Templated amplification of a naphthalenediimide-based receptor from a donor–acceptor dynamic combinatorial library in water†‡

Ho Yu Au-Yeung, Paolo Pengo, G. Dan Pantoş, Sijbren Otto and Jeremy K. M. Sanders*

Received (in Cambridge, UK) 29th September 2008, Accepted 21st November 2008
First published as an Advance Article on the web 9th December 2008
DOI: 10.1039/b816979a

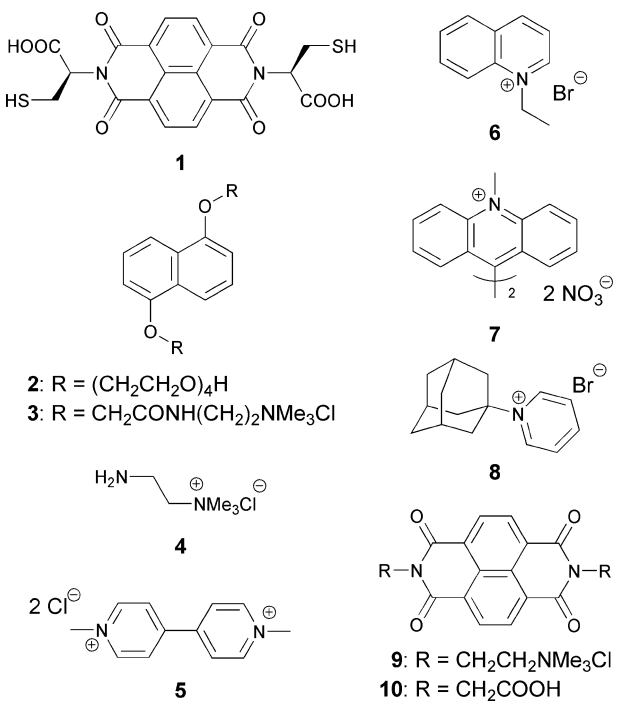
We report a dynamic combinatorial library that, upon binding of an electronically-complementary guest, produces in high yield a tetrameric receptor with flat hydrophobic, electron-deficient surfaces and flexible, water-soluble disulfide-containing linkers; analysis of the dependence of library composition on template concentration gives insight into the binding behaviours of the species involved.

Achieving molecular recognition in water using synthetic receptors remains a major challenge. We report here that the cysteine functionalised naphthalenediimide (NDI) building block, **1** (Scheme 1), is a suitable candidate for use in aqueous dynamic combinatorial libraries[1] (DCLs) and that it can be incorporated efficiently into a large receptor upon binding an electronically-complementary guest. The amino acid moiety in **1** provides both water solubility and access to reversible covalent chemistry,[2] while the NDI core was chosen for its ability to interact with electron-rich aromatic systems.

Interactions between electron-rich donors and electron-poor acceptors tend to have a well-defined geometry due to a combination of van der Waals, local electrostatic, charge transfer and solvophobic effects; for convenience we will call the sum of these effects, donor–acceptor (D–A) interactions. These interactions have been successfully used in directing the formation of supramolecular structures such as pleated aedamers,[3] and topologically complex rotaxanes and catenanes.[4] Previously, we have reported a dynamic combinatorial synthesis of a macrocycle containing a disulfide-linked porphyrin and an electron donor in $CHCl_3$, and have shown that it is stabilised by an NDI guest.[5] Here we present initial results on a DCL of **1** in water.

Dithiol building block **1** was obtained by acid deprotection of the previously reported trityl precursor.[6] DCLs were set up by air oxidation of a 5 mM solution of **1** in water at pH 8.5. The DCL was analysed by reverse phase HPLC/LC-MS after 5 days, when the equilibrium had been reached. At equilibrium, the DCL was found to contain dimeric, trimeric and tetrameric macrocycles in a distribution of approx. 50%, 35% and 15% of library material, respectively (Fig. 1). Two separate peaks are observed (retention time = 19.8 and 23.3 min) with *m*/*z* corresponding to the cyclic trimer. Their similar isotopic pattern and MS/MS fragmentation of the parent molecular ion peaks suggests the existence of two conformers that interconvert slowly.[7]

To probe the effect of electron donors,[8] a new set of DCLs was prepared in the presence of dialkoxynaphthalenes (DN) **2** and **3**. A colour change from yellow to plum (for **2**) or orange (for **3**) was observed, suggesting the presence of D–A interactions between NDI and DN moieties, and indeed, a shift in DCL composition was observed (Fig. 1). For the neutral DN **2**, an increase in total *trimer* concentration was found (amplification factor = 1.4). More pronounced amplification was observed for the cationic DN **3**: around 80% of the library material was found as the *tetramer*, corresponding to an amplification factor of 5.7. The amplification of the higher oligomers occurs at the expense of the *dimer*, whose cavity is probably too small to accommodate the templates. The continuing presence of the trimer (both conformers) and the tetramer in DCLs templated by **3** and **2**, respectively, indicates that these macrocycles interact with the templates, although they are not amplified.



Scheme 1 Dithiol building block **1**, templates **2–9** and model NDI **10**.

University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge, UK CB2 1EW. E-mail: jkms@cam.ac.uk; Fax: +44 (0)1223 336 017

† Electronic supplementary information (ESI) available: Synthetic and DCLs set up methods, *DCLFit* calculations, NMR and UV-Vis data. CCDC 703957 and 703958. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/b816979a

‡ Dedicated to Roeland Nolte on his 65th birthday, with warm thanks for his many contributions to chemistry, his leadership of Chemical Communications, and his friendship.

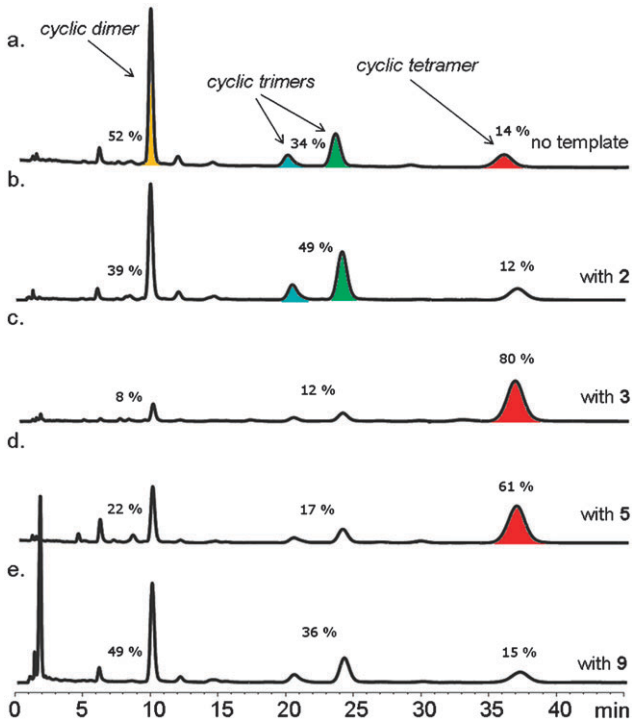


Fig. 1 HPLC traces of the DCLs of 5 mM of **1** (a) without template, and with (b) **2**, (c) **3**, (d) **5** and (e) **9**. All templates were 2.5 mM. The library material distribution is represented as percentages above the corresponding peaks. The amplified species are highlighted.

Templates **2** and **3** both contain a DN unit, yet they elicit different responses from the library, indicating that molecular recognition in water involves more than just D–A interactions. Indeed, the modest amplification induced by **2** indicates that the D–A interactions are relatively weak. Also, the cationic end groups of **3** produced no amplification when used in isolation, *i.e.* the library composition in the presence of **4** is virtually unchanged from the untemplated distribution (see ESI†).

To evaluate the importance of the electron-rich DN function in the recognition, electron-deficient templates **5–9** were tested in DCLs of **1**. Compounds **5**, **6** and **7** amplify the tetramer, though the amplification factors are significantly lower than that observed with **3** (4.4, 2.1, and 2.1, respectively *vs.* 5.7). In contrast, the library distributions with **8** and **9** were very similar to that of the untemplated DCL (Fig. 1 and ESI†). In all the DCLs, regardless of the template used and the amplified species, the amount of the minor trimer is around one-third of the major trimer, with their ratio ranging from 0.28 to 0.37. This suggests that both conformers have similar binding abilities towards all the studied templates.

Based on these experiments, we conclude that the D–A interactions are necessary but not sufficient[9] to induce a large response in DCLs of **1**. Their necessity is supported by the absence of significant tetramer amplification observed in DCLs containing electron-deficient templates **5–9**. However, the DCL response to **2** supports the conclusion that D–A interactions alone are not sufficient to generate amplification in these systems. Electrostatic or hydrophobic effects can also be ruled out as being independently responsible for amplification as indicated by the lack of influence of templates **4**, **8** and **9** on DCLs of **1**. Based on these observations we conclude that the combination of D–A and electrostatic interactions leads to large amplification of the tetramer by **3**, however other factors such as the relative size of the host and the guest are also important.

The DCL templated by **3** was reproduced on a preparative scale, and the tetramer was isolated for ^{1}H NMR and UV-Vis analyses. Complex ^{1}H spectra dominated by broad signals were observed (223–330 K, 500 MHz, CD_3OD), suggesting a high flexibility and the presence of multiple conformations. Addition of **3** to the tetramer solution resulted in the formation of an orange precipitate. Changing the solvent to D_2O–NaOD, where the complex is soluble, again resulted in a ^{1}H spectrum consisting of broad signals (293 K, 500 MHz). An upfield shift of the aromatic protons of the guest **3** by 0.65–0.75 ppm was observed, presumably due to proximity of the aromatic moieties of the host and the guest.

Due to the complexity of the NMR spectra we turned our attention to UV-Vis studies, which proved to be also inconclusive. Mole-ratio plots constructed from the titration of **3** to an aqueous solution (pH 8.5) of $(\mathbf{1})_4$ at two different concentrations (1.1×10^{-5} M and 3.6×10^{-4} M) indicated the presence of a 1 : 2 host–guest complex. Attempts to determine precise association constants at these concentrations failed, possibly because of the complex electronic transitions involving the host and the guest (see ESI†).

To gain more information on the complex equilibria of the system, a series of DCLs at varying concentrations of **3** was used as starting points for HPLC peak-area fitting using *DCLFit*[10] (see ESI†). The simplest model that is able to produce a good fit requires the presence of a series of template-bound library members (Fig. 2). Removing any of the species from the model leads to a poorer fit.[11] The strong amplification of the tetramer by **3** is reflected in the high affinity of the host to the guest, with fitted binding constants of around 10^6 M^{-1} and 10^4 M^{-1} upon binding of the first and the second guest, respectively. In the fitting model, all the association constants are for 1 : 1 complexes, using step-wise binding.

To put the *DCLFit*-determined binding constants in perspective, a UV-Vis titration of guest **3** to NDI model compound **10** was performed in water at pH 8.5. Both the X-ray structure of the co-crystal (Fig. 3) and the Job plot indicate a 1 : 1 binding stoichiometry. Fitting of the UV-Vis titration data to this binding model resulted in a binding constant of 7700 ± 110 M^{-1}, which is in the same order of magnitude as that of the binding of **3** to the NDI dimer. Since only outside binding of **3** to the dimer is anticipated, this binding event is comparable, but not identical, to the binding of **3** to the glycine-derived monomer **10**. Control experiments revealed, as expected, a smaller association constant between **2** and **10**, of around 720 ± 50 M^{-1} (see ESI†).

The relative values of the binding constants obtained from *DCLFit* shed light on the binding modes of **3** to the macrocycles. The association constant for the first binding of **3** to the dimer is comparable with that obtained from the model system. As expected on statistical grounds, the second binding is significantly weaker. On the other hand, the first association constant of **3** to the tetramer (10^6 M^{-1}) is exceedingly high. This value is supported by the interaction of $(\mathbf{1})_4$ with **3** at 1.1×10^{-5} M

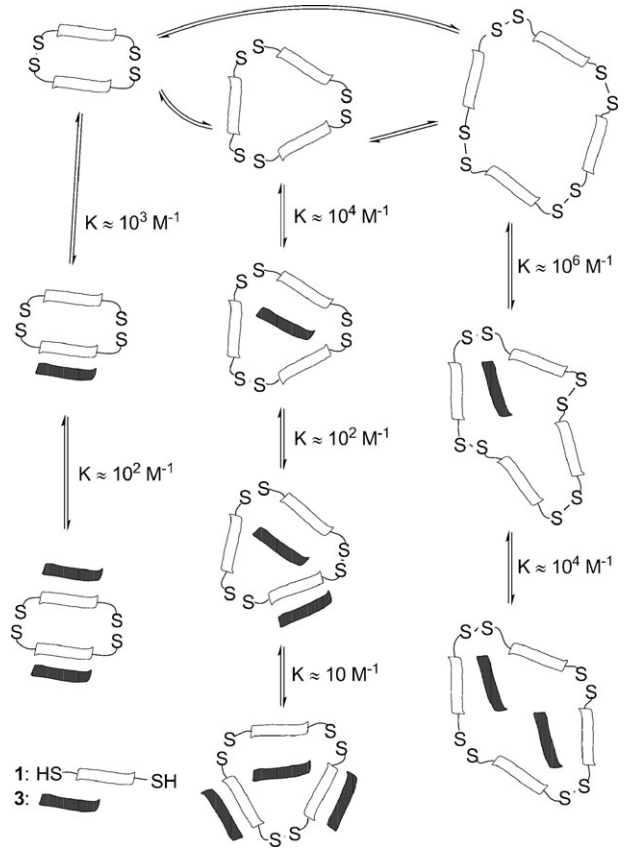


Fig. 2 Model of DCL of **1** in the presence of **3** constructed by *DCLFit* and the respective binding constants obtained from the fitting.

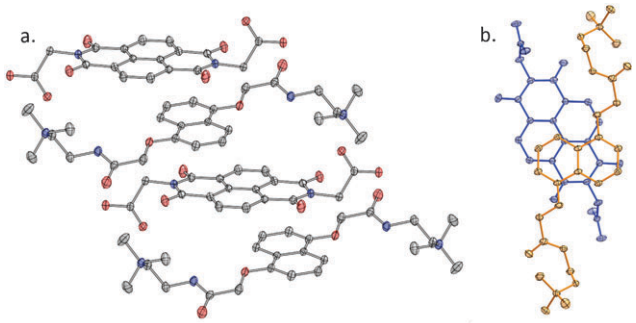


Fig. 3 Side and top views of the X-ray structure of **3**·**10** co-crystal. Ellipsoids are at 50% probability.

observed in UV-Vis studies. This indicates that in the first recognition process, there are at least two NDI components of the tetramer interacting cooperatively with DN **3**. Clearly the flexibility of the macrocycle enables it to adjust its conformation to best suit the template.[2a] The second binding event is characterised by a K_a of 10^4 M^{-1}, implying that the tetramer has the cavity suitably preorganised for an interaction with another DN guest. Similarly, the association constant of 10^4 M^{-1} found for the first binding of **3** to the trimer is indicative of a recognition event involving the DN guest and at least two NDI moieties. The reduction in the orders of magnitudes of the association constants determined for the second and third binding events are sufficient to conclude that the trimer cavity is large enough to fit only one guest.

The fine interplay of the rigid NDI core with the flexible cysteine linkers allows the highly efficient dynamic combinatorial synthesis of receptor molecules that are able to fold in the correct geometry for the recognition of DN guests. The emerging conclusion from these and other observations[2a] from our laboratory is that building blocks that are effective in DCLs combine rigid and flexible subunits and incorporate recognition moieties within the molecular scaffold. *DCLFit* brings insights into otherwise intractable complex equilibria in DCLs. The entropic costs for the high yield synthesis of the tetrameric receptor must be counterbalanced by a very favourable enthalpy of interaction to give the observed large association constant between this receptor and an electron-rich guest molecule. Currently we are investigating the behaviour of libraries containing both donor and acceptor building blocks.

We thank EPSRC, The Royal Society, the Croucher Foundation and Pembroke College for financial support, Dr Fred Ludlow for help with *DCLFit*, Dr Ana Belenguer for maintaining the HPLC facility, and Dr John Davies for solving the X-ray structures.

Notes and references

1 (*a*) S. Ladame, *Org. Biomol. Chem.*, 2008, **6**, 219; (*b*) J.-M. Lehn, *Chem. Soc. Rev.*, 2007, **36**, 151; (*c*) M. M. Rozenman, B. R. McNaughton and D. R. Liu, *Curr. Opin. Chem. Biol.*, 2007, **11**, 259; (*d*) 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; (*e*) B. de Bruin, P. Hauwert and J. N. H. Reek, *Angew. Chem., Int. Ed.*, 2006, **45**, 2660.

2 (*a*) P. T. Corbett, L. H. Tong, J. K. M. Sanders and S. Otto, *J. Am. Chem. Soc.*, 2005, **127**, 8902; (*b*) F. Bulos, S. L. Roberts, R. L. E. Furlan and J. K. M. Sanders, *Chem. Commun.*, 2007, 3092; (*c*) K. R. West, K. D. Bake and S. Otto, *Org. Lett.*, 2005, **7**, 2615.

3 R. S. Lokey and B. L. Iverson, *Nature*, 1995, **375**, 303.

4 (*a*) D. B. Amabilino and J. F. Stoddart, *Chem. Rev.*, 1995, **95**, 2725; (*b*) L. Raehm, D. G. Hamilton and J. K. M. Sanders, *Synlett*, 2002, **11**, 1743; (*c*) O. Š. Miljanić and J. F. Stoddart, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 12966.

5 A. L. Kieran, S. I. Pascu, T. Jarrosson, M. J. Gunter and J. K. M. Sanders, *Chem. Commun.*, 2003, 2674.

6 P. Pengo, G. D. Pantoş, S. Otto and J. K. M. Sanders, *J. Org. Chem.*, 2006, **71**, 7063.

7 Molecular modelling (Hyperchem, v8.0, AM1) supports the existence of at least two trimer conformations similar in energy that interconvert very slowly at room temperature. See ESI†.

8 M. S. Cubberley and B. L. Iverson, *J. Am. Chem. Soc.*, 2001, **123**, 7560.

9 Similar observations that the D–A interaction may not be the primary driving force of association have been made on some catenane systems, for examples see: (*a*) K. N. Houk, S. Menzer, S. P. Newton, F. M. Raymo, J. F. Stoddart and D. J. Williams, *J. Am. Chem. Soc.*, 1999, **23**, 897; (*b*) J. G. Hansen, N. Feeder, D. G. Hamilton, M. J. Gunter, J. Becher and J. K. M. Sanders, *Org. Lett.*, 2000, **2**, 449.

10 R. F. Ludlow, J. Liu, H. Li, S. L. Roberts, J. K. M. Sanders and S. Otto, *Angew. Chem., Int. Ed.*, 2007, **46**, 5762.

11 Reliability of fitted binding constants was probed using a 1000-run "bootstrap" method with varying weights in experimental data input in the fitting. See ESI† for details.