RSC Medicinal Chemistry

ROYAL SOCIETY OF CHEMISTRY

RESEARCH ARTICLE

View Article Online
View Journal | View Issue

Open Access Article. Published on 29 July 2025. Downloaded on 1/20/2026 4:13:46 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Check for updates

Cite this: *RSC Med. Chem.*, 2025, **16**, 4774

Design and synthesis of a chemically diverse, lead-like DNA-encoded library from sequential amide coupling

Cameron E. Taylor,[a] Grace Roper,[a] Rhianna Young,[b] Fredrik Svensson,[c] Andreas Brunschweiger,[d] Sam Butterworth,[e] Andrew G. Leach[e] and Michael J. Waring *[a]

Received 22nd April 2025,
Accepted 24th July 2025

DOI: 10.1039/d5md00350d

rsc.li/medchem

DNA-encoded libraries (DELs) are established as an effective screening strategy to identify protein ligands and offer a cost-effective means of screening large numbers of compounds. However, the synthesis and utilisation of DELs is implemented by relatively few laboratories. Here, we describe the design and synthesis of a medium-sized DEL through simple amide coupling procedures. We provide details of chemistry and enzymatic steps and demonstrate their effectiveness by synthesising 300 thousand and 3 million-member DELs. We demonstrate their integrity through screening against carbonic anhydrase IX and show their chemical diversity through *in silico* comparison with an established high-throughput screening library. The DELs described can be used as a resource to accelerate hit identification for early-phase drug discovery and are available to the academic community for screening.

Introduction

DNA-encoded libraries (DELs) have become a powerful hit-finding method in drug discovery,[1–4] complementary to existing and emerging hit-identification techniques. Libraries potentially consisting of billions of encoded compounds can be synthesised, stored, and screened at significantly lower costs than HTS libraries due to the advantages of DNA encoding. This allows both combinatorial synthesis and the use of libraries as pooled mixtures. Signal amplification by PCR means that screening can be carried out on a very small scale without significant protein demands. Once robust chemistry is established, a large DEL can be synthesised in weeks and screened in days. As a result, DEL approaches are now employed by most large pharmaceutical companies as a core hit-finding capability and are accessible to others *via* specialist providers.[5–8]

Specialist equipment and techniques are required for DEL synthesis, and uptake in smaller companies and academic laboratories is limited partly for this reason. Preparation of very large libraries requires significant capabilities in reagent handling, information capture and logistics, as well as the cost associated with purchasing large numbers of specialised chemical building blocks and coding oligonucleotides. These high initial costs can be a barrier to those new to the field.

Although extremely large (>100 million-member) DELs are increasingly common, the advantages of larger libraries are perhaps overstated. Due to the lack of structural similarities in many commonly used reagent sets, the increase in diversity as the number of monomers is increased is limited. Furthermore, very large libraries have often relied on common scaffolds such as triazines,[9,10] which further limits their structural diversity.[11–13] Another drawback of large libraries is that it is harder to ensure library fidelity as size increases, as compromises have to be made in the validation of the chemical building block couplings. Selections from larger libraries can also be more challenging to sequence reliably since larger numbers of compounds increase signal noise and thus require significantly increased sequencing depth.[14]

Smaller, structurally simpler libraries can avoid some of these difficulties and are potentially more accessible. Chemical diversity can be sufficiently maintained with a judicious choice of DEL synthesis schemes and their constituent monomers. In particular, higher diversity results inherently from using library schemes that couple monomers

[a] *Cancer Research Horizons Newcastle Drug Discovery Unit, Chemistry, School of Natural and Environmental Sciences, Newcastle University, Bedson Building, Newcastle upon Tyne, NE1 7RU, UK. E-mail: mike.waring@ncl.ac.uk*
[b] *Medical School, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK*
[c] *Cancer Research Horizons Therapeutic Innovation, Jonas Webb Building, Babraham Research Campus, Cambridge, CB22 3AT, UK*
[d] *Institut für Pharmazie und Lebensmittelchemie, Lehrstuhl fur Pharmazeutische Chemie, Am Hubland, D-97074 Würzburg, Germany*
[e] *School of Health Sciences, University of Manchester, Oxford Road, Manchester, M13 9PL, UK*

 This journal is © The Royal Society of Chemistry 2025

Open Access Article. Published on 29 July 2025. Downloaded on 1/20/2026 4:13:46 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

together with minimal common substructures (rather than sequential additions to a central scaffold).

To explore these concepts further, we describe the development of a DEL from validation of building block couplings to library synthesis and selection against an exemplar protein, carbonic anhydrase IX (CAIX). We sought to design a lead-like library and investigate its quality by assessment of its physicochemical properties and chemical diversity in comparison with a high-quality traditional high-throughput screening library.

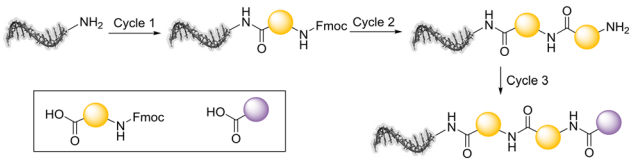
Results and discussion

Library design and validation

A simple, linear, 3-cycle library design was chosen, utilising two readily available building block classes with well-established chemistry to rapidly synthesise a medium-sized DEL. This comprised two cycles of amide coupling of *N*-Fmoc-protected amino acids, each followed by Fmoc deprotection, followed by an amide coupling using capping carboxylic acids (Scheme 1). The amide coupling and Fmoc deprotection steps were optimised to work on a small scale with minimal use of reagents.

After a trial of amide coupling methods, we chose DMTMM-mediated reaction conditions, which gave good conversion across a range of monomers. Validation reactions using optimised conditions were performed in PCR plates using 250 pmol DNA with 630 equivalents of the acid (typically <100 μg per reaction (150 nmol)). This is significantly lower than most reported validation experiments, which use approximately 1.5 mmol.[15–18] Completed reactions were diluted with water, and the crude mixture was analysed by RP-LCMS.

59 *N*-Fmoc amino acids and 116 carboxylic acids were selected based on chemical diversity, desirable functionality and physicochemistry (clog *P* and molecular weight). Each was evaluated for reaction efficiency by coupling to 14 nucleotide DNA headpiece **1** (Fig. 1). The validation experiments showed that a wide range of functional groups and acids were tolerated. Most unprotected aliphatic amines were incompatible as expected, however the majority of anilines coupled well. Alcohols and phenols proved problematic. The presence of very bulky groups α- to the carboxylic acid generally led to poorer conversions. It is hypothesised that the solubility of both the free carboxylate and activated ester play a significant role in determining conversion. However, some acids that were visibly sparingly soluble in DMF did couple with >95% conversion.



Scheme 1 DEL synthesis scheme comprising two rounds of amide-coupled *N*-Fmoc amino acids with a final capping carboxylic acid.

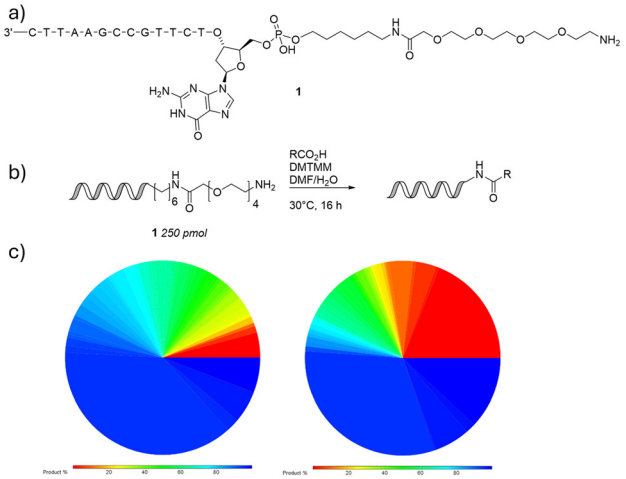


Fig. 1 Summary of building block validations: a) 14 nt DNA headpiece **1** used as a model in this work; b) amide coupling validation reaction; c) observed conversions for *N*-Fmoc amino acids (left) and carboxylic acids (right) determined by AUC integration of RP-LCMS BPC (1050–3000 Da) chromatogram.

The deprotection of DNA-conjugated *N*-Fmoc protecting groups (using 10% aqueous piperidine) is well-documented and typically results in complete deprotection for a wide range of substrates.[15,19,20] Our observations were consistent with the previous work. However, direct precipitation of the reaction mixture typically resulted in low recovery. Little is reported about this in the literature. To investigate isolation recovery, a model system using DNA-coupled-*N*-Fmoc alanine was used. It was supposed that the presence of piperidine was the major reason for the low recovery, so removal of the piperidine was attempted using centrifugal evaporation and buffer exchange *via* size exclusion centrifugal spin filtration.[21,22] A combination of both methods was required for consistently high recovery (Table 1).

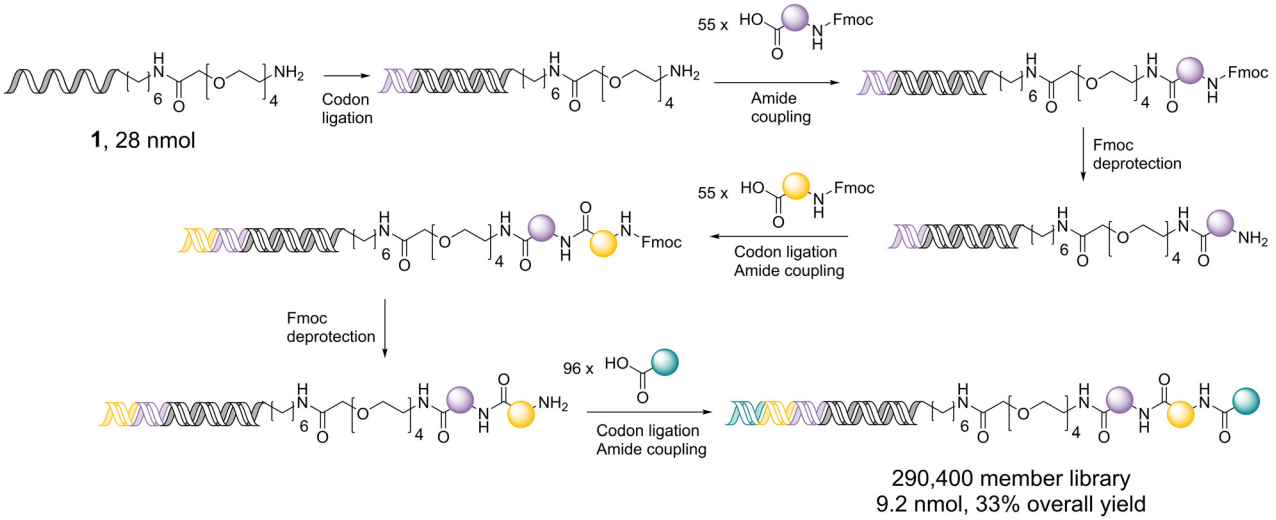
Library synthesis

The validation data was used to guide building block selection, where high conversion was an important consideration for inclusion. We selected 55 *N*-Fmoc amino acids and 96 carboxylic acids for the library synthesis. Two control wells were included in each amide coupling step to simplify potential hit follow-up. The first was subjected to identical reaction conditions without the addition of any carboxylic acid, and in the second, no reagents were added. Including these controls identifies cases in which truncated

Table 1 Percentage recovery from 4 different purification methods after Fmoc deprotection (n = 3)

Conditions	Recovery%
Precipitation only	20 (±11)
Concentration then precipitation	49 (±5)
Buffer exchange, then precipitation	46 (±1)
Concentration, buffer exchange, then precipitation	89 (±2)

Open Access Article. Published on 29 July 2025. Downloaded on 1/20/2026 4:13:46 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.



Scheme 2 300 thousand DEL synthesis comprising two sequential cycles of coupling of 55 *N*-Fmoc amino acids (with Fmoc deprotection) followed by capping with 96 carboxylic acids.

side-products may contribute to activity observed in selections. DNA codons were designed with a Hamming distance of 3, and palindromic or hairpin-forming sequences were removed. 200 barcodes were designed for each cycle number. The full list of barcodes is detailed in the SI.

Library synthesis began with the single-stranded DNA headpiece **1**, which was subjected to two coupling cycles of the 55 *N*-Fmoc amino acids, each with subsequent Fmoc removal, followed by coupling of the 96 carboxylic acids (Scheme 2). An encoding step (ligation of the respective DNA codon sequences) was carried out prior to each amide coupling. The optimised chemistry conditions were used at the same concentration during library synthesis, which we believe to be an important measure to ensure library fidelity. Ligation efficiencies at each stage were assessed by analytical gel electrophoresis of a representative sample of wells (Fig. S7–S9). Precipitation of each ligation reaction was carried out in the plate. Avoiding transferring the samples to individual tubes helped maximise DNA recovery and reduced the risk of handling errors.[23] The precipitate was pelleted by centrifugation, and the supernatant was removed prior to amide coupling in the same plate.

Overall, the final library was produced in 33% yield over the five synthesis and three encoding steps, resulting in 9.2 nmol of the final DEL. This library synthesis used far lower DNA input than the μmol quantities often used, making it an attractive starting point for new projects.[15–17,24]

Selection against CAIX

The library was screened against CAIX as a test case. During affinity selections, we used approx. 1 million copies per compound (500 fmol library) and 4.8 μg His-tagged CAIX as reported by Kunig *et al.*[25–32] Two rounds of selection were performed, and after PCR amplification and Illumina® sequencing, counts for unique DNA barcodes were summed for analysis. Analysis of the sequencing data revealed enrichment of 4-sulfamoylbenzoic acid from the cycle 3 building block set (Fig. 2a). To a lesser degree, imidazole-4-carboxylic acid and 2-hydroxybenzene-1,4-dicarboxylic acid were also enriched (Fig. 2a). There was no significant enrichment of the cycle 1 or 2 building blocks. Primary sulfonamides are well-precedented binding motifs in CAIX inhibitors, in which the sulfonamide group is known to interact directly with the catalytic zinc atom with high tolerance of other functionality within the ligand.[30] Including this inbuilt control to the library confirms that the chemistry steps were successful and the building blocks were correctly encoded.

Imidazole-4-carboxylic acid was also enriched in selections against other targets using His-beads. Therefore, it was assumed that these compounds are false positives that are selected due to the imidazole binding to the NTA-cobalt-based beads directly. The addition of imidazole to block vacant bead sites and herring sperm DNA to outcompete non-specific DNA binding were both investigated. Adding imidazole did not eliminate the observed selection enrichment of the imidazole-containing compounds. However, the background noise was reduced when herring sperm DNA was used as a blocking agent.

As an alternative means to reduce the level of imidazole enrichment, the DEL was *pre-cleared* by incubating with a half portion of beads and the resulting supernatant was incubated with the immobilised protein. The pre-cleared library was selected against CAIX in the presence of herring sperm DNA, resulting in much lower imidazole enrichment and background noise, leading to a clearer trend for enrichment of the sulfonamide-containing compounds (Fig. 2b).

Synthesis of a 3 million-member DEL

The same library scheme (Scheme 1) was expanded to produce a larger DEL. The building block validation experiments were

Open Access Article. Published on 29 July 2025. Downloaded on 1/20/2026 4:13:46 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

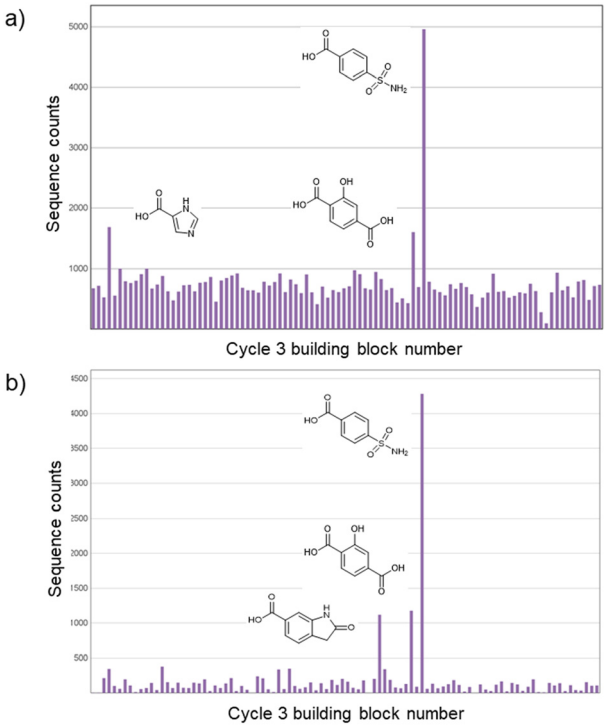


Fig. 2 a) Sequence counts for the cycle 3 monomer for selections of the 300 thousand-member library against CAIX, a) initial selection, b) selections after pre-clearing the library against the beads.

expanded to include 171 *N*-Fmoc amino acids (after property filters). The validation of carboxylic acids was expanded to 227 carboxylic acids (SI). From this, 142 *N*-Fmoc amino acids and 147 carboxylic acids were included in the expanded library. The synthesis proceeded as expected and resulted in a yield of 9 nmol from an initial 56 nmol with good ligation efficiencies (Fig. S1–S3). The library was selected against His-CAIX with library pre-clearing and 4-sulfamoylbenzoic acid was clearly enriched (Fig. 3).

The 3 million-member DEL was assessed for its lead-like properties using established calculated descriptors and compared to the enamine HTS set, which was selected as a

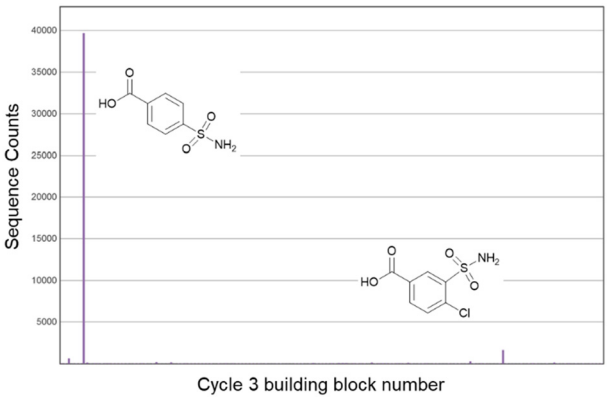


Fig. 3 Sequence counts for the cycle 3 monomer for selection of the 3 million-member DEL against CAIX.

Table 2 Mean and std. dev. values of calculated properties of the DEL (n = 3 089 664) and enamine HTS set (n = 1 367 257)

	DEL	HTS set
MWt	438 ± 69	334 ± 53
log P	1.7 ± 1.3	2.5 ± 1.1
HBD	2.6 ± 0.9	1.1 ± 0.8
TPSA	106 ± 19	69 ± 21
NRB	7.4 ± 2.2	4.5 ± 1.5
NAr	1.8 ± 1	2 ± 0.9

representative, high-quality screening library that is widely used in standard screening approaches. The DEL had overall very good lead-like properties with mean molecular weight (MWt) 438 Da, log P 1.7, hydrogen bond donors (HBD) 2.6, topographical surface areas (TPSA) 106, rotatable bond count (NRB) 7.4 and aromatic ring count (NAr) 1.8 (Table 2).[33–37] These were similar to the enamine HTS library (the DEL is slightly higher in size overall and slightly more polar; no differences were significant).

Due to the combinatorial nature of DELs, it can be difficult to constrain physicochemical properties within a defined range. The combinations of the most extreme building blocks (the largest or most lipophilic in every cycle, for example) may lead to compounds in a DEL occupying a wider range of properties than traditional compound sets that consist of individually prepared sub-libraries, allowing tighter property distributions with fewer compounds outside the ideal range. In the case of this DEL, the property distributions were good overall. The DEL contained more MWt > 500 Da compounds than the HTS set, but this was not a significant portion of the total (Table 2 and Fig. 4). Log P and NAr distributions were similar between the two. HBD and PSA were slightly skewed to higher values in the DEL as a consequence of the inherent amide bonds arising from the synthesis, but the compounds generally populated the desirable range (Fig. 4).

The chemical diversity of the DEL was assessed using a principal components analysis.[38] The DEL had a similar diversity to the HTS with significant overlap in chemical space between the two sets and comparable areas populated exclusively with one or the other (Fig. 5). Principal moments of inertia (PMI) plots showed that the DEL compounds also have similar topology distributions to the HTS set with good representation across the majority of the plots (Fig. 6).[39]

Conclusions

This work demonstrates the facile synthesis of a 3 million-member amide-coupled DNA-encoded library. The optimised library synthesis chemistry conditions are high-yielding and use μg quantities of monomers. These conditions, combined with the validated conversions for 171 *N*-Fmoc amino acids and 227 carboxylic acids, alongside the DNA barcodes, should be transferable to other libraries. This work will significantly reduce the time and cost required to initiate a

Open Access Article. Published on 29 July 2025. Downloaded on 1/20/2026 4:13:46 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

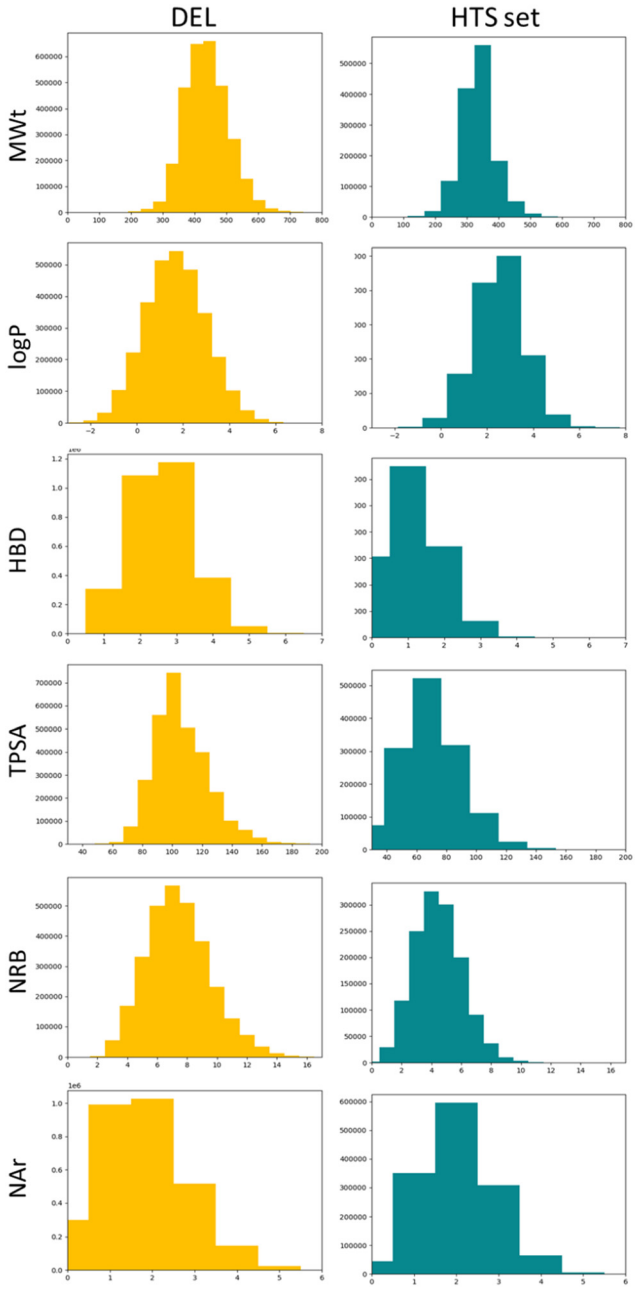


Fig. 4 Calculated property distribution comparisons between the DEL (gold) and the enamine HTS set (teal).

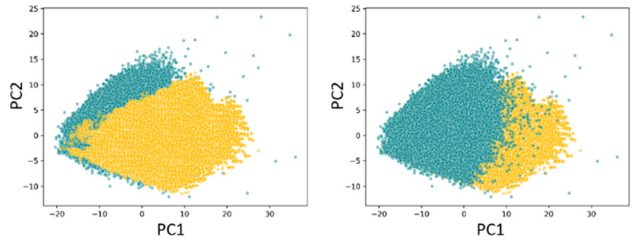


Fig. 5 Chemical diversity principal components analysis of the DEL (gold) compared to the enamine HTS set (teal), left panel with DEL compounds on top, right panel with HTS compounds on top.

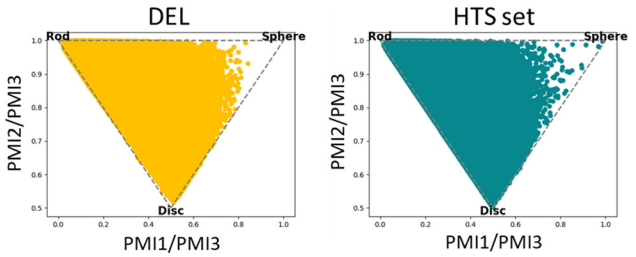


Fig. 6 Principal moments of inertia plots for the DEL (gold) and the enamine HTS set (teal).

similar DEL project. The selection of the synthesised libraries against CAIX demonstrates that the methodology for both the on-DNA chemistry and the corresponding encoding steps were successful. The selection conditions were optimised to reduce noise and the influence of bead-binding compounds; we believe the selection conditions presented will be useful as a general starting point for a DEL screening campaign against almost any immobilised target. Computational analysis showed that the DEL has good lead-like properties and chemical diversity comparable to a high-quality HTS set. This arises from using a library synthesis scheme that employs monomer couplings, not using a common scaffold, and due to the careful selection of monomers. This efficient and user-friendly library methodology will be of great utility to academic groups actively employing DEL technology and those looking to leverage DELs in their research. The DELs presented in this manuscript are available open-access for screening by academic users.

Author contributions

CET carried out the DEL design and synthesis and the selections and co-wrote the manuscript, GR developed conditions for the selections, RY wrote the code that processes the library data, FS carried out the computational analysis of physicochemical properties and chemical diversity, AB cosupervised the project and designed the DNA sequences, SB cosupervised the project, AGL designed the DNA sequences, MJW supervised CET and GR, led the project and co-wrote the manuscript. All authors reviewed the manuscript prior to submission.

Conflicts of interest

There are no conflicts to declare.

Data availability

SI includes full experimental procedures, DNA-encoded library synthesis information, analytical data, NGS analysis information, and assay results. See DOI: **https://doi.org/10.1039/D5MD00350D**.

Full experimental details and data supporting this article are available in the SI.

 This journal is © The Royal Society of Chemistry 2025

Open Access Article. Published on 29 July 2025. Downloaded on 1/20/2026 4:13:46 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Acknowledgements

The research was sponsored by Wellcome Trust (funding for CET, grant ref. 221302/Z/20/Z). We also gratefully acknowledge Cancer Research UK (programme grant to the Newcastle Drug Discovery Group, refs. C2115/A21421 and RCDDRPGMApr2020\100002), Astex Pharmaceuticals (Drug Discovery Alliance, support for GR) and EPSRC MoSMed CDT (ref. EP/S022791/1, studentship award to RY) for funding. We thank K. Götte and E. Detta for useful discussions and advice.

Notes and references

1 N. Favalli, G. Bassi, J. Scheuermann and D. Neri, DNA-Encoded Chemical Libraries – Achievements and Remaining Challenges, *FEBS Lett.*, 2018, 2168–2180.
2 B. Halford, How DNA-Encoded Libraries Are Revolutionizing Drug Discovery, *Chem. Eng. News*, 2019, **95**(25), 28–33.
3 J. Ottl, L. Leder, J. V. Schaefer and C. E. Dumelin, Encoded Library Technologies as Integrated Lead Finding Platforms for Drug Discovery, *Molecules*, 2019, **24**(8), 1629.
4 A. L. Satz, A. Brunschweiger, M. E. Flanagan, A. Gloger, N. J. V. Hansen, L. Kuai, V. B. K. Kunig, X. Lu, D. Madsen, L. A. Marcaurelle, C. Mulrooney, G. O'Donovan, S. Sakata and J. Scheuermann, DNA-Encoded Chemical Libraries, *Nat. Rev. Methods Primers*, 2022, **2**, 3.
5 **https://www.pharmaron.com/services/laboratory-services/laboratory-chemistry/dna-encoded-libraries-del/**.
6 **https://wuxibiology.com/drug-discovery-services/hit-finding-and-screening-services/dna-encoded-library-del/**.
7 **https://www.hitgen.com/en**.
8 **https://www.x-chemrx.com/**.
9 H. Deng, H. O'Keefe, C. P. Davie, K. E. Lind, R. A. Acharya, G. J. Franklin, J. Larkin, R. Matico, M. Neeb, M. M. Thompson, T. Lohr, J. W. Gross, P. A. Centrella, G. K. O'Donovan, K. L. Sargent, K. Van Vloten, S. Mataruse, S. R. Skinner, S. L. Belyanskaya, T. Y. Carpenter, T. W. Shearer, M. A. Clark, J. W. Cuozzo, C. C. Arico-Muendel and B. A. Morgan, Discovery of Highly Potent and Selective Small Molecule ADAMTS-5 Inhibitors That Inhibit Human Cartilage Degradation via Encoded Library Technology (ELT), *J. Med. Chem.*, 2012, **55**, 7061–7079.
10 Y. Ding, S. Belyanskaya, J. L. DeLorey, J. A. Messer, G. Joseph Franklin, P. A. Centrella, B. A. Morgan, M. A. Clark, S. R. Skinner, J. W. Dodson, P. Li, J. P. Marino and D. I. Israel, Discovery of Soluble Epoxide Hydrolase Inhibitors through DNA-Encoded Library Technology (ELT), *Bioorg. Med. Chem.*, 2021, **41**, 116216.
11 H. Zhu, M. E. Flanagan and R. V. Stanton, Designing DNA Encoded Libraries of Diverse Products in a Focused Property Space, *J. Chem. Inf. Model.*, 2019, **59**, 4645–4653.
12 O. Eidam and A. L. Satz, Analysis of the Productivity of DNA Encoded Libraries, *Med. Chem. Commun.*, 2016, **7**, 1323–1327.
13 R. M. Franzini and C. Randolph, Chemical Space of DNA-Encoded Libraries: Miniperspective, *J. Med. Chem.*, 2016, **59**, 6629–6644.
14 A. L. Satz, R. Hochstrasser and A. C. Petersen, Analysis of Current DNA Encoded Library Screening Data Indicates Higher False Negative Rates for Numerically Larger Libraries, *ACS Comb. Sci.*, 2017, **19**(4), 234–238.
15 M. Seydimemet, Y. Yang, Y. Lv, J. Liu, Z. Yan, Y. Zhao, X. Wang and X. Lu, Design, Construction, and Screening of Diversified Pyrimidine-Focused DNA-Encoded Libraries, *ACS Med. Chem. Lett.*, 2023, **14**(8), 1073–1078.
16 X. Wen, M. Zhang, Z. Duan, Y. Suo, W. Lu, R. Jin, B. Mu, K. Li, X. Zhang, L. Meng, Y. Hong, X. Wang, H. Hu, J. Zhu, W. Song, A. Shen and X. Lu, Discovery, SAR Study of GST Inhibitors from a Novel Quinazolin-4(1H)-One Focused DNA-Encoded Library, *J. Med. Chem.*, 2023, **66**(16), 11118–11132.
17 L. Fan, Y. Yu, C. Jayne, J. R. Frost and J. D. Scott, Synthesis of DNA-Encoded Macrocyclic Peptides via Nitrile-Aminothiol Click Reaction, *Org. Lett.*, 2023, **25**(44), 8038–8042.
18 Y. Li, E. Gabriele, F. Samain, N. Favalli, F. Sladojevich, J. Scheuermann and D. Neri, Optimized Reaction Conditions for Amide Bond Formation in DNA-Encoded Combinatorial Libraries, *ACS Comb. Sci.*, 2016, **18**(8), 438–443.
19 Y. Huang, L. Meng, Q. Nie, Y. Zhou, L. Chen, S. Yang, Y. M. E. Fung, X. Li, C. Huang, Y. Cao, Y. Li and X. Li, Selection of DNA-Encoded Chemical Libraries against Endogenous Membrane Proteins on Live Cells, *Nat. Chem.*, 2021, **13**(1), 77–88.
20 R. Ge, Z. Shen, J. Yin, W. Chen, Q. Zhang, Y. An, D. Tang, A. L. Satz, W. Su and L. Kuai, Discovery of SARS-CoV-2 Main Protease Covalent Inhibitors from a DNA-Encoded Library Selection, *SLAS Discovery*, 2022, **27**(2), 79–85.
21 S. He, B. Cao, Y. Yi, S. Huang, X. Chen, S. Luo, X. Mou, T. Guo, Y. Wang, Y. Wang and G. Yang, DNA Precipitation Revisited: A Quantitative Analysis, *Nano Sel.*, 2022, **3**(3), 617–626.
22 C. J. England, T. C. Gray, S. R. L. Malla, S. A. Oliveira, B. R. Martin, G. W. Beall and L. K. Lewis, pH-dependent sedimentation of DNA in the presence of divalent, but not monovalent, metal ions, *Anal. Biochem.*, 2021, **616**, 114099.
23 J. J. Priego, E. De, P. Beato, J. Benavides, A. Adrián Gironda-Martínez, F. Gonzaíez, J. Blas, D. Martín-Ortega, R. Ramón Rama-Garda, J. J. Ezquerra, M. A. Toledo and A. Torrado, On-DNA Palladium-Catalyzed Hydrogenation-like Reaction Suitable for DNA-Encoded Library Synthesis, *Bioconjugate Chem.*, 2021, **32**, 88–93.
24 L. Li, M. Su, W. Lu, H. Song, J. Liu, X. Wen, Y. Suo, J. Qi, X. Luo, Y.-B. Zhou, X.-H. Liao, J. Li and X. Lu, Triazine-Based Covalent DNA-Encoded Libraries for Discovery of Covalent Inhibitors of Target Proteins, *ACS Med. Chem. Lett.*, 2022, **13**, 1574–1581.
25 A. Sannino, E. Gabriele, M. Bigatti, S. Mulatto, J. Piazzi, J. Scheuermann, D. Neri, E. J. Donckele and F. Samain, Quantitative Assessment of Affinity Selection Performance by Using DNA-Encoded Chemical Libraries, *ChemBioChem*, 2019, **20**(7), 955–962.
26 Q. Chen, X. Cheng, L. Zhang, X. Li, P. Chen, J. Liu, L. Zhang, H. Wei, Z. Li and D. Dou, Exploring the Lower Limit of Individual DNA-Encoded Library Molecules in Selection, *SLAS Discovery*, 2020, **25**(5), 523–529.

Open Access Article. Published on 29 July 2025. Downloaded on 1/20/2026 4:13:46 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

27 D. Neri and R. A. Lerner, DNA-Encoded Chemical Libraries: A Selection System Based On Endowing Organic Compounds With Amplifiable Information, *Annu. Rev. Biochem.*, 2018, **87**, 479–502.

28 V. B. K. Kunig, M. Potowski, M. Akbarzadeh, M. Klika Škopić, D. dos Santos Smith, L. Arendt, I. Dormuth, H. Adihou, B. Andlovic, H. Karatas, S. Shaabani, T. Zarganes-Tzitzikas, C. G. Neochoritis, R. Zhang, M. Groves, S. M. Guéret, C. Ottmann, J. Rahnenführer, R. Fried, A. Dömling and A. Brunschweiger, TEAD–YAP Interaction Inhibitors and MDM2 Binders from DNA-Encoded Indole-Focused Ugi Peptidomimetics, *Angew. Chem., Int. Ed.*, 2020, **59**(46), 20338–20342.

29 C. L. Lomelino, C. T. Supuran and R. McKenna, Non-Classical Inhibition of Carbonic Anhydrase, *Int. J. Mol. Sci.*, 2016, **17**, 1150.

30 D. Vullo, M. Franchi, E. Gallori, J. Pastorek, A. Scozzafava, S. Pastorekova and C. T. Supuran, Carbonic Anhydrase Inhibitors: Inhibition of the Tumor-Associated Isozyme IX with Aromatic and Heterocyclic Sulfonamides, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 1005.

31 R. M. Franzini, T. Ekblad, N. Zhong, M. Wichert, W. Decurtins, A. Nauer, M. Zimmermann, F. Samain, J. Scheuermann, P. J. Brown, J. Hall, S. Gräslund, H. Schüler and D. Neri, Identification of Structure–Activity Relationships from Screening a Structurally Compact DNA-Encoded Chemical Library, *Angew. Chem.*, 2015, **127**(13), 3999–4003.

32 S. Oehler, M. Catalano, I. Scapozza, M. Bigatti, G. Bassi, N. Favalli, M. R. Mortensen, F. Samain, J. Scheuermann and D. Neri, Affinity Selections of DNA-Encoded Chemical Libraries on Carbonic Anhydrase IX-Expressing Tumor Cells Reveal a Dependence on Ligand Valence, *Chem. – Eur. J.*, 2021, **27**(35), 8985–8993.

33 T. J. Ritchie and S. J. F. Macdonald, Physicochemical Descriptors of Aromatic Character and Their Use in Drug Discovery, *J. Med. Chem.*, 2014, **57**, 7206–7215.

34 D. F. Veber, S. R. Johnson, H.-Y. Cheng, B. R. Smith, K. W. Ward and K. D. Kopple, Molecular Properties That Influence the Oral Bioavailability of Drug Candidates, *J. Med. Chem.*, 2002, **45**, 2615–2623.

35 C. A. Lipinski, Drug-like Properties and the Causes of Poor Solubility and Poor Permeability, *J. Pharmacol. Toxicol. Methods*, 2000, **44**, 235–249.

36 V. Khanna and S. Ranganathan, Physiochemical Property Space Distribution among Human Metabolites, Drugs and Toxins, *BMC Bioinf.*, 2009, **10**(Suppl 15), S10.

37 S. Prasanna and R. Doerksen, Topological Polar Surface Area: A Useful Descriptor in 2D-QSAR, *Curr. Med. Chem.*, 2008, **16**(1), 21–41.

38 L. B. Akella and D. DeCaprio, Cheminformatics Approaches to Analyze Diversity in Compound Screening Libraries, *Curr. Opin. Chem. Biol.*, 2010, **14**, 325–330.

39 W. H. B. Sauer and M. K. Schwarz, Molecular Shape Diversity of Combinatorial Libraries: A Prerequisite for Broad Bioactivity, *J. Chem. Inf. Comput. Sci.*, 2003, **43**(3), 987–1003.

 This journal is © The Royal Society of Chemistry 2025