nature chemistry

PERSPECTIVE

https://doi.org/10.1038/s41557-021-00877-x

Check for updates

Strategies for developing DNA-encoded libraries beyond binding assays

Yiran Huang[1], Yizhou Li[2,3] ✉ and Xiaoyu Li[1,4] ✉

DNA-encoded chemical libraries (DELs) have emerged as a powerful technology in drug discovery. The wide adoption of DELs in the pharmaceutical industry and the rapid advancements of DEL-compatible chemistry have further fuelled its development and applications. In general, a DEL has been considered as a massive binding assay to identify physical binders for individual protein targets. However, recent innovations demonstrate the capability of DELs to operate in the complex milieu of biological systems. In this Perspective, we discuss the recent progress in using DNA-encoded chemical libraries to interrogate complex biological targets and their potential to identify structures that elicit function or possess other useful properties. Future breakthroughs in these aspects are expected to catapult DEL to become a momentous technology platform not only for drug discovery but also to explore fundamental biology.

Identification of small molecule ligands that can modulate the functions of biological targets is a central task in biological and pharmaceutical sciences. In the 1990s, the advent of high-throughput screening (HTS) made the screening of chemical libraries a routine practice in drug discovery. However, biological display technologies, which include phage[1], messenger RNA[2], yeast[3] and ribosome[4] display libraries, utilize the biological machinery to evolve large libraries of peptides and peptidomimetics. A DNA-encoded chemical library (DEL) may be considered as a union of the two technologies: the synthetic chemotype (chemical compound) is connected with the encoding genotype (DNA tag)[5]; the spatial encoding in HTS is replaced with DNA encoding, which allows the entire library to be synthesized and selected (rather than screened) simultaneously (Fig. 1a). DELs can access a greater chemical space than biological display libraries, especially with the recent expansion of DEL-compatible chemistries[6–11]. DELs can be prepared and selected at a minute scale, which overcomes the throughput limits of traditional HTS and is more affordable and accessible. In 2017, Lerner and Brenner laid out a roadmap on making DEL openly available to the research community, in which they proposed mechanisms to balance the cost, intellectual property and material accessibility for open-source DELs[12]. Today, various types of pre-made DEL kits (DELPro and DELight, https://hits.wuxiapptec.com/delopen; OpenDEL, https://www.hitgen.com/en/capabilities-details-21.html; OpenDEX, https://www.x-chemrx.com; GenDECL, https://www.genscript.com/dna-encoded-chemical-library-kit.html; DyNAbind, https://www.sigmaaldrich.com) and reagent kits (DELEZ, https://www.hitgen.com/en/capabilities-details-22.html) for custom DEL synthesis are available to meet the needs of individual researchers.

The original concept of a DEL was proposed by Brenner and Lerner in 1992 as a way to improve one-bead, one-compound (OBOC) libraries[13]. The concept was quickly turned into reality by Nielsen et al. in 1993[14]. Nearly at the same time, Gallop and co-workers synthesized an ~820,000-member peptide DEL and selected it against an antibody target[15]. These early DELs were in OBOC format, in which each bead contained many copies of one compound and the DNA tag. The bead-based format had a practical limit on library size and did not allow the miniaturized in-solution DEL selections we see today; they also involved the parallel synthesis of the compounds and the DNA tag, which turned out to be highly challenging. Nevertheless, these early works laid out the foundation of DEL, which included the basic principle, the synthesis, encoding, selection and decoding strategies.

This field remained mostly dormant until 2004, when Neri[16], Liu[17], Harbury[18], Winssinger[19] and their respective co-workers independently reported four types of encoded library: the dual-pharmacophore encoded self-assembling chemical library, the DNA-templated synthesis library, the DNA-routing library and the peptide nucleic acid (PNA)-encoded library, respectively. These libraries are in solution and the compounds are encoded by either DNA translation, routing or ligation, and thereby obviate the limitations of OBOC libraries. In 2009, GlaxoSmithKline (GSK) published a seminal work on applying DELs at an industrial scale[20]. Following these landmark works, DELs entered a stage of rapid developments. First, the availability of low-cost, genomic-scale next-generation sequencing (NGS) technologies is a key factor that enabled DELs to reach a multibillion-compound scale; second, the development of DEL-compatible chemistry hugely expanded the chemical space for DELs[6,7,9,21] and third, recent innovations integrated DEL with other legacy and emerging techniques, such as fragment-based drug discovery[16,22–25], dynamic combinatorial library[26–32], diversity-oriented synthesis[33], machine learning[34–36], OBOC libraries[37,38], microfluidics[39–41] and flow cytometry[15,42,43]. The productivity of DELs is demonstrated by the clinical candidates that originated from DEL selections[44]. Notable examples include GSK2256294, an epoxide hydrolase inhibitor to treat pulmonary disease[44,45], and GSK2982772, a first-in-class receptor-interacting protein-1 kinase inhibitor to treat inflammatory diseases, both developed by GSK[44,46]. Recently, X-Chem reported an autotaxin inhibitor X-165 as the clinical candidate for pulmonary fibrosis[47].

Albeit on a massive scale, DEL selection is fundamentally a binding assay, typically performed by incubating the library with a purified protein on a matrix. After washes, the binders are eluted from

[1]Department of Chemistry and State Key Laboratory of Synthetic Chemistry, The University of Hong Kong, Hong Kong SAR, China. [2]Chongqing Key Laboratory of Natural Product Synthesis and Drug Research, School of Pharmaceutical Sciences, Chongqing University, Chongqing, China. [3]Chemical Biology Research Center, School of Pharmaceutical Sciences, Chongqing University, Chongqing, China. [4]Laboratory for Synthetic Chemistry and Chemical Biology Limited, Health@InnoHK, Innovation and Technology Commission, Hong Kong SAR, China. ✉e-mail: yizhouli@cqu.edu.cn; xiaoyuli@hku.hk

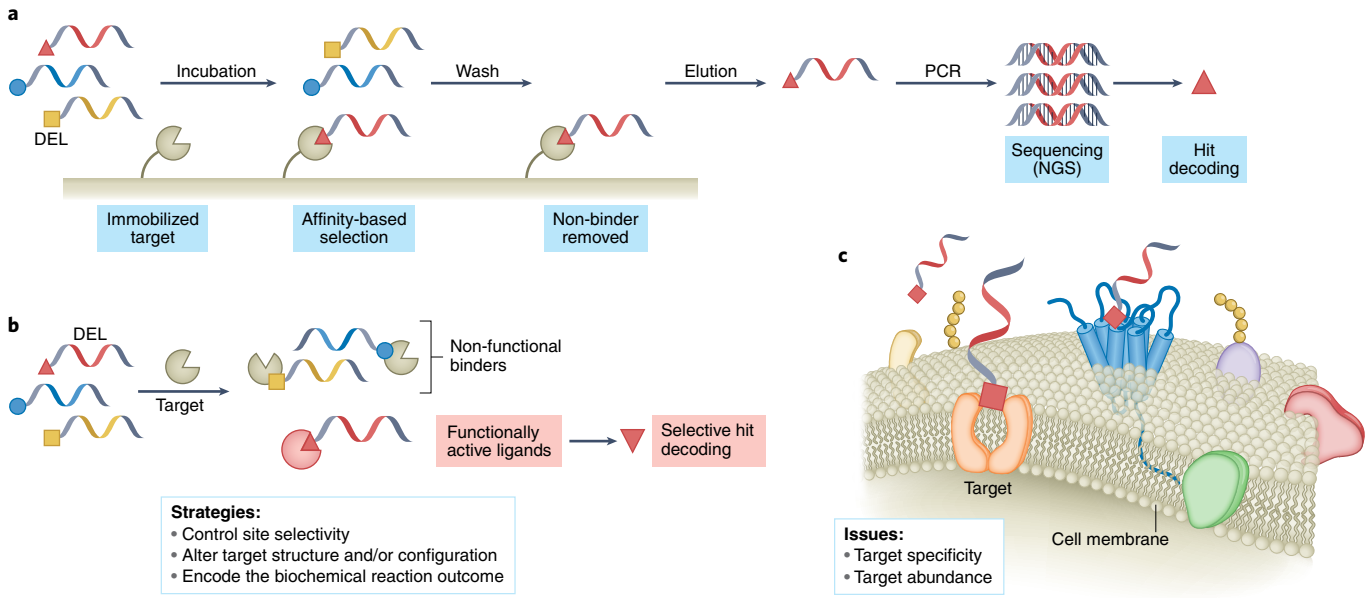

Fig. 1 | General scheme and key considerations in realizing functional DEL assays. a, Scheme for a typical DEL selection. A DEL is incubated with an immobilized target, the non-binders are washed away and the binders are eluted under denaturing conditions for hit identification (PCR and sequencing). **b**, Controlling selection site-selectivity, altering the protein's structure and/or configuration and using DNA to encode the biochemical reaction outcomes are viable strategies for functional DEL selections. Details are provided in Fig. 5. **c**, Target specificity and target abundance are the two major issues of DEL selections against complex biological targets. For example, DEL selections with live cells may be interfered with by other cell surface biomolecules and limited by the low expression level of membrane proteins.

the target and decoded with PCR amplification and NGS (Fig. 1a). Such a simplicity means DELs can be used for almost any target that can be purified and/or immobilized, which include ones without a known ligand or prior structural knowledge. The broad target scope is one of the main reasons that DELs are being widely adopted in drug discovery. Remarkably, DELs have been successfully applied to stabilized membrane proteins[48–50] and a large protein complex by focusing on the domain of interest[51]. However, solid-phase-based selection has two limitations. First, physical binding does not always elicit a biological response, for example, the ligands may not bind to the catalytic site, or the binding affinity may not quantitatively correlate with the activity. A post-selection validation of the biological activity of the hit compounds is always required. Second, proteins may be denatured upon immobilization or intractable to purification and/or immobilization, such as some protein complexes, membrane proteins, live cells; moreover, purified proteins may lose biological features, such as post-translational modifications, peripheral steric hindrance, native charges, complex formation and co-factor binding. Performing DEL selections in a native biological environment is therefore highly desirable, as it may identify biologically more relevant ligands with higher potentials to become drug candidates[52,53].

Reported strategies to address the first limitation include to control the site selectivity, alter the target's configuration and use DNA to encode the outcome of the binding event (Fig. 1b). For the second limitation, achieving target specificity in a complex biological background is the main challenge. DEL selection requires a relatively high target concentration to drive the binding equilibrium; thus, target abundance may be another issue (Fig. 1c). Encouragingly, innovative approaches were developed that not only expanded the target scope but also revealed the potential of DELs as a powerful tool to explore biological systems. Here, we review the evolution of DEL selection methods, but focus on those for complex biological targets; then, we discuss the emerging approaches to realize functional DEL assays; finally, we present our view on the present challenges and future directions of DELs. For more in-depth discussions, recent literature offers many excellent reviews that cover every aspect of DELs[5–11,21,52–59].

Evolution of DEL selection

Early selections. Early DELs were built on a solid phase (OBOC–DELs)[13–15]. The selection was conducted by incubating the library with an in-solution target, and the target-bound beads were identified with a secondary antibody[15]. Today, most DELs are in the solution phase, probably because of the ease of the 'split–mix' synthesis method and that much larger libraries could be prepared. However, OBOC–DELs were recently revived and modernized with sophisticated bead design, novel on-bead chemistry and advanced instrumentation, which has led to many novel applications of DEL[7,8,53]. This is discussed in a later section.

Selection with immobilized targets. Today, DELs are commonly prepared in solution and the selection is performed with an immobilized target based on binding affinity or the target could be incubated with the library first and then immobilized. The selection requires a careful balance of the thermodynamic association and kinetic disassociation parameters to remove the non-binders while retaining the binders. In practice, finding an optimal protocol is difficult due to many confounding factors (DNA copy numbers, washing conditions, buffers, temperature and so on) and is subject to considerable variations among experimenters. Many techniques are employed for better specificity, such as competitive elution, target titration, non-target controls, selection replicates and so on. Quantitative PCR is frequently used as a measure for selection quality control[60–64], to determine the low limit of DNA copy number[61,65] and to model the library population[66]. Computational methods are also useful for researchers to optimize the selection conditions. For example, it was shown that selections over a range of target concentrations could improve the correlation between the equilibrium association constant and the hit compound[67]; by analysing a

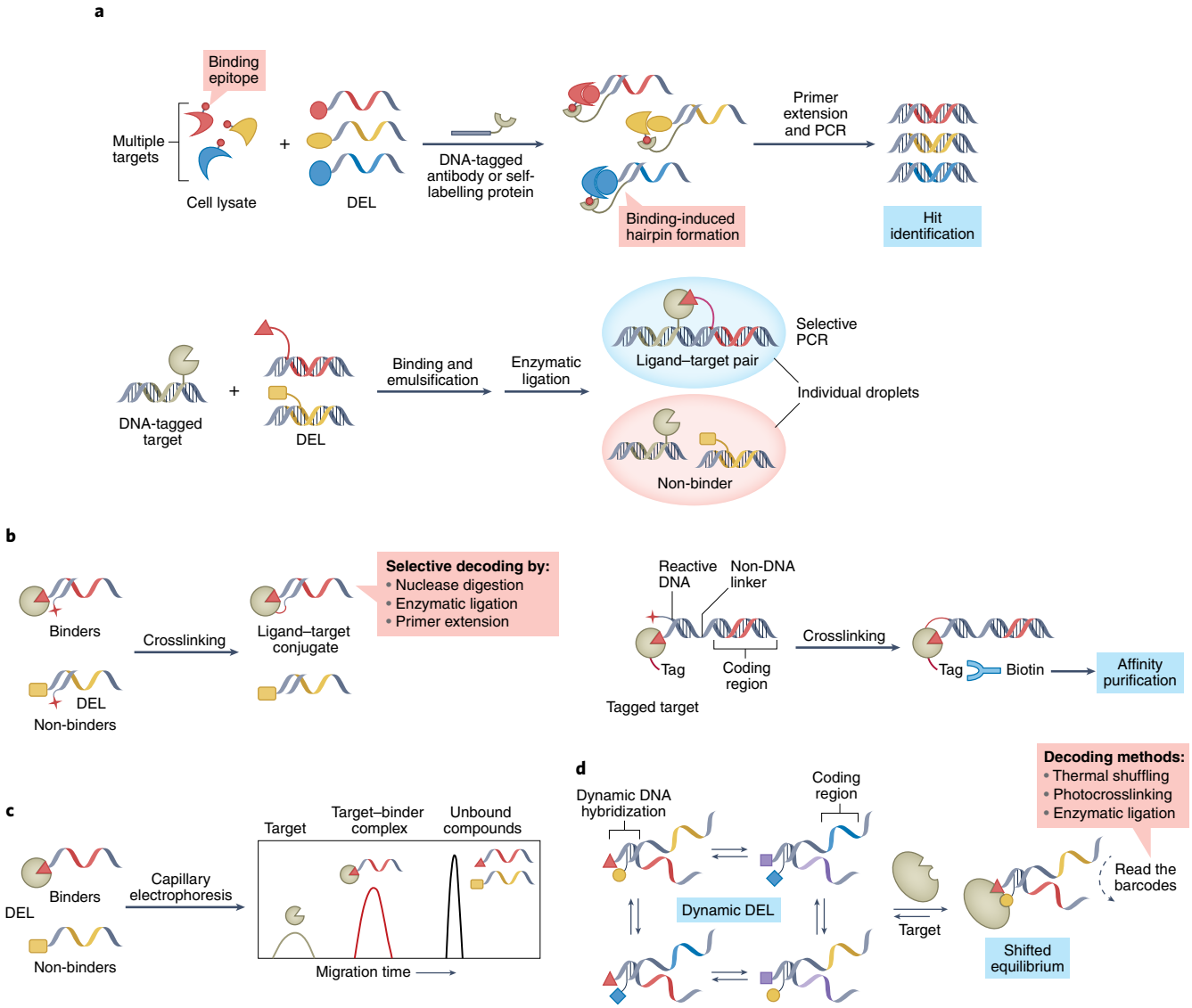

Fig. 2 | In-solution selection methods. a, IDUP (top) and BTE (bottom) both modify the target with a DNA tag, which enables the formation of an amplifiable DNA–target complex on target–ligand binding, whereas the non-binders do not form such a complex (IDUP is through hairpin formation and primer extension; BTE is through water–oil emulsification and ligation). **b**, Crosslinking-based methods establish a covalent linkage between the target and the binders. Left: the DNA tags of the binders are protected by the protein molecule and selectively amplified. Right: the complex of target–binder–crosslinking–DNA is affinity purified for hit identification. **c**, Kinetic CE may directly separate the target-binder complex because of its non-denaturing condition and the size and/or charge differences among the equilibrating species. **d**, The selection of dynamic DELs relies on the formation of short and unstable duplex regions between the two sets of DNA-encoded small molecules; target addition shifts the equilibrium, which could be stabilized or 'frozen' with various techniques; a comparison of the equilibria with and without the target could identify the binding small molecule pairs. Red star, crosslinker.

large set of selection data, a recent study indicated $>10^6$ copies per compound is needed in a selection, whereas the library size that exceeded 10^8 compounds would lead to many false positives[68].

Selection with non-immobilized targets. Conducting DEL selections in solution is necessary to expand the target scope, and the key issue is the differentiation of the binders from the non-binders without physical washes. The methods for in-solution selections can be summarized as follows. (1) For selective amplification and/or enrichment of the binders, Liu and co-workers developed an interaction determination using unpurified proteins (IDUP) method based on the enhanced stability of the pseudo-hairpin structure formed on target-ligand binding[69]. IDUP could operate in a multiplexed format in cell lysates (Fig. 2a, top)[70]. Using an emulsion system, Vipergen developed a binder trap enrichment (BTE) method by isolating individual target–ligand complexes in droplets (Fig. 2a, bottom)[71]; BTE has been used to select against targets in buffer[72] and inside live cells (Fig. 4b)[73]. (2) Covalent crosslinking, which aims to establish a stable linkage between the target and the ligand to be reliably isolated for hit identification. Li and co-workers developed several photocrosslinking-based methods, in which the target–ligand binding complexes were selectively amplified after nuclease digestion, enzymatic ligation or primer extension (Fig. 2b, left)[74–76]. However, these methods are limited to DELs encoded with single-stranded DNAs. The Krusemark group hybridized a reactive DNA strand at a single-stranded DNA segment separated from the encoding DNA, so that the method is suitable for DELs with double-stranded DNAs, the most commonly used DEL format (Fig. 2b, right)[63]. Similar strategies were also used in affinity-based protein profiling (ABPP), in which PNA-encoded probes were used

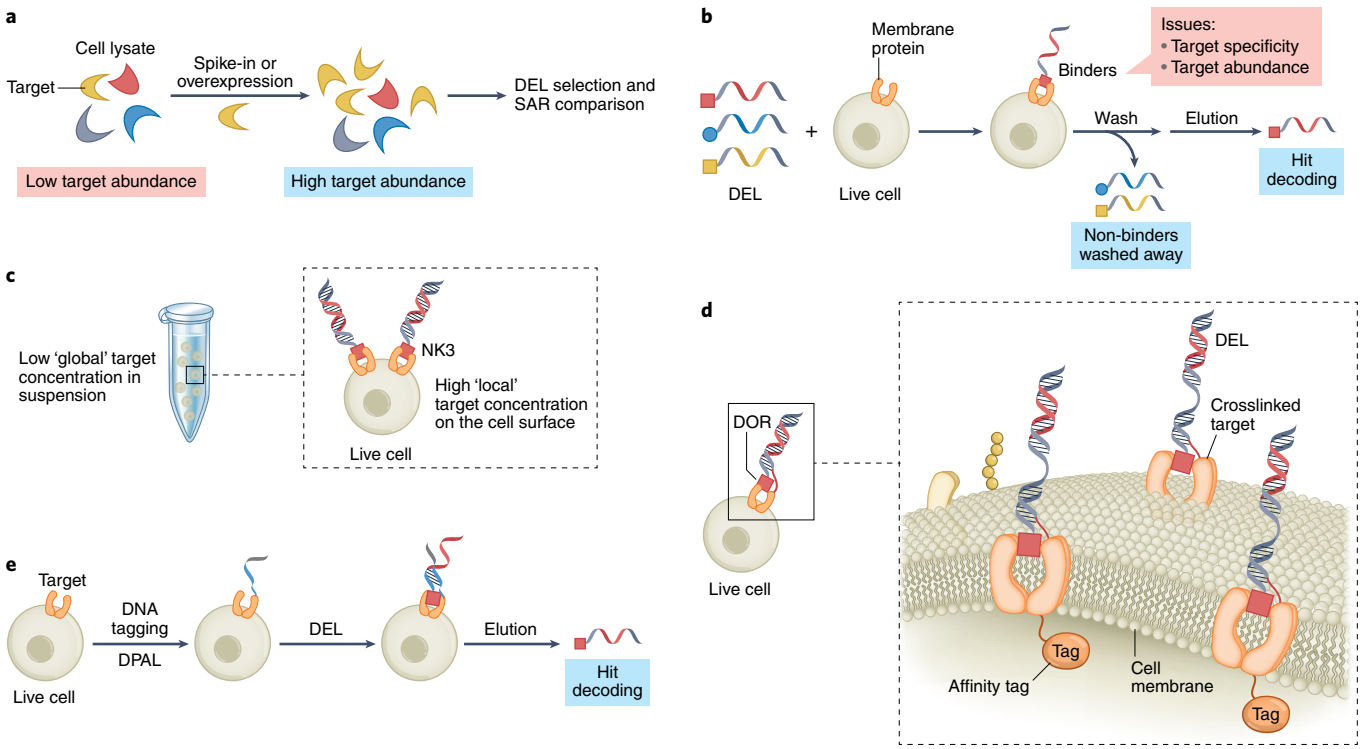

Fig. 3 | Methods for DEL selections in cell lysates and on live cells. a, Spiking in or overexpressing the target protein could realize target specificity and increase the target abundance for in-lysate selections. **b**, In principle, cell-based selection is amenable to the typical affinity-based protocol, as shown in Fig. 1a, but target specificity and target abundance remain to be the issues. **c**, Overexpressing the membrane protein NK3 enabled cell-based DEL selections; notably, the effective 'local' concentration of NK3 on each cell is markedly higher than the 'global' concentration in the cell suspension. **d**, Crosslinking-based DEL selection method (Fig. 2b, right) was used to interrogate the membrane protein target DOR on live cells; the crosslinked target–binder complex was affinity purified for hit deconvolution. **e**, DPAL can specifically install a DNA tag on the target protein on live cells; the DNA tag guides the hybridization of the library specifically to the target, and thereby realize target specificity and increase target abundance. This method does not overexpress or modify the target and can be applied to endogenous membrane proteins.

to interrogate the whole proteome of cells[77,78]; the ABPP probes had an electrophilic 'warhead' that can crosslink to the targets for protein identification. (3) In kinetic separation, the Krylov group pioneered methods to separate the target–ligand complexes from the non-binders by using kinetic capillary electrophoresis (CE)[79,80]. The principle is simple: the target-bound binders have smaller k_{off} (dissociation rate constant) values than those of the non-binders, so the binders are eluted with the protein molecule under the mild conditions of CE (Fig. 2c). However, this method has only been tested with model systems and its effectiveness with large DELs remains to be investigated. (4) For dynamic DELs, the Neri group developed a dual-pharmacophore encoded self-assembling chemical library method[16,22], and the Hamilton group used protein templates to direct the dynamic hybridization of DNA duplexes[81]. The combination of the two concepts was elaborated into several DNA-encoded dynamic libraries[26–32]. With dynamic DELs, non-immobilized targets bind to the ligand pairs and promote the formation of DNA duplexes or three-way junctions, which could be isolated and/or enriched for hit identification (Fig. 2d).

All these methods (except (2)) follow the same principle: irreversibly capture or stabilize the target–binder complex. Studies showed that covalent crosslinking improved the enrichment of both high- and low-affinity ligands compared with those of solid-phase selection[63,64], which allows a target concentration lower than that of the ligand's K_d (equilibrium dissociation constant) value in a selection. The Neri group compared the selections of a macrocycle DEL using on-beads washes, antibody competition or in-solution photocrosslinking, and the photocrosslinking method identified more structure–activity relationship features[82]; similar phenomenon was also observed with reversible, non-covalent 'crosslinkers'[83].

DEL selections against complex targets

Cell lysates. A simple solution to achieve target specificity in cell lysates is to increase the target abundance by spiking-in or overexpression (Fig. 3a). The selection can be compared with a control without the target added-in or overexpressed to distinguish the specific binders. The elevated target concentration also facilitates ligand binding. The IDUP method overexpressed the targets and conducted the selections in cell lysates[69,70]; the selective nuclease digestion method gave an example in lysates (Fig. 2a). In principle, other crosslinking-based methods may also be suitable with cell lysates.

Membrane proteins. Membrane proteins play key roles in the pathobiology of numerous diseases. However, ligand discovery for membrane proteins is notoriously difficult due to their large size, hydrophobicity and the dependence on the lipid bilayer environment of the cell membrane. Previously, DELs were applied to the soluble domains of membrane proteins[22,84–87] and full-length membrane proteins stabilized with detergent[48], nanodisks[49] or mutations[50]. However, selections with membrane proteins should ideally be conducted with live cells (Fig. 3b)[52,53]. This seemingly looks like a straightforward task: many cell lines are adherent and could be directly used in affinity-based selection. For suspension cells, the washes could be done by centrifugation. However, target specificity and concentration are still the main issues. For membrane proteins,

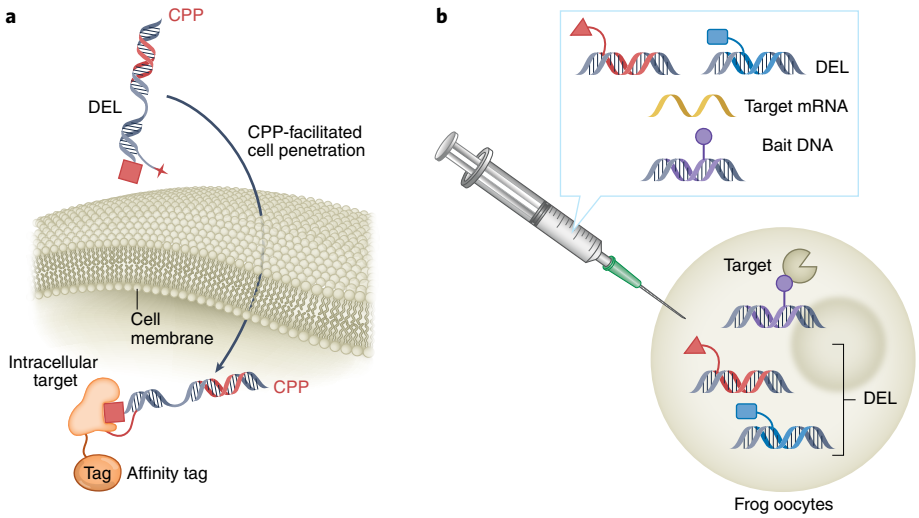

Fig. 4 | Strategies to deliver DELs into live cells for intracellular selections. a, Conjugation of a CPP to the DEL facilitated cell entry; the crosslinking-based method in Fig. 2b (left) was employed for the selection targeting the intracellular proteins. **b**, DEL could be directly injected into very large cells (for example, frog oocytes); the mRNA of the target and a bait DNA were also injected to form the amplifiable target–binder complex, as shown in Fig. 2a (bottom).

target abundance is particularly important. The effective molarity of membrane proteins on live cells could be estimated based on the number of receptors, cell size and volume of the cell suspension. We studied several membrane proteins and found that they have low nanomolar or subnanomolar effective molarities, lower than the micromolar concentration typically needed in DEL selections[88].

Similarly, the target specificity and/or abundance issues could be addressed by protein overexpression. The Bradley group pioneered live-cell-based selections of encoded libraries[89,90]. They incubated a PNA-encoded decapeptide library with cells that overexpressed integrins or a chemokine receptor. After washes, the cell-surface-bound peptides were collected and decoded with a DNA microarray[89,90]. In 2015, GSK reported a DEL selection with a NK3 tachykinin receptor on HEK293 cells[91]. The target overexpression reached ~500,000 receptors per cell, and 10^7 cells were used in a 1.0 ml suspension; thus, the 'global' target concentration was ~8.3 nM, whereas the 'local' effective concentration on each cell should be much higher, as the receptors are concentrated within a much smaller volume (Fig. 3c). This study showed that the comparison with a control without overexpression clearly identified a target-specific structure–activity relationship in the selection fingerprint. In 2019, Krusemark and co-workers reported a DEL selection against the δ-opioid receptor (DOR) on live cells (Fig. 3d)[92]. The DOR was fused with a self-labelling tag, and the selection was performed using their previously reported crosslinking method[63]. After cell lysis, the DOR–ligand conjugate was affinity purified with a biotin probe for hit identification. This study did not estimate the number of the receptors per cell, but the cell density was high (2×10^6 cells per 50 µl), which suggests a high effective concentration of the target. Recently, we reported a selection method without target overexpression or genetic tagging[88]. Using DNA-programmed affinity labelling (DPAL)[93], the target was labelled with a DNA tag. The tag can guide the hybridization of DEL near the protein, and thereby achieve target specificity and a high target concentration (Fig. 3e). Three membrane proteins were tested—folate receptor, carbonic anhydrase 12 and epidermal growth factor receptor—and the nanomolar to micromolar binders were identified. We showed that small molecules and antibodies could be employed for target tagging. Other types of ligand that are used to graft DNAs on cells, such as peptides[94], nanobodies[95] and aptamers[96], may also be used. Notably, the length of the DNA tag determines the affinity increase of the ligands, which could tune the selection stringency. However, a tag longer than ten bases was too stable and resulted in a high background, whereas a tag shorter than six bases would have hybridization-specificity issue[88]. More studies are needed to fine-tune the free energy gain from the hybridization. Finally, Neri and co-workers recently optimized the experimental conditions for cell-based selections[97]. Interestingly, they observed that bivalent ligands improved the enrichment, which suggests that multivalency effect may be another way to drive ligand binding on the cell surface.

Intracellular selection. The ability to conduct DEL selections intracellularly is highly attractive, as it brings the possibility to interrogate proteins in an environment that preserves their biological features and functional and/or structural dynamics. The first intracellular DEL selection was demonstrated by the Krusemark group (Fig. 4a)[92] using the same method for membrane proteins (Fig. 3d), but a cell-penetrating peptide (CPP) was conjugated to the library to facilitate cell penetration. Again, target overexpression and tag fusion were necessary. Two proteins, chromobox protein homologue 7 and dihydrofolate reductase, were tested; the results showed a differential enrichment of compounds whose structures are consistent with a known structure–activity relationship. Vipergen reported an intracellular version of the BTE selection[73]. The DEL was microinjected into a frog oocyte (Fig. 4b), which is >100,000 times bigger than a typical somatic cell. Additionally, an mRNA of the target and a 'bait' DNA, designed to be compatible with the BTE protocol[71], were injected for target overexpression and tag fusion. Three proteins, p38α, ACSS2 and DOCK5, were tested, and a hit validation was conducted for p38α[72].

Library delivery is expected to be the main obstacle for intracellular DEL selections, but it has been circumvented by using CPP or microinjection. Other delivery methods, such as those for antisense oligos, may also be used. It is recommended that quantitative PCR be used to quantify the delivery efficiency and ensure that the copy number of the compounds is above the reliability threshold[61,65]. However, we anticipate other confounding factors for intracellular selections. The crowding effect inside cells may magnify non-specific interactions and increase the background noise. Although fusing the target with a tag should only pull down the specific binders, it is still unclear whether the cellular environment would lead to excessive background binding and reduce the enrichment of the specific binders. The cellular stability of DNA may be another issue, which could lower the DNA copy numbers available for target binding.

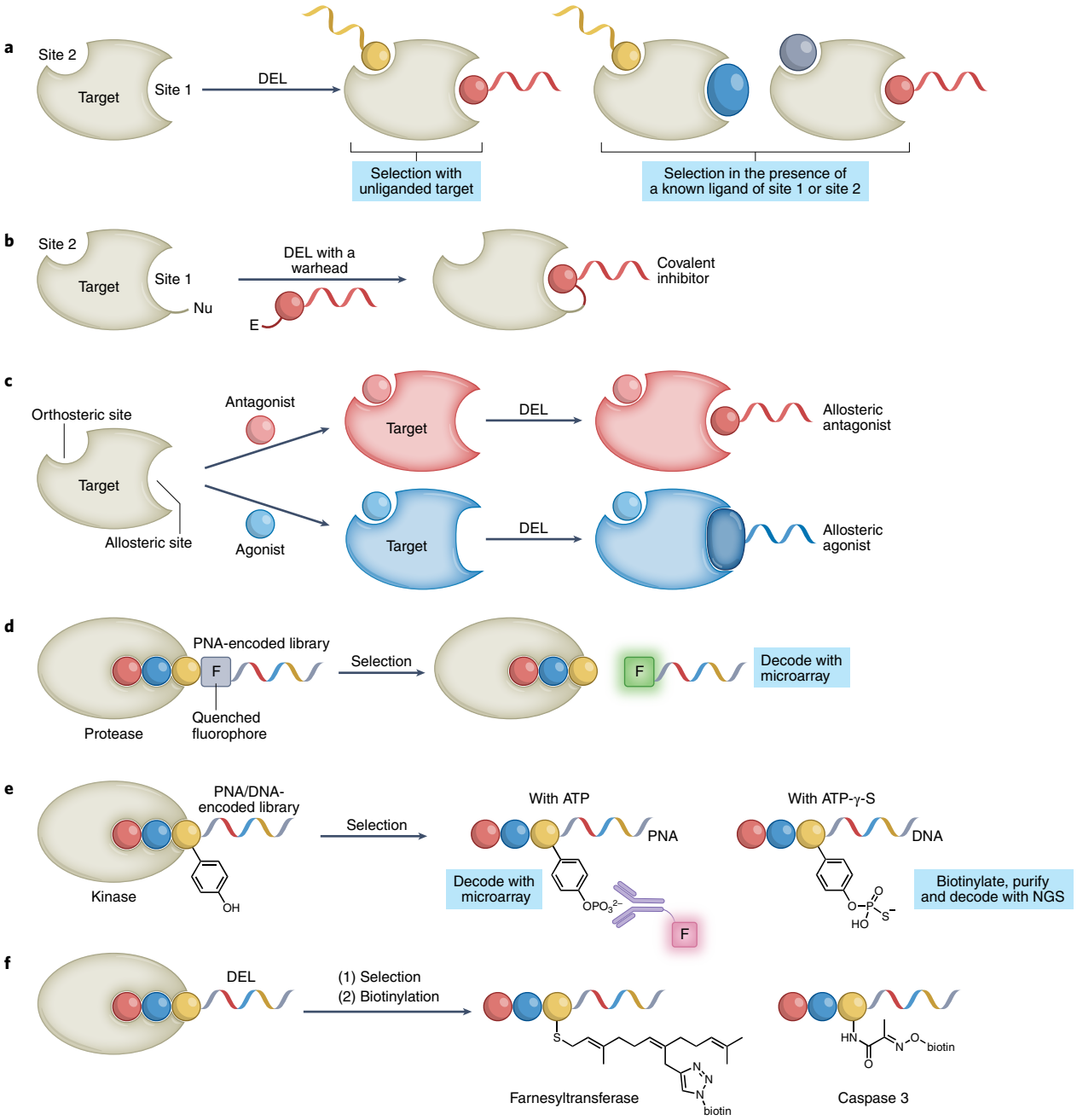

Fig. 5 | Strategies for functional DEL assays. a, Conducting DEL selections with and without a known ligand may identify binders with site selectivity; for example, using a known orthosteric ligand (site 1) may identify allosteric binders (site 2), and vice versa. **b**, In a DEL selection for covalent inhibitors, the nucleophile (Nu) on the protein may direct the library to be sampled at the active site (an electrophile, E) and identify functionally active ligands. **c**, Using a known antagonist or agonist may stabilize the protein structure and/or configuration, and so bias the binding of the allosteric antagonists or agonists. **d**, The protease target may not only bind to but also cleave the peptide substrates in the library; this process also releases the fluorescent PNA tag of the substrate peptides, which could be decoded with a microarray. **e**, The kinase target may phosphorylate the tyrosine motif in the substrate compounds, which could be specifically isolated for hit identification. **f**, Similarly, for farnesyltransferase and caspases 3, the modified substrates may be further biotinylated and isolated for hit identification.

We observed that short, small-molecule-conjugated DNAs were quite stable in cell lysates[93], but the stability of the DNA tags in live cells should be investigated. Furthermore, organelle-specific library delivery and selection is a more challenging, but perhaps more salient, task for future endeavours.

Nucleic acid targets. In principle, DNA/RNAs are suitable DEL targets as they may present folded structures with ligandable sites. However, there is only one report on DEL selections against a G-quartet sequence in the *c-myc* promoter DNA[98]. A recent review discussed the potential of targeting functional RNAs with DELs[53] and one patent described a way to apply DELs to RNAs[99], but there is no peer-reviewed report yet. Obviously, an important issue of targeting DNA and/or RNAs is the DNA tag, which should have no complementarity to the target with minimal folding. Thus, double-stranded DNA-encoded DELs may be more suitable, as double-stranded DNA has a rigid structure and the nucleobases are stably paired. Known DNA- and/or RNA-binding features, for

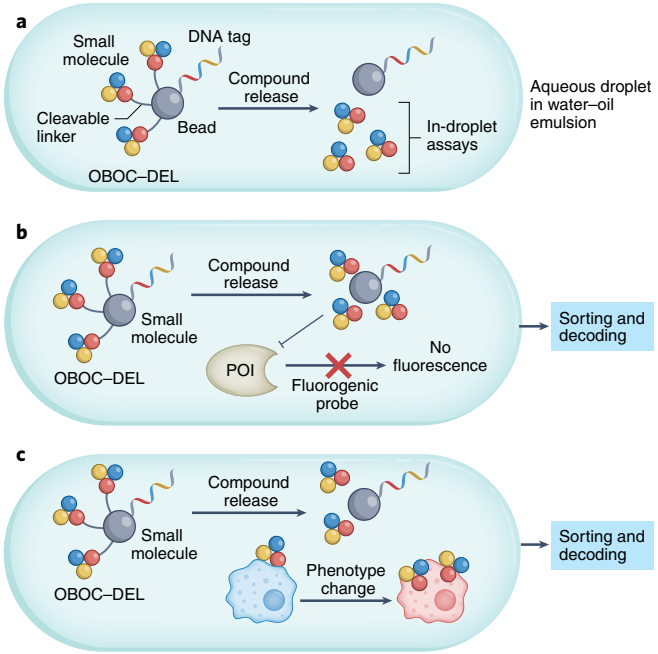

Fig. 6 | OBOC–DELs enable novel functional DEL assays. The cylindrical enclosure represents the droplet in an emulsion. **a**, General design of OBOC–DEL assays; small molecules are released from the beads but confined within the droplet along with the encoding DNA tag. Other necessary components, for example, the target, co-factor, probe and so on, may be included to establish the assay. **b**, Biochemical OBOC–DEL assay. A fluorogenic probe is included in the droplet; if the released compound is an inhibitor, no or low fluorescence would be observed for that droplet. **c**, The proposed concept for a phenotypic OBOC–DEL assay. If the released compound could induce a phenotypic change of the cell, it could be identified with a microscope and the droplet could be isolated for hit identification.

example, positive charges, base intercalators and group binders, may still be incorporated in the library, but care should be given so that they do not overwhelm target-specific interactions. Experimental conditions, for example, the ionic strength, should be carefully controlled to alleviate the charge repulsion between the DNA tag and the target without disrupting the folded structure of the target.

Other targets and outlook. Many proteins exist in the form of protein complexes in cells, and their biological functions closely depend on the composition and dynamics of the complex. DEL selections may potentially identify 'complex-specific' ligands that could modulate the functions of the entire complex. Some protein complexes are stable and may be amenable to affinity-based selections. There is no reported DEL selection with a whole protein complex, but a recent study showed an alternative approach by focusing on the domain of interest in the complex[51]. Recently, we reported the profiling of histone deacetylase complexes using DNA-encoded probes, and suggested the possibility of DEL selection with endogenous protein complexes in cell lysates[100].

Although with limited examples, it is fair to say that DEL selection is possible with complex biological targets. The recently reported DEL selections on and inside live cells are particularly encouraging. However, the research is still in the proof-of-principle stage and mostly limited to model systems. More applications with a wider range of proteins and larger DELs would truly test the generality and performance of DEL selections beyond purified proteins. Moreover, after the ligands were identified, they still needed to be validated with purified proteins. It would be very exciting if cell- and/or lysate-based selections could discover the ligands that preferably bind to the endogenous target, but could not be identified with purified proteins. Finally, whether DEL is applicable to primary cells, organoids and tissue samples remains to be investigated. We expect future innovations will answer these questions and reveal more potential for DELs in biological systems.

DEL selections without a designated target

In ABPP, PNA-encoded probes have been used for protein profiling in cell lysates or even in the whole organism[77]. The probes were not targeting a specific protein but a class of enzymes. Non-encoded OBOC libraries have also been used to target live cells without a specific target. The identified ligands were expected to bind more favourably to a specific cell type[53,101]; in many cases, the receptors responsible for the binding are unknown. Cell-SELEX is a powerful technology for evolving cell-specific aptamers, and the targets of the aptamers could be further identified and recognized as cell-surface biomarkers[102]. In principle, DELs should be suitable for such types of whole-cell selection without a specific target; the identified ligands could be used for drug delivery or diagnostic applications. However, this direction remained largely unexplored. In 2011, the Bradley group used a PNA-encoded library to identify peptides for selective cellular delivery[103]. In 2017, using a two-colour screening strategy, Kodadek and co-workers selected an OBOC–DEL against the sera from the patients of active and latent tuberculosis. The selections did not have a specific target but aimed at probing the IgG antibody difference between the two populations. This study discovered the ligands that can distinguish the two disease states and identified the underlying tuberculosis antigen Ag85B. Kolodny et al. published a hypothesis paper that proposed a cancer theranostic approach by using DELs to identify 'ligand arrays' that bind to whole cancer cells or tissue samples. They proposed that there is no need to know the targets or even the ligands' structures and the DNA tag of the ligand array could be replaced with radiolabels to selectively kill the cancer cells[104]. Although just a concept, it might lead to a new direction for DEL applications in personalized medicine.

Functional DEL assays

In general, DEL is superior to traditional HTS regarding library size, cost and accessibility, but HTS is more versatile and can be adapted to various biochemical and phenotypic assays. Hence, a major direction of DEL research is to develop new methodologies that go beyond binding assays[52,53]. In HTS, the compounds are spatially encoded so that the signal readout can be attributed to the specific compound; thus, for DELs, the question is how to connect the readout with the encoding DNA of the active compounds, not just the binders. We summarized four types of strategy in the following.

Direct the selection to a functionally relevant site. Biologically active compounds require binding to a specific site or interface on the target to exert their functions, for example, to block ligand–receptor interaction, elicit agonistic activities on a receptor or recruit two proteins to form a complex. DEL selection is generally not site selective, but including known ligands in the selection may provide site selectivity (Fig. 5a). Winssinger and co-workers selected a PNA-encoded fragment library with and without ATP against the target HSP70 and identified ATP-competitive ligands[23]. Later, DEL selections with and without ATP or an active site inhibitor identified the ligands that bound to different sites on Bruton's tyrosine kinase[105]. We also used known ligands to direct the selection of dynamic DELs towards the active site of the target[30].

The binding mechanism of covalent inhibitors offers favourable pharmacokinetic properties. As covalent ligands often bind to the sites with properly positioned nucleophilic side chains, selections for covalent ligands may have site selectivity and functional relevance (Fig. 5b). However, this requires synthesizing the library with reactive warheads and the selections are performed under

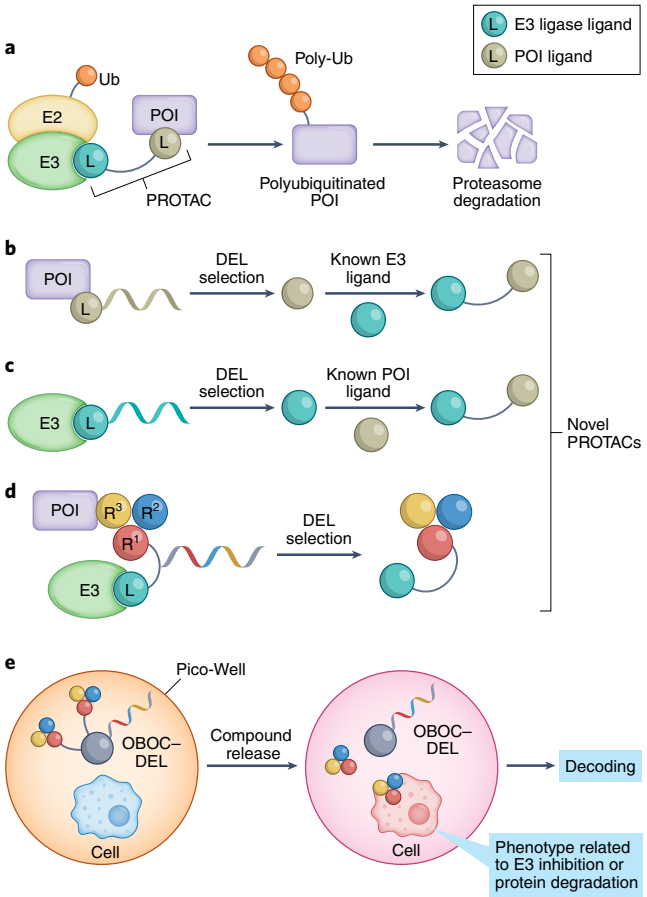

Fig. 7 | Strategies to combine DEL with PROTAC. a, General scheme for PROTAC-mediated target degradation; the POI is ubiquitinated by the E2–E3 ligase complex through a PROTAC molecule; the ubiquitinated POI is further degraded by the proteasome. Ub, ubiquitin. **b**, DEL is used to select for the POI ligand, and then it can be tethered with a known E3 ligase ligand to form the PROTAC. **c**, Conversely, DEL is used to identify the E3 ligase ligand first, and then it is tethered with a known POI ligand to form the PROTAC. **d**, A DEL could be synthesized with a built-in E3 ligase ligand to select for PROTACs directly. **e**, The proposed OBOC–DEL assay for PROTAC. Cells could be engineered with a detectable phenotype specific for E3 ligase binding or degradation profiles of E3 ligases; thus, if the released compound could induce such a phenotype change, the compound is more likely to be an active PROTAC in the cells.

strong conditions, such as using detergent[106] or heating[107], to remove non-covalent binders. Using warhead-modified DELs, covalent inhibitors were identified for c-Jun *N*-terminal kinase[108], 3C protease[107], bromodomains[109], receptor tyrosine kinases[106], MAP2K6 kinase[70] and Bruton's tyrosine kinase[110].

Control the target's configuration. On binding to active ligands, such as agonists or antagonists, a protein's structure can be stabilized in the active or inactive form and the corresponding physical binders are more likely to be active. This can be used to bias the selection towards active ligands (Fig. 5c). By using known orthosteric antagonists in DEL selections, novel allosteric antagonists were identified for membrane protein protease-activated receptor 2 (ref. [50]). Using a known agonist, Lefkowitz and co-workers identified novel allosteric agonists for β_2-adrenoceptor[49]. These studies demonstrated that altering the state of the target protein could facilitate the identification of functionally relevant compounds. If its effectiveness could be proved with live cells, its impact in drug discovery would be tremendous. However, this approach is limited by the nature of the protein and the availability of the known ligands.

Use DNA to encode the outcome of biochemical reactions. A biochemical DEL assay should only amplify the DNA tag of the active compounds, not all the physical binders. This could be achieved by taking advantage of the structural changes of the compounds after the biochemical transformation mediated by the protein. Previous ABPP studies provided several strategies. For example, PNA-encoded peptides were incubated with a protease target. The peptide structure has a fluorogenic motif, so that only the enzyme substrates, not all the binders, generate a fluorescence signal, which can be attributed to individual PNA tags by microarray analysis (Fig. 5d)[111,112]. For kinase, the PNA-encoded peptides, once phosphorylated by the target, could be detected by an antiphospholipid antibody (Fig. 5e)[78,113]. These strategies have already been adopted in biochemical DEL assays. Harbury and co-workers selected a peptide DEL against protein kinase A[114]. The selection was performed in the presence ATP-γ-S (Fig. 5e); thus, the substrates would be thiophosphorylated and affinity purified by using biotinylating reagents; later, Krusemark and co-workers expanded this strategy to farnesyltransferase, caspase and c-Src kinase (Fig. 5f)[62,115]. Although the DNA tag still encoded the compounds' structures, the assays were designed to selectively decode the active compounds; thus, the DNA tag, in effect, encoded the outcome of the biochemical reactions.

OBOC–DELs. In HTS, the signal readout can be easily connected to the compound's identity. By using advanced instrumentation, OBOC–DELs can separate individual beads (that is, individual compounds) in aqueous droplets. Modern OBOC–DELs also feature a photocleavable linker so that the compounds can be released as free ligands for unobstructed 'off-DNA' assays within the droplet[40,41]. The signal readout could be analysed with flow cytometry[15,42,116] to sort the beads for hit decoding (Fig. 6a). These features of OBOC–DELs, supported by robust on-bead chemistry[42,116], have led to novel applications of DELs. Paegel and co-workers developed an OBOC–DEL with a multifunctional linker. The library was selected against cathepsin D within individual droplets that contained the target, a fluorogenic peptide, and an internal standard. Library compounds were photocleaved from the beads and sampled by the target in the same droplet. The droplet with inhibitors showed a low fluorescence due to the inhibition of the target's protease activity (Fig. 6b)[117]. Using a similar approach, the group identified potent inhibitors of the phosphodiesterase autotaxin using a 67,100-member OBOC–DEL[40]. The same library was also used to discover new inhibitors for the receptor tyrosine kinase DDR1 with an in-droplet fluorescence polarization assay[41]. These studies demonstrated that OBOC–DEL combines the advantages of HTS with DEL and is amenable to various formats of biochemical assays.

Recently, the possibility of a phenotypic DEL assay was proposed (Fig. 6c)[52,53]. Again, separating compounds in droplets connects the phenotype change to the compound's identity, and releasing the compound from the beads addresses the cell penetration problem. Previously, non-encoded OBOC libraries were used to interrogate live cells[53,101,118]. Screening of OBOC–DELs with live cells has yet to be reported, but it should be feasible, given the prior success in human sera. The high effective molarity of the compounds within the droplet should also be beneficial[88]. It is conceivable that various formats of phenotypic DEL assays will be developed in the near future. In fact, patent publications already disclose clever designs of functional OBOC–DEL assays for various applications[52]. Phenotypic assays examine the holistic changes of the cells or organisms without prior knowledge of the affected targets and pathways. Target identification is required after the active compounds are identified. Covalent DELs may facilitate the target identification process

by irreversibly capturing the target[53]. For non-covalent DELs, this may be achieved by using crosslinking-based methods (Fig. 2b) or DPAL[93]. However, OBOC–DELs require sophisticated instrumentation and a high technical threshold. Moreover, after being released from beads, compounds may diffuse out of the droplet and result in false positives, which further complicates the library selection process[101]. OBOC–DELs are smaller than solution-phase DELs due to the practical limit of the number of beads that can be handled. In addition, developing a bead-partitioning system that maintains cell viability may be challenging. These issues are surely being addressed and we are confident to see more breakthroughs in near future.

DEL and PROTAC

Proteolysis targeting chimera (PROTAC) is an emerging technology that has attracted huge interest in drug discovery. PROTAC works by hijacking the cell's natural ubiquitin–proteasome system[119]; PROTACs are heterobifunctional molecules that connect the protein of interest (POI) with an E3 ubiquitin ligase, and thereby promote the ubiquitination and proteasome degradation of the POI (Fig. 7a). In principle, DEL should be complementary to PROTAC, and a union of the two may produce major impacts on drug discovery[9,119]. First, DEL selections identify physical binders, which is exactly what PROTAC needs. Traditional biochemical screening will miss the ligands that bind to the target without a biological effect, but they could be identified with DELs and then used for PROTACs. Second, the DNA conjugation site of DEL compounds could be conveniently used as the chimera-linking site. Third, DEL selections often identify many micromolar binders, which are not potent enough to elicit a biological consequence and thus discarded, but such a modest and perhaps transient binding may be sufficient to induce protein degradation[119]. GSK selected a macrocycle DEL and identified novel ligands for VHL, an E3 ubiquitin ligase[120]. X-Chem developed oestrogen receptor α degraders by using DELs to identify oestrogen receptor α binders and then tethering them with known E3 ligands[121]. There is no peer-reviewed publication of a 'DEL–PROTAC' fusion technology yet, but intensive efforts are surely being made in academic and industrial laboratories. Perhaps the most straightforward way is to use DELs to identify binders of the POI, which can be further conjugated with known E3 ligands (Fig. 7b)[121,122]. DEL could also be used to discover new E3 ligands, because only ~1% of the ~600 E3 ligase family proteins have been explored for targeted protein degradation (Fig. 7c)[119]. The E3 ligase ligands could be either inhibitors to increase the POI level or 'harnesses' in PROTACs to decrease the POI level (https://www.nurixtx.com/platform). DEL could be designed with a built-in E3 ligand to identify the POI ligands and assess their ability to stabilize the 'POI–PROTAC–E3 ligase' ternary complex as a predictive measure for their cellular performance (Fig. 7d; https://www.hitgen.com/en/capabilities-details-18.html). Notably, a phenotypic assay based on OBOC–DELs has been proposed to discover novel E3 ligands or degradation profiles of E3 ligases with live cells (Fig. 7e; https://plexium.com/platform-plexium-e3-ligase-drugs). This approach is of particular interest as the identified ligands are more likely to induce POI degradation in cells. It is only a matter of time that we see this 'DEL–PROTAC' field flourish in both drug discovery and academic research.

DEL beyond drug discovery

Although predominantly being used in drug discovery, DELs have been applied in applications beyond ligand identification. The Liu group developed several DNA-encoded reaction discovery systems[123,124], in which the bond-forming events were encoded and identified after the selection. Previously, non-encoded OBOC libraries were used to screen for organocatalysts[125]; the beads that carry an active catalyst could react with a fluorescently labelled reactant for hit identification. This concept was adapted to a DEL format in a model study[126]. Using DNA-templated polymerization, Liu, Hili and their respective co-workers developed several DNA-encoded polymer libraries, in which the polymer sequence is well defined by the DNA template and the folding of the polymers generates a diversity of complex structures with dense functional groups. Notably, the polymer–DELs are evolvable towards the desired properties through iteration[127–129]. Finally, DNA–PNA is extensively used as scaffolds for ligand display to interrogate multivalent protein complexes[51,130,131]. DNA–PNA encodes both the structures and the positions of the displayed ligands. The configuration of the ligand assembly can be controlled by DNA–PNA hybridization to probe the binding pockets on the target or screen for ligand–DNA–PNA constructs that can bind to the complex. Apart from ligand display, the research activities on using DNA to encode other features remain limited to a few reports, probably at least partially due to the overwhelming attention on using DEL in drug discovery. Encouragingly, these 'non-classic' DELs share a similar principle and technical basis to those for regular DELs, and we expect to see more diverse applications of DELs beyond ligand discovery.

Summary and outlook

In the early days, there were serious doubts as to how useful DELs could be. DNA-conjugated small molecules in a DEL were described as analogous to 'a baseball on a light pole'; that is, the size of the small molecule (the baseball) is too small compared with its tag (the light pole). It was believed that the large DNA tag would almost certainly interfere with and overwhelm the small-molecule–protein interaction, and DNA–protein binding would be another serious problem; however, these were proved not to be major issues. The protein–ligand interaction can be reliably preserved in DEL selections, except that the DNA-conjugation site is not available for target binding. DNA–protein interaction can be minimized with a properly designed DNA sequence and well-controlled experiment. Later, DEL faced another challenge: chemical diversity. The properties of the DNA molecule limit what can be synthesized in DELs. For a long time, DELs contained mostly peptides and peptidomimetics with a limited structural diversity and poor likeness to drugs. However, recent developments extended the range of DEL-compatible chemistries, and innovations on the synthesis techniques (for example, reversible solid-phase absorption[132], micelles[133,134], hybrid-phase synthesis[135] and chemically stabilized barcodes[136]) greatly expanded the toolbox[6,7,11,21,56]. Arguably, at present, the diversity of DELs is limited by the availability of the building blocks, rather than by the chemistries that connect them. Today, DELs can access a vast chemical space to discover many types of ligand structures with more desirable properties.

In combinatorial chemistry, the low library quality that results from incomplete synthetic steps is a major problem because truncated products, mis-reacted building blocks and other impurities may interfere with the biochemical assay. DEL selection is based on binding and hit decoding requires a DNA tag; thus, it partially circumvents this problem. However, low library quality is still a major issue for DELs, especially with the large number of building blocks and the inability to purify during library synthesis. Besides improving on-DNA reactions, many other strategies have been developed to address this challenge, such as preselection library purification and/or normalization[137], post-selection data denoising[138], modelling[139,140], statistical analysis[141,142] and hit triaging using affinity selection–mass spectrometry[143].

We envision that the next stage of DEL development is to go beyond identifying binders to creating functional and/or phenotypic assays. Being a mixture, DELs seem to be intrinsically incompatible with any assay that has a holistic signal readout. Encouragingly, recent innovations have demonstrated that this is possible. With spatial separation, OBOC–DELs can operate on the basis of individual compounds and so are highly promising for

realizing functional assays[52,53]. However, OBOC–DELs release the compounds from the encoding beads; thus, they may face similar issues to those of traditional OBOC libraries, such as compound diffusion, interference from the truncated products and variation of the compound concentration in each droplet. It is more desirable that solution-phase DELs be amenable to functional assays. Previous studies took advantage of the structural change during biochemical–catalytic reactions (for example, phosphorylation and peptide cleavage) to connect the signal readout with the active compounds[18,62,111–115]. Further exploiting this strategy in other biochemical transformations (for example, post-translational modifications and ubiquitination) in a cellular setting would lead to more exciting applications, such as selections against signalling pathways. Moreover, the proximity effect is prevalent in biological systems, but has remained unexplored by DELs. Crosslinking-based selection methods relied on the proximity effect between the crosslinker and the target on binding, and they have enabled DEL selections in cell lysates and on or within live cells. This strategy may also be utilized for DELs to interrogate protein–protein interactions or large multivalent protein complexes. Finally, the integration of DEL with other disciplines and technologies would create a great synergy. Besides the proved ones (next-generation sequencing, OBOC, microfluidics, fragment-based drug discovery, dynamic combinatorial library and so on), there are still many others, such as machine learning and/or artificial intelligence[34–36], affinity selection–mass spectrometry[144] and DNA origami[145], that would advance DELs into new paradigms.

Received: 13 July 2021; Accepted: 1 December 2021;
Published online: 4 February 2022

References

1. McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **348**, 552–554 (1990).
2. Wilson, D. S., Keefe, A. D. & Szostak, J. W. The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl Acad. Sci. USA* **98**, 3750–3755 (2001).
3. Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **15**, 553–557 (1997).
4. Hanes, J. & Plückthun, A. In vitro selection and evolution of functional proteins by using ribosome display. *Proc. Natl Acad. Sci. USA* **94**, 4937–4942 (1997).
5. Neri, D. & Lerner, R. A. DNA-encoded chemical libraries: a selection system based on endowing organic compounds with amplifiable information. *Annu. Rev. Biochem.* **87**, 479–502 (2018).
6. Patel, S., Badir, S. O. & Molander, G. A. Developments in photoredox-mediated alkylation for DNA-encoded libraries. *Trends Chem.* **3**, 161–175 (2021).
7. Fitzgerald, P. R. & Paegel, B. M. DNA-encoded chemistry: drug discovery from a few good reactions. *Chem. Rev.* **121**, 7155–7177 (2021).
8. Dickson, P. & Kodadek, T. Chemical composition of DNA-encoded libraries, past present and future. *Org. Biomol. Chem.* **17**, 4676–4688 (2019).
9. Conole, D., J, H. H. & M, J. W. The maturation of DNA encoded libraries: opportunities for new users. *Future Med. Chem.* **13**, 173–191 (2021).
10. Flood, D. T. et al. DNA encoded libraries: a visitor's guide. *Isr. J. Chem.* **60**, 268–280 (2020).
11. Götte, K., Chines, S. & Brunschweiger, A. Reaction development for DNA-encoded library technology: from evolution to revolution? *Tetrahedron Lett.* **61**, 151889 (2020).
12. Lerner, R. A. & Brenner, S. DNA-encoded compound libraries as open source: a powerful pathway to new drugs. *Angew. Chem. Int. Ed.* **56**, 1164–1165 (2017).
13. Brenner, S. & Lerner, R. A. Encoded combinatorial chemistry. *Proc. Natl Acad. Sci. USA* **89**, 5381–5383 (1992).
14. Nielsen, J., Brenner, S. & Janda, K. D. Synthetic methods for the implementation of encoded combinatorial chemistry. *J. Am. Chem. Soc.* **115**, 9812–9813 (1993).
15. Needels, M. C. et al. Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc. Natl Acad. Sci. USA* **90**, 10700–10704 (1993).
16. Melkko, S., Scheuermann, J., Dumelin, C. E. & Neri, D. Encoded self-assembling chemical libraries. *Nat. Biotechnol.* **22**, 568–574 (2004).
17. Gartner, Z. J. et al. DNA-templated organic synthesis and selection of a library of macrocycles. *Science* **305**, 1601–1605 (2004).
18. Halpin, D. R. & Harbury, P. B. DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biol.* **2**, 1015–1021 (2004).
19. Debaene, F., Mejias, L., Harris, J. L. & Winssinger, N. Synthesis of a PNA-encoded cysteine protease inhibitor library. *Tetrahedron* **60**, 8677–8690 (2004).
20. Clark, M. A. et al. Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat. Chem. Biol.* **5**, 647–654 (2009).
21. Song, M. & Hwang, G. T. DNA-encoded library screening as a core platform technology in drug discovery. Its synthetic method development and applications in DEL synthesis. *J. Med. Chem.* **63**, 6578–6599 (2020).
22. Wichert, M. et al. Dual-display of small molecules enables the discovery of ligand pairs and facilitates affinity maturation. *Nat. Chem.* **7**, 241–249 (2015).
23. Daguer, J. P. et al. DNA display of fragment pairs as a tool for the discovery of novel biologically active small molecules. *Chem. Sci.* **6**, 739–744 (2015).
24. Daguer, J. P. et al. DNA-templated combinatorial assembly of small molecule fragments amenable to selection/amplification cycles. *Chem. Sci.* **2**, 625–632 (2011).
25. Barluenga, S. et al. Novel PTP1B inhibitors identified by DNA display of fragment pairs. *Bioorg. Med. Chem. Lett.* **26**, 1080–1085 (2016).
26. Reddavide, F. V., Lin, W., Lehnert, S. & Zhang, Y. DNA-encoded dynamic combinatorial chemical libraries. *Angew. Chem. Int. Ed.* **54**, 7924–7928 (2015).
27. Reddavide, F. V. et al. Second generation DNA-encoded dynamic combinatorial chemical libraries. *Chem. Commun.* **55**, 3753–3756 (2019).
28. Li, G. et al. Design, preparation, and selection of DNA-encoded dynamic libraries. *Chem. Sci.* **6**, 7097–7104 (2015).
29. Zhou, Y. et al. DNA-encoded dynamic chemical library and its applications in ligand discovery. *J. Am. Chem. Soc.* **140**, 15859–15867 (2018).
30. Deng, Y. et al. Selection of DNA-encoded dynamic chemical libraries for direct inhibitor discovery. *Angew. Chem. Int. Ed.* **59**, 14965–14972 (2020).
31. Farrera-Soler, L. et al. PNA-based dynamic combinatorial libraries (PDCL) and screening of lectins. *Bioorg. Med. Chem.* **28**, 115458 (2020).
32. Machida, T. et al. Dynamic cooperative glycan assembly blocks the binding of bacterial lectins to epithelial cells. *Angew. Chem. Int. Ed.* **56**, 6762–6766 (2017).
33. Lenci, E., Baldini, L. & Trabocchi, A. Diversity-oriented synthesis as a tool to expand the chemical space of DNA-encoded libraries. *Biorg. Med. Chem.* **41**, 116218 (2021).
34. Guasch, L., Reutlinger, M., Stoffler, D. & Wichert, M. Augmenting chemical space with DNA-encoded library technology and machine learning. *Chimia* **75**, 105–107 (2021).
35. Martin, A., Nicolaou, C. A. & Toledo, M. A. Navigating the DNA encoded libraries chemical space. *Commun. Chem.* **3**, 127 (2020).
36. McCloskey, K. et al. Machine learning on DNA-encoded libraries: a new paradigm for hit finding. *J. Med. Chem.* **63**, 8857–8866 (2020).
37. Bobers, J. et al. Design of an automated reagent-dispensing system for reaction screening and validation with DNA-tagged substrates. *ACS Comb. Sci.* **22**, 101–108 (2020).
38. Castanon, J. et al. Design and development of a technology platform for DNA-encoded library production and affinity selection. *SLAS Discov.* **23**, 387–396 (2018).
39. MacConnell, A. B., McEnaney, P. J., Cavett, V. J. & Paegel, B. M. DNA-encoded solid-phase synthesis: encoding language design and complex oligomer library synthesis. *ACS Comb. Sci.* **17**, 518–534 (2015).
40. Cochrane, W. G. et al. Activity-based DNA-encoded library screening. *ACS Comb. Sci.* **21**, 425–435 (2019).
41. Hackler, A. L. et al. Off-DNA DNA-encoded library affinity screening. *ACS Comb. Sci.* **22**, 25–34 (2020).
42. Shin, M. H., Lee, K. J. & Lim, H. S. DNA-encoded combinatorial library of macrocyclic peptoids. *Bioconjug. Chem.* **30**, 2931–2938 (2019).
43. Mendes, K. R. et al. High-throughput identification of DNA-encoded IgG ligands that distinguish active and latent *Mycobacterium tuberculosis* infections. *ACS Chem. Biol.* **12**, 234–243 (2017).
44. Arico-Muendel, C. C. From haystack to needle: finding value with DNA encoded library technology at GSK. *MedChemComm* **7**, 1898–1909 (2016).
45. Belyanskaya, S. L. et al. Discovering drugs with DNA-encoded library technology: from concept to clinic with an inhibitor of soluble epoxide hydrolase. *ChemBioChem* **18**, 837–842 (2017).
46. Harris, P. A. et al. 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 (2017).
47. Cuozzo, J. W. et al. Novel autotaxin inhibitor for the treatment of idiopathic pulmonary fibrosis: a clinical candidate discovered using DNA-encoded chemistry. *J. Med. Chem.* **63**, 7840–7856 (2020).
48. Ahn, S. et al. Allosteric 'beta-blocker' isolated from a DNA-encoded small molecule library. *Proc. Natl Acad. Sci. USA* **114**, 1708–1713 (2017).

49. Ahn, S. et al. Small-molecule positive allosteric modulators of the beta$_2$-adrenoceptor isolated from DNA-encoded libraries. *Mol. Pharmacol.* **94**, 850–861 (2018).
50. Brown, D. G. et al. Agonists and antagonists of protease-activated receptor 2 discovered within a DNA-encoded chemical library using mutational stabilization of the target. *SLAS Discov.* **23**, 429–436 (2018).
51. Figuerola-Conchas, A. et al. Small-molecule modulators of the ATPase VCP/p97 affect specific p97 cellular functions. *ACS Chem. Biol.* **15**, 243–253 (2020).
52. Satz, A. L., Kuai, L. & Peng, X. Selections and screenings of DNA-encoded chemical libraries against enzyme and cellular targets. *Bioorg. Med. Chem. Lett.* **39**, 127851 (2021).
53. Kodadek, T., Paciaroni, N. G., Balzarini, M. & Dickson, P. Beyond protein binding: recent advances in screening DNA-encoded libraries. *Chem. Commun.* **55**, 13330–13341 (2019).
54. Goodnow R. A. Jr A *Handbook for DNA-Encoded Chemistry: Theory and Applications for Exploring Chemical Space and Drug Discovery* (John Wiley & Sons, 2014).
55. Reiher, C. A., Schuman, D. P., Simmons, N. & Wolkenberg, S. E. Trends in hit-to-lead optimization following DNA-encoded library screens. *ACS Med. Chem. Lett.* **12**, 343–350 (2021).
56. Shi, Y. et al. DNA-encoded libraries (DELs): a review of on-DNA chemistries and their output. *RSC Adv.* **11**, 2359–2376 (2021).
57. Kunig, V. B. K., Potowski, M., Klika Skopic, M. & Brunschweiger, A. Scanning protein surfaces with DNA-encoded libraries. *ChemMedChem* **16**, 1048–1062 (2021).
58. Madsen, D. et al. An overview of DNA-encoded libraries: a versatile tool for drug discovery. *Prog. Med. Chem.* **59**, 181–249 (2020).
59. Gironda-Martínez, A., Donckele, E. J., Samain, F. & Neri, D. DNA-encoded chemical libraries: a comprehensive review with successful stories and future challenges. *ACS Pharmacol. Trans. Sci.* **4**, 1265–1279 (2021).
60. Li, Y., Zimmermann, G., Scheuermann, J. & Neri, D. Quantitative PCR is a valuable tool to monitor the performance of DNA-encoded chemical library selections. *ChemBioChem* **18**, 848–852 (2017).
61. Sannino, A. et al. Quantitative assessment of affinity selection performance by using DNA-encoded chemical libraries. *ChemBioChem* **20**, 955–962 (2019).
62. Kim, D. et al. Application of a substrate-mediated selection with c-Src tyrosine kinase to a DNA-encoded chemical library. *Molecules* **24**, 2764 (2019).
63. Denton, K. E. & Krusemark, C. J. Crosslinking of DNA-linked ligands to target proteins for enrichment from DNA-encoded libraries. *MedChemComm* **7**, 2020–2027 (2016).
64. Sannino, A. et al. Critical evaluation of photo-cross-linking parameters for the implementation of efficient DNA-encoded chemical library selections. *ACS Comb. Sci.* **22**, 204–212 (2020).
65. Chen, Q. et al. Exploring the lower limit of individual DNA-encoded library molecules in selection. *SLAS Discov.* **25**, 523–529 (2020).
66. Andrade, H. et al. Using a PCR-based method to analyze and model large, heterogeneous populations of DNA. *ChemBioChem* **21**, 1144–1149 (2020).
67. Satz, A. L. DNA encoded library selections and insights provided by computational simulations. *ACS Chem. Biol.* **10**, 2237–2245 (2015).
68. Satz, A. L., Hochstrasser, R. & Petersen, A. C. Analysis of current DNA encoded library screening data indicates higher false negative rates for numerically larger libraries. *ACS Comb. Sci.* **19**, 234–238 (2017).
69. McGregor, L. M., Jain, T. & Liu, D. R. 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 (2014).
70. Chan, A. I., McGregor, L. M., Jain, T. & Liu, D. R. Discovery of a covalent kinase inhibitor from a DNA-encoded small-molecule library × protein library selection. *J. Am. Chem. Soc.* **139**, 10192–10195 (2017).
71. Blakskjaer, P., Heitner, T. & Hansen, N. J. Fidelity by design: YoctoReactor and binder trap enrichment for small-molecule DNA-encoded libraries and drug discovery. *Curr. Opin. Chem. Biol.* **26**, 62–71 (2015).
72. Petersen, L. K. et al. Novel p38 alpha MAP kinase inhibitors identified from YoctoReactor DNA-encoded small molecule library. *MedChemComm* **7**, 1332–1339 (2016).
73. Petersen, L. K. et al. Screening of DNA-encoded small molecule libraries inside a living cell. *J. Am. Chem. Soc.* **143**, 2751–2756 (2021).
74. Zhao, P. et al. Selection of DNA-encoded small molecule libraries against unmodified and non-immobilized protein targets. *Angew. Chem. Int. Ed.* **53**, 10056–10059 (2014).
75. Shi, B., Deng, Y. & Li, X. Polymerase-extension-based selection method for DNA-encoded chemical libraries against nonimmobilized protein targets. *ACS Comb. Sci.* **21**, 345–349 (2019).
76. Shi, B., Deng, Y., Zhao, P. & Li, X. Selecting a DNA-encoded chemical library against non-immobilized proteins using a 'ligate–cross-link–purify' strategy. *Bioconjug. Chem.* **28**, 2293–2301 (2017).
77. Winssinger, N. & Harris, J. L. Microarray-based functional protein profiling using peptide nucleic acid-encoded libraries. *Expert Rev. Proteomics* **2**, 937–947 (2005).
78. Harris, J. L. & Winssinger, N. PNA encoding (PNA = peptide nucleic acid): from solution-based libraries to organized microarrays. *Chem. Eur. J.* **11**, 6792–6801 (2005).
79. Kochmann, S., Le, A. T. H., Hili, R. & Krylov, S. N. Predicting efficiency of NECEEM-based partitioning of protein binders from nonbinders in DNA-encoded libraries. *Electrophoresis* **39**, 2991–2996 (2018).
80. Bao, J. et al. Predicting electrophoretic mobility of protein–ligand complexes for ligands from DNA-encoded libraries of small molecules. *Anal. Chem.* **88**, 5498–5506 (2016).
81. Sprinz, K. I., Tagore, D. M. & Hamilton, A. D. Self-assembly of bivalent protein-binding agents based on oligonucleotide-linked organic fragments. *Bioorg. Med. Chem. Lett.* **15**, 3908–3911 (2005).
82. Onda, Y. et al. A DNA-encoded chemical library based on peptide macrocycles. *Chem. Eur. J.* **27**, 7160–7167 (2021).
83. Dal Corso, A. et al. Affinity enhancement of protein ligands by reversible covalent modification of neighboring lysine residues. *Angew. Chem. Int. Ed.* **57**, 17178–17182 (2018).
84. Kollmann, C. S. et al. 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 (2014).
85. Richter, H. et al. DNA-encoded library-derived DDR1 inhibitor prevents fibrosis and renal function loss in a genetic mouse model of Alport syndrome. *ACS Chem. Biol.* **14**, 37–49 (2019).
86. Xie, J. et al. Selection of small molecules that bind to and activate the insulin receptor from a DNA-encoded library of natural products. *iScience* **23**, 101197 (2020).
87. Favalli, N. et al. Stereo- and regiodefined DNA-encoded chemical libraries enable efficient tumour-targeting applications. *Nat. Chem.* **13**, 540–548 (2021).
88. Huang, Y. et al. Selection of DNA-encoded chemical libraries against endogenous membrane proteins on live cells. *Nat. Chem.* **13**, 77–88 (2021).
89. Svensen, N., Diaz-Mochon, J. J. & Bradley, M. Decoding a PNA encoded peptide library by PCR: the discovery of new cell surface receptor ligands. *Chem. Biol.* **18**, 1284–1289 (2011).
90. Svensen, N., Diaz-Mochon, J. J. & Bradley, M. Encoded peptide libraries and the discovery of new cell binding ligands. *Chem. Commun.* **47**, 7638–7640 (2011).
91. Wu, Z. et al. 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 (2015).
92. Cai, B. et al. Selection of DNA-encoded libraries to protein targets within and on living cells. *J. Am. Chem. Soc.* **141**, 17057–17061 (2019).
93. Li, G. et al. Photoaffinity labeling of small-molecule-binding proteins by DNA-templated chemistry. *Angew. Chem. Int. Ed.* **52**, 9544–9549 (2013).
94. Schroeder, H. et al. Generation of live-cell microarrays by means of DNA-directed immobilization of specific cell-surface ligands. *Angew. Chem. Int. Ed.* **46**, 4180–4183 (2007).
95. Stahl, S. et al. Affibody molecules in biotechnological and medical applications. *Trends Biotechnol.* **35**, 691–712 (2017).
96. Li, L. et al. Aptamer displacement reaction from live-cell surfaces and its applications. *J. Am. Chem. Soc.* **141**, 17174–17179 (2019).
97. Oehler, S. et al. Affinity selections of DNA-encoded chemical libraries on carbonic anhydrase IX—expressing tumor cells reveal a dependence on ligand valence. *Chem. Eur. J.* **27**, 8985–8993 (2021).
98. Litovchick, A. et al. Novel nucleic acid binding small molecules discovered using DNA-encoded chemistry. *Molecules* **24**, 2026 (2019).
99. Blain, J. C. et al. Encoded libraries and methods of use for screening nucleic acid targets. World patent WO2019236644A1 (2019).
100. Zhang, J. et al. Identification of histone deacetylase (HDAC)-associated proteins with DNA-programmed affinity labeling. *Angew. Chem. Int. Ed.* **59**, 17525–17532 (2020).
101. Komnatnyy, V. V., Nielsen, T. E. & Qvortrup, K. Bead-based screening in chemical biology and drug discovery. *Chem. Commun.* **54**, 6759–6771 (2018).
102. Wu, L. et al. Aptamer-based detection of circulating targets for precision medicine. *Chem. Rev.* **121**, 12035–12105 (2021).
103. Svensen, N. et al. Screening of a combinatorial homing peptide library for selective cellular delivery. *Angew. Chem. Int. Ed.* **50**, 6133–6136 (2011).
104. Kolodny, G., Li, X. & Balk, S. Addressing cancer chemotherapeutic toxicity, resistance, and heterogeneity: novel theranostic use of DNA-encoded small molecule libraries. *Bioessays* **40**, e1800057 (2018).
105. Cuozzo, J. W. et al. 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 (2017).

106. Zambaldo, C. et al. Screening for covalent inhibitors using DNA-display of small molecule libraries functionalized with cysteine reactive moieties. *MedChemComm* **7**, 1340–1351 (2016).
107. Zhu, Z. et al. Development of a selection method for discovering irreversible (covalent) binders from a DNA-encoded library. *SLAS Discov* **24**, 169–174 (2018).
108. Zimmermann, G. et al. A specific and covalent JNK-1 ligand selected from an encoded self-assembling chemical library. *Chem. Eur. J.* **23**, 8152–8155 (2017).
109. Daguer, J. P. et al. Identification of covalent bromodomain binders through DNA display of small molecules. *Angew. Chem. Int. Ed.* **54**, 6057–6061 (2015).
110. Guilinger, J. P. et al. Novel irreversible covalent BTK inhibitors discovered using DNA-encoded chemistry. *Bioorg. Med. Chem.* **42**, 116223 (2021).
111. Winssinger, N. et al. PNA-encoded protease substrate microarrays. *Chem. Biol.* **11**, 1351–1360 (2004).
112. Diaz-Mochon, J. J., Bialy, L. & Bradley, M. Dual colour, microarray-based, analysis of 10,000 protease substrates. *Chem. Commun.* **14**, 3984–3986 (2006).
113. Pouchain, D., Diaz-Mochon, J. J., Bialy, L. & Bradley, M. A 10,000 member PNA-encoded peptide library for profiling tyrosine kinases. *ACS Chem. Biol.* **2**, 810–818 (2007).
114. Krusemark, C. J., Tilmans, N. P., Brown, P. O. & Harbury, P. B. Directed chemical evolution with an outsized genetic code. *PLoS ONE* **11**, e0154765 (2016).
115. Jetson, R. R. & Krusemark, C. J. Sensing enzymatic activity by exposure and selection of DNA-encoded probes. *Angew. Chem. Int. Ed* **55**, 9562–9566 (2016).
116. Roy, A., Koesema, E. & Kodadek, T. J. High-throughput quality control assay for the solid-phase synthesis of DNA-encoded libraries of macrocycles. *Angew. Chem. Int. Ed.* **60**, 11983–11990 (2021).
117. MacConnell, A. B., Price, A. K. & Paegel, B. M. An integrated microfluidic processor for DNA-encoded combinatorial library functional screening. *ACS Comb. Sci.* **19**, 181–192 (2017).
118. Borchardt, A. et al. Small molecule-dependent genetic selection in stochastic nanodroplets as a means of detecting protein–ligand interactions on a large scale. *Chem. Biol.* **4**, 961–968 (1997).
119. Schapira, M., Calabrese, M. F., Bullock, A. N. & Crews, C. M. Targeted protein degradation: expanding the toolbox. *Nat. Rev. Drug Discov.* **18**, 949–963 (2019).
120. Zhu, Z. et al. Design and application of a DNA-encoded macrocyclic peptide library. *ACS Chem. Biol.* **13**, 53–59 (2018).
121. Disch, J. S. et al. Bispecific estrogen receptor alpha degraders incorporating novel binders identified using DNA-encoded chemical library screening. *J. Med. Chem.* **64**, 5049–5066 (2021).
122. Andersen, J, N. et al. Degradation of immuno-oncology targets via proprietary PROTAC platform integrating DNA-encoded library technology and rational drug design. *Cancer Res.* **79**, https://doi.org/10.1158/1538-7445.AM2019-1981 (2019).
123. Kanan, M. W. et al. Reaction discovery enabled by DNA-templated synthesis and in vitro selection. *Nature* **431**, 545–549 (2004).
124. Chen, Y., Kamlet, A. S., Steinman, J. B. & Liu, D. R. A biomolecule-compatible visible-light-induced azide reduction from a DNA-encoded reaction-discovery system. *Nat. Chem.* **3**, 146–153 (2011).
125. Krattiger, P., McCarthy, C., Pfaltz, A. & Wennemers, H. Catalyst-substrate coimmobilization: A strategy for catalysts discovery in split-and-mix libraries. *Angew. Chem. Int. Ed.* **42**, 1722–1724 (2003).
126. Hook, K. D., Chambers, J. T. & Hili, R. A platform for high-throughput screening of DNA-encoded catalyst libraries in organic solvents. *Chem. Sci.* **8**, 7072–7076 (2017).
127. Brudno, Y., Birnbaum, M. E., Kleiner, R. E. & Liu, D. R. An in vitro translation, selection and amplification system for peptide nucleic acids. *Nat. Chem. Biol.* **6**, 148–155 (2010).
128. Hili, R., Niu, J. & Liu, D. R. DNA ligase-mediated translation of DNA into densely functionalized nucleic acid polymers. *J. Am. Chem. Soc.* **135**, 98–101 (2013).
129. Kong, D., Yeung, W. & Hili, R. In vitro selection of diversely functionalized aptamers. *J. Am. Chem. Soc.* **139**, 13977–13980 (2017).
130. Yeldell, S. B. & Seitz, O. Nucleic acid constructs for the interrogation of multivalent protein interactions. *Chem. Soc. Rev.* **49**, 6848–6865 (2020).
131. Barluenga, S. & Winssinger, N. PNA as a biosupramolecular tag for programmable assemblies and reactions. *Acc. Chem. Res.* **48**, 1319–1331 (2015).
132. Flood, D. T. et al. Expanding reactivity in DNA-encoded library synthesis via reversible binding of DNA to an inert quaternary ammonium support. *J. Am. Chem. Soc.* **141**, 9998–10006 (2019).
133. Skopic, M. K. et al. Micellar Brønsted acid mediated synthesis of DNA-tagged heterocycles. *J. Am. Chem. Soc.* **141**, 10546–10555 (2019).
134. Hunter, J. H. et al. Highly efficient on-DNA amide couplings promoted by micelle forming surfactants for the synthesis of DNA encoded libraries. *Chem. Sci.* **12**, 9475–9484 (2021).
135. Skopic, M. K. et al. Acid- and Au(I)-mediated synthesis of hexathymidine-DNA-heterocycle chimeras, an efficient entry to DNA-encoded libraries inspired by drug structures. *Chem. Sci.* **8**, 3356–3361 (2017).
136. Potowski, M. et al. Chemically stabilized DNA barcodes for DNA-encoded chemistry. *Angew. Chem. Int. Ed.* **60**, 19744–19749 (2021).
137. Rama-Garda, R. et al. Normalization of DNA encoded library affinity selection results driven by high throughput sequencing and HPLC purification. *Bioorg. Med. Chem.* **40**, 116178 (2021).
138. Komar, P. & Kalinic, M. Denoising DNA encoded library screens with sparse learning. *ACS Comb. Sci.* **22**, 410–421 (2020).
139. Gerry, C. J., Wawer, M. J., Clemons, P. A. & Schreiber, S. L. DNA barcoding a complete matrix of stereoisomeric small molecules. *J. Am. Chem. Soc.* **141**, 10225–10235 (2019).
140. Faver, J. C. et al. Quantitative comparison of enrichment from DNA-encoded chemical library selections. *ACS Comb. Sci.* **21**, 75–82 (2019).
141. Kuai, L., O'Keeffe, T. & Arico-Muendel, C. Randomness in DNA encoded library selection data can be modeled for more reliable enrichment calculation. *SLAS Discov.* **23**, 405–416 (2018).
142. Denton, K. E. et al. Robustness of in vitro selection assays of DNA-encoded peptidomimetic ligands to CBX7 and CBX8. *SLAS Discov.* **23**, 417–428 (2018).
143. Su, W. et al. Triaging of DNA-encoded library selection results by high-throughput resynthesis of DNA-conjugate and affinity selection mass spectrometry. *Bioconjug. Chem.* **32**, 1001–1007 (2021).
144. Prudent, R. et al. Exploring new targets and chemical space with affinity selection-mass spectrometry. *Nat. Rev. Chem.* **5**, 62–71 (2021).
145. Kielar, C. et al. Pharmacophore nanoarrays on DNA origami substrates as a single-molecule assay for fragment-based drug discovery. *Angew. Chem. Int. Ed.* **57**, 14873–14877 (2018).

Acknowledgements

This work was supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (AoE/P-705/16, 17301118, 17111319, 17303220, 17300321 and C7005-20G), and from the National Natural Science Foundation of China (21572014, 21877093, 21907011 and 91953119), the Fundamental Research Funds for the Central Universities (project nos. 2020CQJQY-Z002 and 2021CDJYGRH-002) and Chongqing Research and Frontier Technology (cstc2020jcyj-jqX0009 and cstc2021jcyj-cxttX0002). We also acknowledge the funding support from the Laboratory for Synthetic Chemistry and Chemical Biology under the Health@InnoHK Program launched by the Innovation and Technology Commission, The Government of Hong Kong Special Administrative Region of the People's Republic of China.

Author contributions

Y.H., Y.L. and X.L. contributed to the discussions and co-wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to Yizhou Li or Xiaoyu Li.

Peer review information *Nature Chemistry* thanks Nicolas Winssinger and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.