

Contents

1. Computational Modelling	2
2. General Synthetic and Analytical Methods	11
3. High-Throughput Synthetic Screening	13
4. Synthesis, Purification and Characterisation of Cage Pot OCP3	20
5. References	37

1. Computational Modelling

1.1. Construction and energetic analysis of molecular models

After manually assembly of the cage pot structures, high temperature molecular dynamics (MD) simulations were used to explore the low energy conformations of the molecules. Simulations were run for 50 ns at a time with 5000 structures sampled per run. Calculations were run with MacroModel at 1000 K with a 1 fs timestep. All calculations were carried out with the OPLS3 forcefield,^[1] which we have previously found to be successful for exploring the energy landscapes of porous imine cages.^[2] Each sampled structure was geometry optimised in MacroModel with the Polak-Ribier Conjugate Gradient method and a gradient convergence criterion of 0.05.

Once the lowest energy conformers were located for each molecule, density functional theory (DFT) was applied for a more accurate energetic ranking of the conformations. They were optimised in CP2K^[3] with the PBE functional,^[4] TZVP-MOLOPT basis set^[5] and the Grimme-D3 dispersion correction.^[6] GTH-type pseudopotentials^[7] were used and a plane-wave grid cutoff of 350 Ry. Formation energies per bond were then calculated according to:

$$E_{\text{formation}} = \frac{(E_{\text{cage}} + xE_{\text{water}}) - (mE_{\text{aldehyde}} + nE_{\text{amine}})}{x}$$

where E_{cage} is the energy of the cage formed, E_{water} is the energy of the water molecule formed by the reaction, E_{aldehyde} is the energy of the aldehyde precursor and E_{amine} is the energy of each of the different amines. The number of aldehyde reactants is m , n is the number of each of the amines and x is the number of imine bonds formed, equivalent to the number of water molecules produced in the reaction. We have previously found that formation energies per bond at this level of theory allow comparisons between different systems that contain different molecular fragments and can identify trends in which porous organic cages are most likely to be experimentally realised.^[8]

As well as differences in formation energies of the possible products, the thermodynamic balance between social and narcissistic self-sorting products is also influenced by entropic effects, which are expected to be dominated by the rigid molecule entropy difference due to different numbers of molecules produced by the same number of precursor molecules. For example, the same precursors could produce either 4 product molecules by social self-sorting (4 **OCP1** organic cage pot molecules) or 3 molecules *via* narcissistic self-sorting (2 **CC1** + 1 **[4+4]**). The fewer product molecules in the

narcissistic self-sorted result has an entropic penalty due to greater loss of rigid body translational and rotational entropy of molecules in solution. This is a known effect in drug binding to proteins^[9] and differences between inter- and intra-molecular reaction rates.^[10] The sum of translational and rotational entropy in solution is expected to vary logarithmically with molecular weight and we estimate this contribution using the formula

$$S_{trans+rot} = A + B \ln m$$

where m is the molecular weight and A and B are treated as constants (although their values will depend on molecular shape). We use estimates of these constants given by Murray and Verdonk,^[9] which were fitted to a large set of molecules: $A = 57 \text{ J mol}^{-1} \text{ K}^{-1}$, $B = 5$.

Table S1: Estimated rigid molecule entropy and free energy differences between narcissistic and social self-sorting products. The balanced stoichiometry between narcissistic and social self-sorting products is given in the first two columns and all energies are given for these stoichiometries.

Narcissistic self sorting products	Social self sorting products	DFT calculated formation energy difference (social – narcissistic, kJ mol ⁻¹)	Estimated entropy difference (social – narcissistic, 298 K, J mol ⁻¹ K ⁻¹)	Estimated TΔS, 298 K (social – narcissistic, kJ mol ⁻¹)	Estimated ΔG, 298 K (social – narcissistic, kJ mol ⁻¹)
2 CC1 + 1 [4+4]	4 OCP1	+78.0	+291	+86.8	-8.8
2 CC3 + 1 [4+4]	4 OCP3	+67.2	+300	+89.3	-22.1
2 CC4 + 1 [4+4]	4 OCP4	+126.0	+297	+88.6	+37.4
4 [2+3]-prop + 1 [4+4]	4 OCP-prop	+146.4	-269	-80.1	+226.5
1 CC11 + 1 CC7	4 OCP7	+109.2	+659	+196.4	-87.2

Ignoring the entropy contributions from other terms, such as differences in low frequency intramolecular vibrations between products and solvent ordering around the cages, the entropic contributions are compared to the DFT calculated energy differences in Table S1. For **OCP1**, **OCP3** and

OCP7, our estimate of the entropy difference outweighs the formation energy preference for narcissistic self-sorting. This suggests that social self-sorting is thermodynamically favoured in each of these cases, although the energetic and entropic contributions are finely balanced for **OCP1**. **OCP-prop** is disfavoured entropically and, in the case of **OCP4**, the entropic stabilisation is predicted to be smaller than the energetic preference for narcissistic self-sorting. These results are in excellent agreement with the results of the high-throughput synthetic screening and are in-line with our predictions of the most likely synthesisable cage pots based on formation energy per imine bond.

1.2. Void analysis

Void analysis was carried out using zeo++^[11] and a probe size of 1.42 Å^[12] (equivalent to the kinetic radius of hydrogen gas) to determine the porosity of the structures. Whilst there were some predicted structures containing 1-, 2-, and 3-dimensional void networks when analysing the entire landscape (**OCP3-S**: 0D – 19%, 1D – 34%, 2D – 15%, 3D – 33%; **OCP3-R/S**: 0D – 28%, 1D – 29%, 2D – 16%, 3D – 26%), essentially all of the low energy structures are 0D (Figure S1). Typically, it was also found that the lower density structures in both landscapes have larger cavity diameters (LCD), although this was due to extrinsic voids, and that the pore limiting diameters (PLD) were generally very small (Figure S2). In relation to the lowest energy structures from the CSP, **OCP3-S** had zero-dimensionality with only extrinsic isolated voids (Figure S3), a PLD of 2.8 Å and a LCD of 4.6 Å, whereas **OCP3-R/S** had zero-dimensionality but with a mixture of intrinsic (between window-to-window packed pots) and extrinsic isolated voids (Figure S4), a PLD of 1.2 Å and a LCD of 4.5 Å. There were a few low energy structures with 1D dimensionality that had low density (three structures for **OCP3-R/S**, and a single structure for **OCP3-S**), and had larger PLDs, but this was due to the presence of 1D extrinsic channels (Figure S5-S6).

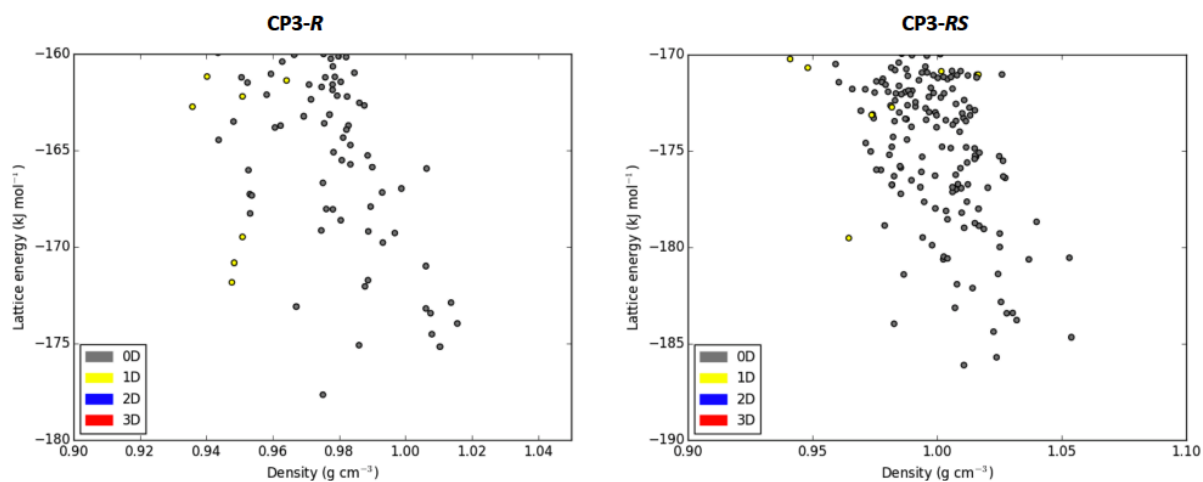


Figure S1: Analysis of the dimensionality of the low energy (within 20 kJ mol⁻¹ of the global minimum) hypothetical crystal structures of **OCP3-S** (left) and **OCP3-RS** (right), analysed with a 1.42 Å probe. Almost all structures are 0-D (grey), with a few low-density structures being 1-D (yellow).

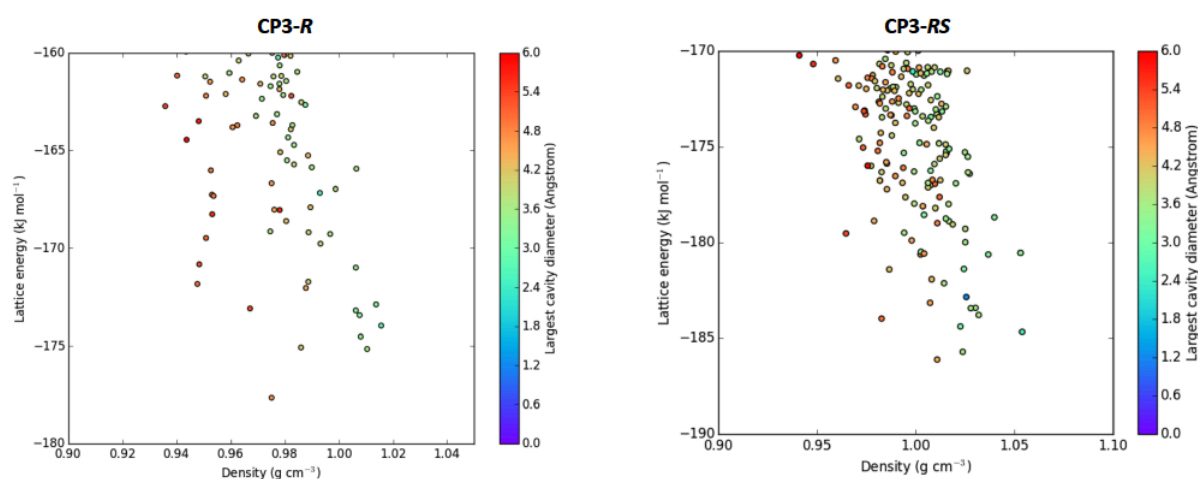


Figure S2: Analysis of the largest cavity diameter of the low energy hypothetical crystal structures of **OCP3-S** (left) and **OCP3-RS** (right).

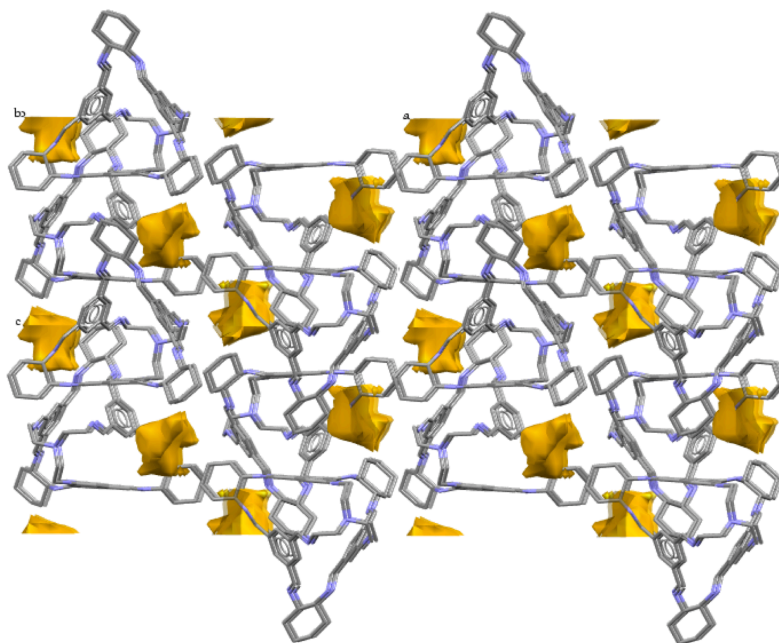


Figure S3: The global minimum structure of **OCP3-S**, with isolated extrinsic voids shown in yellow, calculated with a 1.42 Å probe. The PLD is 2.8 Å and the LCD is 4.6 Å. Carbons are shown in grey, nitrogens in purple and hydrogens omitted.

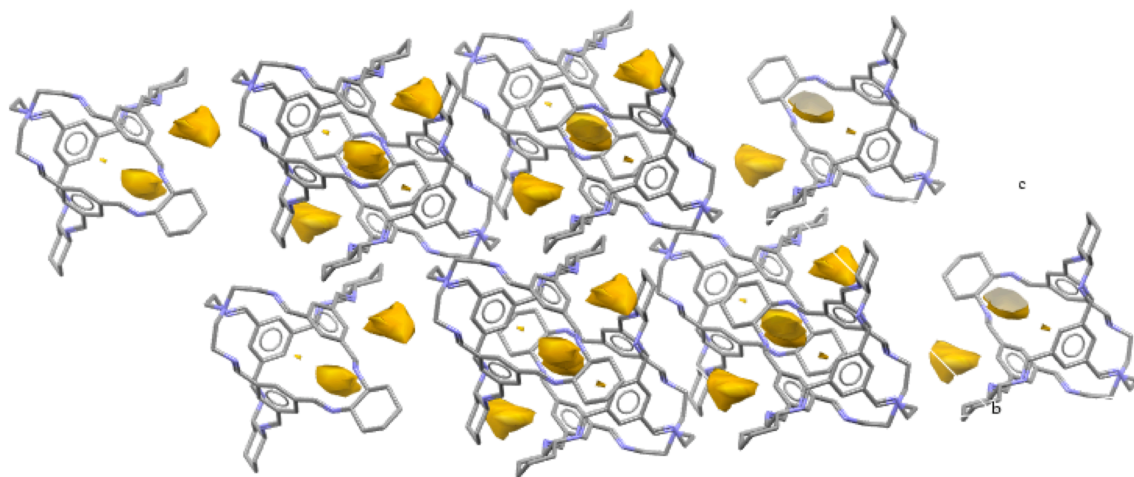


Figure S4: The global minimum structure of **OCP3-RS**, with a mixture of isolated extrinsic voids and intrinsic voids sitting between cages packed window to window. Voids are shown in yellow, calculated with a 1.42 Å probe. The PLD is 1.2 Å and the LCD is 4.5 Å. Carbons are shown in grey, nitrogens in purple and hydrogens omitted.

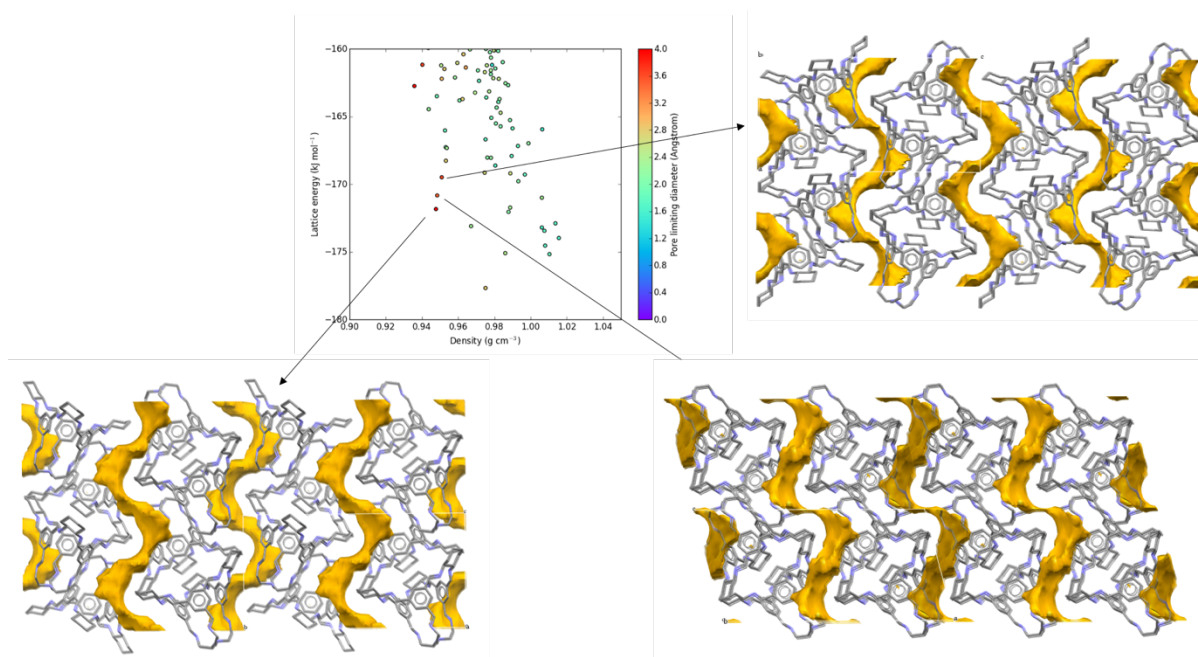


Figure S5: OCP3-S hypothetical crystal structures with one-dimensional extrinsic channels between the pots.

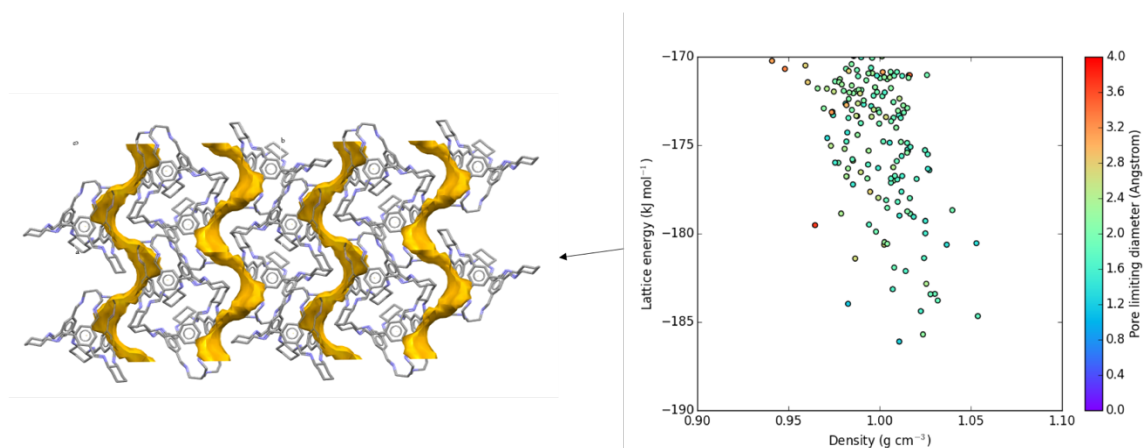


Figure S6: OCP3-RS hypothetical crystal structure with one-dimensional extrinsic channels between the pots.

1.3. Dimer Calculations

Dimer interaction energies were calculated from manually created window-to-window arrangements of molecules, which were optimised using dispersion-corrected density functional theory, using the B3LYP functional, 6-31G** basis set and Grimme's D3 dispersion correction with Becke-Johnson damping (GD3BJ). These calculations were performed using Gaussian 09 revision D.01.^[13] Interaction

energies were calculated as the difference between the DFT-D dimer energy and the energy of the individual optimized monomers.

Table S2: DFT-D interaction energies calculated for isolated dimer pairs of the molecules arranged in a window-to-window packing mode.

First cage in pair	Second cage in pair	Interaction energy (kJ mol ⁻¹)
CC3-S	CC3-S	-175.8
OCP3-S	OCP3-S	-125.4
CC3-R	CC3-S	-197.9
OCP3-R	OCP3-S	-170.9
CC3-S	OCP3-S	-147.8
CC3-R	OCP3-S	-181.2

1.4. Crystal Structure Prediction (CSP)

OCP3 CSP. Crystal structure prediction was performed using a quasi-random sampling procedure, as implemented in our in-house Global Lattice Energy Explorer (GLEE) software.^[14] Single molecules were optimized using DFT at the B3LYP/6-31G** with GD3BJ dispersion correction. These molecular geometries were then held rigid throughout crystal structure generation and lattice energy minimization.

Trial crystal structures were generated with one molecule in the asymmetric unit, generating 5000 crystal structures in each of the 14 Sohncke space groups for enantiomerically pure crystal structures (International space group numbers 1, 4, 5, 18, 19, 20, 76, 78, 92, 96, 144, 145, 169 and 170) and 14 space groups for racemic crystal structures (International space group numbers: 2, 7, 9, 13, 14, 15, 29, 33, 43, 56, 60, 61, 88, 148).

Initial crystal generation involves a low-discrepancy sampling of all structural variables within each space group: unit cell lengths, angles, molecular positions and orientations within the asymmetric unit. Space-group symmetry was then applied, and a geometric test was performed for overlap between molecules. Molecular clashes were removed by lattice expansion. Lattice energy calculations were performed with using DMACRYS2.2.1^[15] with atom–atom repulsion and dispersion interactions modelled with the revised W99 intermolecular potential.^[16] Electrostatic interactions were modelled

using an atomic multipole description of the molecular charge distribution (up to hexadecapole on all atoms) from the B3LYP/6-311G**-calculated charge density using a distributed multipole analysis.^[17] Charge–charge, charge–dipole and dipole–dipole interactions were calculated using Ewald summation; all other intermolecular interactions were summed to a cut-off between molecular centres-of-mass of 30 Å.

Each trial structure was optimized using multiple rigid lattice energy minimizations. During the initial lattice energy minimization point-charges obtained from a MULFIT8 fit of atomic charges to the molecular electrostatic potential generated from the B3LYP/6-311G**-distributed multipole analysis atomic multipoles were used in conjunction with a pressure of 0.01 GPa. The resulting structure was re-optimised with the full multipole electrostatic model without an applied pressure.

Duplicate crystal structures were removed using the crystal structure comparison COMPACK method,^[18] initially within each individual space group, followed by duplicate removal across space groups.

CC3-R(OCP3-S)₄ CSP. Crystal structure prediction was performed on the pentamer unit **CC3-R(OCP3-S)₄**, with the four **OCP3-S** molecules arranged in window-to-window arrangement with a central **CC3-R** molecule. Monomers were held at their isolated-molecule DFT optimized geometries and the pentamer was initially optimized using the revised W99 force field. The **CC3-R(OCP3-S)₄** pentamer was held rigid during crystal structure generation, but all molecules were allowed to move independently during lattice energy minimization of the computer-generated crystal structures. 5000 trial crystal structures were generated using quasi-random sampling within the same 14 Sohnke space groups as used above, followed by lattice energy minimization using the same force field. A large number of these trial crystal structures failed to find a valid local minimum on the lattice energy surface, so that we did not generate a complete set of possible crystal structures of the hypothetical co-crystal. In total, 30,362 successful lattice energy minimizations of trial structures were completed across the 14 space groups.

Although insufficient to be confident that all low energy co-crystal structures had been located, the sampling was sufficient to give an indication of the stability of possible co-crystal structures. To judge the likelihood of co-crystal formation, we compared the lattice energy of the best CSP co-crystal structure (-832.12 kJ/mol per **CC3-R(OCP3-S)₄**) to the calculated lattice energies of the lowest energy individual crystals of **OCP3-S** (-177.66 kJ/mol) and **CC3-R** (-228.79 kJ/mol). The stoichiometrically weighted sum of **OCP3-S** and **CC3-R** crystals (-177.66 x 4 + -228.79 = -939.43 kJ/mol) is much more stable than the best CSP co-crystal, indicating the co-crystallization is unlikely.

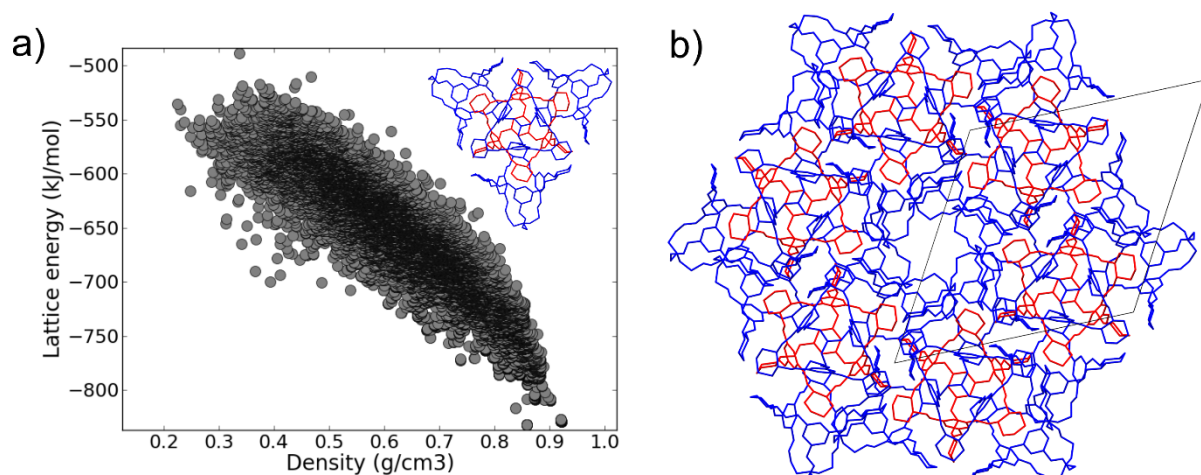


Figure S7: a) Crystal structure prediction results for packing possibilities of the **CC3-R(OCP3-S)₄** pentamer unit (shown as inset, **CC3-R** colored red, **OCP3-S** colored blue). b) Crystal packing in the lowest energy predicted crystal structure of **CC3-R(OCP3-S)₄**. The monoclinic unit cell is shown in black.

Analysis of window-to-window packing

The sets of CSP structures for **OCP3-S** and **OCP3-R/S** were analysed for the presence of window-to-window packing of **OCP3** molecules using the COMPACK algorithm, implemented in the Mercury^[19] crystal structure visualiser. Reference window-to-window dimers (**OCP3-S: OCP3-S** and **OCP3-S: OCP3-R**) were taken from the dimer interaction energy studies (see above) and all CSP structures were compared against these dimers using loose geometric tolerances (40% on atom-atom distances and 40° on angles of interatomic contacts). Hits to the reference **OCP3** dimers were checked visually.

2. General Synthetic and Analytical Methods

Materials: Chemicals were purchased from Sigma-Aldrich, Fluorochem, or Manchester Organics. Solvents were reagent or HPLC grade purchased from Fisher Scientific. All chemicals and solvents were used as received unless specified.

Synthesis: All reactions were stirred magnetically using Teflon-coated stirring bars. Where heating was required, the reactions were warmed using a stirrer hotplate with heating blocks with the stated temperature being measured externally to the reaction flask with an attached probe. Removal of solvents was done using a rotary evaporator.

High-throughput cage discovery: High-throughput automated synthesis was conducted using a Chemspeed Accelerator SLT-100 automated synthesis platform.

IR spectra: Infra-red (IR) spectra were recorded on a Bruker Tensor 27 FT-IR using ATR measurements for solids as neat samples.

NMR Spectra: ^1H Nuclear magnetic resonance (NMR) spectra were recorded using an internal deuterium lock for the residual protons in CDCl_3 ($\delta = 7.26$ ppm) at ambient probe temperature using either a Bruker Avance 400 (400 MHz) or Bruker DRX500 (500 MHz) instrument. NMR data are presented as follows: chemical shift, integration, peak multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad, app = apparent) and coupling constants (J / Hz). Chemical shifts are expressed in ppm on a δ scale relative to δ_{CDCl_3} (7.26 ppm) and coupling constants, J , are given in Hz. ^{13}C NMR spectra were recorded using an internal deuterium lock using CDCl_3 ($\delta = 77.16$ ppm) at ambient probe temperatures using either a Bruker Avance 400 (101 MHz) or Bruker DRX500 (126 MHz) instrument.

Analytical HPLC: HPLC was conducted on a Dionex UltiMate 3000 equipped with a diode array UV detector using a Thermo-Scientific Synchronis C8 column, 150x4.6 mm, 3 μm (SN 10136940, Lot 12459). The mobile phase was isocratic MeOH at a flow rate of 1 mL/min for a 10 min run time, and the column oven temperature was set to 30 $^\circ\text{C}$. The injection volume was 10 μL and the sample concentration was approximately 1 mg/mL. Detection for UV analysis was conducted at 254 nm.

Preparative HPLC: The column used to purify the cage pots was a Thermo Scientific Synchronis C8, 150x50mm, 5 μm (SN 97205-159370, Lot 12105). The mobile phase was isocratic MeOH at a flow rate of 40 mL/min for an 8 min run time, and the column oven temperature was set to 30 $^\circ\text{C}$. The injection volume was 1500 μL and the sample concentration was approximately 20 mg/mL in a DCM/MeOH (1:1) solvent mixture. Detection for UV analysis was conducted at 254 nm.

HRMS: High resolution mass spectrometry (HRMS) was carried out using an Agilent Technologies 6530B accurate-mass QTOF Dual ESI mass spectrometer (MeOH + 0.1% formic acid, 0.25 mL/min, capillary voltage 4000 V, fragmentor 225 V) in positive-ion detection mode.

SCXRD: Data sets were measured on a Rigaku MicroMax-007 HF rotating anode diffractometer (Mo-K α radiation, Kappa 4-circle goniometer, Rigaku Saturn724+ detector); Rigaku frames were converted to Bruker compatible frames using the programme ECLIPSE.^[20] Absorption corrections, using the multi-scan method, were performed with the program SADABS.^[21,22] Structures were solved with SHELXT,^[23] or by direct methods using SHELXS,^[24] and refined by full-matrix least squares on $|F|^2$ by SHELXL,^[23] interfaced through the programme OLEX2.^[25] Unless stated, all non-H atoms were refined anisotropically, and unless stated H-atoms were fixed in geometrically estimated positions and refined using the riding model. Supplementary CIFs, that include structure factors, are available free of charge from the Cambridge Crystallographic Data Centre (CCDC) via www.ccdc.cam.ac.uk/data_request/cif.

PXRD: Laboratory powder X-ray diffraction data were collected in transmission mode on samples held on thin Mylar film in aluminium well plates on a Panalytical X'Pert PRO MPD equipped with a high throughput screening (HTS) XYZ stage, X-ray focusing mirror and PIXcel detector, using Ni-filtered Cu K α radiation. Data were measured over the range 4–40° in ~0.013° steps over 60 minutes.

TGA: Thermogravimetric analysis was carried out using a Q5000IR analyser (TA instruments) with an automated vertical overhead thermobalance. The sample was heated in aluminum pans under nitrogen at a rate of 10 °C/min to 550 °C.

3. High-Throughput Synthetic Screening

High-throughput screening method: All tri-topic aldehyde, tri-topic amine, and di-topic amine precursors (Figure S8) were dissolved in CDCl_3 to make stock solutions ($5 \text{ mg} \cdot \text{mL}^{-1}$) for use in a screen (Table S2). On a Chemspeed Accelerator SLT-100 platform,^[8] the required volume of each amine stock solution, followed by the required volume of each aldehyde stock solution, was added to jacketed reactors (27 mL maximum volume) via liquid dispensing, followed by additional CDCl_3 to make each total volume up to 13 mL (Table S3). The resulting solutions were vortexed at 800 rpm at 65°C for 3 days prior to analysis.

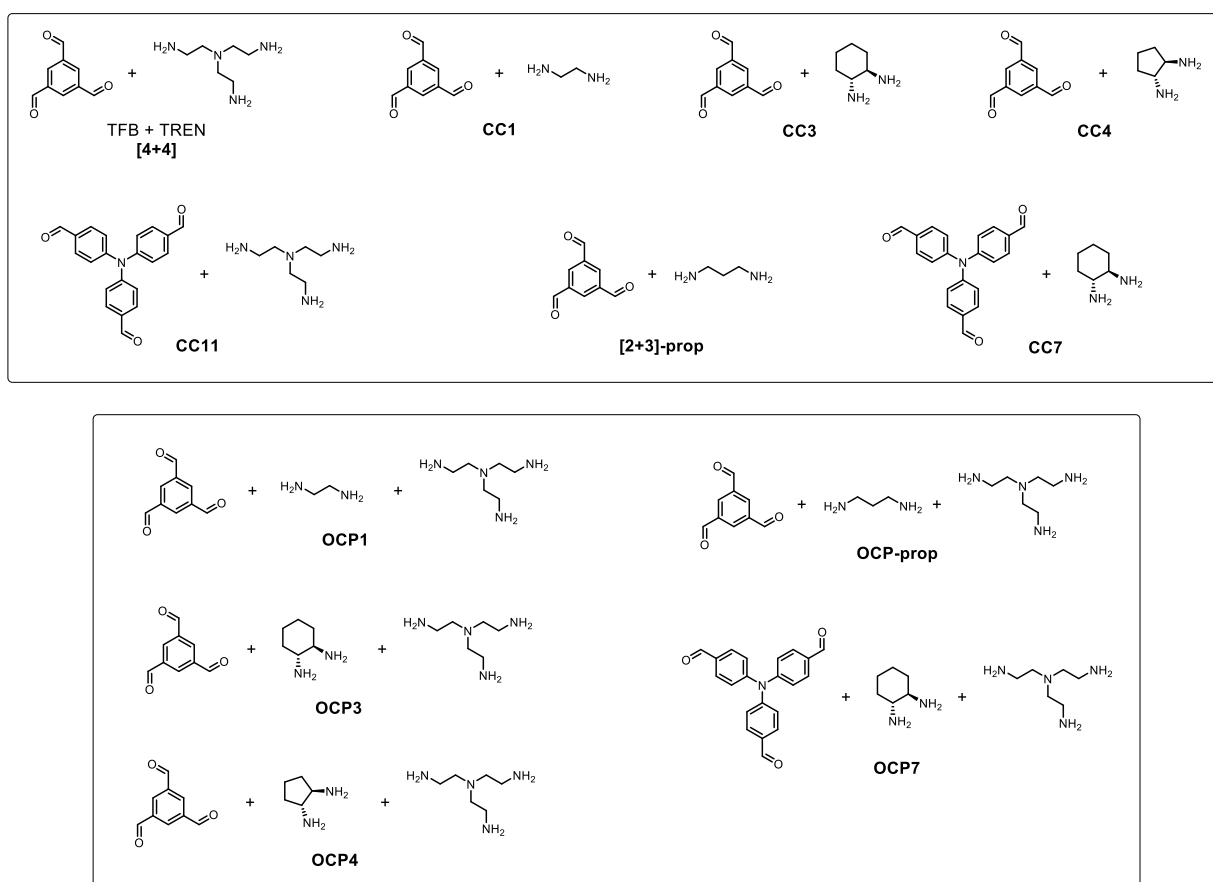
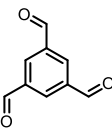
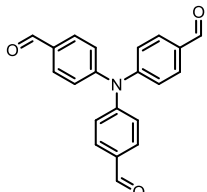
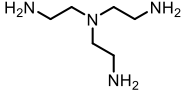
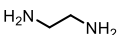


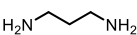


Figure S8: High-throughput screen - binary parent cage combinations (top) and socially self-sorted organic cage pot combinations (bottom).

Table S3: Precursor stock solutions for high-throughput screening

						
Aldehyde 1	Aldehyde 2	Amine 1	Amine 2	Amine 3	Amine 4	Amine 5

Stock Solution Number	Reactant	MW (g.mol ⁻¹)	Stock Solution Concentration (mg. mL ⁻¹)	Stock Solution Concentration (mmol.mL ⁻¹)
1	Aldehyde 1	162.14	5	0.0308
2	Aldehyde 2	329.36	5	0.0152
3	Amine 1	146.24	5	0.0341
4	Amine 2	60.10	5	0.0832
5	Amine 3	114.19	5	0.0438
6*	Amine 4	100.17	5	0.0289
7	Amine 5	74.13	5	0.0674

*Dihydrochloride salt of diamine used, so 0.013 mL (0.0953 mmol, 3.3 eq.) triethylamine added per 1 mL of stock solution formed

Table S4: Reaction volumes used in each reaction on the Chemspeed Accelerator SLT-100

Target Species	Ratio of Aldehyde: Di-topic Amine: Tri-topic Amine used	Aldehyde Reactant	Volume Aldehyde Stock Solution (mL)	Amount of Aldehyde (mmol)	Di-topic Amine Reactant	Amount of Di-topic Amine required (mmol)	Volume Di-topic Amine Stock Solution (mL)	Tri-topic Amine Reactant	Amount of Tri-topic Amine required (mmol)	Volume Tri-topic Amine Stock Solution (mL)	Additional Amount of CDCl ₃ (mL) added (Total Volume = 13 mL)
[4+4]	4 : 0 : 4	1	1.56	0.0481	N/A	-	-	1	0.0481	1.41	10.03
CC1	4 : 6 : 0	1	1.56	0.0481	2	0.0722	0.87	N/A	-	-	10.57
CC3	4 : 6 : 0	1	1.56	0.0481	3	0.0722	1.65	N/A	-	-	9.79
CC4	4 : 6 : 0	1	1.56	0.0481	4	0.0722	2.50	N/A	-	-	8.94
[2+3]-prop	4 : 6 : 0	1	1.56	0.0481	5	0.0722	1.07	N/A	-	-	10.37
CC11	4 : 0 : 4	2	3.17	0.0481	N/A	-	-	1	0.0481	1.41	8.42
CC7	4 : 6 : 0	2	3.17	0.0481	3	0.0722	1.65	N/A	-	-	8.18
OCP1	3 : 3 : 1	1	1.56	0.0481	2	0.0481	0.58	1	0.0161	0.47	10.39
OCP3	3 : 3 : 1	1	1.56	0.0481	3	0.0481	1.10	1	0.0161	0.47	9.87
OCP4	3 : 3 : 1	1	1.56	0.0481	4	0.0481	1.67	1	0.0161	0.47	9.31
OCP-prop	3 : 3 : 1	1	1.56	0.0481	5	0.0481	0.71	1	0.0161	0.47	10.26
OCP7	3 : 3 : 1	2	3.17	0.0481	3	0.0481	1.10	1	0.0161	0.47	8.26

To determine if any cage or pot formation had occurred, crude reaction aliquots were taken for ^1H NMR and HRMS analysis – for a summary of the results see Table S4. Overall, there were two clear hits where a large amount of pot formation was apparent: **OCP3** and **OCP7**, which were also subsequently analysed by HPLC – see Figures S9 to S14 for ^1H NMR, HRMS, and HPLC spectra.

Table S4: Summary of the ^1H NMR and HRMS characterization data from the high-throughput screen on crude reaction solutions.

Target Species	^1H NMR Spectroscopy (CDCl_3)		HRMS (direct injection)						Outcome
	Clear parent cage formation?	Evidence of organic cage pot formation?	Parent cage formed?	Mass ions indicating cage formation	Main mass ion?	Organic cage pot formed?	Mass ions indicating pot formation	Main mass ion?	
[4+4]	✗	N/A	✗	-	-	N/A	-	-	No cage formation
CC1	✓	N/A	✓	$[\text{M}+\text{H}]^+ 793.4246$ $[\text{M}+\text{Na}]^+ 815.4087$	✓	N/A	-	-	CC1 formed
CC3	✓	N/A	✓	$[\text{M}+\text{H}]^+ 1117.7025$ $[\text{M}+2\text{H}]^{2+} 559.3562$	✓	N/A	-	-	CC3 formed
CC4	✓	N/A	✓	$[\text{M}+\text{H}]^+ 1033.6113$	✗	N/A	-	-	CC4 formed
[2+3]-prop	✗	N/A	✓	$[\text{M}+\text{H}]^+ 439.2630$	✓	N/A	-	-	[2+3]-prop cage formed
CC11	✓	N/A	✓	$[\text{M}+2\text{H}]^{2+} 843.9643$ $[\text{M}+3\text{H}]^{3+} 562.9791$ $[\text{M}+4\text{H}]^{4+} 422.4861$	✓	N/A	-	-	CC11 formed
CC7	✗	N/A	✗	$[\text{M}+2\text{H}]^{2+} 893.5037$ $[\text{M}+3\text{H}]^{3+} 596.0048$	✓	N/A	-	-	No sign of [8+12] CC7, possibly [4+6] cage formed
OCP1	✓	✗	✓	$[\text{M}+\text{H}]^+ 793.4261$ $[\text{M}+2\text{H}]^{2+} 397.2154$	✓	Trace	$[\text{M}+\text{H}]^+ 651.3701$	✗	Mainly CC1 formed
OCP3	✓	✓	✓	$[\text{M}+\text{H}]^+ 1117.7036$ $[\text{M}+2\text{H}]^{2+} 559.3569$	✗	✓	$[\text{M}+\text{H}]^+ 813.5120$ $[\text{M}+2\text{H}]^{2+} 407.2595$	✓	Mixture of OCP3 and CC3
OCP4	✓	✗	✓	$[\text{M}+\text{H}]^+ 1033.6107$	✓	✗	-	-	CC4 formed
OCP-prop	✗	✗	✓	$[\text{M}+\text{H}]^+ 439.2629$	✓	✗	-	-	OCP-prop formed
OCP7	✓	✓	✓	$[\text{M}+2\text{H}]^{2+} 843.9698$	✗	✓	$[\text{M}+\text{H}]^+ 1315.7381$ $[\text{M}+2\text{H}]^{2+} 658.3740$ $[\text{M}+3\text{H}]^{3+} 439.2522$	✓	Mixture of OCP7 and CC11

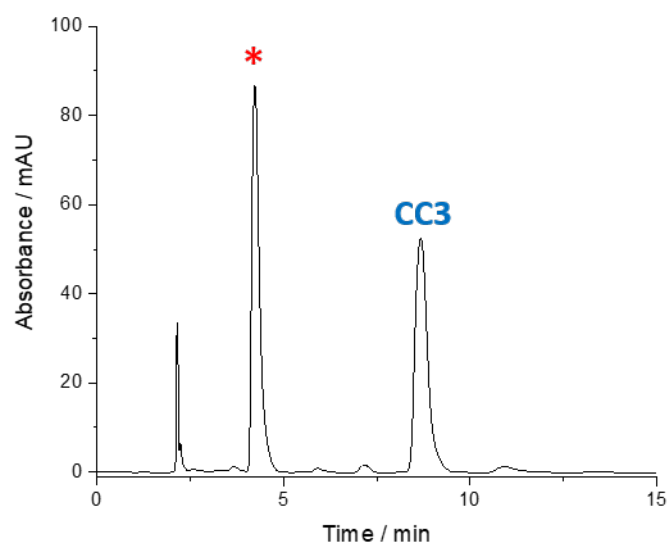


Figure S9: HPLC trace for targeted **OCP3** reaction showing a mixture of organic cage pot **OCP3** (*) and parent cage **CC3**, in an approximate 2:1 ratio, formed during the high-throughput screen.

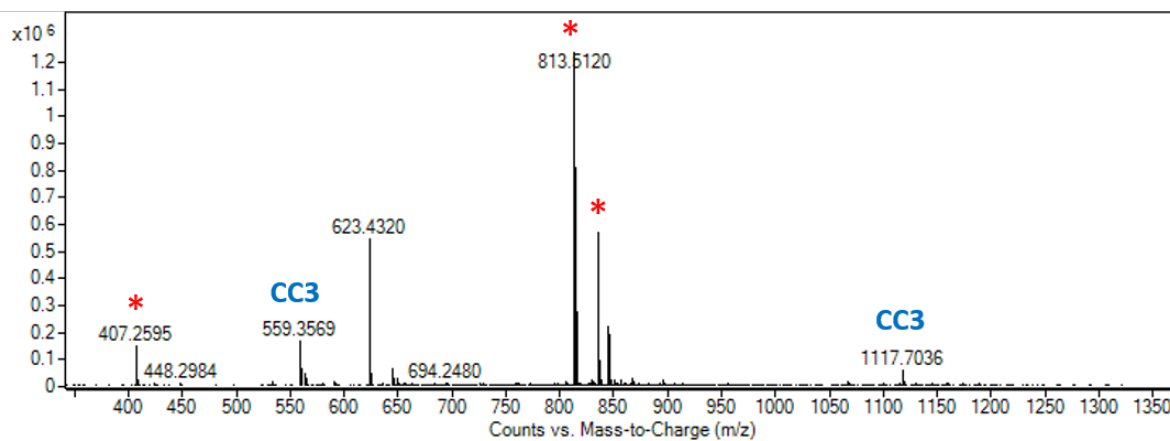


Figure S10: HRMS spectra for targeted **OCP3** reaction, taken directly from the high-throughput screen, indicating clear formation of a mixture of **OCP3** (* - $[M+H]^+$, $[M+Na]^+$ and $[M+2H]^{2+}$) and parent cage **CC3** ($[M+H]^+$ and $[M+2H]^{2+}$).

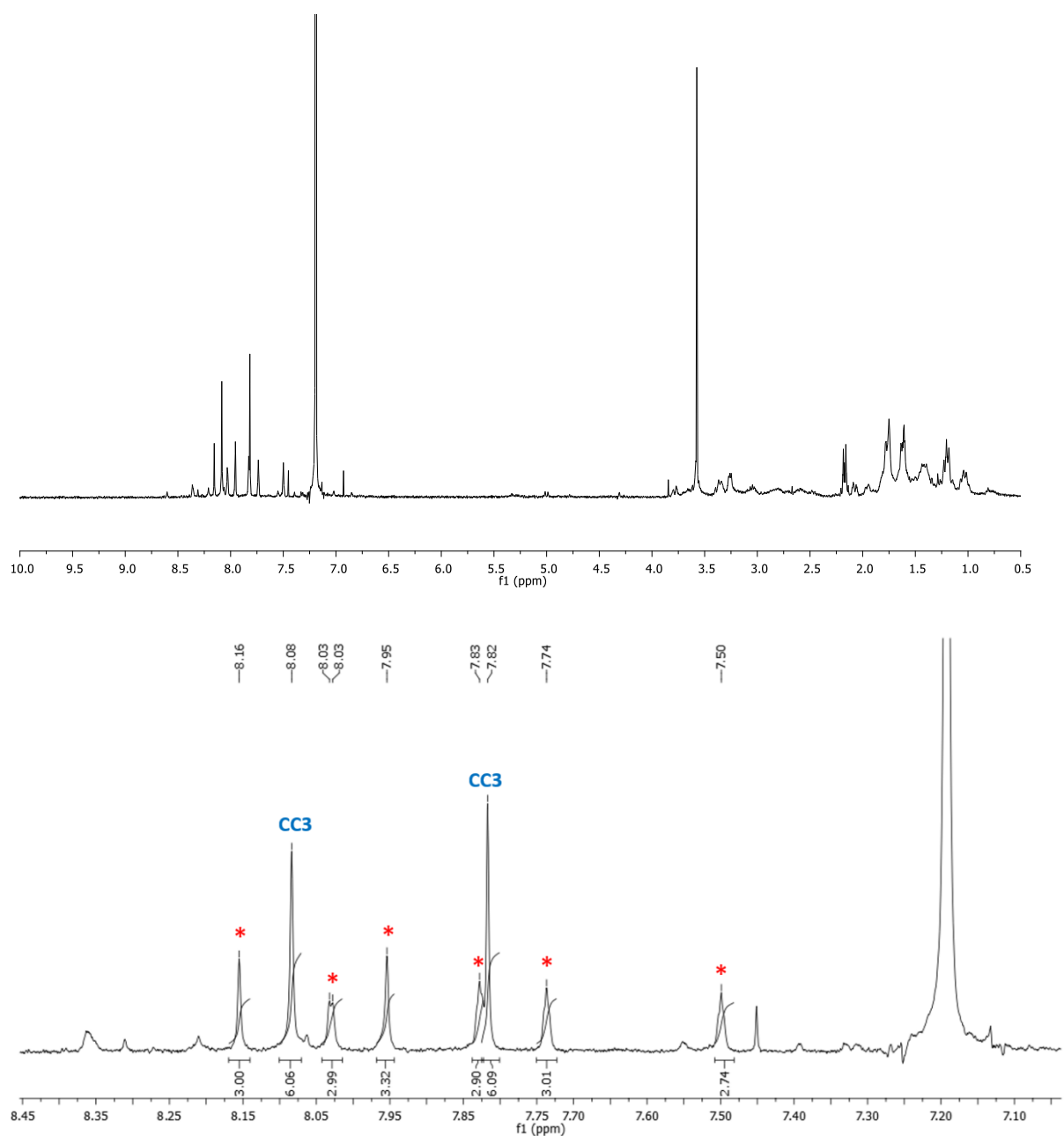


Figure S11: ^1H NMR (CDCl_3) spectra for targeted **OCP3** reaction, taken directly from the high-throughput screen – full spectra (upper), and aromatic region (lower) showing an approximate 2:1 ratio of **OCP3** (*) to parent cage **CC3**.

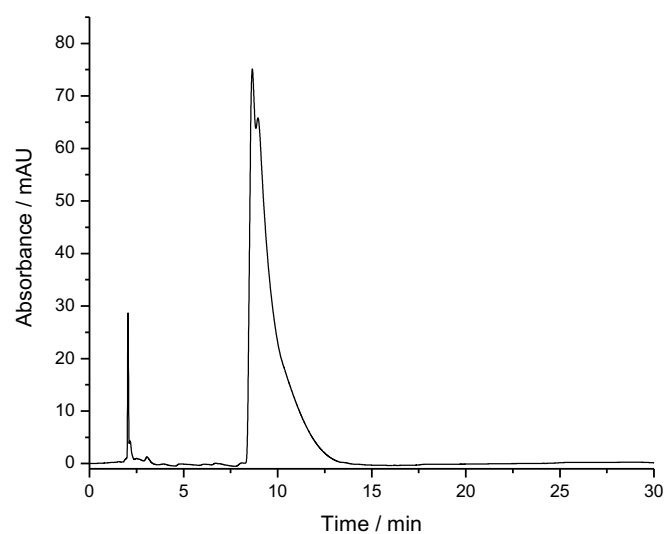


Figure S12: HPLC trace for targeted **OCP7** reaction showing overlapping peaks for organic cage pot **OCP7** and parent cage **CC11**, formed during the high-throughput screen.

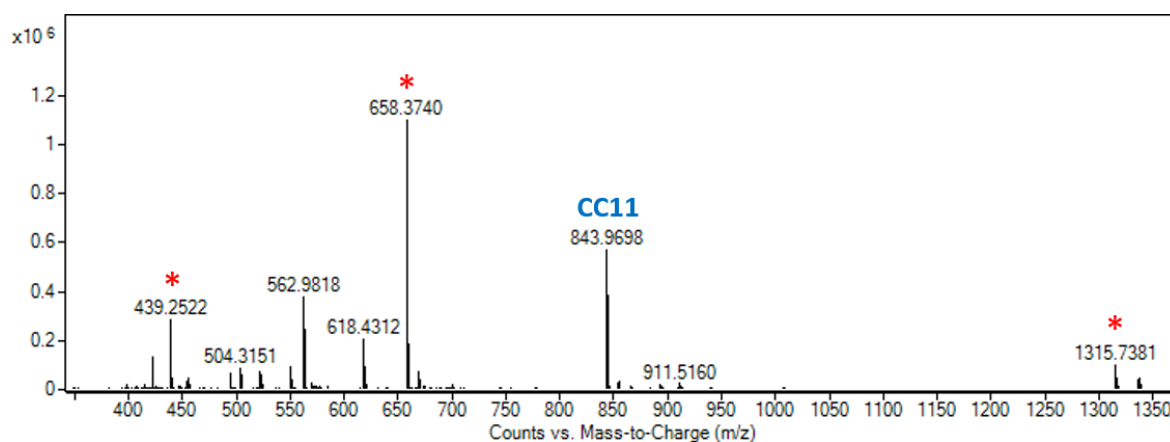


Figure S13: HRMS spectra for targeted **OCP7** reaction, taken directly from the high-throughput screen, indicating clear formation of a mixture of **OCP7** (* - $[M+H]^+$, $[M+2H]^{2+}$ and $[M+3H]^{3+}$) and parent cage **CC11** ($[M+2H]^{2+}$).

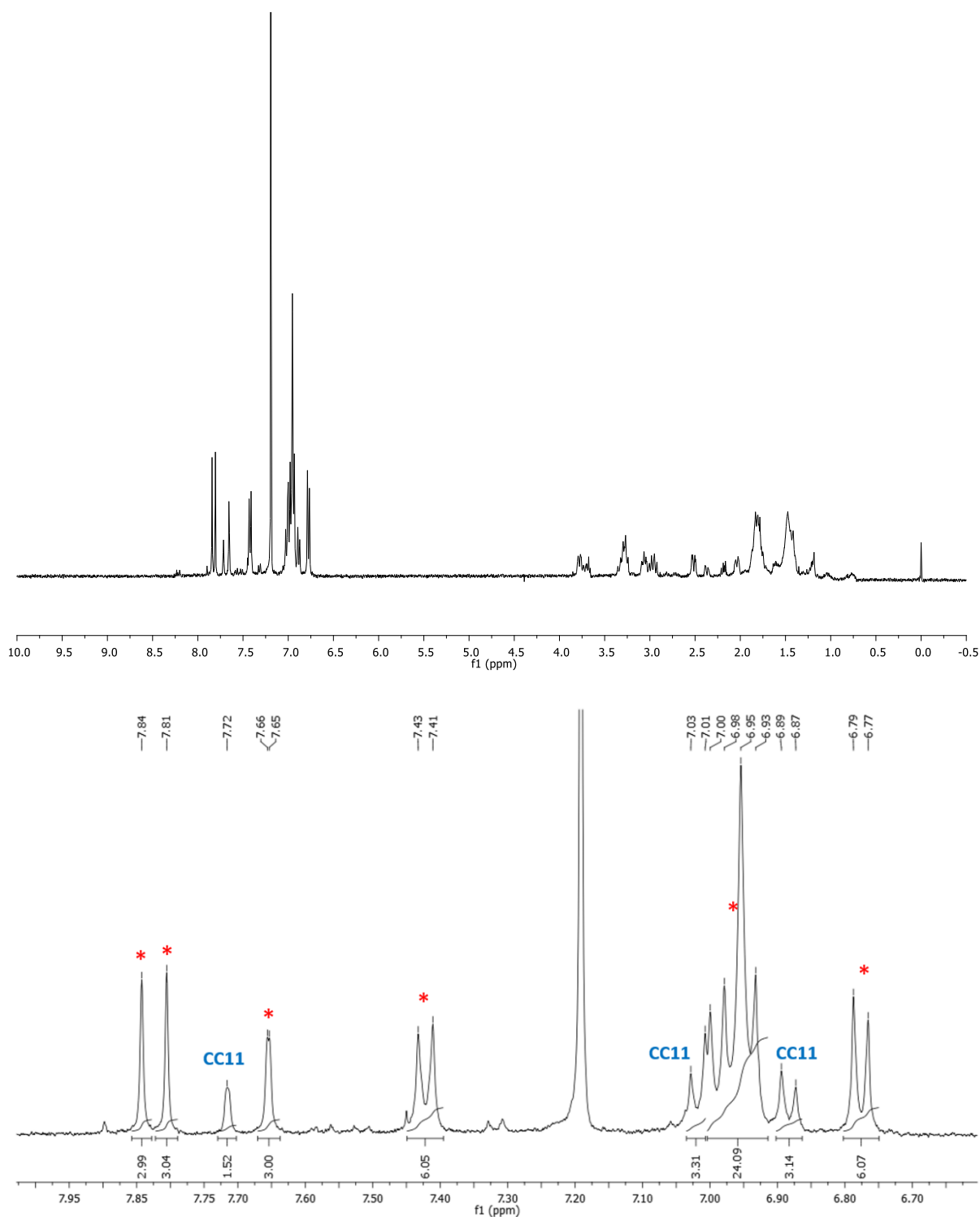
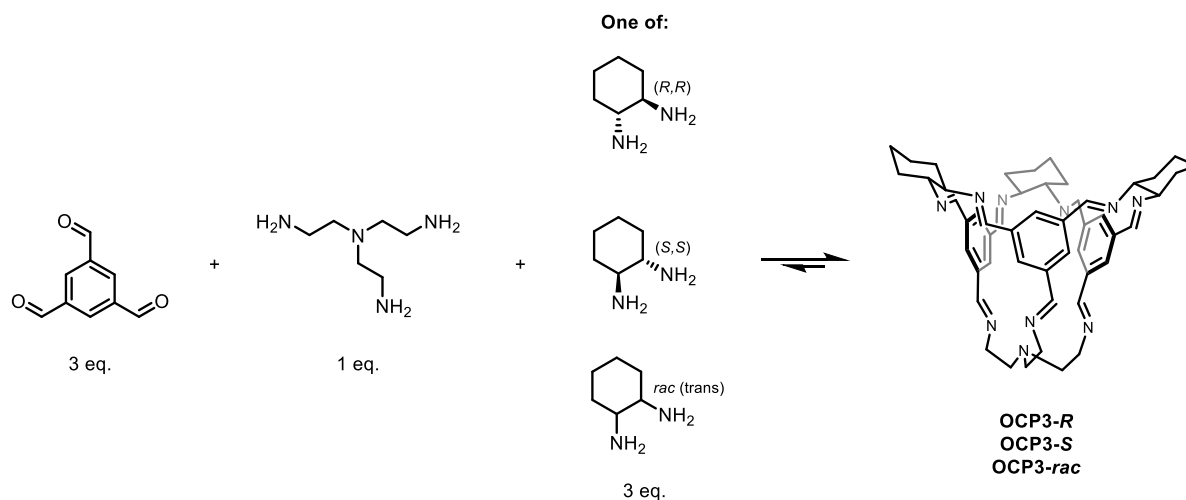


Figure S14: ^1H NMR (CDCl_3) spectra for targeted **OCP7** reaction, taken directly from the high-throughput screen – full spectra (upper), and aromatic region (lower) showing an approximate 8:1 ratio of **OCP7** (*) to parent cage **CC11**.

4. Synthesis, Purification and Characterisation of Cage Pot OCP3



General Procedure for Synthesis of Organic Cage Pots (OCP3-*R*, OCP3-*S*, OCP3-*rac*): To a solution of 1,3,5-triformylbenzene (100 mg, 0.62 mmol, 3.0 eq.) in CHCl₃ (100 mL) was added a solution of tris(2-aminoethyl)amine (30 mg, 0.21 mmol, 1.0 eq.) in CHCl₃ (33 mL), and one of either (1*R*,2*R*)-(-)-1,2-diaminocyclohexane (**OCP3-*R***), (1*S*,2*S*)-(+)-1,2-diaminocyclohexane (**OCP3-*S***) or (±)-*trans*-1,2-diaminocyclohexane (**OCP3-*rac***) (70 mg, 0.62 mmol, 3.0 eq.) in CHCl₃ (33 mL). The resulting solutions were heated at 65 °C and monitored by HPLC and ¹H NMR spectroscopy until no further conversion or equilibration to cage pot was observed (typically 5 days) before being allowed to cool to room temperature.

NB. The ratio of cage:pot formed is directly influenced by the ratio of precursors used, and therefore accurate quantities of each precursor are required to favour and achieve the highest conversion to the socially self-sorted cage pots. On direct concentration of the reaction mixtures, the resulting isolated solid/film was found to be fairly insoluble, suggesting decomposition can occur on isolation. Addition of hexane to the reactions, to precipitate the pots as described below for bulk isolation, enabled the pots to be isolated as impure mixtures with some decomposition, or with an increased quantity of **CC3**, observed. This could be avoided by keeping the pots in solution and isolating analytically pure samples using prep-HPLC as discussed below.

For bulk isolation of crude OCP3-*R* and OCP3-*S*, which seemed less stable to isolation than **OCP3-*rac***, with further formation of **CC3** observed on concentration (see Figure S15): the resulting solutions were carefully concentrated *in vacuo* (20 °C) to approximately half the original volume and hexane (100 mL) was added. The resulting solution was carefully concentrated to ~50 mL, before an additional excess of hexane was added and re-concentrated to ~10 mL. The resulting precipitate was collected by

filtration and dried under vacuum to afford the different cage pots respectively as pale brown solids: **OCP3-R** (44 mg mass recovery (theoretical = 167 mg), 35% purity by HPLC), **OCP3-S** (65 mg mass recovery (theoretical = 167 mg), 19% a/a by HPLC).

For bulk isolation of crude OCP3-rac: The resulting solution was carefully concentrated *in vacuo* (20 °C) to approximately half the original volume and hexane (100 mL) was added, before the resulting solution was concentrated to dryness *in vacuo* to afford the cage pot as a pale brown solid: **OCP3-rac** (quantitative mass recovery (theoretical = 167 mg), 40% a/a by HPLC).

To obtain analytically pure samples for characterisation and crystallisation screens –

The general procedure was repeated, and the reaction allowed to cool to room temperature. HPLC analysis showed reasonably clean crude reactions (see Figure S16), so therefore, 40 mL was removed for direct use in crude crystallisation screens which unfortunately proved unsuccessful. The remaining solution was carefully concentrated *in vacuo* (no heat) to ~15 mL and MeOH (15 mL) was added. The solution was filtered through a 0.2 µm PTFE filter and purified by preparative HPLC. The collected product fractions were combined and carefully concentrated *in vacuo* (20 °C) to <50 mL – HPLC analysis confirmed an improvement in purity for all of the pots from 50–65% a/a to >97% a/a after preparative HPLC (see Figure S16). The total volume was made up to 50 mL by the addition of MeOH, and the solution split for vial-in-vial slow diffusion crystallisation studies with a range of anti-solvents (acetone, diethyl ether, petroleum ether, ethyl acetate, dichloromethane, chloroform, toluene, *p*-xylene, *m*-xylene, *o*-xylene, isopropanol, hexane, pentane, acetonitrile, ethanol, tetrahydrofuran, and 1,4-dioxane), alongside a slow evaporation of the methanol solution. The residual solutions (**OCP3-R** 5 mL, **OCP3-S** 5 mL, **OCP3-rac** 30 mL) had an excess of hexane added and were carefully concentrated *in vacuo* (20 °C) to afford colourless solids which were dried under vacuum to afford the different cage pots respectively: **OCP3-R** (1.9 mg, 97% a/a by HPLC), **OCP3-S** (2.3 mg, 97% a/a by HPLC) and **OCP3-rac** (21 mg, 99% a/a by HPLC).

Spectroscopic data for **OCP3-S**, **OCP3-R**, and **OCP3-rac** was identical (see Figures S17-S25), data was obtained for **OCP3-rac**: IR ($\nu_{\text{max}}/\text{cm}^{-1}$) 2929, 2857, 1646, 1599, 1447, 1377, 1340, 1160, 1028; ^1H NMR (400 MHz, CDCl_3) δ_{H} 8.22 (3H, s), 8.10 (3H, d, J = 1.8 Hz), 8.02 (3H, s), 7.90 (3H, s), 7.81 (3H, s), 7.57 (3H, s), 3.85 (3H, d, J = 12.8 Hz), 3.65 (3H, t, J = 12.1 Hz), 3.51 – 3.35 (6H, m), 3.11 (3H, dd, J = 15.0, 9.1 Hz), 2.14 (3H, dd, J = 12.9, 2.9 Hz), 2.02 (3H, d, J = 10.5 Hz), 1.94 – 1.74 (15H, m), 1.51 (6H, t, J = 9.2 Hz); ^{13}C NMR (101 MHz, CDCl_3) δ_{C} 161.70, 160.79, 160.38, 136.76, 136.57, 136.19, 130.46, 128.68,

128.05, 75.67, 73.75, 59.44, 54.18, 32.45, 31.89, 24.61, 24.58; **HRMS** (ES+) calc. for $C_{51}H_{60}N_{10}$ 812.5002, found $[M+H]^+$ 813.5082, $[M+2H]^{2+}$ 407.2591.

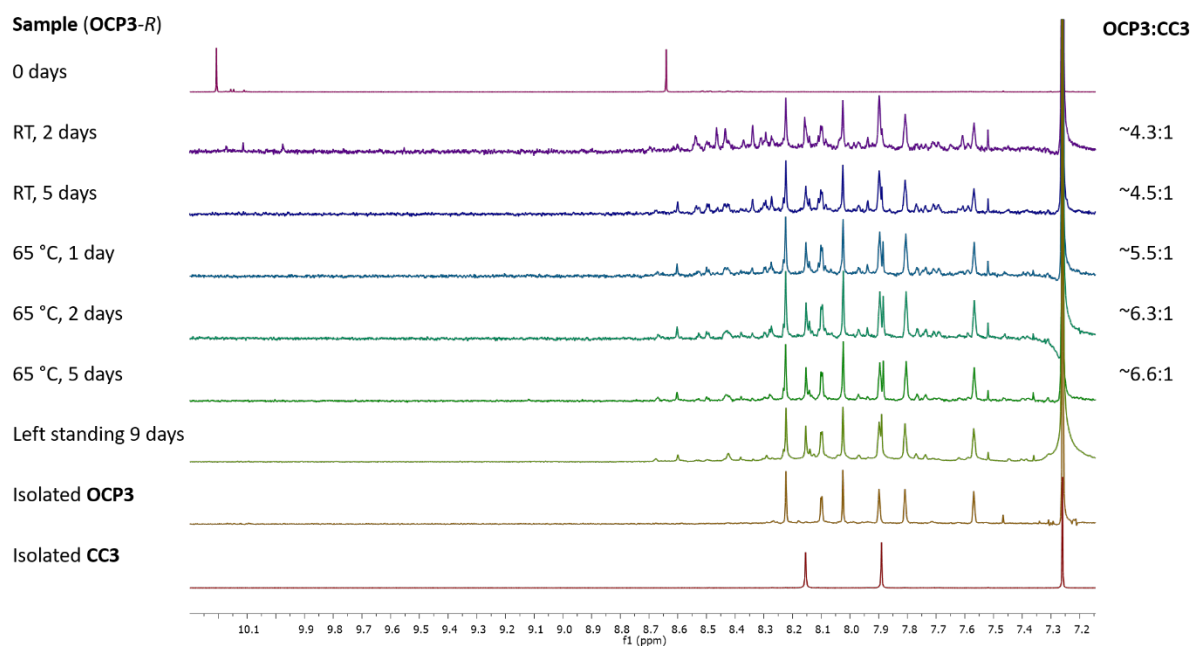


Figure S15: Stacked ^1H NMR spectra (CDCl_3) for the formation of **OCP3-R**, showing an improvement in the ratio of **OCP3:CC3** over time with heating, and a reduction in side-product impurities.

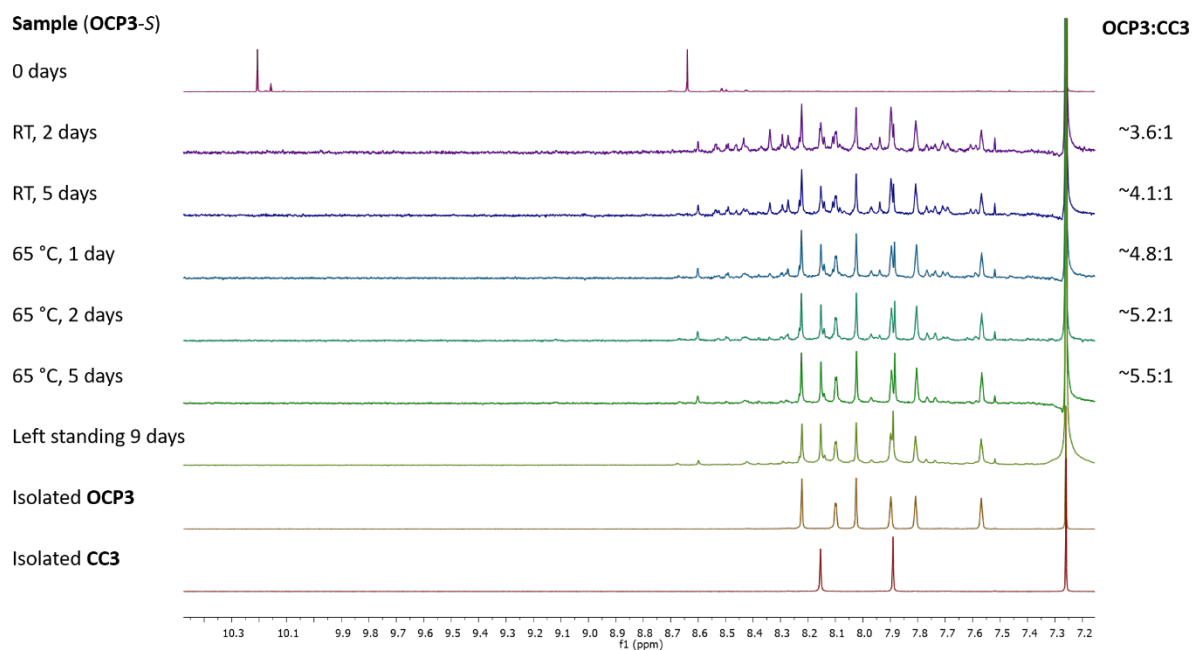


Figure S16: Stacked ^1H NMR spectra (CDCl_3) for the formation of **OCP3-S**, showing an improvement in the ratio of **OCP3:CC3** over time with heating, and a reduction in side-product impurities.

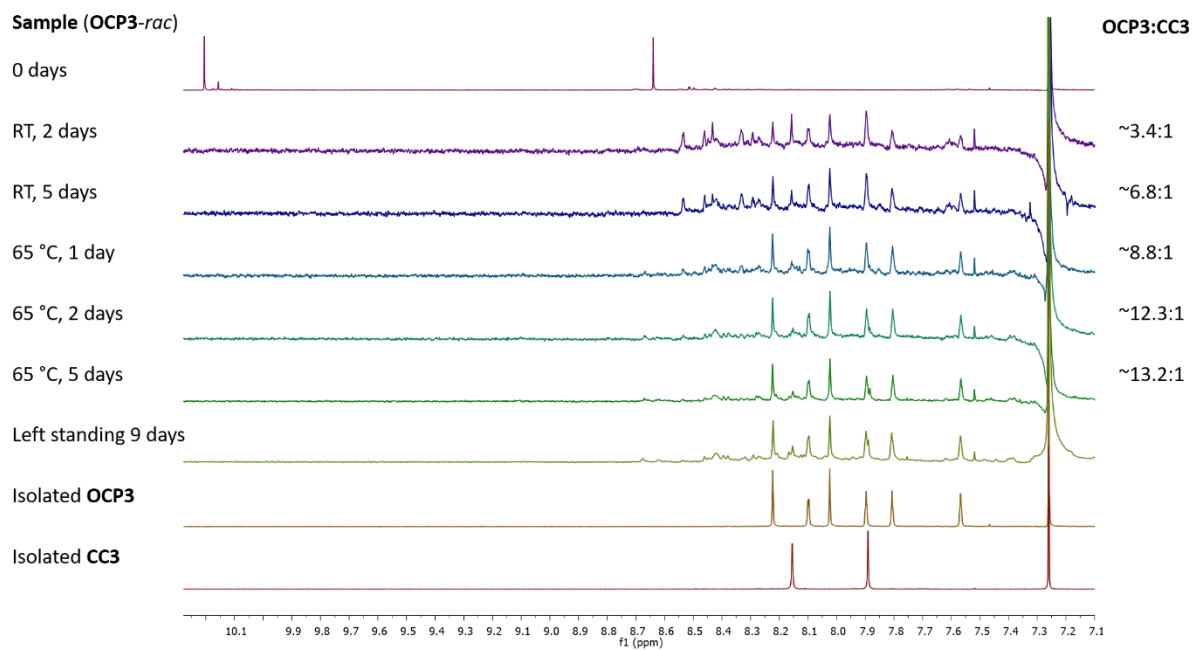


Figure S17: Stacked ^1H NMR spectra (CDCl_3) for the formation of **OCP3-rac**, showing an improvement in the ratio of **OCP3:CC3** over time with heating, and a reduction in side-product impurities.

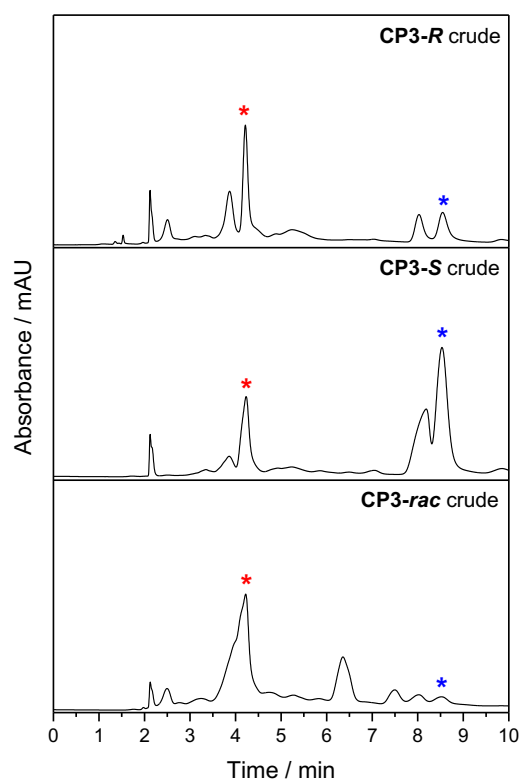


Figure S18: Stacked HPLC spectra for non-purified crude cage pots **OCP3-R** (top), **OCP3-S** (middle) and **OCP3-rac** (bottom) after directly isolating from the crude reaction mixture by solvent exchange with both cage pot (*) and CC3 (*) visible.

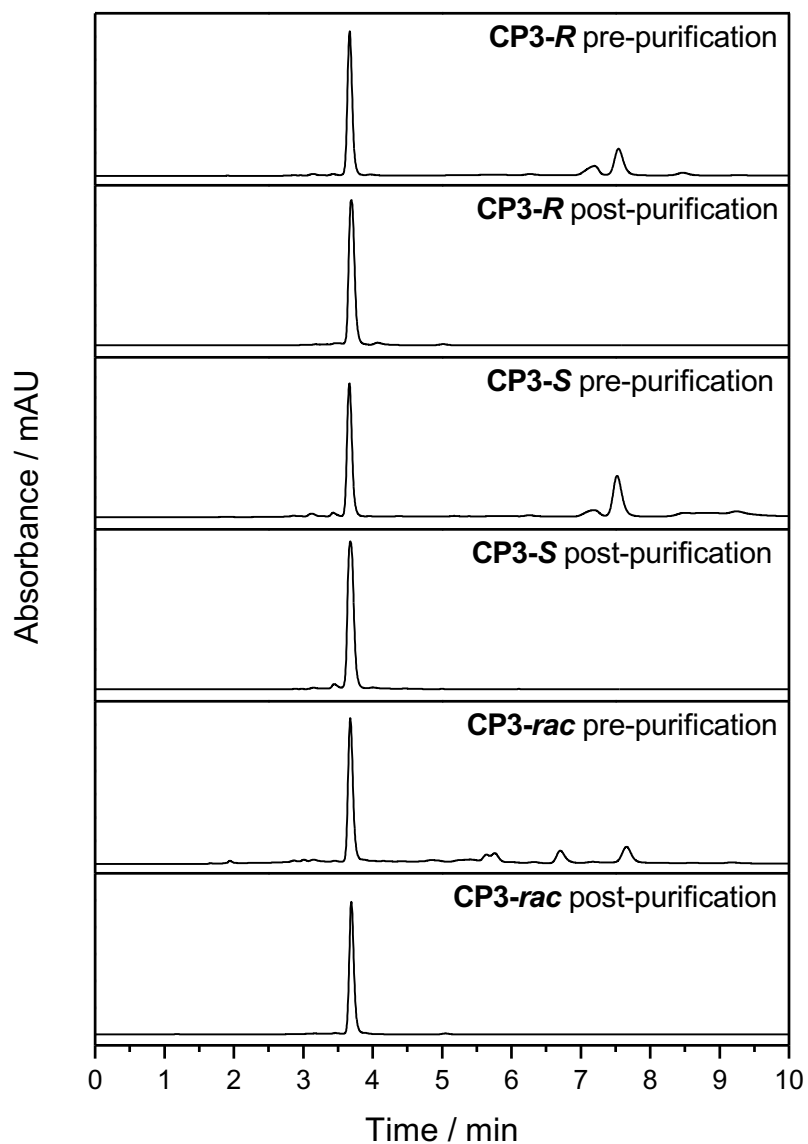


Figure S19: Stacked HPLC traces showing cage pots **OCP3-R**, **OCP3-S**, and **OCP3-rac** at 3.7 min, pre-purification (cage pots not isolated — aliquot taken after concentration of crude reaction mixture and addition of MeOH), and post-purification by preparative HPLC, showing an improvement in purity from 50–65% a/a to >97% a/a.

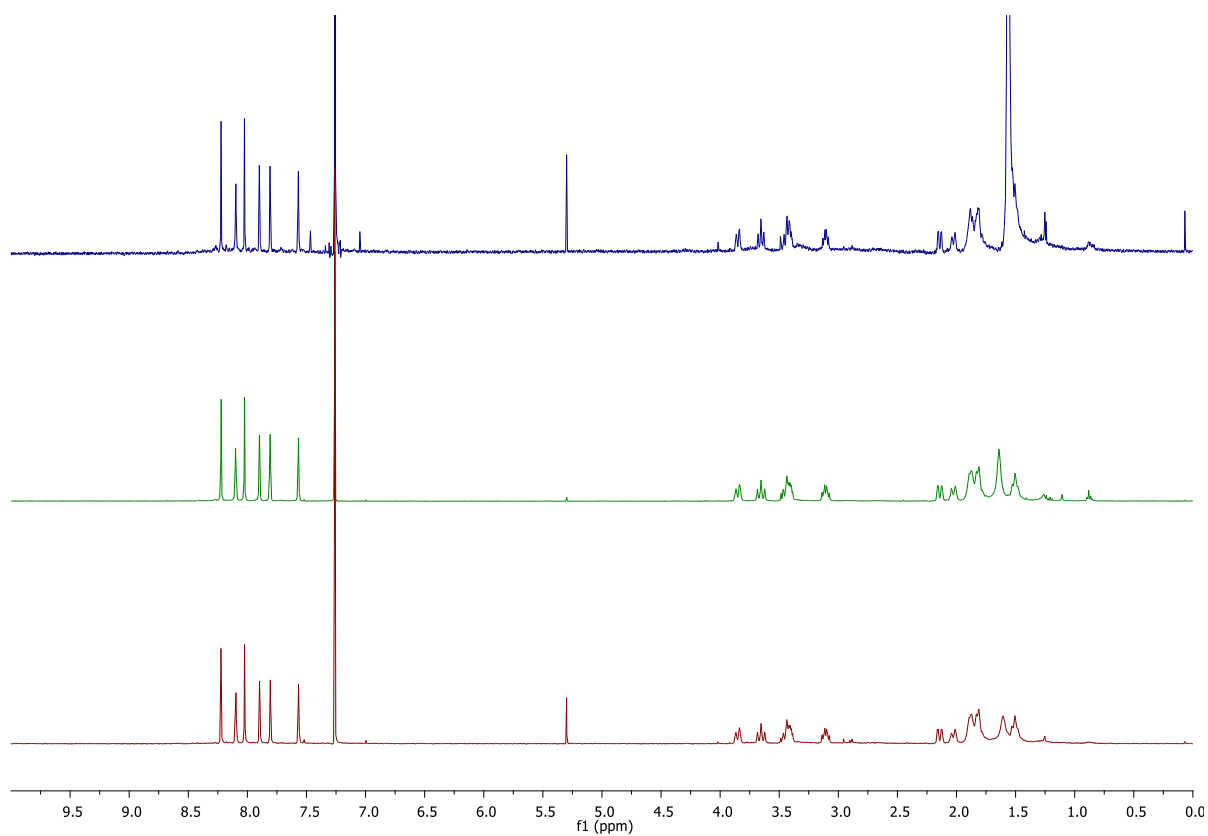


Figure S20: Stacked ^1H NMR spectra (CDCl_3) for purified cage pots **OCP3-R** (top, blue trace), **OCP3-S** (middle, green trace) and **OCP3-rac** (bottom, red trace) after isolation showing identical peak shifts.

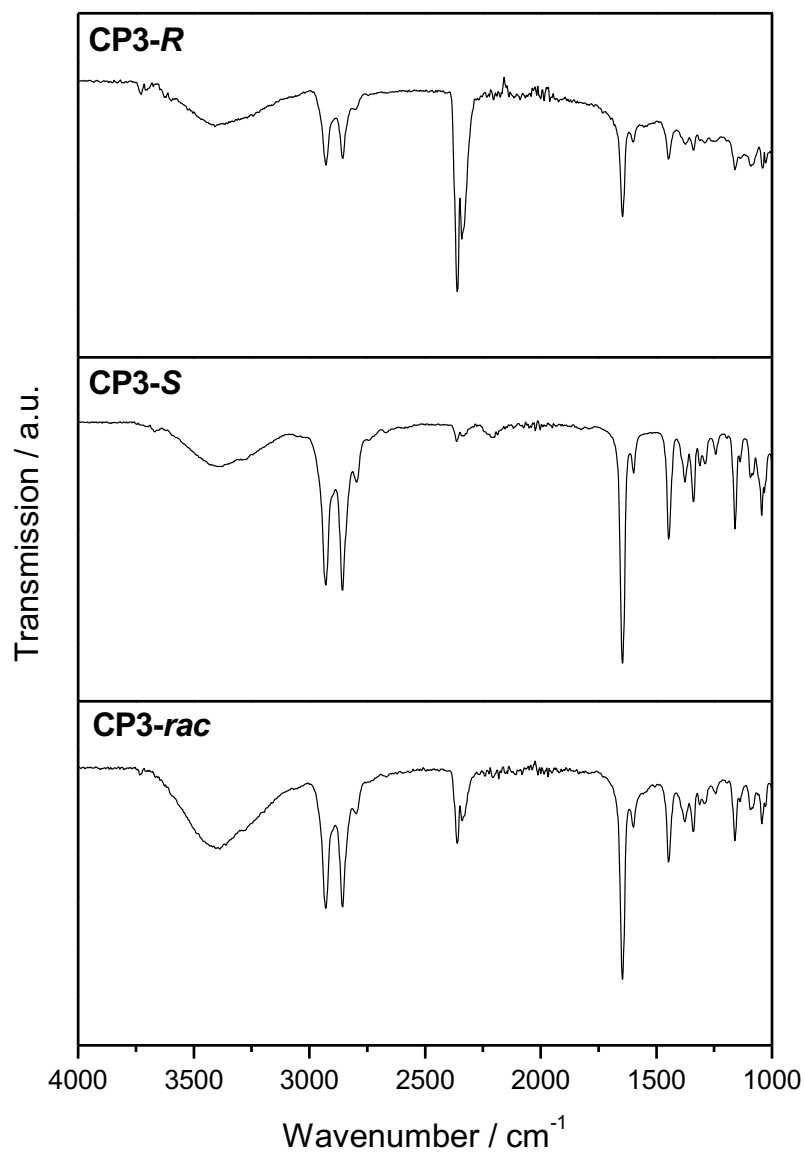


Figure S21: Stacked FTIR spectra of purified cage pots **OCP3-R**, **OCP3-S**, and **OCP3-rac** after isolation.

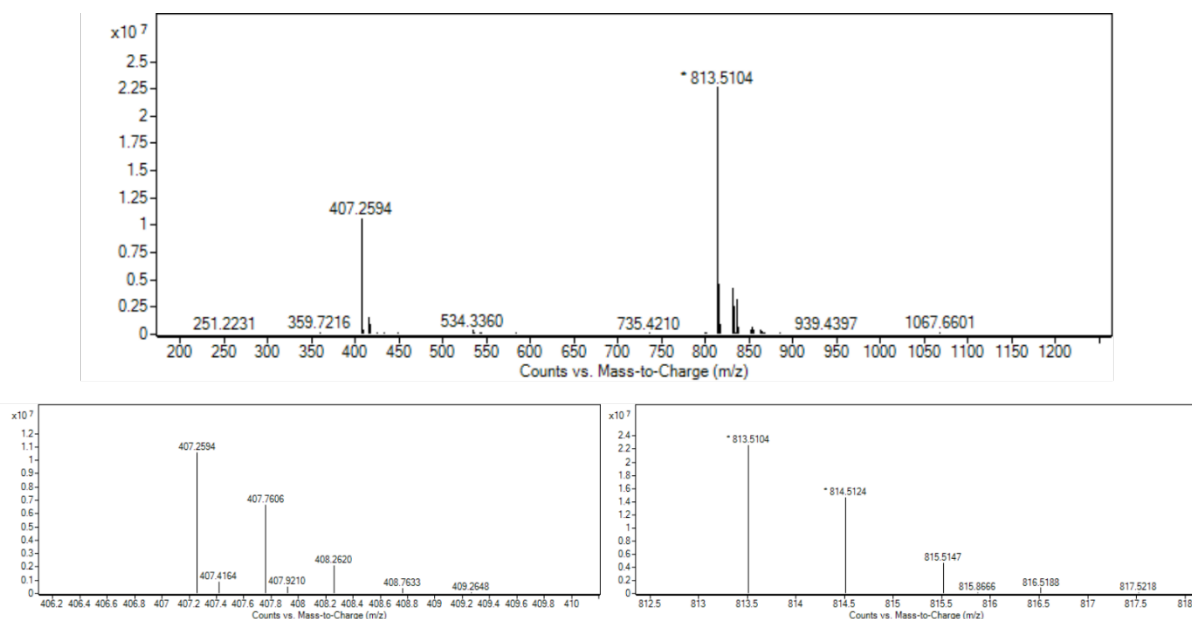


Figure S22: HRMS (ES+) spectra for purified **OCP3-R** — calc. for $C_{51}H_{60}N_{10}$ 812.5002, found $[M+H]^+$ 813.5104, $[M+2H]^{2+}$ 407.2594.

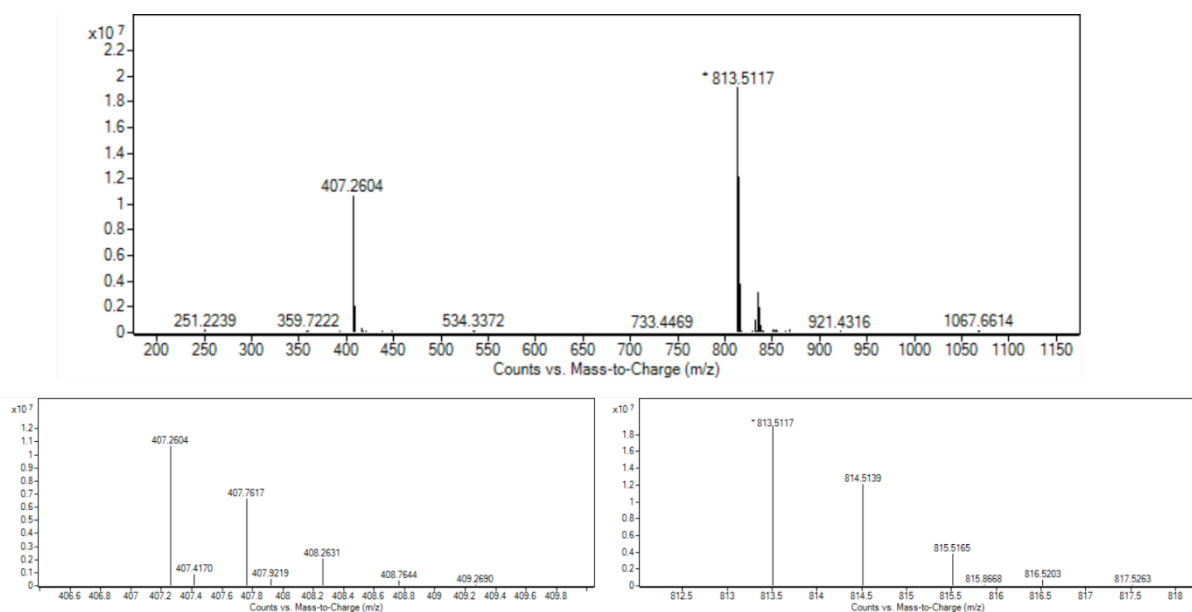


Figure S23: HRMS (ES+) spectra for purified **OCP3-S** — calc. for $C_{51}H_{60}N_{10}$ 812.5002, found $[M+H]^+$ 813.5117, $[M+2H]^{2+}$ 407.2604.

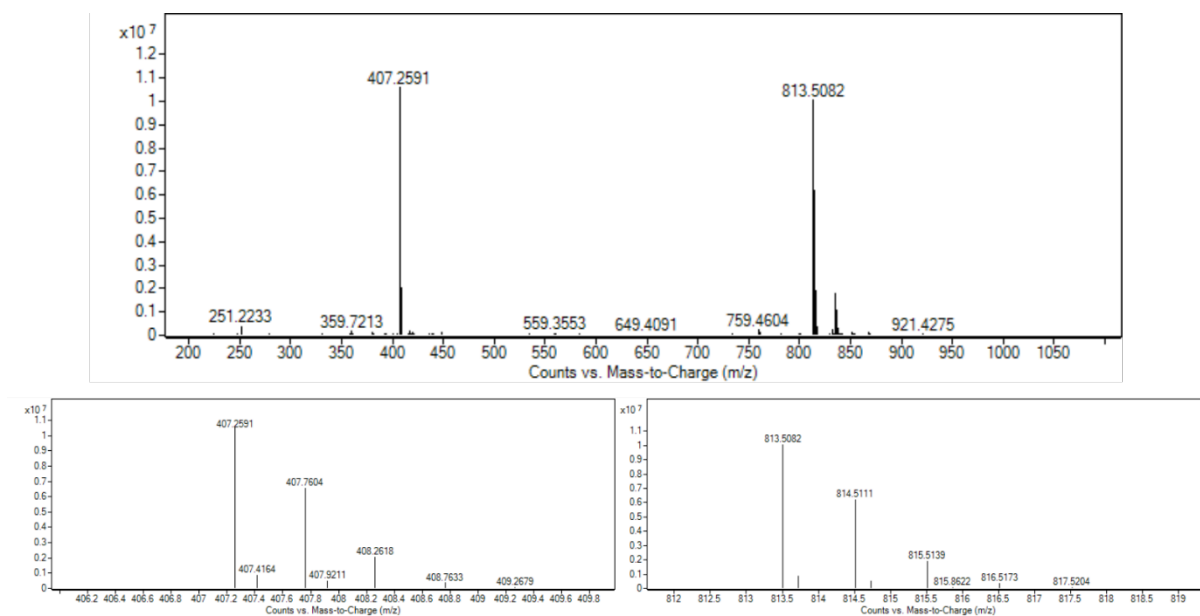


Figure S24: HRMS (ES+) spectra for purified **OCP3-rac** — calc. for $C_{51}H_{60}N_{10}$ 812.5002, found $[M+H]^+$ 813.5082, $[M+2H]^{2+}$ 407.2591.

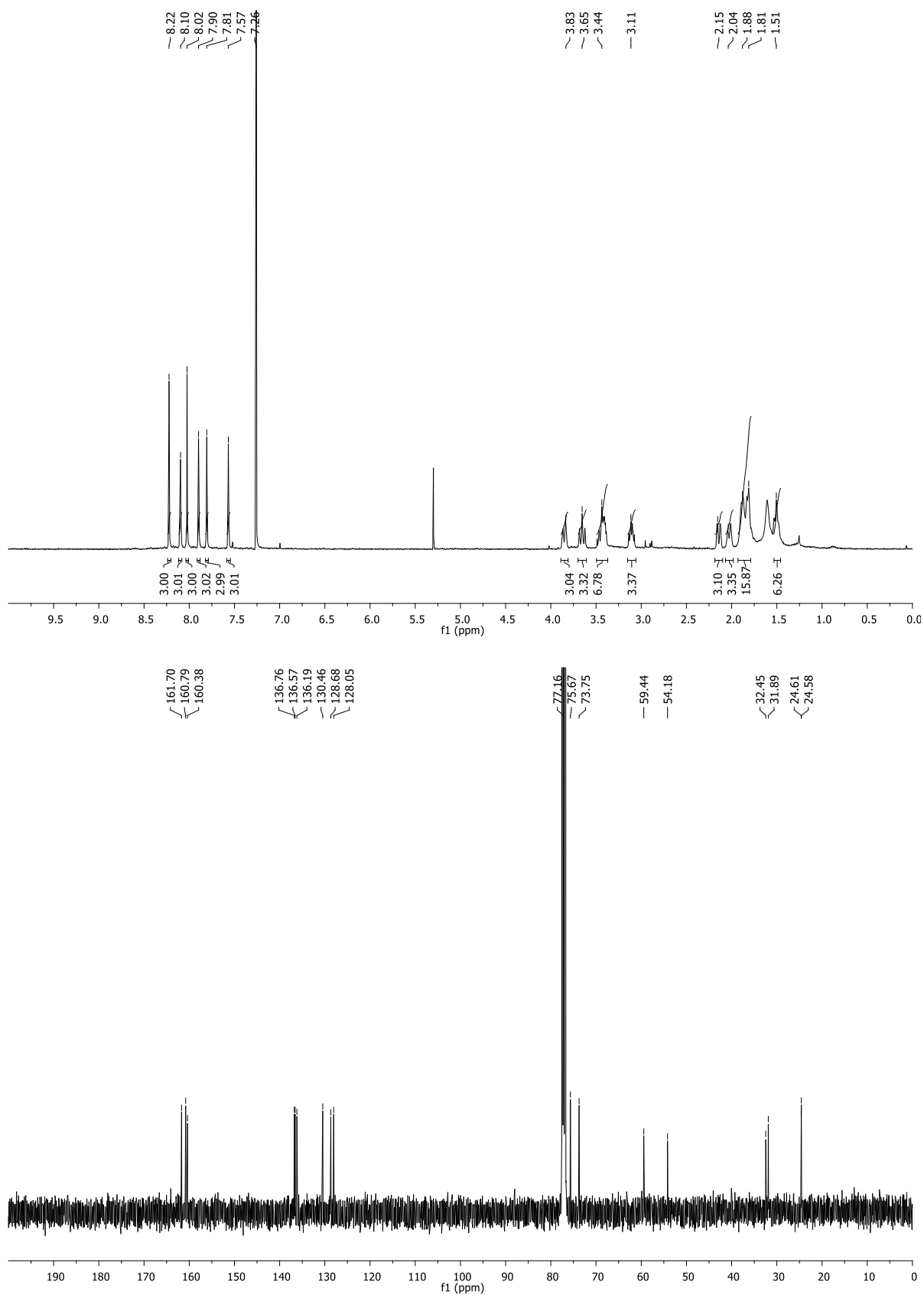


Figure S25: ^1H NMR (CDCl₃; upper) and ^{13}C NMR (CDCl₃; lower) spectra for purified cage pot **OCP3-rac** after isolation.

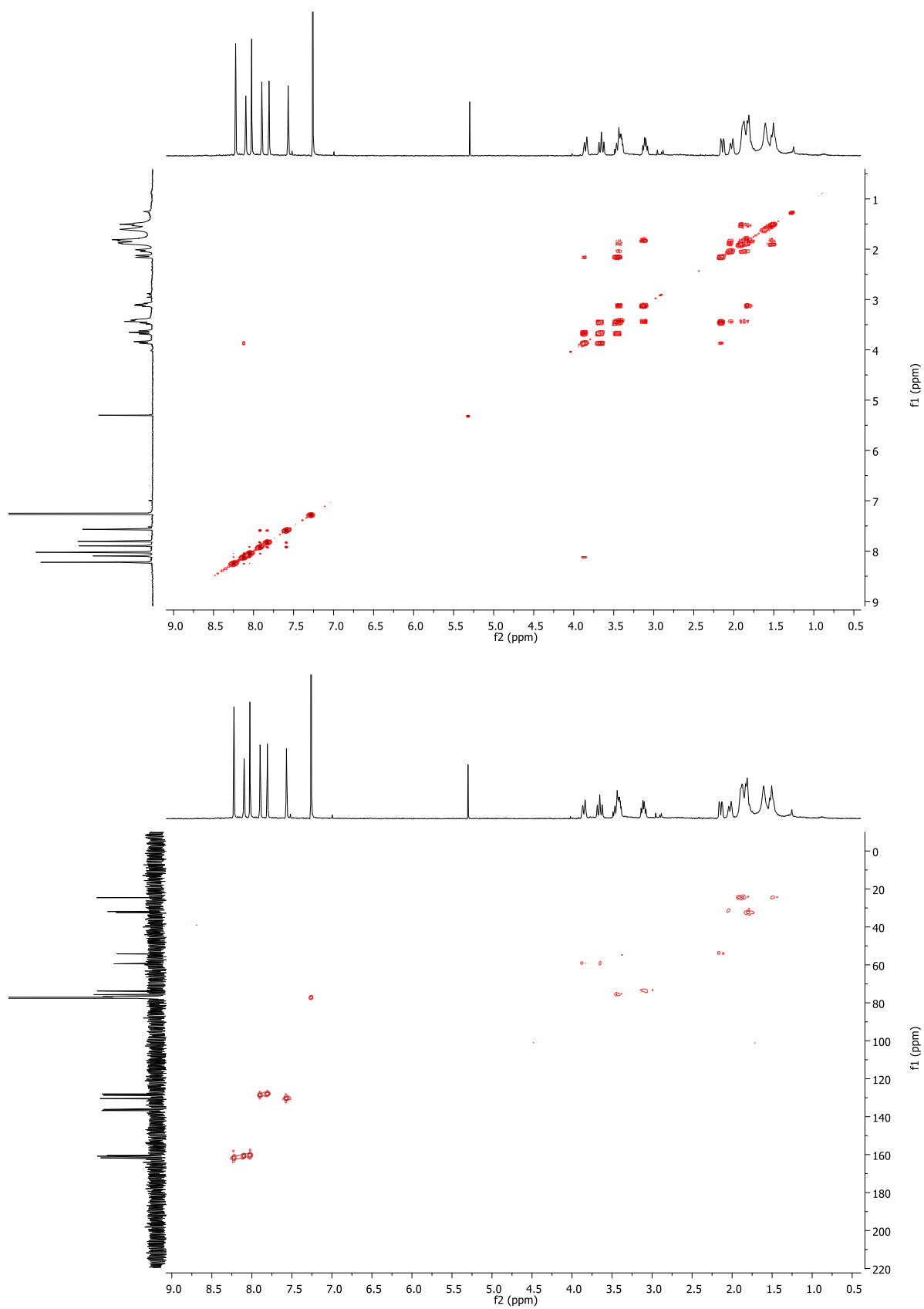


Figure S26: COSY (CDCl_3 ; upper) and HSQC (CDCl_3 ; lower) for purified cage pot **OCP3-rac** after isolation.

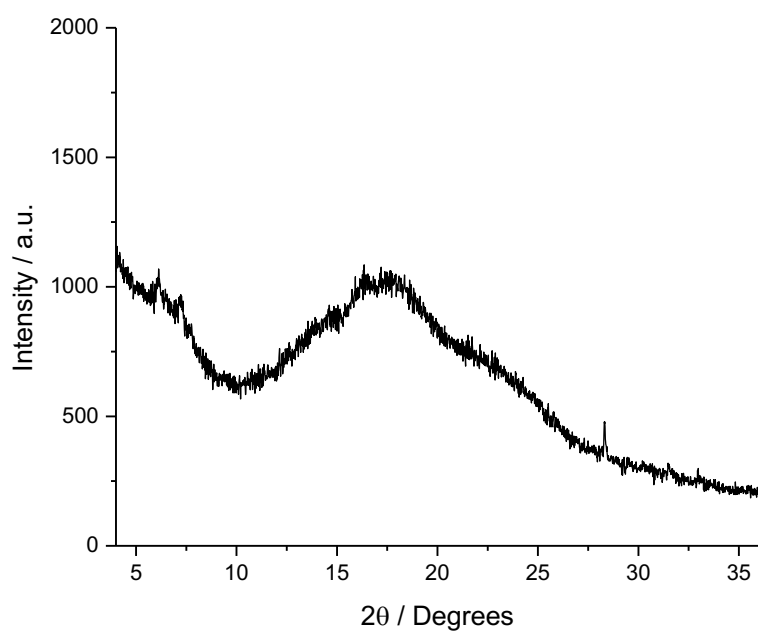


Figure S27: PXRD for purified cage pot **OCP3-rac** after isolation showing the amorphous nature of the isolated bulk solid.

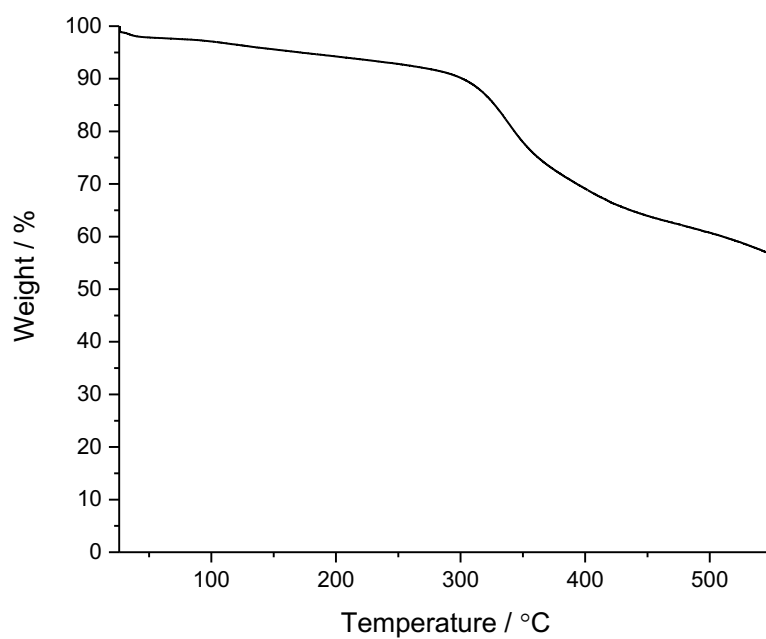


Figure S28: Thermogravimetric (TGA) data for purified cage pot **OCP3-rac** after isolation.

	OCP3-R	OCP3-rac	OCP3-R + OCP3-S
Crystallisation Solvent	MeOH	MeOH	MeOH
Collection Temperature	100 K	150 K	100 K
Formula	C ₅₁ H ₆₀ N ₁₀ , 4(CH ₄ O), 14(H ₂ O)	C ₅₁ H ₆₀ N ₁₀ , 1.5(CH ₄ O), 14.4(H ₂ O)	C ₅₁ H ₆₀ N ₁₀ , 4.25(CH ₄ O), 3.5(H ₂ O)
<i>Mr</i>		1121.40	
Crystal Size (mm)	0.329 x 0.198 x 0.064	0.214 x 0.188 x 0.04	0.321 x 0.105 x 0.07
Crystal System	Hexagonal	Monoclinic	Monoclinic
Space Group	<i>P</i> 6 ₃ 22	<i>C</i> 2/ <i>c</i>	<i>P</i> 2 ₁ / <i>c</i>
<i>a</i> [Å]	16.712(2)	29.543(2)	14.193(2)
<i>b</i> [Å]		16.878(1)	29.912(3)
<i>c</i> [Å]	28.742(3)	26.732(2)	13.706(2)
β [°]		105.675(2)	91.594(3)
<i>V</i> [Å ³]	6952(2)	12833.7(14)	5816(1)
<i>Z</i>	4	8	4
<i>D</i> _{calcd} [g cm ⁻³]	1.129	1.161	1.156
μ [mm ⁻¹]	0.085	0.086	0.078
<i>F</i> (000)	2544	4864	2190
2 θ range [°]	3.15–46.50	2.80–48.21	3.18–52.76
Reflections collected	57055	66637	66147
Independent reflections, <i>R</i> _{int}	3344, 0.0686	10215, 0.0762	11888, 0.0646
Obs. Data [<i>I</i> > 2 σ]	2976	6353	8105
Data / restraints / parameters	3344 / 5 / 251	10215 / 22 / 835	11888 / 7 / 722
Final <i>R</i> 1 values (<i>I</i> > 2 σ (<i>I</i>))	0.0988	0.0732	0.0662
Final <i>R</i> 1 values (all data)	0.1071	0.1227	0.1028
Final <i>wR</i> (<i>F</i> ²) values (all data)	0.2633	0.2439	0.2104
Goodness-of-fit on <i>F</i> ²	1.101	1.039	1.036
Largest difference peak and hole [e.Å ⁻³]	0.52 / -0.29	0.58 / -0.74	0.64 / -0.51
CCDC	1941702	1941703	1941704

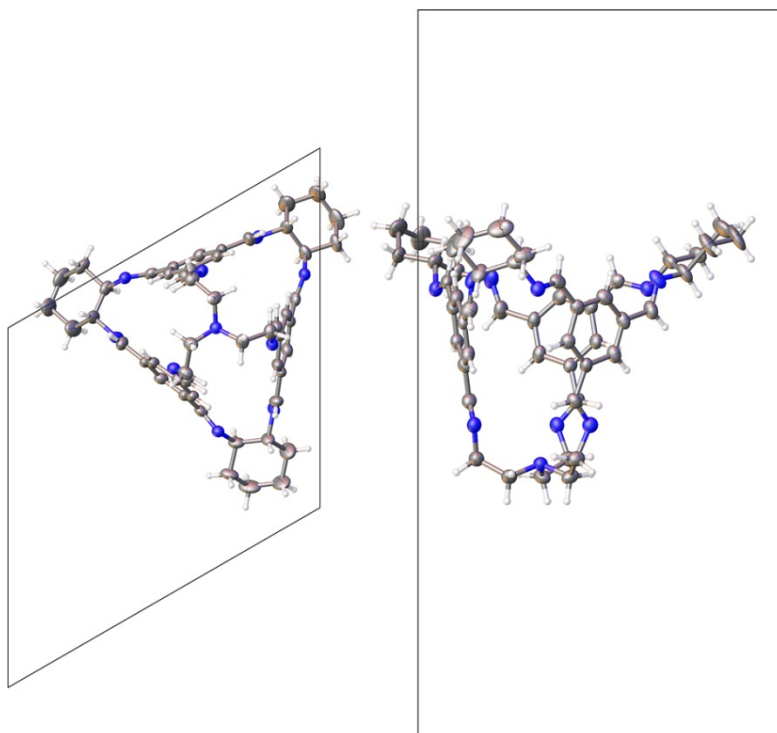


Figure S29: Displacement ellipsoid plot from the solvated single crystal structure, **OCP3-R**, along [100] (left) and [001] (right); **OCP3-R** shown in entirety. Ellipsoids displayed at 50% probability level and disordered solvent omitted for clarity.

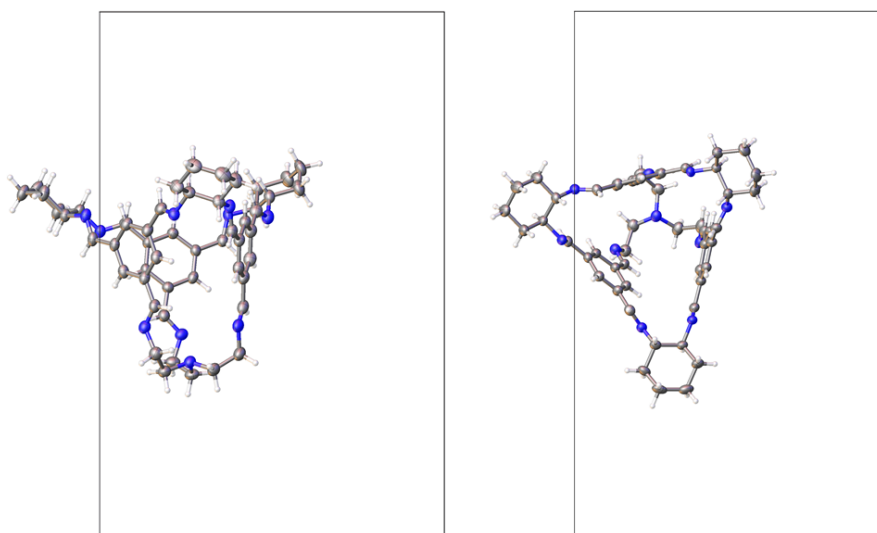


Figure S30: Displacement ellipsoid plot of the asymmetric unit from the solvated single crystal structure, **OCP3-rac** (C2/c), along [100] (left) and [001] (right). Ellipsoids displayed at 50% probability level and disordered solvent omitted for clarity. *NB.* Crystal grown using **OCP3-rac** in MeOH.

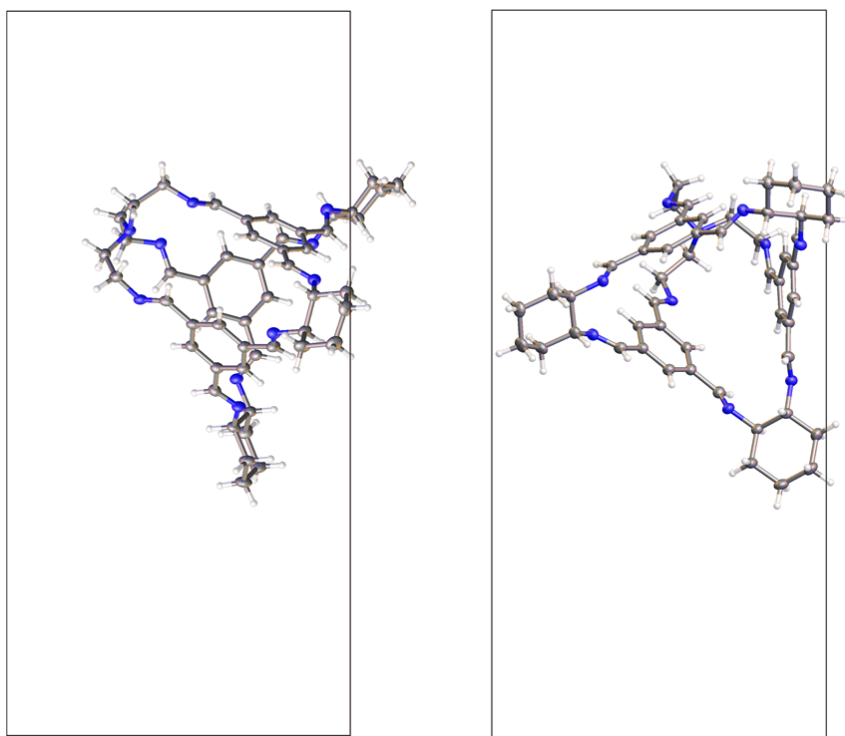


Figure S31: Displacement ellipsoid plot of the asymmetric unit from the solvated single crystal structure, **OCP3-*rac***, along [001] (left) and [100] (right). Ellipsoids displayed at 50% probability level and disordered solvent omitted for clarity. *NB.* Crystal grown using a 1:1 mixture of **OCP3-*R*** and **OCP3-*S*** in MeOH.

5. References

- [1] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, et al., *J. Chem. Theory Comput.* **2016**, *12*, 281–296.
- [2] V. Santolini, M. Miklitz, E. Berardo, K. E. Jelfs, *Nanoscale* **2017**, *9*, 5280–5298.
- [3] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, *Comput. Phys. Commun.* **2005**, *167*, 103–128.
- [4] J. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [5] J. VandeVondele, J. Hutter, *J. Chem. Phys.* **2007**, *127*, DOI 10.1063/1.2770708.
- [6] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*.
- [7] S. Goedecker, M. Teter, *Phys. Rev. B - Condens. Matter Mater. Phys.* **1996**, *54*, 1703–1710.
- [8] R. L. Greenaway, V. Santolini, M. J. Bennison, B. M. Alston, C. Stackhouse, M. A. Little, M. Miklitz, E. G. B. Eden, R. Clowes, A. Shakil, et al., *Nat. Commun.* **2018**, *9*, 2849.
- [9] C. W. Murray, M. L. Verdonk, *J. Comput. Aided. Mol. Des.* **2002**, *16*, 741–753.
- [10] M. I. Page, W. P. Jencks, *Proc. Natl. Acad. Sci.* **1971**, *68*, 1678–1683.
- [11] T. F. Willems, C. H. Rycroft, M. Kazi, J. C. Meza, M. Haranczyk, *Micropor. Mesopor. Mat.* **2012**, *149*, 134–141.
- [12] L. M. Robeson, *J. Membr. Sci.* **1991**, *62*, 165–185.
- [13] M. H. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, and D. J. F. M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J., **2013**.
- [14] D. H. Case, J. E. Campbell, P. J. Bygrave, G. M. Day, *J. Chem. Theory Comput.* **2016**, *12*, 910–924.
- [15] S. L. Price, M. Leslie, G. W. A. Welch, M. Habgood, L. S. Price, P. G. Karamertzanis, G. M. Day, *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478.

- [16] E. O. Pyzer-Knapp, H. P. G. Thompson, G. M. Day, *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, 72, 477–487.
- [17] A. J. Stone, M. Alderton, *Mol. Phys.* **1985**, 56, 1047–1064.
- [18] J. A. Chisholm, S. Motherwell, *J. Appl. Crystallogr.* **2005**, 38, 228–231.
- [19] C. F. Macrae, P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor, M. Towler, J. van de Streek, *J. Appl. Crystallogr.* **2006**, 39, 453–457.
- [20] S. Parsons, **2004**.
- [21] G. Sheldrick, **2008**.
- [22] L. Krause, R. Herbst-Irmer, G. M. Sheldrick, D. Stalke, *J. Appl. Crystallogr.* **2015**, 48, 3–10.
- [23] G. M. Sheldrick, *Acta Crystallogr. Sect. C Struct. Chem.* **2015**, 71, 3–8.
- [24] G. M. Sheldrick, *Acta Crystallogr. Sect. A Found. Crystallogr.* **2008**, 64, 112–122.
- [25] O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard, H. Puschmann, *J. Appl. Crystallogr.* **2009**, 42, 339–341.