Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2017

Electronic Supplementary Information for:

Entropy-driven homochiral self-sorting of a dynamic library

Joan Atcher, Jordi Bujons and Ignacio Alfonso*

Dr. J. Atcher, Dr. J. Bujons and Dr. I. Alfonso
Department of Biological Chemistry and Molecular Modeling. Institute of Advanced Chemistry of Catalonia, IQAC-CSIC
Jordi Girona 18-26, 08034, Barcelona, Spain. Fax: (+34)932045904
e-mail: ignacio.alfonso@iqac.csic.es

TABLE OF CONTENTS

General characteristics

Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka or Merck) and were used without further purification. Flash chromatographic purifications and preparative reversed-phase purifications were performed on Biotage® Isolera Prime™ equipment. TLCs were performed using 6 x 3 cm SiO_2 pre-coated aluminium plates (ALUGRAM® SIL G/UV_{254}).

Reversed-phase high-performance liquid chromatography (RP-HPLC) analyses were performed on a Hewlett Packard Series 1100 (UV detector 1315A) modular system using:

i) For the characterization of (*R*)-**1**-d_4: a reversed-phase X-Terra C_{18} (15 x 0.46 cm, 5 µm) column. CH_3CN + 0.07% TFA and H_2O + 0.1% TFA mixtures at 1 mL/min were used as mobile phase and the monitoring wavelength was set at 254 nm.

ii) For the analysis of the DCLs: a reversed-phase CHIRALPAK® IA (25 x 0.46 cm, 5 µm) chiral column. CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH mixtures at 0.5 mL/min were used as mobile phase. The monitoring wavelength was set at 254 nm and the temperature was set at 25ºC.

Nuclear magnetic resonance (NMR) spectroscopic experiments were carried out on a Varian INOVA 500 spectrometer (500 MHz for ^1H and 125 MHz for ^{13}C) and a Varian Mercury 400 instrument (400 MHz for ^1H and 101 MHz for ^{13}C). The chemical shifts (δ) are reported in ppm relative to trimethylsilane (TMS), and coupling constants (*J*) are reported in Hertz (Hz). Signal assignment was carried out using the necessary 2D NMR spectra including ^1H-^1H gCOSY, ^1H-^{13}C gHSQC and ^1H-^{13}C gHMBC. For describing signals of ^1H NMR spectra, the following abbreviations are used: s = singlet, d = doublet, t = triplet, q = quartet, ABq = AB quartet, dd = doublet of doublets and m = multiplet. The asymmetric carbons are indicated by an asterisk.

Absorbance measurements were performed on a Molecular Devices SpectraMax M at room temperature.

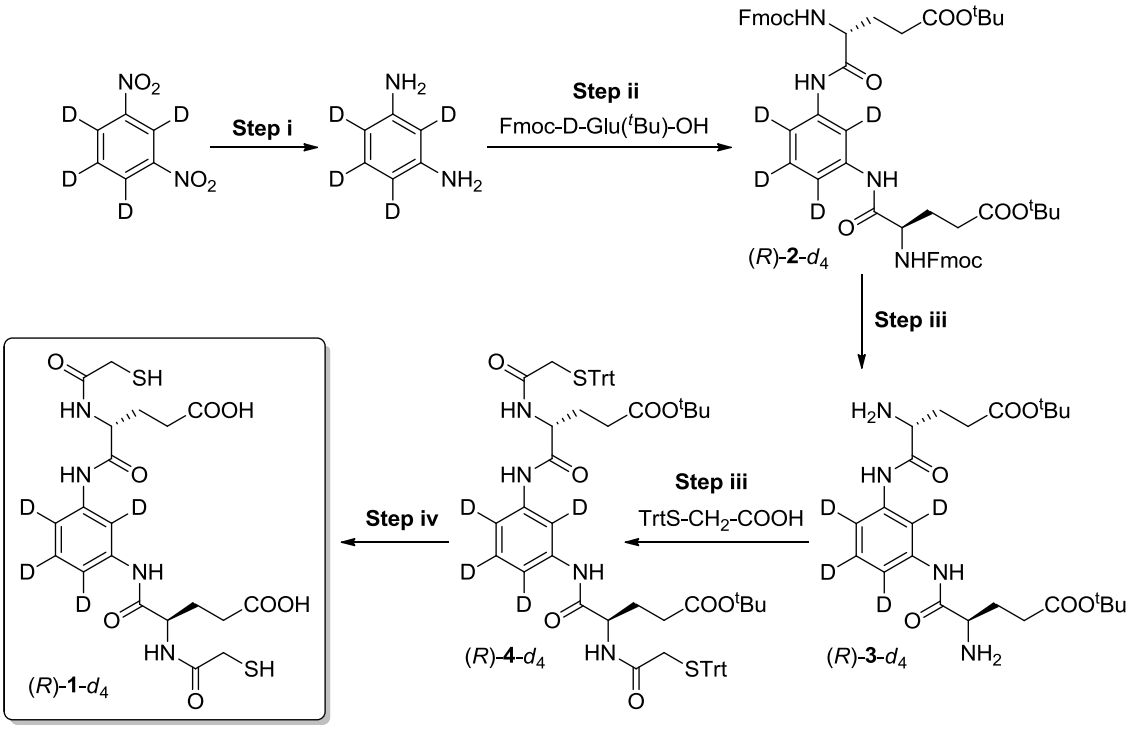
pH measurements were performed at room temperature on a Crison GLP21 pH-meter with the electrodes Crison 50 14T (≥10 mL samples) and PHR-146 Micro (<10 mL samples).

High resolution mass spectrometry (HRMS) analyses were carried out at the IQAC Mass Spectrometry Facility, using UPLC-ESI-TOF equipment: Acquity UPLC® BEH C_{18} 1.7 mm, 2.1x100 mm, LCT Premier Xe, Waters. CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH mixtures at 0.3 mL/min were used as mobile phase.

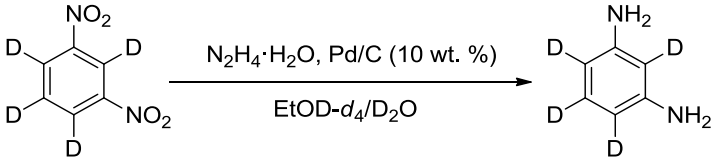
Experimental procedure for the synthesis of the BBs

Tritylsulfanyl acetic acid was prepared as previously described.[1] Compound (*S*)-**1** was also synthesized as we previously reported.[2]

Synthetic scheme

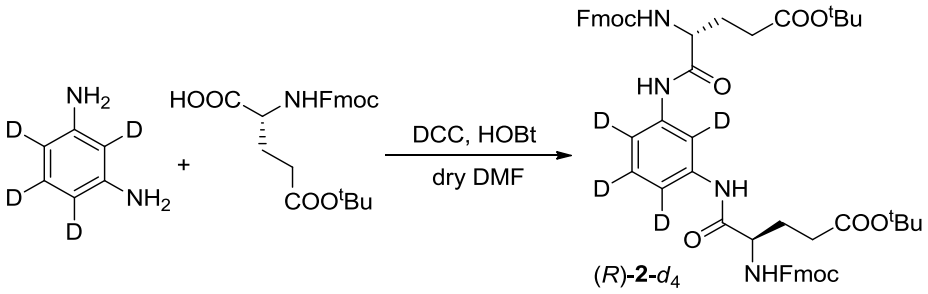


Step i: synthesis of *m*-phenylenediamine-2,4,5,6-d_4



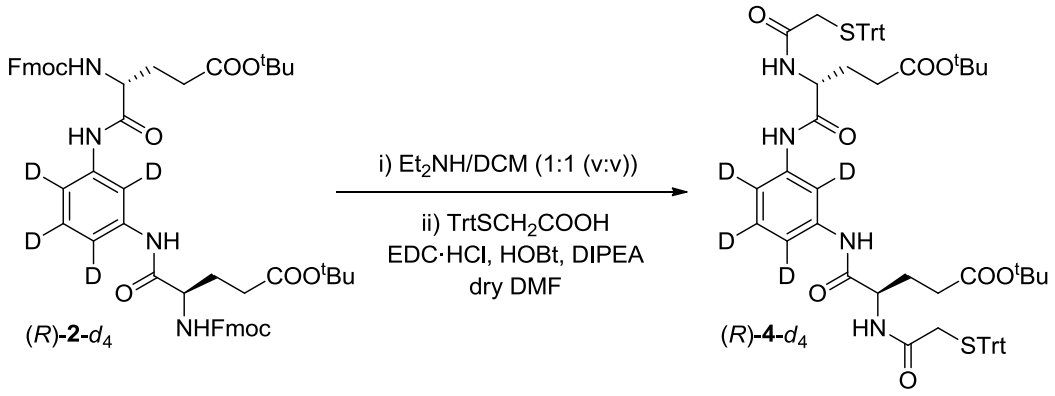
A mixture of 1,3-dinitrobenzene-d_4 (99% atom D, 723 mg, 4.20 mmol), 50 mg of Pd/C (10 wt. %), hydrazine monohydrate (8 mL, 165 mmol), 4.0 mL of EtOD-d_4, and 8.0 mL of D_2O was refluxed with stirring for 1 h, after which complete conversion of the starting material was observed by TLC (Rf of the product in AcOEt/Hexane, 7:3 (v:v): 0.34). After Celite® filtration with abundant EtOH, the filtrate was concentrated under reduced pressure and purified by flash chromatography using hexane/AcOEt as eluent (from 60% to 70% AcOEt) to give 418 mg of *m*-phenylenediamine-2,4,5,6-d_4 (89% yield) as a pale white solid. HRMS (ESI+) calcd. for $C_6H_4D_4N_2$ $[M+H]^+$ (m/z): 113.1011, found: 113.1006. ^{13}C NMR (101 MHz, MeOD-d_4): δ = 149.0 (2 x C_{Ar}), 130.2 (1 x CD_{Ar}), 107.3 (2 x CD_{Ar}), 104.1 (1 x CD_{Ar}).

Step ii: synthesis of (*R*)-**2**-d_4



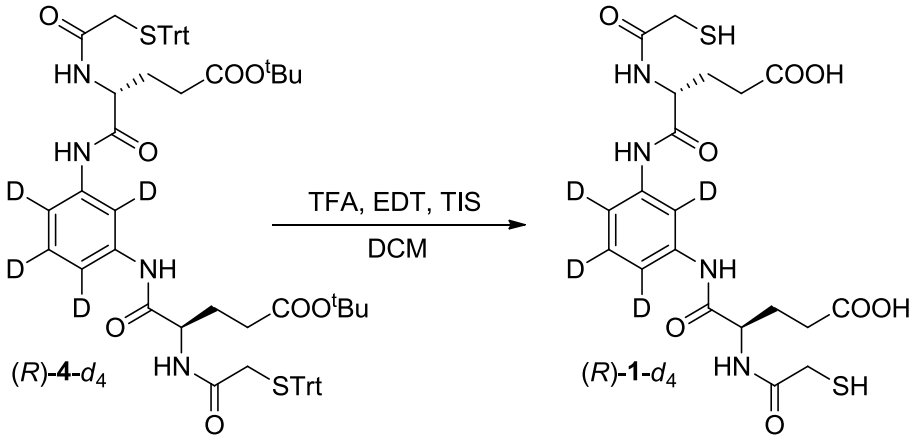
Dicyclohexylcarbodiimide (DCC, 2.41 g, 11.7 mmol) and 1-hydroxybenzotriazole (HOBt, 1.37 g, 10.1 mmol) were added over a solution of Fmoc-D-Glu(^{t}Bu)-OH (3.43 g, 8.05 mmol) in dry DMF (15 mL). The resulting mixture was cooled to 0 °C in an ice-water bath. Then a solution of *m*-phenylenediamine-2,4,5,6-d_4 (362 mg, 3.22 mmol) in dry DMF (10 mL) was added over the mixture through a cannula. The resulting solution was stirred at room temperature for 60 hours in an inert atmosphere of Ar, after which complete conversion of the starting material was observed by TLC (Rf of the product in AcOEt/Hexane, 2:3 (v:v): 0.43). The mixture was filtered, and the filtrate was diluted with AcOEt, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography using hexane/AcOEt as eluent (from 30% to 40% AcOEt) to give 2.66 g of (*R*)-**2**-d_4 (89% yield) as a white solid. HRMS (ESI+) calcd. for $C_{54}H_{54}D_4N_4O_{10}$ $[M+H]^+$ (m/z): 927.4477, found: 927.4473. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.69 (s, 2H, NHCOC*H), 7.74 (d, *J* = 7.6 Hz, 4H, CH_{Ar}), 7.58 (t, *J* = 7.9 Hz, 4H, CH_{Ar}), 7.37 (t, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.32–7.17 (m, 4H, CH_{Ar}), 5.99 (d, *J* = 7.7 Hz, 2H, NHFmoc), 4.50–4.28 (m, 6H: 4H, CH_2 + 2H, C*H), 4.20 (t, *J* = 7.0 Hz, 2H, CH), 2.58–2.43 (m, 2H, CH_2COO^{t}Bu), 2.42–2.28 (m, 2H, CH_2COO^{t}Bu), 2.21–2.08 (m, 2H, CH_2C*H), 2.06–1.88 (m, 2H, CH_2C*H), 1.44 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 173.2 (2 x CO), 169.9 (2 x CO), 156.7 (2 x CO), 143.8 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 138.1 (2 x C_{Ar}), 129.0 (1 x CD_{Ar}), 127.9 (4 x CH_{Ar}), 127.2 (4 x CH_{Ar}), 125.2 (4 x CH_{Ar}), 120.1 (4 x CH_{Ar}), 115.8 (2 x CD_{Ar}), 111.1 (1 x CD_{Ar}), 81.4 (2 x CMe_3), 67.4 (2 x CH_2), 55.2 (2 x C*H), 47.2 (2 x CH), 32.0 (2 x CH_2COO^{t}Bu), 28.2 (6 x CH_3), 24.9 (2 x CH_2C*H).

Step iii: synthesis of (*R*)-**4**-d_4



To a solution of (*R*)-**2**-d_4 (2.30 g, 2.48 mmol) in DCM (25 mL), 25 mL of diethylamine were added. The mixture was stirred at room temperature for 12 hours and then concentrated under reduced pressure, obtaining the corresponding diamine (*R*)-**3**-d_4 as a white impure residue. HRMS (ESI+) calcd. for $C_{24}H_{34}D_4N_4O_6$ $[M+H]^+$ (m/z): 483.3115, found: 483.3091. Tritylsulfanyl acetic acid (1.92 g, 5.74 mmol) was dissolved in dry DMF (20 mL) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC·HCl, 1.65 g, 8.62 mmol), HOBt (1.02 g, 7.55 mmol) and *N*,*N*-diisopropylethylamine (DIPEA, 4.0 mL, 23.0 mmol) were added over the solution. The reaction mixture was cooled down to 0 °C in an ice-water bath and then the crude residue containing (*R*)-**3**-d_4 was added over the mixture. The solution was stirred at room temperature for 24 hours in an inert atmosphere, and the formation of the product was followed by TLC (Rf of the product in AcOEt/Hexane, 2:3 (v:v): 0.20). The mixture was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, and dried under reduced pressure. The residue was purified by flash chromatography using hexane/AcOEt as eluent (from 30% to 40% AcOEt) to give 1.94 g of (*R*)-**4**-d_4 (70% yield over the two steps) as a white solid. HRMS (ESI+) calcd. for $C_{66}H_{66}D_4N_4O_8S_2$ $[M+H]^+$ (m/z): 1115.4959, found: 1115.4965. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.65 (s, 2H, N<u>H</u>COC*H), 7.44–7.37 (m, 12H, CH_{Ar}), 7.32–7.23 (m, 12H, CH_{Ar}), 7.22–7.15 (m, 6H, CH_{Ar}), 6.77 (d, *J* = 7.2 Hz, 2H, CON<u>H</u>C*H), 4.26 (q, *J* = 6.8 Hz, 2H, C*H), 3.13 (ABq, δ_A = 3.12, δ_B = 3.15, *J* = 16.0 Hz, 4H, C<u>H</u>$_2$STrt), 2.50–2.35 (m, 2H, C<u>H</u>$_2$COO^{t}Bu), 2.30–2.16 (m, 2H, C<u>H</u>$_2$COO^{t}Bu), 2.10–1.96 (m, 2H, C<u>H</u>$_2$C*H), 1.85–1.70 (m, 2H, C<u>H</u>$_2$C*H), 1.44 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 172.9 (2 x CO), 169.0 (2 x CO), 168.9 (2 x CO), 144.1 (6 x C_{Ar}), 138.3 (2 x C_{Ar}), 129.6 (12 x CH_{Ar} + 1 x CD_{Ar}), 128.3 (12 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 115.3 (2 x CD_{Ar}), 110.8 (1 x CD_{Ar}), 81.2 (2 x <u>C</u>Me$_3$), 68.1 (2 x <u>C</u>Ph$_3$), 53.5 (2 x C*H), 36.2 (2 x <u>C</u>H$_2$STrt), 31.8 (2 x <u>C</u>H$_2$COO^{t}Bu), 28.2 (6 x CH_3), 27.9 (2 x <u>C</u>H$_2$C*H).

Step iv: Experimental procedure for the synthesis of (*R*)-**1**-d_4



Compound (*R*)-**4**-d_4 (233 mg, 0.209 mmol) was dissolved in DCM (1 mL) and 8.0 mL of trifluoroacetic acid (TFA), triisobutylsilane (TIS, 433 µL, 1.67 mmol) and 1,2-ethanedithiol (EDT, 210 µL, 2,50 mmol) were added rapidly and under stirring. The reaction mixture was stirred at room temperature for 2 hours, after which the solvents were partially evaporated using a N_2 flow. Diethyl ether was added over the reaction mixture and the product was filtered off and washed with diethyl ether. The product was purified by reversed-phase flash chromatography using a mixture of CH_3CN and H_2O (containing 0.07% and 0.1% TFA respectively) as mobile phase (gradient: from 5% to 30% CH_3CN in H_2O), and 49.9 mg of (*R*)-**1**-d_4 (46% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{20}H_{22}D_4N_4O_8S_2$ $[M+H]^+$ (m/z): 519.1516, found: 519.1505. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 4.53 (dd, *J* = 8.7, 5.3 Hz, 2H, C*H), 3.24 (s, 4H, C$\underline{H}_2$SH), 2.45 (t, *J* = 7.5 Hz, 4H, C$\underline{H}_2$COOH), 2.25–2.12 (m, 2H, C$\underline{H}_2$C*H), 2.09–1.95 (m, 2H, C$\underline{H}_2$C*H). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 176.4 (2 x COOH), 173.5 (2 x $\underline{C}OCH_2$), 171.8 (2 x $\underline{C}$OC*H), 139.8 (2 x C_{Ar}), 130.1 (1 x CD_{Ar}), 117.3 (2 x CD_{Ar}), 113.1 (1 x CD_{Ar}), 54.9 (2 x C*H), 31.1 (2 x $\underline{C}H_2$COOH), 28.7 (2 x $\underline{C}H_2$C*H), 28.1 (2 x CH_2SH).

Characterization of (*R*)-**1**-d_4

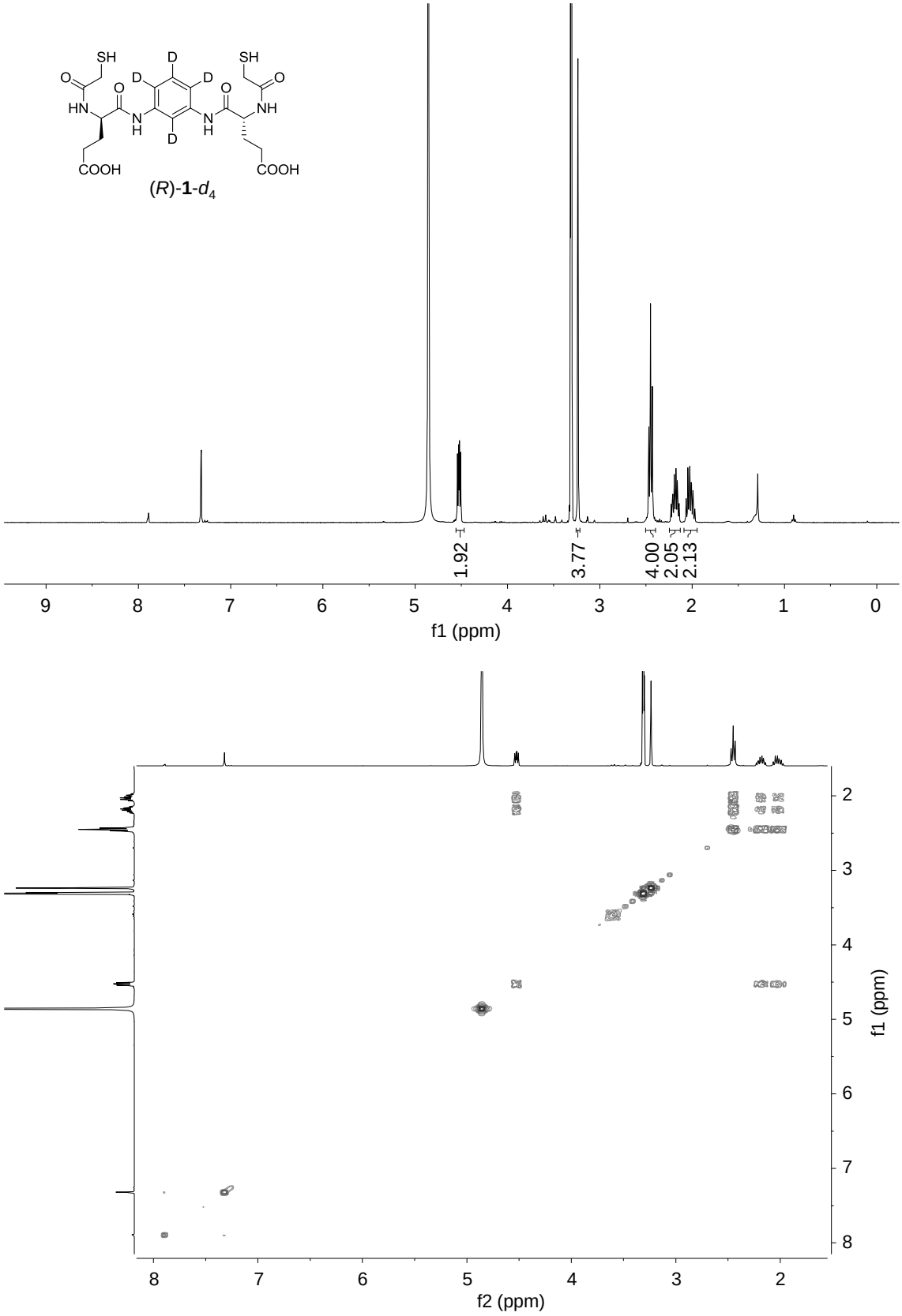


Fig. S1: ^{1}H (400 MHz, 298 K in MeOD-d_4) and ^{1}H-^{1}H gCOSY (400 MHz, 298 K in MeOD-d_4) spectra of (*R*)-**1**-d_4.

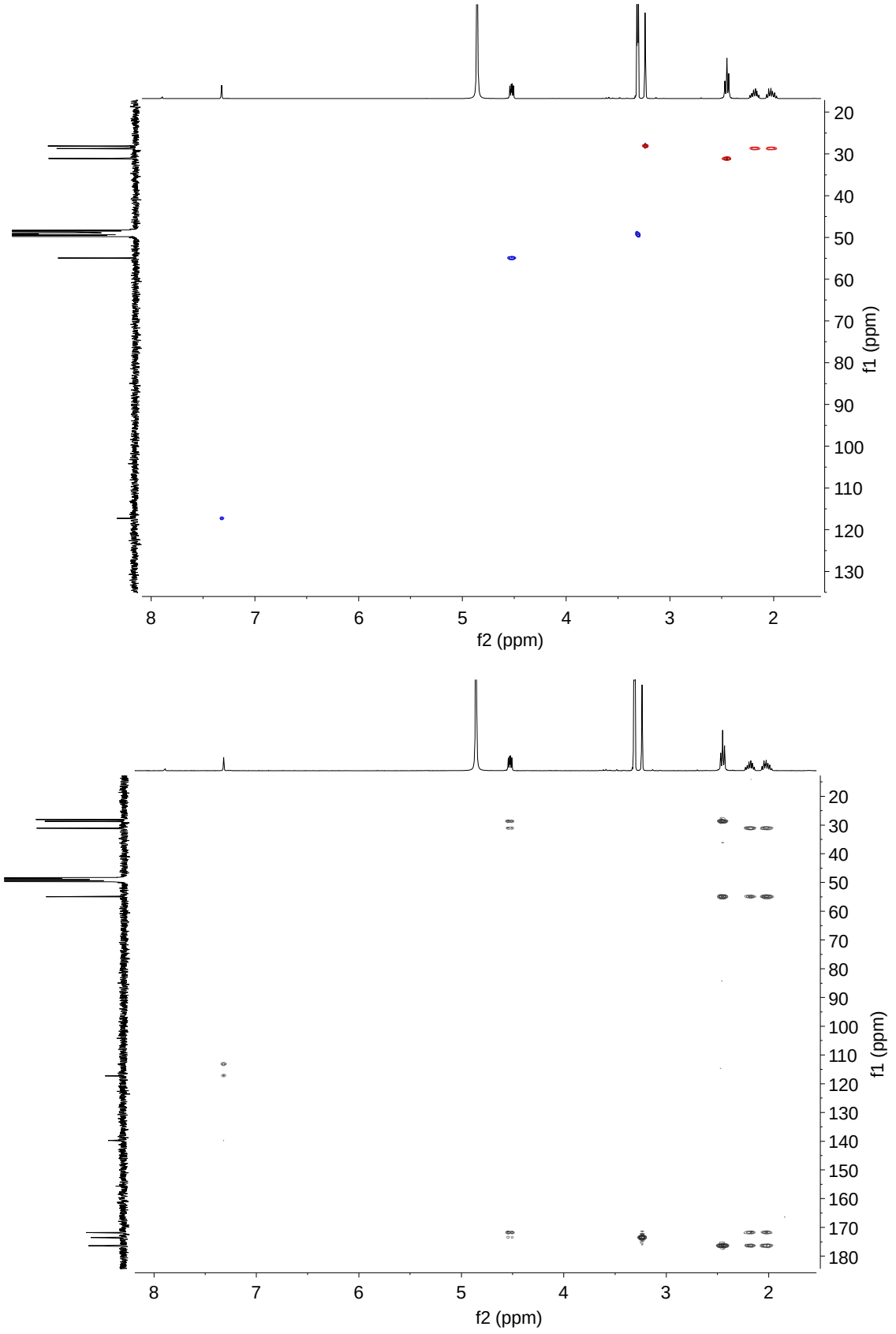


Fig. S2: ^{1}H-^{13}C gHSQC (400 MHz, 298 K in MeOD-d_4) and ^{1}H-^{13}C gHMBC (400 MHz, 298 K in MeOD-d_4) spectra of (*R*)-**1**-d_4.

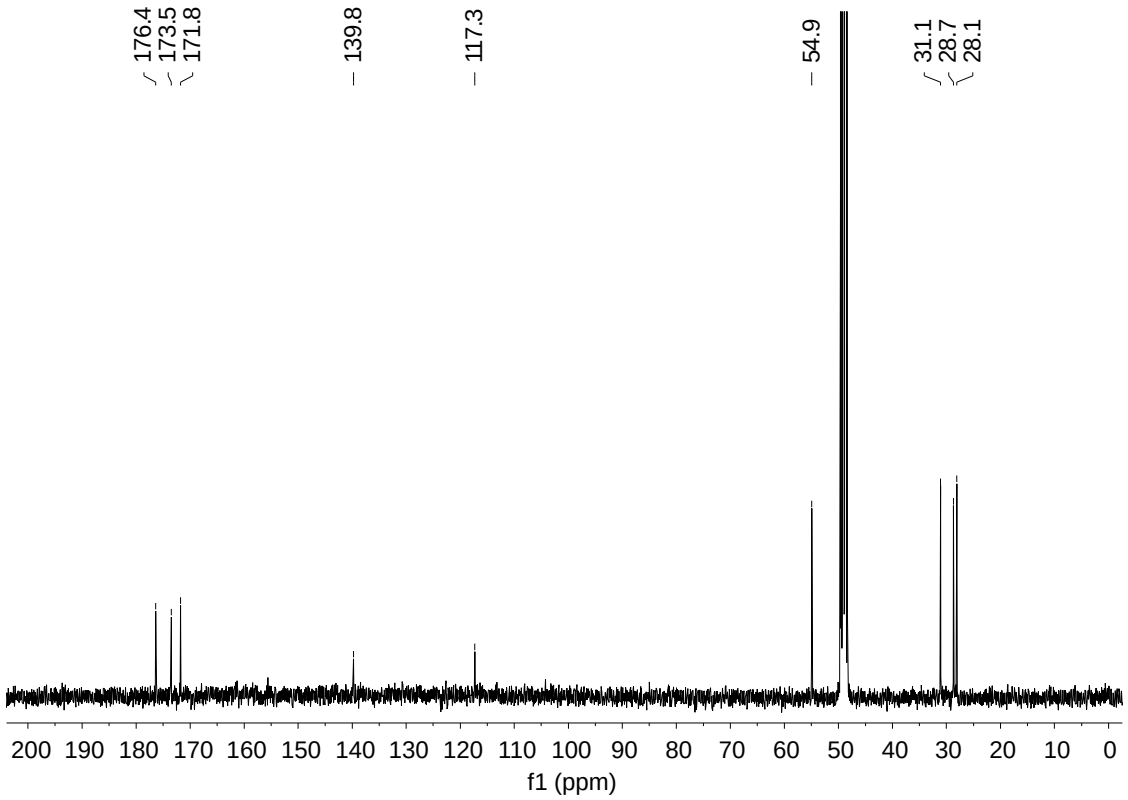


Fig. S3: ^{13}C (101 MHz, 298 K in MeOD-d_4) spectrum of (*R*)-**1**-d_4.

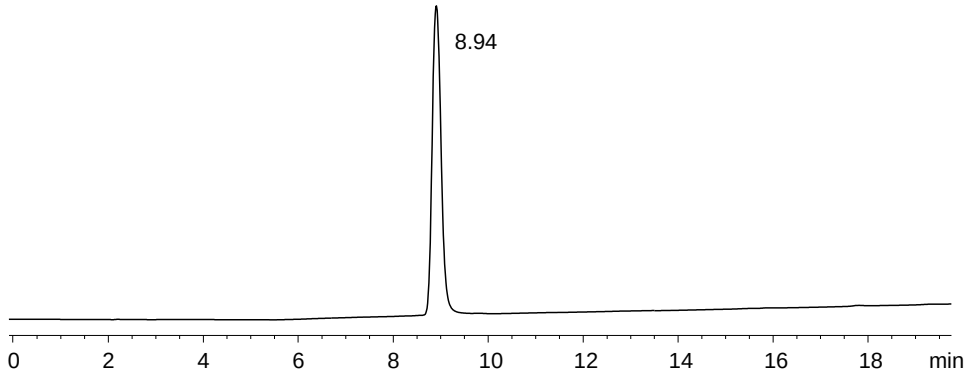


Fig. S4: RP-HPLC analysis of (*R*)-**1**-d_4 (eluent: mixture of CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA; gradient: 2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

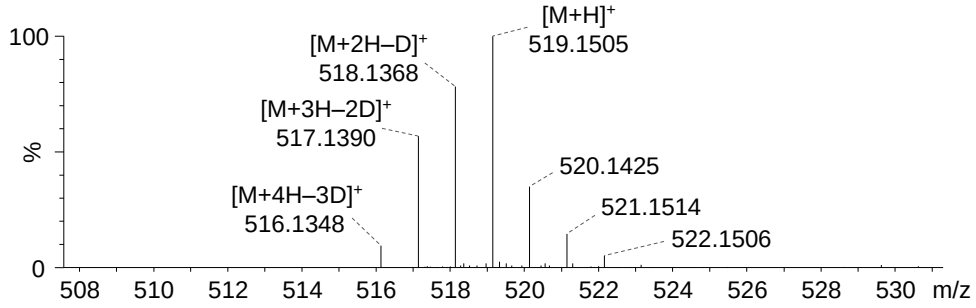


Fig. S5: Experimental ESI-TOF mass spectrum for $[M+H]^+$ of (*R*)-**1**-d_4.

Experimental procedure for the preparation of the DCLs

Individual stocks and *stock mixture*

For each of dithiols (*S*)-**1** and (*R*)-**1**-d_4 a 14.7 mM *individual stock* was prepared in basified water. First of all, the corresponding amount of solid was added to the needed amount of Milli-Q water, obtaining an acidic suspension. Then, a small volume of a 1.0 M TBAOH aqueous solution was added in order to adjust the pH to ~8. As the TBAOH solution was added, the suspension became clearer and eventually all solid was completely dissolved. A *stock mixture* of the two BBs (7.35 mM of each) was prepared by mixing equal volumes of the two *individual stocks*. The pH of the resulting mixture was measured, confirming to be ~8. The stock mixture was stored at –80 °C.

DCLs at different %AN and temperatures

The dynamic libraries were prepared at 0.5 mM of each BB by dilution of the *stock mixture* of (*S*)-**1** + (*R*)-**1**-d_4 (7.35 mM of each) with the corresponding amount of aqueous solution (phosphate buffer (pH 8.0) in the absence and presence of guanidinium chloride for entries 1-2 in Tables S1 and S2, and Milli-Q water for entries 3-9 in Tables S1 and S2) and organic co-solvents (1.2% (v/v) DMSO for entries 1-9 in Tables S1 and S2, and the corresponding volume % of AN for entries 4-9 in Tables S1 and S2). Depending on the solvent composition and temperature, different reaction times were needed to fully oxidize the dithiol building blocks (Table S1). The mixtures were analyzed by chiral-HPLC and the corresponding K_{HS} and $-\Delta\Delta G_i$ values were determined (Table S2).

Table S1: Reaction time of the oxidations carried out at 2, 22 and 45 °C.

entry	sample	χ_{AN}	2 °C	22 °C	45 °C
1	PB+G	0.00	9 d	5 d	— [a]
2	PB	0.00			46 h
3	0%	0.00	10 d	6 d	49 h
4	10%	0.04			53 h
5	30%	0.13			
6	50%	0.26			
7	70%	0.45	11 d	7 d	64 h
8	86%	0.68			
9	92%	0.80			

[a] No reliable quantitative data was obtained at these experimental conditions because of the fast over-oxidation of the library members.

Table S2: K_{HS} constant and $-\Delta\Delta G_i$ energy (kJ·mol^{-1}) values determined for the binary mixtures prepared at different solvent compositions and temperatures. The errors were calculated for a 95% significance.

			2 °C		22 °C		45 °C	
entry	sample	χ_{AN}	K_{HS}	$-\Delta\Delta G_i$	K_{HS}	$-\Delta\Delta G_i$	K_{HS}	$-\Delta\Delta G_i$
1	PB+G	0.00	0.27 ± 0.01	0.16	0.30 ± 0.01	0.42	— [a]	— [a]
2	PB	0.00	0.30 ± 0.01	0.44	0.36 ± 0.01	0.88	0.39 ± 0.01	1.17
3	0%	0.00	0.33 ± 0.01	0.65	0.38 ± 0.01	1.03	0.42 ± 0.01	1.34
4	10%	0.04	0.39 ± 0.01	1.01	0.45 ± 0.02	1.46	0.47 ± 0.01	1.65
5	30%	0.13	0.54 ± 0.01	1.77	0.58 ± 0.01	2.05	0.58 ± 0.01	2.23
6	50%	0.26	0.62 ± 0.01	2.09	0.68 ± 0.01	2.44	0.67 ± 0.01	2.61
7	70%	0.45	0.70 ± 0.02	2.34	0.80 ± 0.02	2.84	0.80 ± 0.01	3.09
8	86%	0.68	1.10 ± 0.04	3.39	1.30 ± 0.08	4.05	1.40 ± 0.05	4.55
9	92%	0.80	2.68 ± 0.06	5.42	2.87 ± 0.09	5.99	3.0 ± 0.2	6.6

[a] No reliable quantitative data was obtained at these experimental conditions because of the fast over-oxidation of the library members.

DCLs at different pHs and concentrations

Six McIlvaine buffer solutions (pH 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5) were prepared as previously reported,[3] by dissolving different amounts of Na_2HPO_4 and citric acid in milli-Q water, and incorporating different amounts of sodium chloride as an inert salt in order to keep the ionic strength of the six solutions constant (0.5 M). For the buffers prepared at pH 6.5 and 7.5 no citrate was used. Instead, these two buffers were prepared dissolving different amounts of Na_2HPO_4 and NaH_2PO_4. Before being used, all buffers were diluted with milli-Q water (two-fold dilution factor) in order to avoid water/AN immiscibility.[4] The dynamic libraries performed at different pH values were prepared at 0.5 mM of each BB by dilution of the *stock mixture* of (*S*)-**1** + (*R*)-**1**-d_4 (7.35 mM of each) with the corresponding amount of aqueous solution (buffers prepared at pH 2.5-7.5) and organic co-solvents (56% (v/v) AN and 1.2% (v/v) DMSO). Different oxidation times were needed depending on the pH of the libraries (Table S3). The mixtures were analyzed by chiral-HPLC and the corresponding K_{HS} and $-\Delta\Delta G_i$ values were determined (Table S3 and Fig. S6).

Table S3: Values of the K_{HS} constant, the $-\Delta\Delta G_i$ energy (kJ·mol^{-1}) and the % amount of Trimers (**1**$_3$), together with the oxidation times required to fully oxidize the binary mixtures prepared at different pH values. The errors were calculated for a 95% significance.

entry	pH	K_{HS}	$-\Delta\Delta G_i$	% Trimers	reaction time
1	2.5	0.38 ± 0.01	1.04	11.8 ± 0.7	8 d
2	3.5	0.47 ± 0.01	1.52	9.9 ± 0.3	7 d
3	4.5	0.52 ± 0.01	1.77	7.3 ± 0.1	
4	5.5	0.34 ± 0.01	1.90	5.3 ± 0.2	6 d
5	6.5	0.56 ± 0.01	1.97	4.2 ± 0.1	
6	7.5	0.56 ± 0.01	1.99	3.7 ± 0.4	5 d

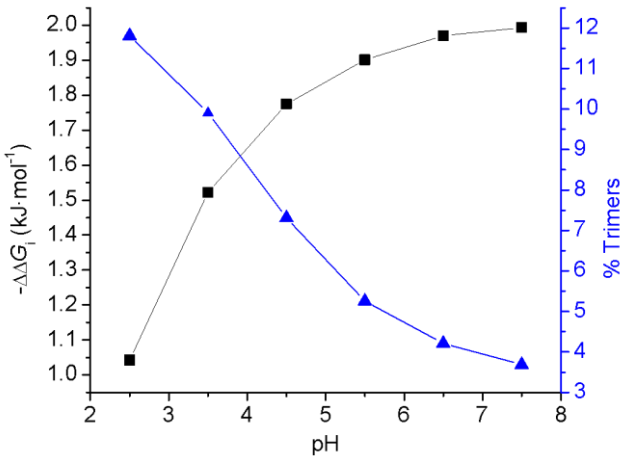


Fig. S6: Plot of the homochiral selectivity energy ($-\Delta\Delta G_i$, black squares) and the % of trimers (blue triangles) *versus* the pH of the medium.

The decrease in the pH clearly slows the oxidation process down. When relatively large amounts of DMSO are used the thiols are rapidly oxidized mainly by reaction with DMSO. However, in the presence of much smaller amounts of DMSO (1.2% (v/v) in this study) the reaction with DMSO is notably slower and the air-oxidation mechanism becomes important. The latter is strongly dependent on the pH, and this is why at low %DMSO the oxidation rate changes with the pH.

The dynamic libraries performed at different concentrations were prepared at 2.0 mM of each BB by dilution of the *stock mixture* with: i) 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO; and ii) TBAOH basified water (pH ~8) with 70% (v/v) AN and 1.2% (v/v) DMSO. The libraries were left to oxidize at 22 °C for 7 and 8 days respectively, until they reached the thermodynamic equilibrium. Finally, the mixtures were analyzed by chiral-HPLC and the corresponding K_{HS} and $-\Delta\Delta G_i$ values were determined in order to be compared with those of the mixtures performed at lower concentration.

Table S4: K_{HS} constant and $-\Delta\Delta G_i$ energy ($kJ \cdot mol^{-1}$) values determined for the binary mixtures performed at 0.5 and 2.0 mM concentration of each BB. The errors were calculated for a 95% significance.

			0.5 mM		2.0 mM	
entry	sample	χ_{AN}	K_{HS}	$-\Delta\Delta G_i$	K_{HS}	$-\Delta\Delta G_i$
1	PB	0.00	0.36 ± 0.01	0.88	0.37 ± 0.01	0.93
2	70%	0.45	0.80 ± 0.02	2.84	0.82 ± 0.01	2.92

As shown in Table S4, both in the presence and absence of AN, the K_{HS} constant has the same value regardless the concentration of the BBs. Since intermolecular processes are intrinsically dependent on the concentration, the observed equal outcome for the experiments performed at different concentration corroborates that no aggregation processes are responsible for the homochiral self-selection.

Identification of the library members

Achiral-HPLC-UV-MS analysis

The HPLC trace of the equimolar binary mixture (Fig. S7a) shows the prevalence of two major peaks. The first one, at the retention time of 30.8 min, corresponds to the dimer joining two BBs with opposed chirality: dimer (*R*,*S*)-**1**$_2$-d_4. The second one, at the retention time of 31.5 min, corresponds to the simultaneous elution of the two homodimers: (*S*,*S*)-**1**$_2$ and (*R*,*R*)-**1**$_2$-d_8. Finally, the minor broad peak at the retention time of 34.7-35.4 min corresponds to the overlap of the four possible trimers (*S*,*S*,*S*)-**1**$_3$, (*R*,*S*,*S*)-**1**$_3$-d_4, (*R*,*R*,*S*)-**1**$_3$-d_8 and (*R*,*R*,*R*)-**1**$_3$-d_{12}.

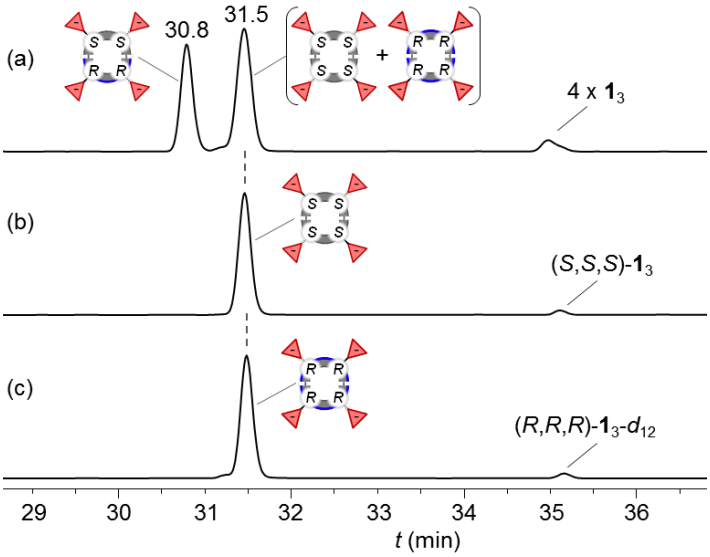


Fig. S7: Achiral-HPLC-UV traces (254 nm) of the DCLs generated from (a) 2.0 mM (*S*,*S*)-**1** + 2.0 mM (*R*,*R*)-**1**-d_4, (b) 2.0 mM (*S*,*S*)-**1**, and (c) 2.0 mM (*R*,*R*)-**1**-d_4; in 20 mM aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO.

The MS analysis provided unambiguous evidences of the assignment of the members of the library. The separately oxidized BBs (*S*)-**1** and (*R*)-**1**-d_4 rendered the isotopic pattern of (*S*,*S*)-**1**$_2$ and (*R*,*R*)-**1**$_2$-d_8 respectively, being the latter 8 m/z units heavier because of the isotopic labeling (Fig. S8a and S8b). Due to partial D/H exchange during the synthesis of *m*-phenylenediamine-d_4, the most intense peak in the isotopic pattern of (*R*,*R*)-**1**$_2$-d_8 (Fig. S8b) does not correspond to the octadeuterated form. By comparison with the isotopic pattern of the two homodimers, the identification of the members of the binary mixture was straightforward: the MS analysis of the heterochiral dimer (*R*,*S*)-**1**$_2$-d_4 formally consists of the weighted isotopic profiles of (*S*,*S*)-**1**$_2$ and (*R*,*R*)-**1**$_2$-d_8 (Fig. S8c), whereas the MS analysis of the peak containing the two simultaneously eluted homochiral dimers ((*S*,*S*)-**1**$_2$ + (*R*,*R*)-**1**$_2$-d_8) contained their superposed isotopic profiles (Fig. S8d). Additionally, the two homochiral dimers in Fig. S7b and S7c showed the same retention time as the second peak in the HPLC trace of the binary mixture (Fig. S7a), in agreement with their assignment.

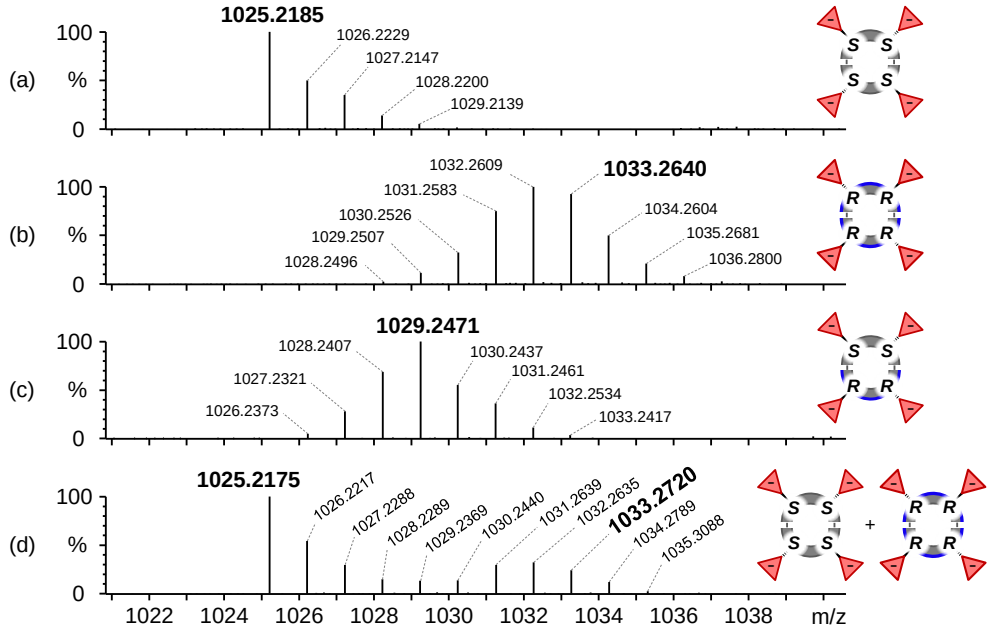


Fig. S8: ESI(+)-TOF isotopic patterns (experimental mass peaks) of (a) (*S*,*S*)-**1**$_2$ (peak at 31.5 min in Fig. S7b), (b) (*R*,*R*)-**1**$_2$-d_8 (peak at 31.5 min in Fig. S7c), (c) (*R*,*S*)-**1**$_2$-d_4 (peak at 30.8 min in Fig. S7a), and (d) (*S*,*S*)-**1**$_2$ + (*R*,*R*)-**1**$_2$-d_8 (peak at 31.5 min in Fig. S7a).

Assignment of the library members in the chiral-HPLC trace

The three previously identified dimers of the binary mixture were assigned in the chiral-HPLC trace by deconvolution experiments. Building blocks (*S*)-**1** and (*R*)-**1**-d_4 were mixed at 0.5 mM each in 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO. Simultaneously, the two BBs were separately oxidized in exactly the same experimental conditions. After 5 days at room temperature, the libraries were analyzed by chiral-HPLC. The peaks of the HPLC trace of the binary mixture were assigned by comparison with the chromatograms of the two separately oxidized building blocks (Fig. S9).

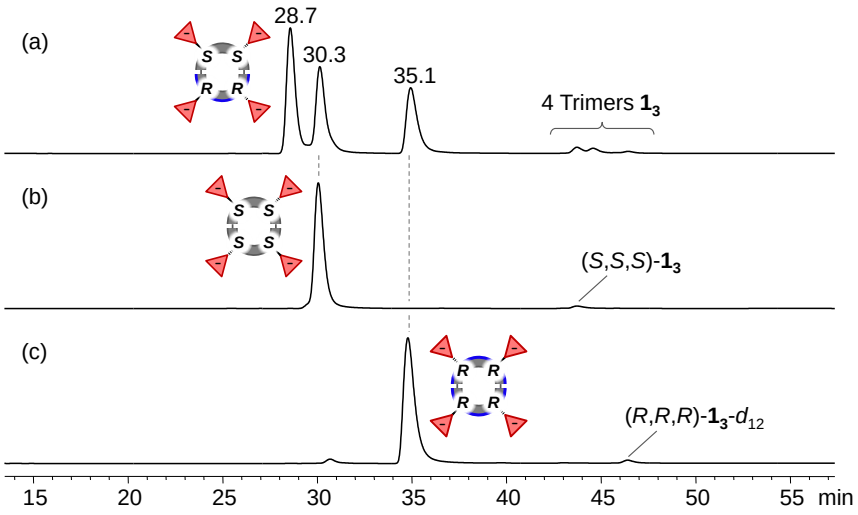


Fig. S9: Chiral-HPLC-UV traces (254 nm) of the DCL generated from (a) (*S*)-**1** + (*R*)-**1**-d_4 (2.0 mM each), (b) (*S*)-**1** (2.0 mM), and (c) (*R*)-**1**-d_4 (2.0 mM); in 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO.

Quantitative evaluation of the DCLs

Homochiral selection exchange constant (K_{HS})

The assumption of equal molar extinction coefficients at 254 nm (ε_{254}) for dithiols (*S*)-**1** and (*R*)-**1**-d_4 is not evident, since the deuterium labeling of (*R*,*R*)-**1**-d_4 is part of the *m*-diamidophenyl chromophore and could have an effect on its electronic properties. Thus, we decided to perform UV absorbance measurements at different concentrations for both (*S*)-**1** and (*R*)-**1**-d_4 in order to determine the corresponding ε_{254} values (Fig. S10). The molar extinction coefficient was calculated as the slope of the least square regression line for the Abs_{254} *vs*. [BB] plot, and the obtained ε_{254} values were found to be the same within the experimental error (in methanol, $17.9 \cdot 10^3$ and $17.6 \cdot 10^3$ $M^{-1} \cdot cm^{-1}$ for (*S*)-**1** and (*R*)-**1**-d_4 respectively).

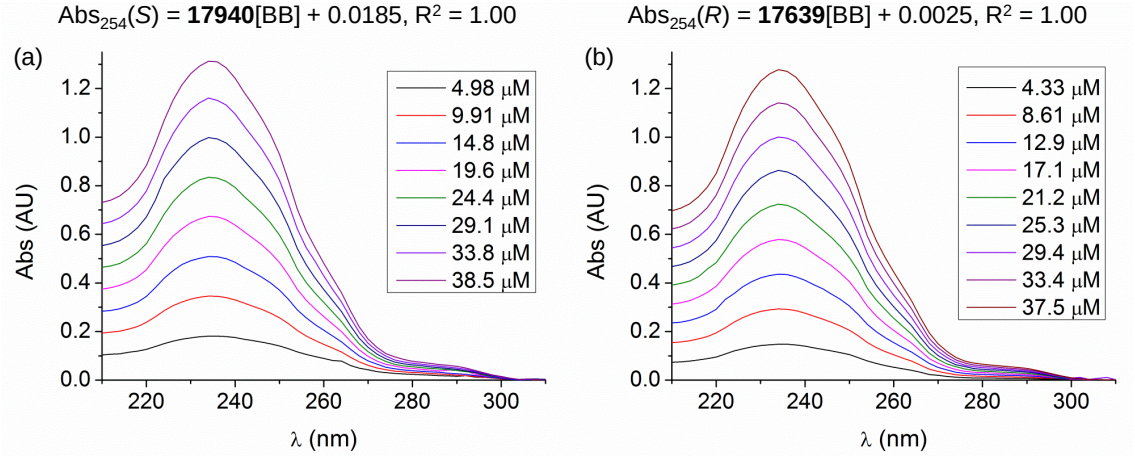


Fig. S10: UV spectra of (*S*)-**1** (a) and (*R*)-**1**-d_4 (b) in methanol at different concentrations, and linear least square equations for λ = 254 nm.

The confirmed equal molar extinction coefficients for the two BBs and their known additive molar absorptivity when forming oligomers, allowed directly using the chiral-HPLC peak areas for calculating the exchange constant. The presence of an unidentified impurity under the (*S*,*S*)-$\mathbf{1}_2$ peak (marked with an asterisk in Fig. S11) prompted us to calculate the K_{HS} by means of Equation S1. Notice that, for equimolar mixtures of (*S*)-**1** and (*R*)-**1**-d_4 the two generated homochiral dimers are necessarily also in equimolar amount.

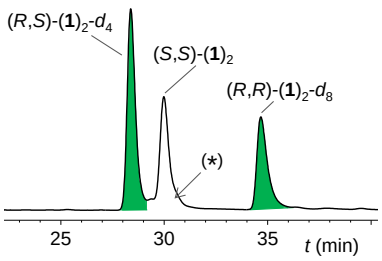


Fig. S11: Chiral-HPLC-UV trace (254 nm) of the DCL generated from (*S*,*S*)-**1** + (*R*,*R*)-**1**-d_4 (0.5 mM each) in 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO. Green peaks are the ones used for calculating the K_{HS} constant and the asterisk corresponds to the unidentified impurity.

$$K_{HS} = \left(\frac{(R,R)\text{-}(\mathbf{1})_2\text{-}d_8}{(R,S)\text{-}(\mathbf{1})_2\text{-}d_4} \right)^2 \qquad \textbf{(Equation S1)}$$

All dynamic libraries were performed in triplicate. The given K_{HS} constants are the corresponding averaged values, and the associated confidence intervals were calculated for a 95% significance by means of the OriginPro 8.1 software.[5]

Van't Hoff analysis

For each of the entries 2-9 in Table S2, nine experimental points (three replicates for each temperatures) were used to obtain a least square regression line for the ln(K_{HS}) *vs*. 1/T plot (Table S5).

Table S5: Slope, interception and r^2 values of the least square regression lines for the ln(K_{HS}) *vs*. 1/T plot. The errors were calculated for a 95% significance.

entry	sample	*slope*	*interception*	r^2
1	PB	–510 ± 92	2.1 ± 0.3	0.96
2	0%	–453 ± 50	1.9 ± 0.2	0.98
3	10%	–370 ± 137	1.8 ± 0.5	0.85
4	30%	–141 ± 59	1.3 ± 0.2	0.82
5	50%	–146 ± 98	1.5 ± 0.3	0.64
6	70%	–289 ± 134	2.1 ± 0.5	0.79
7	86%	–486 ± 129	3.3 ± 0.4	0.92
8	92%	–228 ± 63	3.2 ± 0.2	0.91

The enthalpy and entropy values were calculated as follows: $\Delta\Delta H = -slope \cdot R$; and $\Delta\Delta S = interception \cdot R$, where R is the gas constant (Table S6).

$$\Delta\Delta G_i = \Delta\Delta H_i - T\Delta\Delta S_i \qquad \textbf{(Equation S2)}$$

$$\Delta\Delta G_i = -RT\ln(K_{HS}/0.25) \qquad \textbf{(Equation S3)}$$

Table S6: Thermodynamic parameters obtained from the Van't Hoff analysis (all values are shown in kJ·mol^{-1}). The errors were calculated for a 95% significance.

entry	sample	$\Delta\Delta H_i$	$-T\Delta\Delta S_i$ [a]	$\Delta\Delta G_i$ [a,b]	$\Delta\Delta G_i$ [a,c]
1	PB	4.2 ± 0.8	–5.0 ± 0.8	–0.81	–0.88
2	0%	3.8 ± 0.4	–4.8 ± 0.4	–0.99	–1.03
3	10%	3 ± 1	–4 ± 1	–1.36	–1.46
4	30%	1.2 ± 0.5	–3.2 ± 0.5	–2.01	–2.05
5	50%	1.2 ± 0.8	–3.6 ± 0.8	–2.37	–2.44
6	70%	2 ± 1	–5 ± 1	–2.74	–2.84
7	86%	4 ± 1	–8 ± 1	–3.97	–4.05
8	92%	1.9 ± 0.5	–7.9 ± 0.5	–5.97	–5.99

[a] Values obtained for T = 22 °C. [b] Values obtained by means of Equation S2. [c] Values obtained by means of Equation S3.

Evaluation of the TBA^+ stability

Two NMR samples were prepared emulating the solvents composition and basicity of the binary mixtures but containing sodium acetate instead of the two Glu-based BBs. One sample was prepared by dissolving 10 mM CH_3COONa in D_2O with 1.2% (v/v) DMSO-d_6; and the other by dissolving CH_3COONa (10 mM) in 50% (v/v) CD_3CN with 48.8% (v/v) D_2O and 1.2% DMSO-d_6. TBAOH was used in both cases to adjust the pD to ~8. No higher %AN was tested in order to avoid sodium acetate precipitation.

The ^1H NMR spectrum of the two samples was recorded before and after exposure to 45 °C for 3 days, showing no appreciable variations during this period of time (Fig. S12). Additionally, after being exposed to 45 °C, the two samples were analyzed by UPLC-ESI-TOF and no Bu_3NH^+ was detected in any of them.

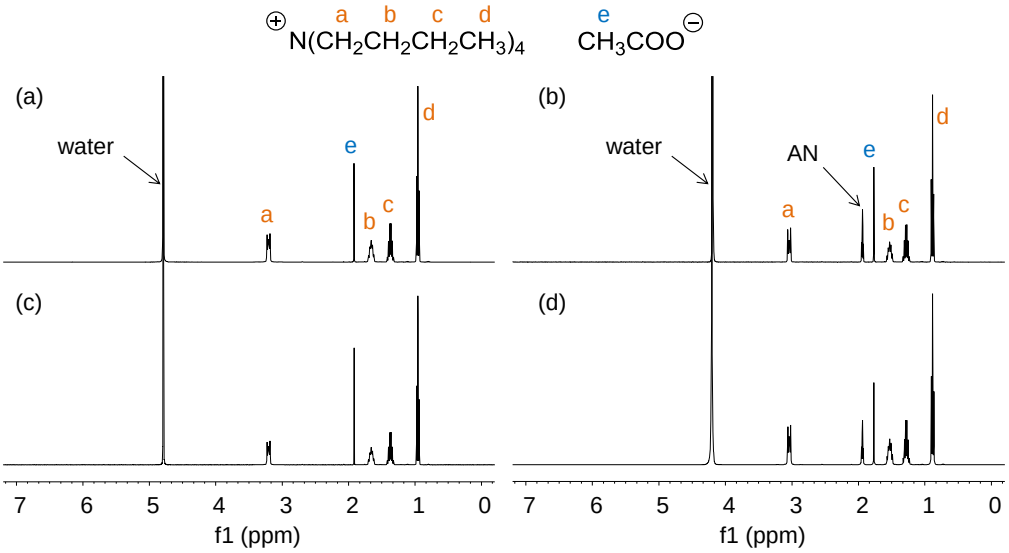


Fig. S12: ^1H NMR (400 MHz, 298 K) spectra of (a) 10 mM CH_3COONa in TBAOH basified D_2O with 1.2% (v/v) DMSO-d_6, (b) 10 mM CH_3COONa in TBAOH basified D_2O with 50% (v/v) AN and 1.2% (v/v) DMSO-d_6, (c) sample (a) after 3 days at 45 °C, and (d) sample (b) after 3 days at 45 °C. The spectra in (a,c) are referenced with the water residual peak (4.79 ppm) and the spectra in (b,d) are referenced with the AN residual peak (1.94 ppm).

Reversibility tests

Two solutions of tris(2-carboxyethyl)phosphine (TCEP, 25 mM) were prepared by dissolving the corresponding amount of TCEP·HCl in: i) Milli-Q water with 1.2% (v/v) DMSO, and ii) Milli-Q water with 92% (v/v) AN and 1.2% (v/v) DMSO. A small volume of NaOH (aq) was added to adjust the pH of both solutions to ~8. The reversibility test was performed at the four extreme conditions of %AN and temperature: 0% AN at 2 °C, 92% (v/v) AN at 2 °C, 0% AN at 45 °C, and 92% (v/v) AN at 45 °C. To the binary mixtures oxidized at these four extreme conditions, a small volume of a TCEP solution (the one prepared with the same %AN) was added. After re-oxidation at the corresponding temperature, the libraries were subsequently analyzed by chiral-HPLC and the K_{HS}^{TCEP} constant (*i.e.* K_{HS} after the addition of TCEP and re-oxidation) was calculated (Table S7).

Table S7: Exchange constants determined for the libraries before (K_{HS}) and after (K_{HS}^{TCEP}) the addition of 0.35 equivalents of TCEP and re-oxidation; at the extreme conditions of %AN (0 and 92% (v/v)) and temperature (2 and 45 °C).

		2 °C			45 °C		
entry	sample	K_{HS}	re-oxidation time	K_{HS}^{TCEP}	K_{HS}	re-oxidation time	K_{HS}^{TCEP}
1	0%	0.33 ± 0.01	3 d	0.31 ± 0.01	0.42 ± 0.01	14 h	0.40 ± 0.01
2	92%	2.68 ± 0.06		2.6 ± 0.1	3.0 ± 0.2		3.0 ± 0.1

After the addition of TCEP and re-oxidation, the exchange constant of the four libraries remained the same within the experimental error. Therefore, the binary mixtures were confirmed to reach the thermodynamic equilibrium at the four extreme conditions. By extension, all intermediate %AN and temperatures were also considered to be under thermodynamic control.

Structural studies

Nuclear magnetic resonance (NMR)

For each of dithiols (*S*)-**1** and (*R*)-**1**-d_4 a 7.00 mM *individual stock* was prepared in TBAOH basified water (pH 7.6) as explained before. From these, two NMR samples were prepared: i) (*S*)-**1** (2.0 mM) in TBAOH basified water (pH 7.6) with 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6 (prepared by mixing 216 µL of the (*S*)-**1** *individual stock* + 525 µL of CD_3CN + 9 µL of DMSO-d_6); and ii) (*S*)-**1** + (*R*)-**1**-d_4 (1.0 mM each) in TBAOH basified water (pH 7.6) with 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6 (prepared by mixing 108 µL of the (*S*)-**1** *individual stock* + 108 µL of the (*R*)-**1**-d_4 *individual stock* + 525 µL of CD_3CN + 9 µL of DMSO-d_6). Both samples were left to oxidize for 5 days at room temperature.

The ^{1}H NMR spectrum of the two samples was recorded on a Varian INOVA 500 spectrometer (500 MHz for ^{1}H, 278 K).

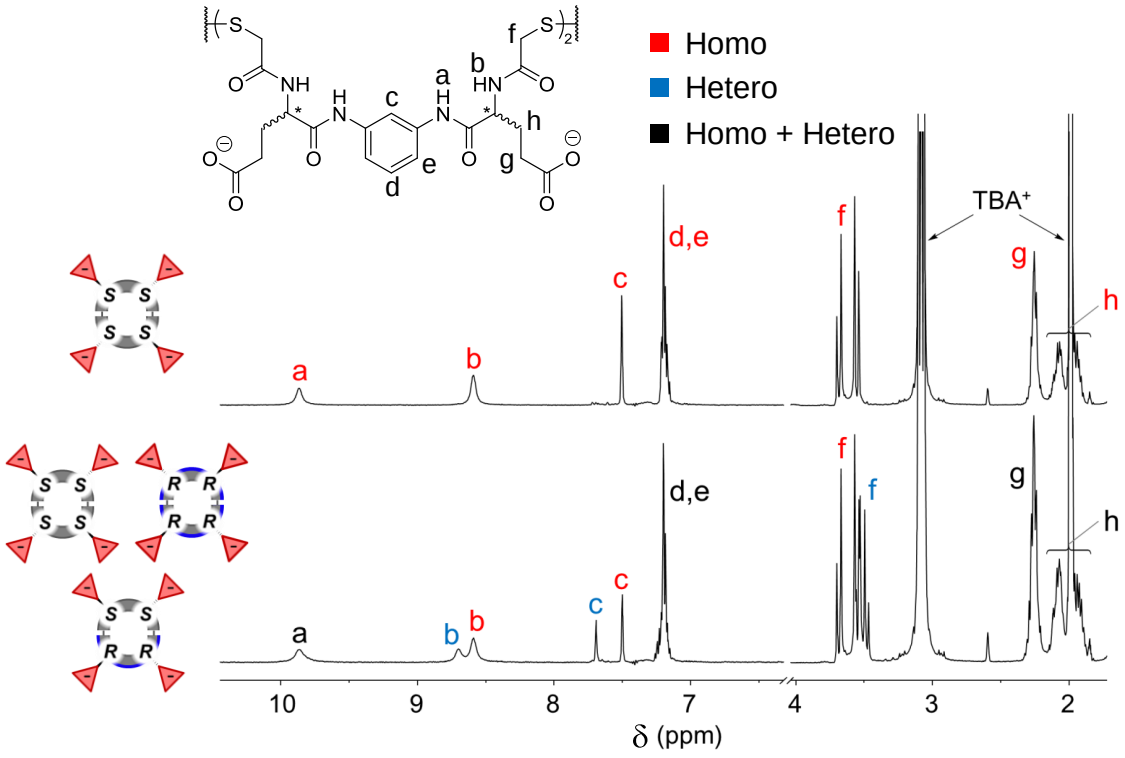


Fig. S13: Partial ^{1}H NMR spectra (500 MHz, in TBAOH basified water at pH 7.6 with 70% CD_3CN and 1.2% DMSO-d_6, 278 K) for the homochiral macrocycle (upper trace) and a mixture of homo- and heterochiral macrocycles (lower trace).

Additionally, only for the sample containing both BBs, gCOSY (Fig. S14a) and 2D ROESY (Fig. S14b) experiments were also performed on the same spectrometer. All spectra were acquired using the same optimized parameters for the Agilent Chempack pulse sequence WATER_ES (with excitation sculping).[6]

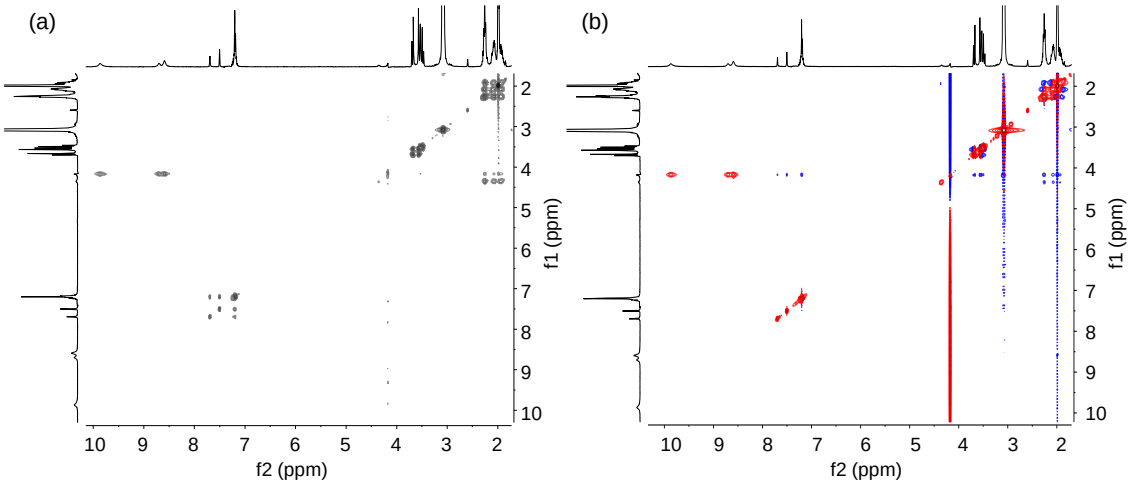
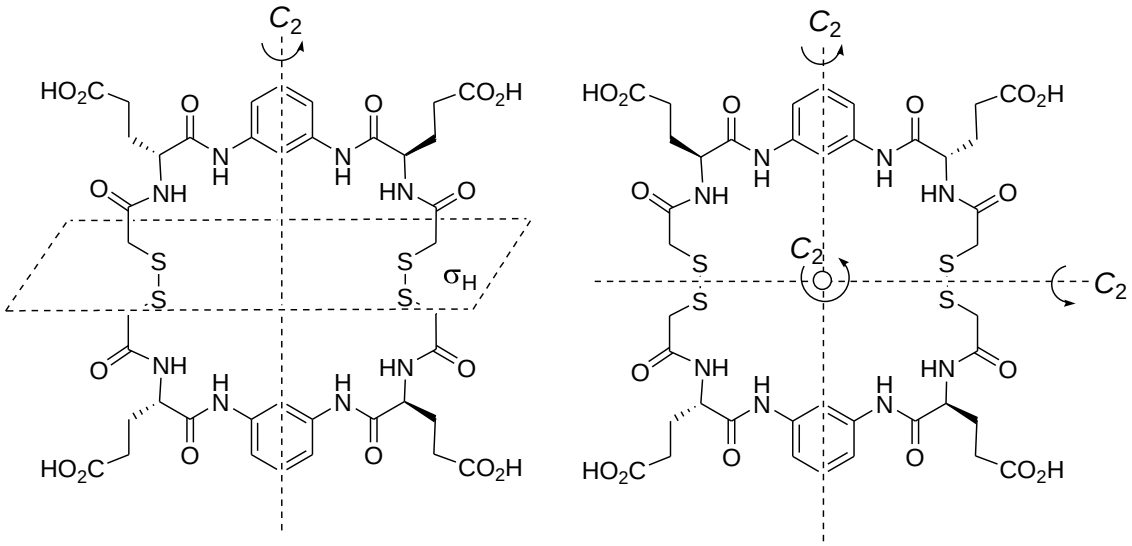


Fig. S14: (a) ^{1}H-^{1}H gCOSY and (b) 2D ROESY (500 MHz, "WATER_ES" water suppression pulse sequence, 278 K) spectra of (*S*)-**1** + (*R*)-**1**-d_4 (1 mM each) oxidized in TBAOH basified water (pH 7.6) containing 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6.

Theoretical estimation of the rotational entropy difference

The rotational entropy associated to a molecule is related to the number of species indistinguishable by rotation and, therefore, it can be calculated as $S = R\ln(\sigma)$, where σ means the symmetry number. If we make the simplification of an average conformation like the ones depicted in Fig. S15, the C_{2h} and D_2 symmetries can be assigned to hetero and homochiral species, respectively. Thus, the entropic contribution to the homo-hetero energetic difference due to their symmetry should be around 1.7 kJ·mol^{-1} (at 22 °C) per molecule, favouring the homochiral macrocycle. For the reaction depicted in Scheme 1c, the entropic contribution due to symmetry would be double, that is 3.4 kJ·mol^{-1} (at 22 °C). However, the molecules are highly flexible in solution, which means that this entropic difference due to symmetry differences must be the upper limit of the real value. Thus, this energy difference is smaller than the one observed experimentally, implying that other entropic factors must contribute to the homochiral selection.



Symmetry number (σ) = 2
$S\sigma = R\ln(2) = 5.763$ J·K^{-1}·mol^{-1}
at 295.15 K, $TS\sigma = 1.7$ kJ·mol^{-1}

Symmetry number (σ) = 4
$S\sigma = R\ln(4) = 11.526$ J·K^{-1}·mol^{-1}
at 295.15 K, $TS\sigma = 3.4$ kJ·mol^{-1}

$T\Delta S\sigma$(homo-hetero) = 1.7 kJ·mol^{-1} (per molecule)

$T\Delta S\sigma$(reaction, 2 molecules) = 3.4 kJ·mol^{-1}

Fig. S15. Schematic representation of the symmetry of both pseudopeptidic macrocycles drawn in an arbitrary symmetric conformation. An estimation of the energy associated to the rotational entropy (at 22 °C) is also shown for each structure, along with the homo-hetero difference.

Molecular simulation methods

All molecular simulations were carried out with the package Schrödinger Suite 2015,[7] through its graphical interface Maestro.[7] The program Macromodel[8] with its default force field OPLS 2005, a modified version of the OPLS-AA force field,[9] and GB/SA water solvation conditions[10] were used for energy minimization. Molecular dynamics simulations were performed with the program Desmond 4.4[11] using the OPLS 2005 force field.

Dimers (*R*,*S*)-$\mathbf{1_2}$ and (*S*,*S*)-$\mathbf{1_2}$ were built within Maestro and minimized with Macromodel as stated above. From these, simulation systems were built using the System Builder of the Maestro-Desmond interface,[12] which automatically assigns parameters to every atom in the system. Each system consisted of one dimer molecule in its completely ionized state, ie. with a total charge of -4, plus 4 Bu_4N^+ counterions placed randomly in a 50 x 50 x 50 Å cubic box of acetonitrile (~1130 AN molecules, ~7100 atoms in total) (Fig. S16). The acetonitrile solvent box was previously set up using the Disordered System Builder module from the Schrödinger Materials Science Suite,[13] and was equilibrated at 300 K and 1.0 bar by a short 1 ns MD run (final density 0.72 g cm^{-3}, in reasonable agreement with the experimental value of 0.77 g cm^{-3} at 300 K[14]). Systems were relaxed by minimization (initial steepest descent followed by LBFGS minimization), first with the solute restrained and then without restrains, until a gradient threshold of 0.1 kcal mol^{-1} $Å^{-1}$ was reached. Then, they were heated stepwise up to 300 K with short MD runs under periodic boundary conditions (PBC) (25 ps at 0.1, 10, 100 and 300 K), and equilibrated for 2 ns at the same temperature and 1.0 bar, in the NPT ensemble. Production MD simulations (100 ns, 2 fs timestep) were performed under the same conditions (PBC, NPT ensemble, 300 K and 1.0 bar) using the Nose-Hoover thermostat method[15] with a relaxation time of 1.0 ps and the Martyna-Tobias-Klein barostat method[16] with isotropic coupling and a relaxation time of 2 ps. Integration was carried out with the RESPA integrator[15b] using time steps of 2.0, 2.0, and 6.0 fs for the bonded, short range and long range interactions, respectively. A cut-off of 15 Å was applied to van der Waals and short-range electrostatic interactions, while long-range electrostatic interactions were computed using the smooth particle mesh Ewald method with an Ewald tolerance of 10^{-9}.[17] Bond lengths to hydrogen atoms were constrained using the Shake algorithm.[18] Coordinates were saved every 40 ps, hence 2500 snapshots (frames) were obtained from each MD run. After removing the solvent molecules and Bu_4N^+ ions from each frame, the 2500 conformations of both $\mathbf{1_2}$ dimers from each MD simulation were clustered in 10 groups based on the atomic RMSD of the heavy atoms, by using the hierarchical clustering with average linkage method as implemented in the Conformer Cluster method included in Maestro. In this way, 10 conformations representative of each cluster were defined for each dimer which covered different folding states, ie. from more folded/packed conformers to essentially extended and intermediately folded ones. In order to improve sampling, new MD simulations were set up using the frames corresponding to each of these 10

conformations as starting points. Thus, the new simulation systems were identical to the original ones (1 dimer molecule and 4 Bu_4N^+ ions, immersed in a 50 x 50 x 50 Å cubic box of acetonitrile) and the simulations were run under the same conditions (PBC, NPT ensemble, 300 K and 1.0 bar) but the simulation time was now increased to 250 ns, implying a total simulation time of 2.5 μs (10 x 250 ns) for each dimer. Coordinates were saved every 50 ps, hence 5000 snapshots were obtained from each MD simulation, summing a total of 50000 snapshots for each dimer. The Simulation Event Analysis application included in the Desmond-Maestro interface was used to determine the time dependence of the number of intramolecular hydrogen bonds and radius of gyration for each dimer and each simulation. Additionally, after removing the solvent molecules from each frame, the water accessible surface areas (ASA) were calculated for each dimer conformation, alone and with the 4 Bu_4N^+ counterions, with the program MOE.[19] The 50000 conformations obtained for each dimer were clustered based on the atomic RMSD of the heavy atoms and using three different clustering methods implemented in the Schrodinger Suite: (1) a simple redundant conformer detection method using energy and RMSD thresholds of 1 kJ mol^{-1} and 1 Å, respectively; (2) a hierarchical clustering with average linkage method that uses the Kelley penalty index[20] as criterion to select an optimal number of clusters; and (3) the same hierarchical clustering method but arbitrarily fixing the number of clusters to 100. Finally, in order to estimate the conformational entropy ($S_{conform}$) of each dimer we used the CENCALC software.[21]

CENCALC uses both trajectory coordinates and topology information to characterize the conformational states of the molecules of interest by discretizing the time evolution of internal rotations. After discretization of the torsion angle evolution, the conformational state of an individual torsion becomes associated with a one-dimensional variable X, and, by extension, the conformational state of a set of M torsion angles $\{A_1;\ldots;A_M\}$ can be described by an M-dimensional variable or vector $\{X_1;\ldots;X_M\}$, whose components specify the discrete conformational state of the different torsions. In principle, the conformational entropy could be estimated directly form the observed relative frequencies of the outcomes x of X, by means of the Shannon expression:

$$S_{conform}(A) = -k_B \sum_x p_x \ln p_x$$

where p_x is an estimation, using the observed relative frequencies, of the probability to obtain the outcome x. However, due to the huge number of potentially accessible conformers for medium-sized and large molecules, the direct application of the Shannon expression would result in large and negatively-biased entropies due to sampling limitations. CENCALC allows using different strategies, namely, the mutual information expansion (MIE), the approximate mutual information expansion (AMIE), the multibody local approximation (MLA) or the correlation-corrected MLA (CC-MLA) methods, to overcome this limitation and calculate the associated entropies. In

addition, all of these methods can be combined with a distance-based cutoff criterion, in order to include only correlation effects among torsion angles whose mean separation is below a predefined cutoff. The rationale behind fixing a cutoff in the calculation relies on the interference due to the growth of false correlation when no cutoff or large cutoff values are used. One could expect that lower $S_{conform}$ values should be obtained at larger cutoff values because of the capture of presumably more physical correlation. However, the use of larger cutoffs does not necessarily result in better entropy estimations because estimations derived from a finite amount of data tend to be considerably biased, resulting in a systematic under estimation of the true entropy, or equivalently, an overestimation of the correlation effects. In other words, as the cutoff increases, the estimated $S_{conform}$ can decrease due to the capture of more physical correlation but also due to the growth of negative bias or false correlation. We used the CC-MLA method since it is the recommended one for medium-sized molecules such as dimers (*R*,*S*)-$\mathbf{1}_2$ and (*S*,*S*)-$\mathbf{1}_2$, and because it allows to determine the best cutoff for a given amount of sampling.

Following, a brief description of the steps required to run the CENCALC software is provided, however, more details about CENCALC should be found in the original article and in the software user manual.[21] First, it was necessary to convert the trajectory files and molecular topologies into AMBER format. Thus, topologies for each dimer were calculated using Antechamber[22] and the GAFF force field,[23] as implemented in the AmberTools 15 package.[24] Each group of 10 simulations for each dimer were merged into a single dcd file (binary trajectory file format) using the VMD software.[25] Then, the *get_tor.py* script included in CENCALC was used to identify the torsions about the rotatable bonds involving only the heavy atoms of each dimer, which are required to characterize the conformational states of each molecule. The output of *get_tor.py* was an input file for the *ptraj* program, included in the AmberTools package. Execution of *ptraj* with this input file resulted in a series of data files containing the evolution of the torsion angles along the MD trajectory and a matrix of average distances among the solute atoms. Then, the *cencalc_prep* program was executed to perform several tasks. First, *cencalc_prep* estimates the probability density functions of the M torsion angles and characterize their maxima and minima critical points. Subsequently, it transforms the initial time series of N real numbers per torsion angle into a set of N integer numbers labeling the conformational states populated by each torsion angle. Thus, as a result a matrix with N rows (ie. the number of MD snapshots) and M columns (ie. the number of rotatable bonds) is generated, where the i-th row is an array of integer numbers that represent the conformational state at the i-th snapshot. In addition, *cencalc_prep* can reduce the dimensionality of this matrix by identifying and removing all the frozen torsions that do not represent conformational changes and do not affect the conformational entropy. Finally, *cencalc_prep* calculates a new distance matrix that contains the mean distances among the center of mass of the two central atoms that are involved in each one of the M torsion angles. All these results were then used as input for *cencalc_omp*, the program that performs the entropy calculation. First, *cencalc_omp*

calculates the probability mass function of the discretized torsion angles by means of the maximum likelihood method fed with the corresponding set of N integer numbers. Next, the program typically computes data points for obtaining an entropy plot with a time step specified in terms of a number of MD snapshots and using one of the implemented entropy methods. At this stage, *cencalc_omp* builds on-the-fly all the subsets of torsion angles that fulfill the corresponding distance-based cutoff criteria, if required. On output, a text file is generated that contains a table summarizing all the calculations and that can be readily imported by spreadsheet software for data analysis.

Molecular simulation results

Comparison of the results from MD simulations on dimers (*R*,*S*)-**1₂** and (*S*,*S*)-**1₂** shows that, on average, the heterochiral dimer has a tendency to form a larger number of intramolecular hydrogen bond (HB) interactions than the homochiral dimer, ie. (*R*,*S*)-**1₂** shows between 6-7 HBs during most of the simulation time, reaching at some points a maximum of 9 HBs, while (*S*,*S*)-**1₂** shows between 5-6 HBs most of the time and a maximum of 8 HBs (Fig. S17 and S20). On the other hand, on average, a smaller radius of gyration (R_g) and a smaller accessible surface area (ASA) are observed for the heterochiral dimer relative to the homochiral one (Fig. S18, S19, S21 and S22). In addition, the observed profiles show simulation periods that alternate between high and low R_g and ASA values. Since smaller R_g and ASA values correlate with more folded conformations and higher values with more extended ones, these profiles indicate that, despite their high flexibility, the structures of both dimers fluctuate between high and low packing states, and that the heterochiral dimer tends to adopt a more packed conformation during longer periods of simulation time than the homochiral one. Fig. S23 shows snapshots from representative simulations of both dimers, illustrating two highly packed states for both. Considering the ASA values as an estimator of the solvation surface of each dimer, the fact that the heterochiral dimer shows a smaller average ASA would suggest that, on average, a smaller number of solvent molecules would be involved in solvation of this isomer. However, each dimer molecule is forming an electrostatic complex with the four Bu_4N^+ counterions, therefore taking also into account the counterions as part of the solute it comes out that the difference in average ASA becomes negligible, thus indicating that, on average, a similar number of solvent molecules would be involved in solvation of both dimers.

After grouping all conformations obtained from the snapshots of the 10 simulations for each dimer, clustering of the two groups of 50000 conformations was performed to group similar conformations and to estimate how many different conformational states are accessible to each dimer and how are they populated. Three different clustering methods were used and consistently all three showed that a larger number of clusters are required to describe the accessible conformations of the homochiral dimer with respect to the heterochiral one (Fig. S24-S26). Thus, for example, using the hierarchical clustering with average linkage method and with selection of optimal number of clusters, a 90 % of the conformations of the heterochiral dimer could be grouped in the 375 most populated clusters, while 628 clusters were required to include the same percentage of population of the homochiral dimer conformations. Similar results were obtained with the other clustering methods, hence, these results suggest that a larger number of conformational states are accessible to (*S*,*S*)-**1₂** under the simulation conditions and, consequently, that the conformational entropy would favor this isomer relative to the heterochiral one.

Estimation of $S_{conform}$ using the CC_MLA method at different cutoff values, using as input the 50000 conformations obtained for both dimers, allowed the determination of

the best cutoff values, ie. the cutoff where the conformational entropy reaches a minimum, which constitutes the lowest upper bound to the exact entropy. Thus, the best cutoff for (*R*,*S*)-$\mathbf{1_2}$ was 8 Å and that for (*S*,*S*)-$\mathbf{1_2}$ was 9 Å (Fig. S27, top). Furthermore, the dependence of $S_{conform}$ at different cutoffs as a function of the number of MD conformations was also studied (Fig. S27, bottom). The results indicated that the entropy values reach a plateau when 30000 or more conformations are considered, suggesting that the sampling for both isomers was enough to reliably estimate their conformational entropy. Thus, at the best cutoff values and considering the whole set of 50000 confomations for each dimer, a $S_{conform}$ (hetero) = 23.6 cal K^{-1} mol^{-1} and a $S_{conform}$ (homo) = 29.9 cal K^{-1} mol^{-1} could be estimated. With these values, it comes out that in terms of conformational entropy the homochiral dimer would be favored by 1.9 kcal mol^{-1} relative to the heterochiral dimer, at the temperature (300 K) at which the simulations were carried out.

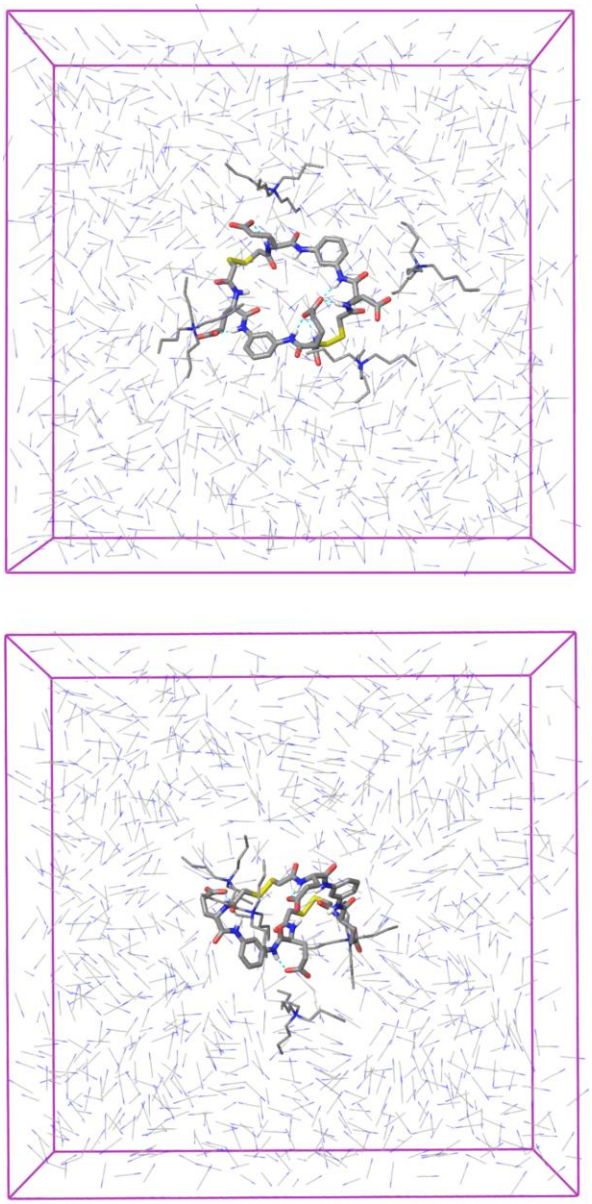


Fig. S16. Molecular dynamics simulation systems for dimers (*R*,*S*)-$\mathbf{1_2}$ (top) and (*S*,*S*)-$\mathbf{1_2}$ (bottom). Both were constituted by a 50 x 50 x 50 $Å^3$ cubic box containing one molecule of each dimer, 4 Bu_4N^+ counterions, and ~1130 acetonitrile molecules.

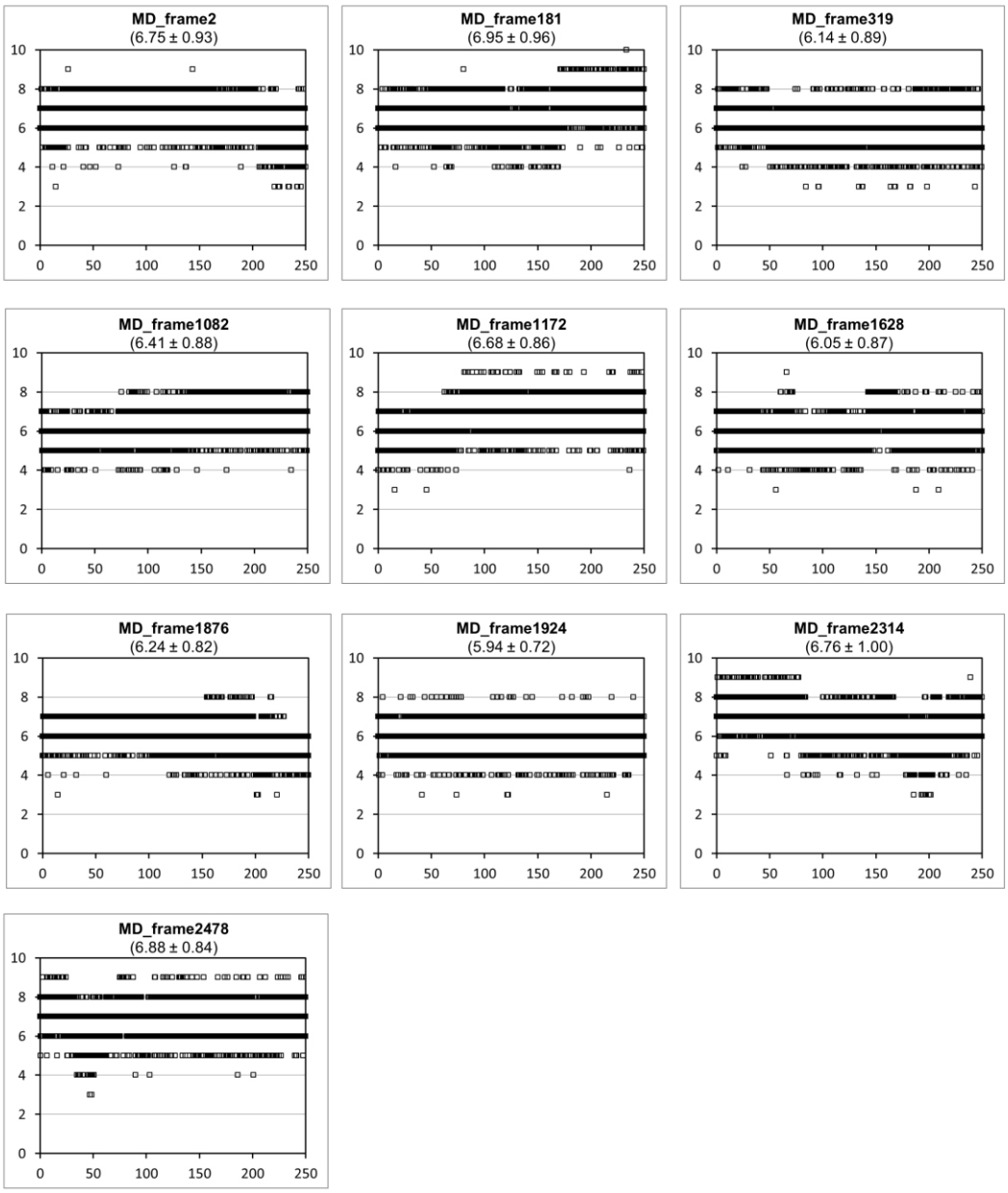


Fig. S17. Summary of results obtained from ten 250 ns MD simulations carried out for the heterochiral dimer (*R*,*S*)-$\mathbf{1}_2$. The graphics represent the number of intramolecular hydrogen bonds *vs* simulation time (ns). Mean ± SD values are provided for each simulation in parenthesis. Global mean ± SD value: 6.48 ± 0.95.

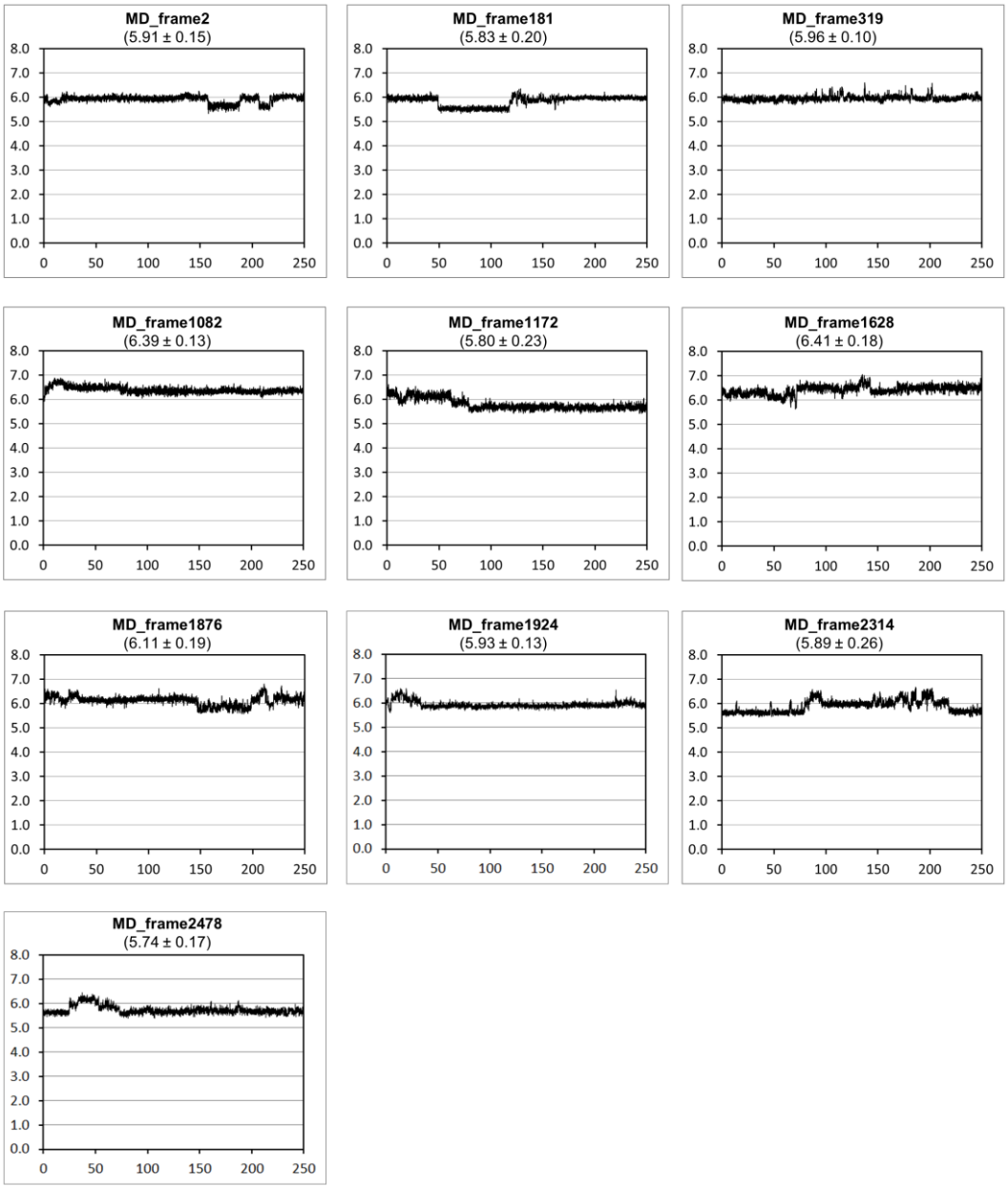


Fig. S18. Summary of results obtained from ten 250 ns MD simulations carried out for the heterochiral dimer (*R*,*S*)-$\mathbf{1_2}$. The graphics represent the radius of gyration (Å) *vs* simulation time (ns). Mean ± SD values are provided for each simulation in parenthesis. Global mean ± SD value: 6.00 ± 0.29 Å.

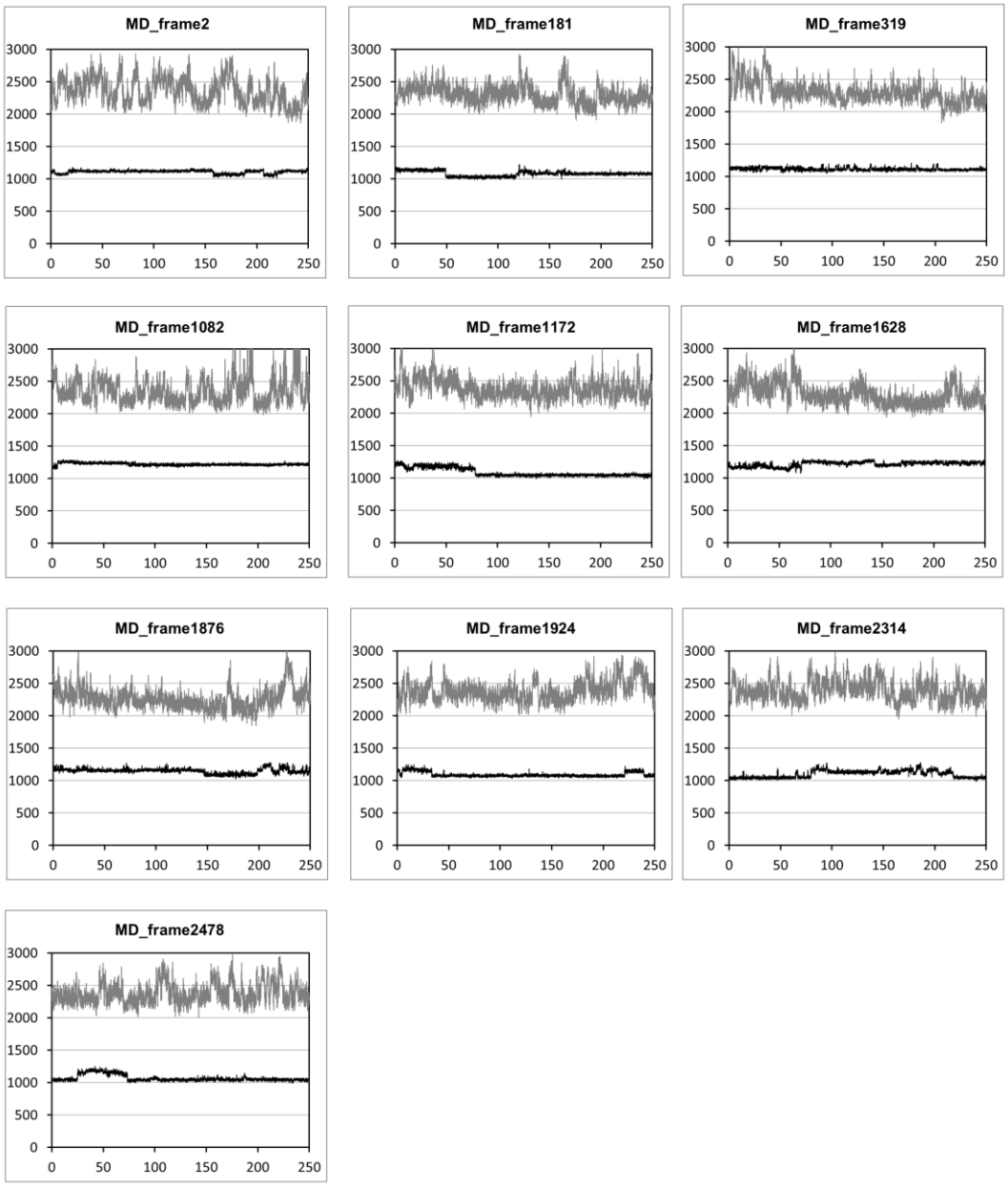


Fig. S19. Summary of results obtained from ten 250 ns MD simulations carried out for the heterochiral dimer (*R*,*S*)-$\mathbf{1_2}$. The graphics represent the water accessible surface areas (ASA, Å^2) of the dimer alone (black) and the dimer plus the 4 Bu_4N^+ counterions (gray), *vs* simulation time (ns). Global mean ASA ± SD values: 1123 ± 66 Å^2 (black lines) and 2338 ± 171 Å^2 (gray lines).

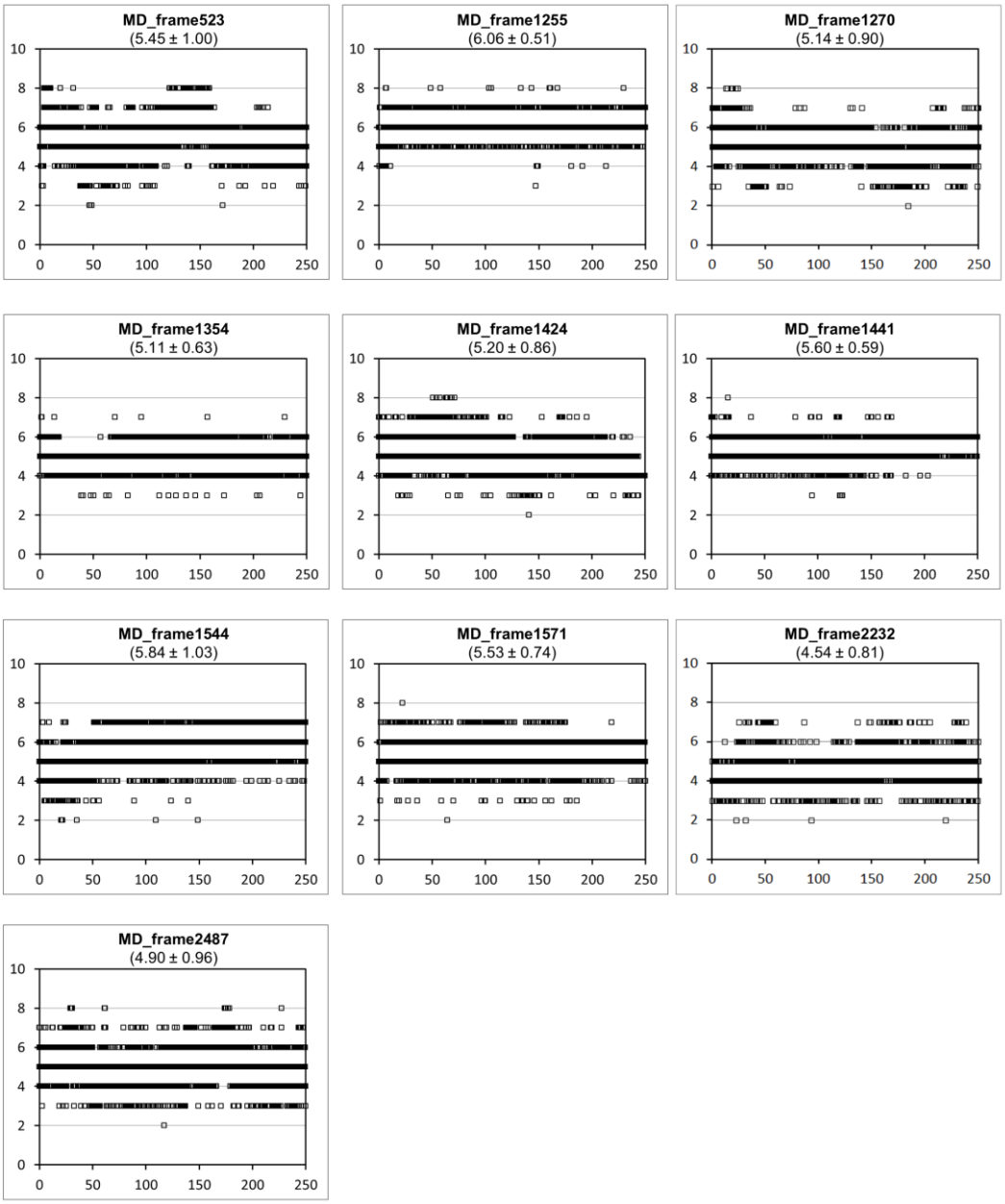


Fig. S20. Summary of results obtained from ten 250 ns MD simulations carried out with the homochiral dimer (*S*,*S*)-$\mathbf{1}_2$. The graphics represent the number of intramolecular hydrogen bonds *vs* simulation time (ns). Mean ± SD values are provided for each simulation in parenthesis. Global mean ± SD value: 5.34 ± 0.93.

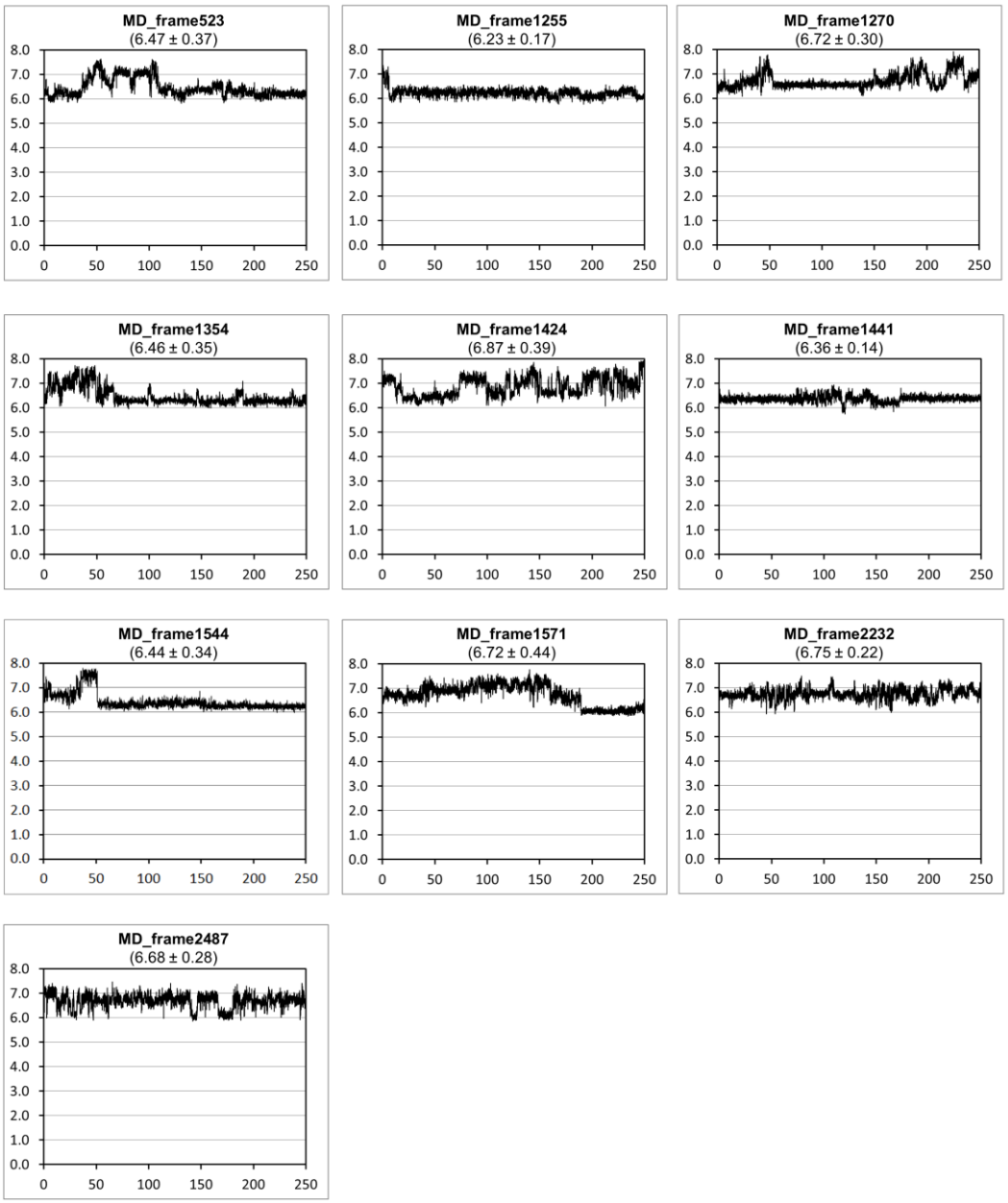


Fig. S21. Summary of results obtained from ten 250 ns MD simulations carried out with the homochiral dimer (*S*,*S*)-$\mathbf{1}_2$. The graphics represent the radius of gyration (Å) *vs* simulation time (ns). Mean ± SD values are provided for each simulation in parenthesis. Global mean ± SD value: 6.57 ± 0.37 Å.

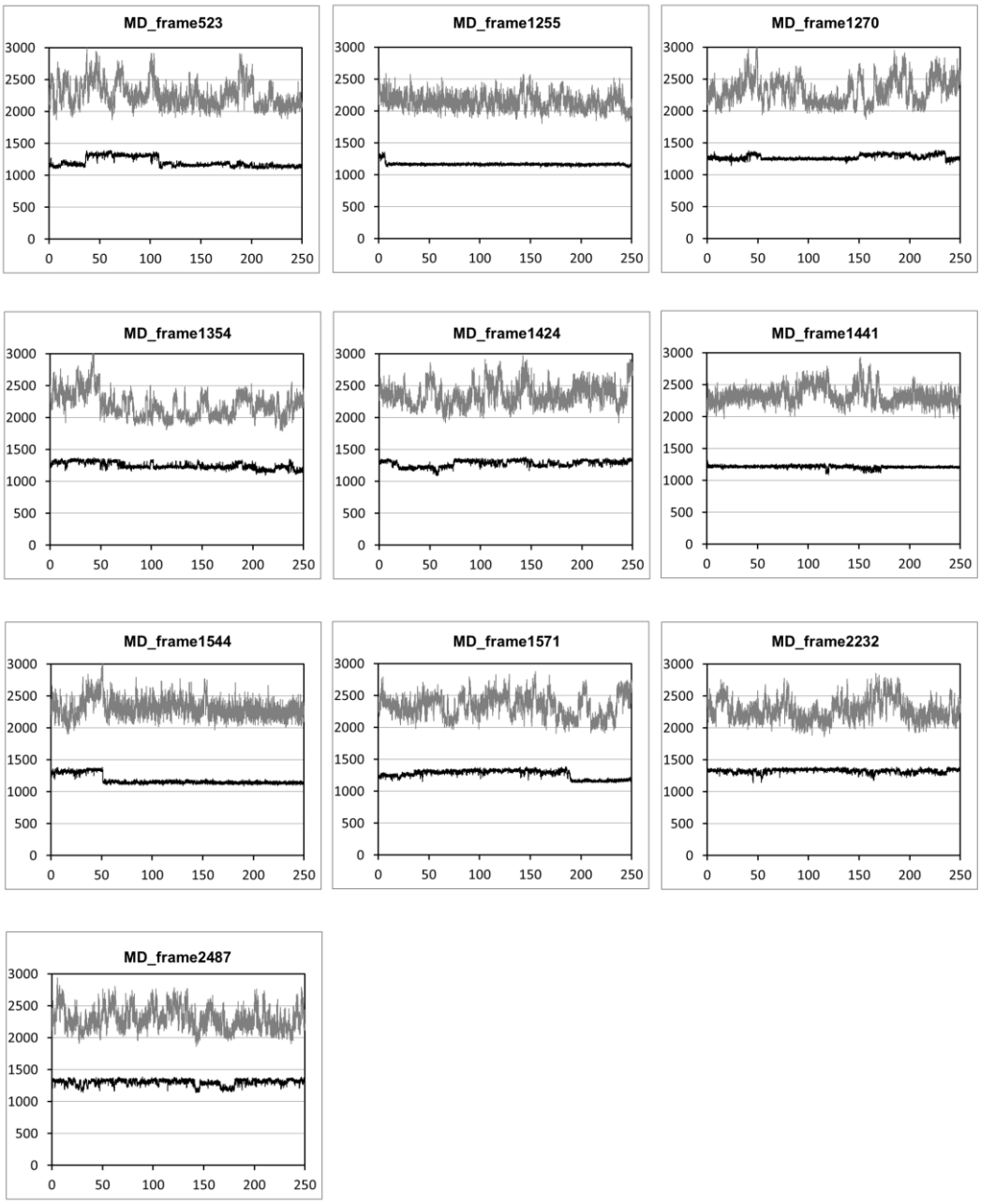


Fig. S22. Summary of results obtained from ten 250 ns MD simulations carried out with the homochiral dimer (*S*,*S*)-$\mathbf{1}_2$. The graphics represent the water accessible surface areas (ASA, Å^2) of the dimer alone (black) and the dimer plus the 4 Bu_4N^+ counterions (gray), *vs.* simulation time (ns). Global mean ASA ± SD values: 1246 ± 72 Å^2 (black lines) and 2284 ± 184 Å^2 (gray lines).

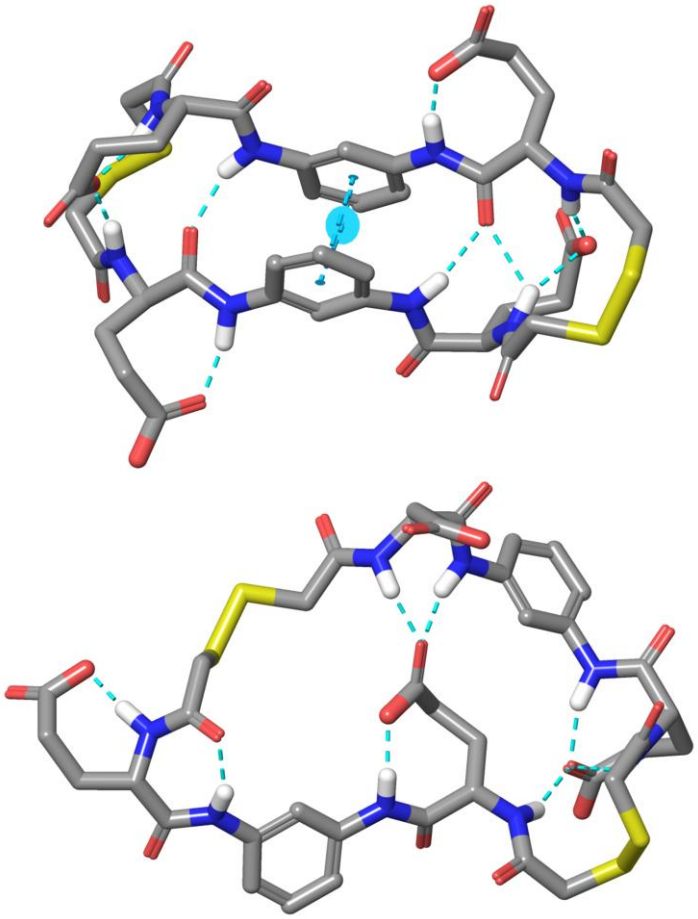


Fig. S23. Snapshots from representative MD simulations of (*R*,*S*)-$\mathbf{1}_2$ (top) and (*S*,*S*)-$\mathbf{1}_2$ (bottom) showing two highly folded conformations with the maximum number of intramolecular HBs. The heterodimer conformation also shows a π-π stacking interaction between the aromatic rings of the molecule.

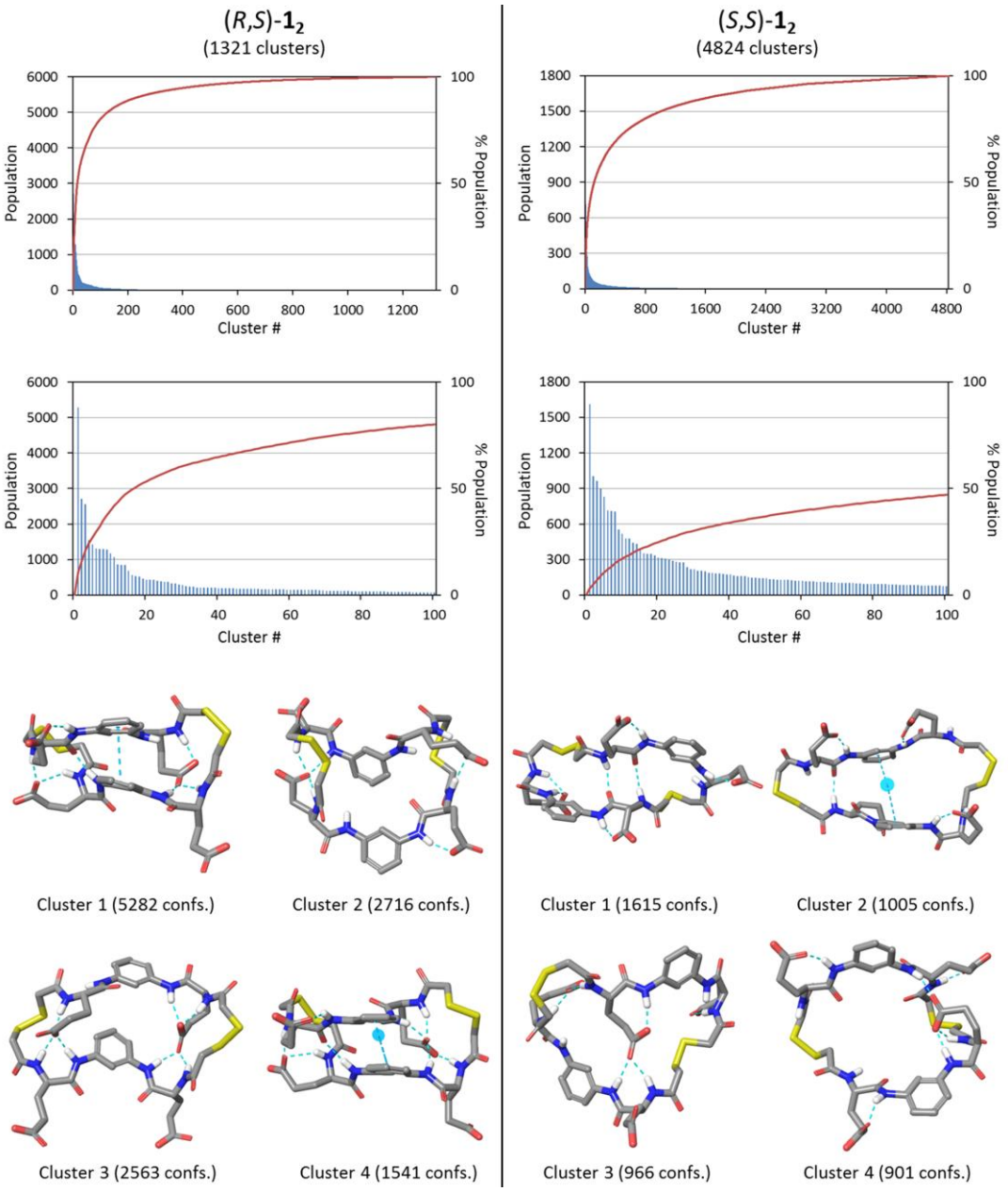


Fig. S24. Results from clustering the 50000 conformations obtained from MD of dimers (*R*,*S*)-**1**$_2$ and (*S*,*S*)-**1**$_2$, by applying the redundant conformer detection method. **Upper panels**: The graphs show the population distribution of conformations among the total number of clusters (given in parenthesis) and among the 100 most populated clusters. The red line curves represent the added % of population. Number of clusters representing: 50 % population, 17 for (*R*,*S*)-**1**$_2$, 122 for (*S*,*S*)-**1**$_2$; 90 % population, 225 for (*R*,*S*)-**1**$_2$, 1661 for (*S*,*S*)-**1**$_2$. **Lower panels**: Representative (closest to centroid) conformations of the four most populated clusters, the number of conformations in each cluster is given in parenthesis. Intramolecular hydrogen bonds and π-π stacking interactions between the aromatic rings are shown with cyan dashed lines.

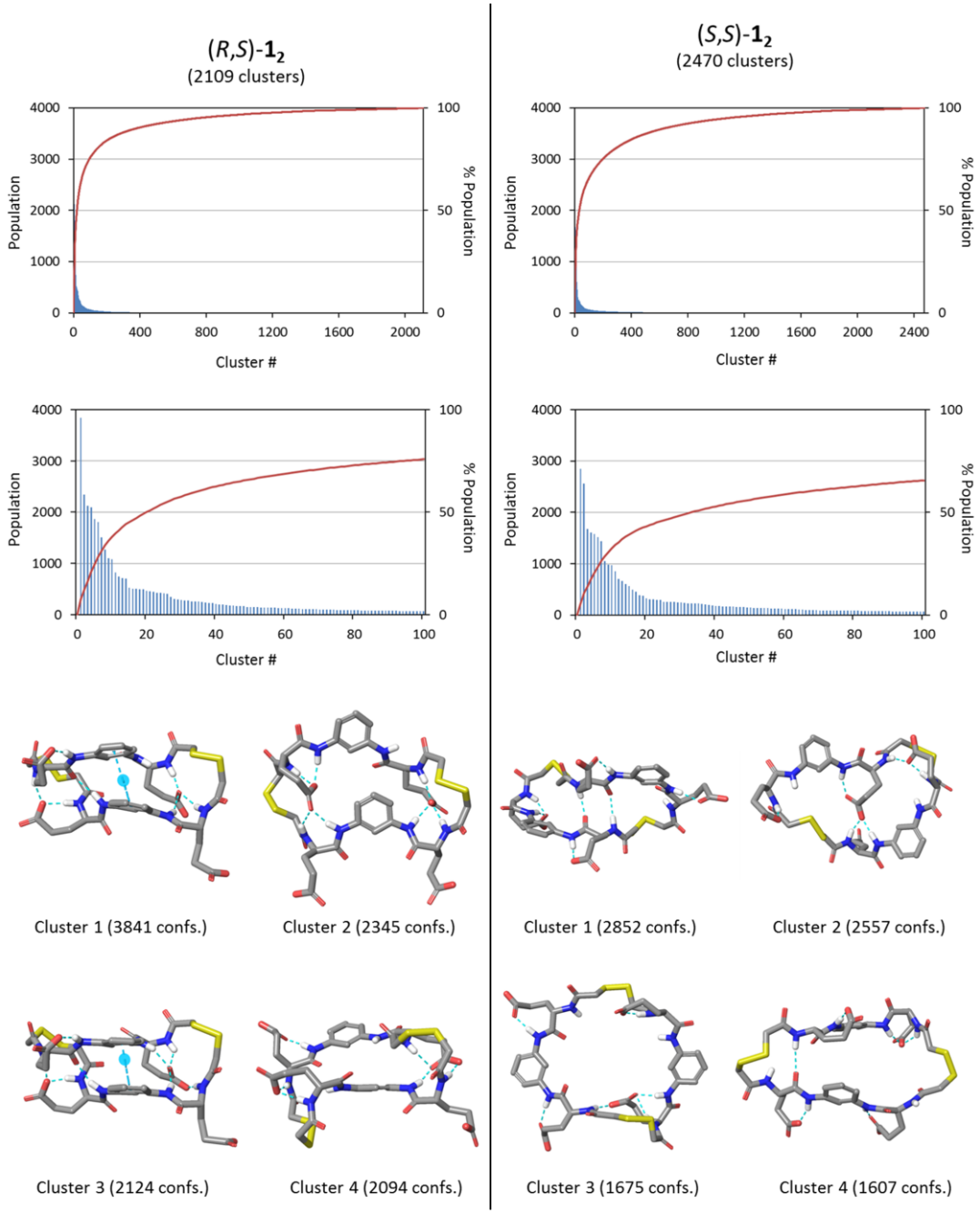


Fig. S25. Results from clustering the 50000 conformations obtained from MD of dimers (*R*,*S*)-**1**$_2$ and (*S*,*S*)-**1**$_2$, by applying a hierarchical clustering with average linkage method and selection of optimal number of clusters. **Upper panels**: The graphs show the population distribution of conformations among the total number of clusters (given in parenthesis) and among the 100 most populated clusters. The red line curves represent the added % of population. Number of clusters representing: 50 % population, 20 for (*R*,*S*)-**1**$_2$, 33 for (*S*,*S*)-**1**$_2$; 90 % population, 375 for (*R*,*S*)-**1**$_2$, 628 for (*S*,*S*)-**1**$_2$. **Lower panels**: Representative (closest to centroid) conformations of the four most populated clusters, the number of conformations in each cluster is given in parenthesis. Intramolecular hydrogen bonds and π-π stacking interactions between the aromatic rings are shown with cyan dashed lines.

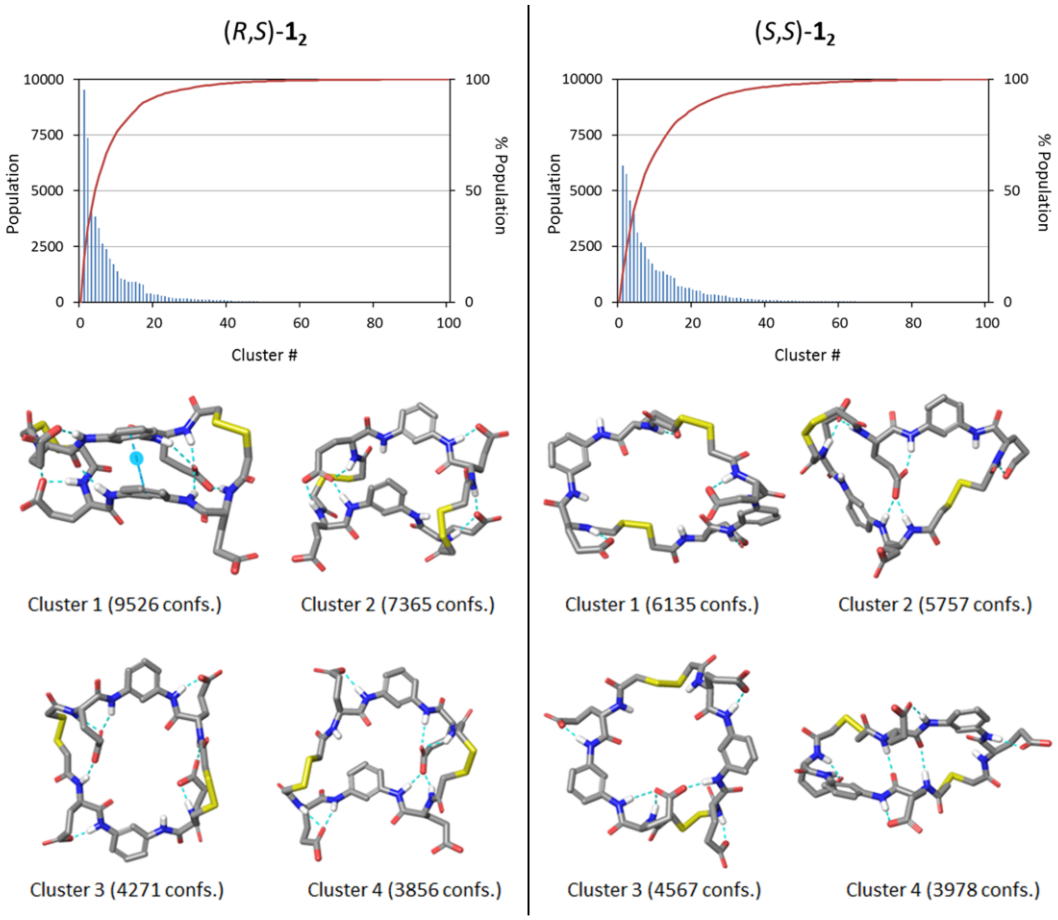


Fig. S26. Results from clustering the 50000 conformations obtained from MD of dimers (*R*,*S*)-$\mathbf{1}_2$ and (*S*,*S*)-$\mathbf{1}_2$, by applying a hierarchical clustering method with 100 clusters. **Upper panels**: The graphs show the population distribution of conformations among the 100 clusters. The red line curves represent the added % of population. Number of clusters representing: 50 % population, 4 for (*R*,*S*)-$\mathbf{1}_2$, 6 for (*S*,*S*)-$\mathbf{1}_2$; 90 % population, 18 for (*R*,*S*)-$\mathbf{1}_2$, 24 for (*S*,*S*)-$\mathbf{1}_2$. **Lower panels**: Representative (closest to centroid) conformations of the four most populated clusters, the number of conformations in each cluster is given in parenthesis. Intramolecular hydrogen bonds and π-π stacking interactions between the aromatic rings are shown with cyan dashed lines.

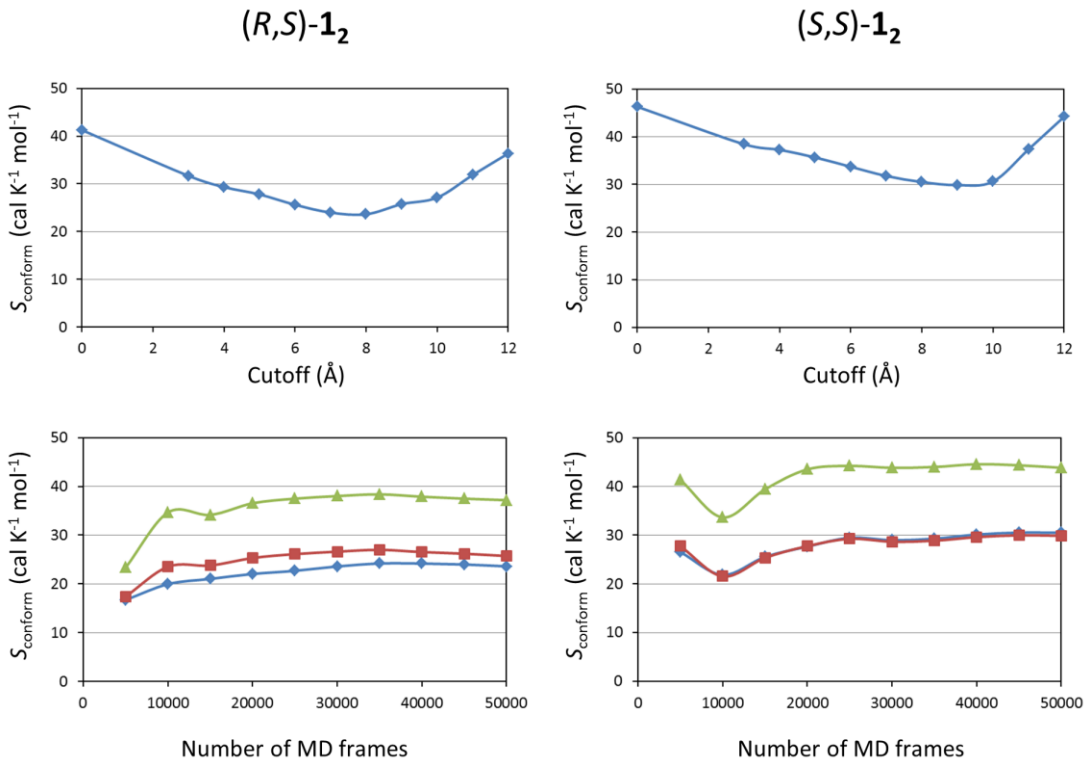


Fig. S27. CC-MLA estimated entropies ($S_{conform}$) for (*R*,*S*)-**1**$_2$ and (*S*,*S*)-**1**$_2$. Top: $S_{conform}$ values calculated from the full set of 50000 MD frames as a function of cutoff; optimum cutoffs for heterochiral and homochiral dimers: 8 and 9 Å, respectively. Bottom: $S_{conform}$ values calculated as a function of number of MD frames considered, at cutoff values of 8 (blue) or 9 Å (red), or without cutoffs (green). Considering the full set of frames: at cutoff = 8 Å, $S_{conform}$ (hetero) = 23.6 cal K^{-1} mol^{-1}, $S_{conform}$ (homo) = 30.5 cal K^{-1} mol^{-1}; at cutoff = 9 Å, $S_{conform}$ (hetero) = 25.7 cal K^{-1} mol^{-1}, $S_{conform}$ (homo) = 29.9 cal K^{-1} mol^{-1}; without cutoff, $S_{conform}$ (hetero) = 37.2 cal K^{-1} mol^{-1}, $S_{conform}$ (homo) = 43.8 cal K^{-1} mol^{-1}.

References

1. A. P. Kozikowski, Y. Chen, A. Gaysin, B. Chen, M. A. D'Annibale, C. M. Suto and B. C. Langley, *J. Med. Chem.*, 2007, **50**, 3054.
2. J. Atcher, A. Moure and I. Alfonso, *Chem. Commun.*, 2013, **49**, 487.
3. P. J. Elving, J. M. Markowitz and I. Rosenthal, *Anal. Chem.*, 1956, **28**, 1179.
4. N. Hafezi and J.-M. Lehn, *J. Am. Chem. Soc.*, 2012, **134**, 12861.
5. Origin (OriginLab, Northampton, MA).
6. T. L. Hwang and A. J. Shaka, *J. Magn. Reson., Ser A*, 1995, **112**, 275.
7. Schrödinger Suite 2015, Schrödinger, LLC, New York, NY, 2015.
8. Schrödinger Maestro v. 10.4, Schrödinger, LLC, New York, NY, 2015.
9. (a) W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225; (b) J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio and R. M. Levy, *J. Comput. Chem.*, 2005, **26**, 1752; (c) G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. B*, 2001, **105**, 6474.
10. W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127.
11. (a) D. E. S. Research, New York, NY, 2015; (b) K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossváry, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan and D. E. Shaw, Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters, Tampa, Florida, 2006.
12. Schrödinger MacroModel v. 11.0, Schrödinger, LLC, New York, NY, 2015.
13. Schrödinger Maestro-Desmond Interoperability Tools v. 4.4, Schrödinger, LLC, New York, NY, 2015.
14. Ł. Ruszczyński, M. Reda, M. Królikowski, M. Gliński and T. Hofman, *J. Chem. Eng. Data*, 2016, **61**, 996.
15. (a) D. J. Evans and B. L. Holian, *J. Chem. Phys.*, 1985, **83**, 4069; (b) G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635.
16. G. J. Martyna, D. J. Tobias and M. L. Klein, *J. Chem. Phys.*, 1994, **101**, 4177.
17. (a) T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089; (b) U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577.
18. V. Kräutler, W. F. van Gunsteren and P. H. Hünenberger, *J. Comput. Chem.*, 2001, **22**, 501.
19. MOE, Chemical Computing Group Inc., Montreal, QC, Canada, v. 2013.08 edn., 2013.
20. L. A. Kelley, S. P. Gardner and M. J. Sutcliffe, *Protein Eng.*, 1996, **9**, 1063.
21. E. Suarez, N. Diaz, J. Mendez and D. Suarez, *J. Comput. Chem.*, 2013, **34**, 2041.
22. J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247.
23. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157.
24. D. A. Case, J. T. Berryman, R. M. Betz, D. S. Cerutti, T. E. I. Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. M. Merz, G. Monard, P. Needham, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, R. Salomon-Ferrer, C. L.

Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, D. M. York and P. A. Kollman, University of California, San Francisco, 2015.
25. W. Humphrey, A. Dalke and K. Schulten, *J. Molec. Graphics*, 1996, **14**, 33.