Organic Cage Dumbbells

Rebecca L. Greenaway,[a]* Valentina Santolini,[b] Filip T. Szczypiński,[b] Michael J. Bennison,[a] Marc A. Little,[a] Andrew Marsh,[a] Kim E. Jelfs[b] and Andrew I. Cooper[a]

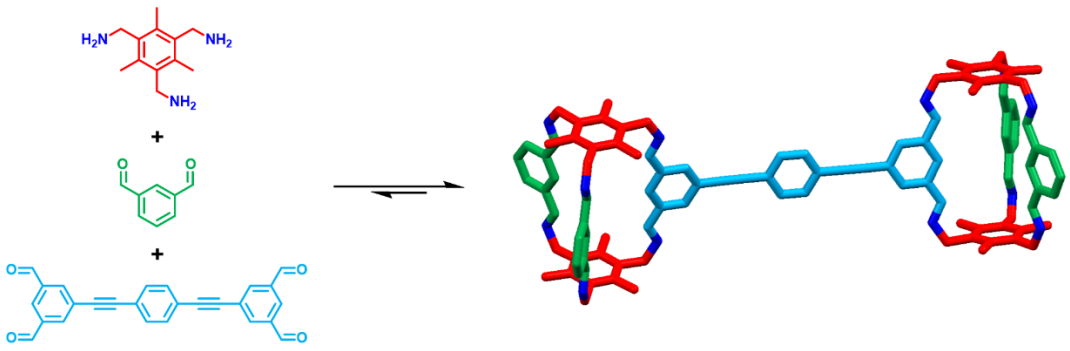
[a]Department of Chemistry and Materials Innovation Factory, University of Liverpool, 51 Oxford Street, Liverpool, L7 3NY, UK; [b]Department of Chemistry, Imperial College London, Molecular Sciences Research Hub, White City Campus, Wood Lane, London, W12 0BZ, UK.

Email: rebecca.greenaway@liverpool.ac.uk

Abstract

Molecular dumbbells with organic cage capping units were synthesised via a multi-component imine condensation between a tri-topic amine and di- and tetra-topic aldehydes. This is an example of self-sorting, which can be rationalised by computational modelling.

Table of Contents



Introduction

Organic cages are self-assembled molecules that are typically formed from two distinct multi-functionalised components using reversible dynamic covalent chemistry.[1–4] Increasing the number of building blocks can lead to a variety of more complex outcomes, including social-sorting into a single mixed assembly, self-sorting into separate discrete species, or the formation of a statistical distribution of assemblies (Fig. 1).[5] To date, there are just a few reports of these types of self-sorting that relate to organic cages; for example, mixtures of three different linkers can lead to self-sorted binary cages,[6–10] a distribution of cage species,[11–14] or, less commonly, socially self-sorted ternary cage assemblies.[15–17] However, these examples typically exploit the use of precursors of the same topicity (*i.e.*, number of reactive functional groups), and they target relatively symmetrical organic cage species, rather than more complex architectures.

Here, we present an example of social-sorting – we used four tri-topic amines, four di-topic aldehydes, and one tetra-topic aldehyde. Self-sorting was observed in solution into two distinct species: socially self-sorted organic cage dumbbells (OCDs) and the corresponding binary 'parent' cage (Scheme 1), alongside the formation of small amounts of insoluble precipitate, which was presumed to be polymeric in nature. To our knowledge, these OCDs are the first example of using self-sorting to form more complex organic cage architectures by covalently connecting two cages together. Furthermore, the consistent formation of a mixture of a dumbbell and a cage was rationalised using computational modelling to compare the formation energy per bond formed.

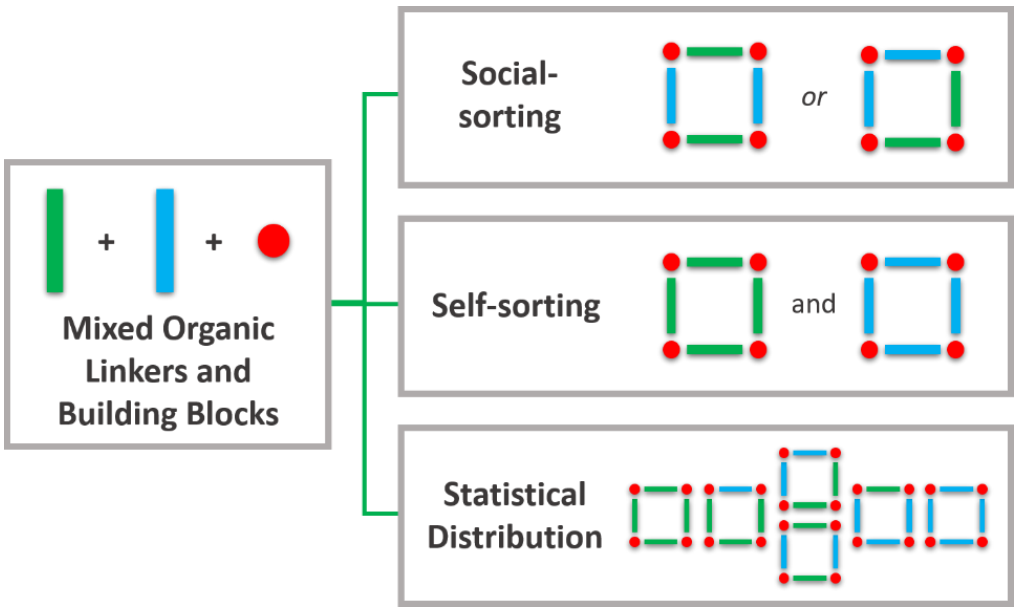
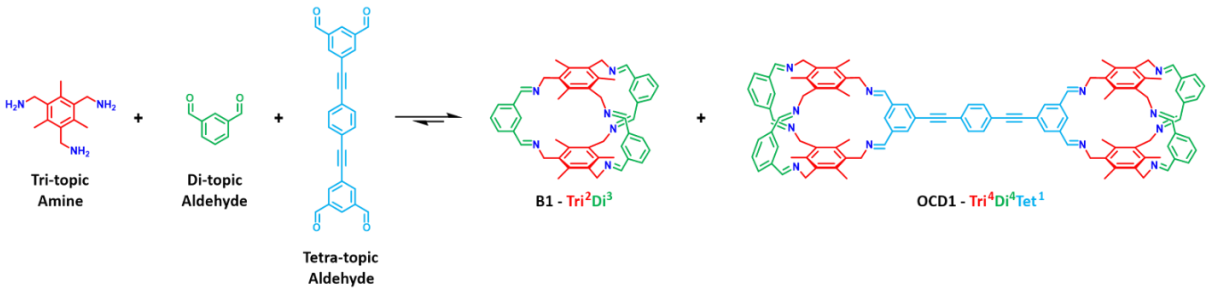


Fig 1. Schematic representation of potential self-assembly outcomes for more than two organic linkers or building blocks: social-sorting leads to the formation of a single mixed assembly, self-sorting results in the formation of separate discrete species, or alternatively, a statistical distribution of mixed species can be formed.

Results and Discussion

Previously, a range of **Tri^{2}Di^3** species have been reported that are formed using an imine condensation strategy – these cage molecules consist of two tri-topic and three di-topic building blocks.[18–21] Following this, we reported a high-throughput workflow that used a robust synthetic method for organic cage discovery, that also included a number of **Tri^{2}Di^3** species.[22,23] While our modelling suggested that organic cages of this topology tend to have cavities that are too small to host gaseous guests based on their kinetic diameters,[22] meaning their porosity is not typically studied, the reduced versions of these small capsular cages, or cryptands/cyclophanes, have been previously studied as binding receptors.[19–21] Furthermore, the **Tri^{2}Di^3** capsular imine cages have a simple and fairly rigid trigonal geometry that lends itself to the design of more complex assemblies. Our aim was to use a multi-component imine condensation to access a socially self-sorted and controlled assembly incorporating more than one cage, instead of self-sorted binary cages. We therefore decided to investigate the one-pot multi-component reaction of a tri-topic amine and di-topic aldehyde (as used in the formation of **B1**, a **Tri^{2}Di^3** cage)[22,24], along with a tetra-topic aldehyde that has the same 1,3-dialdehyde substitution pattern (Scheme 1).



Scheme 1 Reaction scheme for the social-narcissitc self-sorting of a tri-topic amine (red), di-topic aldehyde (green), and tetra-topic aldehyde (blue), to form a mixture of the binary parent cage **B1** (**Tri^{2}Di^3**) and an organic cage dumbbell **OCD1** (**Tri^{4}Di^{4}Tet^1**).

Building on the reaction conditions that we reported for the synthesis of **B1**, which used a moderately high dilution to avoid insoluble polymer formation and modest temperatures to promote conversion,[22] we altered the precursor stoichiometry to include a single tetra-topic aldehyde for every two targeted organic cages, leading to a ratio of 4:1:4 of tri-topic amine:tetra-topic aldehyde:di-topic aldehyde (Table 1, Entry 1). Our hope was that this stoichiometry would favour the formation of a dumbbell architecture, over the large number of other assemblies that can be envisaged. Analysis of the reaction solution by HPLC and high-resolution mass spectrometry (HRMS) indicated clear formation of two species: the binary 'parent' **Tri^{2}Di^3** cage (**B1**, $[M+H]^+$ 709.4085) and a **Tri^{4}Di^{4}Tet^1** species ($[M+2H]^{2+}$ 770.9189), corresponding to an organic cage dumbbell (**OCD1**), in a 76:24 ratio, as determined by HPLC (Fig. 1, and Fig. S3-S4). We then carried out a brief screen to study the effect of varying both the di- and tetra-topic aldehyde molar ratio, in an attempt to favour the formation of **OCD1** (Table 1). We also explored the use of additional equivalents of the tri-topic amine, since we have found previously that this favours complete conversion of the precursors to the desired product (Table 1, Entries 2-3). The use of an increased amount of tri-topic amine was found to

have no significant effect on the observed ratio of cage to dumbbell, so an excess of the amine was utilised in subsequent reactions.

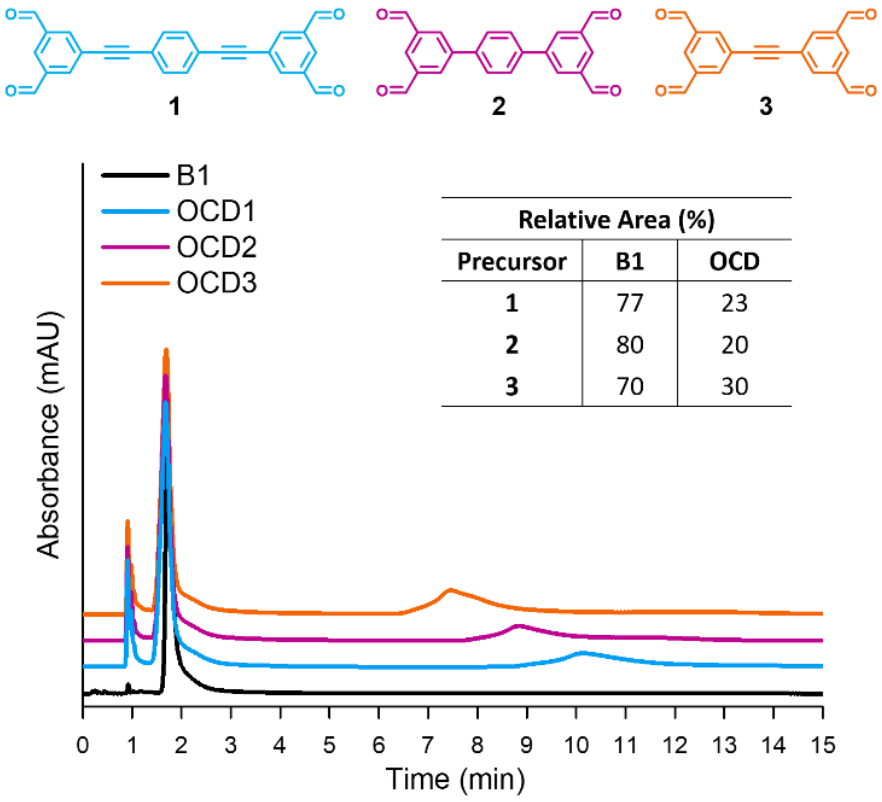
Table 1 Optimisation screen for the formation of organic cage dumbbell **OCD1** – reactions were carried out at 4.65 mM relative to triamine and refluxed for 2–3 days until no further equilibration was observed by HPLC. The ratio of **B1:OCD1** is determined by comparing the relative peak areas.

Entry	Tri-topic Amine (Eq.)	Di-topic Aldehyde (Eq.)	Tetra-topic Aldehyde (Eq.)	Solvent	HPLC Ratio B1:OCD1
1	4	4	1	$CHCl_3$	76:24
2	5	4	1	$CHCl_3$	79:21
3	6	4	1	$CHCl_3$	81:19
4	6	6	1	$CHCl_3$	75:25
5	6	5	1	$CHCl_3$	78:22
6	6	4	2	$CHCl_3$	78:22
7	6	4	3	$CHCl_3$	81:19
8	6	4	1	DCM	77:23
9[a]	5	4	1	DCM	72:28

[a] Reaction carried out on a larger 2 g scale

Next, both the equivalents of di-topic aldehyde (Entries 4-5), and tetra-topic aldehyde (Entries 6-7), were independently varied, increasing the ratio present in the reaction while the other precursor amounts were kept constant. Throughout, a ratio of ~80:20 **B1**:**OCD1** was consistently present in solution as analysed by HPLC, with no appreciable variance apparent based on the precursor feedstock (see Fig. S2). This is thought to be due to the formation of some insoluble polymer in the reaction mixtures, perturbing the thermodynamic equilibration in the solution. In particular, an increased amount of insoluble polymer was apparent in those reactions containing more equivalents of the tetra-topic aldehyde (Fig. S1). Finally, an alternative solvent was investigated (Entry 8), but again, this appeared to have no effect on the formed distribution.

The use of tetra-topic aldehydes of differing lengths (**1**-**3**) as alternative 'struts' was then investigated to determine if it would affect dumbbell formation (Fig. 2). Overall, a similar ratio of cage:dumbbell was always obtained, with LCMS confirming the peak at ~2 min corresponded to cage **B1**, and those between 7-12 min corresponding to **OCD1**-**3** (Fig. S3-S6, ESI†).



Relative Area (%)		
Precursor	B1	OCD
1	77	23
2	80	20
3	70	30

Fig. 1 Tetra-topic precursors **1**-**3** used as the central strut in organic cage dumbbells **OCD1**-**OCD3** along with the corresponding HPLC traces and relative peaks areas of the resulting mixtures of parent cage **B1** (~2 min) and the organic cage dumbbells (~7–12 min). The solvent front is visible at ~1 min.

There was no apparent change in the ratio of dumbbell to cage formation with different feedstock ratios, or when different solvents were used, so the preparation of **OCD1** was scaled-up using a slight excess of tri-topic amine (Table 1, Entry 9). For ease of isolation, dichloromethane was used as the solvent, since it allows a solvent exchange to hexane, with the resulting precipitated cage and dumbbell mixture being collected by filtration. On scale-up, a larger proportion of polymer formation was observed, which was removed from the crude mixture by filtration (28% based on mass recovery). After isolation from the resulting reaction solution by solvent exchange, a 72:28 ratio of **B1:OCD1** was obtained, in a 50% yield based on mass recovery. This mixture proved to have low solubility (~10 $mg.mL^{-1}$), but a small sample of **OCD1** was isolated by preparative HPLC in 93% purity (based on HPLC relative peak areas, Fig. S9) for characterisation (scale-up and purification details, alongside characterisation data of pure **OCD1** can be found in Section 1.3 of the ESI†), and a single crystal was grown from $CHCl_3$/MeCN, confirming the formation of the cage dumbbell structure (Fig. 3). To investigate the stability of the isolated dumbbell, ^1H NMR spectra were recorded in $CDCl_3$ immediately after dissolution and again 4 days later – no change was apparent suggesting that **OCD1** is not susceptible to re-equilibration in solution (Fig. S13).

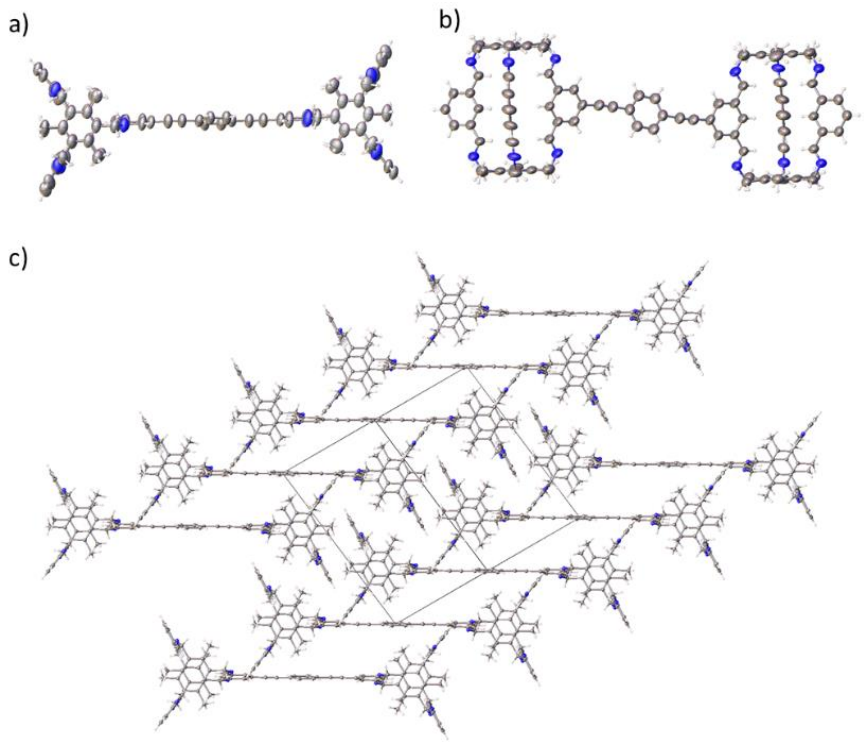


Fig. 3 (a) Top view, and (b) side view, displacement ellipsoid plots from the single crystal structure, **OCD1**·2($CHCl_3$)·10.4(H_2O), crystallised from $CHCl_3$/MeCN. Ellipsoid displayed at 20% probability level; and solvent omitted for clarity. (c) Crystal packing in the X-ray structure of **OCD1** solvate. Ordered and disordered solvent molecules, which fill the voids in the structure between the **OCD1** axels, are omitted for clarity.

We were interested in the chemical composition of the large quantity of insoluble polymer that formed during scale-up, as this may provide some insight into the consistent formation of a similar ratio of cage:dumbbell in solution. Therefore, this polymer, alongside the polymers formed from the reaction of the tri-topic amine with each of the di-topic and tri-topic aldehydes respectively, were synthesised under the same conditions and isolated. Overall, the largest quantity of insoluble polymer was formed when using a mixture of the aldehydes, followed by use of the tetra-topic aldehyde, and very little formed with the di-topic aldehyde which instead predominantly formed **B1** as analysed by HPLC (see Section 1.4 of the ESI†). Subsequent decomposition studies of the three insoluble polymers using trifluoroacetic acid in DMSO, followed by comparison of the ^{1}H NMR spectra, enabled the identification of the aldehyde composition in the polymer formed during the dumbbell formation – a 1:1.3 ratio of di-topic to tetra-topic aldehyde was apparent (Figure S18). We therefore believe that the observed precipitation is predominantly driven by the reduced solubility of the tetra-topic aldehyde derived oligomeric and polymeric species, but the removal of both of the precursors from solution is affecting the thermodynamic equilibration of the cage and dumbbell in solution.

We were also interested as to why there was consistent formation of a mixture of both cage and dumbbell, and so we turned to computational modelling to explore this. Previously, we showed that it is possible to predict the most likely cage topology formed from two precursors by calculating and comparing the formation energies per imine bond for different molecular

assemblies.[22,23,25] These calculations are performed on isolated molecules in the gas phase, which does not consider solvent effects, and hence large energetic differences are needed to predict solution-phase structures with confidence. This process becomes more difficult for complex precursor mixtures, as the number of potential assemblies that might form increases. We therefore decided to consider only the species that were identified experimentally in the reaction solution: that is, the parent cage **B1** and the three dumbbells, **OCD1**-**3**. For each species, we searched for the low-energy conformations using high-temperature molecular dynamics simulations with the OPLS3 force-field[26] before further optimising the geometries at the PBE+D3/TZVP-MOLOPT level (Fig. 4), finally carrying out single point energy calculations with M06-2X/6-311G(3df,3pd).[27–33] Overall, all of the dumbbells had similar DFT formation energies of −17.2, −17.2, −15.7 kJ mol^{-1} per imine bond for **OCD1-3**, respectively, compared to −16.6 kJ mol^{-1} per imine bond for the parent cage **B1**. These similar formation energies rationalise why a product mixture is consistently observed in all three reactions.

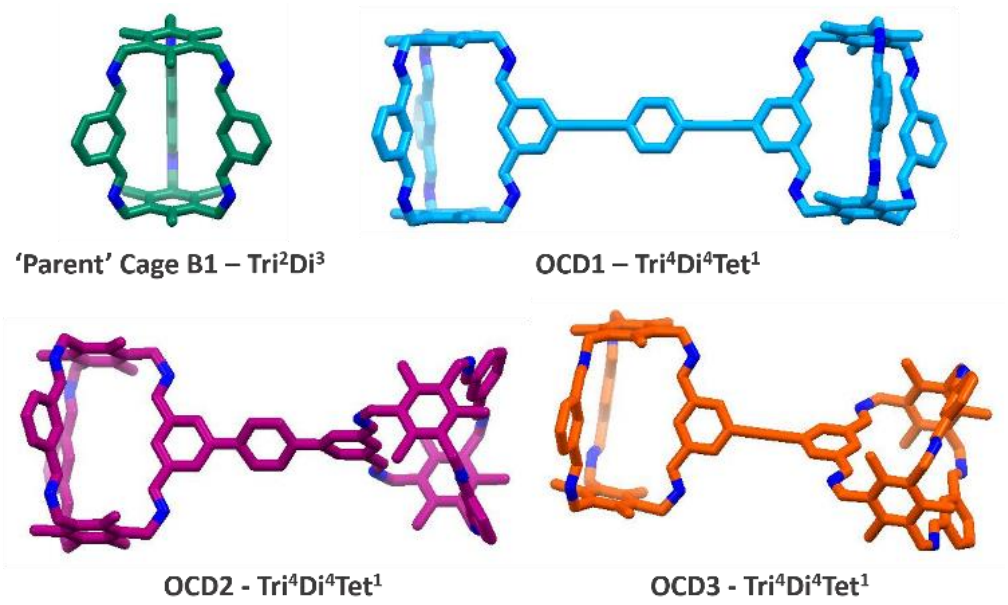


Fig. 4 DFT (PBE-D3/TZVP) optimised structures for the parent **Tri^{2}Di^3** cage **B1**, and each of the **Tri^{4}Di^{4}Tet^1** organic cage dumbbells **OCD1**-**3**. Hydrogens are not shown, nitrogens are in blue and carbons coloured differently for each molecule.

The potential energy surfaces for each of the members of the **OCD1-3** family were found to consist of multiple low-lying minima, which differed only in relative orientation of the two cage-component ends of the dumbbells. We further investigated these orientations by calculating how the OPLS3e force-field[34] potential energy changes with the dihedral angle between the two cages. All relative orientations of the cages are practically degenerate for **OCD1** and **OCD3**, while **OCD2** exhibits a few slightly preferred orientations (Fig. S20, ESI†). Similarly, an extended molecular dynamics simulation at room temperature (Fig. S21, ESI†) showed that **OCD1** and **OCD3** can be found in any relative orientation, as exemplified by the disorder in the single crystal structure of **OCD1**.

Conclusions

In summary, we have demonstrated the synthesis of organic cage dumbbells, where the two capping units are organic cage molecules, using a multi-component social-sorting reaction. These dumbbells were pre-dominantly formed alongside a binary parent cage, which could be rationalized using computational modelling. It has been shown previously that two organic cages can be mechanically interlocked in a cage catenane;[35,36] also, multiple organic cages have been covalently connected in polymers or frameworks.[37,38] To the best of our knowledge, this study is the first example of covalently connecting two cages together in a controlled manner by self-sorting, thus providing proof-of-concept for more complex and controlled architectures involving more than a single cage species. There are only a few other examples of cage-like or supramolecular dumbbells, which incorporate either fullerenes[39] or knots[40] as the capping units, and those strategies involve a coupling of the pre-formed capping units, rather than a one-pot self-assembly. These OCDs might be interesting struts in future rotaxane architectures. Strategies that might be employed in the future to favour dumbbell formation include the use of a pre-formed parent cage and subsequent linker exchange with a tetra-topic aldehyde in a formal transamination reaction,[41] or pre-formation of a mono-functionalised cage and react with a strut post-assembly using a cross-coupling reaction.[37,38]

Acknowledgements

We acknowledge funding from the European Research Council under FP7 (RobOT, ERC Grant Agreement no. 321156 and CoMMaD, ERC Grant No. 758370), and the Engineering and Physical Sciences Research Council (EPSRC) (EP/R005540/1), including the 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 and for a Leverhulme Trust Research Project Grant. R.L.G. thanks the Royal Society for a University Research Fellowship. We thank Diamond Light Source for access to beamlines I19 (CY21726). We also thank Stephen Moss and the MicroBioRefinery for assistance with QTOF-MS and LC-MS measurements.

Conflict of Interest

There are no conflicts to declare.

References

[1] G. Zhang, M. Mastalerz, *Chem. Soc. Rev.* **2014**, *43*, 1934–1947.

[2] J. D. Evans, C. J. Sumby, C. J. Doonan, *Chem. Lett.* **2015**, *44*, 582–588.

[3] T. Hasell, A. I. Cooper, *Nat. Rev. Mater.* **2016**, *1*, 16053.

[4] F. Beuerle, B. Gole, *Angew. Chemie - Int. Ed.* **2018**, *57*, 4850–4878.

[5] K. Ono, N. Iwasawa, *Chem. Eur. J.* **2018**, *24*, 17856–17868.

[6] K. Acharyya, P. S. Mukherjee, *Angew. Chemie Int. Ed.* **2019**, *58*, 8640–8653.

[7] K. Acharyya, P. S. Mukherjee, *Chem. - A Eur. J.* **2014**, *20*, 1646–1657.

[8] K. Acharyya, S. Mukherjee, P. S. Mukherjee, *J. Am. Chem. Soc.* **2013**, *135*, 554–557.

[9] M. Kołodziejski, A. R. Stefankiewicz, J.-M. Lehn, *Chem. Sci.* **2019**, DOI 10.1039/C8SC04598D.

[10] X. Wang, P. Peng, W. Xuan, Y. Wang, Y. Zhuang, Z. Tian, X. Cao, *Org. Biomol. Chem.* **2017**, *16*, 34–37.

[11] S. Jiang, J. T. A. Jones, T. Hasell, C. E. Blythe, D. J. Adams, A. Trewin, A. I. Cooper, *Nat. Commun.* **2011**, *2*, 207.

[12] Q. Wang, C. Zhang, B. C. Noll, H. Long, Y. Jin, W. Zhang, *Angew. Chemie - Int. Ed.* **2014**, *53*, 10663–10667.

[13] S. Lee, A. Yang, T. P. Moneypenny, J. S. Moore, *J. Am. Chem. Soc.* **2016**, *138*, 2182–2185.

[14] J. C. Lauer, W. S. Zhang, F. Rominger, R. R. Schröder, M. Mastalerz, *Chem. - A Eur. J.* **2018**, *24*, 1816–1820.

[15] S. Klotzbach, F. Beuerle, *Angew. Chemie - Int. Ed.* **2015**, *54*, 10356–10360.

[16] D. Beaudoin, F. Rominger, M. Mastalerz, *Angew. Chemie - Int. Ed.* **2017**, *56*, 1244–1248.

[17] R. L. Greenaway, V. Santolini, A. Pulido, M. A. Little, B. M. Alston, M. E. Briggs, G. M. Day, A. I. Cooper, K. E. Jelfs, *Angew. Chemie* **2019**, *131*, 16421–16427.

[18] N. De Rycke, J. Marrot, F. Couty, O. R. P. David, *Tetrahedron Lett.* **2010**, *51*, 6521–6525.

[19] P. Mateus, R. Delgado, P. Brandão, V. Félix, *J. Org. Chem.* **2009**, *74*, 8638–8646.

[20] M. Arunachalam, I. Ravikumar, P. Ghosh, *J. Org. Chem.* **2008**, *73*, 9144–9147.

[21] O. Francesconi, A. Ienco, G. Moneti, C. Nativi, S. Roelens, *Angew. Chemie Int. Ed.* **2006**, *45*, 6693–6696.

[22] R. L. Greenaway, V. Santolini, M. J. Bennison, B. M. Alston, C. Stackhouse, M. A. Little, M. Miklitz, E. G. B. Eden, R. Clowes, A. Shakil, et al., *Nat. Commun.* **2018**, *9*, 2849.

[23] V. Santolini, M. Miklitz, E. Berardo, K. E. Jelfs, *Nanoscale* **2017**, *9*, 5280–5298.

[24] V. Santolini, M. Miklitz, E. Berardo, K. E. Jelfs, *Nanoscale* **2017**, *9*, 5280–5298.

[25] G. Zhu, Y. Liu, L. Flores, Z. R. Lee, C. W. Jones, D. A. Dixon, D. S. Sholl, R. P. Lively, *Chem. Mater.* **2018**, *30*, 262–272.

[26] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, et al., *J. Chem. Theory Comput.* **2016**, *12*, 281–296.

[27] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[28] J. VandeVondele, J. Hutter, *J. Chem. Phys.* **2007**, *127*, 114105.

[29] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456–1465.

[30] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*.

[31] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[32] A. D. McLean, G. S. Chandler, *J. Chem. Phys.* **1980**, *72*, 5639–5648.

[33] M. J. Frisch, J. A. Pople, J. S. Binkley, *J. Chem. Phys.* **1984**, *80*, 3265–3269.

[34] K. Roos, C. Wu, W. Damm, M. Reboul, J. M. Stevenson, C. Lu, M. K. Dahlgren, S. Mondal, W. Chen, L. Wang, et al., *J. Chem. Theory Comput.* **2019**, *15*, 1863–1874.

[35] T. Hasell, X. Wu, J. T. Jones, J. Bacsa, A. Steiner, T. Mitra, A. Trewin, D. J. Adams, A. I. Cooper, *Nat. Chem.* **2010**, *2*, 750–755.

[36] G. Zhang, O. Presly, F. White, I. M. Oppel, M. Mastalerz, *Angew. Chem. Int. Ed.* **2014**, *53*, 5126–5130.

[37] Y. Jin, B. A. Voss, R. McCaffrey, C. T. Baggett, R. D. Noble, W. Zhang, *Chem. Sci.* **2012**, *3*, 874–877.

[38] J.-X. Ma, J. Li, Y.-F. Chen, R. Ning, Y.-F. Ao, J.-M. Liu, J. Sun, D.-X. Wang, Q.-Q. Wang, *J. Am. Chem. Soc.* **2019**, *141*, 3843–3848.

[39] T. Wei, M. E. Pérez-Ojeda, A. Hirsch, *Chem. Commun.* **2017**, *53*, 7886–7889.

[40] O. Lukin, J. Recker, A. Böhmer, W. M. Müller, T. Kubota, Y. Okamoto, M. Nieger, R. Fröhlich, F. Vögtle, *Angew. Chemie - Int. Ed.* **2003**, *42*, 442–445.

[41] E. Vitaku, W. R. Dichtel, *J. Am. Chem. Soc.* **2017**, *139*, 12911–12914.