Published in final edited form as:

Curr Opin Chem Biol. 2015 June ; 26: 55–61. doi:10.1016/j.cbpa.2015.02.010.

Novel selection methods for DNA-encoded chemical libraries

Alix I. Chan^{1,‡}, Lynn M. McGregor^{1,2,‡}, and David R. Liu^{*}

¹Department of Chemistry and Chemical Biology and Howard Hughes Medical Institute, Harvard University, 12 Oxford St, Cambridge, MA 02138

Abstract

Driven by the need for new compounds to serve as biological probes and leads for therapeutic development and the growing accessibility of DNA technologies including high-throughput sequencing, many academic and industrial groups have begun to use DNA-encoded chemical libraries as a source of bioactive small molecules. In this review, we describe the technologies that have enabled the selection of compounds with desired activities from these libraries. These methods exploit the sensitivity of *in vitro* selection coupled with DNA amplification to overcome some of the limitations and costs associated with conventional screening methods. In addition, we highlight newer techniques with the potential to be applied to the high-throughput evaluation of DNA-encoded chemical libraries.

Introduction

The rapidly expanding wealth of genomic, proteomic, and metabolomic data has led to the identification of many biological targets with therapeutic potential. Small molecules that can potently and specifically engage proteins or nucleic acids associated with disease are especially valuable as probes to validate the putative roles of targets in disease progression, or as potential leads for therapeutic development. Conventional efforts to discover small-molecule ligands frequently use high-throughput screening, in which thousands of compounds are individually exposed to a target of interest and assayed for bioactivity. Such screening methods, however, can be costly, time-consuming, and require major instrumentation and specialized expertise.

A complementary approach to evaluating synthetic small-molecule libraries uses *in vitro* selections to rapidly and simultaneously assess the ability of all library members to interact with targets of interest. This approach is especially amenable to the evaluation of large, chemically diverse, DNA-encoded libraries that have been described by several research groups in academia and industry [1-13]. Because DNA can be readily replicated and

© 2015 Published by Elsevier Ltd.

^{*}To whom correspondence should be addressed: drliu@fas.harvard.edu.

²Present address: Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94158

[‡]These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

sequenced, *in vitro* selections of DNA-encoded chemical libraries offer the major advantage of simultaneously evaluating thousands or even millions of compounds for their ability to interact with target proteins in a single experiment. Selections significantly reduce the amount of compound, target protein, time, and cost required to evaluate a library [14,15]. In addition, selections on such libraries can readily yield structure-activity relationships that inform future medicinal chemistry efforts.

In Vitro Selection Methods for Ligand Discovery

The most common selections performed on DNA-encoded chemical libraries are binding selections on purified target proteins (Figure 1a). A target protein is immobilized by binding or covalent attachment to a solid support and then incubated with a DNA-encoded chemical library. Alternatively, the library can be incubated with free protein in solution and captured with target-specific antibodies that are immobilized or tagged with an affinity handle [1,16]. After washing [7], library members capable of binding the target are eluted and subjected to additional round(s) of selection or PCR amplification of their associated DNA templates and massively parallel high-throughput DNA sequencing. Depending on the total library size, putative hits can be identified from amplified sequences either by comparing the sequence abundance to that in the starting library [17] or by fitting observed sequence counts to a negative binomial distribution [8] or to the Poisson distribution [18].

While traditional selections on solid-supported protein targets offer dramatically increased efficiency compared to screens, they must be performed on a purified target protein of interest. Target immobilization may result in artifactual binding or in the loss of native conformational properties that are required for bona fide binding to native targets. In addition, washing and elution steps required for selections using immobilized targets may also remove active library members or fail to result in the isolation of desired species. Interaction-dependent PCR (IDPCR) [19] was developed by Liu and coworkers to address these limitations and enable simultaneous evaluation of binding between all members of combined libraries of DNA-linked targets and compounds in a single solution (Figure 1b). In IDPCR, protein-small molecule binding brings encoding DNA sequences into close proximity and promotes DNA hybridization of a self-priming hairpin. Polymerase-catalyzed primer extension produces a selectively amplifiable DNA sequence encoding both members of the protein-small molecule complex.

The IDPCR approach was extended to detect ligand binding to unpurified proteins in a method called interaction determination using unpurified proteins (IDUP) [20]. By operating in cell lysates, IDUP preserves post-translational modifications and interactions with endogenous binding partners, enabling the study of difficult-to-purify targets and increasing the potential biological relevance of detected interactions. During IDUP, target proteins are associated with DNA oligonucleotide tags either noncovalently using a DNA-linked antibody or covalently using a SNAP-tag. In a model library $\times$ library binding experiment using combined libraries of 262 DNA-linked small molecules and 256 cell lysates expressing SNAP-tagged targets, IDUP enriched all five known interactions highly, despite having affinities varying from 0.2 nM to 3.2 μ M [20]. This method provides an efficient

approach for rapidly evaluating the binding of ligand libraries in cases in which purified proteins are not available or differ significantly from their native cellular counterparts.

Li and coworkers developed another method for assaying binding of small-molecules libraries to unmodified, non-immobilized proteins by DNA-programmed affinity labeling (DPAL) [21]. In DPAL (Figure 1c), unmodified proteins are mixed with small DNA oligonucleotides bearing a 5' azidophenyl photocrosslinking moiety (PC-DNA). When a DNA-tagged small molecule binds the target, DNA hybridization to a complementary region on the short PC-DNA brings the photo-reactive group into proximity with the protein target. UV irradiation leads to covalent attachment of the PC-DNA to the target protein, which can be identified by mass spectrometry. The DPAL technique was adapted to identify binders from libraries of small molecules by taking advantage of hybridization between the PC-DNA and the binding ligands' DNA tags [22]. The DNA tags of non-binding small molecules are digested by Exonuclease I, but the hybridized DNA tags of bound molecules are protected from digestion. In addition, DPAL is compatible with targets in cell lysates, a condition in which targets may closely mimic their native state, and with large excesses of non-binding DNA-linked small molecules. DPAL can also perform iterative rounds of selection to greatly amplify DNA sequences corresponding to binding small molecules.

Photoaffinity probes, such as diazirines, benzophenones, and phenyl azides, stabilize interactions between protein targets and DNA-linked small molecules [21,23]. Using DPAL, Li and coworkers enriched DNA sequences corresponding to interactions as weak as 14 mM. Especially when combined with multivalent ligand display [23], the use of photoaffinity probes could stabilize weak interactions enough to enable DNA-encoded fragment-based screens [24].

In an alternate approach to fragment-based selections, Neri and coworkers developed a DNA-hybridization based approach for performing selections using encoded self-assembled chemical libraries (ESAC) [25]. In an ESAC selection (Figure 1d), a DNA-linked small-molecule ligand is hybridized to a library of DNA-encoded pharmacophores. The hybridized mixture is subjected to *in vitro* selection either to identify fragments that improve binding of a known ligand (affinity maturation) or to discover novel synergistically binding ligands. These fragments are then covalently coupled using a variety of linkers, and the resulting compounds are assayed without the DNA barcodes to discover linker architectures optimal for binding. Using ESAC, Neri and coworkers discovered higher-affinity inhibitors of trypsin [26] and MMP-3 [9].

When DNA-linked antibodies are bound to a protein or protein complex, the DNA tags are brought into close proximity and can be linked by ligation [27] or extension [28] to give a selectively amplifiable DNA sequence. Landegren and coworkers have validated the proximity ligation assay (PLA) for the quantification of biomarkers by high-throughput sequencing [29] and for the analysis of protein-protein interactions and protein subcellular localization or post-translational modification [30,31]. PLA was also applied to the study of small molecules capable of disrupting the interaction between VEGF-A and its receptors VEGFR-1 and VEGFR-2 [32].

Using a microarray to quantify unique DNA tags for each antibody, PLA can evaluate all pairwise protein-protein interactions (PPIs) within a larger set of proteins [33]. Recently, Huang and coworkers applied PLA to interrogate 1204 PPIs in a one-by-one fashion [34], resulting in the identification of hundreds of previously validated PPIs. Tagging individual antibodies with unique DNA barcodes could enable readout of PPIs by high-throughput DNA sequencing and expedite PPI cataloging efforts.

Whereas the previously described methods use DNA hybridization to promote amplification of active library members, *in vitro* selection typically results in spatial separation of active library members from inactive species. To select binding molecules from their DNA-encoded chemical library [1], Vipergen developed a technique called binder trap enrichment (Figure 1e) (<http://vipergen.com>). After exposing a DNA-encoded chemical library to a target of interest, binding pairs are compartmentalized in water and oil emulsion droplets. Statistically, binding ligands are more efficiently co-compartmentalized with the target protein than nonfunctional library members. Enzymatic ligation of the co-compartmentalized oligonucleotides encoding the ligand and target is followed by PCR amplification and high-throughput DNA sequencing to identify the binding entities.

Reaction-Discovery Using DNA-Encoded Chemical Libraries

Interest in the development of novel chemical transformations has led to the development of unbiased reactivity screens based on LC/MS [35], GC/MS [36], and sandwich immunoassay [37]. Reaction discovery using DNA-encoded chemical libraries can dramatically streamline the reaction discovery process by enabling simultaneous evaluation of all possible combinations of potential reactants among substrate library members. The first DNA-encoded reaction discovery schemes by Liu and coworkers relied on solid-phase separation of bond-forming pairs of library members and their encoding DNA from inactive species (Figure 2a) and resulted in the discovery of a Pd(II)-mediated alkynamide-alkene coupling reaction [38], a Au(III)- and acid-catalyzed alkene hydroarylation [39] and a biocompatible, Ru(II)-catalyzed azide reduction induced by visible light [40].

In reactivity-dependent PCR (RDPCR) (Figure 2b), covalent bond formation results in hybridization of a self-priming hairpin that can be extended to encode the identify of both substrates and that is selectively amplified in PCR [41]. RDPCR can also be configured to detect bond cleavage that results in the unmasking of a reactive functional group. In a proof-of-principle example, a DNA-linked peptide reacted with an activated DNA-linked carboxylate only if the peptide was first treated with the protease subtilisin (generating a free amino terminus) [41]. Li and coworkers developed another approach to report covalent bond formation using DNA to bring together substrates that hybridize to a common template (Figure 2c) [42]. After hybridizing the substrates and allowing them to react, a photolabile base in the template strand is cleaved. If covalent bond formation has joined the two substrates, their pendant DNA sequences serve as splints for a ligation reaction that restores the cleaved template strand and enables its PCR amplification.

Potential DNA-Based Assays for Protein-Small Molecule Binding

A variety of novel DNA-based methods for detecting binding between proteins and small molecules rely on the ability of protein-small molecule complexes to prevent interactions between enzymes and encoding DNA. Complexes that prevent digestion by *E. coli* Exonuclease I [43-45] or Exonuclease III [46] could in principle allow detection of DNA sequences encoding active small molecules by high-throughput sequencing, without relying on spatial separation or primer extension to further amplify the codes of active library members. These potential selection methods would not require solid-phase immobilization, washing, or elution steps, but would render the DNA-encoded libraries single-use. Indeed, DPAL, IDUP, and the proximity extension assay all take advantage of protein-small molecule complexes' ability to protect DNA from exonucleases to decrease the signal arising from inactive library members [20,22,28]. Complexes that prevent DNA from interacting with T4 DNA ligase [47] or T7 RNA polymerase [48] could also be adapted to evaluate DNA-encoded chemical libraries. These approaches would have the significant disadvantage, however, of operating by a loss-of-signal mechanism in which the sequences corresponding to inactive compounds are selectively amplified.

Novel DNA sequencing techniques can also drive advances in library evaluation. Colonies are clusters of DNA immobilized in a polyacrylamide gel that can be sequenced in high-throughput by single base extension or ligation-based sequencing [49]. Single-molecular-interaction sequencing (SMI-Seq), an approach developed by Church and coworkers, uses polony sequencing to identify the complexes of DNA-linked proteins by analyzing the degree of co-localization of colonies corresponding to each protein [50]. SMI-Seq was used to evaluate the selectivity of each member of a library of 200 scFv variants for its ability to bind to each member of a library of 55 human antigens.

In the strand-displacement competition assay developed by Gothelf and coworkers [51], a DNA-linked small molecule is hybridized with a complementary oligonucleotide. Protein-small molecule binding decreases the affinity of this duplex, enabling toehold displacement of the complementary strand. In the proof-of-principle study, this toehold displacement resulted in a signal that could be analyzed by PAGE or FRET. In principle, this approach could also be adapted to a high-throughput sequencing-based readout, and to the evaluation of DNA-encoded chemical libraries.

Solid-phase selections have also been adapted to the study of protein-protein interactions. In parallel analysis of translated ORFs (PLATO), Elledge and coworkers use an immobilized protein to capture binding proteins that are linked to their encoding DNA by ribosome display [52]. PLATO could in principle be adapted to evaluate small molecules linked to beads for target identification, or when interfaced with a DNA-encoded chemical library, could potentially enable simultaneous evaluation of all interactions with a set of ribosome-displayed protein targets.

Conclusions

Selections on DNA-encoded chemical libraries have recently resulted in the discovery of new classes of synthetic small-molecule ligands against a variety of protein targets including several associated with human disease [7,17,53-56]. Given the ever-increasing number of proteins, nucleic acids, and metabolites implicated in human disease, innovations in the field of DNA-encoded libraries and their rapid *in vitro* selection are likely to play an increasingly important role in the discovery of small molecules with the potential to probe therapeutically relevant biological pathways or to serve as leads for the development of new medicines. Creative approaches for evaluating DNA-encoded chemical libraries are likely to continue to leverage the ease of designing complementary DNA strands, the extremely high sensitivity afforded by DNA amplification, and the remarkable efficiency of modern high-throughput DNA sequencing to facilitate the increasingly efficient and applicable discovery of novel molecular interactions.

Acknowledgments

This work was supported by the Howard Hughes Medical Institute and the NIH/NIGMS (R01GM065865).

References

**Of outstanding interest

*Of special interest

1. Hansen MH, Blakskjaer P, Petersen LK, Hansen TH, Hojfeldt JW, Gothelf KV, Hansen NJV. A Yoctoliter-Scale DNA Reactor for Small-Molecule Evolution. *J Am Chem Soc.* 2009; 131:1322–1327. [PubMed: 19123795]
2. Tse BN, Snyder TM, Shen Y, Liu DR. Translation of DNA into a Library of 13 000 Synthetic Small-Molecule Macrocycles Suitable for *in Vitro* Selection. *J Am Chem Soc.* 2008; 130:15611–15626. [PubMed: 18956864]
3. Weisinger RM, Wrenn SJ, Harbury PB. Highly parallel translation of DNA sequences into small molecules. *Plos One.* 2012; 7:e28056. [PubMed: 22479303]
4. Clark MA, Acharya RA, Arico-Muendel CC, Belyanskaya SL, Benjamin DR, Carlson NR, Centrella PA, Chiu CH, Creaser SP, Cuzzo JW, et al. Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat Chem Biol.* 2009; 5:647–654. [PubMed: 19648931]
5. He Y, Liu DR. A sequential strand-displacement strategy enables efficient six-step DNA-templated synthesis. *J Am Chem Soc.* 2011; 133:9972–9975. [PubMed: 21657248]
6. He Y, Liu DR. Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker. *Nat Nanotechnol.* 2010; 5:778–782. [PubMed: 20935654]
7. Leimbacher M, Zhang Y, Mannocci L, Stravs M, Geppert T, Scheuermann J, Schneider G, Neri D. Discovery of Small-Molecule Interleukin-2 Inhibitors from a DNA-Encoded Chemical Library. *Chem Eur J.* 2012; 18:7729–7737. [PubMed: 22588840]
8. Buller F, Steiner M, Scheuermann J, Mannocci L, Nissen I, Kohler M, Beisel C, Neri D. High-throughput sequencing for the identification of binding molecules from DNA-encoded chemical libraries. *Bioorg Med Chem Lett.* 2010; 20:4188–4192. [PubMed: 20538458]
- 9*. Scheuermann J, Dumelin CE, Melkko S, Zhang Y, Mannocci L, Jaggi M, Sobek J, Neri D. DNA-Encoded Chemical Libraries for the Discovery of MMP-3 Inhibitors. *Bioconjugate Chem.* 2008; 19:778–785. Neri and coworkers selected a novel binder to MMP-3 from a library of small pharmacophores. Using encoded self-assembling chemical libraries, this hit was optimized to achieve 9.9 μ M affinity for MMP-3.

10. Novoa A, Machida T, Barluenga S, Imberty A, Winssinger N. PNA-encoded synthesis (PES) of a 10 000-member hetero-glycoconjugate library and microarray analysis of diverse lectins. *ChemBioChem*. 2014; 15:2058–2065. [PubMed: 25158314]
11. Kawakami T, Ishizawa T, Fujino T, Reid PC, Suga H, Murakami H. In Vitro Selection of Multiple Libraries Created by Genetic Code Reprogramming To Discover Macrocyclic Peptides That Antagonize VEGFR2 Activity in Living Cells. *ACS Chem Biol*. 2013; 8:1205–1214. [PubMed: 23517428]
12. Schlippe YVG, Hartman MCT, Josephson K, Szostak JW. In vitro selection of highly modified cyclic peptides that act as tight binding inhibitors. *J Am Chem Soc*. 2012; 134:10469–10477. [PubMed: 22428867]
13. Li Y, Zhao P, Zhang M, Zhao X, Li X. Multistep DNA-Templated Synthesis Using a Universal Template. *J Am Chem Soc*. 2013; 135:17727–17730. [PubMed: 24229415]
14. Zhu Z, Cuozzo J. Review Article: High-Throughput Affinity-Based Technologies for Small-Molecule Drug Discovery. *J Biomol Screening*. 2009; 14:1157–1164.
15. Kleiner RE, Dumelin CE, Liu DR. Small-molecule discovery from DNA-encoded chemical libraries. *Chem Soc Rev*. 2011; 40:5707–5717. [PubMed: 21674077]
16. Wrenn SJ, Weisinger RM, Halpin DR, Harbury PB. Synthetic ligands discovered by in vitro selection. *J Am Chem Soc*. 2007; 129:13137–13143. [PubMed: 17918937]
17. Maianti JP, McFedries A, Foda ZH, Kleiner RE, Du XQ, Leissring MA, Tang WJ, Charron MJ, Seeliger MA, Saghatelian A, et al. Anti-diabetic activity of insulin-degrading enzyme inhibitors mediated by multiple hormones. *Nature*. 2014; 511:94–98. [PubMed: 24847884]
18. DeGroot, MH.; Schervish, MJ. Probability and Statistics. Addison-Wesley; 2012.
19. McGregor LM, Gorin DJ, Dumelin CE, Liu DR. 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*. 2010; 132:15522–15524. [PubMed: 20949943]
- 20*. McGregor LM, Jain T, Liu DR. Identification of Ligand-Target Pairs from Combined Libraries of Small Molecules and Unpurified Protein Targets in Cell Lysates. *J Am Chem Soc*. 2014; 136:3264–3270. DNA-linked antibodies or self-labeling protein fusions enable one-pot, solution phase selections of libraries of unmodified proteins in cell lysate against libraries of small molecules. This method provides an efficient approach for evaluating the binding profiles of ligand libraries. [PubMed: 24495225]
21. Li G, Liu Y, Liu Y, Chen L, Wu S, Liu Y, Li X. Photoaffinity Labeling of Small-Molecule-Binding Proteins by DNA-Templated Chemistry. *Angew Chem Int Edit*. 2013; 52:9544–9549.
- 22*. Zhao P, Chen Z, Li Y, Sun D, Gao Y, Huang Y, Li X. Selection of DNA-encoded small molecule libraries against unmodified and non-immobilized protein targets. *Angew Chem Int Ed Engl*. 2014; 53:10056–10059. Photoreactive moieties allow for covalent modification of proteins that bind DNA-linked ligands. Exonuclease I selectively degrades DNA barcodes corresponding to nonbinding small molecules, and the codes of binding molecules are amplified and decoded. [PubMed: 25044298]
23. Li G, Liu Y, Yu X, Li X. Multivalent Photoaffinity Probe for Labeling Small Molecule Binding Proteins. *Bioconjugate Chem*. 2014; 25:1172–1180.
24. Scott DE, Coyne AG, Hudson SA, Abell C. Fragment-Based Approaches in Drug Discovery and Chemical Biology. *Biochemistry*. 2012; 51:4990–5003. [PubMed: 22697260]
25. Melkko S, Scheuermann J, Dumelin CE, Neri D. Encoded self-assembling chemical libraries. *Nat Biotechnol*. 2004; 22:568–574. [PubMed: 15097996]
26. Melkko S, Zhang Y, Dumelin CE, Scheuermann J, Neri D. Isolation of high-affinity trypsin inhibitors from a DNA-encoded chemical library. *Angew Chem Int Edit*. 2007; 46:4671–4674.
27. Fredriksson S, Dixon W, Ji H, Koong AC, Mindrinos M, Davis RW. Multiplexed protein detection by proximity ligation for cancer biomarker validation. *Nat Methods*. 2007; 4:327–329. [PubMed: 17369836]
28. Lundberg M, Eriksson A, Tran B, Assarsson E, Fredriksson S. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Res*. 2011; 39:e102. [PubMed: 21646338]

29. Darmanis S, Nong RY, Vanelid J, Siegbahn A, Ericsson O, Fredriksson S, Bäcklin C, Gut M, Heath S, Gut IG, et al. ProteinSeq: high-performance proteomic analyses by proximity ligation and next generation sequencing. *Plos One*. 2011; 6:e25583. [PubMed: 21980495]
30. Soderberg O, Gullberg M, Jarvius M, Ridderstråle K, Leuchowius KJ, Jarvius J, Wester K, Hydbring P, Bahram F, Larsson LG, et al. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods*. 2006; 3:995–1000. [PubMed: 17072308]
31. Blokzijl A, Friedman M, Pontén F, Landegren U. Profiling protein expression and interactions: proximity ligation as a tool for personalized medicine. *J Intern Med*. 2010; 268:232–245. [PubMed: 20695973]
32. Gustafsdottir SM, Wennstrom S, Fredriksson S, Schallmeiner E, Hamilton AD, Sebti SM, Landegren U. Use of proximity ligation to screen for inhibitors of interactions between vascular endothelial growth factor A and its receptors. *Clin Chem*. 2008; 54:1218–1225. [PubMed: 18499900]
33. Hammond M, Nong R, Ericsson O, Pardali K. Profiling cellular protein complexes by proximity ligation with dual tag microarray readout. *Plos One*. 2012; 7:e40405. [PubMed: 22808155]
34. Chen TC, Lin KT, Chen CH, Lee SA, Lee PY, Liu YW, Kuo YL, Wang FS, Lai JM, Huang CYF. Using an in situ proximity ligation assay to systematically profile endogenous protein-protein interactions in a pathway network. *J Proteome Res*. 2014; 13:5339–5346. [PubMed: 25241761]
35. Beeler AB, Su S, Singleton CA, Porco JA. Discovery of Chemical Reactions through Multidimensional Screening. *J Am Chem Soc*. 2007; 129:1413–1419. [PubMed: 17263426]
36. McNally A, Prier CK, MacMillan DWC. Discovery of an α -Amino C–H Arylation Reaction Using the Strategy of Accelerated Serendipity. *Science*. 2011; 334:1114–1117. [PubMed: 22116882]
37. Quinton J, Kolodych S, Chaumonet M, Bevilacqua V, Nevers MC, Volland H, Gabillet S, Thuéry P, Créminon C, Taran F. Reaction discovery by using a sandwich immunoassay. *Angew Chem Int Ed Engl*. 2012; 51:6144–6148. [PubMed: 22566166]
38. Kanan MW, Rozenman MM, Sakurai K, Snyder TM, Liu DR. Reaction discovery enabled by DNA-templated synthesis and in vitro selection. *Nature*. 2004; 431:545–549. [PubMed: 15457254]
39. Rozenman MM, Kanan MW, Liu DR. Development and Initial Application of a Hybridization-Independent, DNA-Encoded Reaction Discovery System Compatible with Organic Solvents. *J Am Chem Soc*. 2007; 129:14933–14938. [PubMed: 17994738]
- 40*. Chen Y, Kamlet AS, Steinman JB, Liu DR. A biomolecule-compatible visible-light-induced azide reduction from a DNA-encoded reaction-discovery system. *Nature Chem*. 2011; 3:146–153. A Ru(II)-catalyzed, biocompatible azide reduction was discovered using DNA-encoded substrates. This reaction exhibits remarkable chemoselectivity and proceeds under mild conditions. [PubMed: 21258388]
- 41*. Gorin DJ, Kamlet AS, Liu DR. Reactivity-dependent PCR: direct, solution-phase in vitro selection for bond formation. *J Am Chem Soc*. 2009; 131:9189–9191. Reactivity-dependent PCR allows for one-pot, solution phase evaluation of all potential bond-forming or bond-cleaving reactions within libraries of DNA-linked functional groups. [PubMed: 19522494]
- 42**. Li Y, Zhang M, Zhang C, Li X. Detection of bond formations by DNA-programmed chemical reactions and PCR amplification. *Chem Commun*. 2012; 48:9513. DNA-tagged functional groups are annealed to a template containing a photocleaveable linker and only bond-forming events stabilize the cleaved template after irradiation. This strategy allows for re-ligation and decoding of only templates corresponding to reactive substrates.
43. Cao Y, Zhu S, Yu J, Zhu X, Yin Y, Li G. Protein Detection Based on Small Molecule-Linked DNA. *Anal Chem*. 2012; 84:4314–4320. [PubMed: 22503033]
44. Wu Z, Zhen Z, Jiang JH, Shen GL, Yu RQ. Terminal Protection of Small-Molecule-Linked DNA for Sensitive Electrochemical Detection of Protein Binding via Selective Carbon Nanotube Assembly. *J Am Chem Soc*. 2009; 131:12325–12332. [PubMed: 19655753]
45. Xu Y, Jiang B, Xie J, Xiang Y, Yuan R, Chai Y. Terminal protection of small molecule-linked ssDNA for label-free and sensitive. *Talanta*. 2014; 128:237–241. [PubMed: 25059154]
46. Wu Z, Wang H, Guo M, Tang LJ, Yu RQ, Jiang JH. Terminal Protection of Small Molecule-Linked DNA: A Versatile Biosensor Platform for Protein Binding and Gene Typing Assay. *Anal Chem*. 2011; 83:3104–3111. [PubMed: 21438529]

47. Sugita R, Mie M, Funabashi H, Kobatake E. Evaluation of small ligand–protein interaction by ligation reaction with DNA-modified ligand. *Biotechnol Lett.* 2009; 32:97–102. [PubMed: 19705069]
48. Mie M, Sugita R, Endoh T, Kobatake E. Evaluation of small ligand-protein interactions by using T7 RNA polymerase with DNA-modified ligand. *Anal Biochem.* 2010; 405:109–113. [PubMed: 20553866]
49. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science.* 2005; 309:1728–1732. [PubMed: 16081699]
- 50*. Gu L, Li C, Aach J, Hill DE, Vidal M, Church GM. Multiplex single-molecule interaction profiling of DNA-barcoded proteins. *Nature.* 2014; 515:554–557. In SMI-seq, DNA-barcoded proteins are immobilized in a gel, and their interactions with other proteins are determined by the rates of colocalization. This method could be applied to assay binding to other DNA-encoded libraries. [PubMed: 25252978]
- 51*. Zhang Z, Hejesen C, Kjelstrup MB, Birkedal V, Gothelf KV. A DNA-Mediated Homogeneous Binding Assay for Proteins and Small Molecules. *J Am Chem Soc.* 2014; 136:11115–11120. Protein binding to a DNA-linked small molecule reduces the melting temperature of that DNA and its complementary strand. This effect is used to promote toehold displacement visualized by FRET and could be adapted to evaluate DNA-encoded chemical libraries. [PubMed: 25051012]
- 52*. Zhu J, Larman HB, Gao G, Somwar R, Zhang Z, Laserson U, Ciccio A, Pavlova N, Church G, Zhang W, et al. Protein interaction discovery using parallel analysis of translated ORFs (PLATO). *Nat Biotechnol.* 2013; 31:331–334. In PLATO, ribosome-displayed proteins are assayed for binding to an immobilized target protein. This method could be adapted to selections for other binding interactions. [PubMed: 23503679]
53. Clarke J, Wu HC, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol.* 2009; 4:265–270. [PubMed: 19350039]
54. Buller F, Zhang Y, Scheuermann J, Schäfer J, Bühlmann P, Neri D. Discovery of TNF inhibitors from a DNA-encoded chemical library based on diels-alder cycloaddition. *Chem Biol.* 2009; 16:1075–1086. [PubMed: 19875081]
55. Podolin PL, Bolognese BJ, Foley JF, Long E III, Peck B, Umbrecht S, Zhang X, Zhu P, Schwartz B, Xie W, et al. Prostaglandins and Other Lipid Mediators. *Prostag Oth Lipid M.* 2013; 104-105:25–31.
56. Deng H, O'keefe H, Davie CP, Lind KE, Acharya RA, Franklin GJ, Larkin J, Matico R, Neeb M, Thompson MM, et al. Discovery of Highly Potent and Selective Small Molecule ADAMTS-5 Inhibitors That Inhibit Human Cartilage Degradation via Encoded Library Technology (ELT). *J Med Chem.* 2012; 55:7061–7079. [PubMed: 22891645]

Highlights

- DNA-encoded chemical libraries are a source of diverse bioactive compounds
- Current approaches enable one-pot evaluation of all library members
- New methods evaluate libraries against one or more unmodified target proteins
- Related approaches have enabled DNA-encoded reaction discovery
- New DNA-based approaches for detecting protein interactions could be adapted to small-molecule discovery

Figure 1.

Technologies for the *in vitro* selection of DNA-encoded chemical libraries. A) In solid-phase selection, all library members are simultaneously evaluated. Library members that are capable binding of an immobilized target protein are physically separated from inactive library members. B) Interaction determination using unpurified proteins (IDUP) takes advantage of the selective formation of a DNA hairpin between binding species to encode both protein and ligand on the same DNA strand. C) DNA-programmed affinity labeling (DPAL) makes use of an oligonucleotide-coupled photoreactive group (PC-DNA) that covalently labels only proteins bound by a library member. These complexes protect the corresponding DNA tag from digestion and allow its subsequent sequencing and decoding. D) Encoded self-assembling chemical libraries (ESAC) use DNA hybridization to display multiple pharmacophores for interaction with a target protein. Tightly binding pharmacophore combinations guide the synthesis of high-affinity ligands based on the pharmacophore fragments. E) Library selections using binder trap enrichment utilize micelles to co-compartmentalize DNA tags associated with binding species.

Figure 2.

Approaches for DNA-encoded discovery of bond-forming reactions. A) The first reported reactions discovered from DNA-encoded chemical libraries used disulfide cleavage and streptavidin capture to isolate the DNA encoding pairs of substrates that had undergone bond formation under a given set of reaction conditions. B) Reactivity-dependent PCR (RDPCR) is a one-pot method that uses the formation of a self-priming hairpin between coupled library members' DNA barcodes to selectively amplify DNA encoding bond-forming substrate combinations. C) Bond formation can also be detected by installing a photocleavable linker in a DNA template used for reagent hybridization. Only covalently coupled reagents' templates are re-ligated, amplified, and sequenced.