Electronic Supplementary Material (ESI) for Organic Chemistry Frontiers.

Supplementary Information

A biomimetic strategy for the selective recognition of organophosphates in 100% water: synergies of electrostatic interactions, cavity embedment and metal coordination

Solène Collin, Nicolas Giraud, Elise Dumont and Olivia Reinaud

Table of contents

Characterization

Hexyl mono-phosphate CnP

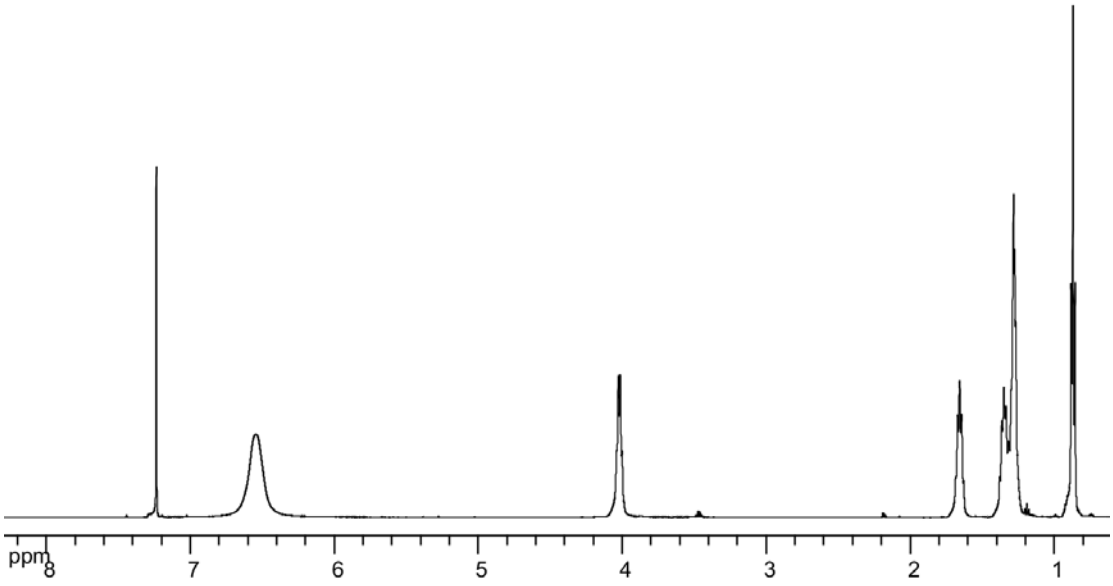


Figure S1. ^{1}H NMR (500 MHz, $CDCl_3$, 300K) of **C6P**. δ (ppm): 6.55 (OH, s, 4H), 4.02 (PO–$\underline{CH_2}$, m, 2H), 1.66 (PO–CH_2–$\underline{CH_2}$, m, 2H), 1.31 ($\underline{(CH_2)_3}$-CH_3, m, 6H), 0.87 (CH_3, t, 3H, J = 7.0 Hz).

WRim$_4$[1]

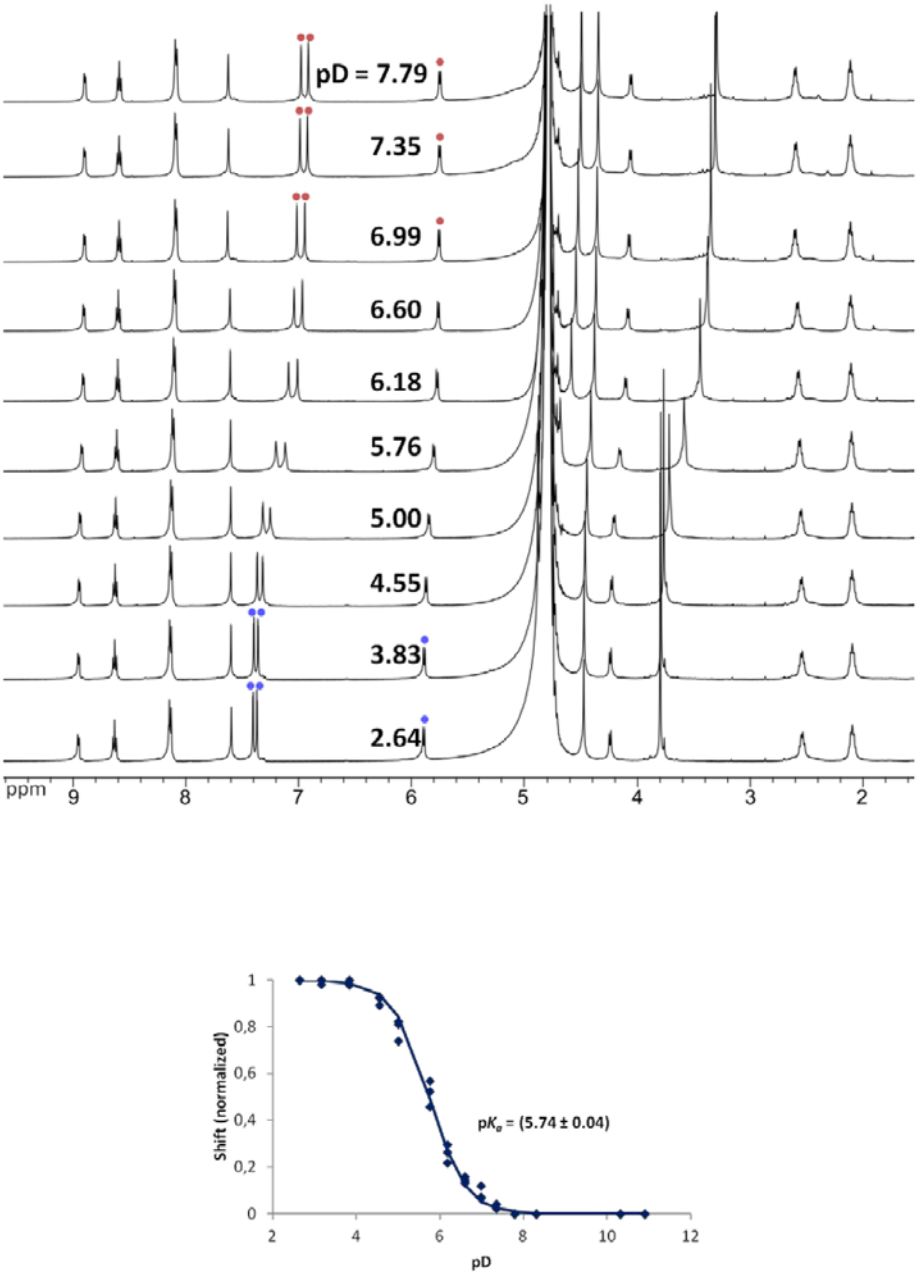


Figure S2. [1] pH study of **WRim$_4$** (3.5 mM, ^{1}H NMR (500 MHz, 300 K, D_2O)). The pD was adjusted with NaOH and HNO_3 0.2 M solutions in D_2O. Top: full ^{1}H NMR spectra. Red dots: fully protonated ligand. Green dots: unprotonated ligand. Bottom: variation of the chemical shifts of the dotted signals according to pH, yielding an average pK_a value of 5.74.

[1] S. Collin, A. Parrot, L. Marcelis, E. Brunetti, I. Jabin, G. Bruylants, K. Bartik, and O. Reinaud, *Chem. Eur. J.* **2018**, *24*, 17964 –17974.

$WRim_4Zn$(C6P)

The complex was generated *in situ* by addition of 1 equiv. zinc nitrate and **C6P** (monohexylphosphate, typically 1.8 equiv.) to a 1 mM solution of **$WRim_4$** in D_2O and characterized by 1D and 2D NMR experiments.

^{1}H NMR (500 MHz, D_2O, pD = 7.4, 300 K) δ (ppm): 8.92 (Pyridine (*ortho*), m, 8H), 8.61 (Pyridine (*para*), t, 4H, *J* = 10.0 Hz), 8.11 (Pyridine (*meta*), m, 8H), 7.65 ($H_{ar,\ down}$, s, 2H), 7.63 ($H_{ar,\ down}$, s, 2H), 7.36 ($H_{Im,\beta}$, s, 2H) , 7.30 ($H_{Im,\alpha}$, s, 2H) , 7.24 ($H_{Im,\beta}$, s, 2H) , 7.06 ($H_{Im,\alpha}$, s, 2H), 6.67 (O–CH_2–O, d, 1H, *J* = 7.0 Hz), 5.70 (O–CH_2–O, m, 4H), 5.05 (s, 1H), 5.02 (s, 1H), 4.74-4.54 (Ar–$\underline{CH_2}$–O, O–$\underline{CH_2}$–Im, m, 10H), 4.29 (Ar–$\underline{CH_2}$–O, O–$\underline{CH_2}$–Im, dd, 2H), 3.89-4.12 (O–CH_2–Im, Ar–CH_2–O, O–CH_2–O, m, 3H), 3.69 (N–CH_3, s, 6H), 3.64 (N–CH_3, s, 6H), 3.44 (C6P (C n°1), m, 2H), 2.62 (CH_2–$\underline{CH_2}$–CH_2, m, 8H), 2.12 (CH–$\underline{CH_2}$, m, 8H), 0.62 (C6P (C n°2), m, 2H), -0.32 (C6P (C n°3-4), m, 4H), -0.62 (C6P (C n°5), m, 2H), -3.27 (C6P (C n°6), t, 3H, *J* = 7.5 Hz).
(The α and β position of the imidazole protons refer to the nitrogen lone pair.)

^{13}C NMR (150 MHz, D_2O, pD = 7.4, 300 K) δ (ppm): 154.0, 153.7, 153.6, 153.3, 145.8, 144.2, 137.5, 128.4, 126.0, 125.8, 123.4, 123.3, 122.9, 122.7, 65.4, 63.3, 62.8, 61.4, 36.4, 33.1, 32.6, 30.9, 30.4, 30.1, 28.8, 25.9, 24.7, 23.5, 22.0, 21.2, 12.9, 7.1.

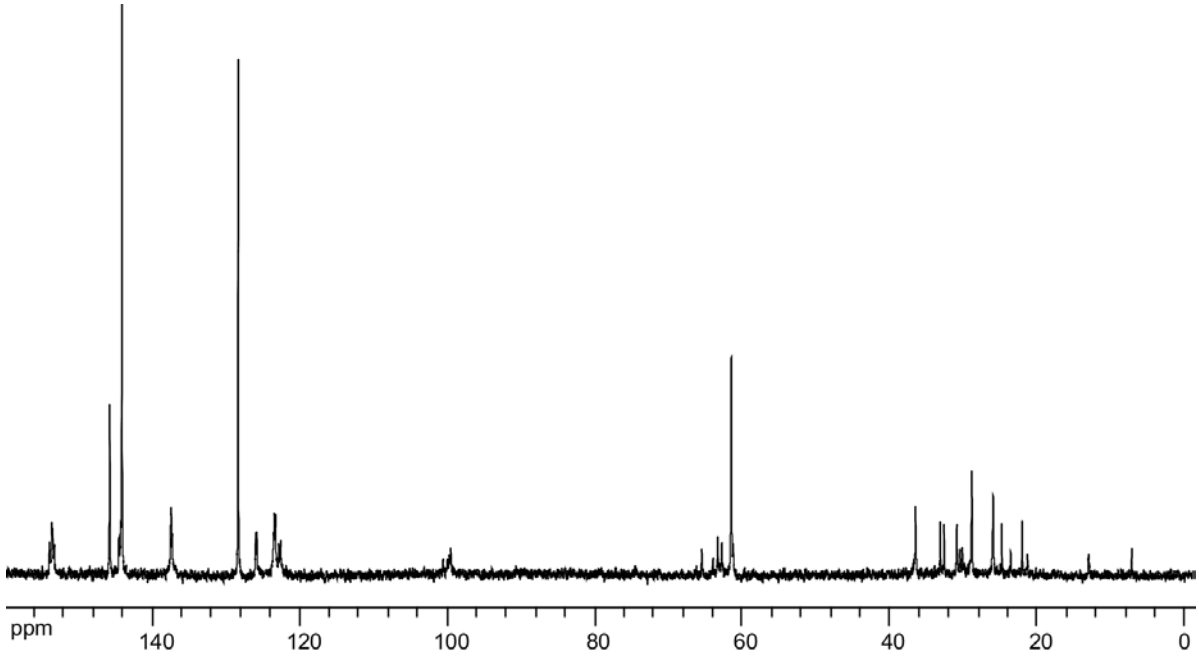


Figure S3. ^{13}C NMR (150 MHz, D_2O, pD = 7.4, 300 K) of **$WRim_4Zn$**(**C6P**).

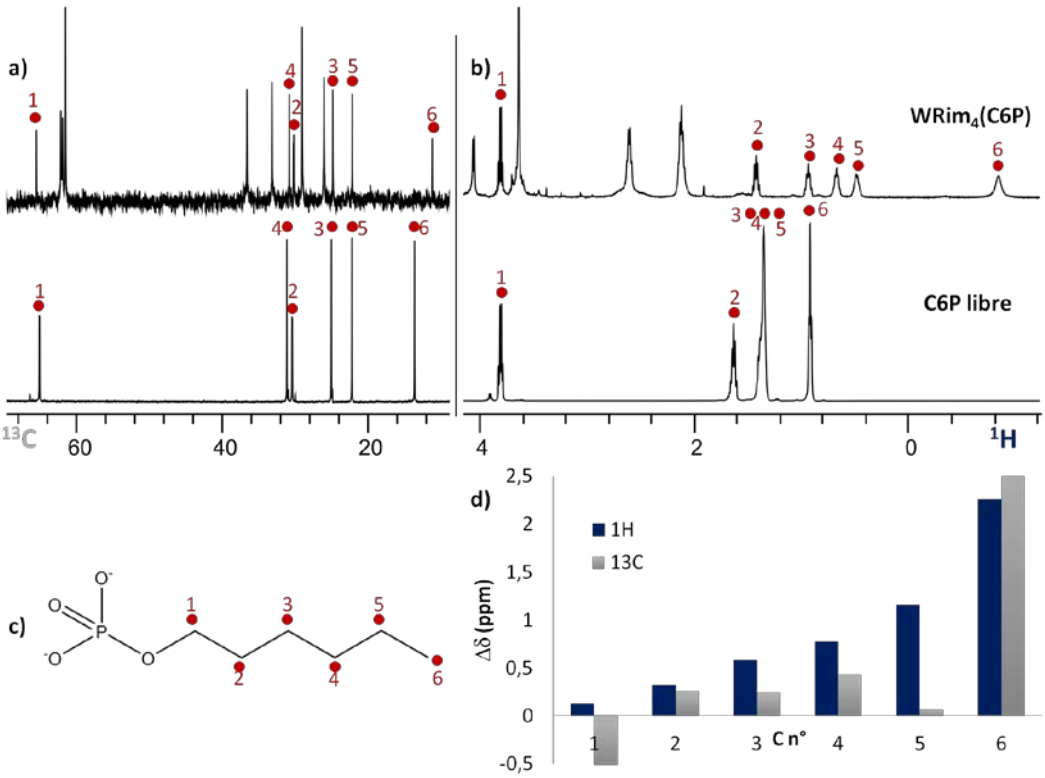


Figure S4. Part of the NMR spectra of **WRim$_4$Zn(C6P)** (pH 6.4) compared to free **C6P** (pH 7.4)**.** Conditions: **WRim$_4$** (3.5 mM), **C6P** 1.8 equiv. pD = 6.4 (top). a) ^{13}C (D_2O, 300 K, 125 MHz) ; b) NMR ^{1}H (D_2O, 300 K, 500 MHz) ; c) Structure of **C6P**; d) Complexation induced shifts for **C6P**.

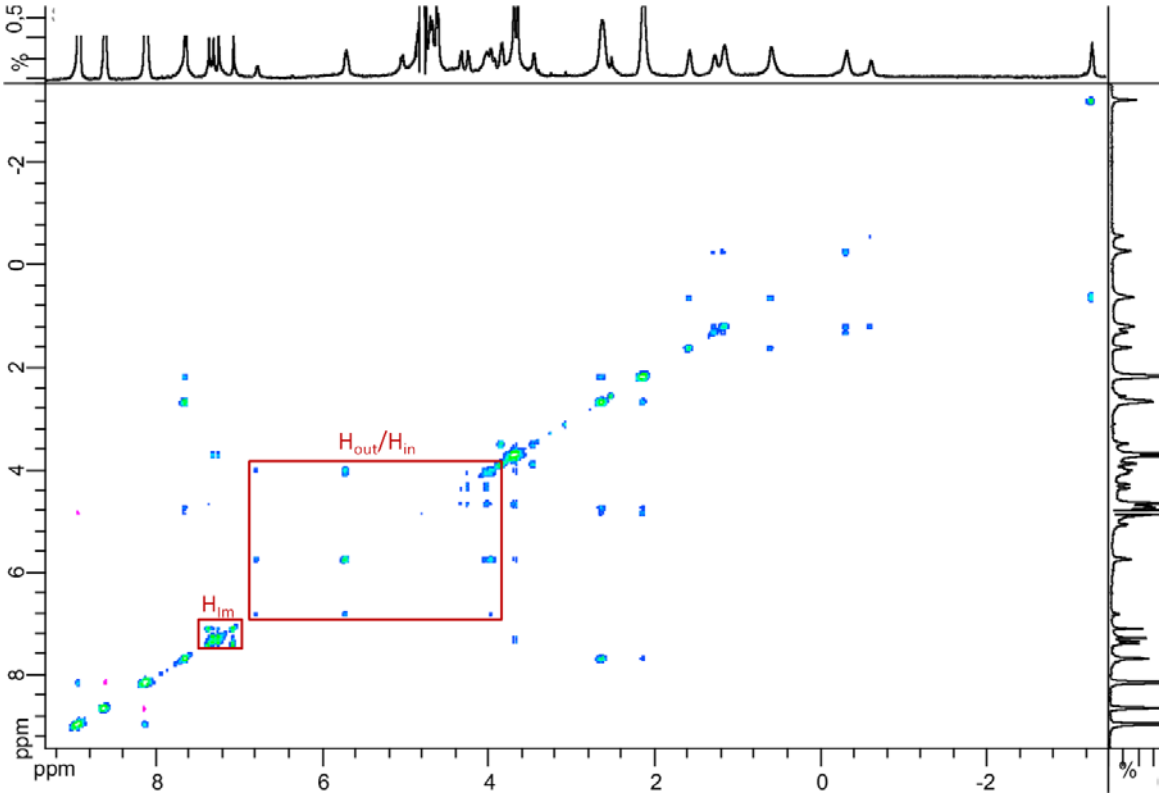


Figure S5. Full NOESY (D_2O, 300 K, 500 MHz) spectrum of **WRim$_4$Zn** (3.5 mM) with 1.8 equiv. **C6P** at pD = 7.4. The red rectangles indicate the regions presented in **Figure S6**. The cross peaks that are observed on this map correspond to either nOe transfer, or chemical exchange.

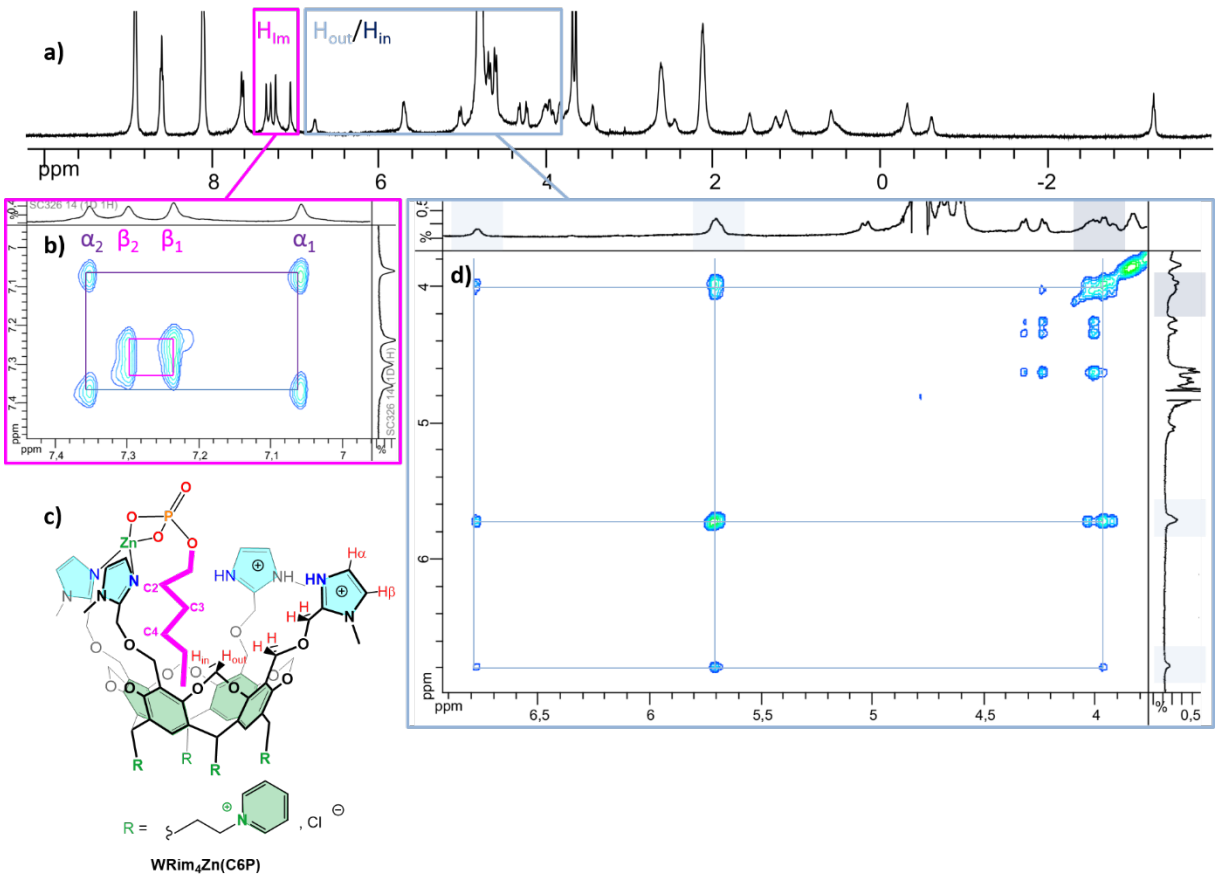


Figure S6. Extracts of the NOESY spectrum displayed in Figure S5. a) Full ^{1}H spectrum; b) NOESY for imidazole signals ($H_{Im,\alpha}$ and $H_{Im,\beta}$), from 7 to 7.4 ppm; c) Structure of complex **WRim$_4$Zn(C6P)** (with non-realistic geometry) and selected protons; d) NOESY extract for methylene bridge protons (H_{in} and H_{out}), from 4 to 7 ppm.

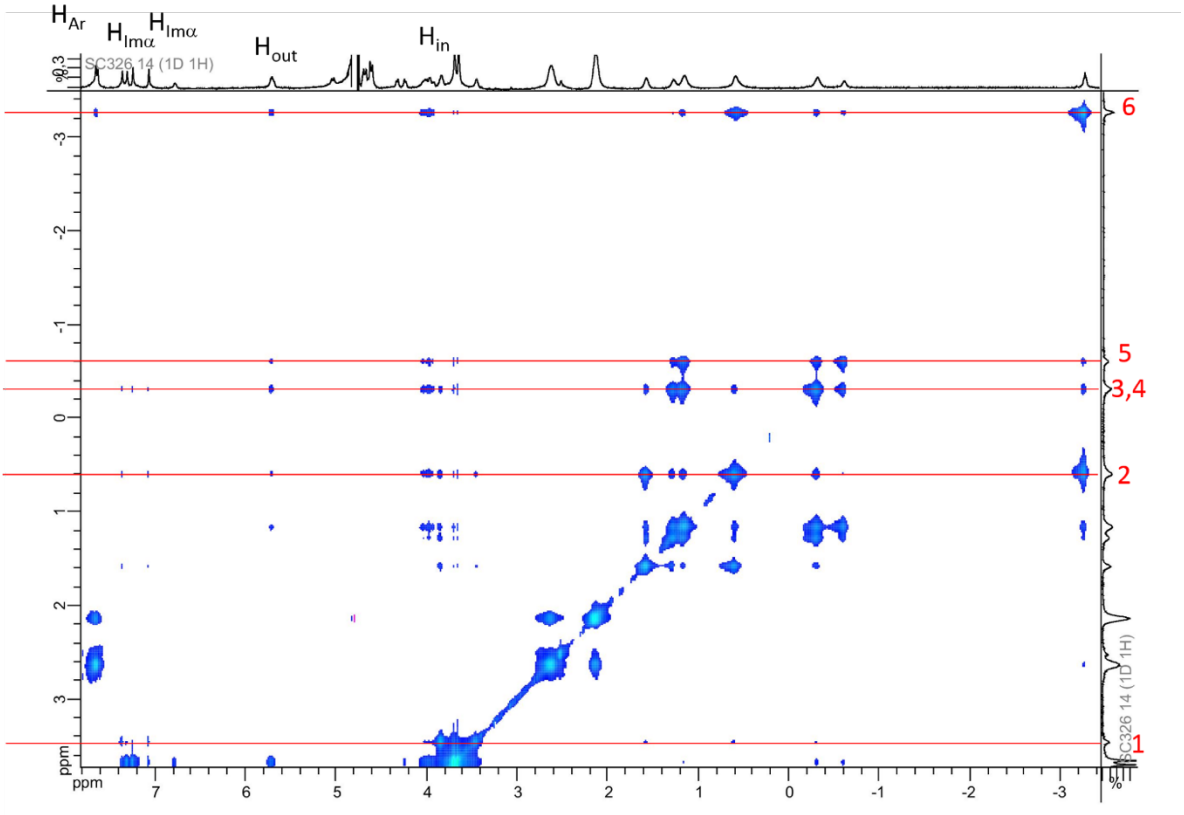


Figure S7. Extract of the NOESY spectrum displayed in **Figure S5** showing the correlation peaks between the host and the guest.

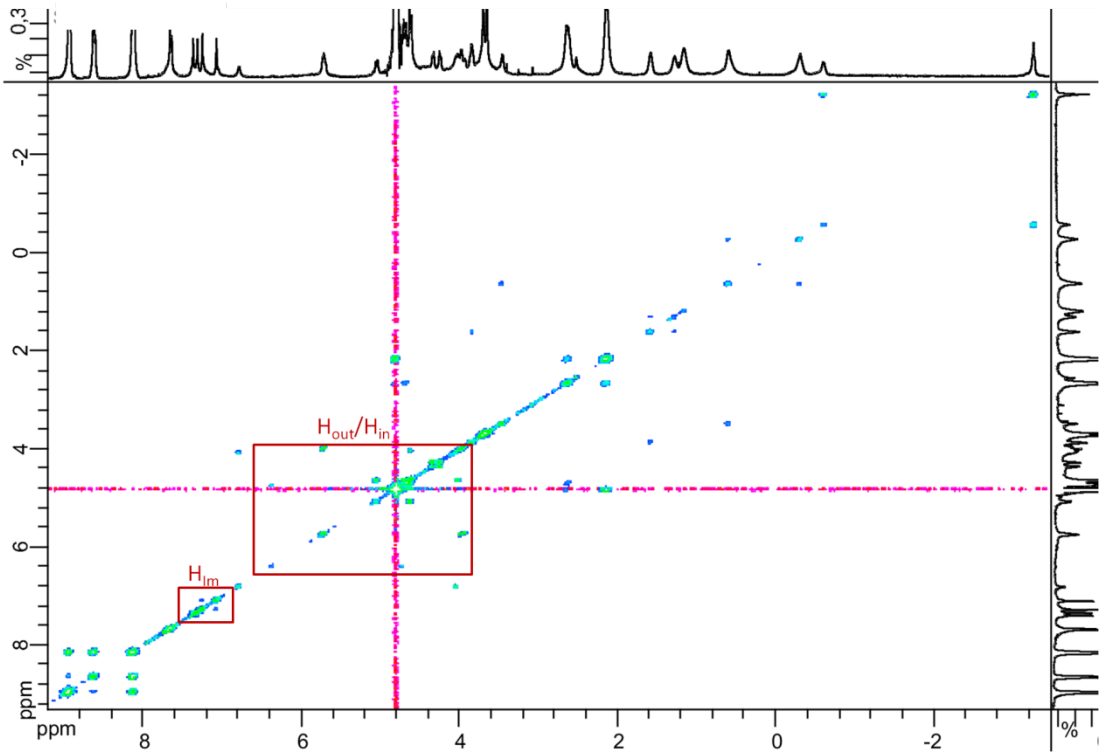


Figure S8. Full COSY (D_2O, 300 K, 500 MHz) spectrum of **WRim$_4$Zn** (3.5 mM) in presence of 1.8 equiv. **C6P** at pD = 7.4. The red rectangles indicate the regions presented in **S9** and **S10**.

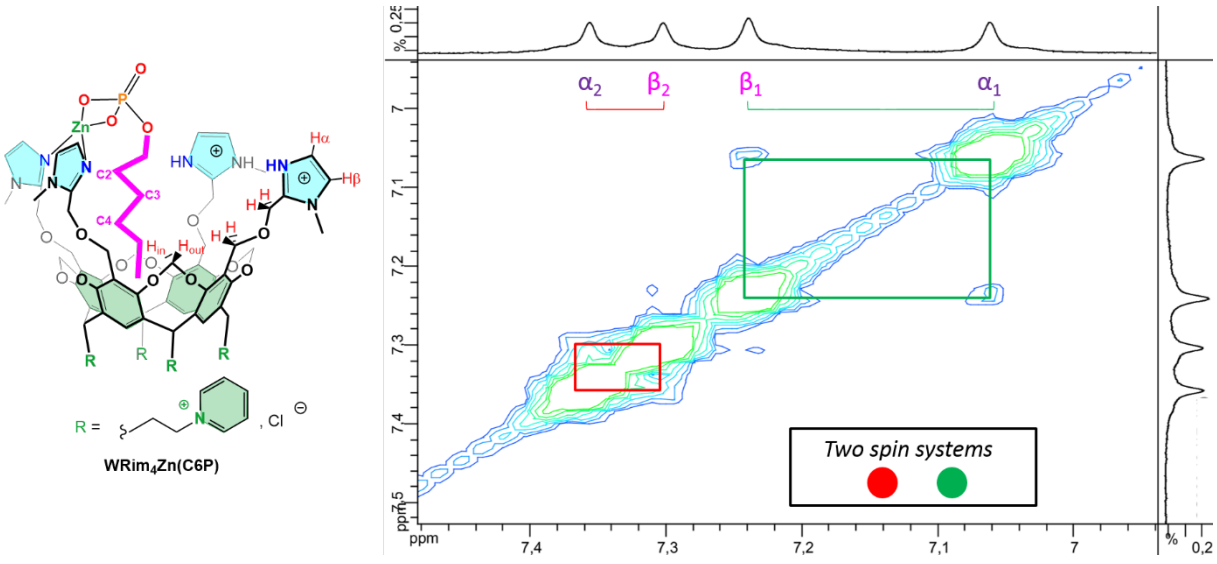


Figure S9. Extract of COSY spectrum of **WRim$_4$Zn** (3.5 mM) in presence of 1.8 equiv. **C6P** at pD = 7.4. ^{1}H NMR (D_2O, 300 K, 500 MHz). Two $H_{Im,\alpha/\beta}$ systems are identified.

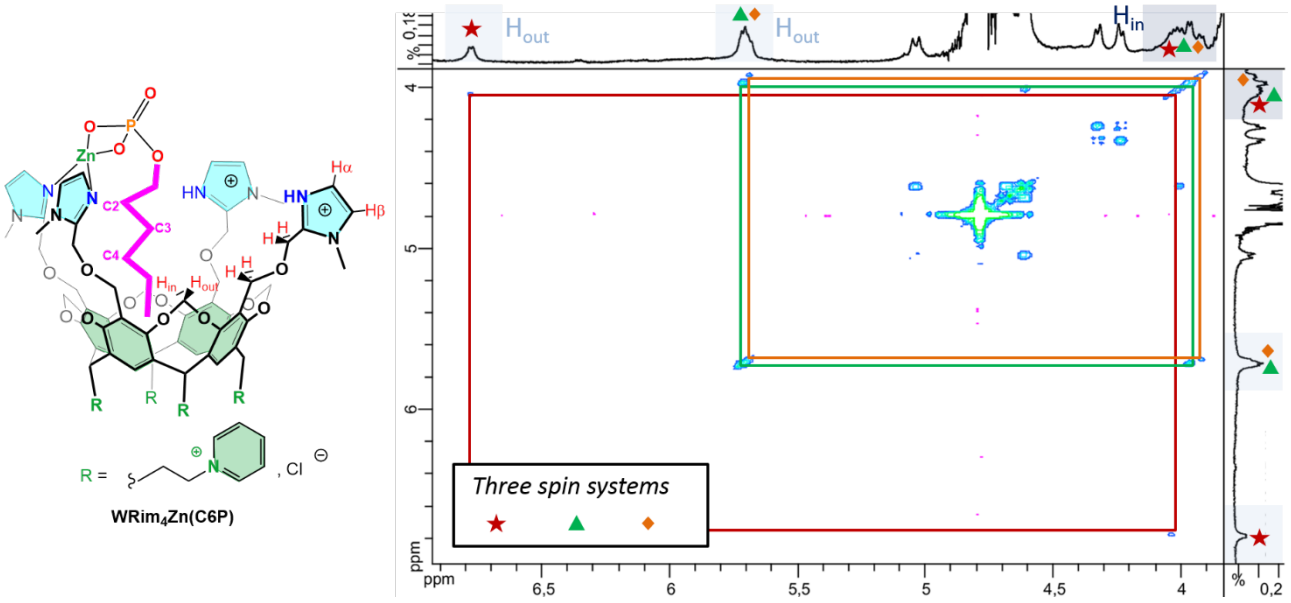


Figure S10. Extract of COSY spectrum of **WRim$_4$Zn** (3.5 mM) in presence of 1.8 equiv. **C6P** at pD = 7.4. ^{1}H NMR (D_2O, 300 K, 500 MHz). Three different H_{in}/H_{out} correlations appear (stars/triangles/diamonds).

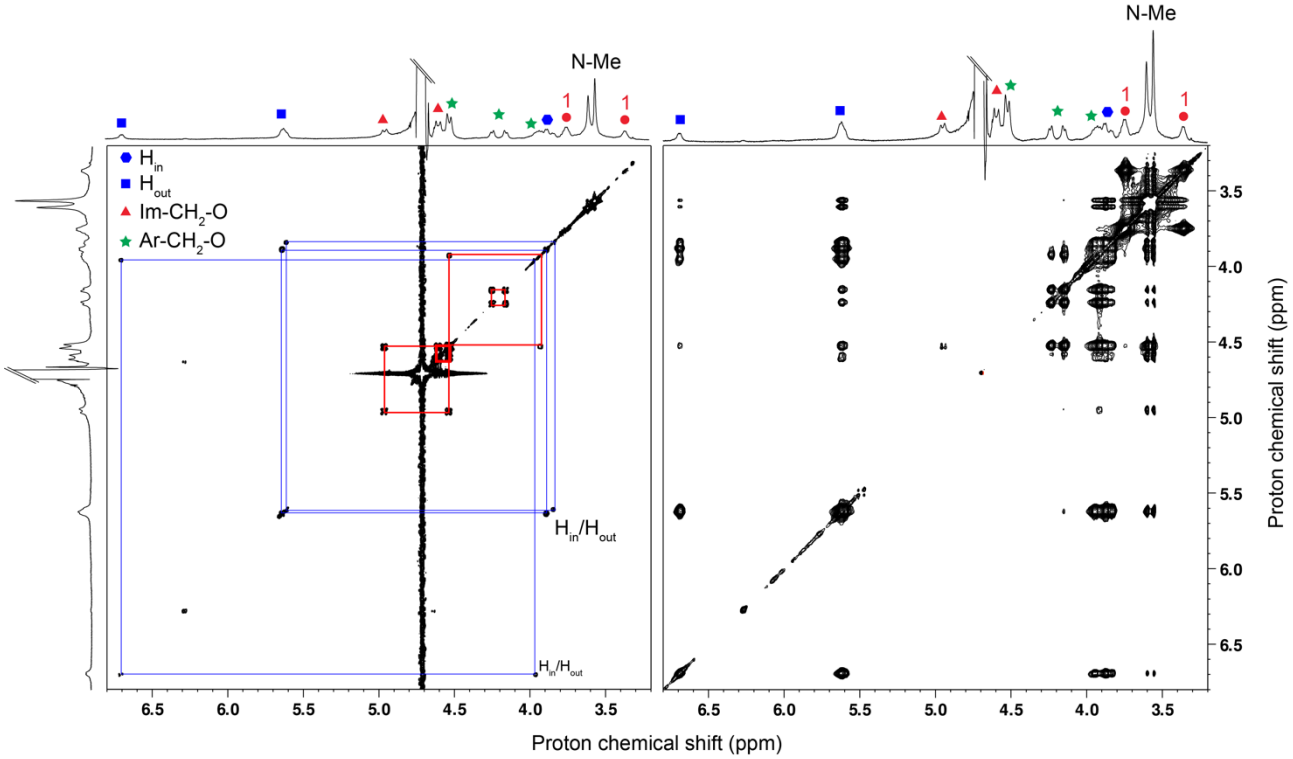


Figure S11. Extracts of the COSY (left) and NOESY (right) spectra of **WRim$_4$Zn** (3.5 mM) in presence of 1.8 equiv. **C6P** at pD = 7.4. The correlations allowing for the assignment of the different CH_2 groups from the ligand WRim$_4$Zn are highlighted on the COSY spectrum.

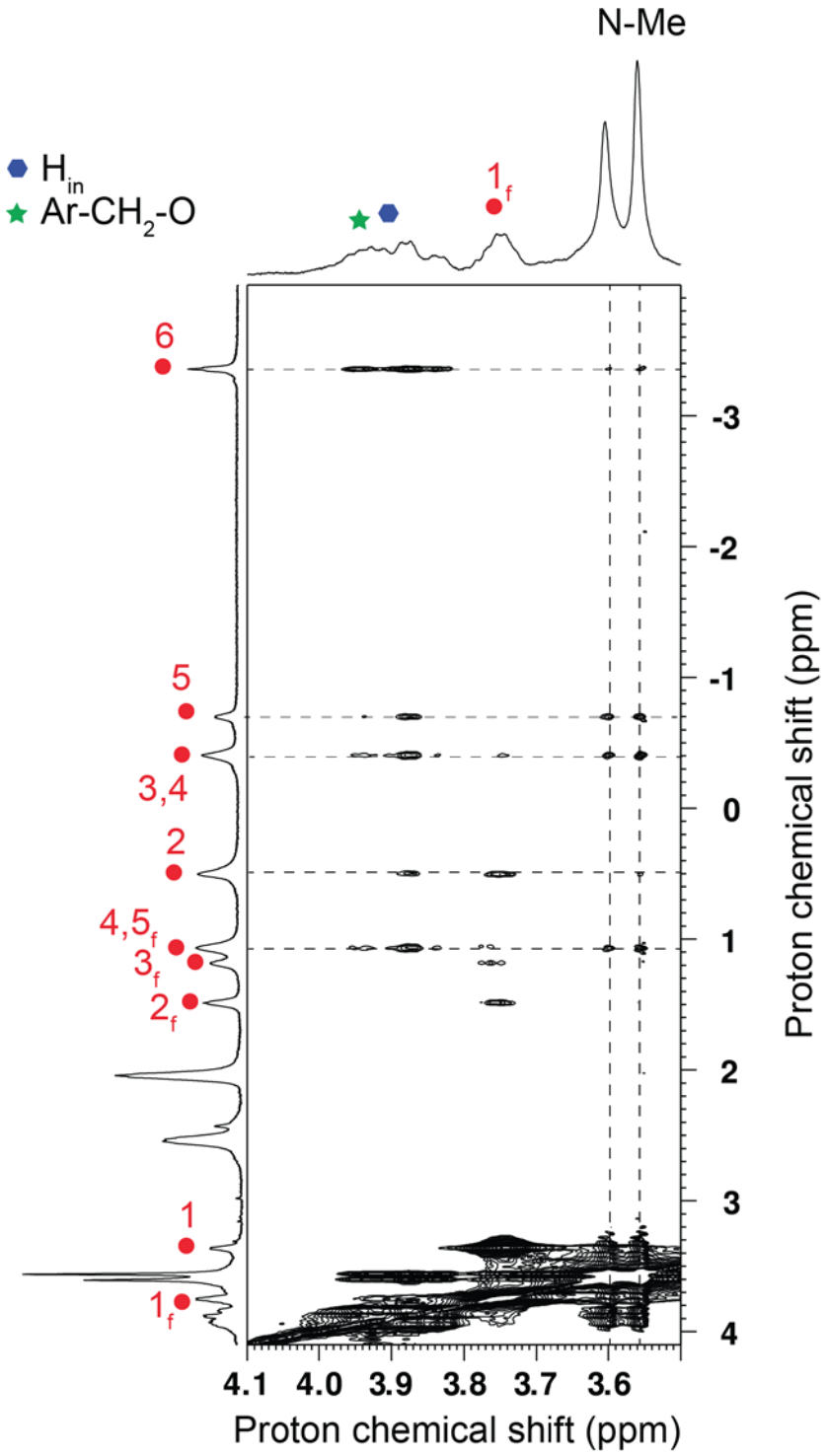


Figure S12. Extract of the NOESY spectrum of **WRim$_4$Zn** (3.5 mM) in presence of 1.8 equiv. **C6P** at pD = 7.4, showing the nOe correlations between protons from **C6P** and *N*-Me protons from **WRim$_4$Zn**, respectively.

$WRim_4 \supset C6P$

The adduct was generated by addition *in situ* of C6P (typically 1.8 equiv.) to a solution of **$WRim_4$** (1 mM) in D_2O and characterized by NMR 1D and 2D experiments.

^1H NMR (500 MHz, D_2O, pD = 6.4, 300 K) δ (ppm): 8.93 (Pyridine (*ortho*), m, 8H), 8.61 (Pyridine (*para*), m, 4H), 8.12 (Pyridine (*meta*), m, 8H), 7.64 ($H_{ar,\ down}$, s, 4H), 7.25 ($H_{Im,\beta}$, s, 4H), 7.16 ($H_{Im,\alpha}$, s, 4H), 5.81 (O–CH_2–O, d, 4H, *J* = 7.5 Hz), 4.71 (Im–CH_2–O, s, 8H), 4.37 (Ar–CH_2–O, s, 8H), 4.05 (O–CH_2–O, d, 4H, *J* = 7.5 Hz), 3.80 (C6P (C n°1), q, 2H, *J* = 7.0 Hz), 3.63 (N–CH_3, s, 12H), 2.60 (CH_2–$\underline{CH_2}$–CH_2, m, 8H), 2.11 (CH–$\underline{CH_2}$, m, 8H), 1.41 (C6P (C n°2), m, 2H), 0.93 (C6P (C n°3), m, 2H), 0.66 (C6P (C n°4), m, 2H), 0.48 (C6P (C n°5), m, 2H), -0.84 (C6P (C n°6), bs, 3H).

^{13}C NMR (150 MHz, D_2O, pD = 6.4, 300 K) δ (ppm): 153.6 ($\underline{C_{Ar}}$-O); 145.9 (pyridine); 144.2 (pyridine); 143.6 ($\underline{C_{Ar}}$-N); 137.7 ($\underline{C_{Ar}}$-CH); 128.4 (pyridine); 123.6 ($\underline{C_{Ar}}$-C_{Ar}-O); 123.4 ($C_{Im,\alpha}$); 122.5 ($C_{Ar,down}$); 122.5 ($C_{Im,\beta}$); 99.8 (C_{bridge}); 65.5 (C6P, C1), 62.1 (O-$\underline{C}H_2$-C_{Ar}-N); 61.9 (CH_2N^+, C_{Ar}-$\underline{C}H_2$-O); 36.5 (CH); 33.0 (NCH_3); 30.6 (C6P, C4); 30.1 (CH$\underline{C}H_2$); 30.0 (C6P, C2); 28.9 ($CH_2\underline{C}H_2CH_2$); 24.7 (C6P, C3); 22.0 (C6P, C5); 11.0 (C6P, C6).

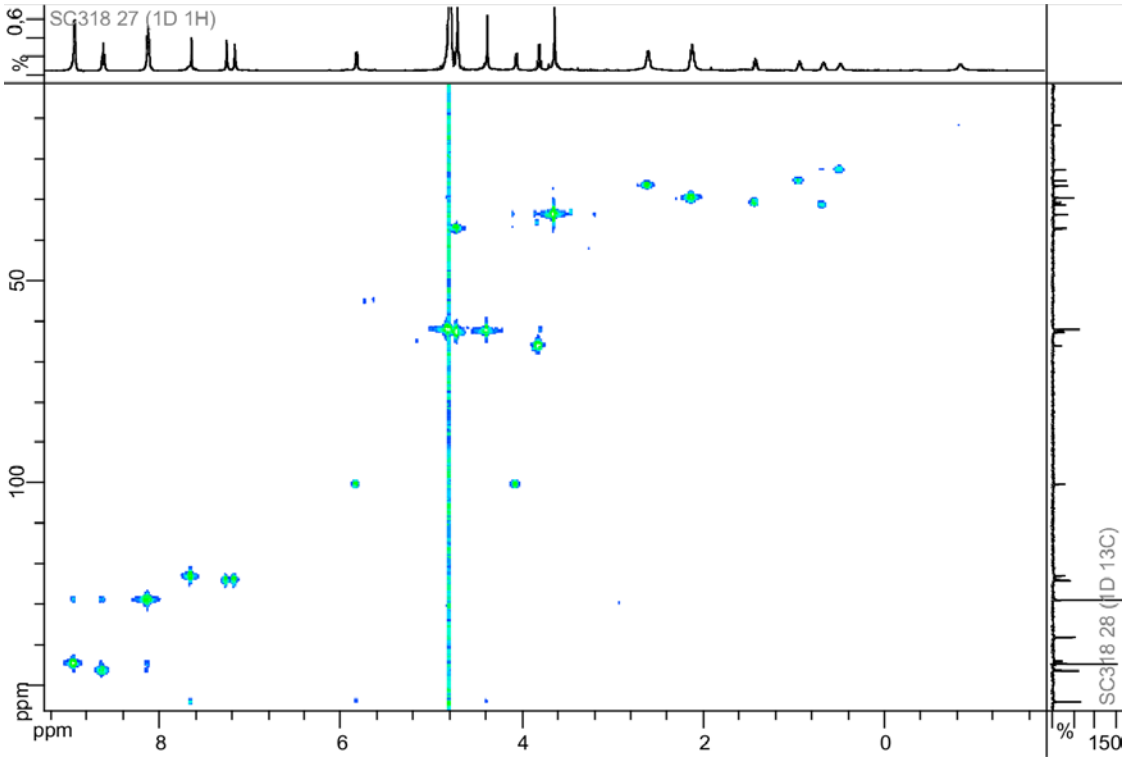


Figure S13. HSQC spectrum (D_2O, 300 K, 500 MHz) of **$WRim_4$** (3.5 mM) with 1.8 equiv. **C6P**, pD = 6.4.

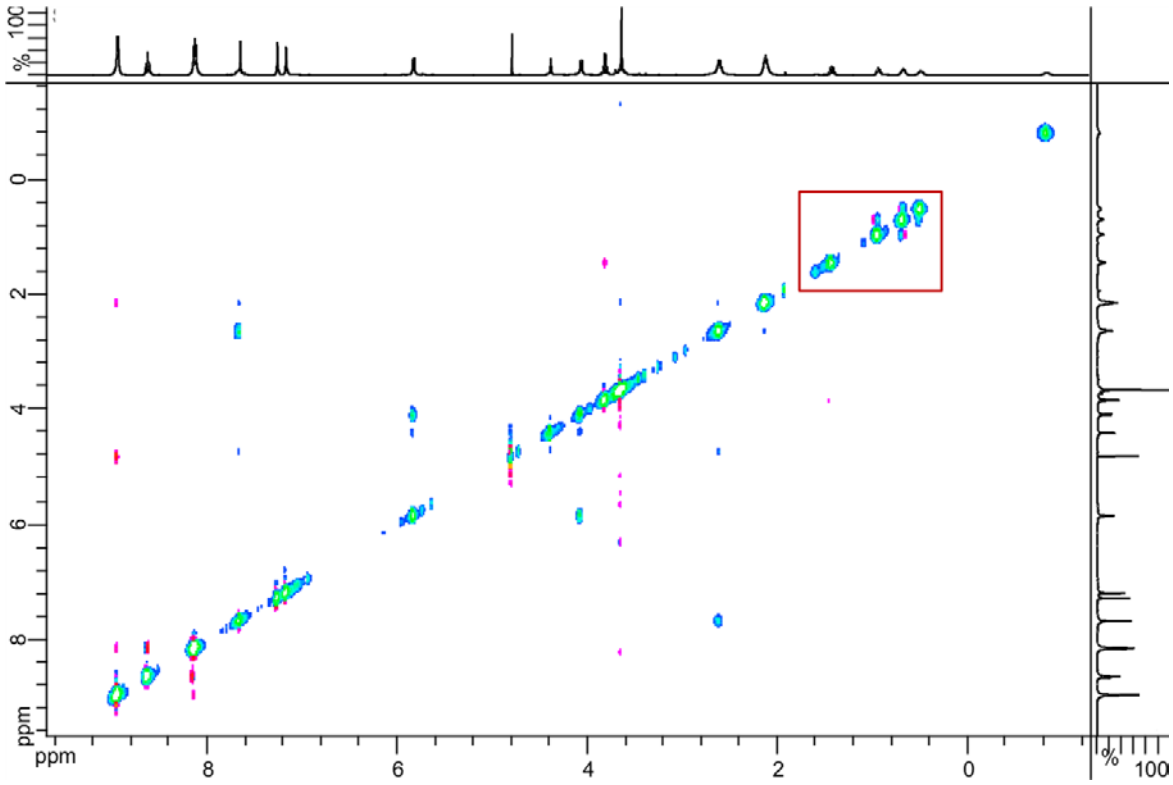


Figure S14. NOESY spectrum (D_2O, 300 K, 500 MHz) of **WRim$_4$** (3.5 mM) with 1.8 equiv. **C6P**, pD = 6.4.

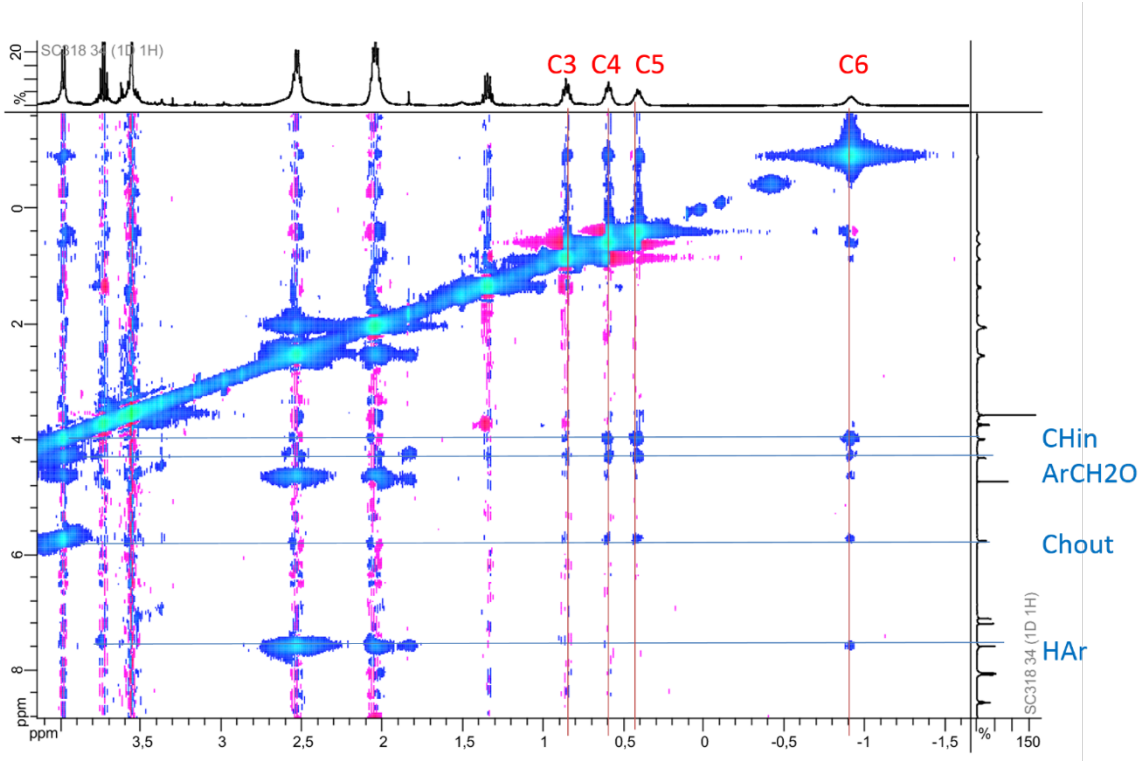


Figure S15. Extract of the NOESY spectrum displayed in Figure S14.

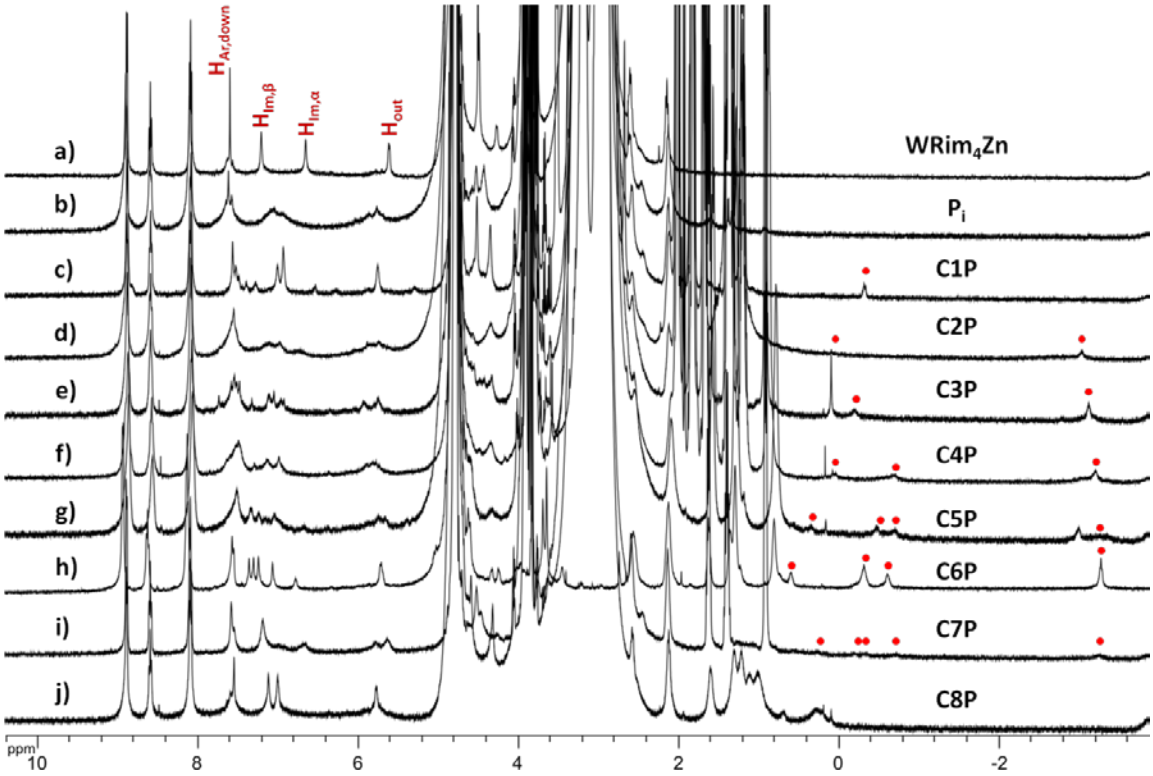


Figure S16. (Full spectra corresponding to Figure 2 in the article). ^{1}H NMR spectra (300 K, 500 MHz) of solutions of complex **WRim$_4$Zn** (1 mM in D_2O, pD = 7.4, HEPES 100 mM) to which various monoalkyl phosphates **CnP** (n = 0 – 8) have been added. a) **WRim$_4$Zn** (1 mM); b) + 9.8 equiv. $H_2PO_4^-$; c) + 11.3 equiv. **C1P**; d) + 15.2 equiv **C2P**; e) + 10.5 equiv. **C3P**; f) + 7.4 equiv. **C4P** ([**WRim$_4$Zn**] = 2 mM); g) + 6.8 equiv. **C5P**; h) + 2.5 equiv. **C6P** ([WRim$_4$Zn] = 3.5 mM, no buffer, pD = 7.4); i) + 1.5 equiv. **C7P**; j) + 3 equiv. **C8P**. The red dots indicate the signals attributed to the alkyl chain of the guest phosphates embedded into the resorcinarene cavity.

ITC titration with C4P

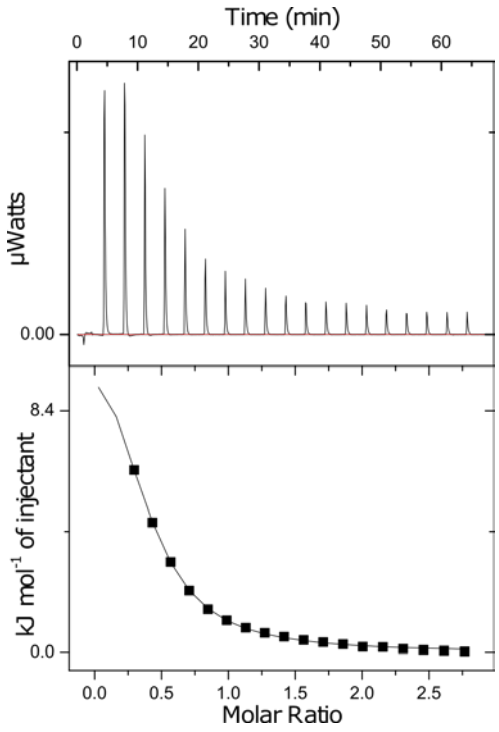


Figure S17. ITC titration of **C4P** (15 mM) by **WRim$_4$Zn** (1 mM) at 293 K, with HEPES 100 mM buffer (pH = 7). The thermodynamic parameters are: K' = (5.0 ± 1.4) x 10^3 M^{-1}, $\Delta H^{0\prime}$ = 11.5 ± 0.2 kJ/mol, $\Delta S^{0\prime}$ = 110 ± 3 $J.K^{-1}.mol^{-1}$, n = 0.69± 0.21.

Full NMR spectra of pH variation studies

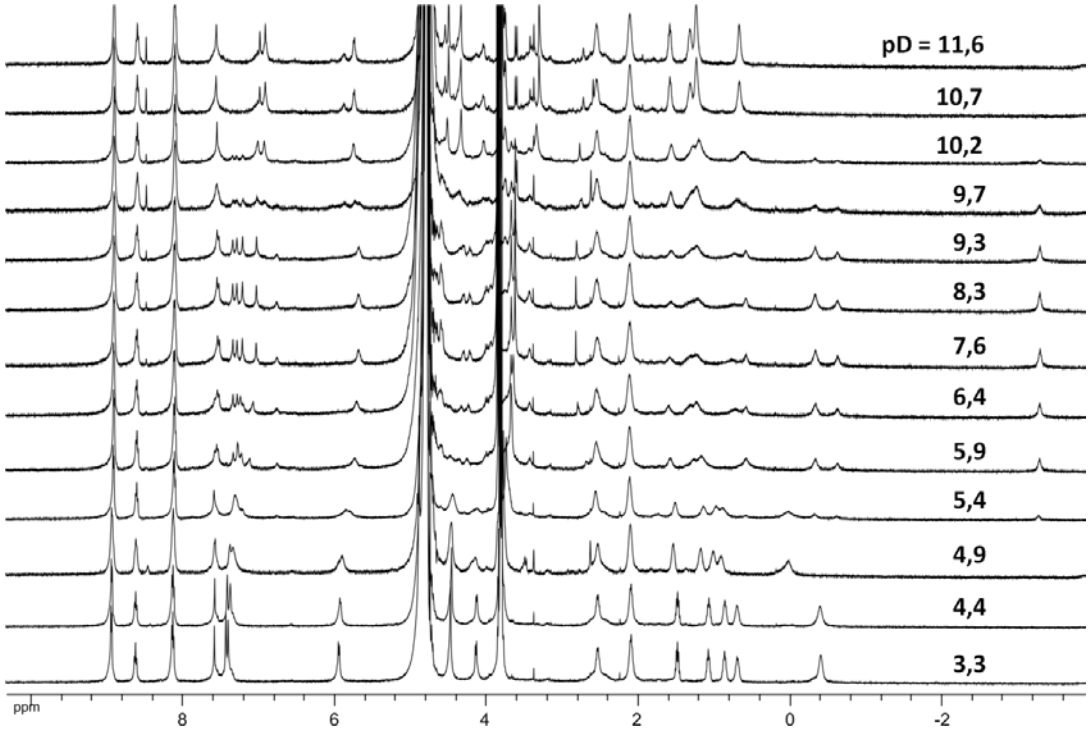


Figure S18. Full ^{1}H NMR spectra (D_2O, 300 K, 500 MHz) for the pD variation study of **WRim$_4$Zn** (3.5 mM) with 1.8 equiv. **C6P**.

Full NMR spectra of titrations

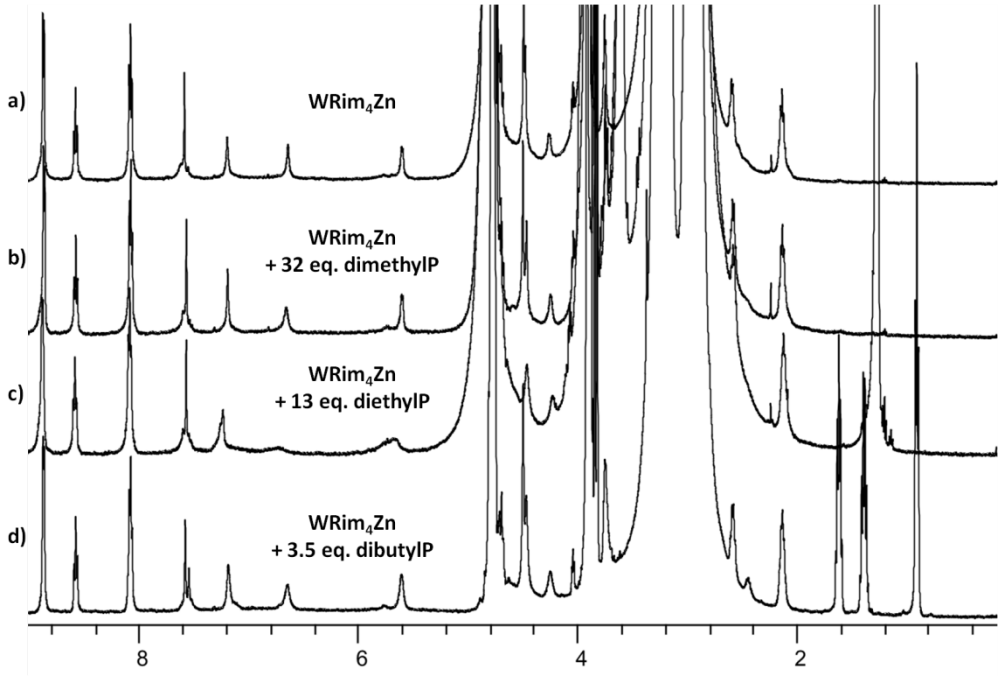


Figure S19. Addition of dialkyl phosphates to **WRim$_4$Zn** (1 mM, buffer HEPES 100 mM) (^{1}H NMR, D_2O, 300 K, 500 MHz). a) **WRim$_4$Zn** (1 mM, pD = 7.4); b) + 32 equiv. dimethylphosphate (pD = 7.4); c) + 13 equiv. diethylphosphate (pD = 6.2); d) + 3.5 equiv. dibutylphosphate (pD = 7.4).

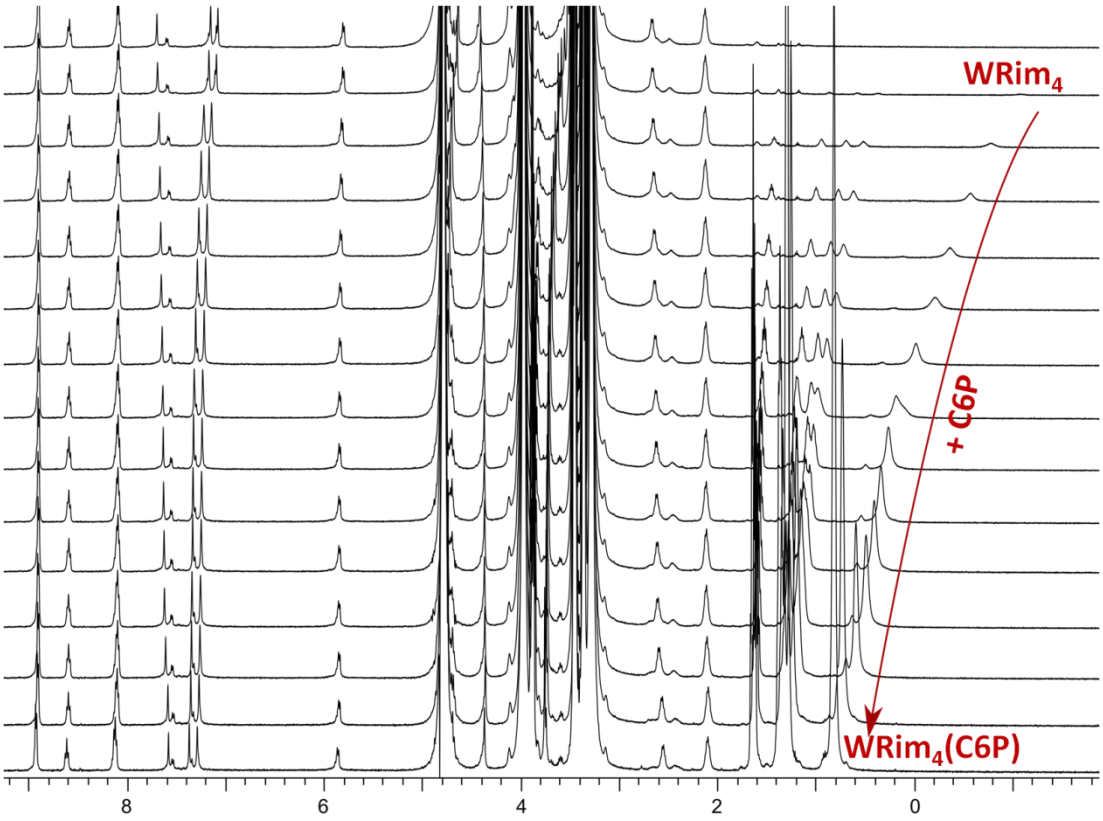


Figure S20. Titration by addition of **C6P** (up to 50 equiv.) to **WRim$_4$** (1 mM) at pD = 5.9 (MES 100 mM buffer). ^{1}H NMR (D_2O, 300 K, 500 MHz).

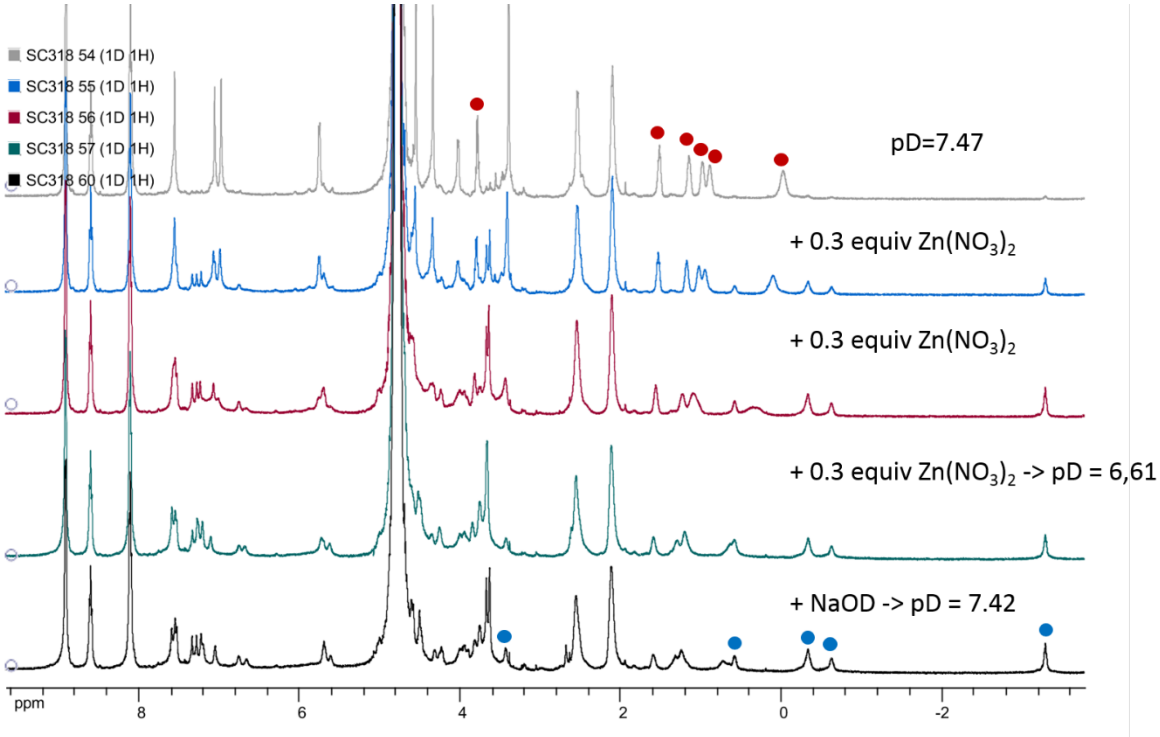


Figure S21. Titration of a solution containing **WRim$_4$** (3.5 mM) + 1.2 equiv. **C6P (**by Zn^{II} showing the progressive formation of **WRim$_4$Zn(C6P)** to the detriment of **WRim$_4$⊃C6P**. ^{1}H NMR (D_2O, 300 K, 500 MHz). The set of guest peaks corresponding to the Zn^{II} host-guest complex **WRim$_4$Zn(C6P)** are denoted in blue, whereas those corresponding to **WRim$_4$⊃C6P** are indicated in red.

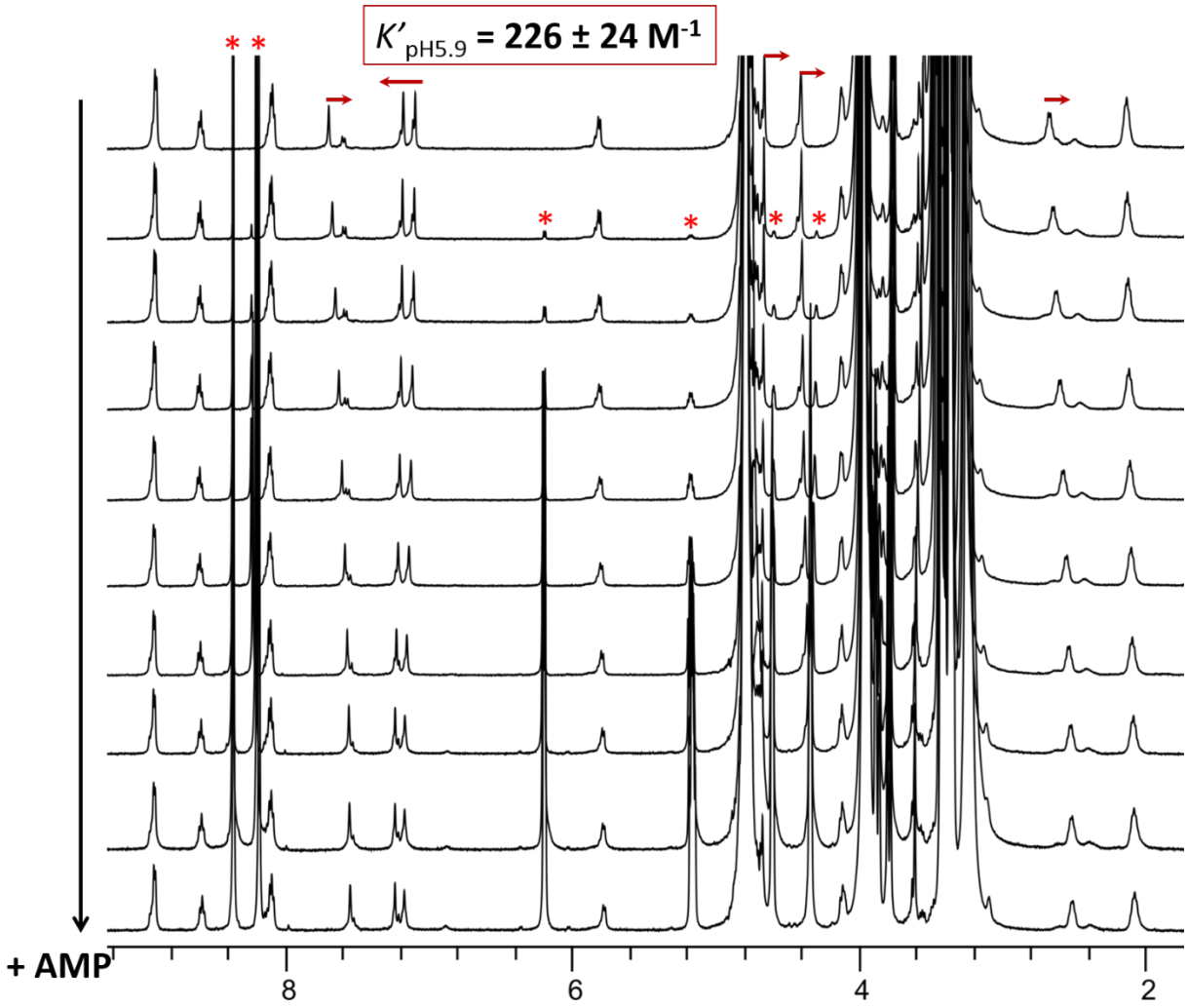
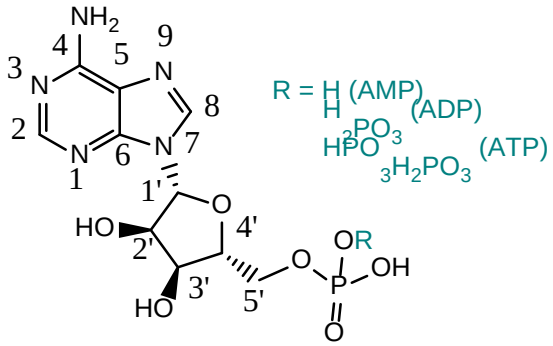


Figure S22. Titration of **WRim$_4$** (1 mM) by **AMP** (up to 40 equiv.) at pD = 5.9 (MES buffer 100 mM in D_2O) monitored by ^1H NMR spectroscopy (300 K, 500 MHz). The red stars indicate the AMP protons, the arrows highlight the shifts observed for the protons belonging to the imidazole moieties and the aromatic bowl core. Top: Structure of the adenosine derivatives. Bottom: full spectra.



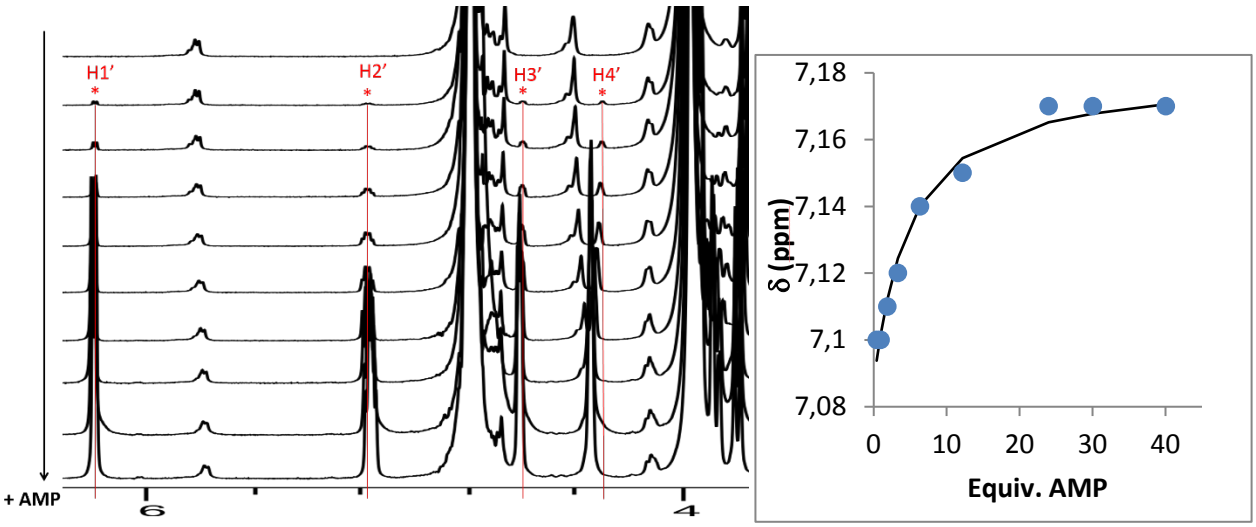


Figure S23. Zoom on Fig. S17 and fitting of the variation of the δ shift of one imidazole proton ($H_{Im\alpha}$) as a function of the amount of added guest for the determination of the associated binding constant, $K'_{pH5.9}$.

Modelling studies

We performed a series of optimizations relying on density functional theory (DFT) to assess the structural changes for n = 1…8. We chose the *meta* hybrid Truhlar's M06-2X, which has proved its performance notably for non-covalent interactions on supramolecular systems. The basis set was the Pople's double-zeta 6-31G(d,p), augmented by diffuse functions on oxygen atoms for a proper description of the negatively-charged monoalkyl phosphate ligands. The aqueous solvent was accounted for using an implicit description, through the polarizable continuum model implemented with the Gaussian 16 Rev B.01 series of programs.[2] This modeling aims at obtaining representative structures along the series of monoalkyl phosphates complexed with **WRim$_4$Zn**, which are shown in Figure S23. We considered the structures with protonated free imidazoles, guided by the NMR evidences. Indeed, our DFT calculations corroborate interactions between the phosphate moiety and the hydrogen at position ε (see Figure S23) for certain residues. Either none or one imidazolium interacts with the phosphate moiety as listed in Table 1: preliminary calculations using different levels of theory to assess the non-covalent interactions involved in the recognition pattern indicate that the system is highly dynamic, with a complex potential energy surface due to many low-energy degrees of freedom.

Relative interaction energies $\Delta\Delta E^{inter}$, listed in Table S1, were evaluated to afford comparison with the experimental evidences. Meaningful comparison to the isolated cage, which would provide absolute interaction energies, cannot be in absence of X-ray structure and also because coordination of one water molecule is also involved for the isolated cage, blurring a direct comparison. Relative interaction energies $\Delta\Delta E^{inter}$ with respect to **WRim$_4$Zn(C3P)** range between +4.1 and +14.1 kcal.mol^{-1}. A dynamical description of the host-guest association, including an explicit solvation pattern, would constitute an interesting perspective to include entropic contributions and provide a direct comparison and quantitative agreement with the ITC measurements.

[2] Gaussian 16, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al. Gaussian, Inc., Wallingford CT, 2016.

Table S1. Relative interaction energies $\Delta\Delta E^{inter}$ evaluated at the M06-2X/6-31G(d,p) level of theory, taking **WRim$_4$Zn(C3P)** as a reference. For **C2P** and **C5P**, the interaction does not exist due to its directionality.

n	1	2	3	4	5	6	7	8
$\Delta\Delta E^{inter}$ (kcal.mol^{-1})	8.5	11.5	0	11.1	14.1	4.1	7.1	10.2
Interaction phosphate...Im (distance O...H)	no	no	yes (1.82 Å)	yes (1.90 Å)	no (3.10 Å)	yes (1.67 Å)	no	no

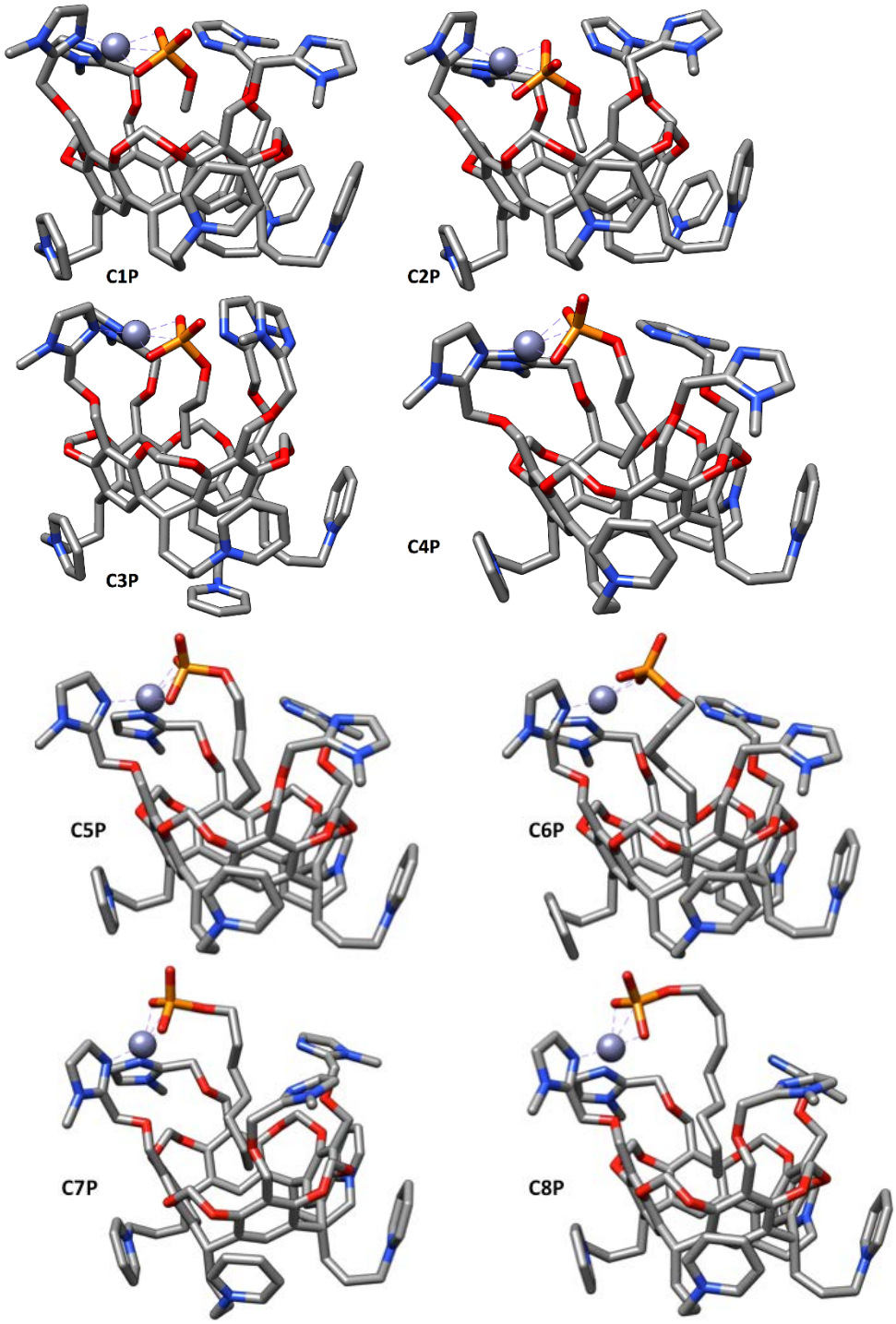


Figure S24. Cartoon representations for the eight complexes **WRim4Zn⊃CnP** optimized by DFT.