Diversified strategy for the synthesis of DNA-encoded oxindole libraries†

Xuan Wang a,b Jiaxiang Liu a Ziqin Yan a Xiaohong Liu a,c,d,e Sixiu Liu a Yanrui Suo a Weiwei Lu a Jinfeng Yue a Kaixian Chen a,c,d,e Hualiang Jiang a,c,d,e Yujun Zhao b,a Mingyue Zheng, b,a Dongcheng Dai b,a and Xiaojie Lu a,b

DNA-encoded library technology (DELT) employs DNA as a barcode to track the sequence of chemical reactions and enables the design and synthesis of libraries with billions of small molecules through combinatorial expansion. This powerful technology platform has been successfully demonstrated for hit identification and target validation for many types of diseases. As a highly integrated technology platform, DEL is capable of accelerating the translation of synthetic chemistry by using on-DNA compatible reactions or off-DNA scaffold synthesis. Herein, we report the development of a series of novel on-DNA transformations based on oxindole scaffolds for the design and synthesis of diversity-oriented DNA-encoded libraries for screening. Specifically, we have developed 1,3-dipolar cyclizations, cyclopropanations, ring-opening of reactions of aziridines and Claisen–Schmidt condensations to construct diverse oxindole derivatives. The majority of these transformations enable a diversity-oriented synthesis of DNA-encoded oxindole libraries which have been used in the successful hit identification for three protein targets. We have demonstrated that a diversified strategy for DEL synthesis could accelerate the application of synthetic chemistry for drug discovery.

Introduction

The advances in molecular biology, in particular genomic sciences, have had a deep impact on the identification of potential disease-related protein targets in drug discovery.¹ Medicinal chemistry programs mostly rely on screening compound collections populated by a range of molecules derived from a set of known and robust chemical reactions. Based on the comprehensive analysis conducted by Dean G. Brown, high-throughput screening dominates early drug discovery, in conjunction with fragment-based screening and knowledge directed screens.² These lead generation strategies have successfully identified potential hit compounds, however the limited chemical space³ and high operation cost⁴ restricts the hit discovery for multiple protein targets and limits the scope of application in small biotech startups and university laboratories due to prohibitive costs. It was through continued innovation by chemists that led to the introduction of DNA-encoded library technology pioneered in 1992 by Brenner and Lerner⁵ and has become a vital tool in the current investigational repertoire.

As an interdisciplinary technology, the integration of combinatorial chemistry, molecular biology, next generation DNA-sequencing and informatic analysis shaped the scene of the DNA-encoded library technology platform and brought it to a new level with larger compound collections, reduced operation costs and screening time. Barcoded small molecules translated from phage display⁶ with a small molecule warhead (phenotype) individually coupled to encoding DNA tags (genotype) serve to retrace the synthetic history. The encoded compound libraries were constructed using a combinatorial strategy followed by successive iterative steps of both chemical transformation and oligonucleotide ligation.⁷ DNA barcodes have facilitated affinity selection⁸ to interrogate vast numbers of compounds in a single pool, exceeding by several orders of magnitude the capacity of the traditional “one-compound per well” screening approach. The approach has dramatically decreased the cost and cycle times, and will become a general technology broadly applied to industry and academia alike.⁹

Nevertheless, every new technology has strengths and weaknesses and DELT is no exception. The encoded compound collections are predominantly composed of limited scaffolds,
focused on peptides and aromatic rings. One reason for this limitation is that the scope of chemical reactions applicable in solution-phase DEL libraries are constrained to robust reactions that preserve the integrity of the oligonucleotides; few new synthetic methodologies can be applied to DEL library construction. To preserve the integrity of the DNA barcodes, chemists have continued to expand the scope of mild chemical transformations and DNA-protecting strategies (solid-phase DEL, reversible absorption of DNA onto resin, micellar catalyst-promoted reactions), which provides new opportunities for DEL library synthesis.

In addition to the directed DNA protection strategy, diversified library construction could be enabled by the identification of mild chemical transformations. As illustrated in Fig. 1, a multitude of innovative chemistries were developed in academia to synthesize drug-like scaffolds, which could be introduced to the encoded library to expand chemical space through innovative designs. This route enables an enhancement of the crosstalk between synthetic and medicinal chemists in industry and academia and an opportunity to impact the future of drug discovery.

In DELT library construction, an ideal scaffold provides mature synthetic methods, enabling synthesis both on-DNA or off-DNA and possessing multiple functional groups and appendage points for DNA-linking along with an abundance of readily available building blocks. Based on an extensive literature analysis, the oxindole core was selected to validate our initial proposal. Firstly, the C3 position of the oxindole could be converted into a number of diverse structures through well-established transformations; secondly, at least two functional groups are appended on these scaffolds; thirdly, oxindole-type compounds are prevalent in nature and have the potential to be transformed into drugs.

As illustrated in Fig. 2, four types of DELs were constructed from oxindoles to validate the concept: DEL-A, the spiropyrrrolizidine centered DEL was built on-DNA using a three-component cyclization reaction from available isatins. DEL-B, the spirocyclopropyl centered DEL was synthesized by a palladium-promoted on-DNA cycloaddition reaction from diazoisatin derivatives. DEL-C, spiroaziridine derivatives were converted from I using an aza-Corey–Chaykovsky reaction, which then underwent a ring-opening reaction. DEL-D, unsaturated centered DEL was synthesized from oxindole by an on-DNA aldol reaction. These DELs arising from a simple oxindole core, spanned spiro, disubstituted and unsaturated structures, elegantly fit the diversity-oriented synthesis (DOS) strategy, which was recently demonstrated by Schreiber, et al.

**Results and discussion**

Our efforts commenced with the investigation of on-DNA three-component reaction that can form highly functionalized spiropyrrrolizidine/pyrrolidine oxindoles. DNA barcodes could be

---

**Fig. 1** The workflow of oxindole-scaffold DNA-encoded library.

**Fig. 2** Design and synthesis of diversified oxindole DNA-encoded libraries. (a) Isatin derivatives were commercially available or synthesized via Sandmeyer isatin synthesis. (b) On-DNA three-component cycloaddition reaction. (c) 3-Diazoisatin derivatives were transformed by 3-N-tosylhydrazone isatin derivatives under basic conditions. (d) PdCl2-promoted on-DNA cyclopropanation reaction. (e) Spiroaziridine isatin derivatives were transformed by aza-Corey–Chaykovsky reaction. (f) On-DNA amidation reaction conditions. (g) On-DNA ring-opening reaction. (h) Oxindole derivatives were commercially available. (i) On-DNA Claisen–Schmidt condensation.
conjugated to the scaffold via three attachment sites: pyrroli-
zidine ring (1), nitrogen of oxindole (2) and aromatic ring (3), which could attenuate the steric hindrance caused by oligonu-
cleotide tags and facilitate target–binder interactions by pre-
senting multiple orientations (Fig. 3A).

As shown in Fig. 3B, DNA-linked acryl functionalized compound 1 was obtained in good conversion with 100 equiv. of isatin and 100 equiv. of proline via a three-component cycliza-
tion reaction at 30 °C. We observed that higher reaction temperatures were required for converting the starting material 1 to the desired products with some of the substituted isatins (Fig. 3B, 4c and 4d). The lack of reactivity in 4e at 80 °C was presumably a consequence of the steric hindrance of the chlor-
ide substitution. Elevated temperatures also caused the formation of side products and DNA decomposition. A repre-
sentative set of experimental results showing the interplay between temperature and isatin derivatives is found in the ESI S2.† 50 °C was chosen as the preferred temperature to balance product formation and minimize side product formation.

Further diversity was explored with cyclic/acyclic α-amino acids (Fig. 3B). Cyclic amino acids such as thiazolidine-4-
carboxylic acid and 4-hydroxyprolilidine-2-carboxylic acid provided moderate conversions under the optimized reaction conditions. Only benzyl-functionalized acyclic amino acid variants provided the desired products at 80 °C, likely due to the decreased stability of the intermediate. DNA-appended acryl-
amide 6 via on-DNA Heck coupling reaction developed previ-
ously were also tested in this annulation reaction (Fig. 3C) and moderate conversions were obtained, proving the possibility of two on-DNA methodologies in sequence. Next, we turned to isatin-functionalized DNA headpieces covalently linked through amide bond formation (Fig. 3D and E). We observed formation of the desired product along with 50% of byproduct, which had a molecular weight of desired product plus 18 Da (ESI S5†).

Several amide bond forming reaction conditions were tested with no improvement in the product ratio. Fortunately, these byproducts didn’t impact the subsequent on-DNA spiro-
annulation reaction under the optimized reaction conditions. In addition, we also carried out the proof-of-concept synthesis to apply this new on-DNA annulation reaction via Heck cross-
coupling reaction,19 Suzuki cross-coupling,20 nitro reduction, reductive amination and amide bond formation. These trans-
formations could be found in the ESI S6.†

Fig. 3 On-DNA three-component cyclization using three DNA-tagging points. (A) Three types of DNA-tagged spiropyrrilizidine centered modes. (B) Optimization of on-DNA three-component cyclization from acrylamide linked DNA oligonucleotides. Scope of substituted isatins, prolines and amino acids. (C) on-DNA reactions from DNA-linked 6. (a) Isatin (200 equiv.), proline (300 equiv.), 80 °C, 3 h. (D) On-DNA cyclization reaction from DNA-linked 8. (b) Acryl ester (200 equiv.), proline (200 equiv.), rt, 3 h. (E) On-DNA cyclization reaction from DNA-linked 10. (c) Acryl ester (200 equiv.), proline (200 equiv.), rt, 3 h. Conversions determined by LCMS.

Fig. 4 On-DNA cyclopropanation reaction.

Fig. 5 On-DNA ring-opening reactions by nucleophiles.
Spirocyclopropyl oxindoles are useful building blocks with a broad spectrum of biological and pharmaceutical applications such as HIV-1 NNRT1 inhibitors and arginine vasopressin inhibitors. Various approaches to access spirocyclopropyl fragments have been established between diazo compounds and electron-deficient olefins using a variety of metal catalysts. In particular, InCl₃ has been demonstrated to promote reaction with 3-diazoindole and electro-deficient olefins in the presence of water. This was the preferred starting point to develop the corresponding DNA-compatible cycloaddition reactions. A range of promoters including Cu, Ag, Pd, Pt, Au, Ru, Ga, Mg, Fe and Zr were tested and PdCl₂ provided 67% of the desired on-DNA spirocyclopropyl oxindolet. Unreacted DNA-conjugated acryl represented the mass balance at 100%, and low DNA recovery was found upon LC-MS analysis (ESI S7† Fig. 4). Unfortunately, lower reaction temperatures also significantly decreased the conversions. Meanwhile, a proof-of-concept library synthesis using ester functionalized diazo-isatin 12b as a bifunctional chemical linker was performed, providing the desired product 13b. After hydrolysis, the resulting carboxylic acid 13c enabled further diversity via acylation but with low conversion (13d, 10% conversion). Although this on-DNA set of reaction conditions was not robust enough for constructing spirocyclopropyl centered DNA-encoded libraries due to the low DNA recovery and suboptimal conversions to the desired products, it opens the door for more future exploration of this type of on-DNA transformation.

Highly strained spiroaziridine oxindoles and 3,3'-disubstituted oxindoles are attracting attention for their interesting bioactivities. We proposed to synthesize the spiroaziridine oxindole scaffolds first followed by DNA barcoding conjugation. DEL-C was constructed via DNA compatible ring-opening reactions to introduce aliphatic amines and phenol building blocks into the library. As illustrated in Fig. 5, the starting point for this exploration was the synthesis of carboxylic acid functionalized...
spiroaziridine oxindoles. From the methodologies established over past decades, we selected an aza-Corey–Chaykovsky reaction\textsuperscript{23} to access the aziridine unit via the addition of sulfur ylides to isatin derived N-tert-butanesulfinyl ketimines. Following oxidation and ester hydrolysis, the desired scaffold 14 was covalently linked to DNA barcoding via DMTMM promoted amide formation. We next turned to investigate the aminolysis and hydrolysis of the aziridine motif through a DNA-compatible ring-opening reaction. Both aliphatic amines and phenols provided moderate to excellent conversions. Weakly nucleophilic aromatic amines showed a lack of reactivity. Meanwhile, we also observed that DNA-conjugated 3,3\textsuperscript{0}-dihydroxyloxindoles were formed as byproducts at elevated temperatures under basic conditions, possibly due to the residual base in the final step. Therefore, a spin filter procedure was required to remove the remaining basic reagents and to maintain neutral conditions. The corresponding DEL construction and validation are described in the next section.

2-Indolinone is a well-known scaffold and present in drugs for cancer treatment.\textsuperscript{16,24} Two tagging points were employed to construct 3-alkenyl oxindole DNA-encoded libraries via a Claisen–Schmidt condensation reaction of substituted indoline-2-ones with various aldehydes in the presence of TEA (Fig. 6). Both aliphatic aldehydes and aromatic aldehydes delivered modest to excellent conversions. However, ketones gave negligible product formation even under forcing basic conditions, possibly due to relatively low nucleophilicity. Depending on the targets to be screened, such libraries may provide potential starting points for covalent drugs. The corresponding DEL construction and validation are described in the next section.

Encoded libraries construction

DNA-encoded libraries A, C and D were constructed by the newly developed DNA-compatible reactions with readily accessible building blocks such as amino acids, amines, borates and aldehydes. The size of the combined libraries was approximately 1.05 million (Fig. 7A). To determine the physicochemical properties of the compounds that compose the three libraries, 6000 compounds were randomly selected to summarize the physicochemical properties as shown in Fig. 7B. The full library molecule (tri-synthon) have a slightly higher molecule weight (ca. 650 Da on average) than Lipinski rule-of-5 compliant

![Fig. 8 Selection outcomes of DELs-A3–5. (A) Selection data analysis of DEL-A3 for BRD4-BD1. (B) Selection data analysis of DEL-A4 for BRD4-BD1. (C) Selection data analysis of DEL-A5 for BRD4-BD1. (D) Selection data analysis of DEL-A3 for Aurora. (E) Selection data analysis of DEL-A4 for P300. (F) Validation of three compounds 21, 22, and 23 by off-DNA resynthesis and biochemical assay data. (G) Validation of two compounds 24 and 25 by off-DNA resynthesis and IC\textsubscript{50} assay data.](image-url)
compounds, while the other properties meet the range of the typical hit compound profiles. Comparing the chemical space of DELs with 10 752 compounds from DrugBank, we applied the Principle Component Analysis (PCA) for both pool and found the overlap between them was more than 90%.

Not surprisingly, the properties of the di-synthon with the oxindole core were more attractive for hit-to-lead optimization. To evaluate whether the divergent isatin based DEL also worked for the screening of biological targets, we performed affinity screenings for three protein targets BRD4-BD1 (Fig. 8A–C), Aurora A (Fig. 8D) and p300 (Fig. 8E). DEL-A library pool was screened against BRD4-BD1 protein using the known inhibitor JQ1 \( ^{25} \) as a positive control. After the affinity screening, the samples were subjected to NGS and the generated data was translated into the corresponding copy number for the data analysis. In order to fully analyze the data among five different isatin libraries (DEL A1–A5), we used the enrichment for the normalization of different DELs\(^{26} \) (see the ESI\(^ \dagger \) for the enrichment calculation). Fig. 8A–E illustrated the three-dimensional view of the corresponding selection outcomes for each individual library. From the data analysis, we found that DEL-A3–5 had obvious features for the BRD4-BD1, and the competition experiment with JQ1 further confirmed that the features are competitive binders likely within the similar binding pocket with JQ1 (ESI S8\(^ \dagger \)). As illustrated in the three cubes (Fig. 8A–C), nascent SAR has been identified for DEL A3–5. From the detailed analysis, we have synthesized the truncated compounds 21–23 for activity confirmation with BRD2, BRD3 and BRD4. All synthesized compounds have moderate Ki values which demonstrated our rationale for the design and synthesis of the diversified oxindole libraries.

For Aurora A, we found the libraries had features with moderate enrichment (<100) in “target” and low in the “no target” control. However, obvious SAR has been identified for the cycle 3 BBs (Fig. 8D). 3_184 and 3_135 have similar structures and the latter was selected for synthesis off-DNA (24). 24 showed weak activity against Aurora A, demonstrating that moderate enrichment also could correlate with binding affinity. For the p300, a potential feature was found in DEL-A4 with good enrichment and the cycle 3 BB 3_147 was selected (Fig. 8E). Truncated structure 25 was synthesized off-DNA and validated in a biochemical assay. 25 showed weak activity against the target, but demonstrated the potential for multi-target selection with our designed DELs.

Conclusion

In summary, based on the oxindole scaffolds, we have developed a series of on-DNA synthetic transformations including 1,3-dipolar cyclization, cyclopropanation, ring-opening of aziridines and Claisen–Schmidt condensations to construct diverse oxindole derivatives. Among these reactions, the cyclopropanation and ring-opening reaction of aziridines are the first reported examples of on-DNA reactions, which opens the door for future exploration and applications. The majority of these transformations enable the diversity-oriented synthesis of DNA-encoded oxindole libraries which have been demonstrated with the design and synthesis of spiropyrrrolizidine-oxindole DEL A libraries. Three protein targets have been screened and active hit compounds were identified and confirmed with moderate to good activities. Expanding the scope of synthetic transformations compatible for DEL library synthesis, particularly based on the oxindole scaffolds, more active hits could be identified as starting points for medicinal chemistry programs. In addition, oxindole libraries based on the novel on-DNA reactions are currently under construction and additional results from DEL selections will be reported in the due course.

Experimental procedures

Full details of synthesis and LC-MS/MS analysis are provided in the ESI\(^ \dagger \).

Author contributions

Jiaxiang Liu, Zi Qin Yan, and Xiaohong Liu contributed equally to this article.

Conflicts of interest

The authors declare no competing interests.

Acknowledgements

We gratefully acknowledge financial support from the National Natural Science Foundation of China (81773634 to M. Z., 21877117 and 91953203 to X. L., 21907103 to X. W., 81673295 to Y. Z.), National Science & Technology Major Project “Key New Drug Creation and Manufacturing Program”, China (Number: 2018ZX09711002 to H. J., No. 2018ZX09711002-005 to X. L.), and “Personalized Medicines—Molecular Signature-based Drug Discovery and Development”, Strategic Priority Research Program of the Chinese Academy of Sciences (XDA12050201 to M. Z.; XDA12040330 and XDA12050405 to X. L.), Shanghai Commission of Science and Technology (18431907100 to X. L.). We also thank Dr Ryan Wurz of Amgen for critical reading.

References

1 (a) M. Schenone, V. Dancik, B. K. Wagner and P. A. Clemons, Nat. Chem. Biol., 2013, 9, 232–240; (b) C. Finan, A. Gaulton, F. A. Kruger, R. T. Lumbers, T. Shah, J. Engmann, L. Galver, R. Kelley, A. Karlsson, R. Santos, J. P. Overington, A. D. Hingorani and J. P. Casas, Sci. Transl. Med., 2017, 9, eaag1166; (c) M. Luo, J. Jiao and R. Wang, BMC Bioinf., 2019, 20, 198.
2 D. G. Brown and J. Bostrom, J. Med. Chem., 2018, 61, 9442–9468.
3 (a) J. L. Reymond and M. Awake, ACS Chem. Neurosci., 2012, 3, 649–657; (b) M. Gonzalez-Medina, J. J. Naveja, N. Sanchez-Cruz and J. L. Medina-Franco, RSC Adv., 2017, 7, 54153–54163.
4 S. Morgan, P. Grootendorst, J. Lexchin, C. Cunningham and D. Greyson, Health Policy, 2011, 100, 4–17.
