nature synthesis

Article

https://doi.org/10.1038/s44160-024-00531-7

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

Received: 25 October 2023

Accepted: 26 March 2024

Published online: 26 April 2024

Check for updates

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[6] ✉ & Andrew I. Cooper[1,2] ✉

Here we report a two-step, hierarchical synthesis that assembles a trigonal prismatic organic cage into a more symmetric, higher-order tetrahedral cage, or 'cage of cages'. Both the preformed [2+3] trigonal prismatic cage building blocks and the resultant tetrahedral [4[2+3]+6]cage molecule are constructed using ether bridges. This strategy affords the [4[2+3]+6]cage molecule excellent hydrolytic stability that is not a feature of more common dynamic cage linkers, such as imines. Despite its relatively high molar mass (3,001 g mol^{-1}), [4[2+3]+6]cage exhibits good solubility and crystallizes into a porous superstructure with a surface area of 1,056 $m^2 g^{-1}$. By contrast, the [2+3] building block is not porous. The [4[2+3]+6]cage molecule shows high CO_2 and SF_6 uptakes due to its polar skeleton. The preference for the [4[2+3]+6]cage molecule over other cage products can be predicted by computational modelling, as can its porous crystal packing, suggesting a broader design strategy for the hierarchical assembly of organic cages with synthetically engineered functions.

The chemical synthesis of complex organic molecules is part of our toolkit to access materials with unique structures and functions[1–5]. Supramolecular self-assembly is a powerful strategy to synthesize molecules comprising a number of separate precursors[6–8]; these assemblies can also be nanometres in size[9,10] or chemically interlocked[11,12]. However, obtaining the desired self-assembly outcomes for more complex molecules quickly becomes synthetically challenging, particularly when the bond-forming chemistry has low reversibility. This creates a dichotomy: the more successful supramolecular reactions often lead to labile, unstable products, and this can limit the scope for applications. This challenge can be tackled by careful tuning of precursor structure and functionality, such as molecular geometry, or by iterative optimization of the synthetic procedures, but the best reaction conditions are often not intuitively obvious.

Some of the earliest supramolecular systems were synthesized by condensing simple bidentate building blocks, such as ethylenediamine and triethylene glycol, to form cryptands and crown ethers, respectively[13]. These molecules inspired the synthesis of larger and more complex architectures. For example, Fujita and co-workers introduced the concept of emergent behaviour in the assembly of large self-assembled macrocyclic products using carefully designed precursors[14]. Such supramolecular design strategies have allowed us

[1]Materials Innovation Factory and Department of Chemistry, University of Liverpool, Liverpool, UK. [2]Leverhulme Research Centre for Functional Materials Design, University of Liverpool, Liverpool, UK. [3]Department of Chemistry, Imperial College London, Molecular Sciences Research Hub, London, UK. [4]Computational Systems Chemistry, School of Chemistry, University of Southampton, Southampton, 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, China. [6]Institute of Chemical Sciences, Heriot-Watt University, Edinburgh, UK. ✉e-mail: m.little@hw.ac.uk; aicooper@liverpool.ac.uk

https://doi.org/10.1038/s44160-024-00531-7

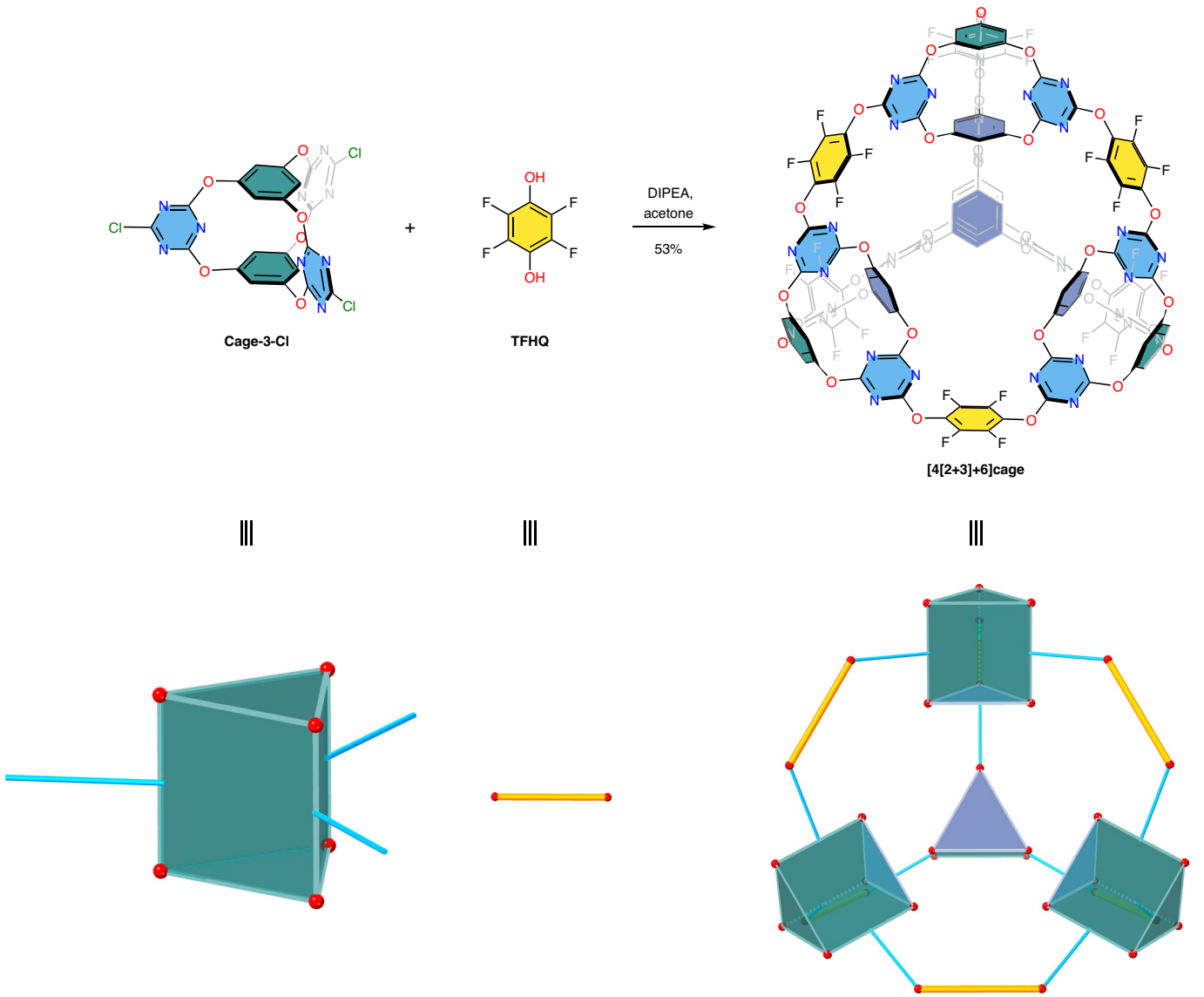

Fig. 1 | Synthetic route for the [4[2+3]+6]cage molecule. The **[4[2+3]+6]cage** molecule was synthesized via the S_NAr reaction between **Cage-3-Cl** and **TFHQ** in the presence of DIPEA. The triangular prism and the yellow sticks in the lower figure scheme represent **Cage-3-Cl** and **TFHQ**, respectively.

to synthesize more complex self-assembled structures and, hence, to unlock new applications[2,15,16]. However, high structural complexity is often accompanied by increased synthetic challenges and lower predictability because of sensitivity to parameters such as the precise bond angles in the precursors[9,14,17].

Postsynthetic modifications have been used previously to enhance the porosity of organic cages[18,19], such as by hooping parts of the cage together[20]. More recently, we and others have used hierarchical assembly strategies to form topologically complex hydrogen-bonded organic frameworks[21,22] and covalently bonded materials, such as covalent organic frameworks[23–26], using three-dimensional organic cages as the building blocks[27]. These studies have shown that cage-based building blocks can assemble into higher-order structures and increase the complexity of the resulting materials, for instance, by controlling network topology and interpenetration, while still offering a degree of structural predictability. In turn, this has afforded cage-based hydrogen-bonded organic frameworks and three-dimensional cage-based covalent organic frameworks with properties such as guest-responsive structural flexibility[23] and self-healing behaviour[28]. However, this hierarchical structuring approach does not appear to have been extended to the preparation of porous organic cage molecules[18,29]: that is, to synthesize larger porous cages from smaller organic cage precursors.

The use of organic cages as precursors to synthesize higher-order porous structures is attractive because it embeds cage molecules, with their own chemical complexity, into larger, hierarchical cages with the potential to create new functions while retaining useful properties such as solution processability[19,27,30]. For example, this strategy might produce porous materials with more sophisticated hierarchical porosities. To tackle this goal, we considered three criteria: (1) geometry—the cage precursors need geometries that can be arranged into a higher-order structure in a useful yield; (2) chemical stability—the chemical bonding in the cages must not be too labile, both to impart stability for applications and also to avoid the dynamic scrambling that might occur, for example, in trying to construct an imine cage from another imine cage[31]; (3) rigidity—the precursors need sufficient rigidity to direct chemical reactivity to the desired product and to ensure that the resultant hierarchical cage is shape persistent and retains its porous structure after removal of solvent from the voids.

To meet these three criteria, we chose a trigonal prismatic [2+3] ether-bridged cage molecule, **Cage-3-Cl**, as the polyhedral building block to construct a hierarchical 'cage of cages' (Fig. 1).

https://doi.org/10.1038/s44160-024-00531-7

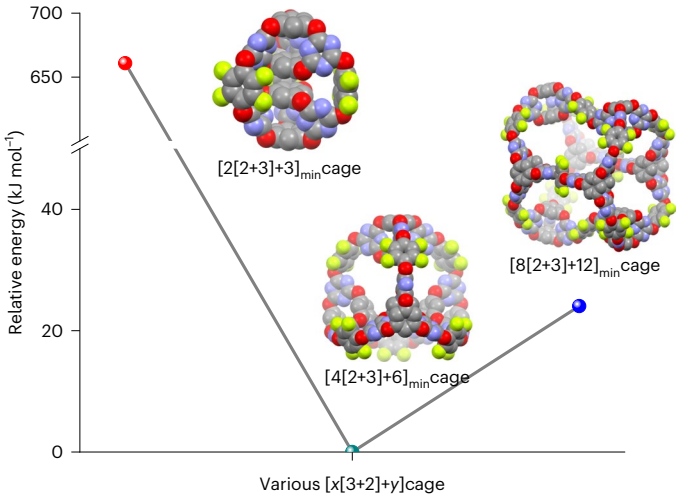

Fig. 2 | Relative DFT energies for minimum-energy configurations for [*x*[2+3]+*y*]cages. *x* = number of **Cage-3-Cl** cages, *y* = number of **TFHQ** linkers. Atom colours: carbon, grey; nitrogen, blue; oxygen, red; fluorine, green. Hydrogen atoms are omitted for clarity. Note the break in the energy scale for the highly strained [2[2+3]+3]cage, which has by far the highest relative energy (660.8 kJ mol^{-1}). The DFT energies indicate that the [4[2+3]+6] stoichiometry is predicted to form a stable, shape-persistent cage structure that has a lower relative energy (24.04 kJ mol^{-1}) than the alternative [8[2+3]+12] topology.

The preconfigured rigid geometry and excellent chemical stability of **Cage-3-Cl** allowed this [2+3] cage to assemble with tetrafluorohydroquinone (**TFHQ**) into the hierarchically structured organic 'cage of cages' compound, **[4[2+3]+6]cage**.

Results and discussion

Nucleophilic aromatic substitution (S_NAr) reactions have been reported to undergo reversible covalent bond formation when using electron-poor aromatic compounds[32–34], while still leading to stable molecular products. Reversible error-correction is important for the formation of complex molecules that must self-sort during the reaction from a variety of possible products. Although the S_NAr reaction has been used in the synthesis of ether-bridged cages, most tend to be [2+3] or [2+4] cage products with small intrinsic cavities[35–37], with the exception of a larger [4+6] ether-linked cage reported by Santos and co-workers[32]. One possible reason for the lack of larger cages synthesized via S_NAr chemistry is the less predictable orientation of the ether bridges compared to the imines and boronate esters for which larger cages are more commonplace[10,38–42].

Previous investigations by our group and others have demonstrated that **Cage-3-Cl** has a highly symmetric and rigid triangular prism geometry both in solution and in the solid state[21,36]. This geometry makes **Cage-3-Cl** an ideal building block for forming higher-order cage molecules, such as molecular barrels[20]. The three residual chlorine atoms exhibit high reactivity[43,44], which is essential for forming ether bridges. We selected **TFHQ** as the linear bridge between **Cage-3-Cl** molecules because the fluorine atoms might afford extra barriers to restrict the rotation of the ether bridges, and might improve the solubility of the resulting cage–cage molecules[36,45].

To explore the available bond angles and the relative flexibility of the ether bridges in possible hierarchical cage products, we performed molecular dynamics (MD) and density functional theory (DFT) calculations. Models were constructed with the supramolecular toolkit (stk) software[46] to predict the most likely reaction products. As shown in Fig. 2, the [4[2+3]+6] stoichiometry is predicted to form a stable, shape-persistent cage structure that exhibits a much lower energy than alternative [2[2+3]+3] and [8[2+3]+12] topologies. The [2[2+3]+3] topology has by far the highest relative energy (660.8 kJ mol^{-1}) due to its highly strained geometry. The [8[2+3]+12] topology has higher relative energy (24.04 kJ mol^{-1}) than the [4[2+3]+6] cage, which suggests that the [4[2+3]+6] topology is the thermodynamically favoured product, although we stress that these calculations do not include any solvent effects. As such, the [8[2+3]+12] topology might also be accessible under other synthesis conditions, whereas we predict that the [2[2+3]+3] topology is not. The *cis*–*trans* configurations of the ether bridges in the hypothetical [8[2+3]+12]cage can result in various positional configurations; all of these structural conformers were predicted to have relative energies that were between 24.0 and 229.1 kJ mol^{-1} higher than the [4[2+3]+6]cage, indicating a strong preference for the [4[2+3]+6] product (Supplementary Information Section 1 and Supplementary Figs. 1–4).

These simulation results suggested that it might be possible to synthesize **[4[2+3]+6]cage** via the S_NAr reaction between **Cage-3-Cl** and **TFHQ** (Fig. 1). We therefore attempted the reaction experimentally, and screened a range of conditions in which we varied the reagent concentration, solvent and base (Supplementary Table 1). From these experiments, we found that the reaction in acetone in the presence of the acid scavenger *N*,*N*-diisopropylethylamine (DIPEA) afforded a new product with the highest yield of 53% after purification. The ^{1}H NMR spectrum for the purified reaction product from the acetone reaction with DIPEA showed two singlets at 7.09 and 6.85 ppm, which we assigned to the two aromatic protons in the [2+3] cage (H_a and H_b; Fig. 3a and Supplementary Fig. 5). The presence of two singlets indicates different environments, which we attribute to one of the protons being more shielded. However, apart from this splitting of the aromatic proton singlet in **Cage-3-Cl**, the NMR spectroscopy data indicated that the resulting product had high symmetry in solution. In the ^{13}C NMR spectrum, we observed three signals in the 174.5–173.1 ppm range (Fig. 3b and Supplementary Fig. 6), which we assigned to the triazine ring carbon atoms. We attribute the characteristic splitting, observed at 142.5 and 140.0 ppm with a coupling constant of 250 MHz, to the coupling between the carbon and fluorine atoms in the **TFHQ** linker (Fig. 3b and Supplementary Fig. 6). We also confirmed the presence of these fluorinated aromatic rings by ^{19}F NMR spectroscopy, observing a singlet at −155.62 ppm (Supplementary Fig. 7), indicating that the fluorine atoms were symmetrically equivalent in solution. We also used high-resolution matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry to analyse the reaction product. We found an ion with a mass-to-charge (*m*/*z*) ratio of 3,002.0756 (Fig. 3c and Supplementary Figs. 8 and 9), which matched well with the theoretical value of [**[4[2+3]+6]cage** + H]$^+$ (3002.0871), indicating the formation of **[4[2+3]+6]cage**.

We next grew crystals for single-crystal X-ray diffraction analysis to confirm the structure of the **[4[2+3]+6]cage** molecule. Slow evaporation of a mixture of acetone/ethanol afforded single crystals suitable for X-ray analysis using synchrotron radiation (Supplementary Fig. 10 and Supplementary Table 2). The synchrotron single-crystal structure, which we refined in the monoclinic $P2_1$ space group, revealed that the **[4[2+3]+6]cage** molecule adopts a tetrahedral topology, where four **Cage-3-Cl** cage molecules serve as the vertices and six **TFHQ** molecules are located as the edges (Fig. 4a). The interior and the exterior aryl caps of the **Cage-3-Cl** cage molecules form a core–shell structure, defining an inner and outer truncated tetrahedron with edge lengths of 6.4 and 13.7 Å, respectively (Fig. 4b). We also calculated the electrostatic potentials for the **[4[2+3]+6]cage** molecule, which showed that the centre of the **[4[2+3]+6]cage** molecule is surrounded by aromatic rings, affording π–π interactions for any guest molecules within the cage (Fig. 4c and Supplementary Information Section 1).

The interior of the cage core exhibits an electron-poor character because of the V-shaped electron-deficient clefts formed by the triazine rings of **Cage-3-Cl** and the fluorine-decorated aromatic rings. This environment might be useful for selective guest molecule separation[47–49]. In the extended crystal structure of this cage of cages, the asymmetric

https://doi.org/10.1038/s44160-024-00531-7

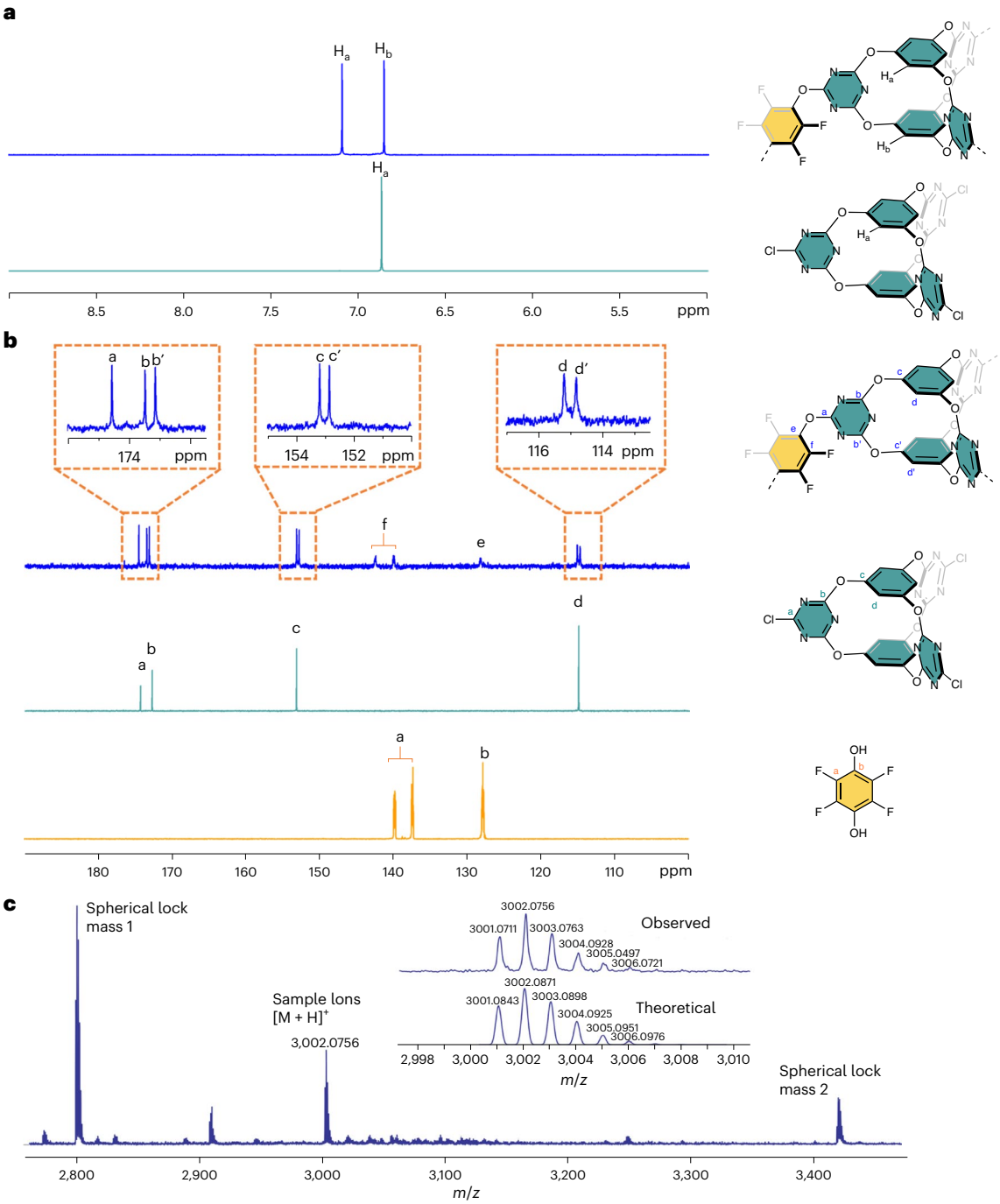

Fig. 3 | NMR spectra and molecular structures of [4[2+3]+6]cage, Cage-3-Cl and TFHQ. a, ^{1}H NMR (400 MHz, acetone-d_6) spectra of **Cage-3-Cl** (green, bottom) and **[4[2+3]+6]cage** (blue, top). **b**, ^{13}C NMR (100 MHz, dioxane-d_8) spectra: **TFHQ** (yellow, bottom), **Cage-3-Cl** (green, middle) and **[4[2+3]+6] cage** (blue, top). Insets: zoom-ins of the boxed regions. The NMR spectra highlight the splitting of peaks due to the formation of a hierarchical 'cage of cages' structure. **c**, High-resolution MALDI-TOF spectrum of **[4[2+3]+6]cage**, showing an ion with an *m/z* ratio of 3,002.0756 assigned to [**[4[2+3]+6]cage** + H]$^+$. Two internal calibrants (Spherical) with *m/z* ratios of 2,979 and 3,423 that bracketed the ion of interest were used to limit the *m/z* error to ±5 ppm.

cell contains one **[4[2+3]+6]cage** molecule, which assembles into a porous supramolecular structure by interacting with 12 neighbouring **[4[2+3]+6]cage** molecules through van der Waals forces (Supplementary Fig. 11). Two of the windows in the **[4[2+3]+6]cage** molecule are narrowed into smaller channels by the **Cage-3-Cl** vertices from neighbouring cage molecules (Fig. 4d,f and Supplementary Fig. 12), yielding three-dimensional interconnected pore channels (Fig. 4d,e). Using Zeo++[50], we calculated that the pore-limiting diameter of the **[4[2+3]+6]cage** crystal structure was 6.4 Å and the largest cavity diameter was 8.9 Å (Supplementary Table 3 and Supplementary Figs. 13–15), suggesting that the structure is microporous. From these calculations, we also determined that voids in the **[4[2+3]+6]cage** crystal structure that are accessible to a 1.65 Å CO_2 probe occupy 32.0% of the unit cell volume (Supplementary Table 3).

There was strong agreement between the predicted structure for the **[4[2+3]+6]cage** molecule and the molecule observed in the crystal structure (Fig. 5). This validates the theoretical predictions, and the close match between the crystal structure prediction (CSP)-predicted structure and experimental crystal structure adds confidence in the crystal structure refinement (Supplementary Fig. 17). The root mean

https://doi.org/10.1038/s44160-024-00531-7

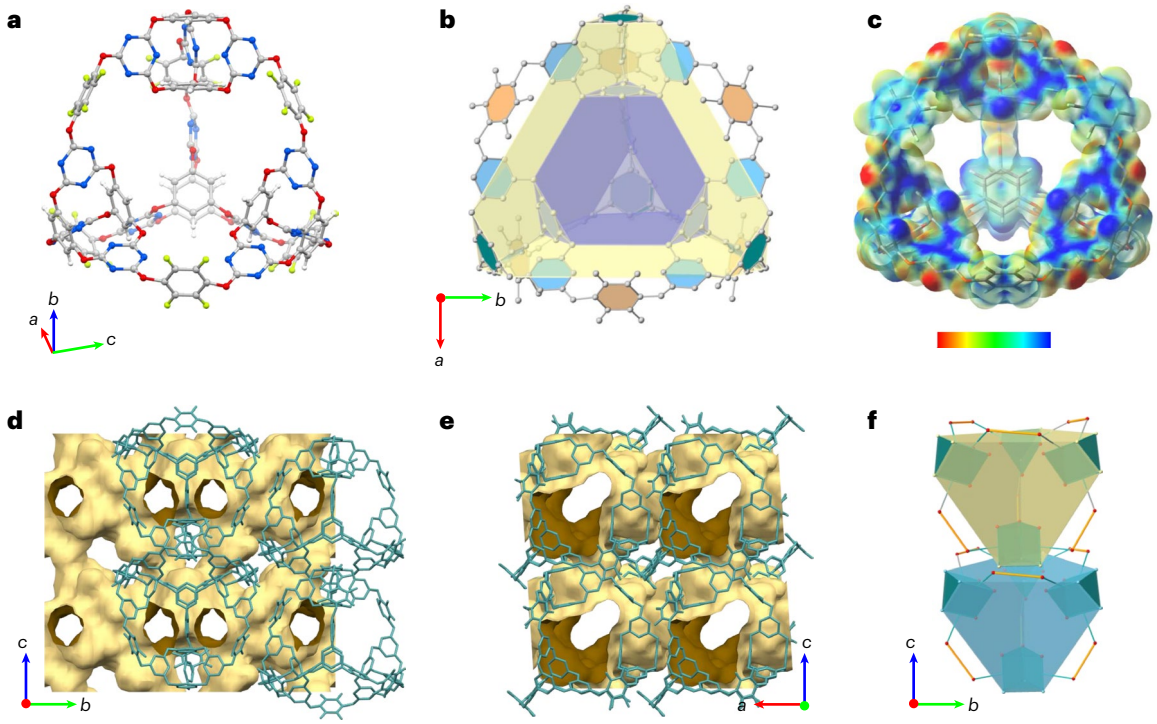

Fig. 4 | Crystal structure of [4[2+3]+6]cage. a, Structure of an individual **[4[2+3]+6]cage** molecule. Atom colours: carbon, grey; hydrogen, white; nitrogen, blue; oxygen, red; fluorine, green. **b**, Representation of the **[4[2+3]+6] cage** molecule using two truncated tetrahedra on the inner and outer aryl caps of the [2+3] **Cage-3-Cl** cage molecules. For clarity, all atoms here are coloured grey. **c**, Electrostatic potential maps of the **[4[2+3]+6]cage** molecule. The red and blue surfaces represent negative and positive regions of potential, respectively. Colour bar, −31.4 to 94.1 kcal mol^{-1}. **d**,**e**, Pore channels in the extended **[4[2+3]+6] cage** crystal structure as viewed along the *a* axis (**d**) and the *b* axis (**e**). For clarity, hydrogen atoms are omitted in **b**, **e** and **f**. The yellow surfaces in **d** and **e** represent the contact surface as measured using a 1.2 Å diameter probe. **f**, Scheme explaining the window splitting in the **[4[2+3]+6]cage** crystal structure along the *a* axis; the window of the lower blue cage is partially occluded by the aryl face of the upper yellow cage.

squared displacement (r.m.s.d.) was calculated as 0.5 Å with a maximum distance between atoms of 1.4 Å. However, the experimental displacement parameters are large due to disorder in the crystal structure (Supplementary Fig. 11a). Further attempts to synthesize the larger [8[2+3]+12] product by varying the reaction conditions were unsuccessful, based on MALDI-TOF analysis of the resulting products (Supplementary Table 1 and Supplementary Fig. 8), in line with the molecular stability predictions (Fig. 2).

In principle, catenation of this cage is possible, given its large intrinsic voids (>10 Å diameter), as observed for considerably smaller imine cages[11]. However, we saw no evidence for catenated cage side-products, either by NMR or by MALDI-TOF characterization.

We next used CSP to explore the solid-state packing of these hierarchical cages. The lattice energy landscape was explored using quasi-random sampling of the crystal packing space with the Global Lattice Energy Explorer (GLEE)[51]. Initial trial structures were generated from rigid molecules and subjected to lattice energy minimization using an empirically parameterized potential with atomic multipole electrostatics[52] (see Supplementary Information Section 4, Supplementary Tables 4 and 5, and Supplementary Figs. 16–25 for full details).

Surprisingly, the CSP landscape for **[4[2+3]+6]cage** (Fig. 6) showed catenated structures, along with the non-catenated cage that was observed experimentally, even though the discrete **[4[2+3]+6]cage** molecule was used for the CSP calculations. Three distinct catenations were identified in the predicted crystal structures: triply interlocked cage dimers (Fig. 6c), singly interlocked cage dimers (Fig. 6d) and singly interlocked one-dimensional (1D) cage chains[12,53] (Fig. 6e). The details of the methods used for catenation detection are provided in Supplementary Information Section 4 and Supplementary Figs. 18–20. All sampled structures within a 197 kJ mol^{-1} energy window from the global energy minimum were found to be catenanes (Supplementary Figs. 21 and 22), indicating a strong thermodynamic preference over the non-catenated cages observed by experiment. To verify the relative energies calculated using the rigid-molecule, force-field approach, a selection of catenated and non-catenated predicted structures were re-evaluated using periodic DFT, which confirmed this greater thermodynamic stability (see Supplementary Information Section 4 for full details).

While the CSP study did not explicitly target catenated structures, the sampled catenated configurations suggest that triply interlocked catenanes (green points, Fig. 6a), in particular, might be much more thermodynamically stable in the solid state. This echoes previous findings for [4+6] imine cages, in which discrete cages were found to transform into triply interlocked catenanes upon exposure to acid, suggesting that the individual cages were the kinetic rather than the thermodynamic product[11]. The absence of catenanes in our experiments might be explained by the much lower reversibility of the ether bonding in the **[4[2+3]+6]cage** molecule, which is not accounted for in the CSP calculations. Prompted by these solid-state CSP results, we also explored the relative thermodynamic stability of catenanes at the molecular level. DFT calculations of catenane dimers showed that the energy difference between the molecular equivalent non-catenated **[4[2+3]+6]cage** dimer and trimer fragments retrieved from the global lowest-energy CSP, and the corresponding triply interlocked catenane molecular fragment was 373.7 kJ mol^{-1} and 324.7 kJ mol^{-1}, respectively, reaffirming strong thermodynamic favour towards the catenane structures.

When we remove the catenated structures from the CSP plot (Fig. 6b and Supplementary Fig. 23), this reveals the observed experimental structure positioned at the bottom of a low-density 'spike' in the energy landscape, approximately 13.6 kJ mol^{-1} higher than the global energy minimum for non-catenated cages. The predicted crystal

https://doi.org/10.1038/s44160-024-00531-7

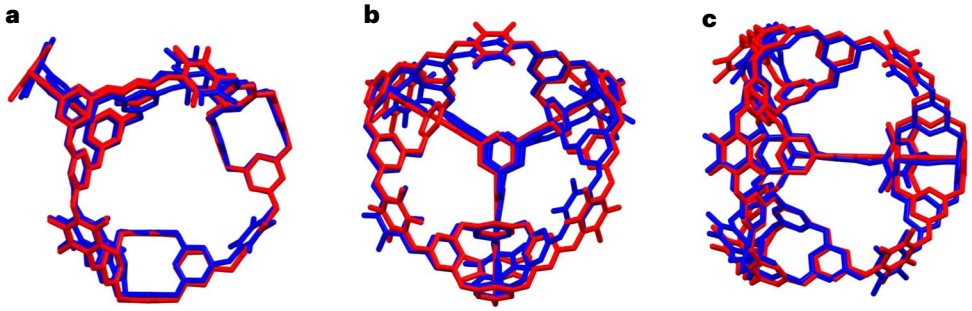

Fig. 5 | Comparison between the predicted [4[2+3]+6]cage model and the experimental single-crystal X-ray diffraction structure. a–c, The predicted structure (red) overlaid with the single-crystal X-ray diffraction structure (blue) is shown as viewed along the *a* (**a**), *b* (**b**) and *c* (**c**) crystallographic axes. The r.m.s.d. was calculated as 0.5 Å with a maximum distance between atoms of 1.4 Å, highlighting the close structural similarity between the predicted and experimental structures.

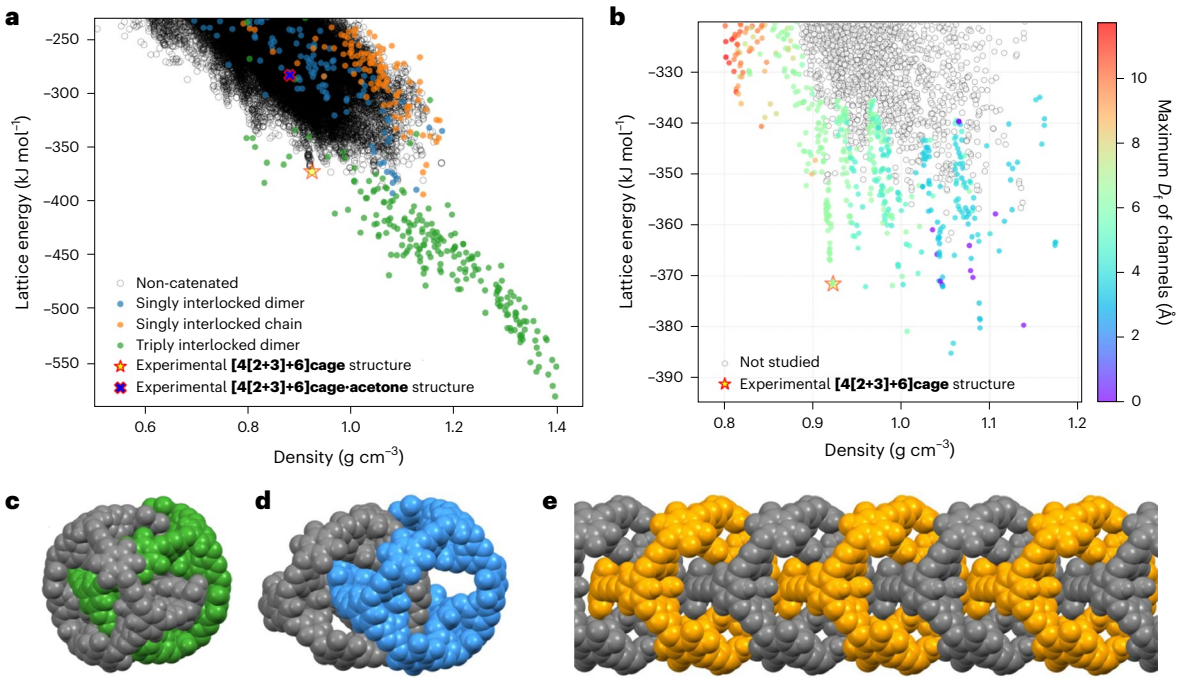

Fig. 6 | Crystal structure prediction for [4[2+3]+6]cage. a, Computational crystal energy landscape of **[4[2+3]+6]cage** with colour-coded categorization based on catenation type: discrete, non-catenated cages (uncoloured circles), triply interlocked cage dimers (green circles), singly interlocked cage dimers (blue) and singly interlocked 1D cage chains (orange). The yellow star and blue cross represent the predicted structures matching the experimentally observed **[4[2+3]+6]cage** crystal structure and **[4[2+3]+6]cage·acetone** solvated structure, respectively. **b**, Energy landscape after removal of the catenated structures, with colour coding based on the diameter of the largest sphere (D_f) capable of freely moving within the crystal structure's channel(s). Channels are found based on their ability to accommodate a CO_2 molecule. $D_f = 0$ corresponds to no channel being found. **c–e**, Atomic structures depicted for examples of a triply interlocked cage dimer (**c**), a singly interlocked cage dimer (**d**) and a singly interlocked 1D cage chain (**e**).

structure reproduces the geometry of the experimentally determined **[4[2+3]+6]cage** crystal structure accurately (Supplementary Fig. 17), confirming that the crystal structure determined by X-ray diffraction corresponds to a low-energy local minimum in lattice energy. The colour coding in this 'non-catenated' crystal structure landscape represents the diameter of the largest sphere capable of unrestricted movement within the crystal structure channels. Channel dimensions are determined based on their capacity to accommodate a CO_2 molecule with a kinetic radius of 1.65 Å (Supplementary Figs. 24 and 25). In the landscape depicted in Fig. 6b, void analysis has been restricted to structures within 20 kJ mol^{-1} of the low-energy edge of the energy-density distribution of structures. Except for a very small number of predicted structures (purple points, Fig. 6b), all investigated structures, including the synthesized structure, show potential for CO_2 uptake. That is, CSP suggests that **[4[2+3]+6]cage** has an intrinsic propensity to be porous in the majority of its potential crystalline packing modes.

Molecular crystals exhibiting permanent porosity in the solid state are attractive for applications such as gas capture, separation and catalysis[18,54]. One successful approach that we and others have developed is to form porous organic crystals by synthesizing cages with prefabricated shape-persistent cavities that are retained after solvents are removed during activation[18,50,54]. Our calculations revealed that the ether bridges in the **[4[2+3]+6]cage** skeleton appeared to be relatively rigid, suggesting shape persistence. We therefore investigated the porosity in the **[4[2+3]+6]cage** crystals using gas sorption analysis. We activated the **[4[2+3]+6]cage** crystals by first exchanging the ethanol and acetone crystallization solvents with diethyl ether or *n*-pentane, which we chose because of their low surface tensions. Then, we removed any residual solvent from the crystals under a dynamic vacuum at room temperature. Subsequent powder X-ray diffraction (PXRD) analysis revealed that the **[4[2+3]+6]cage** crystals retained some crystallinity after being activated using these conditions

https://doi.org/10.1038/s44160-024-00531-7

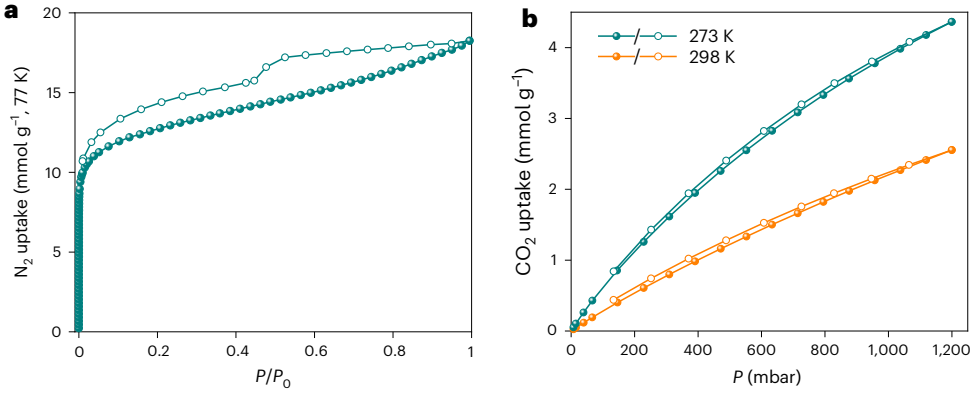

Fig. 7 | Gas sorption isotherms for [4[2+3]+6]cage. a, N_2 sorption isotherms recorded at 77 K showing hysteresis in the desorption isotherm. **b**, CO_2 gas sorption isotherms recorded at 273 K (cyan) and 298 K (orange) showing an uptake capacity of 3.98 mmol g^{-1} at 1 bar and 273 K. Closed and open symbols represent the adsorption and desorption isotherms, respectively.

(Supplementary Fig. 26). The **[4[2+3]+6]cage** crystals activated via the diethyl ether solvent exchange route appeared more crystalline, and this sample was used for the subsequent gas sorption experiments described here.

Nitrogen sorption isotherms recorded at 77 K revealed that the crystalline **[4[2+3]+6]cage** exhibits a type I N_2 sorption isotherm with a relatively high Brunauer–Emmett–Teller surface of 1,056 $m^2 g^{-1}$ (Fig. 6a and Supplementary Figs. 27–29), consistent with a microporous solid and the pore size distribution plot calculated using Zeo++[51] (Supplementary Table 3 and Supplementary Fig. 13). We found that crystalline **[4[2+3]+6]cage** has a CO_2 uptake capacity of 3.98 mmol g^{-1} at 1 bar and 273 K (Fig. 7b and Supplementary Fig. 30). This CO_2 uptake is high compared with other porous organic crystalline materials, such as covalent organic frameworks[55], at comparable temperatures and pressures, and is one of the highest CO_2 uptakes reported to date for a porous organic cage (Supplementary Table 6)[56,57]. The calculated isosteric heat of adsorption of CO_2 on crystalline **[4[2+3]+6]cage** ranges between 21.1 and 23.2 kJ mol^{-1} (Supplementary Fig. 31), which indicates a strong affinity between the adsorbed CO_2 gas and polar **[4[2+3]+6]cage** crystal pores, rationalizing this high uptake capacity. In addition, we found that crystalline **[4[2+3]+6]cage** has a high SF_6 uptake capacity of 3.21 mmol g^{-1} at 1 bar and 273 K (Supplementary Fig. 32). The calculated isosteric heat of adsorption of SF_6 on crystalline **[4[2+3]+6]cage** ranges between 29.2 and 29.5 kJ mol^{-1}, which again indicates a strong affinity between adsorbed SF_6 gas molecules and the **[4[2+3]+6]cage** crystal pores (Supplementary Fig. 33). Analysis of the **[4[2+3]+6]cage** powder after the gas sorption isotherms by PXRD analysis indicated that the material remained crystalline during these measurements (Supplementary Fig. 34).

We also uncovered a second crystal structure of the **[4[2+3]+6]cage** molecule during this study, referred to as **[4[2+3]+6]cage·acetone**, which crystallized from slow evaporation of an acetone-d_6 solution (Supplementary Fig. 35). **[4[2+3]+6]cage·acetone** crystallized in the cubic space group $I\bar{4}3m$ (a = 23.2901(15) Å, V = 12633(2) $Å^3$, Supplementary Table 7) with the ether-bridged cage adopting a perfect tetrahedral geometry in the structure (Supplementary Fig. 36). The **[4[2+3]+6]cage·acetone** lost crystallinity rapidly after being removed from the acetone-d_6 solvent and cracked (Supplementary Fig. 35). We therefore performed single-crystal analysis by sealing a solvated crystal in a borosilicate capillary containing residual acetone-d_6 solvent. However, due to the poorer crystal stability of **4[2+3]+6]cage·acetone**, we did not investigate its solid-state properties further. The instability of this form was further investigated through computational geometry optimization of the crystal structure. Employing the same energy model as used in the CSP study, rigid-molecule geometry optimization of the structure after solvent removal resulted in considerable structural distortion from the original cubic lattice, adopting a monoclinic form, in keeping with the observed experimental instability. Details can be found in Supplementary Information Section 8. The relaxed structure, denoted by a blue cross in the landscape of Fig. 6a, is situated 103 kJ mol^{-1} above the global energy minimum on the landscape of non-catenated structures. This energy difference underscores the crucial role of solvent stabilization in the synthesis of this solvated structure, and can also help to rationalize why this tetrahedral molecular structure was not predicted using gas-phase (that is, solvent-free) DFT calculations (Fig. 5).

For practical applications, gas sorption capacity is not the only criterion. For example, most CO_2 capture applications involve wet or humid gas streams, and hence water stability is important. Many porous organic cage materials, such as imine cages and (particularly) boronate ester cages, are unstable to water. We therefore explored the hydrolytic stability of the **[4[2+3]+6]cage** molecule by immersing the synthesized crystals in water for 12 days. Subsequent analysis of the sample by ^1H NMR spectroscopy revealed that **[4[2+3]+6]cage** remained chemically intact under these conditions (Supplementary Fig. 38). PXRD analysis of the same sample also revealed that the **[4[2+3]+6]cage** crystals retained their crystallinity under these conditions (Supplementary Fig. 39). Hence, both the chemical and crystal structure of **[4[2+3]+6]cage** molecule appear to have good hydrolytic stability.

Conclusion

We report the assembly of a more complex type of porous organic cage—a 'cage of cages'—that was synthesized using a two-step hierarchical self-assembly strategy. In this study, we demonstrate the strategy by assembling four trigonal cages into a larger tetrahedral cage. The resulting **[4[2+3]+6]cage** molecule exhibits excellent stability in water, and crystals of the **[4[2+3]+6]cage** show permanent porosity and a high surface area of 1,056 $m^2 g^{-1}$. The abundance of polar atoms in the cage cavity endows it with high CO_2 and SF_6 uptake capacity. The good solubility of **[4[2+3]+6]cage** in acetone indicates it has the potential to be used as a building block for even more complex structures, such as porous cage co-crystals. More broadly, this illustrates a strategy for hierarchical molecular assembly using computation as a guide to assess the most likely reaction products. For example, it might be possible in the future to design analogous systems where the [2+3] cages contribute discrete, prefabricated porosity into a higher-order, hierarchically porous crystal.

This study also showcases the use of computational design in supramolecular synthesis, both at the molecular level (Fig. 5) and in the solid state (Fig. 6). It is notable that triply interlocked cage catenane dimers emerged as the most stable predicted crystal packings (Fig. 6a). Such catenanes were not observed in experiments, most likely because

they are kinetically disfavoured, but they are nonetheless synthetically plausible because analogous structures have been formed using more reversible [4+6] imine cage-forming reactions[11]. Less obviously, infinite 1D catenated cage chains are also produced in these simulations (Fig. 6e), and in some cases these structures are predicted to have similar lattice energies to the experimentally observed non-catenated cage (Fig. 6a). This highlights how a priori structure predictions have the power to suggest non-intuitive new materials, although it is unclear how one might design a kinetic pathway to these chain structures, even though analogous structures have been observed for less complex macrocycles[53].

Methods

Molecular simulations

Both **Cage-3-Cl** and cage-of-cages models were constructed in Tri2Di3, Tri4Di6 and Tri8Di12 topologies using the stk software[46]. All cages were annealed with an MD simulation at 700 K for 50 ns with a time step of 0.5 fs after a 100 ps equilibration time with the OPLS4 force field as implemented in the Macromodel Suite[58]. Five hundred random configurations from the total MD duration were sampled and energy minimized, with the lowest energy configuration selected for DFT calculations. DFT calculations were performed with CP2K v.2023.1 (ref. 59) software using the generalized gradient approximation theory with the Perdew–Burke–Ernzerhof functional[60] and def2-TZVP basis sets[61]. A planewave cut-off value of 400 Ry and a relative cut-off value of 100 Ry were parameterized to obtain converged energy levels and dispersion interactions were accounted for with Grimme's DFT-D3 approach[62].

The geometries of the **[4[2+3]+6]cage** were then fully optimized by means of the hybrid M06-2X functional in Gaussian16 (ref. 63). The def2-SVP basis set[64,65] was applied for all atoms. No symmetry or geometry constraint was imposed during optimizations. The optimized geometries were verified as local minima on the potential energy surface by frequency computations at the same theoretical level[63].

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

To synthesize **[4[2+3]+6]cage**, DIPEA (61 µl, 0.35 mmol) was dissolved in acetone (25 ml) and purged with N_2 for 10 min. To the acetone solution, a mixture of **Cage-3-Cl** (58.7 mg, 0.1 mmol) and **TFHQ** (27.3 mg, 0.15 mmol) in acetone (6 ml) was added dropwise over 3 h under a N_2 atmosphere. After the addition was complete, the reaction was stirred at room temperature for 36 h. The solvent was then removed by rotary evaporation, and the crude product was purified by column chromatography using acetone/CH_2Cl_2 (10% vol/vol acetone) as eluent to afford **[4[2+3]+6]cage** as a white solid in 53% isolated yield: 40 mg (0.013 mmol). ^{1}H NMR (400 MHz, acetone-d_6): δ (ppm) 7.09 (s, 12H, H_b), 6.85 (s, 12H, H_a); ^{19}F NMR (376 MHz, acetone-d_6): δ (ppm) −155.62; ^{13}C NMR (100 MHz, dioxane-d_8): δ (ppm) 174.5, 173.5, 173.1, 153.2, 152.8, 142.5, 140.1, 140.0, 128.3, 115.2, 114.8. MALDI-TOF $[M + H]^+$, $[C_{120}H_{24}F_{24}N_{36}O_{36} + H]^+$: calculated, 3002.0871; found, 3002.0756.

CSP

CSP involves the following general steps: (1) molecular geometry optimization; (2) trial crystal structure generation; (3) local lattice energy minimization of trial structures; and (4) duplicate removal.

The geometry of the molecular cage was optimized at the B3LYP/6-311 G(d,p) level using Gaussian09 software[66], and the resulting geometry was kept fixed throughout the subsequent steps. Trial crystal structures are generated using the Global Lattice Energy Explorer (GLEE) code[51]. Subsequently, these trial structures undergo lattice optimization while preserving the rigidity of the molecular cage. For this task, we employ an empirically parameterized intermolecular atom–atom exp-6 potential coupled with atomic multipole electrostatics. The force-field parameters are acquired from the FIT force field[67,68]. Atom-centred multipoles up to hexadecapole on each atom were derived from the electron density through DMA, and partial charges (used in early stages of optimization) were fitted to the molecular electrostatic potential generated by these multipoles[69,70]. The overall model is denoted as FIT + DMA.

The search for space groups involves sampling the ten most common space groups for organic crystals along with four trigonal space groups (143, 144, 145 and 146), each with one molecule in the asymmetric unit. A quasi-random method is used to search these selected space groups separately, and valid structures are lattice energy minimized using DMACRYS software[52] in a two-stage protocol. The first stage involves FIT + DMA with partial charges, followed by the second stage with multipole electrostatics. More details can be found in Supplementary Information.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper, its Supplementary Information files, and the Cambridge Crystallographic Data Centre (deposition numbers 2303319 for **[4[2+3]+6]cage** and 2326368 for **[4[2+3]+6]cage·acetone**). The crystal structures and structure factor data can be obtained free of charge from the Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif. The CSP data are available at the University of Southampton Institutional Research Repository at https://doi.org/10.5258/SOTON/D2929 (ref. 71).

References

1. Lehn, J. M. Toward self-organization and complex matter. *Science* **295**, 2400–2403 (2002).
2. McTernan, C. T., Davies, J. A. & Nitschke, J. R. Beyond platonic: how to build metal–organic polyhedra capable of binding low-symmetry, information-rich molecular cargoes. *Chem. Rev.* **122**, 10393–10437 (2022).
3. Guillerm, V. & Eddaoudi, M. The importance of highly connected building units in reticular chemistry: thoughtful design of metal–organic frameworks. *Acc. Chem. Res.* **54**, 3298–3312 (2021).
4. Montà-González, G., Sancenón, F., Martínez-Máñez, R. & Martí-Centelles, V. Purely covalent molecular cages and containers for guest encapsulation. *Chem. Rev.* **122**, 13636–13708 (2022).
5. Lutz, J. F., Lehn, J. M., Meijer, E. W. & Matyjaszewski, K. From precision polymers to complex materials and systems. *Nat. Rev. Mater.* **1**, 16024 (2016).
6. Acharyya, K. & Mukherjee, P. S. Organic imine cages: molecular marriage and applications. *Angew. Chem. Int. Ed.* **58**, 8640–8653 (2019).
7. Deng, H. et al. Multiple functional groups of varying ratios in metal–organic frameworks. *Science* **327**, 846–850 (2010).
8. Pullen, S. & Clever, G. H. Mixed-ligand metal–organic frameworks and heteroleptic coordination cages as multifunctional scaffolds—a comparison. *Acc. Chem. Res.* **51**, 3052–3064 (2018).
9. Fujita, D. et al. Self-assembly of tetravalent Goldberg polyhedra from 144 small components. *Nature* **540**, 563–566 (2016).
10. Koo, J. et al. Gigantic porphyrinic cages. *Chem* **6**, 3374–3384 (2020).
11. Hasell, T. et al. Triply interlocked covalent organic cages. *Nat. Chem.* **2**, 750–755 (2010).
12. Benke, B. P., Kirschbaum, T., Graf, J., Gross, J. H. & Mastalerz, M. Dimeric and trimeric catenation of giant chiral [8+12] imine cubes driven by weak supramolecular interactions. *Nat. Chem.* **15**, 413–423 (2023).
13. Steed, J. W., Turner, D. R. & Wallace, K. J. *Core Concepts in Supramolecular Chemistry and Nanochemistry* (John Wiley, 2007).
14. Sun, Q. F. et al. Self-assembled M24L48 polyhedra and their sharp structural switch upon subtle liqand variation. *Science* **328**, 1144–1147 (2010).
15. Fujita, D. et al. Protein stabilization and refolding in a gigantic self-assembled cage. *Chem* **7**, 2672–2683 (2021).

16. Chakrabarty, R., Mukherjee, P. S. & Stang, P. J. Supramolecular coordination: self-assembly of finite two- and three-dimensional ensembles. *Chem. Rev.* **111**, 6810–6918 (2011).
17. Santolini, V., Miklitz, M., Berardo, E. & Jelfs, K. E. Topological landscapes of porous organic cages. *Nanoscale* **9**, 5280–5298 (2017).
18. Yang, X., Ullah, Z., Stoddart, J. F. & Yavuz, C. T. Porous organic cages. *Chem. Rev.* **123**, 4602–4634 (2023).
19. Mastalerz, M. Porous shape-persistent organic cage compounds of different size, geometry, and function. *Acc. Chem. Res.* **51**, 2411–2422 (2018).
20. Wang, Q. Q. et al. Molecular barrel by a hooping strategy: synthesis, structure, and selective CO_2 adsorption facilitated by lone pair-π interactions. *J. Am. Chem. Soc.* **139**, 635–638 (2017).
21. 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).
22. Han, B. et al. Postsynthetic metalation of a robust hydrogen-bonded organic framework for heterogeneous catalysis. *J. Am. Chem. Soc.* **141**, 8737–8740 (2019).
23. Zhu, Q. et al. 3D cage COFs: a dynamic three-dimensional covalent organic framework with high-connectivity organic cage nodes. *J. Am. Chem. Soc.* **142**, 16842–16848 (2020).
24. Ji, C. et al. Tunable cage-based three-dimensional covalent organic frameworks. *CCS Chem.* **4**, 3095–3105 (2022).
25. Ma, J.-X. et al. Cage based crystalline covalent organic frameworks. *J. Am. Chem. Soc.* **141**, 3843–3848 (2019).
26. Kory, M. J. et al. Gram-scale synthesis of two-dimensional polymer crystals and their structure analysis by X-ray diffraction. *Nat. Chem.* **6**, 779–784 (2014).
27. Giri, A., Sahoo, A., Dutta, T. K. & Patra, A. Cavitand and molecular cage-based porous organic polymers. *ACS Omega* **5**, 28413–28424 (2020).
28. Zhu, Q. et al. Soft hydrogen-bonded organic frameworks constructed using a flexible organic cage hinge. *J. Am. Chem. Soc.* **145**, 23352–23360 (2023).
29. Zhang, G. & Mastalerz, M. Organic cage compounds—from shape-persistency to function. *Chem. Soc. Rev.* **43**, 1934–1947 (2014).
30. Das, S., Heasman, P., Ben, T. & Qiu, S. Porous organic materials: strategic design and structure-function correlation. *Chem. Rev.* **117**, 1515–1563 (2017).
31. Jiang, S. et al. Porous organic molecular solids by dynamic covalent scrambling. *Nat. Commun.* **2**, 207 (2011).
32. Santos, T. et al. Dynamic nucleophilic aromatic substitution of tetrazines. *Angew. Chem. Int. Ed.* **60**, 18783–18791 (2021).
33. Terrier, F. *Modern Nucleophilic Aromatic Substitution* (Wiley, 2013); https://doi.org/10.1002/9783527656141
34. Katz, J. L., Geller, B. J. & Conry, R. R. Synthesis of oxacalixarenes incorporating nitrogen heterocycles: evidence for thermodynamic control. *Org. Lett.* **8**, 2755–2758 (2006).
35. Katz, J. L. et al. Single-step synthesis of D_{3h}-symmetric bicyclooxacalixarenes. *Org. Lett.* **7**, 3505–3507 (2005).
36. Wang, D. X. et al. Versatile anion–π interactions between halides and a conformationally rigid bis(tetraoxacalix[2]arene[2]triazine) cage and their directing effect on molecular assembly. *Chem. Eur. J.* **16**, 13053–13057 (2010).
37. Wang, Z. et al. Multicolor tunable polymeric nanoparticle from the tetraphenylethylene cage for temperature sensing in living cells. *J. Am. Chem. Soc.* **142**, 512–519 (2020).
38. Zhang, G., Presly, O., White, F., Oppel, I. M. & Mastalerz, M. A permanent mesoporous organic cage with an exceptionally high surface area. *Angew. Chem. Int. Ed.* **53**, 1516–1520 (2014).
39. Hong, S. et al. Porphyrin boxes: rationally designed porous organic cages. *Angew. Chem. Int. Ed.* **54**, 13241–13244 (2015).
40. Ivanova, S. et al. Isoreticular crystallization of highly porous cubic covalent organic cage compounds. *Angew. Chem. Int. Ed.* **60**, 17455–17463 (2021).
41. Hähsler, M. & Mastalerz, M. A giant [8+12] boronic ester cage with 48 terminal alkene units in the periphery for postsynthetic alkene metathesis. *Chem. Eur. J.* **27**, 233–237 (2021).
42. Ono, K. et al. Self-assembly of nanometer-sized boroxine cages from diboronic acids. *J. Am. Chem. Soc.* **137**, 7015–7018 (2015).
43. Kory, M. J., Bergeler, M., Reiher, M. & Schlüter, A. D. Facile synthesis and theoretical conformation analysis of a triazine-based double-decker rotor molecule with three anthracene blades. *Chem. Eur. J.* **20**, 6934–6938 (2014).
44. Luo, N., Ao, Y. F., Wang, D. X. & Wang, Q. Q. Exploiting anion–π interactions for efficient and selective catalysis with chiral molecular cages. *Angew. Chem. Int. Ed.* **60**, 20650–20655 (2021).
45. Mooibroek, T. J. & Gamez, P. The *s*-triazine ring, a remarkable unit to generate supramolecular interactions. *Inorganica Chim. Acta* **360**, 381–404 (2007).
46. Turcani, L., Berardo, E. & Jelfs, K. E. stk: a Python toolkit for supramolecular assembly. *J. Comput. Chem.* **39**, 1931–1942 (2018).
47. Itoh, Y. et al. Ultrafast water permeation through nanochannels with a densely fluorous interior surface. *Science* **376**, 738–743 (2022).
48. Yoshizawa, M., Tamura, M. & Fujita, M. Diels–Alder in aqueous molecular hosts: unusual regioselectivity and efficient catalysis. *Science* **312**, 251–254 (2006).
49. Yamashina, M. et al. An antiaromatic-walled nanospace. *Nature* **574**, 511–515 (2019).
50. Willems, T. F., Rycroft, C. H., Kazi, M., Meza, J. C. & Haranczyk, M. Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous Mesoporous Mater.* **149**, 134–141 (2012).
51. Case, D. H., Campbell, J. E., Bygrave, P. J. & Day, G. M. Convergence properties of crystal structure prediction by quasi-random sampling. *J. Chem. Theory Comput.* **12**, 910–924 (2016).
52. 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).
53. Loots, L. & Barbour, L. J. An infinite catenane self-assembled by π···π interactions. *Chem. Commun.* **49**, 671–673 (2013).
54. Tozawa, T. et al. Porous organic cages. *Nat. Mater.* **8**, 973–978 (2009).
55. Ozdemir, J. et al. Covalent organic frameworks for the capture, fixation, or reduction of CO_2. *Front. Energy Res.* **7**, 77 (2019).
56. Kunde, T., Pausch, T. & Schmidt, B. M. Porous organic compounds—small pores on the rise. *Eur. J. Org. Chem.* **2021**, 5844–5856 (2021).
57. Kunde, T., Nieland, E., Schröder, H. V., Schalley, C. A. & Schmidt, B. M. A porous fluorinated organic [4+4] imine cage showing CO_2 and H_2 adsorption. *Chem. Commun.* **56**, 4761–4764 (2020).
58. Lu, C. et al. OPLS4: improving force field accuracy on challenging regimes of chemical space. *J. Chem. Theory Comput.* **17**, 4291–4300 (2021).
59. 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).
60. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
61. VandeVondele, J. & Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **127**, 114105 (2007).

62. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. 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).
63. Frisch, M. J. et al. Gaussian 16 revision a.03.2016 (Gaussian, 2016).
64. 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).
65. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **8**, 1057–1065 (2006).
66. Frisch, M. J. et al. Gaussian 09 revision D.01, C.01 (Gaussian, 2010).
67. Coombes, D. S., Price, S. L., Willock, D. J. & Leslie, M. Role of electrostatic interactions in determining the crystal structures of polar organic molecules. A distributed multipole study. *J. Phys. Chem.* **100**, 7352–7360 (1996).
68. Beyer, T. & Price, S. L. Dimer or catemer? Low-energy crystal packings for small carboxylic acids. *J. Phys. Chem. B* **104**, 2647–2655 (2000).
69. Stone, A. J. Distributed multipole analysis: stability for large basis sets. *J. Chem. Theory Comput.* **1**, 1128–1132 (2005).
70. Ferenczy, G. G. Charges derived from distributed multipole series. *J. Comput. Chem.* **12**, 913–917 (1991).
71. Hafizi, R. & Day, G. M. Supporting data for the journal article "Computationally guided synthesis of a hierarchical [4[2+3]+6] porous organic 'cage of cages'". *University of Southampton Institutional Research Repository* https://doi.org/10.5258/SOTON/D2929 (2024).

Acknowledgements

A.I.C. thanks the Royal Society for a Research Professorship (RSRP\S2\232003). C.Z. acknowledges the China Scholarship Council for financial support (202106745008). R.H. acknowledges the Iridis 5 High Performance Computing facility, and associated support services at the University of Southampton. Via our membership of the UK's HEC Materials Chemistry Consortium, which is funded by the Engineering and Physical Sciences Research Council (EPSRC) (EP/R029431 and EP/X035859), this work used the Archer2 HPC facility. We acknowledge A. Hunter for performing the MALDI-TOF analysis at the National Mass Spectrometry Facility (NMSF) at Swansea University, and Diamond Light Source for access to beamlines I19 (CY30461). We received funding from the EPSRC (EP/V026887/1) and the Leverhulme Trust via the Leverhulme Research Centre for Functional Materials Design. This project has received funding from the European Research Council under the European Union's Horizon 2020 Research and Innovation programme (grant CoMMaD number 758370).

Author contributions

Q.Z. led the experimental work and the synthesis and characterization of the materials. A.I.C., K.E.J. and M.A.L. conceived the idea and modelling strategy with Q.Z. and supervised the project. H.Q. and M.A.L. conducted the single-crystal X-ray diffraction analysis and solved the structure. G.A., K.E.J. and C.Z. performed the molecular simulations. R.H. and G.M.D. performed the CSP, and G.M.D. supervised this part of the project. Q.Z., G.A., R.H., G.M.D., K.E.J., M.A.L. and A.I.C. analysed the data and prepared the paper. All authors discussed the results and contributed to the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44160-024-00531-7.

Correspondence and requests for materials should be addressed to Marc A. Little or Andrew I. Cooper.

Peer review information *Nature Synthesis* thanks Chenfeng Ke, Bernhard Schmidt and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Alison Stoddart, in collaboration with the *Nature Synthesis* team.