Organic Cage Dumbbells

Rebecca L. Greenaway,[a]* Valentina Santolini,[b] Filip T. Szczypiński,[b] Michael J. Bennison,[a] Marc A. Little,[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
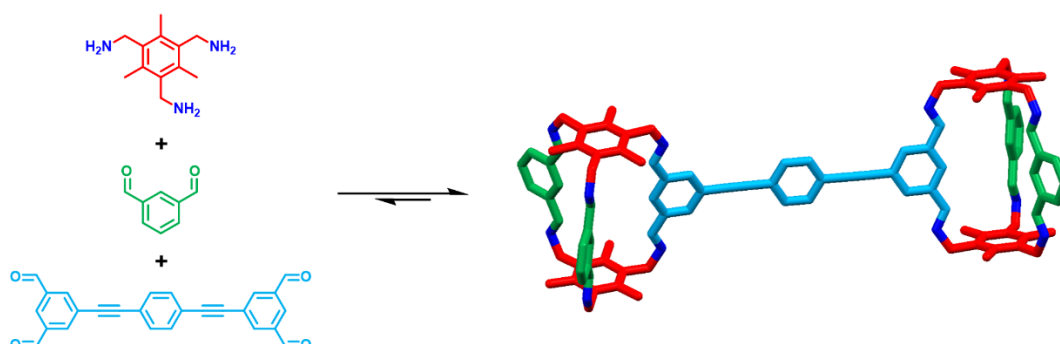
Introduction

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 various more complex outcomes, including social self-sorting into a single mixed assembly, 'scrambling' to form statistical mixtures of assemblies, or narcissistic self-sorting into separate discrete species (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 narcissistically self-sorted binary cages,[6–8] a distribution of cage species,[9–12] or, less commonly, socially self-sorted ternary cage assemblies.[13–15] However, these examples typically exploit the use of precursors of the same topicity (*i.e.,* number of reactive functional groups), and they target individual relatively symmetrical organic cage species, rather than more complex architectures.

Here, we present an example of social-narcissistic self-sorting. We used four tri-topic amines, four di-topic aldehydes, and one tetra-topic aldehyde. Narcissistic self-sorting was observed into two distinct species: socially self-sorted organic cage dumbbells (OCDs) and the corresponding binary 'parent' cage (Scheme 1). 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.
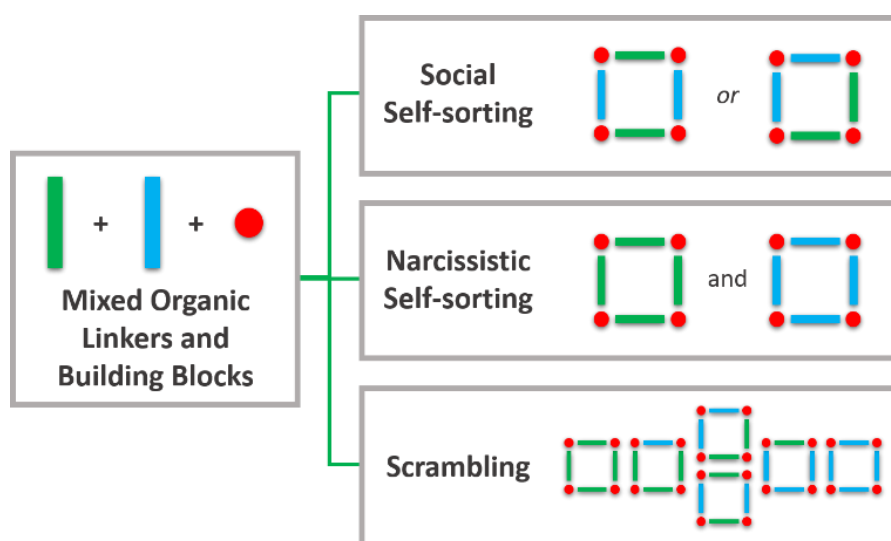

Fig 1. Schematic representation of potential self-assembly outcomes for more than two organic linkers or building blocks: social self-sorting leads to the formation of a single mixed assembly, narcissistic self-sorting results in the formation of separate discrete species, and scrambling leads to a statistical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

Results and Discussion

Previously, we reported a high-throughput workflow for organic cage discovery that included a number of **Tri^{2}Di^3** species – these molecules consist of two tri-topic and three di-topic

building blocks.[16,17] While organic cages of this topology tend to have cavities that are too small to host guests, they have a simple trigonal geometry that lends itself to the design of more complex assemblies. Our aim was to use multi-component imine condensation to access a narcissistically self-sorted and controlled assembly incorporating more than one cage, instead of socially 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)[16,18], along with a tetra-topic aldehyde that has the same 1,3-dialdehyde substitution pattern (Scheme 1).
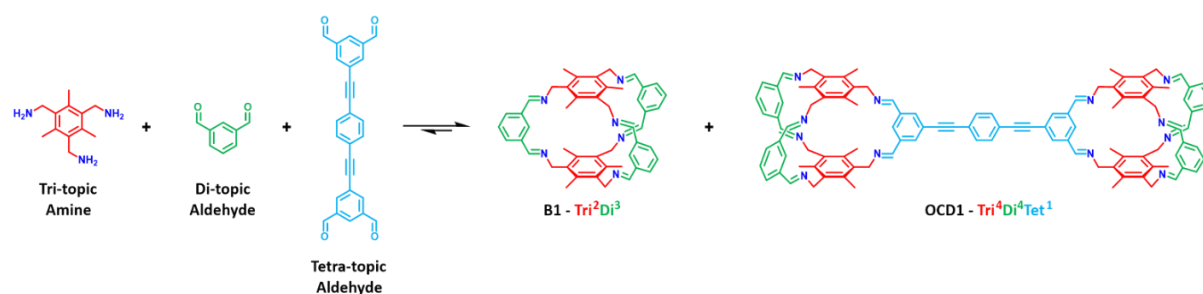
Scheme 1 Reaction scheme for the social-narcissitc self-sorting of a tri-topic amine (red), di-topic aldehyde (green), and tetra-topic aldehyde (blue), to form a mixture of the binary parent cage **B1** (**Tri^{2}Di^3**) and an organic cage dumbbell **OCD1** (**Tri^{4}Di^{4}Tet^1**).

Building on the reaction conditions that we reported for the synthesis of **B1** (Scheme 1),[16] we altered the precursor stoichiometry to include a single tetra-topic aldehyde for every two targeted organic cages, leading to a ratio of 4:4:1 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 crude reaction mixture by HPLC and high-resolution mass spectrometry (HRMS) showed the clean formation of two species: the binary 'parent' **Tri^{2}Di^3** cage (**B1**, $[M+H]^+$ 709.4024) and a **Tri^{4}Di^{4}Tet^1** species ($[M+2H]^{2+}$ 770.9116), corresponding to an organic cage dumbbell (**OCD1**), in a 76:24 ratio, as determined by HPLC (a/a%). 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 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). This 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. 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 whilst the other precursor amounts were kept constant. Throughout, a ratio of ~80:20 **B1:OCD1** was consistently formed as analysed by HPLC, with no appreciable variance apparent based on the precursor feedstock. This could potentially be due to the formation of some insoluble polymer in the reaction mixtures, perturbing the thermodynamic equilibration. Finally, an alternative solvent was investigated (Entry 8), but again, this appeared to have no effect on the formed distribution.

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.

Entry	Tri-topic Amine (Eq.)	Di-topic Aldehyde (Eq.)	Tetra-topic Aldehyde (Eq.)	Solvent	HPLC Ratio B1:OCD1 (a/a%)
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

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. S2-S5, ESI†).
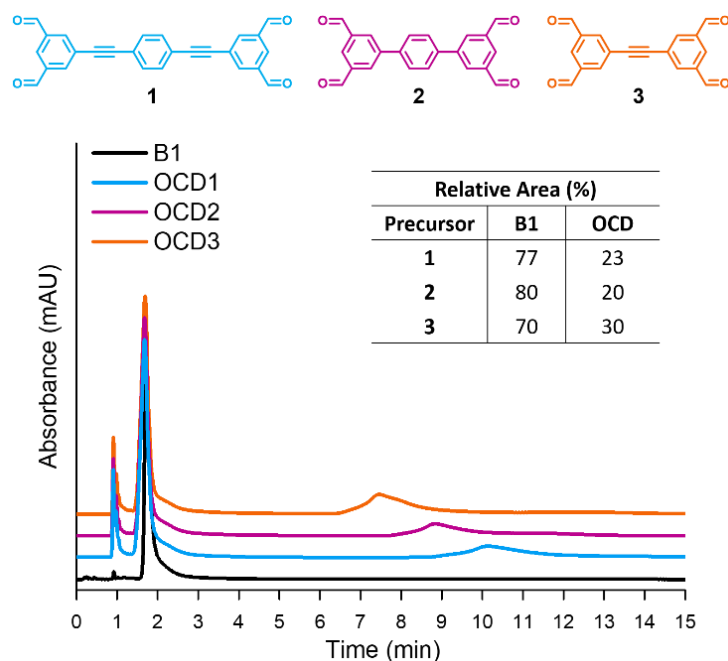

Fig. 1 Tetra-topic precursors **1**-**3** used as the central strut in organic cage dumbbells **OCD1**-**OCD3** along with the corresponding HPLC traces and relative peaks areas of the resulting mixtures of parent cage **B1** (~2 min) and the organic cage dumbbells (~7–12 min).

4

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 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 for characterisation, and a single crystal was grown from CHCl$_3$/MeCN confirming the formation of the cage dumbbell structure (Fig. 3). While only the dumbbell peak was collected, HPLC analysis of the isolated purified solid showed a 7:93 ratio of **B1:OCD1**, suggesting that the isolated dumbbell is susceptible to re-equilibration.
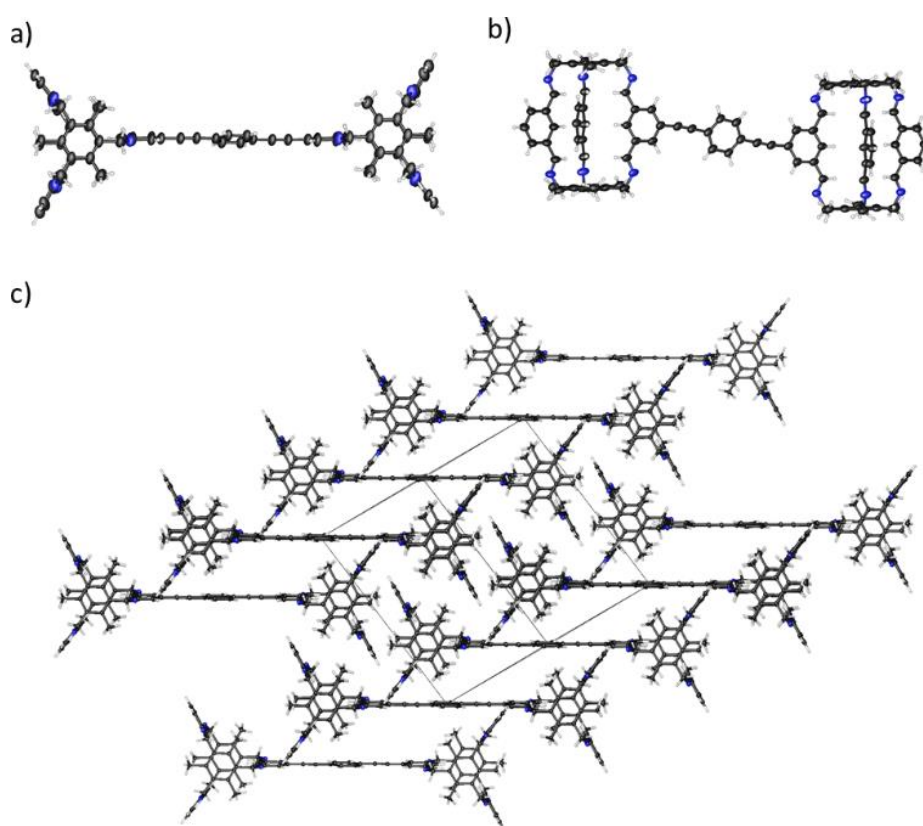

Fig. 3 (a) Top view, and (b) side view, displacement ellipsoid plots from the single crystal structure, **OCD1**·2.5(CHCl$_3$)·9.5(H$_2$O), crystallised from CHCl$_3$/MeCN. Ellipsoid displayed at 30% probability level; and solvent omitted for clarity. (c) Crystal packing in the X-ray structure of **OCD1**·2.5(CHCl$_3$)·9.5(H$_2$O). Ordered and disordered solvent molecules, which fill the voids in the structure between the **OCD1** axels, are omitted for clarity.

We were interested in 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.[16,17,19] 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 formed experimentally: 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[20] 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).[21–27] 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.
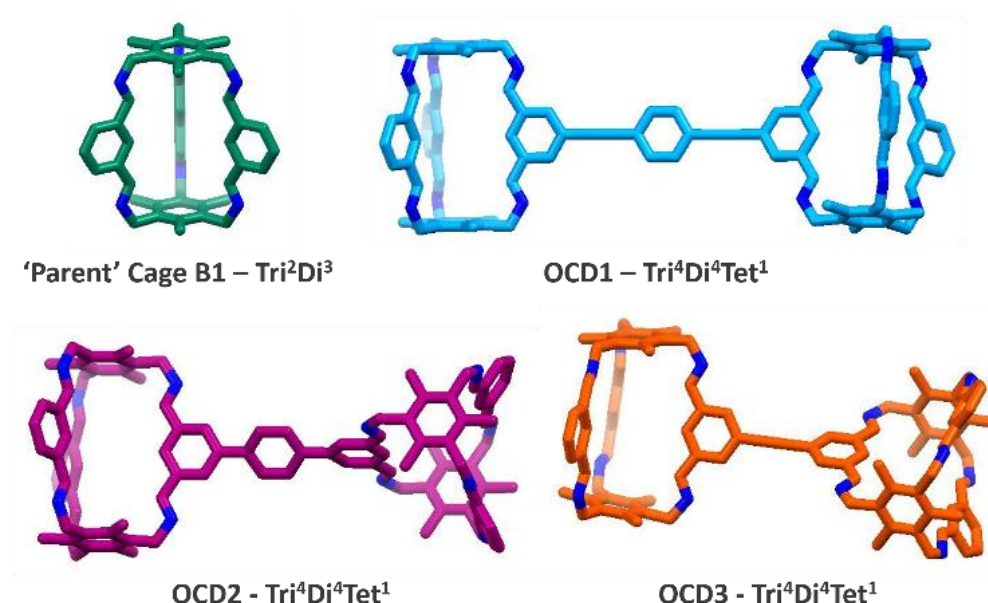

'Parent' Cage B1 − Tri^{2}Di^3 OCD1 − Tri^{4}Di^{4}Tet^1

OCD2 - Tri^{4}Di^{4}Tet^1 OCD3 - Tri^{4}Di^{4}Tet^1

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**.

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[28] 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. S13-S14, ESI†).

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-narcissistic self-sorting reaction. It has been shown previously that two organic cages can be mechanically interlocked in a cage

catenane;[29,30] also, multiple organic cages have been covalently connected in polymers or frameworks.[31,32] 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[33] or knots[34] 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,[35] or pre-formation of a mono-functionalised cage and react with a strut post-assembly using a cross-coupling reaction.[31,32]

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, S. Mukherjee, P. S. Mukherjee, *J. Am. Chem. Soc.* **2013**, *135*, 554–557.

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

[8]   M. Kołodziejski, A. R. Stefankiewicz, J.-M. Lehn, *Chem. Sci.* **2018**, *10*, 1836–1843.

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

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

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

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

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

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

[15]   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.

[16]   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.

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

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

[19]   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.

[20]   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.

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

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

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

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

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

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

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

[28]   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.

[29]   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.

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

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

[32]   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.

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

[34]   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.

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