The Origin of Selectivity in the Complexation of N-Methyl Amino Acids by Tetraphosphonate Cavitands

Article

The origin of selectivity in the complexation of N-methyl amino acids by tetraphosphonate cavitands

Roberta Pinalli, Giovanna Brancatelli, Alessandro Pedrini, Daniela Menozzi, Daniel Hernández, Pablo Ballester, Silvano Geremia, and Enrico Dalcanale

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/jacs.6b04372 • Publication Date (Web): 16 Jun 2016

The origin of selectivity in the complexation of *N*-methyl amino acids by tetraphosphonate cavitands

Roberta Pinalli,[a#] Giovanna Brancatelli,[b#] Alessandro Pedrini,[a] Daniela Menozzi,[a] Daniel Hernández,[c] Pablo Ballester,[c,d] Silvano Geremia*[b] and Enrico Dalcanale*[a]

[a] Department of Chemistry, University of Parma, and INSTM, UdR Parma, Parco Area delle Scienze 17/A, 43124, Parma, Italy

[b] CEB Centre of Excellence in Biocrystallography, Department of Chemical and Pharmaceutical Sciences, University of Trieste, Via L. Giorgieri 1, 34127, Trieste, Italy

[c] Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluís Companys, 23, 08018 Barcelona, Spain

[d] Institute of Chemical Research of Catalonia (ICIQ), The Barcelona Institute of Science and Technology, Avgda. Països Catalans 16, 43007 Tarragona, Spain

ABSTRACT: We report on the eligibility of tetraphosphonate resorcinarene cavitands for the molecular recognition of amino acids. We determined the crystal structure of thirteen complexes of the tetraphosphonate cavitand **Tiiii[H, CH_3, CH_3]** with amino acids. ^1H and ^{31}P NMR experiments and ITC analysis were performed to probe the binding between cavitand **Tiiii[C_3H_7, CH_3, C_2H_5]** or the water-soluble counterpart **Tiiii[$C_3H_6Py^+Cl^-$, CH_3, C_2H_5]** and a selection of representative amino acids. The reported studies and results allowed: (i) to highlight the non-covalent interactions involved in the binding event in each case; (ii) to investigate the ability of tetraphosphonate cavitand receptors to discriminate between the different amino acids; (iii) to calculate the K_a values of the different complexes formed and evaluate the thermodynamic parameters of the complexation process, dissecting the entropic and enthalpic contributions; (iv) to determine the solvent influence on the complexation selectivity. By moving from methanol to water, the complexation changed from entropy driven to entropy opposed, leading to a drop of almost three orders in the magnitude of the K_a. However, this reduction in binding affinity is associated with a dramatic increase in selectivity, since in aqueous solutions only *N*-methylated amino acids are effectively recognized. The thermodynamic profile of the binding does not change in PBS solution. The pivotal role played by cation–π interactions is demonstrated by the linear correlation found between the $\log K_a$ in methanol solution and the depth of $^+N–CH_3$ cavity inclusion in the molecular structures. These findings are relevant for the potential use of phosphonate cavitands as synthetic receptors for the detection of epigenetic modifications of histones in physiological media.

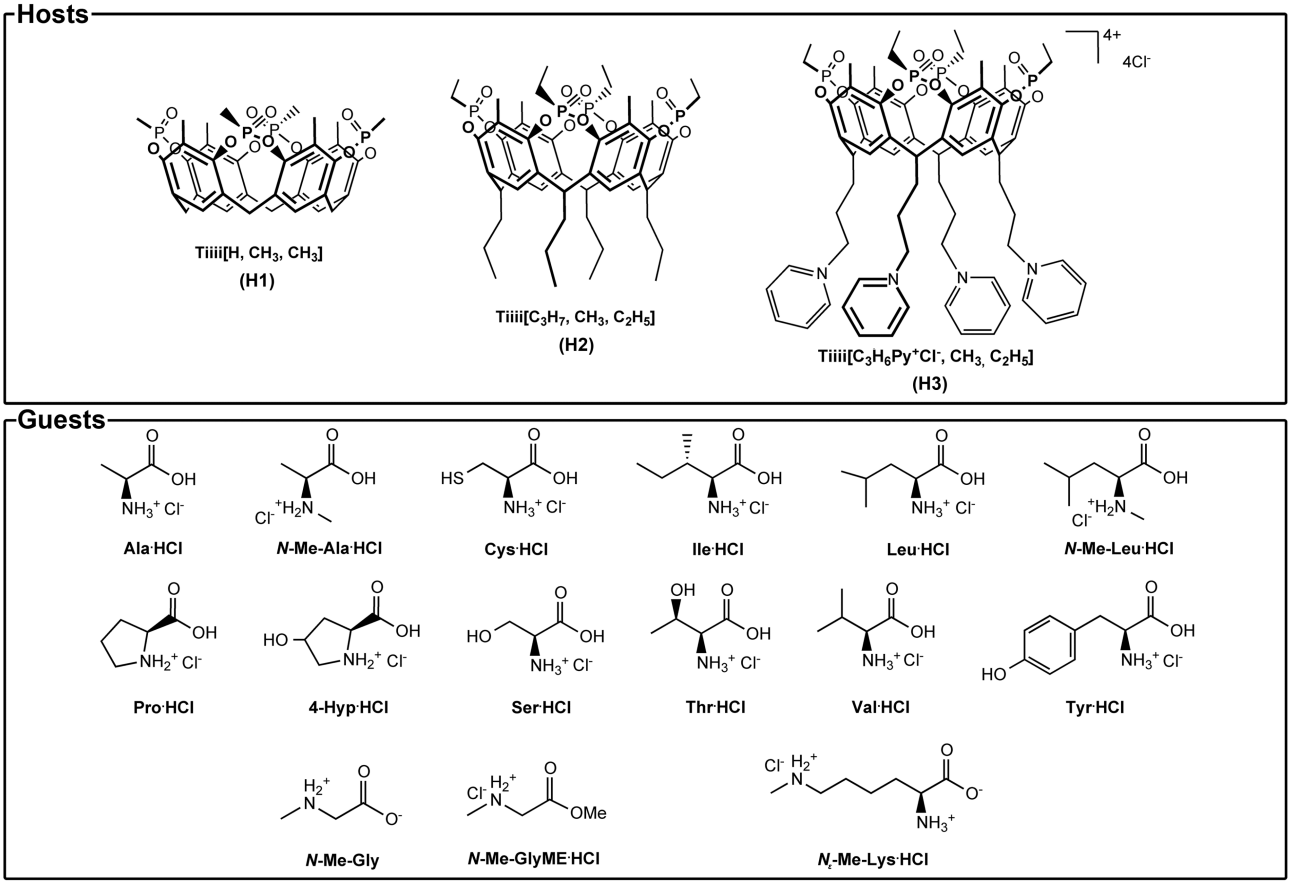
INTRODUCTION

Amino acids (AA), the building blocks of peptides, are appealing targets for molecular recognition. Their biological relevance and chemical diversity make them an ideal playground for testing the complexation ability and selectivity of synthetic receptors.[1] Since under physiological conditions the amino group is usually protonated, amino acids are mainly complexed either as zwitterions or as ammonium ions. In both cases the target is the ammonium cation, whose main potential interaction modes with a given receptor are H-bonds, cation-π and coulombic interactions and salt bridges. The occurrence and strength of these interactions is dictated by the molecular complementarity of the interacting functional groups between host and guest. Amino acid binding has been achieved using many different receptors among which crown ethers,[2] calixarenes,[3] cucurbiturils[4] and molecular tweezers[5] are prominent.

The molecular recognition properties of the above mentioned synthetic receptors toward amino acids have been exploited in protein recognition and protein activity modulation. The binding of lysine and arginine side chains in proteins by molecular tweezers has led to the development of potential new therapeutics for neurodegenerative diseases like Alzheimer and Parkinson.[6] Cucurbit[7]uril (CB[7]) interacts selectively with the *N*-terminal phenylalanine residue of insulin.[7] CB[7] incorporates the hydrophobic portion of the amino acid within the apolar cavity, while stabilizing the nearby cationic group with the carbonyl-lined portals.

In this context, the molecular recognition of *N*-mono-methylated amino acids by abiotic receptors is by far less studied. So far, most of the attention was given to the complexation of side chain permethylated arginine (N_ω)[8] and lysine (N_ε),[4b,9] due to their role in epigenetics. For the recognition of the lysine methylation state, Nature has evolved the "aromatic cage", a desolvated π-electronic rich pocket formed by a preorganized collection of aromatic amino acids residues.[9] As chemical mimic, a water soluble p-sulfonatocalix[4]arene has been used as synthetic pocket to recognize permethylated lysine.[10] The calix[4]arene is also a mediator of protein-protein interactions and potentially could be employed in generating assembly and promoting crystallization.[11] Recently, sarcosine, the *N*-mono-methylated derivative of glycine, has been identified as marker of the aggressive forms of prostate cancer in urine.[12] Tetraphosphonate cavitands were successfully used for the specific recognition of sarcosine in water and urine in presence of glycine as interferent.[13]

Chart 1. Cavitands and amino acids (represented as hydrochloride salts) involved in this study.



Tetraphosphonate cavitands (**Tiiii**)[14] are powerful synthetic receptors presenting remarkable molecular recognition properties toward *N*-methyl ammonium salts. The origin of Tiiii selectivity toward these species, can be attributed to the presence of three interaction modes: (i) N^{+}•••O=P cation–dipole interactions; (ii) cation–π interactions[15] of the $^{+}N–CH_3$ group with the π basic cavity; (iii) two simultaneous hydrogen bonds between two adjacent P=O bridges and the two nitrogen protons. Within the ammonium salt series, the bias toward the mono-methylated species over di- and tri-methylated ones is determined by the number of H-bonds formed with **Tiiii**, while non-methylated ammonium ions are less favored by the lack of cation–π interactions.[16]

Here we report the results derived from a fundamental study on the molecular recognition properties exhibited by phosphonate cavitands towards amino acids (Chart 1). The purpose of the study is to determine, case by case, which are the interactions responsible for amino acid recognition, to quantify the enthalpic and entropic contributions to the overall free binding energy and to evaluate the solvent effect. NMR competition experiments provide a binding strength scale, which complement and support the measured thermodynamic data. The emerging trends deduced from this study will be used as guidelines to predict the binding behavior of phosphonate cavitands toward proteins.

RESULTS AND DISCUSSION

Crystal Structures Determination. We undertook a systematic crystallization program to assess the complexation properties of tetraphosphonate cavitands towards amino acids in the solid state. Both zwitterionic and cationic forms of amino acids, were tested. In this study, **Tiiii[H, CH_3, CH_3]** (host **H1**), a synthetic receptor deprived of substituents at its lower rim, was chosen for its tendency to crystallize. The structures of the 1:1 host-guest complexes with the following amino acids were determined by single crystal X-ray diffraction: Ala, Ala·HCl, Cys, 4-Hyp, (D/L)-Ile·HCl, Leu·HCl, *N*-Me-Ala, *N*-Me-Leu·HCl, Pro, Ser, Thr, Tyr·HCl, Val. Crystallization trials were performed using the vapor diffusion method with sitting drops in Linbro multi-well plates containing trifluoroethanol (TFE) as solvent and PEG300 as precipitant (Table S4, SI).

Most of the single crystals obtained by co-crystallization of the cavitand in the presence of the amino acids were triclinic and only three complexes crystallized in the monoclinic $P2_1$ space group, Thr, *N*-Me-Leu·HCl and Ser, respectively. In the asymmetric unit of each crystal structure, two crystallographically independent host-guest complexes were found, except for Thr and (D/L)-Ile·HCl, which presented one independent complex, and *N*-Me-Ala and Ser, which showed four independent complexes.

A single representative host-guest complex for each crystal structure is shown in Figure 1 (relevant geometric parameters describing the host–guest interactions are reported in Tables S5 and S6). In all the structures containing the cationic forms of the amino acids (Ala·HCl, Leu·HCl, (D/L)-Ile·HCl, *N*-Me-Leu·HCl, Tyr·HCl), the chloride anion has been detected at the bottom of the cavitand lower rim, forming weak C-H···Cl^{-} interactions with the aromatic CH fragments.

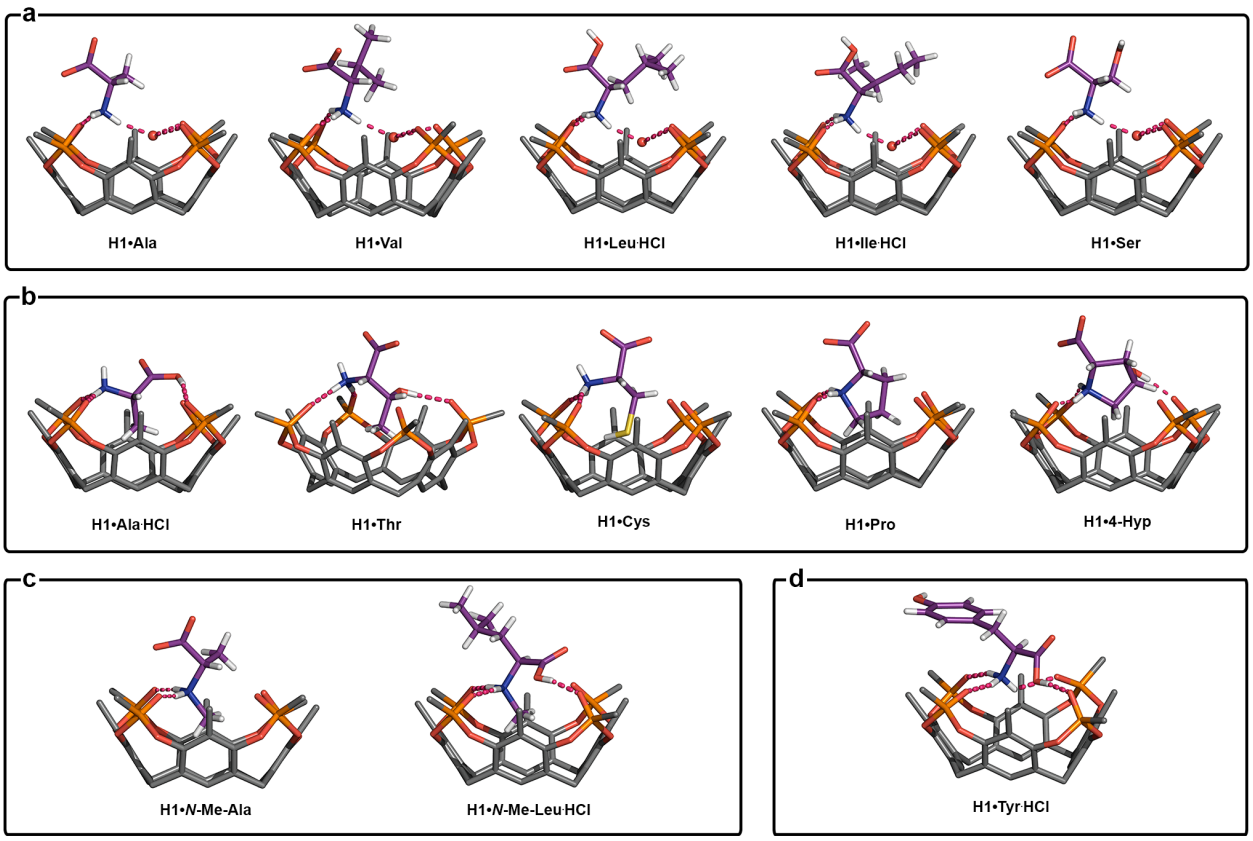


Figure 1. Side views of the molecular structures of the thirteen **H1•AA** complexes. The dashed lines depicted in magenta indicate hydrogen bonding interactions.

The thirteen supramolecular complexes showed different interaction patterns between host and guest, which can be related to the characteristic of amino acid side chain. It has been possible to distinguish four main groups of host-guest complexes, showing common interaction features for the side chains or N-Me substituents:

-not-hosted side chain (Figure 1a): the zwitterionic Ala establishes two hydrogen bonds between the amino acid ammonium group and two adjacent P=O units of the cavitand; the ammonium group completes its H-bond donor ability with a water molecule present inside the cavity of the receptor. The nitrogen atom is above the O mean plane, defined by the oxygen atoms of the P=O units, and the entire amino acid side chain is external to the cavity. Similar behavior is observed for the zwitterionic form of the hydrophobic Val and the hydrophilic Ser amino acids. Analogous weak interactions are observed also for the cationic Leu·HCl and the racemic (D/L)-Ile·HCl amino acids.

-hosted side chain (Figure 1b): the protonated Ala·HCl establishes two hydrogen bonds between the amino acid ammonium group and two oxygen atoms of the adjacent P=O units. Conversely, the ammonium group is above the O mean plane. The side chain of Ala·HCl is inserted in the cavity of the receptor: the methyl group establishes CH-π interactions with the aromatic walls, while the protonated carboxylic group forms a supplementary H-bond with a third P=O group. The zwitterionic Thr shows analogous specific recognition of the side chain with the formation of CH-π interactions and by establishing an additional hydrogen bond between the hydroxyl group and one oxygen atom of a P=O unit. Also the zwitterionic Cys shows a cavitand-side chain amino acid interaction. In this case the thiol SH group is directed inwards the cavity and forms weak sulfur-arene interactions[17] with the aromatic rings. To the best of our knowledge the **H1•Cys** is the first structures of a host-guest complex encapsulating a cysteine. The zwitterionic Pro establishes two H-bonds with both hydrogen atoms of the secondary ammonium group and two oxygen atoms of the adjacent P=O units. In this case the nitrogen atom is almost coplanar to the O mean plane. The hydrophobic five-membered ring is deeply inserted into the cavity and forms cation-π interactions with the aromatic walls of the cavitand. In the zwitterionic 4-Hyp complex the additional H-bond formed between the OH group of 4-Hyp and one P=O unit produces a significant levitation (about 1 Å) of the proline side chain with respect to the bowl-shaped cavity (Table S6).

-hosted *N*-Me (Figure 1c): the secondary ammonium group of the zwitterionic *N*-Me-Ala establishes two hydrogen bonds with two oxygen atoms of the adjacent P=O units, with the nitrogen atom almost coplanar with the mean plane defined by the oxygen atoms of the P=Os. Furthermore, the *N*-methyl group is deeply included in the cavity of the receptor establishing cation-π interactions with the aromatic walls of the cavitand. It is worth noting that these cation-π interactions are stronger that the ones involving the alkyl side chains, because of the more pronounced Lewis acid character of the methyl hydrogens on the methyl ammonium group (see ITC section). In this case the methyl side chain of the modified alanine is oriented outwards from the cavity.

Table 1 Thermodynamic data obtained from ITC experiments for the complexation of 7 AA as guest with cavitand H2 as host, measured in MeOH solution at 298 K. Reported error values correspond to the standard deviation of at least three different measurements.

Entry	Guest	$K_a \pm \sigma_{Ka}$ (M^{-1})	$\Delta H \pm \sigma_{\Delta H}$ ($KJ \cdot mol^{-1}$)	$\Delta G \pm \sigma_{\Delta G}$ ($KJ \cdot mol^{-1}$)	$T\Delta S \pm \sigma_{T\Delta S}$ ($KJ \cdot mol^{-1}$)
1	*N*-Me-Leu•HCl	1.7 (± 0.1)·10^5	-15.9 (± 0.1)	-29.8 (± 0.2)	13.9 (± 0.2)
2	*N*-Me-Ala•HCl	1.4 (± 0.1)·10^5	-20.1 (± 0.2)	-29.4 (± 0.2)	9.3 (± 0.3)
3	Pro•HCl	1.2 (± 0.2)·10^4	-19.8 (± 0.2)	-23.3 (± 0.4)	3.5 (± 0.5)
4	Thr•HCl	2.0 (± 0.3)·10^4	-8.2 (± 0.2)	-24.5 (± 0.4)	16.3 (± 0.5)
5	Ala•HCl	1.1 (± 0.1)·10^4	-7.5 (± 0.7)	-23.0 (± 0.2)	15.5 (± 0.7)
6	Tyr•HCl	1.1 (± 0.1)·10^4	-8.7 (± 0.3)	-23.0 (± 0.2)	14.3 (± 0.4)
7	Cys•HCl	7.3 (± 0.1)·10^3	-9.1 (± 0.1)	-22.0 (± 0.1)	13.0 (± 0.1)

A similar behavior was observed for the cationic *N*-Me-Leu·HCl, with the *i*-butyl side chain protruding outside the cavity. For this complex the two crystallographic independent molecules show different interactions for the protonated carboxylic group: H-bond with a P=O unit and H-bond with a solvent molecule (only the first case is reported).

-Tyr complexation (Figure 1d): protonated Tyr·HCl represents a special case of binding geometry. Tyr forms three hydrogen bonds between its ammonium group and three P=O units, while the fourth P=O unit is able to establish a hydrogen bond with OH fragment of the carboxylic group. The aromatic side chain of Tyr is parallel to the top of the cavity and it is involved in an interesting weak lone pair-π contact with the exo lone pair of electrons of one P=O oxygen atom.[18]

ITC Measurements in Methanol Solution. A comprehensive set of ITC complexation measurements was performed for the binding of seven AA (guests) with the tetraphosphonate cavitand **Tiiii[C_3H_7, CH_3, C_2H_5]** (host **H2**), using methanol as solvent. The seven AA are representative of the three groups defined above that featured non solvent mediated 1:1 complexation in the solid state (b, c, d classes of Figure 1). The ITC experiments allowed the assessment of the thermodynamic parameters ΔH, K_a, ΔG and ΔS associated with the complexation process. The thermodynamic parameters were derived from the fit of the isothermal binding curves. A single-site (monovalent) theoretical model was used to fit all the experimental binding curves. The ITC experiments were performed by using the novel cavitand **H2**, which has been synthesized in one step starting from the corresponding resorcinarene (Scheme S1). This cavitand presents four propyl chains at the lower rim and ethyl substituents on the P=O groups. All these alkyl chains were introduced to impart solubility in the solvents used for the NMR and ITC studies. The small changes in the cavitand structure with respect to the one used in the solid-state studies do not affect significantly the binding properties of the host structure.

In the ITC experiments, all amino acids were used as hydrochloride salts (Figure S17). As shown in Table 1, the inclusion process is both enthalpy and entropy driven for all the analyzed guests and the enthalpy contribution is particular large for entries **1**-**3**. The biggest K_a values were recorded for *N*-Me-Leu·HCl and *N*-Me-Ala·HCl. The other amino acids show K_a values very close to one another and differing by one order of magnitude with respect to those of the *N*-methylated amino acids, with the lowest value recorded being for Cys·HCl. The largest K_a values obtained for entries **1**-**2** highlight the pivotal role of the synergistic presence of multiple interactions. In the case of *N*-Me-Leu·HCl, *N*-Me-Ala·HCl and Pro·HCl all the three possible interactions between host and guest described in the introduction are present. In the specific case of Pro·HCl, entry **3**, the drop of the K_a value is entirely ascribable to a remarkable loss in the entropic contribution, possibly accountable to a decrease in solvent entropy gain associated with the desolvation of the guest upon complexation. This hypothesis is supported by the observation that in aqueous solutions proline tends to form aggregates by hydrophobic interaction of the pyrrolidine ring.[19] In the case of guests **4**-**7** only two interactions are possible, the cation–dipole and the H-bonding ones. The lack of one interaction is reflected in the lower value of the enthalpic contribution that causes the loss of one order of magnitude in the K_a.

Complexation NMR Experiments. Both proton and phosphorous NMR experiments were performed in deuterated methanol at -20°C. The amino acids used in the NMR experiments were the same set used in the ITC measurements. In the NMR experiments, 1 eq of the selected guest dissolved in deuterated methanol was added to an equimolar mM solution of **H2** (host) in the same solvent. Both ^1H and ^{31}P NMR spectra of the mixtures were then recorded. From the analyses of the ^1H and ^{31}P NMR spectra it was possible to monitor the chemical shift changes induced by complexation, that were experienced by the hydrogen atoms of the guest and the phosphorus atoms of the host. As example, in Figure 2, we show the proton NMR spectra of: cavitand **H2** (Figure 2a), *N*-Me-Ala·HCl guest (Figure 2b) and the inclusion complex **H2•*N*-Me-Ala·HCl** (Figure 2c). The fact that the association constant value of **H2•*N*-Me-Ala·HCl,** determined using ITC experiments, is larger than 10^5 M^{-1} warrants its quantitative formation in an equimolar solution of the two binding partner at mM concentration. Diagnostic upfield shifts of the guest hydrogen atom signals were observed, as is expected for the shielding effect caused by its inclusion in the aromatic cavity

of the host. In particular, the $^{+}$N-CH$_3$ resonance moved more than 3 ppm upfield.

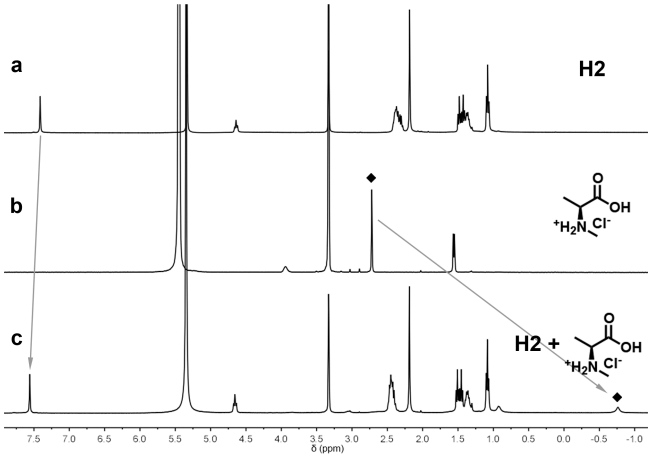


Figure 2. ^{1}H NMR (400 MHz, MeOD, 253 K) spectra of free host **H2** (a); free *N*-Me-Ala·HCl guest (b); 1:1 **H2•*N*-Me-Ala·HCl** complex (c). Signal related to *N*-methyl protons of the guest is labeled (◆).

After the addition of 1 eq. of *N*-Me-Ala·HCl, the ^{31}P resonance of the four P=O groups of the host experienced a downfield shift of 2.7 ppm (from 25.0 to 27.7 ppm, Figure 3). This chemical shift change is a clear indication of the phosphonate groups participation in the guest complexation.

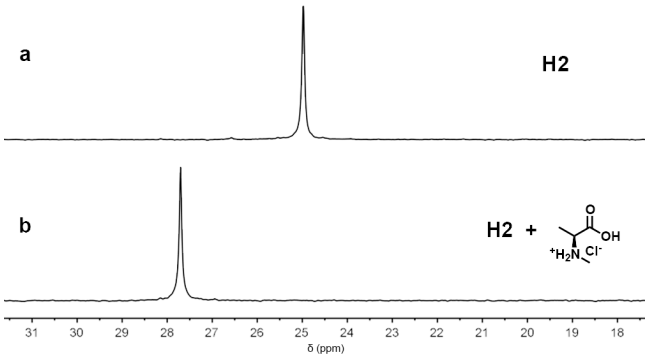


Figure 3. ^{31}P NMR (162 MHz, MeOD, 253 K) spectra of free host **H2** (a) and 1:1 **H2•*N*-Me-Ala·HCl** complex (b).

The same procedure was used to determine the complexation induced shifts experienced by the other six amino acids upon inclusion into the host's cavity. Large upfield shifts were observed for the protons of *N*-Me-Leu·HCl and Pro·HCl upon complex formation, in a similar fashion to what was observed for *N*-Me-Ala·HCl. However, the chemical shift changes induced by the complexation of the other four amino acids with **H2** were less significant.

The analysis of the ^{31}P spectra of the complexes provided more interesting information (Figure 4). Based on the extent of the ^{31}P chemical shift induced changes, the seven amino acids can be divided into two groups: the complexation of *N*-Me-Ala·HCl, Pro·HCl and *N*-Me-Leu·HCl produced the largest downfield shift of the host phosphorus signal ($\Delta\delta$ = 2.0-2.7 ppm, Figure 4f,g,h). By contrast, complexes **H2•Thr·HCl**, **H2•Tyr·HCl**, **H2•Cys·HCl** and **H2•Ala·HCl**, showed a reduced downfield shift of the phosphorous signal of the host in the range of 0.5-0.9 ppm (Figure 4b,c,d,e). This trend nicely matches with the ΔH values scale measured in ITC experiments, summarized in Table 1. The larger ΔH value was determined for the inclusion complex of *N*-Me-Ala·HCl, followed by the values of the Pro·HCl and *N*-Me-Leu·HCl complexes. For the second group of amino acids the enthalpic contribution of the binding is significantly reduced and the entropic term becomes predominant. Taken together, these results highlight that the interactions of the methyl-ammonium group of the amino acid with the P=O groups of the host are enthalpically more favorable that those of a primary ammonium counterpart. Most likely, the inclusion of the *N*-methyl substituent of the guest in the aromatic cavity of the host (hydrophobic binding) is responsible for the enhancement of the additional polar host-guest interactions (cation-π, cation-dipole and charged-hydrogen bonds).

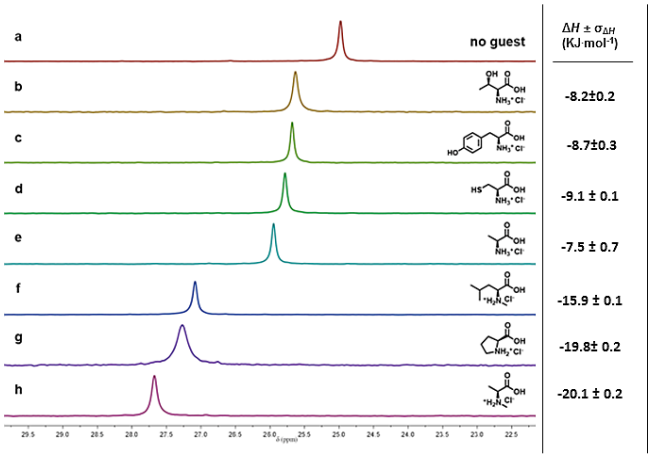


Figure 4. ^{31}P NMR (162 MHz, MeOD, 253 K) spectra of free host **H2** (a), **H2•Thr·HCl** complex (b), **H2•Tyr·HCl** complex (c), **H2•Cys·HCl** complex (d), **H2•Ala·HCl** complex (e), **H2•*N*-Me-Leu·HCl** complex (f), **H2•Pro·HCl** complex (g), **H2•*N*-Me-Ala·HCl** complex (h). For all the complexes a 1:1 stoichiometry was respected.

Competition NMR Experiments. To complete the NMR studies, three pair-wise competition experiments were performed. The results of these experiments allowed a direct assessment of the ratio of the stability constants of the two complexes. Considering the ITC data of Table 1, we decided to study the pair-wise competitions of four amino acids, namely *N*-Me-Ala·HCl, Ala·HCl, Cys·HCl and Pro·HCl. All the experiments were performed by initially adding 1 eq of the amino acid hydrochloride dissolved in deuterated methanol to an equimolar solution of **H2** in the same solvent. The resulting solution mixture was characterized by ^{1}H NMR (see Figures S1-S6) and ^{31}P NMR (Figure 5). Next, 1 eq of the competitive amino acid, also dissolved in deuterated methanol, was added to the same NMR tube, and the ^{1}H and ^{31}P spectra of the final mixture were recorded. To assure that the systems had reached the thermodynamic equilibrium, the pair-wise competition experiments were repeated by reversing the order of addition of the two amino acids.

The pair-wise competition experiment involving *N*-Me-Ala·HCl and Pro·HCl was performed in order to compare two different guests that, although presenting similar ΔH values in their binding, they exhibit different K_a magnitudes. The K_a of *N*-Me-Ala·HCl complex was found to be one order of magnitude larger than that of the corresponding Pro HCl complex (cfr entries 2 and 3, Table 1). The addition of 1 eq of

N-Me-Ala HCl to the quantitatively formed **H2•Pro·HCl** complex caused proline decomplexation and the subsequent binding of *N*-Me-Ala·HCl, as demonstrated by the appearance of the signal corresponding to the bound *N*-methyl group that resonated at -0.77 and the observation of the signals of the protons in free Pro·HCl (Figure S1). Figure 5a (left) shows the corresponding ^{31}P NMR spectra of the pair-wise competition experiment. The addition of 1 eq of *N*-Me-Ala·HCl to the **H2•Pro·HCl** complex also caused the sharpening the phosphorus signal of the receptor and a reduction of its downfield chemical shift. These observations support the exchange of the initially bound amino-acid.

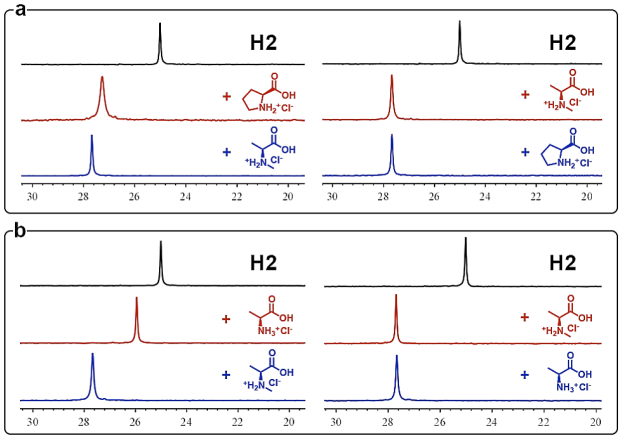


Figure 5. ^{31}P NMR (162 MHz, MeOD, 253 K) spectra: free host **H2** (black), 1:1 complex (red), addition of 1 eq of competitor to the formed complex (blue). The structure of the AA used are shown in the insets.

The reverse experiment was also performed. That is, to the preformed **H2•*N*-Me-Ala·HCl** complex 1 eq of Pro·HCl was added and the resulting mixture was analyzed by NMR spectroscopy. The proton signals for the bound *N*-Me-Ala·HCl did not show noticeable changes in their chemical shifts. In particular, the singlet of the bound *N*-methyl protons resonating at -0.77 ppm was not affected at all (Figure S2). This result evidences the superior thermodynamic stability of the **H2•*N*-Me-Ala·HCl** inclusion complex. In addition, as shown in Figure 5a (right), the phosphorus signal of the bound receptor in the above complex did not experience any change (band shape or chemical shift) upon addition of proline.

The second pair-wise competition experiment was performed between *N*-Me-Ala·HCl and Ala·HCl. This experiment was selected to highlight the importance of the mono-methylation of the ammonium group in adding one additional mode of interaction in the stabilization of the inclusion complex. In fact, alanine has the same molecular structure than *N*-Me-Ala, except for the presence of the CH_3 group on the NH_2 group. The addition of 1 eq of *N*-Me-Ala·HCl to the solution containing the quantitatively formed **H2•Ala·HCl** complex caused the displacement of the alanine by the *N*-Me-Ala counterpart. The ^{1}H NMR spectrum of the mixture shows the diagnostic sharp singlet of the included *N*-methyl protons resonating at -0.77 ppm that is indicative of the **H2•*N*-Me-Ala·HCl** complex formation (Figure S3). The signals of the free alanine protons were also detected. The addition of *N*-Me-Alanine to the **H2•Ala·HCl** complex induced a downfield shift of the receptor phosphorus signal that was also diagnostic of the exchange of the bound guest (Figure 5b, left). Conversely, the addition of one equivalent of Ala·HCl to **H2•*N*-Me-Ala·HCl** complex did not cause noticeable changes to the ^{31}P NMR spectrum of the initial mixture (Figure 5b, right).

The last pair-wise competition experiment involved Cys·HCl, the amino acid that presents the lowest K_a value, and *N*-Me-Ala·HCl. As expected, *N*-Me-Ala·HCl is the preferred guest. The related experiments are reported in SI (Figures S5-S6, S29).

The results of the NMR pair-wise competition experiments demonstrate that the **H2** receptor shows a high selectivity for the binding of *N*-Me-Ala·HCl with respect to the other tested amino acids. This preference derives from the presence of an additional interaction mode between the host and the *N*-methylated guest: the cation–π interactions of the $^+$N–CH_3 group with the aromatic panels of the cavity. These interactions are not present in the complexes of **H2** with alanine and cysteine.

Amino Acids Complexation in Water. Molecular recognition events are highly solvent dependent. Therefore, a further understanding of the complexation processes in solution requires the investigation of solvent effects. Water is an essential biological fluid, which promotes apolar aggregation and complexation processes necessary to sustain all functions of life. Consequently, complexation studies in aqueous media are of special interest since they can directly model molecular recognition events in biological systems.[20] To perform host-guest complexation studies in water, we synthesized the new water-soluble receptor **Tiiii[$C_3H_6Py^+Cl^-$, CH_3, C_2H_5]** (host **H3**). This cavitand, bearing four pyridinium feet providing water solubility, was obtained in four synthetic steps from the tetrahydroxyl footed resorcinarene,[21] with a 48% overall yield (Scheme S2).

The ability of this water-soluble cavitand to complex amino acids in water was probed by means of NMR titration experiments and ITC studies (Figures S1-S16 and S18-S21, S24). In particular, we focused our attention on the complexation of *N*-Me-GlyME·HCl, the methyl ester of *N*-methyl glycine (sarcosine), a biologically relevant amino acid.[12]

The ^{1}H NMR spectra acquired during the titration of the **H3** cavitand in D_2O solution with incremental amounts of *N*-Me-GlyME·HCl showed reduced but detectable downfield shifts for all the signals of the aromatic protons of the receptor (Figure S7). In particular, the singlet assigned to the aromatic proton of the resorcinarene core experienced the most noticeable downfield shift ($\Delta\delta$ = - 0.16 ppm). We observed that, in the early stages of the titration, the protons of the *N*-Me and methylene groups of *N*-Me-GlyME·HCl appeared as two separate broad signals significantly upfield shifted with respect to their chemical shift values for the free guest. In contrast, the methyl ester group resonated as sharp singlet, being only slightly upfield shifted. The incremental addition of *N*-Me-GlyME·HCl gradually shifted downfield all the proton signals detected for the guest, to chemical shift values closer to those of the free guest. Altogether, these observations indicated that the binding of *N*-Me-GlyME·HCl by the cavitand displayed a fast chemical exchange on the NMR timescale and that the *N*-Me and methylene protons of the guest were the most affected by the complexation process. The large upfield shifts suffered by these guest's protons are in agreement with their inclusion in the receptor's cavity, where

they experience a strong magnetic shielding produced by the four phenyl rings defining the resorcinarene cavity. The observed chemical shift changes were mathematically analyzed using a 1:1 binding model and the HypNMR2008 software. The fit of the data was good (Figure S25) and allowed the calculation of a binding constant value of 4.8×10^3 M^{-1} for the 1:1 complex. In addition, the non-linear fit of the titration data, returned the chemical shift values of all the analyzed protons in the complex. These values were used for the calculation of their complexation induced shifts (CIS). CIS values provide fundamental information with respect to the complex binding geometry. Thus, the calculated CIS values for the $^{+}N–CH_3$ and the methylene protons were $\Delta\delta$ = - 3.6 ppm and $\Delta\delta$ = - 1.04 ppm, respectively. These magnitudes locate the $^{+}N–CH_3$ group deeply included into the receptor's cavity and place the methylene group closer to the cavity opening where it experiences a reduced magnetic shielding. The calculated CIS for the protons of the methyl ester group is just $\Delta\delta$ = - 0.16 ppm suggesting its outward orientation with respect to the aromatic cavity of the host.

The complexation of **H3** with *N*-Me-GlyME·HCl was also probed using ^{31}P NMR spectroscopy (Figure S8). The phosphorous signals of the bridging phosphonate groups in the receptor experienced a downfield CIS of $\Delta\delta$ = 2.44 ppm. This result evidences the participation of the phosphonate groups in the complexation of *N*-Me-GlyME·HCl.

We had previously studied the binding of *N*-Me-GlyME·HCl by tetraphosphonate cavitand in methanol using ITC.[13a] For this reason, we decided to determine the thermodynamic parameters of the complexation process of *N*-Me-GlyME·HCl in water using the same technique (Figure S18).

Changing the solvent from methanol to water produced a drop of the complex thermodynamic stability (K_a) of almost two orders of magnitude (cfr entries 1 and 2 of Table 2). Noticeably, the stability constant determined for the **H3•*N*-Me-GlyME·HCl** complex in water using ITC experiments nicely matched the value previously derived from the NMR titrations. The origin of the reduction in binding affinity observed for *N*-Me-GlyME·HCl in water solution was assigned to the change in the sign of the entropic term. The negative ΔS contribution is a distinct feature of the formation of these complexes in water. In the solid-state, one molecule of methanol solvates the aromatic cavity of the cavitand receptor by including its methyl group.

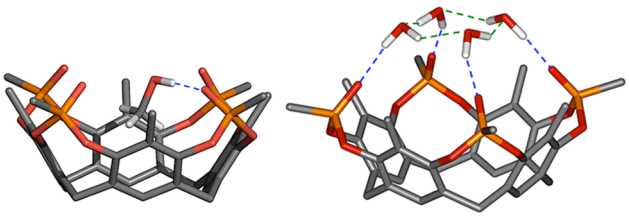


Figure 6. a) Side view of the crystal structure of **H1•MeOH**.[24] The dashed blue lines indicate hydrogen bonding between the cavitands and methanol. b) Side view of the crystal structure of **H1•4H$_2$O**.[22] The dashed lines depicted in blue indicate hydrogen bonding between the cavitand and water, while water-water interactions are depicted in green.

On the contrary, water molecules were not located in the cavity of the receptor (Figure 6). We have shown that cavitand **H1**, in the solid-state, forms a complex with four water molecules in which they act as hydrogen bond donors to the P=O groups (Figure 6b, blue dotted line) and also as hydrogen bond acceptors towards adjacent water molecules (green dotted lines).[22] Each water molecule is more ordered than in the bulk and it is involved in four hydrogen bonds (three shown in Figure 6b plus one in the lattice). Thus, we did not observe high energy water molecules[23] (not hydrogen bonded or partially hydrogen bonded) included in the host cavity, which remained empty. Methanol, instead formed a 1:1 complex with **H1** (Figure 6a), interacting with the cavity *via* H-bonding and CH–π interactions.[24]

In our interpretation, the release of the included methanol to the bulk solution should contribute significantly to the observed entropy gain on complex formation, while in water the cavity is poorly solvated. However, the global picture of complex formation involves several contributions, which compromise the attribution of the observed effect to a single cause. In methanol solution, the entropy cost for the formation of the complex is overcompensated by the entropy gained by the desolvation of the binding partners and solvation of the corresponding complex. Conversely, in water the solvation/desolvation processes do not seem to contribute in a related extent to the binding, confirming the enthalpy driven complexation ("nonclassical hydrophobic effect"),[23] observed in aqueous environment for many other host–guest systems.[25] In particular, the solvation of the cavitand, when complexed by the charged amino acid, should be enhanced in water with respect to methanol.

We also investigated the binding of the sarcosine zwitterion in water using both NMR titrations and ITC experiments (Figures S9, S10 and S21). The results of the ^1H NMR titrations revealed an analogous trend of chemical shift changes as observed for the methyl ester derivative, however, the assessed binding constant value was reduced to a value of $K_a = 1.6 \times 10^3$ M^{-1} (Figure S26).

Likewise, the CIS values determined for the *N*-Me and methylene groups of bound sarcosine zwitterion, $\Delta\delta$ = -2.90 ppm and $\Delta\delta$ = -0.83 ppm respectively, were also reduced by almost 0.7 and 0.2 ppm when compared to those of the methyl ester derivative. The CIS of the phosphorous signal also diminished ($\Delta\delta$ = -1.6). Taken together, these results suggested that the weaker binding of the zwitterionic and neutral species is related to slight modification of the complex binding geometry, in which the amino acid is less buried into the receptor's cavity and forms longer hydrogen bonds with it.

The investigation of the binding process using ITC assigned a $K_a = 1.02 \times 10^3$ M^{-1} to the complex, also in agreement with the NMR estimate. The enthalpic term of the complex was reduced by 5.7 KJ/mol compared to the binding of the methyl ester counterpart *N*-Me-GlyME·HCl. (cfr entry 5 with entry 2 of Table 2). We observed a typical enthalpy-entropy compensation effect, but the entropic term still maintained the negative sign.

Table 2. ITC data in MeOH and water for complexation of *N*-Me-GlyME•HCl and *N*-Me-Gly.

Entry	Guest (Solvent)	$K_a \pm \sigma_{Ka}$ (M^{-1})	$\Delta H \pm \sigma_{\Delta H}$ (KJ·mol^{-1})	$\Delta G \pm \sigma_{\Delta G}$ (KJ·mol^{-1})	$T\Delta S \pm \sigma_{T\Delta S}$ (KJ·mol^{-1})
1	*N*-Me-GlyME·HCl (MeOH)[13a]	6.8 (± 0.5)·10^4	-14.5 (± 0.2)	-28.0 (± 0.7)	13.5 (± 0.1)
2	*N*-Me-GlyME·HCl (H_2O)	3.43 (± 0.03)·10^3	-33.0 (± 0.1)	-20.18 (± 0.04)	-12.8 (± 0.1)
3	*N*-Me-GlyME·HCl (D_2O)	3.75 (± 0.06)·10^3	-35.7 (± 0.25)	-20.39 (± 0.04)	-15.3 (± 0.25)
4	*N*-Me-GlyME·HCl (PBS)	1.89 (± 0.07)·10^3	-35.0 (± 0.2)	-18.71 (± 0.08)	-16.3 (± 0.2)
5	Sarcosine (H_2O)	1.02 (± 0.01)·10^3	-27.3 (± 0.3)	-17.17 (± 0.04)	-10.2 (± 0.3)

The solvent isotopic effect was also investigated in the complexation of *N*-Me-GlyME·HCl in D_2O. Only the NH_2 protons of *N*-Me-GlyME·HCl are expected to exchange for deuterium atoms when dissolved in D_2O, maintaining the rest of the system unperturbed. The K_a value derived from ITC experiments do not change significantly upon moving from water to D_2O (cfr entry 3 with entry 2), suggesting that the variation is mainly due to small changes in hydration shells.

In order to evaluate the effect of the ionic strength in the binding affinity, we decided to perform ITC measurements in phosphate buffer solution (PBS 0.1 M; pH = 7). The presence of physiological salts did not affect significantly the complexation strength. Similarly, the thermodynamic signature of the binding event was maintained (cfr entries 2 and 4). This last result bodes well for the use of phosphonate cavitands in real physiological conditions.

The enthalpy and entropy terms determined for the complexation of *N*-Me-GlyME·HCl with the cavitand in water are indicative of non-classical hydrophobic binding (enthalpy favored, entropy opposed). This fact is also substantiated by the observed increase in enthalpy ($\Delta\Delta H_{D/H}$ = -2.7 kJ/mol) for the *N*-Me-GlyME·HCl complexation in D_2O, resulting from the better hydrogen bonding capabilities of the released D_2O molecules.[23,26] The entropic cost resulted from the tight binding and well defined binding geometry of the complex, as well as from the formation of highly ordered structures of water molecules solvating the complex, which in turn are responsible of the enthalpic gain measured in D_2O solution.

The *N*$_\epsilon$-Methyl-Lysine Case. The complexation of this amino acid at its *N*$_\epsilon$ position is of utmost importance in view of its role in histone epigenetic modifications.[9,10c] Tetraphosphonate cavitand complexation of *N*$_\epsilon$-methyl-lysine is complicated by the presence of two different charged ammonium groups. In fact, crystallization of cavitand **H1** from trifluoroethanol in the presence of *N*$_\epsilon$-Me-Lys afforded a 2:1 host-guest complex structure, which was determined by single crystal X-ray diffraction (Figure 7a). The two binding geometries replicate exactly the ones observed respectively for alanine and its *N*-methyl derivative (Figure 1a,c); the -NH_3^+ ammonium group of peptidic fragment forms H-bonds with the P=O groups and a water molecule, whereas the methyl group on *N*$_\epsilon$ position is hosted in the cavity of an opposing cavitand. Conversely, crystals of the **H3•*N*$_\epsilon$-Me-Lys·HCl** complex grown from water, gave a 1:1 stoichiometry complex in which only the $^+NH_2$–CH_3 group is bound (Figure 7b). The *N*$_\epsilon$-methyl-

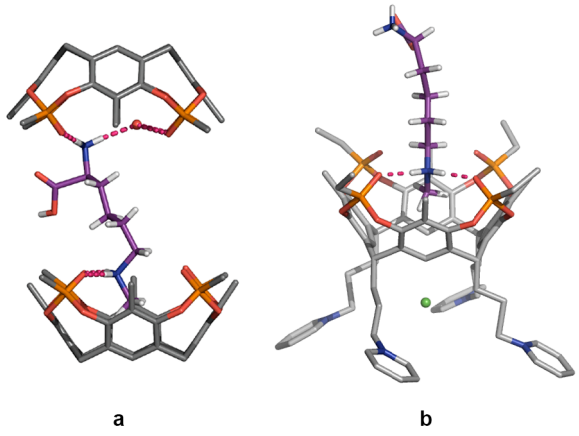


Figure 7. Side view of the molecular structures of **2H1•*N*$_\epsilon$-Me-Lys·HCl** (a) and **H3•*N*$_\epsilon$-Me-Lys·HCl** (b) complexes. The dashed lines depicted in magenta indicate hydrogen bonding interactions.

lysine interacts with the cavity of the host through the usual synergistic three interactions discussed above. The two hydrogen bonds are still present in water thanks to the strong hydrophobic interaction between the electron-rich cavity of the host and the *N*-methyl group of the guest, which brings the hydrogen atoms of the ammonium residue and the P=O groups closer, releasing simultaneously the water molecules that were hydrogen bonded to the phosphonate groups. The distance of -1.40(2) Å of the carbon atom of the hosted ^+N–CH_3 group from the mean plane defined by the oxygen atoms of the P=O groups pointing inwards the cavity (OOp) evidences that the methyl group is deeply inserted in the **H3** cavity. A corresponding value of -1.403(2) Å is also observed for the OOp of the methyl group of the *N*$_\epsilon$-Me-Lys hosted in the **H1** cavity (Table S6).

Firstly, we investigated the complexation process of *N*$_\epsilon$-methyl-lysine hydrochloride with **H2** in methanol solution at 253 and 298 K, moving from slow to fast chemical exchange on the NMR time scale, respectively. The binding process was monitored by acquiring ^1H and ^{31}P NMR spectra in the

presence of incremental amounts of guest (Figures S11-S14). At 298 K, the aromatic signal of the host experienced a monotonic downfield shift. Moreover, during the initial phase of the titration, we observed the emergence of a sharp singlet in the downfield region of the spectra that was assigned to the protons of the *N*-methyl group included in the aromatic cavity of the host (Δδ = -0.88 ppm). Other signals for the protons of the amino acids were not detected due to broadening. The addition of more than one equivalent of *N*-methyl-lysine hydrochloride did not produce further chemical shift changes to the signals of the cavitand both in the ^{1}H NMR and ^{31}P NMR spectra, but induced the broadening of the guest signal resonating highly upfield. From these observations, we assumed that at mM concentration a 1:1 complex with a binding constant larger than 10^4 M^{-1} was formed. The binding geometry of the complex implied the exclusive inclusion of the *N*-methyl group in the aromatic cavity of the host. When the titration was repeated at 253 K (Figure S11), two separate signals assigned to the aromatic protons in the free and the bound host were detected. This observation was indicative that the binding process displayed slow chemical exchange on the NMR timescale at this lower temperature. The obtained results in the NMR titrations did not provide any evidence for the formation of a 2:1 complex.

When the binding process of *N*-methyl-lysine hydrochloride with **H2** in methanol solution was probed using ITC experiments at 298 K, a completely different scenario was revealed. As shown in Figure 8a, the existence of two sigmoidal curves in the binding isotherm evidenced the presence of more than one binding event. The first one had its inflexion point centered at a molar ratio value close to 0.5, while the inflexion point of the second sigmoidal was near a molar ratio value of 1. Not surprisingly, we couldn't fit the titration data to a simple 1:1 binding model. Using the HypDH software, we analyzed the titration data with a binding model that assume the formation of two complexes having 1:1 and 2:1 (host:guest) stoichiometry, respectively. We obtained a good fit to the experimental data. The presence of a complex with stoichiometry different to the simple 1:1 was also confirmed by performing a reversed ITC titration in which the host was incrementally added to the guest solution located in the cell of the calorimeter (Figure S23). In the reverse titration, a clear sigmoidal curve with an inflexion point centered at molar ration close to 1 is visible. The second binding event, the formation of the 2:1 complex, is less visible, but can be inferred from the heat release that takes places even when more than 1.5 eq of the host are added. The mathematical analysis of the reverse titration was also undertaken using the HypDH software and the 1:1, 2:1 binding model affording a remarkable good fit to the experimental values. The thermodynamic data determined for the binding process in methanol are summarized in Table 3. The ITC observations for the complexation process of **H2** and *N*-methyl-lysine hydrochloride in methanol solution are in complete agreement with the formation of the 2:1 complex observed in single crystals grown from trifluoroethanol.

The values of the two binding constants determined for the formation of the 2:1 complex showed a difference of three orders of magnitude. Based on previous binding constant values Next, we probed the interaction of *N*-Me-Lys·HCl with **H3** in water solution using ^{1}H and ^{31}P NMR spectroscopy (Figures S15-S16). The incremental addition of the amino acid to a solution of the receptor in water induced a downfield shift of the aromatic proton of the resorcinarene core (CIS, Δδ = 0.16 ppm). The phosphorus signal of the receptor also moved downfield (CIS, Δδ = 2.26 ppm). In contrast, the proton signals of the guest suffered significant broadening specially in the initial phase of the titration and were more difficult to detect and assign. Nevertheless, the titration data were analyzed mathematically using a 1:1 binding model, the fit was good, and returned a binding constant value of 1.8×10^3 M^{-1} (Figure S27). determined in MeOH, we assigned the first binding event to the inclusion of the *N*-Me residue of the *N*-Me-Lys·HCl in the aromatic cavity of one host and the second binding event to the complexation of its α-NH_3 moiety with a second molecule of **H2**.

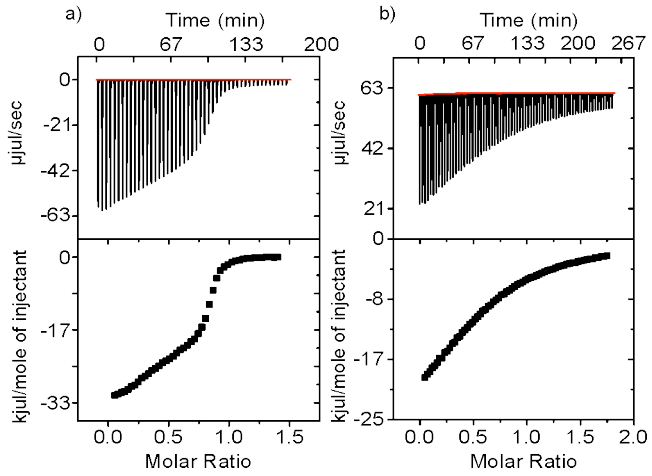


Figure 8. ITC experiments: a) methanol, **H2** = 1.04 mM, *N*-Me-Lys·HCl = 6.74 mM; b) water, **H3** = 1.69 mM, *N*-Me-Lys·HCl = 13.78 mM. The host was placed in the cell and the guest in the syringe. Top trace: raw data for the ITC titration. Bottom trace: binding isotherm of the integrated calorimetric titration data.

The chemical shift values for the host and guest protons in the complex were also determined. The magnitude of the binding constant and the CIS for the host are in good agreement with the values obtained for the complexation of *N*-Me-GlyME·HCl by the same host in water. The CIS values of the *N*-Me group and its α methylene in the bound guest were Δδ = -3.3 and -0.97 ppm, respectively. These values coincide with the ones determined for the *N*-Me-GlyME·HCl complex and suggest a closely related binding geometry for the terminal *N*-Me groups. As expected, the proton α to the primary ammonium group of *N*-Me-Lys·HCl was not affected by the binding. We concluded that in water solution *N*-Me-Lys·HCl and **H3** formed exclusively a 1:1 complex through the inclusion of the terminal *N*-Me group in the host's cavity. This result is fully consistent with the solid state data. The value of the stability constant for this complex determined from the NMR titration data was almost three orders of magnitude lower in water than in methanol.

When the binding of Nε-Me-Lys·HCl and **H3** in water was analyzed using ITC experiments (Figure 8b), the titration data could be fitted to a simple 1:1 binding model. The calculated thermodynamic parameters are summarized in Table 3. The value of the calculated binding constant is in agreement with the one determined using NMR titrations. In close analogy to the binding of **H3** with *N*-Me-GlyME·HCl in water, the complexation event is only enthalpy driven (large negative

ΔH), whilst the entropy contribution is unfavorable (negative $T\Delta S$).

As expected, the modification of the ionic strength of the media did not affect significantly the thermodynamic signature of the *N*-Me-Lys·HCl complex (Entry 3, Table 3). Water wipes out the weaker ammonium binding due to H-bonding observed

Table 3. ITC data for complexation of *N*-Me-Lys with H2 and water-soluble cavitand H3 measured at 298 K. All entries refer to host titration with the guest.

Entry	Solvent	$K_{1:1} \pm \sigma_{Ka}$ (M^{-1})	$\Delta H_{1:1} \pm \sigma_{\Delta H}$ ($KJ \cdot mol^{-1}$)	$\Delta G \pm \sigma_{\Delta G}$ ($KJ \cdot mol^{-1}$)	$T\Delta S \pm \sigma_{T\Delta S}$ ($KJ \cdot mol^{-1}$)	$K_{1:2} \pm \sigma_{Ka}$ (M^{-1})	$\Delta H_{1:2} \pm \sigma_{\Delta H}$ ($KJ \cdot mol^{-1}$)	$\Delta G \pm \sigma_{\Delta G}$ ($KJ \cdot mol^{-1}$)	$T\Delta S \pm \sigma_{T\Delta S}$ ($KJ \cdot mol^{-1}$)
1	MeOH	1.07 (± 0.08)·10^6	-24.43 (± 0.17)	-34.40 (± 0.18)	9.97 (± 0.25)	1.45 (± 0.09)·10^3	-12.72 (± 0.25)	-18.04 (± 0.16)	5.32 (± 0.3)
2	H_2O	1.49 (± 0.06)·10^3	-31.46 (± 1.5)	-18.1 (± 0.1)	-13.36 (± 1.5)	-	-	-	-
3	PBS	1.13 (± 0.02)·10^3	-26.7 (± 0.1)	-17.41 (± 0.05)	-9.33 (± 0.11)	-	-	-	-

in methanol, leaving only the stronger *N*-methyl ammonium complexation driven by hydrophobic cation–π interactions as unique binding mode.

Correlation between K_a and molecular structures. The relevant role of cation–π interactions, played in the selective recognition of the *N*-methylated AA, is also highlighted by the interesting correlation found between the thermodynamic and structural parameters for the complexes where the cavitand hosts the N–CH_x fragment.

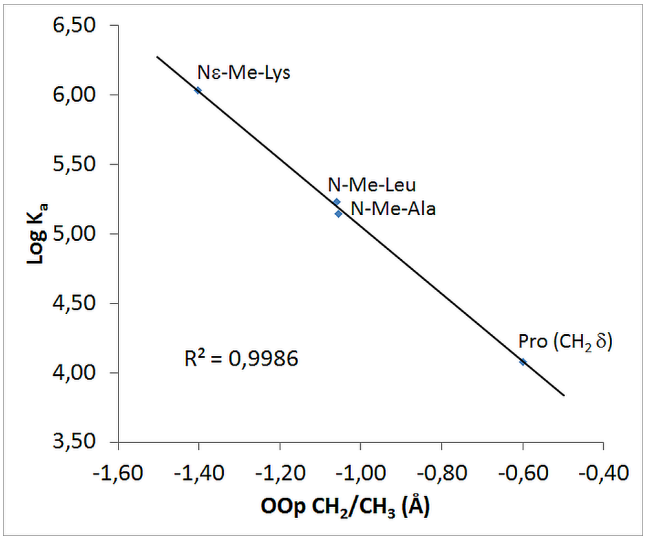


Figure 9. Linear regression curve showing the relationship between the thermodynamic association constant, $\log K_a$ (obtained from ITC experiments in MeOH) and the structural data of cation–π interactions (evaluated as distance [Å] of the carbon atom of the N–CH_x group from the entrance of cavity defined by mean plane of the oxygen atoms of the P=O groups, OOp).

In Figure 9, the $\log K_a$ values obtained from ITC experiments with **H2** in methanol are plotted against the distance of the carbon atom of N–CH_x fragment from the OOp plane observed in the crystal structures of **H1** complexes with *N*-Me-Lys,[27] *N*-Me-Leu, *N*-Me-Ala and Pro. The linear regression curve shows that an increase of the association constant is associated with the deeper insertion of the $^+$N–CH_x group in the cavity. This experimental trend indicates that the cation–π interactions in the solid state, quantified by the distance of the N–CH_x group from the entrance of cavity defined by the oxygen atoms of the P=O groups, have predictive value for the evaluation of the resulting association constants.

CONCLUSIONS

The molecular recognition properties of tetraphosphonate cavitands toward different amino acids have been thoroughly investigated in the solid state *via* X-ray diffraction and in solution by complementary ITC and NMR titration analyses. These experiments were performed both in methanol and aqueous media, in order to evaluate the solvent effect in the binding event. The association constants calculated both through NMR titrations and ITC were perfectly in agreement. The solid state studies revealed that it is possible to distinguish four different complexes' groups based on different interaction modes, depending on the characteristic of amino acid side chain. Complexation experiments in methanol showed that the ^{31}P chemical shifts correlate nicely with the enthalpy contributions to complexation, while competition experiments highlighted a remarkable selectivity among amino acids. The thermodynamic parameters obtained *via* ITC showed that, upon moving from methanol to water, the complex formation changes from an enthalpy-entropy driven process to an enthalpy driven-entropy opposed process.

In our interpretation, the entropy loss experienced in water is attributable to a decrease in solvent entropy gain mainly associated with the desolvation of the guest upon complexation. The thermodynamic profile of complex formation in water is typical of a non-classical hydrophobic effect. We demonstrated that the entropy loss experienced in moving from methanol to water leads to a drop of three orders of magnitude in K_a. However, this drop is associated to a dramatic increase in selectivity, since in aqueous solutions only *N*-methylated amino acids are effectively recognized at mM concentration. Moving from water to PBS buffered solution does not change the thermodynamic signature of the process, confirming that phosphonate cavitands can be used under real physiological conditions as synthetic receptor for mono-methylated amino acids. The overall pictures emerging from these studies point out the pivotal role played by the mono-methylated ammonium ion. Its presence inside the cavity acts as a "hook" to reinforce both cation-dipole and H-bonding interactions, to the point that it revealed to be necessary to ensure complexation in water. The importance of

cation–π interactions is supported by the direct correlation found between the $\log K_a$ in methanol solution and the depth of $^{+}N–CH_3$ cavity inclusion in the molecular structures.

This specificity of binding in water is unique among the synthetic receptors employed for AA recognition. The same complexation protocol applied to the case of *N*-methyl-lysine proved to be highly reproducible both in methanol and water, opening the way for the use of tetraphosphonate cavitands to detect epigenetic modifications of histones. Studies are under way in our lab to transduce this specific recognition event into a readable signal for sensing the *N*-mono-methylation of lysine in aqueous environment, either alone[28] or as part of histone proteins.

ASSOCIATED CONTENT

Supporting Information.
Experimental procedures (Schemes S1-S2); NMR experiments (Figures S1-S16); ITC titrations curves and fittings (Figures S17-S24); NMR CIS curves (Figures S25-S27) and calculated chemical shifts (Tables S1-S3); supporting crystallographic data (Tables S4-S6). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

* E-mail: enrico.dalcanale@unipr.it
* E-mail: sgeremia@units.it

Author Contributions

# These two authors contributed equally to the work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This work was supported by the EU through Project DOGGIES (FP7-SEC-2011-285446). A. P. thanks INSTM for partial support of his scholarship. We acknowledge the Centro Interfacoltà di Misure "G. Casnati" of the University of Parma for the use of NMR and HR-MS facilities.

REFERENCES

(1) Späth, A.; König, B. *Beilstein J. Org. Chem.* **2010,** *6,* No. 32.

(2) (a) Metzger, A.; Gloe, K.; Stephan, H.; Schmidtchen, F. P. *J. Org. Chem.* **1996,** *61,* 2051. (b) Zhang, X. X.; Bradshaw, J. S.; Izatt, R. M. *Chem. Rev.* **1997,** *97,* 3313. (c) Weimann, D. P.; Winkler, H. D. F.; Falenski, J. A.; Koksch, B.; Schalley, C. A. *Nat. Chem.* **2009**, *1*, 573. (d) Späth, A.; König, B. *Tetrahedron* **2010**, *66*, 1859. (e) Chen, Y.; Rodgers, M. T. *J. Am. Chem. Soc.* **2012**, *134*, 5863.

(3) Mutihac, L.; Lee, J. H.; Kim, J. S.; Vicens, J. *Chem. Soc. Rev.* **2011**, *40*, 2777.

(4) (a) Rajgariah, P.; Urbach, A. R. *J. Inclusion Phenom. Macrocyclic Chem.* **2008,** *62,* 251. (b) Gamal-Eldin, M. A.; Macartney, D. H. *Org. Biomol. Chem.* **2013**, *11*, 488-495; (c) Minami, T.; Esipenko, N. A.; Zhang B.; Isaacs, L.; Anzenbacher, P., Jr. *Chem. Commun.* **2014**, 61. (d) Ghale, G.; Nau, W. M. *Acc. Chem. Res.* **2014**, *47*, 2150.

(5) (a) Wehner, M.; Schrader, T.; Finocchiaro, P.; Failla, S.; Consiglio, G. *Org. Lett.* **2000**, *2,* 605. (b) Fokkens, M.; Schrader, T.; Klärner, F.-G. *J. Am. Chem. Soc.* **2005,** *127,* 14415. (c) Klaerner, F.-G.; Schrader, T. *Acc. Chem. Res.* **2013**, *46*, 967.

(6) (a) Attar, A.; Ripoli, C.; Riccardi, E.; Maiti, P.; Li Puma, D. D.; Liu, T.; Hayes, J.; Jones, M. R.; Lichti-Kaiser, K.; Yang, F.; Gale, G. D.; Tseng, C. H.; Tan, M.; Xie, C. W.; Straudinger, J. L.; Klärner, F. G.; Schrader, T.; Frautschy, S. A.; Grassi, C.; Bitan, G. *Brain,* **2012**, *135*, 3735. (b) Bier, D.; Rose, R.; Bravo-Rodriguez, K.; Bartel, M.; Ramirez-Anguita, J. M.; Dutt, S.; Wilch, C.; Klärner, F.-G.; Sanchez-Garcia, E.; Schrader, T.; Ottmann, C. *Nat. Chem.* **2013**, *5*, 234. (c) Lopes, D. H. J.; Attar, A.; Nair, G.; Hayden, E. Y.; Du, Z.; McDaniel, K.; Dutt, S.; Bandmann, H.; Bravo-Rodriguez, K.; Mittal, S.; Klärner, F.-G.; Wang, C.; Sanchez-Garcia, E.; Schrader, T.; Bitan, G. *ACS Chem. Biol.* **2015**, *10*, 1555.

(7) Chinai, J. M.; Taylor, A. B.; Ryno, L. M.; Hargreaves, N. D.; Morris, C. A.; Hart, P. J.; Urbach, A. R. *J. Am. Chem. Soc.* **2011**, *133*, 8810.

(8) James, L. I.; Beaver, J. E.; Rice, N. W.; Waters, M. L. *J. Am. Chem. Soc.* **2013**, *135*, 6450.

(9) Daze, K. D.; Hof, F. *Acc. Chem. Res.* **2013**, *46*, 937.

(10) (a) Beshara, C. S.; Jones, C. E.; Daze, K. D.; Lilgert, B. J.; Hof, F. *ChemBioChem* **2010**, *11*, 63. (b) Florea, M.; Kudithipudi, S.; Rei, A.; González-Álvarez, M. J.; Jeltsch, A.; Nau, W. M. *Chem. Eur. J.* **2012**, *18*, 3521. (c) McGovern, R. E.; Snarr, B. D.; Lyons, J. A.; McFarlane, J.; Whiting, A. L.; Paci, I.; Hof, F.; Crowley, P. B. *Chem. Sci.* **2015**, *6*, 442.

(11) McGovern, R. E.; Fernandes, H.; Khan, A. R.; Power, N. P.; Crowley, P. B. *Nat. Chem*. **2012**, *4*, 527.

(12) Sreekumar, A.; Poisson, L. M.; Rajendiran, T. M.; Khan, A. P.; Cao, Q.; Yu, J.; Laxman, B.; Mehra, R.; Lonigro, R. J.; Li, Y.; Nyati, M. K.; Ahsan, A.; Kalyana-Sundaram, S.; Han, B.; Cao, X.; Byun, J.; Omenn, G. S.; Ghosh, D.; Pennathur, S.; Alexander, D. C.; Berger, A.; Shuster, J. R.; Wei, J. T.; Varambally, S.; Beecher, C.; Chinnaiyan, A. M. *Nature* **2009**, *457*, 910.

(13) (a) Biavardi, E.; Tudisco, C.; Maffei, F.; Motta, A.; Massera, C.; Condorelli, G. G.; Dalcanale, E. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 2263. (b) Valenti, G.; Rampazzo, E.; Biavardi, E.; Villani, E.; Fracasso, G.; Marcaccio, M.; Bertani, F.; Ramarli, D.; Dalcanale, E.; Paolucci, F.; Prodi, L. *Faraday Discuss.* **2015**, *185*, 299.

(14) Pinalli, R.; Suman, M.; Dalcanale, E. *Eur. J. Org. Chem.* **2004**, 451.

(15) According to the electrostatic model, the term cation-π is more appropriate than CH-π to describe the interactions of *N*-methylammonium ions. See: Dougherty, D. A. *Acc. Chem. Res.* **2013**, *46*, 885.

(16) Dionisio, M.; Oliviero, G.; Menozzi, D.; Federici, S.; Yebeutchou, R. M.; Schmidtchen, F. P.; Dalcanale, E.; Bergese, P. *J. Am. Chem. Soc.* **2012**, *134*, 2392.

(17) (a) Senćanski, M.; Došen-Mićović, L.; Šukalović, V.; Kostić-Rajačić, S. *Struct. Chem.* **2015**, *26*, 1139. (b) Daeffler, K. N.-M.; Lester, H. A.; Dougherty, D. A. *J. Am. Chem. Soc.* **2012**, *134*, 14890. (c) Salonen, L. M.; Ellermann, M.; Diederich, F. Angew. *Chem. Int. Ed.* **2011**, *50*, 4808. (d) Biswal, H. S.; Wategaonkar, S. *J. Phys. Chem. A* **2009**, *113*, 12774. (e) Ringer, A. L.; Senenko, A.; Sherrill, C. D. *Protein Science* **2007**, *16*, 2216. (f) Castellano, R. K.; Diederich, F.; Meyer, E. A. *Angew. Chem. Int Ed.* **2003**, *42*, 1210.

(18) Egli, M.; Sarkhel, S. *Acc. Chem. Res.* **2007***, 40*, 197.

(19) Keswani, N.; Kar, K.; Kishore, N. *J. Chem. Thermodyn.* **2010**, *42*, 597.

(20) Smithrud, D. B.; Sanford, E. M.; Chao, I.; Ferguson, S. B.; Carcanague, D. R.; Evanseck, J. D.; Houk, K. N.; Diederich, F. *Pure & Appl. Chem.* **1990**, *62*, 2227.

(21) (a) Mezo, A. R.; Sherman, J. C. *J. Org. Chem.* **1998**, *63*, 6824. (b) Pirondini, L.; Bonifazi, D.; Menozzi, E.; Wegelius, E.; Rissanen, K.; Massera C.; Dalcanale, E. *Eur. J. Org. Chem.* **2001**, 2311.

(22) Massera, C.; Melegari, M.; Ugozzoli, F.; Dalcanale, E. *Chem. Comm.* **2010**, *46*, 88.

(23) Biedermann, F.; Nau, W. M.; Schneider, H.-J. *Angew. Chem. Int. Ed.* **2014**, *53*, 11158.

(24) Maffei, F.; Betti, P.; Genovese, D.; Montalti, M.; Prodi, L.; De Zorzi, R.; Geremia, S.; Dalcanale, E. *Angew. Chem. Int. Ed.* **2011**, *50*, 4654.

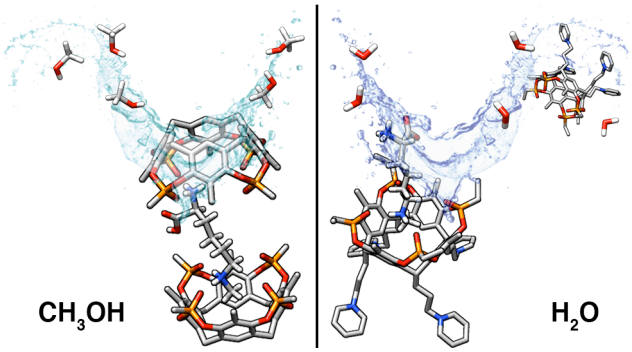
(25) Persch, E.; Dumele, O.; Diederich, F. *Angew. Chem. Int. Ed.* **2015**, *54*, 3290.

(26) Biedermann, F.; Vendruscolo, M.; Scherman, O. A.; De Simone, A.; Nau, W. M. *J. Am. Chem. Soc.* **2013**, *135*, 14879.

(27) For N_ε-Me-Lys·HCl, the first Ka of Table 3 was taken into account, since it corresponds to the binding event of the N_ε-Me residue.

(28) Alessandri, I.; Biavardi, E.; Gianoncelli, A.; Bergese, P.; Dalcanale, E. *ACS Appl. Mater. Interfaces* **2016**, *8*, DOI: 10.1021/acsami.5b08190.

TOC



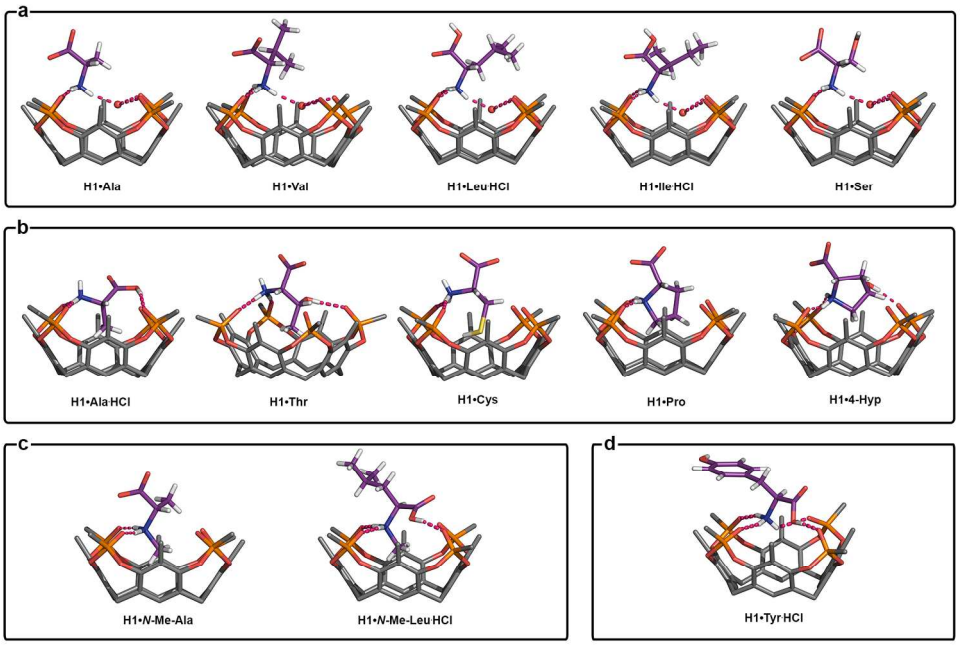


Figure 1. Side views of the molecular structures of the thirteen **H1•AA** complexes. The dashed lines depicted in magenta indicate hydrogen bonding interactions.
180x119mm (300 x 300 DPI)

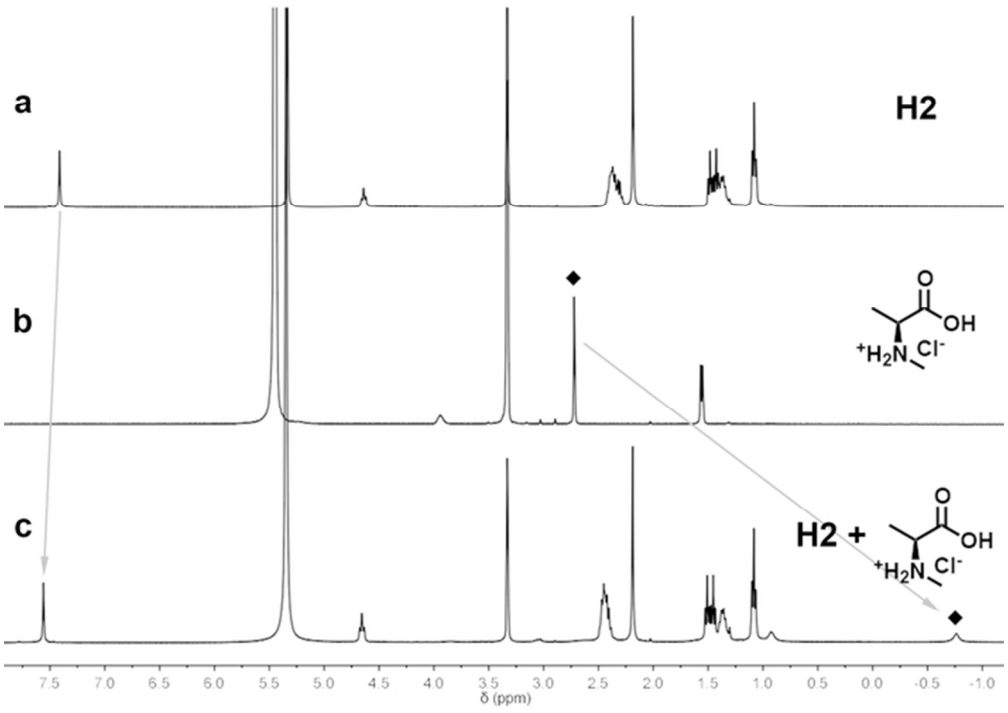


Figure 2. ^{1}H NMR (400 MHz, MeOD, 253 K) spectra of free host **H2** (a); free *N*-Me-Ala·HCl guest (b); 1:1 **H2•*N*-Me-Ala·HCl** complex (c). Signal related to *N*-methyl protons of the guest is labeled (■).
59x42mm (300 x 300 DPI)

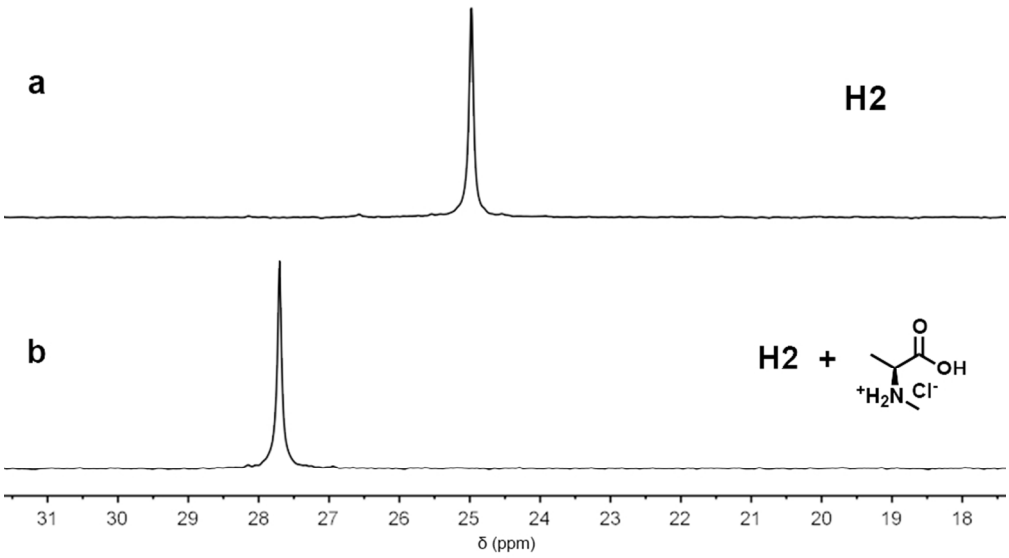


Figure 3. ^{31}P NMR (162 MHz, MeOD, 253 K) spectra of free host **H2** (a) and 1:1 **H2•*N*-Me-Ala·HCl** complex (b).
85x50mm (300 x 300 DPI)

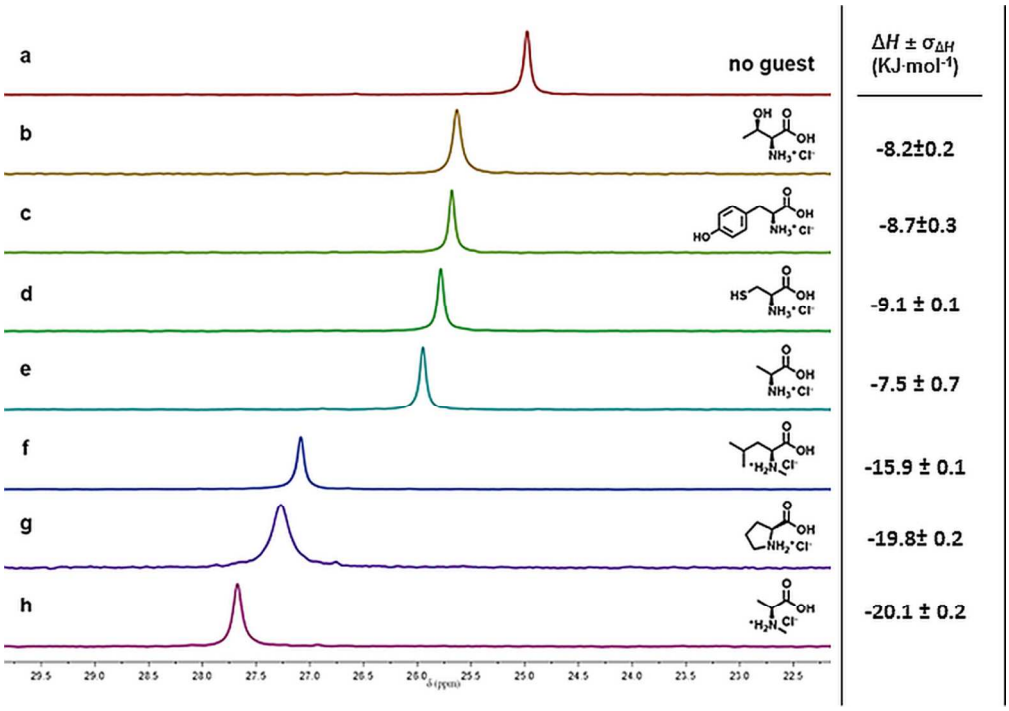


Figure 4. ^{31}P NMR (162 MHz, MeOD, 253 K) spectra of free host **H2** (a), **H2•Thr·HCl** complex (b), **H2•Tyr·HCl** complex (c), **H2•Cys·HCl** complex (d), **H2•Ala·HCl** complex (e), **H2•*N*-Me-Leu·HCl** complex (f), **H2•Pro·HCl** complex (g), **H2•*N*-Me-Ala·HCl** complex (h). For all the complexes a 1:1 stoichiometry was respected.
84x59mm (300 x 300 DPI)

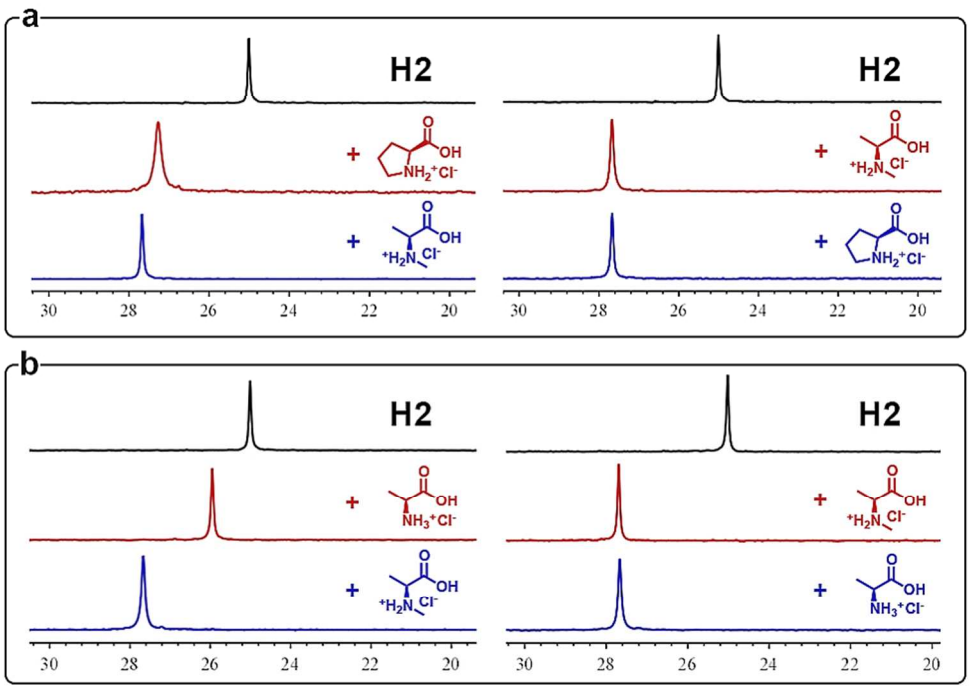


Figure 5. ^{31}P NMR (162 MHz, MeOD, 253 K) spectra: free host **H2** (black), 1:1 complex (red), addition of 1 eq of competitor to the formed complex (blue). The structure of the AA used are shown in the insets.
85x59mm (300 x 300 DPI)

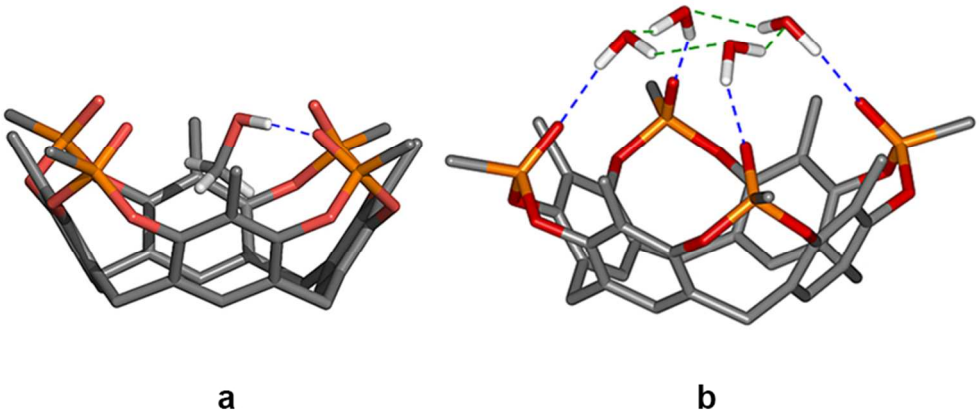


Figure 6. a) Side view of the crystal structure of **H1•MeOH**.[23] The dashed blue lines indicate hydrogen bonding between the cavitands and methanol. b) Side view of the crystal structure of **H1•4H$_2$O**.[21] The dashed lines depicted in blue indicate hydrogen bonding between the cavitand and water, while water-water interactions are depicted in green.
84x44mm (300 x 300 DPI)

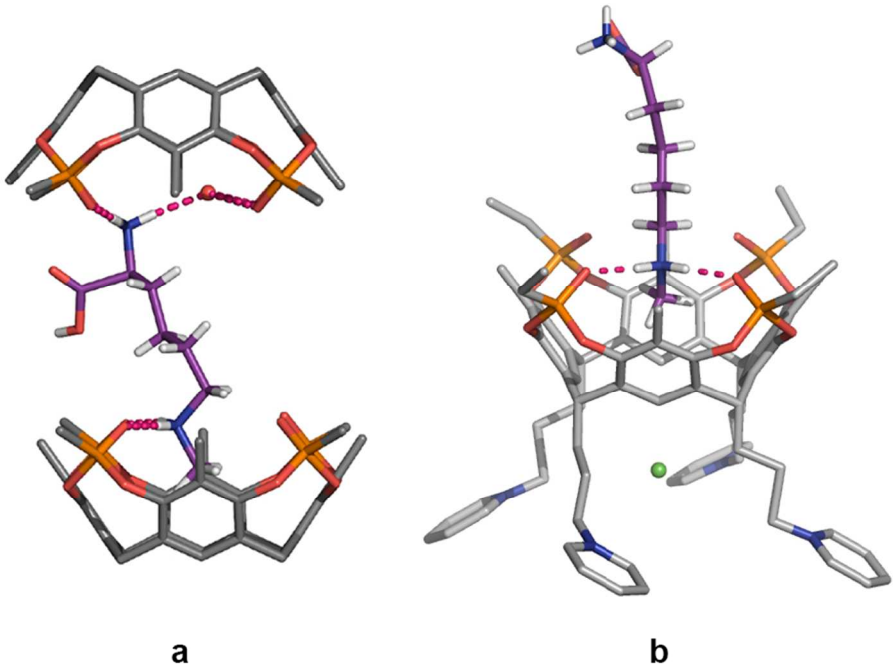


Figure 7. Side view of the molecular structures of **2H1•N_ε-Me-Lys·HCl** (a) and **H3•N_ε-Me-Lys·HCl** (b) complexes. The dashed lines depicted in magenta indicate hydrogen bonding interactions.
84x62mm (300 x 300 DPI)

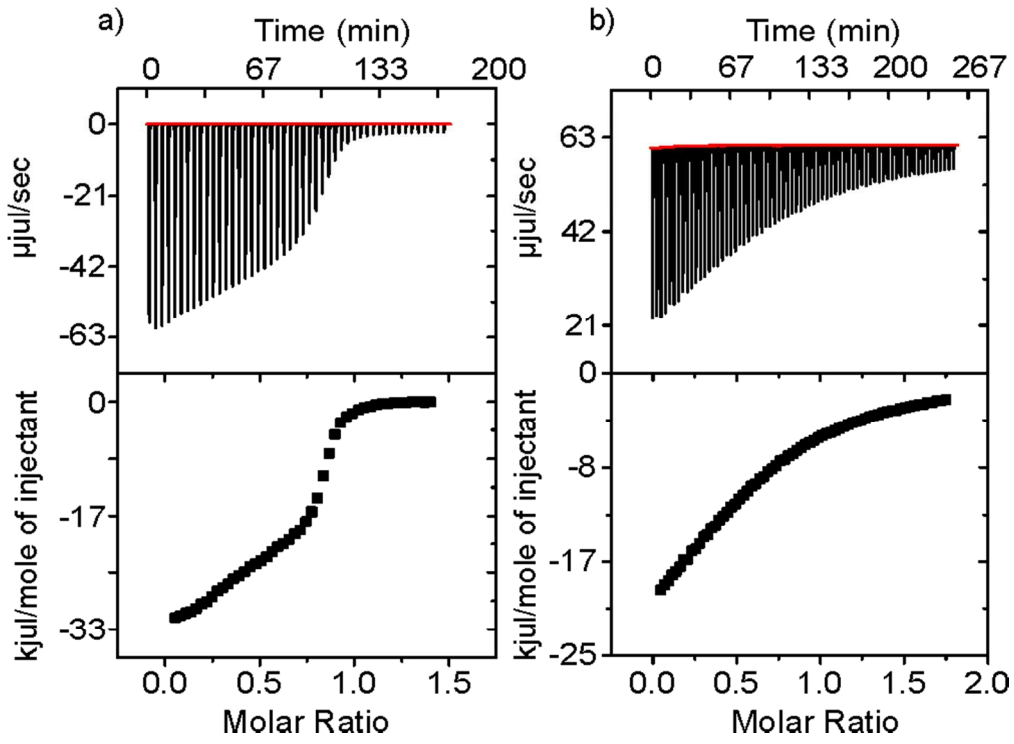


Figure 8. ITC experiments: a) methanol, **H2** = 1.04 mM, N_{ε}-Me-Lys·HCl = 6.74 mM; b) water, **H3** = 1.69 mM, N_{ε}-Me-Lys·HCl = 13.78 mM. The host was placed in the cell and the guest in the syringe. Top trace: raw data for the ITC titration. Bottom trace: binding isotherm of the integrated calorimetric titration data.

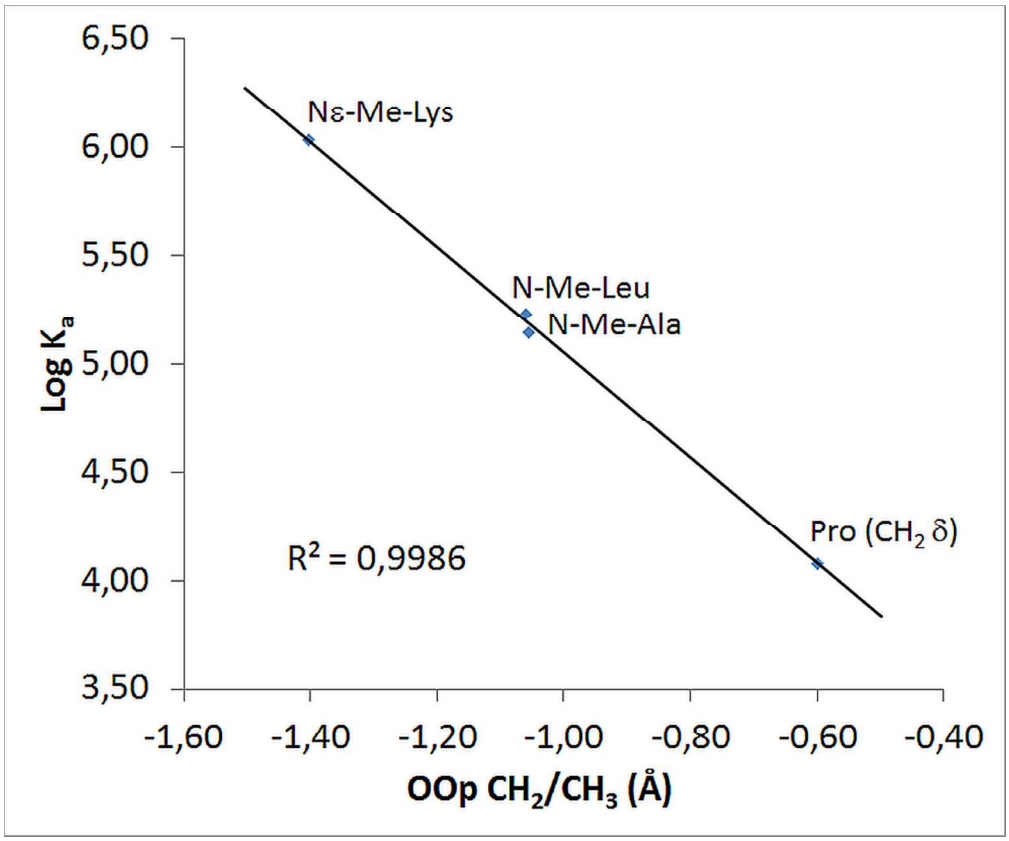


Figure 9. Linear regression curve showing the relationship between the thermodynamic association constant, logK_a (obtained from ITC experiments in MeOH) and the structural data of CH–π interactions (evaluated as distance [Å] of the carbon atom of the N–CH_x group from the entrance of cavity defined by mean plane of the oxygen atoms of the P=O groups, OOp).
83x69mm (300 x 300 DPI)

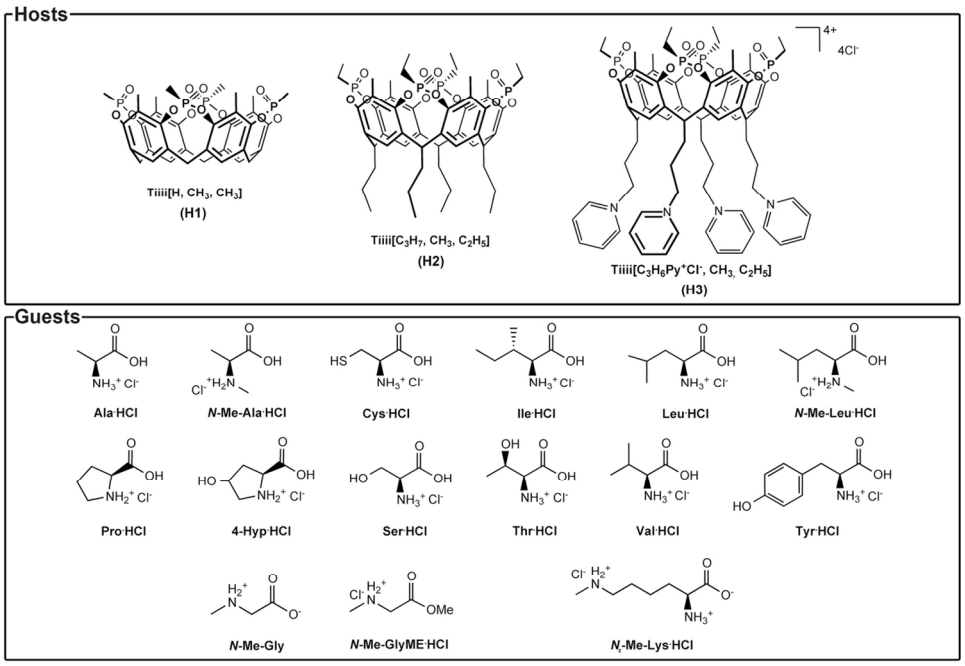
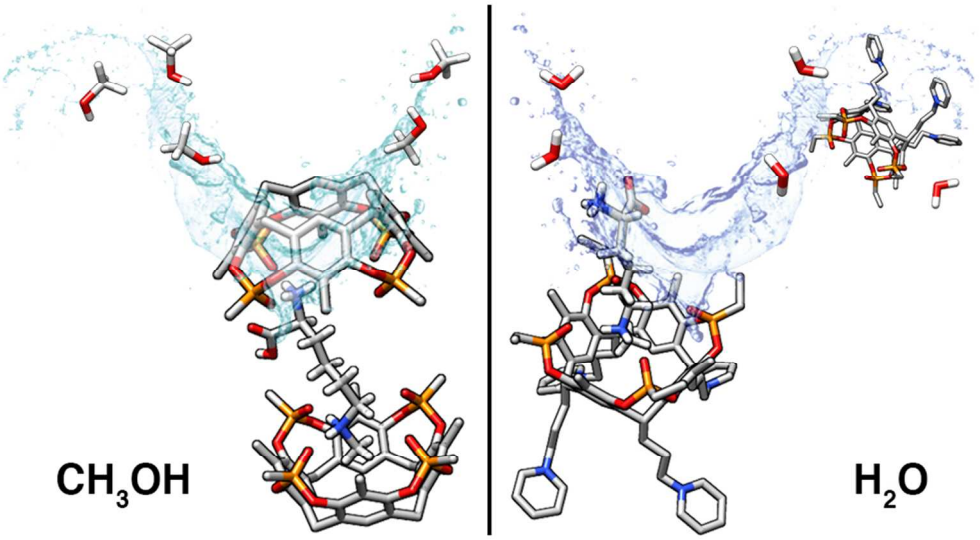


Chart 1. Cavitands and amino acids (represented as hydrochloride salts) involved in this study.
119x80mm (300 x 300 DPI)



TOC
85x47mm (300 x 300 DPI)