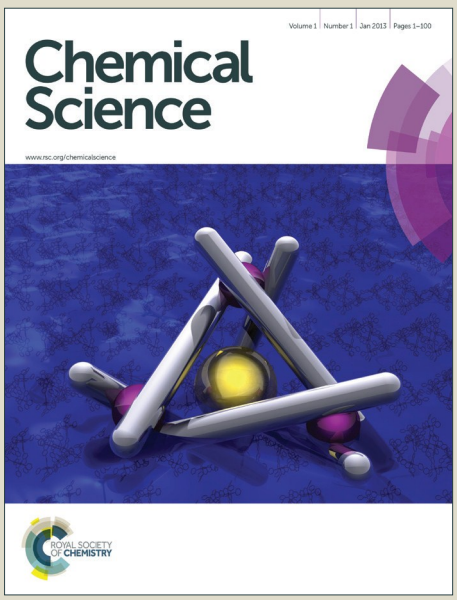
Chemical Science

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemicalscience

Chemical Science

ARTICLE

Design, Preparation, and Selection of DNA-encoded Dynamic Libraries

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Gang Li,[a,‡] Wenlu Zheng,[b,‡] Zitian Chen,[c,‡] Yu Zhou,[b] Yu Liu,[a] Junrui Yang,[c] Yanyi Huang,[c] Xiaoyu Li*[,a,b]

We report a method for the preparation and selection of DNA-encoded dynamic libraries (DEDLs). The library is composed of two sets of DNA-linked small molecules under dynamic exchange through DNA hybridization. Addition of the protein target shifted the equilibrium favouring the assembly of high affinity bivalent binders. Notably, we introduced a novel locking mechanism to stop the dynamic exchange and "freeze" the equilibrium, thereby enabling downstream hit isolation and decoding by PCR amplification and DNA sequencing. Our DEDL approach has circumvented the limitation on library size and realized the analysis and selection of large dynamic libraries. In addition, this method also eliminates the requirement for modified and immobilized target protein.

Introduction

Dynamic combinatorial chemistry (DCC) employs reversible bond formation to create dynamic systems of continuous inter-exchanging chemical entities.[1-4] Built on the principle of DCC, dynamic combinatorial library (DCL) has emerged as an efficient tool for discovering novel ligands against biological targets.[5-8] Compared with static library, DCL has two advantages. First, DCL allows for spontaneous library synthesis based on the inter-conversion of compounds through reversible reactions among building blocks (BBs); the entire library can be synthesized by simply mixing BBs without the need for spatial separation. Second, DCL is adaptive: adding the target imposes the selection pressure to re-distribute BBs favouring the synthesis of target-binding compounds at the expense of non-binding ones.[9-12] Moreover, after reaching a new equilibrium in the presence of the target, the library can be "frozen" by stopping the dynamic exchange (e.g. adding an additive or changing the pH to stop reversible reactions), so that the library population change is preserved and ready for subsequent hit identification.[1, 6] DCL has shown great potential in accelerating the discovery of lead compounds in drug discovery,[5, 6, 13, 14] such as in fragment-based[15-18] and structure-based drug design.[5, 19, 20]

However, DCL is facing a major limitation of low library diversity, mainly resulting from the lack of suitable analytical methods. Typically, chromatographic methods, such as HPLC, are used to resolve DCL and to identify binders by comparing spectra with and without the target,[18, 21-23] but HPLC does not have the capacity to resolve large libraries containing many different compounds.[16, 24] Other methods, such as non-denaturing mass spectrometry,[25] NMR,[26] and spectroscopic methods (UV and fluorescence)[27-29] have been employed in DCL, but the resolution and throughput of these methods are also not sufficient for large libraries. Otto, Miller, and their respective co-workers have developed several elegant approaches capable of analyzing and selecting large DCLs (~10K compounds);[29-32] however, in most cases, DCLs only contain 10~100 compounds. Since the probability of discovering high affinity ligands increases with the library diversity, the limitation on library size has presented a significant obstacle for DCL.[23] New approaches capable of resolving and analyzing large DCLs are still highly desired.

DNA-encoded library (DEL), in which each compound is linked with a unique DNA tag, is another combinatorial library approach employing mixed compounds in library processing.[33-42] In contrast to DCL, due to DNA's high encoding capacity, DELs can contain millions of different compounds;[43-46] library selection can be feasibly decoded by using PCR amplification and DNA sequencing.[47, 48] Therefore, introducing DNA-encoding to DCL could be an effective strategy to address its limitation on library size. Previously, nucleic acids have been successfully used as programmable templates or scaffolds with spatial precision to display ligand combinations interrogating various biological targets.[49-66] The Neri group developed a method named Encoded Self-Assembling Combinatorial (ESAC)

[a] Key Laboratory of Bioorganic Chemistry and Molecular Engineering of the Ministry of Education, Beijing National Laboratory of Molecular Sciences (BNLMS),College of Chemistry and Molecular Engineering, Peking University, Beijing, China 100871. E-mail: xiaoyuli@pku.edu.cn
[b] Key Laboratory of Chemical Genomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen, China 518055
[c] Biodynamic Optical Imaging Centre (BIOPIC) and College of Engineering, Peking University, Beijing, China 100871

† Electronic Supplementary Information (ESI) available: Materials and general methods, experimental details, library selection and sequencing methods and enrichment fold calculations. See DOI: 10.1039/x0xx00000x

‡ These authors contributed equally to this work.

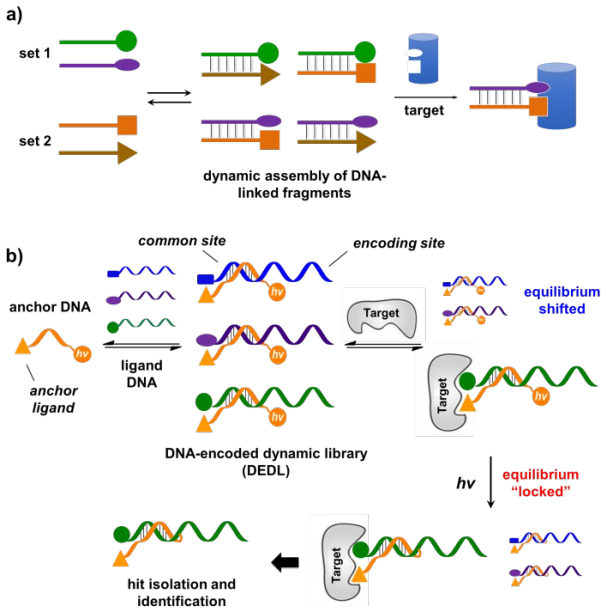


Figure 1: a) Previous work: two sets of DNA-linked fragments form dynamically exchanging duplexes; addition of target enriches high affinity duplexes.[49, 67, 68] b) DNA-encoded dynamic library (DEDL) (this work): An anchor DNA forms dynamic duplexes with multiple ligand DNAs forming the library. Adding the target shifts the equilibrium favouring the formation of high affinity binders. Photo-crosslinker in the anchor DNA locks the shifted equilibrium under irradiation. Crosslinked binders can then be isolated for hit identification via PCR amplification and DNA sequencing.

library, in which two sets of DNA-linked fragments form a static library by combinatorial duplex formation.[65, 66] Hamilton and co-workers introduced dynamic exchange in DNA hybridization, so that the target can shift the equilibrium and enrich high affinity fragment combinations (Figure 1a).[49, 67] Very recently, Zhang and co-workers reported a similar system achieving target-induced enrichment of DNA duplexes.[68] These studies have nicely shown that the principle of dynamic exchange can be applied to DELs; however, more systematic methodology for the preparation and selection of DNA-encoded dynamic library (DEDL) has yet to be developed. Moreover, previous studies require modified and immobilized targets in library selection, which is not compatible with proteins that are difficult to purify or modify, such as membrane proteins.[69,70] Aiming to address these issues, here we report the detailed study of a DEDL system, including library preparation, encoding, selection, hit deconvolution, and notably, a novel "locking" strategy to freeze the equilibrium shift for hit isolation and identification.

Results and discussion

Our strategy is shown in Figure 1b. Libraries BBs are conjugated to different DNA strands (ligand DNAs), all having a common sequence that can form dynamically exchanging duplexes with an "anchor DNA", which is conjugated with an "anchor" molecule. Upon target addition, the equilibrium shifts to form more high affinity bivalent duplexes. Next, the photo-reactive group on the anchor DNA can crosslink the two DNA strands upon irradiation, thereby stopping the dynamic exchange and locking the shifted equilibrium. The distal region on the ligand DNA encodes BB's chemical identity, and the crosslinked duplex can be isolated for hit identification with PCR amplification and DNA sequencing (Figure 1b). By combining features of DEL and DCL, our design allows for the selection of high diversity DCLs to discover synergistic fragments for "affinity maturation" of the anchor molecule.[65, 66, 71, 72]

We first verified that dynamic DNA duplex formation can be affected by the target protein.[49, 68] As shown in Figure 2a, a fluorescein (FAM) and a quencher (DABCYL) were conjugated to two complementary DNAs, respectively; the decrease of fluorescence therefore reports DNA hybridization. The other end of DNA was conjugated with a biotin, a desthiobiotin, or an iminobiotin (Figure 2b). These ligands are well known to bind adjacent pockets on the tetrameric protein streptavidin (SA) with different affinities (K_d: 40 fM, 2.0 nM and 50 nM, respectively).[73] Moreover, we reason that, in order to establish dynamic exchange, DNA duplex should have a melting temperature (Tm) close to the experiment temperature, and it also should be sufficiently long to ensure hybridization specificity; therefore, either 6- or 7-base DNA duplex was chosen in our study

As shown in Figure 2c, for all three ligands, fluorescence decreased significantly in the presence of SA, suggesting the formation of the ternary complex (**i**). In contrast, in control

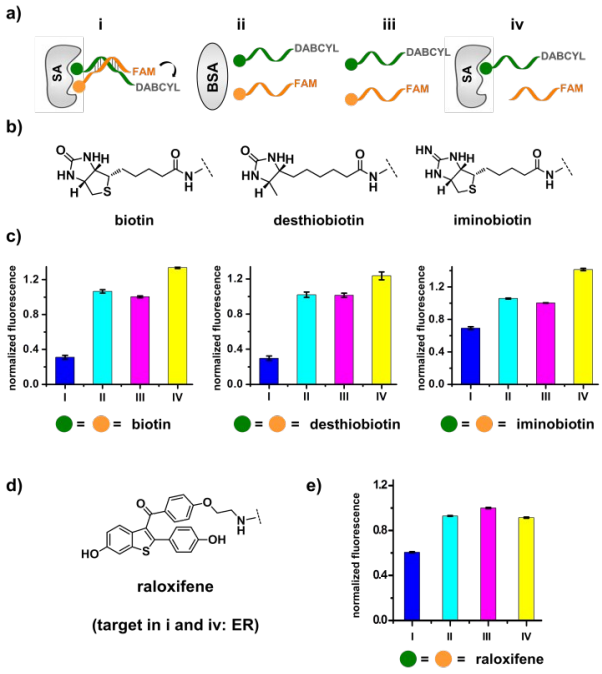


Figure 2: a) Two complementary DNAs conjugated with biotin, desthiobiotin, or iminobiotin and FAM/DABCYL groups were mixed with the target SA (**i**), with BSA (**ii**), no protein (**iii**), or with one ligand omitted (**iv**). b) Structures of the small molecule ligands. c) Fluorescence quenching results. FAM fluorescence values were measured and normalized to (**iii**). Left panel: with biotin; middle panel: with desthiobiotin; right panel: with iminobiotin. d) Structure of the raloxifene ligand. e) Fluorescence quenching results of the raloxifene-ER system. ER was used as the target in (**i**) and (**iv**). DNA: 200 nM each; protein: 400 nM. DNAs and protein were incubated at 30 °C for 1 h before measured by a fluorophotometer. Excitation: 494 nm; emission: 522 nm. Error bars (standard deviation, SD) are based on three replicates of each experiment.

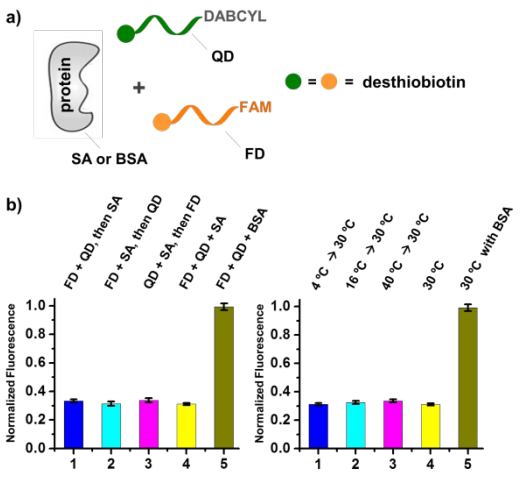


Figure 3: a) Two desthiobiotin-labelled DNAs conjugated with FAM or DABCYL were mixed with SA or BSA at different orders or temperatures; fluorescence decreases were then measured. b) Left panel: data from different mixing orders. Right panel: data from different temperatures; SA was used as the target except in 5 where BSA was used as a negative control. Experiment conditions are the same as in Figure 2 except mixing order and temperature. Error bars (standard deviation, SD) are based on three replicates of each experiment.

experiments with the non-specific protein BSA (bovine serum albumin) (**ii**), without SA (**iii**), or with one ligand omitted (**iv**), little or no fluorescence decrease was observed, indicating the quenching in (**i**) depends on specific bivalent binding to SA. Notably, ~40% quenching was observed for the weak binder iminobiotin (Figure 2c, right panel). Furthermore, we performed similar fluorescence quenching experiments with raloxifene, an estrogen receptor (ER) modulator (Figure 2d);[74] dimeric raloxifene ligands are able to bind the two binding pockets on estrogen receptor dimers.[61, 75, 76] Similar to the biotin ligand series, significant fluorescence decrease was observed in the presence of the specific target ER and the bivalent raloxifene duplex (Figure 2e). In addition, as a thermodynamically-controlled system, an important feature of DCL is that the same state of equilibrium can be reached from different starting points.[77, 78] In order to verify this, we either altered the mixing order or incubated the mixture at 4 °C, 16 °C, 30 °C or 40 °C for 30 min before incubation at 30 °C for another hour (**QD** and **FD**; Figure 3a). We observed that all experiments reached the same equilibrium based on fluorescence readings, proving the dynamic nature of our system (Figure 3b).

Next, we investigated whether the target has shifted the equilibrium to promote the assembly of high affinity duplexes. As shown in figure 4, we mixed a non-fluorescent background DNA (5'-GTCTGC-3'-NH_2; **BD-1**) with a fluorescent ligand DNA (5'-FAM-GTCTGC-3'-ligand; **LD-1**) at 8:1 ratio. Both DNAs dynamically compete for the hybridization to **AD-1**, which is conjugated with an anchor ligand and a DABCYL quencher (5'-ligand-GCAGACT-3'-DABCYL). The bivalent **LD-1**/**AD-1** duplex is expected to have higher affinity for SA than the monovalent **BD-1**/**AD-1**. After mixing the DNAs (**BD-1**/**LD-1**/**AD-1**: 8:1:1) with SA, we observed significant fluorescence quenching for all three ligands, indicating the equilibrium has been shifted to favour the formation of the (**LD-1**/**AD-1**)-SA ternary complex.

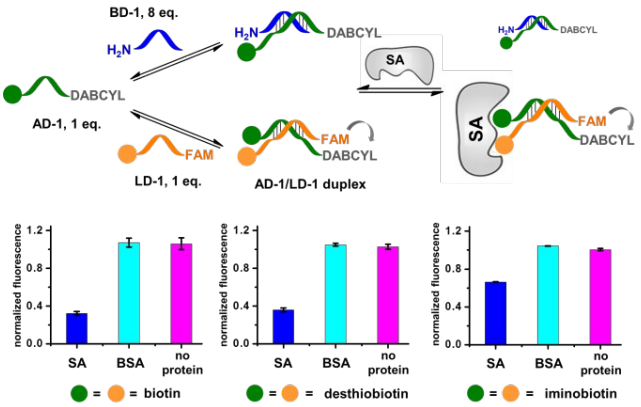


Figure 4: Verification of target-induced equilibrium shift, determined by fluorescence decrease. Fluorescence values were normalized to the "no protein" experiment. AD-1 and LD-1: 200 nM; BD-1: 1.6 μM; proteins: 400 nM. Experiment procedures are the same as in Figure 2. Error bars (standard deviation, SD) are based on three replicates of each experiment

As a negative control, the non-binding BSA did not shift the equilibrium (cyan columns; Figure 4). These results have demonstrated that the target can indeed promote the assembly of high affinity binders.

In selection of DCLs, it is often necessary to stop the dynamic exchange and "freeze" the shifted equilibrium, so that the library population change, induced by the target, can be preserved for further characterization. For example, adding $NaBH_3CN$ to reduce imine to stable amine is a popular method to stop the dynamic imine formation;[21, 22, 79-81] lowering pH can effectively disable disulfide exchange and reversible Michael addition, which optimally occur at basic pH.[16, 17, 24, 31, 78] In this study, we designed a novel photo-crosslinking strategy to stop the dynamic DNA duplex exchange. Photo-crosslinking is kinetically fast and can be imposed/withdrawn conveniently with minimal perturbation to the system.[82] As shown in Figure 5a, psoralen (PS), a photo-crosslinker widely used in nucleic acid crosslinking,[83-86] was conjugated to the 5'-end of a short

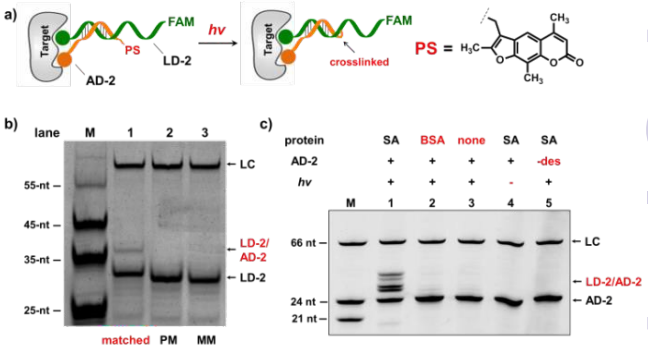
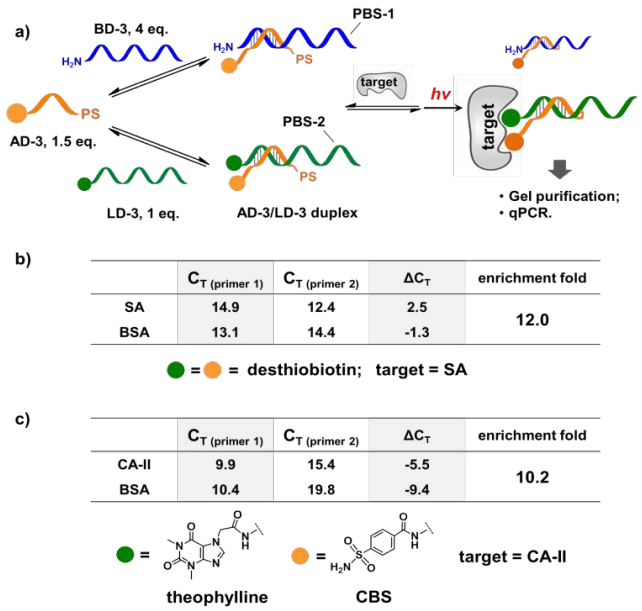


Figure 5: a) Psoralen (PS) was chosen to lock the equilibrium in DEDL. b) Three sets of **AD-2**/**LD-2** were mixed, irradiated, and analysed by denaturing electrophoresis. PM: partially mismatched (2-base mismatch); MM: fully mismatched; LC: a 66-nt DNA loading control. c) 5'-desthiobiotin-labeled **LD-2** and **AD-2** were mixed, irradiated under different conditions, and then analysed by denaturing electrophoresis (18% TBE-Urea denaturing PAGE). Lane 1: with SA; lane 2: with BSA; lane 3: no protein added; lane 4: no irradiation; lane 5: no desthiobiotin on **AD-2**. **AD-2**: 300 nM; **LD-2**: 200 nM; irradiation: 365 nm for 30 s at 30 °C using a UV LED point light system; short irradiation reduces non-specific crosslinking in the background. M: marker; -des: no desthiobiotin.



	C_T (primer 1)	C_T (primer 2)	ΔC_T	enrichment fold
SA	14.9	12.4	2.5	12.0
BSA	13.1	14.4	-1.3	

	C_T (primer 1)	C_T (primer 2)	ΔC_T	enrichment fold
CA-II	9.9	15.4	-5.5	10.2
BSA	10.4	19.8	-9.4	

Figure 6: a) PS-based photo-crosslinking locks the shifted equilibrium for the subsequent hit isolation and qPCR analysis. After gel-purification of crosslinked duplexes, qPCR was performed to determine C_T values and to calculate enrichment fold of the **AD-3/LD-3** duplex. b) Results of the bivalent desthiobiotin-SA system. c) Results of the theophylline/CBS-CA-II system. $\Delta C_T = C_{T\ (primer\ 1)} - C_{T\ (primer\ 2)}$. **AD-3**: 300 nM; **LD-3**: 200 nM; primers: 200 nM. Experimental procedure is the same as in Figure 5. See the Supporting Information for details.

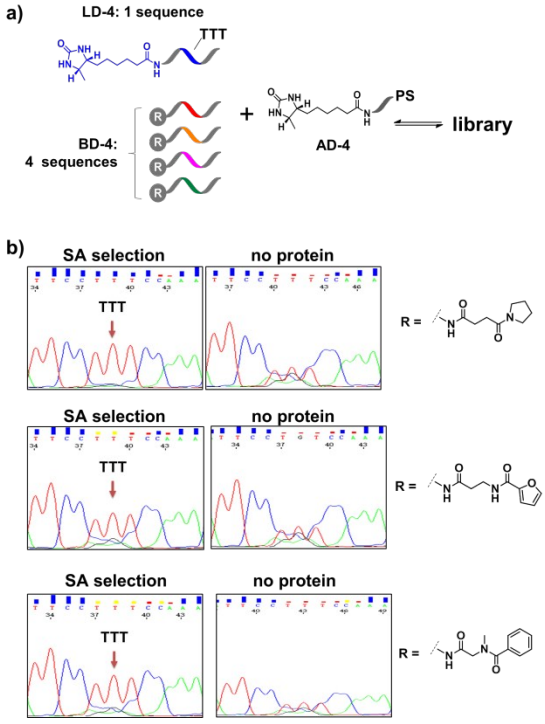


Figure 7: a) Components of the model libraries. Libraries were selected against SA with the same procedure as in Figure 6, except Sanger sequencing was used. b) Sequencing results. Left panels: after SA selection; right panels: control selection without SA. **LD-4**: 200 nM; **BD-4**: 800 nM (total); **AD-4**: 300 nM; SA: 400 nM. See the Supporting Information for details.

7-nt DNA bearing the anchor molecule (**AD-2**). **AD-2** is complementary to the 5'-end of a 24-nt DNA having a ligand and a FAM group (**LD-2**). Moreover, **LD-2** also contains a thymine at the site opposite to PS, which is known to be able to improve the crosslinking efficiency.[87] After DNA incubation and target addition, irradiation triggers the crosslinking between **AD-2** and **LD-2**, thereby stopping strand exchange and locking the equilibrium. The crosslinked **AD-2**/**LD-2** duplex can then be isolated for PCR amplification and DNA sequencing to decode the ligand synergistically binding to the target with the anchor molecule. First, we prepared fully matched, partially mismatched, and fully mismatched **AD-2**/**LD-2** duplexes. These DNAs were mixed, irradiated, and analysed by denaturing electrophoresis. Crosslinked product was only observed with the fully matched DNAs (lane 1; Figure 5b). Next, a set of desthiobiotin-labelled **AD-2** and **LD-2** were subjected to the same procedure; results show that only in the presence of SA was the crosslinking product detected (lane 1; Figure 5c). Multiple bands appeared in lane 1 of Figure 5c; mass analysis confirmed that all are crosslinked duplexes (see the Supporting Information). We hypothesize that the "T" shape of the crosslinked duplex may partially renature in gel, a phenomenon we have observed previously.[88] In all negative controls (with BSA, no protein, no irradiation, no desthiobiotin in **AD-2**; lanes 2-5, Figure 5c), no or very little crosslinking was detected. The product bands were excised, extracted, and quantified. With SA, a 40% crosslinking yield was obtained. Collectively, these results have demonstrated the specificities of PS-based inter-strand DNA crosslinking and its suitability for capturing target-induced duplex formation.

Next, we mixed a background DNA (5'-NH_2, 28-nt; **BD-3**, with a ligand DNA (5'-desthiobiotin, 28-nt; **LD-3**) at 4:1 ratio. **BD-3** and **LD-3** have orthogonal primer binding sites (PBS-1 and PBS-2; Figure 6a). Both **BD-3** and **LD-3** have a 7-base region complementary to a short anchor DNA (3'-desthiobiotin, 7-nt; **AD-3**). These DNAs were mixed at 4:1:1.5 ratio to form the dynamic library. After adding SA, the mixture was irradiated and the crosslinked duplexes were gel- purified for qPCR (quantitative PCR) analysis. qPCR threshold cycle values (C_T's) were determined to calculate the initial copy numbers of **LD-3/AD-3** and **BD-3/AD-3** duplexes with their respective primers.[89, 90] In order to offset possible biases from experimental factors, the library was also subjected to the same procedure (irradiation, gel purification, and qPCR) with the control protein BSA. Enrichment folds were then calculated by comparing results from these two selections (see the Supporting Information for calculation method; Figure S2-S4). As a result, a 12.0-fold enrichment of the high affinity **LD-3/AD-3** duplex was achieved (Figure 6b), which is comparable to typical DCL-based selections.[16, 18, 19, 91, 92] Gel analysis also directly confirmed the enrichment of the crosslinked **LD-3/AD-3** duplex (Figure S5). Moreover, in order to test the generality of our method, we conjugated another pair of ligands, theophylline and CBS to **LD-3** and **AD-3** DNAs, respectively. Theophylline and CBS were found to synergistically bind the target of carbonic anhydrase-II (CA-II) in an ESAC library selection.[66] After mixing with the background DNA **BD-3**, the

formed dynamic library was subjected to the selection against target CA-II and negative control BSA with the same procedure. Results show that a 10.2-fold enrichment of the **LD-3/AD-3** duplex was achieved (Figure 6c). Collectively, these results have demonstrated that the PS-based crosslinking mechanism is suitable for locking and analysing equilibrium shift in DEDL selections.

Encouraged by these results, we further prepared several model DCLs (Figure 7a). These libraries contain a desthiobiotin-labelled ligand DNA (**LD-4**) and 4 background DNAs (**BD-4**'s), all dynamically competing for an anchor DNA (with desthiobiotin, **AD-4**). **BD-4**'s are also conjugated with several small molecules that are not known to bind SA, but representing typical fragment structures in a library. The ligand of desthiobiotin in **LD-4** is encoded by a "TTT" codon, while **BD-4**'s contain varied sequences at the encoding site ("AAG", "GCA", "ACA", "CGC"). These DNAs were mixed at equal ratio to form the library and then selected against SA. After irradiation, "hit compounds" were isolated and decoded with the same procedure as in Figure 6 except Sanger sequencing was used. As shown in figure 7b, in all cases, the "TTT" codon encoding the desthiobiotin in **LD-4** has been enriched markedly by SA due to the high affinity of **LD-4**/**AD-4** duplex (left panels), whereas negative selections (no protein) only generated scrambled sequences at the encoding site (Figure 7b, right panels).

Finally, in order to mimic library diversity, we prepared a model DCL containing 1,024 background DNAs (**BD-5's**), a ligand DNA (with desthiobiotin, **LD-5**), and an anchor DNA (with desthiobiotin**, AD-5**) (Figure 8). **LD-5** and all **BD-5**'s were mixed at equal ratio, realizing a 1,024-fold excess of background DNAs relative to **LD-5**. This library was selected against the target SA and also subjected to a "no-protein" control selection, similar to Figure 6, to control for biases from selection procedures (irradiation, gel purification, PCR, sequencing, etc.). Selection results were decoded by high throughput DNA sequencing (Illumina®). Enrichment folds of selected sequences were plotted against sequence counts to identify "hit compounds" (Figure 8b). Again due to the high affinity of **LD-5**/**AD-5** duplex, the sequence that encodes **LD-5** was distinctly enriched (19.2 folds). In addition, the expected "hit" **LD-5** shows high sequence count ratio after the target selection, while having average count ratio in the control selection, further confirming its target specificity (Figure 8c and S6). It is worth noting that the wide distribution of sequence counts in both target and control selections indicates that sufficient sequencing depth and high library synthesis quality (even distribution of library members)[93] are both important in library selections. Although this model library only has limited chemical diversity, these results have demonstrated our approach's suitability for the selection of large dynamic libraries.

Conclusions

In conclusion, we have developed a DNA-encoded dynamic library (DEDL) approach for the preparation and selection of large dynamic libraries. Notably, we introduced a novel locking

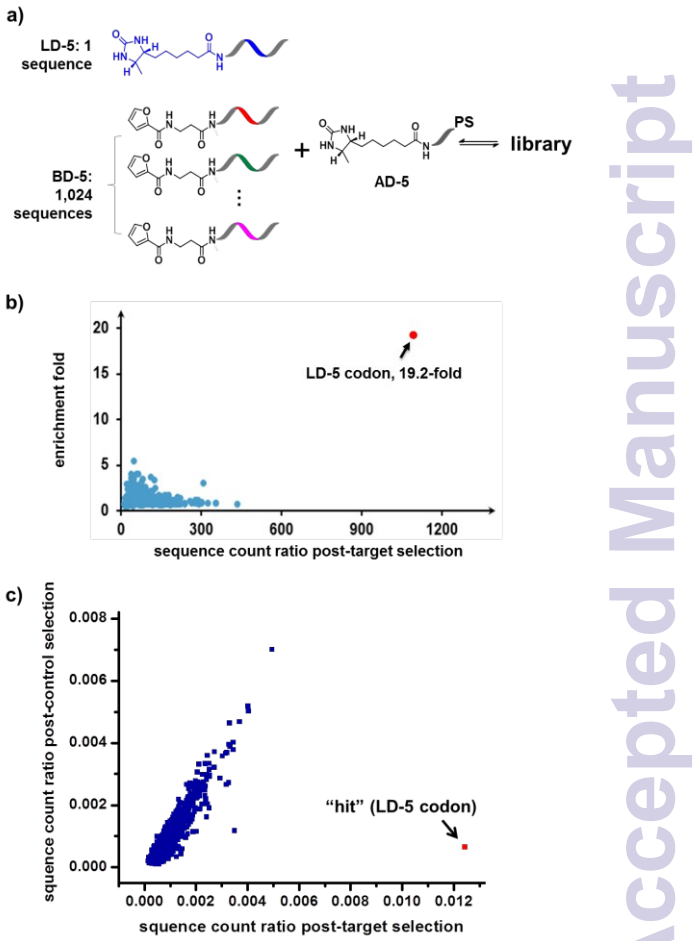


Figure 8: a) Components of the 1,025-member library. b) The library was selected against SA with the same procedure as in Figure 6, except Illumina® sequencing was used. b) Plot of enrichment fold versus sequence count after the target selection. Enrichment fold = (post-target selection fraction)/(post-control selection fraction). c) Plot of sequence count ratios after the control selection (no protein added) versus count ratios after the target selection (with SA). Sequence ratio = (sequence count)/(total sequence count of the library). Each dot represents the DNA sequence corresponding to a library member. "Hit" containing the desired **LD-5** codon is highlighted in red. **LD-5**: 0.19 nM; **BD-5**: 200 nM (total); **AD-5**: 300 nM; SA: 400 nM. Enrichment folds for low-count library members vary widely due to statistical under-sampling. See the Supporting Information for more details on experimental procedure, data analysis and further discussion of the sequencing results.

mechanism, which is able to take a "snapshot photo" of the library equilibrium altered by the target protein, thereby enabling the downstream hit isolation and identification. Second, our method eliminated the requirement of target immobilization and physical washing; therefore, target-induced perturbation of the library equilibrium is better preserved, and unmodified, non-immobilized proteins can be used as targets.[69, 70, 90]

However, the present method only encodes one fragment and thus is limited to the "affinity maturation" of known ligands (the "anchor"),[66, 71, 72] rendering it unsuitable for the *de novo* discovery of synergistic fragment combinations.[38] In contrast, previously nucleic acids have been successfully used as templates to pair DNA/PNA-linked small molecule ligands, therefore enabling the selections of synergistic fragment pairs against biological targets,[50, 52, 53, 57-64] and the strategy of inter-

strand code-transfer also realized the dual-pharmacophore ESAC libraries.[65] These elegant studies highlight the importance of further development of dual-display DNA-encoded dynamic library,[8, 94] which indeed is currently being pursued in our laboratory by using alternative DNA architecture, more efficient crosslinker,[95] and different decoding scheme.[96] We will report the results in due course.

Acknowledgements

This work was supported by Ministry of Science and Technology Basic Research Program (2011CB809100), NSFC (21272016, 21002003, 91013003) and the Doctoral Fund of Ministry of Education of China (20120001110083). We thank the mass spectrometry facility of National Centre for Protein Sciences at Peking University and Dr. Rong Meng for the kind help with mass spectroscopy analysis.

Notes and references

1. P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. Sanders and S. Otto, *Chem. Rev.*, 2006, **106**, 3652-3711.
2. F. B. Cougnon and J. K. Sanders, *Acc. Chem. Res.*, 2011, **45**, 2211-2221.
3. Y. Jin, C. Yu, R. J. Denman and W. Zhang, *Chem. Soc. Rev.*, 2013, **42**, 6634-6654.
4. C. D. Meyer, C. S. Joiner and J. F. Stoddart, *Chem. Soc. Rev.*, 2007, **36**, 1705-1723.
5. M. Mondal and A. K. Hirsch, *Chem. Soc. Rev.*, 2015, **44**, 2455-2488.
6. O. Ramstrom and J. M. Lehn, *Nat. Rev. Drug Discov.*, 2002, **1**, 26-36.
7. S. Otto, R. L. Furlan and J. K. Sanders, *Curr. Opin. Chem. Biol.*, 2002, **6**, 321-327.
8. S. Otto, R. L. Furlan and J. K. Sanders, *Drug Discov. Today*, 2002, **7**, 117-125.
9. J.-M. Lehn, *Angew. Chem. Int. Ed.*, 2015, **54**, 3276-3289.
10. J. M. Lehn, *Chem. Soc. Rev.*, 2007, **36**, 151-160.
11. O. Ramstrom, T. Bunyapaiboonsri, S. Lohmann and J. M. Lehn, *Biochim. Biophys. Acta*, 2002, **1572**, 178-186.
12. B. L. Miller, *Dynamic combinatorial chemistry: in drug discovery, bioorganic chemistry, and materials science*, John Wiley & Sons, 2009.
13. K. R. West, K. D. Bake and S. Otto, *Org. Lett.*, 2005, **7**, 2615-2618.
14. K. R. West and S. Otto, *Curr. Drug Discov. Technol.*, 2005, **2**, 123-160.
15. M. F. Schmidt, A. Isidro‐Llobet, M. Lisurek, A. El‐Dahshan, J. Tan, R. Hilgenfeld and J. Rademann, *Angew. Chem. Int. Ed.*, 2008, **47**, 3275-3278.
16. B. Shi, R. Stevenson, D. J. Campopiano and M. F. Greaney, *J. Am. Chem. Soc.*, 2006, **128**, 8459-8467.
17. D. E. Scott, G. J. Dawes, M. Ando, C. Abell and A. Ciulli, *Chembiochem*, 2009, **10**, 2772-2779.
18. V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano and M. F. Greaney, *Nat. Chem.*, 2010, **2**, 490-497.
19. M. Mondal, N. Radeva, H. Koster, A. Park, C. Potamitis, M. Zervou, G. Klebe and A. K. Hirsch, *Angew. Chem. Int. Ed.*, 2014, **53**, 3259-3263.
20. G. Klebe, *Nat. Rev. Drug Discov.*, 2015, **14**, 95-110.
21. I. Huc and J.-M. Lehn, *Proc. Nat. Acad. Sci. USA*, 1997, **94**, 2106-2110.
22. S. Zameo, B. Vauzeilles and J.-M. Beau, *Angew. Chem. Int. Ed.*, 2005, **44**, 965-969.
23. S. Ladame, *Org. Biomol. Chem.*, 2008, **6**, 219-226.
24. O. Ramström and J. M. Lehn, *Chembiochem*, 2000, **1**, 41-48.
25. M. Demetriades, I. K. Leung, R. Chowdhury, M. C. Chan, M. A. McDonough, K. K. Yeoh, Y. M. Tian, T. D. Claridge, P. J. Ratcliffe and E. C. Woon, *Angew. Chem. Int. Ed.*, 2012, **51** 6672-6675.
26. R. Larsson, Z. Pei and O. Ramstrom, *Angew. Chem. Int. Ed.*, 2004, **43**, 3716-3718.
27. S. M. Voshell, S. J. Lee and M. R. Gagne, *J. Am. Chem. Soc.*, 2006, **128**, 12422-12423.
28. A. Buryak and K. Severin, *Angew. Chem. Int. Ed.*, 2005, **44**, 7935-7938.
29. B. R. McNaughton and B. L. Miller, *Org. Lett.*, 2006, **8**, 1803-1806.
30. R. F. Ludlow and S. Otto, *J. Am. Chem. Soc.*, 2008, **130**, 12218-12219.
31. P. C. Gareiss, K. Sobczak, B. R. McNaughton, P. B. Palde, C. A. Thornton and B. L. Miller, *J. Am. Chem. Soc.*, 2008, **130**, 16254-16261.
32. B. R. McNaughton, P. C. Gareiss and B. L. Miller, *J. Am. Chem. Soc.*, 2007, **129**, 11306-11307.
33. R. M. Franzini, D. Neri and J. Scheuermann, *Acc. Chem. Re.s*, 2014, **47**, 1247-1255.
34. R. E. Kleiner, C. E. Dumelin and D. R. Liu, *Chem. Soc. Rev.*, 2011, **40**, 5707-5717.
35. K. Gorska and N. Winssinger, *Angew. Chem. Int. Ed.*, 2013, **52**, 6820-6843.
36. L. Mannocci, M. Leimbacher, M. Wichert, J. Scheuermann and D. Neri, *Chem. Commun.*, 2011, **47**, 12747-12753.
37. K. Yea, J. Xie, H. Zhang, W. Zhang and R. A. Lerner, *Curr. Opin. Chem. Biol.*, 2015, **26**, 1-7.
38. J. Scheuermann and D. Neri, *Curr. Opin. Chem. Biol.*, 2015, **26**, 99-103.
39. A. I. Chan, L. M. McGregor and D. R. Liu, *Curr. Opin. Chem. Biol.*, 2015, **26**, 55-61.
40. P. Blakskjaer, T. Heitner and N. J. V. Hansen, *Curr. Opin. Chem. Biol.*, 2015, **26**, 62-71.
41. A. D. Keefe, M. A. Clark, C. D. Hupp, A. Litovchick and Y. Zhang, *Curr. Opin. Chem. Biol.*, 2015, **26**, 80-88.
42. G. Li, W. Zheng, Y. Liu and X. Li, *Curr. Opin. Chem. Biol.*, 2015, **26**, 25-33.
43. M. A. Clark, R. A. Acharya, C. C. Arico-Muendel, S. L. Belyanskaya, D. R. Benjamin, N. R. Carlson, P. A. Centrella, C. H. Chiu, S. P. Creaser, J. W. Cuozzo, C. P. Davie, Y. Ding, G. J. Franklin, K. D. Franzen, M. L. Gefter, S. P. Hale, N. J. V. Hansen, D. I. Israel, J. W. Jiang, M. J. Kavarana, M. S. Kelley, C.

S. Kollmann, F. Li, K. Lind, S. Mataruse, P. F. Medeiros, J. A. Messer, P. Myers, H. O'Keefe, M. C. Oliff, C. E. Rise, A. L. Satz, S. R. Skinner, J. L. Svendsen, L. J. Tang, K. van Vloten, R. W. Wagner, G. Yao, B. G. Zhao and B. A. Morgan, *Nat. Chem. Biol.*, 2009, **5**, 647-654.

44. H. Deng, H. O'Keefe, C. P. Davie, K. E. Lind, R. A. Acharya, G. J. Franklin, J. Larkin, R. Matico, M. Neeb, M. M. Thompson, T. Lohr, J. W. Gross, P. A. Centrella, G. K. O'Donovan, K. L. Bedard, K. van Vloten, S. Mataruse, S. R. Skinner, S. L. Belyanskaya, T. Y. Carpenter, T. W. Shearer, M. A. Clark, J. W. Cuozzo, C. C. Arico-Muendel and B. A. Morgan, *J. Med. Chem.*, 2012, **55**, 7061-7079.
45. J. S. Disch, G. Evindar, C. H. Chiu, C. A. Blum, H. Dai, L. Jin, E. Schuman, K. E. Lind, S. L. Belyanskaya, J. Deng, F. Coppo, L. Aquilani, T. L. Graybill, J. W. Cuozzo, S. Lavu, C. Mao, G. P. Vlasuk and R. B. Perni, *J. Med. Chem.*, 2013, **56**, 3666-3679.
46. C. S. Kollmann, X. Bai, C.-H. Tsai, H. Yang, K. E. Lind, S. R. Skinner, Z. Zhu, D. I. Israel, J. W. Cuozzo, B. A. Morgan, K. Yuki, C. Xie, T. A. Springer, M. Shimaoka and G. Evindar, *Bioorg. Med. Chem.*, 2014, **22**, 2353-2365.
47. L. Mannocci, Y. X. Zhang, J. Scheuermann, M. Leimbacher, G. De Bellis, E. Rizzi, C. Dumelin, S. Melkko and D. Neri, *Proc. Nat. Acad. Sci. USA*, 2008, **105**, 17670-17675.
48. R. E. Kleiner, C. E. Dumelin, G. C. Tiu, K. Sakurai and D. R. Liu, *J. Am. Chem. Soc.*, 2010, **132**, 11779-11791.
49. K. I. Sprinz, D. M. Tagore and A. D. Hamilton, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 3908-3911.
50. D. Sil, J. B. Lee, D. Luo, D. Holowka and B. Baird, *ACS Chem. Biol.*, 2007, **2**, 674-684.
51. D. M. Tagore, K. I. Sprinz, S. Fletcher, J. Jayawickramarajah and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2007, **46**, 223-225.
52. K. Gorska, K. T. Huang, O. Chaloin and N. Winssinger, *Angew. Chem. Int. Ed.*, 2009, **48**, 7695-7700.
53. B. A. Williams, C. W. Diehnelt, P. Belcher, M. Greving, N. W. Woodbury, S. A. Johnston and J. C. Chaput, *J. Am. Chem. Soc.*, 2009, **131**, 17233-17241.
54. D. Margulies and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2009, **48**, 1771-1774.
55. D. Margulies and A. D. Hamilton, *J. Am. Chem. Soc.*, 2009, **131**, 9142-9143.
56. K. Gorska, J. Beyrath, S. Fournel, G. Guichard and N. Winssinger, *Chem. Commun.*, 2010, **46**, 7742-7744.
57. C. Scheibe, A. Bujotzek, J. Dernedde, M. Weberb and O. Seitz, *Chem. Sci.*, 2011, **2**, 770-775.
58. M. Ciobanu, K. T. Huang, J. P. Daguer, S. Barluenga, O. Chaloin, E. Schaeffer, C. G. Mueller, D. A. Mitchell and N. Winssinger, *Chem. Commun.*, 2011, **47**, 9321-9323.
59. H. Eberhard, F. Diezmann and O. Seitz, *Angew. Chem. Int. Ed.*, 2011, **50**, 4146-4150.
60. J. P. Daguer, M. Ciobanu, S. Alvarez, S. Barluenga and N. Winssinger, *Chem. Sci.*, 2011, **2**, 625-632.
61. F. Abendroth, A. Bujotzek, M. Shan, R. Haag, M. Weber and O. Seitz, *Angew. Chem. Int. Ed.*, 2011, **50**, 8592-8596.
62. K. T. Huang, K. Gorska, S. Alvarez, S. Barluenga and N. Winssinger, *Chembiochem*, 2011, **12**, 56-60.
63. E. A. Englund, D. Wang, H. Fujigaki, H. Sakai, C. M. Micklitsch, R. Ghirlando, G. Martin-Manso, M. L. Pendrak, D. D. Roberts, S. R. Durell and D. H. Appella, *Nat. Commun.*, 2012, **3**, 614.
64. J. P. Daguer, C. Zambaldo, M. Ciobanu, P. Morieux, S. Barluenga and N. Winssinger, *Chem. Sci.*, 2015, **6**, 739-744.
65. M. Wichert, N. Krall, W. Decurtins, R. M. Franzini, F. Pretto, P. Schneider, D. Neri and J. Scheuermann, *Nat. Chem.*, 2015, **7**, 241-249.
66. S. Melkko, J. Scheuermann, C. E. Dumelin and D. Neri, *Nat. Biotechnol.*, 2004, **22**, 568-574.
67. D. M. Tagore, K. I. Sprinz and A. D. Hamilton, *Supramolecular Chem.*, 2007, **19**, 129-136.
68. F. V. Reddavide, W. Lin, S. Lehnert and Y. Zhang, *Angew. Chem. Int. Ed.*, 2015, **54**, 7924-7928.
69. L. M. McGregor, T. Jain and D. R. Liu, *J. Am. Chem. Soc.*, 2014, **136**, 3264-3270.
70. P. Zhao, Z. Chen, Y. Li, D. Sun, Y. Gao, Y. Huang and X. Li, *Angew Chem Int Ed Engl*, 2014, **53**, 10056-10059.
71. S. Melkko, Y. Zhang, C. E. Dumelin, J. Scheuermann and D. Neri, *Angew. Chem. Int. Ed.*, 2007, **46**, 4671-4674.
72. J. Scheuermann, C. E. Dumelin, S. Melkko, Y. Zhang, L. Mannocci, M. Jaggi, J. Sobek and D. Neri, *Bioconj. Chem.*, 2008, **19**, 778-785.
73. N. Green, *Biochem. J.*, 1966, **101**, 774-780.
74. C. D. Jones, M. G. Jevnikar, A. J. Pike, M. K. Peters, L. J. Black, A. R. Thompson, J. F. Falcone and J. A. Clemens, *J. Med. Chem.*, 1984, **27**, 1057-1066.
75. M. Shan, A. Bujotzek, F. Abendroth, A. Wellner, R. Gust, O. Seitz, M. Weber and R. Haag, *Chembiochem*, 2011, **12**, 2587-2598.
76. K. K. Sadhu, T. Eierhoff, W. Romer and N. Winssinger, *J. Am. Chem. Soc.*, 2012, **134**, 20013-20016.
77. B. Shi and M. F. Greaney, *Chem. Commun., 2005, **7**, 886-888.*
78. S. Ladame, A. M. Whitney and S. Balasubramanian, *Angew. Chem. Int. Ed.* 2005, **44**, 5736-5739.
79. M. Hochgürtel, R. Biesinger, H. Kroth, D. Piecha, M. W. Hofmann, S. Krause, O. Schaaf, C. Nicolau and A. V. Eliseev, *J. Med. Chem.*, 2003, **46**, 356-358.
80. J. C. Leitzel and D. G. Lynn, *Chem. Rec.*, 2001, **1**, 53-62.
81. P. Luo, J. C. Leitzel, Z.-Y. J. Zhan and D. G. Lynn, *J. Am. Chem. Soc.*, 1998, **120**, 3019-3031.
82. G. W. Preston and A. J. Wilson, *Chem. Soc. Rev.*, 2013, **42** 3289-3301.
83. T. Stafforst and D. Hilvert, *Angew. Chem. Int. Ed.*, 2011, **50**, 9483-9486.
84. H. Baigude, Z. Li, Y. Zhou and T. M. Rana, *Angew. Chem. Int. Ed.*, 2012, **51**, 5880-5883.
85. J. E. Hearst, *Annu. Rev. Biophys. Bioeng.*, 1981, **10**, 69-86.
86. U. Pieles and U. Englisch, *Nucleic Acids Res.*, 1989, **17**, 285-299.
87. H. Baigude, Ahsanullah, Z. Li, Y. Zhou and T. M. Rana, *Angew. Chem. Int. Ed*, 2012, **51**, 5880-5883.
88. G. Li, Y. Liu, L. Chen, S. Wu and X. Li, *Angew. Chem. Int. Ed.*, 2013, **52**, 9544-9549.
89. D. J. Gorin, A. S. Kamlet and D. R. Liu, *J. Am. Chem. Soc.*, 2009 **131**, 9189-9191..

90. L. M. McGregor, D. J. Gorin, C. E. Dumelin and D. R. Liu, *J. Am. Chem. Soc.*, 2010, **132**, 15522-15524.
91. M. Demetriades, I. K. H. Leung, R. Chowdhury, M. C. Chan, M. A. McDonough, K. K. Yeoh, Y.-M. Tian, T. D. W. Claridge, P. J. Ratcliffe, E. C. Y. Woon and C. J. Schofield, *Angew. Chem. Int. Ed.*, 2012, **51**, 6672-6675.
92. D. E. Scott, G. J. Dawes, M. Ando, C. Abell and A. Ciulli, *Chembiochem*, 2009, **10**, 2772-2779.
93. A. L. Satz, *ACS Chem. Biol.* 2015, **DOI:** 10.1021/acschem bio.5 b00378.
94. V. Goral, M. I. Nelen, A. V. Eliseev and J. M. Lehn, *Proc. Nat. Acad. Sci. USA*, 2001, **98**, 1347-1352.
95. H. Kashida, T. Doi, T. Sakakibara, T. Hayashi and H. Asanuma, *J. Am. Chem. Soc.*, 2013, **135**, 7960-7966.
96. A. Litovchick, M. A. Clark and A. D. Keefe, *Artif. DNA PNA XNA*, 2014, **5**, e27896.