Article

Selecting DNA-encoded Chemical Library Against Non-immobilized Proteins Using a "Ligate-Crosslink-Purify" Strategy

Bingbing Shi, Yuqing Deng, Peng Zhao, and Xiaoyu Li

Bioconjugate Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.bioconjchem.7b00343 • Publication Date (Web): 25 Jul 2017

Selecting DNA-encoded Chemical Library Against Non-immobilized Proteins Using a "Ligate-Crosslink-Purify" Strategy

Bingbing Shi[†,‡], Yuqing Deng[‡], Peng Zhao[‡,§], and Xiaoyu Li[*,‡]

[†]Key Laboratory of Chemical Genomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, 2199 Lishui Road West, Shenzhen, China 518055

[‡]Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China.

[§]Institute of Nuclear Physics and Chemistry, China Academy of Engineering Physics, 64 Mianshan Road, Mianyang, Sichuan, China 621900

Email: xiaoyuli@hku.hk

ABSTRACT: DNA-encoded chemical library (DEL) has recently emerged and become an important technology platform in biomedical research and drug discovery. DELs containing large numbers of compounds can be prepared and selected against biological targets to discover novel ligands and inhibitors. In practice, DELs are usually selected against purified and immobilized proteins using the typical "bind-wash-elute" protocol; however, selection methods compatible with non-immobilized proteins would be able to greatly expand the target scope of DELs beyond purified proteins to more complex and biologically relevant targets. Using a novel "*ligate-crosslink-purify*" strategy, here we report a method capable of selecting DELs against unmodified and non-immobilized protein targets. In addition, this method has shown excellent capability in identify binders with moderate and weak affinities.

INTRODUCTION

Encoded combinational libraries are powerful discovery tools in biomedical research. Large scale peptide, antibody, protein, and nucleic acid libraries can be routinely built with platforms such as phage display,[1-4] mRNA display,[5, 6] yeast surface display,[7-9] ribosome display,[10-12] SELEX,[13-17] etc. Recently, DNA-encoded chemical library (DEL) has emerged and become an important technology platform for ligand and inhibitor discovery in biomedical research.[18-48] In a DEL, each compound is covalently conjugated with a unique DNA tag, serving as the identifier for the chemical structure of the compound. Therefore, the entire library can be synthesized and processed simultaneously at minute scale. The original concept was proposed and explored by Brenner and Lerner[18-19] and also Gallop[20]. During the next two decades, this field has seen significant developments on library synthesis, selection and deconvolution. Today, DELs containing hundreds of millions to many billions of compounds are routinely prepared and screened against various biological targets.[21, 32, 34-36, 39] Moreover, DEL has also been widely adopted by research institutions and pharmaceutical companies in numerous drug discovery programs;[22-29, 34, 39-43, 49-63] it is not unreasonable to foresee DEL to become a major technology platform complementing traditional HTS in near future.[32, 35, 42, 64]

Typically, selections of DELs are achieved with immobilized protein targets based on binding affinity using the typical "*bind-wash-elute*" protocol: 1) the library is incubated with a matrix-immobilized target, usually a purified protein; 2) non-binding library compounds are removed by multiple washing steps; 3) bound compounds are eluted under strong denaturing conditions and then subjected to PCR amplification and DNA sequencing for hit deconvolution.[65] However, not all proteins are suitable for purification and immobilization, and many immobilized proteins often display altered properties from their soluble form or even a complete loss of activities.[66-70] Therefore, selection methods compatible with non-immobilized proteins would be able to extend the target scope of DEL to more complex biological targets such as protein complexes,[71] membrane proteins,[52, 72-74] and even live cells.[75-77] In addition, selection with unpurified endogenous proteins can also preserve some desirable features such as post-translation modifications and endogenous binding partners.[78]

In a pioneering work, Gallop and co-workers synthesized a heptapeptide DEL immobilized on beads so that soluble protein targets can be directly selected;[20] more recently, Paegel and co-workers have built large-scale peptide DELs on solid phase but with more sophisticated multi-functional linkers.[79-81] Moreover, a variety of strategies have also been developed to select solution-phase DELs against non-immobilized proteins. Liu and co-workers developed methods capable of targeting endogenous proteins in cell lysates using the interaction-dependent PCR.[78, 82, 83] Vipergen reported a Binder Trap Enrichment method based on water-oil emulsion.[27, 84] Biophysical method such as kinetic

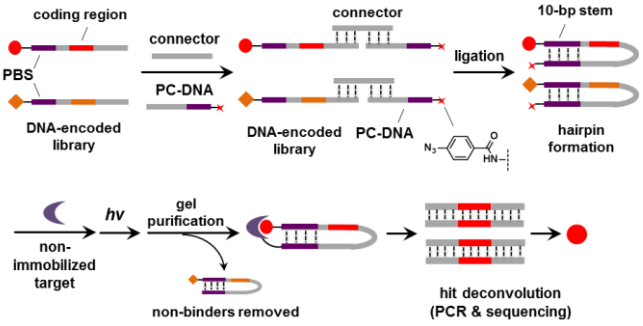


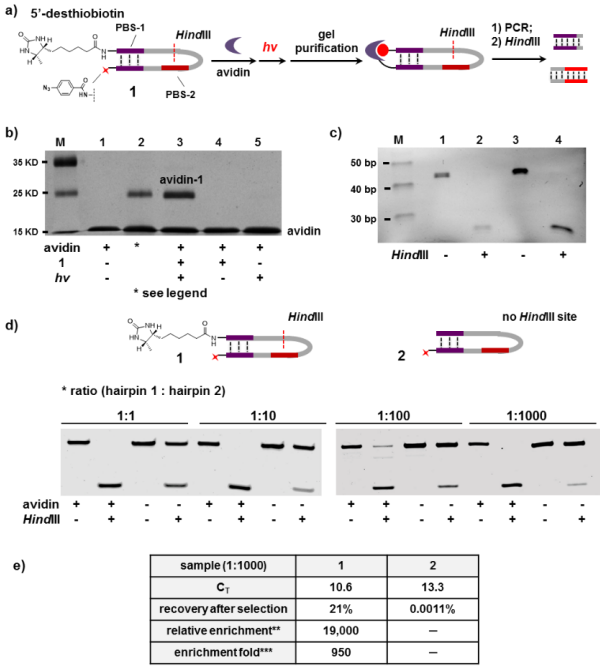
Figure 1. The proposed "*ligate-crosslink-purify*" method: library DNA is ligated with a PC-DNA (photocrosslinker DNA) at the end distal to the small molecule. The PC-DNA has a short sequence complementary to the PBS (primer-binding site), therefore a hairpin structure is formed. After target binding and light irradiation, the target-DNA conjugates formed are gel-purified and separated from non-binders for hit identification by PCR amplification and DNA sequencing.

capillary electrophoresis (KCE) has also been used to partition ligands from non-binders.[85, 86] Recently, Krusemark and co-workers reported a photo-crosslinking-based selection method.[87] In addition, libraries of PNA-encoded small molecule libraries equipped with chemically reactive groups have been used to profiling endogenous enzymes in lysates.[88-91]

Previously we reported a method compatible with unmodified and non-immobilized targets using a terminal protection strategy.[92] In this method, target-binding ligands are crosslinked to the protein, and then exonuclease I (*Exo*I) is used to digest the DNA tags of non-binders while the tags of bound ligands are protected by the target protein. However, this method is limited to DELs with small molecules conjugated at the 3'-end of the library DNA and the use of exonuclease leads to over-digestion of the DNA tags of binding ligands.[92, 93] In order to address these issues, here we report the development and application of a novel selection method that solves these problems and can be directly applied to existing DELs without library redesign or resynthesis.

RESULTS AND DISCUSSION

Our proposed method is based on a "*ligate-crosslink-purify*" strategy. As shown in Figure 1, for a given DEL, a short DNA strand (PC-DNA: photocrosslinker-DNA) can be enzymatically ligated to the library at the end distal to the small molecule. The PC-DNA contains a sequence complementary to the primer binding site (PBS) of the library DNA, thereby forming a hairpin structure and bringing the photo-crosslinker close to the small molecule. The hairpin-shaped library can then be selected against non-immobilized target proteins. After light irradiation, importantly, for binders, a covalent "protein-DNA" conjugate is formed and can be purified and separated from non-binders by electrophoresis, before typical PCR amplification and DNA sequencing are applied for hit deconvolution. For non-binders, they are not crosslinked the target protein and are removed during gel purification



sample (1:1000)	1	2
C_T	10.6	13.3
recovery after selection	21%	0.0011%
relative enrichment**	19,000	–
enrichment fold***	950	–

Figure 2. a) Selection scheme of hairpin **1** with avidin. Conditions: **1**, 2.5 μM; avidin: 2.5 μM; buffer: 0.1 M NaCl, 1x PBS; irradiation: 365 nm for 30 sec. at 0 °C; b) Electrophoresis analysis of the photo-crosslinking reaction (15% SDS-PAGE). M: marker; lane 1 and lane 3-4: reactions are as marked; lane 2: a mixture of avidin and avidin-**1** standard sample, prepared by direct conjugation.[94] c) Denaturing PAGE (15% TBE-Urea) analysis of the PCR products. M: marker; lane 1: PCR products from the avidin-**1** conjugate (lane 2 in b); lane 2: after *Hind*III digestion; lane 3-4: same as lane 1-2 but using a 42-nt positive control DNA with a *Hind*III digestion site. d) Mixtures of **1** and **2** were selected against avidin and digested with the same procedure as in a). PCR products were analyzed with 15% PAGE. e) qPCR analysis of DNA recovery and enrichment in the 1:1000 sample in d). **: (post-selection **1**/**2** ratio)/(pre-selection **1**/**2** ratio). ***: (post-selection percentage of **1**)/(pre-selection percentage of **1**). See Supporting Information and Table S1 for details.

due to the relatively lower molecular weight (Figure 1). Therefore, the main purpose of the ligation step is to establish a stable covalent link between the target protein and the DNA codes of the ligands selected by the target, so that the target-DNA conjugates can be reliably separated from non-binders by gel purification and subjected to further sample processing under various experimental conditions. Moreover, this method avoids the drawbacks of nuclease digestion and can be applied to DELs with either 3'- or 5'-conjugated small molecules.

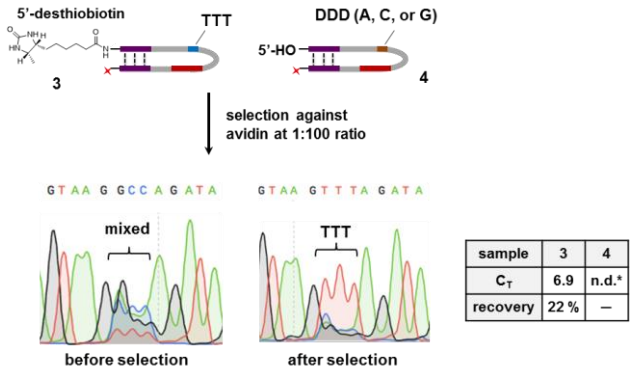
As an initial test, we prepared a 42-nt hairpin DNA (**1**; Figure 2a), modified with a 5'-desthiobiotin and 3'-phenylazide crosslinker. **1** forms a stable hairpin structure with a 10-bp stem (Tm: 54.5 °C in 0.1 M NaCl; Figure S2); it also contains a digestion site specific for restriction enzyme *Hind*III (5'-AAGCTT-3') in the loop. Desthiobiotin is a high affinity ligand for avidin with a K_d of ~2.0 nM.[95] **1** was incubated with avidin (2.5 μM), irradiated under 365 nm for 30 seconds, and then analyzed by electrophoresis. As shown

in Figure 2b, a slower moving band matching the molecular weight of the avidin-**1** conjugate was detected (lane 3). No such band was observed without desthiobiotin or light (lane 4 and 5), proving the formation of the conjugate is specific. The band was excised, extracted, PCR-amplified, and then digested by *Hind*III, and it is shown that the PCR products can be completely digested, confirming the band is the expected avidin-**1** conjugate (lane 1 and 2; Figure 2c).

Next, we mixed **1** with a control DNA without desthiobiotin (**2**; Figure 2d) at various ratios. **2** does not have the *Hind*III site, so that *Hind*III digestion of the PCR products can reveal the ratio of **1** and **2** in the mixture.[82, 96] Mixtures of **1** and **2** were subjected to the same selection, purification, PCR amplification, and digestion procedures as in Figure 2a. It is worth noting that a relatively large range of gel band was excised to ensure the recovery of the protein-DNA conjugate (Figure S3), which is highly important for large libraries where the shifted gel band may not clearly visible. In addition, we have verified by qPCR and Sanger sequencing that excision of a larger gel band did not significantly increase background signals (Figure S4 and Table S2).

As shown in Figure 2d, with avidin, the PCR products of the gel-purified avidin-DNA conjugates were efficiently digested by *Hind*III, even at the ratio of 1:1000, indicating significant enrichment of **1**. Without avidin, the digestion products roughly reflected the original mixing ratio of **1** and **2**. In corroboration, qPCR results showed that nearly full enrichment was obtained for **1** in the 1:1000 sample (see the Supporting Information and Table S1 for details). In addition, qPCR quantitation also showed that the DNA recovery efficiency from gel purification is comparable to similar washing-based methods (Table S1).[87]

In order to further validate that the enriched DNA indeed encodes the small molecule ligand, two 42-nt hairpin DNAs (**3** and **4**; Figure 3a) were prepared. In **3**, the desthiobiotin is encoded by a "TTT" codon in the loop,

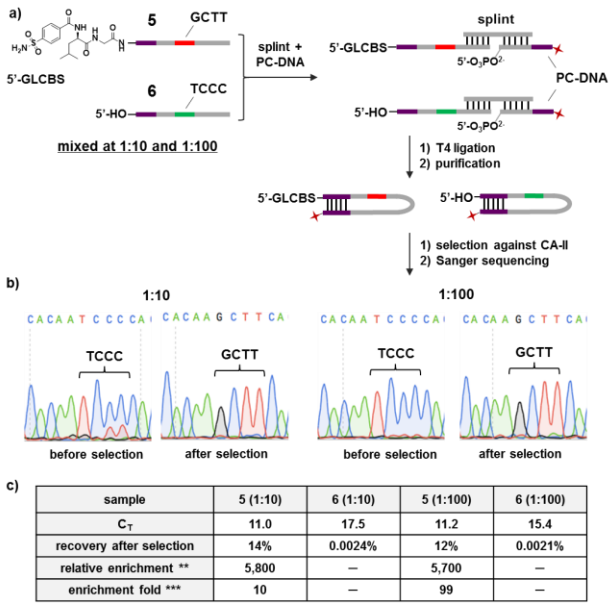


sample	3	4
C_T	6.9	n.d.*
recovery	22 %	–

Figure 3. a) A 1:100 mixture of **3** and **4** was selected against avidin with the same procedure as in Figure 2, except Sanger sequencing was used to analyze the pre-selection mixtures and gel-purified conjugates. b) Sequencing results before (left panel) and after (right panel) the selection. Encoding sites are highlighted. Side table: qPCR analysis results. n.d.*: not determined due to mixed sequences in **4**. See the Supporting Information for details (Figure S6 and Table S1).

while the negative control **4** has a mixed sequence (A, C or G) at the same site. **3** and **4** were mixed at 1:100 and then selected against avidin following the same procedure as in Figure 2. The gel-purified conjugates were analyzed by Sanger sequencing and compared with pre-selection mixtures. As shown in Figure 3b, before selection, as expected, the encoding site had a mixed sequence, while after selection, the "TTT" codon encoding desthiobiotin has been distinctly enriched, proving the capability of this method in identifying specific binders from a large excess of non-binding background. In addition, qPCR analysis showed a similar recovery of **3** to the experiment in Figure 2d (Figure 3, side table; see details in Table S1).

After these preliminary studies, next we tested the "*ligate-crosslink-purify*" strategy to introduce the hairpin structure to a DEL encoded with linear DNA tags. We prepared a 49-nt DNA modified with 5'-GLCBS, a sulfonamide ligand specific for carbonic anhydrase-II (CA-II) (**5**, K_d = 9 nM; Figure 4a);[94] The GLCBS is encoded by a "GCTT" codon. In addition, a hairpin DNA without the ligand of GLCBS but with a "TCCC" codon was also prepared as the negative control (**6**; Figure 4a). Mixtures of **5** and **6** (1:10 and 1:100) were ligated with a 20-nt PC-DNA bearing a 3'-phenylazide and a 5'-phosphate group, respectively, mediated by T4 DNA ligase and a 14- nt splint DNA (Figure S5). The 5'-end of the PC-DNA is complementary to the 5'-PBS of **5** so that a stable hairpin



c)

sample	5 (1:10)	6 (1:10)	5 (1:100)	6 (1:100)
C_T	11.0	17.5	11.2	15.4
recovery after selection	14%	0.0024%	12%	0.0021%
relative enrichment **	5,800	–	5,700	–
enrichment fold ***	10	–	99	–

Figure 4. a) Mixtures of **5** and **6** (1:10 and 1:100) were converted to hairpins via enzymatic ligation with a PC-DNA. Conditions: **5** and **6**, 1.0 μM each; T4 DNA ligase: 10 units in T4 ligase buffer; ligation: 1 hour at 16 °C. Ligation products were gel-purified and selected against CA-II as shown in Figure 1. b) Sanger sequencing results of the pre-selection mixtures and gel-purified CA-II-DNA conjugates. Coding sites are highlighted. See full images in Figure S7 and S8. c) qPCR analysis of DNA recovery and enrichment. **: (post-selection **5**/**6** ratio)/(pre-selection **5**/**6** ratio). ***: (post-selection percentage of **5**)/(pre-selection percentage of **5**). See Table S1 for details.

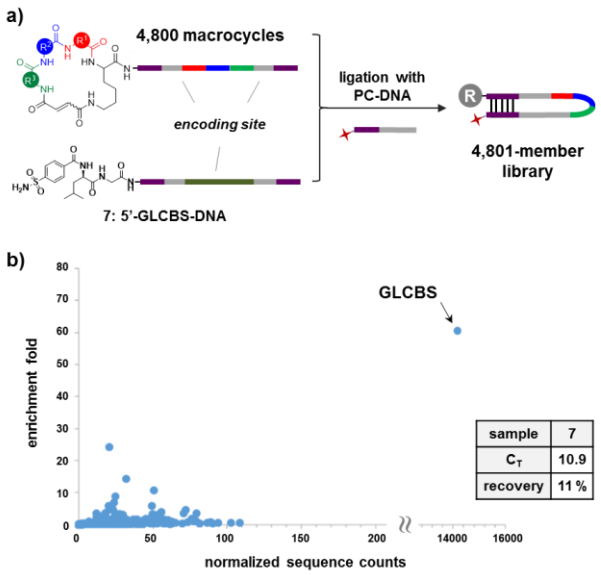


Figure 5. a) A 4,800-member DNA-encoded macrocycle library and a positive control **7** were mixed and ligated with a PC-DNA. The ligated library was selected against CA-II with the same procedure as in Figure 4, except high throughput sequencing was used for hit identification. b) Selection results shown in a scatter plot: enrichment folds (y-axis) vs. post-selection sequence counts (x-axis). Enrichment fold = (post-selection fraction)/(pre-selection fraction).[92, 97] See the Supplementary Information for more details on experimental conditions, macrocycle structures, library DNA sequences, sequencing, and data analysis. Side table: qPCR analysis results for the positive control **7**. Other DNAs were not determined. See the Supporting Information for details.

structure can be formed after ligation (Figure 4a; Tm: 52.7 °C, see details in Figure S2). These "hairpin libraries" were then selected against CA-II; crosslinked target-DNA conjugates were gel-purified to be separated from non-binders; then selected compounds were deconvoluted by Sanger sequencing. Again, before selection, only the dominant "TCCC" codon can be observed; after selection, clear enrichment of the "GCTT" codon encoding the ligand of GLCBS was observed in both experiments (Figure 4b). In addition, qPCR analysis also showed nearly full enrichment of **5** in both cases and similar efficiency of DNA recovery (Figure 4c). This result has demonstrated that the "*ligate-crosslink-purify*" strategy can be straightforwardly applied to DELs by simply ligating a PC-DNA to enable formation of the hairpin structure.

Finally, we tested our method with a DEL containing actual chemical diversity. We prepared a 4,800-member DNA-encoded macrocycle library in which each macrocycle is encoded by a 49-nt DNA tag as previously described.[92, 98] This library was spiked with a DNA-tagged GLCBS (**7**) as an internal positive control (Figure 5a). This 4,801-member library was ligated with a 20-nt PC-DNA to enable hairpin formation, selected against target CA-II, irradiated under 365 nm, and then subjected to gel purification to isolate the CA-II-DNA conjugates. First, qPCR analysis showed a 11% recovery of the positive

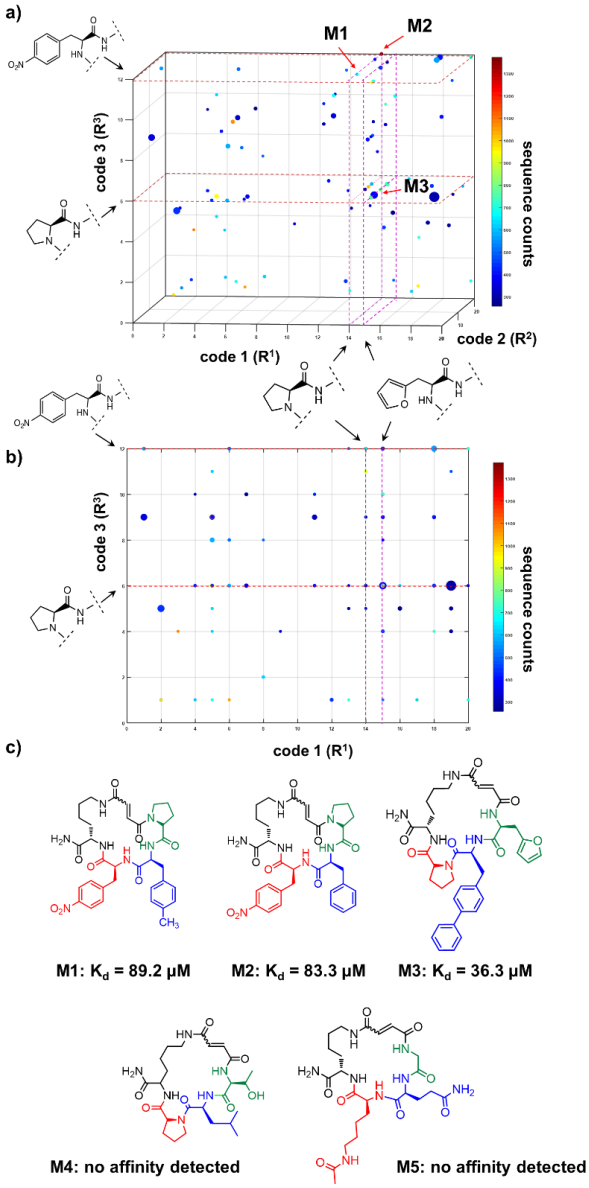


Figure 6. The 4,800-member DNA-encoded macrocycle library was ligated with a 20-nt PC-DNA and then selected against avidin with the same procedure as described in Figure 5. a) Selection results are shown in a 3D scatter plot. Each axis represents a set of building blocks. Size of data points: enrichment folds (larger dot size indicates higher enrichment fold; cut-off: 3.5); color of data points: post-selection sequence counts (cut-off: 250). b) 2D projection of the scatter plot in a) along the y-axis to show the SAR pattern more clearly; c) Structures of three selected "hit" macrocycles (**M1**-**M3**) and two negative control macrocycles (not selected from the selection; **M4** and **M5**). SPR sensorgrams are shown in Figure S9.

control **7**, similar to model libraries (Figure 4c) without chemical diversity (Figure 5b, side table). Next, hit compounds were deconvoluted with high throughput DNA sequencing and the results are shown in form of a scatter plot (Figure 5b). Compounds with relatively high enrichment folds and post-selection sequencing counts are considered potential hits (data points leaning towards the

upper right direction of the plot).[92, 97, 99] As shown in Figure 5b, the positive control **7** was distinctly identified, proving the capability of this method in enriching high affinity binders from chemically diverse libraries. The enrichment fold of **7** appeared to be significantly lower than the positive control **2** in the model system (Figure 2d and 2e), presumably due to the relatively lower affinity of GLCBS and the presence of macrocycle compounds who may also bind and be selected by the target CA-II.

In chemical library selections, moderate and low affinity binders are also valuable since they can be used as starting points for downstream optimization to obtain high quality lead compounds. However, often they are difficult to be identified due to repeated washing steps. Photo-crosslinking, which establishes a stable covalent bond between the target and ligand, may provide opportunities for moderate and weak binders to emerge, such as being demonstrated by the Krusemark group in a recent report.[87] In order to test this and also examine the generality of our method, we selected the same 4,800-member macrocycle library against avidin as the target. We reason that, due to the lack of known avidin-binding structural features in the library,[100] binders with moderate and weak affinities may be identified. After the macrocycle library was selected against non-immobilized avidin and hit deconvolution by high throughput sequencing, we plotted the selection data into a 3D scatter plot (Figure 6a), in which each of the axe represents one respective set of amino acid building blocks in the macrocycle structure (R^1, R^2, and R^3 in Figure 5a). This type of 3D plot enables easy identification of structure-activity relationship (SAR) by observing the "planes" and "lines" formed by selection data points.[49, 52] As shown in Figure 6a, in the avidin selection, several "planes" can be identified, suggesting that these specific building blocks bind more favorably to the target avidin. This pattern can be seen more clearly with the 2D projection along the R^2-axis (Figure 6b).

Next, three macrocycles that are located at the intersections of several "planes" with good enrichment folds or sequence counts were selected (**M1**, **M2**, and **M3**; Figure 6c). These "hits" were resynthesized without the DNA tag using a reported method[97] and then assayed for their binding affinity with surface plasmon resonance (SPR). Results show that they are, as expected, weak binders for avidin with K_d ranging from 36 to 89 μM (Figure 6c and Figure S8). As a specificity control, we have also synthesized and assayed two additional macrocycles (**M4** and **M5**; Figure 6c) that are not selected by avidin; indeed, their binding affinities with avidin are not detectable with SPR (Figure S8). Collectively, the above results have demonstrated the generality and the capability of this selection method in identifying novel binders with relatively weak affinity from chemically diverse DELs.

CONCLUSION

In conclusion, we have developed a novel selection method for DNA-encoded libraries. This method establishes a stable covalent link between the target and the DNA tags of the selected ligands, thereby obviating the limitations arising from exonuclease digestion and also enabling the reliable separation of binders from non-binders by the critical gel purification step.[93] Importantly, this method is suitable for DNA-encoded libraries with small molecules conjugated at either end of encoding DNA strands. Moreover, by ligating a short PC-DNA strand and a subsequent gel purification step, the "*ligate-crosslink-purify*" strategy enables the selection of existing DELs, which may be prepared and encoded with other methods, without the need for library redesign and resynthesis. However, it is worth noting that the formation of the stem-loop structure with variable sequences in the single-stranded region may cause aptamer formation that can be selected by the target; therefore, careful codon design and a negative selection with a "blank library" having the same DNA sequences but without small molecules is necessary to control for selection of aptamers. Since all DELs already have built-in primer binding sites available, we expect wide applications of this method in the ligand discovery against targets previously intractable to DELs.

In principle, this method may also be applicable to selections against targets in cell lysates; however, method improvements to better locate the desired protein-DNA conjugates during gel purification (e.g. using a fluorescently labeled PC-DNA) and to differentiate ligand binding by the target from background proteins (e.g. using negative control selections or cell lysates with overexpressed target[77, 92]) would be instrumental. Finally, this method cannot be directly applied to DELs encoded with double-stranded DNAs (dsDNAs) due to the inability of DNA duplexes to form stem-loop structures; converting dsDNAs to single-stranded DNAs using methods such as biotin-tagging and affinity separation,[101] strand displacement, and Lambda exonuclease digestion,[102, 103] may be explored. We will report our progresses on these two directions in due course.

EXPERIMENTAL PROCEDURES

Syntheses of small-molecule-conjugated DNAs

5'-desthiobiotin labeling. Solid-phase DNA synthesis was performed with the 3'-amino-modifier CPG and 5'-amino-C6 modifier phosphoramidite. 5'-MMT was deprotected with 3% TCA in CH_2Cl_2; the beads were then washed with CH_2Cl_2 and then dried under vacuum. Desthiobiotin (8.7 mg, 40 mmol) was dissolved in 200 μL anhydrous DMF along with 0.9 equiv. HBTU (13.7 mg, 36 mmol) and 1 equiv. HOBt (5.4 mg, 40 mmol). After vortexing for 1 hour at room temperature, the activation mixture was added to the CPG along with 2.3 equiv. DIPEA (15.2 μL, 92 mmol). The suspension was then incubated at 37 °C overnight with agitation. The CPG was washed with DMF (3x 600 μL), CH_3CN (3x 600 μL), and then dried with gentle airstream. The 5'-desthiobiotin-labeled DNA was cleaved with AMA (50:50; 40% aqueous ammonium hydroxide : methylamine, v:v; 55 °C, 55 min) over a dry bath and then purified by HPLC.

4-azidobenzoate-SE synthesis. 4-azidobenzoic acid (76.7 mg, 0.47 mmol) and NHS (54 mg, 0.47 mmol) were dissolved in 1.5 mL anhydrous DMF. *N,N'*-dicyclohexylcarbodiimide (115 mg, 0.56 mmol) was added to the solution

at room temperature. After the mixture was stirred for 8 hours, precipitated urea was removed by filtration. The filtrate was dried by a rotary evaporator. After washing briefly with diethyl ether, the obtained white powder was used directly in DNA labeling.

3'-phenylazide labeling. The 3'-amine, 5'-desthiobiotin-modified DNA was dissolved in 40 μL high purity water and 40 μL phosphate buffer (pH = 7.2). 4-azidobenzoate-SE (3.1 mg, 12 μmol) was dissolved in 40 μL DMSO (final concentration: 0.3 M). The two solutions were mixed and the reaction was incubated with sonication at 55 °C for 1 hour. After brief centrifugation, the reaction mixture was desalted by a NAP-5 column (GE Pharmacia) before the product was purified by HPLC.

5'-GLCBS labeling. GLCBS (4.5 mg, 12 μmol), DCC (2.5 mg, 12 μmol) and NHS (1.4 mg, 12 μmol) were dissolved in 40 μL anhydrous DMSO and then the solution was vortexed for 1 hour at room temperature. The formed urea precipitate was removed by filtration. The filtrate was added to a solution containing 40 μL 5'-amine-modified DNA (20 nmol) and 40 μL phosphate buffer (pH 7.5). The mixture was the maintained with sonication at 37 °C for 1 hour. After centrifugation, the reaction mixture was desalted by a NAP-5 column (GE Pharmacia) before the product was purified by HPLC.

PC-DNA. The synthesis and purification procedure is the same as described above in 3'-phenylazide labeling except that the 5'-phosphate-, 3'-amine-modified DNA was used to conjugate with 4-azidobenzoate-SE.

See the supporting information for MS characterization of DNA-small-molecule conjugates.

DNA sequences

1: 5'-desthiobiotin-CTG AAT TCT CCC GAA AGT AAG CTT AGA TAG GCG AGA ATT CAG-azide-3';

2: 5'-CTG AAT TCT CCC GAA AGT GAA TCC AGA TAG GCG AGA ATT CAG-azide-3';

3: 5'-desthiobiotin-CTG AAT TCT CCC GAA AGT AAG TTT AGA TAG GCG AGA ATT CAG-azide-3'

4: 5'-CTG AAT TCT CCC GAA AGT AAG DDD AGA TAG GCG AGA ATT CAG-azide-3' (D = A, C or G)

5: 5'-GLCBS-CCT GAA TTC CAA AGC CCT CAC AAG CTT CAC AAG ACC CTC CCA AAC TGC C-3'

6: 5'-CCT GAA TTC CAA AGC CCT CAC AAT CCC CAC AAG ACC CTC CCA AAC TGC C-3'

7: 5'-GLCBS-CCT GAA TTC CAA AGC CCT CAC AAT CCC CAC ATC CGT CTC CCA AAC TGC C-3'

Splint-DNA: 5'-TTC GGG AGG CAG TT-3'

PC-DNA: 5'- phosphate-TCC CGA AAG TGG AAT TCA GG-azide -3'

See the supporting information for DNA sequences of PCR primers and high throughput sequencing adapters.

Library selection and selection result deconvolution

Selection. In a typical experiment, 25 pmol of DNA or DNA-encoded library, 1 μL 10x PBS buffer, and 1 μL 1 M NaCl were added to a 200 μL centrifuge tube, before 25 pmol of target protein was added. The mixture was supplemented with water to a final volume of 10 μL, incubated at 4 °C for 1 hour, and then subjected to light irradiation under 365 nm at 0 °C (over ice) for 30 seconds.

PAGE analysis and gel extraction. After photo-crosslinking, 2 μL 5x loading buffer was added to the mixture, which was then heat-denatured at 95 °C for 10 minutes over a dry bath before PAGE analysis (12%-15%). Product bands were excised, extracted by 1x PBS buffer, and purified by ethanol precipitation.

PCR analysis. After ethanol precipitation, the pellet obtained was dissolved with 50 μL water. 10 μL of the solution was added with 5 μL 5 μM PCR primers, 10 μL 5x Phusion HF Reaction Buffer, 1 μL 1 mM Deoxynucleotide Solution Mix and 0.5 μL Phusion HF DNA Polymerase. The solution was then again diluted to 50 μL before PCR amplification. The PCR product was purified by ethanol precipitation.

***Hind*III digestion.** After ethanol precipitation, the pellet obtained was dissolved with 120 μL water. 60 μL of the solution was transferred to a 600 μL centrifuge tube; then 8 μL 10x *Hind*III buffer, 1 μL *Hind*III (20 units, 20 units/μL), and 11 μL water were added to reach a final volume of 80 μL. The reaction mixture was incubated at 37 °C for 2 hours. After digestion, the proteins were heat-denatured at 95 °C for 20 minutes over a dry bath and the DNA products were purified by ethanol precipitation.

T4 DNA ligase-mediated ligation. 200 pmol DNA to be ligated, 200 pmol PC-DNA, and 200 pmol splint DNA were added to a 1.5 mL centrifuge tube, heated to 85 °C for 30 seconds, and cooled to room temperature prior to the addition of 20 μL 10x T4 DNA ligase buffer. 1 μL (350 units) of T4 DNA ligase was added to the reaction solution at 0 °C. Ligation reactions were maintained at 16 °C for 2 hours, before subjected to ethanol precipitation and denaturing PAGE analysis. Product bands were excised, extracted by 1x TE buffer, and purified by ethanol precipitation.

Synthesis, DNA sequences and building block structures of the macrocycle library. the macrocycle library used in this study is the same as the macrocycle library we described previously.[92]

High-throughput sequencing and data analysis. After PCR amplification of the selected library, PCR products were recovered by gel extraction with a gel extraction kit and then directly submitted for high-throughput sequencing. Sequencing experiments were performed on an Illumina® NextSeq 500 sequencer using standard 2 x 75 pair-end sequencing reagent kits and hybridization primers. After sequencing, data were exported for processing. Sequence counts for each library member before and after the selection were tallied to calculate the enrichment fold for each library compound, following the method of previous reports.[49, 92, 97, 104] First, the number of encoding DNA sequence reads for each library compound was counted. Second, the number was divided by the total number of interpretable sequence reads to give the abundance of that library member. Finally, enrichment folds for each library member were obtained by dividing the post-selection

abundance by the pre-selection abundance. Enrichment folds were then plotted against the sequence number. In Figure 6, processed data was plotted into a 3D cube using MATLAB software.

Hit Compound Resynthesis. Selected macrocycles (**M1-M5**) were synthesized without the DNA tag following a reported method.[98] See the supporting information for characterization data of these compounds.

ASSOCIATED CONTENT

Supporting Information. More details of selection procedures, structure of DNAs, characterization data, high-throughput sequencing procedure, full Sanger sequencing data, qPCR analysis and calculation methods, SPR sensorgrams, and other experimental details. The Supporting Information is available free of charge on the ACS Publications website at XXX

AUTHOR INFORMATION

Corresponding Author

* Email: xiaoyuli@hku.hk

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This work was supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No.: AoE/P-705/16 and 106160203), seed funding for basic research (201601159018) and start-up fund from The University of Hong Kong, and a grant from National Natural Science Foundation of China (NSFC, 21572014). The authors thank Dr. Yan Cao at The 2nd Military Medical University, Shanghai, China for assistance on SPR analysis.

REFERENCES

(1) McCafferty, J., Griffiths, A. D., Winter, G., and Chiswell, D. J. (1990) Phage antibodies: filamentous phage displaying antibody variable domains. *Nature 348*, 552-554.

(2) Barbas, C. F., Kang, A. S., Lerner, R. A., and Benkovic, S. J. (1991) Assembly of Combinatorial Antibody Libraries on Phage Surfaces - the Gene-Iii Site. *Proc. Nat. Acad. Sci. USA 88*, 7978-7982.

(3) Kang, A. S., Barbas, C. F., Janda, K. D., Benkovic, S. J., and Lerner, R. A. (1991) Linkage of Recognition and Replication Functions by Assembling Combinatorial Antibody Fab Libraries Along Phage Surfaces. *Proc. Nat. Acad. Sci. USA 88*, 4363-4366.

(4) Lerner, R. A. (2016) Combinatorial antibody libraries: new advances, new immunological insights. *Nat. Rev. Immunol. 16*, 498-508.

(5) Wilson, D. S., Keefe, A. D., and Szostak, J. W. (2001) The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Nat. Acad. Sci. USA 98*, 3750-3755.

(6) Passioura, T., and Suga, H. (2017) A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets. *Chem. Commun. 53*, 1931-1940.

(7) Pepper, L. R., Cho, Y. K., Boder, E. T., and Shusta, E. V. (2008) A decade of yeast surface display technology: Where are we now? *Combinatorial Chem. High Throughput Screening 11*, 127-134.

(8) Boder, E. T., and Wittrup, K. D. (1997) Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol. 15*, 553-557.

(9) Cherf, G. M., and Cochran, J. R. (2015) Applications of Yeast Surface Display for Protein Engineering. *Methods Mol. Biol. 1319*, 155-175.

(10) Hanes, J., and Pluckthun, A. (1997) In vitro selection and evolution of functional proteins by using ribosome display. *Proc. Nat. Acad. Sci. USA 94*, 4937-4942.

(11) Pluckthun, A. (2012) Ribosome display: a perspective. *Methods Mol. Biol. 805*, 3-28.

(12) Zahnd, C., Amstutz, P., and Pluckthun, A. (2007) Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. *Nat. Methods 4*, 269-279.

(13) Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science 249*, 505-510.

(14) Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature 346*, 818-822.

(15) Morris, K. N., Jensen, K. B., Julin, C. M., Weil, M., and Gold, L. (1998) High affinity ligands from in vitro selection: complex targets. *Proc. Nat. Acad. Sci. USA 95*, 2902-2907.

(16) Shangguan, D., Li, Y., Tang, Z. W., Cao, Z. H. C., Chen, H. W., Mallikaratchy, P., Sefah, K., Yang, C. Y. J., and Tan, W. H. (2006) Aptamers evolved from live cells as effective molecular probes for cancer study. *Proc. Nat. Acad. Sci. USA 103*, 11838-11843.

(17) Fang, X. H., and Tan, W. H. (2010) Aptamers Generated from Cell-SELEX for Molecular Medicine: A Chemical Biology Approach. *Acc. Chem. Res. 43*, 48-57.

(18) Brenner, S., and Lerner, R. A. (1992) Encoded Combinatorial Chemistry. *Proc. Nat. Acad. Sci. USA 89*, 5381-5383.

(19) Nielsen, J., Brenner, S., Janda, K.D. (1993) Synthetic Methods for the Implementation of Encoded Combinatorial Chemistry. *J. Am. Chem. Soc. 115*, 9812-9813.

(20) Needels, M. C., Jones, D. G., Tate, E. H., Heinkel, G. L., Kochersperger, L. M., Dower, W. J., Barrett, R. W., and Gallop, M. A. (1993) Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc. Nat. Acad. Sci. USA 90*, 10700-10704.

(21) Kleiner, R. E., Dumelin, C. E., and Liu, D. R. (2011) Small-molecule discovery from DNA-encoded chemical libraries. *Chem. Soc. Rev. 40*, 5707-5717.

(22) Krall, N., Scheuermann, J., and Neri, D. (2013) Small targeted cytotoxics: current state and promises from DNA-encoded chemical libraries. *Angew. Chem. Int. Ed. 52*, 1384-1402.

(23) Franzini, R. M., Neri, D., and Scheuermann, J. (2014) DNA-encoded chemical libraries: advancing beyond conventional small-molecule libraries. *Acc. Chem. Res. 47*, 1247-1255.

(24) Scheuermann, J., and Neri, D. (2015) Dual-pharmacophore DNA-encoded chemical libraries. *Curr. Opin. Chem. Biol. 26*, 99-103.

(25) Keefe, A. D., Clark, M. A., Hupp, C. D., Litovchick, A., and Zhang, Y. (2015) Chemical ligation methods for the tagging of DNA-encoded chemical libraries. *Curr. Opin. Chem. Biol. 26*, 80-88.

(26) Samain, F., and Casi, G. (2015) Small targeted cytotoxics from DNA-encoded chemical libraries. *Curr. Opin. Chem. Biol. 26*, 72-79.

(27) Blakskjaer, P., Heitner, T., and Hansen, N. J. (2015) Fidelity by design: Yoctoreactor and binder trap enrichment for small-molecule DNA-encoded libraries and drug discovery. *Curr. Opin. Chem. Biol. 26*, 62-71.

(28) Chan, A. I., McGregor, L. M., and Liu, D. R. (2015) Novel selection methods for DNA-encoded chemical libraries. *Curr. Opin. Chem. Biol. 26*, 55-61.

(29) Connors, W. H., Hale, S. P., and Terrett, N. K. (2015) DNA-encoded chemical libraries of macrocycles. *Curr. Opin. Chem. Biol. 26*, 42-47.

(30) Li, G., Zheng, W., Liu, Y., and Li, X. (2015) Novel encoding methods for DNA-templated chemical libraries. *Curr. Opin. Chem. Biol. 26*, 25-33.

(31) Zambaldo, C., Barluenga, S., and Winssinger, N. (2015) PNA-encoded chemical libraries. *Curr. Opin. Chem. Biol. 26*, 8-15.

(32) Gura, T. (2015) DNA helps build molecular libraries for drug testing. *Science 350*, 1139-1140.

(33) Salamon, H., Klika Skopic, M., Jung, K., Bugain, O., and Brunschweiger, A. (2016) Chemical Biology Probes from Advanced DNA-encoded Libraries. *ACS Chem. Biol. 11*, 296-307.

(34) Zimmermann, G., and Neri, D. (2016) DNA-encoded chemical libraries: foundations and applications in lead discovery. *Drug Discov. Today*, *21,* 1828-1834.

(35) Mullard, A. (2016) DNA tags help the hunt for drugs. *Nature 530*, 367-369.

(36) Franzini, R. M., and Randolph, C. (2016) Chemical Space of DNA-Encoded Libraries. *J. Med. Chem. 59*, 6629-6644.

(37) Yuen, L. H., and Franzini, R. M. (2016) Achievements, Challenges, and Opportunities in DNA-Encoded Library Research: An Academic Point of View. *ChemBioChem 18*, 829-836.

(38) Malone, M. L., and Paegel, B. M. (2016) What is a "DNA-Compatible" Reaction? *ACS Comb. Sci. 18*, 182-187.

(39) Eidam, O., and Satz, A. L. (2016) Analysis of the productivity of DNA encoded libraries. *MedChemComm 7*, 1323-1331.

(40) Arico-Muendel, C. C. (2016) From haystack to needle: finding value with DNA encoded library technology at GSK. *Medchemcomm 7*, 1898-1909.

(41) Goodnow, R. A., Jr., Dumelin, C. E., and Keefe, A. D. (2017) DNA-encoded chemistry: enabling the deeper sampling of chemical space. *Nat. Rev. Drug Discov. 16*, 131-147.

(42) Lerner, R. A., and Brenner, S. (2017) DNA-Encoded Compound Libraries as Open Source: A Powerful Pathway to NewDrugs. *Angew. Chem. Int. Ed. 56*, 1164-1165.

(43) Shi, B., Zhou, Y., Huang, Y., Zhang, J., and Li, X. (2017) Recent advances on the encoding and selection methods of DNA-encoded chemical library. *Bioorg. Med. Chem. Lett. 27*, 361-369.

(44) Halpin, D. R., and Harbury, P. B. (2004) DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biol. 2*, 1015-1021.

(45) Halpin, D. R., and Harbury, P. B. (2004) DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol. 2*, 1022-1030.

(46) Halpin, D. R., Lee, J. A., Wrenn, S. J., and Harbury, P. B. (2004) DNA display III. Solid-phase organic synthesis on unprotected DNA. *PLoS Biol. 2*, 1031-1038.

(47) Wrenn, S. J., Weisinger, R. M., Halpin, D. R., and Harbury, P. B. (2007) Synthetic ligands discovered by in vitro selection. *J. Am. Chem. Soc. 129*, 13137-13143.

(48) Jetson, R. R., and Krusemark, C. J. (2016) Sensing Enzymatic Activity by Exposure and Selection of DNA-Encoded Probes. *Angew. Chem. Int. Ed. 55*, 9562-9566.

(49) Clark, M. A., Acharya, R. A., Arico-Muendel, C. C., Belyanskaya, S. L., Benjamin, D. R., Carlson, N. R., Centrella, P. A., Chiu, C. H., Creaser, S. P., Cuozzo, J. W. *et al.* (2009) Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat. Chem. Biol. 5*, 647-654.

(50) Seigal, B. A., Connors, W. H., Fraley, A., Borzilleri, R. M., Carter, P. H., Emanuel, S. L., Fargnoli, J., Kim, K., Lei, M., Naglich, J. G. *et al.* (2015) The discovery of macrocyclic XIAP antagonists from a DNA-programmed chemistry library, and their optimization to give lead compounds with in vivo antitumor activity. *J. Med. Chem. 58*, 2855-2861.

(51) Samain, F., Ekblad, T., Mikutis, G., Zhong, N., Zimmermann, M., Nauer, A., Bajic, D., Decurtins, W., Scheuermann, J., Brown, P. J. *et al.* (2015) Tankyrase 1 Inhibitors with Drug-like Properties Identified by Screening a DNA-Encoded Chemical Library. *J. Med. Chem. 58*, 5143-5149.

(52) Wichert, M., Krall, N., Decurtins, W., Franzini, R. M., Pretto, F., Schneider, P., Neri, D., and Scheuermann, J. (2015) Dual-display of small molecules enables the discovery of ligand pairs and facilitates affinity maturation. *Nat. Chem. 7*, 241-249.

(53) Petersen, L. K., Blakskjær, P., Chaikuad, A., Christensen, A. B., Dietvorst, J., Holmkvist, J., Knapp, S., Kořínek, M., Larsen, L. K., Pedersen, A. E. *et al.* (2016) Novel p38α MAP kinase inhibitors identified from yoctoReactor DNA-encoded small molecule library. *MedChemComm 7*, 1332-1339.

(54) Satz, A. L. (2016) Simulated Screens of DNA Encoded Libraries: The Potential Influence of Chemical Synthesis Fidelity on Interpretation of Structure-Activity Relationships. *ACS Comb. Sci. 18*, 415-424.

(55) Deng, H., Zhou, J., Sundersingh, F., Messer, J. A., Somers, D. O., Ajakane, M., Arico-Muendel, C. C., Beljean, A., Belyanskaya, S. L., Bingham, R. *et al.* (2016) Discovery and Optimization of Potent, Selective, and in Vivo Efficacious 2-Aryl Benzimidazole BCATm Inhibitors. *ACS Med. Chem. Lett. 7*, 379-384.

(56) Skopic, M. K., Bugain, O., Jung, K., Onstein, S., Brandherm, S., Kalliokoski, T., and Brunschweiger, A. (2016) Design and synthesis of DNA-encoded libraries based on a benzodiazepine and a pyrazolopyrimidine scaffold. *Medchemcomm 7*, 1957-1965.

(57) Soutter, H. H., Centrella, P., Clark, M. A., Cuozzo, J. W., Dumelin, C. E., Guie, M. A., Habeshian, S., Keefe, A. D., Kennedy, K. M., Sigel, E. A. *et al.* (2016) Discovery of cofactor-specific, bactericidal Mycobacterium tuberculosis InhA inhibitors using DNA-encoded library technology. *Proc. Nat. Acad. Sci. USA 113*, E7880-E7889.

(58) Harris, P. A., Berger, S. B., Jeong, J. U., Nagilla, R., Bandyopadhyay, D., Campobasso, N., Capriotti, C. A., Cox, J. A., Dare, L., Dong, X. *et al.* (2017) Discovery of a First-in-Class Receptor Interacting Protein 1 (RIP1) Kinase Specific Clinical Candidate (GSK2982772) for the Treatment of Inflammatory Diseases. *J. Med. Chem. 60*, 1247-1261.

(59) Cuozzo, J. W., Centrella, P. A., Gikunju, D., Habeshian, S., Hupp, C. D., Keefe, A. D., Sigel, E. A., Soutter, H. H., Thomson, H. A., Zhang, Y. *et al.* (2017) Discovery of a Potent BTK Inhibitor with a Novel Binding Mode by Using Parallel Selections with a DNA-Encoded Chemical Library. *ChemBioChem 18*, 864-871.

(60) Belyanskaya, S. L., Ding, Y., Callahan, J. F., Lazaar, A. L., and Israel, D. I. (2017) Discovering Drugs with DNA-Encoded Library Technology: From Concept to Clinic with an Inhibitor of Soluble Epoxide Hydrolase. *ChemBioChem 18*, 837-842.

(61) Johannes, J. W., Bates, S., Beigie, C., Belmonte, M. A., Breen, J., Cao, S., Centrella, P. A., Clark, M. A., Cuozzo, J. W., Dumelin, C. E. *et al.* (2017) Structure Based Design of Non-Natural Peptidic Macrocyclic Mcl-1 Inhibitors. *ACS Med. Chem. Lett. 8*, 239-244.

(62) Ahn, S., Kahsai, A. W., Pani, B., Wang, Q. T., Zhao, S., Wall, A. L., Strachan, R. T., Staus, D. P., Wingler, L. M., Sun, L. D. *et al.* (2017) Allosteric "beta-blocker" isolated from a DNA-encoded small molecule library. *Proc. Nat. Acad. Sci. USA 114*, 1708-1713.

(63) Satz, A. L., Hochstrasser, R., and Petersen, A. C. (2017) Analysis of Current DNA Encoded Library Screening Data Indicates Higher False Negative Rates for Numerically Larger Libraries. *ACS Comb. Sci. 19*, 234-238.

(64) Lerner, R. A., and Brenner, S. (2017) DNA-Encoded Compound Libraries as Open Source: A Powerful Pathway to New Drugs. *Angew. Chem. Int. Ed. 56*, 1164-1165.

(65) Decurtins, W., Wichert, M., Franzini, R. M., Buller, F., Stravs, M. A., Zhang, Y., Neri, D., and Scheuermann, J. (2016)

Automated screening for small organic ligands using DNA-encoded chemical libraries. *Nat. Protoc. 11*, 764-780.

(66) Spitznagel, T. M., and Clark, D. S. (1993) Surface-density and orientation effects on immobilized antibodies and antibody fragments. *Biotechnology 11*, 825-829.

(67) Lu, B., Smyth, M. R., and O'Kennedy, R. (1996) Oriented immobilization of antibodies and its applications in immunoassays and immunosensors. *Analyst 121*, 29R-32R.

(68) Vijayendran, R. A., and Leckband, D. E. (2001) A quantitative assessment of heterogeneity for surface-immobilized proteins. *Anal. Chem. 73*, 471-480.

(69) Cha, T., Guo, A., and Zhu, X. Y. (2005) Enzymatic activity on a chip: the critical role of protein orientation. *Proteomics 5*, 416-419.

(70) Chen, X., Tan, P. H., Zhang, Y., and Pei, D. (2009) On-bead screening of combinatorial libraries: reduction of nonspecific binding by decreasing surface ligand density. *J. Comb. Chem. 11*, 604-611.

(71) Good, M. C., Zalatan, J. G., and Lim, W. A. (2011) Scaffold proteins: hubs for controlling the flow of cellular information. *Science 332*, 680-686.

(72) Buller, F., Steiner, M., Frey, K., Mircsof, D., Scheuermann, J., Kalisch, M., Buhlmann, P., Supuran, C. T., and Neri, D. (2011) Selection of Carbonic Anhydrase IX Inhibitors from One Million DNA-Encoded Compounds. *ACS Chem. Biol. 6*, 336-344.

(73) Leimbacher, M., Zhang, Y., Mannocci, L., Stravs, M., Geppert, T., Scheuermann, J., Schneider, G., and Neri, D. (2012) Discovery of small-molecule interleukin-2 inhibitors from a DNA-encoded chemical library. *Chem. Eur. J. 18*, 7729-7737.

(74) Kollmann, C. S., Bai, X., Tsai, C. H., Yang, H., Lind, K. E., Skinner, S. R., Zhu, Z., Israel, D. I., Cuozzo, J. W., Morgan, B. A. *et al.* (2014) Application of encoded library technology (ELT) to a protein-protein interaction target: discovery of a potent class of integrin lymphocyte function-associated antigen 1 (LFA-1) antagonists. *Bioorg. Med. Chem. 22*, 2353-2365.

(75) Svensen, N., Diaz-Mochon, J. J., and Bradley, M. (2011) Decoding a PNA encoded peptide library by PCR: the discovery of new cell surface receptor ligands. *Chem. Biol. 18*, 1284-1289.

(76) Svensen, N., Diaz-Mochon, J. J., and Bradley, M. (2011) Encoded peptide libraries and the discovery of new cell binding ligands. *Chem. Commun. 47*, 7638-7640.

(77) Wu, Z., Graybill, T. L., Zeng, X., Platchek, M., Zhang, J., Bodmer, V. Q., Wisnoski, D. D., Deng, J., Coppo, F. T., Yao, G. *et al.* (2015) Cell-Based Selection Expands the Utility of DNA-Encoded Small-Molecule Library Technology to Cell Surface Drug Targets: Identification of Novel Antagonists of the NK3 Tachykinin Receptor. *ACS Comb. Sci. 17*, 722-731.

(78) McGregor, L. M., Jain, T., and Liu, D. R. (2014) Identification of ligand-target pairs from combined libraries of small molecules and unpurified protein targets in cell lysates. *J. Am. Chem. Soc. 136*, 3264-3270.

(79) MacConnell, A. B., McEnaney, P. J., Cavett, V. J., and Paegel, B. M. (2015) DNA-Encoded Solid-Phase Synthesis: Encoding Language Design and Complex Oligomer Library Synthesis. *ACS Comb. Sci. 17*, 518-534.

(80) MacConnell, A. B., Price, A. K., and Paegel, B. M. (2017) An Integrated Microfluidic Processor for DNA-Encoded Combinatorial Library Functional Screening. *ACS Comb. Sci. 19*, 181-192.

(81) Mendes, K. R., Malone, M. L., Ndungu, J. M., Suponitsky-Kroyter, I., Cavett, V. J., McEnaney, P. J., MacConnell, A. B., Doran, T. M., Ronacher, K., Stanley, K. *et al.* (2017) High-throughput Identification of DNA-Encoded IgG Ligands that Distinguish Active and Latent Mycobacterium tuberculosis Infections. *ACS Chem. Biol. 12*, 234-243.

(82) McGregor, L. M., Gorin, D. J., Dumelin, C. E., and Liu, D. R. (2010) Interaction-dependent PCR: identification of ligand-target pairs from libraries of ligands and libraries of targets in a single solution-phase experiment. *J. Am. Chem. Soc. 132*, 15522-15524.

(83) Chan, A. I., McGregor, L. M., Jain, T., and Liu, D. R. (2017) Discovery of a covalent kinase inhibitor from a DNA-encoded smallmolecule library X protein library selection. *J. Am. Chem. Soc. DOI: 10.1021/jacs.7b04880*.

(84) Blakskjaer, P., Christensen, A. B., Hansen, N. J. V., Hansen, T. H., Holmkvist, J., Larsen, L. K., Petersen, L. K., Rasmussen-Dietvorst, J., and SLØK, F. A. (2012). A method for making an enriched library. International Patent No. WO2012041633 A1

(85) Bao, J., Krylova, S. M., Cherney, L. T., Hale, R. L., Belyanskaya, S. L., Chiu, C. H., Arico-Muendel, C. C., and Krylov, S. N. (2015) Prediction of protein-DNA complex mobility in gel-free capillary electrophoresis. *Anal. Chem. 87*, 2474-2479.

(86) Bao, J., Krylova, S. M., Cherney, L. T., Hale, R. L., Belyanskaya, S. L., Chiu, C. H., Shaginian, A., Arico-Muendel, C. C., and Krylov, S. N. (2016) Predicting Electrophoretic Mobility of Protein-Ligand Complexes for Ligands from DNA-Encoded Libraries of Small Molecules. *Anal. Chem. 88*, 5498-5506.

(87) Denton, K. E., and Krusemark, C. J. (2016) Crosslinking of DNA-linked ligands to target proteins for enrichment from DNA-encoded libraries. *Medchemcomm 7*, 2020-2027.

(88) Winssinger, N., Harris, J. L., Backes, B. J., and Schultz, P. G. (2001) From Split-Pool Libraries to Spatially Addressable Microarrays and Its Application to Functional Proteomic Profiling. *Angew. Chem. Int. Ed. 40*, 3152-3155.

(89) Winssinger, N., Ficarro, S., Schultz, P. G., and Harris, J. L. (2002) Profiling protein function with small molecule microarrays. *Proc. Nat. Acad. Sci. USA 99*, 11139-11144.

(90) Harris, J., Mason, D. E., Li, J., Burdick, K. W., Backes, B. J., Chen, T., Shipway, A., Van Heeke, G., Gough, L., Ghaemmaghami, A. *et al.* (2004) Activity profile of dust mite allergen extract using substrate libraries and functional proteomic microarrays. *Chem. Biol. 11*, 1361-1372.

(91) Urbina, H. D., Debaene, F., Jost, B., Bole-Feysot, C., Mason, D. E., Kuzmic, P., Harris, J. L., and Winssinger, N. (2006) Self-assembled small-molecule microarrays for protease screening and profiling. *ChemBioChem 7*, 1790-1797.

(92) Zhao, P., Chen, Z., Li, Y., Sun, D., Gao, Y., Huang, Y., and Li, X. (2014) Selection of DNA-encoded small molecule libraries against unmodified and non-immobilized protein targets. *Angew. Chem. Int. Ed. 53*, 10056-10059.

(93) Zuo, X., Xia, F., Xiao, Y., and Plaxco, K. W. (2010) Sensitive and selective amplified fluorescence DNA detection based on exonuclease III-aided target recycling. *J. Am. Chem. Soc. 132*, 1816-1818.

(94) Li, G., Liu, Y., Chen, L., Wu, S., and Li, X. (2013) Photoaffinity Labeling of Small-Molecule-Binding Proteins by DNA-Templated Chemistry. *Angew. Chem. Int. Ed. 52*, 9544-9549.

(95) Green, N. M. (1966) Thermodynamics of the binding of biotin and some analogues by avidin. *Biochem. J. 101*, 774-780.

(96) Doyon, J. B., Snyder, T. M., and Liu, D. R. (2003) Highly sensitive in vitro selections for DNA-linked synthetic small molecules with protein binding affinity and specificity. *J. Am. Chem. Soc. 125*, 12372-12373.

(97) Kleiner, R. E., Dumelin, C. E., Tiu, G. C., Sakurai, K., and Liu, D. R. (2010) In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *J. Am. Chem. Soc. 132*, 11779-11791.

(98) Tse, B. N., Snyder, T. M., Shen, Y. H., and Liu, D. R. (2008) Translation of DNA into a Library of 13 000 Synthetic Small-Molecule Macrocycles Suitable for in Vitro Selection. *J. Am. Chem. Soc. 130*, 15611-15626.

(99) Cao, C., Zhao, P., Li, Z., Chen, Z., Huang, Y., Bai, Y., and Li, X. (2014) A DNA-templated synthesis of encoded small

molecules by DNA self-assembly. *Chem. Commun. 50*, 10997-10999.

(100) Dumelin, C. E., Scheuermann, J., Melkko, S., and Neri, D. (2006) Selection of streptavidin binders firom a DNA-encoded chemical library. *Bioconjugate Chem. 17*, 366-370.

(101) Stevenson, R., Baxter, H. C., Aitken, A., Brown, T., and Baxter, R. L. (2008) Binding of 14-3-3 proteins to a single stranded oligodeoxynucleotide aptamer. *Bioorg. Chem. 36*, 215-219.

(102) Little, J. W. (1967) An exonuclease induced by bacteriophage lambda. II. Nature of the enzymatic reaction. *J. Biol. Chem. 242*, 679-686.

(103) Kujau, M. J., and Wolfl, S. (1997) Efficient preparation of single-stranded DNA for in vitro selection. *Mol. Biotechnol. 7*, 333-335.

(104) Mannocci, L., Zhang, Y., Scheuermann, J., Leimbacher, M., De Bellis, G., Rizzi, E., Dumelin, C., Melkko, S., and Neri, D. (2008) High-throughput sequencing allows the identification of binding molecules isolated from DNA-encoded chemical libraries. *Proc. Nat. Acad. Sci. USA 105*, 17670-17675.

TOC graphics