Recognition of bio-relevant dicarboxylate anions by an azacalix[2]arene[2]triazine derivative decorated with urea moieties

Miguel Santos,[a,b] Igor Marques,[a] Sílvia Carvalho,[a] Cristina Moiteiro*[b] and Vítor Félix*[a]

[a] Departamento de Química, CICECO and iBiMED, Universidade de Aveiro, 3810-193 Aveiro, Portugal.

[b] Centro de Química e Bioquímica, DQB, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

Contents

Figures index

Tables index

^{1}H and ^{13}C NMR spectra

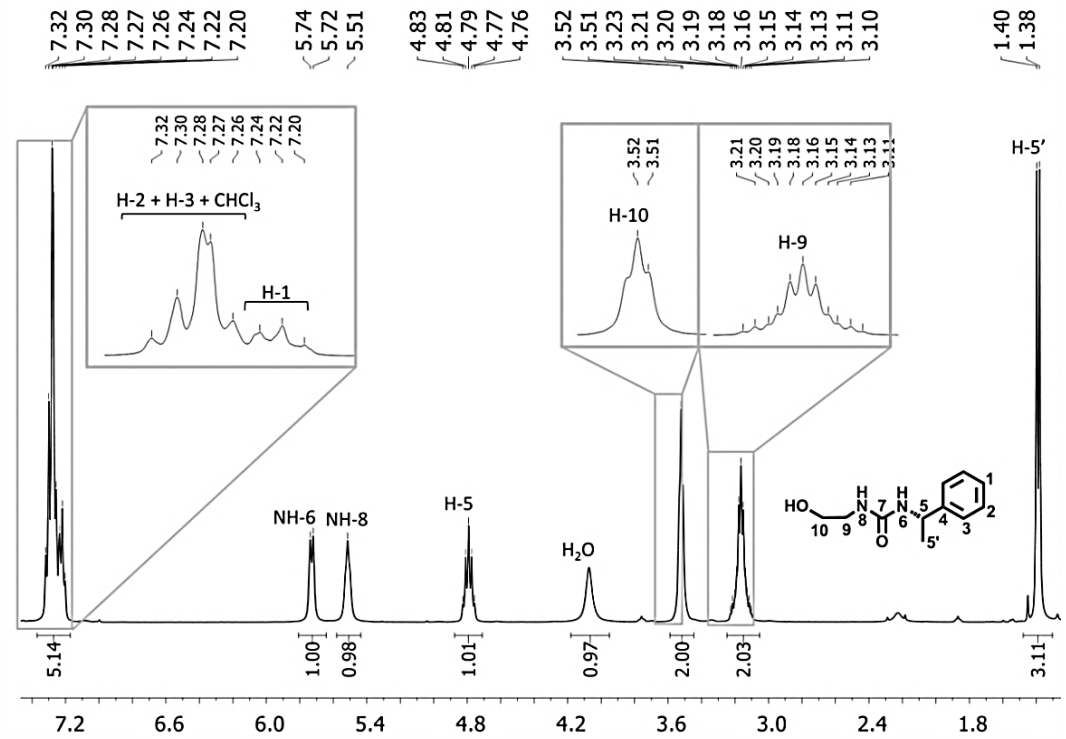

Figure S1. ^{1}H NMR spectrum of **3** in $CDCl_3$.

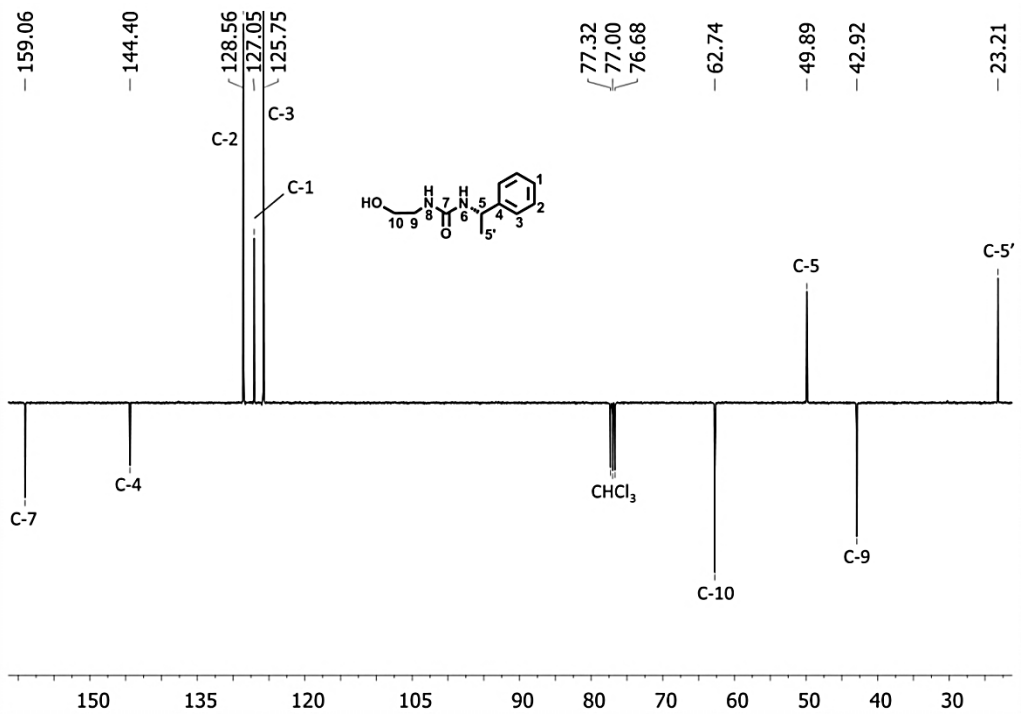

Figure S2. ^{13}C APT NMR spectrum of **3** in $CDCl_3$.

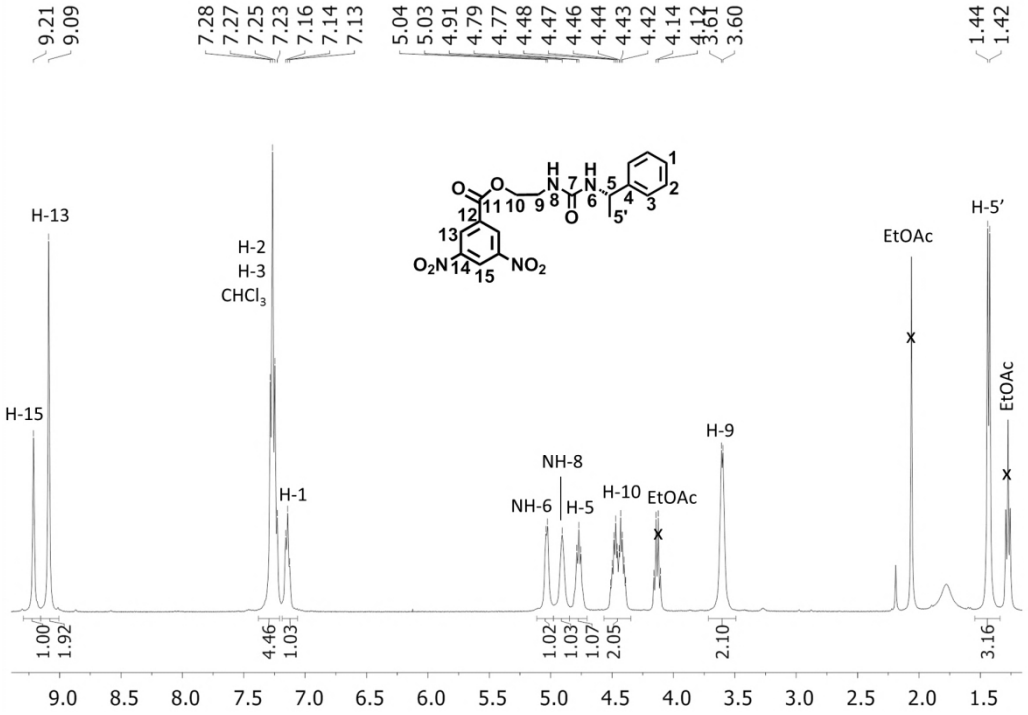

Figure S3. ^{1}H NMR spectrum of **4** in $CDCl_3$.

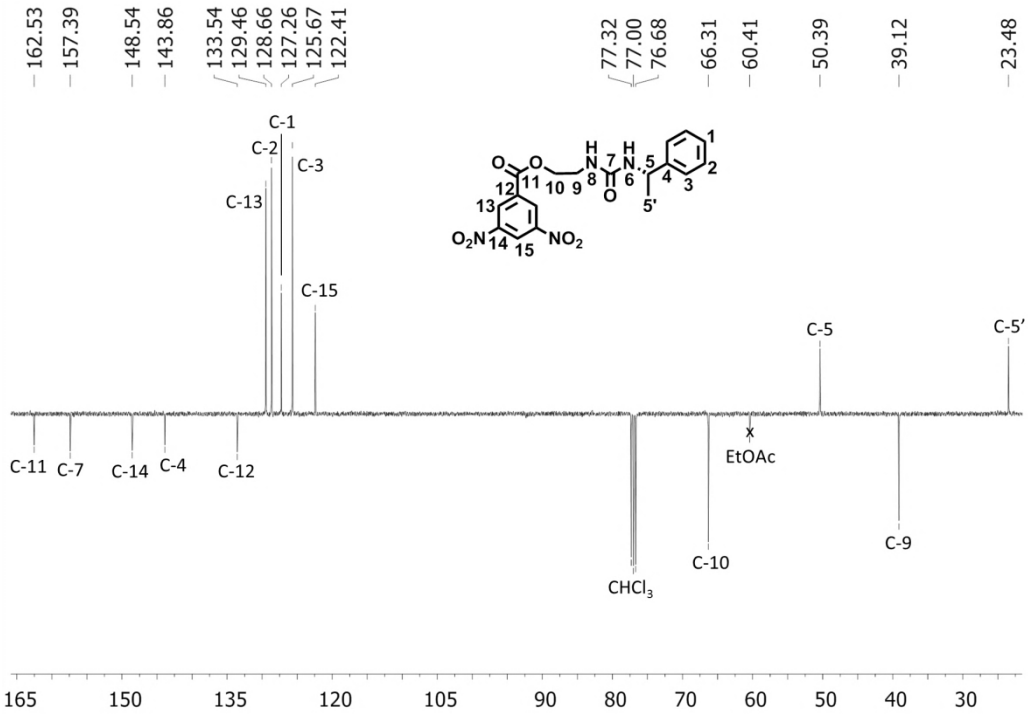

Figure S4. ^{13}C APT NMR spectrum of **4** in $CDCl_3$.

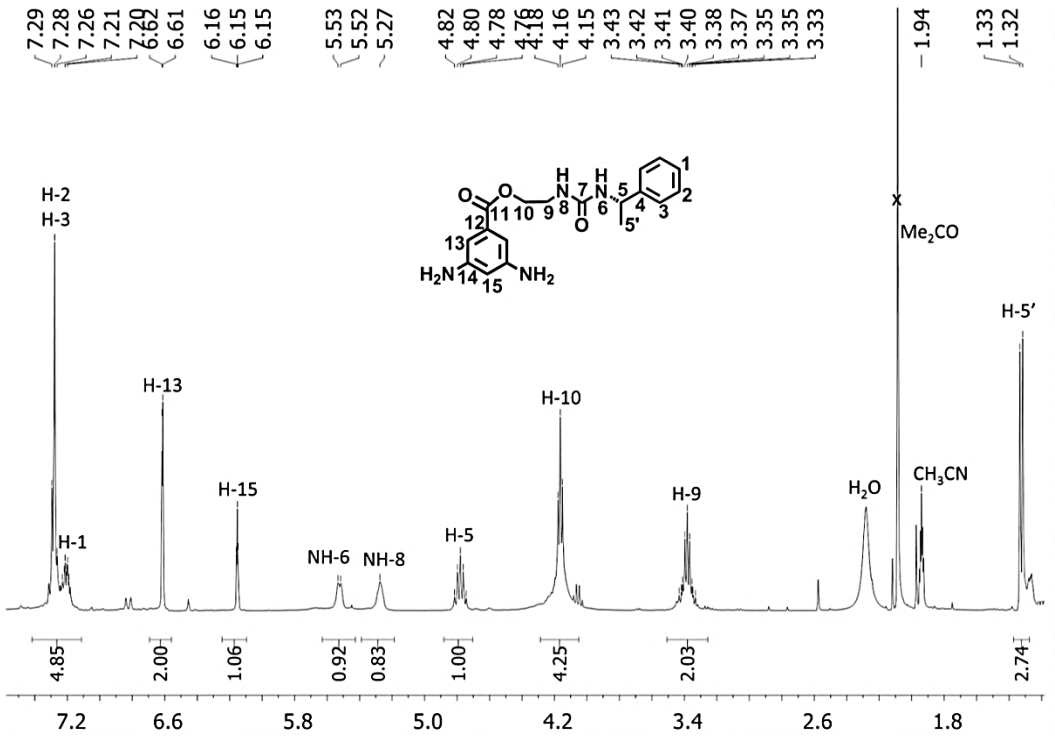

Figure S5. ^{1}H NMR spectrum of **5** in CD_3CN.

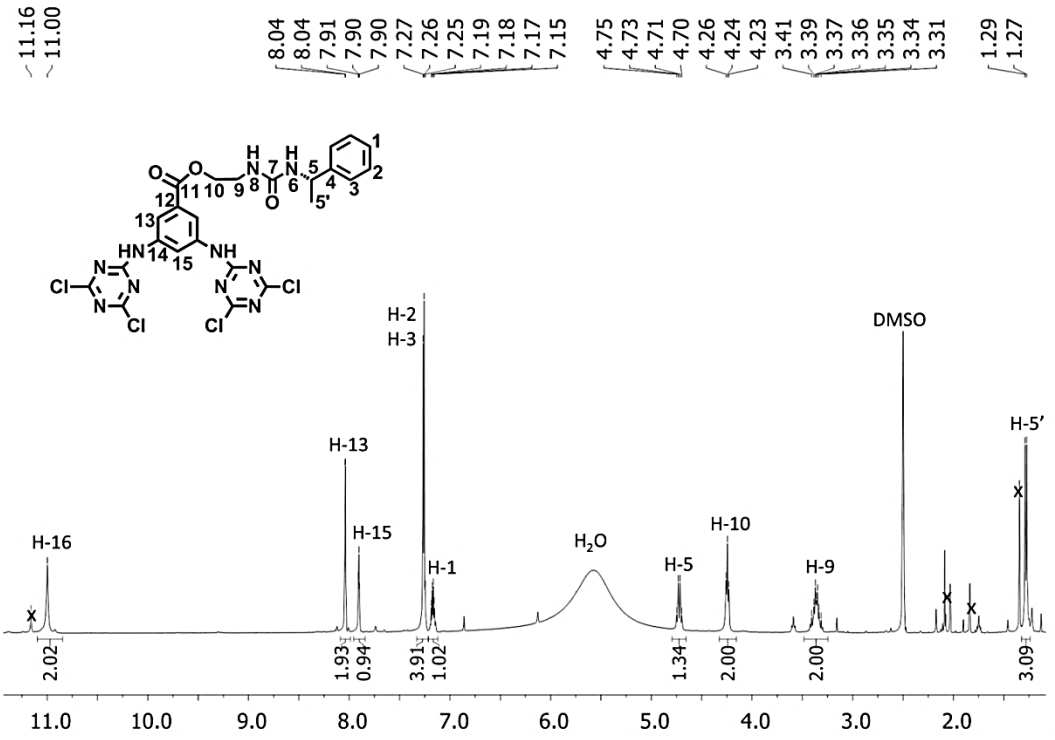

Figure S6. ^{1}H NMR spectrum of **6** in DMSO-d_6.

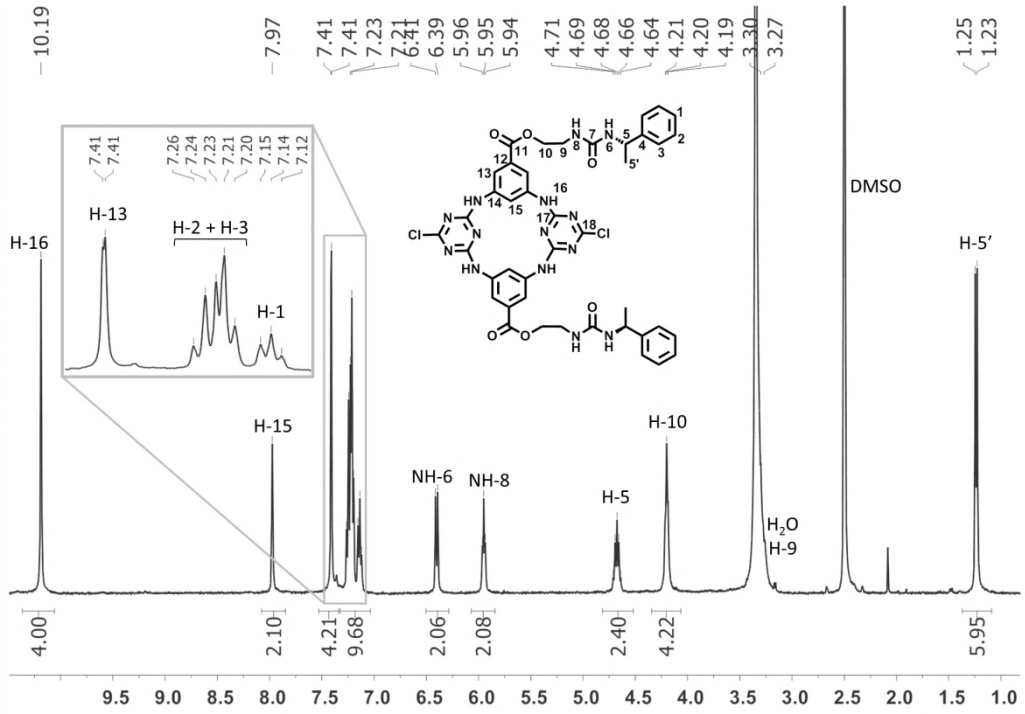

Figure S7. ^{1}H NMR spectrum of **7** in DMSO-d_6.

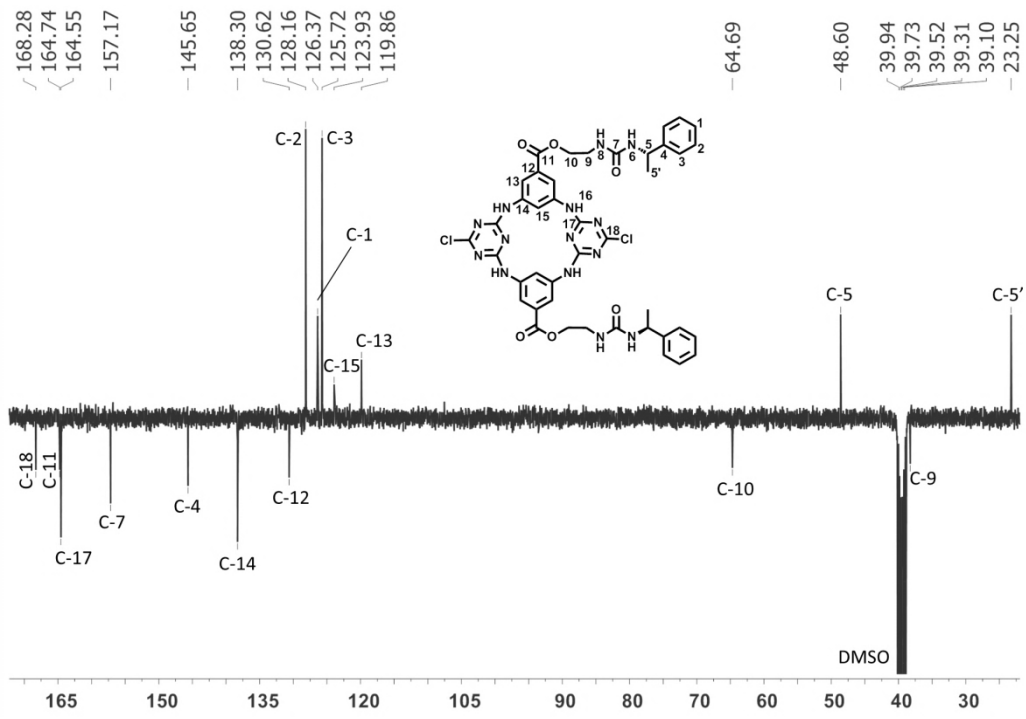

Figure S8. ^{13}C APT NMR spectrum of **7** in DMSO-d_6.

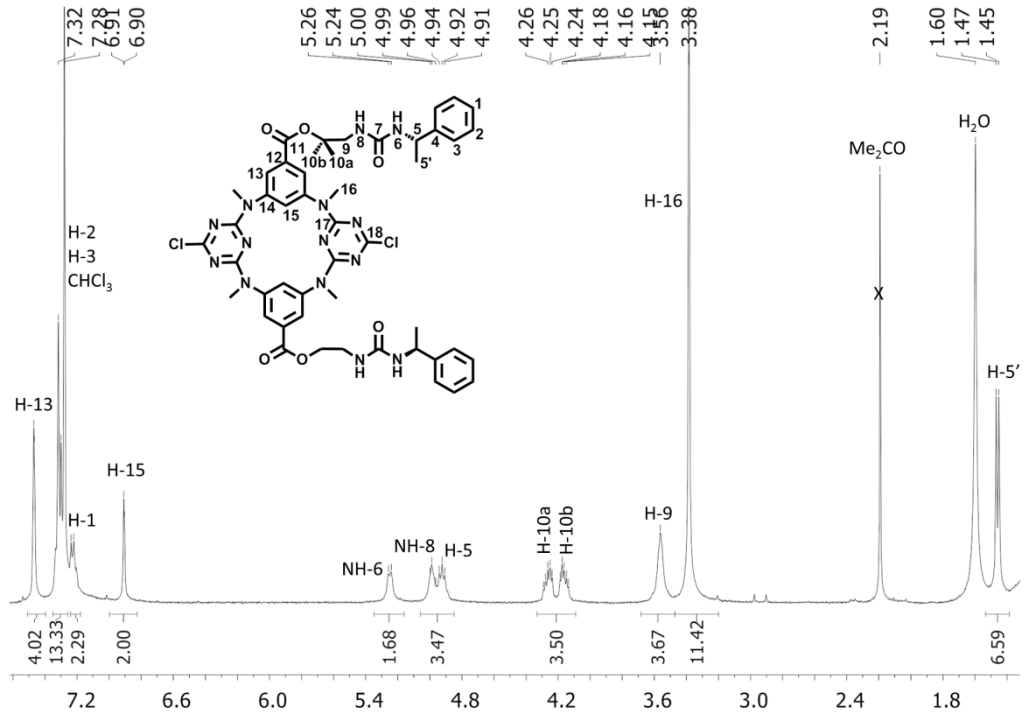

Figure S9. ^{1}H NMR spectrum of **1** in $CDCl_3$.

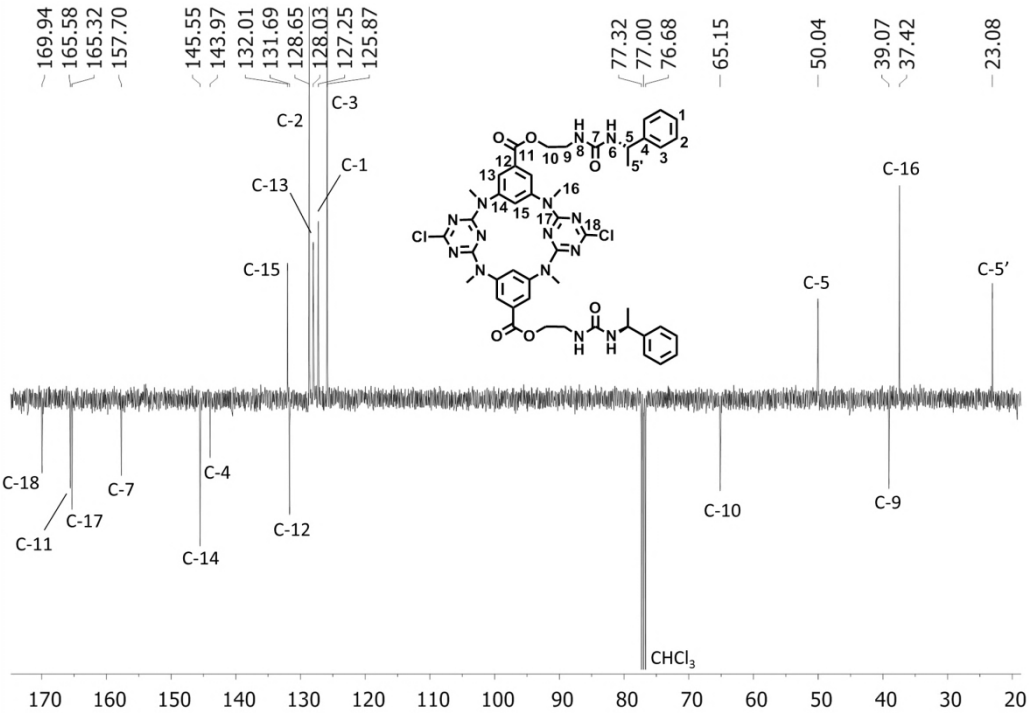

Figure S10. ^{13}C APT NMR spectrum of **1** in $CDCl_3$.

Mass spectra

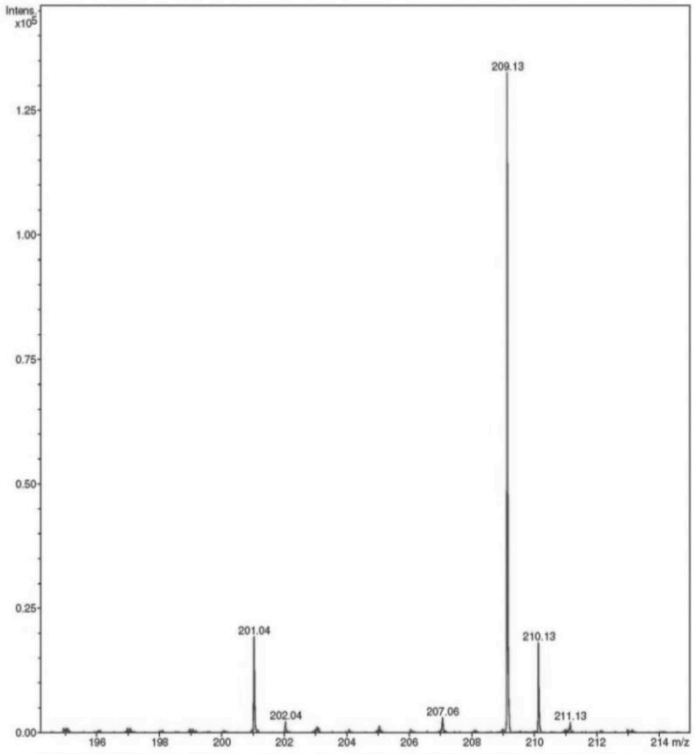

Figure S11. LRMS (ESI) spectrum of **3**.

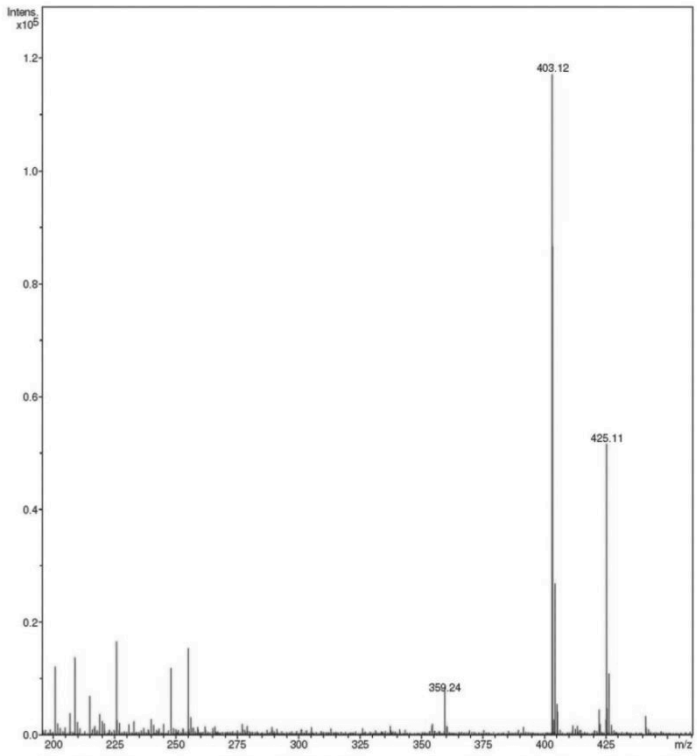

Figure S12. LRMS (ESI) spectrum of **4**.

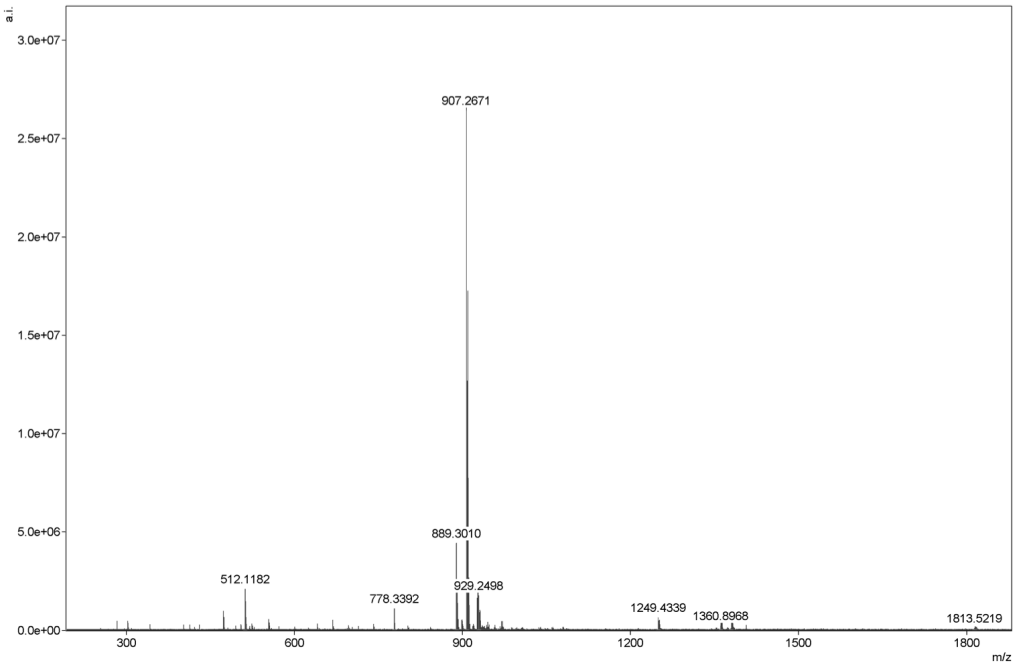

Figure S13. HRMS (ESI) spectrum of **7**.

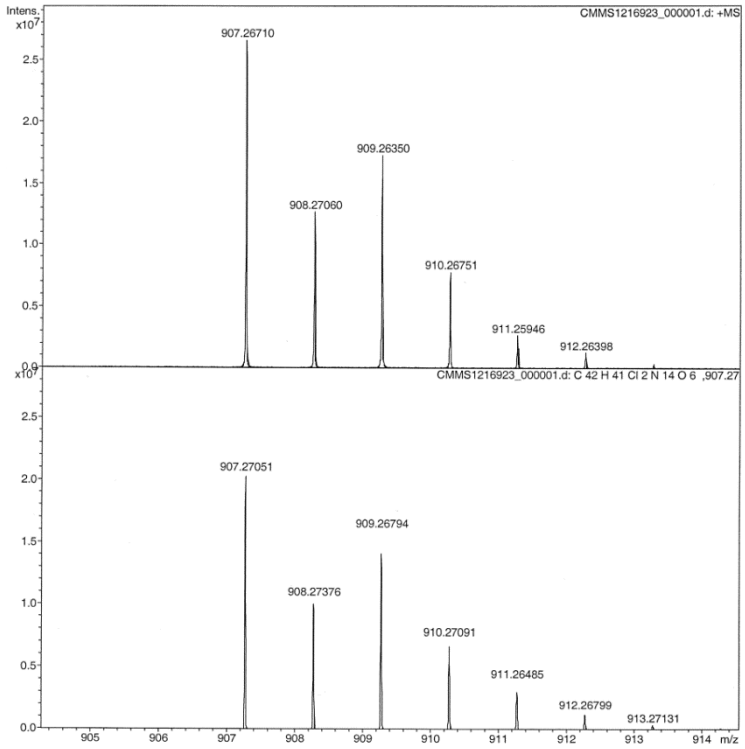

Figure S14. Isotopic distribution of $[M+H]^+$ ion in the HRMS (ESI) spectrum of **7**.

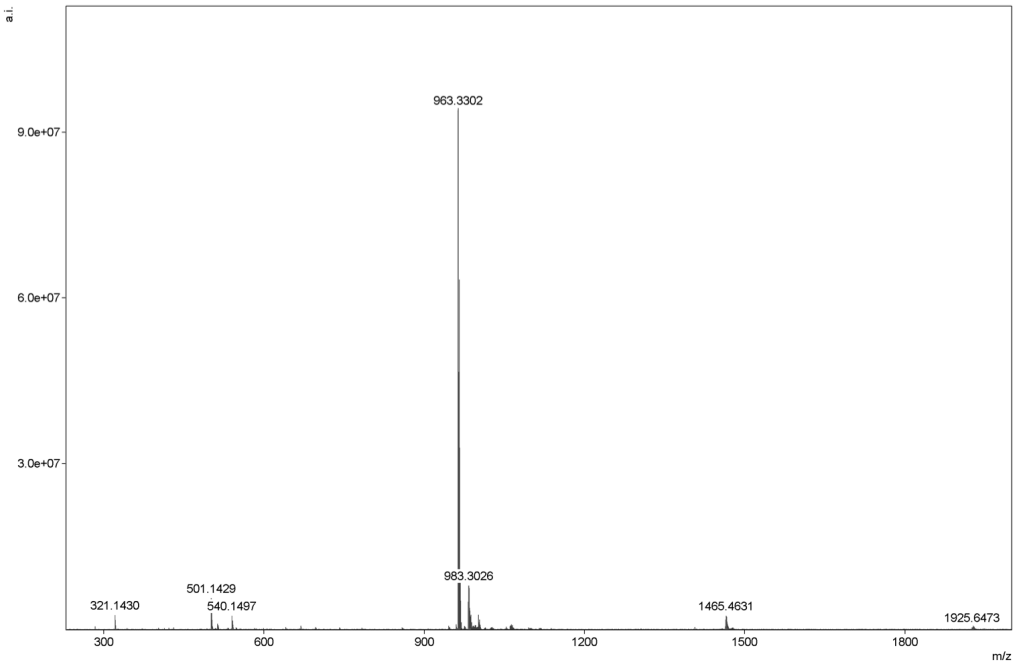

Figure S15. HRMS (ESI) spectrum of **1**.

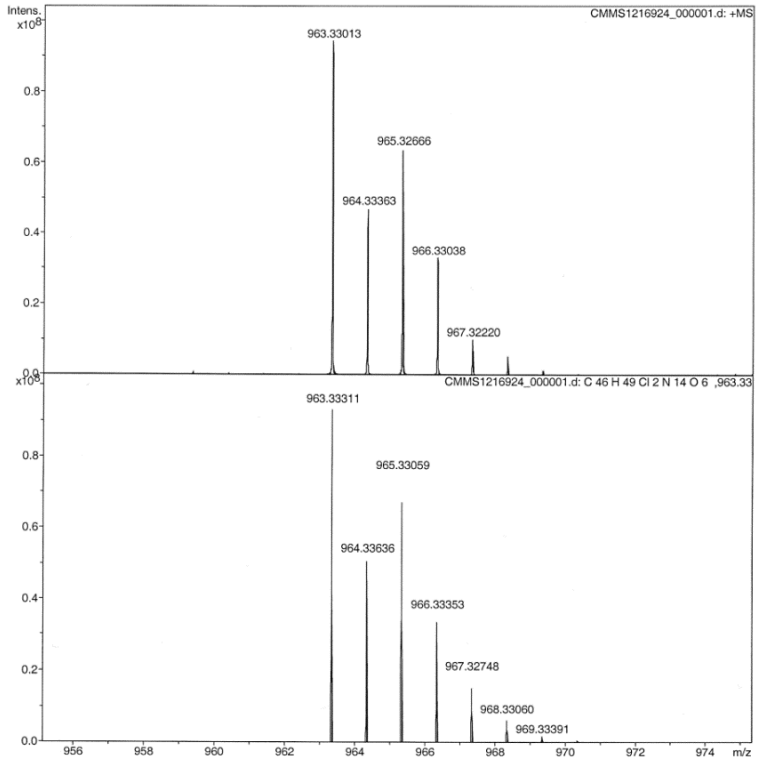

Figure S16. Isotopic distribution of $[M+H]^+$ ion in the HRMS (ESI) spectrum of **1**.

Infrared spectra

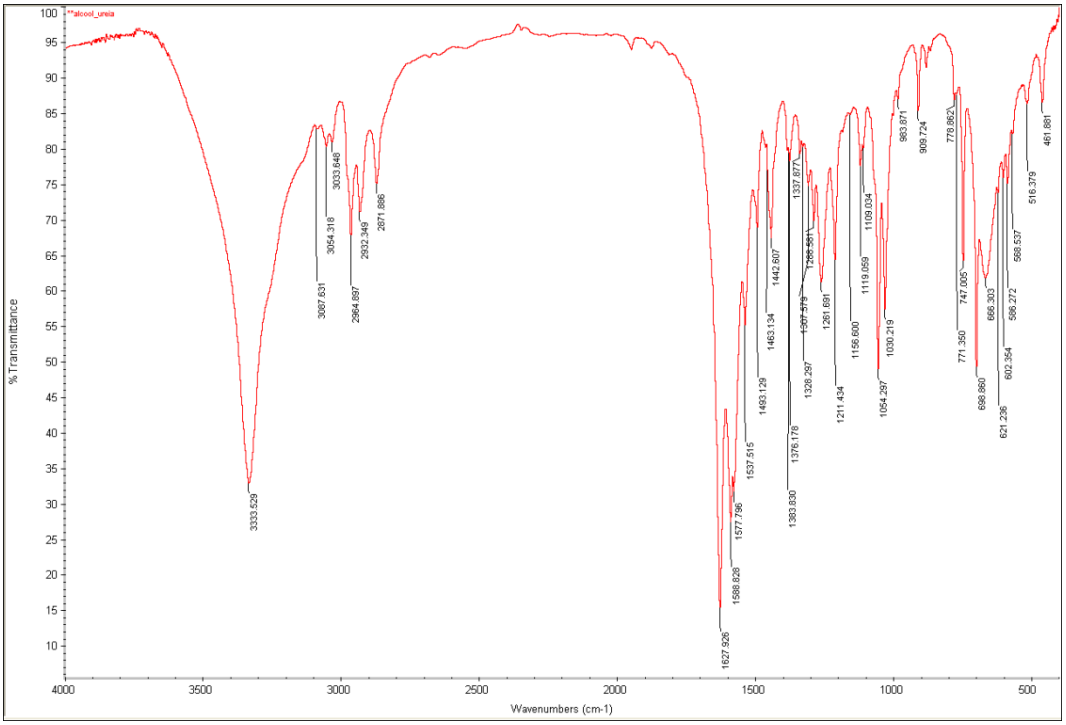

Figure S17. FTIR spectrum of **3** (KBr).

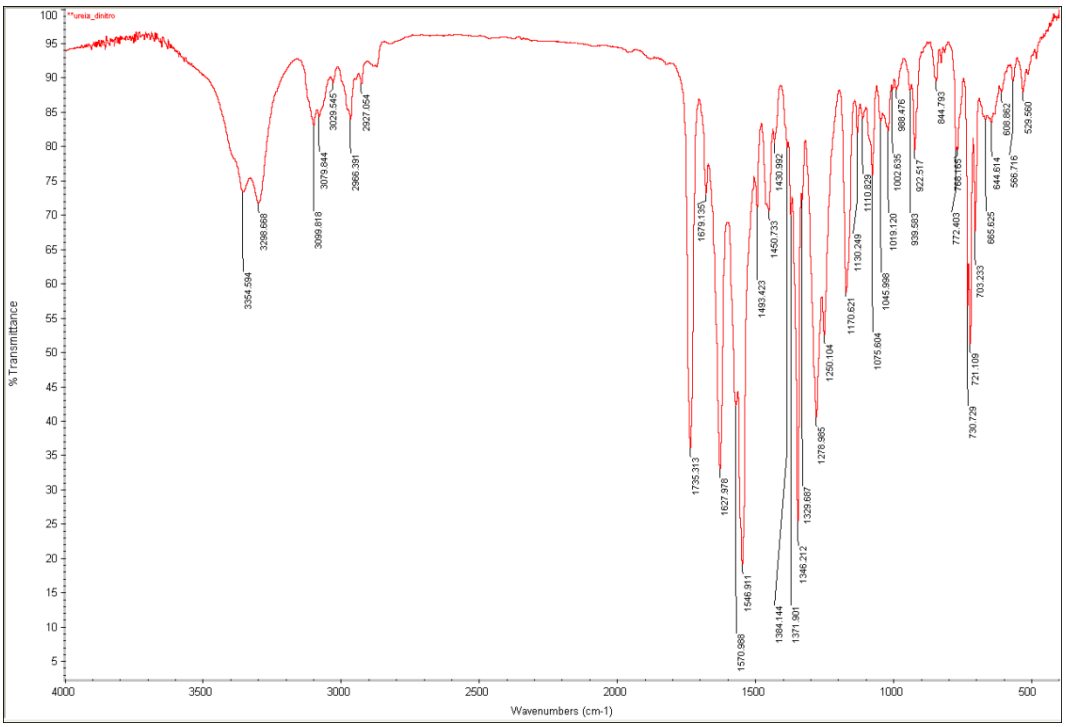

Figure S18. FTIR spectrum of **4** (KBr).

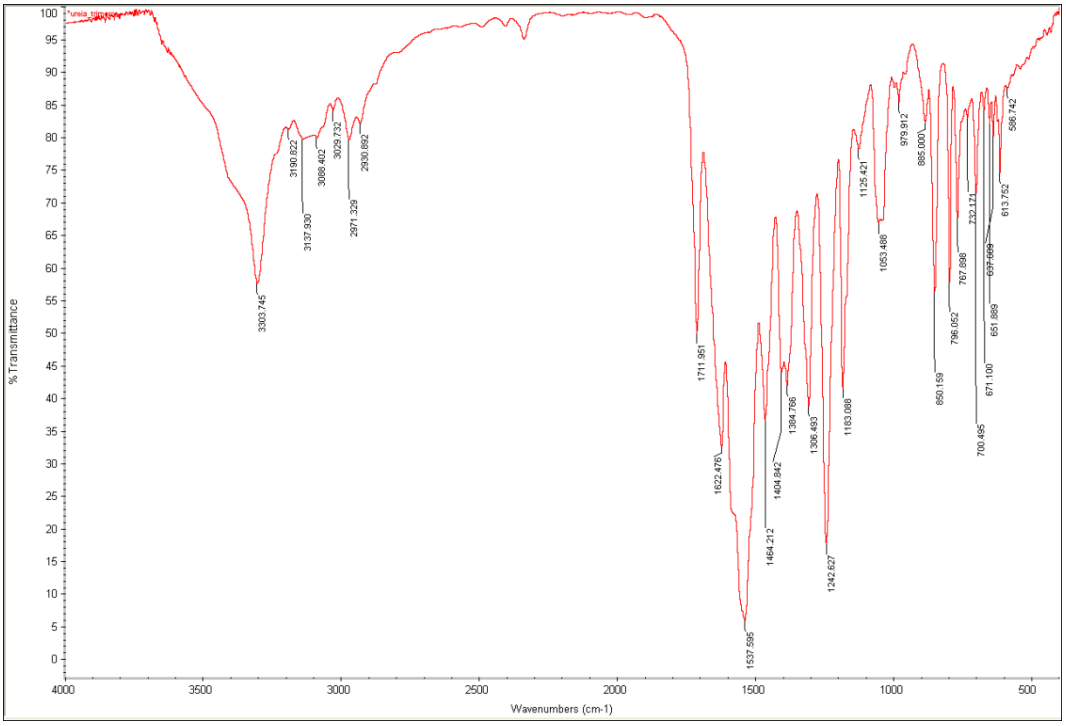

Figure S19. FTIR spectrum of **6** (KBr).

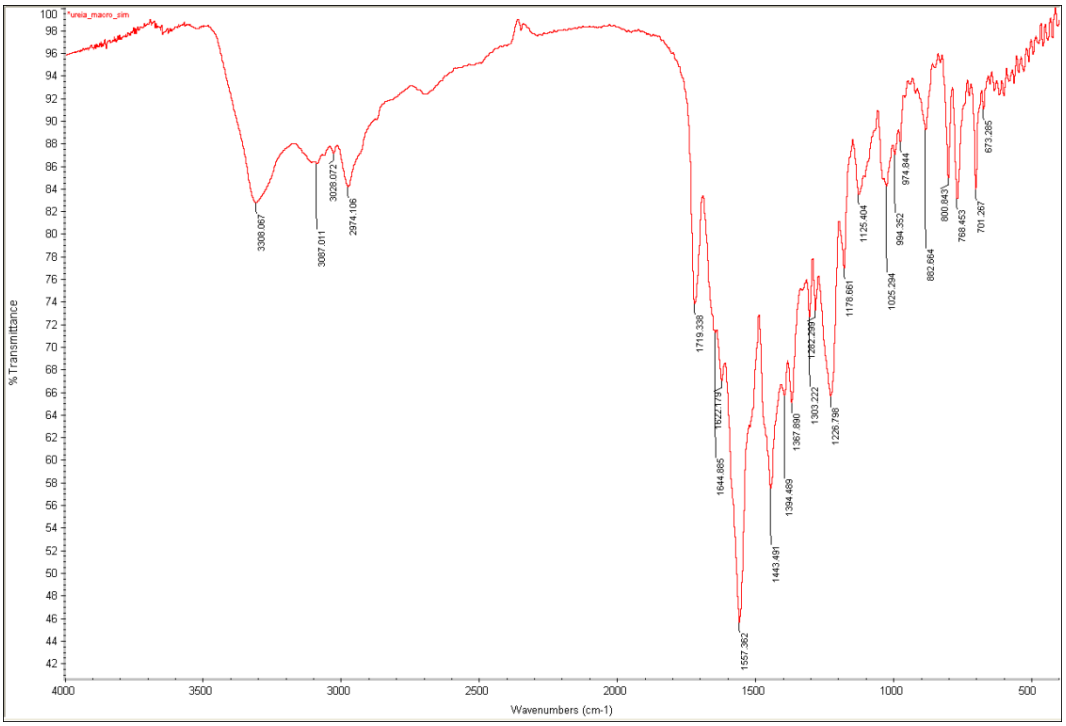

Figure S20. FTIR spectrum of **7** (KBr).

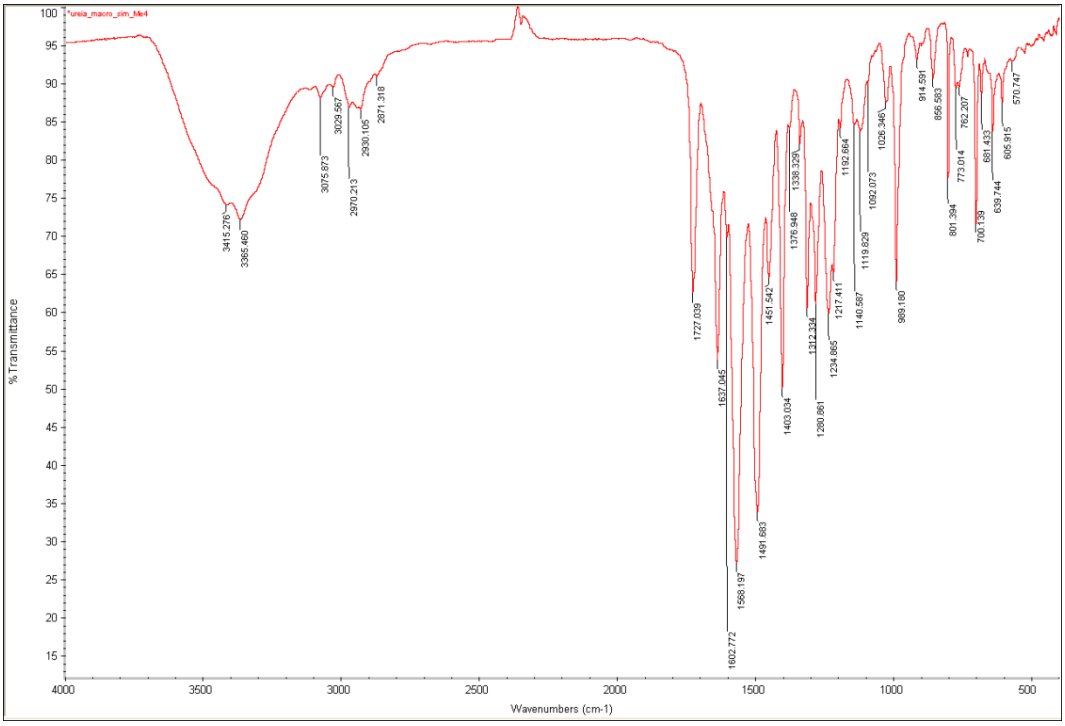

Figure S21.FTIR spectrum of 1 (KBr).

VT ^{1}H NMR

- Tetrazacalix[2]arene[2]triazine **1** in $CDCl_3$

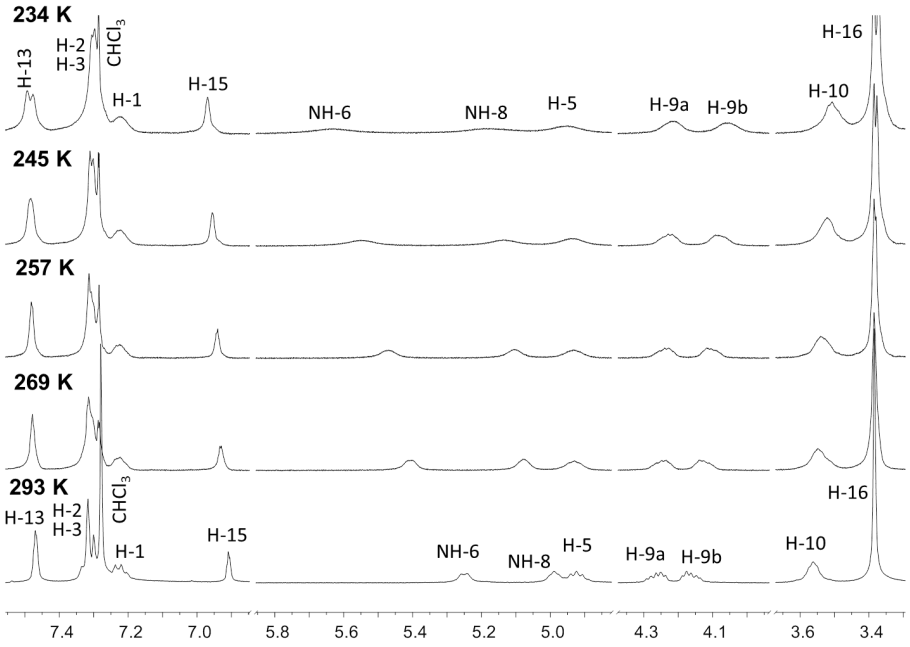

Figure S22. ^{1}H NMR spectra of **1** at 293, 269, 257, 245 and 234 K in $CDCl_3$.

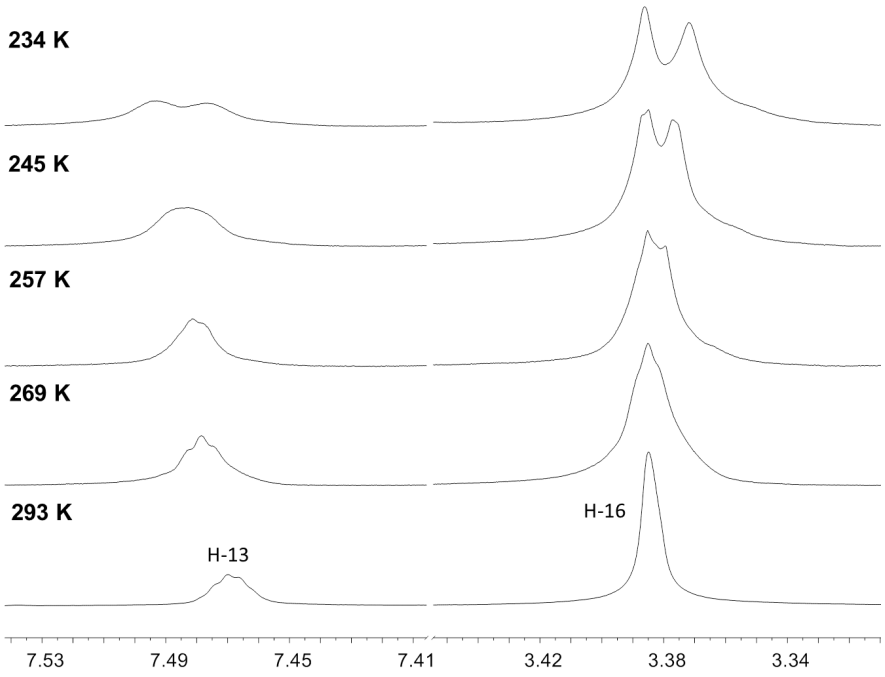

Figure S23. Close-up of H-13 and H-16 signals in the ^{1}H NMR spectra of **1** at 293, 269, 257, 245 and 234 K in $CDCl_3$.

X-ray Crystallography

Table S1. Dimensions of N-H···O hydrogen bonds of **1**.

Hydrogen Bonds [symmetry operation]	H···O (Å)	N···O (Å)	N-H···O (°)
N(38D)-H···O(57D) [-1+*x*,*y*,*z*]	2.27	2.820(8)	121
N(38B)-H···O(57B) [*x*,-1+*y*,*z*]	2.23	2.844(8)	126
N(38A)-H···O(57A) [*x*,-1+*y*,*z*]	2.06	2.823(8)	144
N(38C)-H···O(40D) [*x*,1+*y*,*z*]	2.13	2.913(8)	148
N(41C)-H···O(40D) [*x*,1+*y*,*z*]	2.20	2.964(8)	145
N(55D)-H···O(57C) [1+*x*,*y*,*z*]	2.08	2.872(7)	148
N(55B)-H···O(40A) [1+*x*,1+*y*,*z*]	2.05	2.840(7)	149
N(55A)-H···O(40B)	2.10	2.900(7)	151
N(55C)-H···O(40C) [-1+*x*,*y*,*z*]	2.01	2.820(7)	152
N(58D)-H···O(57C) [1+*x*,*y*,*z*]	2.17	2.951(8)	147
N(58B)-H···O(40A) [1+*x*,1+*y*,*z*]	2.42	3.156(7)	141
N(58A)-H···O(40B)	2.22	3.003(7)	148

[a] The atomic notation scheme used for the urea binding units is given in Figure S24, below.

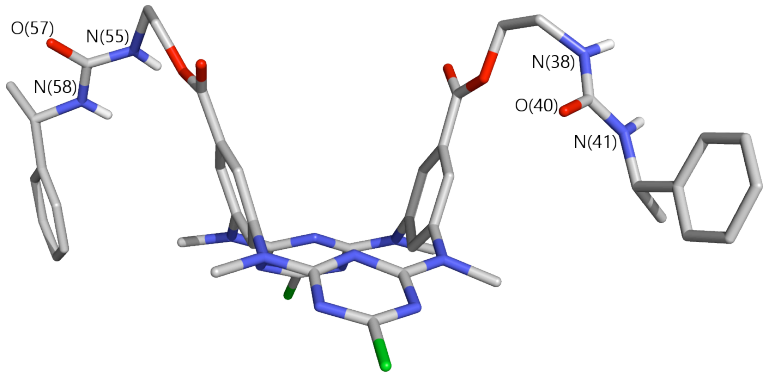

Figure S24. Atomic notation scheme of for the urea binding units of the four independent molecules (A-D) of **1**.

Table S2. Crystal data and structure refinement details for (NMe)4-azacalix[2]arene[2]triazine **1**.

Empirical formula	$C_{46}H_{48}Cl_2N_{14}O_6$
Formula weight	963.88
Temperature/K	150(2)
Crystal system	Triclinic
Space group	*P*1
a/Å	14.9779(6)
b /Å	15.1104(6)
c/Å	23.1992(8)
α/°	86.107(2)
β/°	81.276(2)
δ/°	67.781(2)
V /Å^3	4804.1(3)
Z	4
ρ_{calc} mg/mm^3	1.333
μ/mm^{-1}	0.199
F(000)	2016
Crystal size/mm^3	0.280 × 0.060 × 0.030
Radiation	Mo-Kα (λ = 0.71073)
2 θ range for data collection	4.928 to 51.66
Index ranges	$-18 \le h \le 18$, $-18 \le k \le 16$, $-28 \le l \le 28$
Reflections collected	55221
Independent reflections	28172 [R_{int} = 0.0375, R_{sigma} = 0.0654]
Data/restraints/parameters	28172/3/2473
Goodness-of-fit on F^2	1.037
Flack parameter	0.05(3)
Final *R* indexes [I $\ge 2\sigma$ (I)]	R_1 = 0.0606, wR_2 = 0.1366
Final R indexes [all data]	R_1 = 0.0886, wR_2 = 0.1506
Largest diff. peak/hole /eÅ^{-3}	0.96/-0.36

NMR titrations

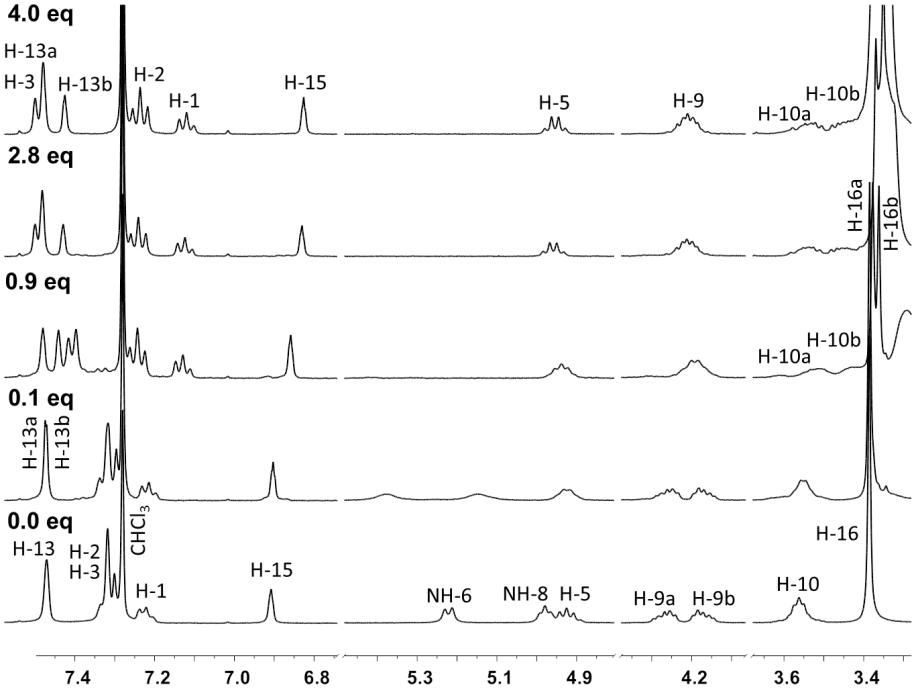

Figure S25. Relevant sections of ^{1}H NMR spectra in $CDCl_3$ of free **1** and upon addition of 0.1, 0.9, 2.8 and 4.0 equivalents of ox^{2-}.

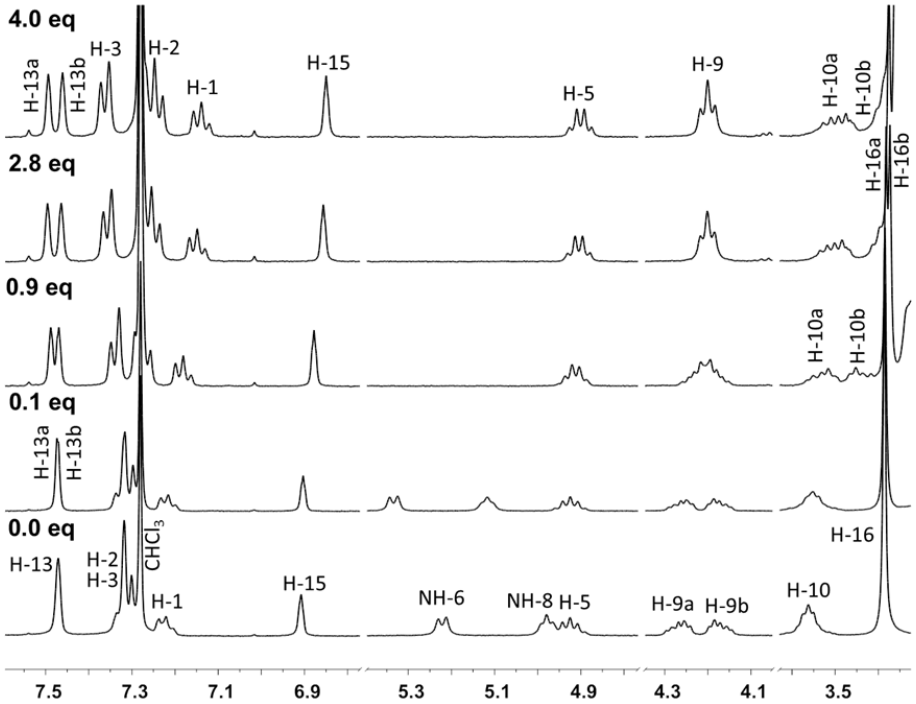

Figure S26. Relevant sections of ^{1}H NMR spectra in $CDCl_3$ of free **1** and upon addition of 0.1, 0.9, 2.8 and 4.0 equivalents of mal^{2-}.

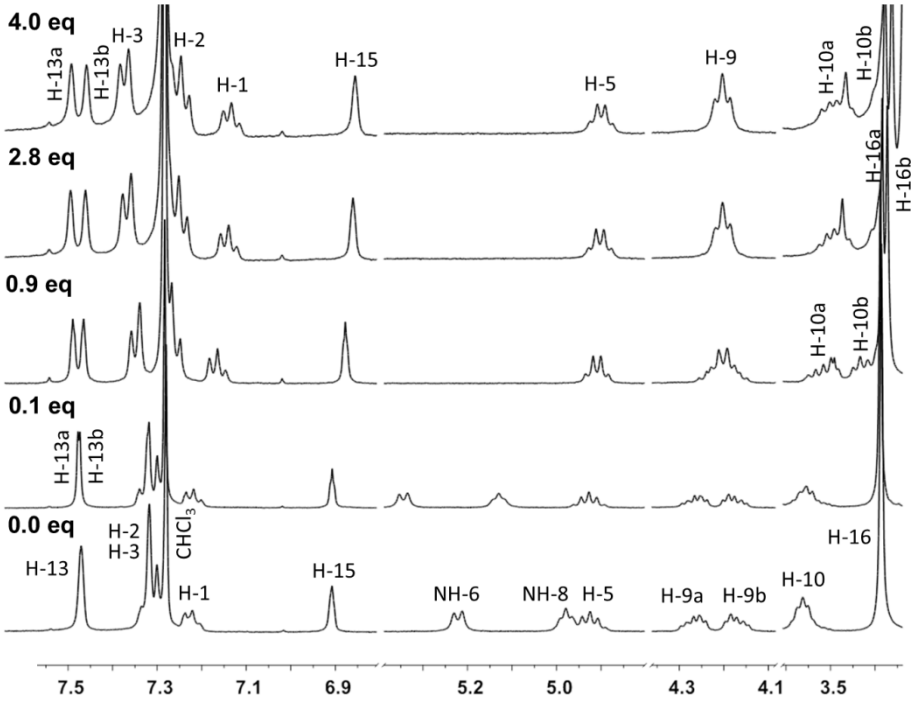

Figure S27. Relevant sections of ^{1}H NMR spectra in $CDCl_3$ of free **1** and upon addition of 0.1, 0.9, 2.8 and 4.0 equivalents of suc^{2-}.

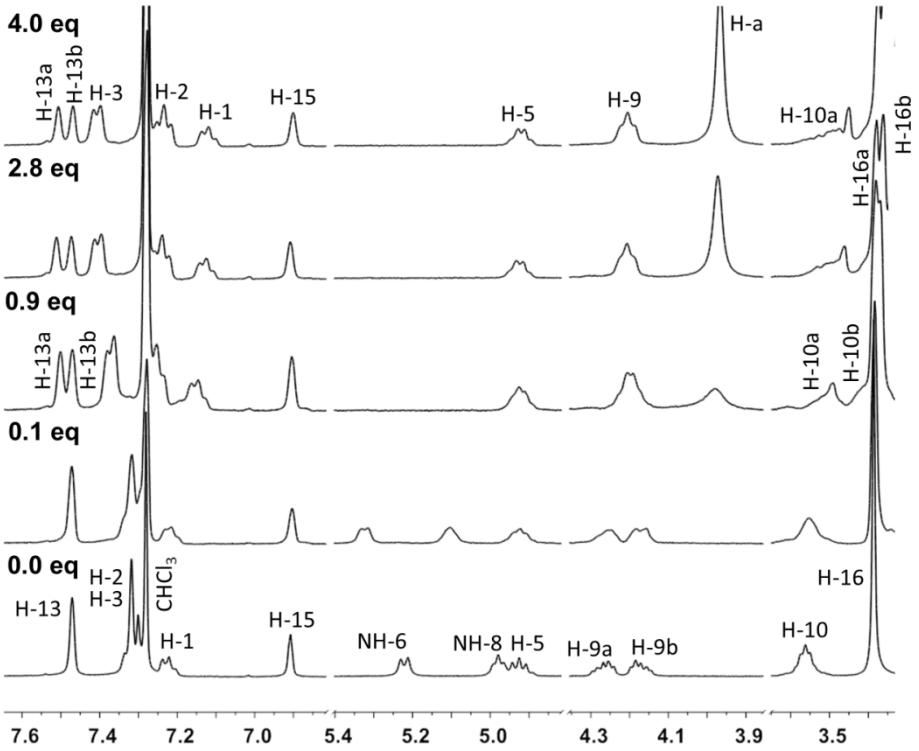

Figure S28. Relevant sections of ^{1}H NMR spectra in $CDCl_3$ of free **1** and upon addition of 0.1, 0.9, 2.8 and 4.0 equivalents of dg^{2-}.

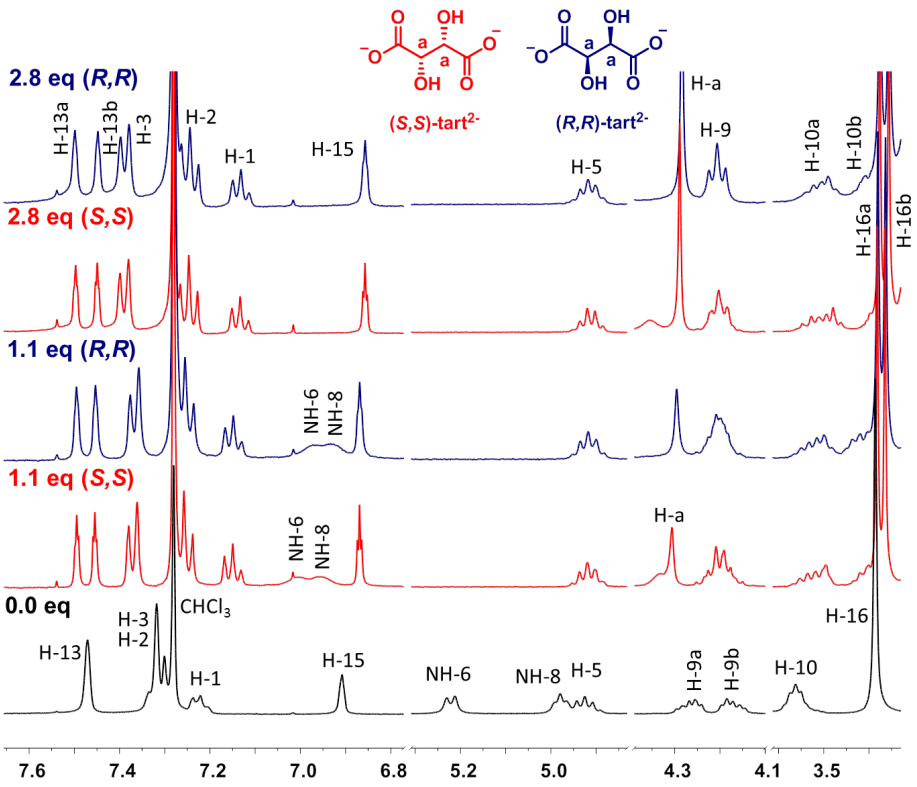

Figure S29. Relevant sections of ^{1}H NMR spectra in $CDCl_3$ of free **1** and upon addition of 1.1 and 2.8 equivalents of (*S*,*S*)- and (*R*,*R*)-tart^{2-}.

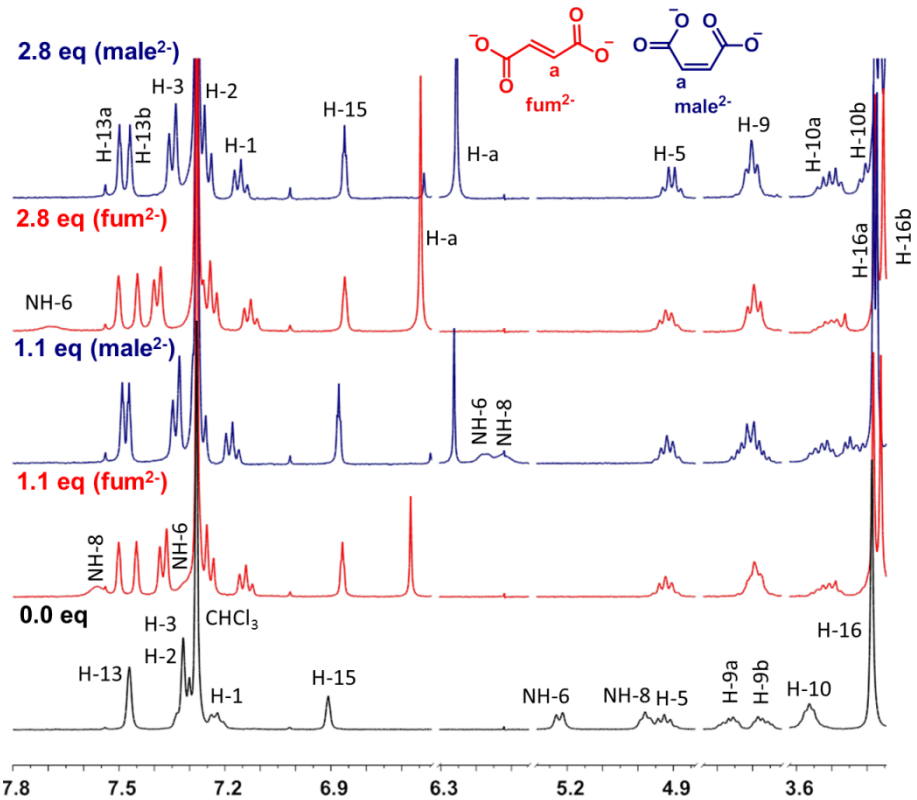

Figure S30. Relevant sections of ^{1}H NMR spectra in $CDCl_3$ of free **1** and upon addition of 1.1 and 2.8 equivalents of fum^{2-} and male^{2-}.

Table S3. Proton resonances used in the determination of K_{ass} for the associations between **1** and the dicarboxylate anion series ox^{2-}, mal^{2-}, suc^{2-}, glu^{2-} and dg^{2-}.

Anions	**ox^{2-}**	**mal^{2-}**	**suc^{2-}**	**glu^{2-}**	**dg^{2-}**
Resonances	H-3	H-1	H-1	H-1	H-1
	NH-6	H-2	H-2	H-2	H-2
	NH-8	H-3	H-3	H-3	H-3
	H-15	H-5	H-5	H-10a	NH-6
		H-5’	H-5’		NH-8
		NH-6	NH-6		H-10a
		NH-8	NH-8		
		H-10a	H-10a		
		H-15	H-15		

Table S4. Proton resonances used in the determination of K_{ass} for the associations between **1** and the dicarboxylate anions (*S,S*)-, (*R,R*)-$tart^{2-}$, fum^{2-} and $male^{2-}$.

Anions	**(*S,S*)-, (*R,R*)-$tart^{2-}$**	**fum^{2-}**	**$male^{2-}$**
Resonances	H-1	H-1	H-1
	H-2	H-2	H-2
	H-3	H-3	H-3
	H-5’	H-5’	H-5’
	H-9a	NH-6	NH-6
	H-10a	NH-8	NH-8
	H-13b	H-13b	H-9a
	H-15	H-15	H-9b
	H-a	H-a	H-10a
			H-10b
			H-13b
			H-15
			H-a

Job plots

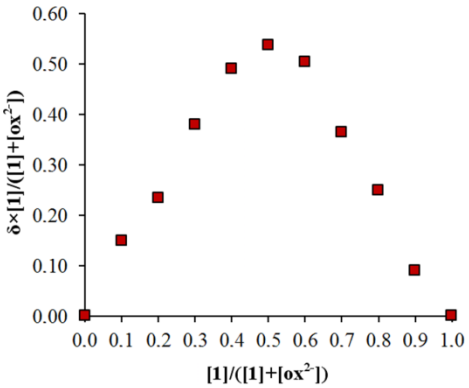

Figure S31. Job plot for **1**·ox^{2-} in $CDCl_3$.

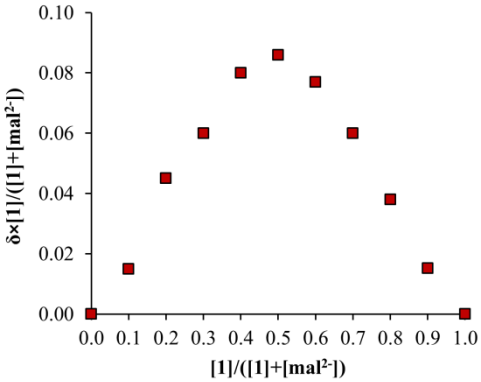

Figure S32. Job plot for **1**·mal^{2-} in $CDCl_3$.

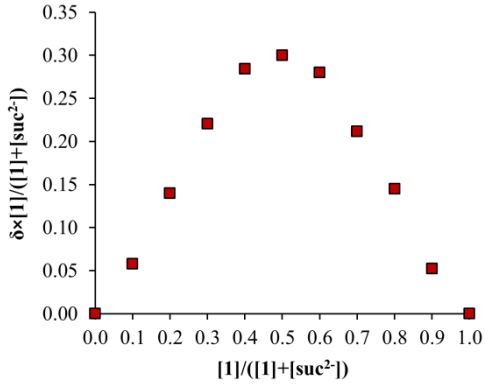

Figure S33. Job plot for **1**·suc^{2-} in $CDCl_3$.

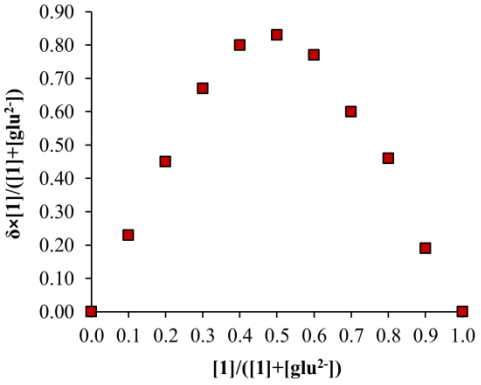

Figure S34. Job plot for **1**·glu^{2-} in $CDCl_3$.

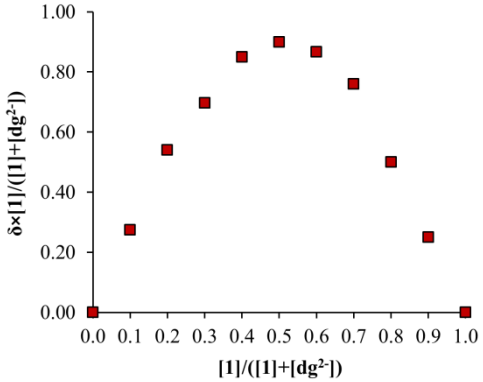

Figure S35. Job plot for **1**·dg^{2-} in $CDCl_3$.

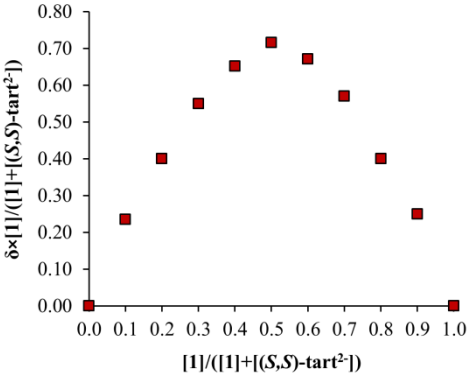

Figure S36. Job plot for **1**·(*S*,*S*)-$tart^{2-}$ in $CDCl_3$.

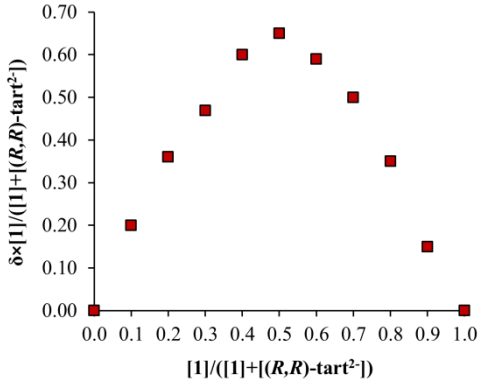

Figure S37. Job plot for **1**·(*R*,*R*)-$tart^{2-}$ in $CDCl_3$.

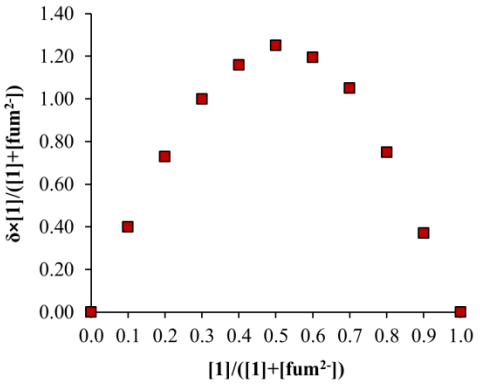

Figure S38. Job plot for **1**·fum^{2-} in $CDCl_3$.

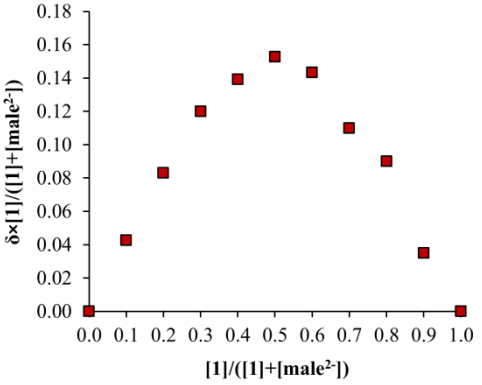

Figure S39. Job plot for **1**·$male^{2-}$ in $CDCl_3$.

Molecular dynamics

Table S5. Receptor's $N_{urea}\cdots N_{urea}$ distances for the associations between **1** and ox^{2-}, mal^{2-}, suc^{2-}, glu^{2-}, dg^{2-}, fum^{2-} and $male^{2-}$.

Anion association	Replicate	Average ± SD	Range [min ; max]
ox^{2-}	R1 (50 ns)	7.703 ± 0.684	[5.260 ; 9.367]
	R2 (50 ns)	7.760 ± 0.596	[5.530 ; 9.747]
	R3 (50 ns)	7.755 ± 0.724	[5.279 ; 10.034]
	All replicates (150 ns)	*7.739 ± 0.671*	*[5.260 ; 10.034]*
mal^{2-}	R1 (50 ns)	8.499 ± 0.662	[5.690 ; 10.445]
	R2 (50 ns)	8.756 ± 0.648	[5.908 ; 10.793]
	R3 (50 ns)	7.771 ± 0.818	[5.163 ; 10.614]
	All replicates (150 ns)	*8.342 ± 0.826*	*[5.163 ; 10.793]*
suc^{2-}	R1 (50 ns)	9.033 ± 0.711	[5.771 ; 11.261]
	R2 (50 ns)	9.252 ± 0.720	[5.972 ; 11.403]
	R3 (50 ns)	9.454 ± 0.541	[6.405 ; 11.666]
	All replicates (150 ns)	*9.246 ± 0.685*	*[5.771 ; 11.666]*
glu^{2-}	R1 (50 ns)	9.805 ± 1.109	[6.165 ; 12.421]
	R2 (50 ns)	10.344 ± 0.679	[6.586 ; 12.351]
	R3 (50 ns)	10.286 ± 0.775	[6.039 ; 13.032]
	All replicates (150 ns)	*10.145 ± 0.907*	*[6.039 ; 13.032]*
dg^{2-}	R1 (50 ns)	10.294 ± 1.087	[6.307 ; 12.405]
	R2 (50 ns)	11.391 ± 0.335	[9.406 ; 12.615]
	R3 (50 ns)	10.542 ± 0.609	[6.249 ; 12.619]
	All replicates (150 ns)	*10.743 ± 0.881*	*[6.249 ; 12.619]*
fum^{2-}	R1 (50 ns)	9.683 ± 0.690	[6.000 ; 12.233]
	R2 (50 ns)	9.468 ± 0.715	[6.357 ; 11.776]
	R3 (50 ns)	9.658 ± 0.518	[7.105 ; 11.395]
	All replicates (150 ns)	*9.603 ± 0.654*	*[6.000 ; 12.233]*
$male^{2-}$	R1 (50 ns)	9.401 ± 0.529	[6.558 ; 11.368]
	R2 (50 ns)	5.254 ± 1.211	[2.874 ; 12.158]
	R3 (50 ns)	9.411 ± 0.494	[6.549 ; 11.211]
	R4 (50 ns)	7.228 ± 2.641	[2.868 ; 11.136]
	R5 (50 ns)	9.340 ± 1.306	[3.505 ; 11.313]
	R6 (50 ns)	9.437 ± 0.734	[5.346 ; 11.342]
	R7 (50 ns)	5.463 ± 2.119	[2.856 ; 15.866]
	All replicates (350 ns)	*7.933 ± 2.332*	*[2.856 ; 15.866]*

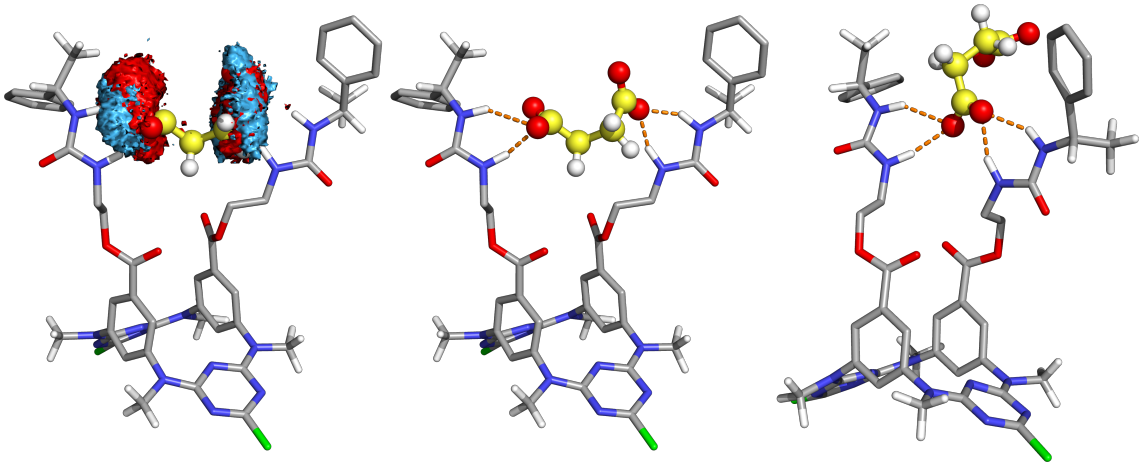

Figure S40. 3-D histogram for the third MD run of **1**·suc^{2-} (left), and the respective type A (center) and type B (right) binding scenarios. Each carboxylate unit (shown either in red or blue) is represented by the positions occupied by its oxygen atoms, 5 or more times.

Table S6. Average $^{-}O_2C\cdots CO_2^{-}$ distances collected for the associations between **1** and ox^{2-}, mal^{2-}, suc^{2-}, glu^{2-}, dg^{2-}, $male^{2-}$ and fum^{2-}.

Anion association	Average ± SD	Range [min ; max]
ox^{2-}	2.907 ± 0.049	[2.694 ; 3.111]
mal^{2-}	3.498 ± 0.083	[3.126 ; 3.865]
suc^{2-}	4.692 ± 0.419	[3.424 ; 5.313]
glu^{2-}	5.812 ± 0.210	[4.642 ; 6.376]
dg^{2-}	5.908 ± 0.253	[4.186 ; 6.510]
fum^{2-}	5.048 ± 0.064	[4.747 ; 5.333]
$male^{2-}$	4.316 ± 0.099	[3.767 ; 4.776]

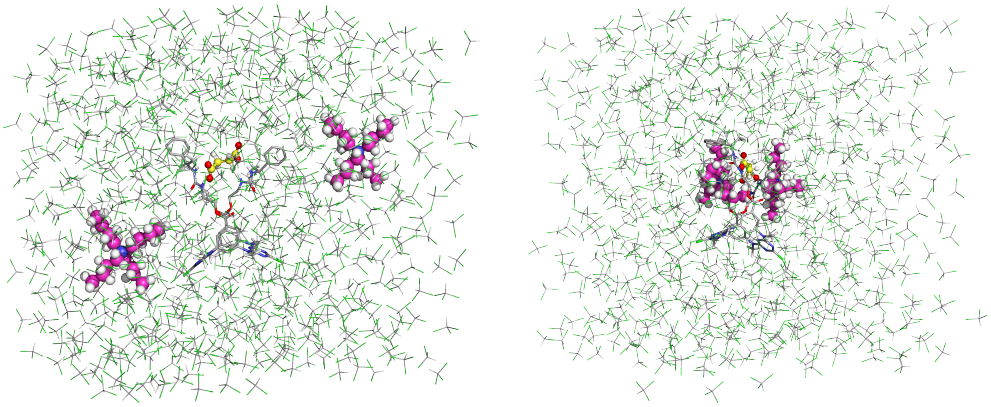

Figure S41. MD simulation cell of the system with the suc^{2-} anion complex, showing the position of the TBA counter-ions in the beginning (left) and at the end of the equilibration stage (right). The TBA molecules are shown in space-filling mode, with the carbon atoms in magenta, while the solvent molecules are shown in lines. Remaining details as given in Figure S40.

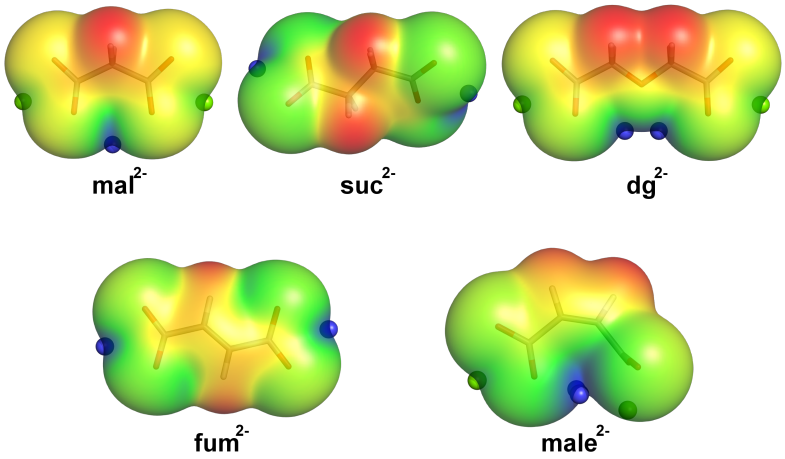

Figure S42. Electrostatic potential mapped on the molecular electron density surface (0.001 electrons per Bohr3) of mal^{2-}, suc^{2-}, dg^{2-}, fum^{2-} and $male^{2-}$. The $V_{S,min}$ and $V_{S,carb}$ are respectively represented by blue and green dots on the molecular surfaces. The colour scales range from blue to red as follows: mal^{2-} [-242.8; -158.5 kcal mol^{-1}]; suc^{2-} [-207.7; -152.0 kcal mol^{-1}]; dg^{2-} [-222.4; -135.2 kcal mol^{-1}]; fum^{2-} [-206.7 ; -160.3 kcal mol^{-1}]; and $male^{2-}$ [-229.8 ; -132.9 kcal mol^{-1}].

Table S7. Additional force field parameters used to model the tetraazacalix[2]arene[2]triazine derivatives.

Bonds	r_{eq} (Å)	K_f (kcal mol^{-1} $Å^{-2}$)		
NH-CX	351.80	1.440		
NH-CT	462.60	1.355		
Angles	θ_{eq} (deg)	K_f (kcal mol^{-1} rad^{-2})		
nb-CT-NH	73.450	116.950		
NH-CX-ca	67.440	120.130		
hn-NH-CT	49.360	116.130		
hn-NH-CX	46.780	116.130		
CT-NH-CX	63.670	123.53		
CT-NH-c3	64.740	117.770		
CX-NH-c3	63.010	117.770		
CT-NH-ca	64.560	127.460		
CX-NH-ca	62.510	127.460		
Dihedral angles	IDIVF	V_n (kcal mol^{-1})	γ (deg)	n
X -CX-NH-X	4	14.5	180.0	2
X -CT-NH-X	2	9.6	180.0	2
van der Waals	ε (kcal mol^{-1})	σ (Å)		
CT	1.9080	0.0860		
CX	1.9080	0.0860		
NH	1.8240	0.1700		

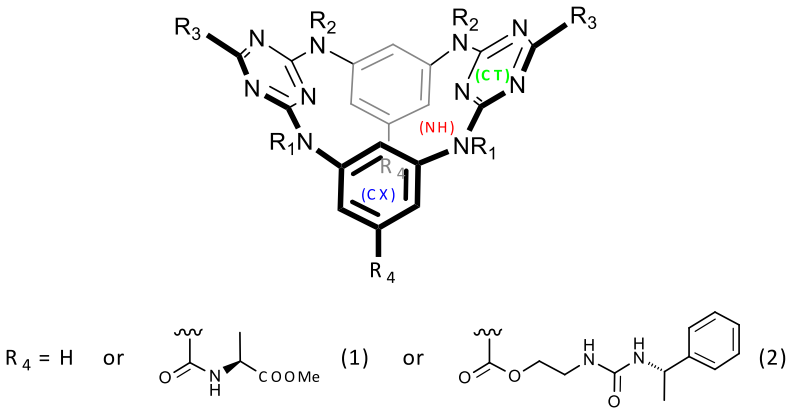

Figure S43. Sketch of tetraazacalix[2]arene[2]triazine derivatives showing the atom types' positions. The R_1, R_2 and R_3 substituents are given in Table S8.

Table S8. RMSD values (Å) between the single crystal structures of tetraazacalix[2]arene[2]triazine derivatives and gas-phase structures obtained MM energy minimization with GAFF default parameters (GFD) and optimized parameters (OPT), respectively.

Macrocycle	Substituents				RMSDs [a), b)]			
	R_1	R_2	R_3	R_4	GFD		OPT	
1 [c)]	Me	Me	Cl	(2)	1.120	*0.182*	2.48	*0.412*
FAYZUB [d)]	H	H	Cl	H	0.990	*0.895*	0.493	*0.372*
FAZBAK [d)]	H	Me	Cl	H	0.264	*0.258*	0.225	*0.231*
FAZBIS [d)]	Me	Me	Cl	H	0.398	*0.279*	0.286	*0.232*
YESRIY [d)]	PMB [e)]	H	Cl	H	1.743	*0.655*	0.312	*0.256*
YESROE [d)]	H	Bn [f)]	Cl	H	0.786	*0.542*	0.434	*0.174*
YESREU [d)]	PMB [e)]	Bn [f)]	Cl	H	1.291	*0.524*	0.497	*0.200*
YESRUK [d)]	PMB [e)]	PMB [e)]	NMe_2	H	1.538	*0.568*	1.242	*0.236*
YESSAR [d)]	Bn [f)]	Bn [f)]	Cl	H	0.683	*0.329*	0.615	*0.165*
VUMNOH [d)]	H	H	OMe	H	0.584	*0.535*	0.203	*0.203*
NAQDEQ [d)]	H	H	Cl	(1)	1.818	*0.649*	0.572	*0.218*
L2 [g)]	Me	Me	Cl	(1)	1.168	*0.253*	1.418	*0.430*
BIVCUG [d)]	H	H	$N(C_6H_{13})_2$	H	1.150	*0.605*	0.559	*0.280*
L3 [g)]	Me	Me	$N(C_6H_{13})_2$	H	2.102	*0.574*	0.419	*0.135*
BIVCIU [d)]	CH_2CO_2Me	CH_2CO_2Me	Cl	H	0.759	*0.294*	0.636	*0.163*
BIVCOA [d)]	CH_2CO_2Me	CH_2CO_2Me	NMe_2	H	1.642	*0.584*	0.525	*0.300*

[a)] All RMSD values calculated excluding the hydrogen atoms; [b)] Values in italic are for tetraazacalix[2]arene[2]triazine plataform; [c)] This work; [d)] CSD REFCODES; [e)] *p*-metoxi-benzene; [f)] benzyl; [g)] unpublished results.