## COMMUNICATION

# Crystal structure of a complex between β-glucopyranose and a macrocyclic receptor with dendritic multicharged water solubilizing chains†

Pradeep K. Mandal,[ab] Brice Kauffmann,[cde] Harry Destecroix,[f] Yann Ferrand,[ab] Anthony P. Davis*[f] and Ivan Huc*[ab]

**Using commercial screens for crystallization of biomolecules and taking advantage of the use of racemic crystallography allowed the production of X-ray quality single crystals and the elucidation at 1.08 Å resolution of the solid state structure of a difficult target: the complex between glucopyranose and a water soluble macrocyclic receptor equipped with dendritic multianionic solubilizing chains.**

Solid state structure elucidation by single crystal X-ray crystallography has been an invaluable tool to many scientific disciplines and particularly so to supramolecular chemistry. However, in some contexts, the use of X-ray crystallography has been limited by perceived experimental difficulties with respect to crystal growth. In the following we present the successful elucidation of the crystal structure of a host–guest complex that combines several of these perceived difficulties and show that they can nevertheless be overcome. The underlying message is that current tools and methods may achieve more than one may expect and that one should definitely not refrain from undertaking screening for conditions favourable to crystallogenesis.

Among criteria that may represent a (mis-)perceived difficulty for crystal growth is the fact that the supramolecular object of interest may be soluble only in water. Developing supramolecular chemistry in water represents an important and topical subject,[1] including very challenging carbohydrate recognition.[2] Fortunately, methods for crystallogenesis from water have already been developed for biomacromolecules that are often larger,

more complex, and available in smaller quantities than supramolecular systems.[3] Organic chemists may not always be familiar with these methods that differ from those used to crystallize lipophilic supramolecules. Yet their efficiency is well-demonstrated. Furthermore, a number of large and complex supramolecular objects have already been crystallized from water though not always by using commercial screens developed for proteins.[4] Thus, that water as a solvent would in itself constitute a difficulty is probably a misperception. On the contrary, the possibility to vary salt nature and concentration, pH, and the availability of innumerable additives provide advantages to working in water over organic solvents.

Other criteria perceived as making crystal growth difficult include a high charge density (but again DNA would constitute a striking counter-example), the presence of dendritic segments, and the presence of saccharides. Concerning the latter two, it is true that very few dendrimers have been crystallized and that these were early generations.[5] Furthermore, after decades of research on synthetic saccharide receptors, the first examples of crystal structures of host–guest complexes with neutral unsubstituted monosaccharides have only been reported in 2015.[6,7] Finally, low binding affinity also makes it more difficult to crystallize a complex.

In the following, we report the successful growth of single crystals suitable for X-ray diffraction and the resolution of a solid-state structure that combines all of the apparent inconveniences mentioned above: solubility in water, multicharged state, presence of dendritic groups, binding of an unsubstituted mono-saccharide, and low binding affinity. The selective recognition of β-D-glucopyranose D-**2** by hexaanionic macrocycle **1** (Fig. 1) in water has previously been reported.[8] The 1 : 1 binding constant of 56 L mol$^{-1}$ and the observed selectivity for an all equatorial sugar can be considered to be very good from the perspective of saccharide recognition in water.[2] However, this low value did not bode well for growing crystals of a complex. Because of this and the other features of the system, and despite the existence of crystal structures of related organic soluble anthracene-based host,[9] crystallization was not attempted at the time and a solution

$^a$ *Univ. Bordeaux, CBMN (UMR 5248), IECB, 2 rue Robert Escarpit, F-33600 Pessac,*
 *France. E-mail: i.huc@iecb.u-bordeaux.fr*

$^b$ *CNRS, CBMN (UMR 5248), F-33600 Pessac, France*

$^c$ *Univ. Bordeaux, Institut Européen de Chimie et Biologie (UMS 3033),*
 *F-33600 Pessac, France*

$^d$ *CNRS, IECB, UMS 3033, F-33600 Pessac, France*

$^e$ *INSERM, IECB, US 001, F-33600 Pessac, France*

$^f$ *School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK.*
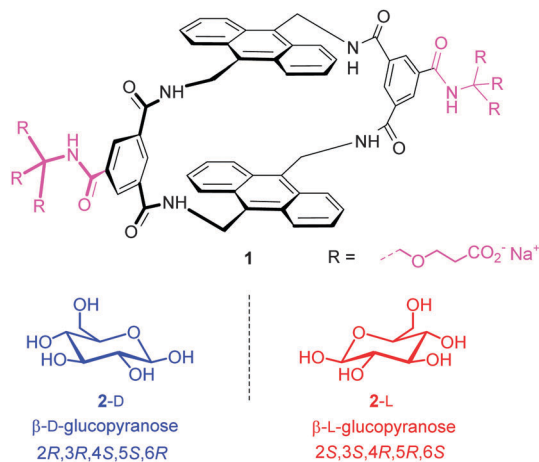 *E-mail: Anthony.Davis@bristol.ac.uk*

**Fig. 1** Structures of the macrocyclic anthracene based receptor **1** and of the two enantiomers of β-glucopyranose **2**.

structure of a complex with methyl β-D-glucoside, which has a slightly larger affinity than β-D-glucose, was proposed based on NMR data.[8]

Encouraged by recent successes at elucidating the crystal structure of several receptor–saccharide complexes from organic solvents,[6] thanks to the use of racemic crystallography, we eventually attempted the crystallization of **1** ⊃ **2**. Racemic crystallography exploits the tendency of racemic solutions to crystallize as racemic crystals, and thus to offer possibilities of packing and crystal growth additional to those of enantiopure conditions. Racemic crystals have access to chirality inverting symmetry operations such as centres and planes of symmetry, and tend to be denser than their chiral counterparts, an observation known as the Wallach rule. The reasons for this close-packing and whether it is truly related to the frequent occurrence of racemates were debated.[10] In any case, racemic crystallization has been long known for small organic molecules,[11] and is advantageous for helical aromatic foldamers,[8,12] β-peptides,[13] peptides,[14] small proteins,[15] nucleic acids,[16] and, as it turns out, saccharides. Thus, a solution of anthracene-based receptor **1** in pure water was exposed to a large excess of a racemic mixture of the two enantiomers of β-glucopyranose. The solution used for crystallization was made of 1.5 mM of **1** and 0.5 M of D/L-glucose, a large excess of guest sufficient to saturate the host despite the low affinity. In order to find suitable crystallization conditions, commercial sparse-matrix screens were used. Crystallization trials were performed using the standard vapour diffusion method at 293 K. X-ray quality crystals (see Fig. S1, ESI†) were grown in aqueous hanging drops prepared by mixing the **1** ⊃ (**2**-D/**2**-L) solution with an equal volume of a crystallization reagent composed of 100 mM HEPES buffer (pH 7.5), 200 mM CaCl$_2$ and 28% polyethylene glycol (PEG) 400. The presence of calcium in favourable conditions is not surprising as this metal ion is often seen in the structures of nucleic acids[17] and other polyanionic species.[18] Importantly, no crystals were obtained under these conditions if a single enantiomer of β-glucopyranose was used.

The crystals obtained from the **1** ⊃ (**2**-D/**2**-L) solution diffracted X-rays at moderate resolution, and a full data set was collected

and processed to 1.08 Å in space group $P2_1/c$. Table S1 (ESI†) summarizes the crystal data and refinement parameters. Experimental details are also presented in the ESI.† The structure revealed a network of dodecameric clusters of macrocycles bridged by calcium cations having a pseudo $C_3$ symmetry through their long axis (Fig. 2a, see ESI,† for views of the packing of the clusters). This multi-anionic-receptor/calcium coordination network is reminiscent of structures of molecular organic frameworks (MOFs), among which Ca$^{2+}$ is not a common metal ion.[19] Other than coordination bonds, the structure is held together by inter-molecular hydrogen bonds and aromatic π–π stacking. Each dodecameric arrangement is composed of a central core including six guest-free imbricated macrocycles. The top and bottom parts of the dodecamer each contain three macrocycle-glucose host–guest complexes. As the structure is centrosymmetric, each set of three receptors include either β-D-glucopyranose (blue, Fig. 2a–c) or β-L-glucopyranose (red, Fig. 2a–c). The dodecameric assembly is held together via 15 bridging calcium cations coordinated either to isophthalamide carbonyls of **1** or to its side chain carboxylates. Five out of the six solubilizing chains of free receptors are involved in interactions with calcium cations. Unexpectedly, the sixth chain is sandwiched in-between anthracene moieties of a neighbour macrocycle (in orange in Fig. 2b, d and 3o–q) and thus occupies the sugar binding site.

Detailed analysis of the moderate resolution structure provided key insights into the mode of interactions between the macrocycle and the carbohydrate. As predicted from the receptor design, the all-equatorial β anomer of **2** lies flat in the macrocycle sandwiched in between the two parallel anthracenyl moieties (Fig. 3c–e). A distance as low as 2.5 Å between the axial hydrogens of the guest and the aromatic surfaces of the host was measured. The top view shows that the position of the sugar between peripheral benzenic rings of anthracene is offset with respect to the center of the complex. Thus, two degenerate binding modes are available. The four equatorial hydroxyls of β-glucopyranose were found to form strong hydrogen bonds with NH protons of the isophthalamide pillars of the macrocycle ($d$(O···HN) ≤ 2.11 Å). The C$_6$ primary hydroxyl was shown to point outside the cavity thus not interacting with the receptor.

This structure shows notable differences from the NOE-based solution structure of the complex with methyl β-D-glucoside.[8] In the latter structure, the C$_1$ methoxy group was found to point out of the receptor while the C$_6$ hydroxy group was involved in hydrogen bonding. However, it is unsurprising that a different geometry is adopted for glucose **2**, as the steric and hydrogen-bonding requirements of this substrate are dissimilar. Meanwhile, the binding geometry of methyl β-D-glucoside is accessible to **2** and its existence in solution cannot be excluded. Other important structural features revealed in the solid state concern inter-actions between macrocycles. In addition to aromatic stacking between isophthalamide and anthracene rings, we observe the intercalation of some propionic acid side chains within the cavity of a neighbour macrocycle (Fig. 3b, 3o and 3q). The carboxylate group is inserted parallel to the anthracenyl moieties and is doubly hydrogen bonded to the NH amides of one macrocycle pillar. NMR studies indicate that **1** self-associates at
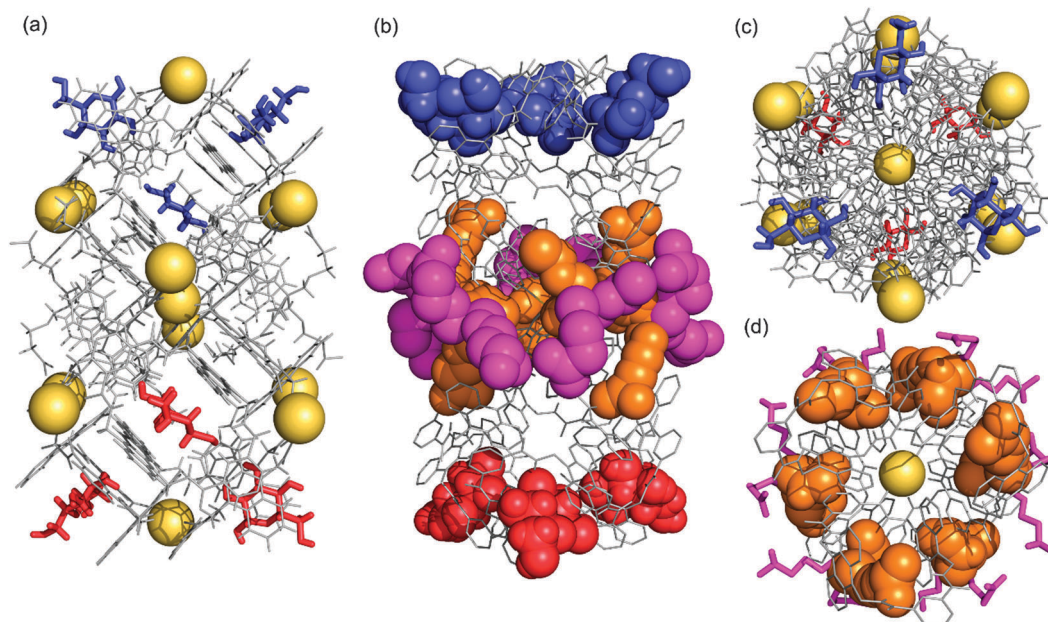
**Fig. 2** Molecular packing in the crystal structure of **1**⊃**2**. (a) Side view of the dodecameric metallo-organic cluster of twelve molecules of **1**. The macrocycles are shown in light grey tube representation whereas D-β-glucopyranose and L-β-glucopyranose are coloured in blue and red, respectively. Some calcium cations are shown as gold CPK spheres. (b) Other view of the dodecameric cluster highlighting the insertion of some carboxylate side chains in the cavities of molecules of **1** that do not contain **2**. These chains are shown in orange CPK whereas other side chains are shown in magenta. Molecules of **2** are shown in CPK representation with the same colour code as in (a). Calcium ions have been omitted for clarity. (c) Top view of the cluster showing the pseudo $C_3$ symmetry of the aggregate. (d) Top view of the six central macrocycles of the dodecamer (*i.e.* peripheral macrocycles that host **2** have been removed) showing the arrangement of the intercalated carboxylate side chains. The side chains are coloured as in (b).
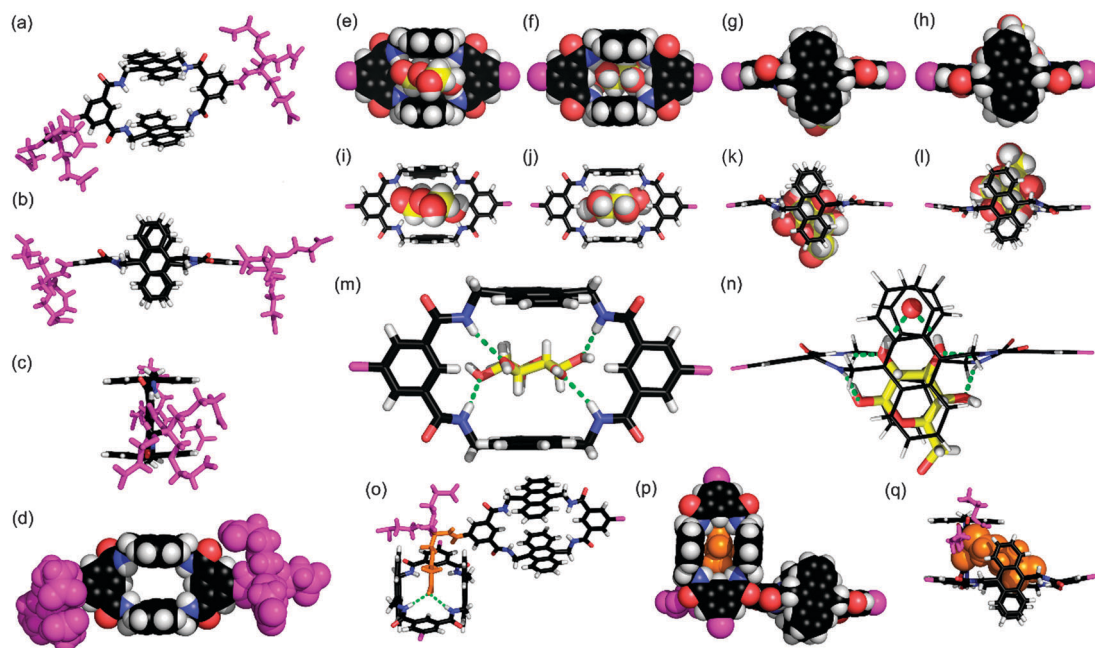


**Fig. 3** Crystal structure of **1**⊃**2**. (a) front view, (b) top view, (c) side view, and (d) back view of macrocycle **1** shown in tube (a–c) or CPK (d) representations. The dendritic side chains are coloured in magenta. (e and i) front view, (f and j) back view, (g and k) top view and (h and l) bottom view of the structure of host–guest complex **1**⊃**2**-D. The carbohydrate is shown in CPK representation in all cases whereas the macrocycle is shown either in CPK (e–h) or in tube (i–l) representations. Enlarged front (m) and top (n) views of **1**⊃**2**-D in tube representation show the array of host–guest intermolecular hydrogen bonds. In (n) an additional water molecule sandwiched in between the two anthracenyl moieties of the receptor is involved in two interactions with glucose. In (e–n) the solubilizing side chains have been omitted for clarity. (o–q) Views of the macrocycle carboxylate side chain (orange) intercalated in the empty cavity a neighbour macrocycle.

high concentrations in solution,[8] and this interaction suggests a possible mode of aggregation.

## Notes and references

1 (a) E. A. Kataev and C. Müller, *Tetrahedron*, 2014, **70**, 137; (b) A. Dalla Cort, G. Forte and L. Schiaffino, *J. Org. Chem.*, 2011, **76**, 7569; (c) A. P. Davis, *Nature*, 2010, **464**, 169; (d) S. Kubik, *Chem. Soc. Rev.*, 2009, **38**, 585; (e) G. V. Oshovsky, D. N. Reinhoudt and W. Verboom, *Angew. Chem., Int. Ed.*, 2007, **46**, 2366; (f) E. A. Meyer, R. K. Castellano and F. Diederich, *Angew. Chem., Int. Ed.*, 2003, **42**, 1210.

2 (a) T. J. Mooibroek, J. M. Casas-Solvas, R. L. Harniman, C. M. Renney, T. S. Carter, M. P. Crump and A. P. Davis, *Nat. Chem.*, 2016, **8**, 69; (b) N. P. Barwell, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2009, **48**, 7673; (c) Y. Ferrand, E. Klein, N. P. Barwell, M. P. Crump, J. Jimenez-Barbero, C. Vicent, G. J. Boons, S. Ingale and A. P. Davis, *Angew. Chem., Int. Ed.*, 2009, **48**, 1775; (d) Y. Ferrand, M. P. Crump and A. P. Davis, *Science*, 2007, **318**, 619; (e) E. Klein, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2005, **44**, 298.

3 (a) A. McPherson, *Introduction to Macromolecular Crystallography*, Hoboken, Wiley-Blackwell, 2nd edn, 2009; (b) R. Giegé, B. Lorber and A. Théobald-Dietrich, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1994, **50**, 339; (c) A. Ducruix and R. Giegé, in *Crystallization of Nucleic Acids and Proteins. A Practical Approach*, ed. A. Ducruix and R. Giegé, Oxford, IRL Press/Oxford Univ. Press, 1992.

4 (a) T. Sawada, M. Yoshizawa, S. Sato and M. Fujita, *Nat. Chem.*, 2009, **1**, 53; (b) C. Zhao, Q.-F. Sun, W. M. Hart-Cooper, A. G. Di Pasquale, F. D. Toste, R. G. Bergman and K. N. Raymond, *J. Am. Chem. Soc.*, 2013, **135**, 18802; (c) J. L. Bolliger, T. K. Ronson, M. Ogawa and J. R. Nitschke, *J. Am. Chem. Soc.*, 2014, **136**, 14545.

5 (a) B. Huang and J. R. Parquette, *Org. Lett.*, 2000, **2**, 239; (b) O. Lukin, D. Schubert, C. M. Müller, W. B. Schweizer, V. Gramlich, J. Schneider, G. Dolgonos and A. Shivanyuk, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 10922; (c) R. E. Bauer, V. Enkelmann, U. M. Wiesler, A. J. Berresheim and K. Müllen, *Chem. – Eur. J.*, 2002, **8**, 3858.

6 N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil and I. Huc, *Nat. Chem.*, 2015, **7**, 334.

7 For an earlier example with a substituted saccharide, see: M. Mazik, H. Cavga and P. G. Jones, *J. Am. Chem. Soc.*, 2005, **127**, 9045.

8 C. Ke, H. Destecroix, M. P. Crump and A. P. Davis, *Nat. Chem.*, 2012, **4**, 718.

9 (a) J. J. Gassensmith, J. M. Baumes, J. Eberhard and B. D. Smith, *Chem. Commun.*, 2009, 2517; (b) J. M. Baumes, I. Murgu, A. Oliver and B. D. Smith, *Org. Lett.*, 2010, **12**, 4980.

10 C. P. Brock, W. B. Schweizer and J. D. Dunitz, *J. Am. Chem. Soc.*, 1991, **113**, 9811.

11 For a review see: J. Jacques, A. Collet and S. H. Wilen, *Enantiomers, racemates and resolutions*, Krieger, Malabar, 1994.

12 (a) Y. Ferrand, A. M. Kendhale, B. Kauffmann, A. Grélard, C. Marie, V. Blot, M. Pipelier, D. Dubreuil and I. Huc, *J. Am. Chem. Soc.*, 2010, **132**, 7858; (b) G. Lautrette, B. Kauffmann, Y. Ferrand, C. Aube, N. Chandramouli, D. Dubreuil and I. Huc, *Angew. Chem., Int. Ed.*, 2013, **52**, 11517; (c) M. Kudo, V. Maurizot, B. Kauffmann, A. Tanatani and I. Huc, *J. Am. Chem. Soc.*, 2013, **135**, 9628; (d) M. Kudo, V. Maurizot, H. Masu, A. Tanatani and I. Huc, *Chem. Commun.*, 2014, **50**, 10090; (e) C. Dolain, H. Jiang, J.-M. Léger, P. Guionneau and I. Huc, *J. Am. Chem. Soc.*, 2005, **127**, 12943.

13 (a) M. Lee, J. Shim, P. Kang, I. A. Guzei and S. H. Choi, *Angew. Chem., Int. Ed.*, 2013, **52**, 12564; (b) J.-H. Eom, R. Jeong, J. Gong, R. W. Driver and H.-S. Lee, *Bull. Korean Chem. Soc.*, 2015, **36**, 2583.

14 (a) C. Toniolo, C. Peggion, M. Crisma, F. Formaggio, X. Shui and D. S. Eggleston, *Nat. Struct. Biol.*, 1994, **1**, 908; (b) W. R. Patterson, D. H. Anderson, W. F. DeGrado, D. Cascio and D. Eisenberg, *Protein Sci.*, 1999, **8**, 1410; (c) Z. Hayouka, N. C. Thomas, D. E. Mortenson, K. A. Satyshur, B. Weisblum, K. T. Forest and S. H. Gellman, *J. Am. Chem. Soc.*, 2015, **137**, 11884.

15 (a) L. E. Zawadzke and J. M. Berg, *Proteins: Struct., Funct., Genet.*, 1993, **16**, 301; (b) S. W. Wukowitz and T. O. Yeates, *Nat. Struct. Biol.*, 1995, **2**, 1062; (c) B. L. Pentelute, Z. P. Gates, V. Tereshko, J. L. Dashnau, J. M. Vanderkooi, A. Kossiakoff and S. B. H. Kent, *J. Am. Chem. Soc.*, 2008, **130**, 9695; (d) T. O. Yeates and S. B. H. Kent, *Annu. Rev. Biophys.*, 2012, **41**, 41; (e) D. E. Mortenson, K. A. Satyshur, I. A. Guzei, K. T. Forest and S. H. Gellman, *J. Am. Chem. Soc.*, 2012, **134**, 2473.

16 (a) M. Doi, M. Inoue, K. Tomoo, T. Ishida, T. Ueda, M. Akagi and H. Urata, *J. Am. Chem. Soc.*, 1993, **115**, 10432; (b) W. Rypniewski, M. Vallazza, M. Perbandt, S. Klussman, L. J. DeLucas, C. Betzel and V. A. Erdmann, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2006, **62**, 659; (c) P. K. Mandal, G. W. Collie, B. Kauffmann and I. Huc, *Angew. Chem., Int. Ed.*, 2014, **53**, 14424; (d) P. K. Mandal, G. W. Collie, S. C. Srivastava, B. Kauffmann and I. Huc, *Nucleic Acids Res.*, 2016, DOI: 10.1093/nar/gkw367.

17 (a) F. A. Hays, A. T. Teegarden, Z. J. R. Jones, M. Harms, D. Raup, J. Watson, E. Cavaliere and P. S. Ho, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 7157; (b) M. P. H. Lee, G. N. Parkinson, P. Hazel and S. Neidle, *J. Am. Chem. Soc.*, 2007, **129**, 10106; (c) N. Narayana and M. A. Weiss, *J. Mol. Biol.*, 2009, **385**, 469; (d) M. A. Boerneke, S. M. Dibrov, J. Gu, D. L. Wyles and T. Hermann, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15952; (e) P. K. Mandal, S. Venkadesh and N. Gautham, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 2011, **67**, 1506; (f) T. Chatake and T. Sunami, *J. Inorg. Biochem.*, 2013, **124**, 15; (g) M. W. Lau and A. R. Ferré-D'amaré, *Nat. Chem. Biol.*, 2013, **9**, 805.

18 S. E. Cnudde, M. Prorok, Q. Dai, F. J. Castellino and J. H. Geiger, *J. Am. Chem. Soc.*, 2007, **129**, 1586.

19 (a) L. J. Wang, H. Deng, H. Furukawa, F. Gándara, K. E. Cordova, D. Peri and O. M. Yaghi, *Inorg. Chem.*, 2014, **53**, 5881; (b) D. L. Reger, A. Leitner, P. J. Pellechia and M. D. Smith, *Inorg. Chem.*, 2014, **53**, 9932.