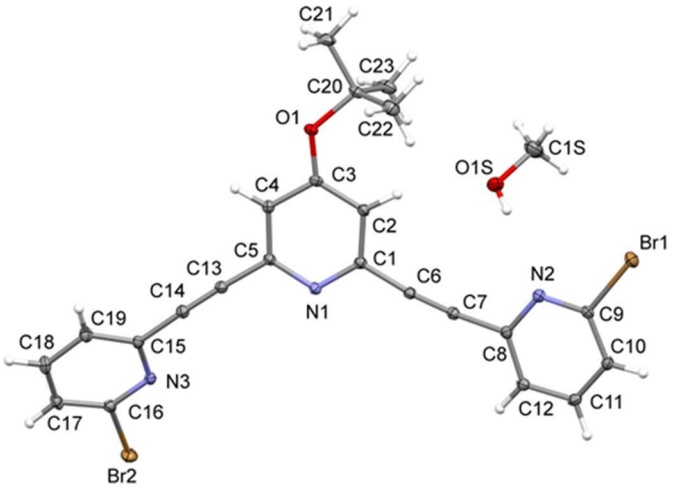


Figure S4. Ortep view of **TBrM** with displacement ellipsoids drawn at the 20% probability level.

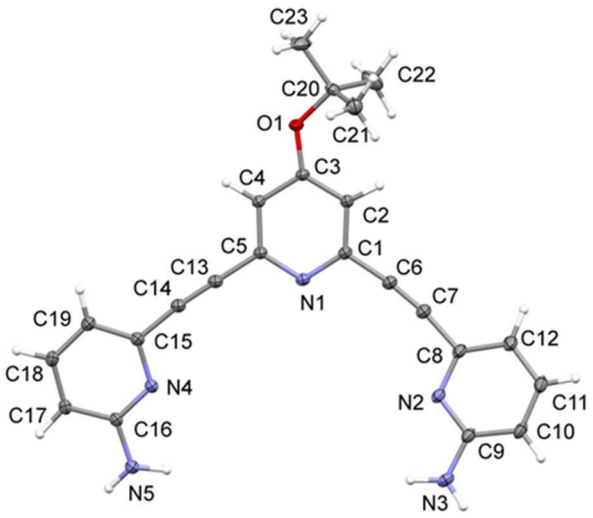


Figure S5. Ortep view of **TN** with displacement ellipsoids drawn at the 20% probability level.

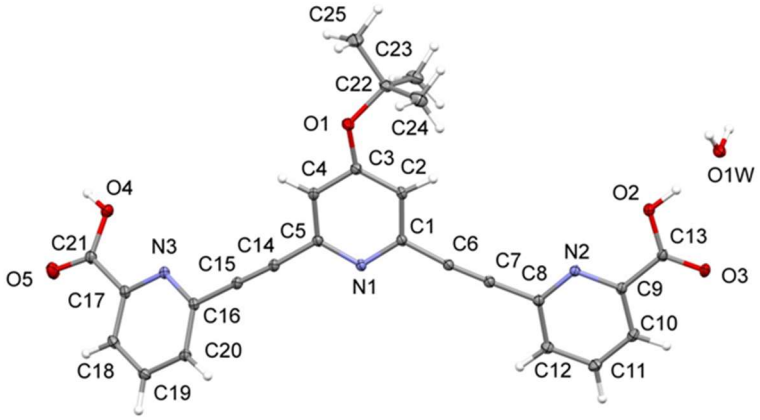


Figure S6. Ortep view of **TA** with displacement ellipsoids drawn at the 20% probability level.

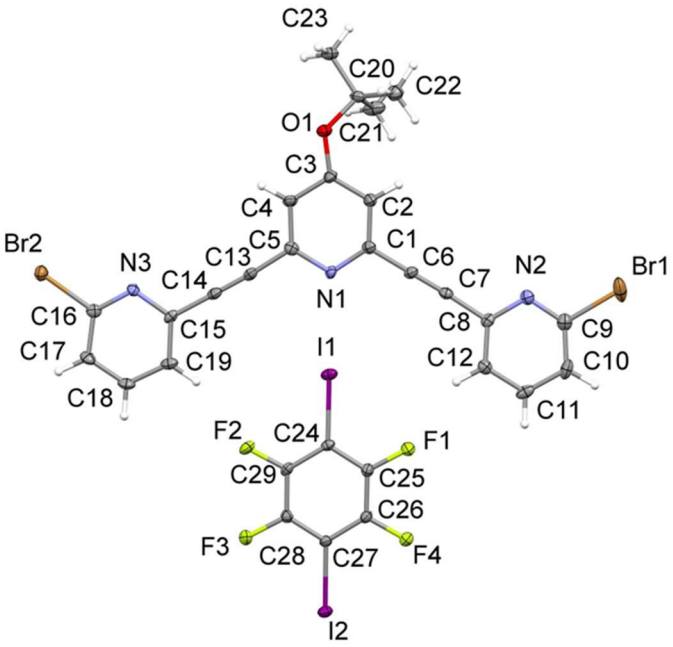


Figure S7. Ortep view of **TBr-DITFB** with displacement ellipsoids drawn at the 20% probability level.

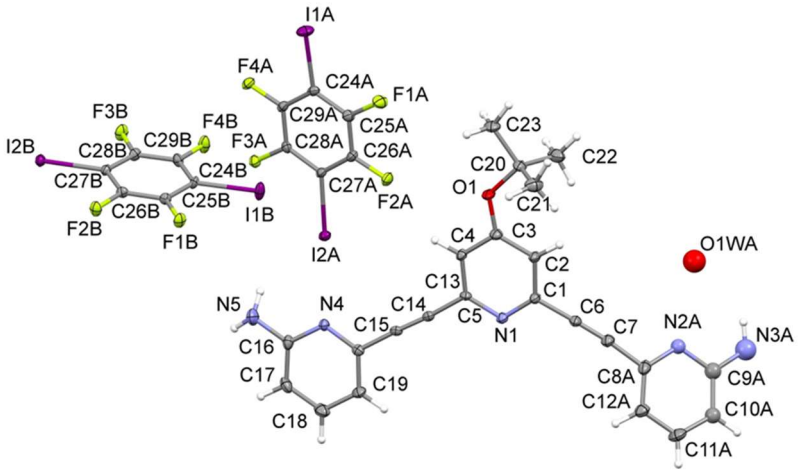


Figure S8. Ortep view of **TN-DITFB** with displacement ellipsoids drawn at the 20% probability level.

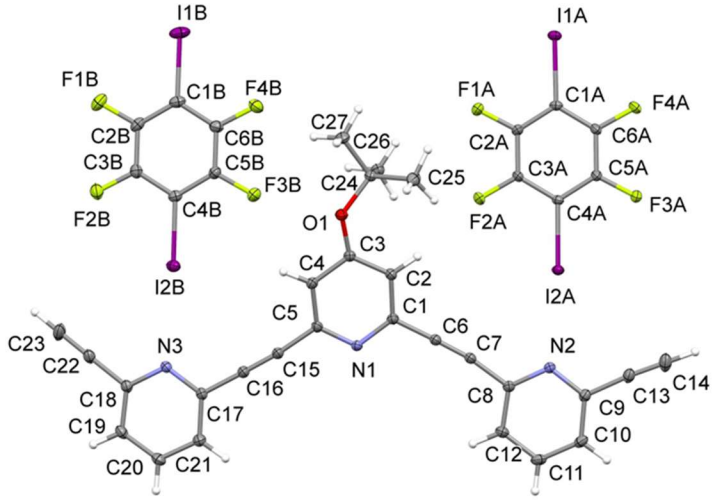


Figure S9. Ortep view of **TT-DITFB** with displacement ellipsoids drawn at the 20% probability level.

X-ray Crystallography for Section 4: Fluorine-pyridine ligand

The crystal structures of $[Ag(TF)_2]PF_6$ was determined by X-ray diffraction methods. Crystal data and experimental details for data collection and structure refinement are reported in Table 7. Intensity data and cell parameters were recorded at 190(2) K on a Bruker D8 Venture PhotonII diffractometer equipped with a CCD area detector, using the MoKα radiation (λ = 0.71073 Å). The raw frame data were processed using SAINT and SADABS to yield the reflection data file.[41] The structures were solved by Direct Methods using the SIR97 program[42] and refined on F_o^2 by full-matrix least-squares procedures, using SHELXL-2014[43] in the WinGX suite v.2014.1.[44] All non-hydrogen atoms were refined with anisotropic atomic displacements; the hydrogen atoms were included in the refinement at idealized geometry (C-H 0.95-0.99 Å) and refined "riding" on the corresponding parent atoms with $U_{iso}(H)$ set to $1.2U_{eq}(C)$ and $1.5U_{eq}(C_{met})$. The weighting schemes used in the last cycle of refinement was $w = 1/ [\sigma^2 F_o^2 + (0.1467P)^2]$, where $P = (F_o^2 + 2F_c^2)/3$.

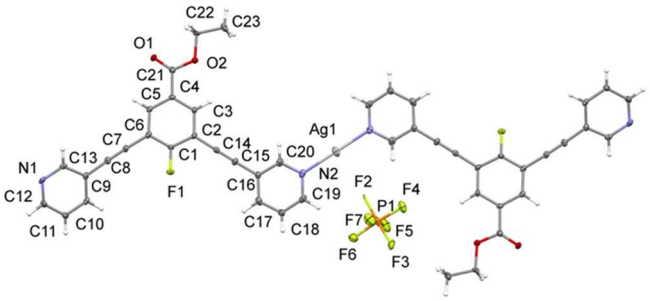


Figure S10. Ortep view of $[Ag(TF)_2]PF_6$, with displacement ellipsoids drawn at the 20% probability level. The atoms which are not labelled are generated by symmetry (-x, -y, -z).

Table S6. Crystallographic data for $[Ag(TF)_2]PF_6$.

	$[Ag(TF)_2]PF_6$
Formula	$C_{46}H_{30}F_8N_4O_4PAg$
Formula weight	257.50
Crystal system	Monoclinic
Space group	*C2/c*
a/Å	51.2626(3)
b/Å	3.9700(4)
c/Å	21.2553(4)
β/°	91.829(1)
V/Å^3	4323.5(4)
Z	4
D_c/g cm^{-3}	1.526
F(000)	2000
μ/mm^{-1}	0.586
$\theta_{min,max}$/°	2.052, 23.414
Reflections collected	15442
Independent reflections	3141 (Rint = 0.1950)
Observed reflections	1553
Data/restr./param.	3141/ 0 / 323
S[a]	1.003
R [Fo>4σ(Fo)][b], wR_2[b]	0.0759, 0.2032
$\Delta\rho_{min,max}$/e Å^{-3}	0.956, -0.693

[a]Goodness-of-fit S = $[\Sigma w(F_o^2-F_c^2)^2/(n-p)]1/2$, where n is the number of reflections and p the number of parameters. [b]$R_1 = \Sigma \| F_o | - | F_c \| / \Sigma | F_o |$, $wR_2 = [\Sigma[w(F_o^2 F_c^2)^2]/\Sigma[w(F_o^2)^2]]^{1/2}$.

X-ray Crystallography for Section 5: Coordination polymers of zinc with pyridine-carboxylic ligands

PXRD spectra were measured at room temperature with a diffractometer X'TRA (THERMO) using a Bragg-Brentano geometry θ-θ and Cu Kα radiation (40kV-30mA), equipped with a solid state Si(Li) detector. The samples were placed in a zero-background sample holder made of silica. The spectra were collected in a 2θ 3-40 ° range, with steps of 0.05°, 2 seconds per step. Some of the spectra were collected with a Bruker D2 diffractometer (Cu radiation) in collaboration with Dr.ssa Luciana Mantovani, using similar experimental conditions.

The crystal structures of **complex1** (**C1**), **complex2** (**C2**), **CP1** and **CP3** were determined by X-ray diffraction methods. Crystal data and experimental details for data collection and structure refinement are reported in Tables S7 and S8. Intensity data and cell parameters were recorded at 190(2) K on a Bruker Apex II diffractometer equipped with a CCD area detector, using the MoKα radiation (λ= 0.71073 Å) for **C1**, **C2** and **CP1**, while for **CP3** data were collected at 100(2) K at the Synchrotron ELETTRA (CNR Trieste, strada statale 14, Area Science Park, 34149, Basovizza, Trieste). The raw frame data were processed using SAINT and SADABS to yield the reflection data file.[41] The structures were solved by Direct Methods using the SIR97 program[42] and refined on F_o^2 by full-matrix least-squares procedures, using SHELXL-2014[43] in the WinGX suite v.2014.1.[44]

All non-hydrogen atoms were refined with anisotropic atomic displacements, except when disorder was present. The hydrogen atoms were included in the refinement at idealized geometry (C-H 0.95-0.99 Å, O-H 0.82 Å) and refined "riding" on the corresponding parent atoms with

U_{iso}(H) set to 1.2U_{eq}(C) and 1.5U_{eq}(O, C_{met}). When possible, the H atoms of water molecules were found in the difference Fourier map. The weighting schemes used in the last cycle of refinement were $w = 1/ [\sigma^2 F_o^2 + (0.0656P)^2 + 1.3010P]$, $w = 1/ [\sigma^2 F_o^2 + (0.0618P)^2 + 1.3392P]$, $w = 1/ [\sigma^2 F_o^2 + (0.0372P)^2 + 4.5618P]$ and $w = 1/ [\sigma^2 F_o^2 + (0.2000P)^2]$ for **C1**, **C2**, **CP1** and **CP3**, respectively, where $P = (F_o^2 + 2F_c^2)/3$.

Table S7. Crystallographic data for **[Zn(HPDC)$_2$]·4H$_2$O** (**C1**) and **[Zn(PDC)(H$_2$O)$_2$]·(C2).**

	C1	**C2**
Formula	$C_{14}H_{16}N_2O_{12}Zn$	$C_7H_7NO_6Zn$
Formula weight	471.67	266.51
Crystal system	Triclinic	Orthorhombic
Space group	*P-1*	*Pbcm*
a/Å	10.815(1)	10.4661)
b/Å	13.361(2)	12.695(2)
c/Å	14.426(2)	7.1870(1)
α/°	64.724(2)	-
β/°	78.981(2)	-
γ/°	88.767(2)	-
V/Å^3	1845.9(4)	954.9(2)
Z	4	4
D_c/g cm^{-3}	1.697	1.854
F(000)	968	536
μ/mm^{-1}	1.4	2.580
$\theta_{min,max}$/°	1.594, 27.565	1.946, 30.508
Reflections collected	22119	13241
Independet reflections	8512 (*R*int = 0.0457)	1574 (*R*int = 0.0364)
Observed reflections	6358	1348
Data/restr./param.	8512 / 0 / 513	1574 / 0 / 105
S[a]	1.008	1.003
R [Fo>4σ(Fo)][b], w*R*$_2$[b]	0.0465, 0.1083	0.0339, 0.0969
$\Delta\rho_{min,max}$/e Å^{-3}	0.780, -0.731	0.954, -0.710

[a]Goodness-of-fit S = [Σw(F_o^2-F_c^2)^2/ (n-p)]1/2, where n is the number of reflections and p the number of parameters. [b]R_1 = Σ‖F_o| - |F_c‖/Σ|F_o|, wR_2 = [Σ[w(F_o^2 F_c^2)^2]/Σ[w(F_o^2)^2]]$^{1/2}$.

Table S8. Crystallographic data for **{[Zn(PDC)(H$_2$O)]·0.5H$_2$O}$_n$** (**CP1**) and **{[Zn(TA)(H$_2$O)]·H$_2$O·MeOH·DMF}$_n$** (**CP3**)

	CP1	**CP3**
Formula	$C_7H_6NO_{5.5}Zn$	$C_{27}H_{32}N_4O_9Zn$
Formula weight	257.50	621.95
Crystal system	Monoclinic	Monoclinic
Space group	*C2/c*	*C2/c*
a/Å	21.600(6)	21.4595(6)
b/Å	5.658(2)	21.0853(4)
c/Å	15.754(4)	13.7037(4)
α/°	-	-
β/°	115.009(4)	107.242(3)
γ/°	-	-
V/Å^3	1744.8(9)	5922.0(3)
Z	8	8
D_c/g cm^{-3}	1.960	1.329
F(000)	1032	2592
μ/mm^{-1}	2.816	0.886
$\theta_{min,max}$/°	2,081, 28.391	1.823, 32.935
Reflections collected	11498	57557
Independent reflections	2177 (*R*int = 0.0419)	9850 (*R*int = 0.0419)
Observed reflections	1852	8292
Data/restr./param.	2177 / 0 / 144	9850 / 0 / 352
S[a]	1.003	1.008
R [Fo>4σ(Fo)][b], wR_2[b]	0.0326, 0.0745	0.0985, 0.3338
$\Delta\rho_{min,max}$/e Å^{-3}	0.614, -0.610	2.426, -1.546

[a]Goodness-of-fit S = [Σw(F_o^2-F_c^2)^2/ (n-p)]1/2, where n is the number of reflections and p the number of parameters. [b]R_1 = Σ‖F_o| - |F_c‖/Σ|F_o|, wR_2 = [Σ[w(F_o^2 F_c^2)^2]/Σ[w(F_o^2)^2]]$^{1/2}$.

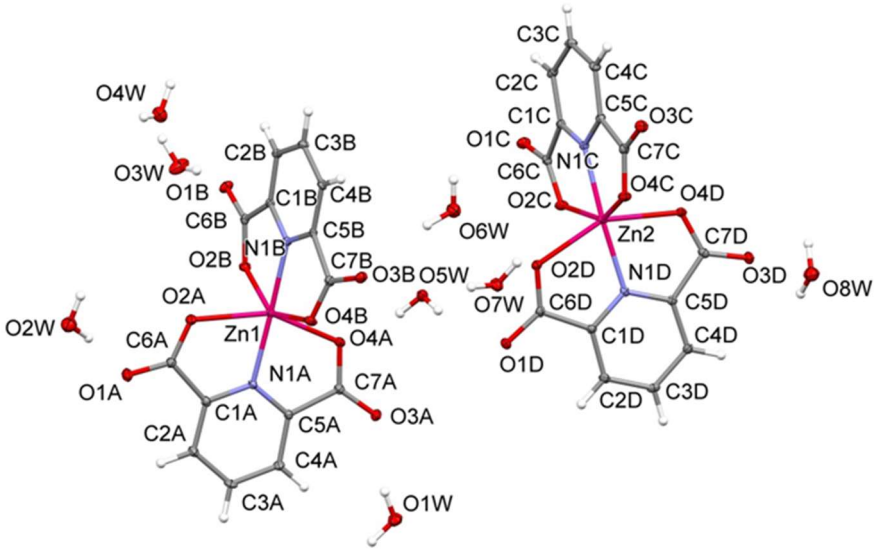


Figure S11. Ortep view of the two independent molecules of **C1**; displacement ellipsoids are drawn at the 20% probability level.

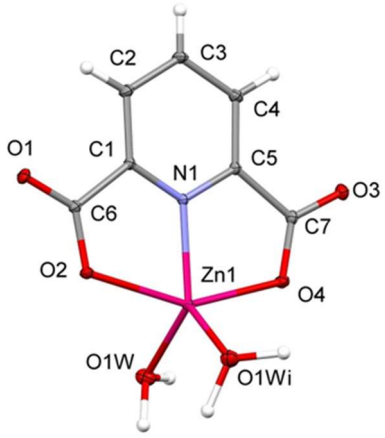


Figure S12. Ortep view of **C2** highlighting the coordination around the metal centre; displacement ellipsoids are drawn at the 20% probability level. i = x, y, 1/2-z.

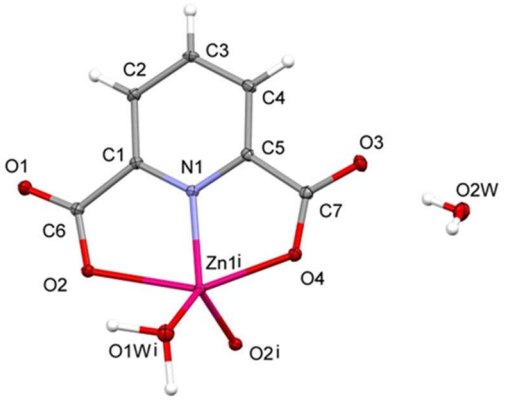


Figure S13. Ortep view of **CP1** showing the coordination sphere around the metal centre; displacement ellipsoids are drawn at the 20% probability level. i = 3/2-x, 1/2+y, 3/2-z.

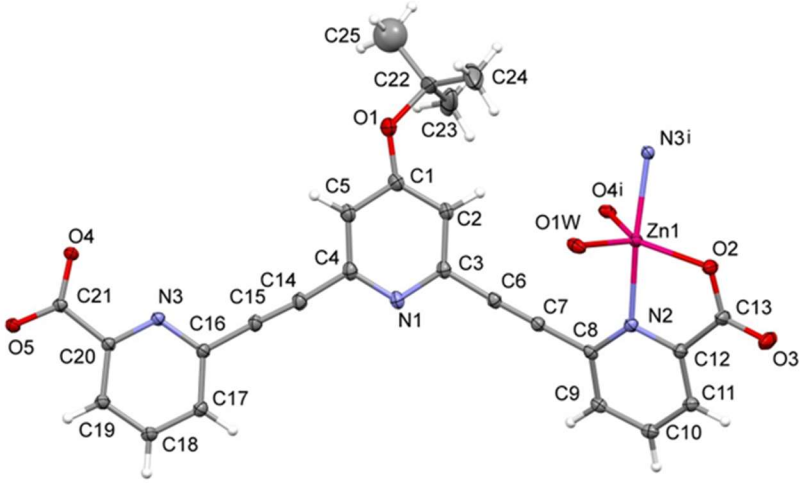


Figure S14. Ortep view of **CP3** highlighting the coordination sphere around the metal centre; displacement ellipsoids are drawn at the 20% probability level. Solvent molecules have been omitted for clarity. i = 1/2-x, -1/2+y, 3/2-z.

9. Bibliography

1 P. Liao, B. W. Langloss, A. M. Johnson, E. R. Knudsen, F. S. Tham, R. R. Julian and R. J. Hooley, *Chem. Commun.*, 2010, **46**, 4932–4934.

2 K. J. Kilpin, M. L. Gower, S. G. Telfer, G. B. Jameson and J. D. Crowley, *Inorg. Chem.*, 2011, **50**, 1123–1134.

3 C. Gütz, R. Hovorka, G. Schnakenburg and A. Lützen, *Chem. - A Eur. J.*, 2013, **19**, 10890–10894.

4 D. P. August, G. S. Nichol and P. J. Lusby, *Angew. Chemie - Int. Ed.*, 2016, **55**, 15022–15026.

5 L. Wang, Z. Zhang, X. Jiang, J. A. Irvin, C. Liu, M. Wang and X. Li, *Inorg. Chem.*, 2018, **57**, 3548–3558.

6 J. R. Li, A. A. Yakovenko, W. Lu, D. J. Timmons, W. Zhuang, D. Yuan and H. C. Zhou, *J. Am. Chem. Soc.*, , DOI:10.1021/ja1080794.

7 D. Preston, S. M. McNeill, J. E. M. Lewis, G. I. Giles and J. D. Crowley, *Dalt. Trans.*, 2016, **45**, 1–29.

8 F. Kaiser, A. Schmidt, W. Heydenreuter, P. J. Altmann, A. Casini, S. A. Sieber and F. E. Kuehn, *Eur. J. Inorg. Chem.*, 2016, **2016**, 5181.

9 A. Schmidt, V. Molano, M. Hollering, A. Pöthig, A. Casini and F. E. Kühn, *Chem. - A Eur. J.*, 2016, **22**, 2253–2256.

10 J. Han, A. Schmidt, T. Zhang, H. Permentier, G. M. M. Groothuis, R. Bischoff, F. E. Kühn, P. Horvatovich and A. Casini, *Chem. Commun.*, , DOI:10.1039/C6CC08937B.

11 V. Martí-Centelles, A. L. Lawrence and P. J. Lusby, *J. Am. Chem.*

Soc., 2018, **140**, 2862–2868.

12 J. E. M. Lewis, C. J. McAdam, M. G. Gardiner and J. D. Crowley, *Chem. Commun.*, 2013, **49**, 3398–3400.

13 D. Preston, A. Fox-Charles, W. K. C. Lo and J. D. Crowley, *Chem. Commun.*, 2015, **51**, 9042–9045.

14 D. Preston, J. E. M. Lewis and J. D. Crowley, *J. Am. Chem. Soc.*, 2017, **139**, 2379–2386.

15 B. Som, S. R. Salpage, J. Son, B. Gu, S. G. Karakalos, M. D. Smith and L. S. Shimizu, *CrystEngComm*, 2017, **19**, 484–491.

16 K. Roy, C. Wang, M. D. Smith, M. B. Dewal, A. C. Wibowo, J. C. Brown, S. Ma and L. S. Shimizu, *Chem. Commun*, 2011, **47**, 277–279.

17 S. Dawn, M. B. Dewal, D. Sobransingh, M. C. Paderes, A. C. Wibowo, M. D. Smith, J. A. Krause, P. J. Pellechia and L. S. Shimizu, *J. Am. Chem. Soc.*, 2011, **133**, 7025–7032.

18 S. R. Salpage, Y. Xu, B. Som, A. J. Sindt, M. D. Smith and L. S. Shimizu, *RSC Adv.*, 2016, **6**, 98350–98355.

19 O. B. Berryman, C. A. Johnson, L. N. Zakharov, M. M. Haley and D. W. Johnson, *Angew. Chemie - Int. Ed.*, 2008, **47**, 117–120.

20 C. N. Carroll, O. B. Berryman, C. A. Johnson, L. N. Zakharov, M. M. Haley and D. W. Johnson, *Chem. Commun.*, 2009, 2520–2522.

21 J. M. Engle, P. S. Lakshminarayanan, C. N. Carroll, L. N. Zakharov, M. M. Haley and D. W. Johnson, *Cryst. Growth Des.*, 2011, **11**, 5144–5152.

22 J. M. Engle, C. N. Carroll, D. W. Johnson and M. M. Haley, *Chem. Sci.*, 2012, **3**, 1105.

23 J. M. Engle, P. S. Singh, C. L. Vonnegut, L. N. Zakharov, D. W. Johnson and M. M. Haley, *CrystEngComm*, 2014, **16**, 3703–3706.

24 C. L. Vonnegut, A. M. Shonkwiler, L. N. Zakharov, M. M. Haley and D. W. Johnson, *Chem. Commun*, 2016, **52**, 9506–9509.

25 A. Marie, S. Riel, D. A. Decato, J. Sun, C. J. Massena, M. J. Jessop and O. B. Berryman, *Chem. Sci.*, 2018, 5828–5836.

26 H. Abe, H. Kurokawa, Y. Chida and M. Inouye, *J. Org. Chem.*, 2011, **76**, 309–311.

27 K. Raatikainen and K. Rissanen, *Chem. Sci.*, 2012, **3**, 1235–1239.

28 W. X. Wu, H. Wang and W. J. Jin, *Cryst. Growth Des.*, 2018, acs.cgd.8b01021.

29 H.-C. Zhou, J. R. Long and O. M. Yaghi, *Chem. Rev.*, 2012, **112**, 673–674.

30 S. R. Batten, N. R. Champness, X.-M. Chen, J. Garcia-Martinez, S. Kitagawa, L. Öhrström, M. O'Keeffe, M. P. Suh and J. Reedijk, *CrystEngComm*, 2012, **14**, 3001.

31 Y. V. Kaneti, J. Tang, R. R. Salunkhe, X. Jiang, A. Yu, K. C.-W. Wu and Y. Yamauchi, *Adv. Mater.*, 2017, **29**, 1604898.

32 D. Farrusseng, *Metal-organic frameworks : applications from catalysis to gas storage*, Wiley-VCH, 2011.

33 N. Stock and S. Biswas, *Chem. Rev.*, 2012, **112**, 933–969.

34 S. M. Cohen, *Chem. Rev.*, 2012, **112**, 970–1000.

35 V. Bernales, M. A. Ortuño, D. G. Truhlar, C. J. Cramer and L. Gagliardi, *ACS Cent. Sci.*, 2018, **4**, 5–19.

36 J. L. Howard, Q. Cao and D. L. Browne, *Chem. Sci.*, 2018, **9**, 3080–3094.

37 J. L. Do and T. Friščić, *ACS Cent. Sci.*, 2017, **3**, 13–19.

38 S. L. James, C. J. Adams, C. Bolm, D. Braga, P. Collier, T. Friščić, F. Grepioni, K. D. M. Harris, G. Hyett, W. Jones, A. Krebs, J. Mack, L. Maini, A. G. Orpen, I. P. Parkin, W. C. Shearouse, J. W. Steed and D. C. Waddell, *Chem. Soc. Rev.*, 2012, **41**, 413–447.

39 T. Friščić, D. G. Reid, I. Halasz, R. S. Stein, R. E. Dinnebier and M. J. Duer, *Angew. Chemie Int. Ed.*, 2010, **49**, 712–715.

40 T. Friščić and L. Fábián, *CrystEngComm*, 2009, **11**, 743.

41 (a) SADABS Bruker AXS; Madison, Wisconsin, USA, 2004; SAINT, Software Users Guide, Version 6.0; Bruker Analytical X-ray Systems, Madison, WI, 1999; (b) G. M. Sheldrick, SADABS v2.03: Area-Detector Absorption Correction. University of Göttingen, Germany, 1999.

42 A. Altomare, M. C. Burla, M. Camalli, G. L. Cascarano, C. Giacovazzo, A. Guagliardi, A. G. G. Moliterni, G. Polidori, R. Spagna, *J. Appl. Crystallogr.* 1999, **32**, 115-119.

43 G. M. Sheldrick, *Acta Cryst.* 2008, **A64**, 112-122.

44 WinGX, L. J. Farrugia, *J. Appl. Crystallogr.*, 1999, **32**, 837-838.

45 SQUEEZE, P. v.d. Sluis and A. L. Spek, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 1990, **46**, 194-201.

CHAPTER 2

Tetra-phosphonate cavitands as ligands for metal coordination

1. Macrocyclic compounds as ligands for metal complexation

1.1. Complexes, cages and capsules

Coordination complexes based on macrocyclic ligands (Figure 1), have been widely studied and found application in many fields, including mimic biological systems,[1–3] catalysis,[4–6] molecular machines[7] and gas adsorption and storage.[8] Different macrocycle classes have been employed for this purpose, including calixarenes, resorcinarene-based cavitands and cyclotriveratrylene-based cavitands.

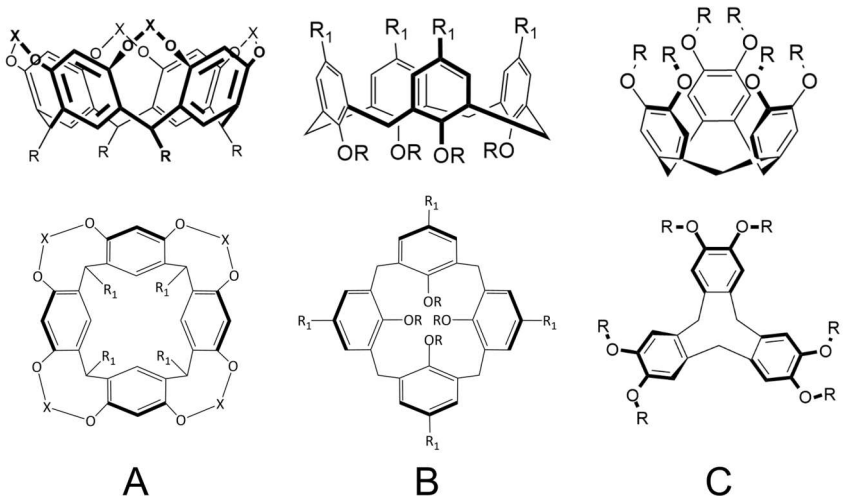


Figure 1. Examples of macrocyclic compounds used as building blocks for metal coordination; side views (top panels) and top views (bottom panels). Generic (A) resorcinarene-cavitand. X = P=OEt, CH_2, Quinoxaline, etc. (B) Calix[4]arene and (C) cyclotriveratrylene.

Examples with calixarenes are ubiquitous, but the work of the Reinaud group is of particular high-profile. In their numerous reports,[1,9] the authors use calix[4]arenes and calix[6]arenes to mimic metalloenzymes (Figure 2). They synthesized a library of calixarenes functionalized in order to bind zinc and copper and studied the host-guest chemistry of the resulting funnel complexes. Their research focused on how the macrocyclic ligand affected metal coordination properties, taking into account 1st, 2nd and 3rd coordination spheres and hydrophobic effect. Furthermore, in a recent paper,[2] they achieved water solubility by employing a resorcinarene-based cavitand.

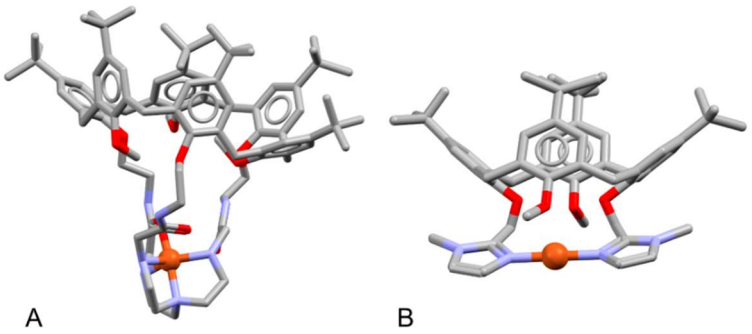


Figure 2. (A) Calix[6]arene and (B) calix[4]arene-copper (I) funnel complexes (CCDC number 1541407 and 138488, respectively). The flexibility of the calixarene scaffold is an ideal characteristic in enzyme mimic as it can aid guest binding. Copper (I) is shown as an orange sphere. Counter ions, split conformations and solvent molecules are omitted for clarity.

As expected, coordination complexes based on macrocyclic ligands have been exploited as catalysts.[6] Matt and co-workers have extensively worked on the functionalization of resorcinarene cavitands to bind nickel, palladium and platinum.[10–12] They successfully carried out coupling reactions catalysed by cavitand-palladium complexes.[11] Resorcinarene-

cavitands have also been proven useful in oxidation reactions. In an example from the Hooley group, the grafting of a cavitand-iron (II) complex on a silica surface resulted in an heterogeneous catalyst that oxidizes cyclic hydrocarbons.[4] The reaction was performed under mild conditions, in aqueous media with hydrogen peroxide as the oxidant. Deep cavitands were also exploited to mimic biocatalytic oxidations.[5,13]

Recently, the Kobayashi group proved how coordination-driven self-assembly can be employed to achieve highly sophisticated molecular machines.[7] The authors reported a molecular gear constructed on a tetra(4-pyridyl) cavitand, with each pyridine coordinated to a rhodium-porphyrin (Figure 3). The gear rotates both in solution and in the solid state by changes of temperature, as proved by NMR and XRD studies.

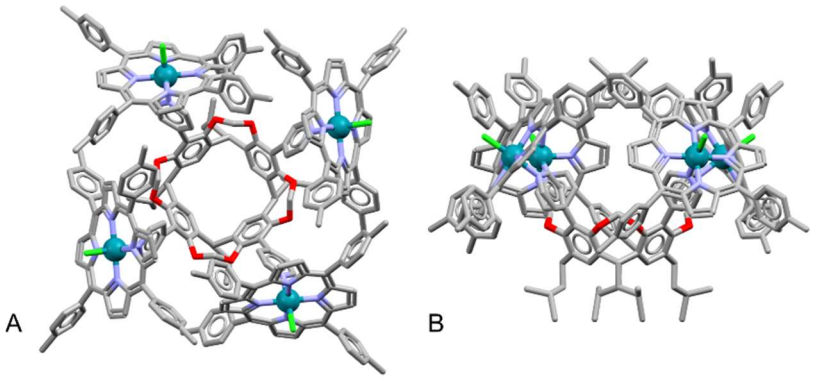


Figure 3. Tetra(4-pyridyl) cavitand-Rh-porphyrin molecular gear (CCDC number 1507633). Views (A) and (B) are rotated of 90°, respectively. Rhodium atoms are represented as teal blue spheres. Solvent molecules and H atoms are omitted for clarity.

Resorcinarene cavitands are ideal to construct dimeric rigid cages and capsules.[14] In the past, our group reported various pyridine-functionalized

cavitands that self-assembled with palladium (II), platinum (II) and rhenium (I) yielding the corresponding dimeric cage (Figure 4).[15–17] In a more recent case, the Kobayashi group showed a hybrid capsule,[18] assembled by a set of hydrogen-bonds and metal-ligand coordination bonds. Using NMR experiments, they were able to identify the hemi capsule, proving the assembly mechanism.

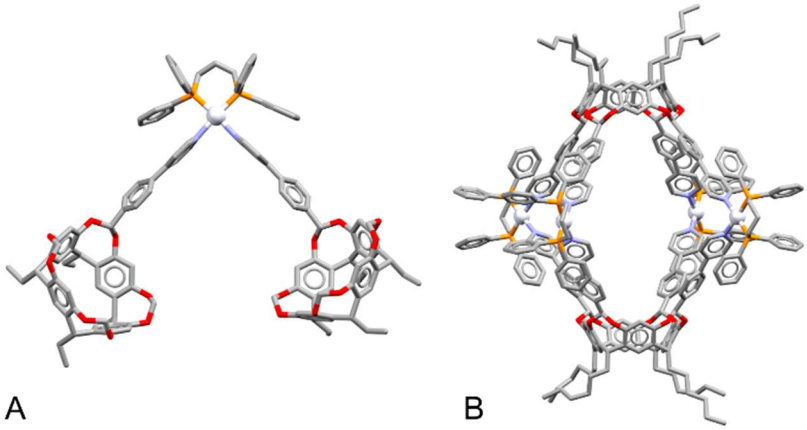


Figure 4. Cavitand-Pd(II) (A) dimeric capsule and (B) cage (CCDC number 252894 and 205091, respectively). Palladium atoms are represented as white spheres. Solvent molecules, H atoms and counter-ions are omitted for clarity.

1.2. The development of frameworks

In some cases, the 3D assembly of the complex plays a crucial role for the resulting properties of the material. Recently, Jian-Fang Ma and collaborators showed a strategy to self-assemble deep cavitands and cages

from a resorcinarene cavitand (Figure 5).[8] These complexes showed highly porous assemblies in the solid state. The high porosity of the crystals is partly due to the cavities of the macrocycles which result empty. The obtained network retained structural flexibility and, upon gas adsorption, was able to convert from a narrow-pore to an open-pore conformation. This outstanding example highlights the importance of developing a crystalline porous material exploiting cavitands.

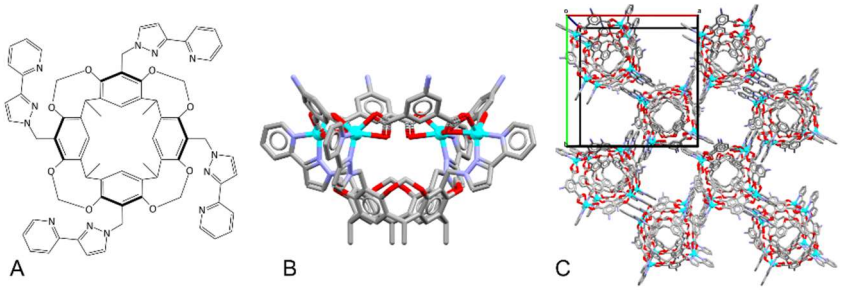


Figure 5. (A) The cavitand used as ligand in ref 8. (B) Structure of one of the Zn^{2+} complexes obtained (CCDC number 1559338). Zn^{2+} cations are represented as cyan spheres. (C) 3D assembly in the solid state. Solvent molecules and H atoms are omitted for clarity.

In fact, a major target in this field is the development of a MOF-like material employing macrocyclic receptors. The resulting hybrid material possesses two levels of porosity, one from the network and one from the cavities of the macrocycles themselves. During guest uptake, recognition properties are enhanced by the cavities and the material is ideal for sensing applications. Calixarenes appeared to be useful building blocks in this context (Figure 6).[19,20] A. Schaate and co-workers created a MOF from an acid-functionalized calix[4]arene and zinc (II) (Figure 6A).[20] They showed

how the macrocycle cavities are accessible through the pores system. Their material was employed for sensing NO_2.

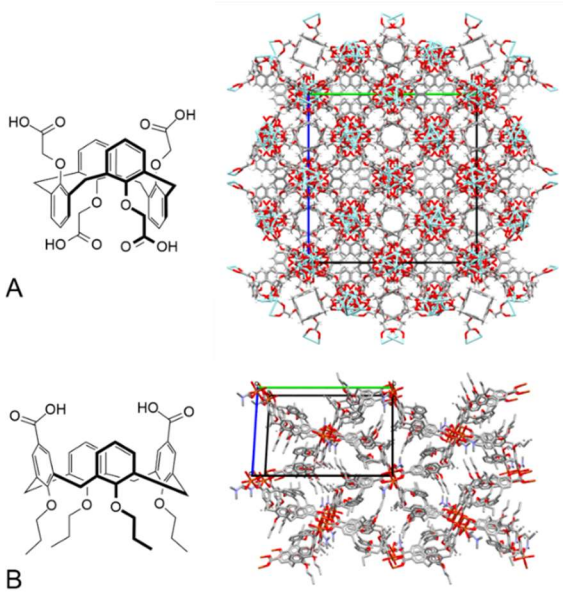


Figure 6. MOFs created from (A) a 1,3-alternate calix[a]arene ligand and zinc (II) (CCDC number 1830436) and (B) a cone calix[4]arene ligand and copper (II) (CCDC number 866579). Solvent molecules and H atoms are omitted for clarity.

While many examples with calixarenes, cucurbiturils and pillarenes have been reported so far,[21] examples involving resorcinarenes and cavitands are still rare. Nevertheless, strategies can in principle be extended to any macrocyclic building block.[21,22] Various groups have been working on the idea of using resorcinarenes and cavitands as ligands for MOF construction. The group of Jian-Fang Ma has been quite active in this field. They employed a functionalized resorcin[4]arene to develop metal organic frameworks (Figure 7).[23,24] The MOF showed good catalytic

properties. In these cases, the flexibility of the ligand aids self-assembly and the formation of different coordination geometries.

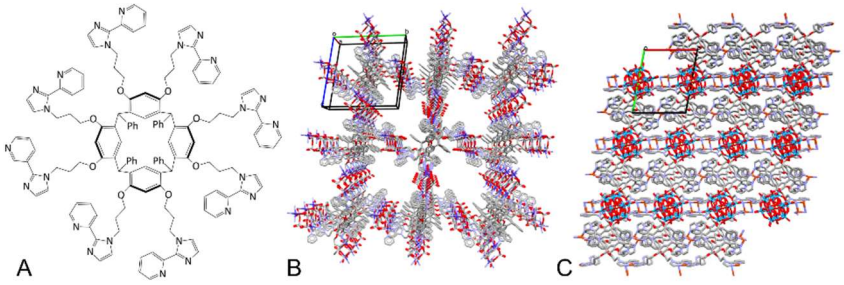


Figure 7. (A) Functionalized resorcin[4]arene ligand and MOFs obtained in the presence of (B) polyoxovanadate (CCDC number 1556497) and (C) polyoxometalate and copper (I).

The same group reported copper coordination polymers obtained from resorcinarene-based cavitands,[25,26] (Figure 8). The ligands were functionalized either at the upper or lower rim with 1,2,4-triazole[26] and carboxylic acid,[25] respectively. The coordination polymers showed interesting ion exchange and luminescence properties.

Cavitands can also be used as templates for the construction of a MOF, as shown by the Holman group.[27] The same concept can be employed in a mechanochemical synthesis.[28] This strategy leads to the incorporation of the cavitand in a 3D framework without the need for synthetic functionalization.

The complexes and coordination polymers described so far employ methylene bridged cavitands, which do not have specific recognition properties besides those provided by a π-basic cavity. Our hypothesis is that these systems would greatly benefit from the presence of an additional interaction site in the scaffold of the macrocycle. In our studies we

employed tetra-phosphonate cavitands, which provide peculiar guest recognition properties.[29–32]

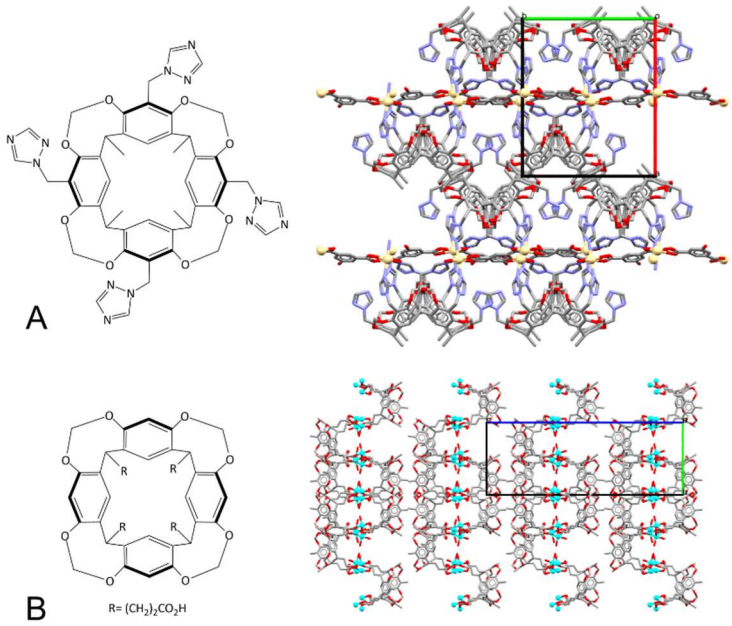


Figure 8. Coordination polymers created from (A) triazol-functionalized cavitand and cadmium (II) and (B) carboxylic acid-functionalized cavitand and zinc (II) (CCDC numbers 1426511 and 1012008, respectively). Solvent molecules and counter-ions are omitted for clarity.

2. The chosen approach: tetra-phosphonate cavitands

Tetra-phosphonate cavitands are bowl-shaped macrocycles constructed on a resorcin[4]arene scaffold. The resorcinarene hydroxyls are bridged with phosphonate groups pointing inward the cavity (Figure 9).

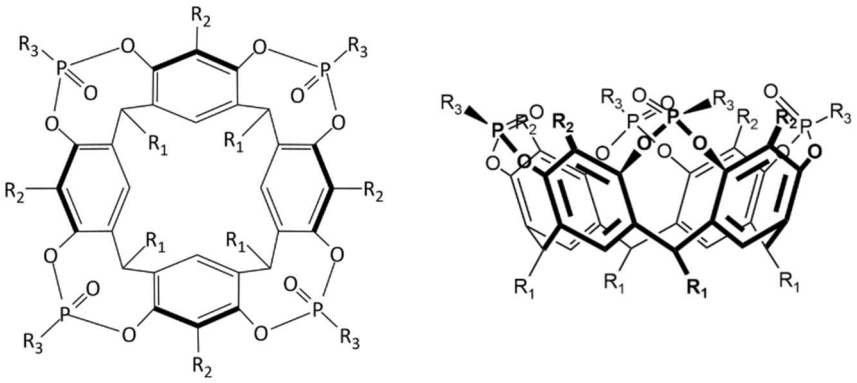


Figure 9. Top-view (left) and side-view (right) of a generic tetra-phosphonate cavitand.

The P=O groups provide an anchoring point for H-bond, dipole-dipole and cation-dipole interactions. Furthermore, the electron-rich cavity interacts with guests *via* C-H···π interactions. Hence, tetra-phosphonate cavitands are excellent receptors for methylated species, methyl-ammonium and methyl-pyridinium salts (Figure 10).[29–32]

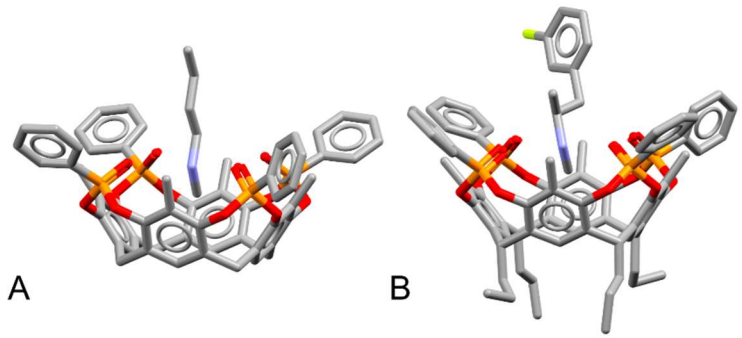


Figure 10. Tetra-phosphonate cavitands complexing (A) butyl-methyl-ammonium (CCDC number 731070) and (B) the illicit drug 3-fluoromethamphet-amine (3-FMA) (CCDC number 978408).

Metals can coordinate tetra-phosphonate cavitands through the P=O moieties. Indeed, our group showed the formation of dimeric capsules with barium, calcium, zinc,[33] lead and copper (Figure 11).[34]

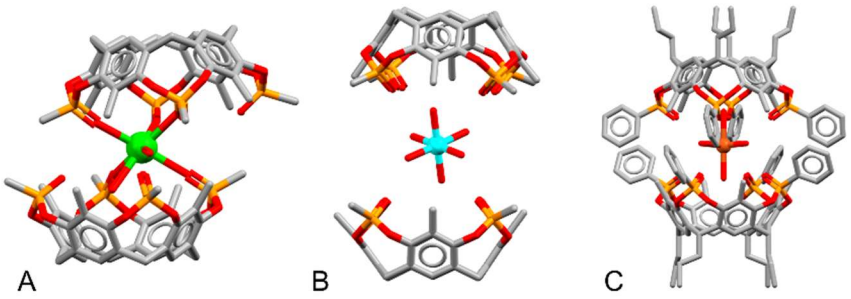


Figure 11. Tetra-phosphonate cavitands forming dimeric capsules with barium (CCDC number 767369), zinc (CCDC number 767371) and copper (CCDC number 1539301).

In the present work, we decided to functionalize tetra-phosphonate cavitands with moieties suitable for metal coordination, leaving the P=O groups free for guest complexation.

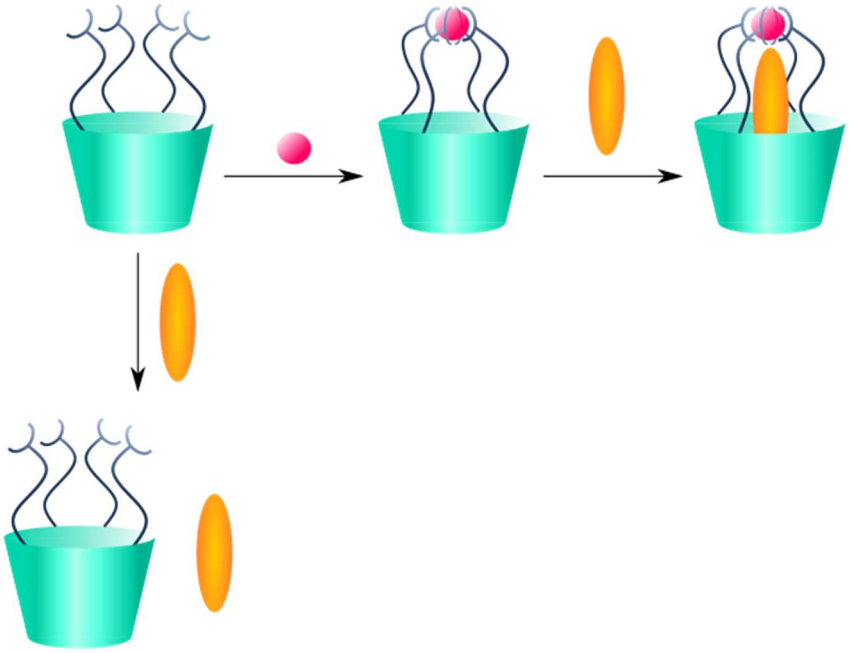


Figure 12. A functionalized cavitand (top left) forms a capsule upon metal (pink sphere) coordination. The guest (yellow oval) will be complexed by the capsule, but not by the pristine cavitand.

The ligands synthesized have been studied from two different points of view. Initially we aimed at creating a capsule suitable for guest complexation (Figure 12). The capsule would be ideal for guests that can be complexed by the cavity and coordinated by the metal. The affinity would be enhanced compared to the one of the cavitand itself. Our second goal was the study of a 3D assembly of these complexes and the development of a framework or of coordination polymers (Figure 13).

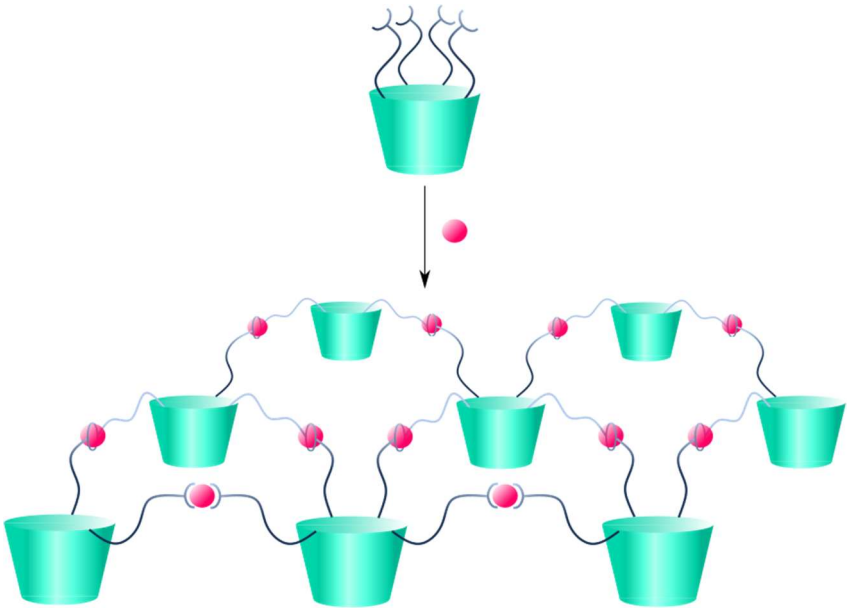


Figure 13. Metal (pink sphere) coordination of an appropriately functionalized cavitand (top) results in a coordination polymer.

2.1. Design of the ligands

Two ligands were synthetized (Figure 14), bearing hydroxyl and pyridine moieties, respectively. Functionalization has been achieved at the upper rim (group R_2 in Figure 9) following a similar strategy. The alkyl substituent on the phosphorous (group R_3 in Figure 9) has been chosen to be ethyl, to avoid steric clash with the upper rim chains. To increase solubility, the lower rim of the cavitands is functionalized with a C_3 alkyl chain (group R_1 in Figure 9).

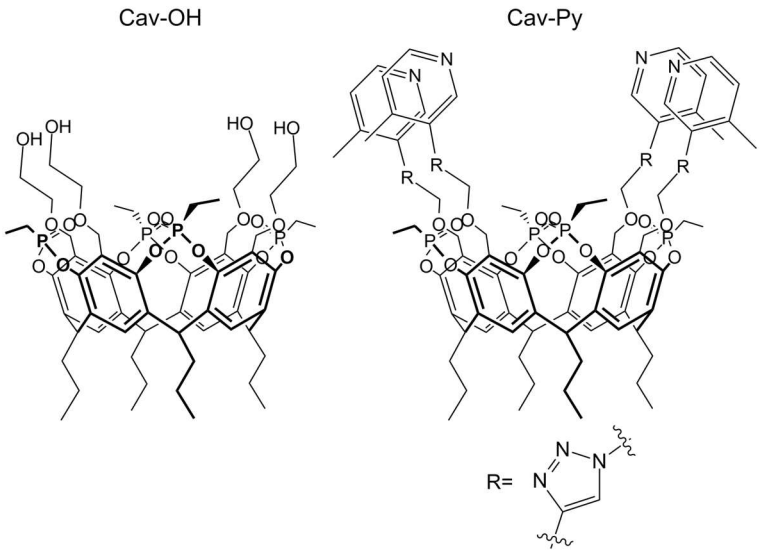


Figure 14. Molecular structure of the two ligands, **Cav-OH** (left) and **Cav-Py** (right).

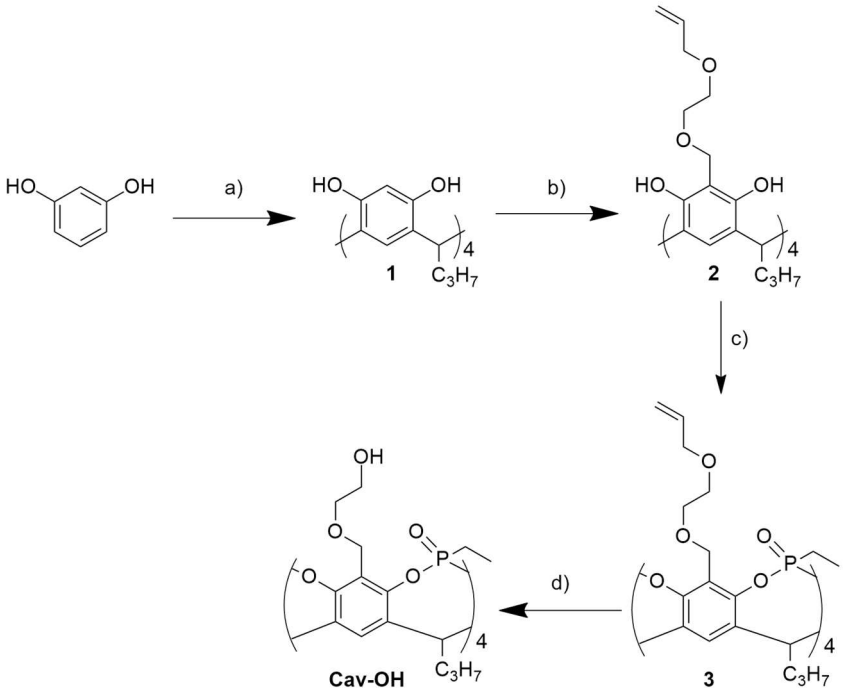
The functionalization of tetra-phosphonate cavitands can be synthetically challenging. One of the main difficulties relies in purification steps. Column chromatography is not always feasible, because tetra-phosphonate cavitands are readily binding Na^+, K^+ and Ca^{2+} ions, even from silica gel. These complexes are formed during elution, overcomplicating purification. For this reason, reactions have been optimized to very high yields or to obtain the product by precipitation or recrystallization.

In the initial phase of this project, we studied the behaviour of our cavitands as ligands to metals. The information obtained will aid in the development of a coordination polymer or a metal organic framework based on tetra-phosphonate cavitand ligands.

3. Upper rim hydroxyl functionalized tetra-phosphonate cavitand (Cav-OH)

3.1. Synthesis

Cav-OH was synthesized from commercially available reagents in 4 steps with an overall yield of 3% (Scheme 1). The synthesis of resorcinarene **1** was achieved by condensation of resorcinol with butyraldehyde. The macrocycle precipitated from the reaction mixture, driving the equilibrium towards the desired product. The functionalization at the upper rim was achieved through a Mannich reaction in the presence of an alcohol.[35] The iminodiacetic acid-catalysed reaction of **1** with 2-allyloxyethanol and formaldehyde afforded resorcinarene **2** with four allyl-protected hydroxyl groups at the apical positions. The yield of this step resulted quite low, only 9%, and the presence of non-reacted resorcinarene **1** was detected by ESI-MS on the crude. Nevertheless, this functionalization strategy was chosen due to the ease of purification. Separation of resorcinarene **2** by column chromatography on silica gel was fairly easy, unlike what is generally observed with resorcinarenes. Indeed, functionalization at the upper rim with long chains increased the molecule hydrophobicity, facilitating purification by chromatography. Allyl-protection of the hydroxyl functions was necessary because of the following bridging step. Phosphonate bridges were obtained through nucleophilic substitution with dichloroethylphosphine, followed by *in-situ* oxidization with hydrogen peroxide. The reaction is stereospecific and only the product **3**, with the four phosphonates pointing inward the cavity, was isolated. Pd-catalysed deprotection, using 1,3-dimethylbarbituric acid as scavenger, gave tetrahydroxyl cavitand **Cav-OH**.



Scheme 1. Synthesis of **Cav-OH**. a) Butyraldehyde, HCl 37%, MeOH, 50°, 7 days, 50%; b) 2-allyloxyethanol, CH_2O aq. 37%, iminodiacetic acid, acetonitrile, 84°C, 18 h, 9%; c) (1) $EtPCl_2$, Py, 80°C, 3 h; (2) H_2O_2 35%, r.t, 1 h, 74% (over two steps); d) $Pd(PPh_3)_4$, 1,3-dimethylbarbituric acid, CH_2Cl_2/MeOH, r.t, 24 h, 90%.

The product was characterized by ^{1}H-NMR, ^{31}P-NMR spectroscopy and ESI-MS spectrometry.

Firstly, we studied **Cav-OH** metal coordination in solution by NMR spectroscopy. After a search in the Cambridge Structural Database, lanthanum (III), barium (II) and caesium (I) were chosen due to their ability of coordinating hydroxyls. Solution and solid state characterizations

showed how the formation of a coordination complex with lanthanum is highly solvent-dependent.

3.2. Lanthanum: solvent-mediated complex formation

Lanthanum complexes of **Cav-OH** were assembled by mixing equimolar amounts of the ligand and $LaCl_3$ in acetonitrile or acetone. Initially, ^{1}H and ^{31}P 1D-NMR were measured. Chemical shift perturbation at the glycol moieties confirmed complex formation. A more accurate analysis of resonances shifts suggested how solvent played a key role in complex formation (Figure 15 and Figure 16).

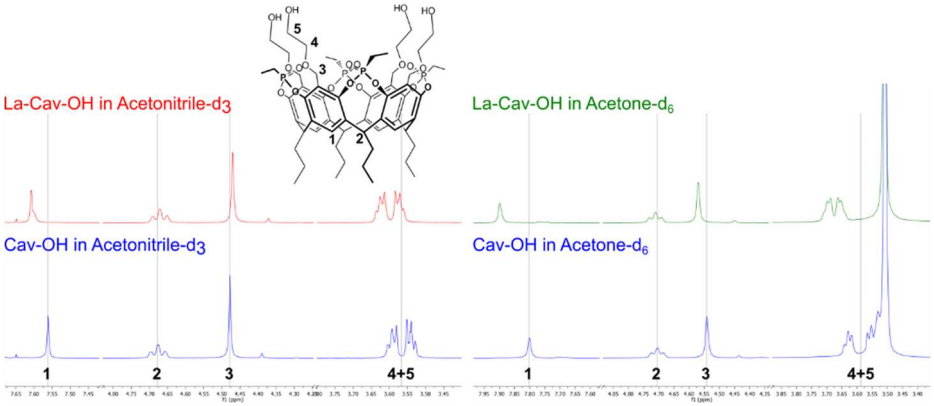


Figure 15. Portions of the ^{1}H-NMR spectra for **Cav-OH** ligand and lanthanum complex in acetonitrile-d_3 (left) and acetone-d_6 (right). The panels highlight signals of protons 1 to 5 as shown in the molecule diagram.

Generally, resonances of protons showed wider shifts in acetone than acetonitrile (Figure 16). The glycol moiety signals (H_4 and H_5) shifted

downfield in both solvents (Figure 15), suggesting lanthanum coordination at this site. Notably, the methylene connecting the aromatic scaffold to the upper rim chain (H_3) shifted upfield in acetonitrile and downfield in acetone (Figure 15). This difference pointed at two different complexes forming in the two solvents. Chemical shift perturbations were generally larger in acetone, suggesting a more stable complex. Most likely, more than one upper-rim hydroxyl is involved in lanthanum coordination. The signals of aromatic protons (H_1) shifted upfield as well, suggesting a change at the cavity site (Figure 16). A possibility would be the complexation of La^{3+} cations by the electron-rich cavity. Nevertheless, this option is unlikely, as corroborated by previous observations from our group.[34] Phosphonate groups are likely to interact with metal cores through second sphere coordination. The signal of the phosphonate groups was shifted in the ^{31}P NMR spectra (Figure 17). This last observation, combined with the wide shift of the ethyl substituent on the phosphonate group (H_6, Figure 16) corroborates involvement of P=O moieties in complex formation.

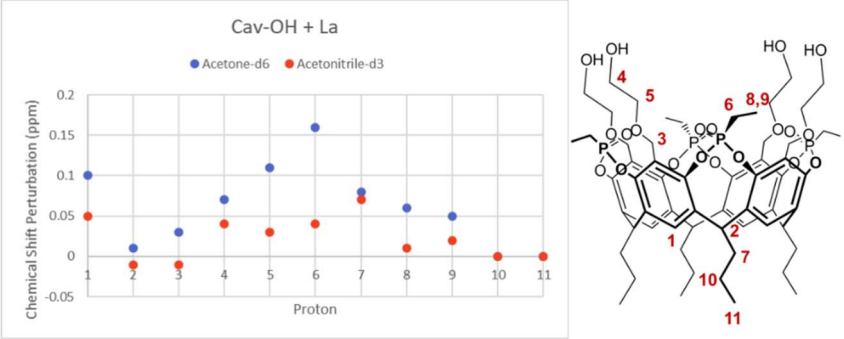


Figure 16. Plot of chemical shift perturbation (y axis) for **Cav-OH** resonances in the presence of lanthanum in acetonitrile (red dots) and acetone (blue dots). Each proton (x axis) is marked with a number as shown in the molecule diagram (right). Shifts are generally wider in

acetone. The trend of chemical shift perturbation is different, with H_2 and H_3 shifting upfield in acetonitrile and downfield in acetone.

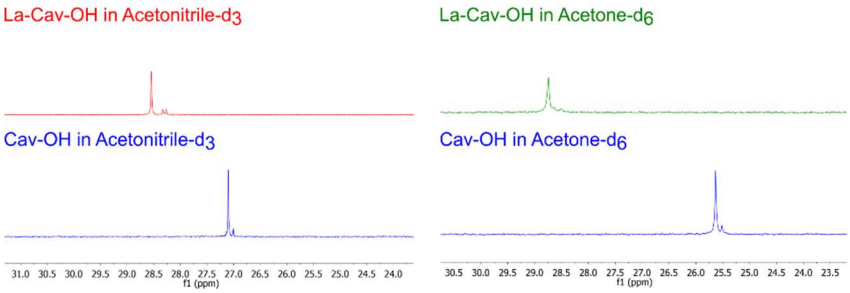


Figure 17. Comparison between the ^{31}P-NMR spectra of **Cav-OH** and the La**-Cav-OH** complex in acetonitrile-d_3 (left) and acetone-d_6 (right). The P=O signal is shifted of ≈ 1.5 ppm in acetonitrile and of ≈ 3 ppm in acetone.

Crystals were obtained by slow evaporation of a solution of the complex in acetonitrile. Single-crystal diffraction analysis revealed the formation of a 1:1 ligand:metal complex of general formula **[(Cav-OH)La(CH_3CN)(H_2O)$_7$]Cl_3 ·3CH_3CN·3H_2O** (Figures 18 and 32). In the complex, the four upper-rim chains of the cavitand are all stretched out in the plane perpendicular to the aromatic scaffold (Figure 18.A), and, surprisingly, only one hydroxyl group is coordinated by lanthanum. The coordination sphere around the metal is completed by one acetonitrile and seven water molecules. The three chloride counterions are located in the lattice, stabilized by a network of H-bonds with the free and coordinated water molecules, and with the OH groups of the hydroxyl chains.

As can be seen in Figure 18.A., the solvent plays a crucial role in the supramolecular structure of the complex, being responsible for the formation of a dimer. Indeed, the acetonitrile molecule coordinated by lanthanum *via* the free electron pair on the CN is at the same time

encapsulated in the cavity of a symmetry-related cavitand through C-H···π interactions [C-H···Cg1: 3.813(3) Å and 156.17(8)°; C-H···Cg2: 3.880(4) Å and 126.99(8)°. Cg1 and Cg2 are the centroids of the aromatic rings C1B-C6B and C1C-C6C, respectively]. Two of the water molecules coordinated by lanthanum form H-bonds directly with the P=O groups of the symmetry-related cavitand [O4W-H7W···O3D^{i}: 2.713(4) Å and 171.0(9)°; O7W-H13W···O3A^{i}: 2.729(5) Å and164.9(8)°. i = 1-x, -y, 1-z]. Hence, this second sphere coordination around the metal appears to facilitate dimerization over formation of a monomeric capsule. These sets of interactions taking place in the solid state match nicely with the chemical shift perturbation highlighted by ^{1}H and ^{31}P-NMR measurements.

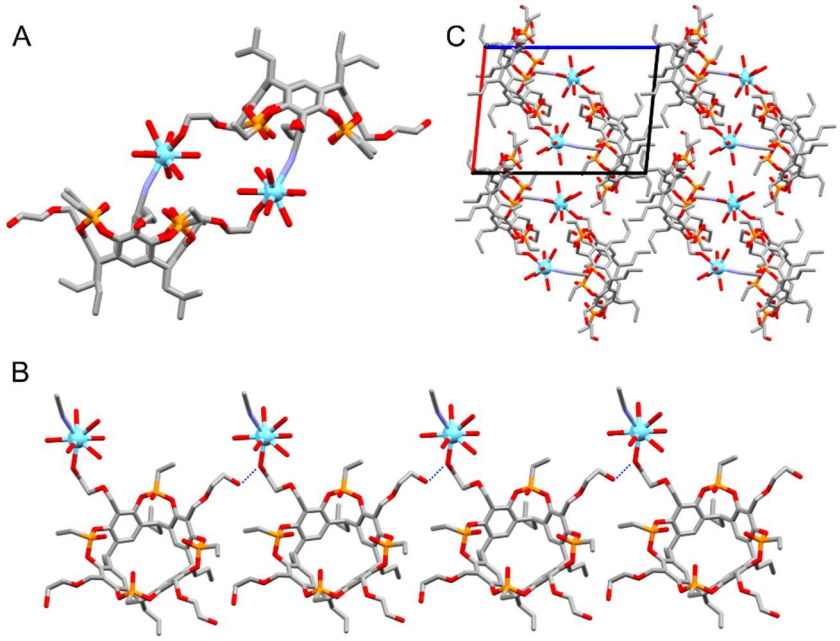


Figure 18. (A) Single-crystal X-ray analysis of La-**Cav-OH** revealed the formation of a supramolecular dimer driven by acetonitrile. (B) In the

lattice, adjacent complexes form supramolecular chains along the *b*-axis direction held together by H-bonds between hydroxyl functionalities (blue dashes). (C) The view of the packing in the *ac* plane highlights layers of ligands and metals. In all the panels, **Cav-OH** is represented as sticks, while La^{3+} as cyan spheres. Solvent molecules, H atoms and Cl^- counter-ions are omitted for clarity.

H-bonding plays also an important role in the 3D assembly. Indeed, H-bonds between coordinated and free hydroxyl arms [O5A-H5OA···O5B^{ii}: 2.646(4) Å and 170.7(9)°. ii = x, y+1, z] yield chains of the complex along the *b*-axis direction (Figure 18.B and 18.C).

Diffusion Ordered Spectroscopy (DOSY) NMR was measured to investigate the complex further. DOSY-NMR is a valid technique for the investigation of supramolecular assemblies in solution.[36] Such technique provides a diffusion coefficient value (D). Assuming that the species is spherical, D relates to the hydrodynamic radius (r) through the Stokes-Einstein equation ($D = k_bT/6\pi\eta r$). The value of r accounts for the whole solvation sphere of the species, namely for coordinated solvent molecules. For this reason, the hydrodynamic radius of the pristine ligand was measured in both solvents and used as a comparison (Table 1). The data was acquired on a 600 JEOL spectrometer and processed with delta software (See Experimental).

In acetonitrile, the diameter of the complex (≈14 Å) was double of the one of the ligand itself (≈7 Å), which is consistent with the formation of a dimeric species. This observation remarkably matches the solid state data, confirming the presence of the dimeric assembly in solution. Being much bulkier than acetonitrile, acetone cannot be encapsulated by the cavity. Hence, the complex in acetone was expected to be a discrete monomeric

species. Comparison between the hydrodynamic radius (Table 1) of the complex and the ligand corroborated this hypothesis. The ligand diameter (≈8 Å) is comparable with the diameter of the complex (≈10 Å). The difference is most likely due to the presence of lanthanum.

Table 1. Diffusion coefficient and hydrodynamic radius for **Cav-OH** and its lanthanum complex in acetonitrile and acetone.

	Diffusion Coefficient	Hydrodynamic Radius r*
Cav-OH in acetonitrile	$8.4\ 10^{-10}\ m^2s^{-1}$	7 Å
La-Cav-OH in acetonitrile	$4.0\ 10^{-10}\ m^2s^{-1}$	14 Å
Cav-OH in acetone	$8.2\ 10^{-10}\ m^2s^{-1}$	8 Å
La-Cav-OH in acetone	$6.1\ 10^{-10}\ m^2s^{-1}$	10 Å
La-Cav-OH in acetonitrile 10% acetone	$3.88\ 10^{-10}\ m^2s^{-1}$	14 Å

*r is calculated based on the Stokes-Einstein equation. See the main text and the experimental section for details.

A small quantity of acetone (10%) in acetonitrile was enough to trigger the assembly interconversion. The D coefficient for a solution of the complex in acetonitrile was measured *via* DOSY-NMR. The experiment was repeated after addition of 10% acetone. After the addition, the value of D decreased toward the value found for the complex in acetone (Table 1). Most likely, acetone disrupts the second sphere coordination by H-bonding to the water molecules coordinated by lanthanum. This mechanism breaks the dimer apart and prompts the assembly interconversion.

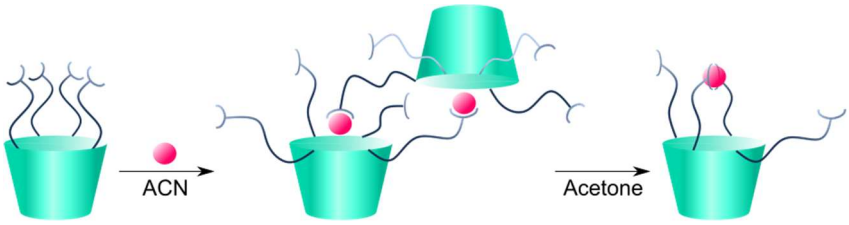


Figure 19. Cartoon of assembly interconversion.

Furthermore, the complex was fully dis-assembled in a protic solvent like methanol. When exposed to methanol, the La**-Cav-OH** crystals dissolved and recrystallized yielding the cavitand alone (Figure 20), with methanol held inside the cavity by H-bonds with the P=O groups and by C-H···π interactions with the aromatic walls. The crystal structure is stabilized by weak H-bonding interactions involving the OH groups of the upper rim chains.

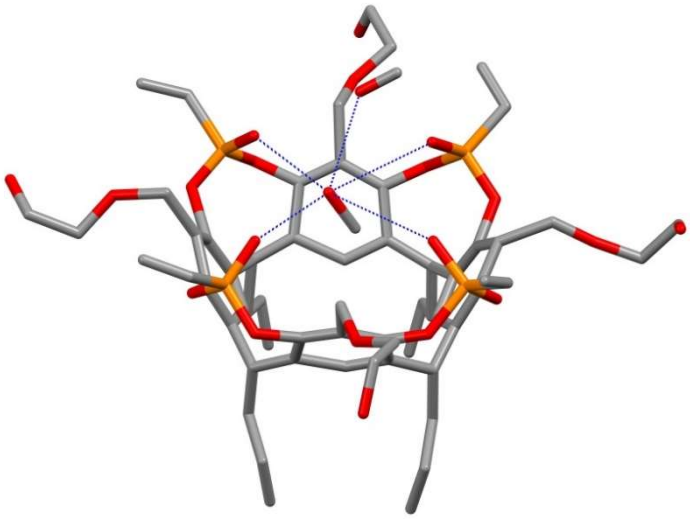


Figure 20. Molecular structure of **Cav-OH** crystallized from methanol. The H-bond pattern is shown as blue dashed lines.

3.3 Barium and caesium: second sphere coordination and encapsulation in the cavity

Solutions of **Cav-OH** and $Cs(NO_3)_2$ or $BaCl_2$ in 1:1 stoichiometry were prepared in deuterated acetonitrile and acetone. Addition of 1% of H_2O was necessary to favour solubility. ^{1}H and ^{31}P spectra were recorded and compared to the one of **Cav-OH**.

The ^{1}H-NMR spectrum in acetonitrile of **Cav-OH** in the presence of barium (Figure 21) did not show dramatic chemical shift perturbations of any of the resonances (Figure 25). Nevertheless, a significant change in the multiplicity was observed for the glycol moiety signals (H_4 and H_5). Furthermore, in the ^{31}P-NMR spectrum (Figure 22), the signal of the phosphonate groups shifted and broadened almost beyond detection. These two observations suggested complex formation and a pivotal role of the P=O groups. The most likely situation is a complex in which barium is coordinated to, at least, one hydroxyl moiety, with water molecules H-bonded to phosphonate groups completing its coordination sphere. Furthermore, barium interacts with at least one P=O group in a fluxional complex, as suggesting from the broadening and shifting of the P signal. This model is consistent with the dimeric capsule reported in the case of non-functionalized tetra-phosphonate cavitand and barium(II).[33]

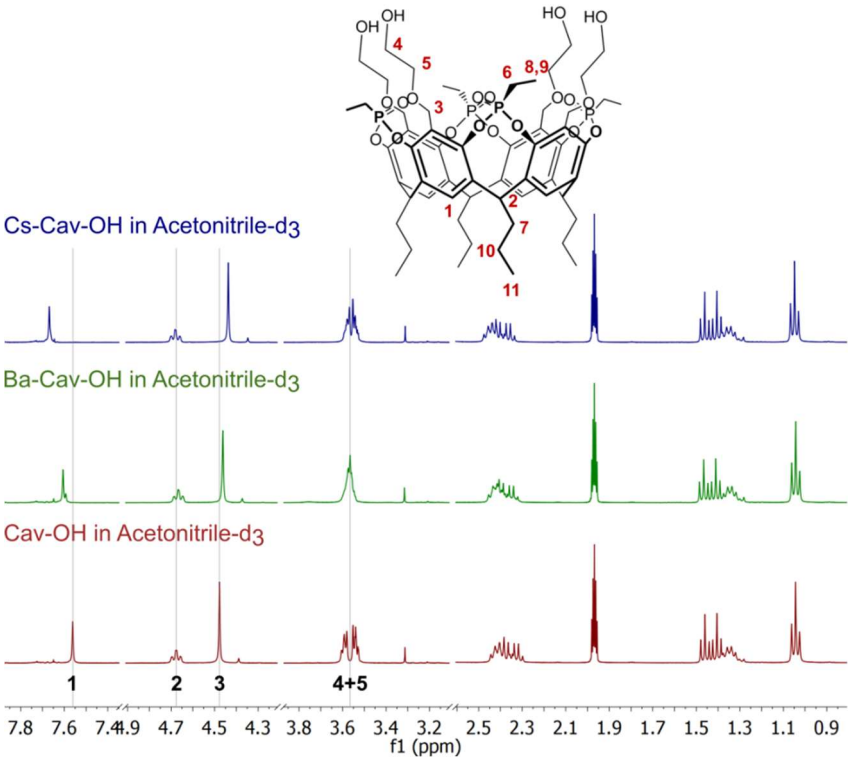


Figure 21. ^{1}H-NMR spectra in acetonitrile of **Cav-OH** alone (red) and in the presence of equimolar amounts of barium(II) (green) and caesium(I) (blue). The signals of protons 1 to 5 are highlighted as shown in the molecule diagram.

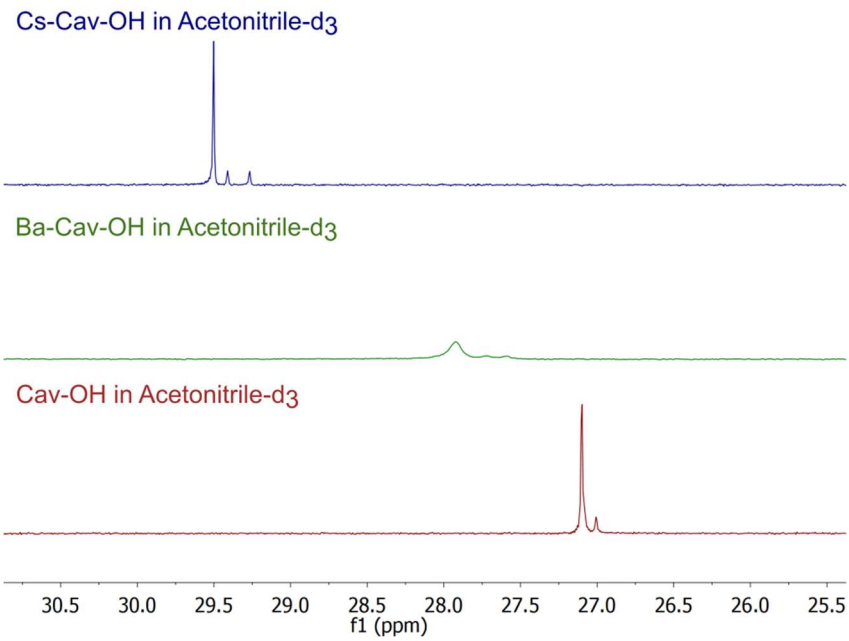


Figure 22. ^{31}P-NMR spectra in acetonitrile of **Cav-OH** pristine (red) and in the presence of equimolar amounts of barium (green) and caesium (blue). The P=O signal is shifted of ≈1 ppm and broadened in the presence of barium and shifted of ≈2.5 ppm in the presence of caesium.

Larger chemical shift perturbations were observed both for ^{1}H and ^{31}P signals in acetone (Figures 23, 24 and 25). Complex formation appears more facile in this solvent. Furthermore, the ethyl substituent on phosphorous showed a large shift in acetone (Figures 23 and 25), probably because the interaction of P=Os with the metal is stronger in this solvent. Moreover, H_1 on the aromatic scaffold showed a ≈0.15 ppm shift in acetone (Figure 23). The complex formed in acetone is similar to the one obtained in acetonitrile, with barium interacting with at least one chain and one P=O. Nevertheless, encapsulation of Ba^{2+} in the cavity is not to be excluded in acetone, due to the large H_1 chemical shift perturbation.

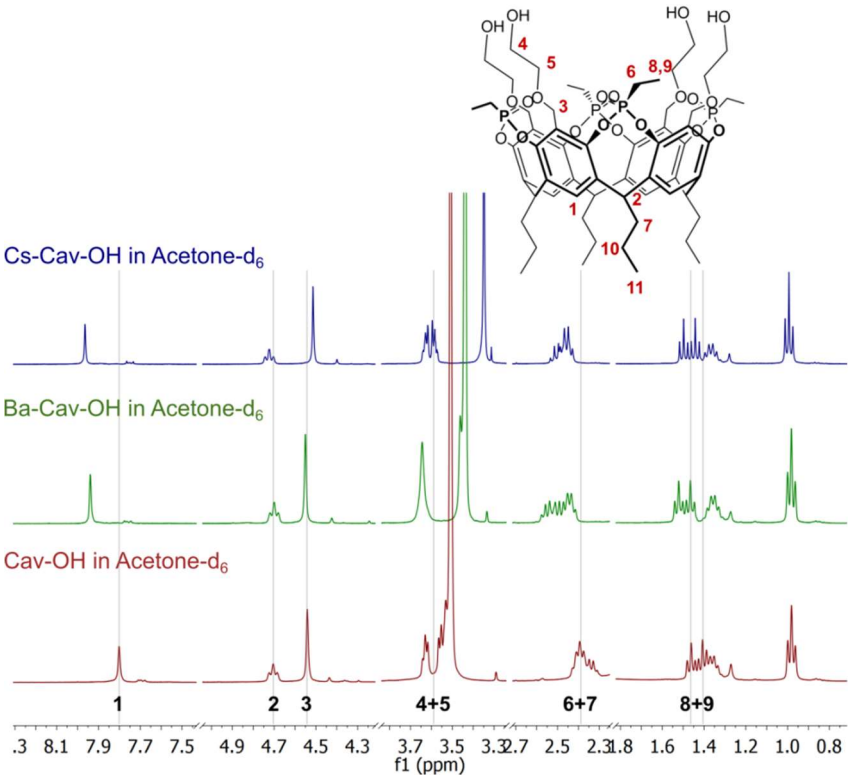


Figure 23. ^{1}H-NMR spectra in acetone of **Cav-OH** pristine (red) and in the presence of equimolar amounts of barium (II) (green) and caesium (I) (blue). The signals of protons 1 to 9 are highlighted as shown in the molecule diagram.

The presence of caesium did not affect the signal of the glycol moiety, neither in acetonitrile nor in acetone, suggesting that no coordination was occurring at this site (Figures 21, 23, 26). On the contrary, the protons close to the aromatic scaffold, namely H_1, H_3, $H_{6\text{-}9}$, showed chemical shift perturbation. Hence, caesium is probably sitting inside the cavity of **Cav-OH** and interacting with P=Os. Large shifts of the phosphorous signals in the

^{31}P-NMR spectra (Figure 22 and 24) seems to corroborate this hypothesis. This situation is consistent with previously reported data.[37]

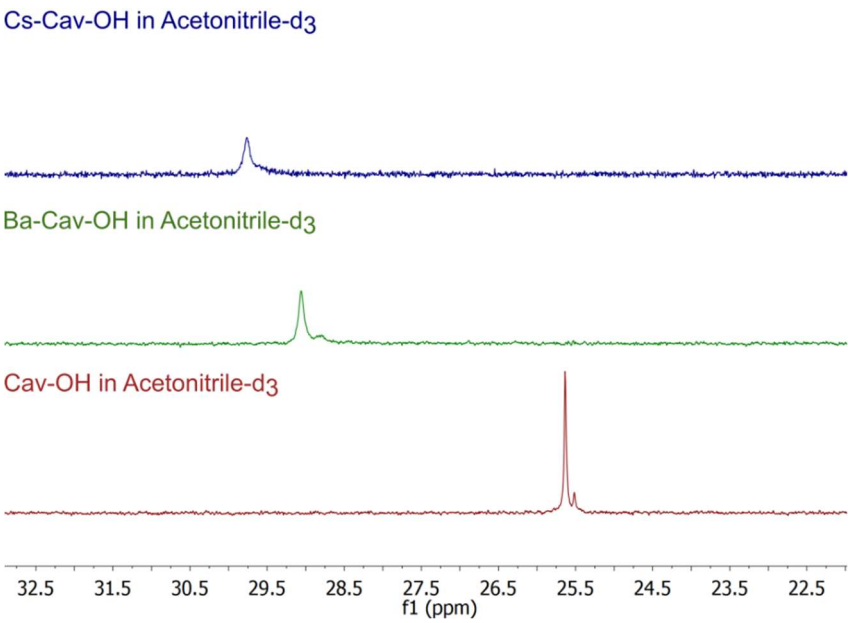


Figure 24. ^{31}P-NMR spectra in acetonitrile of **Cav-OH** (red) and in the presence of equimolar amounts of barium (II) (green) and caesium (I) (blue). P=O signals shifted of ≈3.2 ppm and broadened in the presence of barium and shifted of ≈4 ppm and broadened in the presence of caesium.

As in the case of lanthanum and barium, the analysis of chemical shift perturbation (Figure 26) suggested that the coordination of caesium is much stronger in acetone than acetonitrile. The trend of chemical shift perturbations for barium and caesium complexes (Figures 25, 26) appeared similar in both acetonitrile and acetone.

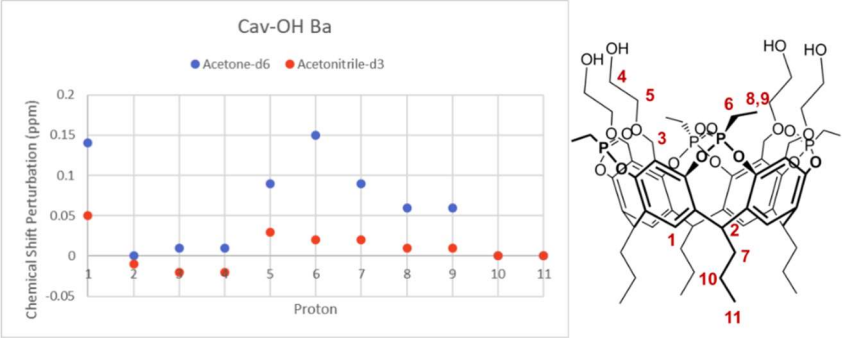


Figure 25. Plot of the chemical shift perturbation (y axis) for **Cav-OH** resonances in the presence of barium in acetonitrile (red dots) and acetone (blue dots). Each proton (x axis) is marked with a number as shown in the molecule diagram (right). See main text for details.

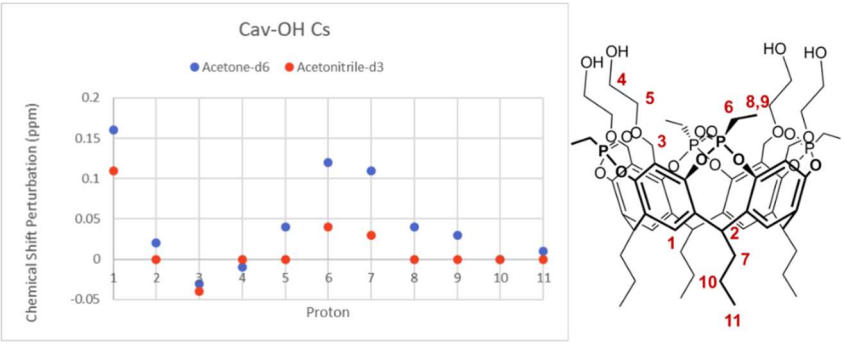
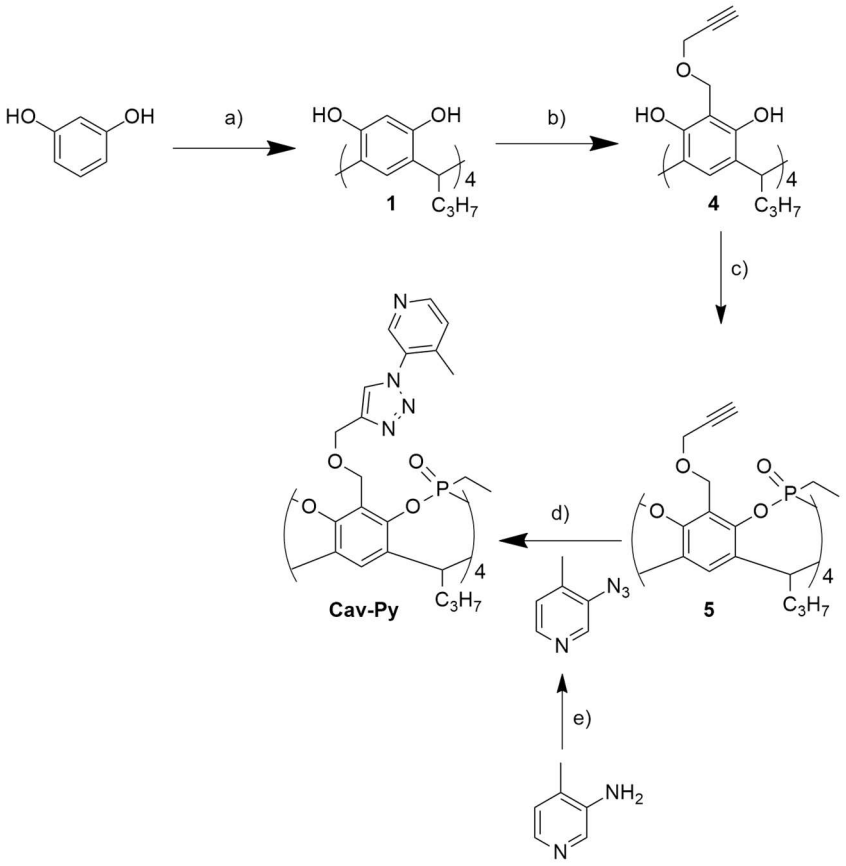


Figure 26. Plot of the chemical shift perturbation (y axis) for **Cav-OH** resonances in the presence of caesium in acetonitrile (red dots) and acetone (blue dots). Each proton (x axis) is marked with a number as shown in the molecule diagram (right). Shifts are generally wider in acetone, with a similar trend in both solvents. See main text for details.

4 Upper rim pyridine functionalized tetra-phosphonate cavitand (Cav-Py)

4.1 Synthesis

Cav-Py was obtained following a strategy (Scheme 2) similar to the one employed for **Cav-OH** (Scheme 1). The synthesis foresees four steps starting from commercially available reagents, with an overall yield of 5%. The starting resorcinarene **1** was obtained as explained in Paragraph 3.1. The Mannich reaction was again employed to functionalize the four apical positions of the aromatic scaffold. The reaction was performed in similar conditions to obtain intermediate **2** (Paragraph 3.1) but in the presence of propargylic alcohol instead of 2-allyloxyethanol. As in the case of resorcinarene **2**, upper rim functionalization enabled easy separation of intermediate **4** from the reaction mixture by column chromatography. Bridging of the phenols of resorcinarene **4** with dichloroethylphosphine and the subsequent *in situ* oxidization with hydrogen peroxide afforded the tetraphosphonate cavitand **5**. Pyridine functionalization was achieved through a click reaction (Huisgen cycloaddition) between cavitand **6** and 3-azido-4-methylpyridine. 3-azido-4-methylpyridine was previously synthesized *via* Sandmeyer reaction from commercially available 3-amino-4-methylpyridine. The click reaction was performed in the presence of a copper catalyst generated *in situ* and afforded **Cav-Py** with a quantitative yield. The product was characterized via ^{1}H and ^{31}P-NMR and ESI-MS.



Scheme 2. Synthesis of **Cav-Py**. a) Butyraldehyde, HCl 37%, MeOH, 50°, 7 days, 50%; b) propargyl alcohol, CH_2O aq. 37%, iminodiacetic acid, acetonitrile, 84°C, 18 h, 14%; c) (1) $EtPCl_2$, Py, 80°C, 3 h; (2) H_2O_2 35%, r.t, 1 h, 77% (over two steps); d) sodium ascorbate, $CuSO_4$, DMF, 80°., 24 h, quant. e) (1) H_2SO_4 98%, 55°, 5'; (2) $NaNO_2$ aq., 0°, 20'; (3) urea, 0°, 20'; (4) NaN_3 aq., rt, 4h.

Pyridines are excellent ligands to a variety of metals. In this study, we mainly focused on platinum (II) and zinc (II), as these two metals are ideal to pursue our initial objective (Paragraph 2, Figures 12 and 13). Platinum has a coordination number of 4 in the large majority of its complexes, adopting

a square planar geometry. In addition, platinum complexes are often stable in a variety of conditions. For these reasons, platinum is an ideal candidate to form molecular capsules, as its coordination is quite predictable. On the other hand, zinc coordination number varies from 4 to 6 and complexes often have variable geometries. These characteristics, combined with chemical stability of the resulting complex, make zinc ideal for the preparation of coordination polymers or metal organic frameworks.

4.2 Platinum: capsules versus polymer formation

We employed three different platinum complexes as precursors to generate Pt-**Cav-Py** complexes: *trans*-$Pt(ACN)_2Cl_2$, *cis*-$Pt(ACN)_2Cl_2$ and $Pt(ACN)_4(CF_3SO_3)_2$. In all of these complexes, acetonitrile is a labile ligand and can be easily substituted by pyridines. To perform this reaction, we mixed equimolar amounts of **Cav-Py** ligand and platinum precursors in chloroform at room temperature. ^1H NMR was measured to monitor the formation of a complex.

The spectra of **Cav-Py** in the presence of *trans*-$Pt(ACN)_2Cl_2$ (Figure 27) showed a pronounced chemical shift perturbation in the aromatic region.

Pyridine signals were shifted up to ≈1 ppm, indicating platinum coordination. Furthermore, the signals were much sharper. The platinum complex spectrum showed more signals in the aromatic region than **Cav-Py** (Figures 27 and 28). COSY and TOCSY NMR spectroscopy confirmed the presence of two homologues sets of signals (Figure 28). One set of signals had lower intensities than the other, consistent with metal coordination.

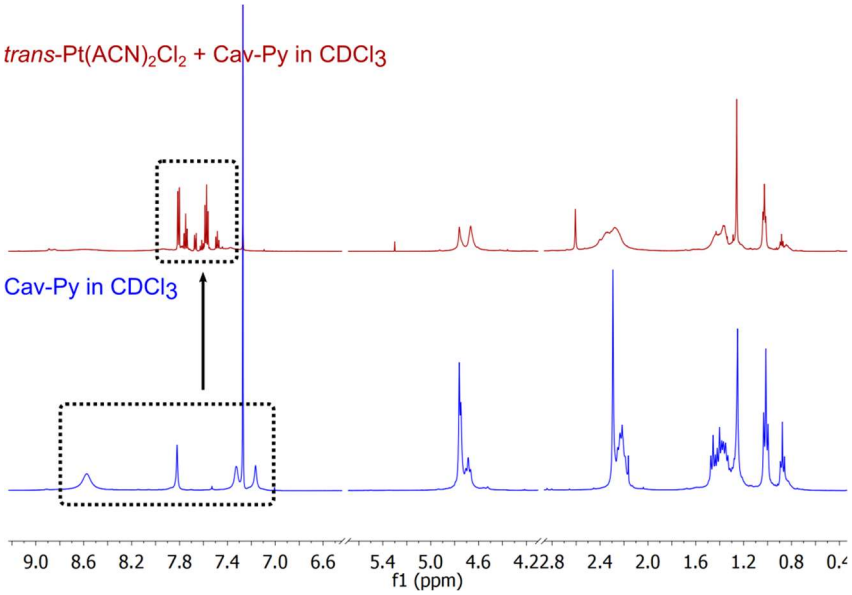


Figure 27. ^{1}H spectra of **Cav-Py** pristine (blue) and in the presence of *trans*-$Pt(ACN)_2Cl_2$ (red). The aromatic region is indicated with a rectangular black box.

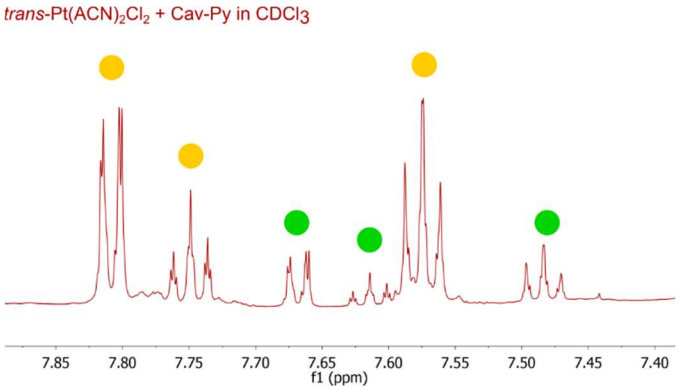
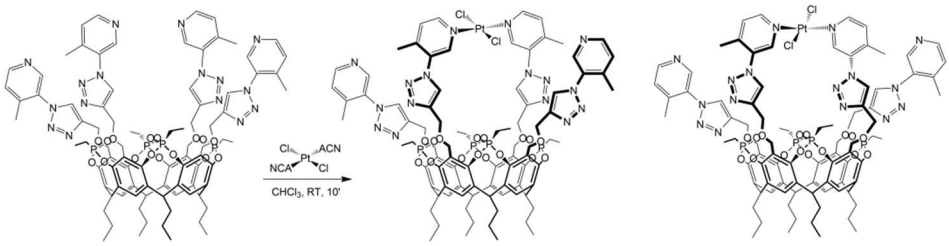


Figure 28. ^{1}H-NMR spectrum of the *trans*-Pt-**Cav-Py** complex, zoom on pyridine spectral region. Two sets of signals (highlighted in yellow and green) were identified by 2D NMR.

The NMR data suggested the formation of a 1:1 complex, in which two pyridines substitute the two acetonitrile ligands on the platinum precursor (Scheme 3). Chlorines are retained on platinum and the overall complex charge is neutral (Scheme 3). Two isomers (pseudo-cis and pseudo-trans) are formed (Scheme 3) involving two adjacent and two opposite pyridines.



Scheme 3. Formation of a *trans*-Pt-**Cav-Py** complex.

The same experiment was repeated for *cis*-$Pt(ACN)_2Cl_2$ and $Pt(ACN)_4(CF_3SO_3)_2$. However, the ^{1}H-NMR spectra (Figure 29) appeared broadened, especially in the aromatic region. Such spectra suggest the formation of coordination polymers.

This hypothesis is supported by the fact that the formation of a polymer over a capsule strongly depends on the coordination geometry. Most likely, only *trans*-$Pt(ACN)_2Cl_2$ has the correct geometry to yield a 1:1 capsule, while *cis*-$Pt(ACN)_2Cl_2$ and $Pt(ACN)_4(CF_3SO_3)_2$ favour polymerization.

Various techniques have been used to prompt crystallization. Nevertheless, the small dimensions of the crystals obtained has for the time being hampered further structural characterization of the complexes.

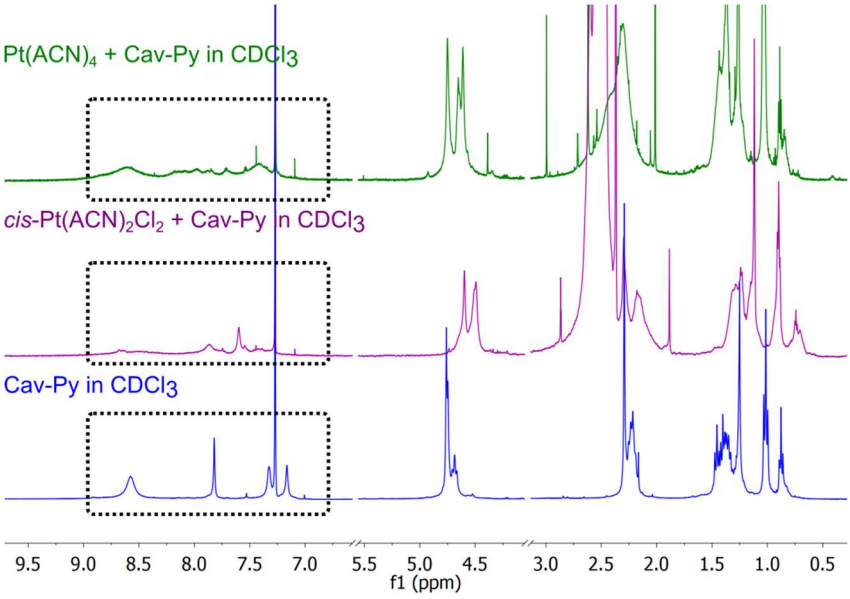


Figure 29. ^{1}H spectra of **Cav-Py** pristine (blue) and in the presence of *cis*-$Pt(ACN)_2Cl_2$ (purple) and $Pt(ACN)_4(CF_3SO_3)_2$ (green). The aromatic region is indicated with a rectangular black box.

4.3 Zinc: toward the formation of coordination polymers via solvothermal reaction and mechanochemistry

In the first experiment, DMSO or chloroform solutions of **Cav-Py** and $Zn(NO_3)_2$ were mixed at room temperature. This approach did not produce any chemical shift perturbation in the ^{1}H-NMR spectra, indicating that coordination had not occurred. Therefore, we decided to proceed with two harsher synthetic methodologies, that is solvothermal reaction and liquid assisted grinding.

Since one of the goals of the project was to obtain a coordination polymer or a MOF, the solvothermal synthesis seemed the most obvious technique to employ. Due to the small quantities of **Cav-Py** available, optimization of the reaction resulted quite difficult. However, after screening different conditions, we obtained a crystalline material. The solvothermal synthesis was performed in DMF, in the presence of four equivalents of $Zn(NO_3)_2$ and two equivalents of terephthalic acid as an ancillary ligand. Orange crystals were isolated by centrifugation and dried in vacuum. The crystallinity of the material was proved by PXRD (Figure 30), but the crystals were too small for single-crystal X-ray diffraction analysis. Further characterization of the product is, at the time being, on-going.

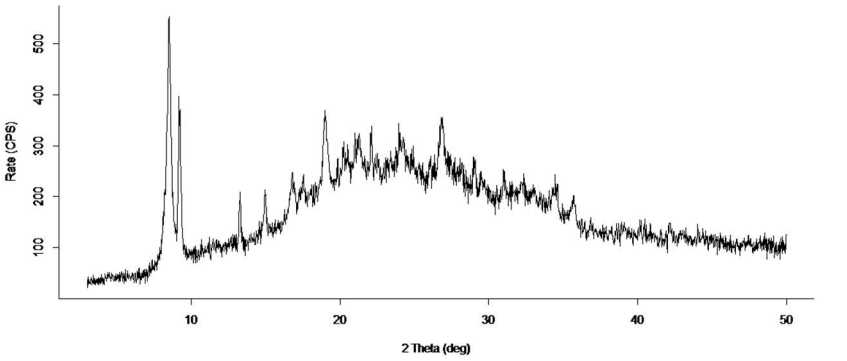


Figure 30. PXRD of the product from solvothermal reaction between **Cav-Py**, zinc nitrate and terephthalic acid. The baseline is noisy, due to the small quantities of product available.

Liquid Assisted Grinding (LAG) is a mechanochemical technique for the synthesis of coordination polymers and MOFs (see chapter 1, paragraph

4).[38,39] The reaction is performed by grinding reagents together and exploits the energy obtained by friction.

Manual LAG of **Cav-Py** with 2 equivalents of zinc nitrate was carried out in the presence of methanol. The reagents were ground together until complete evaporation of the solvent. The reaction crude was partly soluble in methanol. Hence, two fractions were isolated based on methanol solubility: soluble and insoluble. It was not possible to dissolve the methanol-insoluble fraction in other common organic solvents nor in water. However, colourless needle crystals suitable for XRD of the soluble fraction were obtained by slow evaporation from methanol over two days.

Single-crystal diffraction analysis revealed the formation of a centrosymmetric, 2:2 ligand:metal dimeric complex of general formula **$[(Cav\text{-}Py)Zn(H_2O)_4]_2(NO_3)_6$** (Figure 31.A). In the complex, zinc shows an octahedral O4N2 geometry, being surrounded by four water molecules, (two of which occupy the axial positions) and by two pyridine ligands in *cis*. These belong to two cavitands related by a centre of symmetry, so that their cavities point in opposite directions. The overall result is that each cavitand employs only two of the four pyridine arms to bind two distinct metal ions, while the other two remain non-coordinated. As observed for **Cav-OH**, the four upper rim chains stretch out in the plane perpendicular to the aromatic scaffold, making the whole dimer roughly planar.

Second sphere coordination is observed also in this case (Figure 31.B). Indeed, three of the four water molecules coordinated by zinc are simultaneously H-bonded to the four phosphonate groups of a symmetry-related **Cav-Py**. These interactions are pivotal in determining the overall 3D assembly. As can be seen from Figure 32, each complex acts as H-bond

donor (with the aquo cation) and acceptor (with the phosphonate cavitand) towards two adjacent dimers, located above and below its mean plane.

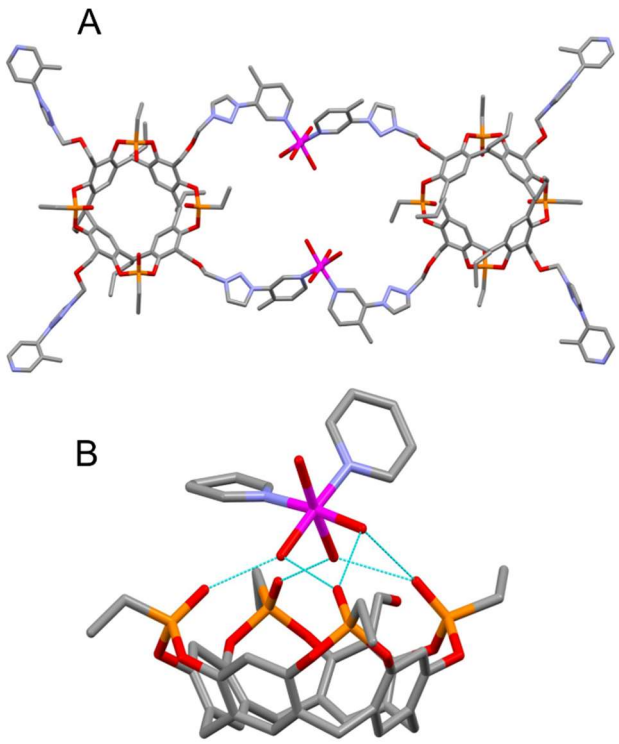


Figure 31. (A) The molecular structure of the cationic part of Zn-**Cav-Py**; it consists of a 2:2 ligand:metal dimeric complex. Colour code: Zinc, pink; phosphorus, orange: nitrogen, light blue; oxygen, red; carbon, grey. (B) Zoom on the coordination sphere of the Zn(II) cation. The water molecules coordinated by the metal are H-bonded to the phosphonate groups of a symmetry-related cavitand.

The crystal structure of the Zn-**Cav-Py** dimer raised questions regarding the reaction path using LAG. Even if two equivalents of $Zn(NO_3)_2$

were employed for one equivalent of cavitand, the resulting complex shows a 1:1 stoichiometry. Nevertheless, there is no apparent reason why the two free pyridines should not be coordinated. A hypothetical explanation is that the 2:2 complex is an intermediate in the LAG reaction which should, possibly, lead to a coordination polymer, namely the insoluble fraction. Unfortunately, the small amount obtained for this fraction thwarted the possibility of studying it through PXRD analysis.

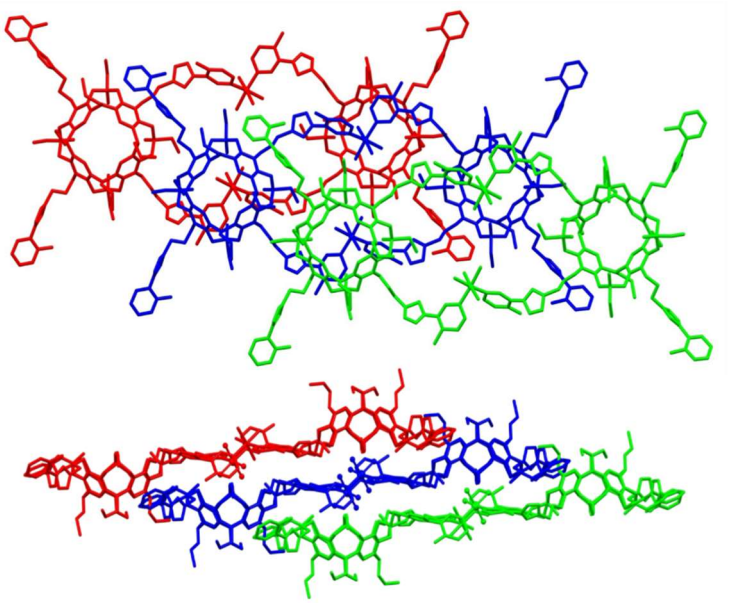


Figure 32. Crystal structure of Zn-**Cav-Py**. The 2:2 dimeric complexes (shown in red, blue and green) are held together by H-bonding involving the aquo cation and the P=O groups of the cavitands. The two view are rotated by 90°.

To test this hypothesis, the reaction was performed with an excess of $Zn(NO_3)_2$, with longer reaction times. At the time being, characterization of the product is ongoing.

5 Conclusions and future perspective

Macrocyclic receptors are a promising class of ligands for the development of coordination polymers and MOFs.[19–21,23–26,40] To our knowledge, tetra-phosphonate cavitands have not yet been used for such application. Here we report preliminary studies of *ad-hoc* functionalized tetra-phosphonate cavitands as ligand for metals. Functionalization was achieved at the apical position with hydroxyl or pyridine moieties.

Ba^{2+} and Cs^{+} coordinate **Cav-OH**, bearing four hydroxyl groups at the upper rim, in solution, as indicated by NMR characterization; in the case of La^{3+}, the coordination was also proved in the solid state. The assembly of the lanthanum complex is strongly solvent-dependent and can be interconverted. **Cav-OH** complexes appear to be discrete entities, with a low tendency of forming coordination polymers.

Pt^{2+} coordination to **Cav-Py** was also studied in solution using NMR spectroscopy, and appears to be strongly influenced by the geometry of the platinum complex used as a precursor: *trans*-$Pt(ACN)_2Cl_2$ seems to facilitate capsule formation, while *cis*-$Pt(ACN)_2Cl_2$ and $Pt(ACN)_4(CF_3SO_3)_2$ led to less resolved and broader spectra, possibly due to polymer formation. The small dimensions of the crystals obtained has for the moment hampered further structural characterization.

Zn^{2+}-**Cav-Py** systems have led to the most promising results from the perspective of the preparation of coordination polymers. Crystalline insoluble materials were obtained from both solvothermal reaction and liquid assisted grinding. Structural characterization is ongoing.

6 Acknowledgments

We would like to acknowledge Dr. Alessandro Pedrini (University of Milano Bicocca) for advice and useful criticism. Dr. Domenico Acquotti (CIM, University of Parma) and Dr. Stefano Ghelli (SpinLab) are thanked for assistance with NMR spectroscopy. Dr. Stefano Canossa is thanked for synchrotron data collection. Letizia Rizzo is thanked for assistance with **Cav-Py** synthesis.

7 Experimental

Synthesis

Resorcinarene [C_3H_7, H] (1): Resorcinol (15.0 g, 0.14 mol) was dissolved in 50 mL of MeOH and a 37% solution of HCl (20 mL, 0.24 mol) was slowly added at 0 °C. At the same temperature a solution of butyraldehyde (12.3 mL, 0.14 mol) in 30 mL of MeOH was added dropwise. The mixture was allowed to warm to room temperature over 30 min and heated at 50 °C for 7 days. After cooling, water (500 mL) was added and the precipitate was filtered, washed with water and dried. The crude was recrystallized twice from MeOH affording resorcinarene **1** as a white solid (y: 50%).

^{1}H NMR (acetone-d_6, 400 MHz): δ (ppm) = 8.45 (s, 8H, O**H**), 7.60 (s, 4H, Ar$\mathbf{H_{up}}$), 6.25 (s, 4H, Ar$\mathbf{H_{down}}$), 4.35 (t, 4H, J=8 Hz, C**H**CH_2), 2.30 (q, 8H, J=8 Hz, CHC$\mathbf{H_2}CH_2$), 1.33 (sext, 8H, J=8 Hz, $CH_2C\mathbf{H_2}CH_3$), 0.90 (t, 12H, J=8 Hz, $CH_2C\mathbf{H_3}$).

ESI-MS: m/z = 655.7 $[M-H]^-$.

Resorcinarene [C_3H_7, $CH_2OCH_2CH_2OAllyl$] (2): To a suspension of resorcinarene **1** (0.5 g, 0.76 mmol) in 8 mL of acetonitrile, 2-allyloxyethanol (3.26 mL, 30.45 mmol) was added, followed by a 37% aqueous solution of formaldehyde (0.286 mL, 3.81 mmol). Iminodiacetic acid (0.051g, 0.38 mmol) was added and the mixture was refluxed for 18 h. After cooling, chloroform (100 mL) was added and the organics were separated and washed with water (3 × 100 mL). The solvent was removed under reduced pressure and flash column chromatography (gradient from Hex/EtOAc 7:3 to Hex/EtOAc 1:1) afforded pure resorcinarene **2** as a white solid (y: 9%).

^{1}H NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.60 (s, 8H, O**H**); 7.19 (s, 4H, Ar$\mathbf{H_{down}}$); 5.99-5.91 (m, 4H, $OCH_2C\mathbf{H}=CH_2$); 5.32-5.24 (m, 8H, $CH_2CH=C\mathbf{H_2}$); 4.84 (s, 8H,

ArC**H$_2$**O); 4.35 (t, 4H, ArC**H**CH$_2$, J=8 Hz); 4.05 (dt, 8H, OC**H$_2$**CH=CH$_2$, J=6 Hz, J=1 Hz); 3.72-3.70 (m, 8H, ArCH$_2$OC**H$_2$**CH$_2$O); 3.63-3.61 (m, 8H, ArCH$_2$OCH$_2$C**H$_2$**O); 2.30 (q, 8H, C**H$_2$**CH$_2$CH$_3$, J=7 Hz); 1.32 (sext, 8H, CH$_2$C**H$_2$**CH$_3$, J=7 Hz); 0.99 (t, 12H, CH$_2$CH$_2$C**H$_3$**, J=7 Hz).

ESI-MS: m/z = 1130.6 $[M+NH_4]^+$; 1135.6 $[M+Na]^+$.

Tetra-phosphonate cavitand [C$_3$H$_7$, CH$_2$OCH$_2$CH$_2$OAllyl, Et] (3): To a solution of resorcinarene 2 (0.170 g, 0.15 mmol) in 8 mL of pyridine, dichloroethylphosphine (0.07 mL, 0.67 mmol) was added. The mixture was heated at 80 °C for 3 h. After cooling, 2 mL of aqueous 35% H_2O_2 were added at 0 °C and the mixture was stirred for 1 h. The reaction was quenched with water (100 mL) and the precipitate was filtered, washed with water and dried. Cavitand **3** was obtained as a white solid (y: 74%).

^{1}H NMR ($CDCl_3$, 400 MHz): δ = (ppm) 7.13 (s, 4H, Ar**H$_{down}$**); 5.99-5.91 (m, 4H, OCH$_2$C**H**=CH$_2$); 5.32-5.24 (m, 8H, CH$_2$CH=C**H$_2$**); 4.73-4.62 (m, 12H, ArC**H**CH$_2$+ArC**H$_2$**O); 4.02 (d, 8H, OC**H$_2$**CH=CH$_2$, J=6 Hz); 3.69-3.67 (m, 8H, ArCH$_2$OC**H$_2$**CH$_2$O); 3.63-3.61 (m, 8H, ArCH$_2$OCH$_2$C**H$_2$**O); 2.26-2.20 (m, 16H, P(O)C**H$_2$**CH$_3$+C**H$_2$**CH$_2$CH$_3$, J=7 Hz); 1.53-1.32 (m, 20H, P(O)CH$_2$C**H$_3$**+CH$_2$C**H$_2$**CH$_3$); 1.03 (t, 12H, CH$_2$CH$_2$C**H$_3$**, J=7 Hz).

^{31}P NMR ($CDCl_3$, 162 MHz): δ (ppm) = 22.3 (s, P=O).

ESI-MS m/z = 1426.6 $[M+NH_4]^+$; 1431.6 $[M+Na]^+$.

Tetra-phosphonate cavitand [C$_3$H$_7$, CH$_2$OCH$_2$CH$_2$OH, Et] (Cav-OH): Cavitand **3** (0.378 g, 0.27 mmol) was dissolved in 20 mL of a degassed mixture of CH_2Cl_2/MeOH (1:3), followed by the addition of $Pd(PPh_3)_4$ (0.031 g, 0.027 mmol). 1,3-dimethylbarbituric acid (0.335 g, 2.15 mmol) was added and the reaction mixture was stirred at room temperature for 24 h. The reaction was monitored by ^{31}P-NMR. The solvent was removed under reduced

pressure and the crude was recrystallized from toluene twice affording pure **Cav-OH** as yellowish solid (y: 90%).

^{1}H NMR (CD_3OD, 400 MHz): δ (ppm) = 7.55 (s, 4H, Ar**H_{down}**); 4.74 (t, 4H, ArC**H**CH_2, J=8 Hz); 4.59 (s, 8H, ArC**H_2**O); 3.67-3.65 (m, 8H, $ArCH_2OC$**H_2**CH_2OH); 3.56-3.54 (m, 8H, $ArCH_2OCH_2C$**H_2**OH); 2.47-2.29 (m, 16H, P(O)C**H_2**CH_3+C**H_2**CH_2CH_3, J=7 Hz); 1.55-1.34 (m, 20H, $P(O)CH_2C$**H_3**+CH_2C**H_2**CH_3); 1.08 (t, 12H, CH_2CH_2C**H_3**, J=7 Hz).

^{31}P NMR (CD_3OD, 162 MHz): δ (ppm) = 25.8 (s, P=O).

^{13}C NMR (CD_3OD, 600 MHz): δ (ppm) = 145; 145; 135; 124; 123; 72; 63; 61; 36; 33; 29; 21; 18; 17; 13; 5.

ESI-MS m/z = 1266.6 $[M+NH_4]^+$.

Resorcinarene [C_3H_7, CH_2OCH_2CCH] (4): To a solution of resorcinarene **1** (300 mg, 0.46 mmol) propargylyc alcohol (1 mL, 18.28 mmol) was added, followed by a 37% aqueous solution of formaldehyde (0.170 mL, 2.29 mmol). Iminodiacetic acid (30.41 mg, 0.23 mmol) was added and the mixture was refluxed for 18 h. After cooling, chloroform (100 mL) was added and the organics were separated and washed with water (3 × 100 mL). The solvent was removed under reduced pressure and flash column chromatography (gradient from Hex/EtOAc 8:2 to Hex/EtOAc 7:3) afforded pure resorcinarene **4** as a white solid (y: 14%).

^{1}H NMR ($CDCl_3$, 300 MHz): δ (ppm) = 8.33 (s, 8H, O**H**); 7.21 (s, 4H, Ar**H_{down}**); 4.93 (s, 8H, ArC**H_2**OCH_2CCH); 4.35 (t, 4H, ArC**H**CH_2, J=6 Hz); 4.25 (s, 8H, $ArCH_2OC$**H_2**CCH); 2.53 (s, 4H, $ArCH_2OCH_2CC$**H**); 2.23-2.16 (m, 8H, C**H_2**CH_2CH_3); 1.40-1.27 (m, 8H, CH_2C**H_2**CH_3); 1.03-0.98 (m, 12H, CH_2CH_2C**H_3**).

^{13}C NMR ($CDCl_3$, 300 MHz): δ (ppm) = 150; 124; 123; 109; 77; 76; 67; 58; 36; 33; 21; 14.

ESI-MS m/z = 927.6 $[M+H]^+$; 946.6 $[M+NH_4]^+$; 951.54 $[M+Na]^+$.

Tetra-phosphonate cavitand [C_3H_7, CH_2OCH_2CCH, Et] (5): To a solution of resorcinarene **4** (80 mg, 0.086 mmol) in 5 mL of pyridine, dichloroethylphosphine (0.04 mL, 0.38 mmol) was added. The mixture was heated at 80 °C for 3 h. After cooling, 1 mL of aqueous 35% H_2O_2 was added at 0 °C and the mixture was stirred for 1 h. The reaction was quenched with water (100 mL) and the precipitate was filtered, washed with water and dried. Cavitand **5** was obtained as a white solid (y: 77%).

^{1}H NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.69 (s, 4H, Ar**H_{down}**); 4.65-4.62 (m, 12H, ArC**H_2**OCH_2CCH+ ArC**H**CH_2); 4.16 (s, 8H, $ArCH_2OC$**H_2**CCH); 2.46 (s, 4H, $ArCH_2OCH_2CC$**H**); 2.35-2.28 (m, 8H, P(O)C**H_2**CH_3); 1.55-1.45 (m, 12H, P(O)CH_2C**H_3**); 1.38-1.36 (m, 8H, CH_2C**H_2**CH_3); 1.05-1.01 (m, 12H, CH_2CH_2C**H_3**).

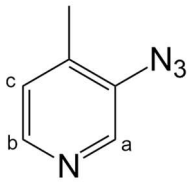
^{31}P NMR ($CDCl_3$, 162 MHz): δ (ppm) = 24.4 (s, P=O).

MALDI-TOF m/z = $C_{64}H_{76}NaO_{16}P_4$ $[M+Na]^+$ predicted: 1247.4, found: 1247.4; $C_{64}H_{76}KO_{16}P_4$ $[M+K]^+$ predicted: 1263.4, found: 1263.3.

3-Azido-4-methylpyridine: A mixture of 0.49 mL of H_2SO_4 98% and 2.80 mL of distilled water was heated to 55°. 3-Amino-4-methylpyridine (300 mg, 2.77 mmol) was added and the mixture was stirred at 55° for 5 minutes. The solution was firstly cooled to room temperature and then to 0°. A solution of $NaNO_2$ (230 mg, 3.33 mmol) in 3 mL of water was added dropwise and the mixture was reacted for another 20 minutes. Urea (33 mg, 0.56 mmol) was added and the mixture was stirred at 0° for other 20 minutes. A solution of NaN_3 (216 mg, 3.33 mmol) in 2 mL of water was added dropwise at 0°. The reaction mixture was allowed to warm to room temperature and further

reacted for 4 h. The reaction was quenched with a saturated solution of $NaHCO_3$ and pH of the mixture was confirmed to be basic. The organic phase was extracted with chloroform (3 x 100 mL), the organic layers were combined and dried over Na_2SO_4. The solvent was removed under reduced pressure and column chromatography (gradient from DCM to DCM/EtOAc 7:3) afforded the product as a yellow oil (y: 72%).

^{1}H NMR ($CDCl_3$, 300 MHz): δ (ppm) = 8.34 (s, 1H, **H_a**); 8.19 (d, 1H, **H_b**, J=4 Hz); 7.03 (d, 1H, **H_c**, J=4 Hz); 2.16 (s, 3H, C**H_3**).



^{13}C NMR ($CDCl_3$, 300 MHz): δ (ppm) = 145; 140; 125; 16.

ESI-MS m/z = 135.2 $[M+H]^+$.

IR: 2106.6 cm^{-1} ($-N_3$).

Tetra-phosphonate cavitand [C_3H_7, CH_2OCH_2TriazolePy, Et] (Cav-Py): To a solution of cavitand **5** (68 mg, 0.06 mmol) in 10 mL of DMF $CuSO_4$ (18 mg, 0.06 mmol) was added, followed by sodium ascorbate (18 mg, 0.09 mmol) and 3-azido-4-methylpyridine (80 mg, 0.60 mmol). The reaction was refluxed for 24 h. The volatiles were removed under reduced pressure. The crude was dissolved in DCM (100 mL) and the organic phase was washed with water (3 x 100 mL) and with a saturated solution of EDTA (5 x 100 mL). The organic layer was collected, dried over Na_2SO_4 and the solvent was evaporated under reduced pressure affording **Cav-Py** (quantitative yield).

^{1}H NMR ($CDCl_3$, 300 MHz): δ (ppm) = 8.58 (s, 4H, **H_a**); 7.83 (s, 4H, **$H_{triazole}$**); 7.34 (s, 4H, **H_b**); 7.18 (s, 4H, **H_c**); 4.73-4.71 (m, 20H, ArC**H_2**O + OC**H_2**Triazole + ArC**H**CH_2); 2.30 (s, 12 H, PyC**H_3**); 2.26-2.18 (m, 16H, P(O)C**H_2**CH_3 + CHC**H_2**CH_2); 1.47-1.36 (m, 20H, P(O)CH_2C**H_3** + CH_2C**H_2**CH_3); 1.01 (t, 12H, CH_2CH_2C**H_3**).

^{31}P NMR ($CDCl_3$, 162 MHz): δ (ppm) = 22.9 (s, P=O).

ESI-MS m/z = 893.2 $[M+H+Na]^{2+}$; 1763.3 $[M+H]^{+}$; 1785.6 $[M+Na]^{+}$.

Cav-OH complexes characterization in solution

Materials: $LaCl_3{\cdot}7H_2O$, $BaCl_2$ and $CsNO_3$ were purchased from Sigma.

NMR samples: NMR samples were prepared by dissolving **Cav-OH** and an equimolar amount of $LaCl_3$ or $BaCl_2$ or $CsNO_3$ in acetonitrile-d_3 or acetone-d_6. 1% of water was added to favour dissolution. 5% of acetone was also added in the case of the assembly interconversion experiment.

DOSY-NMR: DOSY-NMR spectra were acquired at 20 °C on a 600 MHz JEOL spectrometer. The spectra were acquired with a bbp_dste_led sequence with gradient spoiling, using 32 scans. The gradient strength was logarithmic incremented of 16 steps from 0 to 0.55 T/m. Diffusion time was set to 0.1 s, delta was 1.2 ms and the relaxation delay was 7 s. Diffusion coefficient values were obtained from fitting peak intensity decays in Delta 5.3.0 software using the curve analysis option.

La-Cav-OH crystallization: Slow evaporation of an acetonitrile + 1% water solution of the complex afforded plate-like transparent crystals over 3 days.

Cav-Py complexes characterization in solution

Materials: *trans*-$Pt(ACN)_2Cl_2$ and *cis*-$Pt(ACN)_2Cl_2$ were purchased from Sigma. $Pt(ACN)_4(CF_3SO_3)_2$ was prepared according to a literature procedure.[41]

NMR: NMR samples were prepared by dissolving **Cav-Py** and an equimolar amount of *trans*-$Pt(ACN)_2Cl_2$ or *cis*-$Pt(ACN)_2Cl_2$ or $Pt(ACN)_4(CF_3SO_3)_2$ in $CDCl_3$. Spectra were acquired at 25° on a 600 MHz JEOL spectrometer.

LAG reaction: Cav-Py (15 mg, 0.0085 mmol) was transferred to an agate mortar. $Zn(NO_3)_2{\cdot}6H_2O$ (4 mg, 0.013 mmol) was added, followed by 10 µL of methanol. The reagents were manually ground together until complete solvent evaporation (15 minutes). The crude was collected and two fractions were separated based on solubility in methanol.

Soluble fraction of Zn-Cav-Py crystallization: Slow evaporation of a methanol solution of the soluble fraction from LAG afforded needle transparent crystals in two days.

X-ray Crystallography

The crystal structures of La-**Cav-OH**, **Cav-OH** and Zn-**Cav-Py** were determined by X-ray diffraction methods. Crystal data and experimental details for data collection and structure refinement are reported in Table S1. Intensity data and cell parameters were recorded at 190(2) K on a Bruker D8 Venture PhotonII diffractometer equipped with a CCD area detector, using the MoKα radiation (λ = 0.71073 Å) for La- **Cav-OH** and **Cav-OH**, and the CuKα radiation (λ = 1.54178) for **Zn-Cav-Py**. The raw frame data were processed

using SAINT and SADABS to yield the reflection data file.[41] The structures were solved by Direct Methods using the SIR97 program[42] and refined on F_o^2 by full-matrix least-squares procedures, using SHELXL-2014[43] in the WinGX suite v.2014.1.[44] In the case of La-**Cav-OH** all non-hydrogen atoms were refined with anisotropic atomic displacements, except when disorder was present. The hydrogen atoms were included in the refinement at idealized geometry (C-H 0.95-0.99 Å, O-H 0.82 Å) and refined "riding" on the corresponding parent atoms with $U_{iso}(H)$ set to $1.2U_{eq}(C)$ and $1.5U_{eq}(O, C_{met})$. When possible, the H atoms of water molecules were found in the difference Fourier map. The weighting scheme used in the last cycle of refinement was $w = 1/ [\sigma^2 F_o^2 + (0.0662P)^2 + 4.1299P]$, where $P = (F_o^2 + 2F_c^2)/3$.

Crystals of **Cav-OH** and Zn-**Cav-Py** diffracted poorly, and the data were not good enough to refine the structures; however, it was possible to determine the species which were formed and their bond connectivity. Main crystallographic data:

Cav-OH. Monoclinic, *C*2/m, a = 14.068(1) Å, b = 23.176(2) Å, c = 24.274(2) Å, β = 93.678(3)°, V = 7898.1(11) Å^3, Z = 4. Theta range for data collection: 2.322 to 17.284°. Reflections collected / unique: 32510 / 2496 [R(int) = 0.1176].

Zn-**Cav-Py**. Triclinic, *P*-1, a = 13.503(3) Å, b = 16.215(3) Å, c = 24.413(5) Å, α= 71.255(9)°, β = 86.985(9)°, γ = 79.410(10)°, V = 4975(2) Å^3, Z = 2. Theta range for data collection: 1.911 to 36.571°. Reflections collected / unique: 17698 / 4627 [R(int) = 0.1594].

Table S1. Crystallographic data for La-**Cav-OH**

Formula	$C_{68}H_{116}Cl_3LaN_4O_{30}P_4$
Formula weight	1838.78
Crystal system	Triclinic
Space group	*P*-1
a/Å	14.6516(2)
b/Å	15.3311(3)
c/Å	20.0375(5)
α/°	92.916(1)
β/°	95.482(1)
γ/°	92.327(1)
V/Å^3	4469.9(2)
Z	2
D_c/g cm^{-3}	1.366
F(000)	1920
μ/mm^{-1}	0.715
$\theta_{min,max}$/°	2.501-28.285
Reflections collected	59535
Independent reflections	20927 [R(int) = 0.0519]
Observed reflections	16270
Data/restr./param.	20927 / 6 / 1073
S[a]	1.029
R[$Fo>4\sigma(Fo)$][b]	0.0490
wR_2[b]	0.1208
$\Delta\rho_{min,max}$/e Å^{-3}	1.624, -1.341

[a]Goodness-of-fit S = $[\Sigma w(F_o^2-F_c^2)^2/ (n-p)]1/2$, where n is the number of reflections and p the number of parameters. [b]$R_1 = \Sigma \| F_o | - | F_c \| / \Sigma | F_o |$, $wR_2 = [\Sigma[w(F_o^2 F_c^2)^2]/\Sigma[w(F_o^2)^2]]^{1/2}$.

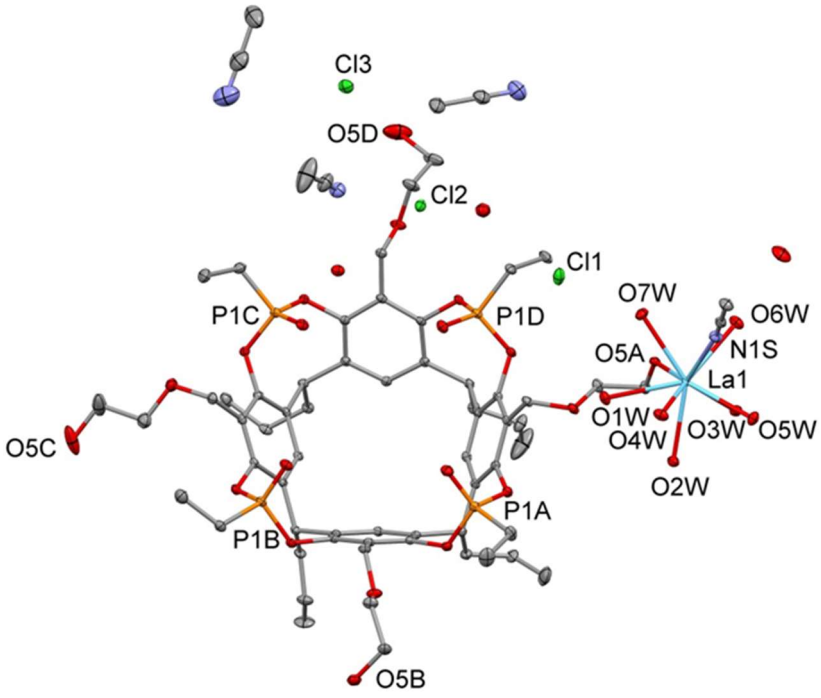


Figure S1. Ortep view of La-**Cav-OH** with partial atom labelling scheme and displacement ellipsoids drawn at the 20% probability level. H atoms have been omitted for clarity.

8 Bibliography

1 J.-N. Rebilly, B. Colasson, O. Bistri, D. Over and O. Reinaud, *Chem. Soc. Rev. Chem. Soc. Rev*, 2015, **44**, 467–489.

2 S. Rat, J. Gout, O. Bistri and O. Reinaud, *Org. Biomol. Chem.*, 2015, **13**, 3194–3197.

3 G. De Leener, D. Over, C. Smet, D. Cornut, A. G. Porras-Gutierrez, I. López, B. Douziech, N. Le Poul, F. Topić, K. Rissanen, Y. Le Mest, I. Jabin and O. Reinaud, *Inorg. Chem.*, 2017, **56**, 10971–10983.

4 J. Hong, K. E. Djernes, I. Lee, R. J. Hooley and F. Zaera, *ACS Catal.*, 2013, **3**, 2154–2157.

5 D. Vidal, M. Costas and A. Lledó, *ACS Catal.*, 2018, **8**, 3667–3672.

6 N. Natarajan, E. Brenner, D. Sémeril, D. Matt and J. Harrowfield, *Eur. J. Org. Chem.*, 2017, **2017**, 6100–6113.

7 M. Nakamura, K. Kishimoto, Y. Kobori, T. Abe, K. Yoza and K. Kobayashi, *J. Am. Chem. Soc.*, 2016, **138**, 12564–12577.

8 W.-Y. Pei, G. Xu, J. Yang, H. Wu, B. Chen, W. Zhou and J.-F. Ma, *J. Am. Chem. Soc.*, 2017, **139**, 7648–7656.

9 N. Le Poul, Y. Le Mest, I. Jabin and O. Reinaud, *Acc. Chem. Res.*, 2015, **48**, 2097–2106.

10 T. Chavagnan, D. Sémeril, D. Matt, L. Toupet, J. Harrowfield and R. Welter, *Eur. J. Inorg. Chem.*, 2016, **2016**, 497–502.

11 L. Monnereau, H. El Moll, D. Sémeril, D. Matt and L. Toupet, *Eur. J.*

Inorg. Chem., 2014, **2014**, 1364–1372.

12 H. El Moll, D. Sémeril, D. Matt, M.-T. Youinou and L. Toupet, *Org. Biomol. Chem.*, 2009, **7**, 495–501.

13 M. Mettry, M. P. Moehlig and R. J. Hooley, *Org. Lett.*, 2015, **17**, 1497–1500.

14 R. Pinalli, E. Dalcanale, F. Ugozzoli and C. Massera, *CrystEngComm*, 2016, **18**, 5788–5802.

15 F. Fochi, P. Jacopozzi, E. Wegelius, K. Rissanen, P. Cozzini, E. Marastoni, E. Fisicaro, P. Manini, R. Fokkens and E. Dalcanale, *J. Am. Chem. Soc.*, 2001, **123**, 7539–7552.

16 R. Pinalli, V. Cristini, V. Sottili, S. Geremia, M. Campagnolo, A. Caneschi and E. Dalcanale, *J. Am. Chem. Soc.*, 2004, **126**, 6516–6517.

17 E. Menozzi, M. Busi, R. Ramingo, M. Campagnolo, S. Geremia and E. Dalcanale, *Chem. Eur. J.*, 2005, **11**, 3136–3148.

18 Y. Nito, H. Adachi, N. Toyoda, H. Takaya, K. Kobayashi and M. Yamanaka, *Chem. Asian J.*, 2014, **9**, 1076–1082.

19 S. P. Bew, A. D. Burrows, T. Düren, M. F. Mahon, P. Z. Moghadam, V. M. Sebestyen and S. Thurston, *Chem. Commun.*, 2012, **48**, 4824.

20 M. Schulz, A. Gehl, J. Schlenkrich, H. A. Schulze, S. Zimmermann and A. Schaate, *Angew. Chem. Int. Ed.*, 2018, **57**, 12961–12965.

21 H. Zhang, R. Zou and Y. Zhao, *Coord. Chem. Rev.*, 2015, **292**, 74–90.

22 M. Moradi, L. G. Tulli, J. Nowakowski, M. Baljozovic, T. A. Jung and P. Shahgaldian, *Angew. Chem. Int. Ed.*, 2017, **56**, 14395–14399.

23 B.-B. Lu, J. Yang, Y.-Y. Liu and J.-F. Ma, *Inorg. Chem.*, 2017, **56**, 11710–11720.

24 B. B. Lu, J. Yang, G. B. Che, W. Y. Pei and J. F. Ma, *ACS Appl. Mater. Interfaces*, 2018, **10**, 2628–2636.

25 W. Jiang, H. Zhang, J. Yang, Y.-Y. Liu, H.-Y. Liu and J.-F. Ma, *CrystEngComm*, 2014, **16**, 9638–9644.

26 Y.-J. Hu, J. Yang, Y.-Y. Liu, S. Song and J.-F. Ma, *Cryst. Growth Des.*, 2015, **15**, 3822–3831.

27 J. R. Ramirez, H. Yang, C. M. Kane, A. N. Ley and K. T. Holman, *J. Am. Chem. Soc.*, 2016, **138**, 12017–12020.

28 I. Brekalo, C. M. Kane, A. N. Ley, J. R. Ramirez, T. Friščić and K. T. Holman, *J. Am. Chem. Soc.*, 2018, **140**, 10104–10108.

29 E. Biavardi, M. Favazza, A. Motta, I. L. Fragalà, C. Massera, L. Prodi, M. Montalti, M. Melegari, G. G. Condorelli and E. Dalcanale, *J. Am. Chem. Soc.*, 2009, **131**, 7447–7455.

30 M. Dionisio, G. Oliviero, D. Menozzi, S. Federici, R. M. Yebeutchou, F. P. Schmidtchen, E. Dalcanale and P. Bergese, *J. Am. Chem. Soc.*, 2012, **134**, 2392–2398.

31 E. Biavardi, C. Tudisco, F. Maffei, A. Motta, C. Massera, G. G. Condorelli and E. Dalcanale, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 2263–8.

32 E. Biavardi, S. Federici, C. Tudisco, D. Menozzi, C. Massera, A. Sottini, G. G. Condorelli, P. Bergese and E. Dalcanale, *Angew. Chem. Int. Ed.*, 2014, **53**, 9183–9188.

33 M. Melegari, C. Massera, F. Ugozzoli and E. Dalcanale, *CrystEngComm*, 2010, **12**, 2057.

34 R. Pinalli, E. Dalcanale, K. Misztal, R. Lucentini, F. Ugozzoli and C. Massera, *Inorg. Chim. Acta*, 2018, **470**, 250–253.

35 M. Urbaniak and W. Iwanek, *Tetrahedron*, 2006, **62**, 1508–1511.

36 L. Avram and Y. Cohen, *Chem. Soc. Rev.*, 2015, **44**, 586–602.

37 P. Delangle, J.-C. Mulatier, B. Tinant, J.-P. Declercq and J.-P. Dutasta, *Eur. J. Org. Chem.*, 2001, **2001**, 3695.

38 A. L. Garay, A. Pichon and S. L. James, *Chem. Soc. Rev.*, 2007, **36**, 846.

39 N. Stock and S. Biswas, *Chem. Rev.*, 2012, **112**, 933–969.

40 F. L. Thorp-Greenwood, T. K. Ronson and M. J. Hardie, *Chem. Sci.*, 2015, **6**, 5779–5792.

41 O. F. Wendt, N.-F. K. Kaiser and L. I. Elding, *J. Chem. Soc. Dalt. Trans.*, 1997, 4733–4738.

42 (a) SADABS Bruker AXS; Madison, Wisconsin, USA, 2004; SAINT, Software Users Guide, Version 6.0; Bruker Analytical X-ray Systems, Madison, WI, 1999; (b) G. M. Sheldrick, SADABS v2.03: Area-Detector Absorption Correction. University of Göttingen, Germany, 1999.

43 A. Altomare, M. C. Burla, M. Camalli, G. L. Cascarano, C. Giacovazzo, A. Guagliardi, A. G. G. Moliterni, G. Polidori, R. Spagna, *J. Appl. Crystallogr.*, 1999, **32**, 115-119.

44 G. M. Sheldrick, *Acta Cryst*. 2008, **A64**, 112-122.

45 WinGX, L. J. Farrugia, *J. Appl. Crystallogr*., 1999, **32**, 837-838.

CHAPTER 3

Molecular recognition between methylated-lysine and 3 synthetic receptors in a model protein*

*This work has been partially carried out in the group of Prof. Peter B. Crowley at National University of Ireland (Galway).

Part of this chapter is published in Angewandte Chemie[1]

1. Supramolecular Chemistry with proteins

Supramolecular building blocks, namely macrocyclic hosts, have been used to aid protein assemblies and to control proteins' behaviour.[2] Proteins are ubiquitous in nature and take part in a wealth of biological processes. Traditionally, proteins have been categorized based on their individual function as catalysts, signal transducers, structural building blocks, etc. Nevertheless, a clear majority of processes involves highly ordered assemblies, where proteins assemble in complex molecular machines.[3–5] To do this, proteins interact with each other through their surfaces, yielding protein-protein interaction networks.[6,7] Recognition is highly specific and often involves simultaneously hydrophobic effect, hydrogen bond and electrostatic interactions. These very same interactions are the ones driving complexation and self-assembly of supramolecular building blocks.[8]

Supramolecular tools can be used to alter and control protein behaviours. Macrocyclic hosts bind to proteins' surfaces, aiding or impeding assemblies or masking hotspots' reactivity.[2] One of the fundamental characteristics of supramolecular chemistry is reversibility, since systems are based on non-covalent interactions. Therefore, protein behaviours can be controlled in a reversible way. For example, dis-assembly of constructs can be triggered once the target is reached. It is also possible to inhibit a biological process by non-covalently binding the hotspot and masking its reactivity. Since the binding is reversible, the process can be re-initiated in response to a determined stimulus.

Various macrocyclic hosts can be used to target specific side chains on protein surface. Selectivity of the recognition is influenced by many

factors, such as accessibility, neighbouring residues, overall polarity of the surroundings. A deep understanding of the complexation is needed in the design of binding sites. Here is where structural characterization comes into play. Both NMR and X-ray crystallography can provide insights on complexation modes.

1.1 AA Recognition

The first step in the understanding of protein-ligand interactions is the characterization at a small molecules level. Therefore, many examples have been provided relative to macrocyclic hosts-amino acids complexes. Calixarenes, cavitands, cucurbiturils and molecular tweezers have proven very useful in this context. They all provide a preorganized structure, a hydrophobic cavity and accessibility to chemical modifications. Their structural differences depend on cavity volume, peripheral groups, portal size and flexibility upon complexation. Such characteristics differ their complexation behaviour and recognition of amino acid elements.

Calixarenes are oligomers obtained by phenol and formaldehyde condensation. The most common homologues are calix[4] and [6], with 4 and 6 phenol rings respectively. The lower and upper rim can be functionalized improving selectivity and broadening the application range.

Calix[4]arenes are the least flexible and are often found in a cone-shaped conformation. Calix[4]arenes can be functionalized with different moieties in order to bind both anionic and cationic residues. Metallic complexes of the naphthyridine calix[4] conjugates can recognize cysteine, histidine, aspartic acid and glutamic acid.[9,10]

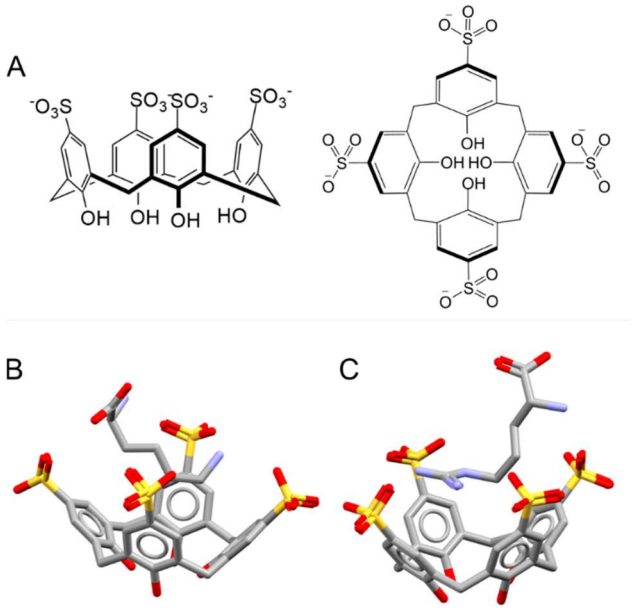


Figure 1. (A) **sclx$_4$** in the cone conformation, side and top views. (B) The **sclx$_4$**-lysine complex (CCDC no. 140210) highlights interaction of the guest with aromatic walls and sulfonates. (C) In the **sclx$_4$**-arginine complex (CCDC no. 245594) arginine is disposed in a planar conformation, with the guanidinium H-bonding to sulfonates. Solvent molecules are omitted for clarity.

The sulfonate derivative of calix[4]arene, namely p-sulfonatocalix[4]arene (**sclx$_4$**) (Figure 1.A), recognizes the cationic amino acids lysine and asparagine.[11,12] X-ray characterization showed that the side chain of lysine was encapsulated in the cavity (Figure 1.B).[11] Cation-π and C-H-π interactions between Lys-C$^{\zeta}$ and the aromatic walls of **sclx$_4$** were observed. In the arginine complex (Figure 1.C), the side chain sits into **sclx$_4$** and is planar from C$^{\gamma}$ to the guanidinium.[12] The guanidinium was hydrogen bonding with the rim sulfonates. Furthermore, **sclx$_4$** binding to methylated

lysine and arginine resulted in micromolar affinity toward methyl-lysine and millimolar affinity toward methyl-arginine.[13,14] **Sclx$_4$** binds the methyl-lysine (LysMe$_n$) series with an affinity that increases with the degree of methylation.[13] A 70-fold selectivity for trimethyl-lysine (LysMe$_3$) over the natural amino acid was reported.

As **sclx$_4$**, also resorcinarene-based receptors showed good recognition abilities toward methyl-lysine.[15–18] Resorcinares are cup-shaped receptors obtained by the condensation of resorcinol and aldehydes. The eight hydroxyl groups at the upper rim can be bridged with various functionalities yielding cup shaped cavitands. In the case of P=O bridging, tetra-phosphonate cavitands (**Tiiii**) are obtained bearing 4 P=O moieties pointing inward the cavity. **Tiiii** can be functionalized both at the upper and lower rim by chemical modification. Their complexation toward LysMe$_n$ relies on cation-π, CH-π, cation-dipole and H-bond interactions. The highest affinity was observed for mono-methyl-Lysine (LysMe)[18] yielding a micromolar K_d in methanol and millimolar in buffered water.[17] This affinity can be explained considering the H-bond pattern. The secondary ammonium in LysMe can H-bond to two of the P=Os simultaneously, thanks to the two N-Hs. X-ray characterization of the complex[17] confirmed this hypothesis, showing encapsulation of the methyl-ammonium. Distances were compatible with two H-bonds between the N-Hs and P=Os. **Tiiii** can also complex dimethyl-lysine (LysMe$_2$) and lysine with a millimolar affinity. The bulkier LysMe$_3$ is too big to fit in the cavity resulting in no complexation. These characteristics were employed in sensing of methylated histone tail peptides with a T-rex nanoparticle-**Tiiii** device.[18]

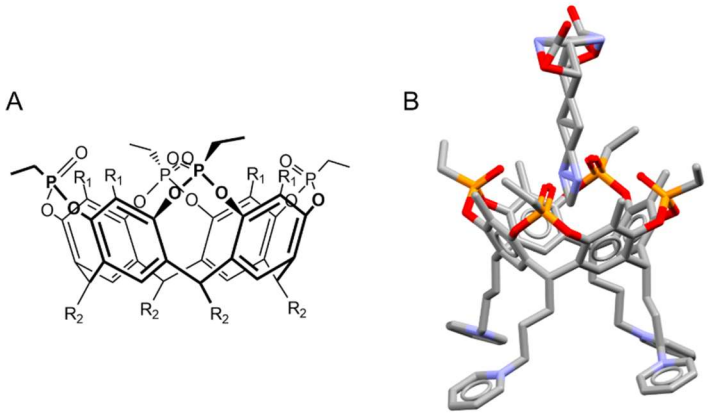


Figure 2. (A) General structure of tetra-phosphonate cavitand (**Tiiii**). The ligand can be functionalized in position R_1 and R_2. (B) The **Tiiii**-LysMe complex (CCDC no. 1471940) shows full encapsulation of the methyl-ammonium. Solvent molecules and chloride counterion are omitted for clarity. Both LysMe split conformations are shown.

The doughnut shaped cucurbituril family was used as well in the recognition of aa and peptides.[19–21] Cucurbiturils (**Q*n***) are relatively rigid macrocycles comprising *n* glycouril units and *2n* methylene bridges, where the diameter and volume of the cavity increase with *n*.[22,23] **Q*n*** have shown complexation properties with very high affinity constants in water.[23] Their complexation abilities rely mostly on the energy gained by realising high energy water from the cavity upon complexation.[24]

Similarly to **sclx$_4$**, **Q7** can recognize methylated lysine and arginine and other basic amino acids.[13,20] With a cavity of 279 Å^3, **Q7** is also suitable for complexing aromatic amino acids. Phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) form complexes with **Q7** with an affinity in the low micromolar range.[19,25] **Q8** (Figure 3.A) is large enough (Vol = 479 Å^3) to

accommodate two guests, yielding ternary complexes.[23] **Q8** can form homo-ternary complexes with the aromatic Phe (Figure 3.B) and Trp but also with the aliphatic Isoleucine (Ile).[26,27] Hetero-ternary complexes are formed with auxiliary guests like methyl viologen.[28]

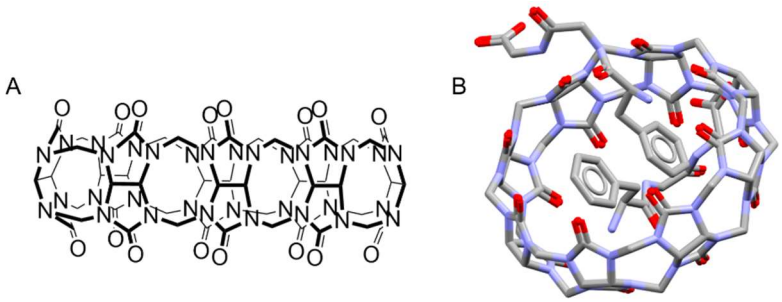


Figure 3. (A) Side view of **Q8**. (B) **Q8** size enables encapsulation of two Phe side chains simultaneously, yielding the homo-ternary complex (CCDC no. 628235). Solvent molecules are omitted for clarity.

Schrader and co-workers designed a series of molecular tweezers comprising norbornadiene and benzene rings alternating units (Figure 4).[29] The tweezer possesses a hydrophobic cavity and is relatively flexible. Various functionalities can be introduced to the central ring. Phosphate functionalization renders the tweezer both water soluble and specific for binding to lysine and arginine. Molecular tweezers complex Lys thanks to hydrophobic effect, CH-π, electrostatic and H-bond interactions.[30,31] The authors predicted the structure of the complex (Figure 4.C) through molecular dynamics simulations since it was not possible to obtain single crystal of the complex.[31]

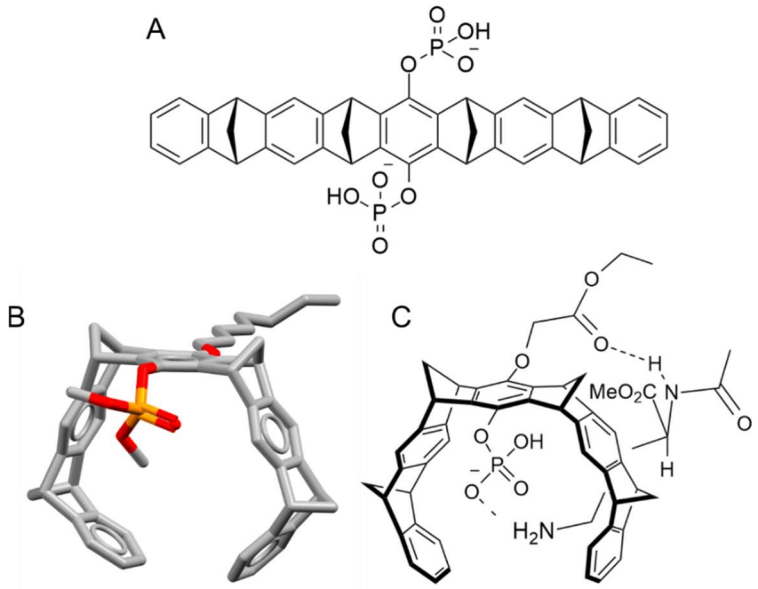


Figure 4. (A) The symmetric di-phosphate molecular tweezer **CRL01**. (B) Un-symmetric mono-phosphate molecular tweezer crystal structure. Solvent molecules are omitted for clarity. (C) Scheme of the molecular tweezer-AcLysOMe complex as seen in ref 23.

1.2 Protein Recognition

Recognition of protein surface elements provides information on how protein interact with each other in biologically relevant processes. Nevertheless, it has so far been a challenge in life sciences to recognize elements on the protein surface. Synthetic supramolecular receptors offer a valid alternative to small molecules or peptide-based methods. Synthetic receptors can be used as probes to understand how molecular recognition works in water[32] and can be helpful in the development of new drug target.

The surface of a protein bears a variety of functional groups and is therefore extremely complex. In the rational design of ligands, it is necessary to account not only for the target side chain itself, but also for its surroundings. Indeed, neighbouring amino acids play often a role in the interaction. Moreover, the target amino acid might be included in a positive or negative patch. We can gain the necessary knowledge by analysing supramolecular protein recognition events. Structural data, especially X-ray single crystal characterization, help in elucidating the role of neighbouring residues. On a protein surface there is, almost certainly, more than one possible binding site for the ligand. Understanding if there is a preferred binding site and the reason behind this selectivity provides information on the specific requirements. Such information will aid in the design of stronger and more specific ligands.

Up to now few crystal structures of protein-macrocycle complexes have been provided, even if the number of examples is increasing.

In 2012, **sclx$_4$** was shown to complex cytochrome-c (cyt-c).[33] X-ray crystallography revealed 3 binding sites: Lys89, Lys22 and Lys4 (Figure 5.B). As in the free amino acid case,[11] the complex involved encapsulation of the side chain (Figure 5.A). The side chain was bent in order to form salt bridges with the rim sulfonates. Analysis of the packing in the crystal structure allowed the calculation of a contact area of 200 Å^2, compatible with the HSQC-NMR data. The combination of solution and X-ray study led to a model in which **sclx$_4$** is exploring cyt-c surface. This transient interaction leads to a camouflage of the positive protein surface. A similar mechanism was demonstrated for human ubiquitin and supramolecular anions.[34]

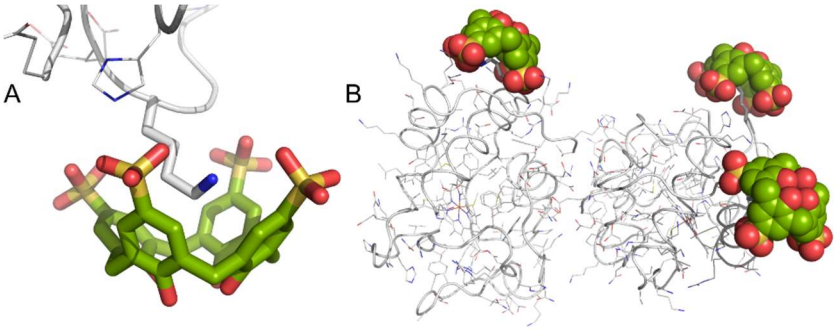


Figure 5. (A) The **sclx$_4$**-Lys complex in cyt-c at Lys4 (PDB code 3TYI). Complexation mode is comparable to the **sclx$_4$**-lysine complex (Figure 1) The protein is represented as ribbon, **sclx$_4$** and Lys4 side chain are represented as sticks. (B) The asymmetric unit comprises two molecules of cyt-c and three molecules of **sclx$_4$** binding to Lys4, Lys89 and Lys22. **Sclx$_4$** are represented as spheres.

However, **sclx$_4$** binds selectively to methylated lysine (KMe$_n$) over regular Lysine.[13,35,36] A structure of **sclx$_4$** in complex with dimethyl-Lysozyme (Lysozyme-LysMe$_2$)[36] illustrated the interactions taking place (Figure 6). Unlike the **sclx$_4$**-Lys complex,[11,33] the ammonium in LysMe$_2$ was positioned in the middle of the cavity (Figure 6.A). One methyl was pointing toward the core of **sclx$_4$** maximizing CH-π interactions. **Sclx$_4$** acts a four-walled aromatic cage mimicking the binding site of histone reading proteins. Among the six LysMe$_2$ in Lysozyme, only Lys116Me$_2$ is complexed by **sclx$_4$**. Steric accessibility is higher than for other LysMe$_2$ and the neighbouring aspartate Asn106 hydrogen bonds to **sclx$_4$** sulfonates. Binding also occurred at Arg14 (Figure 6.B). The complex resembles the one observed for the aa,[12] with the guanidinium unit being planar inside **sclx$_4$** cavity.

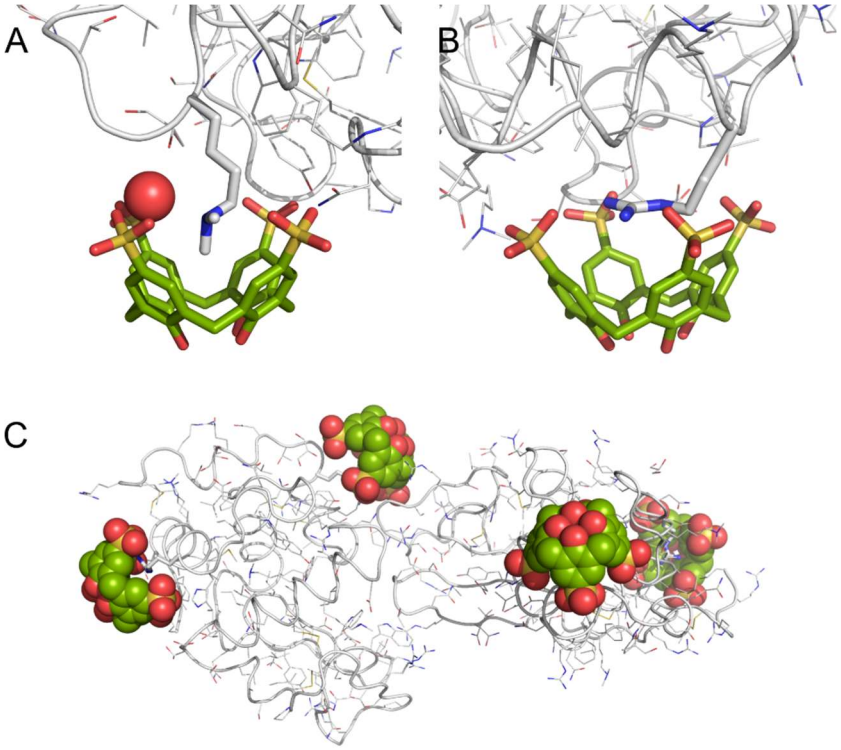


Figure 6. The Lysozyme-LysMe_2 – **sclx$_4$** complex (PDB code 4N0J). (A) Binding at Lys116Me_2 involved encapsulation of the side chain, with one methyl pointing towards the core of **sclx$_4$**. A water molecule (red sphere) persisted in solvating the ammonium. (B) Arg14 is also complexed by **sclx$_4$** with a binding mode resembling the aa complex. (C) The asymmetric unit comprises two molecules of Lysozyme-LysMe_2 and four molecules of **sclx$_4$**.

The molecular tweezer **CLR01** was shown to modulate the behaviour of 14-3-3 adapter proteins as a result of binding to solvent exposed lysine side chains.[37,38] In 2013, a crystal structure of the 14-3-3σ-**CLR01** complex was reported.[37] In the structure the tweezer binds to a single surface-exposed

lysine (Lys214) out of the 14 possibilities. This selectivity was explained by the authors accounting for the additional interactions by neighbouring residues. In fact, H-bonds were observed between **CLR01** phosphates and the surrounding amino acids. Additionally, the protein shielded **CLR01** from bulk water via three apolar residues. Lys214 is flanked between Tyr218, Thr217 and Leu218 providing an optimal fit. The authors illustrated how the other accessible lysines are not bound because of the more cationic chemical environment that can divert **CLR01**. Lys214 is close to the binding site for signalling proteins such as C-RAF. **CLR01** can therefore impede the binding or generally change the dynamics of these partner proteins.

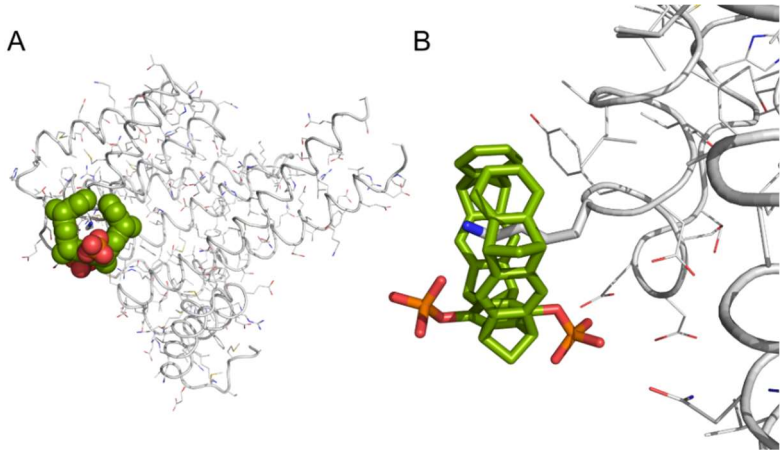


Figure 7. (A) The asymmetric unit in the **CLR01**-14-3-3 complex comprised one protein molecule and one ligand molecule (PDB code 5OEH). High selectivity for Lys214 was observed. (B) Detailed view of the Lys214 - **CLR01** complex. Lys214 is interacting with **CLR01** via hydrophobic effect, CH-π, electrostatic and H-bond interactions.

Q7 was shown to bind to human Insulin and other proteins bearing N-terminal aromatic residue.[39,40] X-ray crystallography of the insulin-**Q7**

complex[40] revealed encapsulation of N-terminal phenylalanine (Phe1) side chain (Figure 8). Insulin N-terminus is flexible enough to partially unfold aiding complexation. The complex is additionally stabilized by H-bonds between Phe1 NH_2 and the rim carbonyls of **Q7**. Again, the neighbouring residues played a role and H-bonds were observed between **Q7** carbonyls and Valine2 and Asparagine3.

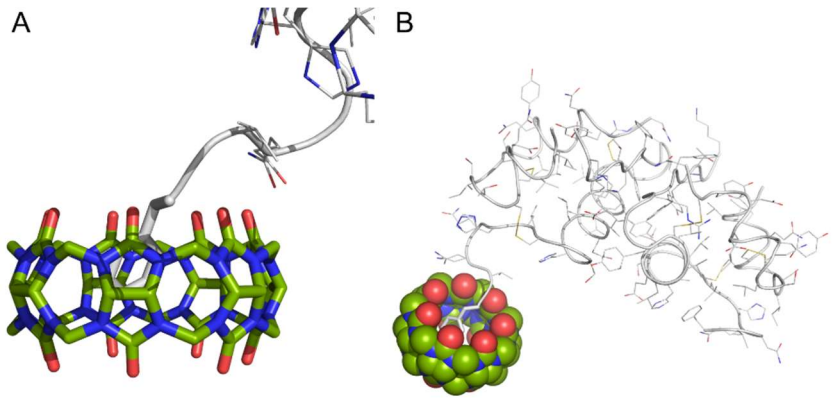


Figure 8. (A) Detailed view of the N-terminal Phe - **Q7** complex in Insulin (PDB code 3Q6E). N-terminal Phe side chain is fully encapsulated in **Q7** cavity, while the upper rim carbonyls are H-bonding to the backbone NH. (B) The asymmetric unit comprises one molecule of **Q7** and two molecules of Insulin. **Q7** is complexing the N-terminal Phe on chain A.

1.3 Protein Assembly

Well-ordered protein aggregates can be found in actin filaments, viral capsids, tubulins and flagella and in the more complexed nucleosomes and

ribosomes. Nature design proteins and nucleic acid chains to specifically and dynamically interact and assemble yielding complex and elaborate structures with specific functionalities.[41,42] Inspired by nature, scientists have been trying to rationally design and control protein and nucleic acids assemblies.[43] Hence, protein assembly emerged as an inspiring field of study, as it leads to sophisticated biomaterials. In addition, knowledge on assembly formation can also be useful in better understanding many diseases. For example, viruses, like Ebola virus, assemble in complex structures.[44] Moreover, neurodegenerative diseases like Alzheimer and Parkinson are generated by abnormal protein assemblies.[45,46]

While the field of programmable DNA self-assembly is fairly developed, the situation with proteins is at a quite early stage. Proteins are attractive building blocks because of their sophisticated structures and broad functions. Nevertheless, handling proteins is complicated by their instability to pH, temperature, ionic strength and solvent conditions. Moreover, the characteristics making proteins attractive building blocks like heterogeneity, flexibility and complexity, are also the ones making rational design more challenging.

Quite some research effort has been done with this aim, with computational design as a widely employed approach.[47–49] This strategy involves cloning the designed aa sequence in a vector for expression in bacteria (often E. Coli). Another approach is the employment templates such as peptides, like coiled coils, or self-assembled DNA, like DNA origami. Chemical strategies have also proven very useful, with host-guest recognition playing an important role. A wealth of structural evidence is available for protein assembly, by supramolecular receptors.[37,50–54]

Tight, selective binding to N-terminal aromatic residues[26,39,40] has paved the way for the development of **Q8**-mediated dimerization,[55] polymerization[56] as well as ternary complex formation.[50] With its ability of encapsulating two guests contemporary, **Q8** is the ideal building block for protein dimerization.[50,55] A versatile approach is the incorporation of N-terminal FGG (Phenylalanine Glycine Glycine) motif that yields to a surface exposed Phe, available for **Q8** binding. This approach was used to dimerize 14-3-3 aptamer protein.[50] Fusion of the FGG motif to 14-3-3 binding epitopes (ERα) led to a peptide that is simultaneously complex by **Q8** (at the FGG site) and by 14-3-3 (at the ERα site). This enabled the construction of a binary protein assembly platform. The FGG-**Q8** complex was successfully employed to create nanowires[56] and nanosprings,[57] demonstrating the potential of the ternary complex.

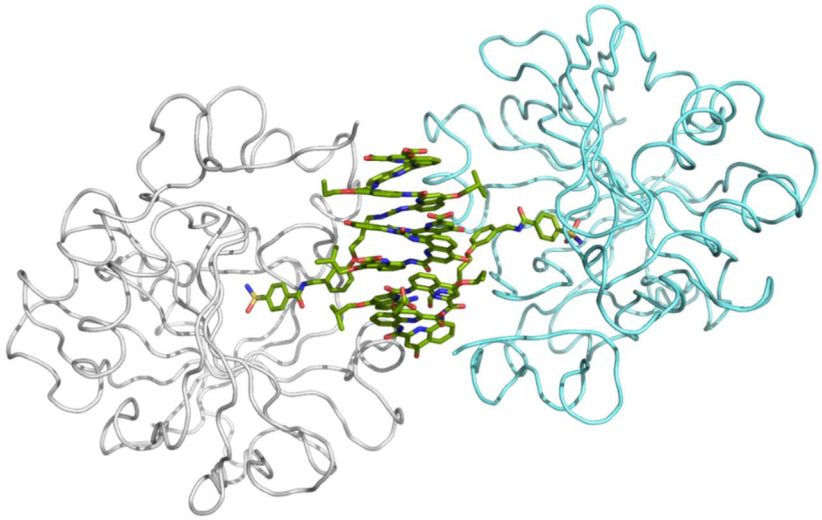


Figure 9. Foldamer mediates HCA dimerization (PDB code 5LVS). The two HCA chains are grey and cyan ribbon. The foldamer is represented as green sticks.

Dimerization via foldamer binding was also reported on the model protein human carbonic anhydrase II (HCA).[54] A crystal structure showed the formation of a HCA_2-$foldamer_3$ complex. Complex formation was mainly driven by hydrophobicity of the foldamer and by the interactions taking place between the foldamer and HCA side chains.

Also calixarenes have been extensively employed for protein assembly.[51–53,58] P-phosphonato-calix[6]arene (**pclx$_6$**) promoted dimerization of the model protein cyt-c, both in solution and at the solid state.[52] X-ray crystallography of the complex highlighted a self-dimer of **pclx$_6$** bringing together two cyt-c molecules. Interestingly, **pclx$_6$** was nestling in a well-defined area on cyt-c surface.

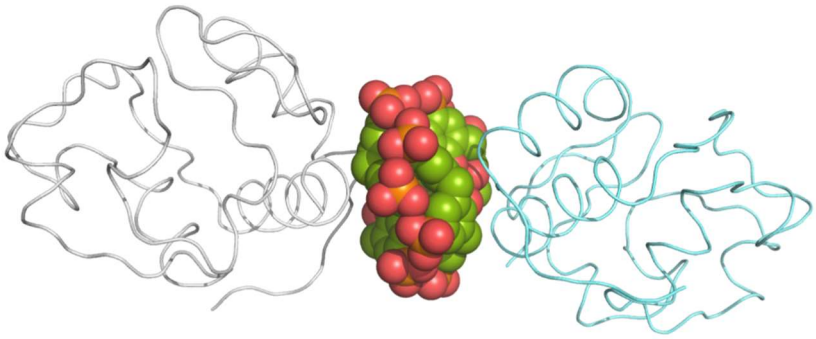


Figure 10. Cyt-c dimerization mediated by a **pclx$_6$** homodimer (PDB code 5LYC). The two cyt-c molecules are grey and cyan respectively. **Pclx$_6$** dimer is represented as green spheres.

The larger sulfonato-calix[8]arene (**sclx$_8$**) triggered tetramerization of cyt-c in solution.[53] The complex dis-assembled at >2 equivalents of **sclx$_8$**, mimicking the biological process of autoregulation. This example remarkably shows how assemblies mediated by synthetic receptors can be controlled via chemical stimuli.

In this context, there's an increasing need of new host-guest motifs to be employed for protein recognition and assembly. In this chapter, I will discuss the binding of three synthetic receptors to di-methylated lysine in a model protein.

2. Lysine Methylation

Proteins are governed by several post translational modifications, namely chemical alterations, of exposed residues. Mono-, di- and tri-methylation of the lysine ammonium (N^{ζ}) (Figure 11) are common post translational modifications that occur most notably in histones, with vast ramifications for gene expression.[59,60] The discovery that hundreds of non-histone proteins are methylated at lysine side chains broadened the research interest.[61–63] Furthermore, a non-regulated degree of methylation has been associated with various forms of cancer.[64]

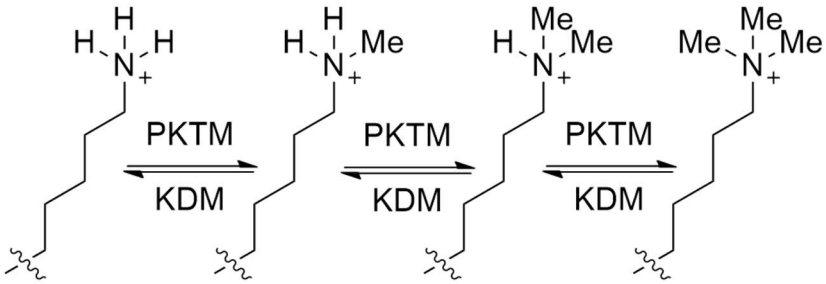


Figure 11. Lysine methylation is controlled by the balanced activity of the enzymes protein lysine methyltransferases (PKTMs) and demethylases (KDMs).

Methylated lysines ($LysMe_n$) present a unique hotspot for recognition by reader proteins that possess an aromatic cage motif. Consequently, synthetic receptors that recognize $LysMe_n$ hold great potential as probes to study the biological systems and as inhibitors of protein-protein interactions.[14,16,18,35,36,65,66]

We decided to study the complexation abilities of three synthetic receptors (Figure 12) to dimethyl-lysine in a model protein. The three

receptors chosen were water soluble tetra-phosphonate cavitand, p-sulfonatocalix[4]arene and cucurbit[7]uril. Their affinity for $LysMe_n$ has been studied before at the amino acids level,[11,13,17] but only in the case of **sclx$_4$** on a model protein.[36] All the three receptors possess a hydrophobic cavity with comparable size. Nevertheless, their complexation abilities rely on a different set of interactions, as explained in the paragraph 1.1.

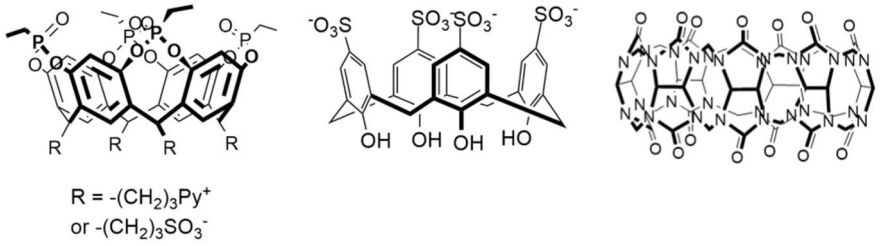


Figure 12. The three receptors employed in this study. From the left water soluble **Tiiii**, **sclx$_4$** and **Q7**.

3. Choice of the model system

Ralstonia solanacearum lectin (RSL) was chosen as the model system, because of its high stability and ease of purification. RSL is an extensively characterized trimer with a six-bladed β-propeller topology and a pI~6.5.[67–69] It comprises 3 homologues monomers and 6 sugar binding sites. RSL has 4 possible methylation sites for each monomer, 3 lysines (Lys25, Lys34 and Lys83) and the N-terminal serine (Ser1) (Figure 13).

RSL and ^{15}N labelled RSL samples were produced by over-expression in *E. Coli* according to literature procedure.[67,68]

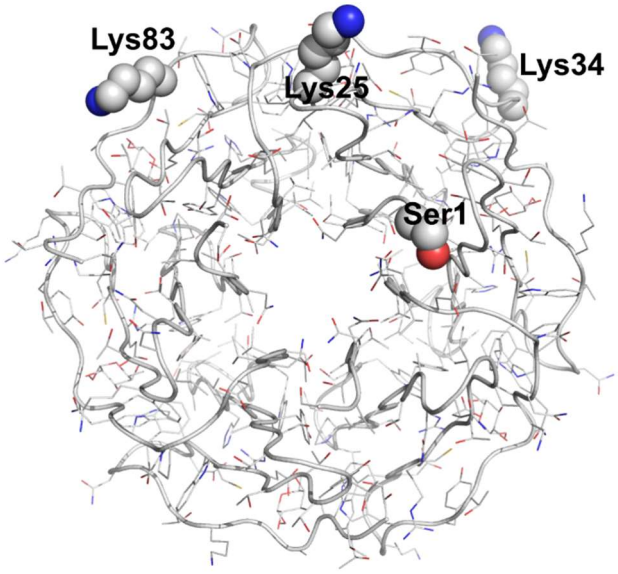


Figure 13. RSL trimer (PDB code 2BT9). At chain A, the four possible methylation sites (Ser1, Lys25, Lys34 and Lys83) side chains are represented as spheres.

Methylation of lysines and N-termini was achieved by following an established protocol (Figure 14).[70,71] The procedure consists in a reductive methylation, carried out in the presence of formaldehyde and of a reducing agent, namely sodium cyanoborohydride or dimethylamino borane complex (DMAB). This protocol yields mono-methyl and di-methyl lysine depending on the stoichiometric amounts of the reagents.

We made various attempts at mono-methylating lysine residues, but the product was not detected by ESI mass spectrometry. A complex mixture of methylated products was present instead, containing native, mono- and di-methyl lysines. The impurity of the sample led us to employ only dimethyl-RSL (RSL-LysMe$_2$) in the study. In fact, by using an excess of formaldehyde it is possible to fully convert RSL to RSL-LysMe$_2$.

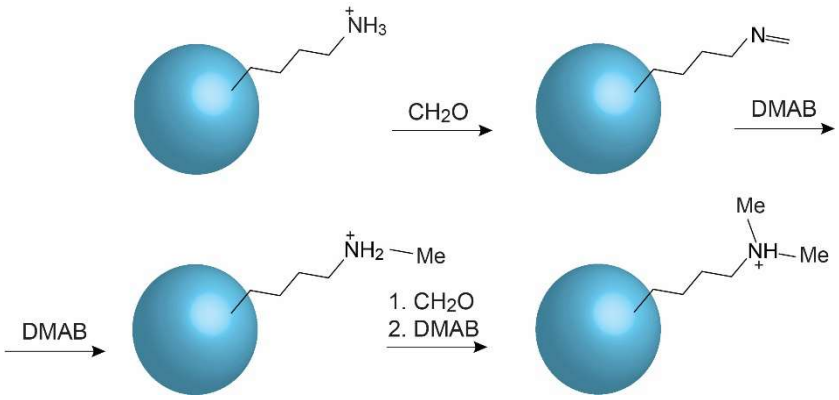


Figure 14. Scheme of the methylation reaction. The protein is represented as a blue sphere.

Full methylation was confirmed by ESI-MS yielding Lys25Me$_2$, Lys34Me$_2$, Lys83Me$_2$ and the N-terminus Ser1Me$_2$.

^{1}H-^{15}N HSQC NMR was measured to confirm the success of the reaction (Figure 15). Backbone amide NHs resonances were assigned by comparison with native RSL (Figure 15). Overall, resonances didn't appear dramatically shifted, confirming that the modification didn't affect RSL folding. Lysines showed minor chemical shift perturbation, consistent with the modification. The N-terminus, i.e. Ser1, is not detectable my NMR, as well as Ser2. Therefore, the closest probe to N-terminus is Valine 3 (Val3). Chemical shift perturbation at Val3 corroborates methylation of the N-termini.

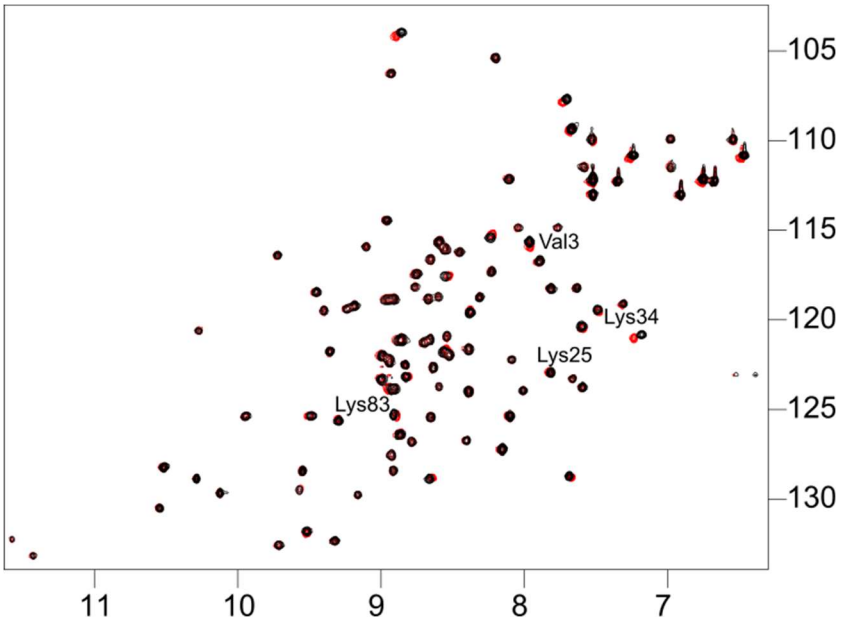


Figure 15. Comparison of the overlaid ^{1}H-^{15}N HSQC spectra of native RSL (black contours) and modified RSL (red) revealed minor chemical shift perturbations due to dimethylation.

4. Tetra-phosphonate cavitands

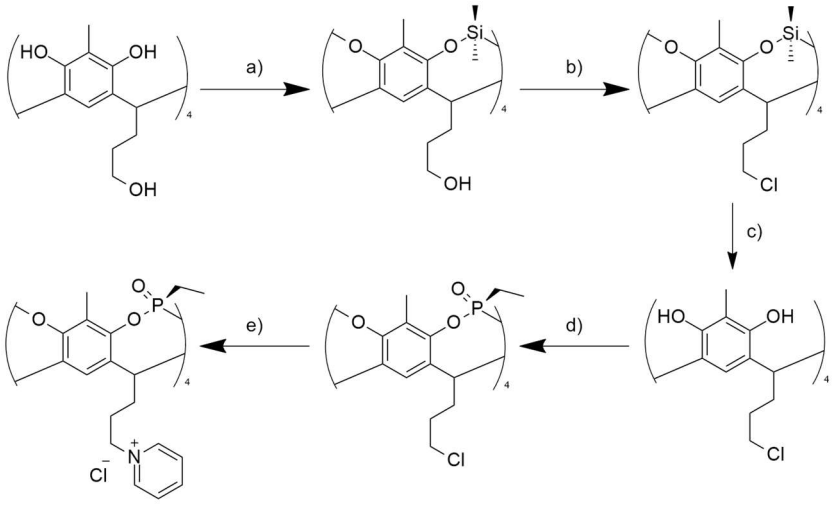
Our group has already highlighted the ability of tetra-phosphonate cavitands to recognize methylated lysine.[17,18] A sensing device was developed grafting **Tiiii** on T-rex nanoparticles.[18] Nevertheless, binding has never been tested on a protein. Hence, we synthesised two water soluble cavitands and tested their recognition of RSL-LysMe_2.

4.1 Synthesis of water soluble cavitands

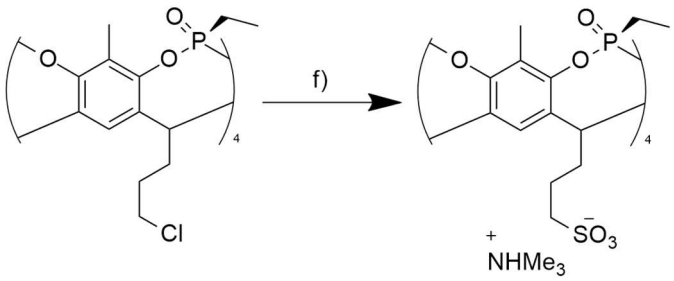
Two cavitands were synthesized, bearing ionic groups to provide water solubility (Figure 12). Pyridinium (**Tiiii-1**) and sulfate (**Tiiii-2**) have been picked as functionalities, in order to test the effect of cationic or anionic substituent on protein complexation.

Tiiii-1 was obtained following a previously reported procedure.[17] A brief scheme of the reactions is shown in Scheme 1. The synthesis was carried out starting from readily available OH-footed resorcinarene. Protection of phenolic OHs with silane was necessary in the following chlorination of lower rim OHs. After the chlorination, phenolic hydroxyls were deprotected with aqueous solution of HF. Deprotection was followed by phosphonate bridging, yielding the four P=O groups pointing inward the cavity. The last step was the substitution of chlorides to achieve the pyridinium salt.

Tiiii-2 was easily obtained by changing the last step of the synthesis (Scheme 2). Instead of pyridine, sulphur trioxide trimethylamine complex was used in the substitution, yielding the sulfonate salt.



Scheme 1. Synthesis of water-soluble tetra-phosphonate cavitand **Tiiii-1**. a) $SiMe_2Cl_2$, Pyridine, 100°C, 3h, 85%; b) $SOCl_2$, DMF cat., toluene, 55°C, 12 h, quant.; c) HF aq., DMF, 50°C, 12 h, quant.; d) (1) $EtPCl_2$, Pyridine, 70°C, 3h; (2) H_2O_2 35%, r.t., 30 min, 57% (over two steps); e) Pyridine, 110°C, 72 h, quant.



Scheme 2. Synthesis of water-soluble tetra-phosphonate cavitand **Tiiii-2**. f) $NMe_3{\cdot}SO_3$, Pyridine, DMF, 50°C, 6 h, quant.

4.2 Tiiii-RSL-LysMe2 complexation in solution

Recognition of **Tiiii-1** on the di-methylated model protein (RSL-$LysMe_2$) was tested initially in solution. A 40 mM stock solution of **Tiiii-1** was titrated against a 0.1 mM solution of ^{15}N-RSL-$LysMe_2$. Each titration step was monitored by ^{1}H-^{15}N HSQC NMR. Chemical shift perturbations were followed as they indicate complexation events. In Figure 16, panels with zooms on $LysMe_2$ resonances are shown. **Tiiii-1** seemed to interact significantly only with $Lys34Me_2$, the most solvent accessible dimethyl-lysine. Also, neighbouring residues showed chemical shift perturbation, indicating were the cavitand might sit.

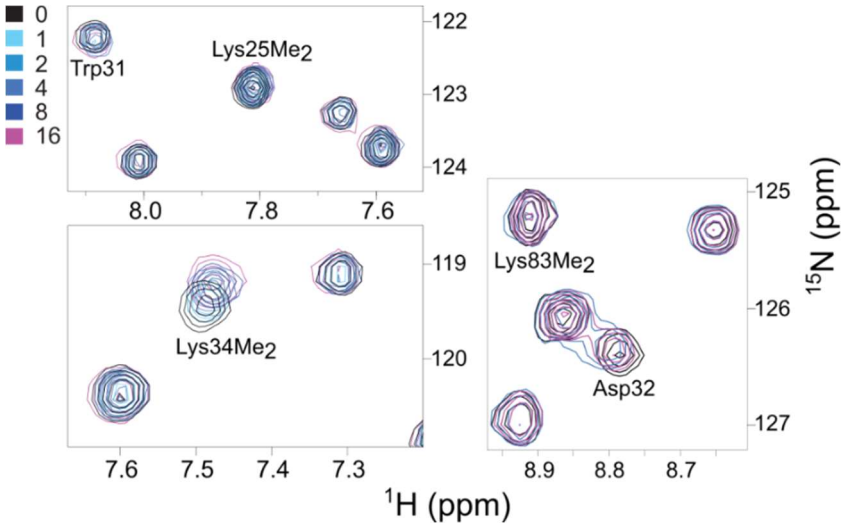


Figure 16. Spectral regions from the overlaid ^{1}H-^{15}N HSQC spectra of RSL-$LysMe_2$ in the presence of 0-16 eq **Tiiii-1**. Each panel provides data for at least one of the possible binding sites. Significant chemical shift perturbations were observed for $Lys34Me_2$ only.

Analysis of Δδ as a function of ligand concentration yielded binding curves that fit to a Kd ≈ 4.6 mM.

Resonances showing chemical shift perturbation were mapped on the surface of RSL-LysMe$_2$ (Figure 17). The map provides a qualitative idea of **Tiiii-1** binding surface, which is centred on Lys34Me$_2$.

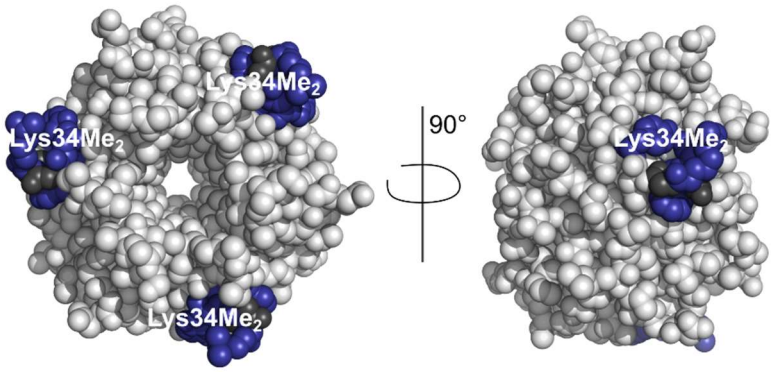


Figure 17. Space filling representation of RSL-LysMe2 highlighting **Tiiii-1** binding surface. Trp31, Asp32, Lys34Me$_2$, Tyr37, for which the amide resonance had a significant chemical shift perturbation, are in blue. Gly33 and Gly35, not detectable by NMR, are dark grey.

Anionic **Tiiii-2** binding to RSL-LysMe$_2$ was tested as well. Since solubility of **Tiiii-2** was lower than **Tiiii-1**, it was only possible to obtain a 5 mM stock solution. Therefore, the titration was complicated by the low concentration reached. It was necessary to halve RSL-LysMe$_2$ concentration and to use a double number of scans.

The titration resulted in chemical shift perturbations equivalent to the ones observed for the cationic **Tiiii-1** (Figure 18). This observation suggests that the binding event occurring is very similar or identical. A change in the

overall charge of the tetra-phosphonate cavitand didn't affect RSL-LysMe$_2$ complexation.

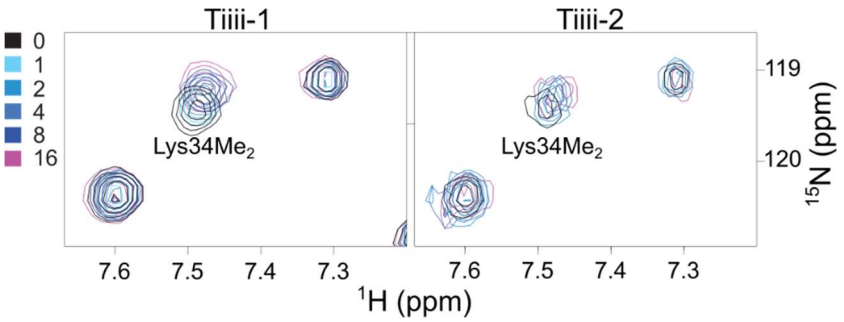


Figure 18. Lys34Me$_2$ amide resonance from the overlaid ^{1}H-^{15}N HSQC spectra of RSL-LysMe$_2$ in the presence of 0-16 eq **Tiiii-1** (left) and 0-16 eq **Tiiii-2** (right). Almost identical upfield perturbations were observed.

4.3 Crystallization trials

Crystallization trials were initiated employing a sparse matrix screen (Jena JCSG++, Oryx 8 robot) and in manual screens using the hanging drop technique. Crystals were tested in-house on a Rigaku diffractometer. The first diffraction images suggested a unit cell too small to be compatible with the presence of a protein, but consistent instead with the ligand alone. Not surprisingly the crystals grew from conditions containing either ammonium or calcium, which can be complexed by the cavitand. Additional crystallization trials were initiated, avoiding ammonium or calcium in the buffer components, but these conditions didn't result in crystallization.

5. P-sulfonatocalix[4]arene

Crowley group has extensively employed **sclx$_4$** for protein recognition and assembly.[33,36,51] Furthermore they showed how calixarenes can act as molecular glue facilitating protein crystallization.[58] Since protein crystallization remains the bottleneck in structure determination, several techniques have been developed to improve crystallization processes.[72–74] The use of small molecules as molecular glues is a reliable strategy to aid crystal growth.[75–77] Calixarenes are ideal molecular glue candidates, since they can interact simultaneously with more than one protein molecule. All of the examples reported so far involved anionic calixarenes binding to highly cationic proteins.[33,51–53,58] Here, we show that a neutral protein, when dimethylated, co-crystallizes with **sclx$_4$**.

5.1 Sclx$_4$ – RSL-LysMe$_2$ complexation in solution

A 95 mM **sclx$_4$** stock solution was titrated against RSL-LysMe$_2$. Similarly to **Tiiii**, the presence of **sclx$_4$** resulted in upfield perturbation of Lys34Me$_2$ backbone amide resonance (Figure 26, Paragraph 6). Neighbouring resides Trp31, Asp32, Tyr37 and Thr38 were affected as well. Thanks to the ligand high water solubility, it was possible to reach up to 64 equivalents. Several resonances shifted in the presence of a large amount of **sclx$_4$** (> 8 eq), including Lys83Me$_2$ and Val3, the closest detectable residue to the N-terminus. This observation suggested a wide binding surface at high equivalents of the ligand.

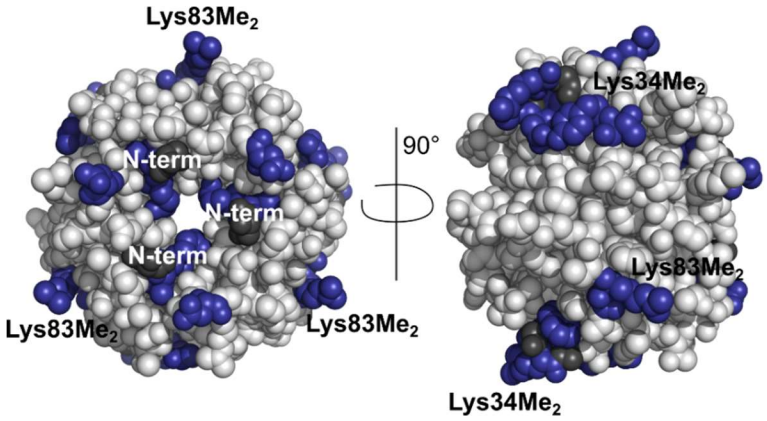


Figure 19. Space filling representation of RSL-LysMe$_2$ highlighting **sclx$_4$** binding surface. Val3, Gln4, Asn23, Trp31, Asp32, Lys34Me$_2$, Tyr37, Thr38, Lys83Me$_2$ and Gly84, for which the amide resonance had a significant chemical shift perturbation, are in blue. Ser1Me$_2$, Ser2, Gly33 and Gly35, not detectable by NMR, are grey.

5.2 Sclx$_4$ - RSL-LysMe$_2$ structure determination

Large crystals were easily obtained in a sparse matrix screen (Jena JCSG++, Oryx 8 robot) in various conditions. Condition B1 (Figure 20), containing 0.8 M ammonium sulfate, 0.1 M sodium citrate pH 4.0, yielded ~700 µM crystals that diffracted to a 1.5 Å resolution. X-ray data collection was performed in-house on a Rigaku diffractometer. The structure was solved in space group $P2_12_12_1$ by molecular replacement using PDB entry 2bt9 as a search model.

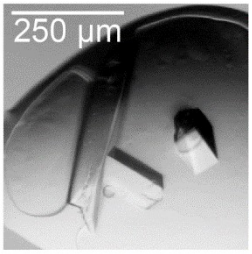


Figure 20. Large co-crystals of dimethylated RSL and **sclx$_4$** were obtained in 0.8 M ammonium sulfate, 0.1 M sodium citrate, pH 4.0. The control drop, without **sclx$_4$**, did not yield crystals or precipitates.

The asymmetric unit comprised a trimer of RSL-LysMe$_2$ and 6 calixarenes (Figure 21). The presence of dimethylated lysines and S1 was clear in the initial electron density maps. Surprisingly **sclx$_4$** complexation was asymmetric and occurred at the sidechain of Lys83Me$_2$ on chain C and on the 3 methylated N-termini, Ser1Me$_2$. Two additional calixarenes were refined whose cavities contained a PEG fragment and water respectively.

Lys83Me$_2$ was buried in the cavity with both methyl substituents sitting in the core of **sclx$_4$** (Figure 22). As described previously for the LysMe$_2$-**sclx$_4$** complex,[36] one of the methyls is pointing towards the centre of **sclx$_4$** cavity. The interactions established in the complex suggested a binding mode similar to Lysozyme-LysMe$_2$. Distances between methyl carbon (C^{η}) and centroids of the phenylene rings (3.7-3.9 Å) suggested CH-π interactions. One water molecule was refined in the calixarene cavity and mediated salt bridge between ammonium N^{ξ} and sulfonates ($O^{W}\cdots N^{\xi}$ ~2.8 Å). Binding occurred in a different geometry compared to the dimethylated Lysozyme case. The plane of the rim sulfonates intersected the C^{δ}-C^{ε} bond. The cavity resulted more open, due to a different angle formed by two of the phenyl rings and the methylene bridges.

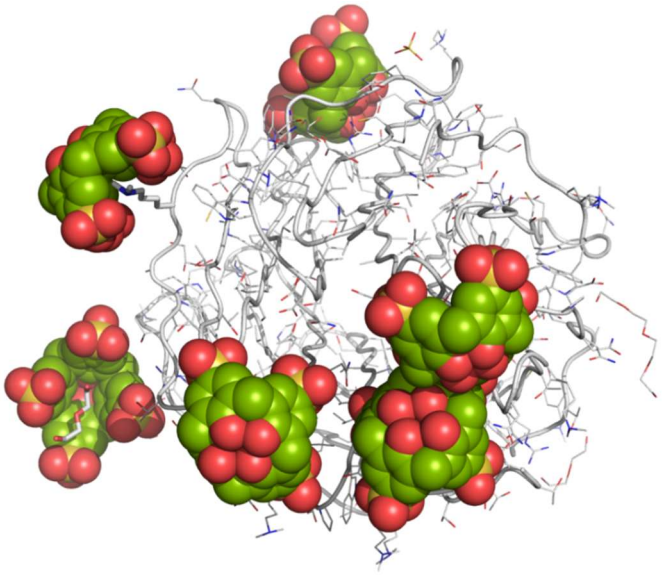


Figure 21. The asymmetric unit comprises a trimer of RSL-LysMe$_2$ and 6 molecules of **sclx$_4$**. C.Lys83Me$_2$ and an entrapped PEG fragment are represented as sticks.

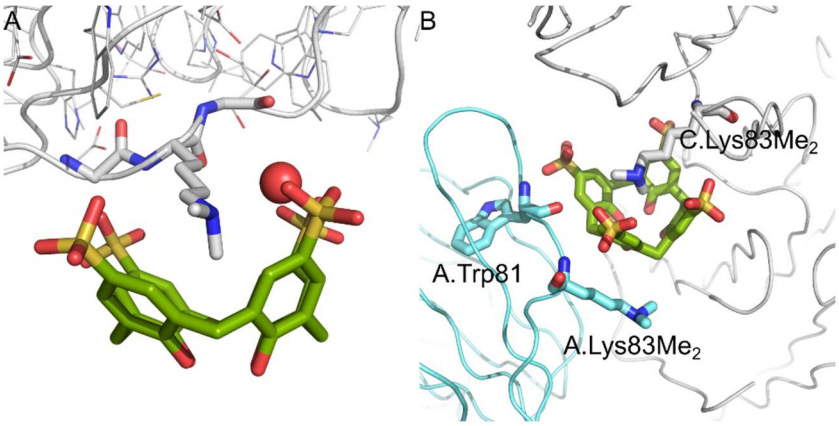


Figure 22. Detailed view of the **sclx$_4$** - LysMe$_2$ complex. (A) C.Lys83Me$_2$ is encapsulated in **sclx$_4$** cavity, while a water molecule (red sphere) persisted

in solvating the ammonium. Lys83Me$_2$ and the two adjacent residues backbone are represented as sticks. (B) **Sclx$_4$** is acting as a molecular glue between two RSL-LysMe$_2$ trimers, grey and cyan respectively. Residues H-bond to the sulfonates are highlighted as sticks.

As previously noticed, **sclx$_4$** had a tendency of interacting with LysMe$_2$ backbone NH (SO···NH 2.8 Å) (Figure 22.B). **Sclx$_4$** mediated the interface between two protein trimers, by interacting with backbone NHs of Lys83Me$_2$ and Trp81 chain A from the symmetry related molecule (SO···NH 2.85-3.1 Å).

A new interaction mode was observed at the N-termini, where **sclx$_4$** was bound to dimethylated serine (Figure 23). Here the methylated ammonium is again encapsulated in the cavity. Complexation occurred in a different fashion at the three homologues chains (Figure 23).

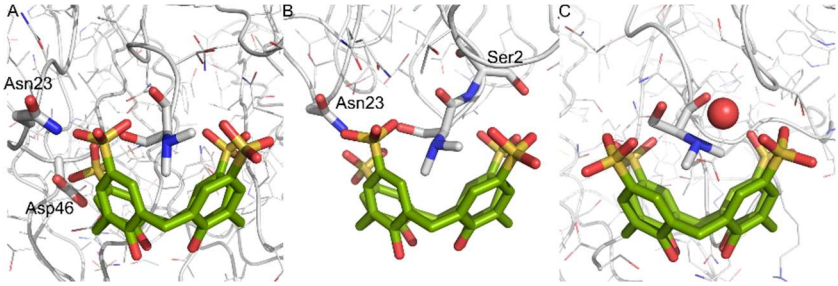


Figure 23. Different complexation modes of N-terminal SerMe$_2$ by **sclx$_4$** in chain A, B and C. Residues interacting with **sclx$_4$** are highlighted as sticks (see main text for details). A retained water molecule at C.Ser1Me$_2$ site is shown as a red sphere.

At chain A (Figure 23.A), one methyl was pointing towards the centre of **sclx$_4$** with the ammonium forming salt bridge with one of the sulfonates (SO···N ~2.8 Å). **Sclx$_4$** formed H-bond with Ser1Me2 hydroxyl (SO···OH 2.8 Å),

backbone NH (SO···NH 3.1 Å) and carbonyl (SO···O=C 3.1 Å). As a consequence, Ser1Me2 resulted completely masked by the calixarene. Additional interactions with the protein surface were provided by H-bonds between sulfonates and Asn23 and Asp46 side chains.

A similar situation was observed at chain B (Figure 23.B), but in this case a sulfonate was bound to Ser2 hydroxyl group (SO···OH 2.6 Å). The presence of this calixarene strongly influenced the packing. **Sclx$_4$** induced a structural rearrangement in the loop 11-14 from the symmetry related molecule. Two split conformations were refined, since the calixarene had 70% occupancy.

The complexation at chain C (Figure 23.C) appeared quite different due to the presence of a retained solvent molecule. The dimethylammonium was disposed in a fork conformation and remained solvated by a water molecule. The same water molecule H-bonded also to Ser1Me$_2$ hydroxyl and carbonyl and mediated an interaction with the sulfonate.

Furthermore, at N-termini we observed different calixarene conformations (Figure 23). Flexibility of **sclx$_4$** allows the ligand to adapt to the substrate when binding.

The two **sclx$_4$** bound to N-termini A and C assembled with the PEG-containing **sclx$_4$** (Figure 24.A). Distances were consistent with π-π stacking (~3.2 Å) between the aromatic walls. The calixarene cluster introduced a new major surface at N-termini. The cluster mediated interfaces between three RSL-LysMe$_2$ trimers and was therefore directly involved in crystal packing (Figure 24.B). As reported previously,[58] calixarene surface decoration may be a special case of patchy particle model.[78] The new

surface provided in fact a "sticky patch" necessary for protein crystallization.

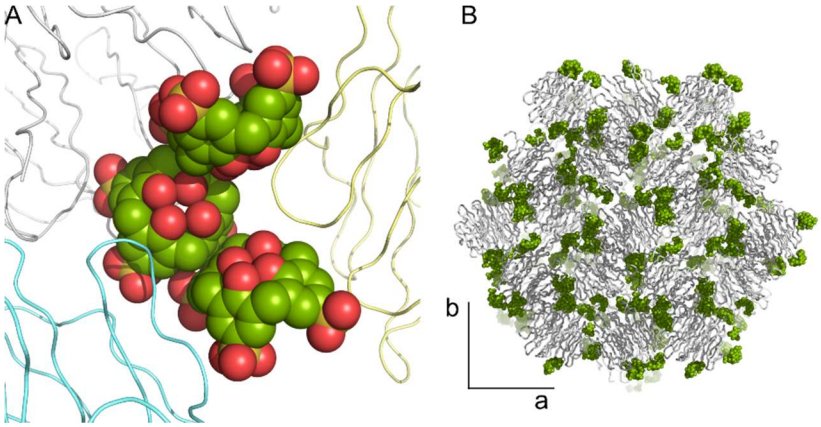


Figure 24. (A) The calixarene cluster mediates the interface between three RSL-LysMe$_2$ trimers coloured grey, cyan and yellow. (B) Crystal packing in the RSL-LysMe$_2$ – **sclx$_4$** complex. Protein and ligand are represented as ribbon and spheres respectively. Cell axis a and b are indicated.

Furthermore, every calixarene on the surface of RSL-LysMe$_2$ was interacting with more than one protein trimer, facilitating packing contacts. The resulting crystals were massive compared to the ones of ligand free RSL-LysMe$_2$ grown in control experiments. It is tempting to conclude that **sclx$_4$** acted as a molecular glue aiding the crystal growth.[33,75–77] The protein-calixarene interface, involving salt bridge, cation-π and CH-π bonds, resulted in interesting assembly in the solid state (Figure 24.B).

5.3 Solution- solid state comparison

At first sight, we were surprised by the **sclx$_4$** complexation mode. We expected complexation to occur at Lys34Me$_2$, as this was the most affected site during NMR titration. Instead, Lys83Me$_2$ and N-termini were complexed by the calixarene. To explain this behaviour, it is important to note that crystals were grown from a 1.5 mM RSL-LysMe$_2$ solution and a 10-fold excess of ligand. An excess of **sclx$_4$** resulted in a very large binding surface, as shown in the NMR titration. A graphical comparison of chemical shift perturbation and calixarene binding is provided in Figure 25.

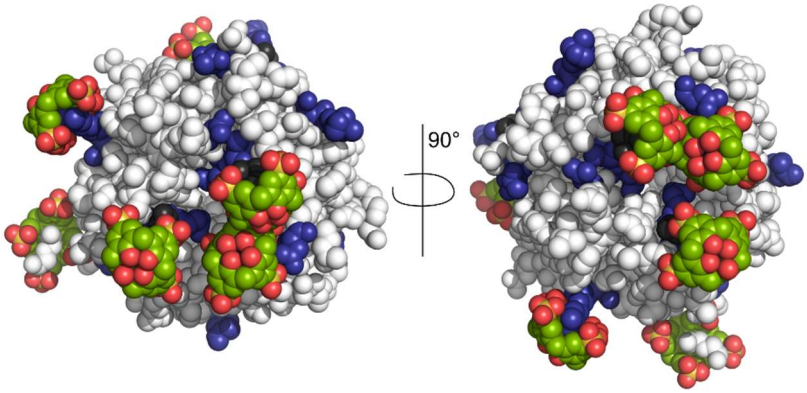


Figure 25. Scxl$_4$ binding and the chemical shift perturbation map are compared. Calixarenes are represented as green spheres. Residues showing chemical shift perturbation and non-detectable residues are blue and grey respectively. The two panels are related by a 90° rotation.

Notably, residues showing csp are involved in calixarene binding. Moreover, Ser1Me$_2$ and Ser2 are not detectable by NMR due to her high mobility. The closest probe to the N-terminus is Val3. Its resonance was shifted upon addition of high equivalents of **sclx$_4$** suggesting complexation

at this site. We can hypothesize that even at lower equivalents Ser1Me_2 was affected, but higher quantities of the ligand are needed to observe a shift of Val3 resonance. The selectivity for Lys83Me_2 over the more solvent exposed Lys34Me_2 could be justified considering the additional H-bond between **sclx$_4$** and the lysine's backbone NH. Still, it is not clear why Lys34Me2 was not complexed. Probably the residue's high flexibility impedes **sclx$_4$** binding and consequent protein-protein interface formation.

6. Cucurbit[7]uril

While **Q7**–$LysMe_n$ interactions have been characterized for the amino acids,[13] there are no reports of **Q*n***–$LysMe_n$ binding in a protein. To address this gap, we characterized the interactions of **Q7** with a protein that contains dimethyllysine ($LysMe_2$) and thereby provide a stepping stone to new applications of **Q7** in controlled protein interactions.

6.1 Q7 complexation in solution

Complex formation with **Q7** was investigated by ^{1}H-^{15}N HSQC-monitored titrations in 20 mM potassium phosphate, 50 mM NaCl, pH 6.0 (Figures 26). A 3.8 mM Q7 stock solution was titrated against RSL-$LysMe_2$ and resulted in large upfield perturbations of the $Lys34Me_2$ amide resonance (Figure 26).

The resonances of neighbouring residues Trp31 (indole, not shown), Asp32 and Tyr37 were affected also. No chemical shift changes occurred at the other possible binding sites ($Ser1Me_2$, $Lys25Me_2$ or $Lys83Me_2$). Control experiments with native RSL indicated that there was no binding to **Q7** under these conditions (Figure 27). An analysis of $\Delta\delta$ as a function of the ligand concentration yielded shallow binding curves that fit to a K_d of ~1 mM (Figure 26). It was not possible to achieve saturation in the NMR experiments given the limited solubility of **Q7**.

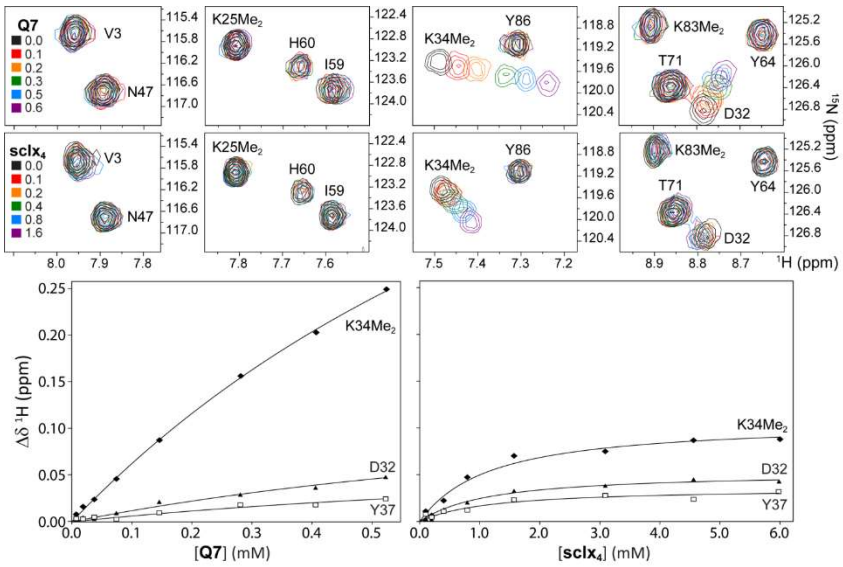


Figure 26. Upper panels. Spectral regions from overlaid ^{1}H-^{15}N HSQC spectra of RSL-LysMe$_2$ in the presence of 0-0.6 mM **Q7** or 0-1.6 mM **sclx$_4$**. Each panel provides data on at least one of the four potential binding sites. Resonance V3 is a reporter for Ser1Me$_2$ (the N-terminus). Significant chemical shift perturbations were observed for Lys34Me$_2$ and adjacent residues only. **Lower panels.** NMR-derived binding curves for complex formation between RSL-LysMe$_2$ and **Q7** or **sclx$_4$**.

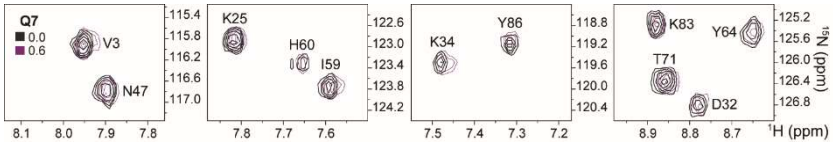


Figure 27. Spectral regions from the overlaid ^{1}H-^{15}N HSQC spectra of native RSL in the presence of 0.0 or 0.6 mM **Q7**.

Attempts to obtain thermodynamic information by ITC were thwarted by the lack of binding heats (Figure 28). Standard titrations of **Q7** into RSL-

$LysMe_2$ did not yield usable data. A reverse titration of 5.4 mM RSL-$LysMe_2$ into 0.1 mM **Q7** also gave negligible heat changes. These results appeared to be consistent with the low affinity determined by NMR.

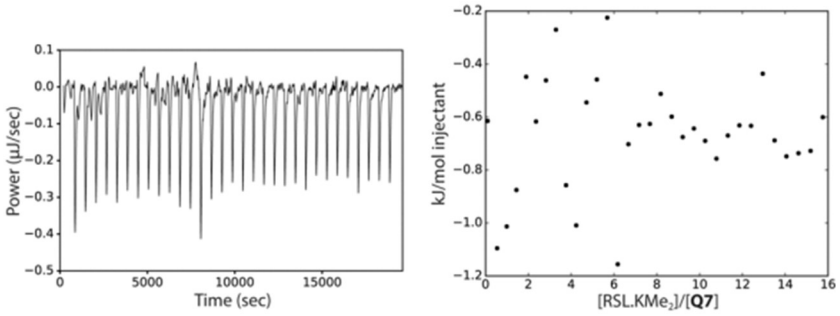


Figure 28. Representative ITC data showing baseline adjusted heats for injection of RSL-$LysMe_2$ (5.3 mM) into **Q7** (0.1 mM).

6.2 Q7-RSL-$LysMe_2$ structure determination

Further insights into **Q7**-$LysMe_2$ host-guest complexation were obtained by X-ray crystallography. Crystallization was achieved by using a sparse matrix screen (Jena JCSG++). Simple solutions containing ~20 % PEG and a buffer (over the pH range 5 - 8) were sufficient to prompt crystal growth. Cubic crystals were obtained when methyl-α-L-fucoside (MeFuc), a ligand to RSL, was included. Rod-shaped crystals also grew in the presence or absence of **Q7** (Figure 29).

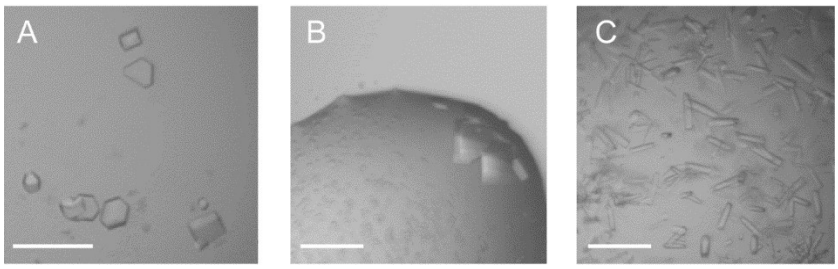


Figure 29. Representative crystals of RSL-LysMe$_2$ (A, B) in complex with **Q7** and (C) in the **Q7**-free form. The scale bar is 200 µm. A, B and C correspond to entries 1-3 in Table 1.

X-ray data collection was performed both in-house (Rigaku) and at the APS synchrotron (Argonne National Laboratory). The structures were solved by molecular replacement and the presence of **Q7** was clear in the electron density maps (Figure 30).

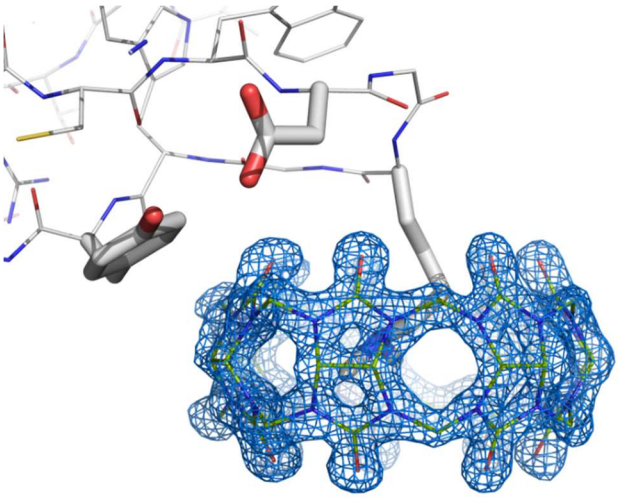


Figure 30. 2Fo - Fc electron density map (at 1.3 Å resolution and contoured at 1.5 σ) of **Q7** bound to Lys34Me$_2$ in $C222_1$. Binding site side chains Asp32, Lys34Me$_2$ and Tyr37 are represented as sticks.

The rod-shaped crystals proved to be devoid of **Q7**. Two crystal structures of RSL-LysMe$_2$ in the **Q7**-bound and one structure in the **Q7**-free state are reported (Table 1). The latter structure provided useful details on the binding site in the absence of **Q7**.

Table 1. Crystal structures of RSL.KMe$_2$ in Q7-bound and -free forms

Space Group	Resolution (Å)	Precipitant and Buffer	Notes
*C*222$_1$	1.3	20 % PEG 3350 0.2 M Na^+ malonate pH 7.0	Symmetric **Q7** binding MeFuc free
*F*432	2.4	25 % PEG 3350 0.1 M Bis-Tris pH 5.5	Asymmetric **Q7** binding MeFuc bound
*P*63	1.6	20 % PEG 3350 0.2 M K^+ formate pH 7.3	No **Q7** present MeFuc free

The **Q7**-bound structures crystallized in the *C*222$_1$ or *F*432 space groups, with an asymmetric unit that comprised one RSL-LysMe$_2$ trimer and three or two **Q7** ligands, respectively. Of the four potential binding sites, **Q7** complexation occurred exclusively at Lys34Me$_2$, consistent with the NMR observations. The high selectivity of **Q7** for Lys34Me$_2$ can be rationalized in terms of side chain accessibility.[36] The accessible surface area of the LysMe$_2$ side chains (calculated as an average from all chains in

the three crystal structures) was 160, 240 and 190 $Å^2$ for $Lys25Me_2$, $Lys34Me_2$ and $Lys83Me_2$, respectively. $Lys34Me_2$ was the most exposed side chain due to its location in a loop (residues 31-37), while $Lys25Me_2$ and $Lys83Me_2$ are in β-strands. Furthermore, $Lys34Me_2$ is flanked by Gly33 and Gly35 which confer steric accessibility and backbone mobility.

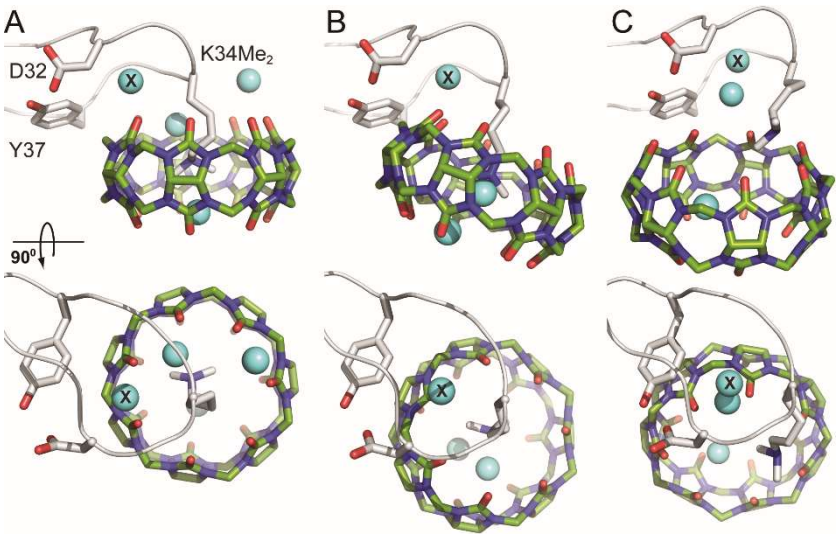


Figure 31. X-ray crystallography reveals complexation of **Q7** at $Lys34Me_2$ in RSL-$LysMe_2$. The binding mode of **Q7** was significantly different in space groups (A) $C222_1$ and (B) $F432$. (C) The distal interaction at $Lys34Me_2$ in chain B of $F432$. Only the loop residues 31-37 are shown and oriented identically in each structure. Side chains Asp32, $Lys34Me_2$ and Tyr37 are represented as sticks. Residues 33 and 35 are Gly. Cyan spheres are water molecules and X denotes the conserved water (See Figure 33 for hydrogen bonding pattern). The upper and lower panels are related by a 90° rotation.

Interestingly, three different modes of **Q7**-$Lys34Me_2$ binding were observed (Figure 31). In the $C222_1$ structure, ~220 Å2 of $Lys34Me_2$ was buried

in the cavity with both methyl substituents sitting in the central plane of **Q7**. In this orientation the Lys C^{γ}-C^{δ} bond was intersected by the plane of the rim carbonyl oxygens (Figure 31.A). While all three Lys34Me_2 sites in the RSL-LysMe_2 trimer were similar, the data at 1.3 Å resolution permitted model building with alternate conformations of Lys34Me_2 in chains B and C (Figure 32). These conformations suggest that the side chain retained some mobility inside **Q7** and that the cavity was incompletely filled by LysMe_2. This packing deficiency may contribute to the lower affinity of **Q7**-LysMe_2 with respect to **Q7**-LysMe_3.[13]

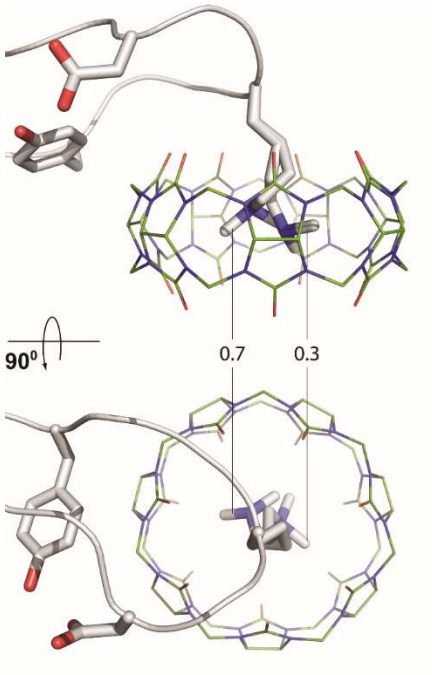


Figure 32. Alternate conformations of Lys34Me_2 inside **Q7** in $C222_1$ (chain B). The fractional occupancies are indicated. Water molecules were omitted for clarity.

The *F*432 structure grew from conditions that contained a lower **Q7**:protein ratio compared to $C222_1$ (Table 1) and ligand binding was asymmetric with respect to the RSL-LysMe$_2$ trimer. At chain A, **Q7** bound Lys34Me$_2$ (Figure 31.B) in a fashion similar to that in the $C222_1$ structure. However, less of the side chain (~190 Å^2) was buried and the plane of the **Q7** rim carbonyl oxygens intersected the Lys C^{δ}-C^{ε} bond. At chain B, Lys34Me$_2$ formed a distal interaction (~130 Å^2 buried surface) with **Q7** (Figure 31.C). This binding mode involved ion-dipole bonds between the ammonium group and two rim carbonyl oxygens ($N^{\zeta}\cdots O{=}C$ ~3.0 Å). The opposite carbonyl portal formed a similar distal interaction with Ser1Me$_2$ of a symmetry related molecule. At chain C, the Lys34Me$_2$ was devoid of ligand and the electron density for this side chain was poor, indicative of disorder.

The complete burial of the cationic dimethylammonium group in the hydrophobic cavity of **Q7** raises interesting questions regarding solvation.[36] Indeed, two water molecules were refined together with the Lys34Me$_2$ side chain in the **Q7** cavity (Figure 31). In the $C222_1$ structure, one water formed a hydrogen bond with the Lys34Me$_2$ ammonium group ($O^{W}\cdots N^{\zeta}$ ~2.7 Å), while simultaneously hydrogen bonded to two rim carbonyl oxygens ($O^{W}\cdots O{=}C$ <2.5 Å, Figure 31.A). The second water was ~3.4 Å from the closest methyl substituent of the dimethylammonium group, and hydrogen bonded to a lower rim carbonyl, suggesting that the positive charge is partially dissipated to the electronegative rim. Similar interactions were observed in the *F*432 structure although in this case the water bonded to N^{ζ} was fully buried and did not interact with the rim carbonyls (Figure 31.B). These observations are further intriguing since the cavity of **Q7** can accommodate eight water molecules, the release of which provides the driving force for guest binding.[23,24,79] This interpretation is based on studies of organic ammonium ions, with encapsulation of the hydrophobic portion

inside **Q7** and ion-dipole interactions between the ammonium ion(s) and the rim carbonyls. In the case of $LysMe_2$ encapsulation by **Q7**, six of the waters were displaced and two remaining waters solvated the tertiary ammonium ion inside the cavity (Figure 31).

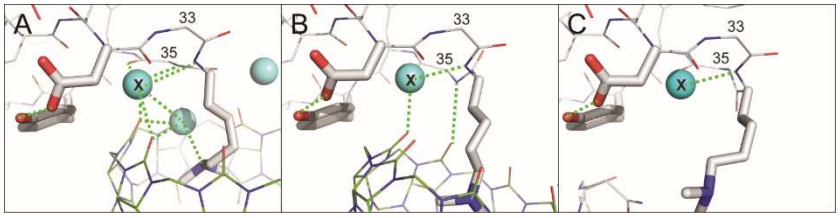


Figure 33. Putative hydrogen bonding network (green dotted lines) between water X, the protein backbone and **Q7** at $Lys34Me_2$ in **(A)** $C222_1$ and **(B)** $F432$. Panel **(C)** shows the equivalent site in the **Q7**-free structure. Note the altered conformation of the $Lys34Me_2$ - Gly35 amide bond in all three structures. The Asp32-Tyr37 hydrogen bond is shown also. The C^{α} of Gly33 and Gly35 are labelled.

Water also played a role in mediating **Q7** binding to the protein surface. A conserved water molecule (denoted X, Figures 31 and 33) was refined in each type of binding site and was present also in the **Q7**-free structure. In the $C222_1$ structure, water X was within hydrogen bond distance of three amide NH groups (Asp32, $Lys34Me_2$ and Gly35), the carbonyl of Gly35 and two **Q7** rim carbonyls (Figure 33.A). In the $F432$ structure, water X could hydrogen bond with three amides (Asp32, Gly33 and $Lys34Me_2$) and one **Q7** rim carbonyl (Figure 33.B). Interestingly, the loop 31-37 had a slightly different conformation at this site such that the amide NH of Gy35 was flipped out and pointed towards a rim carbonyl to form a unique protein-**Q7** hydrogen bond ($N^{\alpha}\cdots O{=}C$ ~3.45 Å). In the distal interaction mode (Figure

33.C) a second water molecule completed the hydrogen bond network between water X and **Q7**.

Despite the high molecular weight (~1.2 kDa) and large surface area (~980 Å2) of **Q7**, only 3 or 4 residues on the protein surface were involved directly in ligand complexation. In addition to Lys34Me_2, neighbouring residues Asp32, Tyr37 were the only side chains to form van der Waals contacts with **Q7**. Notably, these two side chains are linked by a hydrogen bond ($O^{\eta}\cdots O^{\delta}$ ~2.6 Å) and retained similar positions in all of the binding sites even though the loop, Lys34Me_2 and **Q7** adopted different conformations (Figures 32 and 34). Apparently, the interaction of **Q7** and the phenol of Tyr37 were less important than other packing interactions.

Striking examples of protein architectures were observed in the crystal packing. The *C*222$_1$ and *F*432 structures involved trimeric and tetrameric clusters of **Q7**, respectively, suggesting the potential to use the **Q7**-LysMe_2 complex as a pivot point for protein assembly (Figures 34 and 35).

Qn clusters are well-established in the literature, involving at least one CH···O=C bond between pairs of macrocycles,[80–82] and a supramolecular triangle was reported recently for **Q8**.[81] In *C*222$_1$ (Figure 35.A), sheets of trigonally-arrayed RSL-LysMe_2 trimers are arranged around **Q7** trimers (Figure 34.A). In addition to six CH···O=C interactions, each **Q7** acted as a bidentate ligand to a central sodium ion. The cation was complexed by ion-dipole bonds from two of the rim carbonyls ($Na^{+}\cdots$O=C ~2.5 Å) resulting in an octahedral coordination geometry. Salts of alkali metals can increase **Qn** solubility through coordination of the rim carbonyls.[23] It is tempting to conclude that sodium was critical to the growth of the *C*222$_1$ crystal considering that **Q7**-free crystals grew from similar conditions in which sodium malonate was replaced by potassium formate (Table 1).

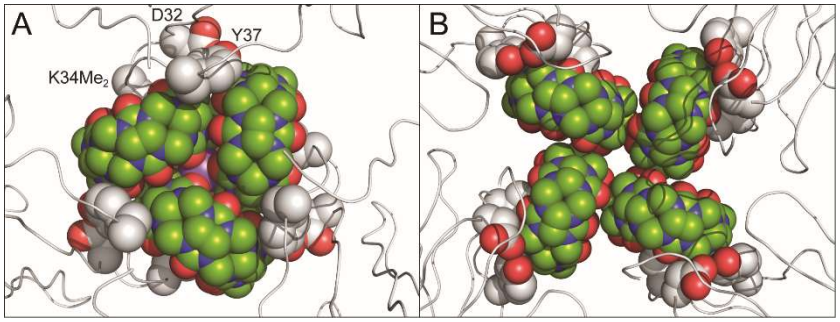


Figure 34. Trimeric and tetrameric cucurbituril clusters in space groups **(A)** *C*222$_1$ and **(B)** *F*432 in which **Q7-Q7** packing buried 150 or 120 Å^2 of ligand surface, respectively. In *C*222$_1$ the trimeric **Q7** formed an octahedral complex with a sodium cation (central, purple sphere). Side chains Asp32, Lys34Me$_2$ and Tyr37 are represented as spheres.

*F*432 is a rare space group characteristic of protein cages such as ferritin.[83] Remarkably, **Q7** at chain A mediated four-fold symmetric junctions to form a porous assembly of interlocked cages with 24 RSL-LysMe$_2$ trimers disposed at the vertices of a regular octahedron (Figures 34.B and 35.B). This cage-like assembly has an internal diameter of ~6 nm, comparable to that of ferritin.[83] This observation suggests the possibility of using **Q*n*** to generate macromolecular containers.[81,82,84]

In addition to the lower **Q7**:protein ratio, the crystallization pH was also lower (relative to *C*222$_1$, Table 1) and may have favoured cage formation as the net charge on the protein switched from anionic to cationic (RSL pI ~6.8). Interestingly, the **Q7**-**Q7** interfaces (Figure 34) buried ~300 and ~240 Å^2 of surface (per **Q7**) in the *C*222$_1$ and *F*432 structures, while **Q7**-protein contacts buried ~240 Å^2 of the ligand surface. It can be assumed that the relatively low water solubility of **Q7** is conducive to the formation of **Q7**-mediated protein architectures. Furthermore, the **Q7** clusters at protein-

protein interfaces may be an extension on the theme of macrocyclic molecular glues for protein assembly and crystallization.[52,53,58,85]

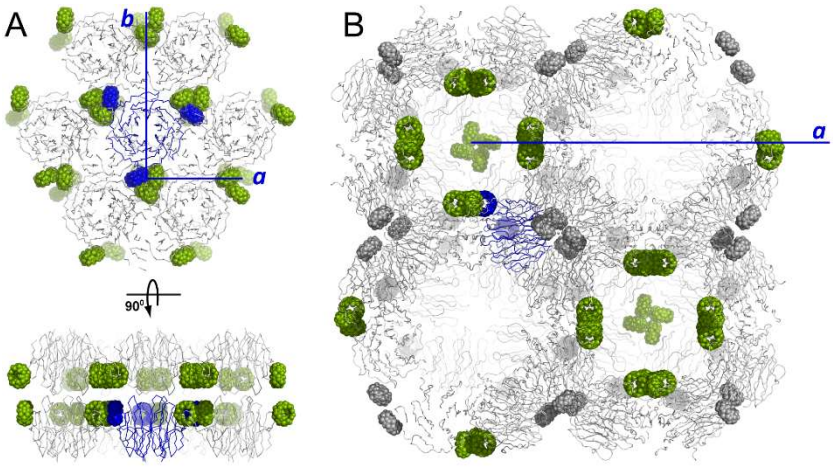


Figure 35. Crystal packing in space groups **(A)** $C222_1$ and **(B)** $F432$. Note the sheet assembly in $C222_1$ with **Q7** trimers mediating the packing. In $F432$ two interlocked cage assemblies are mediated *via* **Q7** tetramers (green). The distal binding **Q7** is grey. Protein and ligand are shown as ribbons and spheres, respectively. The asymmetric unit and the unit cell axes are indicated in blue.

7. Conclusion and future perspective

Supramolecular building blocks are increasingly popular as receptors for protein binding and assembly.[2,33,34,36–38,40,50–55,58,85] Examples include protein oligomerization mediated by calixarenes[51–53] and foldamers.[54]

Here, we demonstrated binding of a different class of receptors, tetra-phosphonate cavitands, to a dimethylated model protein in solution. Binding constants were in the millimolar range, suggesting weak binding. The neat charge on the receptor didn't affect binding to the most solvent exposed dimethyl-lysine. The difficulties in obtaining crystals thwarted further structural characterization.

Anionic calixarenes have proven particularly useful for the assembly of cationic proteins.[51–53] Here, I have discussed recognition and assembly of a neutral protein (RSL) mediated by **sclx$_4$**. Calixarene binding appeared asymmetric in the solid state. A wide binding area was identified, corroborated by solution state NMR experiments. The calixarene acted as a molecular glue and mediated protein-protein interfaces, suggesting a strategy for impeding binding sites. Recognition of methylated N-termini led to a calixarene cluster introducing a new major surface at this site. The cluster provided a "sticky patch" for protein crystallization, aiding crystal growth.

The binding of **Q*n*** to N-terminal aromatic residues is an established route to controlled protein interactions.[40,50,55,86,87] Programmable assembly of dimers[55] and polymers[87] can be achieved via the combination of **Q8** with proteins that bear an N-terminal phenylalanine. This repertoire has been expanded now to include **Q7**-LysMe$_2$ complexation and assembly. Contrary to amino acid studies[13] we observed a modest affinity for the **Q7**-LysMe$_2$

interaction in the model protein RSL-LysMe$_2$. Nevertheless, the affinity was sufficient to effect protein assembly in the solid state, with **Q7**-**Q7** packing playing pivotal role. The poor water solubility of **Q7** is apparently advantageous in this regard. The pronounced selectivity of **Q7** for the most exposed LysMe$_2$ points to a simple strategy of engineered protein assemblies based on **Q7**-LysMe$_2$ complexation. Finally, the tighter **Q7**-LysMe$_3$ interaction[13] suggests that this motif can be employed to greater advantage than **Q7**-LysMe$_2$ in protein assembly.

8. Acknowledgments

Prof. Peter Crowley is thanked for support and advice, Dr. Martin Rennie and Dr. Pawel Antonik are thanked for assistance in structure solution and NMR data collection respectively. We acknowledge Dr. Amir Khan for X-ray data collection. B. Harhen and R. Doohan are thanked for assistance with mass spectrometry and NMR spectroscopy, respectively.

9. Experimental

Materials: **Tiiii-1** was prepared as reported previously.[17] **sclx$_4$** was purchased from TCI. **Q7** was purchased from Sigma.

Tiiii-2: To a solution of tetrachloride footed cavitand Tiiii [C_3H_6Cl, CH_3, Et] (IV in ref 16) (150 mg, 0.140 mmol) in anhydrous DMF, sulphur trioxide trimethylamine complex (234 mg, 0.168 mmol) was added. The solution was stirred at 50 °C for 6 h. The solvent was evaporated under reduced pressure and the residue was dissolved in MeOH. The solution was kept at 0 °C for 1 h, after which precipitation occurred. The precipitate was filtered and dried in vacuum, affording **Tiiii-2** as a white solid (quantitative yield).

^{1}H NMR (400 MHz, D_2O): δ (ppm) 7.72 (s, 4H, Ar**H**); 4.70 (t, J = 7 Hz, 4H, ArC**H**); 4.21 (t, J = 6 Hz, 8H, $CHCH_2CH_2C\mathbf{H_2}OSO_3^-$); 3.76 (s, $NH(C\mathbf{H_3})_3^+$); 3.24 (s, $NH(C\mathbf{H_3})_3^+ \cdot SO_3^-$); 2.67-2.63 (m, 8H, $C\mathbf{H_2}CH_2CH_2OSO_3^-$); 2.46 (dd, ^3J = 8 $^2J_{P\text{-}H}$ = 19 Hz, 8H, $P(O)C\mathbf{H_2}CH_3$); 2.13 (s, 12H, $ArC\mathbf{H_3}$); 1.80-1.74 (m, 8H, $CH_2C\mathbf{H_2}CH_2OSO_3^-$); 1.47 (dt, ^3J = 8, $^3J_{P\text{-}H}$ = 22 Hz, 12H, $P(O)CH_2C\mathbf{H_3}$).

^{31}P NMR (162 MHz, D_2O): δ (ppm) 29.0 (s, P=O). **ESI-MS:** m/z = 347.2 [M - $4NHMe_3^+$] $^{4-}$; 482.9 [M - $3NHMe_3$]$^{3-}$.

Protein production: Unlabelled and ^{15}N-labelled RSL were produced and purified as described previously.[67,68] Dimethylation of the amino groups was achieved by using formaldehyde and dimethylaminoborane complex according to published methods.[70,71] Mass analysis (Agilent 6460 Triple Quadrupole LC/MS) of RSL (9726 Da) and the dimethylated protein (9838 Da) verified complete dimethylation of all 3 lysines and the N-terminus.

Stock solutions: Stock solutions of the ligands were prepared in water at pH 6.0 at a concentration of 40 mM, 5 mM and 95 mM for **Tiiii-1**, **Tiiii-2** and **sclx$_4$**

respectively, while the cucurbituril concentration was determined by UV/vis titration with cobaltocenium to be 3.8 mM.[88]

NMR spectroscopy: ^{15}N-labelled RSL-LysMe$_2$ samples were prepared to a concentration of 0.1 mM (or 0.05 mM in the case of **Tiiii-2** titration) in 20 mM potassium phosphate, 50 mM sodium chloride, 1.2 mM α-methyl-L-fucoside, 10 % D_2O at pH 6.0. ^{1}H-^{15}N HSQC spectra of ^{15}N-labelled RSL-LysMe$_2$ were acquired at 30 °C on a 600 MHz Varian NMR spectrometer with a HCN cold probe. Titrations were performed by the addition of µL aliquots of ligands stock solutions. pH was confirmed to be constant before and after each titration step. The spectra were acquired with 8 scans (or 16 scans in the case of Tiiii-2 titration) and 64 increments, processed in NMR Pipe and analysed in CCPN.

NMR-derived binding curves: Binding isotherms were obtained by plotting the magnitude of the chemical shift change (Δδ) as a function of the ligand concentration. The data for resonances Asp32, Lys34Me$_2$, and Tyr37 were fit globally to a 1:1 binding model (Equation below),

$$\Delta\delta = \Delta\delta_{\max}\frac{[A_T] + [B_T] + K_d - \sqrt{([A_T] + [B_T] + K_d)^2 - 4[A_T][B_T]}}{2[B_T]}$$

where [A_T] and [B_T] are the ligand and protein concentration, respectively. Parameter errors were estimated using Markov Chain Monte Carlo methods as implemented in PyMC.[89]

Isothermal Titration Calorimetry: RSL-LysMe$_2$ was dialysed into 20 mM potassium phosphate, 50 mM NaCl, pH 6.0 and concentrated to 5.4 mM. A **Q7** stock was diluted ~40-fold to 0.1 mM with the dialysis buffer. Samples were degassed prior to the titration. Measurements were made at 25° C using a Standard Volume NanoITC with a Hastelloy cell (TA Instruments).

RSL-LysMe$_2$ was injected into **Q7** in ~8 µL aliquots with a delay of 600 seconds between injections.

Co-crystallization of dimethylated RSL and sclx$_4$: Crystallization experiments were performed at 20 °C by using a sparse matrix screen (JCSG++, Jena Bioscience) in sitting drop vapour diffusion format, prepared with an Oryx 8 robot (Douglas Instruments). Crystals were obtained from a mixture of 1.5 mM sugar-free dimethylated RSL and 15 mM **sclx$_4$** under a variety of conditions. The crystals used for X-ray diffraction were grown in 0.8 M ammonium sulfate, 0.1 M sodium citrate, pH 4.0 (JCSG++ condition B1).

X-ray diffraction and structure determination: Crystals were transferred to the reservoir solution supplemented with 20 % glycerol as cryoprotectant. X-ray diffraction data were collected on a Rigaku MicroMax-007 generator with R-axis IV++ detector. Diffraction data were integrated and scaled using iMosflm.[90] The structure was solved by molecular replacement in Phaser[91] using PDB 2bt9 as a search model. The presence of **sclx$_4$** was clear in the unbiased electron density map. Iterative model building and refinement in COOT[92] and REFMAC[93] were performed until no further improvements in R_{free} or electron density were obtained.

Table S1: X-ray data collection, processing and refinement statistics of dimethylated RSL bound to **sclx$_4$**[a]

Data Collection	
Light source	Rigaku
Wavelength, Å	1.5418
Space group	$P2_12_12_1$
Resolution, Å	45.94-1.60 (1.63-1.60)
Unique reflections	42,850 (1,899)
Multiplicity	6.8 (5.2)
I/σ	24.1 (3.4)
Completeness, %	99.5 (90.2)
R_{merge},[b] %	5.1 (39.8)
Solvent content, %	55
Refinement	
R_{factor}, %	16.0
R_{free}, %	18.1
rmsd[c] bonds, Å	0.0085
rmsd angles, °	1.42
# molecules in a.u.	
Protein monomers	3
sclx$_4$	6
glycerol	6
PEG fragment	6
sulfate	1
water	382
Mean *B*-factor, Å^2	17.1
Ramachandran analysis[d]	
% residues in favoured regions	96.3
% residues in allowed regions	100
PDB code	6GL5

[a]Values in parentheses correspond to the highest resolution shell

[b]$R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl) \rangle / \Sigma_{hkl} \Sigma_i I_i(hkl)$

[c]root mean square deviation calculated for non-hydrogen protein atoms

[d]Determined in Molprobity

Co-crystallization of dimethylated RSL and Q7: Crystallization experiments were performed at 20 °C. A sparse matrix screen (JCSG++, Jena Bioscience) was performed in sitting drop vapour diffusion format (3-drop MRC Crystallization Plate) with an Oryx 8 robot (Douglas Instruments). A mixture of 1.5 mM RSL-LysMe_2, 1.1 mM **Q7** and 18 mM α-methyl-L-fucoside yielded crystals in 25 % PEG 3350, 0.1 M BIS-TRIS, pH 5.5 (JCSG++ condition H3). Crystals were obtained also from a mixture of 1.0 mM RSL-LysMe_2 and 2.3 mM **Q7** by using the hanging drop vapour diffusion method, in 20 % PEG 3350 and 0.2 M sodium malonate pH 7.0 (JCSG++ condition G6). Crystals devoid of **Q7** were obtained in 20 % PEG 3350 and 0.2 M potassium formate JCSG++ condition A10).

X-ray diffraction and structure determination: Crystals were transferred to their respective reservoir solutions supplemented with 2 % additional PEG and 25 % glycerol as cryoprotectants and cryo-cooled in liquid N_2. X-ray diffraction data were collected on the NE-CAT 24-ID-C beamline at the Advanced Photon Source (APS), Argonne, Illinois, USA and in-house on a Rigaku MicroMax-007 generator with R-axis IV++ detector. Diffraction data were integrated and scaled using XDS[94] (APS data) or iMosflm[90] (Rigaku data). The structures were solved by molecular replacement in Phaser[91] using PDB 2bt9 as a search model. **Q7** was clear in the unbiased electron density map. Grade Web Server (http://grade.globalphasing.org) was used to generate the **Q7** coordinates and refinement restraints. Iterative model building and refinement in COOT[92] and REFMAC[93] was performed until no further improvements in R_{free} or electron density were obtained.

Table S2: X-ray data collection, processing and refinement statistics for RSL-LysMe$_2$ in **Q7**-bound (space groups *C*222$_1$ and *F*432) and **Q7**-free (*P*6$_3$) forms[a]

Data Collection			
Light source	APS	Rigaku	APS
Beam line	24-ID-C	-	24-ID-C
Wavelength, Å	0.9791	1.5418	0.9791
Space group	***C*222$_1$**	***F*432**	***P*6$_3$**
Unit cell	a = 50.3 Å		a = b = 122.3 Å
	b = 87.2 Å	a = b = c = 200.6 Å	c = 36.4 Å
	c = 146.6 Å	α = β = γ = 90°	α = β = 90°
	α = β = γ = 90°		γ = 120°
Resolution, Å	43.60-1.28	70.91-2.42	40.03-1.60
	(1.30-1.28)	(2.51-2.42)	(1.62-1.60)
Unique reflections	78,616 (3,749)	13,677 (1,416)	41,544 (1,938)
Multiplicity	4.5 (4.0)	8.9 (8.6)	6.6 (6.2)
I/σ	18.8 (1.9)	13.2 (4.0)	12.2 (1.2)
Completeness, %	95.0 (92.4)	99.8 (100.0)	99.6 (95.9)
$R_{meas,}$[b] %	5.0 (81.3)	17.5 (57.2)	12.4 (167.8)
Half-set correlation ($CC_{1/2}$)	0.999 (0.706)	0.991 (0.885)	0.997 (0.417)
Solvent content, %	54.9	56.8	39.15

Refinement			
R_{factor}, %	12.4	17.6	16.9
R_{free}, %	15.1	23.0	19.9
rmsd[c] bonds, Å	0.009	0.010	0.007
rmsd angles, °	1.33	1.40	1.27
# molecules in a.u.			
Protein monomers	3	3	4
Q7	3	2	-
Glycerol	8	1	9
methyl-α-L-fucoside	-	6	-
Sodium	1	-	-
Water	382	243	314
Mean B-factor (Å^2)	18.5	15.8	21.5
Ramachandran analysis,[d] % residues			
in favoured regions	98.3	98.3	98.1
in allowed regions	100.0	99.6	100
PDB id	6F7W	6F7X	6F7Y

[a]Values in parentheses correspond to the highest resolution shell

[b]$R_{meas} = \sqrt{(n/n-1)}\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$

[c]root mean square deviation calculated for non-hydrogen protein atoms

[d]Determined in Molprobity

10. Bibliography

1 F. Guagnini, P. M. Antonik, M. L. Rennie, P. O'Byrne, A. R. Khan, R. Pinalli, E. Dalcanale and P. B. Crowley, *Angew. Chemie Int. Ed.*, 2018, **57**, 7126–7130.

2 S. Van Dun, C. Ottmann, L. G. Milroy and L. Brunsveld, *J. Am. Chem. Soc.*, 2017, **139**, 13960–13968.

3 L. J. Sweetlove and A. R. Fernie, *Nat. Commun.*, 2018, **9**, 2136.

4 T. Lei, P. Zhang, X. Zhang, X. Xiao, J. Zhang, T. Qiu, Q. Dai, Y. Zhang, L. Min, Q. Li, R. Yin, P. Ding, N. Li, Y. Qu, D. Mu, J. Qin, X. Zhu, Z.-X. Xiao and Q. Li, *Nat. Commun.*, 2018, **9**, 1876.

5 J. M. Dunce, O. M. Dunne, M. Ratcliff, C. Millán, S. Madgwick, I. Usón and O. R. Davies, *Nat. Struct. Mol. Biol.*, 2018, **1**.

6 H. Jeong, S. P. Mason, A.-L. Barabási and Z. N. Oltvai, *Nature*, 2001, **411**, 41–42.

7 J.-F. Rual, K. Venkatesan, T. Hao, T. Hirozane-Kishikawa, A. Dricot, N. Li, G. F. Berriz, F. D. Gibbons, M. Dreze, N. Ayivi-Guedehoussou, N. Klitgord, C. Simon, M. Boxem, S. Milstein, J. Rosenberg, D. S. Goldberg, L. V. Zhang, S. L. Wong, G. Franklin, S. Li, J. S. Albala, J. Lim, C. Fraughton, E. Llamosas, S. Cevik, C. Bex, P. Lamesch, R. S. Sikorski, J. Vandenhaute, H. Y. Zoghbi, A. Smolyar, S. Bosak, R. Sequerra, L. Doucette-Stamm, M. E. Cusick, D. E. Hill, F. P. Roth and M. Vidal, *Nature*, 2005, **437**, 1173–1178.

8 J.-M. Lehn, *Science*, 2002, **295**, 2400–3.

9 J. P. Chinta, A. Acharya, A. Kumar and C. P. Rao, *J. Phys. Chem. B*, 2009, **113**, 12075–12083.

10 R. K. Pathak, J. Dessingou and C. P. Rao, *Anal. Chem.*, 2012, **84**, 8294–8300.

11 M. Selkti, A. W. Coleman, I. Nicolis, N. Douteau-guével, F. Villain, A. Tomas, C. De Rango, C. Ep, F. De Pharmacie, U. P. V, P. Cedex, C. Upr, P. Vercors, L. Cedex, L. De Physique, C. Pharmaceutique, U. Paris-sud and C. Cedex, *Chem. Comm.*, 2000, **2**, 161–162.

12 A. Lazar, D. Silva, A. Navaza, A. W. Coleman, C. Umr, *Chem. Comm.*, 2004, **2**, 2162–2163.

13 M. A. Gamal-Eldin and D. H. Macartney, *Org. Biomol. Chem.*, 2013, **11**, 488–495.

14 C. S. Beshara, C. E. Jones, K. D. Daze, B. J. Lilgert and F. Hof, *ChemBioChem*, 2010, **11**, 63–66.

15 E. Biavardi, C. Tudisco, F. Maffei, A. Motta, C. Massera, G. G. Condorelli and E. Dalcanale, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 2263–8.

16 H. Peacock, C. C. Thinnes, A. Kawamura and A. D. Hamilton, *Supramol. Chem.*, 2016, **28**, 575–581.

17 R. Pinalli, G. Brancatelli, A. Pedrini, D. Menozzi, D. Hernández, P. Ballester, S. Geremia and E. Dalcanale, *J. Am. Chem. Soc.*, 2016, **138**, 8569–8580.

18 N. Bontempi, E. Biavardi, D. Bordiga, G. Candiani, I. Alessandri, P. Bergese and E. Dalcanale, *Nanoscale*, 2017, **9**, 8639–8646.

19 E. Kovalenko, M. Vilaseca, M. Díaz-Lobo, A. N. Masliy, C. Vicent and V. P. Fedin, *J. Am. Soc. Mass Spectrom.*, 2016, **27**, 265–276.

20 T. Minami, N. A. Esipenko, B. Zhang, L. Isaacs and P. Anzenbacher, *Chem. Commun.*, 2014, **50**, 61–63.

21 J. W. Lee, M. H. Shin, W. Mobley, A. R. Urbach and H. I. Kim, *J. Am. Chem. Soc.*, 2015, **137**, 15322–15329.

22 J. W. Lee, S. Samal, N. Selvapalam, H. J. Kim and K. Kim, *Acc. Chem. Res.*, 2003, **36**, 621–630.

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

24 F. Biedermann, V. D. Uzunova, O. A. Scherman, W. M. Nau and A. De Simone, *J. Am. Chem. Soc.*, 2012, **134**, 15318–15323.

25 A. R. Urbach and V. Ramalingam, *Isr. J. Chem.*, 2011, **51**, 664–678.

26 L. M. Heitmann, A. B. Taylor, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.*, 2006, **128**, 12574–12581.

27 J.-M. Yi, Y.-Q. Zhang, H. Cong, S.-F. Xue and Z. Tao, *J. Mol. Struct.*, 2009, **933**, 112–117.

28 P. Rajgariah and A. R. Urbach, *J. Incl. Phenom. Macrocycl. Chem.*, 2008, **62**, 251–254.

29 F. G. Klärner and T. Schrader, *Acc. Chem. Res.*, 2013, **46**, 967–978.

30 S. Dutt, C. Wilch, T. Gersthagen, P. Talbiersky, K. Bravo-Rodriguez, M. Hanni, E. Sánchez-García, C. Ochsenfeld, F.-G. Klärner and T. Schrader, *J. Org. Chem.*, 2013, **78**, 6721–6734.

31 S. Dutt, C. Wilch, T. Gersthagen, C. Wölper, A. A. Sowislok, F.-G. Klärner and T. Schrader, *European J. Org. Chem.*, 2013, **2013**, 7705–7714.

32 P. S. Cremer, A. H. Flood, B. C. Gibb and D. L. Mobley, *Nat. Chem.*, 2017, **10**, 8–16.

33 R. E. Mcgovern, H. Fernandes, A. R. Khan, N. P. Power and P. B. Crowley, *Nat. Chem.*, 2012, **4**, 527-533.

34 M. Mallon, S. Dutt, T. Schrader and P. B. Crowley, *ChemBioChem*, 2016, **17**, 774–783.

35 S. A. Minaker, K. D. Daze, M. C. F. Ma and F. Hof, *J. Am. Chem. Soc.*, 2012, **134**, 11674–11680.

36 R. E. McGovern, B. D. Snarr, J. A. Lyons, J. McFarlane, A. L. Whiting, I. Paci, F. Hof and P. B. Crowley, *Chem. Sci.*, 2015, **6**, 442–449.

37 D. Bier, R. Rose, K. Bravo-Rodriguez, M. Bartel, J. M. Ramirez-Anguita, S. Dutt, C. Wilch, F. G. Klärner, E. Sanchez-Garcia, T. Schrader and C. Ottmann, *Nat. Chem.*, 2013, **5**, 234–239.

38 D. Bier, S. Mittal, K. Bravo-Rodriguez, A. Sowislok, X. Guillory, J. Briels, C. Heid, M. Bartel, B. Wettig, L. Brunsveld, E. Sanchez-Garcia, T. Schrader and C. Ottmann, *J. Am. Chem. Soc.*, 2017, **139**, 16256–16263.

39 W. Li, A. T. Bockus, B. Vinciguerra, L. Isaacs and A. R. Urbach, *Chem. Commun.*, 2016, **52**, 8537–8540.

40 J. M. Chinai, A. B. Taylor, L. M. Ryno, N. D. Hargreaves, C. A. Morris, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.*, 2011, **133**, 8810–8813.

41 E. Uemura, T. Niwa, S. Minami, K. Takemoto, S. Fukuchi, K. Machida, H. Imataka, T. Ueda, M. Ota and H. Taguchi, *Sci. Rep.*, 2018, **8**, 678.

42 Y. Zhu, W. L. Wang, D. Yu, Q. Ouyang, Y. Lu and Y. Mao, *Nat. Commun.*, 2018, **9**, 1360.

43 Q. Luo, C. Hou, Y. Bai, R. Wang and J. Liu, *Chem. Rev.*, 2016, **116**, 13571–13632.

44 E. Pavadai, B. S. Gerstman and P. P. Chapagain, *Sci. Rep.*, 2018, **8**, 9776.

45 G. Matsumoto, T. Inobe, T. Amano, K. Murai, N. Nukina and N. Mori, *Sci. Rep.*, 2018, **8**, 9585.

46 L. Palazzi, E. Bruzzone, G. Bisello, M. Leri, M. Stefani, M. Bucciantini and P. Polverino de Laureto, *Sci. Rep.*, 2018, **8**, 8337.

47 Y. T. Lai, E. Reading, G. L. Hura, K. L. Tsai, A. Laganowsky, F. J. Asturias, J. A. Tainer, C. V. Robinson and T. O. Yeates, *Nat. Chem.*, 2014, **6**, 1065–1071.

48 E. Sasaki, D. Böhringer, M. Van De Waterbeemd, M. Leibundgut, R. Zschoche, A. J. R. Heck, N. Ban and D. Hilvert, *Nat. Commun.*, 2017, **8**, 1–10.

49 J. B. Bale, S. Gonen, Y. Liu, W. Sheffler, D. Ellis, C. Thomas, D. Cascio, T. O. Yeates, T. Gonen, N. P. King and D. Baker, *Science*, 2016, **353**, 389–94.

50 P. J. de Vink, J. M. Briels, T. Schrader, L. G. Milroy, L. Brunsveld and C. Ottmann, *Angew. Chemie - Int. Ed.*, 2017, **56**, 8998–9002.

51 R. E. McGovern, A. A. McCarthy and P. B. Crowley, *Chem. Commun.*, 2014, **50**, 10412–10415.

52 M. L. Rennie, A. M. Doolan, C. L. Raston and P. B. Crowley, *Angew. Chemie - Int. Ed.*, 2017, **56**, 5517–5521.

53 M. Rennie, G. Fox, J. Pérez and P. B. Crowley, *Angew. Chemie Int. Ed.*, 2018, **57**, 13764–13769.

54 M. Jewginski, T. Granier, B. Langlois D'Estaintot, L. Fischer, C. D. Mackereth and I. Huc, *J. Am. Chem. Soc.*, 2017, **139**, 2928–2931.

55 H. D. Nguyen, D. T. Dang, J. L. J. Van Dongen and L. Brunsveld, *Angew. Chemie - Int. Ed.*, 2010, **49**, 895–898.

56 C. Hou, J. Li, L. Zhao, W. Zhang, Q. Luo, Z. Dong, J. Xu and J. Liu, *Angew. Chemie - Int. Ed.*, 2013, **52**, 5590–5593.

57 C. Si, J. Li, Q. Luo, C. Hou, T. Pan, H. Li and J. Liu, *Chem. Commun.*, 2016, **52**, 2924–2927.

58 J. M. Alex, M. L. Rennie, S. Volpi, F. Sansone, A. Casnati and P. B. Crowley, *Cryst. Growth Des.*, 2018, **18**, 2467–2473.

59 C. A. Musselman, M. E. Lalonde, J. Côté and T. G. Kutateladze, *Nat. Struct. Mol. Biol.*, 2012, **19**, 1218–1227.

60 M. Luo, *Chem. Rev.*, 2016, **116**, 13571-13632.

61 S. M. Carlson and O. Gozani, , DOI:10.1101/cshperspect.a026435.

62 J. Lee, L. Perez, Y. Liu, H. Wang, R. J. Hooley and W. Zhong, *Anal. Chem.*, 2018, **90**, 1881–1888.

63 J. Małecki, V. K. Aileni, A. Y. Y. Ho, J. Schwarz, A. Moen, V.

Sørensen, B. S. Nilges, M. E. Jakobsson, S. A. Leidel and P. Ø. Falnes, *Nucleic Acids Res.*, 2017, **45**, 4371.

64 X. Yi, X.-J. Jiang, X.-Y. Li and D.-S. Jiang, *Am. J. Transl. Res.*, 2015, **7**, 2159–75.

65 J. E. Beaver and M. L. Waters, *ACS Chem. Biol.*, 2016, **11**, 643–653.

66 I. N. Gober and M. L. Waters, *J. Am. Chem. Soc.*, 2016, **138**, 9452–9459.

67 N. Kostlánová, E. P. Mitchell, H. Lortat-Jacob, S. Oscarson, M. Lahmann, N. Gilboa-Garber, G. Chambat, M. Wimmerová and A. Imberty, *J. Biol. Chem.*, 2005, **280**, 27839–27849.

68 P. M. Antonik, A. N. Volkov, U. N. Broder, D. Lo Re, N. A. J. Van Nuland and P. B. Crowley, *Biochemistry*, 2016, **55**, 1195-1203.

69 P. M. Antonik, A. M. Eissa, A. R. Round, N. R. Cameron and P. B. Crowley, *Biomacromolecules*, 2016, **17**, 2719-2725.

70 W. R. Rypniewski, H. M. Holden and I. Rayment, *Biochemistry*, 1993, **32**, 9851–9858.

71 S. T. Larda, M. P. Bokoch, F. Evanics and R. S. Prosser, *J. Biomol. NMR*, 2012, **54**, 199-209.

72 K. L. Longenecker, S. M. Garrard, P. J. Sheffield and Z. S. Derewenda, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2001, **57**, 679–688.

73 P. J. Sheffield and Z. S. Derewenda, *Acta Cryst D*, 2001, **D57**, 679–688.

74 T. S. Walter, C. Meier, R. Assenberg, K.-F. Au, J. Ren, A. Verma, J. E. Nettleship, R. J. Owens, D. I. Stuart and J. M. Grimes, *Structure*, 2006, **14**, 1617–1622.

75 H. P. Morgan, W. Iain, P. A. M. Michels and A. Linda, *Acta Cryst. F*, 2010, **F66**, 215–218.

76 A. Mcpherson, C. Nguyen, R. Cudney and S. B. Larson, *Cryst. Growth Des.*, 2011, **11**, 1469–1474.

77 A. Bijelic and A. Rompel, *Acc. Chem. Res.*, 2017, **50**, 1441–1448.

78 I. Staneva and D. Frenkel, *J. Chem Phys.*, 2015, **143**, 194511.

79 Y. Wang, J. R. King, P. Wu, D. L. Pelzman, D. N. Beratan and E. J. Toone, *J. Am. Chem. Soc.*, 2013, **135**, 6084–6091.

80 S. Lim, H. Kim, N. Selvapalam, K. J. Kim, S. J. Cho, G. Seo and K. Kim, *Angew. Chemie - Int. Ed.*, 2008, **47**, 3352–3355.

81 D. Bardelang, K. Banaszak, H. Karoui, A. Rockenbauer, M. Waite, K. Udachin, J. A. Ripmeester, C. I. Ratcliffe, O. Ouari and P. Tordo, *J. Am. Chem. Soc.*, 2009, **131**, 5402–5404.

82 D. Bardelang, K. A. Udachin, D. M. Leek, J. C. Margeson, G. Chan, C. I. Ratcliffe and J. A. Ripmeester, *Cryst. Growth Des.*, 2011, **11**, 5598–5614.

83 B. Maity, S. Abe and T. Ueno, *Nat. Commun.*, 2017, **8**, 14820.

84 Y. M. Jeon, J. Kim, D. Whang and K. Kim, *J. Am. Chem. Soc.*, 1996, **118**, 9790–9791.

85 A. M. Doolan, M. L. Rennie and P. B. Crowley, *Chem. - A Eur. J.*, 2018,

24, 984–991.

86 J. Murray, J. Sim, K. Oh, G. Sung, A. Lee, A. Shrinidhi, A. Thirunarayanan, D. Shetty and K. Kim, *Angew. Chemie - Int. Ed.*, 2017, **56**, 2395–2398.

87 S. Yi and A. E. Kaifer, *J. Org. Chem.*, 2011, **76**, 10275–10278.

88 A. Patil, D. Huard and C. J. Fonnesbeck, *J. Stat. Softw.*, 2010, **35**, 4, 1-81.

89 T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell and A. G. W. Leslie, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 271–281.

90 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658–674.

91 P. Emsley and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2004, **60**, 2126–2132.

92 G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 355–367.

93 W. Kabsch and IUCr, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 125–132.

Appendix A: List of Abbreviations

THF	Tetrahydrofuran
EtOH	Ethanol
MeOH	Methanol
AcOEt	Ethyl Acetate
Hex	Hexane
DCM	Dichloromethane
HB / H-Bond	Hydrogen Bond
XB / X-Bond	Halogen Bond
LAG	Liquid Assisted Grinding
CP	Coordination Polymer
MOF	Meral Organic Framework
NMR	Nuclear Magnetic Resonance
TLC	Thin Layer Chromathography
GC	Gas Chromatography
PXRD	Powdre X-Ray Diffraction
XRD	X-Ray Diffraction
ITC	Isothermal Titration Calorimetry
Csp	Chemical Shift Perturbation

Appendix B: Materials and Methods for the Synthetic Parts

Materials

All commercial reagents were used as received. Unless otherwise stated, reactions were conducted in flame-dried glassware under an atmosphere of argon using anhydrous solvents (either freshly distilled or treated on molecular sieves). Silica column chromatography was performed using silica gel 60 (Fluka 230-400 mesh ASTM), or silica gel 60 (MERCK 70-230 mesh). Preparative TLC were performed on silica (Sigma) glass plates.

Products Characterizations

- **NMR measurements:** ^{1}H NMR spectra were obtained using a Bruker AVANCE 300 (300 MHz) or a Bruker AVANCE 400 (400 MHz) spectrometer at 25 °C. All chemical shifts (δ) were reported in ppm relative to the proton resonances resulting from the incomplete deuteration of the NMR solvents. ^{13}C NMR spectra were obtained using a Bruker AVANCE 300 (75 MHz) or a Bruker AVANCE 400 (100 MHz) spectrometer. All chemical shifts (δ) were reported in ppm relative to the carbon resonances of the NMR solvents. ^{31}P NMR spectra were obtained using a Bruker AVANCE 400 (162 MHz) spectrometer. All chemical shifts (δ) were recorded in ppm relative to external 85% H_3PO_4 at 0.00 ppm. ^{19}F NMR spectra were obtained using a Bruker AVANCE 400 (376 MHz) spectrometer. All chemical shifts (δ) were recorded in ppm relative to external CCl_3F at 0.00 ppm.

- **Electrospray Ionization Mass Spectrometry (ESI-MS):** MS Measurements Electrospray ionization ESI-MS experiments were

performed on a Waters ZMD spectrometer equipped with an electrospray interface.

- **MALDI-MS Measurements:** MALDI was performed on an AB SCIEX MALDI TOF-TOF 4800 Plus (matrix: α–cyano-4- hydroxycinnamic acid).

- **GC-MS Measurements:** GC spectra were recorded on a Agilent Technologies 6890N Network GC System.

- **PXRD Measurements:** PXRD spectra were recorded at room temperature on a X'TRA THERMO diffractometer with a Bragg-Brentano geometry, Cu Kα radiation (40kV per 30mA) and a solid state Si(Li) detector. Samples were prepared on a zero-background sample holder. Spectra were collected in a 2θ range 3-40° with 0.05° steps and 2s collection time.

Ringraziamenti
Acknowledgments

Questa pagina, che non leggerà mai nessuno, è dedicata ai ringraziare tutte le persone che lavorativamente e non hanno contribuito a questo piccolo successo.

This page, that no one will ever read, is dedicated to thanking every person, that somehow contributed to this small victory.

Ringrazio tantissimo Chiara, che con la sua pazienza e gentilezza (e occasionalmente con l'offerta di cioccolato), mi ha guidata durante questi 3 anni. Ringrazio tantissimo anche Enrico, per tutte le possibilità che mi ha dato. Grazie a entrambi per essere stati i miei tutor e per i preziosi insegnamenti e consigli.

I thank Chiara that, with her patience and kindness (and occasionally with the offer of chocolate), has been mentoring me during these 3 years. I thank Enrico for all the possibilities he gave me. I would like to acknowledge both of them for being my tutors and for their precious advice.

Vorrei ringraziare i miei compagni del Dalcanale group per essere stati un grande sostegno durante il dottorato. In particolare, grazie a Jonathan per aver invaso la mia vita e per tutto il divertimento che ne è derivato, a Anjali per ogni avventura vissuta insieme, a Andreas per non aver mai fatto mancare storie interessanti, e ad Arek per aver condiviso la nostra follia. Grazie per essere stati dei colleghi ma anche degli amici, per avermi portata lontano e avermi sempre riportata a casa. Grazie a Alessandro, Martina e Ilaria per non avermi mai fatto mancare i loro consigli e supporto.

I would like to thank my lab-mates from the Dalcanale group for the constant support during these years. In particular, thanks to Jonathan for

invading my life and for all the fun he provided, thanks to Anjali for all the adventures together, thanks to Andreas for always providing interesting stories, and thanks to Arek for sharing our craziness. Thanks for being my colleagues but also my friends, for taking me far away and always bringing me back home. Thanks to Alessandro, Martina and Ilaria for their constant support and advice.

Grazie a Peter per tutto quello che mi ha insegnato. Un ringraziamento a tutto il Crowley group, Martin, Kiefer, Ciara, Jimi, Gosia e Srinu per essere stati i migliori colleghi che potessi desiderare. Non voglio scrivere nessun ringraziamento per Pawel qui, perchè lo ringrazio già abbastanza di persona.

I would like to acknowledge Peter for everything he has taught me. Thanks to all the Crowley group, Martin, Kiefer, Ciara, Jimi, Gosia and Srinu for being the best colleagues I could ever wish for. I don't want to write anything for Pawel here, because I thank him enough every day.

Vorrei davvero ringraziare i miei studenti, Silvia, Alice, Tommaso, Letizia e Danilo, per tutto quello che ho imparato lavorando con loro e per la motivazione che mi hanno dato.

I would like to thank the students I supervised, Silvia, Alice, Tommaso, Letizia and Danilo, for everything I've learnt working with them and all the motivation they gave me.

Fuori dal laboratorio, ringrazio tantissimo la mia famiglia, senza la quale non avrei combinato un granchè nella vita.

A tutti i miei amici, a Parma e a casa, grazie per avermi sopportata durante tutti questi anni e per esserci sempre. Un super-grazie a Claudia, Margherita e Anna, senza le quali sarei molto persa.

Outside of the lab, I would like to really thank my family, without which I wouldn't have achieved much in life. To all my friend, in Parma and at home, thanks for standing by my side during these years. Super-thanks to Claudia, Margherita and Anna, because I would be lost without them.