SUPPORTING INFORMATION

The dark side of disulfide-based dynamic combinatorial chemistry

Mélissa Dumartin,[a] Jean Septavaux,[a,b] Marion Donnier-Maréchal,[a] Emeric Jeamet,[a] Elise Dumont,[c,d] Florent Perret,*[a] Laurent Vial,*[a] and Julien Leclaire*[a]

[a]*Univ Lyon, Univ Lyon 1, CNRS, INSA, CPE, ICBMS, F-69622 Lyon, France. Emails : florent.perret@univ-lyon1.fr; laurent.vial@univ-lyon1.fr; julien.leclaire@univ-lyon1.fr*

[b]*Secoya Technologies, Louvain-La-Neuve, 1348 Belgium.*

[c]*ENS Lyon, Univ Lyon 1, CNRS, Laboratoire de Chimie, F-69364, France.*

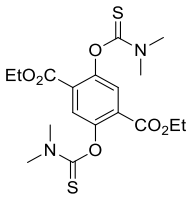
[d]*Institut Universitaire de France, 5 rue Descartes, 75005 Paris, France.*

General methods

All reagents were commercially available and used as received.

Synthesis

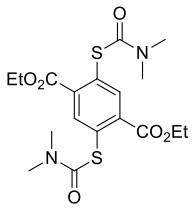
2,5-dimercaptoterephtalic acid **M** was synthesized according to a previously reported procedure from Vial *et al.*[1]



Diethyl 2,5-bis((dimethylcarbamothioyl)oxy)terephthalate

To a solution of diethyl 2,5-dihydroxyterephthalate (10 g, 39 mmol, 1 eq) in dry DMA (100 mL) was added dropwise at 0°C a solution of dimethylcarbamothioic chloride (19 g, 157 mmol, 4 eq) and DABCO (18 g, 157 mmol, 4 eq) in dry DMA (50 mL). The mixture was stirred under nitrogen at room temperature for 24 h. The precipitate was filtrated and washed extensively with water (4 x 250 mL). The solid was dried under vacuum to give a white powder (17 g, 99 %).

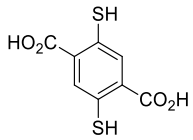
^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.72 (s, 2H), 4.30 (q, J = 7.1 Hz, 4H), 3.45 (s, 6H), 3.39 (s, 6H), 1.33 (t, J = 7.1 Hz, 6H); ^{13}C NMR (75 MHz, $CDCl_3$): δ = 187.1, 163.1, 150.6, 128.6, 127.8, 61.7, 43.4, 39.1, 14.2; melting point: 211 °C.



Diethyl 2,5-bis((dimethylcarbamoyl)thio)terephthalate

Diethyl 2,5-bis((dimethylcarbamothioyl)oxy)terephthalate (3 g, 6.88 mmol) was heated under neat conditions at 210°C for 2 h. The solid was cooled at 90°C. Then, ethanol was added (90 mL) and the mixture was cooled at 0°C overnight. The precipitate was washed with cooled ethanol (2 x 45 mL) and dried under vacuum to give a beige powder (2.46 g, 82 %).

^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.09 (s, 2H), 4.33 (q, J = 7.1 Hz, 4H), 3.14 (s, 6H), 3.04 (s, 6H), 1.35 (t, J = 7.1 Hz, 6H); ^{13}C NMR (75 MHz, $CDCl_3$): δ = 165.4, 165.3, 138.8, 137.3, 131.0, 61.7, 37.1, 14.2; melting point: 140 °C.



2,5-dimercaptoterephthalic acid **M**

A solution of diethyl 2,5-bis((dimethylcarbamoyl)thio)terephthalate (2.6 g, 6.06 mmol, 1 eq) in degassed 1.3 M KOH (5.8 g, 104 mmol, 17 eq) in EtOH/H_2O (1:1, 80 mL) was refluxed under an inert atmosphere for 3 h. The reaction mixture was cooled in ice, and concentrated HCl (15 mL) was added until pH 1. A bright yellow precipitate was formed, filtered, and washed extensively with water, yielding compound **M** as a yellow solid (1.215 g, 95 %).

^{1}H NMR (300 MHz, DMSO-d6): δ = 8.06 (s, 2H); ^{13}C NMR (75 MHz, DMSO-d6): δ = 166.7, 133.2, 130.0; melting point: degradation over 387 °C.

Dynamic combinatorial libraries

Dynamic combinatorial screening was conducted as follows: a solution of building block **M** (4 mM) and a chosen template (1 mM **T2**-**T4** in 200 mM TRIS at pH 7.4, or 200 mM ammonium acetate) were allowed to cyclize through the formation of disulfide bridges and equilibrate under stirring in presence of air. Identification and quantification of the oligomers were conducted by HPLC-MS on an Agilent 1260 Infinity Quaternary LC system equipped with a quaternary pump, autosampler, heated column compartment (35°C), diode array detector and a triple quadrupole mass spectrometer Agilent 6490. Elution was conducted at 0.8 mL/min on an Agilent Eclipse XDB C8, 4.6x150mm, 3µm column (T = 35°C) using the following gradient:

Time (min)	***Water + 0.05% TFA***	***Acetonitrile + 0.05% TFA***
0	80%	20%
5	60%	40%
8	0%	100%
20	0%	100%

Table S1. HPLC gradient used for the analysis of the dynamic combinatorial libraries. UV/Vis monitoring: 254nm; Library volume injected: 5µL.

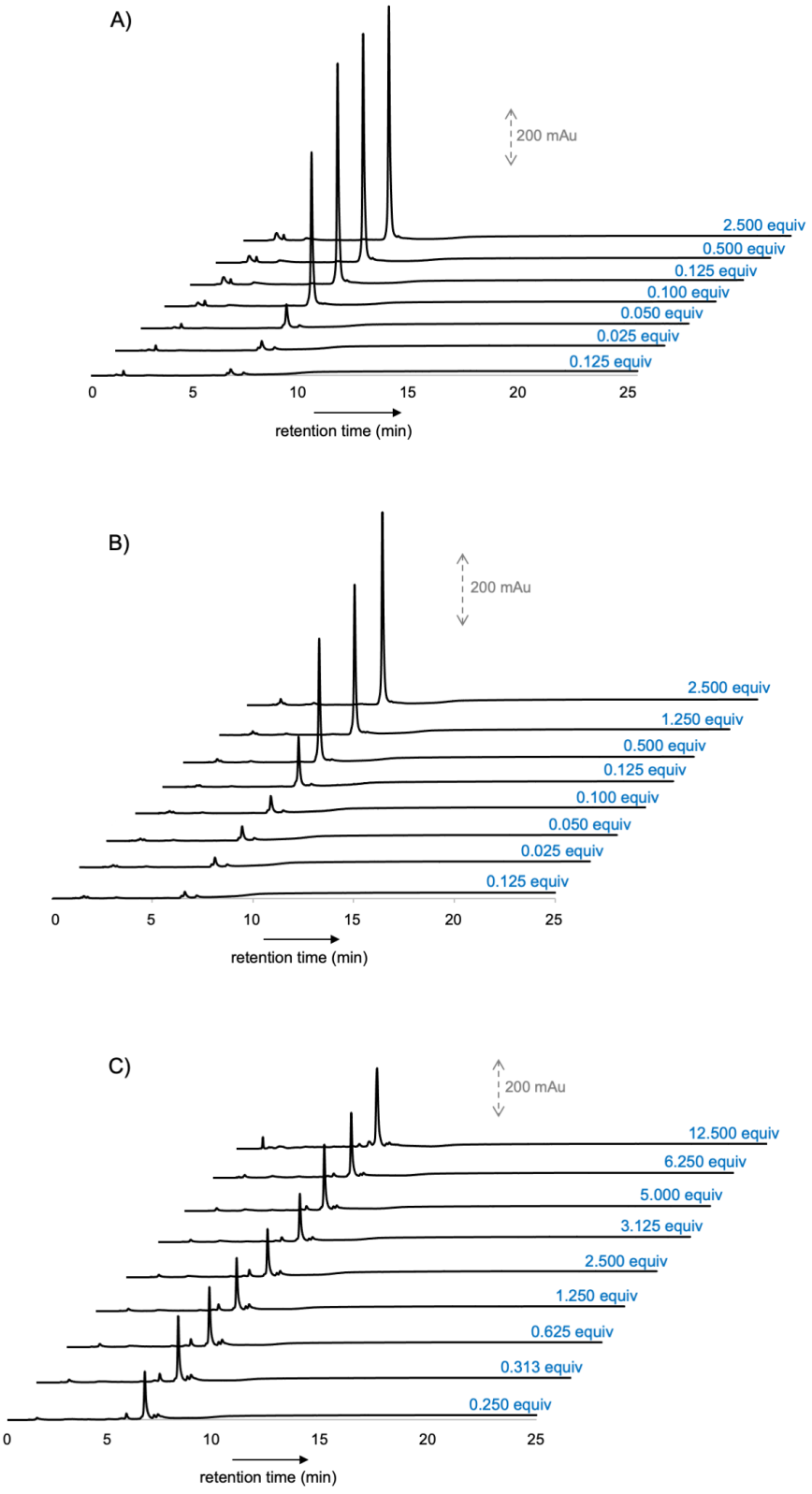


Figure S1. HPLC-UV monitoring of the conversion of **M** (4mM) in **M_4-a** with increasing amounts of cadaverine (A), butylamine (B), and ammonium acetate (C).

Mass spectra were acquired in ultrascan mode by using a drying temperature of 400°C, a drying gas flow of 11 L/min, and a capillary voltage of 2000 V (ESI-).

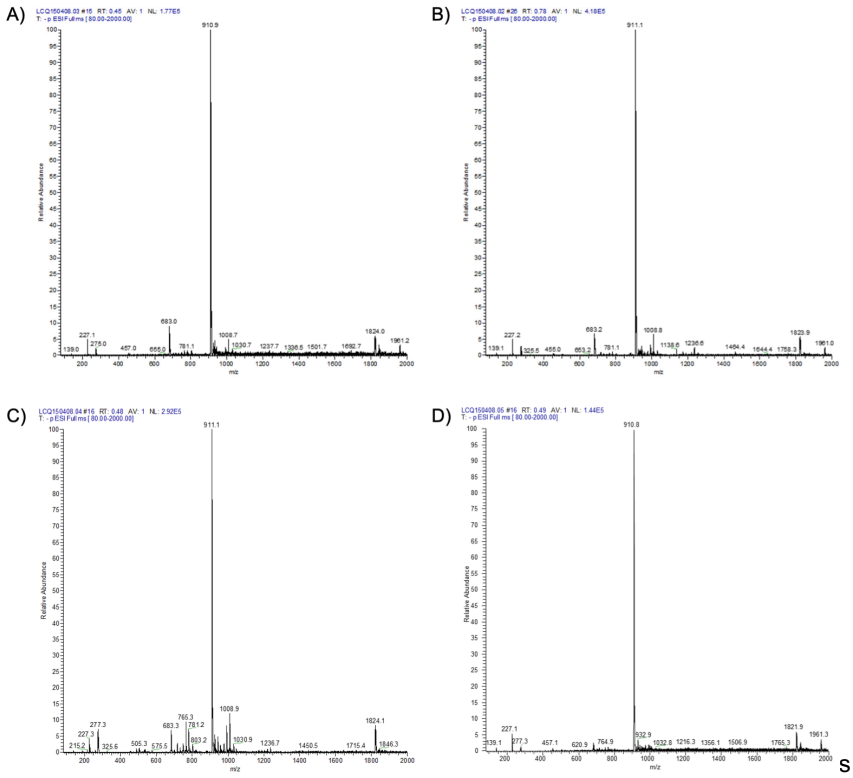


Figure S2. ESI-MS spectra (negative ion mode) of the following *ex-situ* libraries: (A) **L1'** templated by **T1**, (B) **L2'** templated by **T2**, (C) **L3'** templated by **T3**, (D) **L4'** templated by **T4**. $(M_4\text{-}H)^-$: m/z = 910.8 (calc.); $(M_3\text{-}H)^-$: m/z = 682.8 (calc.)

Templates and buffers were removed upon addition of a few milliliters of TFA to the libraries. The resulting yellow precipitate was filtered off, taken up in borate buffer at pH 9, and then TFA was added. The precipitate was collected by centrifugation, suspended in Milli-Q water, collected by centrifugation again, and dried under vacuum overnight for subsequent analysis by NMR. ^{1}H NMR and DOSY spectra were recorded in heavy water at pH 7.4 at 293 K on a spectrometer operating at 300 MHz. 3-(trimethylsilyl)proprionate-2,2,3,3-d4 signal was used as internal standard. DOSY experiments were driven on a 500MHz NMR Bruker Advance III spectrometer (BBFO or BBI probes with standard 50G/cm Z gradient). In order to avoid convection due to inhomogeneous temperature in the samples, the height of liquid was limited to approximatively 37mm. The standard ledbpgp1s pulse sequence was used, and the gradient pulses δ and delay Δ were adjusted to reach an attenuation of approximatively 95% at maximum gradient. Pulses were calibrated prior acquisition. Processing was performed using the Topspin package and Dynamics center.

F2 (ppm)	***D (m2/s)***	***error***
8.146	1.59e-10	1.300e-11
8.096	1.59e-10	9.053e-12
8.002	8.45e-11	1.825e-11
7.993	2.50e-10	3.035e-11
7.956	1.62e-10	7.086e-12
7.801	1.57e-10	4.650e-12
7.631	1.51e-10	6.076e-12

Table S2. Diffusion coefficients extracted from the DOSY NMR analysis of the library $\mathbf{L_1}$ made from monomer **M** in presence of ammonium acetate **T1** as template.

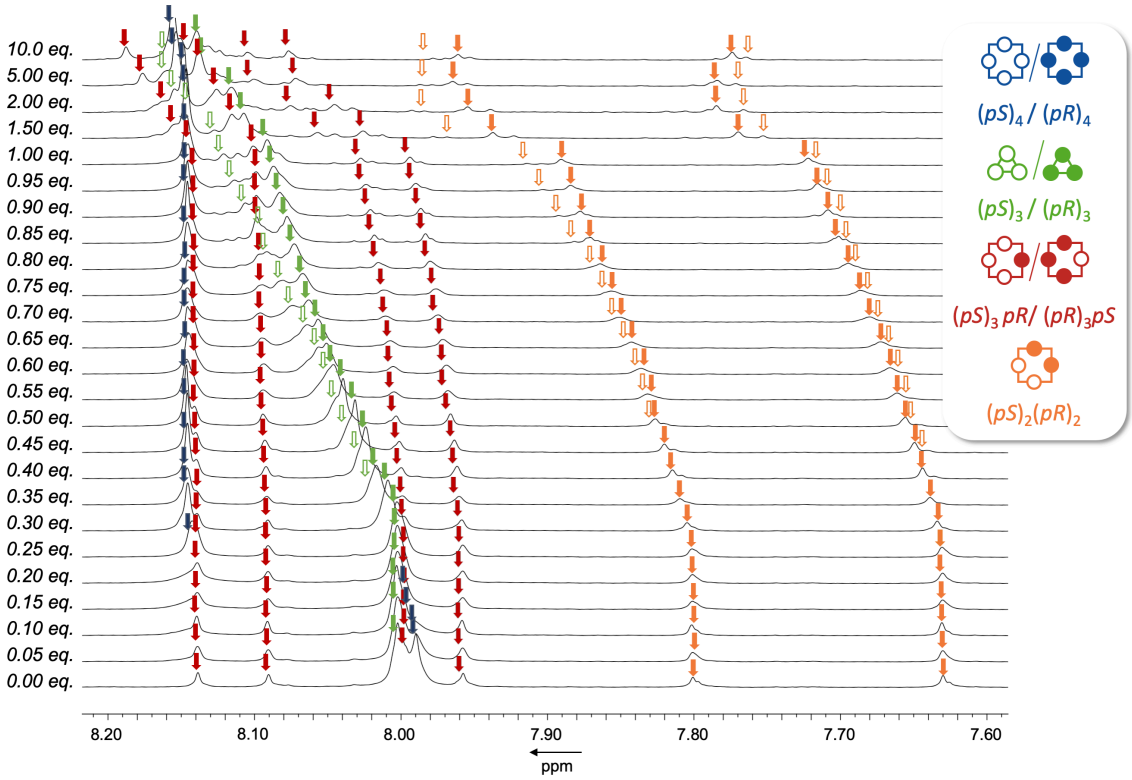


Figure S3. Proton NMR titration (selected spectra) with cadaverine **T3** of the static library **L₁'** made from monomer **M** in presence of ammonium acetate as template. For some proton signals, symmetry breaking of the host induced splitting into major (isochronous, *full arrows*) peaks and minor (individual, *hollow arrows*) peaks.

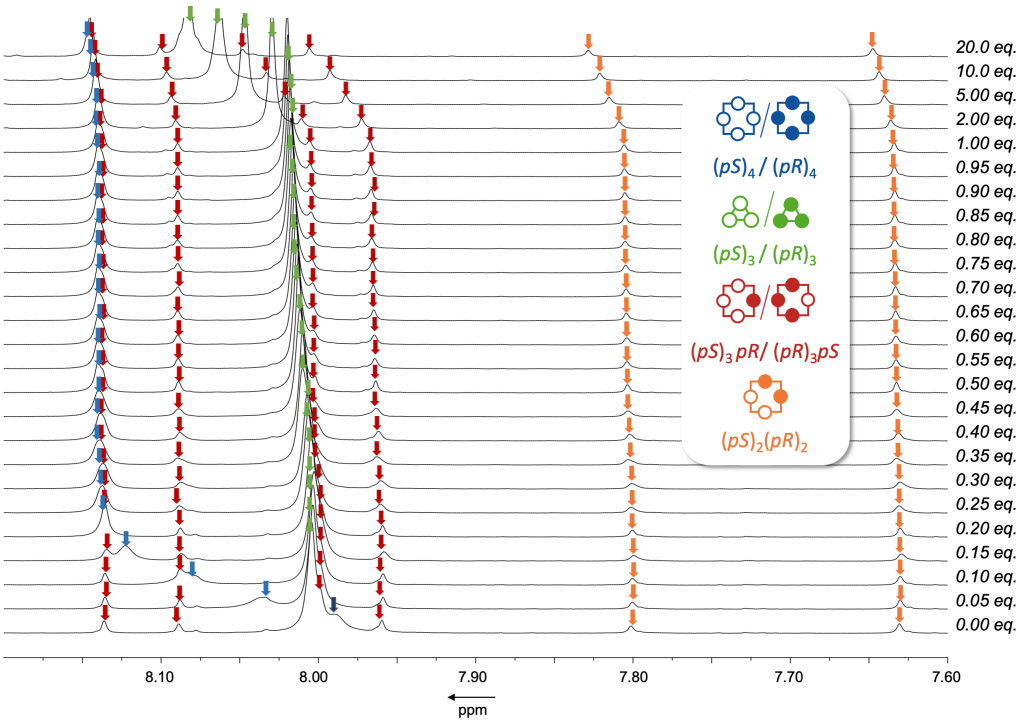


Figure S4. Proton NMR titration (selected spectra) with butylamine **T2** of the static library **L₂'** made from monomer **M** in presence of butylamine as template.

Curve fitting

The apparent association constants of dyn[n]arenes of *(pR)$_3$pS/(pS)$_3$pR*, *(pR)$_2$/(pS)$_2$* and *(pR)$_3$/(pS)$_3$* configuration with cadaverine were obtained by curve fittings on the downfield displacements of the ^{1}H NMR signals of the abovementioned host species upon incremental addition of the guest molecule. Herein, *(pR)$_3$/(pS)$_3$* displays two ^{1}H signal, *(pR)$_3$pS/(pS)$_3$pR* displays eight signals and *(pR)$_2$/(pS)$_2$* displays four signals.

The chemical system used for the ^{1}H NMR study of the association is a mixture of hosts in different proportions displaying distinct behaviours. The multiple hosts compete for a limited amount of guest. Hereafter, for the sake of conciseness, *(pR)$_4$/(pS)$_4$*, *(pR)$_3$pS/(pS)$_3$pR*, *(pR)$_2$/(pS)$_2$* and *(pR)$_3$/(pS)$_3$* have been attributed the subscripts 1, 2, 3 and 4 respectively. Each binding event association constant could be expressed as follow:

$$K_1^1 = \frac{[H1G]}{[H1][G]}, \quad K_2^1 = \frac{[H2G]}{[H2][G]}, \quad K_2^2 = \frac{[H2G_2]}{[H2G][G]}, \quad K_3^1 = \frac{[H3G]}{[H3][G]}, \quad K_3^2 = \frac{[H3G_2]}{[H3G][G]}, \quad K_4^1 = \frac{[H4G]}{[H4][G]}$$

With binding stoichiometry set as 1:2 for *(pR)$_3$pS/(pS)$_3$pR*:cadaverine and *(pR)$_2$/(pS)$_2$*:cadaverine and 1:1 otherwise. The mass balance of the species gives the following equations:

$$[H1]_0 = [H1] + [H1G]$$
$$[H2]_0 = [H2] + [H2G] + [H2G_2]$$
$$[H3]_0 = [H3] + [H3G] + [H3G_2]$$
$$[H4]_0 = [H4] + [H4G]$$
$$[G]_0 = [G] + [H1G] + [H2G] + [H3G] + [H4G] + 2[H2G_2] + 2[H3G_2]$$

With $[X]_0$ being the known concentration of species X prior to binding. By substituting the expressions of the association constants, we obtain:

$$[H1]_0 = [H1] + K_1^1[H1][G]$$
$$[H2]_0 = [H2] + K_2^1[H2][G] + K_2^1K_2^2[H2][G]^2$$
$$[H3]_0 = [H3] + K_3^1[H3][G] + K_3^1K_3^2[H3][G]^2$$
$$[H4]_0 = [H4] + K_4^1[H4][G]$$
$$[G]_0 = [G](1 + K_1^1[H1] + K_2^1[H2] + K_3^1[H3] + K_4^1[H4] + 2K_2^1K_2^2[H2][G] + 2K_3^1K_3^2[H3][G])$$

K_1^1 has been previously measured as 3.11*10^7. The concomitant fitting of modelled speciation to obtain the remaining association constants was performed using a Matlab script. The system of equation could be solved using the *fsolve* Matlab function for any set of $[K_2^1, K_2^2, K_3^1, K_3^2, K_4^1]$, giving the speciation of the system at equilibrium. The *lsqcurvefit* Matlab function is used to adjust $[K_2^1, K_2^2, K_3^1, K_3^2, K_4^1]$ by fitting the speciation to the ^{1}H NMR data. Ccoefficients of determination for each signal range from 0.986 to 0.999. The same methodology was applied for different binding stoichiometries. The present scenario displayed the most satisfactory fitting and was therefore selected.

The same procedure has been applied with butylamine as a guest molecule. A scenario involving stoichiometric complexes for all macrocycles displayed the most satisfactory average coefficient of determination R^2=0.946.

MATLAB scripts are available at https://github.com/lovial/ChemSci

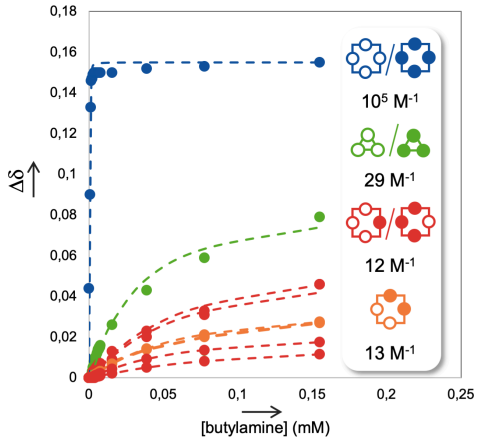


Figure S5. Concomitant fitting of the experimental chemical shifts for each host NMR signal as the concentration of butylamine increases, and the corresponding association constants $K_{\text{bind}}^{\text{app}}$.

Molecular Dynamics simulations

All molecular dynamic simulations and post-processing were performed with the Amber 12 Molecular Dynamics software package.[2] The force-field parameters were taken from parm99, while the parameters for the cage were generated using the generalized AMBER force field GAFF.[3] Each compound was previously build using the Spartan software and their geometries were optimized with the Gaussian09 suite of programs software using density functional theory at the B3LYP level of theory with the double-zeta 6-31G(d,p) basis set. The different parameters were generated with antechamber and parmcheck subprograms, and atomic point charges were computed using the RESP protocol. Notably the latter ensures a repartition of charges that respects the fourfold symmetry of the host (equivalence of charge of the carboxylates). The guest was inserted at specific position around the host using the xleap module, which turn out to be stable along our simulations. This strategy was adopted since non-biaised trajectories did not lead to insertion of the guest association, probably denoting a small free energy barrier. Ammonium cations (NH_4^+) were added in order to neutralize the systems, which were immersed in a truncated octahedral water TIP3P water box containing ~8000 water molecules (10 Å buffer).[4] Each system was first minimized in a 5000 steps simulation, including 2500 steps of steepest descent. Then, a thermalization step was performed to heat each system from 0 to 300 K in 30 ps. The temperature was kept constant during the following steps using Langevin thermostat with a collision frequency γln of 1 ps^{-1}. A 100 ps equilibration run was performed in NPT conditions. Finally, a 100 ns production was executed with constant pressure. After this series of molecular dynamics simulations, a cluster analysis was performed using the cpptraj module of AMBER to extract representative structures, and binding free energies were extracted using the MM/GBSA (molecular mechanics/generalized Born surface area) method to assess the relative association strengths (with the internal and external dielectric constants set to 1 and 80, respectively). The salt concentration was assigned to 0.1M. Trajectories were visualized with the VMD software.[5]

References

[1] L. Vial, R. F. Ludlow, J. Leclaire, R. Pérez-Fernández, S. Otto, *J. Am. Chem. Soc.*, 2006, **128**, 10253–10257.

[2] D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman (2012), AMBER 12, University of California, San Francisco.

[3] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.

[4] P. Mark, L. Nilsson, *J. Phys. Chem. A*, 2001, **105**, 9954-9960.

[5] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33−38.