Supplementary Information for
Computationally guided synthesis of a hierarchical [4[2+3]+6] porous organic 'cage of cages'

Qiang Zhu,[1,2] Hang Qu,[1] Gokay Avci,[3] Roohollah Hafizi,[4] Chengxi Zhao,[1,5] Graeme M. Day,[4] Kim E. Jelfs,[3] Marc A. Little*[1'6] and Andrew I. Cooper*[1,2]

[1] Materials Innovation Factory and Department of Chemistry, University of Liverpool, 51 Oxford Street, Liverpool, L7 3NY UK.

[2] Leverhulme Research Centre for Functional Materials Design, University of Liverpool, 51 Oxford Street, Liverpool, L7 3NY UK.

[3] Department of Chemistry, Imperial College London, Molecular Sciences Research Hub, White City Campus, Wood Lane, London, W12 0BZ, UK.

[4] Computational Systems Chemistry, School of Chemistry, University of Southampton, SO17 1BJ, UK

[5] Key Laboratory for Advanced Materials and Joint International Research Laboratory of Precision Chemistry and Molecular Engineering, Feringa Nobel Prize Scientist Joint Research Center, Frontiers Science Center for Materiobiology and Dynamic Chemistry, Institute of Fine Chemicals, School of Chemistry and Molecular Engineering East China University of Science and Technology, Shanghai 200237, China.

[6] Institute of Chemical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS UK.

e-mail: m.little@hw.ac.uk; aicooper@liverpool.ac.uk

Supporting Information

Section 1 Materials and methods

1.1 Materials

All reagents were obtained from Sigma-Aldrich, TCI Europe, Fisher, and Alfa Aesar and used as received. Anhydrous solvents were purchased from Acros Organics and used without further purification. All gases for sorption analysis were supplied by BOC at a purity of ≥99.9 %.

1.2 Methods

1.2.1 Molecular simulation

Both cage and cage-of-cages models were constructed in **Tri^{2}Di^{3}**, **Tri^{4}Di^{6}**, **Tri^{8}Di^{12}** topologies using the supramolecular toolkit (*stk*) software.[1] All cages were annealed with a molecular dynamics simulation at 700 K for 50 ns with a timestep of 0.5 fs after a 100 ps equilibration time with OPLS4 force field as implemented in the Macromodel Suite.[2] Five hundred random configurations from the total MD duration were sampled and energy minimised, with the lowest energy configuration selected for Density Functional Theory (DFT) calculations.

DFT calculations were performed with CP2K 2023.1[3] version software using the Generalised Gradient Approximation (GGA) theory with the PBE functional[4] and def2-TZVP basis sets[5]. A planewave cut-off value of 400 Ry and a relative cut-off value of 100 Ry were parameterised to obtain converged energy levels and dispersion interactions were accounted for with Grimme's DFT-D3 approach[6].

Structural properties of the cages, accessible surface area, non-accessible surface area, the largest cavity diameter, pore limiting diameter and porosity were calculated using the Zeo++ software[7]. A N_2 size probe was used to calculate the accessible surface area and a He-sized probe was used to calculate the porosity. Pore channels were visualised with Mercury software[8].

The geometries of the **[4[2+3]+6]cage** were then fully optimised by means of the hybrid M06-2X functional[9] in Gaussian16.[10] The def2-SVP basis set[11,12] was applied for all atoms. No symmetry or geometry constraint was imposed during optimisations. The

optimised geometries were verified as local minima on the potential energy surface by frequency computations at the same theoretical level.[10]

1.2.2 NMR

NMR spectra were recorded on a Bruker 400 NMR spectrometer at 400 MHz (^{1}H), 376 MHz (^{19}F) and 100 MHz (^{13}C) and referenced against the residual ^{1}H, ^{19}F or ^{13}C signal of the solvent.

1.2.3 MALDI-TOF MS

The molecular mass analysis was conducted on a Bruker Ultraflex MALDI TOF/TOF mass spectrometer using a reflection mode at Swansea University. Before analysis, we removed the reaction solvent by evaporation, the residual was suspended in acetone (1 mL), and the samples were analysed by MALDI-TOF using 2,5-dihydroxybenzoic acid as the matrix. A linear calibration method using two internal lock masses of the calibrant SphericalTM that bracketed the ions of interest (at m/z 2979 and m/z 3423) was used with m/z error limits of +/- 5ppm.

1.2.4 Powder X-ray diffraction (PXRD)

Laboratory powder X-ray diffraction (PXRD) data patterns were collected in transmission mode on samples held on thin Mylar film in aluminium well plates on a Panalytical Empyrean diffractometer, equipped with a high throughput screening (HTS) XYZ stage, X-ray focusing mirror, and PIXcel detector, using Cu-Kα radiation. For HT screening, PXRD patterns were measured over the 2θ range 1–40° in 0.013° steps over 10 minutes. Capillary PXRD patterns were collected on powdered samples loaded in borosilicate glass capillaries that were spun to improve averaging. Capillary PXRD patterns were collected on powdered samples loaded in borosilicate glass capillaries, and the capillaries were spun to improve averaging.

1.2.5 Single crystal X-ray diffraction (SC-XRD)

SCXRD data for **[4[2+3]+6]cage** was measured at beamline I19, Diamond Light Source, Didcot, UK using silicon double crystal monochromated synchrotron radiation (λ = 0.6889 Å, Pilatus 2M detector) and data reduction was performed using xia2. SCXRD data for **[4[2+3]+6]cage·acetone** was measured on a XtaLAB Synergy R, DW system rotating anode diffractometer (Cu-Kα radiation, λ = 1.54184 Å, 4-circle

goniometer, HyPix-Arc 100 detector), and data reduction was performed using CrysAlisPro. Structures were solved with SHELXT[13] and refined by full-matrix least squares on $|F|^2$ by SHELXL,[14] interfaced through the programme OLEX2.[15] All non-H atoms were refined anisotropically and all H-atoms were fixed in geometrically estimated positions and refined using the riding model. For full refinement details, see Tables S2 and S7.

1.2.6 Gas sorption Analysis

Surface areas were measured by nitrogen sorption at 77.3 K. Powder samples were degassed offline, followed by degassing on the analysis port under vacuum at 298 K for 15 hours. Isotherm measurements were performed using a Micromeritics ASAP 2020 characterisation analyser. CO_2, SF_6 and CH_4 were measured with the same procedure at 273 K and 298 K, using a Micromeritics ASAP.

1.2.7 Optical microscopy images

Optical microscopy images for **[4[2+3]+6]cage** and **4[2+3]+6]cage·acetone** were recorded using an Olympus BX53 Microscope with 10X objective lenses and an Olympus DP26 digital colour camera under a reflection model.

Section 2 Molecular simulations

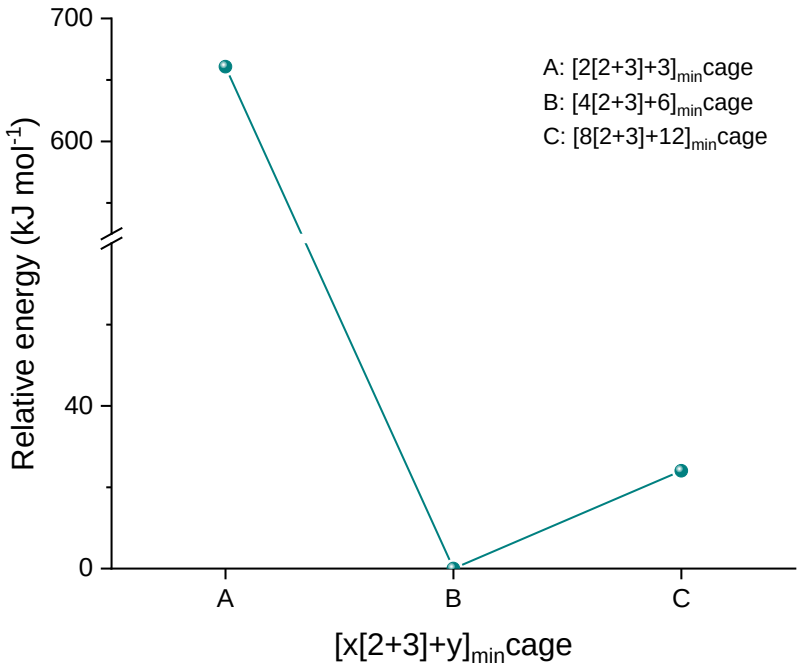


Figure S1 Relative DFT energies for the minimum energy configurations for the [x[2+3]+y]cages.

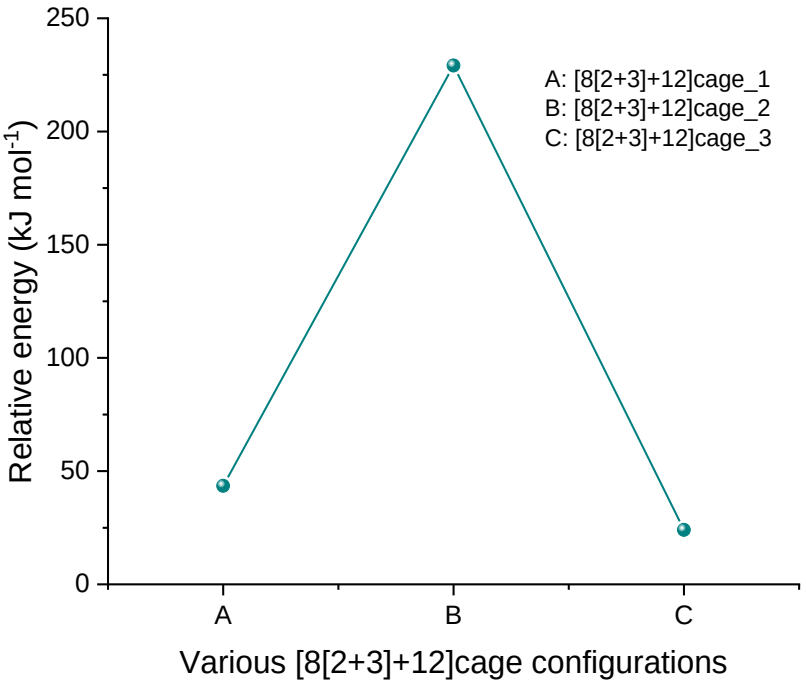


Figure S2 Relative energy values for randomly sampled ether *cis-trans* configurations for the [8[2+3]+12]cages.

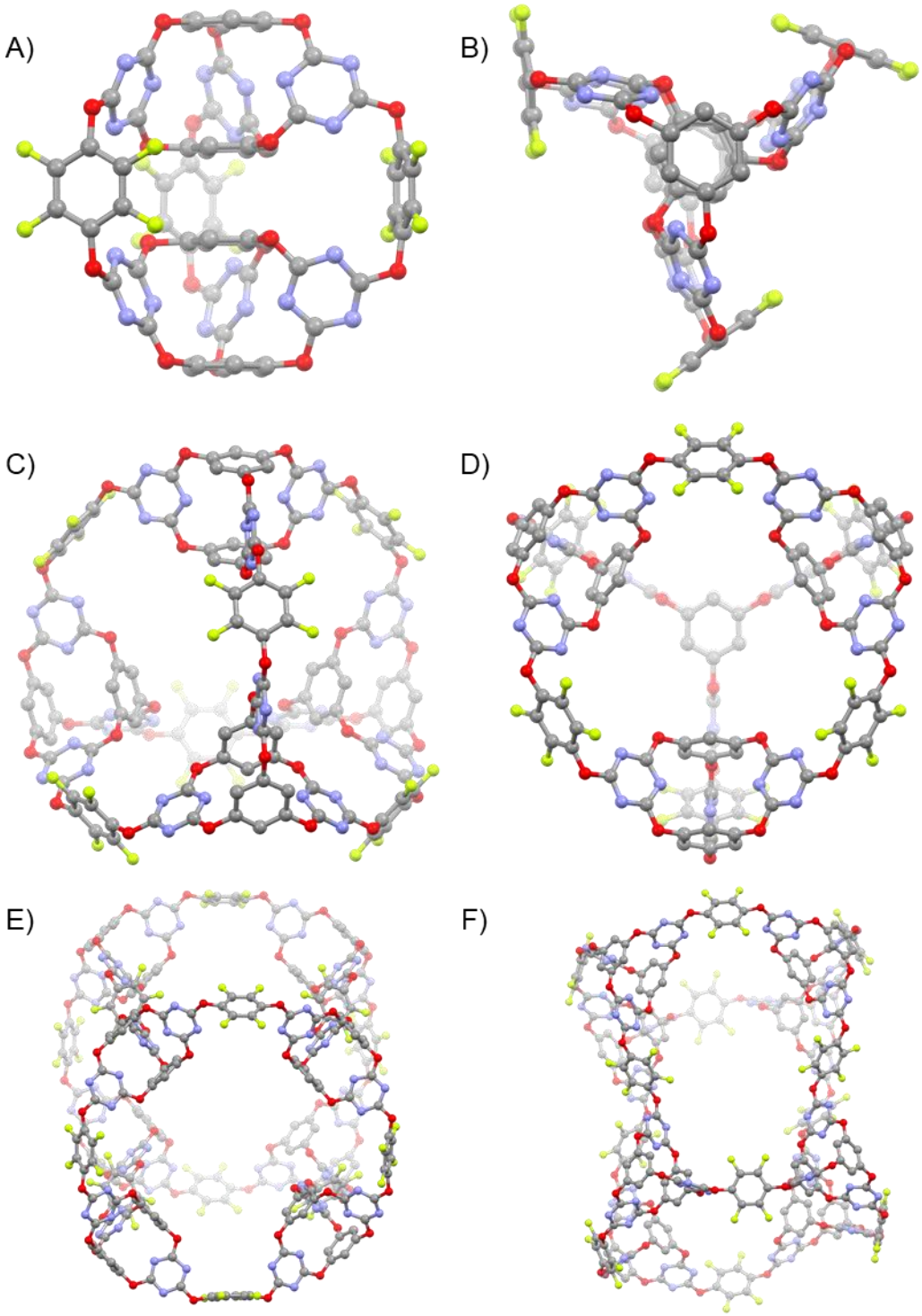


Figure S3 Structural representations of hypothetical **[2[3+2]+3]cage** (A, B), **[4[3+2]+6]cage** (C, D), and hypothetical **[8[3+2]+12]cage** (E, F). Atoms colour, C: grey; N: blue; O: red and F: green.

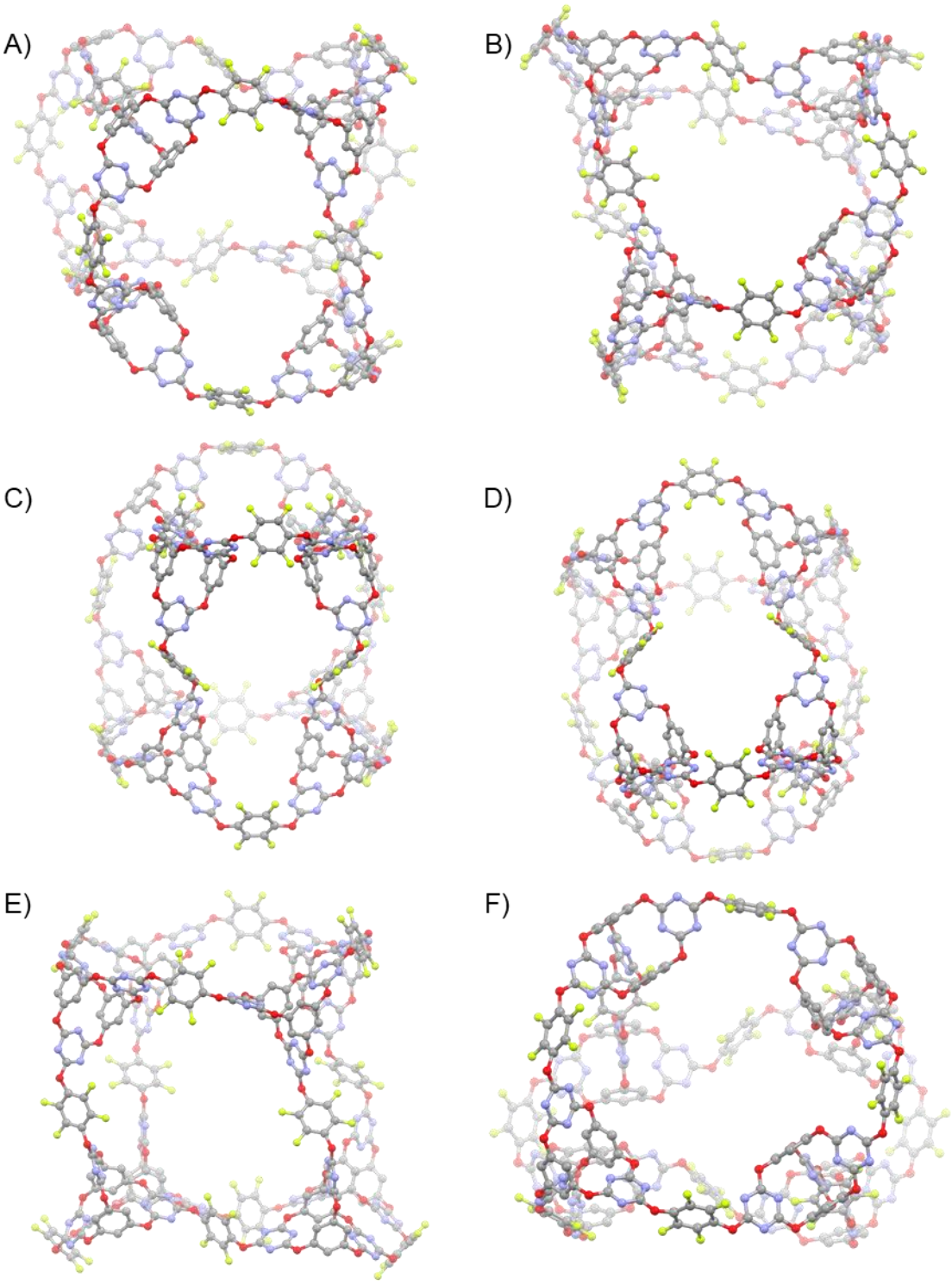
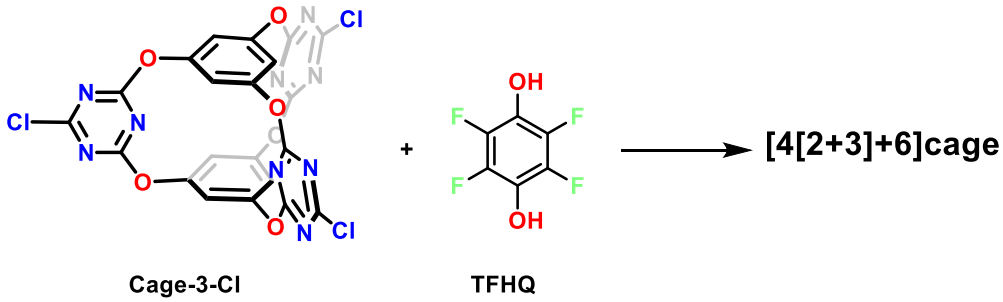


Figure S4 Structural conformer representations of the **[8[3+2]+12]cage_1** (A, B), **[8[3+2]+12]cage_2** (C, D), and **[8[3+2]+12]cage_3** retrieved from random snapshots within **[8[3+2]+12]cage** MD simulation trajectories (E, F). Atoms colour, C: grey; N: blue; O: red and F: green.

Section 3 Synthetic procedures

Synthesis of [4[2+3]+6]cage



Scheme S1 Synthetic route for **[4[2+3]+6]cage**. **Cage-3-Cl** was first synthesised according to previous methods.[16]

Table S1 Reaction conditions screened during the study[a]. All crude reaction products were analysed by MALDI-TOF and purified by column chromatography to calculate the product yield.

Entry	Solvent	Base	Conc. (mM)	Yield (%)
1[b]	acetone	DIPEA	3.2	53
2	CH_3CN	DIPEA	3.2	15
3	THF	DIPEA	3.2	21
4[c]	acetone	DIPEA	8.3	43
5[d]	acetone	DIPEA	1.6	33
6	acetone	TEA	3.2	42
7[e]	acetone	K_2CO_3	3.2	51

[a] Summary of the standard reaction conditions: Cage-3-Cl (0.05 mmol) and TFHQ (0.075 mmol) dissolved in 3 mL of solvent were added dropwise (3 mL h^{-1}) to a mixture of base (0.175 mmol) in the same solvent (12 mL) under N_2 at room temperature. [b] All the regent quantities were doubled compared to the standard reaction conditions. [c] DIPEA was dissolved in acetone (3 mL) before addition to the reaction. [d] Cage-3-Cl (0.1 mmol) and TFHQ (0.15 mmol) were dissolved in acetone (12.5 mL) before addition to the reaction. [e] The mixture was not applied to MALDI-TOF analysis.

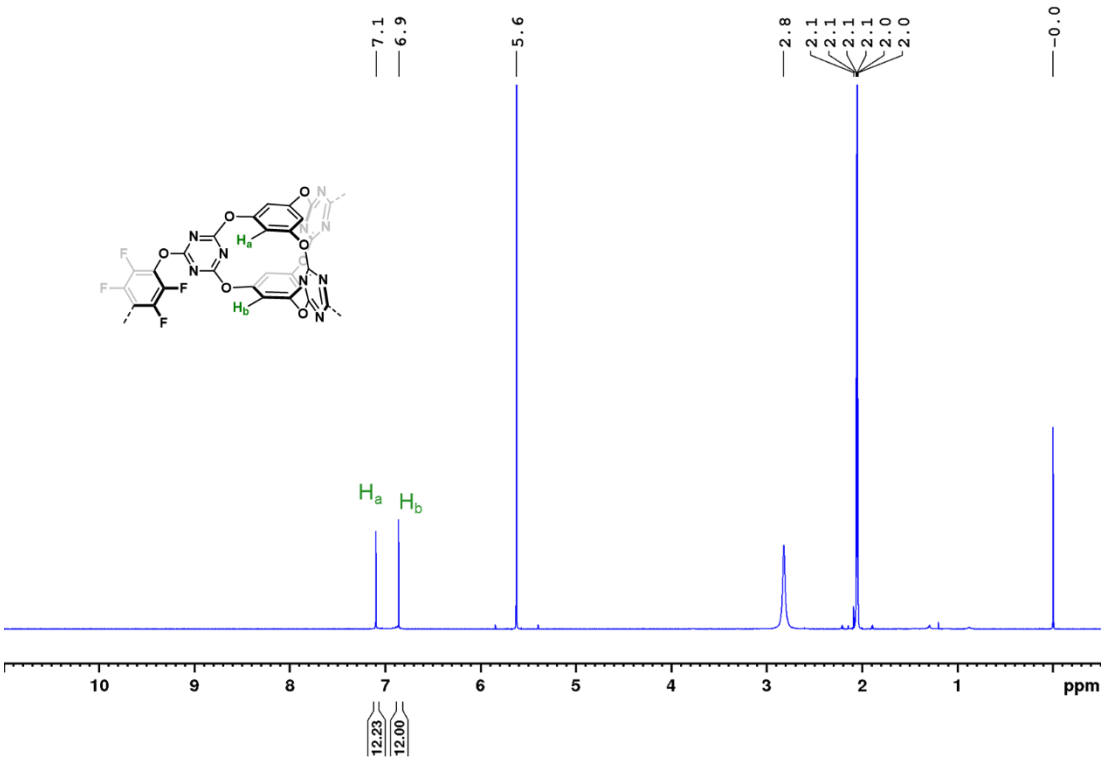


Figure S5 ^{1}H NMR (400 MHz, acetone-d_6) spectrum of **[4[2+3]+6]cage**.

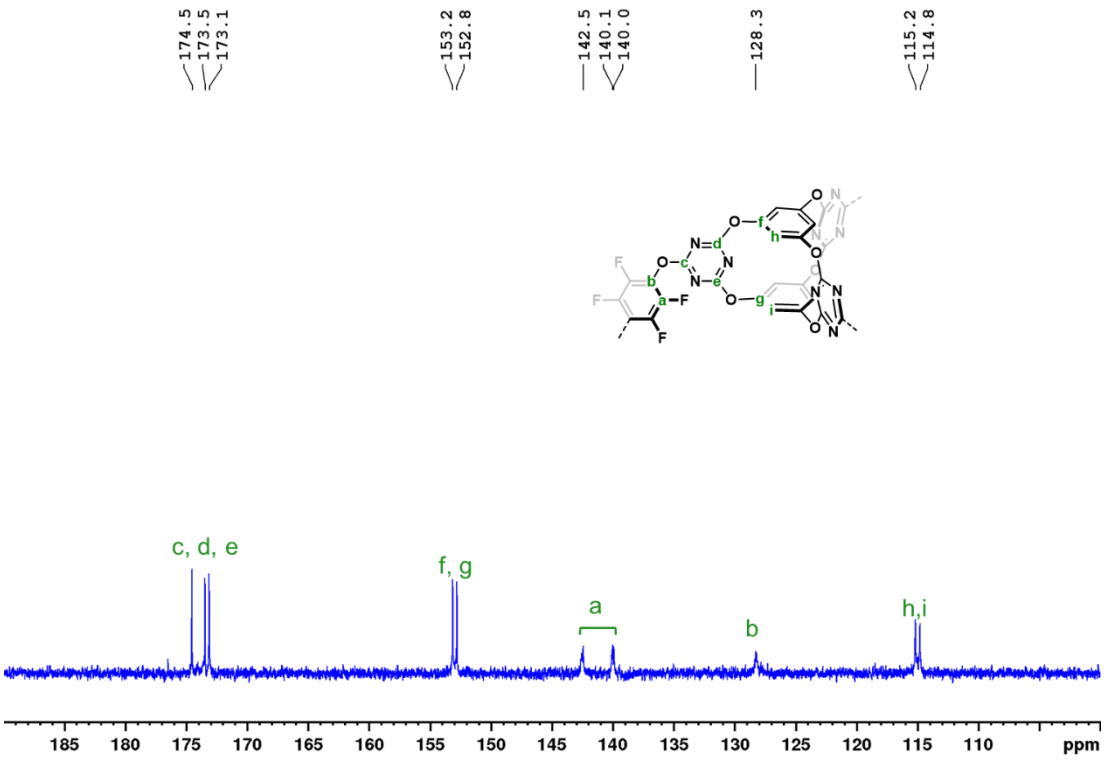


Figure S6 ^{13}C NMR (400 MHz, dioxane-d_8) spectrum of **[4[2+3]+6]cage**.

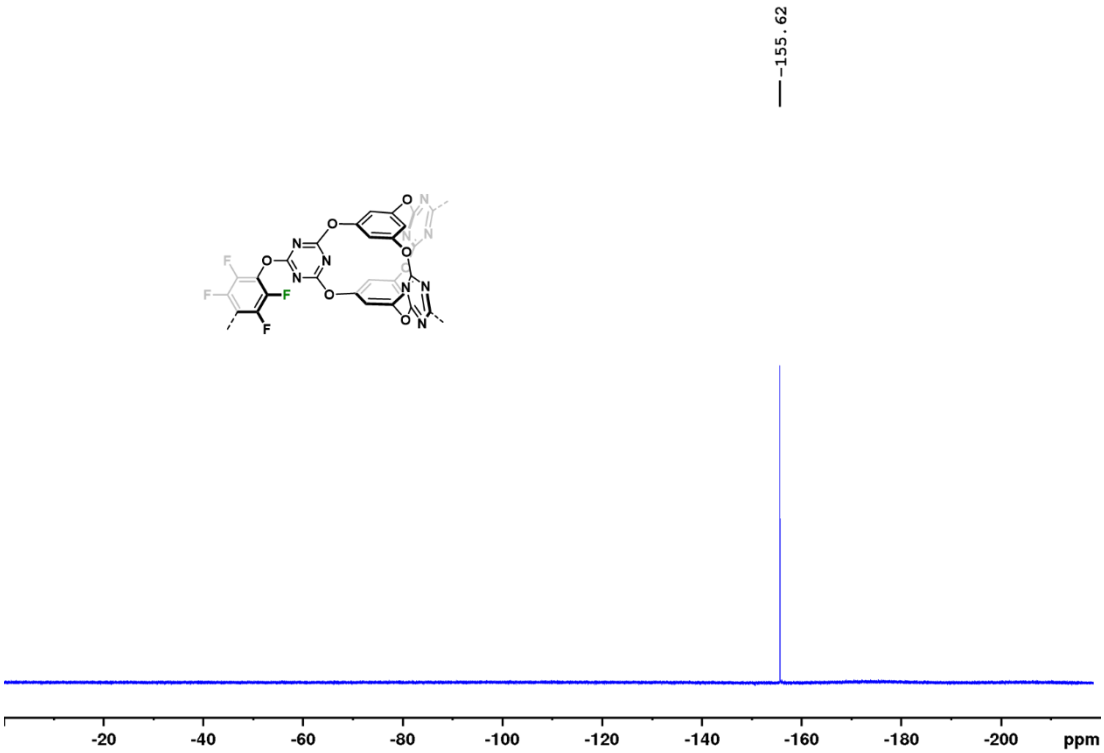


Figure S7 ^{19}F NMR (376 MHz, acetone-d_6) spectrum of **[4[2+3]+6]cage**.

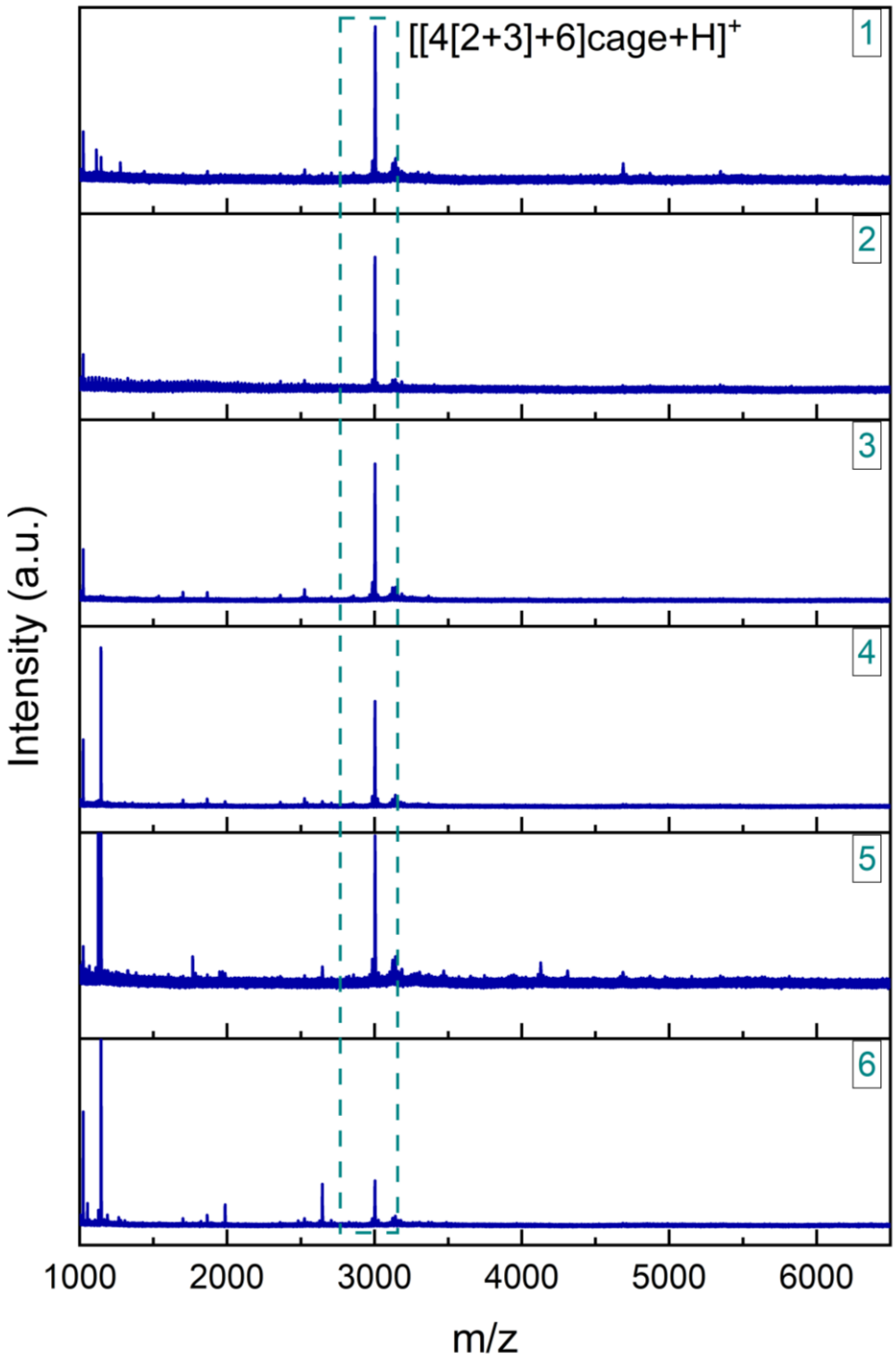


Figure S8 MALDI-TOF spectra of the resulting product from the reactions that used the experimental conditions and reagent quantities listed in Table S1. Before analysis, we removed the reaction solvent by evaporation, the residual was suspended in acetone (1 mL), and the samples were analysed by MALDI-TOF using 2,5-dihydroxybenzoic acid as the matrix. The teal dotted box insert highlights the peak were assigned to **[4[2+3]+6]cage**.

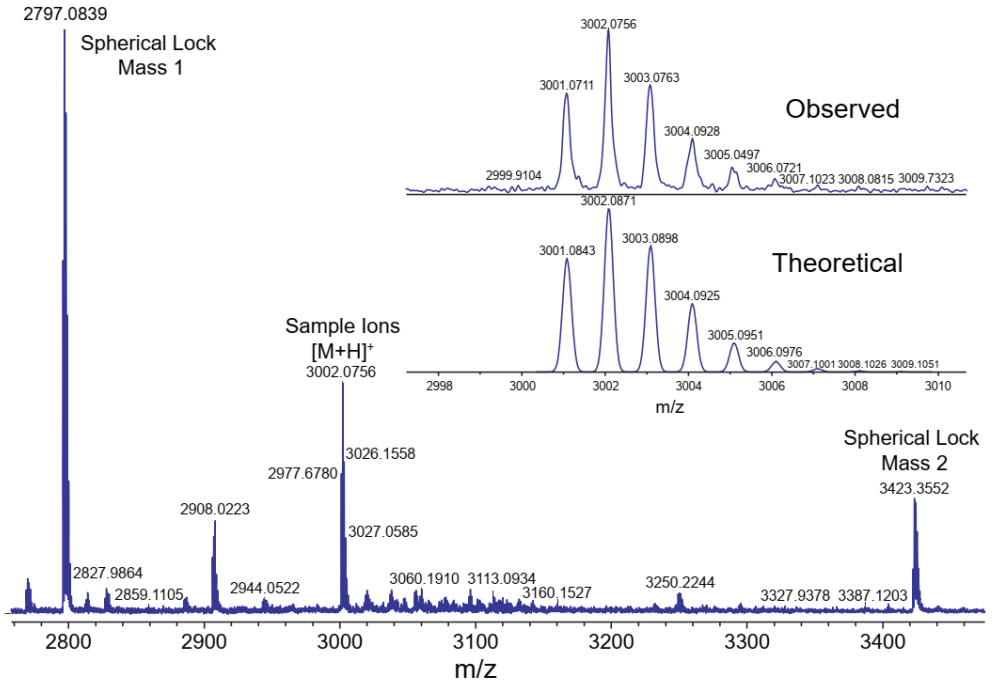


Figure S9 Further interpreted high-resolution MALDI-TOF spectrum of **[4[2+3]+6]cage**. A linear calibration method using two internal lock masses of the calibrant SphericalTM that bracketed the ions of interest (at m/z 2979 and m/z 3423) was used with m/z error limits of +/- 5ppm.

Crystallisation of [4[2+3]+6]cage

[4[2+3]+6]cage (4 mg) was dissolved in acetone (4 mL). The solution was filtered into a 14 mL vial through a 0.2 μm syringe filter and layered by EtOH (1 mL). The vial was covered with a cap, which has a small hole on the top, and placed at room temperature, affording rhombic crystals suitable for single crystal X-ray diffraction analysis.

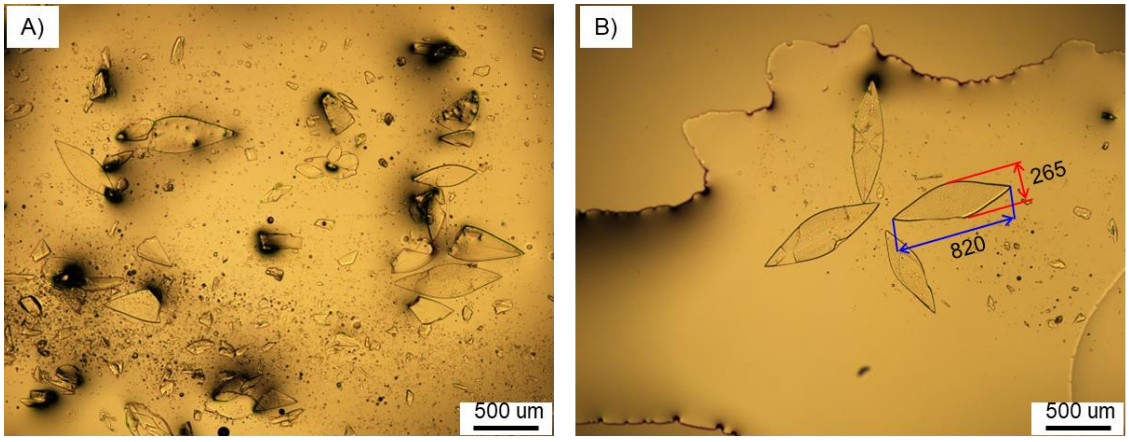


Figure S10 Optical images of the **[4[2+3]+6]cage** crystals. From these images, the size of one crystal was measured to be 820 × 265 × 20 μm.

Table S2 Single crystal refinement details for **[4[2+3]+6]cage**.

Name	**[4[2+3]+6]cage**
Crystallization solvent	Acetone-Ethanol
Wavelength/ Å	Synchrotron, 0.6889
Formula	$C_{120}H_{24}F_{24}N_{36}O_{36}$
Weight	3001.75
Crystal system	monoclinic
Space group	$P2_1$
***a* (Å)**	16.2851(7)
***b* (Å)**	37.1746(17)
***c* (Å)**	16.5847(10)
***β* (°)**	93.767(5)
***V* (Å^3)**	10018.6(9)
ρ_{calcd} (g cm^{-1})	0.995
Z	2
***T* (K)**	100
μ (mm^{-1})	0.084
***F*(000)**	3000
2θ range (°)	2.124 – 45.002
Reflections collected	38701
Independent reflections	24380
Observed data ($I > 2\sigma(I)$)	8932
Data / restraints / parameters	24380 / 2127 / 1633
Final R_1 values ($I > 2\sigma(I)$)	0.1041
Final R_1 values (all data)	0.1491
wR_2 (all data)	0.2895
Goodness-of-fit on $\|F\|^2$	1.088
CCDC	2303319

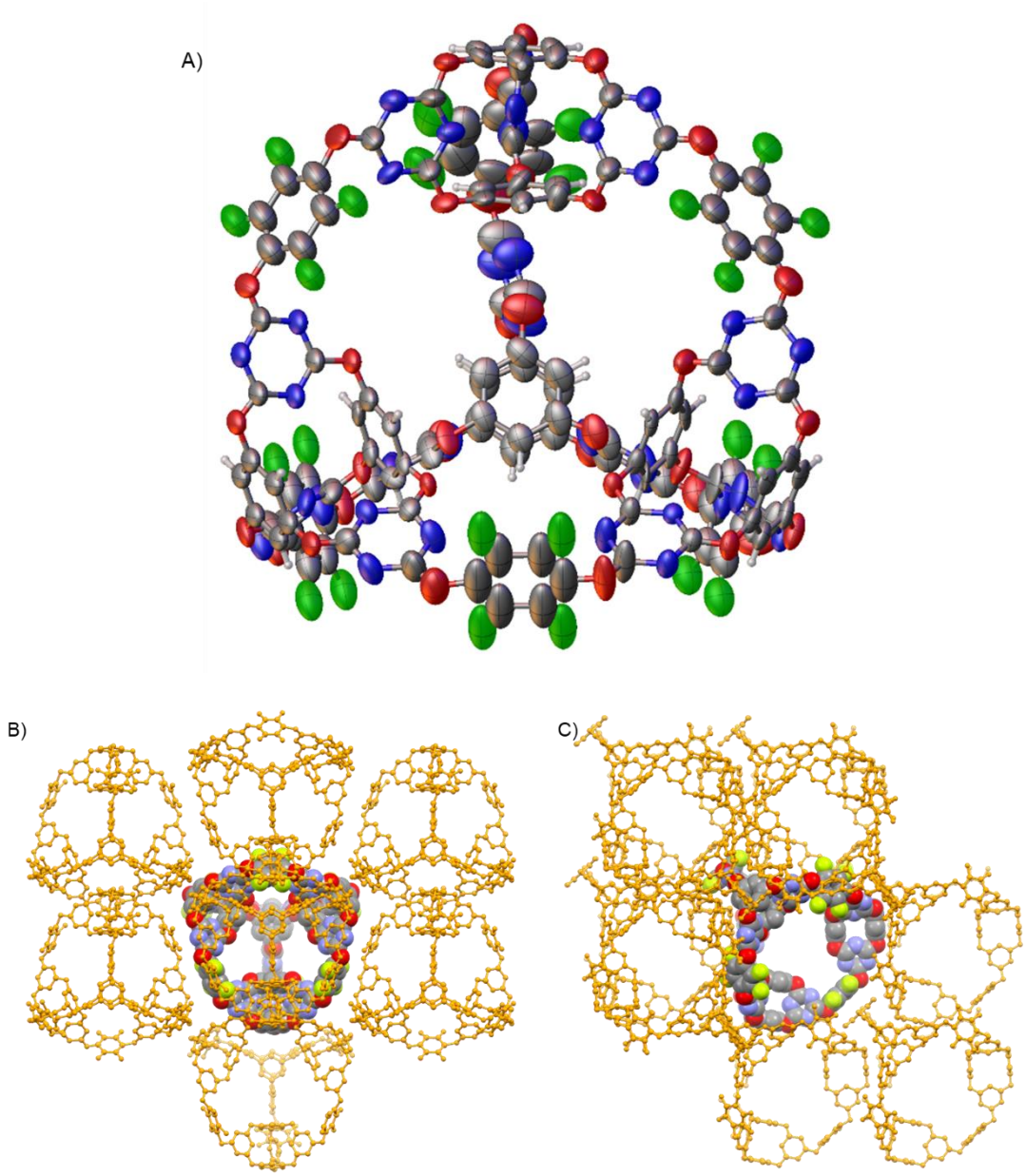


Figure S11 (A) Displacement ellipsoid plot from the single crystal structure of **[4[2+3]+6]cage** viewed along the *c*-axis. Ellipsoids are displayed at a 30% probability level due to disorder. Crystal packing of **[4[2+3]+6]cage** viewed along the *a*-axis (B) and the *b*-axis (C).

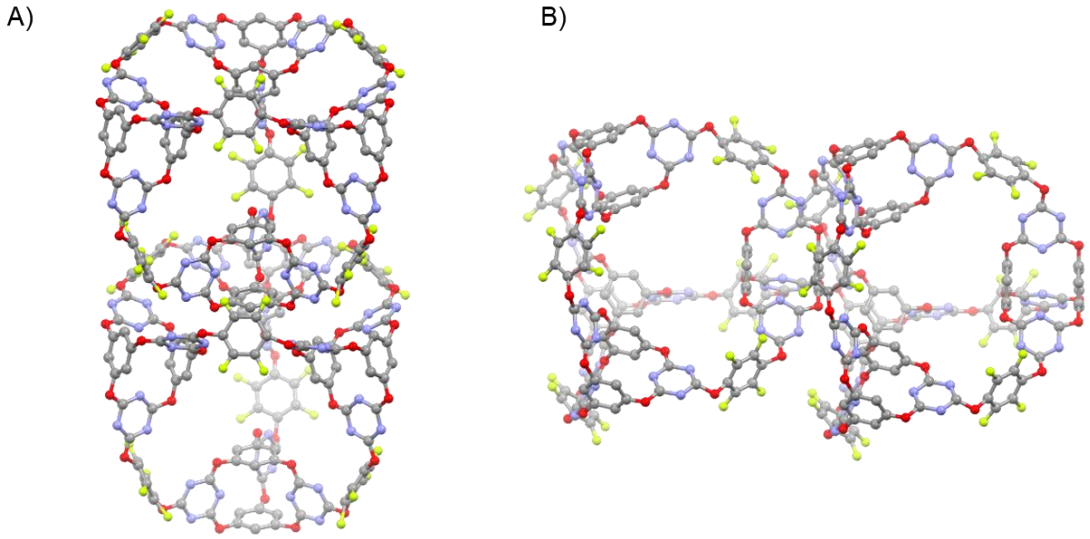


Figure S12 Illustration of the cage window splitting by neighbouring **[4[2+3]+6]cage** molecules, as viewed along the a-axis (A) and the b-axis (B).

Table S3 Summary of the predicted gas sorption properties of **[4[2+3]+6]cage** that were calculated using Zeo++[7] and the listed probe radii.

Probe Molecule	N_2	CO_2	-[a]
Probe Radius (Å)	1.87	1.65	0.5
Accessible Surface Area ($m^2\ g^{-1}$)	1359	1659	3266
Probe occupiable accessible volume ($Å^3$)	2935	3202	4448
Probe occupiable accessible volume fraction (%)	29.3	32.0	44.4

[a] The 0.5 A Radius is used as a hypothetical radius to probe the general coverage.

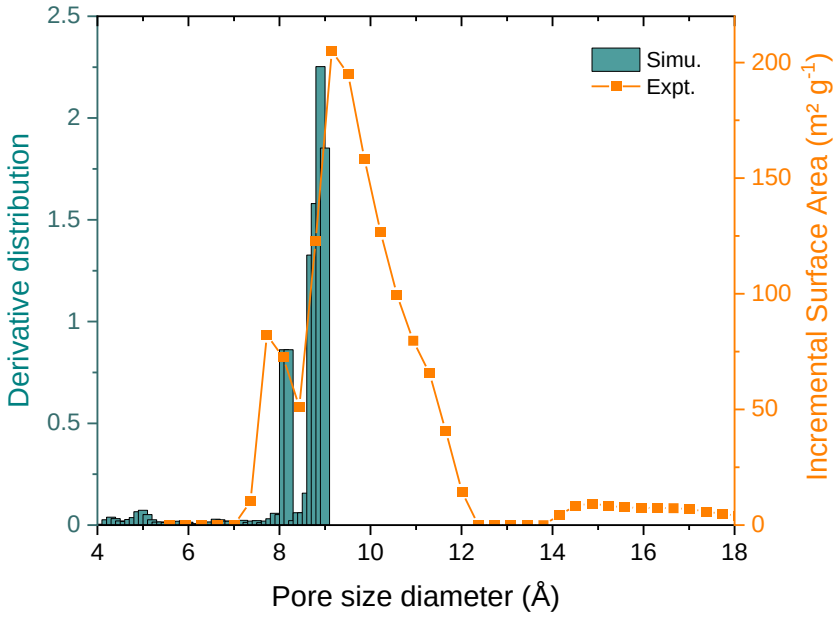


Figure S13 Plots showing the simulated pore size distribution calculated using Zeo++ (teal-coloured bars) and the experimental pore size distribution calculated from the N_2 adsorption isotherm (orange-coloured squares and line).

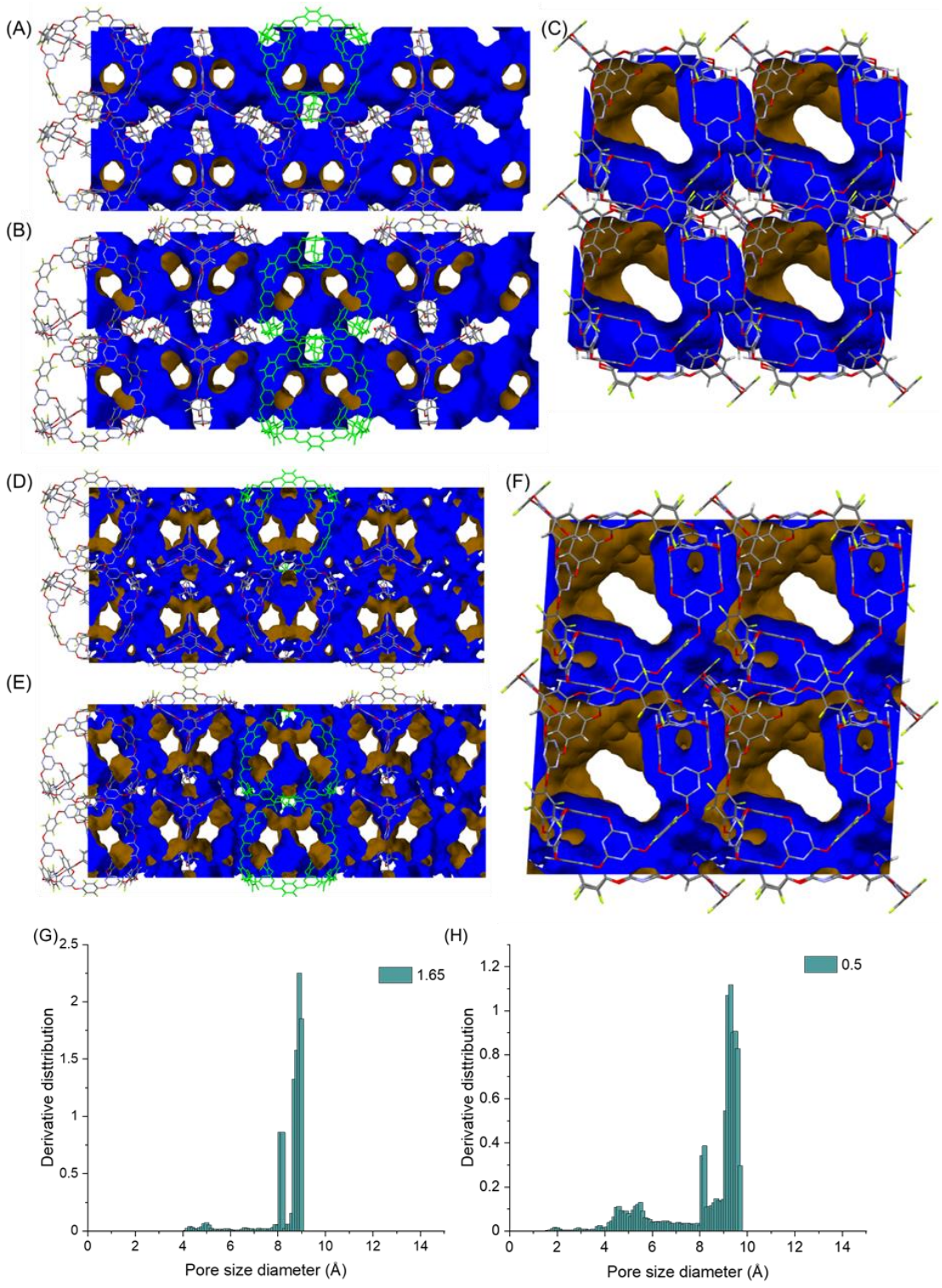


Figure S14 Porosity analysis of the **[4[2+3]+6]cage** single crystal structure using different probe radii. (A, B, and C) Contact surfaces using a 1.65 Å probe. (D, E, and F) Contact surfaces using a 0.5 Å probe. Inside contact surfaces: blue, outside contact surfaces: brown. Simulated pore size distributions using various probes (G) 1.65 Å and (H) 0.5 Å.

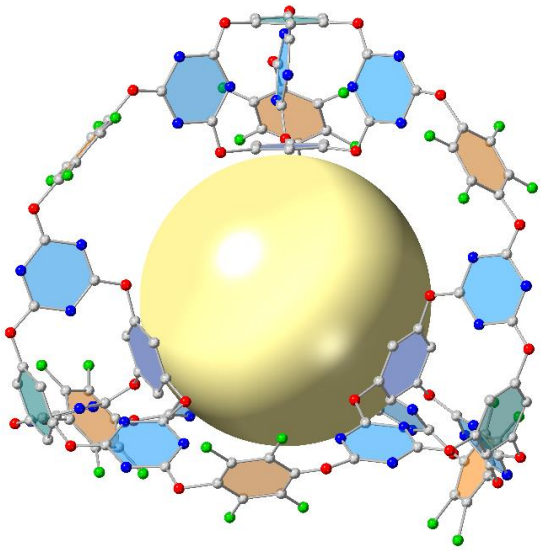


Figure S15 Illustration of the voids within the **[4[2+3]+6]cage** cavity. The yellow ball represents a sphere with a radius of 5.5 Å.

Section 4 Crystal structure prediction (CSP)

CSP involves the following general steps: (i) molecular geometry optimization; (ii) trial crystal structure generation; (iii) local lattice energy minimization of trial structures; and (iv) duplicate removal.

The geometry of the molecular cage was optimised at the B3LYP/6-311G(d,p) level using the Gaussian09 software[17], and the resulting geometry was kept fixed throughout the subsequent steps. Trial crystal structures are generated using the Global Lattice Energy Explorer (GLEE) code[18]. Subsequently, these trial structures undergo lattice optimisation while preserving the rigidity of the molecular cage. For this task, we employ an empirically parametrised intermolecular atom-atom exp-6 potential coupled with atomic multipole electrostatics. The force field parameters are acquired from the FIT force field[19,20]. Atom-centered multipoles up to hexadecapole are derived from the electron density through distributed multipole analysis (DMA), and partial charges are fitted to these multipoles[21,22]. The overall model is denoted as FIT+DMA.

The search for space groups involves sampling the 10 most common space groups for organic crystals along with the four trigonal space groups (143, 144, 145, and 146), each with one molecule in the asymmetric unit. A quasi-random method is employed to search these selected space groups separately, and valid structures are lattice energy minimised using the DMACRYS software[23] in a two-stage protocol. The first stage involves FIT+DMA with partial charges, followed by the second stage with multipole electrostatics. The search terminates after generating at least 10,000 valid structures in each space group. Table S4 presents the space groups included in the Crystal Structure Prediction (CSP) search and the corresponding number of structures sampled from each group. Post lattice energy minimisation, clustering of structures is performed by comparing Powder X-ray Diffraction (PXRD) patterns. This clustering aims to eliminate duplicates to evaluate the completeness of the search.

Table S4. List of space groups that were sampled in this study.

Space group name	Space group number	Number of valid structures to generate
P 1	1	10,000
P -1	2	10,000
$P2_12_12_1$	19	10,000
$Pna2_1$	33	10,000
$P2_1$	4	40,000
C2	5	10,000
Cc	9	10,000
Pbca	61	10,000
$P2_1/c$	14	10,000
C2/c	15	10,000
P3	143	10,000
$P3_1$	144	10,000
$P3_2$	145	10,000
R3	146	10,000

The energy landscape of the **[4[2+3]+6]cage** is shown in Figure S16, where the experimental **[4[2+3]+6]cage** crystal structure and the solvent-removed **[4[2+3]+6]cage·acetone** structure are represented by a yellow star and a blue cross, respectively. The identification of these structures involved a geometric overlay of 30 molecules from both the experimental and predicted structures using COMPACK[24]. An overlay of the predicted crystal structure and the original **experimental [4[2+3]+6]cage** structure is shown in Figure S17. Despite exhibiting very similar geometries, the **experimental [4[2+3]+6]cage** structure is slightly away from the minimum on the potential energy surface of the distributed multipole analysis-based force field that was used in the crystal structure prediction. After rigid-body geometry optimization, all the 30 molecules demonstrated an exceptional alignment, achieving an RMSD in atomic positions of less than 0.05 Å. The stable synthesized structure exhibits an energy elevation of 210 kJ mol^{-1} relative to the global minimum, while the other structure shows a higher energy level, 300 kJ mol^{-1} above the global minimum. Details of computational study of this structure are provided in section S7. Interestingly, periodic

DFT calculations validate the ranking of these structures, prompting further inquiry into the substantial energy differences. Upon visually inspecting the ten lowest-energy structures, we found that all these structures exhibit catenation. This underscores the significance of exploring catenation within the predicted crystal structures generated by CSP.

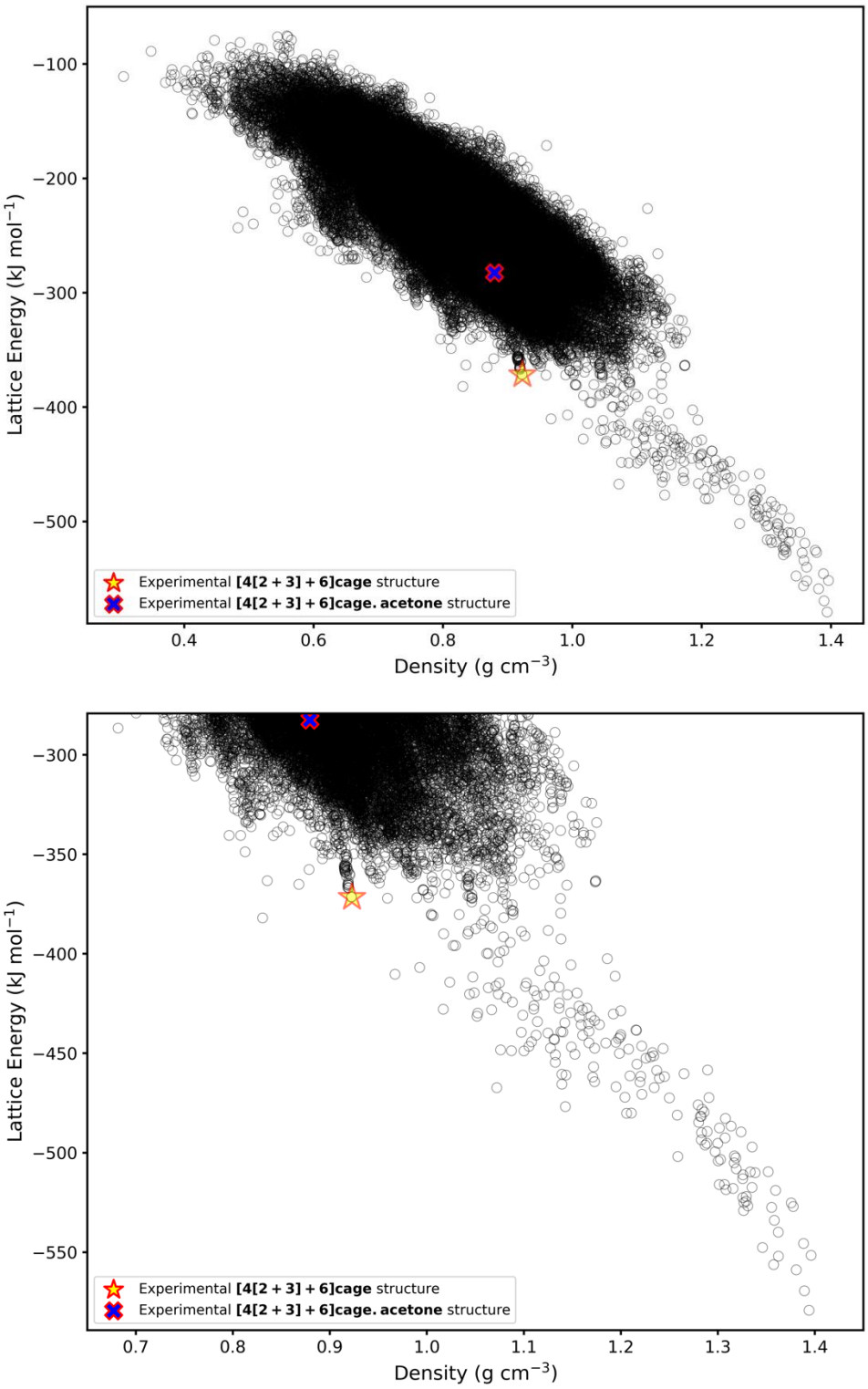


Figure S16 Initial (top) full and (bottom) low energy part of energy landscape of **[4[2+3]+6]cage**.

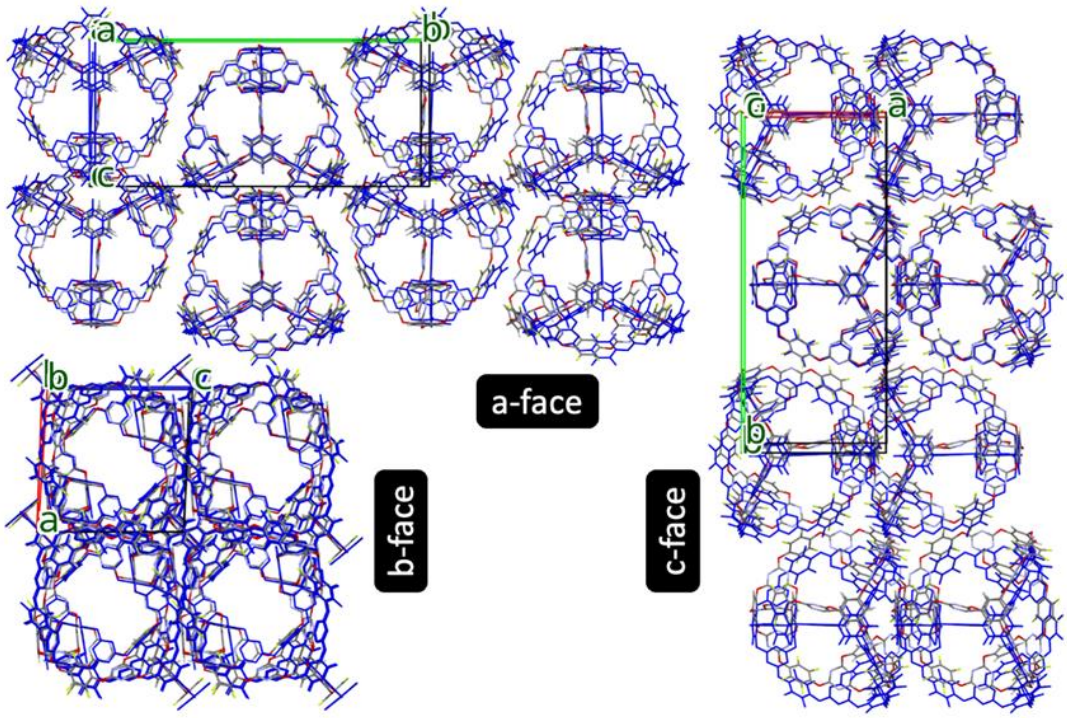


Figure S17 Overlay of the predicted crystal structure (blue, lattice energy minimised using the exp-6 force field + atomic multipoles model) and the original **experimental [4[2+3]+6]cage** crystal structure (atoms coloured by element). Top, bottom left, and right panels show views down the lattice directions perpendicular to lattice cell vectors **a**, **b**, and **c**, respectively.

To detect catenated structures, we leveraged the highly symmetric and rigid structure of the **[4[2+3]+6]cage**, and employed the distance between centroids of molecular cages as a primary criterion for catenation analysis. The methodology, depicted schematically in Figure S18, outlines the steps involved in calculating the minimum centroid distance: (1) Each cage molecule within the unit cell is projected onto its centroid; (2) the unit cell is standardized; (3) the minimum periodic image is determined in all directions, and finally; (4) the smallest distance between centroids is computed for all the centroids within the central unit cell. This systematic approach provides a foundation for evaluating catenation in the context of the **[4[2+3]+6**]cage's symmetric and rigid molecular structure.

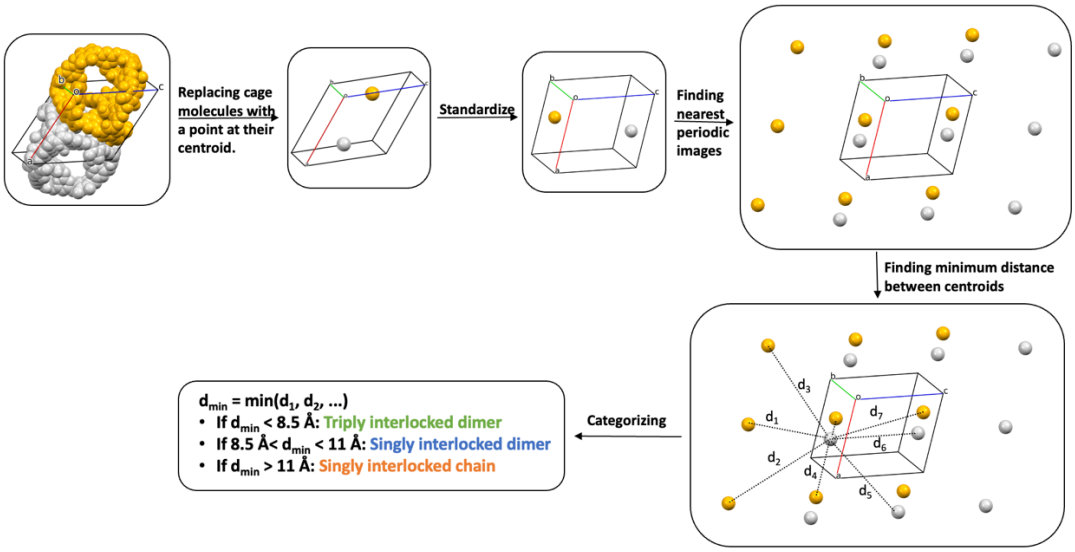


Figure S18 Schematic of catenation detection procedure

In Figure S19, the energy landscape is depicted with a color-coded representation based on the minimum distance between centroids of cages. We found that catenation occurs when the distance between centroids falls below an empirically determined threshold of 14.2 Å, offering a robust indicator for identifying catenation regions within the crystal structure. The analysis based on centroid distances unveiled three distinct types of catenated structures: triply interlocked dimers, characterized by a centroid distance of less than 8.5 Å (Figure S20b), singly interlocked dimers with a centroid distance between 11 and 14.2 Å (Figure S20c), and a singly interlocked 1-D chain of cages exhibiting centroid distances between 8.5 and 11 Å (Figure S20d) were observed. No exceptions to these catenation criteria were noted in our study. Figure S20a colour-codes the landscape based on the type of catenation, providing a visual representation of the diverse catenated structures present in the CSP landscape. While our approach has proven successful for detecting catenation for this rigid cage of cages, we acknowledge alternative criteria might be needed where symmetry and rigidity are absent.

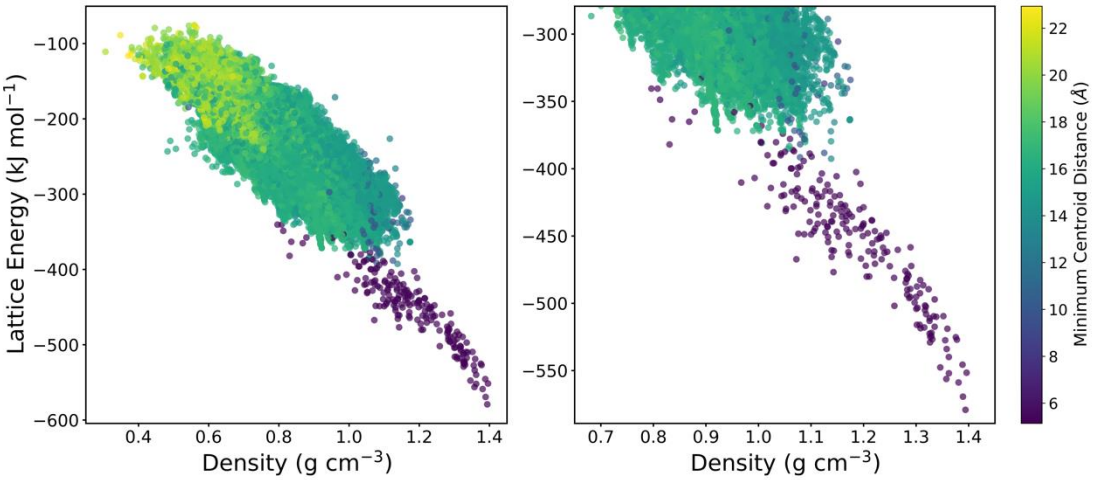


Figure S19 (left) Full and (right) low energy part of energy landscape color-coded by the minimum distance between centroids of cages in the crystal structure.

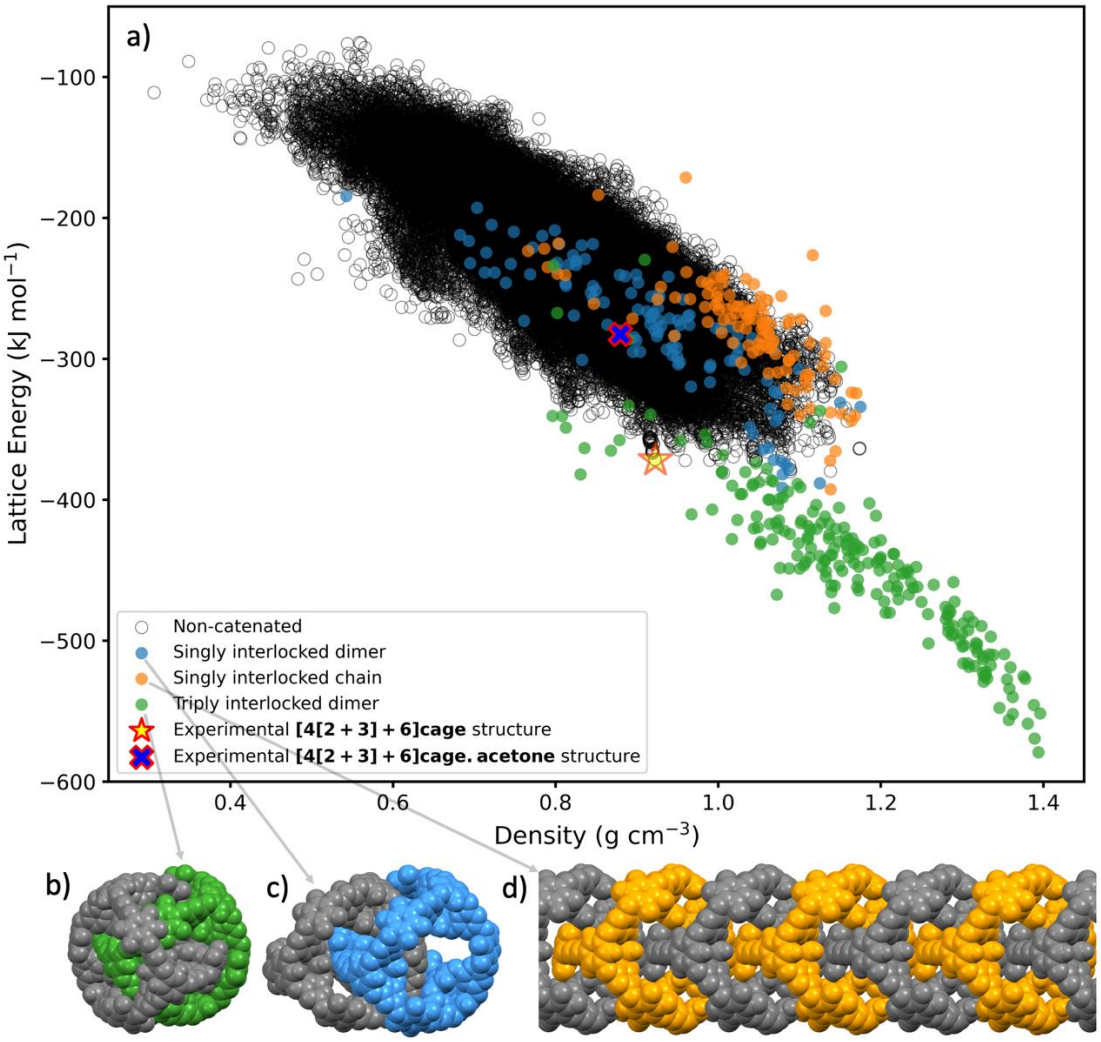


Figure S20 a) Energy landscape color-coded by structures' type of catenation, along with atomic structure of b) triply interlocked dimer, c) singly interlocked dimer, and d) singly interlocked chain.

To validate the energy preferences of catenated structures through periodic DFT calculations, we selected 21 structures from the landscape to cover catenated, non-catenated, and experimental structures. This set comprised 10 lowest lying structures on the landscape (all of which contain triply interlocked catenated dimers), two structures from the lowest lying singly interlocked dimers and chains, five low-energy non-catenated structures, and experimental structures. The positions of these selected structures on the energy landscape are illustrated in Figure S21. Periodic density functional theory, as implemented in the VASP software package[25–29], was used for these calculations. The PBE exchange correlation functional, complemented by Grimme's D3(BJ) dispersion correction[30], was employed. All calculations utilized the projector augmented wave method with standard supplied pseudopotentials. A plane wave cut-off of 500 eV was employed across all calculations, and K-point sampling was carried out using a regular K-point mesh with a K-spacing of at least 0.31415 $Å^{-1}$.

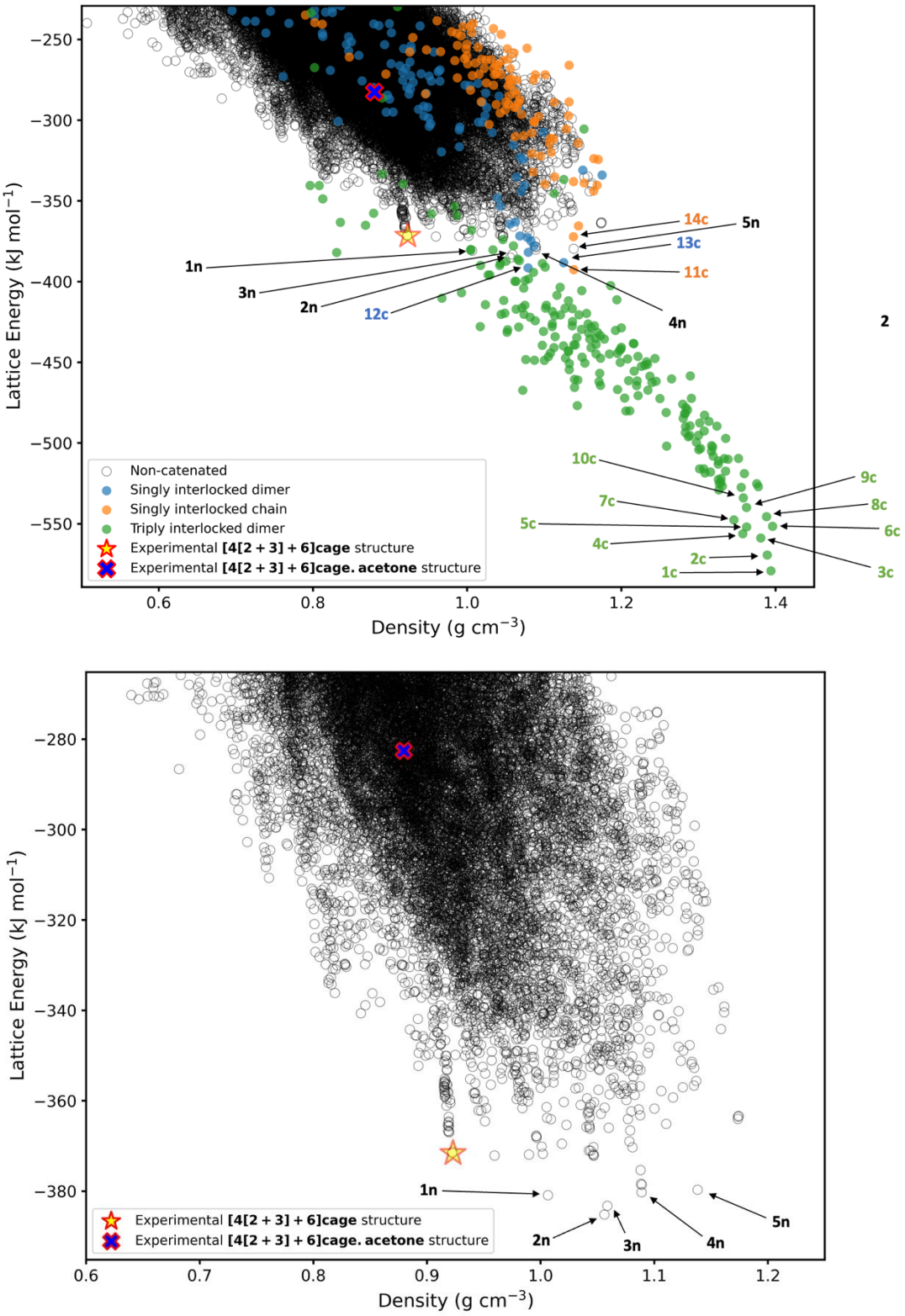


Figure S21 Selected structures for periodic DFT calculations are depicted in the top panel, while the bottom panel illustrates the position of non-catenated selected structures within the 'non-catenated' subset of the energy landscape.

In Table S5, we present the energy differences between selected structures and the global minimum for both FIT+DMA and periodic DFT calculations. It is noteworthy that the global minimum remains consistent for both levels of theory. In FIT+DMA calculations, the experimental **[4[2+3]+6]cage** structure exhibits energy elevations of 210.8, 21.0, and 20.1 kJ mol^{-1} relative to the most stable catenated structure with triply interlocked dimers, singly interlocked dimers, and singly interlocked chains, respectively. The same ranking is observed in periodic DFT calculations, with corresponding values of 132.1, 19.5, and 11.6 kJ mol^{-1}, respectively. This consistent ranking across both calculation methods clearly shows the robustness of the predicted energy landscape in FIT+DMA calculations and confirms the thermodynamic favourability of catenated structures.

Table S5 Energy of chosen structure in Figure S21 at FIT+DMA and periodic DFT levels of theory

ID	Space group	Density (g cm^{-3})	Relative FIT+DMA (kJ mol^{-1})	Relative VASP (kJ mol^{-1})
1c	2	1.3988	0.0	0
2c	2	1.3941	3.2	2.7
3c	2	1.3893	13.0	7.1
4c	2	1.3809	23.5	10.4
5c	2	1.3576	26.1	12.6
6c	15	1.3625	30.4	13.7
7c	2	1.3962	30.8	20.0
8c	2	1.3461	34.7	16.3
9c	2	1.3884	36.8	18.6
10c	2	1.3624	42.4	30.0
11c	19	1.1384	189.8	112.6
12c	15	1.0791	190.9	120.5
13c	15	1.1252	194.1	126.6
14c	61	1.1379	210.2	125.3
1n	15	1.0065	200.5	118.0
2n	14	1.0562	197.2	109.8
3n	14	1.0588	199.2	111.8
4n	33	1.0889	202.2	131.3
5n	14	1.1383	202.7	101.4
Experimental [4[2+3]+6]cage Structure	4	0.9224	210.8	132.1
Experimental [4[2+3]+6]cage·acetone structure	4	0.8796	299.8	197.0

To investigate the occurrence of catenation within specific space groups, we have color-coded catenated structures on the landscape by their respective space groups in Figure S22. This representation delineates triply interlocked dimers, singly interlocked dimers, and singly interlocked chains with circles, triangles, and stars, respectively. While a

majority of the lowest lying structures are observed in space group 2 (P-1), no discernible pattern emerges for triply interlocked dimers and singly interlocked chains. Structures featuring singly interlocked dimers appear mostly in space groups 5 (C2) and 15 (C2/c). We stress that our sampling methodology did not specifically target catenated structures; rather, they emerged from the predictions. It might also be possible to discover even lower-energy catenated structures through more focused sampling.

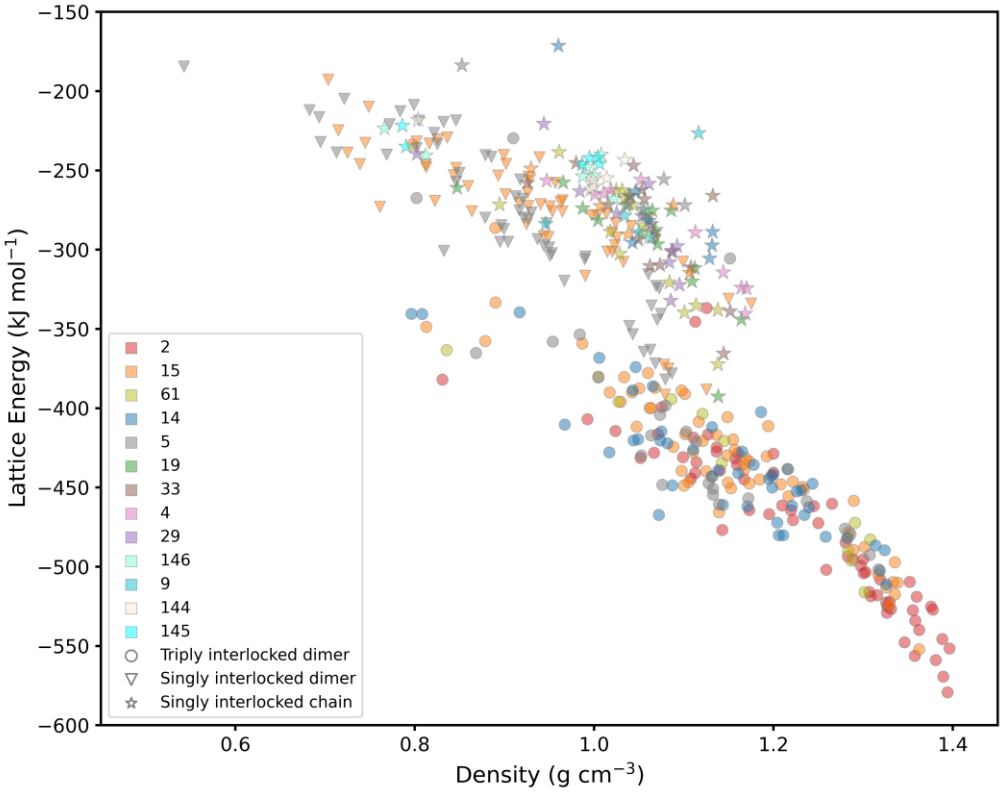


Figure S22 Catenated subset of energy landscape color-coded by their space group. Circles, triangles, and stars show triply interlocked dimers, singly interlocked dimers, and singly interlocked chains, respectively.

Despite catenated structures dominating the bottom 197 kJ mol^{-1} energy range in the lattice energy landscape, these seem to be kinetically inaccessible by experiment. We therefore also excluded catenated structures from the conformational search landscape. As depicted in Figure S23, following the removal of catenated structures, the synthesized structure resides at the bottom of one of the 'spikes'.

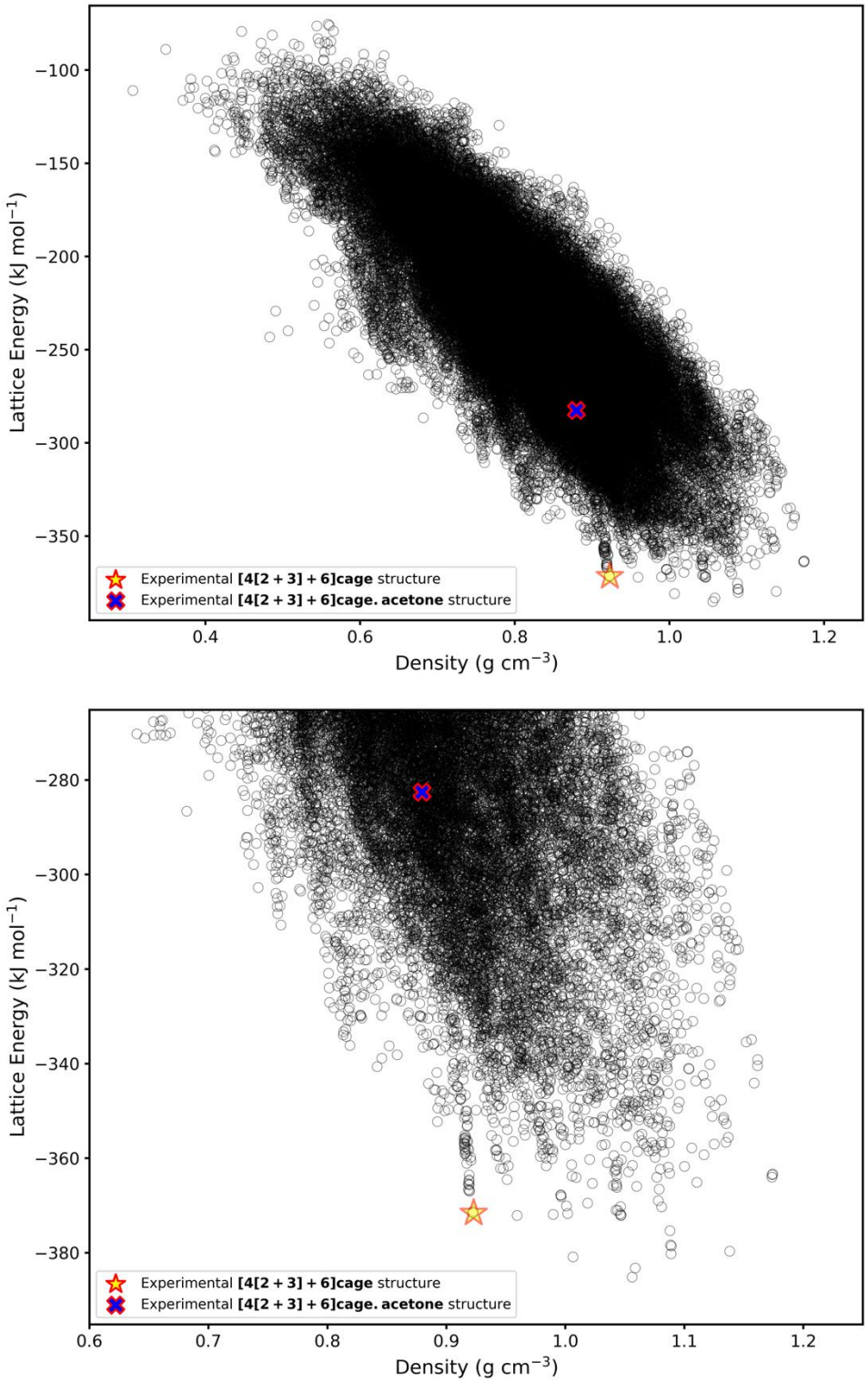


Figure S23 Post-catenated removal (top) full and (bottom) low energy part of energy landscape of **[4[2+3]+6]cage**.

In Figure S24, we visualize the most porous looking facet of the structures identified in the bottom panel of Figure S21, all of which were characterized by a lower porosity level than the experimental structure. Zeo++[7] was employed to find channels within the

predicted structures capable of accommodating a CO_2 molecule, considering its kinetic radius as 1.65 Å (Table S3). We focused on structures within 20 kJ mol^{-1} from the low-energy edge of the landscape, and analysed the maximum dimension of these channels and the diameter (D_f) of the largest sphere capable of free movement within them. The left panel of Figure S24 colour-codes the "edge band" structures based on their maximum channel dimensions, where zero signifies structures devoid of channels suitable for a CO_2 molecule. Furthermore, the right panel of Figure S24 displays the landscape color-coded by D_f. Notably, in the low-energy region, it becomes evident that the experimental **[4[2+3]+6]cage** structure is the lowest-lying structure with both 3D channels and a pore size exceeding 6 Å, reinforcing its distinct energetic and structural characteristics within this context.

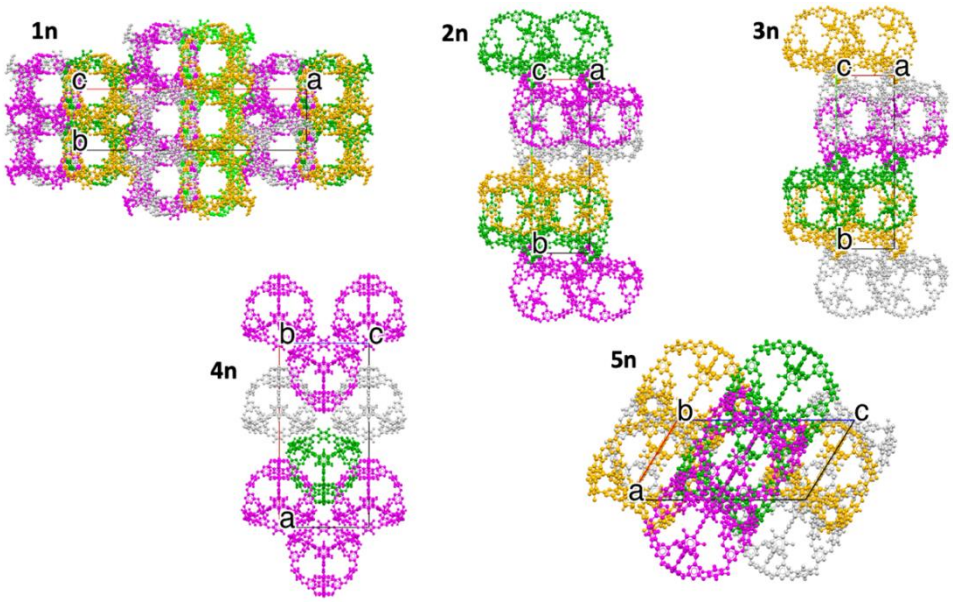


Figure S24 Crystal structure of predicted low energy structures. Cage molecules are color-coded by their symmetry operation. The most porous looking face of each structure has been presented.

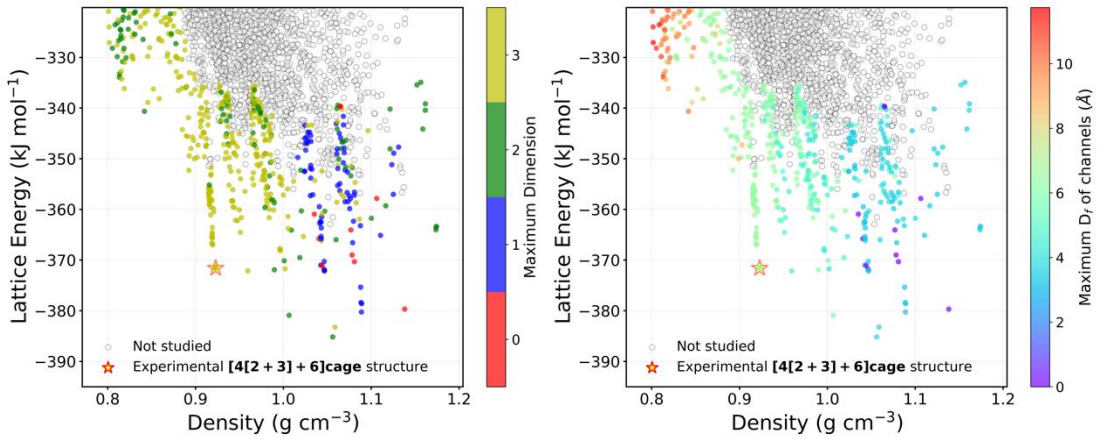


Figure S25 (left) Maximum dimension of channels capable of accommodating a CO_2 molecule and (right) the diameter (d_f) of the largest sphere capable of free movement within them.

Section 5 Gas sorption analysis

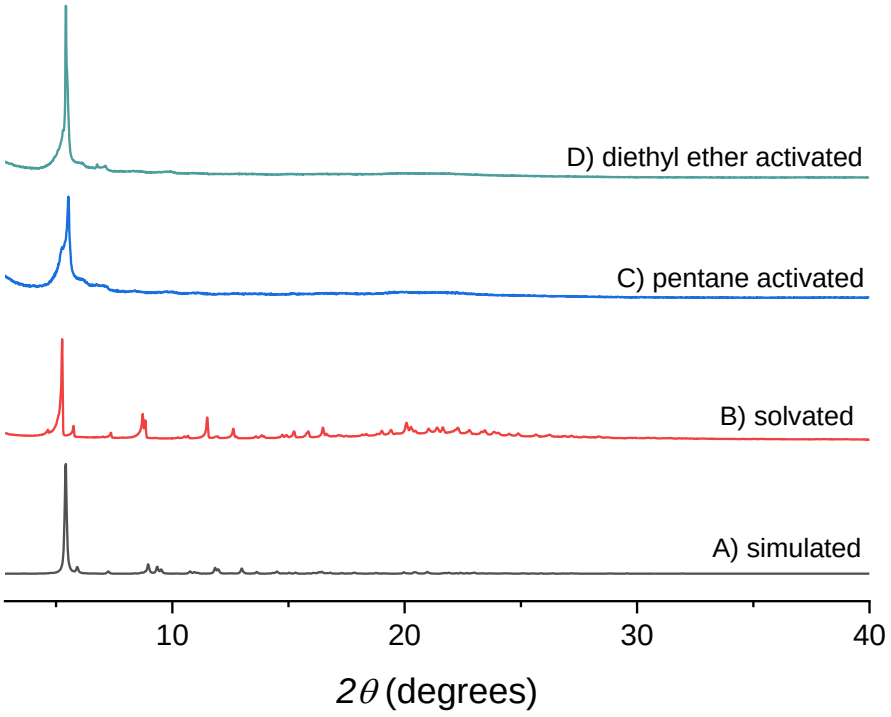


Figure S26 Comparison of the PXRD patterns of **[4[2+3]+6]cage** crystals. (A) simulated from the solvated single crystal structure of **[4[2+3]+6]cage**. (B) solvated **[4[2+3]+6]cage** crystals immersed in the crystallisation solvents (EtOH/acetone). (C) **[4[2+3]+6]cage** crystals after solvent exchange with n-pentane and activation under dynamic vacuum. (D) **[4[2+3]+6]cage** crystals after solvent exchange with diethyl ether and activation under dynamic vacuum.

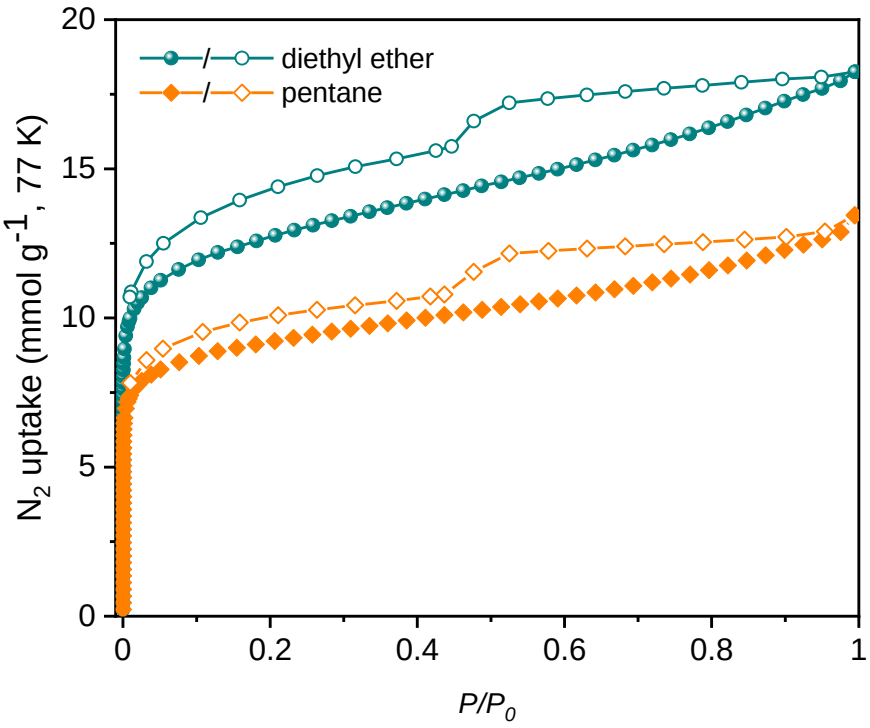


Figure S27 N_2 sorption isotherms of **[4[2+3]+6]cage** recorded at 77 K. The isotherms were recorded using **[4[2+3]+6]cage** samples that were evacuated under dynamic vacuum after exchanging the acetone and ethanol crystallisation solvents with n-pentane (square symbols) or diethyl ether (circle symbols).

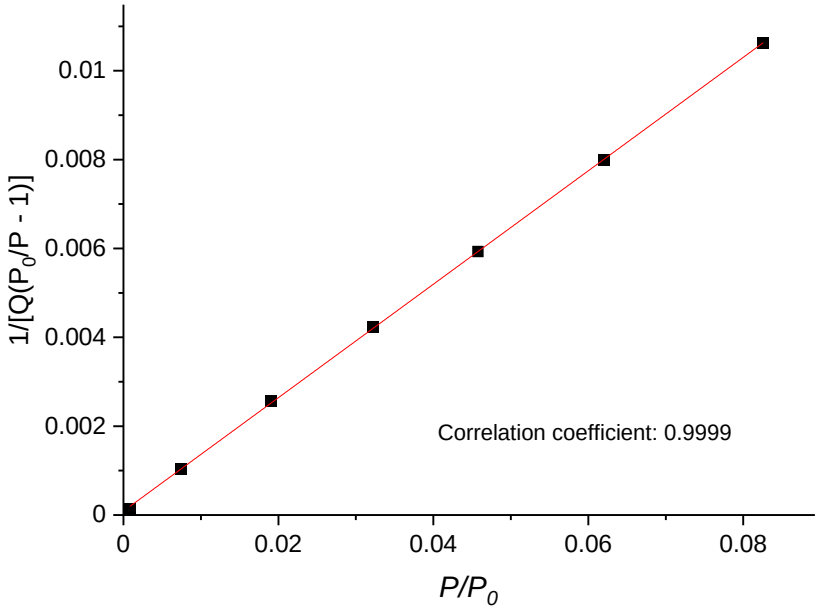


Figure S28 t-Plot for BET surface area calculation. The sample was activated by the solvent exchange with pentane ten times prior to evacuation and sorption analysis.

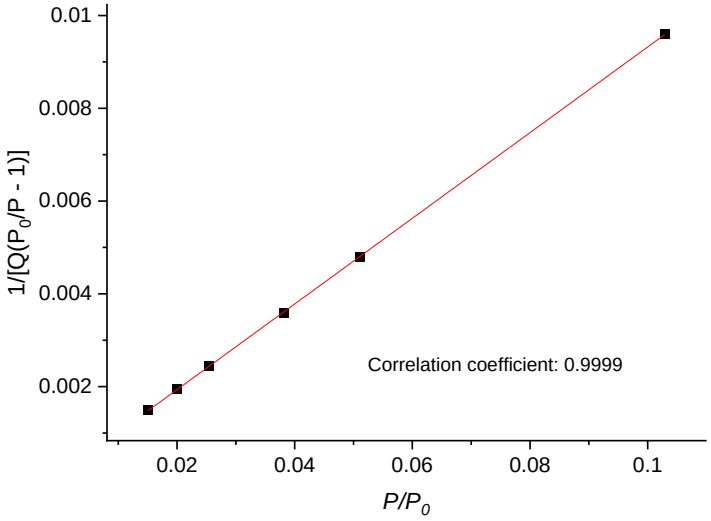


Figure S29 t-Plot for BET surface area calculation. The sample was activated by the solvent exchange with diethyl ether ten times prior to evacuation and sorption analysis.

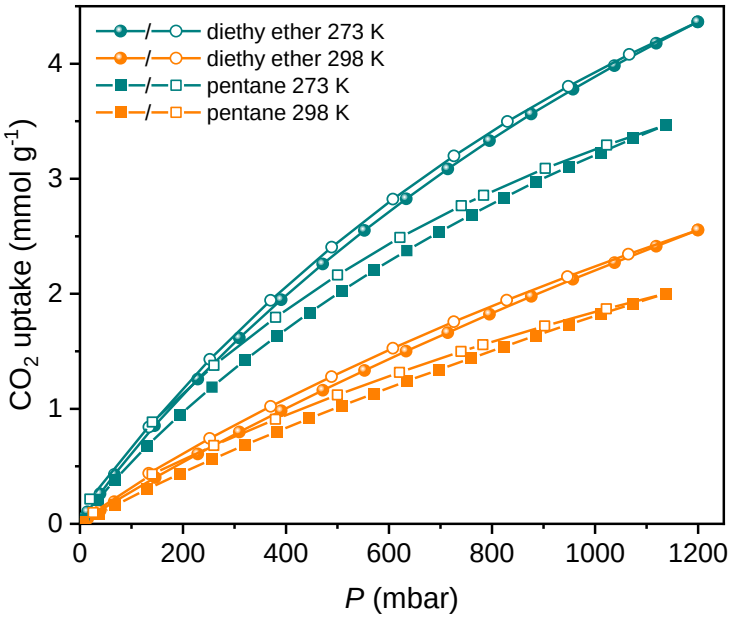


Figure S30 CO_2 sorption isotherms at 273 K (orange) and 298 K (cyan). The isotherms were recorded using **[4[2+3]+6]cage** samples that were evacuated under dynamic vacuum after exchanging the acetone and ethanol crystallisation solvents with n-pentane (square symbols) or diethyl ether (circle symbols). Adsorption points are shown using closed symbols; desorption points are shown using open symbols.

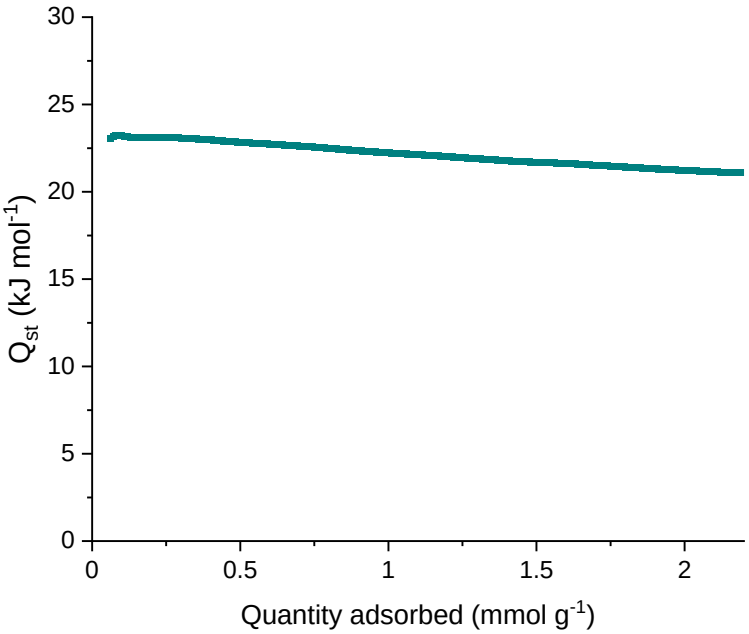


Figure S31 Isosteric heat of adsorption of CO_2 on crystalline **[4[2+3]+6]cage**, calculated using the Clausius–Clapeyron equation and the CO_2 adsorption isotherms of **[4[2+3]+6]cage** collected at 273 K and 298 K.

Table S6 Summary of CO_2 uptakes in various porous organic cages.

Entry	POCs	linkage	SA_{BET} (m^2 g^{-1})	CO_2 uptake (mmol g^{-1})[a]	References
1	cage 2	imine to quinoline	698	1.15	[31]
2	CC15α	imine	2.7	1.30	[32]
3	Fe-PB	imine	409[b]	1.33	[33]
4	CC14α	imine	320	1.57	[32]
5	PB-1	imine	1370	1.66	[34]
7	CC3-S/CC15-R cocrystal	imine	13	1.84	[32]
8	3-Et-Et	imine	71	1.87	[35]
9	BTPOC	imine	605[b]	1.96	[36]
10	CC3α	imine	409	2.01	[32]
10	CC3-RS	imine	598	2.01	[37]
12	cage 9	amide	102	2.05	[38]
12	cage 7	amide	398	2.05	[38]
14	cage 8	amide	260	2.07	[38]
15	cage 2	imine	1375	2.10	[39]
16	cage 4	amide	275	2.14	[38]
17	CC3-RScore/CC19-RSshell – 3 μm	imine	539	2.26	[37]
18	CC19-RS	imine	500	2.39	[37]
19	CC3	imine	624	2.47	[40]
20	CC19-RScore/CC3-RSshell – 3 μm	imine	523	2.55	[37]
21	cage 1	ether	432	2.8	[41]
22	CC2	imine	533	3.00	[40]
23	3-Et-H	imine	443	3.14	[35]
24	PB-2	imine	935[b]	3.31	[34]
25	**[4[2+3]+6]cage**	**ether**	**1056**	**3.98**	**This work**
26	FC1	imine	536	4.20	[42]

[a] All CO_2 uptake data here were collected at 273 K, 1 bar.

[b] Apparent BET surface area was calculated based on the CO_2 sorption isotherm at 195 K, rather than a nitrogen isotherm at 77 K.

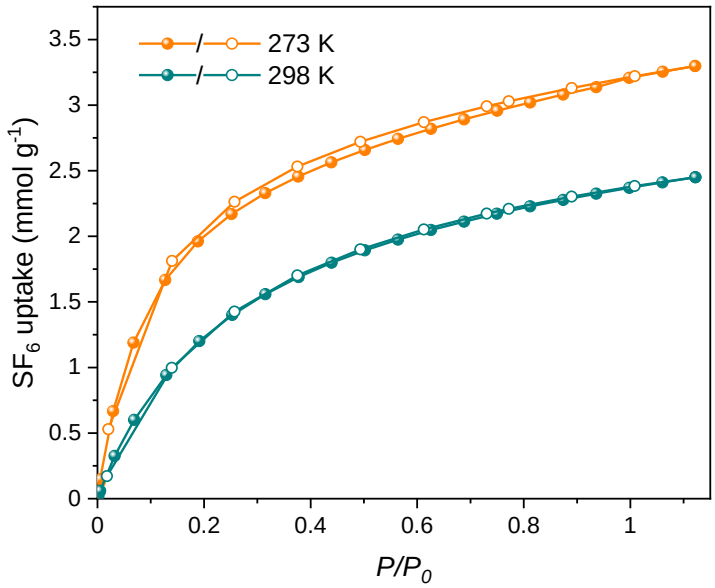


Figure S32 SF_6 sorption isotherms of **[4[2+3]+6]cage** recorded at 273 K (orange) and 298 K (cyan). The isotherms were recorded using a **[4[2+3]+6]cage** sample that was evacuated under dynamic vacuum after exchanging the acetone and ethanol crystallisation solvents with diethyl ether. Adsorption points are shown using closed symbols; desorption points are shown using open symbols.

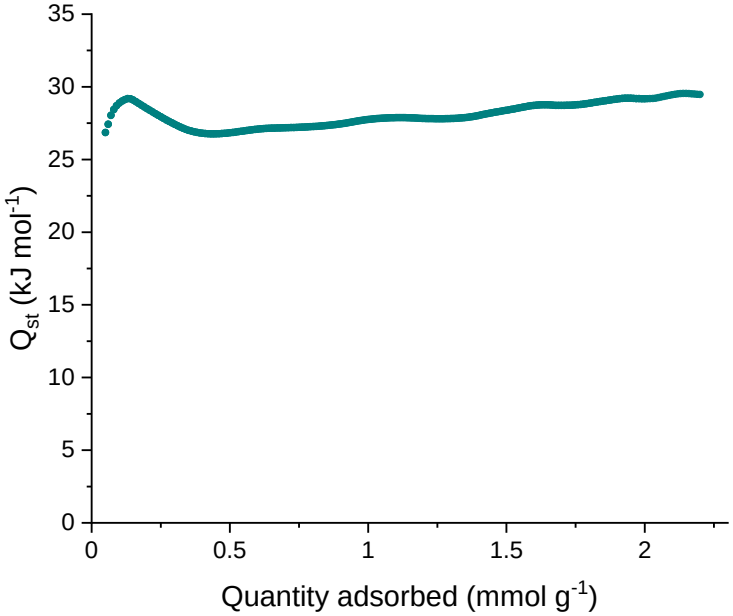


Figure S33 Isosteric heat of adsorption of SF_6 on crystalline **[4[2+3]+6]cage**, calculated using the Clausius–Clapeyron equation and the SF_6 adsorption isotherms of **[4[2+3]+6]cage** collected at 273 K and 298 K.

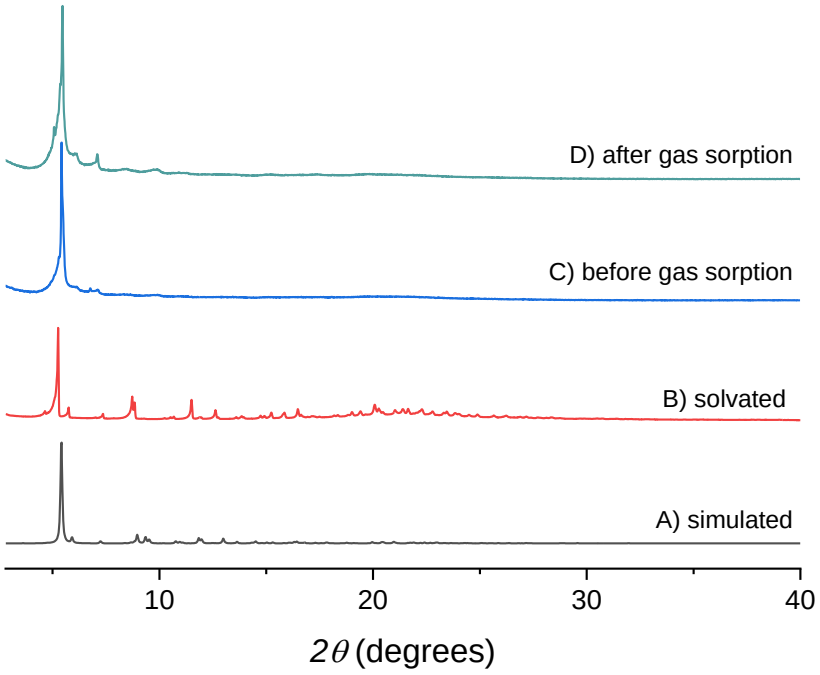


Figure S34 Comparison of the PXRD patterns of **[4[2+3]+6]cage** crystals. (A) simulated from the solvated single crystal structure of [**4[2+3]+6]cage**. (B) solvated **[4[2+3]+6]cage** crystals immersed in the crystallisation solvents (EtOH/acetone). **[4[2+3]+6]cage** crystals (C) before and (D) after gas sorption. The sample was activated by the solvent exchange with diethyl ether and activation under dynamic vacuum.

Section 6 Additional crystal structure information

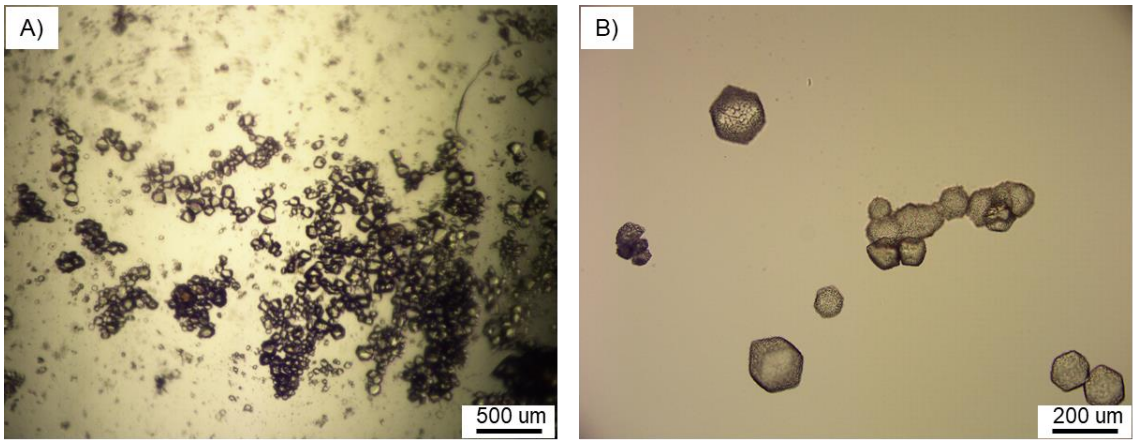


Figure S35 Crystallisation of **[4[2+3]+6]cage·acetone.** Optical microscope images of **[4[2+3]+6]cage·acetone** crystals captured (A) in acetone-d_6 and (B) after being dropped to the paraffin oil within one minute. To grow the crystals, **[4[2+3]+6]cage** (4 mg) was dissolved in acetone-d_6 (4 mL). The vial was covered with a cap, which has a small hole on the top, and placed at room temperature, affording hexagonal crystals suitable for single crystal X-ray diffraction analysis.

Table S7 Single crystal refinement details for **[4[2+3]+6]cage·acetone**.

Name	**[4[2+3]+6]cage·acetone**
Crystallization solvent	Acetone-d_6
Wavelength/ Å	Cu-Kα (λ = 1.54184)
Formula	$C_{120}H_{24}F_{24}N_{36}O_{36}$
Weight	3001.75
Crystal system	cubic
Space group	$I\bar{4}3m$
***a* (Å)**	23.2901(15)
***V* (Å^3)**	12633(2)
ρ_{calcd} (g cm^{-1})	0.789
Z	2
***T* (K)**	210
μ (mm^{-1})	0.632
***F*(000)**	3000
2θ range (°)	5.37–164.47
Reflections collected	33715
Independent reflections	2492
Observed data ($I > 2\sigma(I)$)	1202
Data / restraints / parameters	2492 / 148 / 106
Final R_1 values ($I > 2\sigma(I)$)	0.0426
Final R_1 values (all data)	0.0712
wR_2 (all data)	0.1584
Goodness-of-fit on $\|F\|^2$	0.930
CCDC	2326368

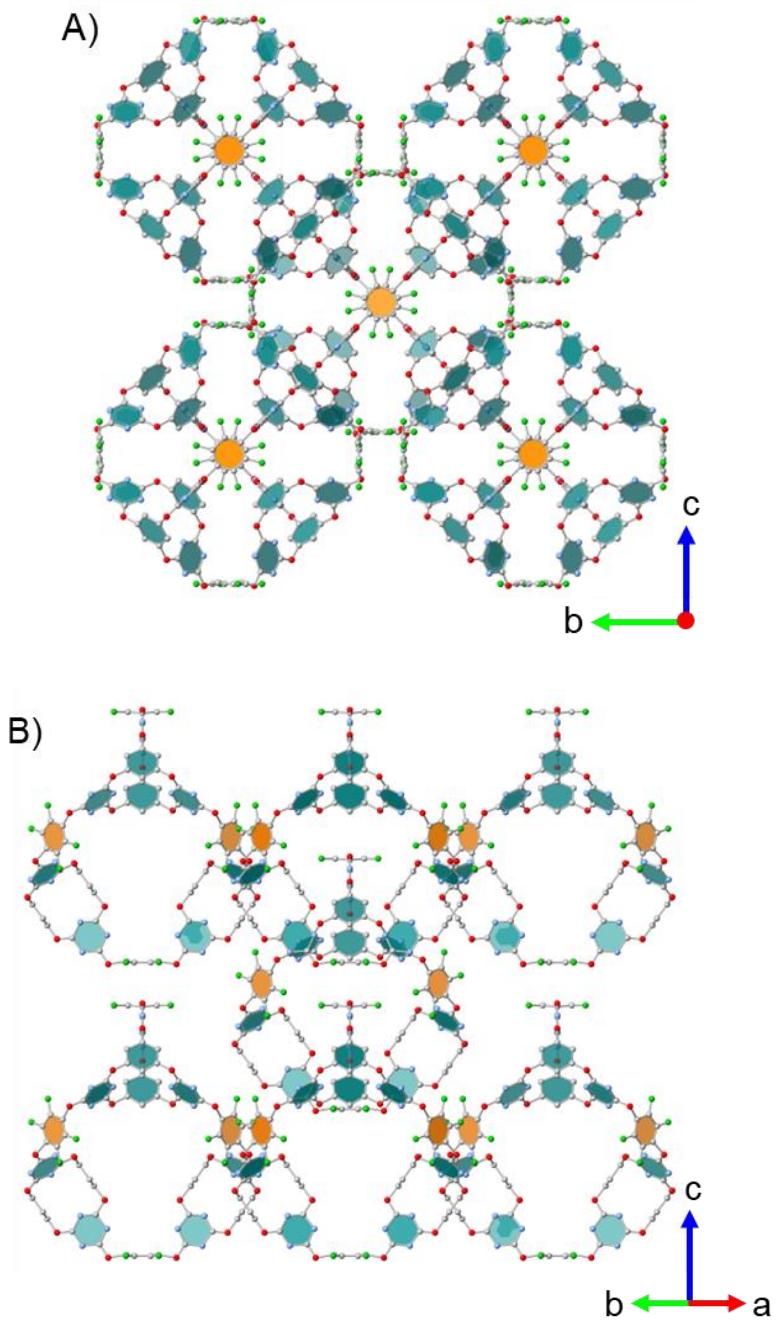


Figure S36 Crystal packing in the SCXRD structure of **[4[2+3]+6]cage·acetone** shown along the (A) *a*-axis and (B) [110] plane. Atom colours: C: grey, N: blue, O: red, and F: green. H atoms are omitted.

Section 7 Lattice energy minimization of the [4[2+3]+6]cage·acetone crystal structure

To assess the stability of the solvated **[4[2+3]+6]cage·acetone** structure, a rigid-body geometry optimization was conducted using DMACRYS software. Similar to previous procedures, the geometry of the molecular cage underwent optimization at the B3LYP/6-311G(d,p) level through Gaussian09 software[17], with the resulting geometry held constant during lattice energy minimization. The evaluation of energy utilized the FIT+DMA method, as detailed earlier. Figure S37 illustrates the solvent-removed **[4[2+3]+6]cage·acetone** structure before and after rigid-body lattice energy minimization. As detailed in Table S8, the solvated structure undergoes a loss of cubic symmetry, transforming into a monoclinic structure. The cell volume experiences a 10% reduction, leading to an 11% increase in density. This observation substantiates the experimental finding that the **[4[2+3]+6]cage·acetone** structure becomes destabilized upon the removal of solvent molecules.

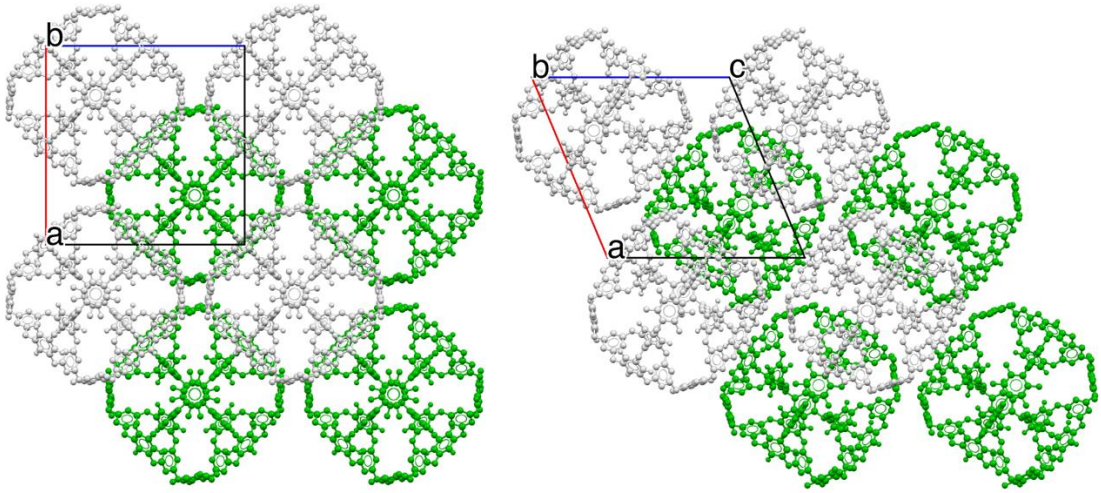


Figure S37 Crystal structure of the solvent-removed **[4[2+3]+6]cage·acetone** structure before (left) and after (right) rigid body lattice energy minimization.

Table S8 Crystal structure parameters of the experimental solvent-removed **[4[2+3]+6]cage·acetone** structure before and after lattice energy minimization.

	Experimental [4[2+3]+6]cage·acetone structure	**Relaxed [4[2+3]+6]cage·acetone structure**
Crystal System	Cubic	Monoclinic
Space group	217 ($I\bar{4}3m$)	4 ($P2_1$)
a (Å)	23.2901	22.9664
b (Å)	23.2901	23.0892
c (Å)	23.2901	23.1725
α (°)	90	90
β (°)	90	67.2627
γ (°)	90	90
Unit cell Volume (Å^3)	12633	11332.9
Density (g cm^{-3})	0.789	0.8796

Section 8 Hydrolytic stability study of [4[2+3]+6]cage

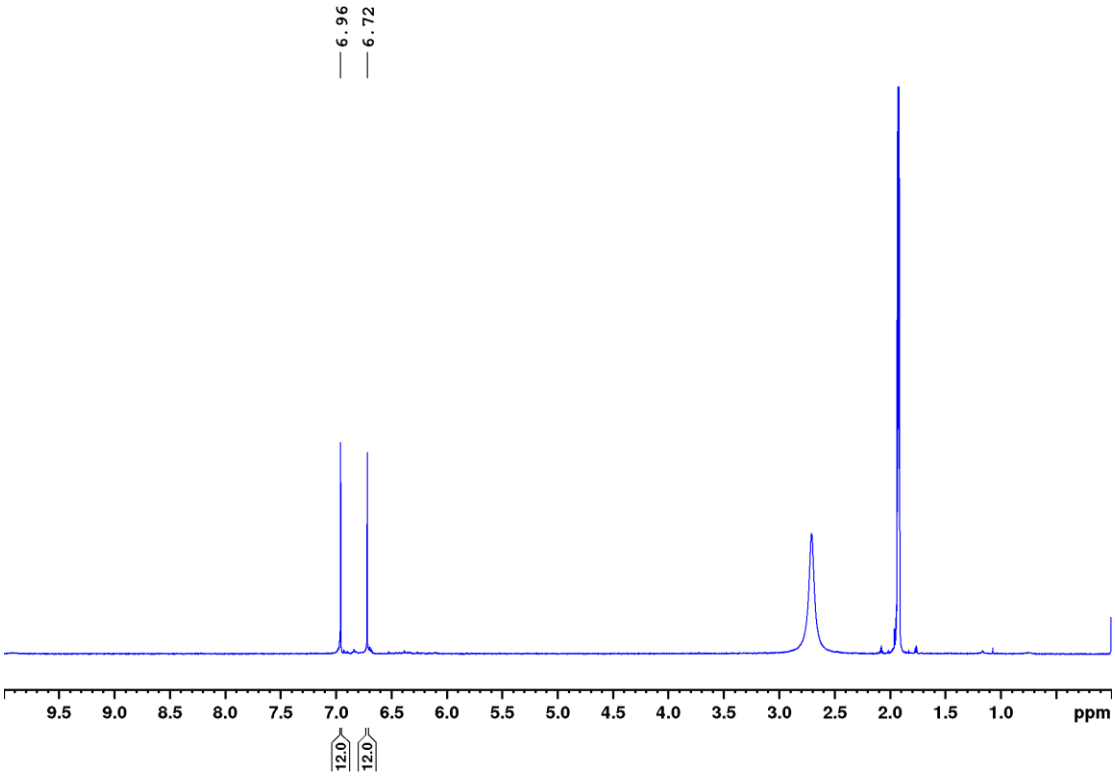


Figure S38 ^{1}H NMR (400 MHz, acetone-d_6) spectrum of **[4[2+3]+6]cage** recorded after the sample had been immersed in water for 12 days.

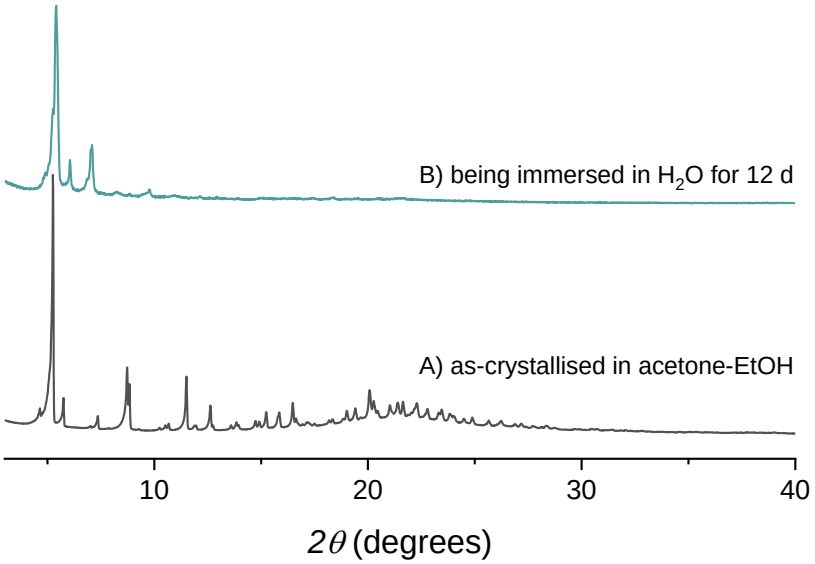


Figure S39 Comparison of PXRD patterns of **[4[2+3]+6]cage** crystal recorded: (A) in crystallisation solvents; (B) after being immersed in H_2O for 12 days.

References

1. Turcani, L. *et al.* stk: A python toolkit for supramolecular assembly. *J. Comput. Chem.* **39**, 1931–1942 (2018).
2. Lu, C. *et al.* OPLS4: Improving force field accuracy on challenging regimes of chemical space. *J. Chem. Theory Comput.* **17**, 4291–4300 (2021).
3. Kühne, T. D. *et al.* CP2K: An electronic structure and molecular dynamics software package -Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **152**, 194103 (2020).
4. Perdew, J. P. *et al.* Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
5. VandeVondele, J. & Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **127**, 114105 (2007).
6. Grimme, S. *et al.* A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
7. Willems, T. F. *et al.* Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous Mesoporous Mater.* **149**, 134–141 (2012).
8. Macrae, C. F. *et al.* Mercury 4.0 : from visualization to analysis, design and prediction. *J. Appl. Crystallogr.* **53**, 226–235 (2020).
9. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theor. Chem. Acc.* **120**, 215–241 (2008).
10. Frisch, M. J. *et al.* Gaussian 16 revision a. 03. 2016; gaussian inc. *Wallingford CT* **2**, 4 (2016).
11. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
12. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **8**, 1057–1065 (2006).

13. Sheldrick, G. SHELXT. *Acta Cryst.* **A71**, 3–8 (2015).
14. Sheldrick, G. SHELXL. *Acta Cryst.* **C71**, 3–8 (2015).
15. Dolomanov O. V. *et al.* OLEX2. *J. Appl. Cryst.* **42**, 339–341 (2009).
16. Zhu, Q. *et al.* Analogy powered by prediction and structural invariants: computationally led discovery of a mesoporous hydrogen-bonded organic cage crystal. *J. Am. Chem. Soc.* **144**, 9893–9901 (2022).
17. Frisch, M. J. *et al.* Gaussian09 Revision D.01, Gaussian Inc. Wallingford CT. *Gaussian 09 Revision C.01* at (2010).
18. Case, D. H. *et al.* Convergence properties of crystal structure prediction by quasi-random sampling. *J. Chem. Theory Comput.* **12**, 910–924 (2016).
19. Coombes, D. S. *et al.* Role of electrostatic interactions in determining the crystal structures of polar organic molecules. A distributed multipole study. *J. Phys. Chem.* **100**, 7352–7360 (1996).
20. Beyer, T. & Price, S. L. Dimer or catemer? low-energy crystal packings for small carboxylic acids. *J. Phys. Chem. B* **104**, 2647–2655 (2000).
21. Stone, A. J. Distributed multipole analysis: Stability for large basis sets. *J. Chem. Theory Comput.* **1**, 1128–1132 (2005).
22. Ferenczy, G. G. Charges derived from distributed multipole series. *J. Comput. Chem.* **12**, 913–917 (1991).
23. Price, S. L. *et al.* Modelling organic crystal structures using distributed multipole and polarizability-based model intermolecular potentials. *Phys. Chem. Chem. Phys.* **12**, 8478–8490 (2010).
24. Chisholm, J. A. & Motherwell, S. COMPACK: A program for identifying crystal structure similarity using distances. *J. Appl. Crystallogr.* **38**, 228–231 (2005).
25. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
26. Kresse, G. & Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metalamorphous- semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251–14269 (1994).
27. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
28. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B - Condens. Matter Mater.*

Phys. **54**, 11169–11186 (1996).

29. Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B - Condens. Matter Mater. Phys.* **59**, 1758, (1999).
30. Grimme, S. *et al.* Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).
31. Alexandre, P. E. *et al.* A robust porous quinoline cage: transformation of a [4+6] salicylimine cage by povarov cyclization. *Angew. Chem. Int. Ed.* **59**, 19675–19679 (2020).
32. Slater, A. G. *et al.* Computationally-guided synthetic control over pore size in isostructural porous organic cages. *ACS Cent. Sci.* **3**, 734–742 (2017).
33. Smith, P. T. *et al.* Iron porphyrins embedded into a supramolecular porous organic cage for electrochemical CO_2 reduction in water. *Angew. Chem. Int. Ed.* **57**, 9684–9688 (2018).
34. Hong, S. *et al.* Porphyrin boxes: rationally designed porous organic cages. *Angew. Chem. Int. Ed.* **54**, 13241–13244 (2015).
35. Lauer, J. C. *et al.* Shape-Persistent [4+4] imine cages with a truncated tetrahedral geometry. *Chem. - A Eur. J.* **24**, 1816–1820 (2018).
36. Liu, C. *et al.* Porous organic cages for efficient gas selective separation and iodine capture. *Chem. Eng. J.* **428**, 131129 (2022).
37. Jiang, S. *et al.* Core–shell crystals of porous organic cages. *Angew. Chem. Int. Ed.* **57**, 11228–11232 (2018).
38. Bhat, A. S. *et al.* Transformation of a [4+6] salicylbisimine cage to chemically robust amide cages. *Angew. Chem. Int. Ed.* **58**, 8819–8823 (2019).
39. Mastalerz, M. *et al.* A salicylbisimine cage compound with high surface area and selective CO_2/CH_4 adsorption. *Angew. Chem. Int. Ed.* **50**, 1046–1051 (2011).
40. Tozawa, T. *et al.* Porous organic cages. *Nat. Mater.* **8**, 973–978 (2009).
41. Zhang, C. *et al.* A porous tricyclooxacalixarene cage based on tetraphenylethylene. *Angew. Chem. Int. Ed.* **54**, 9244–9248 (2015).
42. Kunde, T. *et al.* A porous fluorinated organic [4+4] imine cage showing CO_2 and H_2 adsorption. *Chem. Commun.* **56**, 4761–4764 (2020).