Computationally-Inspired Discovery of an Unsymmetrical Porous Organic Cage

Enrico Berardo,[†,§] Rebecca L. Greenaway,[‡,§] Lukas Turcani,[†] Ben M. Alston,[‡] Michael J. Bennison,[‡] Marcin Miklitz,[†] Rob Clowes,[‡] Michael E. Briggs,[‡] Andrew I. Cooper,[‡] Kim E. Jelfs[†*]

[†] Department of Chemistry, Imperial College London, South Kensington, London, SW7 2AZ, United Kingdom

[‡] Department of Chemistry and Materials Innovation Factory, University of Liverpool, 51 Oxford Street, Liverpool, L7 3NY, United Kingdom

[§]E.B. and R.L.G. contributed equally.

Email: k.jelfs@imperial.ac.uk

Abstract

A completely unsymmetrical porous organic cage was synthesized from a C_{2v} symmetrical building block that was identified by a computational screen. The cage was formed through a 12-fold imine condensation of a tritopic C_{2v} symmetric trialdehyde with a di-topic C_2 symmetric diamine in a [4+6] reaction. The cage was rigid and microporous, as predicted by the simulations, with an apparent Brunauer-Emmett-Teller surface area of 578 m^2 g^{-1}. The reduced symmetry of the tritopic building block relative to its topicity meant there were 36 possible structural isomers of the cage. Experimental characterization suggests a single isomer with 12 unique imine environments, but techniques such as NMR could not conclusively identify the isomer. Computational structural and electronic analysis of the possible isomers was used to identify the most likely candidates, and hence to construct a 3-dimensional model of the amorphous solid. The rational design of unsymmetrical cages using building blocks with reduced symmetry offers new possibilities in controlling the degree of crystallinity, porosity, and solubility, of self-assembled materials.

Introduction

Porous molecular materials, such as porous organic cages (POCs), are a sub-class of porous materials that lack the 3-dimensional extended bonding found in networks such as zeolites, metal-organic frameworks (MOFs), and covalent organic frameworks (COFs). POCs are molecules that have intrinsic porosity as a result of an internal cavity that can be accessed through multiple windows.[1,2] In the solid state, these molecules can pack together to afford a 1-, 2-, or 3-D pore network that passes through the internal cage cavities; the intrinsic porosity of the cage may be augmented by extrinsic voids located between the cages. POCs have shown potential in applications such as chemical separations,[3] sensing[4,5] and as porous liquids.[6] Because of their molecular nature, one can control both crystal phase, the degree of crystallinity and, hence, the porosity by post-synthetic processing.[7–12] This can be done by varying the solvent, the crystallization temperature, the mixing rate, or by modifying the desolvation process. For example, cages can be directed to form 3-D interconnected pore networks simply by changing the crystallization solvent,[13] or made intentionally amorphous using techniques such as freeze-drying.[7]

Conventionally, porous materials—both molecular solids and frameworks—tend to be synthesized from highly symmetric precursors to generate crystalline structures. The use of building blocks with reduced symmetry to construct such materials results in a larger number of possible structural isomers, and thus a greater phase space of structural possibilities. This makes purposeful design based on chemical knowledge much more challenging than for more symmetrical precursors. Mukherjee *et al*. have recently observed a self-selection process between multiple structural isomers when an unsymmetrical ditopic building block was employed for the synthesis of imine cages,[14] but such processes are very hard to anticipate. In the area of frameworks, Zhang *et al*. showed that it was possible to obtain COFs with heterogeneous pore structures by using a C_{2v} tritopic building block.[15] These studies suggest that introducing asymmetric building blocks, or reducing the overall symmetry of the building blocks, has the potential to create increased structural complexity and hence access to novel properties. For porous molecular materials, reduced symmetry has the potential to increase porosity of the materials, by frustrating efficient packing, or to dramatically improve solubility.

However, this increased structural complexity for less symmetrical materials presents a major challenge for purposeful design, for experimental characterization methods, and for atomistic simulations. Figure 1a illustrates the structural isomerism that is generated for an assembly when we reduce the precursor symmetry. By reducing the symmetry of a tritopic building block on each vertex of a tetrahedral structural model to C_{2v}, the number of potential structural isomers increases. In the presence of a strong driving force (thermodynamic or kinetic), synthesis might favor a specific structural isomer. In the absence of such a strong preference, a statistical mixture of isomers may be expected.[14] It is extremely difficult to pre-judge which of the possible isomers, if any, will be favored by experiment. As a result, it is difficult to design these less symmetrical systems, for example by making analogies with known polyhedral cages.

Here we present the formation of a POC that was synthesized from a tritopic aldehyde building block of reduced symmetry (C_{2v}) with a ditopic amine building block. This building block was suggested by a large-scale computational screen. This demonstrates the power of such strategies because we would not otherwise have selected this reduced symmetry precursor based on existing chemical knowledge. Compared to related cages formed from more symmetrical precursors, this cage is unsymmetrical and, hence, highly soluble (Fig. 1b). The structure of the unsymmetrical cage could not be determined definitively by experiment, and further computational modelling was therefore used

to assist in its identification. As such, computation was key to both the genesis *and* the characterization of this unusual porous organic cage.

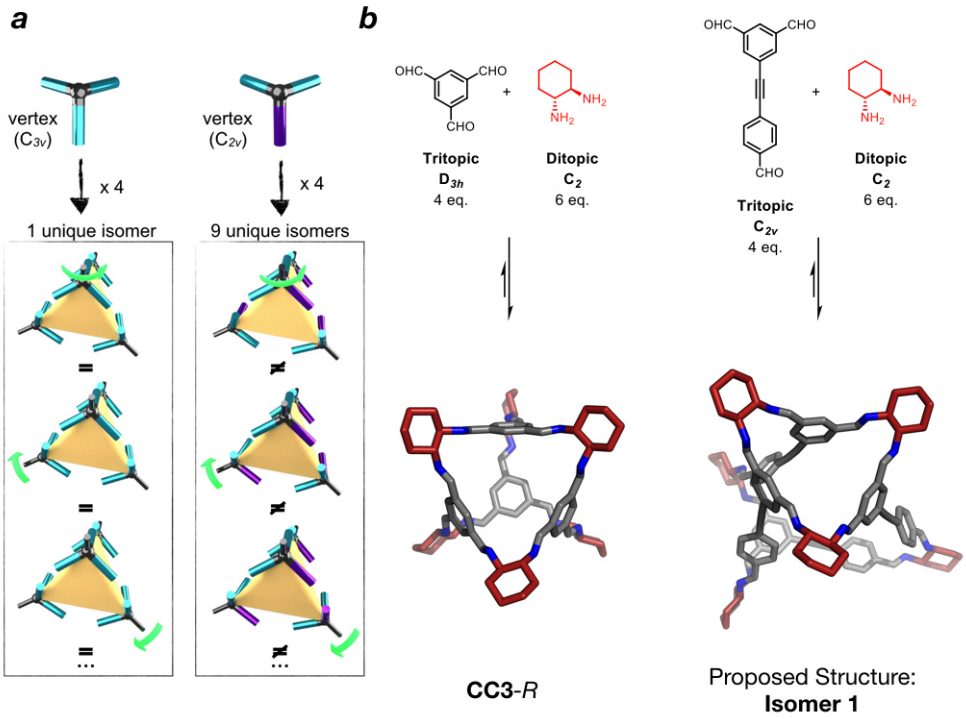


Figure 1. Structural complexity generated by cage precursors with reduced symmetry. (a) Scheme for a tetrahedral topology with a tritopic precursor on each of its four vertices. With a C_{3v} symmetric precursor, a 120° rotation about the C_3 axis of the tetrahedron results in no new structural isomers of the assembly. By contrast, with a C_{2v} symmetric precursor, each rotation creates a new structural isomer of the assembly: a total of 9 symmetry-inequivalent isomers results for this example; (b) Formation of symmetrical **CC3** using a tritopic D_{3h} symmetric building block, and the proposed structure of the unsymmetrical cage species formed using the tritopic C_{2v} building block. The D_{3h} and C_{2v} symmetry assignment of the tritopic precursors only describes the symmetry of the substituted aromatic cores and does not take into account the aldehyde orientation.

Results and Discussion

This work started with a computational screen where we assembled 10,000 combinations of tritopic aldehydes and ditopic amines from a subset of the Reaxys database[16] into [4+6] cages. From this screen, we identified 5-(4-formylphenyl)-isophthalaldehyde (trialdehyde **1**) as an interesting potential precursor. This screening procedure used our supramolecular toolkit (*stk*)[17] to automate the cage assembly and then to test candidate cages for "shape-persistence" by running a short gas phase molecular dynamics simulation, checking if an internal cavity was retained. If the candidate cage contained a cavity that was large enough to accommodate a nitrogen molecule (kinetic diameter 3.84 Å), and all four windows had a spherical diameter at least large enough to accommodate a hydrogen molecule (kinetic diameter 2.8 Å), then the cage was deemed shape-persistent. After discarding those cages that lacked shape persistence and collapsed (~95% of combinations), we then inspected the results for promising synthetic candidates. Shape-persistent cages formed from trialdehyde **1** caught our attention: these were not structures that we would have otherwise designed, and they offered the

potential to form an unsymmetrical cage. This precursor also led to predicted shape-persistent cages by *in silico* combination with different diamine partners (3 shape-persistent combinations predicted in total). Of the 84 potential diamine partners selected from the Reaxys database, the C_2 symmetric (1*R*,2*R*)-cyclohexane-1,2-diamine ((*R,R*)-CHDA) was particularly promising and commercially available. This diamine has also been used previously to synthesize a number of POCs, including **CC3**, formed with the tritopic aldehyde, (Fig. 1b),[18] and **TCC1-3**, formed with tetratopic aldehydes.[19]

A high-throughput (HT) experimental screen of trialdehyde **1** with 16 diamine partners (see further details in the Supporting Information, section 3), also identified (*R,R*)-CHDA (Fig. 1b) as a promising cage-forming partner, with the formation of a [4+6] cage indicated by high-resolution mass spectrometry (HRMS) (Fig. S5).

Optimization and scale-up of the reaction between trialdehyde **1** and (*R,R*)-CHDA resulted in the formation of a [4+6] cage molecule with 96% mass recovery and 91% (a/a) purity as analyzed by HPLC. Based on our previous experience, the observed stability of the cage to solvent removal suggested that the cage might be shape-persistent, because cages that lack shape-persistence quite often also exhibit some chemical decomposition along with collapse of the cage cavity.[20,21] A single peak in the HPLC suggested the formation of a single cage species (Fig. S12). However, the ^{1}H NMR spectra was very complex by comparison with spectra usually observed for symmetrical cages (Fig. 2b, Fig. S45-S46): for example, the [4+6] cage **CC3** (Fig. 1b) displays a single chemical shift for all 12 imine protons and a single chemical shift for the 12 aromatic protons. This complex NMR suggested that the reaction with trialdehyde **1** either forms a mixture of cage isomers, or a single highly unsymmetrical cage isomer. To aid characterization, further purification was carried out by preparative-HPLC to afford the unsymmetrical cage with >98% (a/a) purity. Despite the complexity of the ^{1}H NMR spectra, which could suggest the presence of cage isomers, the integration of the peaks as whole numbers before and after purification suggested instead that a single cage had formed, rather than a mixture of species. To further confirm this, diffusion NMR was used to determine both the size of the cage in solution and to establish whether a mixture of cages with varying shapes (and sizes) had formed. The measured solvodynamic diameter of the cage was 18.6 Å, which is consistent, by comparison with previous measurements, with a [4+6] cage containing an internal cavity (Table S11). A single size of cage species was suggested because all peaks had identical diffusion co-efficients, as illustrated by the ^{1}H DOSY spectra (Fig. 2d, Table S4).

We next sought to gain more insight into the cage structure. Discrete peaks for most of the 'structurally equivalent' carbons in the cage were observed in the ^{13}C NMR spectra (Fig. 2c) in particular, different shifts were observed for all 12 imine carbons as well as the 12 carbons on the cyclohexyl ring adjacent to the imines. We also observed discrete resonances for each of the other aromatic and quaternary carbons originating from the trialdehyde **1**. The use of 2D NMR (for COSY, HSQC and HMBC spectra see Fig. S15-S17) assisted in the assignment of both the ^{1}H and ^{13}C NMR spectra, alongside allowing the identification of four individual 'faces' in the cage species (Fig. S18-S21), with 1D-NOEs enabling the assignment of the 12 different corresponding imine protons (see Fig. S23-S27). Furthermore, there was evidence of connectivity between these faces by through-space interactions observed in the 1D-NOEs – see Fig. S28. This suggests the formation of a single [4+6] cage that has no symmetry (C_1 symmetry).

Multiple attempts at crystallizing the unsymmetrical cage from different solvents were unsuccessful (for details see the Supporting Information), with powder X-ray diffraction (PXRD) indicating the formation of an amorphous solid (Fig. S31). This is likely due to the lack of an extended packing motif - using comparable crystallization conditions, a series of porous organic cages with tetrahedral,[22] or

trigonal prismatic symmetry,[23] with strong 3-D and 1-D packing motifs respectively, can be routinely crystallised. This indicates that our low symmetry precursor results in a cage that hampers crystallization.

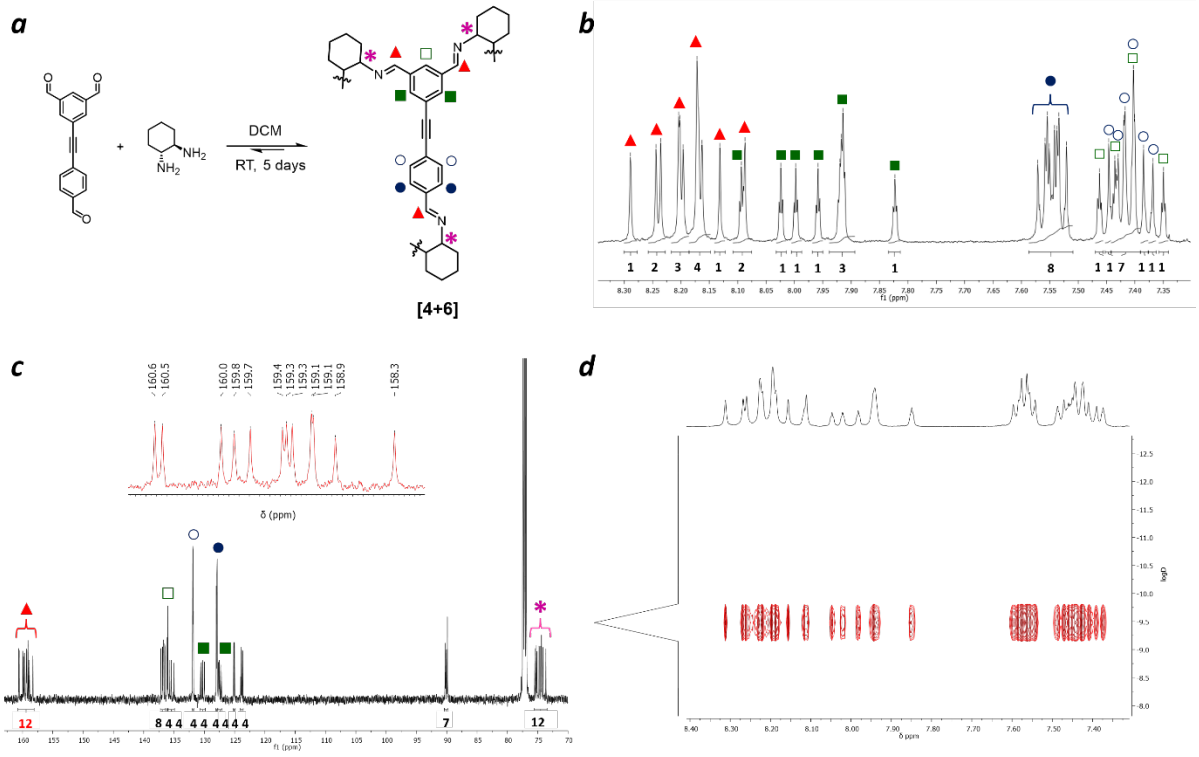


Figure 2 Experimental characterization of the unsymmetrical cage. (a) Reaction scheme; (b) ^{1}H NMR spectra of the isolated cage, showing the aromatic region with the different peak shifts for the protons originating from trialdehyde 1 indicating different chemical environments, with the integrations suggesting the presence of a single species; (c) ^{13}C NMR spectra showing the unsymmetrical nature of the cage inset (red) shows the 12 different shifts for the imine carbons; (d) ^{1}H DOSY spectra of the isolated cage suggesting a single size of cage species.

The amorphous solid was found to be microporous, with an apparent Brunauer-Emmett-Teller surface area (SA_{BET}) of 578 m^2 g^{-1}. The N_2 and H_2 uptake was 13.1 mmol g^{-1} and 4.7 mmol g^{-1}, respectively (1 bar, 77 K) (Fig. S33). This suggested to us the presence of an intrinsic cavity in the cage, based on our prior experience of amorphous cage solids formed from more symmetrical cages where the cavity collapses to form a non-porous material.[20]

Taking the experimental data as a whole suggests the formation of a single, completely unsymmetrical, shape-persistent [4+6] cage. To our knowledge, this is the first example of its kind among POCs. The cage was found to be extremely soluble (>400 mg/mL in $CHCl_3$) compared to previously reported symmetrical cages (see Fig. S35), for example, **CC3** has ~3 mg/ml solubility in $CHCl_3$.[24] This is presumably because of disruption of the lattice packing in this highly unsymmetrical molecule, and highlights the potential advantage to exploring lower symmetry cage molecules.

Without a single crystal to definitively solve the structure of the cage molecule, we again turned to computational modeling, following the computational pipeline summarized in Figure 3. There are two possible cage topologies, $\mathbf{Tri^4Di^6}$ and $\mathbf{Tri_2^4Di^6}$,[25] that would lead to the [4+6] molecular mass ion

identified by HRMS ([M+2H]$^{2+}$ 759.9211 g mol^{-1}; Fig. S29). The **Tri^{4}Di^6** molecule would consist of four tritopic and six ditopic building blocks. The subscript "2" in the $\mathbf{Tri_2^4Di^6}$ topology nomenclature indicates that the tritopic building blocks are connected twice to another tritopic block (see image in Fig. 3 and Supporting Information Section 5 for a full description).[25] For each of these two topologies, we generated all possible structural isomers by rotation of the C_{2v} tritopic building block around the topology's C_3 rotation axis. This corresponds to 3^4=81 structural isomers for each of the 2 topologies. The resulting 162 isomers were then reduced by removing symmetry equivalent structures to afford 36 unique isomers, of which 9 were obtained from the **Tri^{4}Di^6** topology and 27 from the $\mathbf{Tri_2^4Di^6}$ topology (Figure 3). For details on the methods used to determine structurally unique isomers, see Supplementary Information Section 5.3.

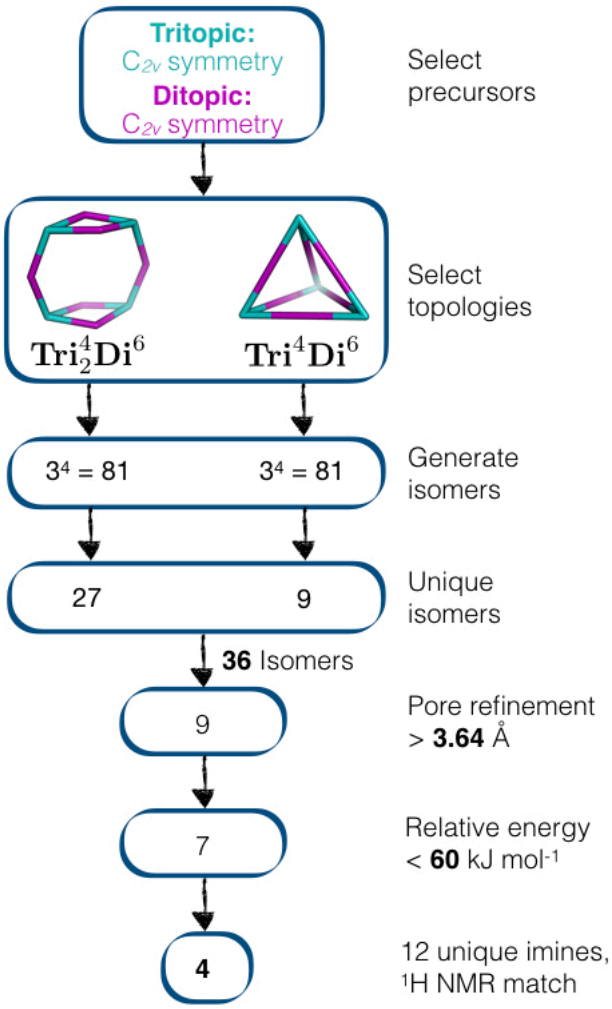


Figure 3 Computational pipeline for the generation and screening of isomers. 162 structural isomers were assembled in 2 topologies. Through a combination of experimental data and computational predictions, these were filtered down to 4 unique isomers. Teal and purple connectors correspond to tritopic and ditopic building blocks, respectively.

To obtain good structural models for these 36 cage isomers, we first geometry optimized each model using the OPLS3 force field,[26] which we have previously found to reproduce the structure and energetics of cage molecules.[25,27] We then used simulated annealing to search for low energy

conformers (see Supporting Information Section 5.4. for full details). The resulting lowest energy conformer for each isomer was then re-optimised at Density Functional Theory (DFT) level (B3LYP-D3/Def2-TZVP),[28–33] and finally a single point M06-2X/Def2-TZVP calculation was conducted to obtain accurate total energies.[33,34] A polarizable continuum model with a dielectric value of 8.93 for dichloromethane was employed in the last step to reproduce the electrostatic influence of solvent on the final energy of the various isomers.

The final structures of the 36 cage isomers were then investigated for structural properties such as pore size and average weighted diameter. Both properties were calculated with the use of pyWindow,[35] a python package we have developed for the structural analysis of discrete porous molecules (see the Supporting Information for full details). To identify the most likely structural isomer of our synthesized cage, we applied the following criteria to successively narrow down the possibilities: (i) remove cages that are collapsed or that have cavities smaller than 3.64 Å, which corresponds to the kinetic diameter of N_2, on the basis that sorption analyses suggest that the molecule contains an intrinsic void large enough to accommodate N_2; (ii) remove candidates that fall outside the energy range of 60 kJ mol^{-1} from the lowest energy remaining isomer—this assumes that the lowest energy isomers are the most likely to be obtained experimentally (*i.e.*, thermodynamic control),[36] but by having a large window, we also consider that there may be kinetic bottlenecks in the synthesis;[37] (iii) select only the isomers that are completely unsymmetrical and that therefore contain 12 unique imine environments, to match NMR observations; and, finally; (iv) select the structural isomer with the lowest mean absolute error (MAE) between calculated and experimental ^{1}H NMR chemical shifts.

The first criterion, pore size, reduced the number of potential isomers from 36 to 9, which ranged in predicted pore diameter from 4.2 to 7.2 Å. The next criterion, relative energy, reduced the number of potential isomers from 9 to 7. Consideration of the number of unique imine environments left 4 potential isomers. The structural and electronic properties for the final selected isomers are shown in Table 1 and Fig. 4. A full discussion of how we determined the number of unique imine environments can be found in the Supporting Information Section 5.3.

Table 1 Simulated properties of the four selected cage isomers.

Cage isomer	Pore diameter (Å)	Average weighted diameter (Å)	Relative energy (kJ mol^{-1})†	MAE ^{1}H NMR (ppm)
1	6.8	16.1	18	0.085
2	6.7	16.7	39	0.113
3	7.2	16.6	40	0.104
4	5.4	16.3	47	0.116

†Energy is relative to the lowest energy isomer among the 36 initial isomers (not shown in table).

All four of the selected cage isomers possessed the **Tri**^{4}**Di**^{6} topology, since the $\mathbf{Tri}_2^4\mathbf{Di}^6$ topology tended to lead to assemblies with higher symmetry. The calculated average weighted diameters for isomers **1**–**4** range from 16.1 to 16.3 Å. This is slightly lower than the experimental solvodynamic diameter of 18.6 Å, but these differences between computed and measured diameters are consistent with observations for other organic cages for which we know the structure (Table S11).

For the fourth criterion, we compared the experimental ^{1}H NMR chemical shifts with ^{1}H NMR shielding tensors calculated at the WP04/DGTZVP level of theory in the solvated phase ($CDCl_3$) for the remaining 4 isomers.[38,39] This level of theory, followed by scaling through the use of statistical parameters, was shown by Benassi to perform exceptionally well for a test set of molecules.[40]

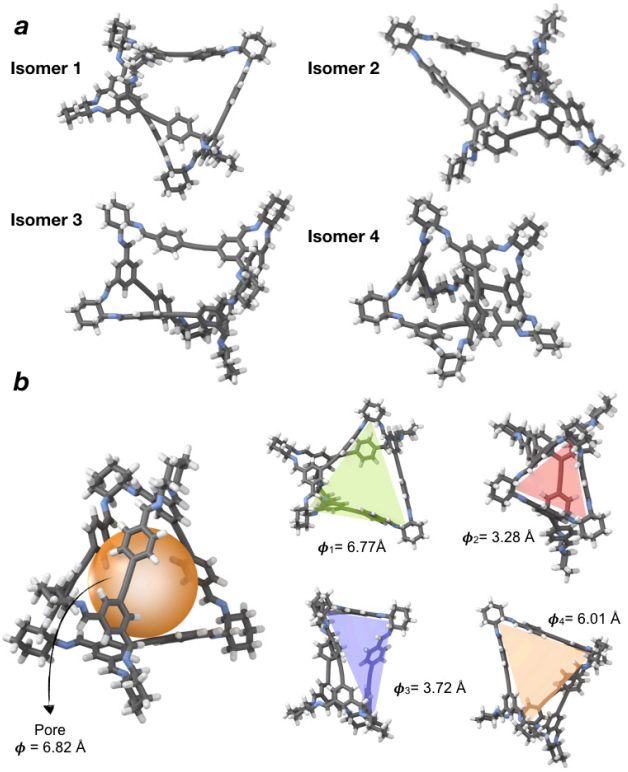


Figure 4 (a) DFT optimized structures for the 4 most likely isomers of the experimentally synthesized unsymmetrical cage. The isomers are identified by numbers **1**–**4**, with **1** being the lowest energy isomer and **4** the highest. Carbon, nitrogen, and hydrogen atoms are represented as grey, blue, and white sticks, respectively; **(b)** Structural properties of isomer 1. The 6.8 Å diameter pore is shown as an orange sphere on the left. On the right the diameter of the largest sphere that fits in each of the 4 windows for **1** is given. Transparent triangles have been used to highlight each distinct window for visualization purposes only.

We were only able to consider a single conformation here, rather than the recommended conformational averaging,[40] due to the large size of these molecules (216 atoms). To determine the goodness of the fit for each isomer relative to the experimental values, we calculated the mean absolute error (MAE) for each isomer,[41] as shown in Table 1. The predicted MAE values for the ^{1}H NMR chemical shifts indicate that our fits (~0.10 ppm) are in line with the errors obtained for much smaller molecular systems.[40,41] The analysis of the calculated MAE suggests that among the four candidate isomers, isomer **1** has the best fit with the experimental ^{1}H NMR spectrum (MAE 0f 0.085 ppm), albeit by a small margin. Isomer **1**, shown in Fig. 4b, has a large internal cavity (fitting a sphere of 6.8 Å diameter), four unsymmetrical windows of diameter 6.8, 6.0, 3.7, and 3.3 Å, an average weighted diameter of 16.1 Å, and the lowest calculated energy (18 kJ mol^{-1}) of the final four selected isomers.

Finally, we then compared the experimentally determined configuration of the meta-substituted imines to those in the four candidate isomers. Using a combination of 2D NMR (COSY) and 1D-NOEs, it was possible to assign an 'up'/'down' configuration on each individual face (see Fig. S19, S23-24 and

S26-27). On comparison of this with the four isomers, isomer **1** was the only cage that demonstrated the same configuration (see Fig. S48). We therefore suggest that the most likely identity of our experimental cage is isomer **1**. Further information for the remaining 3 isomers can be found in the Supporting Information. To verify that isomer **1** would indeed form a porous system in the solid state, we generated an amorphous model from its molecular structure. Briefly, we loaded a simulation cell with 100 molecules at low density, before conducting a 21-step relaxation procedure[42] using a bespoke forcefield, CSFF[43] (see the Supporting Information Section 5.8. for full details).[11,44] The resulting amorphous structure is shown in Fig. 5.

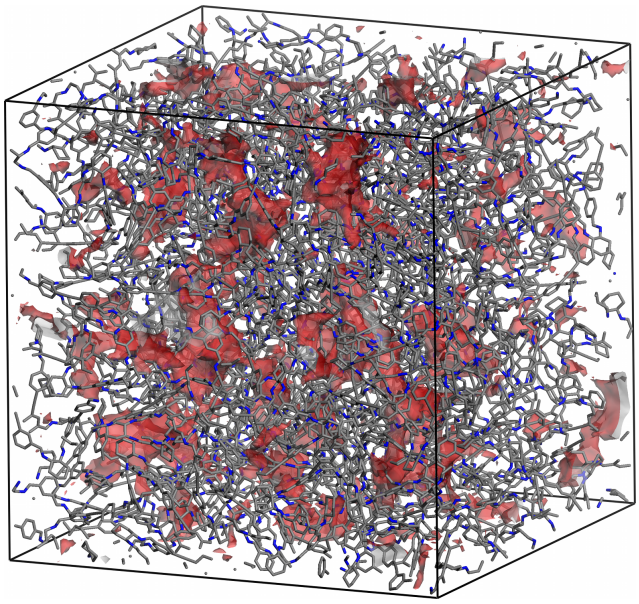


Figure 5 Amorphous model of isomer **1**. Voids within the structure are shown in red, calculated with a probe of radius 1.55 Å.

To evaluate the accessible surface area for this model, we used a probe radii corresponding to the van der Waals radius of N_2 (1.55 Å) and to its kinetic radius (1.82 Å). In this amorphous model, whilst the largest cavity diameter is 8.38 Å, which is larger than that of the cage (6.82 Å), showing the potential benefits of generating extrinsic porosity in a system through poor packing. The calculated pore limiting diameter of 3.09 Å is close to the diameter of N_2: it is therefore reasonable to expect that thermal fluctuations can open up pores that are not formally accessible in static models, and we have shown previously that such "dynamic" porosity occurs during molecular dynamics simulations of other POC systems.[45–47] Here, we calculated a surface area of 456 and 837 $m^2\ g^{-1}$ with probe radii of 1.82 and 1.55 Å, respectively, which spans the observed SA_{BET} of 578 $m^2\ g^{-1}$. This further validates that isomer **1** is a plausible identity for this cage. For further details and structural properties for the amorphous model of isomer **1**, see Supplementary Information Table S10 and Fig. S44 in Section 5.8.

Intriguingly, when we analyzed the cage molecules in the amorphous model, we found that 83% of the potential intrinsic void volume inside the cages was lost due to interpenetration from neighboring cages. This was due to the large windows in isomer **1**; as such, the majority of the porosity in this model comes from extrinsic voids, not from cage cavities. Reassessing our earlier selection criterion that the isomer must have an internal cavity (diameter >3.64 Å), we found that the same four final candidates are still found in the absence of this criterion (Fig. 6).

We also note that all four final candidates have at least two windows that are large enough for penetration by cyclohexanes from neighboring cages: as such, we believe that it is highly likely that the porosity in this cage results primarily from extrinsic voids, even if our selection of isomer **1** proves incorrect. This does not mean the shape-persistent intrinsic voids are unimportant, however, because they likely promote a greater degree of interconnectivity between the extrinsic pores.

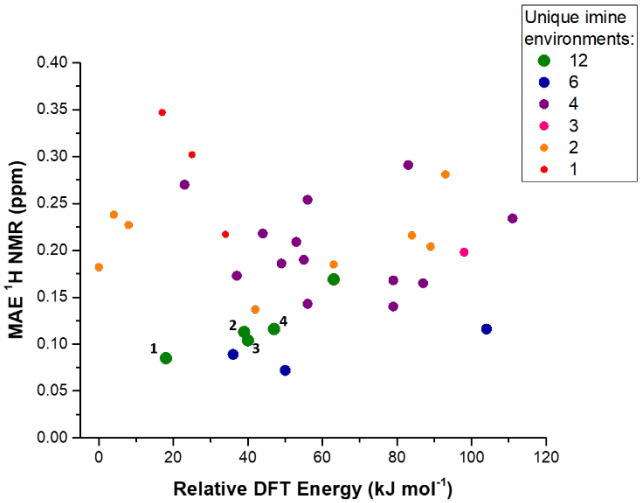


Figure 6 Comparison of the 36 possible cage isomers, taking into account the final criteria: (i) relative DFT energy <60 kJ mol^{-1}, (ii) 12 unique imine environments, and (iii) low MAE between the experimental and calculated ^{1}H NMR spectra. The final four candidate isomers **1-4** are labelled, with isomer **1** having both the lowest relative energy and lowest mean absolute error (MAE) of the ^{1}H NMR shifts.

Conclusions

In summary, we have successfully demonstrated the dynamic covalent synthesis of a unique, fully unsymmetrical **Tri^{4}Di^6** imine cage. We would not have designed this cage without the results of computational screening, which inspired the choice of the less symmetrical precursor. The amorphous solid obtained was porous to both N_2 and H_2, and had greatly increased solubility compared to both symmetrical cages and also most 'scrambled' cage mixtures, the latter of which have been used to generate porous liquids.[48] We show that simulations can complement experimental results and provide atomistic insights when it is not possible to obtain definitive structural information by crystallography. By exploiting the close synergy between molecular structure prediction, experimental synthesis, and structural modelling, we were able to tackle the challenging problem of postulating a structure for a cage that has 162 possible isomers.

Acknowledgement

We acknowledge funding from the European Research Council under FP7 (RobOT, ERC Grant Agreement No. 321156; CoMMaD, ERC Grant Agreement No. 758370), and the Engineering Research Council and Physical Sciences Research Council (EPSRC) (EP/M017257/1, EP/Poo5543/1, EP/N004884/1), including the UK's HEC Materials Chemistry Consortium (EP/L000202/1) for time on

the UK supercomputer, ARCHER**.** K.E.J. thanks the Royal Society for a University Research Fellowship. E.B. thanks Dr. Henry Rzepa for his help with the simulation of NMR chemical shifts and for fruitful discussions. We thank Dr. Marc Little for discussions relating to the crystallization of the unsymmetrical cage, the MicroBioRefinery for assistance with QTOF-MS measurements, and Dr. Konstantin Luzyanin for assistance with the NMR measurements.

Supporting Information

The supporting information contains the full high throughput synthetic and experimental details, gas sorption, TGA, and NMR data, together with computational methodology details, structural and electronic properties for all the 36 isomers, structures of all the 36 DFT optimized isomers and the amorphous structure model.

Conflicts of interest

There are no conflicts of interest to declare.

References

(1) Hasell, T.; Cooper, A. I. Porous Organic Cages: Soluble, Modular and Molecular Pores. *Nat. Rev. Mater.* **2016**, *1*, 16053.

(2) Beuerle, F.; Gole, B. Covalent Organic Frameworks and Cage Compounds: Design and Applications of Polymeric and Discrete Organic Scaffolds. *Angew. Chemie - Int. Ed.* **2018**, *57* (18), 4850–4878.

(3) Zhang, J.-H.; Xie, S.-M.; Chen, L.; Wang, B.-J.; He, P.-G.; Yuan, L.-M. Homochiral Porous Organic Cage with High Selectivity for the Separation of Racemates in Gas Chromatography. *Anal. Chem.* **2015**, *87* (15), 7817–7824.

(4) Brutschy, M.; Schneider, M. W.; Mastalerz, M.; Waldvogel, S. R. Direct Gravimetric Sensing of GBL by a Molecular Recognition Process in Organic Cage Compounds. *Chem. Commun.* **2013**, *49* (75), 8398–8400.

(5) Acharyya, K.; Mukherjee, P. S. A Fluorescent Organic Cage for Picric Acid Detection. *Chem. Commun.* **2014**, *50* (99), 15788–15791.

(6) Giri, N.; Del Pópolo, M. G.; Melaugh, G.; Greenaway, R. L.; Rätzke, K.; Koschine, T.; Pison, L.; Gomes, M. F. C.; Cooper, A. I.; James, S. L. Liquids with Permanent Porosity. *Nature* **2015**, *527*, 216.

(7) Hasell, T.; Chong, S. Y.; Jelfs, K. E.; Adams, D. J.; Cooper, A. I. Porous Organic Cage Nanocrystals by Solution Mixing. *J. Am. Chem. Soc.* **2012**, *134* (1), 588–598.

(8) Patil, R. S.; Banerjee, D.; Zhang, C.; Thallapally, P. K.; Atwood, J. L. Selective CO2 Adsorption in

a Supramolecular Organic Framework. *Angew. Chem. Int. Ed.* **2016**, *55* (14), 4523–4526.

(9) Tian, J.; Ma, S.; Thallapally, P. K.; Fowler, D.; McGrail, B. P.; Atwood, J. L. Cucurbit[7]uril: An Amorphous Molecular Material for Highly Selective Carbon Dioxide Uptake. *Chem. Commun.* **2011**, *47* (27), 7626–7628.

(10) Tian, J.; Thallapally, P. K.; Dalgarno, S. J.; McGrail, P. B.; Atwood, J. L. Amorphous Molecular Organic Solids for Gas Adsorption. *Angew. Chem. Int. Ed.* **2009**, *48* (30), 5492–5495.

(11) Jiang, S.; Jelfs, K. E.; Holden, D.; Hasell, T.; Chong, S. Y.; Haranczyk, M.; Trewin, A.; Cooper, A. I. Molecular Dynamics Simulations of Gas Selectivity in Amorphous Porous Molecular Solids. *J. Am. Chem. Soc.* **2013**, *135* (47), 17818–17830.

(12) Schneider, M. W.; Lechner, L. G.; Mastalerz, M. Uniform Porous Nanospheres of Discrete Shape-Persistent Organic Cage Compounds. *J. Mater. Chem.* **2012**, *22* (15), 7113–7116.

(13) Hasell, T.; Culshaw, J. L.; Chong, S. Y.; Schmidtmann, M.; Little, M. A.; Jelfs, K. E.; Pyzer-Knapp, E. O.; Shepherd, H.; Adams, D. J.; Day, G. M.; et al. Controlling the Crystallization of Porous Organic Cages: Molecular Analogs of Isoreticular Frameworks Using Shape-Specific Directing Solvents. *J. Am. Chem. Soc.* **2014**, *136* (4), 1438–1448.

(14) Acharyya, K.; Mukherjee, P. S. Shape and Size Directed Self-Selection in Organic Cage Formation. *Chem. Commun.* **2015**, *51* (20), 4241–4244.

(15) Zhu, Y.; Wan, S.; Jin, Y.; Zhang, W. Desymmetrized Vertex Design for the Synthesis of Covalent Organic Frameworks with Periodically Heterogeneous Pore Structures. *J. Am. Chem. Soc.* **2015**, *137* (43), 13772–13775.

(16) Riniker, S.; Landrum, G. A. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *J. Chem. Theory Comput.* **2015**, *55* (12), 2562–2574.

(17) Turcani, L.; Berardo, E.; Jelfs, E. K. STK: A Python Toolkit for Supramolecular Assembly. ChemRxiv 2018.

(18) Tozawa, T.; Jones, J. T. A.; Swamy, S. I.; Jiang, S.; Adams, D. J.; Shakespeare, S.; Clowes, R.; Bradshaw, D.; Hasell, T.; Chong, S. Y.; et al. Porous Organic Cages. *Nat. Mater.* **2009**, *8*, 973.

(19) Slater, A. G.; Little, M. A.; Pulido, A.; Chong, S. Y.; Holden, D.; Chen, L.; Morgan, C.; Wu, X.; Cheng, G.; Clowes, R.; et al. Reticular Synthesis of Porous Molecular 1D Nanotubes and 3D Networks. *Nat. Chem.* **2016**, *9*, 17.

(20) Jelfs, K. E.; Wu, X.; Schmidtmann, M.; Jones, J. T. A.; Warren, J. E.; Adams, D. J.; Cooper, A. I. Large Self-Assembled Chiral Organic Cages: Synthesis, Structure, and Shape Persistence. *Angew. Chem. Int. Ed.* **2011**, *50* (45), 10653–10656.

(21) Briggs, M. E.; Jelfs, K. E.; Chong, S. Y.; Lester, C.; Schmidtmann, M.; Adams, D. J.; Cooper, A. I. Shape Prediction for Supramolecular Organic Nanostructures: [4 + 4] Macrocyclic Tetrapods. *Cryst. Growth Des.* **2013**, *13* (11), 4993–5000.

(22) Jones, J. T. A.; Hasell, T.; Wu, X.; Bacsa, J.; Jelfs, K. E.; Schmidtmann, M.; Chong, S. Y.; Adams, D. J.; Trewin, A.; Schiffman, F.; et al. Modular and Predictable Assembly of Porous Organic Molecular Crystals. *Nature* **2011**, *474*, 367–371.

(23) Slater, A. G.; Little, M. A.; Pulido, A.; Chong, S. Y.; Holden, D.; Chen, L.; Morgan, C.; Wu, X.; Cheng, G.; Clowes, R.; et al. Reticular Synthesis of Porous Molecular 1D Nanotubes and 3D Networks. *Nat. Chem.* **2017**, *9* (1), 17–25.

(24) Slater, A. G.; Little, M. A.; Briggs, M. E.; Jelfs, K. E.; Cooper, A. I. A Solution-Processable Dissymmetric Porous Organic Cage. *Mol. Sys. Des. Eng.* **2018**.

(25) Santolini, V.; Miklitz, M.; Berardo, E.; Jelfs, K. E. Topological Landscapes of Porous Organic Cages. *Nanoscale* **2017**, *9* (16), 5280–5298.

(26) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281–296.

(27) Jelfs, K. E.; Eden, E. G. B.; Culshaw, J. L.; Shakespeare, S.; Pyzer-Knapp, E. O.; Thompson, H. P. G.; Bacsa, J.; Day, G. M.; Adams, D. J.; Cooper, A. I. In Silico Design of Supramolecules from Their Precursors: Odd-Even Effects in Cage-Forming Reactions. *J. Am. Chem. Soc.* **2013**, *135* (25), 9307–9310.

(28) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.

(29) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789.

(30) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200–1211.

(31) 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.* **2010**, *132*.

(32) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8* (9), 1057–1065.

(33) 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.* **2005**, *7* (18), 3297–3305.

(34) 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.* **2008**, *120* (1), 215–241.

(35) Miklitz, M.; Jelfs, K. E., pyWindow. *ChemRxiv*, **2018**, DOI: 10.26434/chemrxiv.6850109.

(36) Greenaway, R. L.; Santolini, V.; Bennison, M. J.; Alston, B. M.; Stackhouse, C.; Little, M. A.; Miklitz, M.; Eden, E. G. B.; Clowes, R.; Shakil, A.; et al. High-Throughput Discovery of Organic Cages and Catenanes Using Computational Screening Fused with Robotic Synthesis. *Nat. Commun.* No. 2018, 1–27.

(37) Moneypenny, T. P.; Yang, A.; Walter, N. P.; Woods, T. J.; Gray, D. L.; Zhang, Y.; Moore, J. S. Product Distribution from Precursor Bite Angle Variation in Multitopic Alkyne Metathesis: Evidence for a Putative Kinetic Bottleneck. *J. Am. Chem. Soc.* **2018**, *140* (17), 5825–5833.

(38) Wiitala, K. W.; Cramer, C. J.; Hoye, T. R. Comparison of Various Density Functional Methods for Distinguishing Stereoisomers Based on Computed 1H or 13C NMR Chemical Shifts Using Diastereomeric Penam β-Lactams as a Test Set. *Magn. Reson. Chem.* **2007**, *45* (10), 819–829.

(39) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. Optimization of Gaussian-Type Basis Sets for Local Spin Density Functional Calculations. Part I. Boron through Neon, Optimization Technique and Validation. *Can. J. Chem.* **1992**, *70* (2), 560–571.

(40) Benassi, E. Benchmarking of Density Functionals for a Soft but Accurate Prediction and Assignment of 1H and 13C NMR Chemical Shifts in Organic and Biological Molecules. *J. Comp. Chem.* **2017**, *38* (2), 87–92.

(41) Willoughby, P. H.; Jansma, M. J.; Hoye, T. R. A Guide to Small-Molecule Structure Assignment through Computation of (1H and 13C) NMR Chemical Shifts. *Nat. Protoc.* **2014**, *9*, 643.

(42) Abbott, L. J.; Hart, K. E.; Colina, C. M. Polymatic: A Generalized Simulated Polymerization Algorithm for Amorphous Polymers. *Theor. Chem. Acc.* **2013**, *132* (3), 1334.

(43) Holden, D.; Jelfs, K. E.; Cooper, A. I.; Trewin, A.; Willock, D. J. Bespoke Force Field for Simulating the Molecular Dynamics of Porous Organic Cages. *J. Phys. Chem. C* **2012**, *116* (31), 16639–16651.

(44) Evans, J. D.; Huang, D. M.; Hill, M. R.; Sumby, C. J.; Sholl, D. S.; Thornton, A. W.; Doonan, C. J. Molecular Design of Amorphous Porous Organic Cages for Enhanced Gas Storage. *J. Phys. Chem. C* **2015**, *119* (14), 7746–7754.

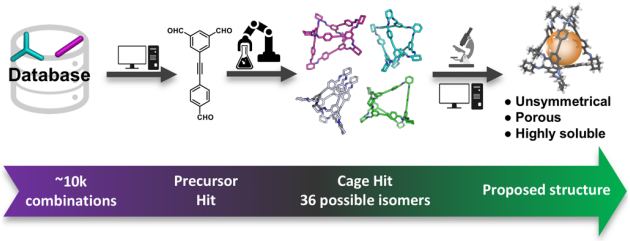
(45) Holden, D.; Chong, S. Y.; Chen, L.; Jelfs, K. E.; Hasell, T.; Cooper, A. I. Understanding Static, Dynamic and Cooperative Porosity in Molecular Materials. *Chem. Sci.* **2016**, *7* (8), 4875–4879.

(46) Holden, D.; Jelfs, K. E.; Trewin, A.; Willock, D. J.; Haranczyk, M.; Cooper, A. I. Gas Diffusion in a Porous Organic Cage: Analysis of Dynamic Pore Connectivity Using Molecular Dynamics Simulations. *J. Chem. Phys. C* **2014**, *118* (24), 12734–12743.

(47) Manurung, R.; Holden, D.; Miklitz, M.; Chen, L.; Hasell, T.; Chong, S. Y.; Haranczyk, M.; Cooper, A. I.; Jelfs, K. E. Tunable Porosity through Cooperative Diffusion in a Multicomponent Porous Molecular Crystal. *J. Phys. Chem. C* **2015**, *119* (39), 22577–22586.

(48) Greenaway, R. L.; Holden, D.; Eden, E. G. B.; Stephenson, A.; Yong, C. W.; Bennison, M. J.; Hasell, T.; Briggs, M. E.; James, S. L.; Cooper, A. I. Understanding Gas Capacity, Guest Selectivity, and Diffusion in Porous Liquids. *Chem. Sci.* **2017**, *8* (4), 2640–2651.

Graphical TOC Entry



Computationally inspired and rationalised discovery of a completely unsymmetrical organic cage, which was both porous and highly soluble.