Chemical Science

RSC Publishing

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2013, **4**, 1825

Sequential and phototriggered supramolecular self-sorting cascades using hydrogen-bonded motifs†

Maria L. Pellizzaro,[a] Kelly A. Houton[a] and Andrew J. Wilson*[ab]

Received 10th December 2012
Accepted 8th February 2013

DOI: 10.1039/c3sc22194f

www.rsc.org/chemicalscience

A series of hydrogen-bonding motifs are shown to be capable of both high-fidelity and promiscuous molecular recognition behaviour. This gives rise to self-sorting and therefore well defined product distributions for up to four sequential phases of building block composition. Inclusion of a hydrogen-bonding motif that becomes capable of molecular recognition only upon photo-cleavage, extends the number of phases in the cascade to five. This supramolecular system thus reproduces multiple features of biological signalling cascades including the ability to switch between successive states comprising multiple well-defined complexes and triggered modification of molecular recognition preferences.

Introduction

A hallmark of nature is the ability to assemble multiple functional entities in defined locations and at specific times. Such assemblies modulate complex adaptive networks that control biological processes *e.g.* intracellular signalling cascades.[1] An emerging theme in supramolecular chemistry is to mimic aspects of this astonishing complexity and control; notable recent achievements include the identification of molecular recognition motifs[2] and assemblies[3,4] capable of self-sorting,[5] the development of adaptive supramolecular architectures that change structure in response to chemical[6] or physical stimuli[7] and the realisation of kinetically controlled self-sorting[8] and self-assembly.[9] Alongside this, the emergence of "*systems chemistry*"[10–12] from pioneering work on dynamic combinatorial chemistry[13] and studies on molecular amplification[14] has stimulated widespread interest in more closely mimicking natural self-assembly processes.

In a generalised signalling cascade, multiple molecular recognition events happen *simultaneously* and *sequentially*, requiring individual components to interact selectively with one partner at one point during a cascade and then with another partner at a different point in the process. For these events to occur *simultaneously*, building blocks that are capable of orthogonal molecular recognition[15] (*i.e.* self-sorting) are required whereas for a process to occur *sequentially*, building blocks need to be promiscuous in their molecular recognition behaviour (Fig. 1). In biology these two opposing needs are met by compartmentalisation (or more broadly phase separation), expression levels of individual components and triggering events (*e.g.* phosphorylation/dephosphorylation) with a signal ultimately emerging from an ensemble of partially bound states. In considering a supramolecular signalling cascade, several consecutive narcissistic[16] and/or social[17] non-integrative self-sorting[5] molecular recognition events are needed to generate successive states comprising multiple well-defined complexes/assemblies.[18] Although cascades[19,20] exploiting integrative[20,21] self-sorting are known they usually involve transitions between single rather than, multiple assemblies. The current manuscript describes a supramolecular ensemble capable of reproducing this latter property and therefore several key features of a biological signalling cascade. Such behaviour has been demonstrated using cucurbit[*n*]uril and cyclodextrin based hosts[22] whilst parallel preparation of and switching between different metal complexes comprising multiple components is established.[23] These prior examples exploit shape complementarity and multiple non-covalent interactions whereas this manuscript makes use of only linear arrays of hydrogen-bonds for which high-fidelity orthogonal recognition[24] and therefore sorting[25,26] is known to be challenging. We

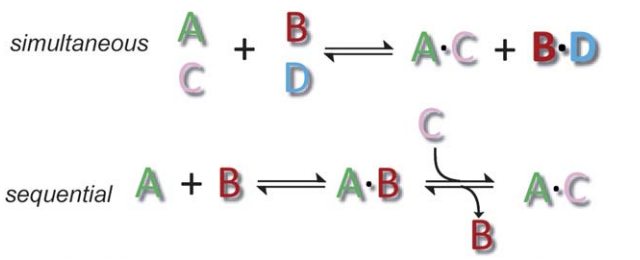

Fig. 1 Representation of key molecular recognition process required for signalling.

[a]*School of Chemistry, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK. E-mail: A.J.Wilson@leeds.ac.uk; Fax: +44(0)113 3436565; Tel: +44(0)113 3431409*

[b]*Astbury Centre for Structural Molecular Biology, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK*

† Electronic supplementary information (ESI) available: Synthetic procedures, characterisation, details of binding studies and titration curves, methods used to determine fidelity and additional speciation/fidelity plots. See DOI: 10.1039/c3sc22194f

firstly describe a cascade of self-sorting events involving hydrogen-bonding motifs (linear arrays of hydrogen-bonds)[27] where the supramolecular products obtained are dependent upon the presence or absence of individual members. We then illustrate that such self-sorting events can be triggered by UV activation of a masked hydrogen-bonding motif which is present throughout the self-sorting experiment.

Results and discussion

Our initial objective of this study was to identify hydrogen-bonding motifs such as ureidoimidazole **1** (UIM) and amido-isocytosine (AIC) **2** (Fig. 2),[28–30] capable of orthogonal recognition for the coded assembly of polymers.[25,31] Linear arrays participate in high-affinity molecular recognition through a sequence of donor and acceptor atoms. They have served as a powerful tool with which to study and understand co-operative hydrogen-bond mediated molecular recognition.[27,32–34]

Of relevance to this work, Zimmerman and co-workers have performed a series of studies that have focused on the development of linear arrays capable of high-fidelity molecular recognition (defined as the ratio of the concentration of desired complexes to total concentration of complexes).[35] Notably, the ureidoguanosine·diamidonaphthyridine (UG·DAN) complex[36] forms with high affinity but minimal interference from competitor motifs.[24] In contrast, the ureidopyrimidinone (UPy) motif *e.g.* **3** (ref. 37) has poor fidelity molecular recognition because it forms strong self-complementary homodimers and heterodimers with DAN motifs *e.g.* **4**.[38] This was further illustrated by Li and co-workers, who observed low fidelity in self-sorting experiments using hydrazide, UPy and DAN motifs.[26] Our efforts to elaborate orthogonal hydrogen-bonding motifs employed a number of different strategies. Of our designs only the triply hydrogen-bonded heterodimers *e.g.* UIM·AIC **1·2** that we reported previously[28–30] gave well-defined and strongly bonded complexes in chloroform. We therefore explored if UIM·AIC **1·2** would assemble in the presence of UPy·DAN **3·4** (Fig. 3).

Pleasingly, the NMR spectrum of an equimolar mixture of these four components indicated the simultaneous formation of both heterodimers as the predominant products (Fig. 3a–e) although other unidentified complexes are observed in small quantities at higher concentrations (Fig. 3c – red boxes). Notably, the NMR spectrum did not change from 10–1 mM indicating the fidelity of the system (Fig. 3c–e) was respectable over a range of concentrations. This was supported by using the known binding constants (Table 1) for all the dimers to calculate the speciation (Fig. 3f) and resultant fidelity (Fig. 3g), for an equimolar mixture of the components over a range of concentrations. These calculations indicate a fidelity of ~0.7 above 10 mM.

Given the results described above and the known promiscuous binding behaviour of UPy **3**, we saw the self-sorting of UIM·AIC **1·2** and UPy·DAN **3·4** as the basis upon which to develop a synthetic artificial signalling cascade. The first step in this process was to determine pertinent binding constants for all the components; in earlier work, we had previously illustrated that AIC **2** forms a relatively weak interaction with DAN **4**,[39] whilst the remaining binding constants were taken from the literature or determined by ^{1}H NMR titration. The binding constants for all possible combinations of **1–4** as homo and heterodimers were measured (Table 1).

We then performed a series of experiments to exemplify a molecular recognition cascade (Fig. 4a – *Path I*). Starting with the homodimer of UPy **3**, addition of UIM **1** disrupts assembly of the high affinity heterodimer to form a new and distinctive species in the ^{1}H NMR spectrum proposed to be **1·3** (Fig. 4b and c). Although the UPy homodimer **3·3** is formed with higher K_{dim}, UIM·UPy **1·3** heterodimerisation is favoured because the overall number of hydrogen bonds in the system is greater. Further addition of AIC **2** to the mixture releases UPy **3** from its complex with **1** to generate UPy·UPy **3·3** and UIM·AIC **1·2**

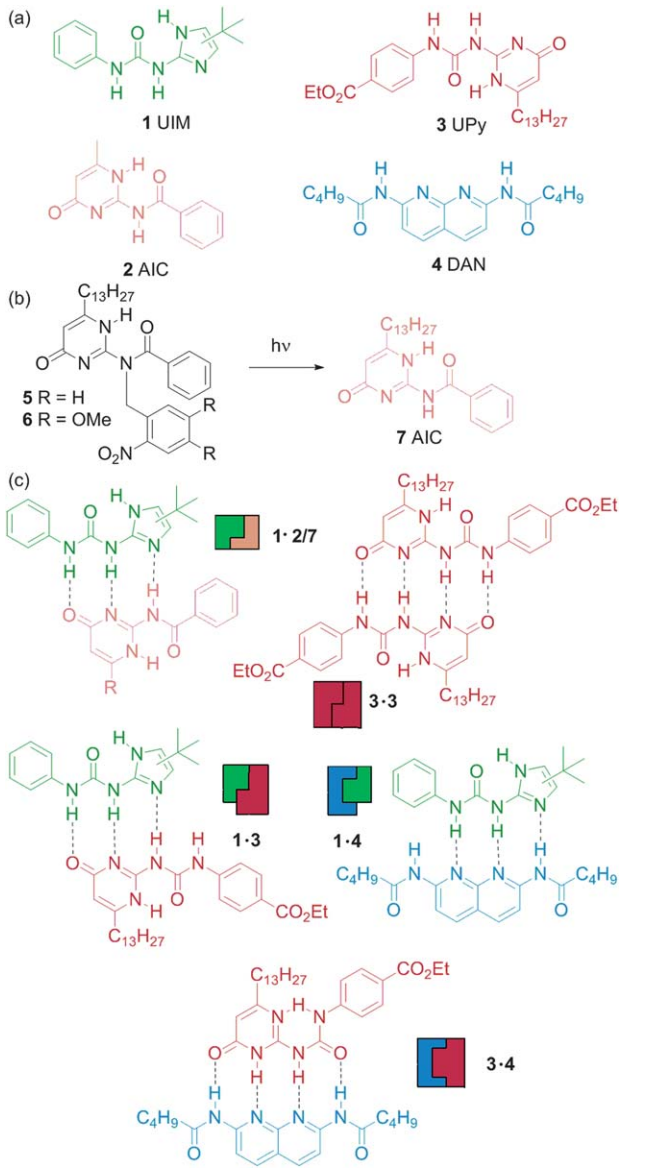

Fig. 2 Self-sorting hydrogen-bonding motifs (a) structures of compounds **1–4**, (b) the photoinduced conversion of **5**/**6** to **7** and (c) proposed structures of hydrogen-bonded dimers.

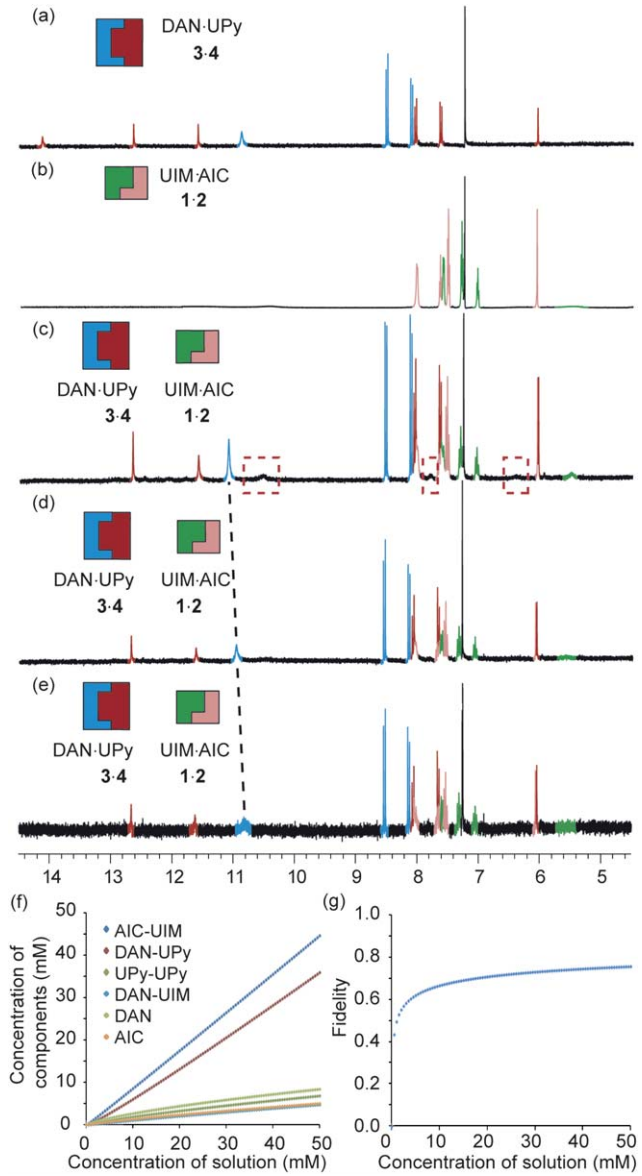

Fig. 3 Self-sorting experiments with compounds **1–4**, (a–e) 300 MHz ^{1}H NMR spectra ($CDCl_3$, 293 K) (a) complex formed between UPy **3**·DAN **4** at 10 mM, (b) complex formed between UIM **1**·AIC **2** at 10 mM, (c) complexes formed between UIM **1**·AIC **2**, and UPy **3**·DAN **4** at 10 mM (dashed red boxes denote "undesired" complexes) (d) UIM **1**·AIC **2**, and UPy **3**·DAN **4** at 5 mM (e) UIM **1**·AIC **2**, and UPy **3**·DAN **4** at 1 mM.

Table 1 Binding constants determined by ^{1}H NMR titration ($CDCl_3$) for all combinations of hydrogen-bonding motifs **1–4** studied in this work

Complex	K_a or K_{dim} (M^{-1})
UIM·UIM **1·1** (ref. 30)	10 ± 2
AIC·UIM **1·2** (ref. 28)	$(3.3 \pm 1.6) \times 10^4$
UIM·UPy **1·3**	$(6.8 \pm 1.0) \times 10^4$
UIM·DAN **1·4** (ref. 39)	$(2.1 \pm 0.1) \times 10^3$
AIC·AIC **2·2** (ref. 30)	3.6 ± 0.3
AIC·UPy **2·3**	<1
AIC·DAN **2·4**	<1
UPy·UPy **3·3** (ref. 37)	6×10^7
UPy·DAN **3·4** (ref. 40)	5×10^6
DAN·DAN **4·4** (ref. 41)	3.6 ± 0.3

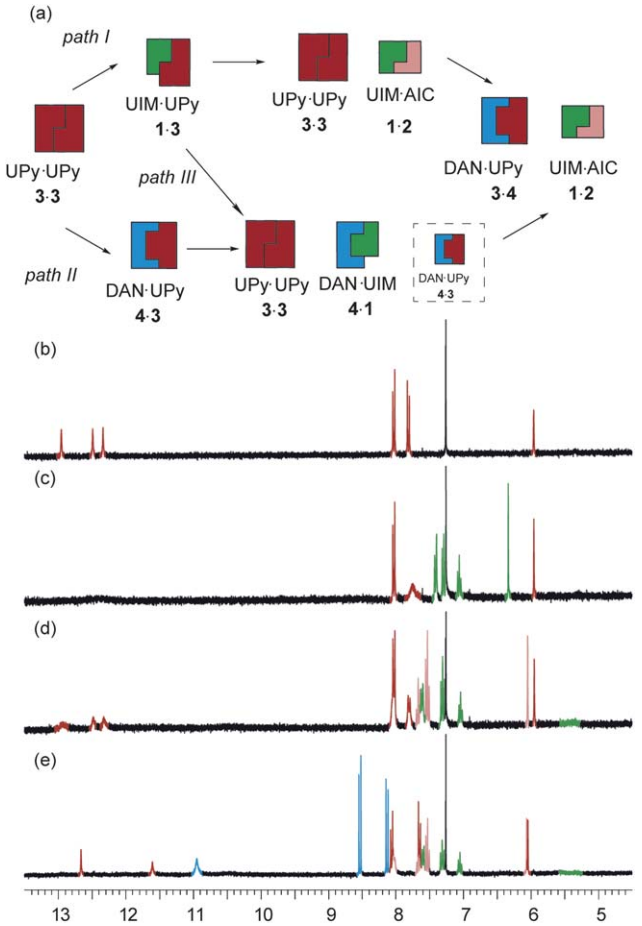

Fig. 4 Signaling cascade using hydrogen-bonding motifs **1–4** (a) a schematic depicting complexes that form upon addition of different components to the supramolecular ensemble (b–e) 300 MHz NMR spectra (10 mM, $CDCl_3$, 293 K) of the signaling cascade following *path I* (b) UPy **3** (c) UPy **3** and UIM **1**, (d) UPy **3**, UIM **1** and AIC **2**, (e) UPy **3**, UIM **1**, AIC **2** and DAN **4**.

(Fig. 4d). Finally, addition of DAN **4** changes the product distribution again – although a plethora of alternative partners are available to DAN **4** at this point it sequesters UPy **3** to form UPy·DAN **3·4** and does not interfere with **1·2** (Fig. 4e). This is noteworthy given that the binding constants for UIM·AIC **1·2** and UIM·DAN **1·4** are comparable whilst UPy **3** dimerisation is the strongest isolated interaction in the entire system. It appears that maximising non-covalent interactions in the entire system drives the product distribution and the individual affinities are unimportant.

The speciation and fidelity were calculated at each stage during the process for a starting concentration of 10 mM (see ESI†). The calculations suggest quite poor fidelity for the first step of the process ~0.25, but much improved fidelity for step two (~0.9) and step three (~0.7). These results are instructive; the ^{1}H NMR spectrum from the first stage (Fig. 4c) is well-defined, indicating that even for partially bound states, well-defined "outputs" can be generated as is the case in cell-signalling. In addition it should be noted that whilst the system requires components capable of promiscuous behaviour, it benefits from at least one member (in this case AIC **2**)

recognising a limited number of other components within the system. Indeed, fidelity is improved in its presence (see ESI†). Significantly, when the order of addition of hydrogen-bonding motif is changed, different intermediate complexes are 'expressed' (Fig. 4a – *Path II* and *III*, see also ESI†). Where DAN **4** is added prior to AIC **2**, a new complex **1**·**4** is found in the system that is not accessible following *Path I*. Furthermore, the presence of this species illustrates how addition of components to a system can indirectly affect the molecular recognition of other components, hence addition of AIC **2** in *Path III* induces a change in the binding behaviour of both UPy **3** and DAN **4** without competing for the former. As for *Path I*, the speciation and fidelity plots in *Path I* and *II* reveal one phase in the cascade that is of lower fidelity (see ESI†).

We finally considered methods by which it might be possible to trigger a change in product distribution using a chemical reaction or stimulus. Such a triggering event bears some analogy to removal of a functional group (*e.g.* dephosphorylation). The modulation of the molecular recognition properties of hydrogen-bonding motifs in a reversible manner is not well established; to our knowledge, only a few examples have been described.[42,43] We therefore used an irreversible reaction – the light induced cleavage of an *o*-nitrobenzyl group[44] to illustrate the proof of concept. Our efforts to attach a simple *o*-nitrobenzyl group to the oxygen of 2-hydroxy-4-methyl-6-aminopyrimidine resulted instead in alkylation of the 6-amino group (to be discussed elsewhere) which upon acylation gave compound **5** (see ESI†). In our hands light activated cleavage of the *o*-nitrobenzyl amide was unsuccessful returning only starting material **5**. We therefore turned to a more labile *o*-nitro-4,5-dimethoxybenzyl group, obtaining **6** *via* the same route used for **5** (see ESI†). Irradiation of **6** (standard laboratory TLC lamp operating at 365 nm) proceeded cleanly in around 2–3 h giving the unmasked AIC **7**.

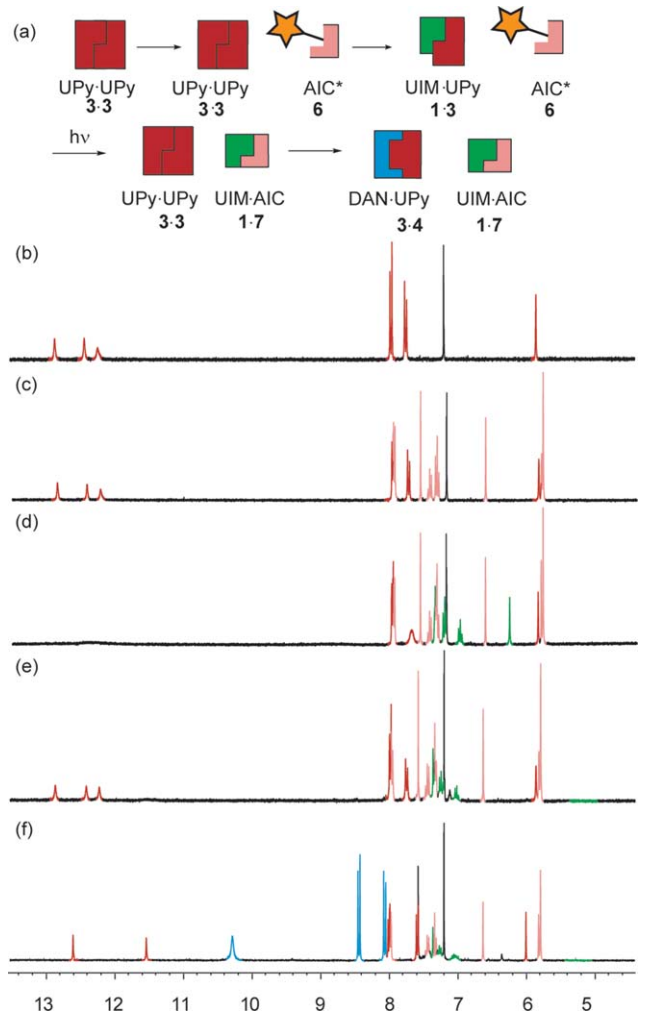

Fig. 5 Signalling cascade using hydrogen-bonding motifs **1**, **3**, **4** and **6**/**7** (a) a schematic depicting complexes that form upon addition or excitation of different components to the supramolecular ensemble (b–f) 300 MHz NMR spectra (10 mM, $CDCl_3$, 293 K) of the signalling cascade depicted in (a). (b) UPy **3** (c) UPy **3** and AIC **6**, (d) UPy **3**, AIC **6** and UIM **1**, (e) UPy **3**, AIC **6**, UIM **1** resulting from 4 h irradiation at 365 nm (f) UPy **3**, AIC **6**, UIM **1** and DAN **4**.

As illustrated in Fig. 5, incorporation of this reaction within the process works smoothly. Addition of AIC **6** to UPy **3** causes no change to its molecular recognition behavior – this is not unsurprising as AIC **2** without an alkyl group on the amide was shown not to bind to UPy **3**. Addition of UIM **1** to the mixture results in disassembly of the UPy **3** homodimer and formation of a heterodimer **1** : **3**, confirming the inability of AIC **6** to participate in hydrogen-bonding (note that at this stage its preferred binding partner is present). After irradiation with UV light (standard laboratory TLC lamp operating at 365 nm), the ^{1}H NMR spectrum is significantly different – the unmasked AIC **7** is now free to interact with UIM **1** and UPy **3** is released to reform its homodimer. Finally addition of DAN **4** completes the cycle.

Conclusions

In conclusion, our results demonstrate that simple hydrogen-bonding motifs can be used to construct complex supramolecular systems that self-sort in a sequential manner depending upon the composition of the system and that it is possible to do this in a light dependent manner. Such a property reproduces multiple aspects of complex biological signalling processes. Our future work will focus on developing supramolecular synthons whose molecular recognition can be reversibly controlled so as to facilitate temporal control over product distribution, and, on the inclusion (within the cascade) of multi-component supramolecular complexes that are amenable to allosteric manipulation of product distribution.

Notes and references

1 J. Nelson, *Structure and Function in Cell Signalling*, Wiley, 2009.
2 A. Wu and L. Isaacs, *J. Am. Chem. Soc.*, 2003, **125**, 4831–4835.
3 R. Kramer, J. M. Lehn and A. Marquis-Rigault, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 5394–5398.
4 B. H. Northrop, Y.-R. Zheng, K.-W. Chi and P. J. Stang, *Acc. Chem. Res.*, 2009, **42**, 1554–1563.
5 M. M. Safont-Sempere, G. Fernández and F. Würthner, *Chem. Rev.*, 2011, **111**, 5784–5814.

6 S. Liu, P. Y. Zavalij, Y.-F. Lam and L. Isaacs, *J. Am. Chem. Soc.*, 2007, **129**, 11232–11241.
7 P. J. Lusby, P. Müller, S. J. Pike and A. M. Z. Slawin, *J. Am. Chem. Soc.*, 2009, **131**, 16398–16400.
8 P. Mukhopadhyay, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2006, **128**, 14093–14102.
9 O. Chepelin, J. Ujma, P. E. Barran and P. J. Lusby, *Angew. Chem.*, 2012, **124**, 4270–4273.
10 R. F. Ludlow and S. Otto, *Chem. Soc. Rev.*, 2008, **37**, 101–108.
11 D. Jiao, F. Biedermann, F. Tian and O. A. Scherman, *J. Am. Chem. Soc.*, 2010, **132**, 15734–15743.
12 Z. Qi, P. Malo de Molina, W. Jiang, Q. Wang, K. Nowosinski, A. Schulz, M. Gradzielski and C. A. Schalley, *Chem. Sci.*, 2012, **3**, 2073–2082.
13 P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders and S. Otto, *Chem. Rev.*, 2006, **106**, 3652–3711.
14 A. Vidonne and D. Philp, *Eur. J. Org. Chem.*, 2009, 593–610.
15 C.-H. Wong and S. C. Zimmerman, *Chem. Commun.*, 2013, **49**, 1679–1695.
16 P. N. Taylor and H. L. Anderson, *J. Am. Chem. Soc.*, 1999, **121**, 11538–11545.
17 A. Shivanyuk and J. Rebek, *J. Am. Chem. Soc.*, 2002, **124**, 12074–12075.
18 V. E. Campbell, X. de Hatten, N. Delsuc, B. Kauffmann, I. Huc and J. R. Nitschke, *Nat. Chem.*, 2010, **2**, 684–687.
19 I. A. Riddell, M. M. J. Smulders, J. K. Clegg, Y. R. Hristova, B. Breiner, J. D. Thoburn and J. R. Nitschke, *Nat. Chem.*, 2012, **4**, 860.
20 M. M. J. Smulders, A. Jiménez and J. R. Nitschke, *Angew. Chem., Int. Ed.*, 2012, **51**, 6681–6685.
21 W. Jiang, H. D. F. Winkler and C. A. Schalley, *J. Am. Chem. Soc.*, 2008, **130**, 13852–13853.
22 S. Ghosh, P. Mukhopadhyay and L. Isaacs, *J. Syst. Chem.*, 2010, **1**, 6.
23 J. Nitschke, *Acc. Chem. Res.*, 2007, **40**, 103–112.
24 T. Park, E. M. Todd, S. Nakashima and S. C. Zimmerman, *J. Am. Chem. Soc.*, 2005, **127**, 18133–18142.
25 S. K. Yang, A. V. Ambade and M. Weck, *J. Am. Chem. Soc.*, 2010, **132**, 1637–1645.
26 X. Zhao, X. Z. Wang, X. K. Jiang, Y. Q. Chen, Z. T. Li and G. J. Chen, *J. Am. Chem. Soc.*, 2003, **125**, 15128–15139.
27 A. J. Wilson, *Soft Matter*, 2007, **3**, 409–425.
28 M. L. Pellizzaro, A. M. McGhee, L. C. Renton, M. G. Nix, J. Fisher, W. B. Turnbull and A. J. Wilson, *Chem.–Eur. J.*, 2011, **17**, 14508–14517.
29 A. Gooch, A. M. McGhee, M. L. Pellizzaro, C. I. Lindsay and A. J. Wilson, *Org. Lett.*, 2011, **13**, 240–243.
30 A. M. McGhee, C. Kilner and A. J. Wilson, *Chem. Commun.*, 2008, 344–346.
31 S.-L. Li, T. Xiao, C. Lin and L. Wang, *Chem. Soc. Rev.*, 2012, **41**, 5950–5968.
32 S. C. Zimmerman and P. S. Corbin, *Struct. Bonding*, 2000, **96**, 63–94.
33 H.-B. Wang, B. P. Mudraboyina and J. A. Wisner, *Chem.–Eur. J.*, 2012, **18**, 1322–1327.
34 B. A. Blight, C. A. Hunter, D. A. Leigh, H. McNab and P. I. T. Thomson, *Nat. Chem.*, 2011, **3**, 246–250.
35 E. M. Todd, J. R. Quinn, T. Park and S. C. Zimmerman, *Isr. J. Chem.*, 2005, **45**, 381–389.
36 T. Park, S. C. Zimmerman and S. Nakashima, *J. Am. Chem. Soc.*, 2005, **127**, 6520–6521.
37 F. H. Beijer, R. P. Sijbesma, H. Kooijman, A. L. Spek and E. W. Meijer, *J. Am. Chem. Soc.*, 1998, **120**, 6761–6769.
38 T. F. A. de Greef, G. Ercolani, G. B. W. L. Ligthart, E. W. Meijer and R. P. Sijbesma, *J. Am. Chem. Soc.*, 2008, **130**, 13755–13764.
39 M. L. Pellizzaro, S. A. Barrett, J. Fisher and A. J. Wilson, *Org. Biomol. Chem.*, 2012, **10**, 4899–4906.
40 G. B. W. L. Ligthart, H. Ohkawa, R. P. Sijbesma and E. W. Meijer, *J. Am. Chem. Soc.*, 2005, **127**, 810–811.
41 P. S. Corbin, S. C. Zimmerman, P. A. Thiessen, N. A. Hawryluk and T. J. Murray, *J. Am. Chem. Soc.*, 2001, **123**, 10475–10488.
42 M. Herder, M. Patzel, L. Grubert and S. Hecht, *Chem. Commun.*, 2011, **47**, 460–462.
43 Y. Li, T. Park, J. K. Quansah and S. C. Zimmerman, *J. Am. Chem. Soc.*, 2011, **133**, 17118–17121.
44 C. G. Bochet, *J. Chem. Soc., Perkin Trans. 1*, 2002, 125–142.