

Efficient Multigram Procedure for the Synthesis of Large Hydrazone-linked Molecular Cages

Olav Vestrheim,^{a,b} Mica E. Schenkelberg,^{a,b} Qingsheng Dai,^{b,c} and Severin T. Schneebeli^{a,b,*}

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Covalently linked molecular cages can provide significant advantages (including, but not limited to enhanced thermal and chemical stability) over metal-linked coordination cages. Yet, while large coordination cages can now be created routinely, it is still challenging to create chemically robust, covalently linked molecular cages with large internal cavities. This fundamental challenge has made it difficult, for example, to introduce endohedral functional groups into covalent cages to enhance their practical utility (e.g., for selective guest recognition or catalysis), since the cavities would have simply been filled up with such endohedral functional groups in most cases. Here we now report the synthesis of some of the largest known covalently linked molecular tetrahedra. Our new covalent cages all contain 12 peripheral functional groups, which keep them soluble. They are formed from a common vertex, which aligns the hydrazone functions required for the hydrazone linkages with atropisomerism. While we previously reported this vertex as a building block for the smallest member of our hydrazone-linked tetrahedra, our original synthesis was not feasible to be carried out on the larger scales required to successfully access the larger tetrahedra. To overcome this synthetic challenge, we now present a greatly improved synthesis of our vertex, which only requires a single chromatographic step (compared to 3 chromatographic purification steps, which were needed for the initial synthesis). Our new synthetic route enabled us to create a whole family of molecular cages with increasing size (all linked with hydrolytically stable hydrazone bonds), with our largest covalent cage featuring *p*-quarterphenyl linkers and the ability to encapsulate a hypothetical sphere of approximately 3 nm in diameter. These results now open up the possibility to introduce functional groups required for selective recognition and catalysis into chemically robust covalent cages (without blocking the cavities of the covalent cages).

Introduction

Covalently linked molecular cages have gained a significant amount of attention in the last few decades as they can function as selective receptors and catalysts inspired by enzymes.^{1–12} A great variety of covalent cages have been created over the last decade, but one major drawback for many current structures is the limited stability in aqueous environments.^{6,13} This is especially an issue with molecular cages assembled using dynamic imine or boronate ester linkages, which are prone to hydrolysis.^{3,4,6,14–17}

Several strategies to access hydrolytically stable covalent cages have been devised, based on (i) alkyne metathesis (which leads to robust ethynylene linkages)^{18,19}, (ii) hydrazone^{6,20–23} or oxime²⁴ links, (iii) conversion of imine bonds into more stable amide bonds^{8,25,26}, or (iv) use of irreversible carbon-carbon^{27,28} or carbon-oxygen²⁹ bonds to form the covalent cages. Our own strategy has been to create robust covalent cages with hydrazone links, which show superior hydrolytic stability, and can also act as excellent hydrogen bonding acceptors to enable selective polymerization catalysis within the cavities of the molecular cages.¹³ However, in our original report, we were limited to the synthesis of a relatively small tetrahedral hydrazone cage with simple phenyl units as the linkers due to the fairly complex vertex synthesis, which involved 3 chromatographic purification steps. Here, we now provide a new, readily scalable synthetic pathway to access our unique vertex *syn-9*, which precisely orients three hydrazone groups in three dimensions to direct the assembly of hydrazone-linked tetrahedral cages (Scheme 1). This

new synthesis, coupled with the ability to selectively attach 12 peripheral solubilizing groups to our molecular cages, has enabled us to create a whole family of tetrahedral molecular cages, with successively larger size using linkers of increasing length (Scheme 2).

Results and discussion

In our previously reported vertex synthesis¹³, the peripheral triglyme functional groups –R were attached to the vertices of the tetrahedron at the aldehyde stage, rendering every compound afterwards in the original synthesis an oil with relatively high molecular weight (Scheme 1). The physical properties of these intermediates thus prohibited purification with scalable techniques such as recrystallization or distillation, requiring purification with expensive chromatographic techniques at nearly every step. In our new synthesis, we were now able to selectively deprotect all three

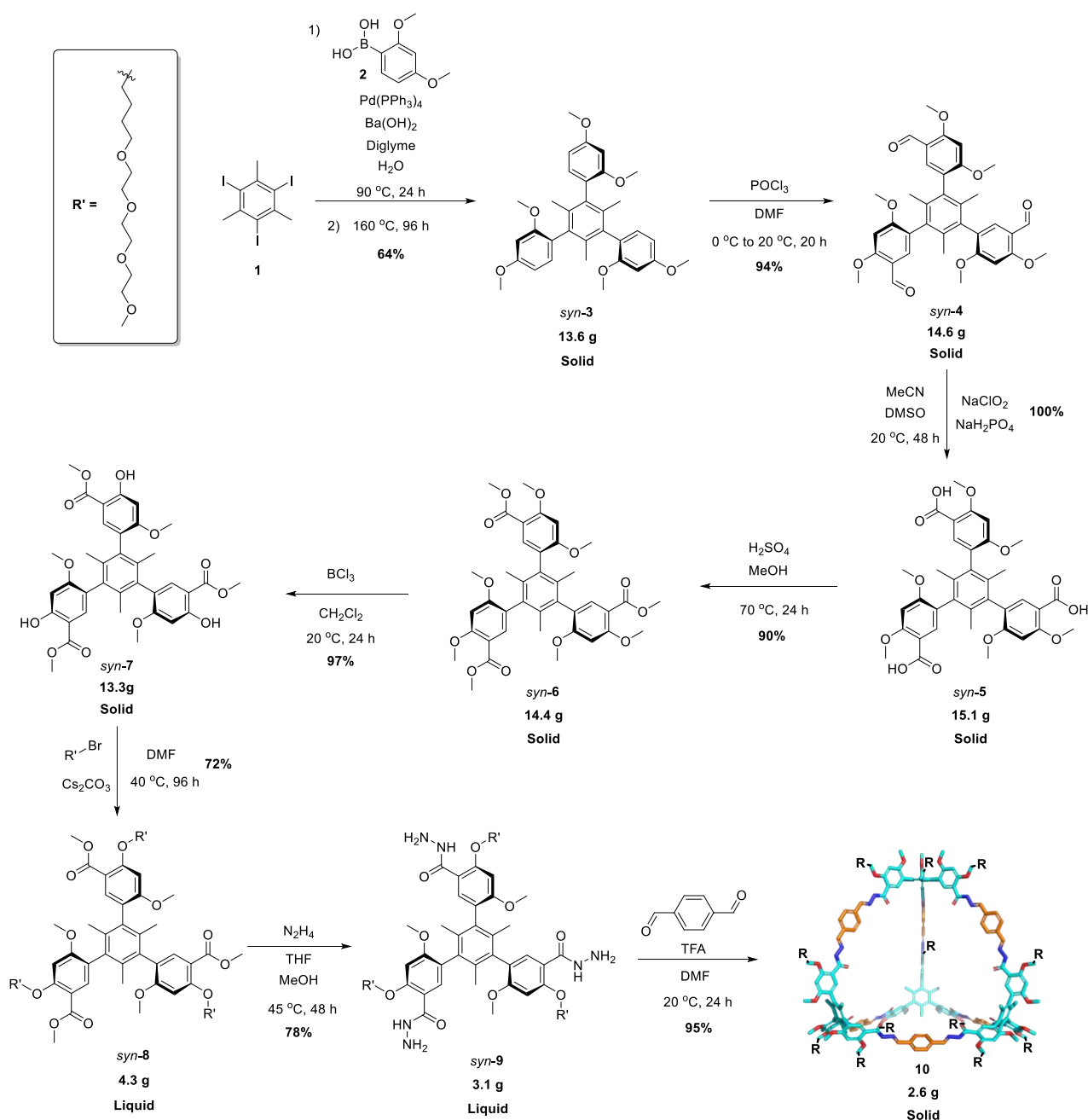
^a Departments of Industrial & Physical Pharmacy and Chemistry, Purdue University, 575 Stadium Mall Drive, West Lafayette, IN 47907.

^b Department of Chemistry and Materials Science Program, University of Vermont, 82 University Place, Burlington, VT 05405, USA.

^c This paper describes experimental work performed by Qingsheng Dai but she is unaware that it has been submitted for publication as we have no contact details for her. Qingsheng Dai, therefore, does not take any responsibility for the submission.

* Corresponding author email: schneebeli@purdue.edu

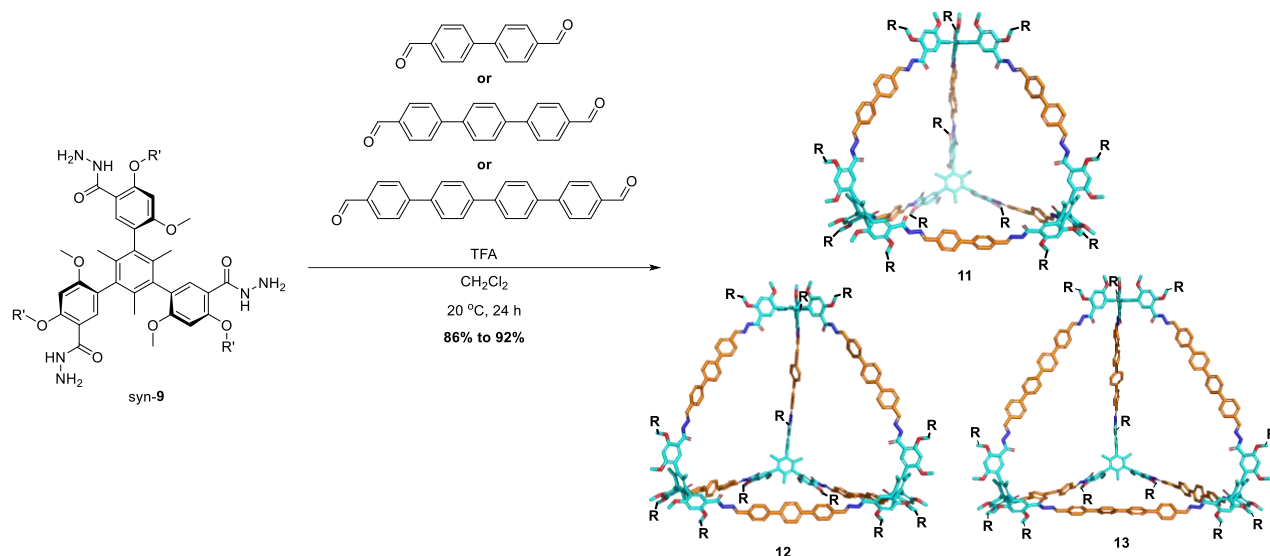
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x



Scheme 1. Our newly developed synthetic procedure for the hydrazone-linked molecular tetrahedron **10**. While **10** was reported previously by us,¹³ our new synthetic procedure to access it is much more scalable and now requires only a single chromatographic purification step. Compounds *syn-5*, *syn-6* and *syn-7* are solids, which enabled us to purify these compounds with simple precipitation/recrystallization techniques to allow for simple and cost-effective scale up. *Syn-5*, *syn-6* and *syn-7* are new compounds. $R' = -(CH_2)_4(OCH_2CH_2)_3OCH_3$. $R = -(CH_2)_3(OCH_2CH_2)_3OCH_3$.

ortho-methoxyl groups of the triester vertex *syn-6* with BCl_3 in CH_2Cl_2 (Scheme 1). This provides for an efficient late-stage diversification strategy, since the peripheral triglyme functional groups are now introduced at the latest possible stage in the synthesis. Therefore, our new synthetic strategy provides the vertex *syn-7* as a common late-stage intermediate for a large family of molecular cages with different peripheral functionalization patterns, simplifying access to cages with different $-R$ groups. However, the biggest advantage of our new late-stage functionalization approach is that all synthetic

intermediates up to compound *syn-7* are now crystalline solids, which can be readily purified with simple precipitation/recrystallization techniques. Specifically, we discovered that we could simply precipitate compounds *syn-5*, *syn-6*, and *syn-7* in water and filter them off without any further purification required. The only compound in our new synthesis that still requires purification by column chromatography is the triglyme-functionalized vertex *syn-8*.



Scheme 2. Synthesis of three new hydrazone-linked covalent cages, **11** with a biphenyl linker, **12** with a *p*-terphenyl linker and **13** with a *p*-quaterphenyl linker. R' = -(CH₂)₄(OCH₂CH₂)₃OCH₃. R = -(CH₂)₃(OCH₂CH₂)₃OCH₃. The structures of the molecular cages represent DFT-optimized models with R = -H. See the supplementary information for computational details.

With our scalable procedure to access the universal vertex, **syn-9**, we were able to significantly extend (Scheme 2) the size of the molecular cages, with the largest containing a cavity volume of approximately 17 nm³ and an internal height of about 33 Å (Figure 1). Our new structures (especially compound **13**, which contains 6 *p*-quaterphenyl linkers), represent some of the largest tetrahedral

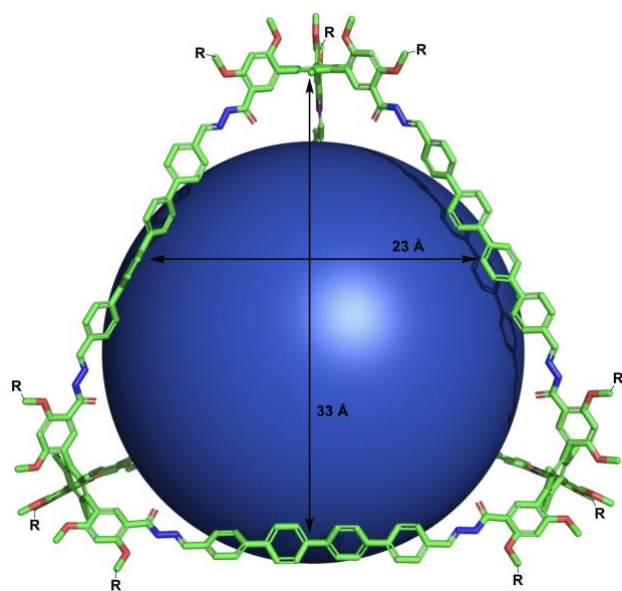


Figure 1. DFT-optimized structure of the *p*-quaterphenyl-linked molecular cage **13**. The sphere illustrating the cavity of the cage measures 3.2 nm in diameter, with a sphere volume of 17 nm³. R = -(CH₂)₃(OCH₂CH₂)₃OCH₃ in the synthesized structure, while R = -H in the DFT model.

porous organic cages created^{13, 30-33} to date (Scheme 2). Notably, the largest cage (**13**) can now be synthesized with only a single chromatographic purification step, which provides a practical path to gram-scale quantities of the common vertex **syn-9** (Scheme 1).

Notably, while the synthesis of tetrahedral cages and cages with large pore openings in general is often complicated by intercatenation³⁴⁻³⁶, we did not observe intercatenation in our synthesis. The glyme-derived solubilizing chains attached to all the vertices stabilize the cages in their monomeric forms in solution, which helps to prevent intercatenation. With the successful formation of the larger cavities, we can now gain access to a much greater selection of potential targets for selective recognition and catalysis inside the tetrahedral cages, since the larger cages contain cavities large enough for endohedral functionalization. In addition, as a tetrahedral geometry provides the largest openings of any polyhedral shape³⁷, our cages will provide the possibility of larger polymers with bulky side chains to gain access to the cavity of the cage, with the potential for processive catalytic functionalization.¹³

Method Description

The synthesis of the hydrazone linked molecular tetrahedron **10** started with a Suzuki reaction between the commercially available 1,3,5-triiodomesitylene (**1**) and 2,4-dimethoxyphenylboronic acid (**2**) which produced a mixture of the *syn* and *anti*-isomers of compound **3**. By simply heating this crude mixture of *syn* and *anti-3*, we then achieved a solid-state-driven amplification of the desired *syn*-atropoisomer (13.6 g, 64% yield). While we reported this unique solid-state driven amplification of a minor atropostereoisomer previously³⁸, we now scaled up and simplified the corresponding procedure significantly to over 10 g in a single batch, and also eliminated the need for chromatographic purification of the crude *syn/anti*-mixture of compound **3** (see also the experimental details section). The overall yield for the first synthetic step was slightly lower (64% compared to 85% reported previously on smaller scale³⁸)

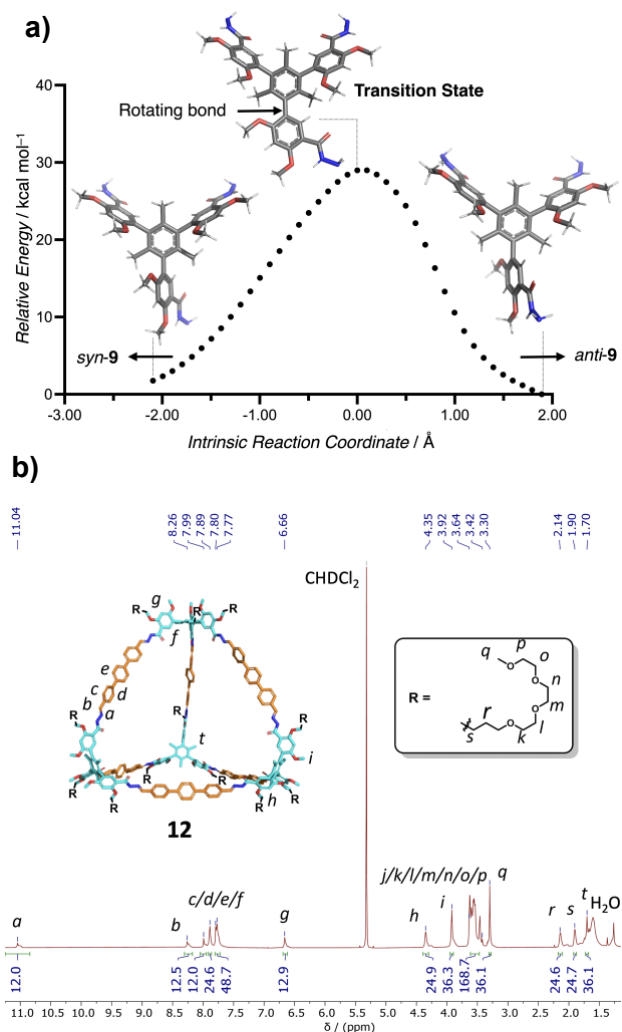


Figure 2. a) Intrinsic reaction coordinate scan for the inversion of *syn*-9 to *anti*-9, obtained at the B3LYP/LACVP* level of theory. At the B3LYP/aug-cc-PVDZ//B3LYP/LACVP* level of theory, the activation energy for *syn*-9 to *anti*-9 conversion was calculated to be 28.5 kcal/mol, which explains why the *syn*-configuration of our vertex is stable, even under mild heating. b) Partial ¹H NMR spectrum (800 MHz, CD₂Cl₂, 298 K) of the molecular cage 12. The full spectrum is shown in the supplementary information, along with additional characterization data (¹³C NMR spectra, mass spectra, and ¹H DOSY NMR spectra) for cages 10, 11, 12, and 13.

with our scaled-up, chromatography free procedure, which simply washes away other minor side-products of the Suzuki-coupling reaction. Nevertheless, our new synthesis places priority on the simplicity and ease of purification by simple precipitation to allow for greater scalability of this key synthetic step. Following the large-scale synthesis of *syn*-3, *syn*-3 was formylated using phosphorus oxychloride to reach *syn*-4 (14.6 g, 94% yield).¹³ Next, a Pinnick oxidation was used to convert the aldehyde functions to carboxylic acids, producing *syn*-5 (15.1 g, 100% yield). The carboxylic acid functional groups were subsequently esterified to methyl esters, to yield compound *syn*-6 (14.4 g, 90% yield). Finally, the methoxyl

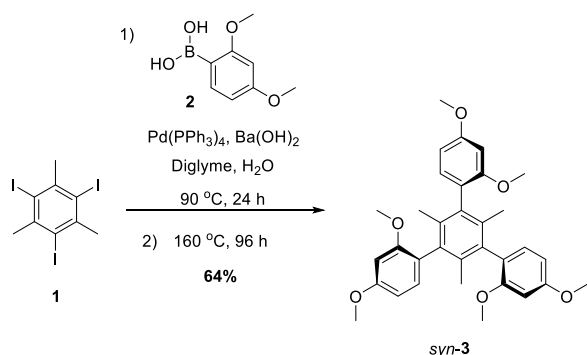
groups *ortho* to the aldehydes of the methyl ester functions were removed selectively with BCl₃ to afford the tris-phenol *syn*-7 (13.3 g, 97%), which can serve as a common late-stage intermediate for the synthesis of a large family of future molecular cages with different peripheral functional groups. Ultimately, an optimized large-scale procedure to convert *syn*-7 into our original, hydrazone-linked tetrahedron 10¹³ was also developed by (i) alkylation of *syn*-7 with triglyme functions, followed by (ii) conversion of the methyl ester functional groups into the tris-hydrazide *syn*-9 by nucleophilic acyl substitution with hydrazine. Importantly, the *syn*-configuration proved stable during all these chemical transformations, which is made possible by the high barrier of rotation around the atropisomeric phenyl-phenyl single bonds. In this work, we calculated the rotational barrier with density functional theory (DFT). We first optimized the structure of the transition state for the *syn*-9 to *anti*-9 conversion at the B3LYP/LACVP* level (see the supplementary information for the computational details). Next, we performed an intrinsic reaction coordinate (IRC) coordinate scan (Figure 2a), starting from the transition state, and then optimized the geometries of *syn*-9 and *anti*-9. Finally, we calculated the activation energy for the *syn*-9 to *anti*-9 interconversion at the B3LYP/aug-cc-PVDZ//B3LYP/LACVP* level which provided a value of 28.5 kcal/mol for the activation energy (with zero-point energy corrections calculated at the B3LYP/LACVP* level). At the same level of theory, we found that the optimized structure of *anti*-9 lies 2.3 kcal/mol lower in energy than *syn*-9. The relatively high rotational barrier of nearly 29 kcal/mol explains why the *syn*-vertex is configurationally stable during the gram-scale synthetic protocol.

Starting from *syn*-9, we were then able to form the previously reported¹³ molecular cage 10 (Scheme 1), as well as three new, larger molecular cages (11, 12, and 13, Scheme 2), by reacting *syn*-9 with a variety of linear dialdehyde linkers as detailed in the experimental procedures section. Notably, our new procedure to form the molecular cage 10 now uses DMF instead of CH₂Cl₂ as the solvent, which allowed us to simply precipitate the molecular cage 10 with water from the crude reaction mixture. Furthermore, the solvent volume required for molecular cage formation was also reduced by a factor of 6 compared to our original report¹³, resulting in a much more cost-effective synthesis of the molecular cage 10¹³ (2.6 g, 95% yield). Closely related procedures were then utilized to synthesize the larger covalent cages 11, 12, and 13 as detailed in the Experimental Procedures section.

The three new cages, 11, 12 and 13 were characterized by ¹H, ¹³C (¹H), and ¹H DOSY NMR spectroscopy (see: Figure 2b and Figures S18–S26) along with mass spectrometry (see: Figures S27–S33). For the biphenyl-linked tetrahedron 11, the [M + 5H]⁵⁺ and [M + 6H]⁶⁺ ions were observed (see: Figures S27–S29) with high resolution electrospray ionization (ESI) mass spectrometry, while the molecular masses of the larger cages, 12 and 13 were verified with matrix-assisted laser desorption ionization (MALDI) mass spectrometry. In the MALDI spectra of 12 and 13 (see Figures S30–S33) both the sodium and potassium ion adducts were observed. Further characterization data, in-line with the DFT-predicted structures (see: Scheme 1, Scheme 2, and Figure 1) of the molecular cages, was obtained from ¹H DOSY NMR spectroscopy, which provided diffusion coefficients of 2.8 × 10⁻¹⁰ m² s⁻¹, 2.1 × 10⁻¹⁰ m² s⁻¹, 2.0 × 10⁻¹⁰ m² s⁻¹, and 1.4 × 10⁻¹⁰ m² s⁻¹ for the molecular cages 10, 11, 12, and 13, respectively. Based on the Stokes-Einstein Equation, which assumes a spherical model (see the supplementary information for details), these diffusion coefficients translate into approximate solvodynamic

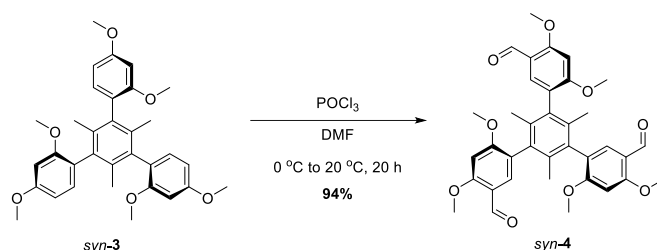
radii of 1.9, 2.5, 2.7, and 3.8 nm, respectively, which is in qualitative agreement with the expected sizes of the molecular cages, especially given that the triglyme solubilizing chains and weakly bound solvent molecules are expected to add additional bulk to the molecular cages. Overall, it is very challenging to completely eliminate convection from DOSY NMR experiments. To bring convection to an absolute minimum, we used Shigemi NMR tubes (which greatly reduce convection due to a reduced solvent volume) and utilized a pulse sequence with convection compensation for all the ^1H DOSY NMR experiments. Notably, the diffusion coefficient measured for **10** obtained in this work, is slightly reduced ($2.8 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$), compared to the diffusion coefficient, which we measured previously¹³ for **10** ($5.6 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$) with a regular 5 mm diameter NMR tube. This finding is consistent with the fact that there is less convection to be expected with a smaller sample volume (i.e. with the Shigemi NMR tube).¹³ Taken together, the relative diffusion coefficients calculated from the ^1H DOSY NMR spectra for the three new cages **11**, **12**, and **13** as well as for our original tetrahedron **10** clearly indicate an increasing solvodynamic radius from the smallest to the largest cage (see also Table S1), which is in qualitative agreement with the DFT-optimized structures of the molecular cages.

Experimental Procedures

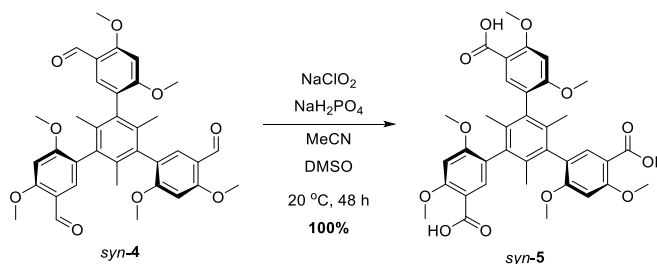


Syn-3.³⁸ 1,3,5-triiodo-2,4,6-trimethylbenzene (**1**, 20.0 g, 40.2 mmol), 2,4-dimethoxyphenylboronic acid (**2**, 35.5 g, 195 mmol), and barium hydroxide octahydrate (110 g, 348 mmol) were suspended in a mixture of diglyme (500 mL) and water (100 mL) in a 1000 mL round bottom flask charged with a magnetic stir bar. After purging the suspension with argon for 3 minutes, tetrakis(triphenylphosphine) palladium(0) (2.32 g, 2.01 mmol) was added to the reaction flask followed by further purging with Argon for 5 additional minutes, at which point the flask was capped with a silicon stopper and fitted a balloon filled with argon. The reaction mixture was then heated to 90°C and stirred for 24 h. After 24 h, the reaction mixture was allowed to cool to room temperature and filtered through celite. The filtrate was then poured into deionized water (3 L) and left to precipitate overnight. The precipitate was filtered off using a 350 mL sintered funnel and transferred to a 500 mL Erlenmeyer flask. The CH_2Cl_2 was evaporated by adding the flask to a hotplate and setting it to 50°C . When the mixture stopped boiling, ethanol (300 mL) was added to the flask, the hotplate was set to 80°C and the mixture was brought to boil. After boiling for approximately 2 minutes the flask was taken off the hotplate and allowed to cool to room temperature over 24 h. The precipitate was filtered off leaving a colourless solid. The solid was added to three 20 mL vials and heated to 160°C for

96 h, converting all the compound to the *syn*-confirmation. As the compound heated up it would melt into a liquid and after about 24 h of heating it became a solid again. After heating, the compound was again transferred to an Erlenmeyer flask which was added to a hotplate. Ethanol (150 mL) was then added to the flask, the hotplate was set to 80°C , and the mixture was brought to a boil. After boiling for approximately 2 minutes the flask was taken off the hotplate and allowed to cool to room temperature over 24 h. The precipitate was filtered off providing **syn-3** (13.6 g, 64%) as a colorless solid. The characterization data matched the data reported³⁸ in the literature: ^1H NMR (500 MHz, CDCl_3) δ (ppm); 7.04 – 6.99 (m, 3H), 6.58 – 6.52 (m, 6H), 3.84 (s, 9H), 3.70 (s, 9H), 1.71 (s, 9H). ^{13}C (^1H) NMR (125 MHz, CDCl_3) δ (ppm); 159.85, 158.07, 135.50, 135.04, 131.73, 124.04, 104.64, 99.45, 55.82, 55.41, 18.85. HRMS (ESI+) m/z : calcd for $\text{C}_{33}\text{H}_{36}\text{O}_6$ [$\text{M} + \text{H}$] $^+$ 529.2590, found 529.2604.

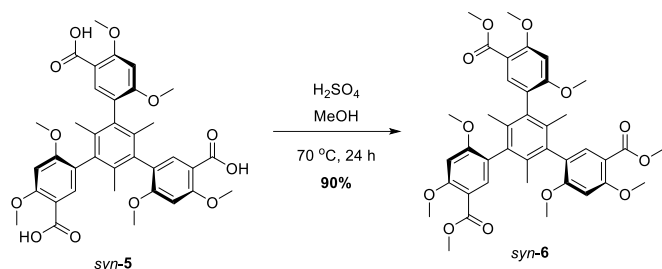


Syn-4. **Syn-3** (13.3 g, 25.2 mmol) was dissolved in anhydrous DMF (500 mL) followed by the addition of phosphorus oxychloride (60 mL, 644 mmol) at room temperature. After 2 h, another batch of phosphorus oxychloride (20 mL, 215 mmol) was added, followed by the addition of three new batches of phosphorus oxychloride (20 mL, 215 mmol) every 24 h until completion. When the reaction showed full conversion, the mixture was quenched by pouring it into 3 L of ice water. The mixture was left overnight to allow for the product to completely precipitate out. The precipitated solid was filtered off using a 350 mL sintered funnel, washed with deionized water, and left to dry under high vacuum overnight to afford **syn-4** as a colorless solid (14.6 g, 94%). The characterization data matched the data reported¹³ in the literature: ^1H NMR (500 MHz, CDCl_3) δ (ppm); 10.30 (s, 3H), 7.54 (s, 3H), 6.52 (s, 3H), 3.99 (s, 9H), 3.86 (s, 9H), 1.65 (s, 9H). ^{13}C NMR (125 MHz, CDCl_3) δ (ppm); 188.34, 163.28, 163.02, 135.10, 134.55, 132.67, 123.67, 118.70, 94.81, 55.93, 55.85, 18.81. HRMS (ESI+) m/z : calcd for $\text{C}_{36}\text{H}_{36}\text{O}_9$ [$\text{M} + \text{H}$] $^+$ 613.2438, found 613.2452.

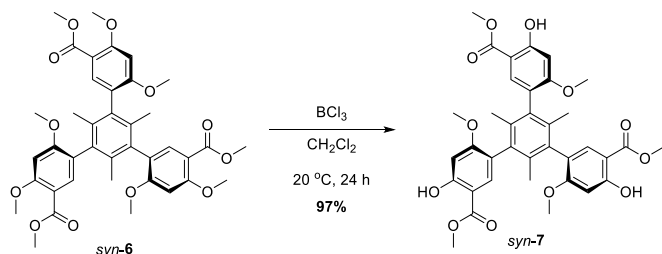


Syn-5. **Syn-4** (14.0 g, 22.9 mmol) was dissolved in a 500 mL of a 1:1 vol% MeCN/DMSO mixture in a 1000 mL round-bottomed flask. Next, the flask was placed in an ice bath following the addition of sodium phosphate monobasic (32.9 g, 274 mmol) and sodium chlorite (tech., nominally, 80%, 24.8 g, 274 mmol) in deionized water (150 mL). After an hour of cooling, the ice bath was removed, and the reaction was allowed to stir at room temperature for 48 h at which point the solution was poured into 3 L of 1% HCl in water and was allowed to precipitate out for 24 h. The precipitate was filtered off using a 350

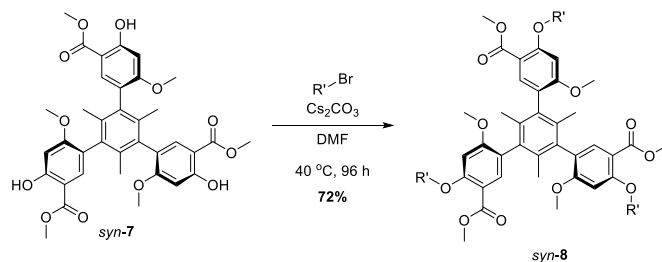
mL sintered funnel and washed with deionized water. The precipitated solid was dried under vacuum overnight providing **syn-5** a colourless solid (15.1 g, 100%). ^1H NMR (500 MHz, DMSO-*d*₆) δ (ppm); 12.17 (s, 3H), 7.37 (s, 3H), 6.79 (s, 3H), 3.91 (s, 9H), 3.85 (s, 9H), 1.57 (s, 9H). ^{13}C (^1H) NMR (125 MHz, DMSO-*d*₆) δ (ppm); 166.42, 160.54, 160.00, 134.69, 134.09, 133.89, 121.30, 112.11, 96.65, 55.95, 55.73, 18.58. HRMS (ESI+) *m/z*: calcd for C₃₆H₃₆O₁₂ [M + Na]⁺ 683.2104, found 683.2106.



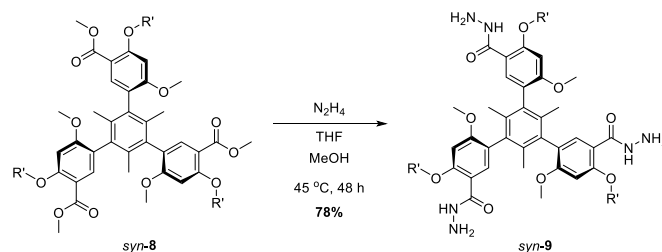
Syn-6. **Syn-5** (15.0 g, 22.7 mmol) was dissolved in MeOH (750 mL) in a 1000 mL round-bottomed flask, followed by the dropwise addition of concentrated sulfuric acid (28.4 mL, 534 mmol). A reflux condenser was attached, and the reaction mixture was stirred at reflux (at around 70 °C) for 24 h. The solution was then poured into 3 L of deionized water and allowed to precipitate overnight. The white precipitate was filtered off and washed with deionized water and then dried under vacuum overnight to provide **syn-6** as a colorless solid (14.4 g, 90%). ^1H NMR (500 MHz, CDCl₃) δ (ppm); 7.62 (s, 3H), 6.56 (s, 3H), 3.98 (s, 9H), 3.84 (s, 9H), 3.83 (s, 9H), 1.67 (s, 9H). ^{13}C (^1H) NMR (125 MHz, CDCl₃) δ (ppm); 166.18, 161.34, 160.87, 135.36, 135.26, 134.74, 122.75, 111.78, 95.96, 56.29, 55.78, 51.74, 18.82. HRMS (ESI+) *m/z*: calcd for C₃₉H₄₂O₁₂ [M + H]⁺ 703.2755, found 703.2755.



Syn-7. **Syn-6** (14.2 g, 20.2 mmol) was dissolved in CH₂Cl₂ (500 mL) in a 1000 mL round bottom flask. Next, a 1 M Boron trichloride solution in CH₂Cl₂ (90 mL, 90 mmol) was added by syringe to the reaction mixture under an argon atmosphere. After stirring at room temperature for 24 h, deionized water (50 mL) was added to quench the reaction. The flask with the resulting biphasic mixture was attached to a rotary evaporator and the CH₂Cl₂ was evaporated leading the product to precipitate out. The product was filtered off using a 350 mL sintered funnel, providing a grey solid. The precipitate was added to a 150 mL beaker along with 100 mL of ethanol, the ethanol mixture was brought to a boil under stirring and allowed to cool back down to room temperature. The colorless precipitate was filtered off and dried under vacuum overnight to afford **syn-7** (13.3 g, 97%). ^1H NMR (500 MHz, CDCl₃) δ (ppm); 10.94 (s, 3H), 7.52 (s, 3H), 6.55 (s, 3H), 3.90 (s, 9H), 3.80 (s, 9H), 1.68 (s, 9H). ^{13}C (^1H) NMR (125 MHz, CDCl₃) δ (ppm); 170.65, 163.24, 163.14, 135.40, 134.75, 132.39, 122.97, 105.18, 99.43, 55.88, 52.07, 18.81. HRMS (ESI+) *m/z*: calcd for C₃₆H₃₆O₁₂ [M + H]⁺ 661.2285, found 661.2286.

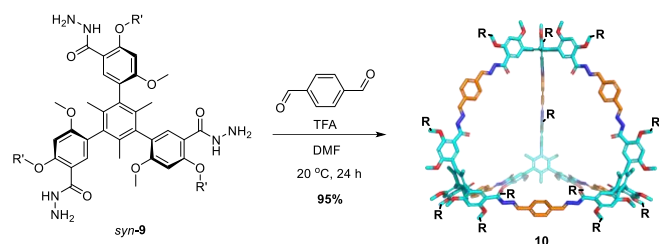


Syn-8. **Syn-7** (3.00 g, 4.54 mmol) and cesium carbonate (7.40 g, 22.7 mmol) were added to a 500 mL round-bottomed flask along with DMF (180 mL) followed by the addition of 15-bromo-2,5,8,11-tetraoxapentadecane (4.35 g, 14.5 mmol). The reaction mixture was heated to 40 °C and stirred for 72 h. Then, the reaction mixture was transferred to a 1000 mL separatory funnel and 200 mL of deionized water was added. Finally, the mixture was extracted with EtOAc (3 × 100 mL), the organic phases were combined in the separatory funnel and the reaction was washed four times with 150 mL of 20% brine solution to remove any remaining DMF. The organic phase was then dried over anhydrous Na₂SO₄, filtered, and evaporated under reduced pressure. The crude product was purified by flash column chromatography over silica gel using a Teledyne CombiFlash® Rf+ chromatography system (eluent: 10 vol% MeOH in CH₂Cl₂) to afford **syn-8** as a pale-yellow oil (4.32 g, 72%). The characterization data matched the data reported¹³ in the literature: ^1H NMR (500 MHz, CDCl₃) δ (ppm); 7.60 (s, 3H), 6.53 (s, 3H), 4.11 (t, *J* = 6.3 Hz, 6H), 3.83 (s, 9H), 3.80 (s, 9H), 3.67 – 3.53 (m, 42H), 1.96 (m, 6H), 1.85 (m, 6H), 3.37 (s, 9H), 1.66 (s, 9H). ^{13}C (^1H) NMR (125 MHz, CDCl₃) δ (ppm); 166.25, 161.22, 160.31, 135.30, 135.20, 134.78, 122.80, 112.07, 97.05, 72.06, 71.00, 70.75, 70.72, 70.65, 70.18, 68.98, 59.15, 55.77, 51.62, 26.29, 26.25, 18.82. HRMS (ESI+) *m/z*: [M + NH₄]⁺ calcd 1332.7128, found 1332.7105.



Syn-9. **Syn-8** (4.00 g, 3.04 mmol) was dissolved in THF (85 mL) and MeOH (170 mL) in a 500 mL round bottom flask. Next, the reaction mixture was degassed with argon and hydrazine hydrate (47.7 mL, 1.52 mol) was added. The reaction was heated to 45 °C and allowed to stir for 48 h under an argon atmosphere. After 48 h the solvent was evaporated under reduced pressure and the resulting mixture was transferred to a 500 mL separatory funnel and extracted with CH₂Cl₂ (3 × 50 mL). The combined organic layers were washed with water (3 × 75 mL) to remove excess hydrazine, dried over anhydrous Na₂SO₄, filtered using a 150 mL sintered funnel, concentrated under reduced pressure, and dried *in vacuo* overnight to afford **syn-9** as a yellow oil (3.11 g, 78%). The characterization data matched the data reported¹³ in the literature: ^1H NMR (500 MHz, CDCl₃) δ (ppm); 8.81 (s, 3H), 7.91 (s, 3H), 6.51 (s, 3H), 4.18 (t, *J* = 6.6 Hz, 6H), 3.77 (s, 9H), 3.67 – 3.52 (m, 42H), 3.35 (s, 9H), 2.00 (m, 6H), 1.80 (m, 6H), 1.64 (s, 9H). ^{13}C (^1H) NMR (125 MHz, CDCl₃) δ (ppm); 166.57, 160.45, 157.46, 135.17, 135.00, 134.61, 124.05, 112.84, 96.29, 72.02, 70.72, 70.71, 70.70, 70.66, 70.62, 70.27, 69.17, 59.14,

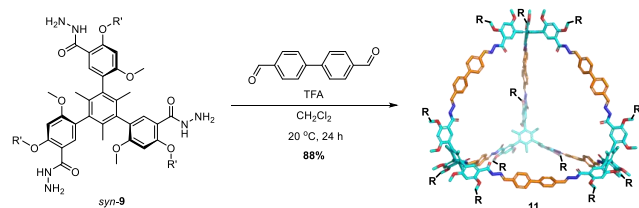
55.94, 26.32, 26.28, 18.85. HRMS (ESI+) m/z : $[M + H]^+$ calcd 1315.7176, found 1315.7200.



Multigram-scale synthesis of the hydrazone-linked Tetrahedron 10.

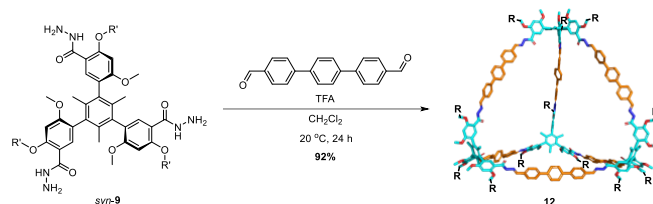
Syn-9 (2.41 g, 1.83 mmol) was dissolved in DMF (400 mL), and the solution was degassed with argon. Next, terephthalaldehyde (369 mg, 2.75 mmol) and trifluoroacetic acid (140 μ L, 1.83 mmol) were added consecutively, and the reaction mixture was stirred at room temperature under an argon atmosphere for 48 h. After 48 h, the mixture was poured into 2 L of a 1% aqueous solution of sodium bicarbonate and allowed to precipitate for 24 h at which point the precipitate was filtered off using a 350 mL sintered funnel. The yellow precipitate was dissolved in CH_2Cl_2 , filtered again, evaporated to dryness under reduced pressure, and dried under high vacuum to provide the hydrazone-linked tetrahedron **10** as a yellow solid (2.55 g, 95%). The characterization data matched the data reported¹³ in the literature: ^1H NMR (500 MHz, CD_2Cl_2) δ (ppm); 11.08 (s, 12H), 8.30 (s, 12H), 8.00 (s, 12H), 7.84 (s, 24H), 6.65 (s, 12H), 4.34 (t, J = 6.6 Hz, 24H), 3.90 (s, 36H), 3.63–3.45 (m, 168H), 3.30 (s, 36H), 2.13 (m, 24H), 1.89 (m, 24H), 1.68 (s, 36H). ^{13}C (^1H) NMR (125 MHz, CD_2Cl_2) δ (ppm); 161.84, 161.04, 157.81, 146.43, 136.41, 135.35, 134.93, 128.12, 124.53, 113.59, 96.25, 72.28, 70.96, 70.94, 70.89, 70.85, 70.76, 70.71, 69.98, 59.00, 56.21, 26.84, 26.78, 18.84. HRMS (ESI+) m/z : $[M + 5\text{H}]^{5+}$ calcd 1,170.9964, found 1,170.9982.

General Procedure for the Synthesis of the Larger Molecular Tetrahedra. **Syn-9** (1 eq) and a dicarbaldehyde (1.5 eq) was dissolved in CH_2Cl_2 (1.5 mM solution) followed by the addition of trifluoroacetic acid (1.0 eq). Next, the reaction mixture was stirred at room temperature for 24 h. After 24 h, the solvent was evaporated under reduced pressure. The resulting crude product was dried *in vacuo* and washed with methanol to provide the hydrazone-linked tetrahedra **11**, **12**, and **13**.

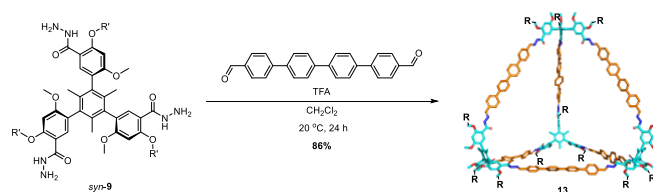


Hydrazone-linked Tetrahedron 11. Following the general procedure for the synthesis of the molecular tetrahedra with **syn-9** (30.7 mg, 23.3 μ mol) and [1,1'-biphenyl]-4,4'-dicarbaldehyde (7.4 mg, 35.0 μ mol), we obtained **11** as a yellow solid (32 mg, 88%). ^1H NMR (500 MHz, CD_2Cl_2) δ (ppm); 11.04 (s, 12H), 8.26 (s, 12H), 7.99 (s, 12H), 7.88 to 7.87 (m, 24H), 7.78 to 7.77 (m, 24H) 6.65 (s, 12H), 4.34 (t, J = 6.6 Hz, 24H), 3.90 (s, 36H), 3.62–3.47 (m, 168H), 3.29 (s, 36H), 2.12 (m, 24H), 1.89 (m, 24H), 1.69 (s, 36H). ^{13}C (^1H) NMR (200 MHz, CD_2Cl_2) δ (ppm); 161.94, 161.09, 157.88, 146.63, 141.79, 135.61, 135.01,

134.28, 128.42, 127.47, 124.45, 113.51, 96.21, 72.27, 70.93, 70.88, 70.76, 70.72, 69.97, 58.99, 56.22, 50.79, 26.90, 26.80, 18.85. HRMS (ESI+) m/z : $[M + 5\text{H}]^{5+}$ calcd 1262.4349, found 1262.4342.



Hydrazone-linked Tetrahedron 12. Following the general procedure for the synthesis of the molecular tetrahedra with **syn-9** (25.5 mg, 19.4 μ mol) and [1,1':4',1''-terphenyl]-4,4''-dicarbaldehyde (8.33 mg, 29.1 μ mol), we obtained **12** as a yellow solid (30 mg, 92%). ^1H NMR (800 MHz, CD_2Cl_2) δ (ppm); 11.04 (s, 12H), 8.26 (s, 12H), 7.99 (s, 12H), 7.88 to 7.87 (m, 24H), 7.78 to 7.77 (m, 24H) 6.65 (s, 12H), 4.34 (t, J = 6.6 Hz, 24H), 3.90 (s, 36H), 3.62–3.47 (m, 168H), 3.29 (s, 36H), 2.12 (m, 24H), 1.89 (m, 24H), 1.69 (s, 36H). ^{13}C (^1H) NMR (200 MHz, CD_2Cl_2) δ (ppm); 161.94, 161.10, 157.92, 146.65, 142.14, 139.83, 135.62, 135.37, 135.07, 134.03, 128.39, 127.79, 127.50, 124.44, 113.52, 96.20, 72.28, 70.97, 70.95, 70.90, 70.78, 70.74, 69.97, 59.00, 56.23, 30.10, 26.93, 26.83, 18.86. MS (MALDI, DCTB) m/z : $[M + \text{Na}]^+$ calcd 6786, found 6785.



Hydrazone-linked Tetrahedron 13. Following the general procedure for the synthesis of the molecular tetrahedra with **syn-9** (25.0 mg, 19.0 μ mol) and [1,1':4',1''':4'''-quaterphenyl]-4,4'''-dicarbaldehyde (10.3, 28.5 μ mol), we obtained **13** as a yellow solid (29 mg, 86%). ^1H NMR (800 MHz, CD_2Cl_2) δ (ppm); 11.02 (s, 12H), 8.24 (s, 12H), 8.00 (s, 12H), 7.92 to 7.84 (m, 24H), 7.82 to 7.64 (m, 72H) 6.66 (s, 12H), 4.35 (s, 24H), 3.93 (s, 36H), 3.64–3.38 (m, 168H), 3.31 (s, 36H), 2.14 (m, 24H), 1.90 (m, 24H), 1.72 (s, 36H). ^{13}C (^1H) NMR (200 MHz, CD_2Cl_2) δ (ppm); 162.05, 161.13, 157.94, 146.50, 142.30, 140.07, 139.65, 135.59, 135.35, 135.09, 133.98, 128.37, 127.76, 127.55, 124.29, 113.38, 96.18, 72.26, 72.17, 70.95, 70.93, 70.88, 70.76, 70.72, 70.33, 69.93, 59.07, 58.99, 56.22, 56.16, 50.75, 30.08, 26.91, 26.81, 26.64, 26.20, 18.89. MS (MALDI, DCTB) m/z : $[M + \text{Na}]^+$ calcd 7242, found 7241.

Conclusions

In summary, we invented a facile new synthesis of a covalent, hydrazone-linked molecular tetrahedron, which allowed us to produce the catalytically-active¹³ tetrahedron for the first time on a multigram scale in a single batch. Our new procedure requires a minimal amount of purification as it relies heavily on precipitation in water rendering the synthesis much more practical for large scale. In the nine synthetic steps performed,

only one compound required purification by column chromatography. Furthermore, our new synthesis provides access to a common late-stage intermediate, which will allow us to readily expand the family of tetrahedral molecular cages in the future to tune the peripheral functional groups for selective recognition and catalysis. With our new, more efficient procedure we were then able to synthesize three new larger cages, which are represent some of the largest covalent molecular tetrahedra reported to date.^{39, 40}

Acknowledgements

We thank Dr. M. Sharafi for helpful discussions, B. O'Rourke for performing high-resolution mass spectrometry, and Dr. H. Mo for assistance with the DOSY NMR experiments. The large-scale synthesis of the *syn*-**9** was supported by the NSF (CAREER Award CHE-1848444/CHE-2317652) and by the USDA (NIFA grant 2018-07583). The synthesis of the larger molecular cages was supported by the NIH (NIGMS grant 1R35GM147579-01). Q.D. was supported by a UVM Chemistry Summer Undergraduate Research Fellowship provided by the Women in Chemistry Fund.

Conflict of Interests Statement

S.T.S. has a pending patent application related to sensing and catalytic applications of the tetrahedral molecular cages reported in this paper.

References

1. P. S. Bols and H. L. Anderson, Template-Directed Synthesis of Molecular Nanorings and Cages, *Acc. Chem. Res.*, 2018, **51**, 2083-2092.
2. M. Mastalerz, Porous Shape-Persistent Organic Cage Compounds of Different Size, Geometry, and Function, *Acc. Chem. Res.*, 2018, **51**, 2411-2422.
3. K. Acharyya and P. S. Mukherjee, Organic Imine Cages: Molecular Marriage and Applications, *Angew. Chem., Int. Ed.*, 2019, **58**, 8640-8653.
4. J. C. Lauer, W.-S. Zhang, F. Rominger, R. R. Schroeder and M. Mastalerz, Shape-Persistent [4+4] Imine Cages with a Truncated Tetrahedral Geometry, *Chem. - Eur. J.*, 2018, **24**, 1816-1820.
5. E. G. Percastegui, T. K. Ronson and J. R. Nitschke, Design and Applications of Water-Soluble Coordination Cages, *Chem. Rev.*, 2020, **120**, 13480-13544.
6. M. Yang, F. Qiu, E.-S. M. El-Sayed, W. Wang, S. Du, K. Su and D. Yuan, Water-stable hydrazone-linked porous organic cages, *Chem. Sci.*, 2021, **12**, 13307-13315.
7. T. Hasell, M. Schmidtman and A. I. Cooper, Molecular Doping of Porous Organic Cages, *J. Am. Chem. Soc.*, 2011, **133**, 14920-14923.
8. M. Liu, M. A. Little, K. E. Jelfs, J. T. A. Jones, M. Schmidtman, S. Y. Chong, T. Hasell and A. I. Cooper, Acid- and Base-Stable Porous Organic Cages: Shape Persistence and pH Stability via Post-synthetic "Tying" of a Flexible Amine Cage, *J. Am. Chem. Soc.*, 2014, **136**, 7583-7586.
9. R. L. Greenaway, V. Santolini, M. J. Bennison, B. M. Alston, C. J. Pugh, M. A. Little, M. Miklitz, E. G. B. Eden-Rump, R. Clowes, A. Shakil, H. J. Cuthbertson, H. Armstrong, M. E. Briggs, K. E. Jelfs and A. I. Cooper, High-throughput discovery of organic cages and catenanes using computational screening fused with robotic synthesis, *Nat. Commun.*, 2018, **9**, 1-11.
10. T. Hasell and A. I. Cooper, Porous organic cages: soluble, modular and molecular pores, *Nat. Rev. Mater.*, 2016, **1**, 16053.
11. D. Fujita, Y. Ueda, S. Sato, N. Mizuno, T. Kumasaka and M. Fujita, Self-assembly of tetravalent Goldberg polyhedra from 144 small components, *Nature*, 2016, **540**, 563-566.
12. R. Breslow, Artificial enzymes, *Science*, 1982, **218**, 532.
13. M. Sharafi, K. T. McKay, M. Ivancic, D. R. McCarthy, N. Dudkina, K. E. Murphy, S. C. Rajappan, J. P. Campbell, Y. Shen, A. R. Badireddy, J. Li and S. T. Schneebeli, Size-Selective Catalytic Polymer Acylation with a Molecular Tetrahedron, *Chem*, 2020, **6**, 1469-1494.
14. N. Christinat, R. Scopelliti and K. Severin, Multicomponent assembly of boronic acid based macrocycles and cages, *Angew. Chem., Int. Ed.*, 2008, **47**, 1848-1852.
15. M. Hutin, G. Bernardinelli and J. R. Nitschke, An iminoboronate construction set for subcomponent self-assembly, *Chem. - Eur. J.*, 2008, **14**, 4585-4593.
16. B. Icli, N. Christinat, J. Tonnemann, C. Schuttler, R. Scopelliti and K. Severin, Synthesis of Molecular Nanostructures by Multicomponent Condensation Reactions in a Ball Mill, *J. Am. Chem. Soc.*, 2009, **131**, 3154-3155.
17. G. Rao and M. Philipp, Boronic acid catalyzed hydrolyses of salicylaldehyde imines, *J. Org. Chem.*, 1991, **56**, 1505.
18. S. Lee, A. Yang, T. P. Money Penny and J. S. Moore, Kinetically Trapped Tetrahedral Cages via Alkyne Metathesis, *J. Am. Chem. Soc.*, 2016, **138**, 2182-2185.
19. Q. Wang, C. Zhang, B. C. Noll, H. Long, Y. Jin and W. Zhang, A Tetrameric Cage with D_{2h} Symmetry through Alkyne Metathesis, *Angew. Chem., Int. Ed.*, 2014, **53**, 10663-10667.
20. G. Wu, Y. Chen, S. Fang, L. Tong, L. Shen, C. Ge, Y. Pan, X. Shi and H. Li, A Self-Assembled Cage for Wide-Scope Chiral Recognition in Water, *Angew. Chem., Int. Ed.*, 2021, **60**, 16594-16599.
21. Z. Lin, T. J. Emge and R. Warmuth, Multicomponent Assembly of Cavitand-Based Polyacylhydrazone Nanocapsules, *Chem. - Eur. J.*, 2011, **17**, 9395.
22. X. Zheng, Y. Zhang, G. Wu, J.-R. Liu, N. Cao, L. Wang, Y. Wang, X. Li, X. Hong, C. Yang and H. Li, Temperature-dependent self-assembly of a purely organic cage in water, *Chem. Commun.*, 2018, **54**, 3138-3141.
23. H. Li, H. Zhang, A. D. Lammer, M. Wang, X. Li, V. M. Lynch and J. L. Sessler, Quantitative self-assembly of a purely organic three-dimensional catenane in water, *Nat. Chem.*, 2015, **7**, 1003-1008.
24. L. Shen, N. Cao, L. Tong, X. Zhang, G. Wu, T. Jiao, Q. Yin, J. Zhu, Y. Pan and H. Li, Dynamic Covalent Self-Assembly Based on Oxime Condensation, *Angew. Chem., Int. Ed.*, 2018, **57**, 16486-16490.
25. J. L. Culshaw, G. Cheng, M. Schmidtman, T. Hasell, M. Liu, D. J. Adams and A. I. Cooper, Dodecaamide Cages: Organic 12-Arm Building Blocks for Supramolecular Chemistry, *J. Am. Chem. Soc.*, 2013, **135**, 10007-10010.

26. E. Martinez-Ahumada, D. He, V. Berryman, A. Lopez-Olvera, M. Hernandez, V. Jancik, V. Martis, M. A. Vera, E. Lima, D. J. Parker, A. I. Cooper, I. A. Ibarra and M. Liu, SO₂ Capture Using Porous Organic Cages, *Angew. Chem., Int. Ed.*, 2021, **60**, 17556-17563.
27. A. Avellaneda, P. Valente, A. Burgun, J. D. Evans, A. W. Markwell-Heys, D. Rankine, D. J. Nielsen, M. R. Hill, C. J. Sumby and C. J. Doonan, Kinetically Controlled Porosity in a Robust Organic Cage Material, *Angew. Chem., Int. Ed.*, 2013, **52**, 3746-3749.
28. H. Duan, Y. Li, Q. Li, P. Wang, X. Liu, L. Cheng, Y. Yu and L. Cao, Host-Guest Recognition and Fluorescence of a Tetraphenylethene-Based Octacationic Cage, *Angew. Chem., Int. Ed.*, 2020, **59**, 10101-10110.
29. C. Zhang, Z. Wang, L. Tan, T.-L. Zhai, S. Wang, B. Tan, Y.-S. Zheng, X.-L. Yang and H.-B. Xu, A porous tricyclooxacalixarene cage based on tetraphenylethylene, *Angew. Chem., Int. Ed.*, 2015, **54**, 9244-9248.
30. D. Chakraborty and P. S. Mukherjee, Recent trends in organic cage synthesis: push towards water-soluble organic cages, *Chem. Commun. (Cambridge, U. K.)*, 2022, **58**, 5558-5573.
31. G. Monta-Gonzalez, F. Sancenon, R. Martinez-Manez and V. Martí-Centelles, Purely covalent molecular cages and containers for guest encapsulation, *Chem. Rev. (Washington, DC, U. S.)*, 2022, **122**, 13636-13708.
32. J. G. Betancourth, J. A. Castano, R. Visbal and M. N. Chaur, Versatility of the Amino Group in Hydrazone-based Molecular and Supramolecular Systems, *Eur. J. Org. Chem.*, 2022, **2022**, e202200228.
33. M. Mastalerz, Porous Shape-Persistent Organic Cage Compounds of Different Size, Geometry, and Function, *Accounts of Chemical Research*, 2018, **51**, 2411-2422.
34. L. Shen, N. Cao, L. Tong, X. Zhang, G. Wu, T. Jiao, Q. Yin, J. Zhu, Y. Pan and H. Li, Dynamic Covalent Self-Assembly Based on Oxime Condensation, *Angewandte Chemie International Edition*, 2018, **57**, 16486-16490.
35. H. Li, H. Zhang, A. D. Lammer, M. Wang, X. Li, V. M. Lynch and J. L. Sessler, Quantitative self-assembly of a purely organic three-dimensional catenane in water, *Nature Chemistry*, 2015, **7**, 1003-1008.
36. Z. Lin, T. J. Emge and R. Warmuth, Multicomponent Assembly of Cavitand-Based Polyacylhydrazone Nanocapsules, *Chemistry – A European Journal*, 2011, **17**, 9395-9405.
37. K. Mahata, P. D. Frischmann and F. Würthner, Giant Electroactive M4L6 Tetrahedral Host Self-Assembled with Fe(II) Vertices and Perylene Bisimide Dye Edges, *Journal of the American Chemical Society*, 2013, **135**, 15656-15661.
38. M. Sharafi, J. P. Campbell, S. C. Rajappan, N. Dudkina, D. L. Gray, T. J. Woods, J. Li and S. T. Schneebeli, Crystal-Packing-Driven Enrichment of Atropoisomers, *Angew. Chem., Int. Ed.*, 2017, **56**, 7097-7101.
39. G. Montà-González, F. Sancenón, R. Martínez-Máñez, and V. Martí-Centelles, Purely Covalent Molecular Cages and Containers for Guest Encapsulation, *Chem. Rev.* **2022**, **122**, 13636–13708. Note: This review article failed to cite our clearly relevant prior work of purely covalent molecular cages (see: ref. 13) for polymer recognition and novel catalytic applications, which we assume was an inadvertent oversight on behalf of the authors.
40. X. Yang, Z. Ullah, J. F. Stoddart, and C. T. Yavuz, Porous Organic Cages, *Chem. Rev.* **2023**, **123**, 4602–4634.

