Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2018

Supporting Information for

Quantified Structural Speciation in Self-sorted $Co^{II}_6L_4$ Cage Systems

Felix. J. Rizzuto, Marion Kieffer, Jonathan R. Nitschke

Table of Contents

1. Materials and methods

All reagents were purchased from commercial sources and used as received without further purification. 2-Formyl-1,10-phenanthroline,[1] subcomponent **D**,[2] $Co(OTf)_2$[3] and $Co(NTf_2)_2 \cdot 6H_2O$[3] were prepared by literature procedures.

1.1 Nuclear Magnetic Resonance (NMR)

NMR spectra were recorded using a 400 MHz Avance III HD Smart Probe NMR spectrometer. Chemical shifts for ^1H and ^{19}F are reported in ppm on the δ scale; ^1H NMR spectra are referenced to the residual solvent peak, ^{19}F NMR spectra are referenced against C_6F_6 at –164.9 ppm. Wide sweep paramagnetic NMR spectra were recorded in the analogue digitisation mode with a spectral width (SW) of 372.98 ppm, a transmitter frequency offset (O1P) of 100.00 ppm and the line width set to 10.0 Hz. Due to the experimental difficulties associated with collecting NMR data for ^1H nuclei with vastly different relaxation times, differences between the measured and theoretical integration values were in some cases observed. While the paramagnetic nature of the complex precluded complete assignment of the proton environments, we propose that through-bond proximity of the proton environment to each Co^{II} centres dictates the extent of downfield shifting of each signal, akin to previous reports.[3,4]

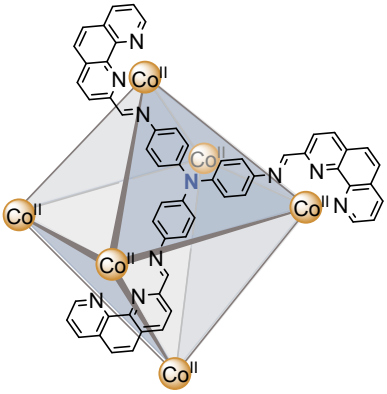
1.2 Mass spectrometry (MS)

Low resolution electrospray ionisation (LR-ESI) mass spectrometry was undertaken on a Micromass Quattro LC mass spectrometer (cone voltage 10-30 eV; desolvation temp. 313 K; ionization temp. 313 K) infused from a Harvard syringe pump at a rate of 10 μL min^{-1}. High resolution electrospray ionisation mass spectrometry (HRMS-ESI) was performed on a Waters LCT Premier Mass Spectrometer featuring a Z spray source with electrospray ionisation and modular LockSpray interface. Travelling Wave Ion Mobility Quadrupole Time-of-Flight mass spectra were collected on a Waters Vion IMS QTof Mass Spectrometer equipped with XS Ion Optics and the QuanTof2 detection system.

2. Cage synthesis and characterisation

All reactions were conducted under benchtop conditions in either an NMR tube or small vial. Anoxic conditions were not observed to alter the spectra collected for any species; no oxidation from Co^{II} to Co^{III} was observed for any species described herein, neither during nor after assembly.

2.1 Synthesis and characterisation of 1



Triamine **A** (1.02 mg, 3.52×10^{-6} mol, 4 equiv), 2-formylphenanthroline (2.20 mg, 1.06×10^{-5} mol, 12 equiv) and $Co(OTf)_2$ (1.89 mg, 5.28×10^{-6} mol, 6 equiv) were combined in CD_3CN (0.5 mL) and heated at 70 °C overnight. Spectra were collected after cooling to room temperature.

^{1}H NMR (400 MHz, 298 K, CD_3CN): δ 279.2, 216.5, 124.9, 114.3, 113.7, 50.9, 34.4, 15.1, –16.2 ppm.

LR-ESI-MS [charge, calculated for $Co_6(C_{57}H_{36}N_{10})_4(CF_3SO_3)_{12}$]: *m/z* = 1247.3 [**1**$(OTf)_8^{4+}$, 1247.5], 968.1 [**1**$(OTf)_7^{5+}$, 968.2], 781.9 [**1**$(OTf)_6^{6+}$, 782.0], 648.9 [**1**$(OTf)_5^{7+}$, 649.0], 549.1 [**1**$(OTf)_4^{8+}$, 549.2].

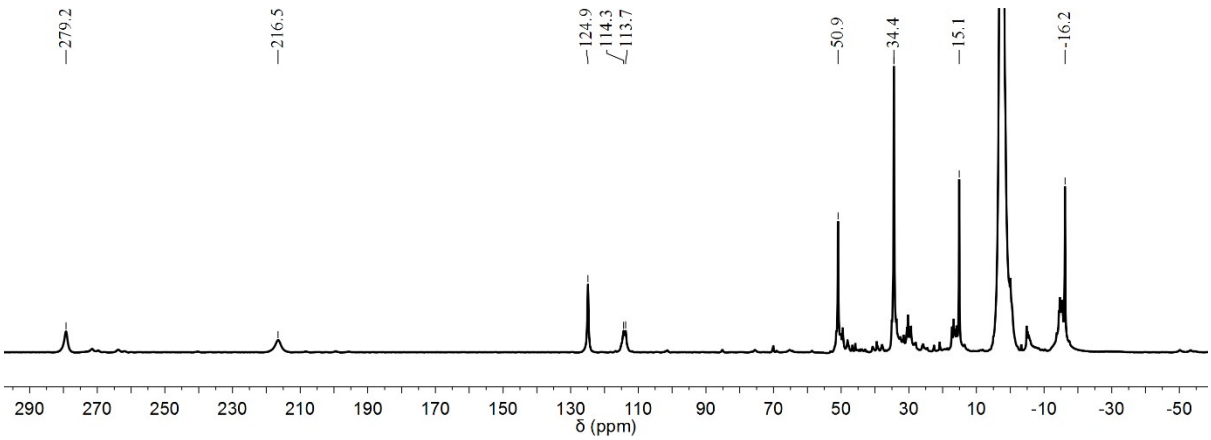


Figure S1. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of **1**$(OTf)_{12}$.

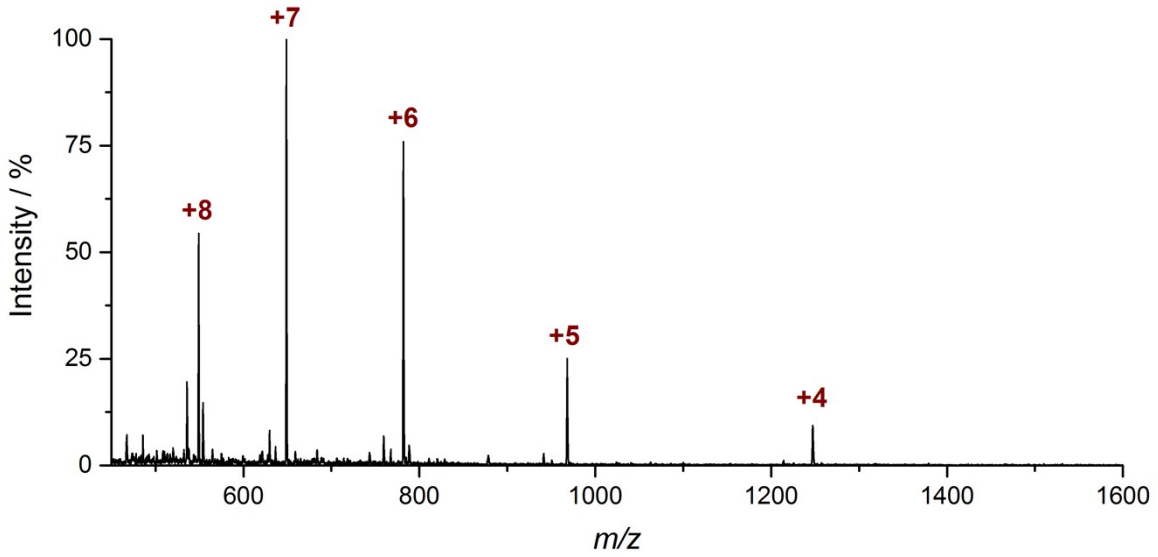


Figure S2. Low resolution ESI mass spectrum of **1**$(OTf)_{12}$.

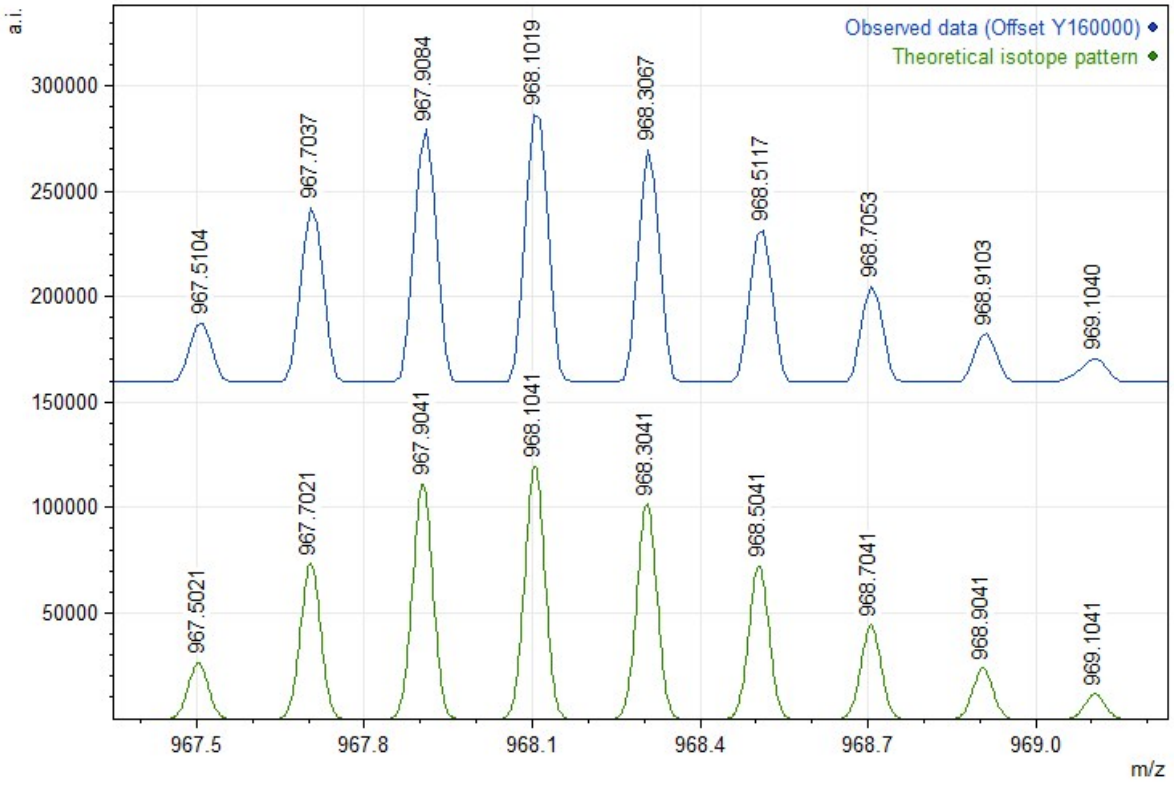
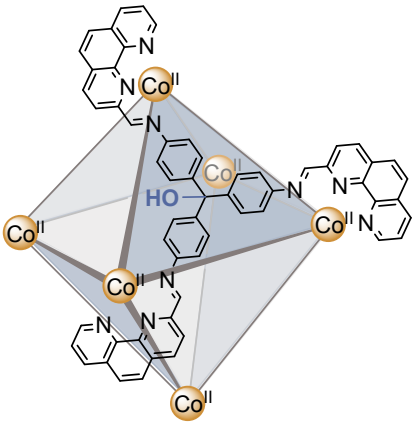


Figure S3. High resolution ESI mass spectrum of **1**$(OTf)_{12}$ showing the observed z = +5 charge (top), compared to the theoretical isotope pattern (bottom).

2.2 Synthesis and characterisation of 2



Triamine **B** (1.47 mg, 4.80 × 10^{-6} mol, 4 equiv), 2-formylphenanthroline (3.00 mg, 1.44 × 10^{-5} mol, 12 equiv) and $Co(NTf_2)_2 \cdot 6H_2O$ (5.26 mg, 7.20 × 10^{-6} mol, 6 equiv) were combined in CD_3CN (0.5 mL) and heated at 70 °C overnight. Spectra were recorded upon cooling to room temperature.

^{1}H NMR (400 MHz, 298 K, CD_3CN): δ 215.5, 146.0, 87.5, 42.2, 32.8, 30.4, 25.4, 4.2, 2.8, –6.5, –34.4 ppm.

LR-ESI-MS [charge, calculated for $Co_6(C_{58}H_{37}N_9O)_4(NC_2F_6S_2O_4)_{12}$]: *m/z* = 1163.7 [**2**$(NTf_2)_7^{5+}$, 1163.7], 922.9 [**2**$(NTf_2)_6^{6+}$, 923.1], 751.0 [**2**$(NTf_2)_5^{7+}$, 751.2], 622.1 [**2**$(NTf_2)_4^{8+}$, 622.3].

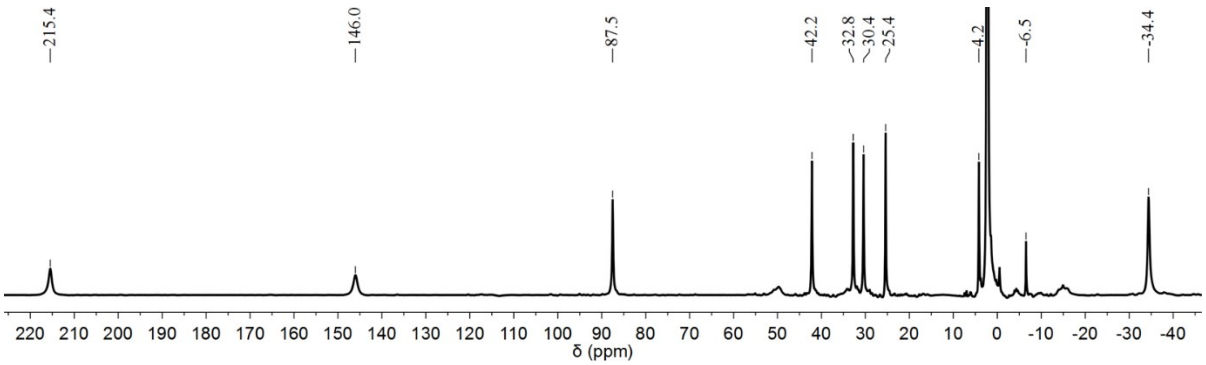


Figure S4. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of **2**$(NTf_2)_{12}$.

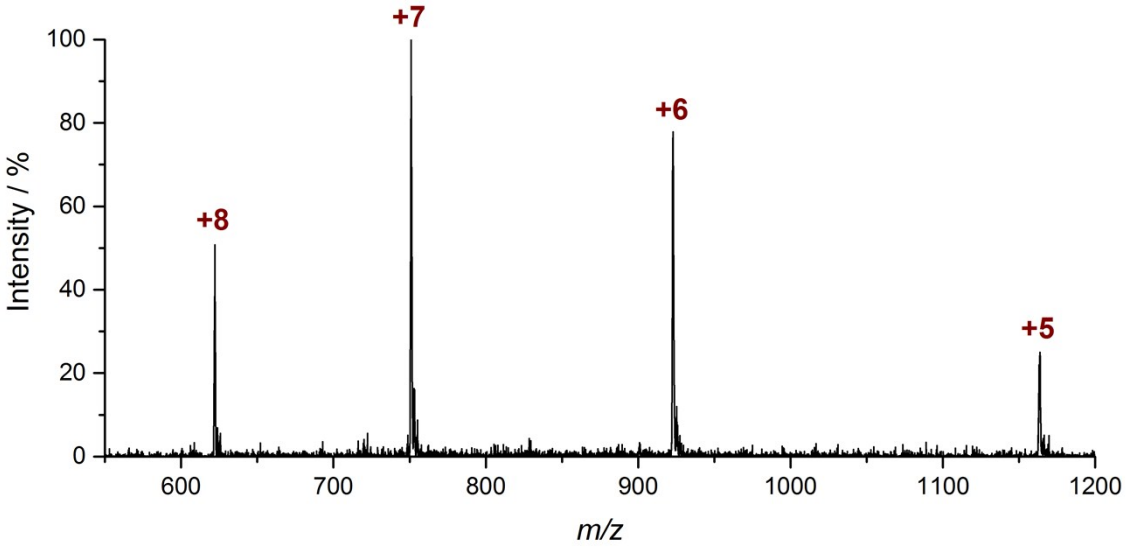


Figure S5. Low resolution ESI mass spectrum of **2**$(NTf_2)_{12}$.

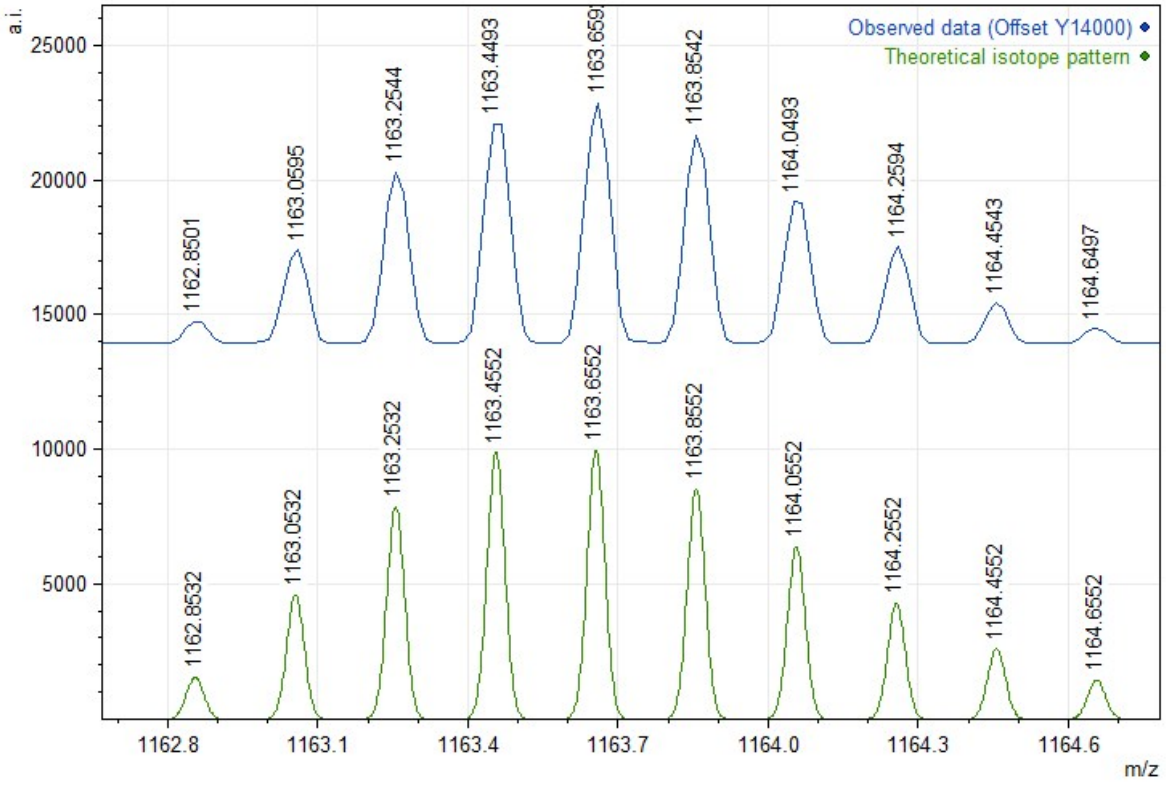
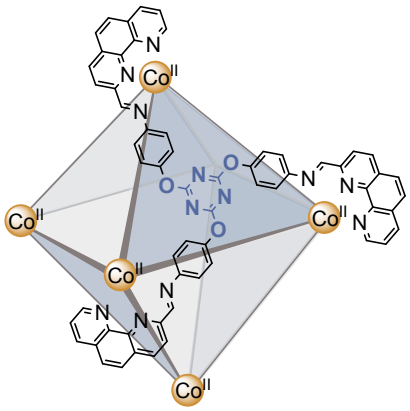


Figure S6. High resolution ESI mass spectrum of **2**$(NTf_2)_{12}$ showing the observed z = +5 charge (top), compared to the theoretical isotope pattern (bottom).

2.3 Synthesis and characterisation of 3



Triamine **C** (1.93 mg, 4.80 × 10^{-6} mol, 4 equiv), 2-formylphenanthroline (3.00 mg, 1.44 × 10^{-5} mol, 12 equiv) and $Co(OTf)_2$ (2.57 mg, 7.20 × 10^{-6} mol, 6 equiv) were combined in CD_3CN (0.5 mL) and heated at 70 °C overnight. Spectra were recorded after cooling to room temperature.

^{1}H NMR (400 MHz, 298 K, CD_3CN): δ 264.7, 200.5, 113.4, 49.6, 35.9, 30.2, 18.7, –0.6, –13.9 ppm.

LR-ESI-MS [charge, calculated for $Co_6(C_{60}H_{36}N_{12}O_3)_4(CF_3SO_3)_{12}$]: *m/z* = 1057.3 [**3**$(OTf)_7^{5+}$, 1057.8], 856.4 [**3**$(OTf)_6^{6+}$, 856.7], 712.8 [**3**$(OTf)_5^{7+}$, 713.0], 605.1 [**3**$(OTf)_4^{8+}$, 605.2].

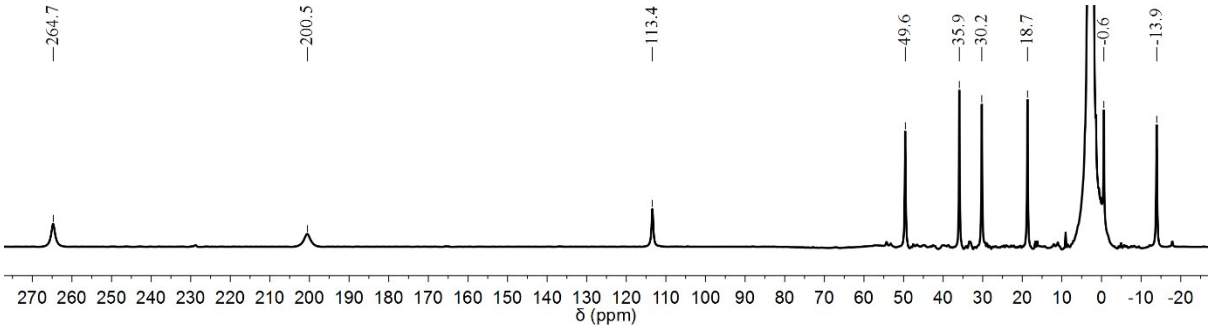


Figure S7. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of **3**$(OTf)_{12}$.

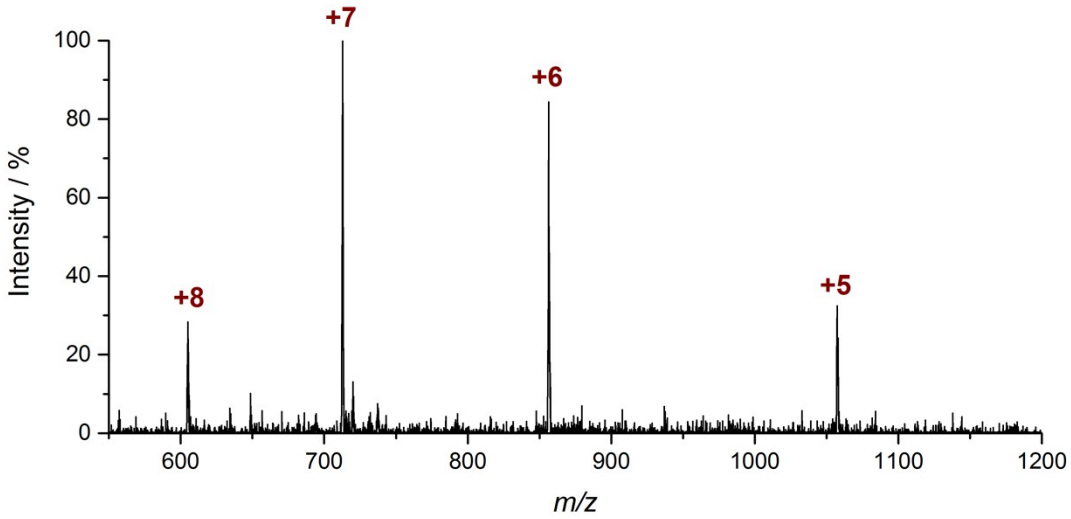


Figure S8. Low resolution ESI mass spectrum of **3**$(OTf)_{12}$.

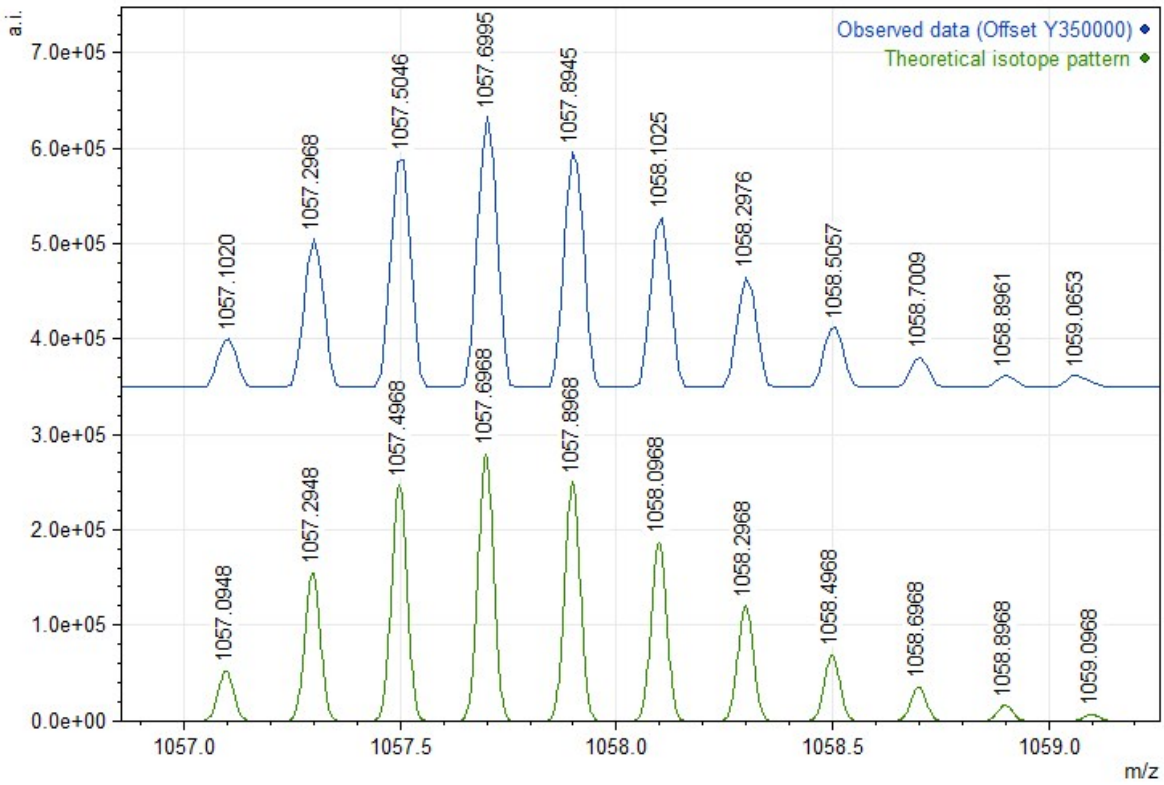
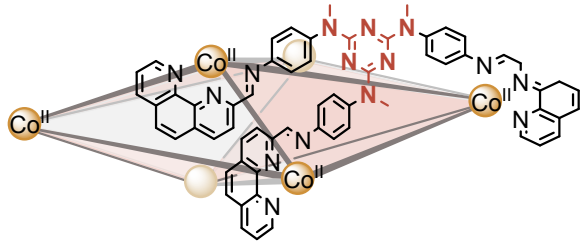


Figure S9. High resolution ESI mass spectrum of **3**$(OTf)_{12}$ showing the observed z = +5 charge (top), compared to the theoretical isotope pattern (bottom).

2.4 Synthesis and characterisation of 4



Triamine **D** (1.41 mg, 3.20 × 10^{-6} mol, 4 equiv), 2-formylphenanthroline (2.00 mg, 9.61 × 10^{-6} mol, 12 equiv) and $Co(OTf)_2$ (1.71 mg, 7.20 × 10^{-6} mol, 6 equiv) were combined in CD_3CN and heated at 70 °C overnight. Upon cooling, Et_2O (10 mL) was added and the solution cooled in a fridge for 24 h. The suspension was centrifuged, the supernatant decanted and the solid dried *in vacuo* to yield **4** as an orange powder.

^{1}H NMR (400 MHz, 298 K, CD_3CN): δ 265.1, 200.0, 186.3, 166.0, 118.6, 112.1, 91.0, 65.1, 59.6, 49.4, 39.8, 37.3, 36.1, 35.9, 33.7, 30.6, 27.8, 27.6, 25.8, 18.7, 18.1, 17.0, 12.0, 7.6, 7.3, 5.5, 4.5, –0.4, –1.5, –9.2, –11.1, –37.8, –60.5 ppm.

^{19}F NMR (376 MHz, 298 K, CD_3CN): δ –78.1 (OTf^- in fast exchange with the cavity) ppm.

LR-ESI-MS [charge, calculated for $Co_6(C_{57}H_{36}N_{10})_4(CF_3SO_3)_{12}$]: *m/z* = 1398.5 [**4**$(OTf)_8^{4+}$, 1398.7], 1089.0 [**4**$(OTf)_7^{5+}$, 1089.1], 882.7 [**4**$(OTf)_6^{6+}$, 882.8], 735.4 [**4**$(OTf)_5^{7+}$, 735.4].

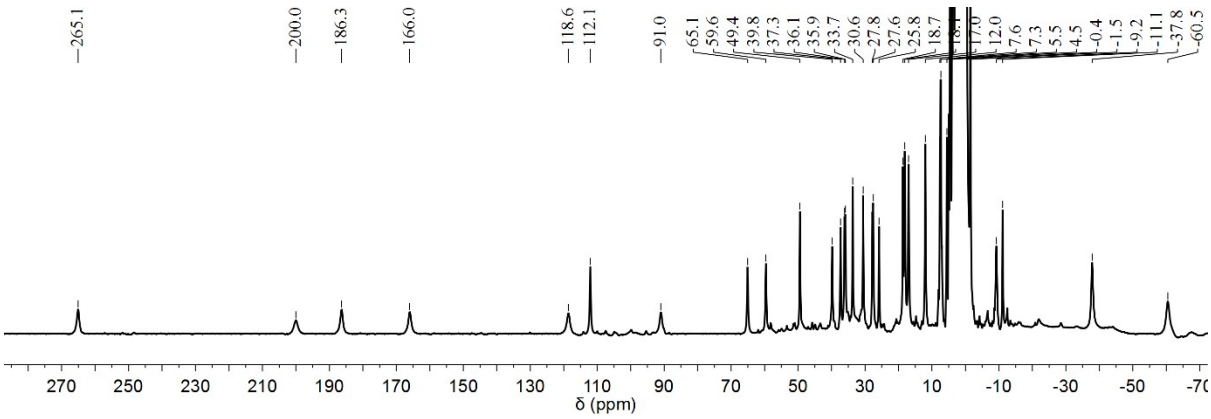


Figure S10. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of **4**$(OTf)_{12}$.

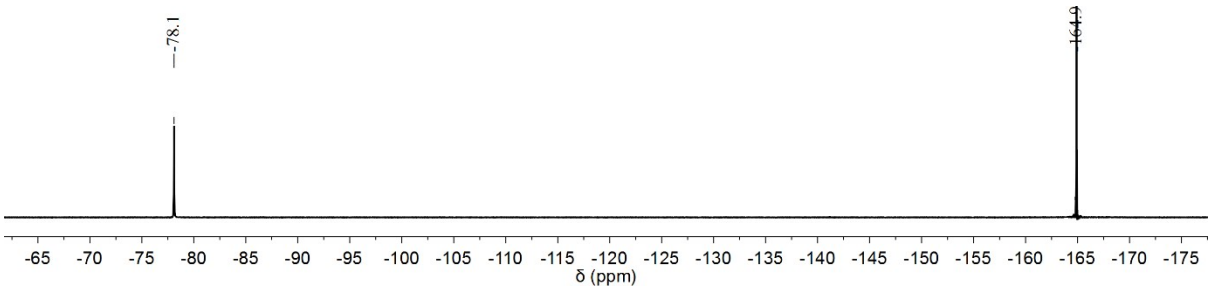


Figure S11. ^{19}F NMR spectrum (376 MHz, 298 K, CD_3CN) of **4**$(OTf)_{12}$.

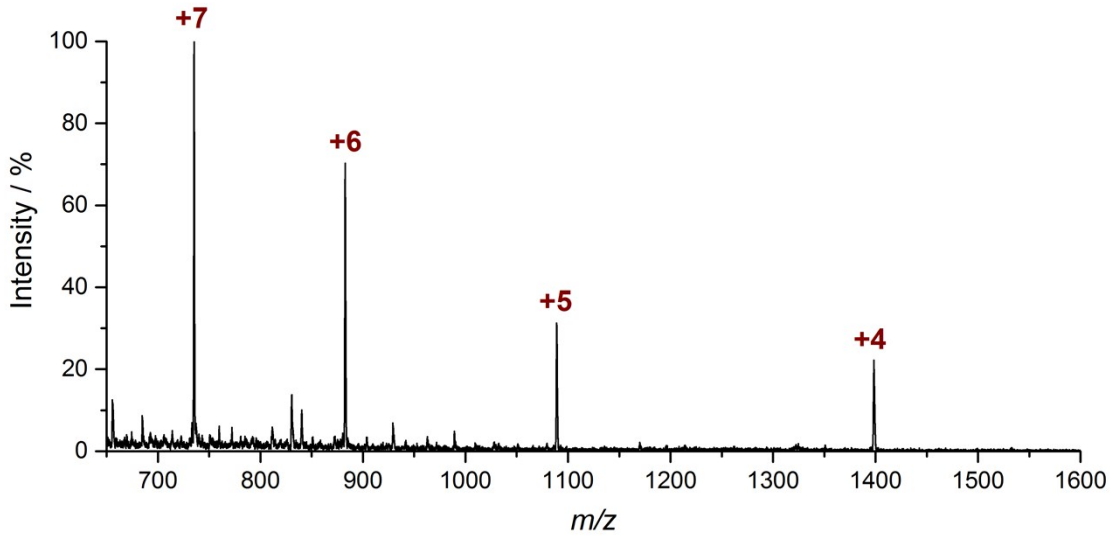


Figure S12. Low resolution ESI mass spectrum of **4**$(OTf)_{12}$.

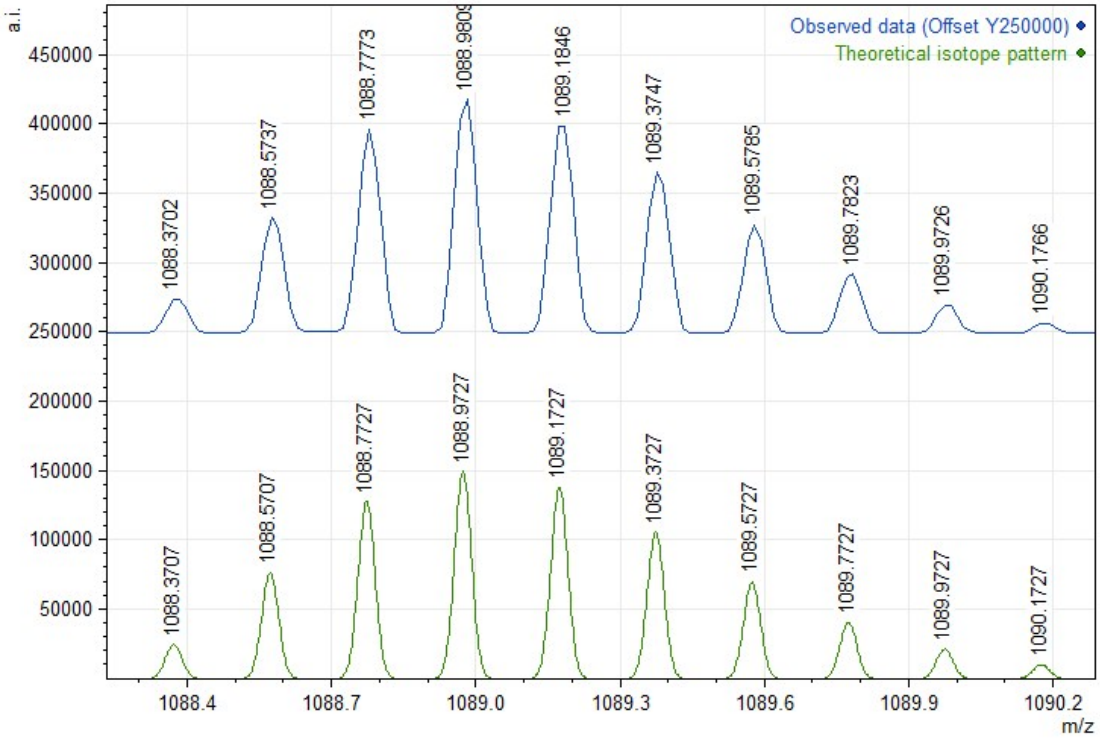


Figure S13. High resolution ESI mass spectrum of **4**$(OTf)_{12}$ showing the observed z = +5 charge (top), compared to the theoretical isotope pattern (bottom).

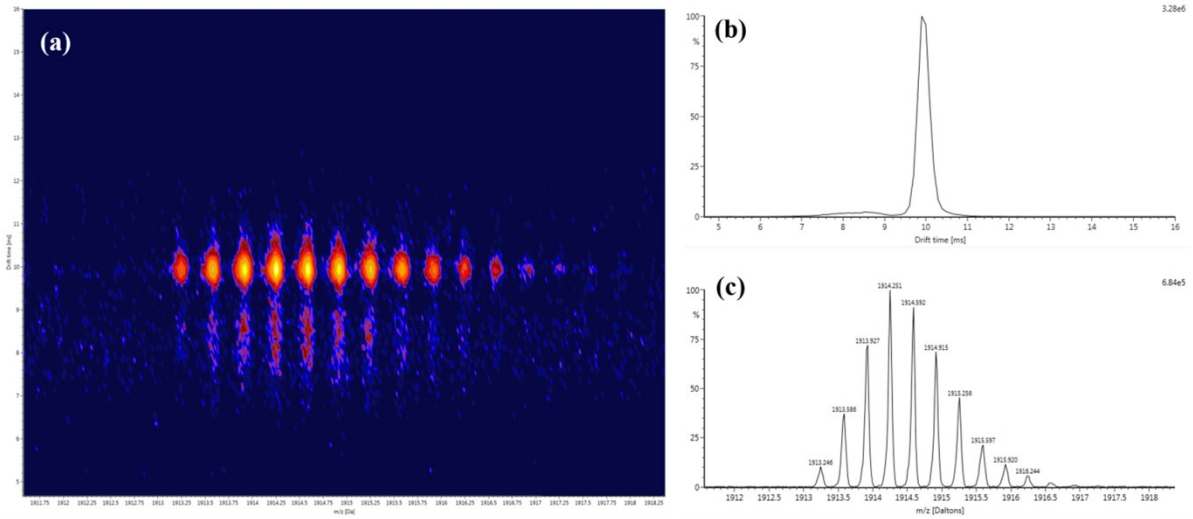


Figure S14. High resolution ion mobility mass spectra of the z = +4 charge fragment of **4**$(OTf)_{12}$. (a) Ion mobility and (b) drift time spectra showing one major species (10 ms). (c) High resolution ESI mass spectrum of the fragment.

3. Crystallography

Data were collected at Beamline I19 of Diamond Light Source employing silicon double crystal monochromated synchrotron radiation (0.6889 Å) with ω scans at 100(2) K. Data integration and reduction for **1** were undertaken with Crysalis PRO;[5] data integration and reduction for **4** were undertaken with xia2.[6,7] Multi-scan empirical absorption corrections were applied to the data using using SADABS[8] for **1** and xia2[6,7] for **4**. Subsequent computations were carried out using the WinGX-32[9] graphical user interface. Structures were solved by direct methods using SHELXT-2013[10] then refined and extended with SHELXL-2013.[10] In general, non-hydrogen atoms with occupancies greater than 0.5 were refined anisotropically. Carbon-bound hydrogen atoms were included in idealised positions and refined using a riding model. Disorder was modelled using standard crystallographic methods including constraints, restraints and rigid bodies where necessary. In all cases, molecular formulas were determined firstly from the required number of charge-balancing anions, and then confirmed from the number of electrons identified in the disordered portion of the crystal by SQUEEZE.[11] The amount of solvent quoted in each formula is only that which could be assigned directly from the electron density map. Crystallographic data along with specific details pertaining to the refinements follow. Crystallographic data have been deposited with the CCDC (CCDC 1568164 and 1568165 for **1** and **4**, respectively).

3.1 Crystal structure of $1 \cdot 5BF_4 \cdot 7OTf \cdot 7MeCN \cdot 0.25iPr_2O$

Formula $C_{250.50}H_{168.50}B_5Co_6F_{41}N_{47}O_{21.25}S_7$, M 5587.86, Triclinic, P□ (#2), a 21.2508(3), b 21.6182(2), c 39.5929(4) Å, α 93.2080(10), β 101.6240(10), γ 118.5210(10)°, V 15407.3(3) Å^3, D_c 1.204 g cm^{-3}, Z 2, crystal size 0.130 by 0.080 by 0.040 mm, colour purple, habit block, temperature 100(2) Kelvin, λ(synchrotron) 0.6889 Å, μ(synchrotron) 0.409 mm^{-1}, $T(\text{SADABS})_{\text{min,max}}$ 0.5776, 0.7446, $2\theta_{\text{max}}$ 42.62, *hkl* range –22 22, –22 22, –41 41, N 142112, N_{ind} 37722(R_{merge} 0.0617), N_{obs} 30531(I > 2σ(I)), N_{var} 3468, residuals* $R1(F)$ 0.1132, $wR2(F^2)$ 0.2782, GoF(all) 1.058, $\Delta\rho_{\text{min,max}}$ –0.756, 1.568 e$^-$ Å^{-3}.

*$R1 = \Sigma||F_o| - |F_c||/\Sigma|F_o|$ for $F_o > 2\sigma(F_o)$; $wR2 = (\Sigma w(F_o^2 - F_c^2)^2/\Sigma(wF_c^2)^2)^{1/2}$ all reflections

$w=1/[\sigma^2(F_o^2)+(0.1000P)^2+25.0000P]$ where $P=(F_o^2+2F_c^2)/3$ $w=1-\exp[-20.0000(\sin\theta/\lambda)^2)]$

Specific refinement details

Crystals of **1·5BF$_4$·7OTf·7MeCN·0.25*i*Pr$_2$O** were grown by slow diffusion of diisopropyl ether into a CD_3CN solution of **1**$(OTf)_{12}$ containing excess *n*$BuNBF_4$. The crystals employed rapidly lost solvent after removal from the mother liquor and rapid handling prior to flash cooling in the cryostream was required to collect data. Despite these measures and the use of high intensity synchrotron radiation, few reflections at greater than 0.95 Å resolution were observed. Nevertheless, the quality of the data is more than sufficient to establish the connectivity of the structure. Due to the less than ideal resolution, bond lengths and angles within pairs of organic ligands were restrained to be similar to each other (SAME) and thermal parameter restraints (SIMU, DELU) were applied to all non-metal atoms to facilitate anisotropic refinement. Ligand-based atoms that still displayed thermal parameters greater than 0.5 were further refined to approximate isotropic behaviour (ISOR).

All four triflate anions and two tetrafluoroborate anions were modelled as disordered over two sites. On two triflate anions, EADP restrains were applied to the disordered parts to facilitate anisotropic refinement. A small portion of disorder (<10%) was identified on a phenanthroline ring (C340-C352) that could not be satisfactorily modelled despite numerous attempts. All solvent and anion molecules displayed a high degree of thermal motion; these molecules were modelled to approximate isotropic behaviour (ISOR). The diisopropyl ether molecule showed a high degree of thermal motion and was modelled with isotropic thermal parameters.

The remaining triflate anions present in the asymmetric unit could not be successfully resolved despite numerous attempts at modelling, including the use of rigid bodies. Consequently, the SQUEEZE[11] function of PLATON[12] was employed to remove the contribution of the electron density associated with the remaining anions and further highly disordered solvent molecules.

3.2 Crystal structure of 4·5.5$CB_{11}H_{12}$·6.5OTf·2.5MeCN·*i*Pr$_2$O

Formula $C_{273}H_{263.50}B_{60.50}Co_6F_{19.50}N_{62.50}O_{20.50}S_{6.50}$, *M* 6334.42, Triclinic, P□□ (#2), *a* 25.0303(4), *b* 26.7592(5), *c* 34.5514(7) Å, *α* 68.1130(17), *β* 88.8092(17), *γ* 66.5154(14)°, *V* 19472.2(7) Å^3, D_c 1.080 g cm^{-3}, *Z* 2, crystal size 0.040 by 0.015 by 0.010 mm, colour orange, habit prism, temperature 100(2) Kelvin, *λ*(synchrotron) 0.6889 Å, *μ*(synchrotron) 0.320 mm^{-}

[1], T(xia2)$_{min,max}$ 0.9934, 1.0108, 2θ_{max} 33.36, *hkl* range –20 20, –22 21, –28 28, N 38843, N_{ind} 22663(R_{merge} 0.0412), N_{obs} 15201(I > 2σ(I)), N_{var} 4178, residuals* $R1(F)$ 0.1179, $wR2(F^2)$ 0.3464, GoF(all) 1.146, $\Delta\rho_{min,max}$ –1.038, 1.055 e$^-$ Å^{-3}.

*$R1 = \Sigma||F_o| - |F_c||/\Sigma|F_o|$ for $F_o > 2\sigma(F_o)$; $wR2 = (\Sigma w(F_o^2 - F_c^2)^2/\Sigma(wF_c^2)^2)^{1/2}$ all reflections $w=1/[\sigma^2(F_o^2)+(0.2000P)^2+30.0000P]$ where $P=(F_o^2+2F_c^2)/3$

Specific refinement details

Crystals of **4·5.5CB$_{11}$H$_{12}$·6.5OTf·2.5MeCN·*i*Pr$_2$O** were grown by slow diffusion of diisopropyl ether into a CD_3CN solution of **4**$(OTf)_{12}$ containing *ca.* 12 equivalents of $CsCB_{11}H_{12}$. The crystals employed rapidly lost solvent after removal from the mother liquor and rapid handling prior to flash cooling in the cryostream was required to collect data. Despite these measures and the use of high intensity synchrotron radiation, few reflections at greater than 1.2 Å resolution were observed. We furthermore observed rapid degradation of the crystal under synchrotron radiation, producing a poor data/parameter ratio and poor completeness of the data, reflected in the checkCIF report. Nevertheless, the quality of the data is more than sufficient to establish the connectivity of the structure. Due to the less than ideal resolution, bond lengths and angles within pairs of organic ligands were restrained to be similar to each other (SAME) and thermal parameter restraints (SIMU, DELU) were applied to all non-metal atoms to facilitate anisotropic refinement. Ligand-based atoms that still displayed thermal parameters greater than 0.5 were further refined to approximate isotropic behaviour (ISOR).

The centrally-bound triflate molecule was heavily disordered and modelled over two sites with bond length (DFIX) and thermal parameter (EADP) restraints. Another triflate anion was modelled over three positions. Three carborate anions showed a significant amount of thermal motion; the SAME command was applied to all unique carborate anions to approximate icosahedral symmetry and realistically model these anions. Given the 12-fold-symmetric disorder of the carbon atom in $CB_{11}H_{12}^-$, all twelve atoms were assigned as boron; there was no indication that one of the atoms was carbon outright. One $CB_{11}H_{12}^-$ anion, two triflate anions and one acetonitrile molecule were modelled with half occupancy.

The remaining triflate anions present in the asymmetric unit could not be successfully resolved despite numerous attempts at modelling, including the use of rigid bodies. Consequently, the SQUEEZE[11] function of PLATON[12] was employed to remove the

contribution of the electron density associated with the remaining anions and further highly disordered solvent molecules.

4. Voidoo calculations

In order to determine the available void space within **4**, a VOIDOO[13] calculation was performed on the crystal structure of **4**, with anions and solvent molecules removed. A virtual probe with the minimum radius, such that it would not exit the cavity of the structure, was employed. To obtain an approximation of the void of **3**, an MM3 molecular model (based on the structural data of **1**) was employed; however, the apertures of **3** were too large for probe radii <10 Å. The apertures of **3** were thus capped with ligands of $\mathbf{L}^{C}$, enabling a probe radius akin to that employed for the calculation on **4** to be used. The following parameters were changed from their default values in both calculations, following a previously published procedure.[14]

Probe radius: 2.0 Å

Primary grid spacing: 0.1

Maximum number of volume-refinement cycles: 30

Minimum size of secondary grid: 1

Grid for plot files: 0.2

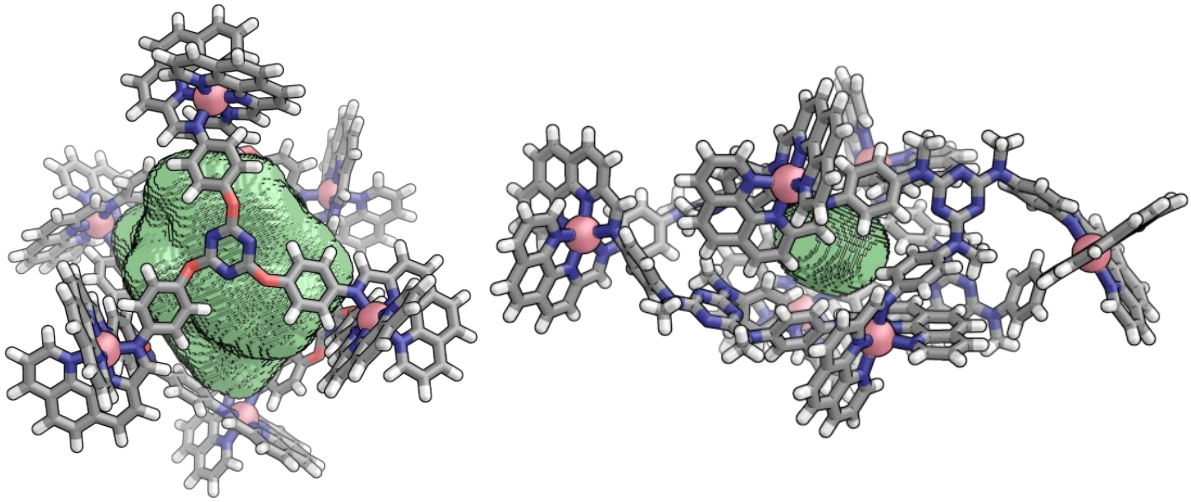


Figure S15. Void space of a molecular model of **3** (left) and the crystal structure of **4** (right), calculated using VOIDOO. The volume of the void of **3** was determined to be 1448 ± 5 $Å^3$; the volume of the void of **4** was determined to be 98 ± 1 $Å^3$. Voids are shown in green.

5. Sorting experiments

5.1 General procedure

Two triamines (1.60×10^{-6} mol of each triamine, 4 equiv of each triamine), 2-formylphenanthroline (2.00 mg, 9.61×10^{-6} mol, 24 equiv) and $Co(OTf)_2$ (1.71 mg, 4.80×10^{-6} mol, 12 equiv) were combined in CD_3CN (0.5 mL) and heated at 70 °C overnight. ^1H NMR and mass spectra were collected upon cooling to room temperature.

5.2 Subcomponents A and B

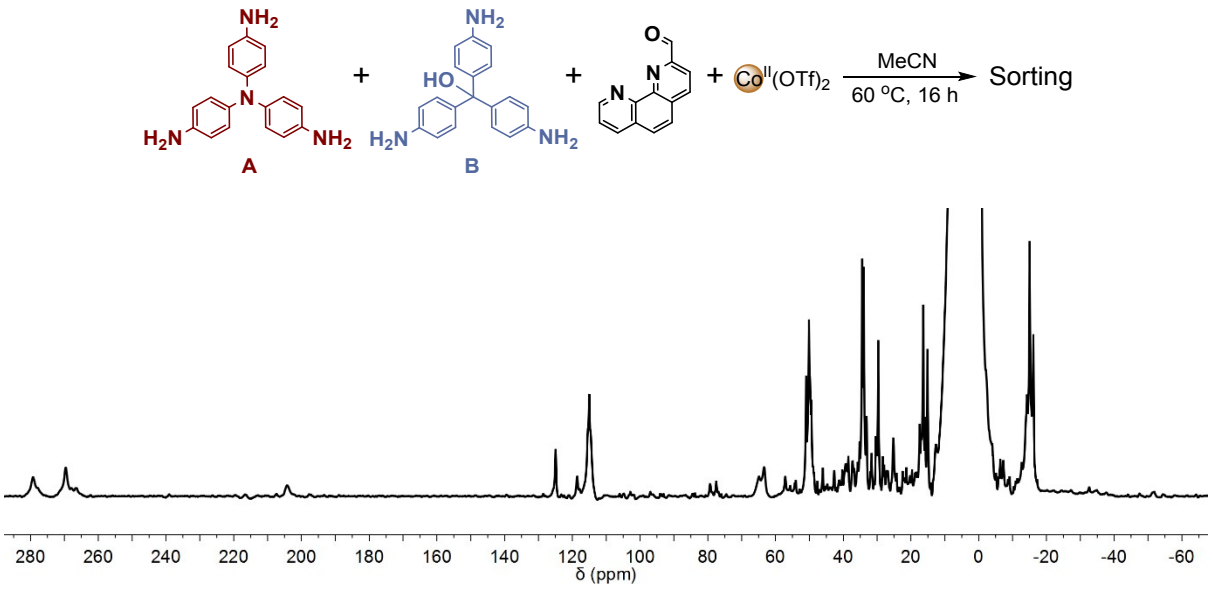


Figure S16. Wide sweep ^1H NMR spectrum (400 MHz, 298 K, CD_3CN) of the mixture of cages generated when both **A** and **B** were used in the self-assembly process.

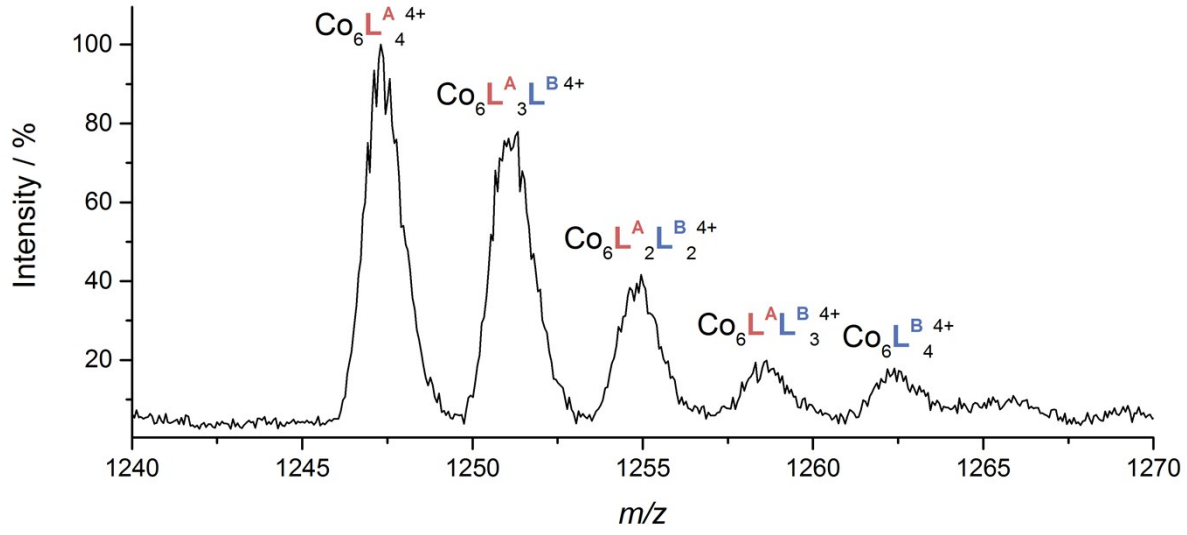


Figure S17. ESI mass spectrum of the mixture of cages generated when both **A** and **B** were used in the self-assembly process, showing the $z = +4$ peaks.

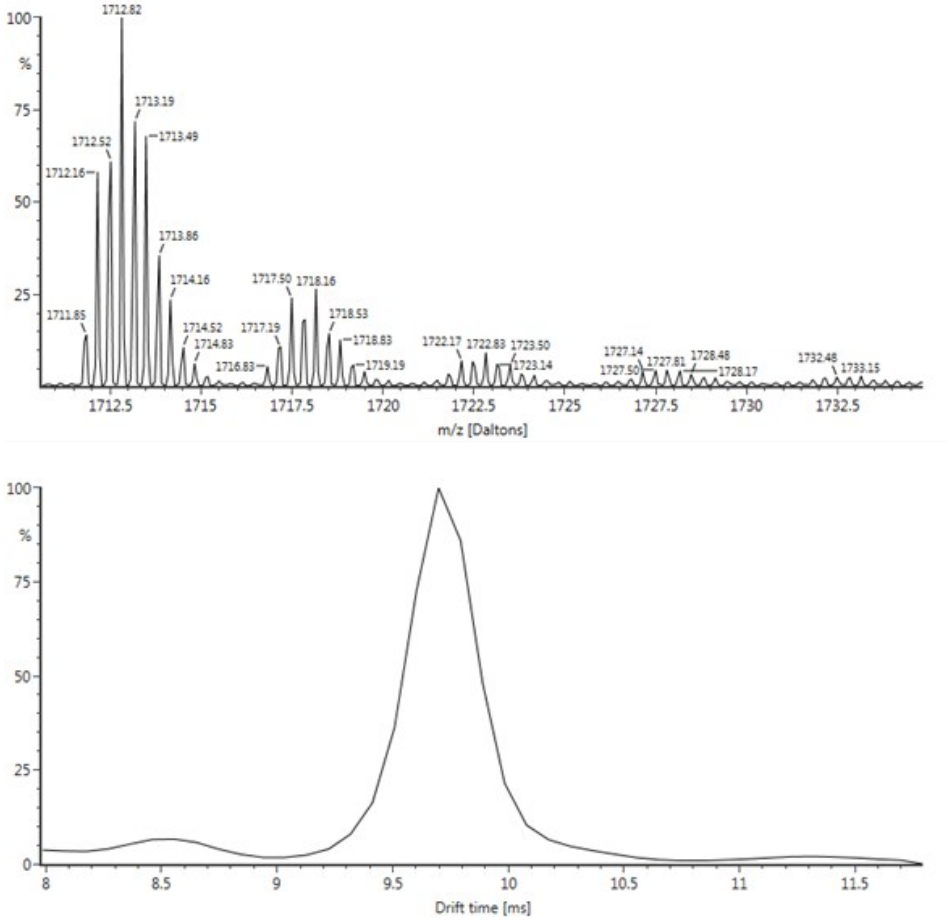


Figure S18. IM-MS data for cages generated when both **A** and **B** were used in the self-assembly process, showing the z = +3 peaks (top). The assembly resolved a single, broad drift time, indicating similar sizes between the sorted species.

5.3 Subcomponents A and C

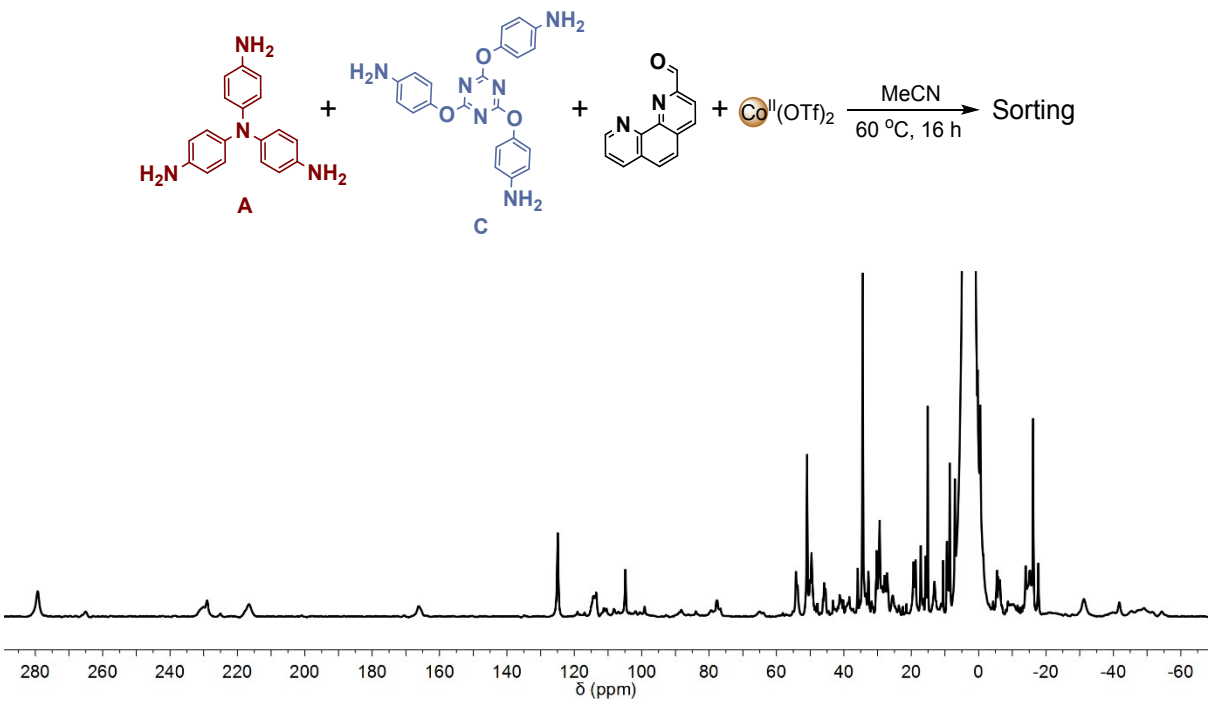


Figure S19. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of the mixture of cages generated when both **A** and **C** were used in the self-assembly process.

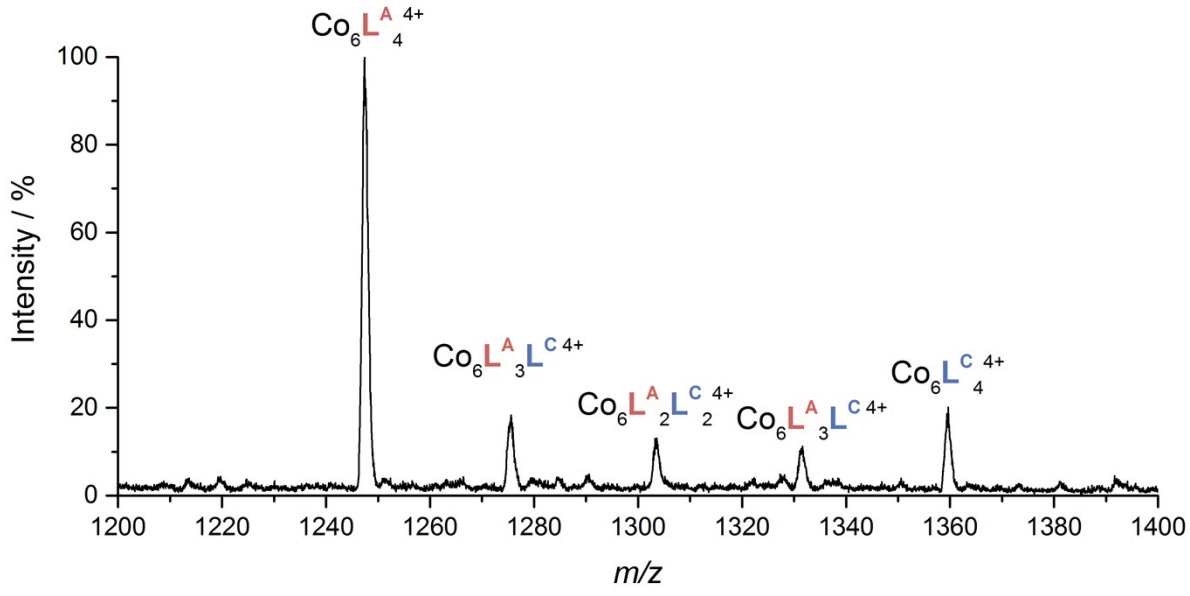


Figure S20. ESI mass spectrum of the mixture of cages generated when both **A** and **C** were used in the self-assembly process, showing the z = +4 peaks.

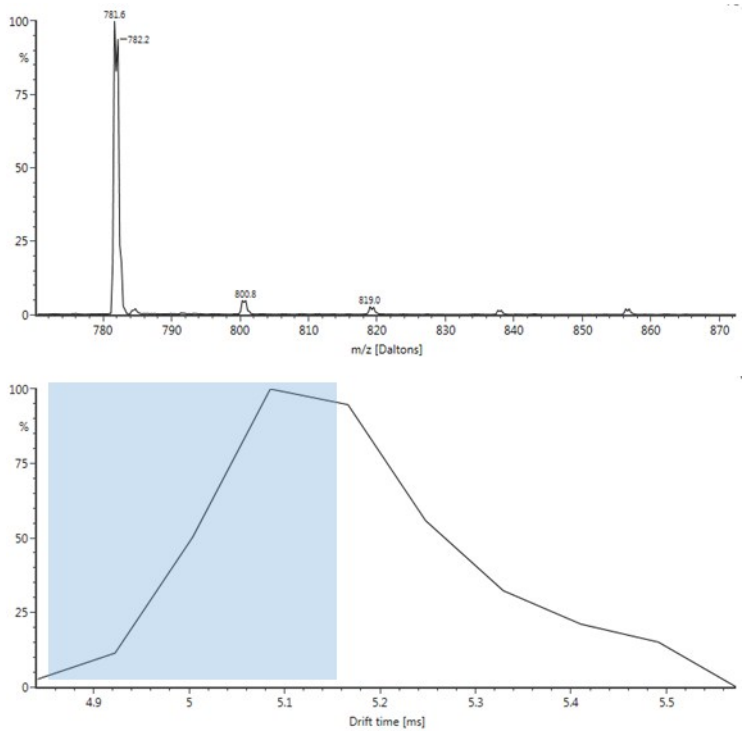


Figure S21. IM-MS data for cages generated when both **A** and **C** were used in the self-assembly process, showing the z = +6 peaks (top). The species resolved with broad, asymmetric drift times, indicative of multiple species sizes in the assembly for this charge fragment set.

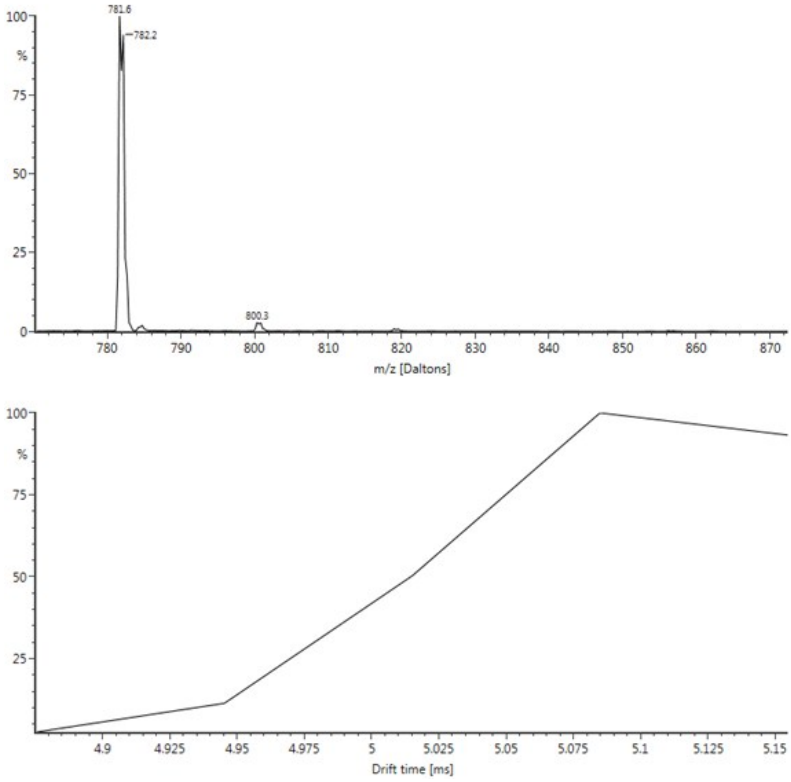


Figure S22. When only the blue region in the drift time plot in Figure S21 was examined (*i.e.* the species with faster drift times), only peaks corresponding to species with a greater proportion of subcomponent **A** were observed.

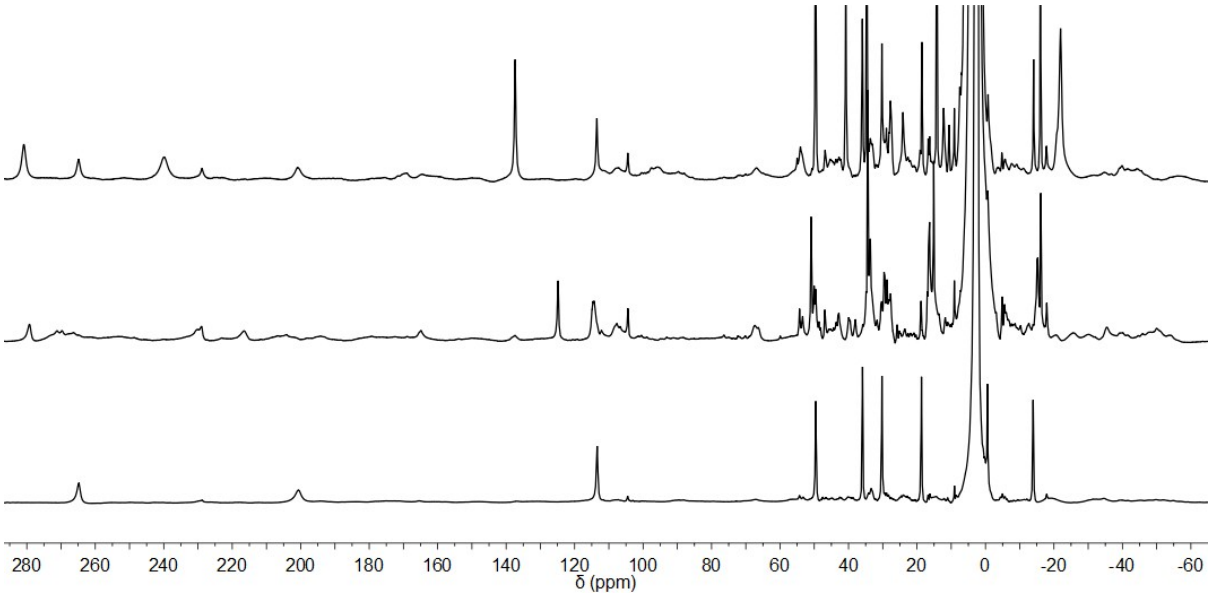


Figure S23. Wide sweep ^{1}H NMR spectra (400 MHz, 298 K, CD_3CN) of homoleptic **3** (bottom), the mixture of cages generated when both **A** and **C** were used in the self-assembly process (middle) and the spectrum following the addition of nBu$_4$BPh$_4$ (10 equiv) and heating at 70 °C over 3 days (topmost). Simplification of the spectrum to two prominent imine environments (>240 ppm) was observed, corresponding to **3** and $BPh_4^-\cdot$**1**. The presence of other signals <220 ppm indicates that some heteroleptic species remain.

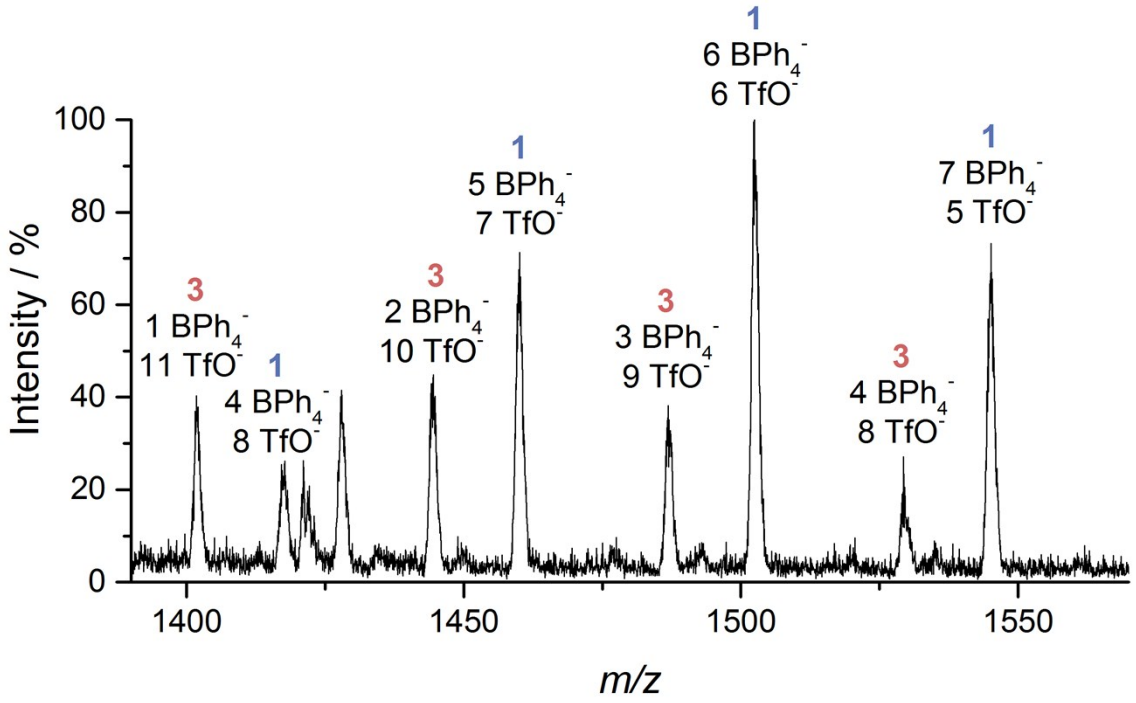


Figure S24. ESI mass spectrum of $Co^{II}_6\mathbf{L}^{A}_{x}\mathbf{L}^{C}_{4-x}$ to which was added nBu$_4$BPh$_4$ (10 equiv) and heated at 70 °C for three days. Peaks attributed to homoleptic **1** and **3** were the only species that could be identified. Only the +4 charge fragments have been displayed for clarity.

5.4 Subcomponents A and D

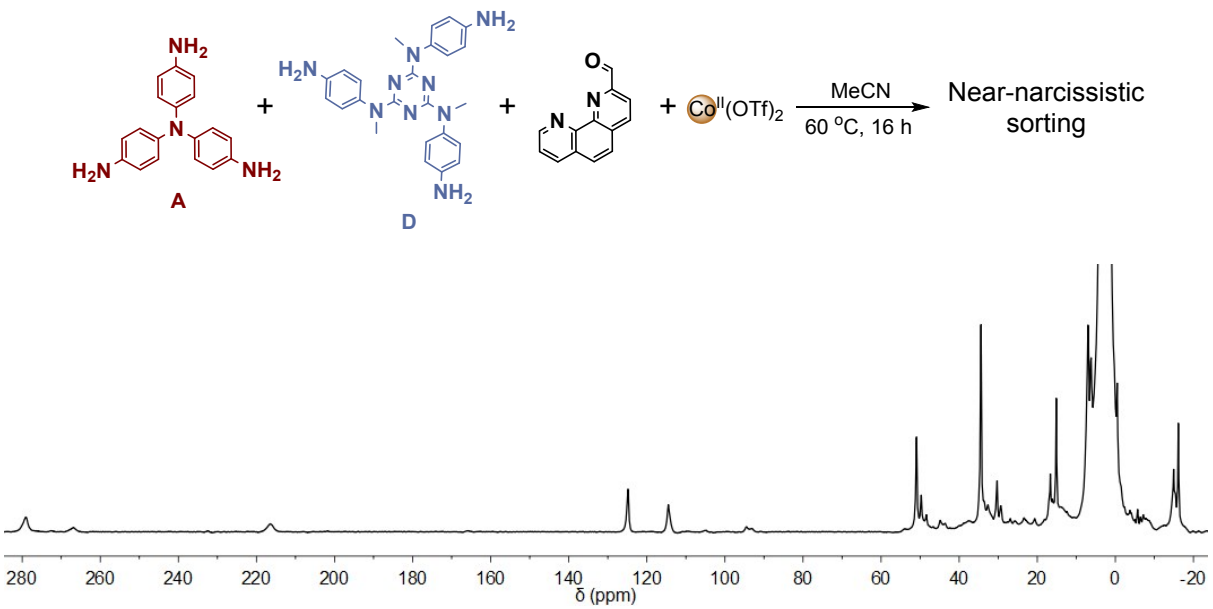


Figure S25. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of the mixture of cages generated when both **A** and **D** were used in the self-assembly process.

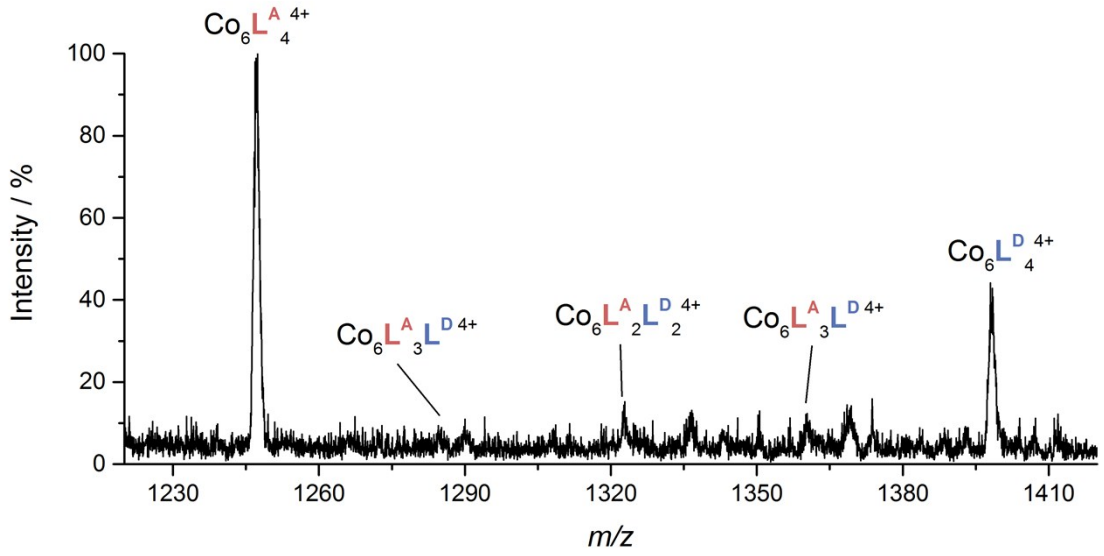


Figure S26. ESI mass spectrum of the mixture of cages generated when both **A** and **D** were used in the self-assembly process, showing the z = +4 peaks. Heteroleptic species were almost completely absorbed into the baseline of the spectrum.

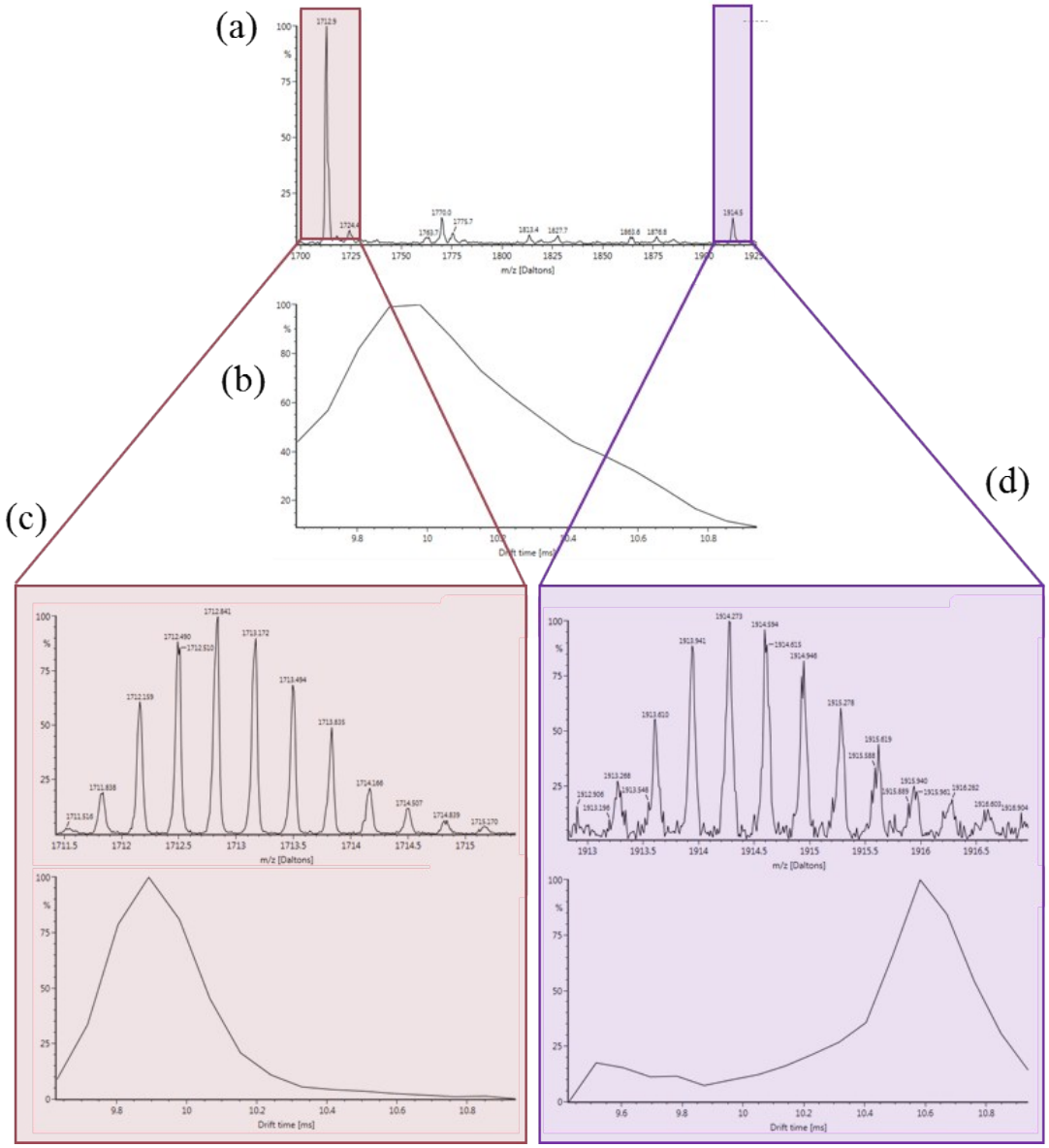


Figure S27. IM-MS data for cages generated when both **A** and **D** were used in the self-assembly process, showing the z = +3 peaks of all sorting possibilities (a). The sample resolved two broad, overlapped drift time regions at 9.6-10.4 ms and 10.4-11 ms, which may be assigned to the octahedron and scalenohedral geometries, respectively (b). We were able to assign these regions by examining the *m/z* regions of the purely homoleptic cages, which revealed two unique drift times. Peaks attributed to heteroleptic species were too weak in this case to examine, reflective of the near-narcissistic sorting of this assembly.

5.5 Subcomponents B and C

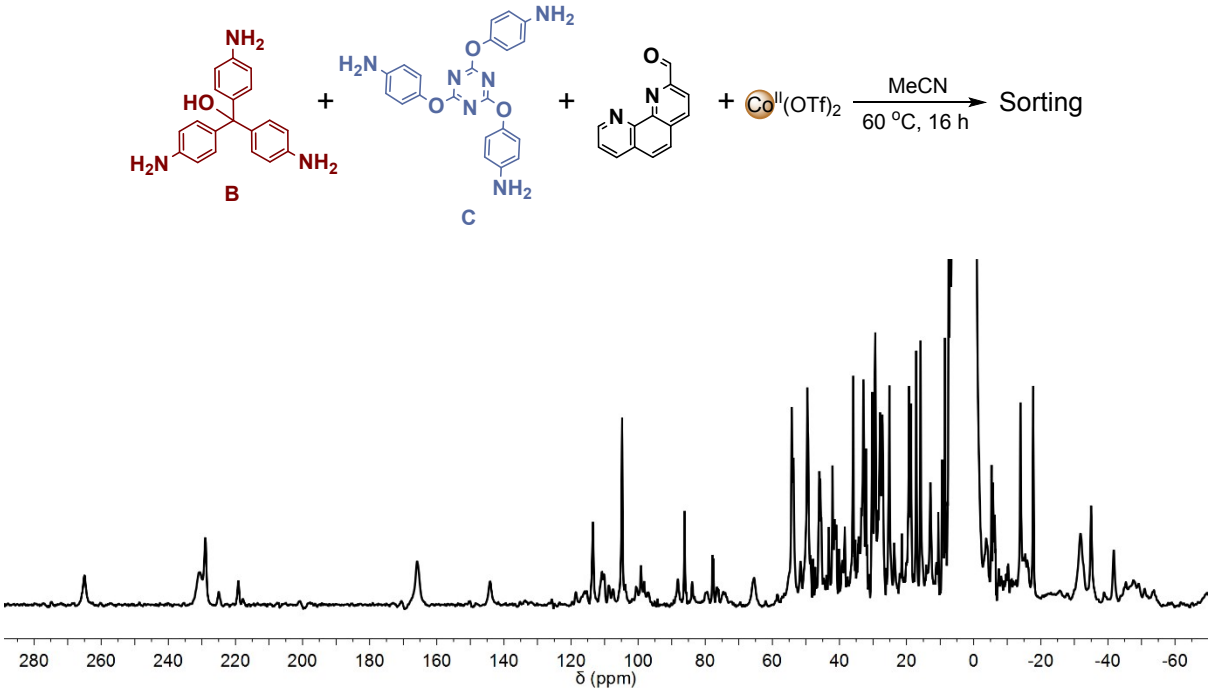


Figure S28. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of the mixture of cages generated when both **B** and **C** were used in the self-assembly process.

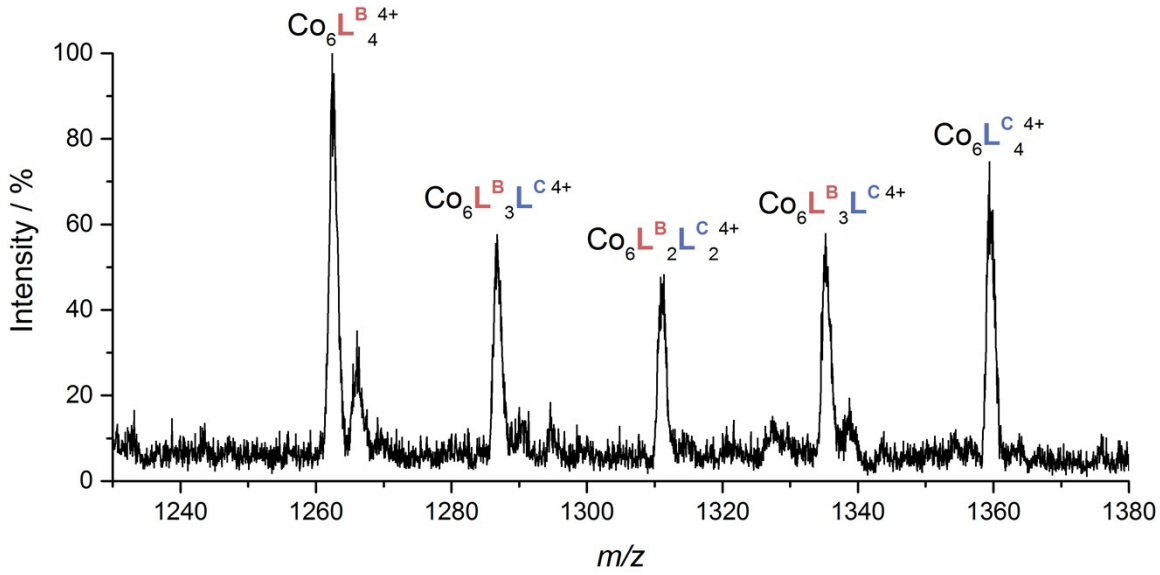


Figure S29. ESI mass spectrum of the mixture of cages generated when both **B** and **C** were used in the self-assembly process, showing the z = +4 peaks.

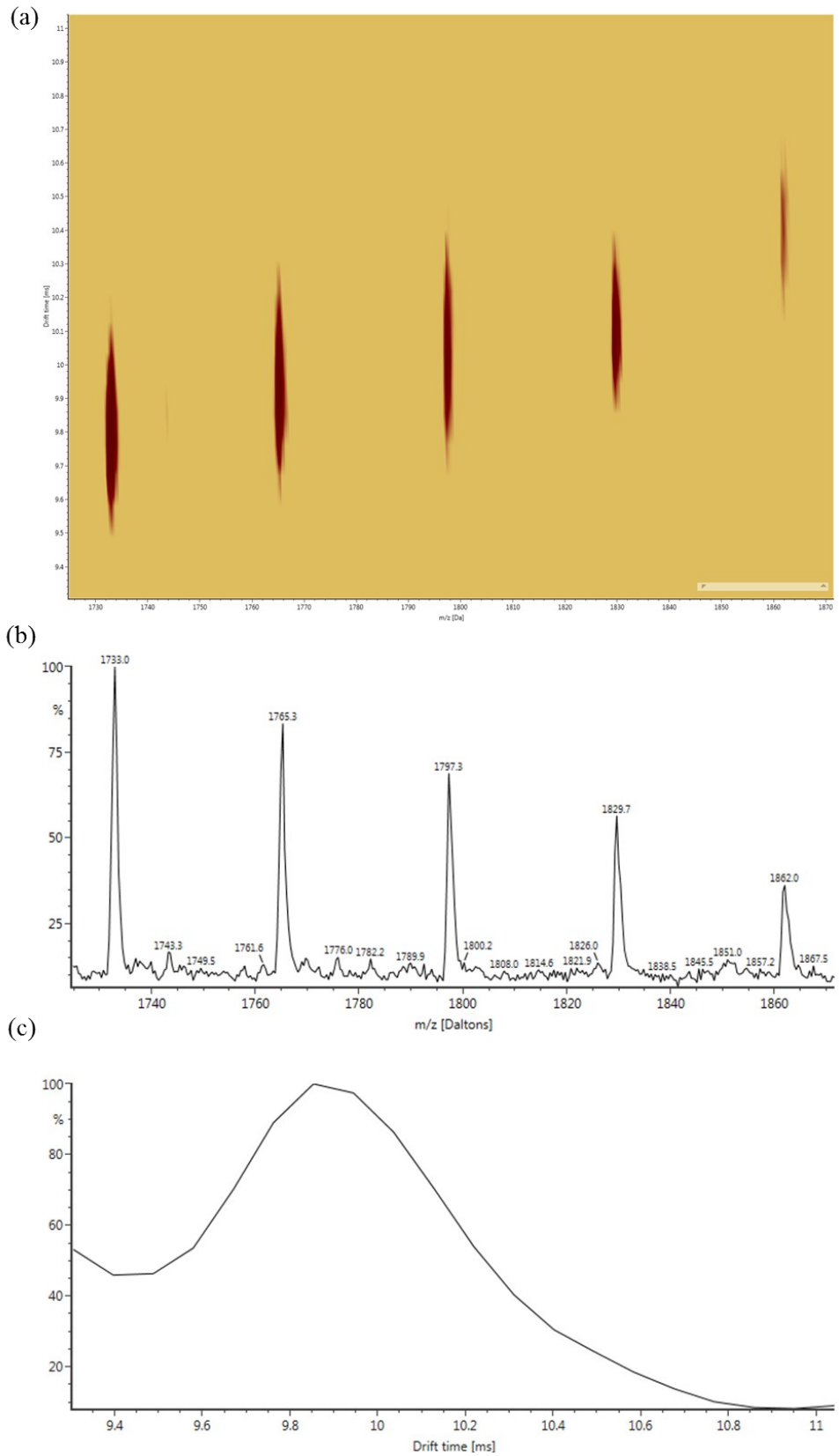


Figure S30. IM-MS data for cages generated when both **B** and **C** were used in the self-assembly process, showing the z = +3 peaks of all sorted possibilities (a,b). The sample resolved a single, broad drift time region at 9.6-10.4 ms for this segment of the spectrum, indicative of the octahedral geometry of all complexes (c).

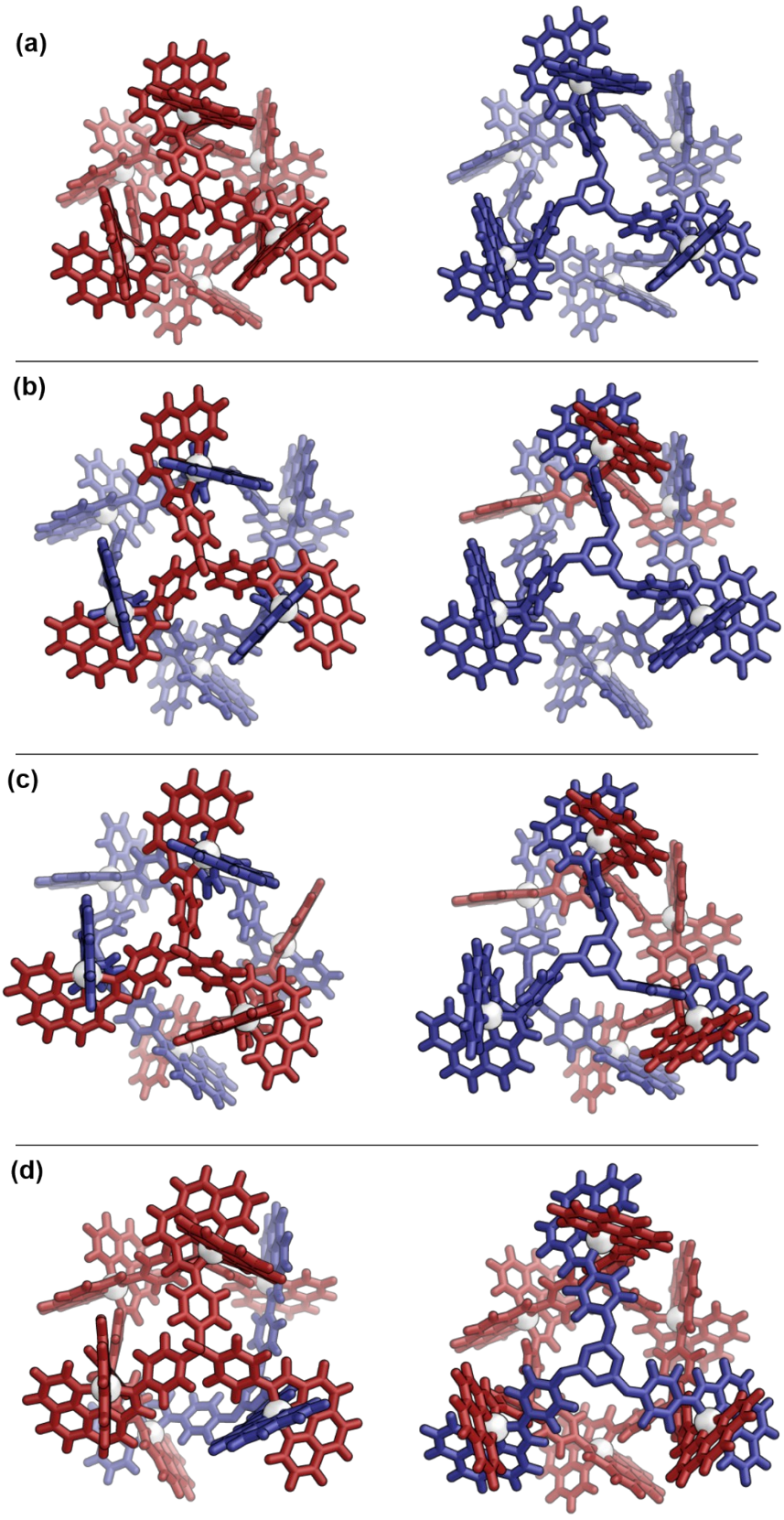


Figure S31. MM3 molecular models of the (a) homoleptic and (b-d) heteroleptic complexes obtained when both **B** and **C** were used in the self-assembly process. All species are modelled as the all-Λ enantiomer.

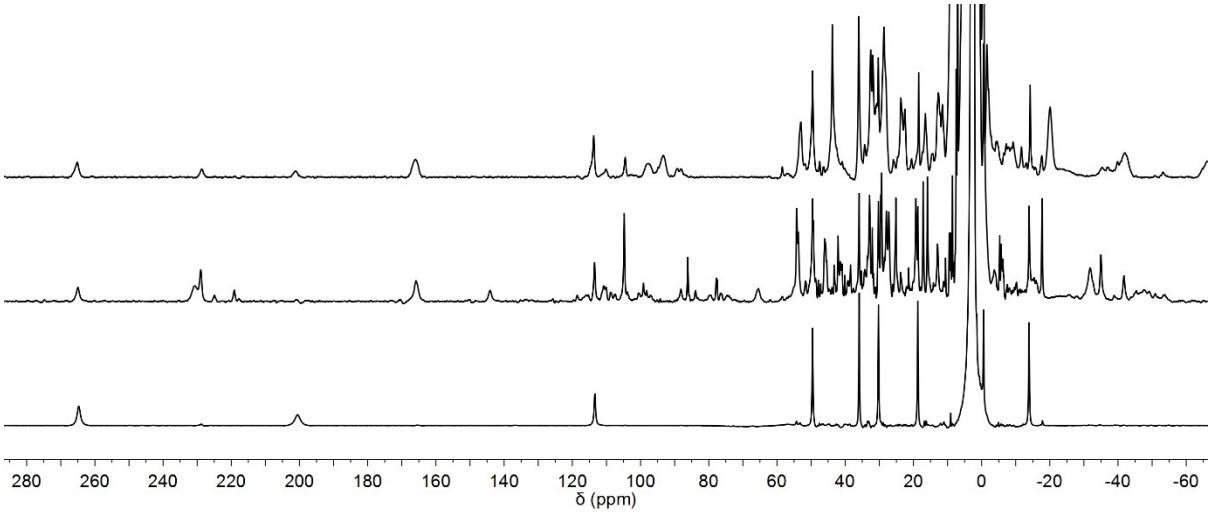


Figure S32. Wide sweep ^{1}H NMR spectra (400 MHz, 298 K, CD_3CN) of homoleptic **3** (bottom), the mixture of cages generated when both **B** and **C** were used in the self-assembly process (middle) and the spectrum following the addition of nBu$_4$BPh$_4$ (10 equiv) and heating at 70 °C overnight (topmost). Simplification of the spectrum to only two distinct imine environments (>220 ppm) was observed, corresponding to **3** and BPh_4^-⊂**2**. The presence of other signals <120 ppm indicates that some heteroleptic species remain.

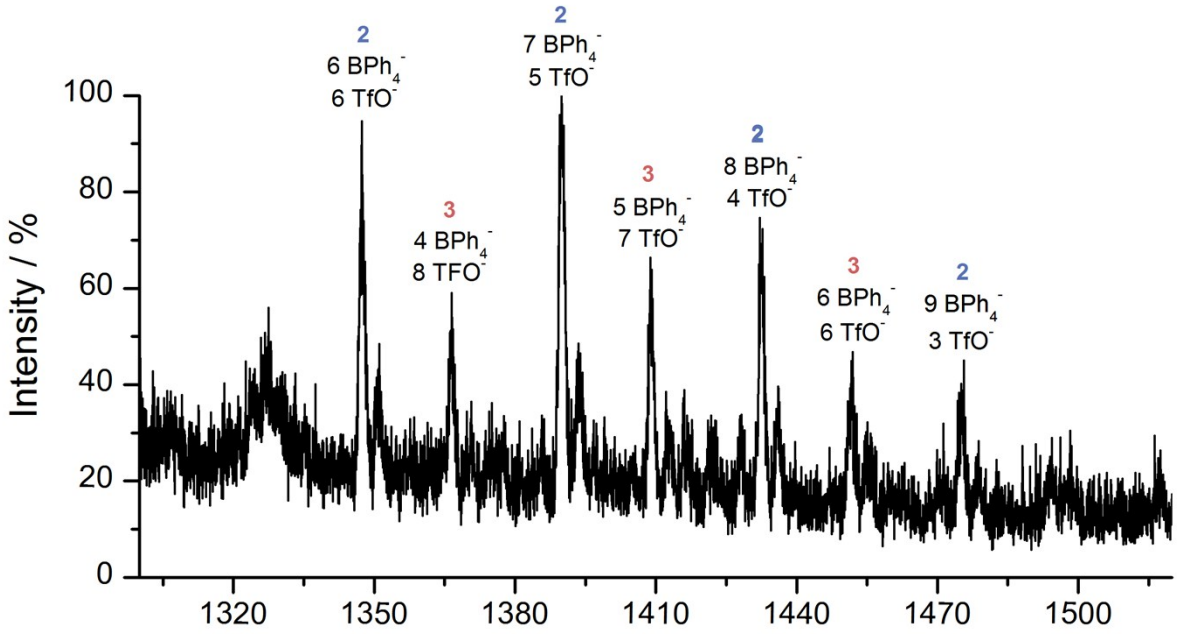


Figure S33. ESI mass spectrum of $Co^{II}_6\mathbf{L}^{\mathbf{B}}_x\mathbf{L}^{\mathbf{C}}_{4-x}$ to which was added nBu$_4$BPh$_4$ (10 equiv) and heated at 70 °C overnight. Peaks attributed to homoleptic **2** and **3** can be identified as the prominent species. Only the +4 charge fragments have been displayed for clarity.

5.6 Subcomponents B and D

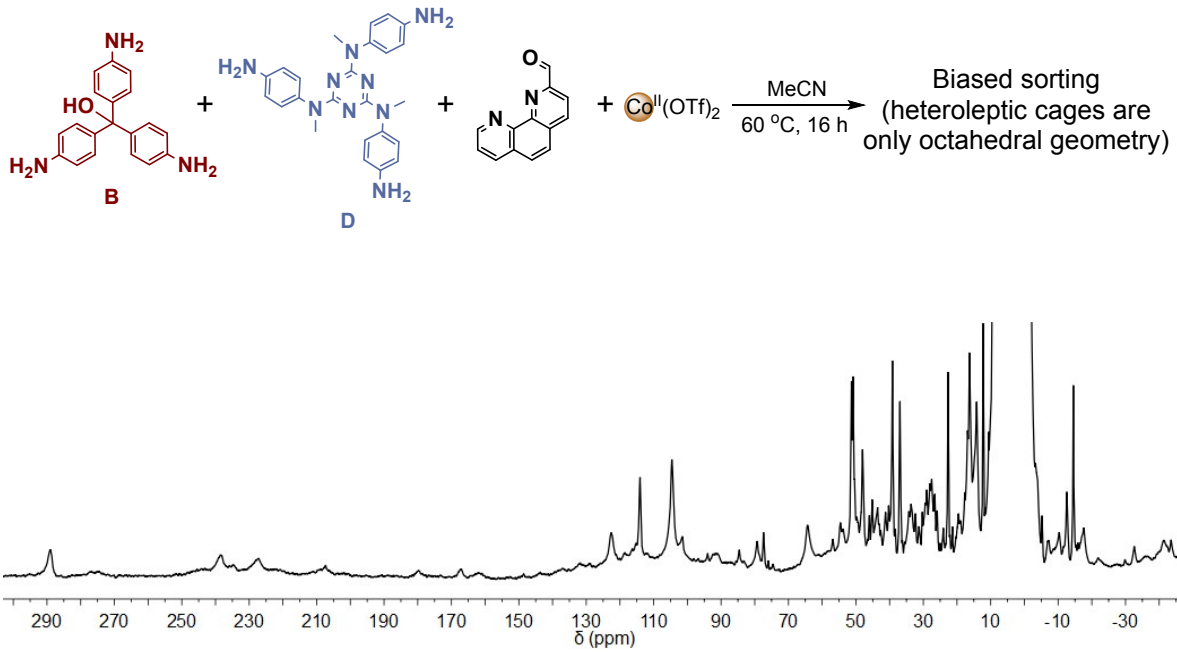


Figure S34. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of the mixture of cages generated when both **B** and **D** were used in the self-assembly process.

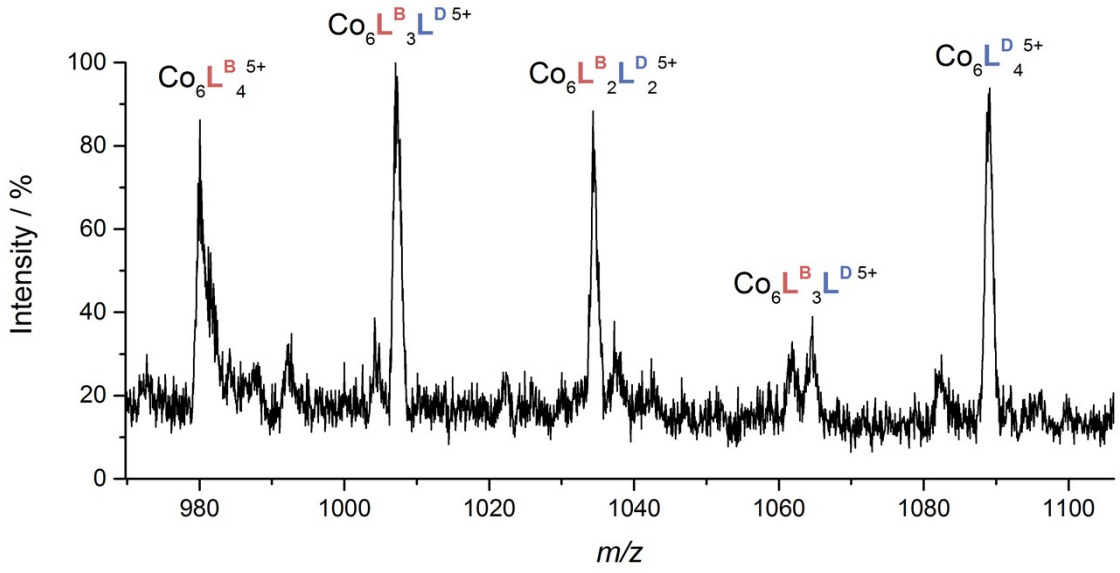


Figure S35. ESI mass spectrum of the mixture of cages generated when both **B** and **D** were used in the self-assembly process, showing the z = +5 peaks.

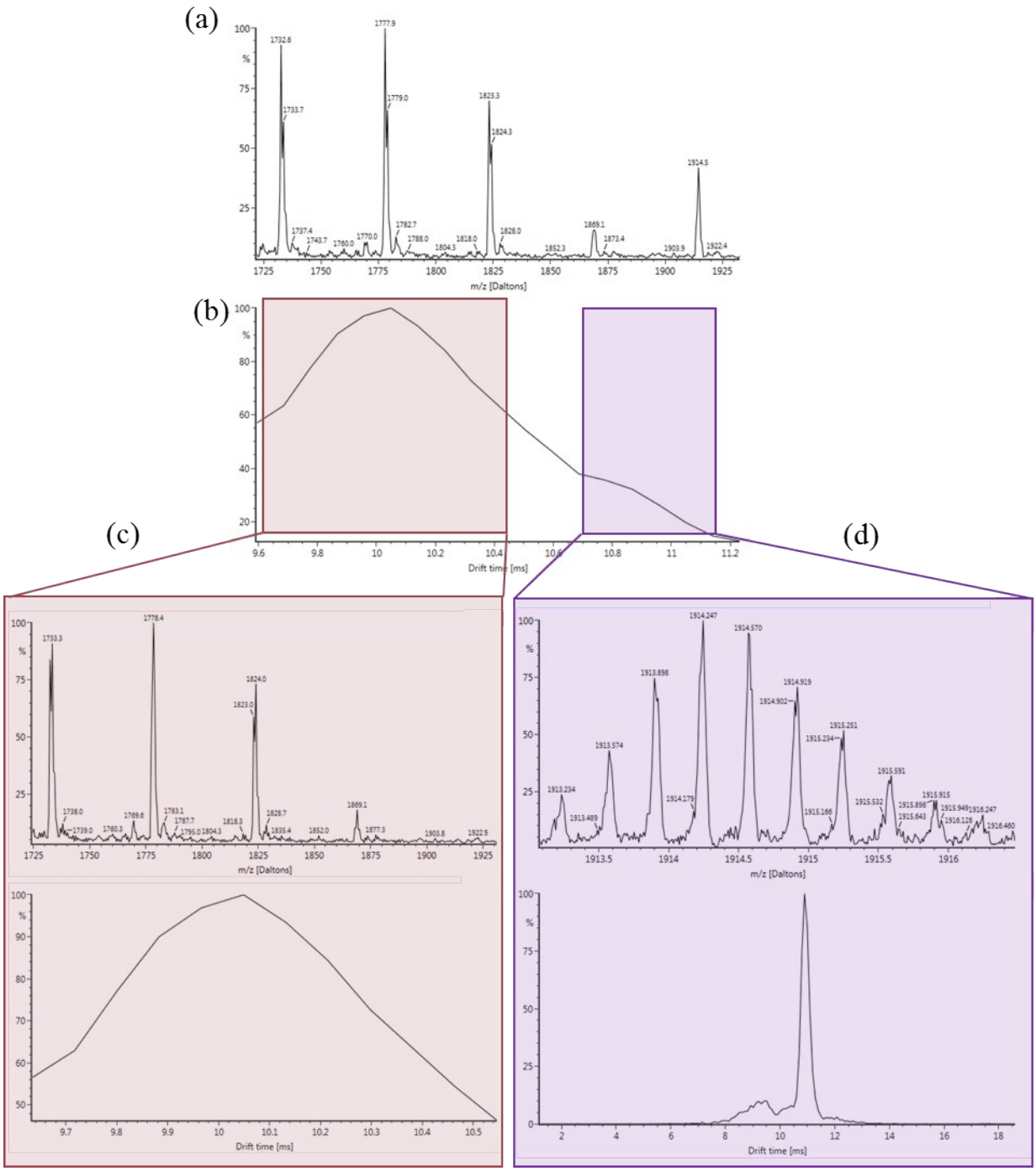


Figure S36. IM-MS data for cages generated when both **B** and **D** were used in the self-assembly process, showing the z = +3 peaks of all sorted possibilities (a). The sample resolved two distinct, broad drift time regions at 9.6-10.6 ms and 10.6-11.2 ms, which may be assigned to the octahedron and scalenohedral geometries, respectively (b). When these two regions were examined separately, we noted that all species comprised of some portion of **B** displayed octahedral geometry (c). Likewise, when the scalenohedron range was examined, only **4** was observed (d). The data indicates that all heteroleptic complexes are octahedral in geometry; the only scalenohedron observed was homoleptic **4**.

5.7 Subcomponents C and D

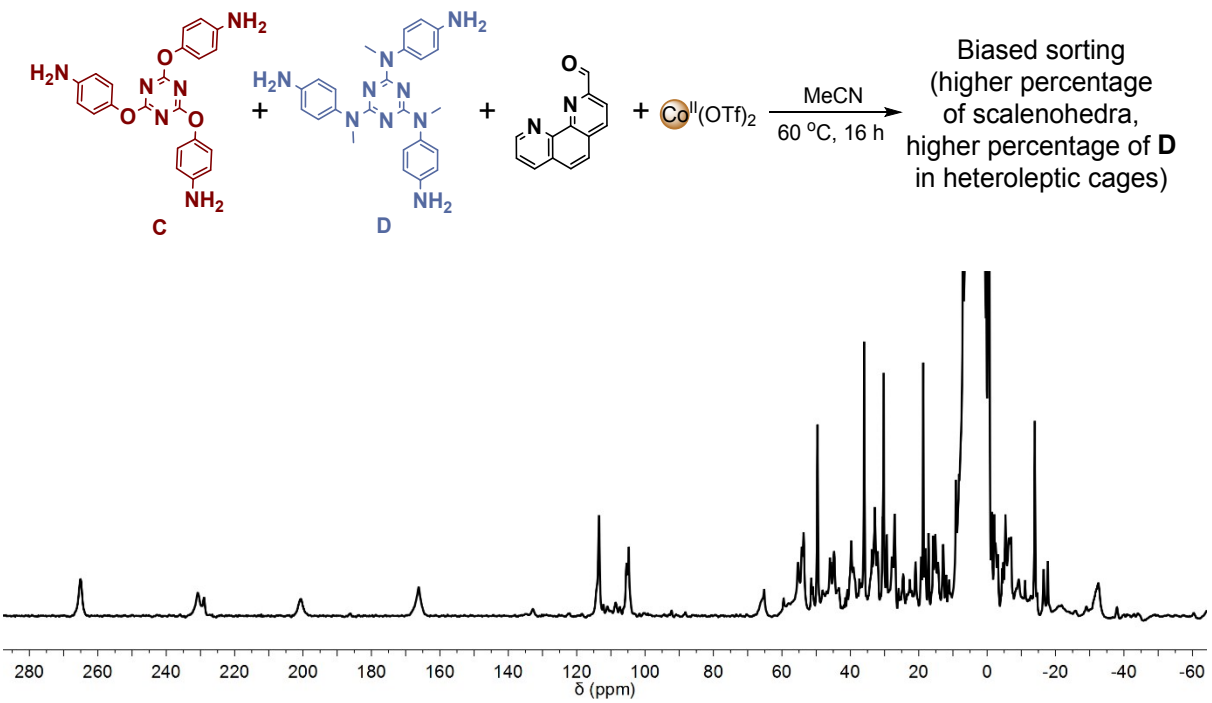


Figure S37. Wide sweep ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of the mixture of cages generated when both **C** and **D** were used in the self-assembly process.

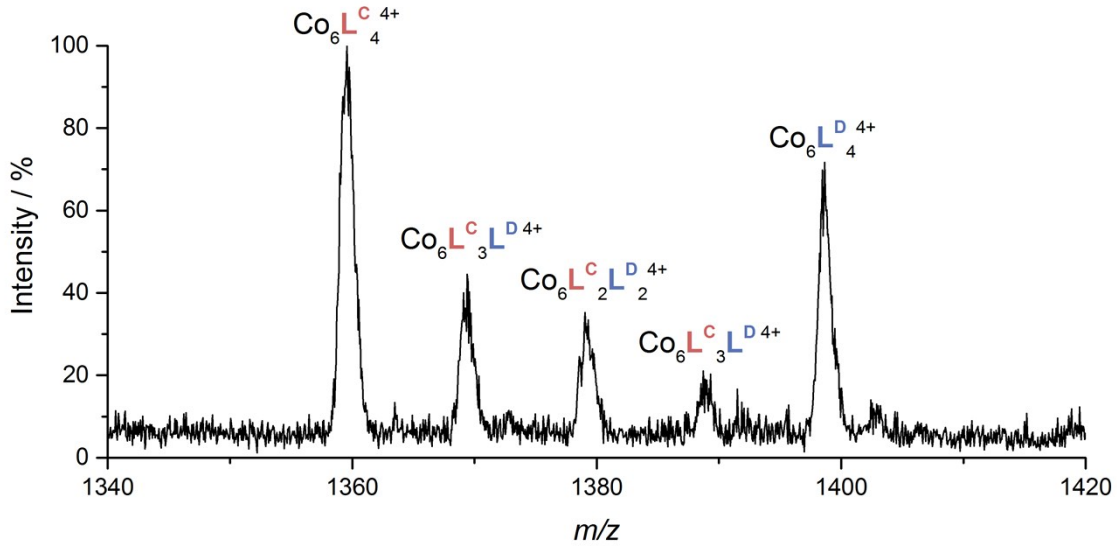


Figure S38. ESI mass spectrum of the mixture of cages generated when both **C** and **D** were used in the self-assembly process, showing the z = +4 peaks.

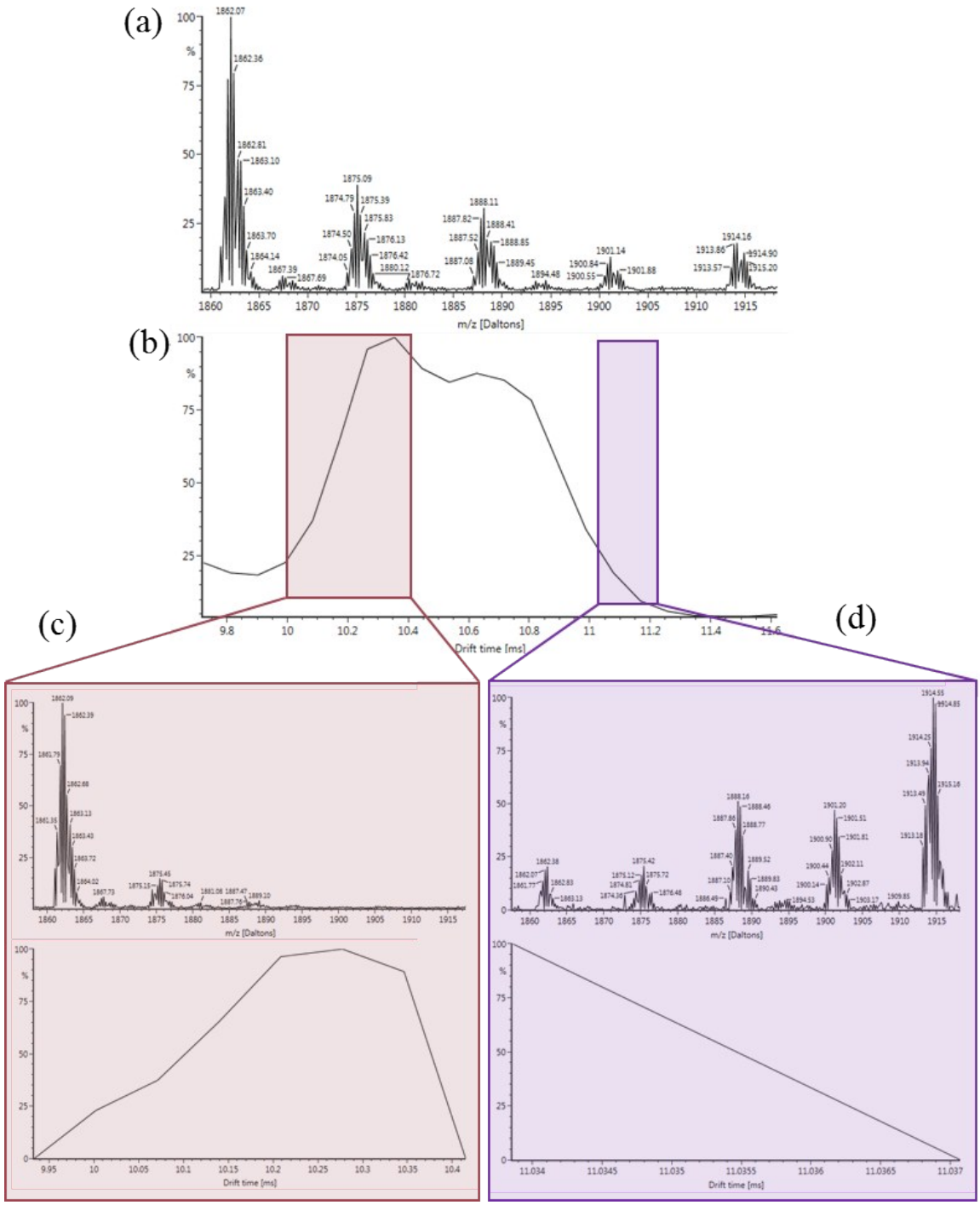


Figure S39. IM-MS data for cages generated when both **C** and **D** were used in the self-assembly process, showing the z = +3 peaks of all sorted possibilities (a). The sample resolved two distinct, broad drift time regions at 10-10.5 ms and 10.6-11.2 ms, which we assign to the octahedron and scalenohedral geometries, respectively (b). When these regions were examined separately, we noted that the octahedron region was comprised of species with a higher proportion of **C** (c). Likewise, when the scalenohedron range was examined, species with a higher proportion of **D** were observed (d). This latter spectrum is poorly resolved due to the significant overlapping of the two drift time regions.

6. Quantification of sorting distributions

Stock solutions of **1**, **2**, **3** and **4** (0.89 mM) were respectively prepared from triamine **A**, **B**, **C** and **D** (3.20×10^{-6} mol, 4 equiv), 2-formylphenanthroline (2.00 mg, 9.61×10^{-6} mol, 12 equiv) and $Co(OTf)_2$ (1.72 mg, 4.81×10^{-6} mol, 6 equiv), heated at 70 °C overnight in CH_3CN (0.9 mL). Once cooled to room temperature, 0.1 mL of two homoleptic cage solutions were combined, making solutions of two purely homoleptic complexes: **1** and **2**, **1** and **3**, **1** and **4**, **2** and **3**, **2** and **4** and **3** and **4**. These mixtures were heated at 70 °C overnight, after which equilibrium was reached and the self-sorted output was formed. Similarly, freshly added mixtures of homoleptic cages were prepared by mixing 0.1 mL of the homoleptic cage solutions directly before spectral measurement.

Both fresh and equilibrated mixtures were analysed by low resolution electrospray ionisation (LR-ESI) mass spectrometry on a Micromass Quattro LC mass spectrometer (cone voltage 22 eV; desolvation temp. 313 K; ionization temp. 313 K) infused from a Harvard syringe pump at a rate of 10 μL min^{-1}. The data used for analysis were averaged over a 30 collection scans, to minimise instrument error.

For all mixtures, the +4, +5 and +6 charge states were used for analysis: lower charge states were outside the spectral window of the instrument, and higher charge states overlapped with too many fragments to be incorporated into the analysis. Peaks corresponding to both homoleptic and heteroleptic species were observed in the MS spectra of equilibrated mixtures, whereas no peaks matching with heteroleptic species were observed in the freshly combined mixtures.

Each region of the MS spectrum (corresponding to either a homoleptic or heteroleptic cage) was integrated for both equilibrated and fresh mixtures. No peaks for heteroleptic species were observed in the freshly combined mixtures. To compensate for fragmentation in the heteroleptic regions, the integrals for regions corresponding to heteroleptic cages in freshly combined mixtures were subtracted from the integrals of the corresponding regions in the spectra of the equilibrated mixtures. Under the assumption that all homoleptic species consumed are converted into heteroleptic species (no free ligand is released), integrals for each charge state were summed to unity to facilitate comparison (Figure S39-S44).

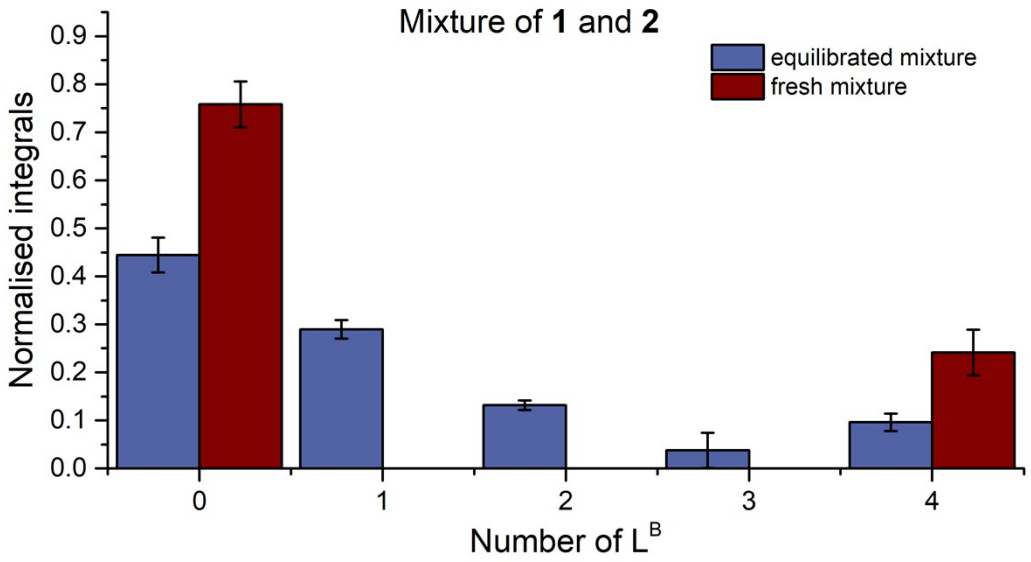


Figure S40. Normalised integrals for equilibrated (blue) and fresh (red) mixtures of **1** and **2**. The error bars represent the standard deviation of the relative amounts of each species between the different charge states observed in the ESI-MS.

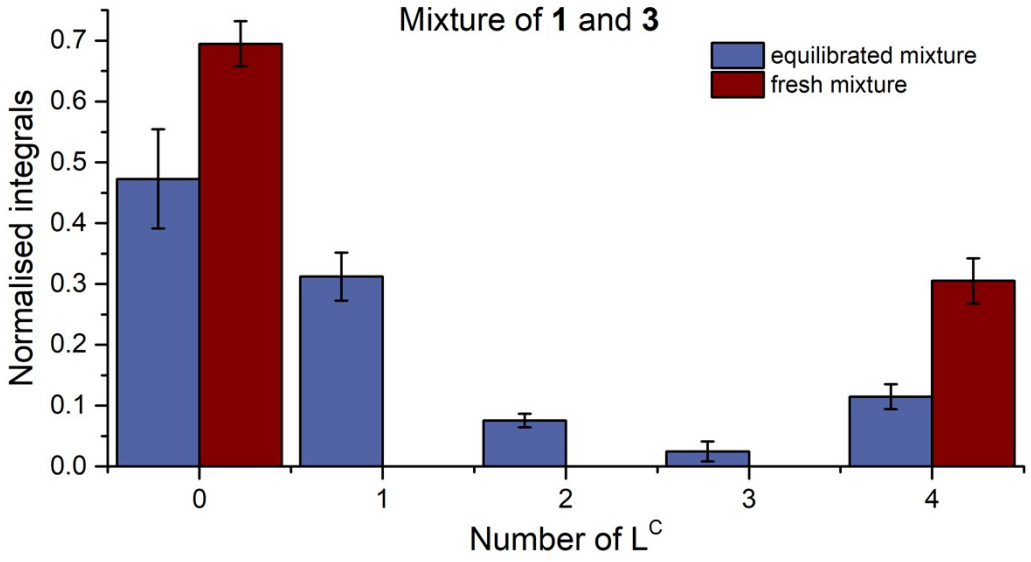


Figure S41. Normalised integrals for equilibrated (blue) and fresh (red) mixtures of **1** and **3**. The error bars represent the standard deviation of the relative amounts of each species between the different charge states observed in the ESI-MS.

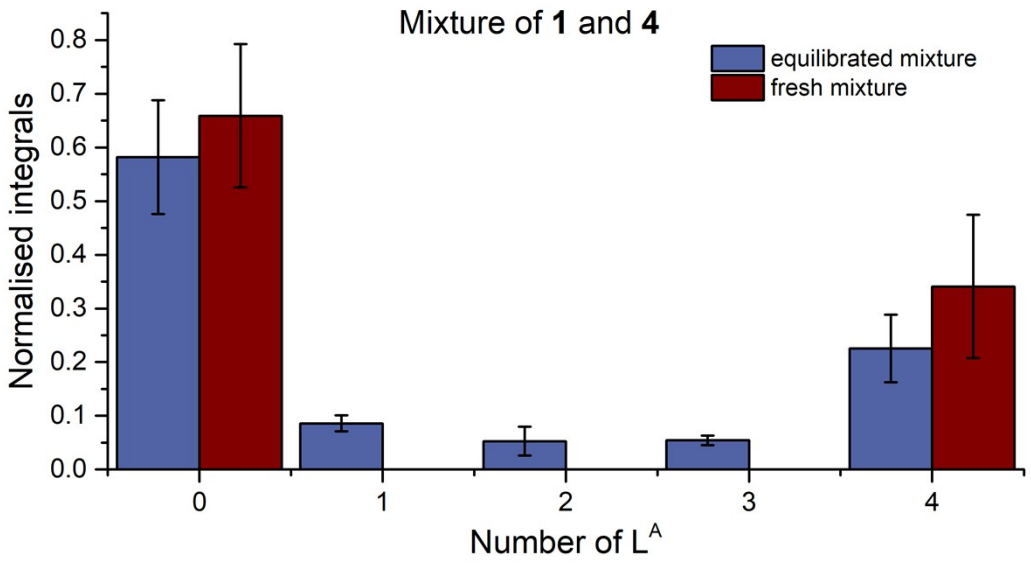


Figure S42. Normalised integrals for equilibrated (blue) and fresh (red) mixtures of **1** and **4**. The error bars represent the standard deviation of the relative amounts of each species between the different charge states observed in the ESI-MS.

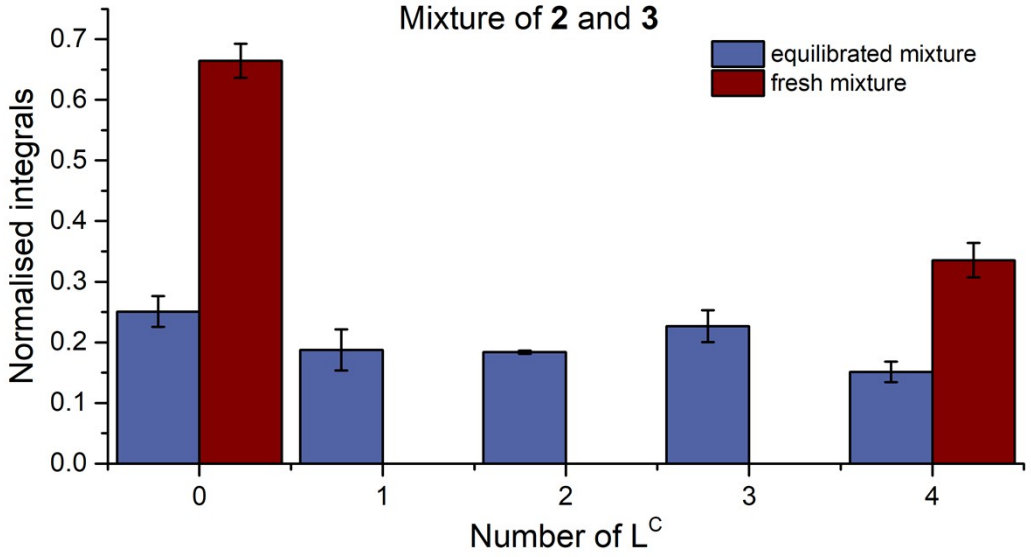


Figure S43. Normalised integrals for equilibrated (blue) and fresh (red) mixtures of **2** and **3**. The error bars represent the standard deviation of the relative amounts of each species between the different charge states observed in the ESI-MS.

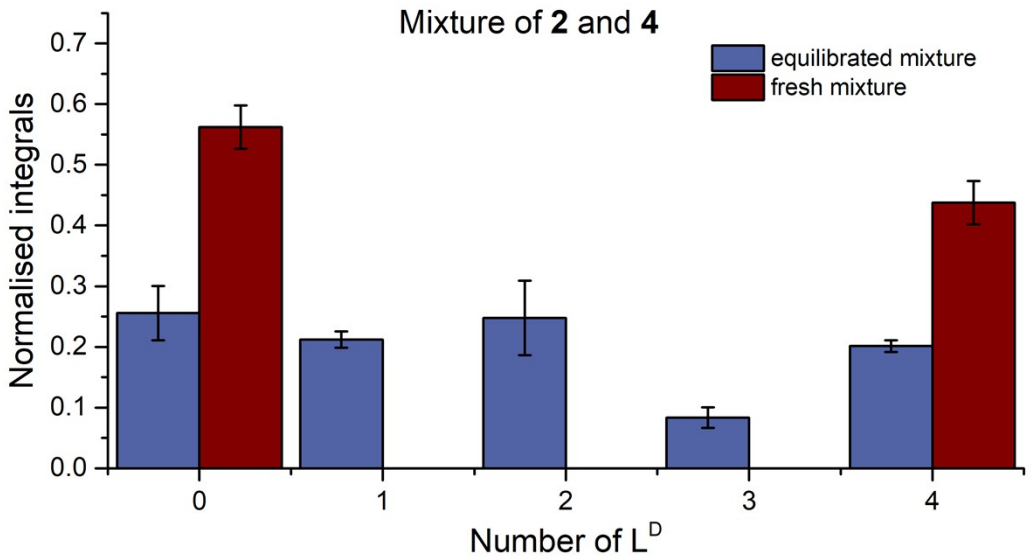


Figure S44. Normalised integrals for equilibrated (blue) and fresh (red) mixtures of **2** and **4**. The error bars represent the standard deviation of the relative amounts of each species between the different charge states observed in the ESI-MS.

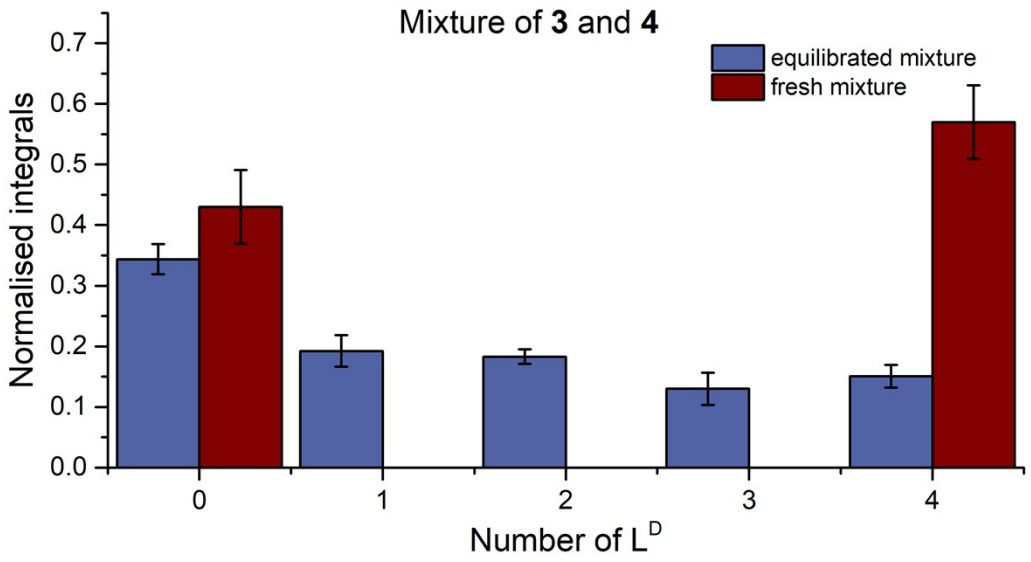


Figure S45. Normalised integrals for equilibrated (blue) and fresh (red) mixtures of **3** and **4**. The error bars represent the standard deviation of the relative amounts of each species between the different charge states observed in the ESI-MS.

No comparison between structures incorporating different ratios of ligands can be made, as the variation of ionisation and detection between species could be large. However, comparison between amounts of homoleptic species in both equilibrated and fresh mixtures can be made, as they are structurally identical, and thus exhibit identical ionisation

characteristics. This method eliminates effects associated with different detection intensities of charged complexes with different structures and different ligand configurations.

For example, in a self-sorted mixture of **A** and **B** (generating $Co^{II}\mathbf{L}^{\mathbf{A}}_{x}\mathbf{L}^{\mathbf{B}}_{4-x}$), the percentage of **1** and **2** remaining in the equilibrated mixture, as compared to **1** and **2** in the fresh mixture, is linked to the proportion of homoleptic species consumed by the generation of heteroleptic species. The decrease in homoleptic species thus represents the percentage of heteroleptic species in solution, assuming that all subcomponents generate cages. Likewise, the individual percentage decrease in either **1** or **2** indicates the proportion of that subcomponent integrated into heteroleptic cages, and can be calculated directly from the decrease observed in **1** or **2** in the equilibrated mixtures, as opposed to **1** or **2** in the fresh mixtures.

The percentage of homoleptic species remaining after equilibration of the mixtures can be obtained by comparing the normalised integrals of identical homoleptic species in the equilibrated ($I_{x(eq)}$) and the fresh mixtures ($I_{x(fresh)}$). The percentage of each homoleptic species (**X** = **1**, **2, 3**, or **4**) remaining in the equilibrated mixtures is given by:

$$\%X(eq) = \frac{I_{X(eq)}}{I_{X(fresh)}} * 100$$

The values obtained for all mixtures are given in the table below:

Table S1: Percentage of homoleptic species remaining after equilibration of the mixtures of homoleptic cages.

Mixtures	*Percentage of homoleptic species remaining*	
1* and *2	Percentage of **1**	Percentage of **2**
	59 ± 3	40 ± 8
1* and *3	Percentage of **1**	Percentage of **3**
	68 ± 8	37 ± 2
1* and *4	Percentage of **1**	Percentage of **4**
	89 ± 7	69 ± 11
2* and *3	Percentage of **2**	Percentage of **3**
	38 ± 5	46 ± 9
2* and *4	Percentage of **2**	Percentage of **4**
	45 ± 5	46 ± 5
3* and *4	Percentage of **3**	Percentage of **4**
	81 ± 11	27 ± 4

Similarly, the distribution between heteroleptic and homoleptic species is obtained by measuring the overall % of homoleptic species remaining after, as compared to before,

equilibration. The quantity of heteroleptic species is the reciprocal of this percentage. For example, the distribution between species in a mixture of **1** and **2** is given by:

$$\%Homoleptic(1,2) = \frac{\%1(eq) + \%2(eq)}{2}$$

$$\%Heteroleptic(1,2) = 100 - \%Homoleptic$$

The values obtained for all mixtures are given in the table below:

Table S2: Distribution between homoleptic and heteroleptic species after equilibration of the mixtures of homoleptic cages.

Mixtures	*Percentage of homoleptic species*	*Percentage of heteroleptic species*
1* and *2	49 ± 6	51 ± 6
1* and *3	53 ± 5	47 ± 5
1* and *4	79 ± 9	21 ± 9
2* and *3	42 ± 7	58 ± 7
2* and *4	46 ± 5	54 ± 5
3* and *4	54 ± 7	46 ± 7

Under the assumption that all homoleptic species consumed are converted into heteroleptic species during equilibration, the percentage of ligand **A**, **B**, **C** or **D** incorporated into the heteroleptic cages of the equilibrated mixtures can thus be quantified. This quantity is the % decrease of the corresponding homoleptic species in the equilibrated mixture, as compared to that in the fresh mixture. For example, in a mixture of **1** and **2**, the percentage of **A** incorporated in the heteroleptic species is given by:

$$\%A(hetero) = \frac{(100 - \%1(eq))}{(100 - \%1(eq)) + (100 - \%2(eq))} * \%Heteroleptic(1,2)$$

The ligand composition of heteroleptic cages can be obtained from the percentage of each ligands incorporated in heteroleptic species by cross multiplication. For examples, in a mixture of **1** and **2**, the ligand composition of heteroleptic cages is given by:

$$\%A(composition) = \frac{\%A(hetero)}{\%A(hetero) + \%B(hetero)} * 100$$

The values obtained for all mixtures are given in the table below:

Table S3: Ligand composition of heteroleptic species after equilibration of the mixtures of homoleptic cages.

Mixtures	*Ligand composition of heteroleptic species*	
1 *and* ***2***	Percentage of **A**	Percentage of **B**
	41 ± 3	59 ± 8
1 *and* ***3***	Percentage of **A**	Percentage of **C**
	34 ± 8	66 ± 2
1 *and* ***4***	Percentage of **A**	Percentage of **D**
	27 ± 17	73 ± 25
2 *and* ***3***	Percentage of **B**	Percentage of **C**
	53 ± 4	47 ± 7
2 *and* ***4***	Percentage of **B**	Percentage of **D**
	50 ± 5	50 ± 5
3 *and* ***4***	Percentage of **C**	Percentage of **D**
	21 ± 12	79 ± 4

To complement our ESI-MS analysis, we compared our results to product distributions measured by ^{1}H NMR spectroscopy. The ^{1}H NMR spectra of the sorted reaction mixtures were complex: many homoleptic signals were significant broadened, and only select regions of the spectra could be used in each case. To calculate the approximate distribution of homoleptic vs. heteroleptic cages, the integrals of signals corresponding to the two homoleptic species were compared against those of all other signals in the region (the 'heteroleptic' signals). Only three spectral sets could be reliably analysed in this way (Figures S45-S47), all of which produce values in line with those reported in Table S2.

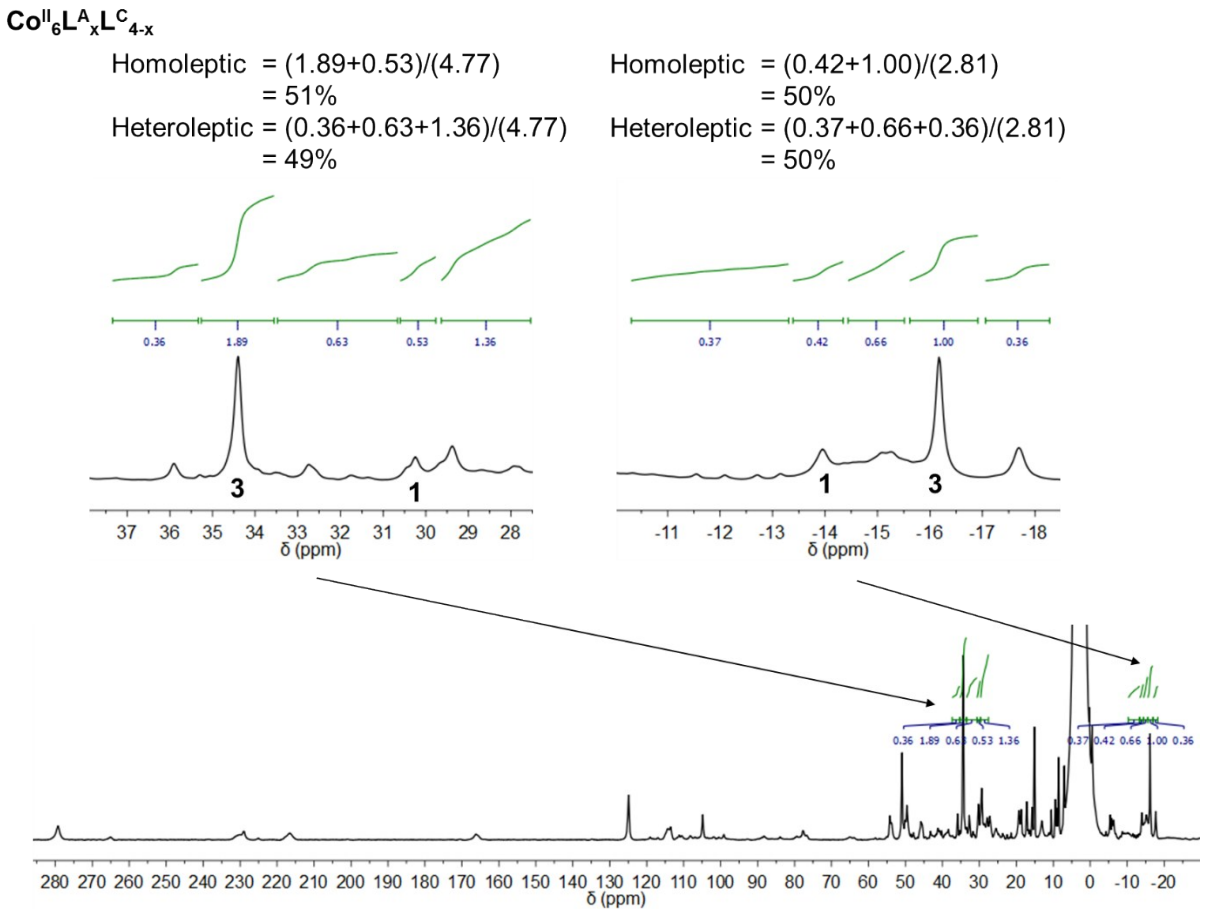


Figure S46. Integration of the proton signals of homoleptic **1** and **3** (from the sorting of **A** and **C**) vs. all other signals attributed to similar proton environments (i.e. the heteroleptic signals) over two regions of the ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of $Co^{II}_6\mathbf{L}^{A}_{x}\mathbf{L}^{C}_{4-x}$. The quantities are within error of the values determined by ESI-MS (Table S2).

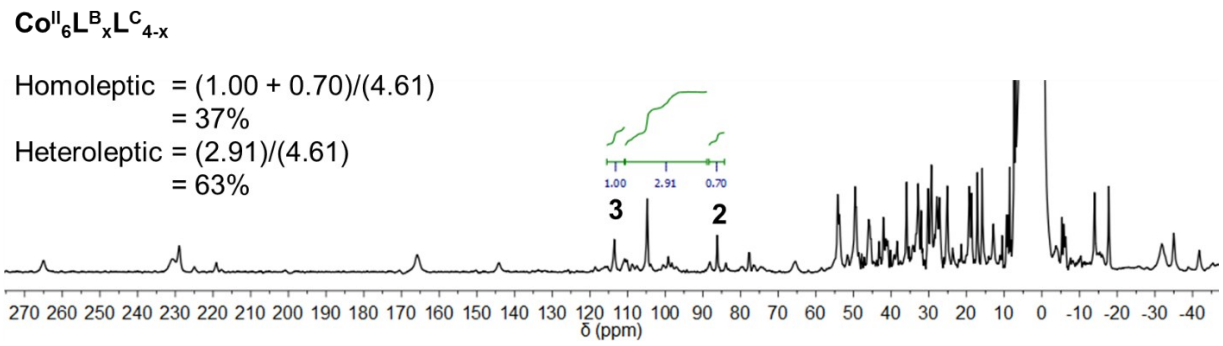


Figure S47. Integration of the proton signals of homoleptic **2** and **3** (from the sorting of **B** and **C**) vs. all other signals attributed to similar proton environments (i.e. the heteroleptic signals) over one region of the ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of $Co^{II}_6\mathbf{L}^{B}_{x}\mathbf{L}^{C}_{4-x}$. The quantities indicate preference for the formation of heteroleptic over homoleptic species, in line with the values determined by ESI-MS (Table S2).

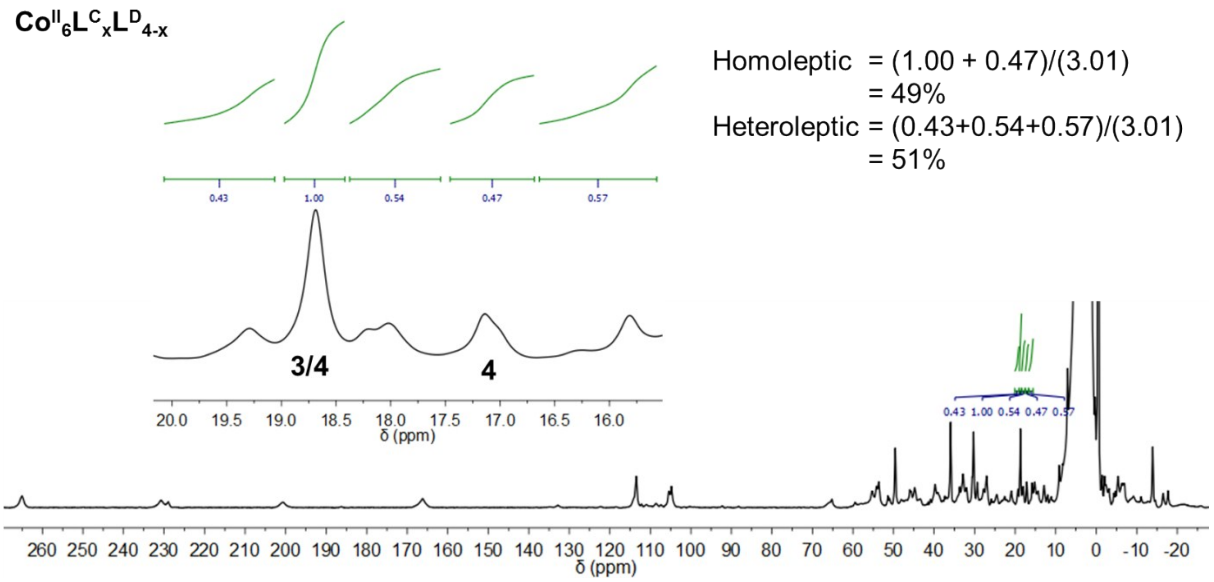


Figure S48. Integration of the proton signals of homoleptic **3** and **4** (from the sorting of **C** and **D**) vs. all other signals attributed to similar proton environments (i.e. the heteroleptic signals) over one region of the ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of $Co^{II}_6\mathbf{L}^C_x\mathbf{L}^D_{4-x}$. The quantities are within error of the values determined by ESI-MS (Table S2).

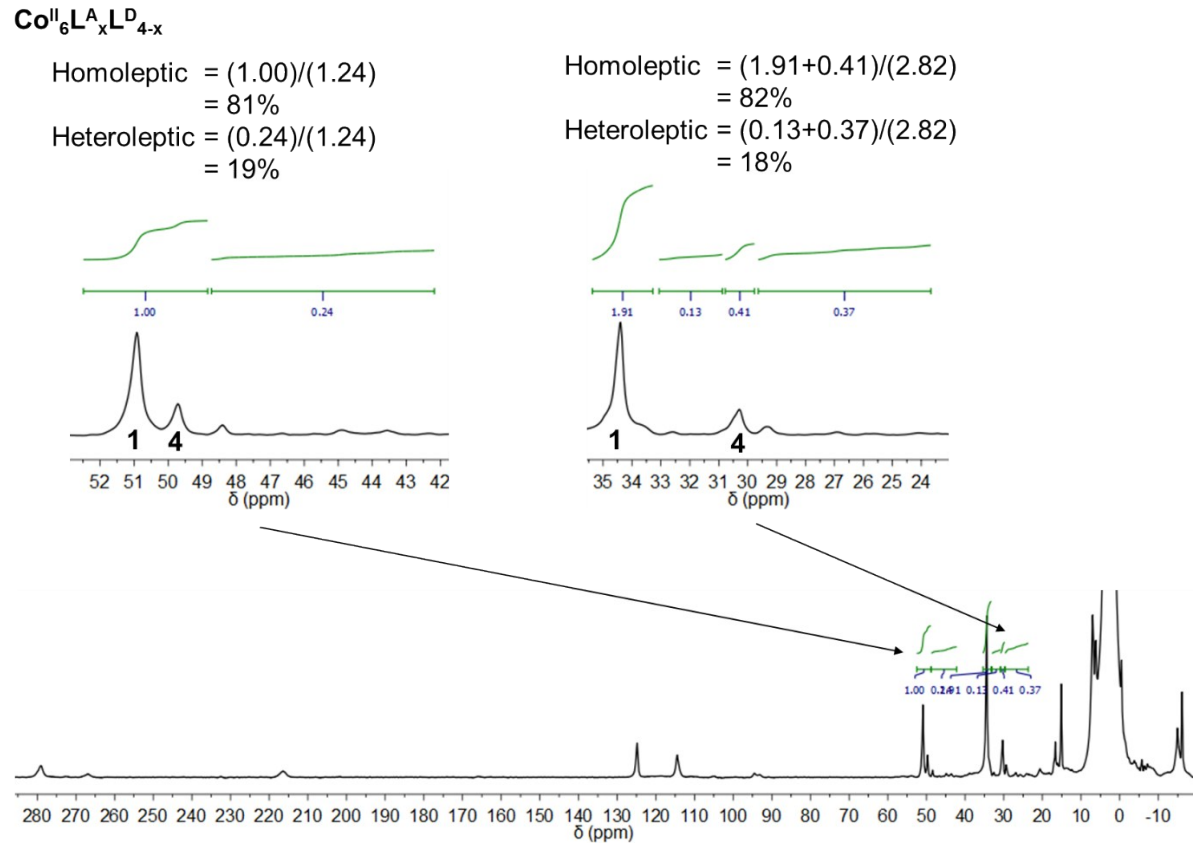


Figure S49. Integration of the proton signals of homoleptic **1** and **4** (from the sorting of **A** and **D**) vs. all other signals attributed to similar proton environments (i.e. the heteroleptic signals) over one region of the ^{1}H NMR spectrum (400 MHz, 298 K, CD_3CN) of $Co^{II}_6\mathbf{L}^A_x\mathbf{L}^D_{4-x}$. As only one signal for **4** could be identified reliably, the percentage of homoleptic species in this mixture is greater than that calculated using this method; likewise, the proportion of heteroleptic species is less than that calculated here. The quantities are within error of the values determined by ESI-MS (Table S2).

Table S4: Distribution between homoleptic and heteroleptic species after equilibration of the mixtures of homoleptic cages, comparing the values obtained by ESI-MS and NMR integration.

Analytical Method	*ESI-MS*		*NMR*	
Mixtures	*Percentage of homoleptic species*	*Percentage of heteroleptic species*	*Percentage of homoleptic species*	*Percentage of heteroleptic species*
1 *and* ***2***	49 ± 6	51 ± 6	Too broad to conclude	
1 *and* ***3***	53 ± 5	47 ± 5	51	49
1 *and* ***4***	79 ± 9	21 ± 9	>81	<19
2 *and* ***3***	42 ± 7	58 ± 7	37	63
2 *and* ***4***	46 ± 5	54 ± 5	Too broad to conclude	
3 *and* ***4***	54 ± 7	46 ± 7	49	51

7. References

1 M. Zhang, R. Gao, X. Hao and W.-H. Sun, *J. Organomet. Chem.*, 2008, **693**, 3867-3877.

2 J. L. Bolliger, T. K. Ronson, M. Ogawa and J. R. Nitschke, *J. Am. Chem. Soc.*, 2014, **136**, 14545-14553.

3 I. A. Riddell, M. M. J. Smulders, J. K. Clegg, Y. R. Hristova, B. Breiner, J. D. Thoburn and J. R. Nitschke, *Nat. Chem.*, 2012, **4**, 751-756.

4 C. S. Wood, T. K. Ronson, A. M. Belenguer, J. J. Holstein and J. R. Nitschke, *Nat. Chem.*, 2015, **7**, 354-358.

5 Agilent, *Crysalis PRO*, Agilent Technologies, Ltd., Yarnton, Oxfordshire, England, 2014.

6 P. R. Evans and G. N. Murshudov, *Acta Crystallographica Section D*, 2013, **69**, 1204-1214.

7 G. Winter, *Journal of Applied Crystallography*, 2010, **43**, 186-190.

8 Bruker, *SADABS*, Bruker AXS, Inc., Madison, Wisconsin, USA, 2001.

9 L. Farrugia, *J. Applied Cryst.*, 2012, **45**, 849-854.

10 G. M. Sheldrick, *Acta Crystallogr. Sect. A*, 2015, **71**, 3-8.

11 A. L. Spek, *Acta Crystallogr. Sect C*, 2015, **71**, 9-18.

12 A. L. Spek, *Acta Crystallogr. Sect. D*, 2009, **65**, 148-155.

13 G. J. Kleywegt and T. A. Jones, *Acta Crystallogr. Sect. D*, 1994, **50**, 178-185.

14 Y. R. Hristova, M. M. J. Smulders, J. K. Clegg, B. Breiner and J. R. Nitschke, *Chem. Sci.*, 2011, **2**, 638-641.