Acta Pharmaceutica Sinica B 2024;14(2):492−516

Chinese Pharmaceutical Association
Institute of Materia Medica, Chinese Academy of Medical Sciences

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com

REVIEW

Evolution of chemistry and selection technology for DNA-encoded library

Peixiang Ma[a,†], Shuning Zhang[a,b,†], Qianping Huang[b], Yuang Gu[b], Zhi Zhou[c], Wei Hou[d,*], Wei Yi[c,*], Hongtao Xu[b,*]

[a]*Shanghai Key Laboratory of Orthopedic Implants, Department of Orthopedics, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai 200025, China*
[b]*Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China*
[c]*Guangzhou Municipal and Guangdong Provincial Key Laboratory of Molecular Target & Clinical Pharmacology, The NMPA and State Key Laboratory of Respiratory Disease, Guangzhou Medical University, Guangzhou 511436, China*
[d]*College of Pharmaceutical Science and Institute of Drug Development & Chemical Biology, Zhejiang University of Technology, Hangzhou 310014, China*

Received 29 June 2023; received in revised form 11 September 2023; accepted 21 September 2023

KEY WORDS

DNA-encoded library;
Drug discovery;
High-throughput selection;
DNA-compatible synthesis;
In vitro evolution;
Chemical central dogma

Abstract DNA-encoded chemical library (DEL) links the power of amplifiable genetics and the non-self-replicating chemical phenotypes, generating a diverse chemical world. In analogy with the biological world, the DEL world can evolve by using a chemical central dogma, wherein DNA replicates using the PCR reactions to amplify the genetic codes, DNA sequencing transcripts the genetic information, and DNA-compatible synthesis translates into chemical phenotypes. Importantly, DNA-compatible synthesis is the key to expanding the DEL chemical space. Besides, the evolution-driven selection system pushes the chemicals to evolve under the selective pressure, *i.e.,* desired selection strategies. In this perspective, we summarized recent advances in expanding DEL synthetic toolbox and panning strategies, which will shed light on the drug discovery harnessing *in vitro* evolution of chemicals *via* DEL.

*Corresponding authors.
E-mail addresses: houwei@zjut.edu.cn (Wei Hou), yiwei@gzhmu.edu.cn (Wei Yi), xuht@shanghaitech.edu.cn (Hongtao Xu).
†These authors made equal contributions to this work.

Peer review under the responsibility of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences.

https://doi.org/10.1016/j.apsb.2023.10.001

1. Introduction

The selection of specific ligands including micromolecules and small molecules that bind to a target (protein, DNA, or RNA) of interest is a fundamental activity in basic research. Annotation of the complex biological processes often needs specific ligands that bind to macromolecules, thereby enabling their visualization, quantification, tracing, and function studies[1]. Similarly, the drug discovery campaigns usually begin with the selection of a specific ligand that binds and modulates a validated target[2].

In 1985, George P. Smith[3] proposed the display of polypeptide on filamentous phage. The polypeptides were genetically fused at the N-terminal of the filamentous phage's minor coat protein pIII. Shortly afterward, Sir Gregory Winter[4] and Richard Lerner[5] independently reported the synthesis of large combinatorial antibody libraries and the use of phage display to identify specific antibody[6,7]. In these phage display systems, based on the biological central dogma (Fig. 1a), the bacterial viral particle features display a potential function (the polypeptides or antibodies displayed on the phage surface serve as a binding phenotype), while simultaneously containing the corresponding genetic information (genotype) in the modified bacterial phage genome, thereby the genotype and phenotype are packed in a single phage. This procedure is logical and is in full analogy to the clonal expansion of antigen-specific B cells in the human immune system, where genotype and phenotype (immunoglobulin on the B cell surface) are linked in the cell[8].

The antibody phage display technology has substantially influenced basic research and drug discovery. On the other hand, it would also be highly desirable if small molecular ligands to protein targets of interest, like the biotin−(streptavidin) avidin system[9], were readily available. The advent of DNA-encoded combinatorial libraries of polypeptides and antibodies has provided an important inspiration for the genesis of DNA-encoded chemical libraries (DEL).

In the 1992 landmark article, Richard Lerner and Sydney Brenner[10] proposed the encoding of chemicals with DNA barcodes. They envisaged the simultaneous synthesis of diverse polypeptides and the corresponding encoding DNA barcodes *via* orthogonal chemistry and split-and-pool procedure on the beads. In the DEL system, a "chemical central dogma" dominates the process, wherein the oligonucleotides serve as amplifiable barcodes (genotype, identifiers) of the corresponding peptides (phenotype). However, it doesn't need the biosynthetic system required in the phage display, yeast display[11], and ribosome display systems[12]. Instead, a chemical connection between phenotype and genotype was used to mimic the biosynthetic

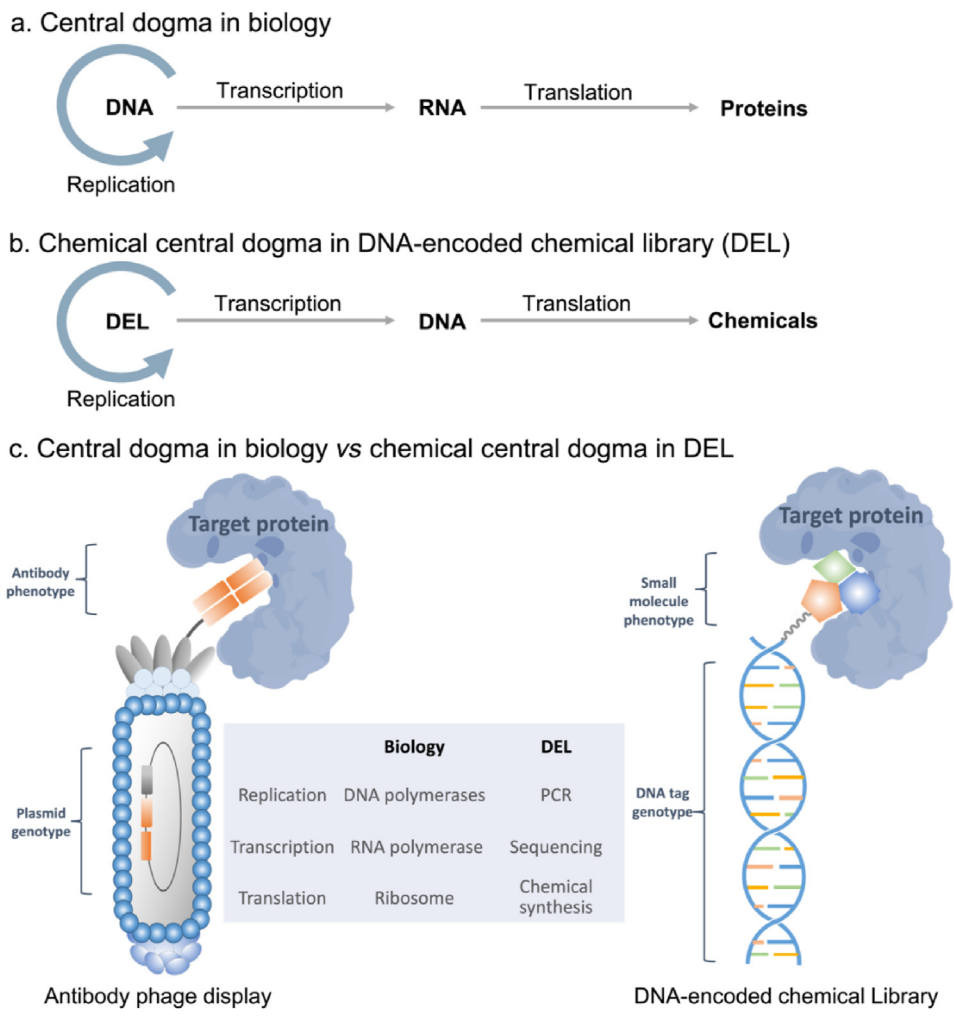


Figure 1 Overview of DNA-encoded library (DEL) technology. (a) Central dogma in biology. (b) Chemical central dogma in DEL. (c) Central dogma in biology *vs*. chemical central dogma in DEL.

linkage between oligonucleotides and proteins (Fig. 1b and c). The DNA-compatible synthesis is the key to expanding the DEL chemical space. Significantly, they also anticipated the replacement of the beads by a generic chemical linker.

In 1993, the groups of Needels[13] and Janda[14] independently exemplified Lerner and Brenner's DEL concept *via* the synthesis of peptide DELs and the successful DEL selection against the antibody epitopes. In 2004, the groups of Neri[15], Liu[16], and Harbury[17] independently reported the construction of DELs in a bead-free fashion (Fig. 2a), and the use of affinity-based DEL selection, which is similar to the panning procedure of phage display libraries. To date, DEL has become an emerging powerful hit selection technique in chemical biology and medicinal chemistry. Single-pharmacophore DEL[18], dual-pharmacophore DEL or the encoded self-assembling chemical (ESAC) library[15], and the recent tri-pharmacophore DEL[19] are the main types of DEL based on the number and/or assembly of pharmacophores (Fig. 2b).

Satisfactorily, a large number of hits have been identified from the DELs for target-based drug discovery (Table 1), as well as for the design of chemical probes[20]. Notably, some drug candidates derived from DEL hits have been tested in clinical trials for various diseases[21−27]. As a bench-top technology, it obviates the need for costly high throughput screening (HTS) infrastructures, thus can easily be accessed by both biotech startups and academic laboratories, thereby playing a key role in bridging chemistry and biology[28].

Importantly, DNA-compatible synthesis and DEL selection technology are two cornerstones in DEL, the former is the key to expanding the DEL chemical space, while the latter is responsible for efficient hit discovery from DEL. Consequently, advancing the frontiers of DEL strategy demands the ongoing development of DNA-compatible synthesis to expand the chemical space continuously, and the development of innovative selection technology to enable efficient panning under different screening requirements (*e.g.*, membrane protein, protein−protein interaction), thus realizing DEL's full potential[28−39]. In this perspective, we will introduce the origin and principle of DEL, provide special attention to the latest developments on the emerging DEL chemistry and selection strategies that will shape future research activities, and comment on challenges, perspectives, and opportunities in future DEL research.

2. The evolution of DEL chemistry

In DEL synthesis, although the DNA barcode enables efficient affinity-based selection of compound libraries with high capacity, it imposes idiosyncratic requirements and formidable challenges to on-DNA synthetic methodology development. The genetic information must be preserved during the reaction, therefore, achieving efficient and high-fidelity on-DNA chemical synthesis is not an easy task.

Generally, on-DNA reactions must be operated at high dilution (∼1 mmol/L), with ≥10% water as the co-solvent, under mild conditions (pH 4−14, 25−90 °C), on small scales (∼25 nmol), avoidance of oxidants and many Lewis acids, and have high chemo-selectivity for modification of the conjugated small molecule rather than the DNA tag. Despite the enormous

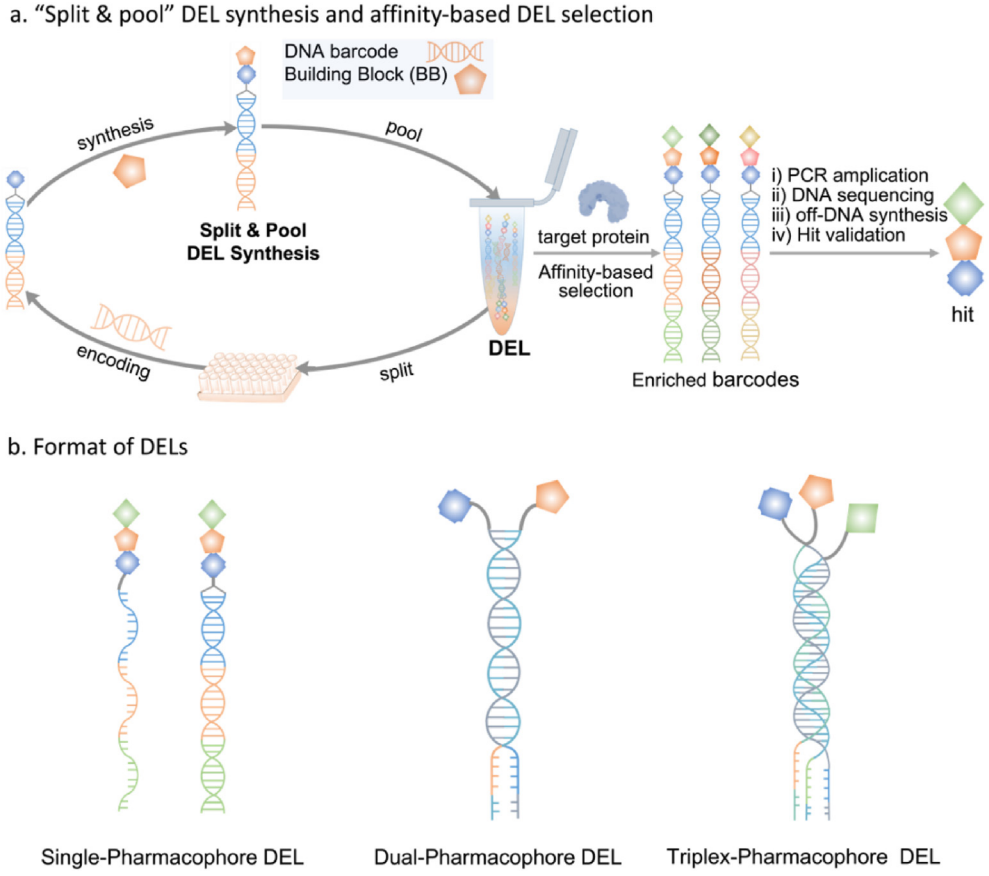
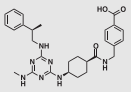
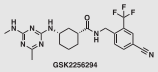
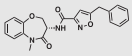
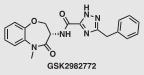
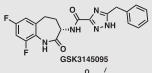
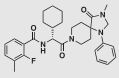
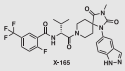
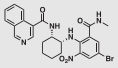
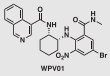
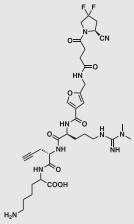
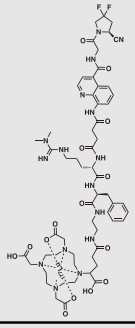


Figure 2 DEL synthesis, selection, and formats. (a) "Split & Pool" DEL synthesis and affinity-based DEL selection. (b) Format of DELs.

Table 1 Clinical candidates selected from DELs.

Target	DEL selection	Hits	Clinical candidates	Clinical phase
Soluble epoxide hydrolase (sEH)[22,23]	Size: 9.0×10^8, target protein was labelled with biotin, and selected three rounds		 GSK2256294	Phase I/II, aneurysmal subarachnoid haemorrhage, insulin sensitivity
Receptor-interacting protein kinase 1 (RIPK1)[21,26]	Size: 7.7×10^9, target protein was labelled with His-GST, and selected three rounds		 GSK2982772	Phase I/II, mild-to-moderate plaque-type psoriasis, rheumatoid arthritis, ulcerative colitis
			 GSK3145095	Phase I/II, ulcerative colitis
Autotaxin/ENPP2[24]	Size: 2.5×10^8, target protein was tagged with Flag, and selected two rounds		 X-165	Approved for phase I, attenuated induced lung fibrosis
SARS-CoV-2 3CLpro[27]	Size: 49×10^9, target protein was tagged with 6 × His-tag		 WPV01	Phase I, SARS-CoV-2
Fibroblast activation protein (FAP)[25]	Size: 5.0×10^4, target protein was biotinylated			^{177}Lu-OncoFAP (therapy) will be investigated in a Phase I clinical trial expected to start in 2H 2023

challenges, recently, many new strategies and/or approaches have been elegantly developed to expand the toolkit and landscape of DEL chemistry. In this section, we will provide an overview of DEL chemistry with emphasis on the emerging DEL chemistry and approaches including modern photo-promoted reaction, C−H activation & functionalization, bioinspired click selenylation, micellar-mediated reactions, the reversible adsorption to solid support (RASS)-based approach and the substrate activation strategy.

2.1. *Conventional DEL chemistry*

The synthesis of first-generation DELs mainly focused on the ever-increasing compound numbers that have been synthesized from large sets of starting materials *via* a small set of well-established building block (BB) connection reactions, including amine capping reactions[40−42], metal-catalyzed (mediated) cross-couplings[43−52], copper-catalyzed azide−alkyne cycloaddition (CuAAC)[53−55], heterocycles synthesis[56−61], Fmoc and Boc deprotections, and azide and/or nitro reduction reactions, etc. (Fig. 3). Generally, most of these robust DEL chemistries fall in the range of the top 20 most frequently used chemical reactions in medicinal chemistry, and they have supported the fundamental chemical space of DEL. At present, the on-DNA versions of these classic reactions are still evolving. For example, the Gironda-Martínezer group reported an efficient on-DNA diazo-transfer reaction, which used imidazole-1-sulfonyl azide tetrafluoroborate salt to convert a wide range of primary aliphatic amines into their corresponding azides[54]. Besides, our group has employed DNA-linked (hetero)aryl fluorosulfates (−OSO_2F) as versatile electrophiles, and developed mild and efficient on-DNA Suzuki, Sonogashira, and Buchwald cross-coupling reactions[45].

Despite these successes, to be sure, there are still some obstacles that need to be solved, mainly including, i) these classical reactions mainly create amides, 1,2,3-triazoles, diaryls, and C(sp^2)−N linkages (*via* either Buchwald cross-coupling or aromatic nucleophilic substitution), thus producing DELs in a planar and C(sp^2) rich fashion, which would no doubt limit the diversity and chemical space of DEL; ii) the use of transition metal-catalysts combined with elevated temperatures in cross-coupling cause DNA degradation, resulting poor DNA fidelity; iii) the employing aqueous solvent systems excludes many reactions involving reagents that are unstable to water. Consequently, the development of new and mild DNA-compatible reactions will not only help to reduce DNA degradation, enrich DEL synthetic toolbox, and expand chemical space, but also ultimately lead to more drug-like libraries.

2.2. *Emerging DEL chemistries and approaches*

Recently, several emerging DEL chemistries and/or approaches such as the photo-catalyzed reactions, C−H activation & functionalization, bioinspired click selenylation, micellar-mediated reactions, the RASS-based reactions and substrate activation strategy have greatly advanced the DEL compatible chemistry

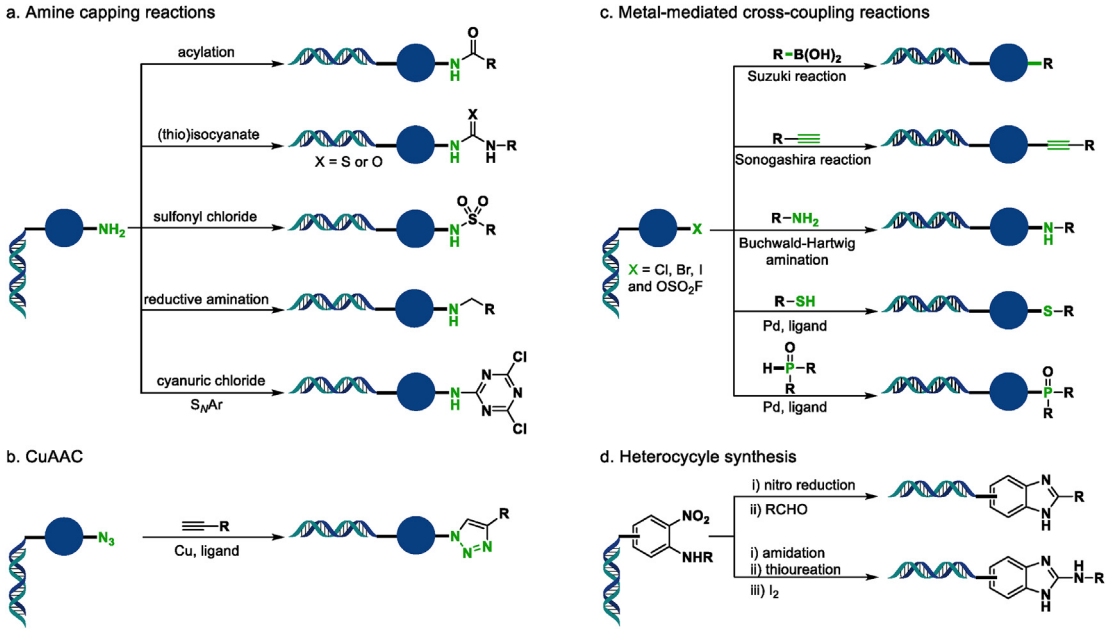


Figure 3 Representative conventional DEL chemistry. (a) Amine capping reactions. (b) CuAAC. (c) Metal-mediated cross-coupling reactions. (d) On-DNA heterocycle synthesis.

such as the synthesis of challenging C(sp^3)−C(sp^3), C(sp^2)−C(sp^3), C−X bonds as well as several privileged heterocycles.

2.2.1. *Photo-promoted reactions*

DNA-compatible photo-promoted chemical transformation, pioneered by Liu, Flanagan, and Molander laboratories, is an attractive synthesis, that forges chemical bonds *via* mainly radical intermediates (Fig. 4). Recently, taking advantage of the mild reaction conditions of photocatalyzed transformations, several elegant DNA-compatible reactions have been well developed, which not only provides a diverse set of novel scaffolds that are challenging in classical on-DNA synthesis, but also greatly expanded the scope of the available BBs[62,63].

In 2011, Liu et al.[64] reported the first photoinduced DNA-compatible reaction, wherein efficient reduction of both aromatic and aliphatic azides was achieved *via* Ru/photoredox dual promotion. Notably, this reduction is highly selective and the disulfide bond was not affected, whereas tris(2-chloroethyl) phosphate (TCEP) reduced both the azide and disulfide groups (Fig. 4a).

Surprisingly, the photoinduced DEL chemistry did not enter a burst period until 2018, when Flanagan et al.[65] reported a Giese addition of photoinduced radicals from *N*-Boc protected α-amino acids to DNA-conjugated acrylamide and styrene to construct C(sp^3)−C(sp^3) bond by using [Ir(dF(CF_3)ppy)$_2$(bpy)]PF_6 as photocatalyst (PC). This reaction tolerates a broad array of structurally diverse radical precursors (RPs), including all of the 20 canonical amino acids (Fig. 4b)[65,66].

Afterward, some photoinduced Giese addition reactions using different alkyl RPs have been reported one after another. In 2022, Molander and coworkers[67] developed a fast, general, and robust method by employing alkyl halogen (Br or I) as the RP. Notably, 1°, 2° and 3° alkyl halogens are all competent coupling partners, and highly functionalized bicyclo[1.1.1]pentanes (BCPs), adamantly, glycosyl and pyrazole group are all well tolerated. Owing to its mild and rapid reaction conditions, this reaction shows good DNA fidelity with 63% amplifiable DNA maintained (Fig. 4c). Independently, the Hu group applied diazo compounds as RPs. This transformation shows excellent functional group tolerance, even for complex natural products and drug molecules (Fig. 4d)[68]. Besides, other RPs including tertiary amine[69,70], 2,4,6-trimethoxybenzaldehyde imine[71] and sodium *O*-alkyl carbonodithioate[72] are also exploited as effective RPs for photoinduced Giese addition. These reactions greatly enriched the toolbox for on-DNA forging the challenging C(sp^3)−C(sp^3) bond.

In 2019, Molander and colleagues[73] reported a photoredox polar crossover defluorinative alkylation of trifluoromethyl alkenes to form *gem*-difluoroalkenes, which generally act as a more metabolically stable isostere of carbonyl group (Fig. 5a). A series of alkyl RPs were feasible for this on-DNA cross-coupling when combined with a suitable PC, including bis(catecholato)silicates (4CzlPN), 4-alkyl-1,4-dihydropyridines (DHPs, 4CzlPN), *N*-Boc/Fmoc α-amino acids ([Ir (dF(CF_3)ppy)$_2$ (bpy)]PF_6) and alkyl iodides (4CzlPN). Notably, when employing Hantzsch ester to replace PC, the hydroalkylation product could be selectively generated *via* a radical formation-addition and hydrogen atom termination pathway through electron donor−acceptor (EDA) complex activation. Moreover, 1°, 2° and 3° alkyl redox-active esters (RAEs) are all competent coupling partners (Fig. 5b)[74]. Satisfactorily, these reaction conditions show high confidence in the fidelity based on qPCR amplification and next-generation sequencing.

Simultaneously to their defluorinative alkylation work, Molander et al.[73] developed a Ni/photoredox dual catalytic C(sp^2)−C(sp^3) cross-coupling between DNA-conjugated aryl halides (Br and I) and a series of radical progenitors, including 4-alkyl-1,4-dihydropyridines (4CzlPN), *N*-Boc α-amino acids ([Ir (dF(CF_3)ppy)$_2$(bpy)]PF_6), alkyl bromides ([Ir(ppy)$_2$(dtbpy)]PF_6) and α-trimethylsilylamine ([Ir(dFCF_3ppy)$_2$(bpy)]PF_6). Functional groups such as glycoside, tryptophan, tetrahydropyrane, hydroxy, and *tert*-butyl proline were all well tolerated, albeit some cases achieved only medium yields (**6b−6f**, 50%−63%, Fig. 6a).

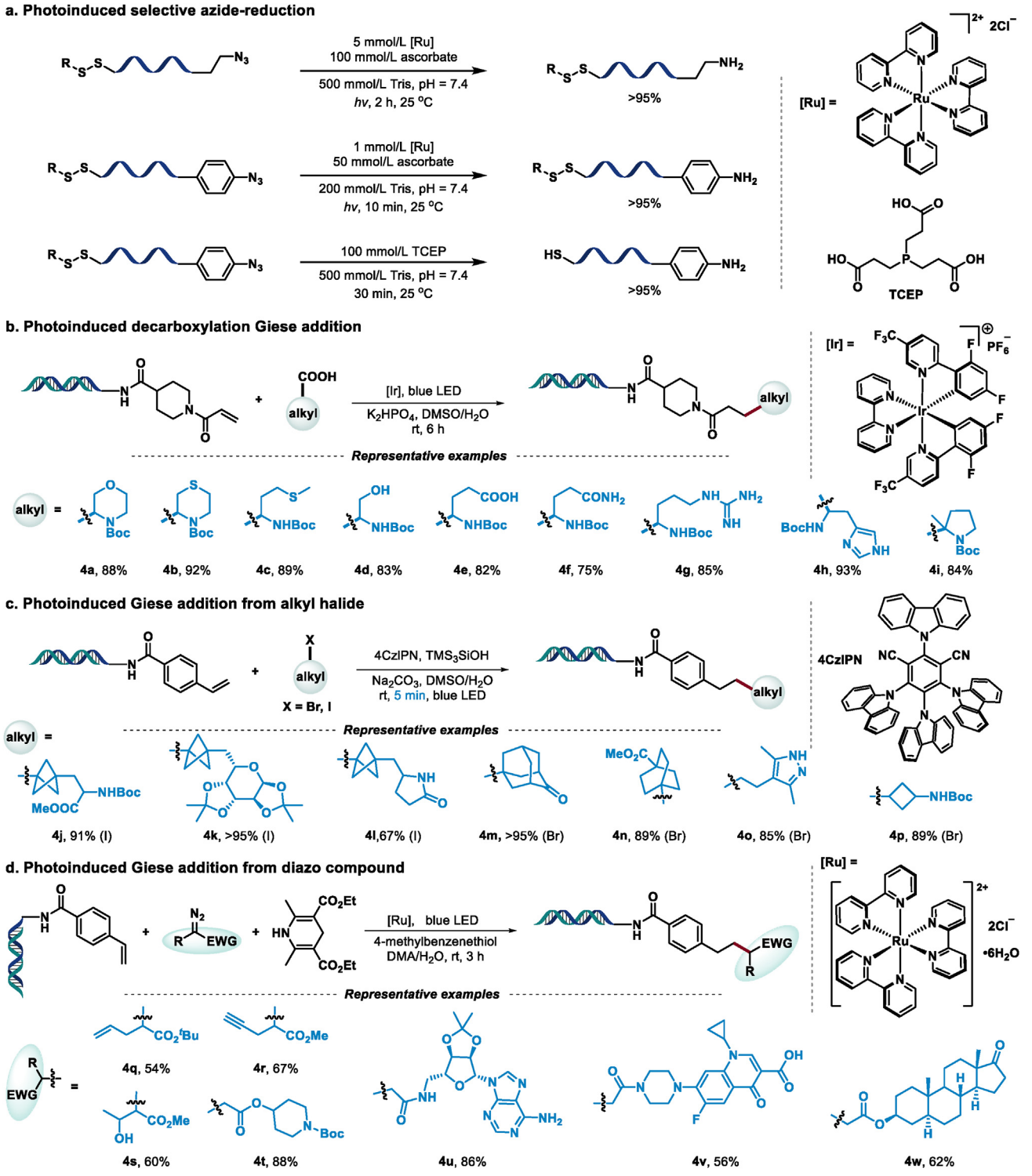


Figure 4 Photoinduced azide-reduction and Giese addition. (a) Photoinduced selective azide-reduction. (b) Photoinduced decarboxylation Giese addition. (c) Photoinduced Giese addition from alkyl halide. (d) Photoinduced Giese addition from the diazo compound.

Significantly, this reaction represents the first metallaphotoredox approach on-DNA. Almost at the same time, Flanagan et al. independently achieved this coupling by using bis(carboxamidine) pyridine as a ligand to reduce the dehalogenation and TMSOH as a sacrificial reductant to improve yields[75]. Apart from C(sp^2)–C(sp^3) cross-coupling, the Molander group also reported a Ni/photoredox dual catalytic C(sp^2)–C(sp^2) bond formation reaction for the direct C—H carbofunctionalization of medicinally-relevant heterocycles through the photoreduction of DNA-tagged (hetero) aryl halides. Notably, this method is amenable to a broad scope of medicinally-relevant heterocyclic building blocks, including quinoline, isoquinoline, indole, pyrrole, caffeine, and dipeptide

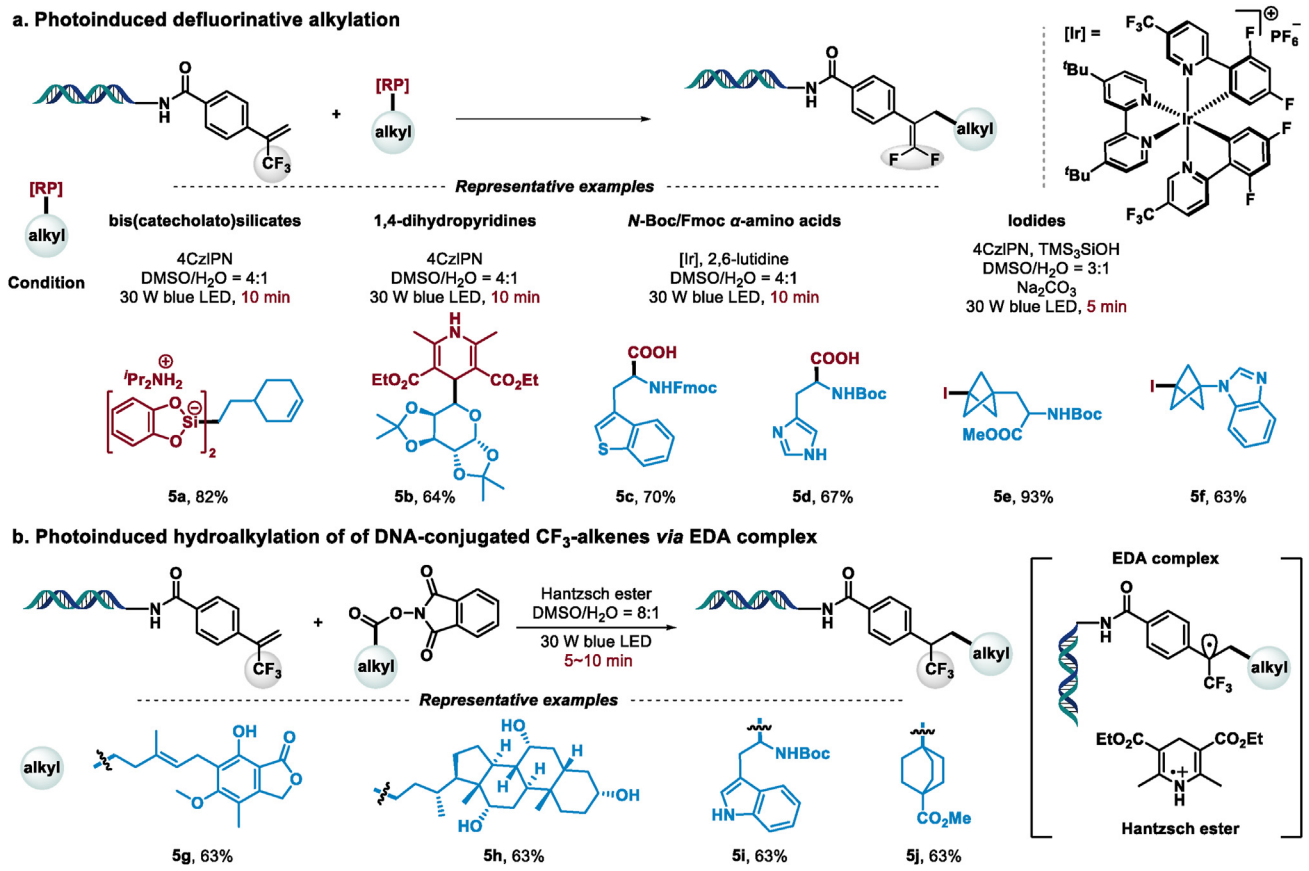


Figure 5 Photoinduced defluorinative alkylation and hydroalkylation of DNA-conjugated CF_3−alkenes. (a) Photoinduced defluorinative alkylation. (b) Photoinduced hydroalkylation of DNA-conjugated CF_3−alkenes *via* EDA complex.

(Trp−Pro), thus extending the chemical space greatly for DEL synthesis (Fig. 6b)[76].

In 2020, Kölmel et al.[77] achieved a photocatalytic [2 + 2] cycloaddition reaction between DNA-conjugated styrenes and cinnamates, which efficiently produced highly substituted cyclobutanes containing a set of medicinal chemistry important heterocycles with good DNA fidelity. Notably, a three-step on-DNA synthesis consisting of amide condensation, photocatalytic [2 + 2] cycloaddition, and reductive amination was achieved in good overall yield, highlighting its great application potential in DEL synthesis (Fig. 7a). In 2022, Hu group reported a photocatalytic [2 + 1] cycloaddition reaction between DNA-conjugated styrene and diazo compounds. This method shows excellent functional group tolerance, providing corresponding cyclopropanation products in moderate to high yields (Fig. 7b)[68].

In summary, photocatalyzed reaction has great potential in DNA-compatible reaction development. Particularly, it provides an efficient way to access the challenging C(sp^2)−C(sp^3), C(sp^3)−C(sp^3) and also C(sp^2)−C(sp^2) bonds under room temperature. In addition, these reactions have greatly expanded the choice of BBs, which is critical for the split and pool DEL synthesis.

2.2.2. *DNA-compatible C−H activation and functionalization*

C−H activation has become a powerful and versatile strategy for the synthesis and modification of valuable BBs. Intermediates and drug molecules without the need to pre-functionalize the parental molecules in this strategy, thus traditional inert C−H bond could serve as a reactive functionality, making "multi-functional BBs" more readily accessible. Consequently, these merits have attracted chemists to incorporate this strategy into DEL synthesis.

In 2018, Lu et al.[78] reported the first DNA-compatible C−H activation reaction between DNA-conjugated acrylamides and aromatic acids (C(sp^2)−C(sp^2)) (Fig. 8a). In addition, the reverse version (DNA-linked benzoic acid and ethyl acrylate) of this reaction was also feasible, albeit it produced a spot of bis-olefination by-product. In 2020, our group achieved the first on-DNA C−H selenylation of privileged indole skeleton under Rh(III) promotion by designing benzoselenazolone (BSEA) as a hydrostable and bifunctional selenylation reagent (C(sp^2)−Se)[79]. It provides an efficient procedure for the synthesis of biologically important selenium-containing DEL (*Se*DEL) (Fig. 8b). Soon after, we developed an Ir(III)-mediated C−H activation/[4 + 2] annulation reaction between DNA-conjugated benzoic acid and *gem*-difluoromethylene alkynes, affording DNA-conjugated difluorinated isocoumarins efficiently (Fig. 8c)[80]. In addition, we further reported a Rh(III)-mediated [4 + 3] annulation of DNA-conjugated *N*-methoxy amides with *gem*-difluorocyclopropenes as innovative *β*-monofluorinated three C(sp^2) sources, which directly generated seven-numbered 2*H*-azepin-2-one frameworks (Fig. 8d)[81].

To efficiently incorporate C(sp^3) center into DEL from the readily available starting materials, Yu et al. successfully translated their previous C(sp^3)−H arylation efforts into the DEL setting. The C(sp^3)−H arylation reactions of DNA-linked aromatic iodides with simple aliphatic acids[82], *α*-amide acids, and oximes were developed with satisfactory DNA stability. Interestingly, the involved ligands could not only improve the conversions

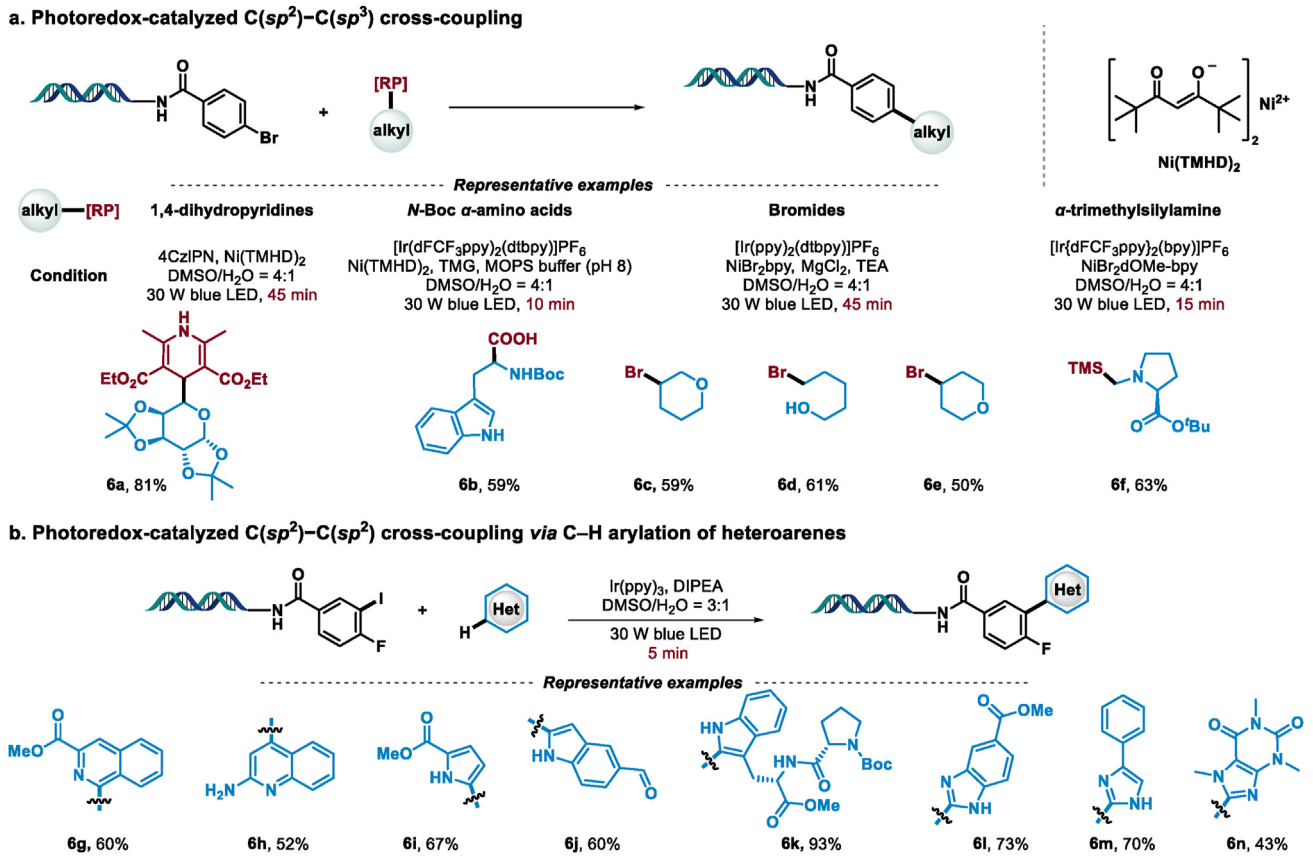


Figure 6 Photoredox promoted C(sp^2)−C(sp^3) and C(sp^2)−C(sp^2) cross-coupling. (a) Photoredox catalyzed C(sp^2)−C(sp^3) cross-coupling. (b) Photoredox catalyzed C(sp^2)−C(sp^2) cross-coupling *via* C−H arylation of heteroarenes.

but also decrease the degradation of DNA. The potential applications of these reactions in DEL synthesis were demonstrated by a three cycles rehearsal including C(sp^3)−H arylation of free acid (73%), amidation (71%), and C(sp^3)−H arylation of oxime (28%), delivering the final product (14.5% total yield) rich in linkage and structural element diversity (Fig. 9a).

Apart from the C−H activation strategy, in 2020, Lu group[83] demonstrated a photoinduced cross-dehydrogenative coupling

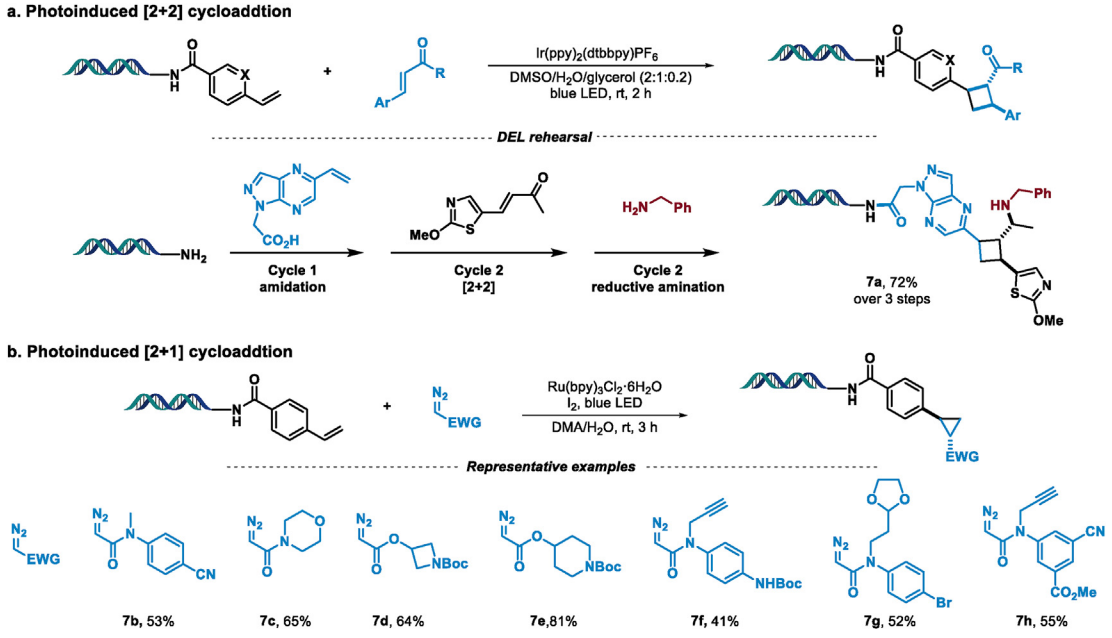


Figure 7 Photoinduced cycloaddition reaction. (a) Photoinduced [2 + 2] cycloaddition. (b) Photoinduced [2 + 1] cycloaddition.

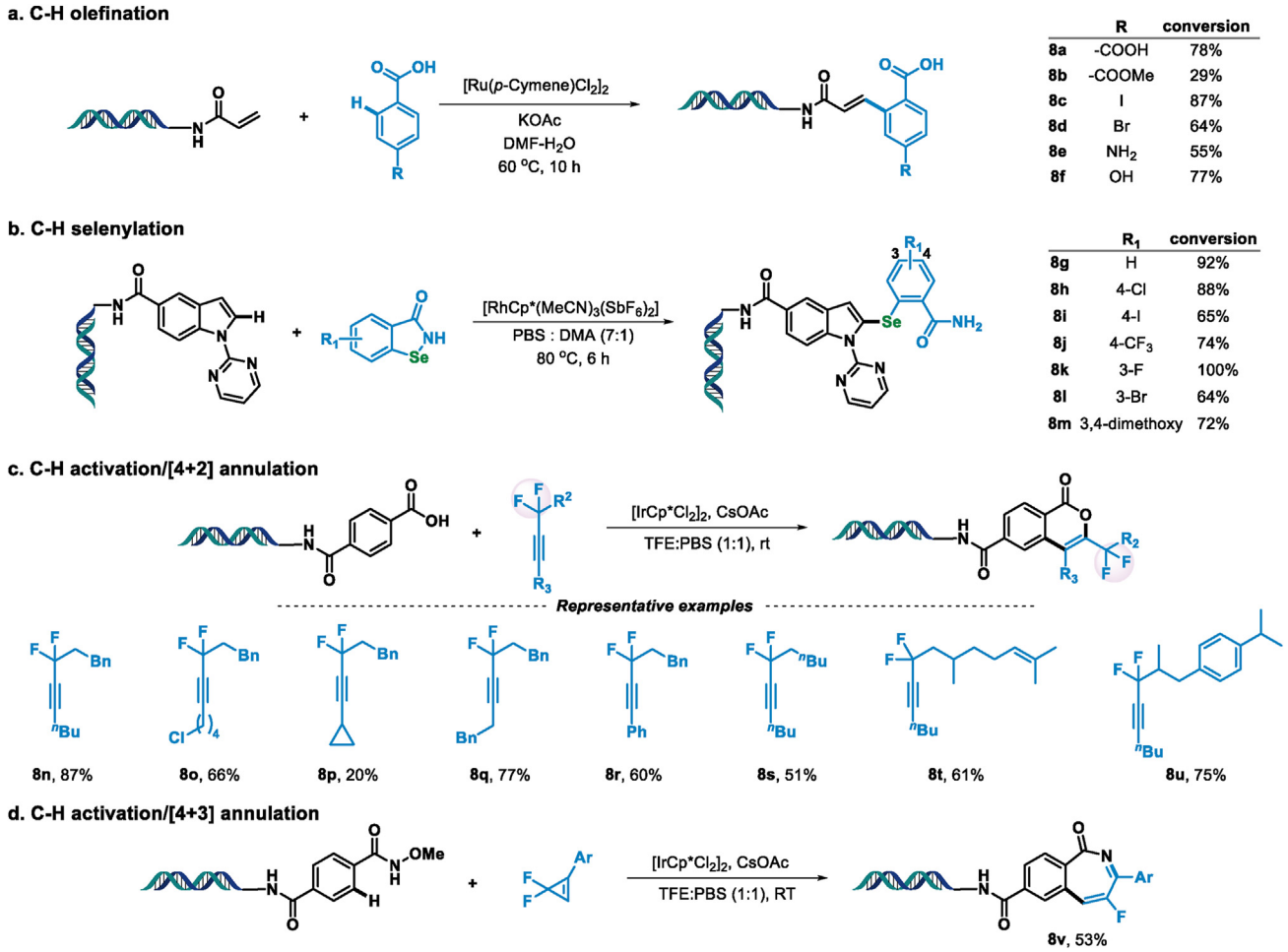


Figure 8 DNA-compatible C(sp^2)−H activation. (a) C−H olefination. (b) C−H selenylation. (c) C−H activation/[4 + 2] annulation. (d) C−H activation/[4 + 3] annulation.

(CDC) reaction for C(sp^3)−H functionalization of DNA tagged tetrahydroisoquinolines (THIQs) under mild and open-air conditions (Fig. 9b). The iminium cation formed *via* photo-promoted oxidation of the phenyl-substituted tertiary amine of THIQ was the deductive key intermediate of this transformation.

Overall, the C−H activation approach has become an emerging strategy in DEL synthesis from C−C bond formation and C−Se bond formation to the on-DNA synthesis of heterocycles. In addition, in this strategy traditional inert C−H bond could serve as a reactive group, making "multi-functional BBs" to be more readily accessible, which has great merits in the diversification of DELs.

2.2.3. *Bioinspired click selenylation*

Click chemistry is a powerful molecular assembly strategy for rapid functional discovery[84]. However, the know click chemistry has focused predominately on the employing of carbon or sulfur as the connecting centers[85]. Therefore, the development of click reaction with new connecting linkage is of great importance for both expanding the toolbox of click chemistry and for new functional discovery. Organo-Se species is a series of emerging core motifs widely found in endogenous proteins and small molecules. In recent years, Se-containing small molecules have attracted considerable attention in many research areas, particularly in the realm of medicinal chemistry, for their various promising biological activities[86,87]. Therefore, the development of new click chemistry using biological important Se as the connecting center would be appealing[88].

Inspired by the cellular biochemical reaction of BSEA and the Cys residue of proteins, our group has developed the first click selenylation reaction between indole and BSEA by using the LUMO activation strategy (Fig. 10a)[89]. This click selenylation reaction is robust, modular, predictable, and highly site-selective, and it features mild and simple reaction system (promoted by non-metallic $B(C_6F_5)_3$), thereby can be applied for the on-microplate parallel synthesis and *Se*DEL synthesis. Later, we achieved the first sulfoximine pharmacophore-directed and rhodium (III)-catalyzed C−H/N−H annulative selenylation directly with simple elemental Se, which provides a novel class of organo-Se species, namely benzothiaselenazole-1-oxides (BTSAs) (Fig. 10b)[90]. Notably, we demonstrated that BTSA can be recognized as a new kind of click selenylation reagent, which also enabled efficient on-microplate parallel synthesis and *Se*DEL synthesis.

Besides, some other selenylation methods are also reported. In 2021, based on the strategy of reversible adsorption to solid support (RASS), the Zhang group developed a visible-light promoted C(3)−H selenylation reaction between DNA-conjugated indole and diselenide *via* an EDA complex under open-air conditions, producing DNA-conjugated selenides in medium to excellent yields (Fig. 10c). Notably, the results of qPCR analysis

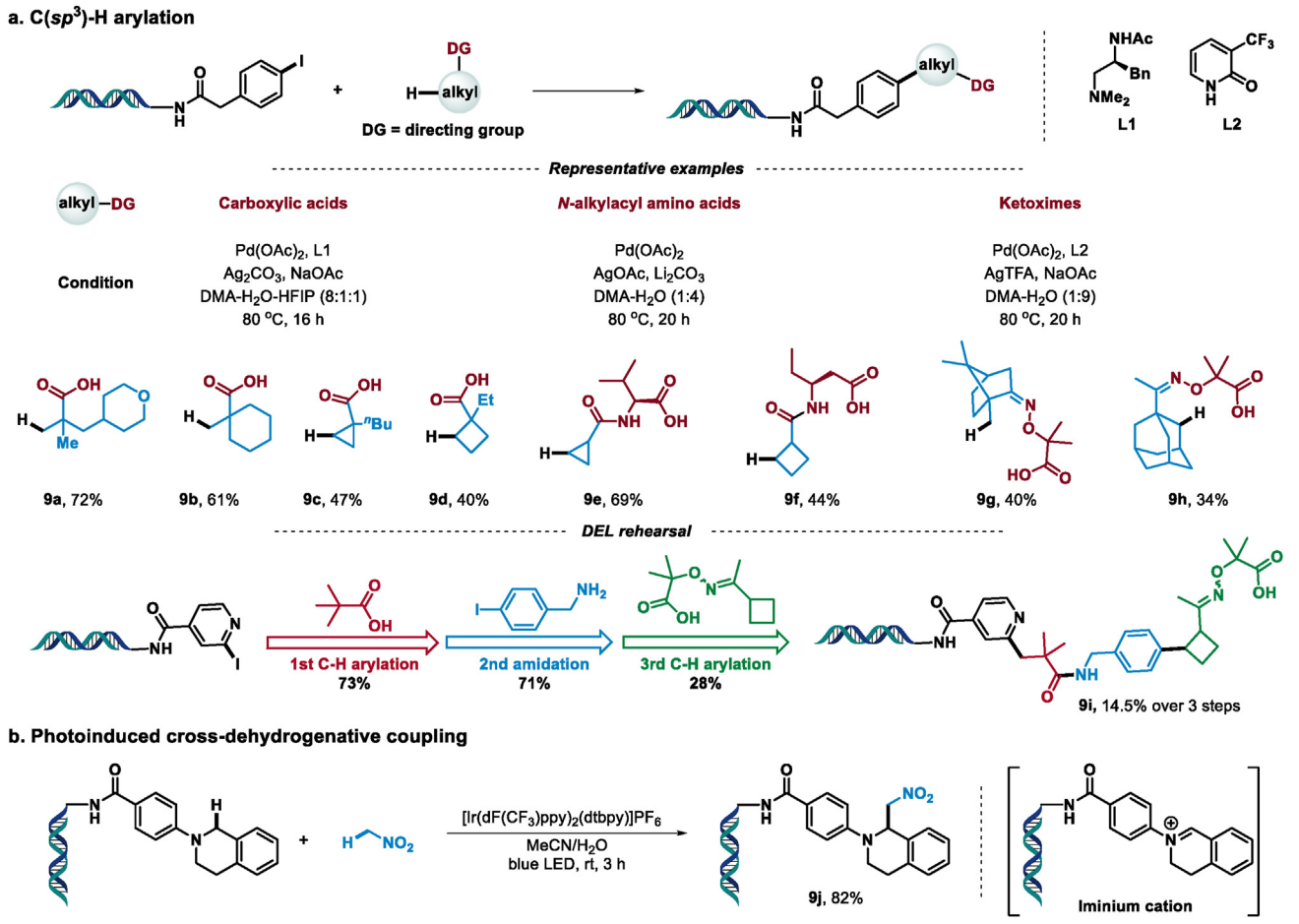


Figure 9 DNA-compatible C(sp^3)−H functionalization. (a) C(sp^3)−H arylation. (b) Photoinduced cross-dehydrogenative coupling.

and next-generation sequencing indicated no apparent DNA damage during this photoredox reaction[91]. In addition, Li and colleagues reported an in-solution oxidative selenylation of DNA-conjugated electron-rich arene (ERA) by using iodine/BSA (bovine serum albumin) as the oxidation system (Fig. 10c)[92]. This method features a transition-metal-free process and a broad scope of Se source, including diselenide, selenol, selenyl chloride, sulfonyl selenol, and Ebselen.

Overall, the mild and efficient on-DNA click selenylation reaction of BSEA and BTSA as well as other methods have greatly expanded the chemical space of *Se*DEL, and thus will be helpful to accelerate Org-Se hit-discovery endeavor in the future.

2.2.4. *Micellar-mediated reactions*

Generally, acidic conditions (pH < 4) entail the risk of DNA depurination, thus ruling out the use of many protic acid-catalyzed reactions such as Brønsted acids-catalyzed heterocycle synthesis in the DEL setting. To address this obstacle, Brunschweiger et al.[93] incorporated micellar catalysis into DEL synthesis. They proposed that the amphiphiles associate in water with nanometer-sized micelles that are characterized by a hydrophilic corona and a hydrophobic core (Fig. 11). The micelles can form micro-heterogeneous systems to solubilize hydrophobic chemicals, in which the hydrophobic core serves as nanoreactor to package all the reactants including the lipophilic part of the DNA-conjugated substrate, while the hydrophilic DNA is in the aqueous phase to be protected from chemical modification and degradation.

Following this assumption, Brunschweiger et al.[93] designed and synthesized a sulfonic acid-substituted block copolymer (**M1**, Fig. 11b, sulfonic acid substituted at hydrophobic core) as an amphiphile that could associate at very low concentrations to minimize the fraction of free catalysts to interact with DNA in the water phase. By using **M1**, they developed the first Povarov reaction of DNA-conjugated benzaldehyde, which efficiently produced DNA-conjugated tetrahydroquinoline (**12a**, >90%) with no significant DNA damage (Fig. 12a). On the contrary, block copolymer **M2** (Fig. 11c), with $-SO_3H$ group locating in the hydrophilic part of the amphiphile, showed worse catalytic behavior and reproducibility. Besides, the block copolymer **M3** (Fig. 11d) without covalently linked sulfonic acid failed to mediate this transformation due to the absence of a catalytically active moiety. Similarly, the Gröebke−Blackburn−Bienaymé reaction of DNA-conjugated benzaldehyde with 2-aminopyridine and isocyanide also worked well[93], efficiently affording the DNA-conjugated 3-aminoimidazo-[1,2-*a*]pyridine by using sulfonic acid-substituted block copolymer **M1** as the acidic nanoreactor (Fig. 12b). Moreover, oxidation conditions are usually excluded from DEL synthesis for the potential DNA damage by 8-oxoA or 8-oxoG formation. To address this, they synthesized a bipyridine ligand containing amphiphilic block copolymer (**M4**, Fig. 11e) to immobilize copper(I) (Fig. 12c)[93]. Remarkably, the

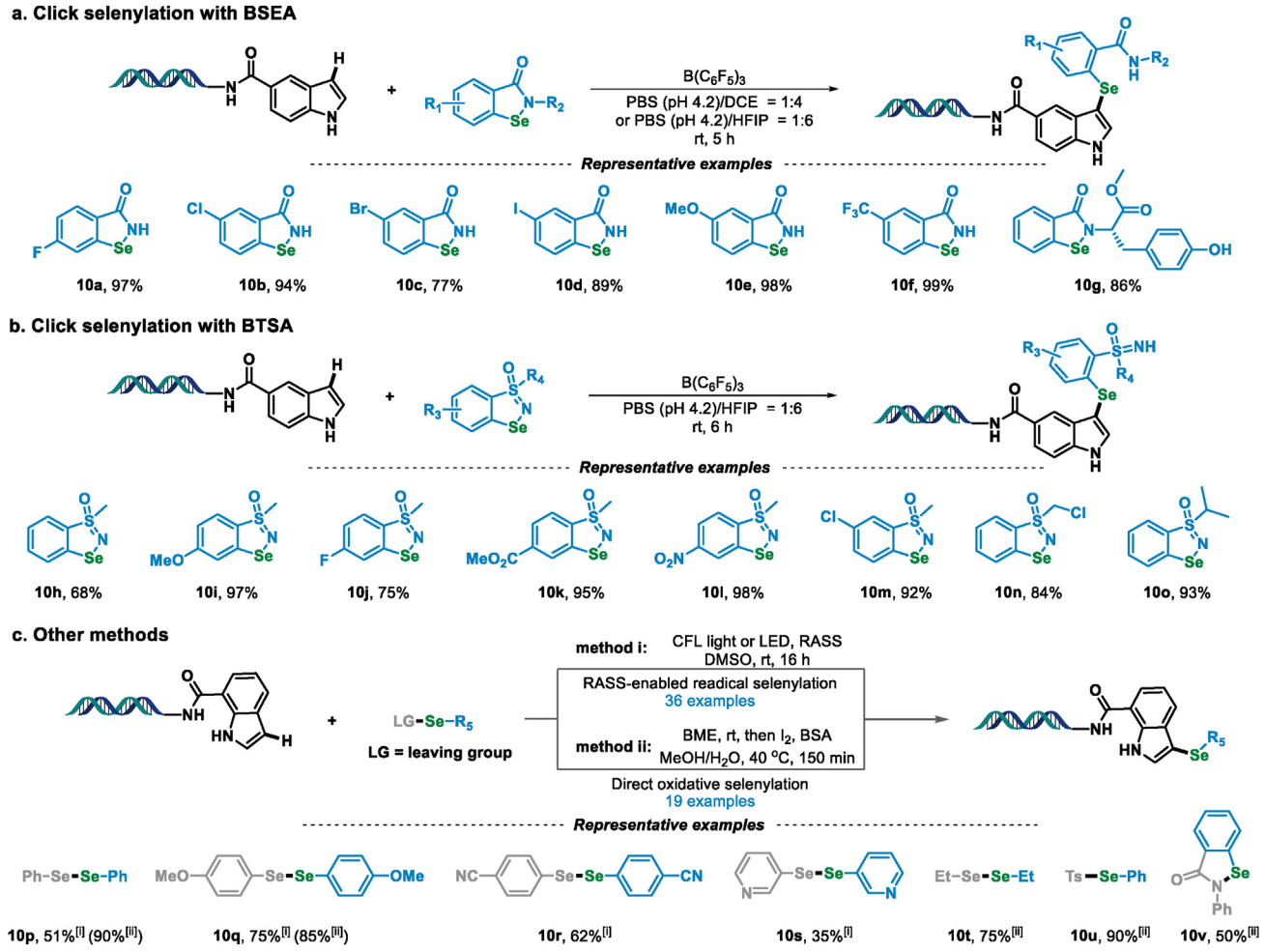


Figure 10 DNA-compatible C(3)−H selenylation of indole. (a) Click selenylation with BSEA. (b) Click selenylation with BTSA. (c) Other methods.

micellar-bipyridine/CuBr/TEMPO system showed a time- and micelle-dependent oxidation of DNA-conjugated benzyl alcohol without detectable formation of 8-oxopurines or overoxidation to the carboxylic acid.

Soon afterward, Waring et al.[94] reported the use of micellar catalytic strategy in optimizing on-DNA Suzuki−Miyaura coupling reaction by using commercially available surfactant TPGS-750-M. The best condition led to exceptionally efficient reaction across the full range of tested substrates with good DNA fidelity, including some more desirable but problematic heterocyclic substrates (**12g−12l**, 81%−100%, Fig. 12d).

Overall, micelles-mediated reaction has shown great potential in reducing DNA damage. In addition, it also provides a strategy to conduct on-DNA chemical synthesis in organic solvents or nano-heterogeneous systems, thereby having the potential to access unprecedented molecular diversity.

2.2.5. Reversible adsorption to solid support (RASS) approach

The aqueous reaction conditions have excluded many reactions that involve water-sensitive reagents or intermediates. Bringing DNA from aqueous condition to organic solvents is a useful tactic to incorporate such reactions into the DEL chemistry toolbox. As a result, Brunschweiger group has developed the controlled porous glass (CPG) solid support strategy to conduct the reaction in organic solvent and encoded after cleavage from the solid support[95]. Besides, Paegel and Kodadek et al.[96,97] also described the water-free DNA-encoded solid-phase synthesis (DESPS) based on the one-bead-one-compound (OBOC) strategy by using polystyrene-polyethylene glycol beads (PS−PEG Tentagel). However, both strategies require the DNA tags covalently linked to the solid supports, the additional covalent bonding and cleavage steps may to some extent reduce their practicality.

In this context, Dawson, Baran, and colleagues developed a reversible adsorption to solid support (RASS) strategy for DEL synthesis to bring DNA into organic solvents, thus expanding the scope of DNA-compatible reactions[40]. A mixed mode polystyrene strong anion exchange resin containing a butyl quaternary ammonium moiety (Phenomenex, *Strata-XA* resin, Fig. 13a) was identified as an appropriate solid support, which incorporates both hydrophobic and electrostatic interactions to anchor DNA, allowing the use of neat organic solvents. The workflow of a RASS includes: i) binding the DNA-conjugated substrate to resin→ii) wash decant→iii) conducting the desired reaction→iv) quench wash→v) eluted from the resin→vi) precipitation with cold EtOH→vii) enzyme encoding.

Under the RASS setting, the kinetic issue related to highly diluted conditions in the water faded, enabling the rapid development of a Ni-catalyzed C(sp^2)−C(sp^3) decarboxylative cross-

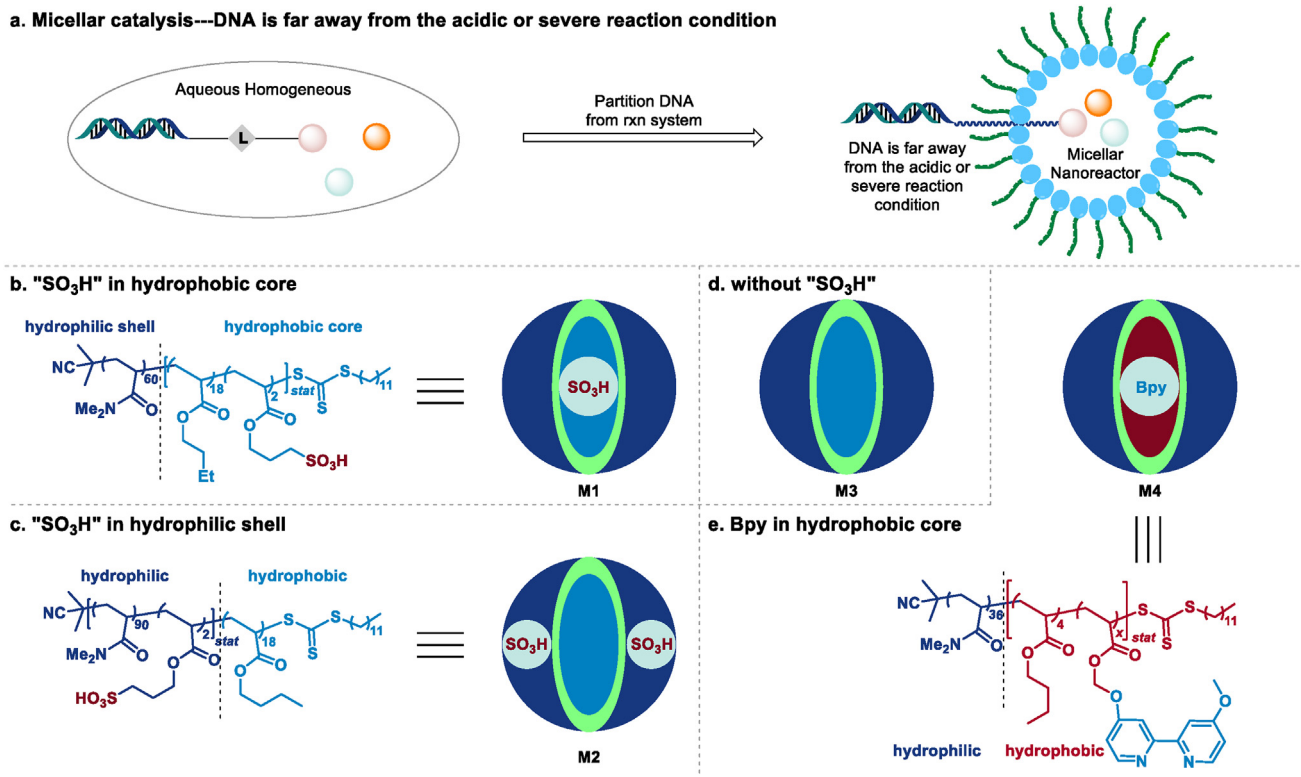


Figure 11 Graphic description of micellar-mediated reactions. (a) On-DNA micellar catalysis—partition DNA from the acidic or severe reaction condition. (b) Block copolymer with "SO_3H" in the hydrophobic core. (c) Block copolymer with "SO_3H" in a hydrophilic shell. (d) Block copolymer without the "SO_3H" group. (e) Block copolymer with "bipyridine" in the hydrophobic core.

coupling with broad substrate scope (Fig. 13b)[40]. Besides, RASS also offers a unique opportunity to site-isolate DNA, thereby rendering it amenable in electrochemical transformation. Accordingly, the first Ni/electrochemical amination of DNA-conjugated aryl iodides with alkyl primary amine, alkyl secondary amine, heteroaryl primary amine, and pyrrolidinone (amide) was achieved[40]. Although about half of the tested cases gave <50% conversions, it opens a door to translate electrochemistry into DEL synthesis (Fig. 13c). Moreover, using the RASS strategy, the authors achieved improved on-DNA reductive aminations between DNA tagged amines and a series of carbonyl compounds. This newly-established $B(OH)_3$-mediated reductive amination exhibited expanded generality in both aqueous and RASS contexts, some challenging substrates including acyclic ketones, benzylic carbonyls, and sterically hindered carbonyls (adamantyl) were also competent (Fig. 13d). Notably, to illustrate the application potential in DEL build of the developed three reactions in RASS setting, the authors conducted a three-cycle synthesis including cross-coupling/deprotection, reductive amination, and finally electrochemical amination from a DNA tagged aryl-iodide substrate, generating a product, rich in linkage diversity with 9% yield over 4 steps. This rehearsal indicated that the RASS strategy is amenable in multiple syntheses (Fig. 13e).

Based on these good results, the same group further developed a suite of reactions for the construction of medicinally relevant C−S, C−P, and S−N bonds on DNA *via* a RASS approach[98]. For instance, i) the first on DNA C(sp^2)−S and C(sp^2)−P cross-coupling between DNA-conjugated aryl iodide and mercaptan or thiophenol, and phosphinic chlorides under a homogeneous Ni/Si system in DMA. Notably, phosphinic chlorides are extremely sensitive to water, even 1% water quenches the C(sp^2)−P bond formation reaction on the RASS (Fig. 14a); ii) The on-DNA synthesis of sulfones *via* a nucleophilic substitution of alkyl or aryl sodium sulfinate salts to DNA-tagged alkyl bromides. Alternatively, sulfones can also be obtained *via* two steps, which consist of a nucleophilic substitution by mercaptan or thiophenol, followed by electrochemical oxidation mediated by TEMPO (Fig. 14b); and iii) the on-DNA synthesis of sulfonamide *via* oxidative coupling between sulfinates and DNA-tagged amines (Fig. 14b)[98].

At about the same time, Berst et al. independently reported a similar strategy which was defined as the "catch-and-release" to circumvent potential solvent-dependent swelling issues[99]. A cationic, **a**mphiphilic PEG-based **p**olymer (Fig. 15a) was used to perform chemical **t**ransformations on immobilized DNA conjugates under **a**nhydrous **c**onditions (APTAC). Under the APTAC setting, a water-sensitive nucleophilic addition to DNA-conjugated ketone by using benzaldehyde as the alkyl carbanion equivalent was achieved in 75% conversion (Fig. 15b). In addition, the synthesis of medium ring and spirocyclic heterocycles by using the Sn amine protocol (SnAP) in dichloromethane was also realized. Although the conversions may be poor *via* a single reaction cycle, one of the advantages of reaction on solid support is the ability to rapidly perform successive reaction cycles to improve conversion. Significantly, no obvious DNA degradation was observed in the SnAP protocol (53% of amplifiable DNA after reaction and 59% DNA recovery release from the resin without reaction). Overall, the RAAS and APTAC approaches enable the

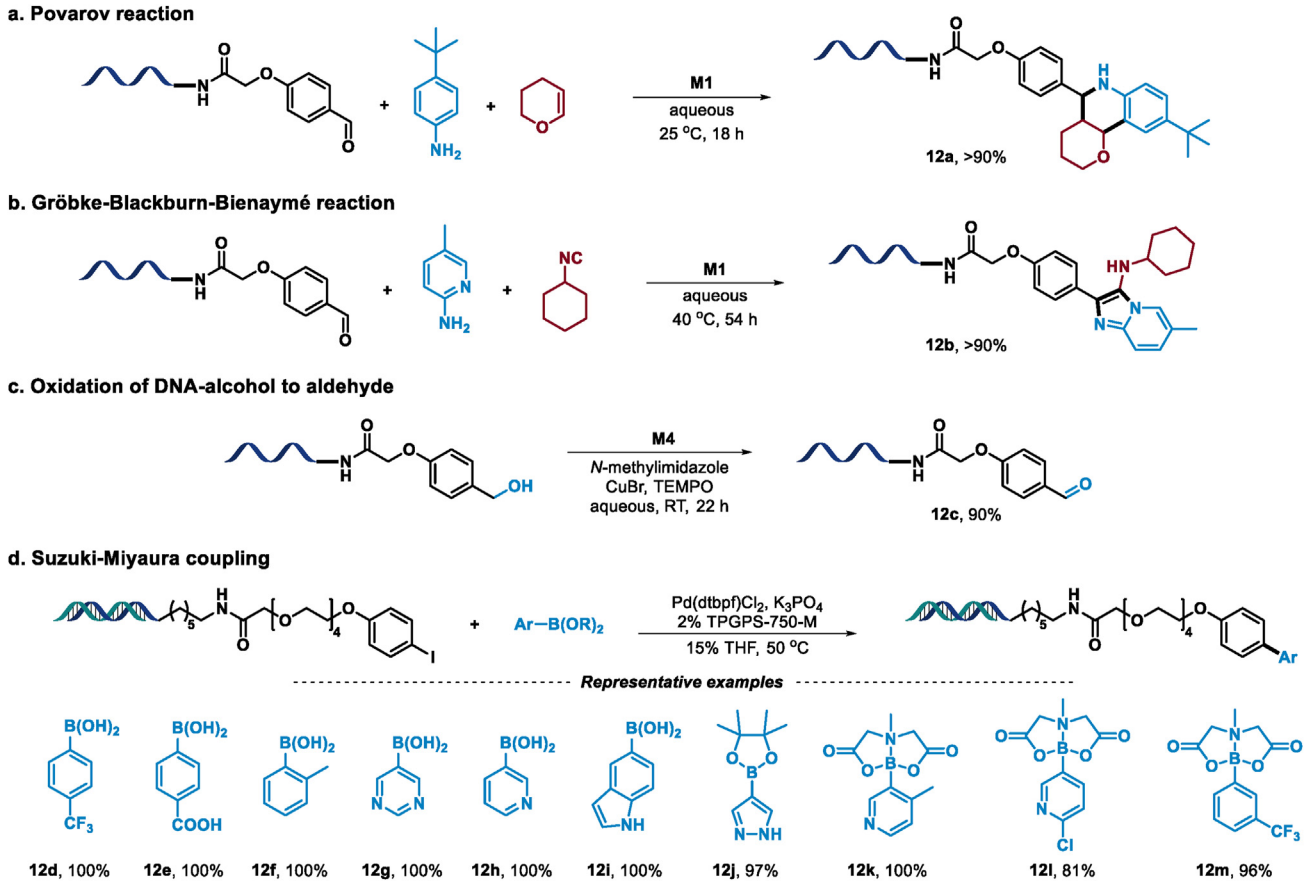


Figure 12 Micellar-mediated reactions. (a) Povarov reaction. (b) Gröbke−Blackburn−Bienaymé reaction. (c) Oxidation of DNA-alcohol to aldehyde. (d) Suzuki−Miyaura coupling.

on-DNA synthesis in neat organic solvents, which will facilitate the water-sensitive reaction development. However, the DNA recovery is only about ∼30% in some cases, which may result from reaction workup and precipitation procedures[40]. Moreover, the amount of resin will increase significantly when a large number of BBs are used, which will make the "split and pool" synthesis impractical. These issues should be circumvented to meet the requirement of practical DEL synthesis.

2.2.6. *Substate activation strategy*

Adopting the substrate activation strategy, and using highly reactive reagents as substrates can effectively avoid harsh conditions such as high temperature and transition-metal addition, thus facilitating the development of DNA-compatible reactions with high DNA fidelity.

In 2020, Schreiber et al.[100] employed the *in situ* generated 6-membered azacyclic allene as the highly reactive substrate and developed a series of strain-promoted cycloaddition reactions for diversity-oriented synthesis of structurally diverse rigid core structures in a C(sp^3) rich fashion. Olefins, N-substituted pyrroles, and 1,3-dipoles undergo [2 + 2], [3 + 2], and [4 + 2] reactions smoothly at room temperature (Fig. 16a). Notably, CsF is the only additive needed for *in situ* allene formation. Under high fluoride concentrations (substantially more activating agent than that needed for efficient allene formation), a full-length DEL showed *ca*. 80% remaining amplifiable materials, highlighting the excellent DNA fidelity of these DNA-compatible cycloaddition reactions.

Although diaryl ethers are privileged motifs in pharmaceuticals and natural products, the on-DNA synthesis of C(sp^2)−O is challenging due to the inherently weak nucleophilic activity and poor reactivity of phenols. Recently, using the substrate activation strategy, our group achieved a metal-free and open-air on-DNA arylation of phenols[101]. The highly reactive diaryliodonium salts (DAIs) were used as Ar^+ source to replace Ar−X (X = halogen, OTf, OFs, activated by metal to form Ar−M−X, Ar^-) in conventional cross-coupling, thus converting the cross-coupling pathway into concerted nucleophilic aromatic substitution (CS_NAr) pathway. The desired diarylethers were efficiently prepared on DNA by simply regulating the pH to 9.4. Notably, a DEL library showed good DNA fidelity with *ca*. 86% remaining amplifiable materials under this arylation condition. This reaction can also be applied to the late-stage on-DNA modification of tyrosine-contained peptide (**16k**) and to synthesize DNA-tagged analogs of market drug sorafenib (**16l**), a pan-kinase inhibitor. Moreover, DAIs are also suitable for on-DNA arylation of oximes, affording *N*-aryl nitrones with good DNA fidelity (Fig. 16b)[101].

Overall, these positive results indicate that substrate activation is an effective strategy for developing mild and efficient DNA-compatible reactions with high fidelity.

3. Evolution-based DEL selection strategies

DEL was originally inspired by the phage display library, which was an evolution-driven selection system to isolate biological

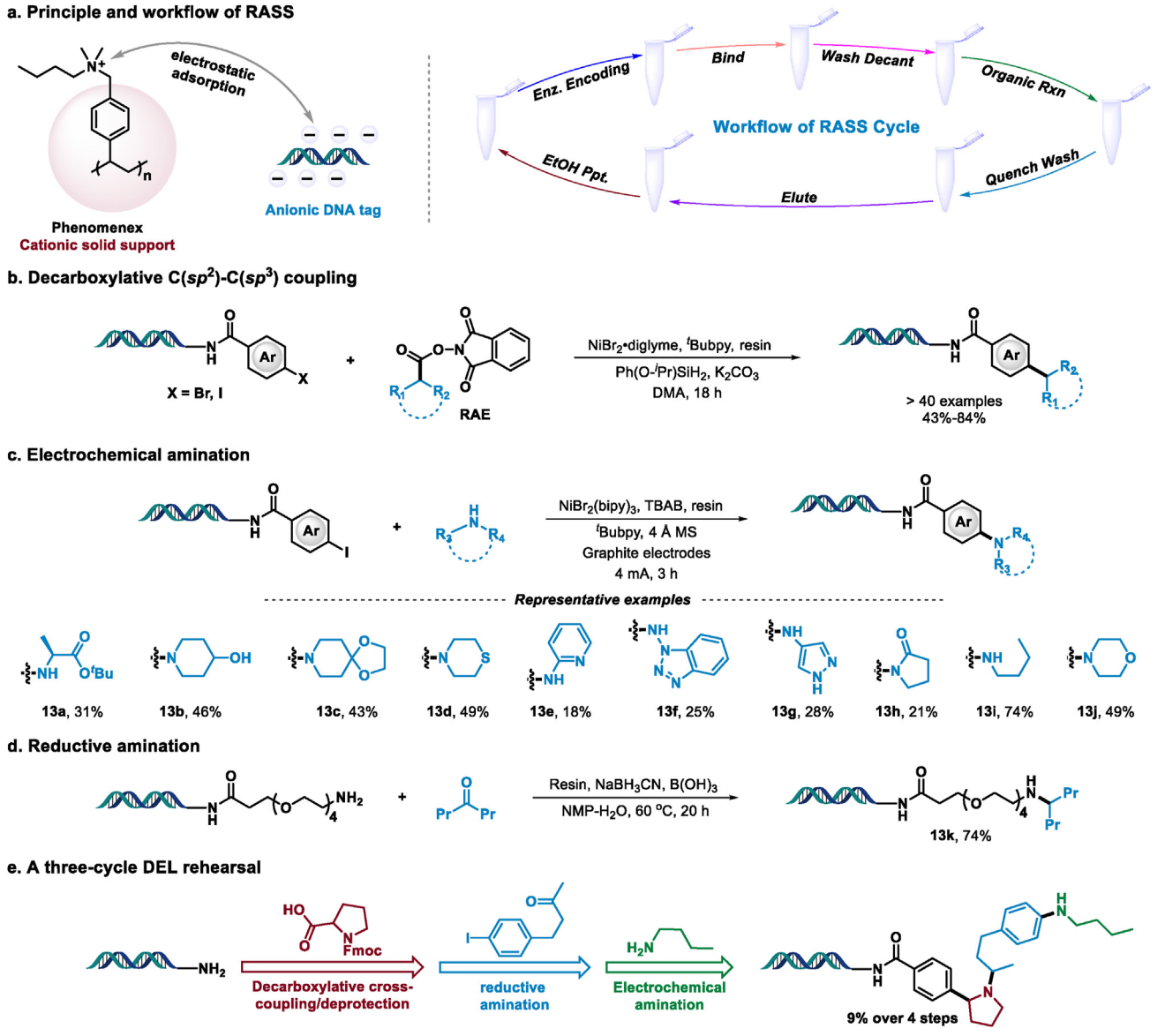


Figure 13 RASS-enabled reactivity expanding. (a) Principle and workflow of RASS. (b) Decarboxylative C(sp^2)−C(sp^3) coupling. (c) Electrochemical amination. (d) Reductive amination. (e) A three-cycle DEL rehearsal.

molecules such as antibodies and peptides[28]. Therefore, the selection system was intrinsically analogous to the evolution system, *i.e.,* selective pressures from the environment were applied to pan out the survivors over time[102]. In the biological system, rapidly replicating organisms, such as a bacteriophage[5], mammalian virus[103,104] or a yeast[105,106] allow scientists to perform selections on the bench in a couple of days. However, due to the time of chemical synthesis, the iteration is limited for the selection of DELs. In addition, the main challenge for the selection in solution is how to distinguish the binders and nonbinders. Most ligands non-covalently bind to the target and their affinities are limited, *i.e.,* they can dissociate during the following step. Therefore, increasing the effective target, and ligand concentration and slowing down the dissociation of DEL molecules are the fundamental issues to solve in the selection methodology development.

3.1. *Affinity-based selections against targets on solid phase*

The most direct selective pressures are the direct interaction between target proteins and small molecules (Fig. 17). Along the line of biomolecule selection, DEL selection is normally considered to be a massive binding assay conducted over an immobilized protein to identify the physical binders using the typical bind-wash-elute procedure (Fig. 17a)[37]. The target proteins of interest are immobilized on the solid supports to capture the binders from the DEL in solution. Magnetic beads were separated by the magnetic field, while the sepharose was isolated by centrifuge or filters. Nickel[18,107] or cobalt[108] chelate affinity resin was applied to capture his-tag recombinant proteins. Streptavidin-coated beads can capture the biotinylated protein, which does not have recombinant tags, and cyanogen bromide (NCBr)-activated

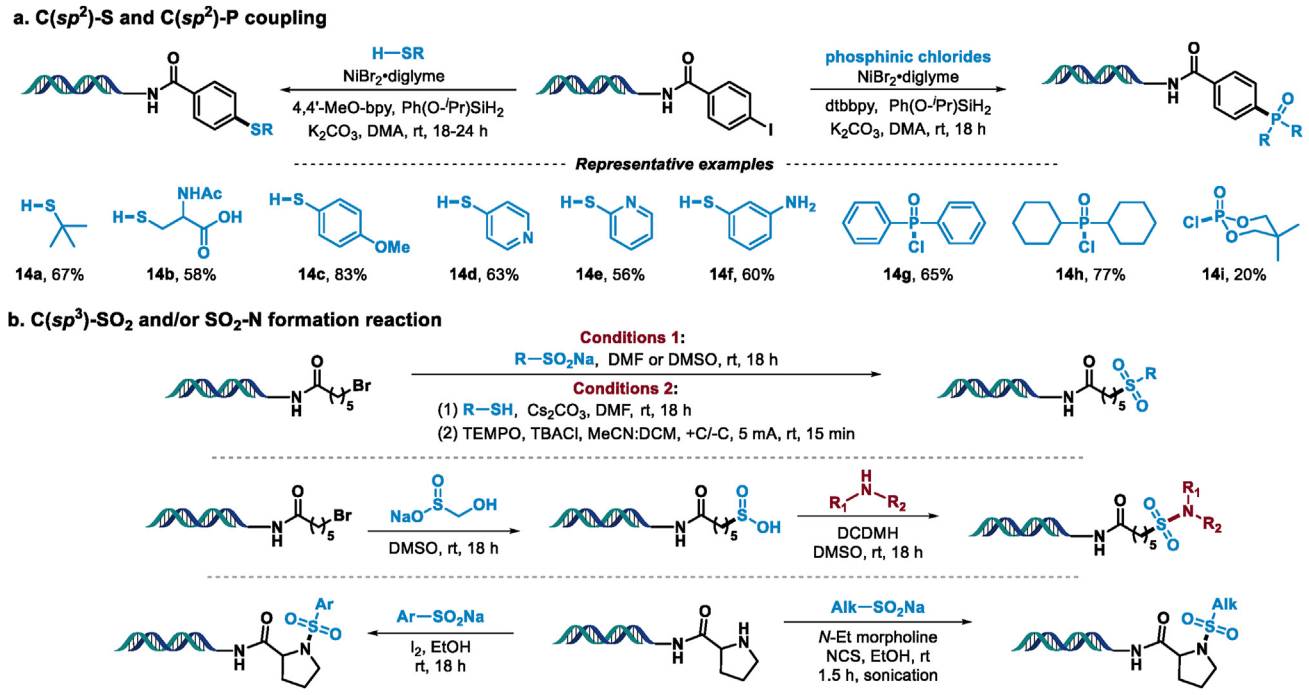


Figure 14 RASS-enabled S/P—C and S—N bond formation reactions. (a) C(sp^2)−S and C(sp^2)−P coupling. (b) C(sp^3)−SO_2 and/or SO_2—N formation reaction.

sepharose can covalently immobilize the proteins *via* their primary amines in a mild reaction condition, thus can be broadly applied to different types of proteins[109,110]. To obtain good recovery and well-resolved selection fingerprints, multiple experimental parameters (*e.g.*, protein capture procedures, detergent type, detergent concentrations, washing conditions, and elution conditions) should be optimized[111]. The activity and structure of the target protein are key to successful DEL selection, the protein

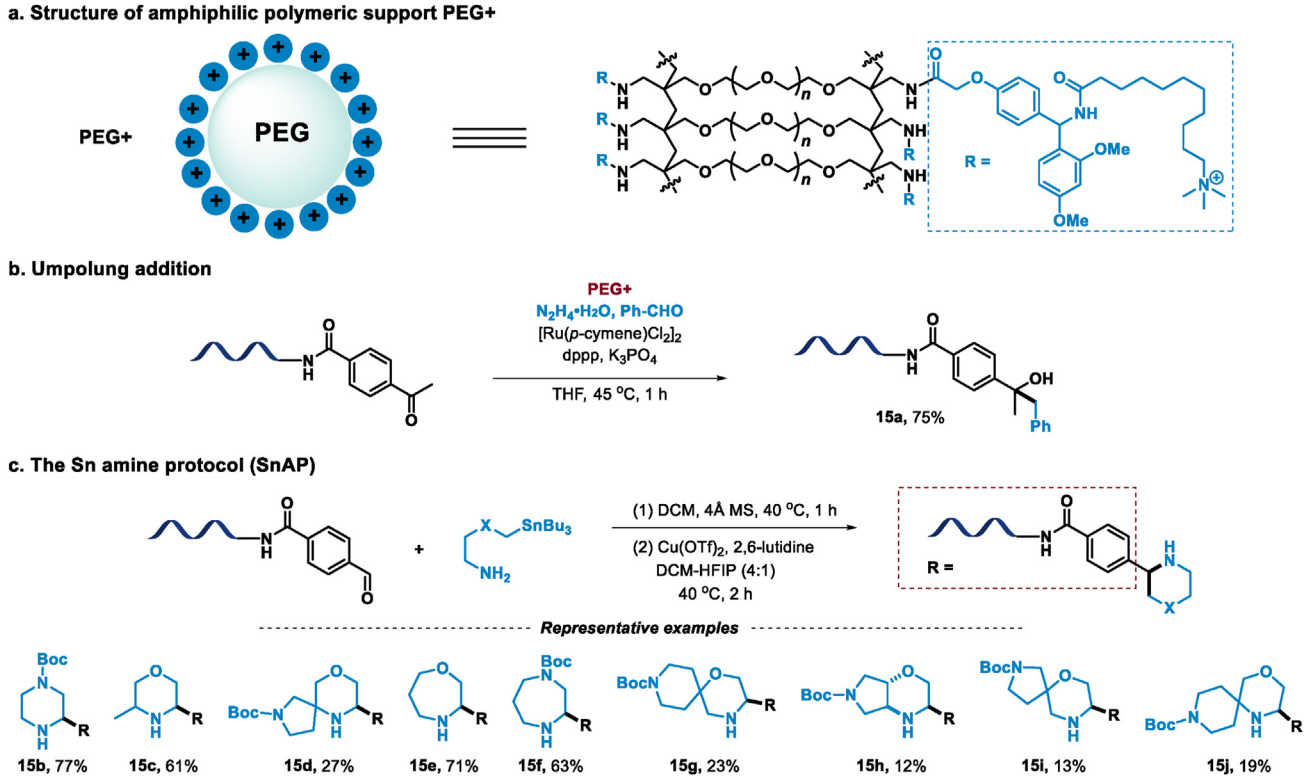


Figure 15 APTAC-enabled transformations. (a) Umpolung addition. (b) The Sn amine protocol (SnAP).

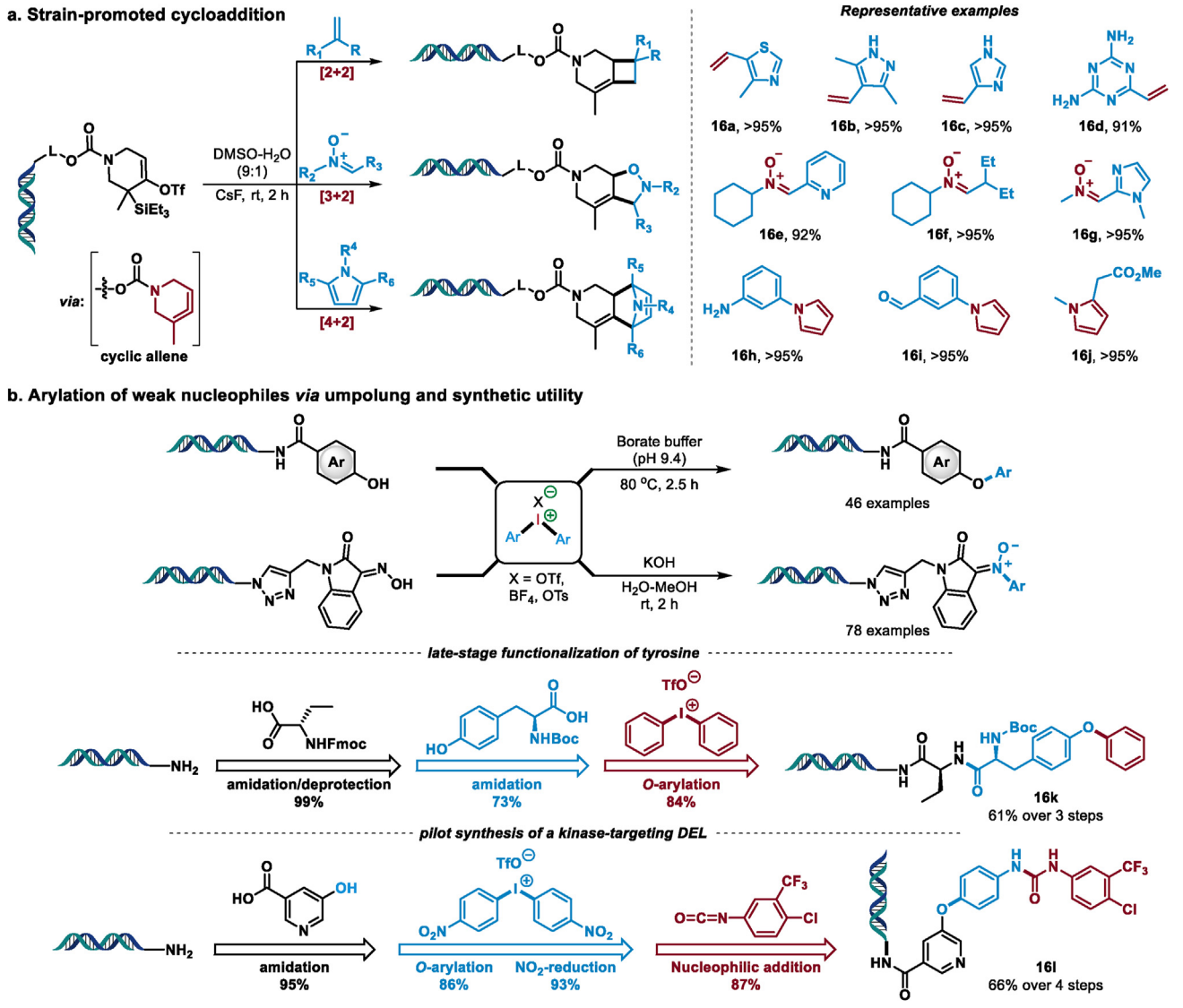


Figure 16 Substate activation strategy-enabled transformations. (a) Strain-promoted cycloaddition. (b) Arylation of weak nucleophiles *via* umpolung and synthetic utility.

capture requires a mild condition, *i.e.,* pH shift from the protein isoelectric point (pI), Good's buffer to avoid precipitation, and reasonable salt concentrations. To minimize the undesired interactions between DEL molecules and the target protein, the immobilized material, sheared salmon sperm DNA was supplemented into the selection buffer[111]. Remarkably, unspecific interaction between DEL molecules and solid phase can't be ignored, especially for the non-enzyme proteins. Normally enzymes have active binding sites, which provide plenty of hydrogen bonding, electrostatic interactions, and hydrophobic interactions to hold the enzyme and substrate together in the complex.

Currently, most targets are non-enzyme, *e.g.,* G-protein coupled receptors (GPCRs), nuclear receptors, and ion channels. These targets functions *via* protein−protein interactions (PPI). Advanced proteomics studies predicted 650,000 PPIs in the organism, much more than the number of proteins (20,000)[112]. Therefore, the discovery of PPI modulators attracting more and more attentions[113,114]. However, PPIs were regarded as "undruggable", because protein−protein interfaces were much larger (*ca.* 1000−2000 Å^2) and flatter compared to the enzyme active sites (*ca.* 300−500 Å^2)[115]. The surface pockets at druggable sites are a key feature for the druggable protein[116], and PPIs with buried surface area (BSA) less than 2000 Å^2 were more likely to be inhibited by small molecules, while PPIs with larger BSA values were typically inhibited by peptides[117], all of which could be annotated with DNA tag in the DEL. The large and flat surface makes it difficult to provide strong interaction, especially in the small molecule fragment discovery stage, which is mainly probed by tracking the chemical shift differences using "SAR by NMR"[118] or soaking with a collection of organic solvents using "multiple solvent crystal structures"[119]. The coupled protein or matrix on the solid phase, *e.g.,* streptavidin on streptavidin beads, could also interact with small molecules and therefore generate background signals in the DEL selection. For the enzyme modulator selection, the impact of the solid phase is small. For instance, in the kinase inhibitors selection, selection results using Dynabeads TALON or IMAC resins are similar, albeit different amine preferences were observed in the different methods[18]. However,

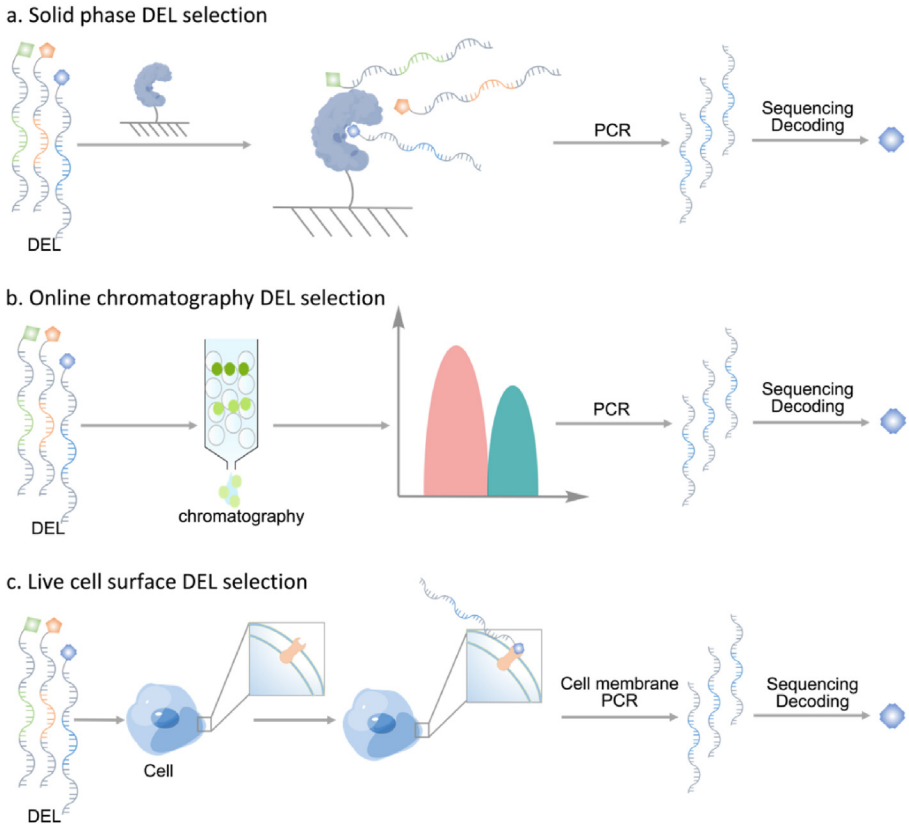


Figure 17 Affinity-based selections against targets on solid phase. (a) Solid phase DEL selection. (b) Online chromatography DEL selection. (c) Live cell surface DEL selection.

the impact could be more remarkable in the PPI modulator selection. In the insulin receptor modulator selection, the maximum enrichment was similar with or without immobilized insulin receptor using cobalt beads or streptavidin beads[108]. In this case, function-guided DEL selection could moderate the impact of the solid phase. In the function-guided selection, a functional binder, such as the intrinsic interaction protein or known small molecule, competes with the binding molecules and elutes them off the target protein. Following this concept, insulin elutes DEL molecules activating the insulin receptor[108] and SPD304 elutes DEL molecules inhibiting the TNF-α function[120]. Parallel DEL selection with InhA, InhA:NAD$^+$, and InhA:NADH discovered the compounds bind adjacent to the NADH cofactor and adopt a variety of conformations[121]. A non-ATP competitive inhibitor was discovered by comparing enrichments of DEL under different conditions[122]. For example, SB-735204 completely eliminated the NK3-specific DEL features in the selected DEL35 population (Fig. 17c)[123].

Another key parameter is the buffer composition, especially for the membrane protein. The correct membrane protein conformation requires the correct membrane environment. *In vitro* studies usually use detergents, amphipols, and nanodiscs to solubilize the membrane proteins. Different membrane protein prefers different solubilizing reagents. For instance, an allosteric "β-blocker" for β_2AR was identified using the detergent *n*-dodecyl-β-D-maltoside (DDM)[124]. Agonists and antagonists were discovered for the protease-activated receptor 2 (PAR2) solubilized in lauryl maltose neopentyl glycol and cholesteryl hemisuccinate[125]. In comparison with DDM, maltose neopentyl glycol (MNG), and amphipole A8-35, CXCR2 in lipid nanodisc was shown to be stable and suitable for DEL selection[126]. Most DEL elution using the batch mode, elaborate fractions will facilitate the selection and decoding procedures. Recently, online chromatography selection of the DEL library using immobilized Halo-tagged angiotensin II type I receptor (AT1R) resulted in seven phenolic acid derivatives (Fig. 17b)[127]. However, the solid supports-based methods could not avoid the artifacts induced by the protein immobilization.

3.2. *Affinity-based selection against targets in solution*

The ideal selection strategy is the *in situ* cellular selection. For the intracellular target, the main obstacle is the delivery of DEL into the cell. One option is cell-penetration peptide, which can bring the DEL library into the cells[128]. The DEL library is engineered to contain a reactive group on each component. After cytosolic delivery of a reactive DEL, the halo-tagged protein covalently binds to the ligand DNA by affinity labeling. Cells are then lysed, and all target proteins are captured on beads. The second option is direct injection of the DEL library into the cells. For example, an injection performed in the giant frog oocyte, 100,000 times bigger than normal somatic cell[129]. However, most cells are not big enough to accept the DEL injection. The frog genome is different from the human genome, thus this method is difficult to apply widely. Another solution is to use photochemistry annotating the target protein with DNA during ligand interaction (Fig. 18a). Photocrosslink group was attached to the *ss*DNA to generate the photoreactive *ss*DNA (PC-DNA), which has multiple applications in the DNA-programmed affinity labelling (DPAL)[37]. The key design in DPAL is two functional DNAs, *i.e.,* a binding probe (BP) tagging the ligand and a capture probe (CP) encompassing a

photo-cross-linker and an affinity tag. BP guides CP close to the target protein and cross-links the DNA with the protein. The protein−DNA complex can be isolated by capillary electrophoresis, gel electrophoresis, or affinity-based pull-down coupled with mass spectrometry[130]. After cross-link, the non-binding DNAs can be degraded by the exonuclease I (Exo I), and the binders can be decoded, however, over-digestion perturbs its further application[37,131]. Therefore, the method was further improved to use ligase, resulting in the "ligate-crosslink-purify" approach (Fig. 18b)[132]. Except for ligase, another enzyme, polymerase, was applied to replicate the DNA information[133]. In addition to the crosslink with the target protein, fusion enzymes with the target protein can also modify the adjacent DNA to distinguish them from the unbound DNAs[34]. For instance, the engineered small biotin ligase (UltraID)[134] label the NH_2-DNA with biotin and further pull-down by streptavidin beads (Fig. 18c)[135].

However, due to the low density of membrane protein and the complexity of the cell membrane microenvironment, the ligand discovery for membrane protein in cells is challenging. Overexpression is the most straightforward solution to the low density. In the NK3 ligand selection, the engineered cell reached 500,000 NK3 receptors per cell and generated the effective identification of novel binders[123]. High expression levels in cells are presumably required to simulate the high concentrations of target protein on the solid phase. The signal enhancement strategy, such as fluorescence-labeled oligonucleotide, facilitates this stringent screening. Furthermore, DPAL was improved for ligand discovery for endogenous membrane protein in living cell[136]. Firstly, DPAL delivers DNA tags to membrane proteins. A known ligand (small molecule or antibody) of the target is conjugated with a DNA strand as the binding probe (BP). The BP forms a duplex with a photoreactive capture probe (CP). The BP/CP duplex engages the target and irradiation triggers the target crosslinking. After displacement probe (DP) removes the BP, the DNA tag originally from the CP guided the subsequent DEL selection. In this strategy, crosslink solved two hurdles for the membrane protein selection, target specificity and target concentration. The crosslink methods have been successfully applied for the ligand identification of δ-opioid receptor (DOR). DOR fused with a SNAP-tag is overexpressed on HEK293T cells. The DOR captures the binders from the DEL library, afterwards, cross-linking fixed the interaction. After cell lysis, the DOR−binder complex was purified *via* the fusion SNAP protein and sent to decoding[128].

The microfluidic technique is a promising technique for ligand selection in living cells. It is not limited to one cell but also can extend to multiple cells, which could generate more phenotypes and be applied in the functional selection. Traditional DELs are difficult to use in phenotypic screens because they cannot be employed for biochemical assays, difficult to enter the cells, and difficult to segregate different molecules in the library[137]. One-bead one-compound (OBOC) DEL could be a solution to the above issues, although the library size is much smaller than the

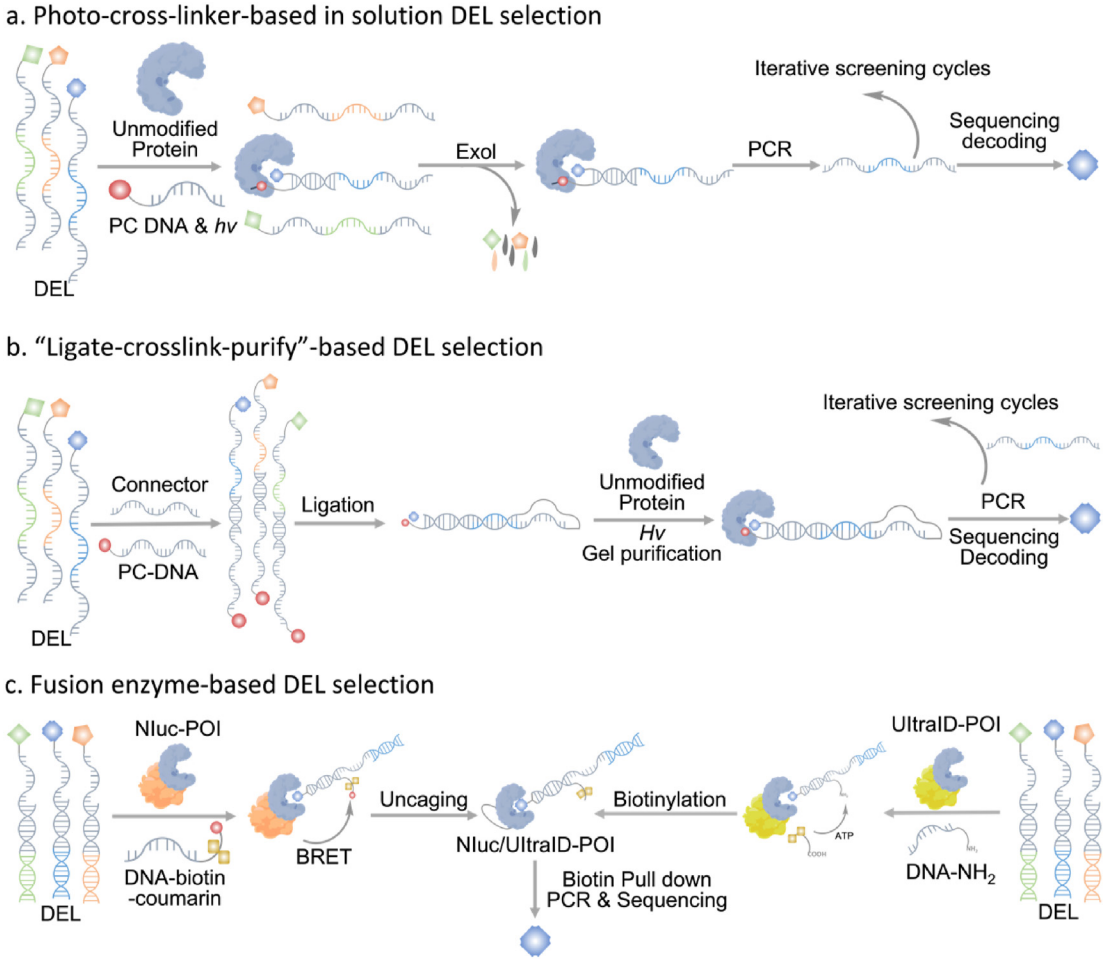


Figure 18 Affinity-based selections against targets in solution. (a) Photo-cross-linker-based in solution DEL selection. (b) "Ligate-crosslink-purify"-based DEL selection. (c) Fusion enzyme-based DEL selection.

traditional DEL. In the microfluidic system, targets, assay substrate, and DEL are injected from different channels and encapsulated into one droplet. The compound is photocleaved and released to the droplet, and the droplets are further sorted and sequenced. This platform has identified ligands for phosphodiesterase autotaxin and DDR1[138,139] and recently expanded to RNA binder[140,141]. Combined with the bacterial cytotoxicity screen, this strategy has identified two low-micromolar inhibitors of *Bacillus subtilis*[142].

3.3. Sequence integration-based DEL selections

In situ DEL selection against the native protein in cells is an ideal strategy for chemical evolution. In other words, the DEL selection is a procedure to map the chemical functions. The chemical information links to the target protein information in the chemical function map. Solid phase separation is a way to isolate the protein information, alternatively, if the target protein can add its information to the DNA sequence of the chemical, the function map also works. Interaction-dependent PCR (IDPCR) integrates codes identifying the ligand and target into one selectively amplifiable DNA molecule[143]. This concept was first demonstrated by tagging the single-strand DNA (*ss*DNA) to the target protein and chemicals *via* NHS chemistry. When the ligand binds to the target, the *ss*DNAs from the target will form a duplex with the one from the ligand *via* the complementary regions on each strand. The resulting hairpin could serve as a starting point for selective extension. The resulting amplified DNA contains both information of ligand and target (Fig. 19a). However, this is only done in purified protein. The second generation was improved to perform in the unpurified cell lysates. The target of interest was either fused with SNAP (Fig. 19b), which can covalently link the SNAP-tagged DNA or non-covalently recognized by DNA-tagged antibody. Remarkably, this method enriched corresponding sequences to five known ligand−target pairs, indicating its potential power in the library × library selection[144].

The cellular binder trap enrichment (cBTE) approach also employed fusion protein to capture the DNA, which is further integrated with the DEL DNA *via* ligation (Fig. 19a). The ligated DNA contains both information from the target and the ligand[129]. The mRNA encoding target-prey (carbonic anhydrase IX) fusion protein is injected into the oocyte for expression. DEL members bind to the target inside the cell. After cell lysis, the complex is isolated by emulsion. T4 DNA ligase links the target DNA and the DEL DNA, resulting in the information integration of the target and ligand. Ligand with single-digit nanomolar potencies for p38α is identified using this system.

3.4. Evolution of DEL selection data processing

Selection data analysis for hit picking is one of the key issues in DEL research, as the purpose of DEL selection is to find drug-like hit compounds against the target of interest. The major challenge in the data analysis of DEL selection is that the results can be plagued by false negatives and noisy data, especially for large DELs where the library quality is compromised by the truncated and/or side products during library synthesis[145].

Previously many statistical methods have been developed to address these issues. Neri et al.[146] first reported the use of the negative binomial distribution to analyze the sequencing data. Satz[147] developed a data aggregation approach to accurate analysis the selection data of DELs that contain significant quantities of truncates. Paegel group[148] disclosed the use of the Poisson distribution to calculate the false negative rate. Kuai and co-workers described the estimation of the Poisson confidence interval-based data normalization and enrichment calculation for the classical DEL selection[149]. In addition, Faver et al.[150] reported a normalized *z*-score enrichment metric for the quantitative comparison of compound enrichment from parallel DEL selections. Different from the previously reported negative binomial or Poisson distribution statistics which assumed an equal distribution of library members to model screening data, Gerry et al.[151]

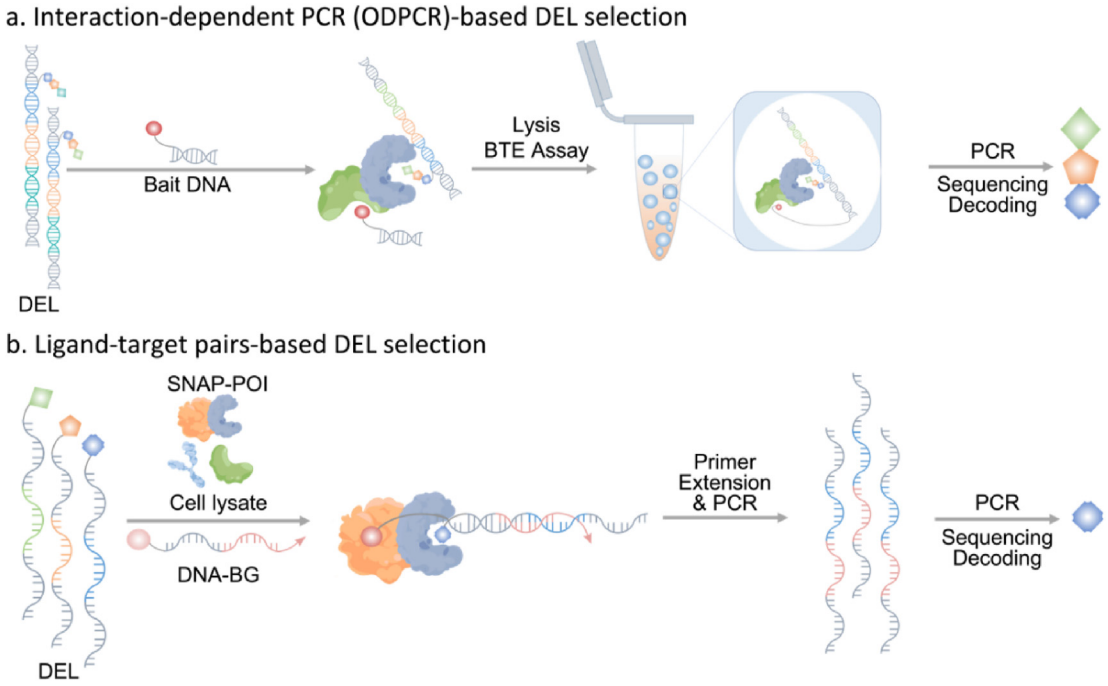


Figure 19 Sequence integration-based DEL selections. (a) Interaction-dependent PCR (ODPCR)-based DEL selection. (b) Ligand-target pairs-based DEL selection.

described a statistical framework that takes into account the nonuniform abundance of DEL members, and developed a normalized barcode fold-change score to rank binders by modelling the pre- and post-selection read counts *via* a superposition of multiple Poisson distributions.

Beyond the classical statistical methods, the emergence of "big data" and machine learning (ML) tools have been demonstrated to be promising approaches for DEL dataset processing. For instance, Komár et al.[152] developed a dubbed "deldenoiser" method based on sparse learning to recover and rank the true potential binders from the background noise. Besides, Lim et al.[153] reported a regression approach to learn enrichments of individual DEL molecules by using a custom negative-log-likelihood loss function that can efficiently denoise DEL data and visualize the learned structural activity relationships (SARs) trends. These ML methods have facilitated the analysis of DEL selection data with purified proteins, however, approaches for the processing of the much noisier cell-based selection data have not yet been reported. Recently, Li's group[154] described an ML-based method to process cell-based DEL selection datasets. They used the Maximum A Posteriori (MAP)-based enrichment metric to denoise the datasets, thereby facilitating to obtain high-confidence enrichment values. Additionally, the combination of deep learning and the MAP loss function has been demonstrated to be able to reduce false positive hits and identify true binders with reliable SAR from the noisy cell-based selection datasets.

DEL selection datasets offer unprecedentedly large real-world drug screening information, which constitutes a requisite for the implementation of ML-based hit or lead generation. McCloskey et al.[155] have applied the concept of ML to DEL selection datasets by screening high-affinity binders from commercially available compounds. They used DEL selection datasets against Erα, sEH, and c-KIT to train the computational model. After that, the model was used to virtually screen large commercially available chemical libraries. Numerous novel hits were identified, and interestingly, some of them displayed little structural resemblance to the DEL hits. Similarly, Torng et al.[156] used carbonic anhydrase IX (CAIX)-DEL screening datasets to train ML models generating novel hits with different structural features of the original DEL hits. Among the 152 novel potential hits that were generated by ML, 70% displayed submicromolar *in vitro* enzymatic inhibitory activities ($IC_{50} < 1$ μmol/L). The best lead candidate displayed low nanomolar affinity to CAIX ($K_D = 5.7$ nmol/L) and can selectively target CAIX high-expression tumors in mice. This result indicated the synergy between DEL and ML for the efficient generation of novel hits and candidates for *in vivo* tumor targeting.

In summary, although these ML techniques have just been reported very recently, their lasting impact on the DEL field is appealing. With the fast evaluation of the appropriate algorithms, DEL will play a key role in the success of AI-aided drug discovery.

4. Conclusion remarks and future perspectives

DNA-compatible synthesis and DEL selection technology are two cornerstones of DEL. The former is the key to expanding the DEL chemical space, while the latter is responsible for efficient hit selection from DEL. Consequently, the endeavor to develop novel DNA-compatible chemical reactions and selection technology has been conceived as particularly important to propel the technology development and will continue to be the frontiers of the DEL realm. This review summarized the latest developments in emerging DEL chemistry and selection technology that have been elegantly developed in the past few years.

For DEL chemistry, emerging protocols including photocatalyzed reactions, C−H activation & functionalization, click selenylation, micellar-mediated reactions, the RASS-based reactions, and substrate activation strategy have preliminarily demonstrated their feasibility in the on-DNA synthesis of skeletons rich in C(sp^3) centers, Se containing compounds, diaryl ethers, and some privileged heterocycles. Although these DNA-compatible synthetic methods are in their infancy, they have taken an important first step and opened some new fields for extending the diversity of the chemical space. As for selection technology, the advancement of evolution-based DEL selection strategies, including the affinity-based selections against targets on solid phase, and the emerging affinity-based selection against targets in solution and sequence-integration-based selections continuously improve selection methods and show an ever-expanding target profile.

Overall, we believe that DEL technology will serve as a multifunctional basic platform that eventually fertilizes the discovery of chemical biology tools in basic research and novel therapeutics for disease treatment in the future. To advance DEL technology in basic science and drug discovery, there remain several issues and new directions for future research.

i) The chemical reactions in the DEL technology are generally carried out on a very small scale. The final products in these small-scale reactions are typically confirmed by liquid chromatography−mass spectrometry (LC/MS) analysis. Usually, a DNA headpiece conjugated final product, which was synthesized by amide coupling was used as a reference to verify the new on-DNA bond formation reaction. This is logical in new DEL chemistry development and is suitable for the 1st round of combinatorial DEL synthesis. However, in the 2nd and 3rd rounds of DEL synthesis, due to the dramatic increase in the number of compounds, the yields of individual products cannot be determined. To this end, the development of modular and predictable DNA-compatible chemical reactions[88], high-throughput library quality analysis methods[157], and machine learning (ML) techniques to select better building blocks would be of great importance for practical DEL synthesis[158].

ii) DNA fidelity is an important indicator for evaluating the feasibility of an on-DNA reaction in DEL synthesis. Therefore, we propose that every newly developed on-DNA reaction should be tested for its DNA fidelity. In addition, the DNA recovery rate also needs to be tested for that it is related to the molecular abundance of DEL. Importantly, although the RAAS and APTAC approaches enable the on-DNA synthesis in neat organic solvents, the DNA recovery is only about ∼30% in some cases, indicating that they still need to be further optimized for actual DEL construction. Notably, the photoinduced radical coupling involves the combination of photo-catalysis with the RASS strategy, which shows no observable DNA damag[91], indicating this joint strategy may facilitate the development of more DNA-compatible reactions with high DNA fidelity.

iii) In recent years, the photo-promoted DNA compatibility reaction has developed rapidly, not only because it has achieved the introduction of many fragments rich in C(sp^3) center, but also because of its mild and fast reaction conditions leading to high DNA fidelity. Among them, most reactions require the

addition of photocatalysts, while strategies such as electron-donor-acceptor (EDA) that do not require photocatalysts may further reduce the impact of metal catalysts on DNA fidelity, which should be paid more attention in the future. For example, some photocatalytic off-DNA reactions activated *via* EDA complex that are not sensitive to water and air, are of great chance to be adopted into on DNA reactions[159,160]. In addition, developing complementary photo-promoted reactions that use easily available raw materials (*e.g.*, expanding from α-amino acids to ordinary acids) is of great significance.

iv) Although some on-DNA synthetic transformations have been established, they are far from universal. Generally, they may perform well on simple model substrates but often fail when applied to polar and complex substrates. Therefore, it is necessary to test the feasibility of some polar groups (*e.g.,* carboxyl, hydroxy, phenol, amide, amine, heterocycle) in new DNA-compatible synthesis development, which is important for generating compound sets in a more drug-likeness fashion[161]. Besides, considering that the on-DNA reaction profiles can vary wildly over a set of similar BBs, even the reactivity of the same BB is variable when conducted by different laboratories or practitioners. It is important to validate individual BBs and reoptimize the reaction conditions before a DEL synthesis.

v) Despite the recent innovations in DNA-compatible reactions that have greatly advanced DEL technology, the diversity of DELs is also limited by the availability of multifunctional BBs, especially the bifunctional and/or trifunctional core BBs. Therefore, the development of novel methods such as high through-put solid-phase synthesis[162] and on-microwell plate parallel synthesis that enable high through-put BBs synthesis will be appealing. In addition, as seen from the published paper, there are usually only limited hits that have been off-DNA synthesized and validated. This is not enough to find an idea hit, therefore the development of high through-put hit synthesis and validation methods, such as modular synthesis and artificial intelligence (AI), robotic platform automats synthesis will be appealing.

vi) Traditional DEL selections mainly focused on the affinity, rather than the activity/function in a biological system. Affinity-based strategies allow batch mode selection of huge numbers of molecules, but non-functional binder confirmation is time-consuming due to the lack of function indication. Further development of phenotypic selection using novel design of DNA encoding strategy, novel sequencing techniques, cross-linking, fusion protein, microfluidic system, and chromatography will speed up the functional ligand discovery. Similar to the development of biomolecule selection, the targets will not be limited to protein but expands to RNA, whole-cell, and mini-ecosystems.

Acknowledgments

Financial support was provided by the National Natural Science Foundation of China (grant numbers 22177073, 21977070, 21907085, and U19A2011), the Natural Science Foundation of Shanghai, China (grant numbers 21ZR1442900 and 23ZR1437600), the Natural Science Foundation of Zhejiang Province, China (grant number LY22H300001), Shanghai Frontiers Science Center of Degeneration and Regeneration in Skeletal System, and Shanghai Key Laboratory of Orthopedic Implants (grant number KFKT202207, China) for financial support.

Author contributions

Hongtao Xu, Wei Yi, Wei Hou, and Peixiang Ma conceived the project and jointly wrote the manuscript based on an intense literature survey conducted by Shuning Zhang, Qianping Huang, Yuang Gu, and Zhi Zhou. Hongtao Xu, Peixiang Ma, Shuning Zhang, and Qianping Huang made the figures. Wei Yi, Wei Hou, Zhi Zhou, and Yuang Gu reviewed and edited the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Schreiber SL, Kotz JD, Li M, Aubé J, Austin CP, Reed JC, et al. Advancing biological understanding and therapeutics discovery with small-molecule probes. *Cell* 2015;**161**:1252−65.
2. Folmer RHA. Integrating biophysics with HTS-driven drug discovery projects. *Drug Discov Today* 2016;**21**:491−8.
3. Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 1985;**228**:1315−7.
4. Mccafferty J, Griffiths AD, Winter G, Chiswellt DJ. Phage antibodies: filamentous variable domains. *Nature* 1990;**348**:552−4.
5. Huse WD, Sastry L, Iverson SA, Kang AS, Alting-Mees M, Burton DR, et al. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* 1989;**246**:1275−81.
6. Kang AS, Barbas CF, Janda KD, Benkovic SJ, Lerner RA. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc Natl Acad Sci U S A* 1991;**88**:4363−6.
7. Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD, Winter G. By-passing immunization: human antibodies from V-gene libraries displayed on phage. *J Mol Biol* 1991;**222**:581−97.
8. Winter G, Milstein C. Man-made antibodies. *Nature* 1991;**349**:293−9.
9. Wilchek M, Bayer EA, Livnah O. Essentials of biorecognition: the (strept)avidin−biotin system as a model for protein−protein and protein−ligand interaction. *Immunol Lett* 2006;**103**:27−32.
10. Brenner S, Lerner RA. Encoded combinatorial chemistry. *Proc Natl Acad Sci U S A* 1992;**89**:5381−3.
11. Gai SA, Wittrup KD. Yeast surface display for protein engineering and characterization. *Curr Opin Struct Biol* 2007;**17**:467−73.
12. Vinet L, Zhedanov A. *Ribosome display and related technologies*, vol. 805. New York, NY: Springer New York; 2012.
13. Needels MC, Jones DG, Tate EH, Heinkel GL, Kochersperger LM, Dower WJ, et al. Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc Natl Acad Sci U S A* 1993;**90**:10700−4.
14. Nielsen J, Brenner S, Janda KD. Synthetic methods for the implementation of encoded combinatorial chemistry. *J Am Chem Soc* 1993;**115**:9812−3.
15. Melkko S, Scheuermann J, Dumelin CE, Neri D. Encoded self-assembling chemical libraries. *Nat Biotechnol* 2004;**22**:568−74.
16. Gartner ZJ, Tse BN, Grubina R, Doyon JB, Snyder TM, Liu DR. DNA-templated organic synthesis and selection of a library of macrocycles. *Science* 2004;**305**:1601−5.
17. Halpin DR, Harbury PB. DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biol* 2004;**2**:1015−21.
18. Clark MA, Acharya RA, Arico-Muendel CC, Belyanskaya SL, Benjamin DR, Carlson NR, et al. Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat Chem Biol* 2009;**5**:647−54.

19. Cui M, Nguyen D, Gaillez MP, Heiden S, Lin W, Thompson M, et al. Trio-pharmacophore DNA-encoded chemical library for simultaneous selection of fragments and linkers. *Nat Commun* 2023;**14**:1481.
20. Salamon H, Klika Škopić M, Jung K, Bugain O, Brunschweiger A. Chemical biology probes from advanced DNA-encoded libraries. *ACS Chem Biol* 2016;**11**:296−307.
21. Harris PA, Berger SB, Jeong JU, Nagilla R, Bandyopadhyay D, Campobasso N, 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* 2017;**60**: 1247−61.
22. Ding Y, Belyanskaya S, DeLorey JL, Messer JA, Joseph Franklin G, Centrella PA, et al. Discovery of soluble epoxide hydrolase inhibitors through DNA-encoded library technology (ELT). *Bioorg Med Chem* 2021;**41**:116216.
23. Luther JM, Ray J, Wei D, Koethe JR, Hannah L, DeMatteo A, et al. GSK2256294 dcreases sEH (soluble epoxide hydrolase) activity in plasma, muscle, and adipose and reduces F2-isoprostanes but does not alter insulin sensitivity in humans. *Hypertension* 2021;**78**: 1092−102.
24. Cuozzo JW, Clark MA, Keefe AD, Kohlmann A, Mulvihill M, Ni H, et al. Novel autotaxin inhibitor for the treatment of idiopathic pulmonary fibrosis: a clinical candidate discovered using DNA-encoded chemistry. *J Med Chem* 2020;**63**:7840−56.
25. Puglioli S, Schmidt E, Pellegrino C, Prati L, Oehler S, De Luca R, et al. Selective tumor targeting enabled by picomolar fibroblast activation protein inhibitors isolated from a DNA-encoded affinity maturation library. *Chem* 2023;**9**:411−29.
26. Harris PA, Marinis JM, Lich JD, Berger SB, Chirala A, Cox JA, et al. Identification of a RIP1 kinase inhibitor clinical candidate (GSK3145095) for the treatment of pancreatic cancer. *ACS Med Chem Lett* 2019;**10**:857−62.
27. Hou N, Shuai L, Zhang L, Xie X, Tang K, Zhu Y, et al. Development of highly potent non-covalent inhibitors of SARS-CoV-2 3CLpro. *bioRxiv* 2022. Available from: https://doi.org/10.1101/2022.08.10.503531.
28. Lerner RA, Brenner S. DNA-encoded compound libraries as open source: a powerful pathway to new drugs. *Angew Chem Int Ed* 2017; **56**:1164−5.
29. Á Furka. Combinatorial technology revitalized by DNA-encoding. *MedComm* 2021;**2**:481−9.
30. Satz AL. What do you get from DNA-encoded libraries? *ACS Med Chem Lett* 2018;**9**:408−10.
31. Dixit A, Barhoosh H, Paegel BM. Translating the genome into drugs. *Acc Chem Res* 2023;**56**:489−99.
32. Neri D, Lerner RA. DNA-encoded chemical libraries: a selection system based on endowing organic compounds with amplifiable information. *Annu Rev Biochem* 2018;**87**:479−502.
33. Fitzgerald PR, Paegel BM. DNA-encoded chemistry: drug discovery from a few good reactions. *Chem Rev* 2021;**121**:7155−77.
34. Dockerill M, Winssinger N. DNA-encoded libraries: towards harnessing their full power with Darwinian evolution. *Angew Chem Int Ed* 2023;**62**:e202215542.
35. Huang Y, Li Y, Li X. Strategies for developing DNA-encoded libraries beyond binding assays. *Nat Chem* 2022;**14**:129−40.
36. Zhao G, Huang Y, Zhou Y, Li Y, Li X. Future challenges with DNA-encoded chemical libraries in the drug discovery domain. *Expet Opin Drug Discov* 2019;**14**:735−53.
37. Song Y, Li X. Evolution of the selection methods of DNA-encoded chemical libraries. *Acc Chem Res* 2021;**54**:3491−503.
38. Matsuo B, Granados A, Levitre G, Molander GA. Photochemical methods applied to DNA encoded library (DEL) synthesis. *Acc Chem Res* 2023;**56**:385−401.
39. Götte K, Chines S, Brunschweiger A. Reaction development for DNA-encoded library technology: from evolution to revolution? *Tetrahedron Lett* 2020;**61**:151889.
40. Flood DT, Asai S, Zhang X, Wang J, Yoon L, Adams ZC, et al. Expanding reactivity in DNA-encoded library synthesis *via* reversible binding of DNA to an inert quaternary ammonium support. *J Am Chem Soc* 2019;**141**:9998−10006.
41. Li Y, Gabriele E, Samain F, Favalli N, Sladojevich F, Scheuermann J, et al. Optimized reaction conditions for amide bond formation in DNA-encoded combinatorial libraries. *ACS Comb Sci* 2016;**18**: 438−43.
42. Li K, Qu Y, An Y, Breinlinger E, Webster MP, Wen H, et al. DNA-Compatible copper-catalyzed oxidative amidation of aldehydes with non-nucleophilic arylamines. *Bioconjugate Chem* 2020;**31**: 2092−7.
43. Ding Y, Clark MA. Robust suzuki−miyaura cross-coupling on DNA-linked substrates. *ACS Comb Sci* 2015;**17**:1−4.
44. Li JY, Huang H. Development of DNA-compatible Suzuki−Miyaura reaction in aqueous media. *Bioconjugate Chem* 2018;**29**:3841−6.
45. Xu H, Ma F, Wang N, Hou W, Xiong H, Lu F, et al. DNA-encoded libraries: aryl fluorosulfonates as versatile electrophiles enabling facile on-DNA Suzuki, Sonogashira, and Buchwald reactions. *Adv Sci* 2019;**6**:1901551.
46. Lu X, Roberts SE, Franklin GJ, Davie CP. On-DNA Pd and Cu promoted C−N cross-coupling reactions. *Medchemcomm* 2017;**8**: 1614−7.
47. de Pedro Beato E, Priego J, Gironda-Martínez A, González F, Benavides J, Blas J, et al. Mild and efficient palladium-mediated C−N cross-coupling reaction between DNA-conjugated aryl bromides and aromatic amines. *ACS Comb Sci* 2019;**21**:69−74.
48. Chen YC, Faver JC, Ku AF, Miklossy G, Riehle K, Bohren KM, et al. C−N coupling of DNA-conjugated (hetero)aryl bromides and chlorides for DNA-encoded chemical library synthesis. *Bioconjugate Chem* 2020;**31**:770−80.
49. Ruff Y, Berst F. Efficient copper-catalyzed amination of DNA-conjugated aryl iodides under mild aqueous conditions. *Medchemcomm* 2018;**9**:1188−93.
50. Ji Y, Dai D, Luo H, Shen S, Fan J, Wang Z, et al. C−S coupling of DNA-conjugated aryl iodides for DNA-encoded chemical library synthesis. *Bioconjugate Chem* 2021;**32**:685−9.
51. Zhang Y, Chen W, Tan T, Gu Y, Zhang S, Li J, et al. Palladium-catalyzed one-pot phosphorylation of phenols mediated by sulfuryl fluoride. *Chem Commun* 2021;**57**:4588−91.
52. Ma F, Li J, Zhang S, Gu Y, Tan T, Chen W, et al. Metal-catalyzed one-pot on-DNA syntheses of diarylmethane and thioether derivatives. *ACS Catal* 2022;**12**:1639−49.
53. Qu Y, Wen H, Ge R, Xu Y, Gao H, Shi X, et al. Copper-mediated DNA-compatible one-pot click reactions of alkynes with aryl borates and TMS-N_3. *Org Lett* 2020;**22**:4146−50.
54. Gironda-Martínez A, Neri D, Samain F, Donckele EJ. DNA-compatible diazo-transfer reaction in aqueous media duitable for DNA-encoded chemical library synthesis. *Org Lett* 2019;**21**: 9555−8.
55. Ede S, Schenk M, Bierer D, Weinmann H, Graham K. Improved diazo-transfer reaction for DNA-encoded chemistry and its potential application for macrocyclic del-libraries. *Molecules* 2021;**26**:1790.
56. Pan K, Yao Y, Zhang Y, Gu Y, Wang Y, Ma P, et al. Enolate-azide [3 + 2]-cycloaddition reaction suitable for DNA-encoded library synthesis. *Bioconjugate Chem* 2023;**34**:1459−66.
57. Su L, Feng J, Peng T, Wan J, Fan J, Li J, et al. Synthesis of multifunctional 2-aminobenzimidazoles on DNA *via* iodine-promoted cyclization. *Org Lett* 2020;**22**:1290−4.
58. Zhang J, Li X, Wei H, Li Y, Zhang G, Li Y. Sequential DNA-encoded building block fusion for the construction of polysubstituted pyrazoline core libraries. *Org Lett* 2021;**23**:8429−33.
59. Xiong H, Gu Y, Zhang S, Lu F, Ji Q, Liu L, et al. Iridium-catalyzed C−H amidation of *s*-tetrazines. *Chem Commun* 2020;**56**:4692−5.
60. Yuen J, Chai J, Ding Y. Condensation of DNA-conjugated imines with homophthalic anhydride for the synthesis of isoquinolones on DNA. *Bioconjugate Chem* 2020;**31**:2712−8.
61. Ma F, Li J, Zhang S, Gu Y, Tan T, Chen W, et al. DNA-encoded libraries: hydrazide as a pluripotent precursor for on-DNA synthesis of various azole derivatives. *Chem Eur J* 2021;**27**:8214−20.

62. Matsuo B, Granados A, Levitre G, Molander GA. Photochemical methods applied to DNA encoded library (DEL) synthesis. *Acc Chem Res* 2023;**56**:385−401.
63. Patel S, Badir SO, Molander GA. Developments in photoredox-mediated alkylation for DNA-encoded libraries. *Trends Chem* 2021;**3**:161−75.
64. Chen Y, Kamlet AS, Steinman JB, Liu DR. A biomolecule-compatible visible-light-induced azide reduction from a DNA-encoded reaction-discovery system. *Nat Chem* 2011;**3**:146−53.
65. Kölmel DK, Loach RP, Knauber T, Flanagan ME. Employing photoredox catalysis for DNA-encoded chemistry: decarboxylative alkylation of α-amino acids. *ChemMedChem* 2018;**13**:2159−65.
66. Wu R, Gao S, Du T, Cai K, Cheng X, Fan J, et al. Exploring aldol reactions on DNA and applications to produce diverse structures: an example of expanding chemical space of DNA-encoded compounds by diversity-oriented synthesis. *Chem Asian J* 2020;**15**:4033−7.
67. Yen-Pon E, Li L, Levitre G, Majhi J, McClain EJ, Voight EA, et al. On-DNA hydroalkylation to introduce diverse bicyclo[1.1.1]pentanes and abundant alkyls *via* halogen atom transfer. *J Am Chem Soc* 2022;**144**:12184−91.
68. Fu X, Tang J, Hua R, Li X, Kang Z, Qiu H, et al. Functionalization of DNA-tagged alkenes with diazo compounds *via* photocatalysis. *Org Lett* 2022;**24**:2208−13.
69. Shan J, Ling X, Liu J, Wang X, Lu X. DNA-encoded C−H functionality *via* photoredox-mediated hydrogen atom transformation catalysis. *Bioorg Med Chem* 2021;**42**:116234.
70. Wu R, Du T, Sun W, Shaginian A, Gao S, Li J, et al. Functionalization of DNA-tagged alkenes enabled by visible-light-induced C−H activation of *N*-aryl tertiary amines. *Org Lett* 2021;**23**:3486−90.
71. Shen Y, Yang G, Huang W, Shaginian A, Lin Q, Wan J, et al. Photoredox deaminative alkylation in DNA-encoded library synthesis. *Org Lett* 2022;**24**:2650−4.
72. Chheda PR, Simmons N, Schuman DP, Shi Z. Photoredox-mediated deoxygenative alkylation of DNA-tagged alkenes with activated alcohols. *Org Lett* 2022;**24**:9514−9.
73. Phelan JP, Lang SB, Sim J, Berritt S, Peat AJ, Billings K, et al. Open-air alkylation reactions in photoredox-catalyzed DNA-encoded library synthesis. *J Am Chem Soc* 2019;**141**:3723−32.
74. Badir SO, Lipp A, Krumb M, Cabrera-Afonso MJ, Kammer LM, Wu VE, et al. Photoredox-mediated hydroalkylation and hydroarylation of functionalized olefins for DNA-encoded library synthesis. *Chem Sci* 2021;**12**:12036−45.
75. Kölmel DK, Ratnayake AS, Flanagan ME. Photoredox cross-electrophile coupling in DNA-encoded chemistry. *Biochem Biophys Res Commun* 2020;**533**:201−8.
76. Krumb M, Kammer LM, Badir SO, Cabrera-Afonso MJ, Wu VE, Huang M, et al. Photochemical C−H arylation of heteroarenes for DNA-encoded library synthesis. *Chem Sci* 2022;**13**:1023−9.
77. Kölmel DK, Ratnayake AS, Flanagan ME, Tsai MH, Duan C, Song C. Photocatalytic [2 + 2] cycloaddition in DNA-encoded chemistry. *Org Lett* 2020;**22**:2908−13.
78. Wang X, Sun H, Liu J, Dai D, Zhang M, Zhou H, et al. Ruthenium-promoted C−H activation reactions between DNA-conjugated acrylamide and aromatic acids. *Org Lett* 2018;**20**:4764−8.
79. Xu H, Gu Y, Zhang S, Xiong H, Ma F, Lu F, et al. A chemistry for incorporation of selenium into DNA-encoded libraries. *Angew Chem Int Ed* 2020;**59**:13273−80.
80. Gao H, Lin S, Zhang S, Chen W, Liu X, Yang G, et al. Gem-difluoromethylene alkyne-enabled diverse C−H functionalization and application to the on-DNA synthesis of difluorinated isocoumarins. *Angew Chem Int Ed* 2021;**60**:1959−66.
81. Xu H, Chen W, Bian M, Xu H, Gao H, Wang T, et al. Gem-difluorocyclopropenes as versatile β-monofluorinated three-*sp*2 carbon sources for Cp*Rh(III)-catalyzed [4 + 3] annulation: experimental development and mechanistic insight. *ACS Catal* 2021;**11**:14694−701.
82. Fan Z, Zhao S, Liu T, Shen PX, Cui ZN, Zhuang Z, et al. Merging C(sp^3)−H activation with DNA-encoding. *Chem Sci* 2020;**11**:12282−8.
83. Wen X, Duan Z, Liu J, Lu W, Lu X. On-DNA cross-dehydrogenative coupling reaction toward the synthesis of focused DNA-encoded tetrahydroisoquinoline libraries. *Org Lett* 2020;**22**:5721−5.
84. Kolb HC, Finn MG, Sharpless KB. Click chemistry: diverse chemical function from a few good reactions. *Angew Chem Int Ed* 2001;**40**:2004−21.
85. Dong J, Krasnova L, Finn MG, Sharpless KB. Sulfur(VI) fluoride exchange (SuFEx): another good reaction for click chemistry. *Angew Chem Int Ed* 2014;**53**:9430−48.
86. Hou W, Xu H. Incorporating selenium into heterocycles and natural products−from chemical properties to pharmacological activities. *J Med Chem* 2022;**65**:4436−56.
87. Hou W, Dong H, Zhang X, Wang Y, Su L, Xu H. Selenium as an emerging versatile player in heterocycles and natural products modification. *Drug Discov Today* 2022;**27**:2268−77.
88. Hou W, Dong H, Yao Y, Pan K, Yang G, Ma P, et al. Clickable selenylation−a paradigm for seleno-medicinal chemistry. *ChemMedChem* 2022;**17**:e202200324.
89. Xu H, Wang Y, Dong H, Zhang Y, Gu Y, Zhang S, et al. Selenylation chemistry suitable for on-plate parallel and on-DNA library synthesis enabling high-throughput medicinal chemistry. *Angew Chem Int Ed* 2022;**61**:e202206516.
90. Zhou Z, Gu Y, Wu L, Wang Y, Xu H, Ma L, et al. Organo-Se BTSAs-enabled performance: from racemic and asymmetric synthesis to click chemistry application. *Chem* 2023;**9**:3335−46.
91. Lin B, Lu W, Chen ZY, Zhang Y, Duan YZ, Lu X, et al. Enhancing the potential of miniature-scale DNA-compatible radical reactions *via* an electron donor−acceptor complex and a reversible adsorption to solid support strategy. *Org Lett* 2021;**23**:7381−5.
92. Yang S, Zhao G, Gao Y, Sun Y, Zhang G, Fan X, et al. In-solution direct oxidative coupling for the integration of sulfur/selenium into DNA-encoded chemical libraries. *Chem Sci* 2022;**13**:2604−13.
93. Škopić MK, Götte K, Gramse C, Dieter M, Pospich S, Raunser S, et al. Micellar Brønsted acid mediated synthesis of DNA-tagged heterocycles. *J Am Chem Soc* 2019;**141**:10546−55.
94. Graham JS, Hunter JH, Waring MJ. Micellar Buchwald-Hartwig coupling of aryl and heteroarylamines for the synthesis of DNA-encoded libraries. *J Org Chem* 2021;**86**:17257−64.
95. Škopić MK, Salamon H, Bugain O, Jung K, Gohla A, Doetsch LJ, 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* 2017;**8**:3356−61.
96. MacConnell AB, McEnaney PJ, Cavett VJ, Paegel BM. DNA-encoded solid-phase synthesis: encoding language design and complex oligomer library synthesis. *ACS Comb Sci* 2015;**17**:518−34.
97. Pels K, Dickson P, An H, Kodadek T. DNA-compatible solid-phase combinatorial synthesis of β-cyanoacrylamides and related electrophiles. *ACS Comb Sci* 2018;**20**:61−9.
98. Flood DT, Zhang X, Fu X, Zhao Z, Asai S, Sanchez BB, et al. RASS-enabled S/P−C and S−N bond formation for DEL synthesis. *Angew Chem Int Ed* 2020;**59**:7377−83.
99. Ruff Y, Martinez R, Pellé X, Nimsgern P, Fille P, Ratnikov M, et al. An amphiphilic polymer-supported strategy enables chemical transformations under anhydrous conditions for DNA-encoded library synthesis. *ACS Comb Sci* 2020;**22**:120−8.
100. Westphal MV, Hudson L, Mason JW, Pradeilles JA, Zécri FJ, Briner K, et al. Water-compatible cycloadditions of oligonucleotide-conjugated strained allenes for DNA-encoded library synthesis. *J Am Chem Soc* 2020;**142**:7776−82.
101. Xu H, Tan T, Zhang Y, Wang Y, Pan K, Yao Y, et al. Metal-free and open-air arylation reactions of diaryliodonium salts for DNA-encoded library Synthesis. *Adv Sci* 2022;**9**:2202790.

102. DeBenedictis EA, Chory EJ, Gretton DW, Wang B, Golas S, Esvelt KM. Systematic molecular evolution enables robust biomolecule discovery. *Nat Methods* 2022;**19**:55−64.
103. Berman CM, Papa LJ, Hendel SJ, Moore CL, Suen PH, Weickhardt AF, et al. An adaptable platform for directed evolution in human cells. *J Am Chem Soc* 2018;**140**:18093−103.
104. English JG, Olsen RHJ, Lansu K, Patel M, White K, Cockrell AS, et al. VEGAS as a platform for facile directed evolution in mammalian cells. *Cell* 2019;**178**:748−61.
105. Crook N, Abatemarco J, Sun J, Wagner JM, Schmitz A, Alper HS. *In vivo* continuous evolution of genes and pathways in yeast. *Nat Commun* 2016;**7**:13051.
106. Ravikumar A, Arzumanyan GA, Obadi MKA, Javanpour AA, Liu CC. Scalable, continuous evolution of genes at mutation rates above genomic error thresholds. *Cell* 2018;**175**:1946−57.
107. Li J, Li Y, Lu F, Liu L, Ji Q, Song K, et al. A DNA-encoded library for the identification of natural product binders that modulate poly (ADP-ribose) polymerase 1, a validated anti-cancer target. *Biochem Biophys Res Commun* 2020;**533**:241−8.
108. Xie J, Wang S, Ma P, Ma F, Li J, Wang W, et al. Selection of small molecules that bind to and activate the insulin receptor from a DNA-encoded library of natural products. *iScience* 2020;**23**:101197.
109. Dumelin CE, Trüssel S, Buller F, Trachsel E, Bootz F, Zhang Y, et al. A portable albumin binder from a DNA-encoded chemical library. *Angew Chem Int Ed* 2008;**47**:3196−201.
110. Mannocci L, Melkko S, Buller F, Molnàr I, Gapian Bianké JP, Dumelin CE, et al. Isolation of potent and specific trypsin inhibitors from a DNA-encoded chemical library. *Bioconjugate Chem* 2010;**21**:1836−41.
111. Decurtins W, Wichert M, Franzini RM, Buller F, Stravs MA, Zhang Y, et al. Automated screening for small organic ligands using DNA-encoded chemical libraries. *Nat Protoc* 2016;**11**:764−80.
112. Azzarito V, Long K, Murphy NS, Wilson AJ. Inhibition of α-helix-mediated protein−protein interactions using designed molecules. *Nat Chem* 2013;**5**:161−73.
113. Xiang H, Zhou M, Li Y, Zhou L, Wang R. Drug discovery by targeting the protein−protein interactions involved in autophagy. *Acta Pharm Sin B* 2023;**13**:4373−90.
114. Wu D, Li Y, Zheng L, Xiao H, Ouyang L, Wang G, et al. Small molecules targeting protein−protein interactions for cancer therapy. *Acta Pharm Sin B* 2023;**13**:4060−88.
115. Arkin MR, Tang Y, Wells JA. Small-molecule inhibitors of protein−protein interactions: progressing toward the reality. *Chem Biol* 2014;**21**:1102−14.
116. Johnson DK, Karanicolas J. Druggable protein interaction sites are more predisposed to surface pocket formation than the rest of the protein surface. *PLoS Comput Biol* 2013;**9**:e1002951.
117. Ran X, Gestwicki JE. Inhibitors of protein−protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Curr Opin Chem Biol* 2018;**44**:75−86.
118. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 1996;**274**:1531−4.
119. Mattos C, Bellamacina CR, Peisach E, Pereira A, Vitkup D, Petsko GA, et al. Multiple solvent crystal structures: probing binding sites, plasticity and hydration. *J Mol Biol* 2006;**357**:1471−82.
120. Wang S, Shi X, Li J, Huang Q, Ji Q, Yao Y, et al. A small molecule selected from a DNA-encoded library of natural products that binds to TNF-α and attenuates inflammation i*n vivo*. *Adv Sci* 2022;**9**:2201258.
121. Soutter HH, Centrella P, Clark MA, Cuozzo JW, Dumelin CE, Guie MA, et al. Discovery of cofactor-specific, bactericidal Mycobacterium tuberculosis InhA inhibitors using DNA-encoded library technology. *Proc Natl Acad Sci U S A* 2016;**113**:E7880−9.
122. Cuozzo JW, Centrella PA, Gikunju D, Habeshian S, Hupp CD, Keefe AD, et al. Discovery of a potent BTK inhibitor with a novel binding mode by using parallel selections with a DNA-encoded chemical library. *Chembiochem* 2017;**18**:864−71.
123. Wu Z, Graybill TL, Zeng X, Platchek M, Zhang J, Bodmer VQ, 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* 2015;**17**:722−31.
124. Ahn S, Kahsai AW, Pani B, Wang QT, Zhao S, Wall AL, et al. Allosteric "beta-blocker" isolated from a DNA-encoded small molecule library. *Proc Natl Acad Sci U S A* 2017;**114**:1708−13.
125. Brown DG, Brown GA, Centrella P, Certel K, Cooke RM, Cuozzo JW, 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* 2018;**23**:429−36.
126. Ma P, Xu H, Li J, Lu F, Ma F, Wang S, et al. Functionality-independent DNA encoding of complex natural products. *Angew Chem Int Ed* 2019;**58**:9254−61.
127. Liang Q, He J, Zhao X, Xue Y, Zuo H, Xu R, et al. Selective discovery of GPCR ligands within DNA-encoded chemical libraries derived from natural products: a case study on antagonists of angiotensin II type I receptor. *J Med Chem* 2021;**64**:4196−205.
128. Cai B, Kim D, Akhand S, Sun Y, Cassell RJ, Alpsoy A, et al. Selection of DNA-encoded libraries to protein targets within and on living cells. *J Am Chem Soc* 2019;**141**:17057−61.
129. Petersen LK, Christensen AB, Andersen J, Folkesson CG, Kristensen O, Andersen C, et al. Screening of DNA-encoded small molecule libraries inside a living Cell. *J Am Chem Soc* 2021;**143**:2751−6.
130. Le ATH, Krylova SM, Beloborodov SS, Wang TY, Hili R, Johnson PE, et al. How to develop and prove high-efficiency selection of ligands from oligonucleotide libraries: a universal framework for aptamers and DNA-encoded small-molecule ligands. *Anal Chem* 2021;**93**:5343−54.
131. Zhao P, Chen Z, Li Y, Sun D, Gao Y, Huang Y, et al. Selection of DNA-encoded small molecule libraries against unmodified and non-immobilized protein targets. *Angew Chem Int Ed* 2014;**53**:10056−9.
132. 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. *Bioconjugate Chem* 2017;**28**:2293−301.
133. Shi B, Deng Y, Li X. Polymerase-extension-based selection method for DNA-encoded chemical libraries against nonimmobilized protein targets. *ACS Comb Sci* 2019;**21**:345−9.
134. Kubitz L, Bitsch S, Zhao X, Schmitt K, Deweid L, Roehrig A, et al. Engineering of ultraID, a compact and hyperactive enzyme for proximity-dependent biotinylation in living cells. *Commun Biol* 2022;**5**:657.
135. Cai B, Mhetre AB, Krusemark CJ. Selection methods for proximity-dependent enrichment of ligands from DNA-encoded libraries using enzymatic fusion proteins. *Chem Sci* 2023;**14**:245−50.
136. Huang Y, Meng L, Nie Q, Zhou Y, Chen L, Yang S, et al. Selection of DNA-encoded chemical libraries against endogenous membrane proteins on live cells. *Nat Chem* 2021;**13**:77−88.
137. Kodadek T, Paciaroni NG, Balzarini M, Dickson P. Beyond protein binding: recent advances in screening DNA-encoded libraries. *Chem Commun* 2019;**55**:13330−41.
138. Cochrane WG, Malone ML, Dang VQ, Cavett V, Satz AL, Paegel BM. Activity-based DNA-encoded library screening. *ACS Comb Sci* 2019;**21**:425−35.
139. Hackler AL, FitzGerald FG, Dang VQ, Satz AL, Paegel BM. Off-DNA DNA-encoded library affinity screening. *ACS Comb Sci* 2020;**22**:25−34.
140. Gibaut QMR, Akahori Y, Bush JA, Taghavi A, Tanaka T, Aikawa H, et al. Study of an RNA-focused DNA-encoded library informs design of a degrader of a r(CUG) repeat expansion. *J Am Chem Soc* 2022;**144**:21972−9.
141. Benhamou RI, Suresh BM, Tong Y, Cochrane WG, Cavett V, Vezina-Dawod S, et al. DNA-encoded library *versus* RNA-encoded library

selection enables design of an oncogenic noncoding RNA inhibitor. *Proc Natl Acad Sci U S A* 2022;**119**:e2114971119.
142. Cochrane WG, Fitzgerald PR, Paegel BM. Antibacterial discovery *via* phenotypic DNA-encoded library screening. *ACS Chem Biol* 2021;**16**:2752−6.
143. 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**:15522−4.
144. 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**:3264−70.
145. Satz AL, Hochstrasser R, Petersen AC. Analysis of current DNA encoded library screening data indicates higher false negative rates for numerically larger libraries. *ACS Comb Sci* 2017;**19**:234−8.
146. 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**:1075−86.
147. Satz AL. Simulated screens of DNA encoded libraries: the potential influence of chemical synthesis fidelity on interpretation of structure−activity relationships. *ACS Comb Sci* 2016;**18**:415−24.
148. MacConnell AB, Paegel BM. Poisson statistics of combinatorial library sampling predict false discovery rates of screening. *ACS Comb Sci* 2017;**19**:524−32.
149. 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* 2018;**23**:405−16.
150. Faver JC, Riehle K, Lancia DR, Jbj Milbank, Kollmann CS, Simmons N, et al. Quantitative comparison of enrichment from DNA-encoded chemical library selections. *ACS Comb Sci* 2019;**21**:75−82.
151. Gerry CJ, Wawer MJ, Clemons PA, Schreiber SL. DNA barcoding a complete matrix of stereoisomeric small molecules. *J Am Chem Soc* 2019;**141**:10225−35.
152. Kómár P, Kalinić M. Denoising DNA encoded library screens with sparse learning. *ACS Comb Sci* 2020;**22**:410−21.
153. Lim KS, Reidenbach AG, Hua BK, Mason JW, Gerry CJ, Clemons PA, et al. Machine learning on DNA-encoded library count data using an uncertainty-aware probabilistic loss function. *J Chem Inf Model* 2022;**62**:2316−31.
154. Hou R, Xie C, Gui Y, Li G, Li X. Machine-learning-based data analysis method for cell-based selection of DNA-encoded libraries. *ACS Omega* 2023;**8**:19057−71.
155. McCloskey K, Sigel EA, Kearnes S, Xue L, Tian X, Moccia D, et al. Machine learning on DNA-encoded libraries: a new paradigm for hit finding. *J Med Chem* 2020;**63**:8857−66.
156. Torng W, Biancofiore I, Oehler S, Xu J, Xu J, Watson I, et al. A deep learning approach for the discovery of tumor-targeting small organic ligands from DNA-Encoded Chemical Libraries. *ACS Omega* 2023;**8**:25090−100.
157. Roy A, Koesema E, Kodadek T. High-throughput quality control assay for the solid-phase synthesis of DNA-encoded libraries of macrocycles. *Angew Chem Int Ed* 2021;**60**:11983−90.
158. Wang DY, Wen X, Xiong CD, Zhao JN, Ding CY, Meng Q, et al. Non-transition metal-mediated diverse aryl-heteroatom bond formation of arylammonium salts. *iScience* 2019;**15**:307−15.
159. Cabrera-Afonso MJ, Granados A, Molander GA. Sustainable thioetherification *via* electron donor−acceptor photoactivation using thianthrenium salts. *Angew Chem Int Ed* 2022;**61**:e202202706.
160. Granados A, Cabrera-Afonso MJ, Escolano M, Badir SO, Molander GA. Thianthrenium-enabled sulfonylation *via* electron donor−acceptor complex photoactivation. *Chem Catal* 2022;**2**:898−907.
161. Santanilla AB, Regalado EL, Pereira T, Shevlin M, Bateman K, Campeau L, et al. Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science* 2015;**347**:443−8.
162. Zambaldo C, Geigle SN, Satz AL. High-throughput solid-phase building block synthesis for DNA-encoded libraries. *Org Lett* 2019;**21**:9353−7.