RESEARCH ARTICLE

Photochemical synthesis of an epigenetic focused tetrahydroquinoline library†

Adam I. Green[a] and George M. Burslem [iD] *[ab]

Discovery of epigenetic chemical probes is an important area of research with potential to deliver drugs for a multitude of diseases. However, commercially available libraries frequently used in drug discovery campaigns contain molecules that are focused on a narrow range of chemical space primarily driven by ease of synthesis and previously targeted enzyme classes (*e.g.*, kinases) resulting in low hit rates for epigenetic targets. Here we describe the design and synthesis of a compound collection that augments current screening collections by the inclusion of privileged isosteres for epigenetic targets.

Introduction

The screening of libraries of small molecules underpins many drug discovery efforts.[1] Most commonly, the chemistry used to construct these compound collections is selected to be easily expanded in throughput to provide hundreds or thousands of compounds.[2] This allows for a large number of compounds to be tested but can mean compounds included in such libraries have simple structures or lack diversity. Certain motifs are overrepresented because of their facile incorporation into high-throughput synthesis workflows. An analysis of 66 submissions describing clinical candidates to the *Journal of Medicinal Chemistry* between 2016 and 2017 revealed that 30% targeted kinases, 17% GPCRs and 9% epigenetic targets,[3] yet epigenetic bioisosteres are infrequently included in screening collections.[4] This is compounded by the prediction that much of purchasable chemical space is targeted towards proteases, GPCRs, and kinases.[5] This deficiency in diversity and lack of bioisostere inclusion means that hit rates are often low against epigenetic targets.[6] Consequently, we[4] and others,[7] have called for bespoke screening collections to be designed and synthesized to address these deficiencies.

Drug discovery often utilizes privileged chemical motifs, or bioisosteres, as a starting point to identify chemical matter for a specific target class. Examples include hinge-binding motifs for kinases[8] and acetyl lysine mimics for bromodomains.[9] While effective chemotypes for kinases are frequently included in commercially available compound collections similarly effective chemotypes are not included for epigenetic targets. This is particularly true for privileged chemotypes that target acetyl- and methyl- lysine reader, writer, and eraser proteins which are of significant interest for the treatment of cancers, inflammation, and other diseases.[4,10,11] Examples of highly effective chemotypes include 3,5-dimethylisoxazoles[12] and 1,2,4-triazoles[13] for bromodomains (acetyl lysine readers), and *N*-alkyl amines for methyl lysine readers[14] (Fig. 1A).

One issue with screening libraries is the narrow toolkit of reactions that has typically been used to synthesize small molecules, limiting the types of scaffolds that can be accessed for biological testing.[15] However, recently developed photochemical cyclizations have enabled the synthesis of new molecular entities.[16–19] Specifically, the synthesis of tetrahydroquinolines is of particular interest as the scaffold is a highly effective and versatile pharmacophore, with over 10 000 bioactive molecules containing this pharmacophore reported in the ChEMBL database (Fig. 1B).[20] An example of their synthesis has been reported *via* a photochemically induced radical annulation between maleimides and *N*-alkyl anilines.[21] This reaction is enabled by the formation of an electron donor–acceptor (EDA) complex between the two reactants. EDA complexes are Lewis acid–base complexes that can absorb visible light and undergo single-electron transfer (SET) to generate radical intermediates.[22] The main benefit of EDA-mediated reactions is that they do not require a catalyst, or anything beyond the two starting materials, and are therefore amenable for the rapid synthesis of small molecule screening collections. Whilst this EDA-mediated scaffold synthesis has been previously reported, its application to

[a] Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA.
E-mail: George.Burslem@Pennmedicine.upenn.edu
[b] Department of Cancer Biology and Epigenetics Institute, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA
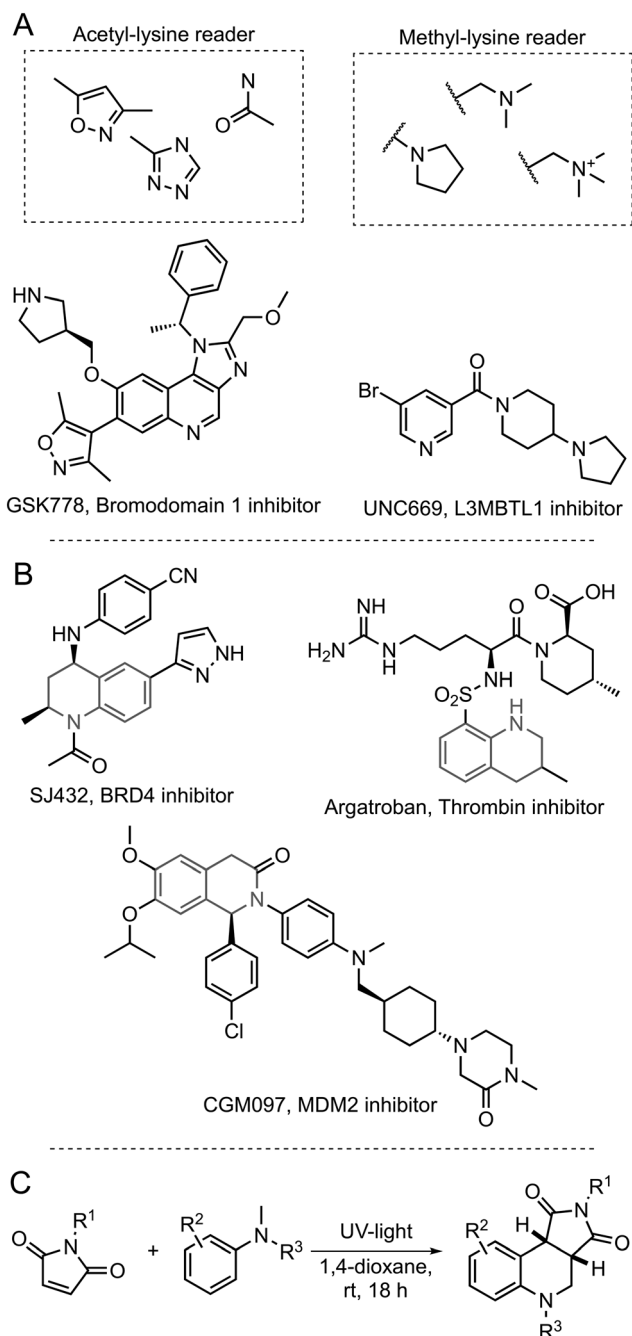
Fig. 1 Exemplar epigenetic bioisosteres and ligands containing a tetrahydroquinoline pharmacophore highlighted in green. A: Examples of chemotypes employed as part of inhibitors of acetyl-lysine readers and methyl-lysine readers.[23–25] B: Examples of bioactive molecules with tetrahydroquinoline or tetrahydroisoquinoline scaffolds.[26–28] C: EDA-Mediated photochemical annulation of maleimides with N,N-dimethylaniline to yield a scaffold with a tetrahydroquinoline core.[21]

Fig. 2 Maleimide building blocks used in the synthesis of the epigenetic focused tetrahydroquinoline library.

Results and discussion

We envisaged that by introducing chemotypes from epigenetic ligands to the maleimide substrates it would be possible to prepare a library of compounds that target epigenetic reader, writer, or eraser protein domains. Therefore, three maleimides were synthesized from their respective primary amines *via* a condensation reaction with maleic anhydride or *cis*-5-norbornene-*endo*-2,3-dicarboxylic anhydride followed by a retro Diels-Alder reaction[29] to function as substrates for the photochemical annulation. Maleimides bearing a 3,5-dimethylisoxazole (**1**) to target bromodomains, an ethyl ester (**2**) that could be used as a precursor for the synthesis of other functionalized products, and an N-alkyl pyrrolidine (**3**) to target methyl-lysine reader domains were obtained in moderate yields (Fig. 2).

The maleimide substrates were then reacted with a range of N-methylanilines in the EDA-mediated α-alkylamino radical annulation to create the epigenetic focused compound collection (Scheme 1). A range of substituted anilines were employed to add extra utility to the compound collection. For example, 4-fluoro-N,N-dimethylaniline afforded the fluorinated products **6**, **12** and **18** that could be tested in ^{19}F ligand-observed NMR bioassays. Similarly, N-benzyl-N-methylaniline afforded products **5** and **11** that could be subsequently deprotected to reveal a secondary amine.

To further expand the range of epigenetic chemotypes found in the library and include other derivatizations, compound **5** was deprotected by hydrogenation to reveal the secondary aniline **19** in a good yield, which was decorated *via* two exemplar acylations (Scheme 2). An acetylation with acetyl chloride in dichloromethane afforded **20** bearing the N-acetyltetrahydroquinoline pharmacophore found in several bromodomain 4 (BRD4) ligands.[30] An amide coupling with 5-methylisoxazole-3-carboxylic acid using N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC) and dimethylaminopyridine (DMAP) in dichloromethane afforded compound **21**. Additionally, compound **16** was hydrolyzed using lithium hydroxide in tetrahydrofuran and water to afford the acid **22** and subsequently converted to the hydroxamic acid **23** yielding a putative HDAC inhibitor. Collectively, this yielded a small and focused collection of compounds bearing a range of privileged epigenetic chemotypes.

the synthesis of biologically relevant compound collections has not. We therefore sought to construct a bespoke collection of small molecules bearing chemotypes known to bind epigenetic targets based upon this tetrahydroquinoline scaffold (Fig. 1C).[21]
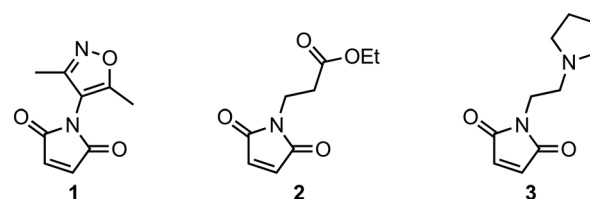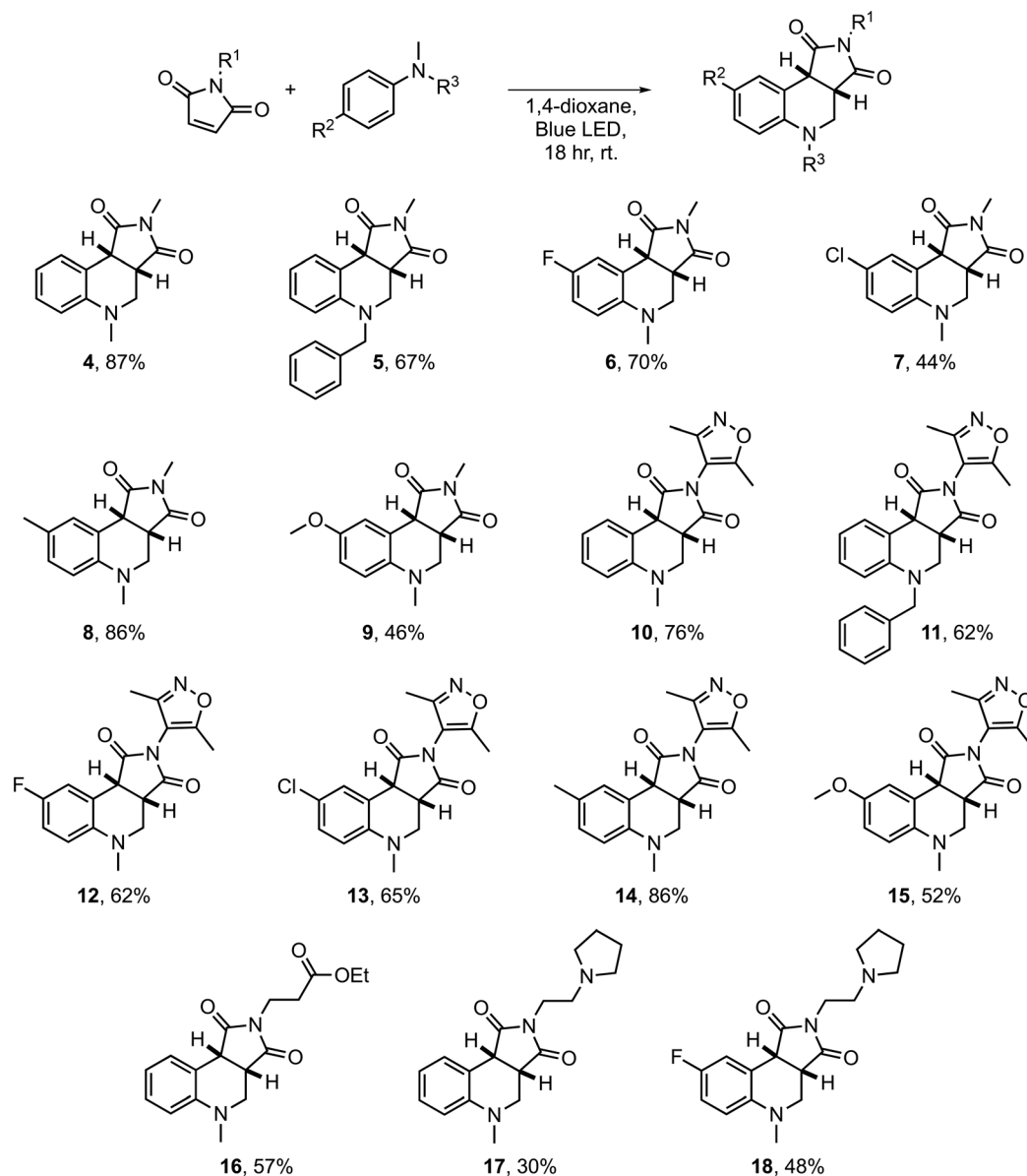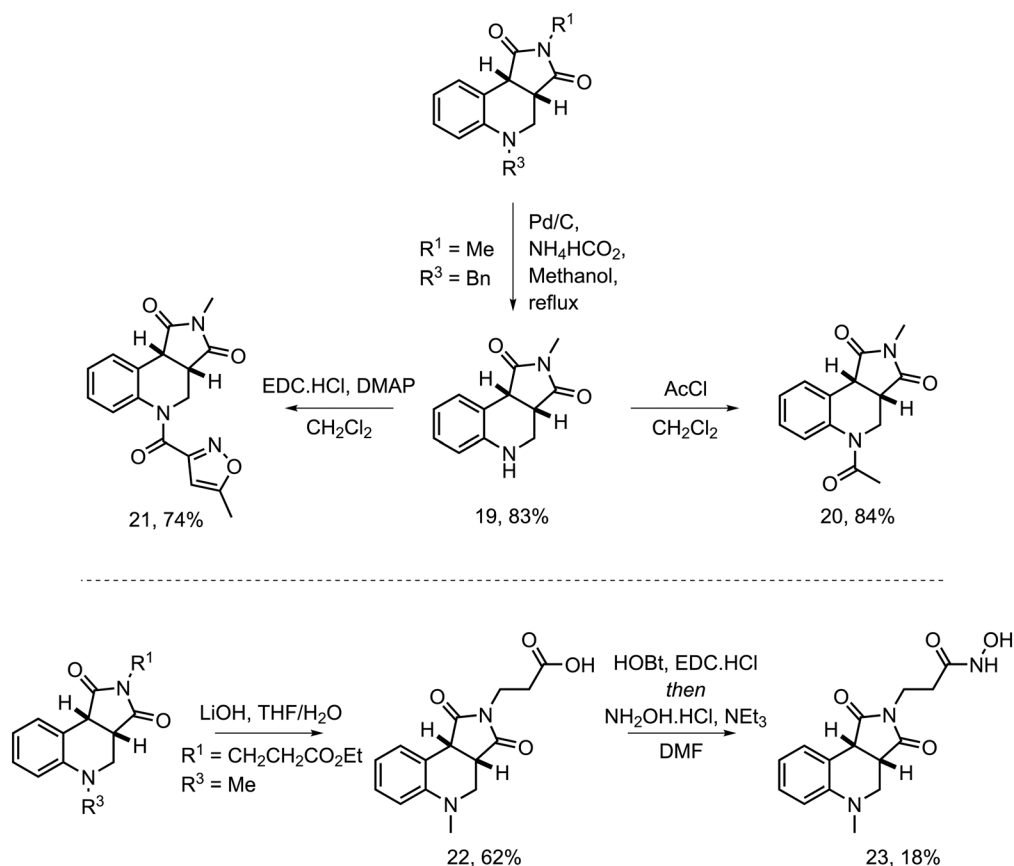
Scheme 1   Synthesis of the tetrahydroquinoline scaffold and epigenetic focused analogues.

Cheminformatics

To demonstrate the relevance and value of our epigenetic focused library, the molecular properties of the compounds were calculated (Fig. 3). Most of the compounds (19 out of 20, 95%) possess lead like properties of $A \log P < 3$ and molecular weight <350 Da, and all the molecules have $A \log P$ between 0 and 4 and weigh under 400 Da making them ideal starting points for optimization *via* medicinal chemistry.[31] Analysis of the principal moment of inertia (PMI) for the focused library shows that all the molecules are shifted from the rod-like region of the plot where many commercially available compounds reside, and some are shifted from the rod–disc axis, demonstrating promising 3D structures.

The tetrahydroquinolinedione core scaffold is also unique across several databases including ChEMBL, the NIH Molecular Libraries Small Molecule Repository (MLSMR), and a commercially available epigenetic library. The molecular fingerprint of compound **4** was compared by Tanimoto similarity to the same databases, where similar compounds typically have values ≥ 0.7.[32] Searching the MLSMR database identified 19 molecules with similarity ≥ 0.6 (0.004%) and no molecules ≥ 0.7 (approx. 440 000 molecules). Conducting the same analysis with a commercially available epigenetic library found no molecules with similarity ≥ 0.6 (approx. 10 000 molecules). Finally, a comparison of compound **10** to BRD4 ligands deposited to ChEMBL found only four molecules with similarity ≥ 0.6 (0.1%) and no molecules ≥ 0.7 (3832 molecules) demonstrating that unique molecules can still be obtained despite enriching our focused library with privileged chemotypes for epigenetic targets.

Scheme 2 Derivatizations of the tetrahydroquinoline scaffold.

While substructure searches and Tanimoto similarity ranking are useful tools for identifying molecules with new chemical structures, they do not quantify similarities in 3-dimensional shape. We conducted a Rapid Overlay of Chemical Structures (ROCS) analysis of compound **10** *versus*

BRD4 ligands deposited in ChEMBL and compound **17** *versus* lethal 3 malignant brain tumor-like protein 1 (L3MBTL1) ligands. An ensemble of conformers was generated in OMEGA for each ligand in ChEMBL and compared to the molecular shape of compound **10** or **17**. Generally, ROCS
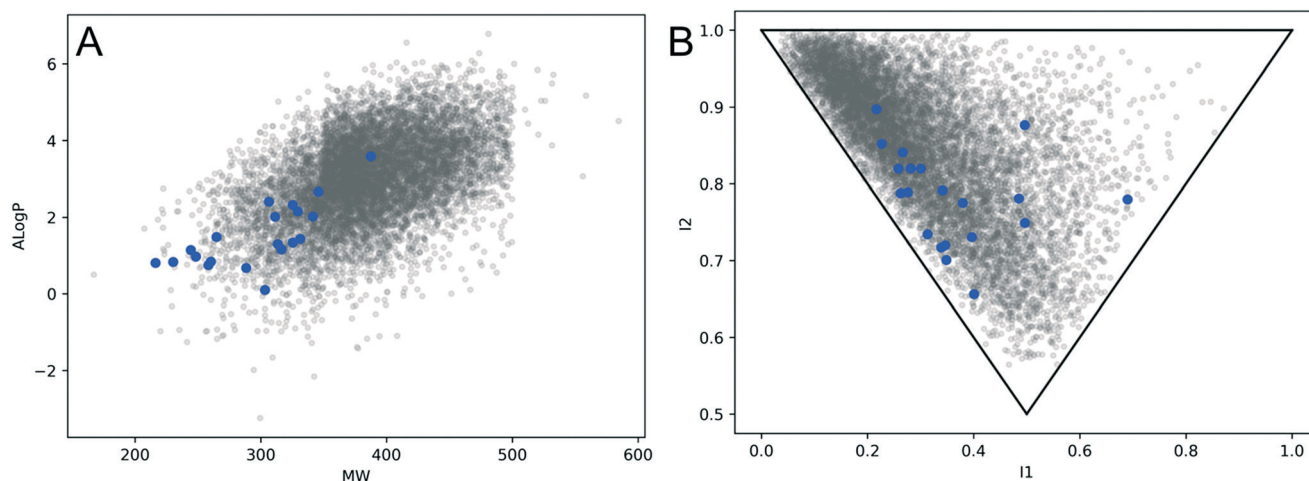

Fig. 3 Comparison of molecular properties for this compound collection (blue) and a commercially available epigenetic compound collection (grey). A: Molecular weight *vs.* *A* Log *P*. B: Principal moment of inertia plot. The commercially available epigenetic library is shown in grey and our focused library described is shown in blue. *A* Log *P* and principal moment of inertia (PMI) were calculated in RDKit. PMI represents the three-dimensional shape of a molecule where rod-like molecules are in the top left, disc-like molecules are at the lower middle, and sphere-like molecules are in the top right of the plot.
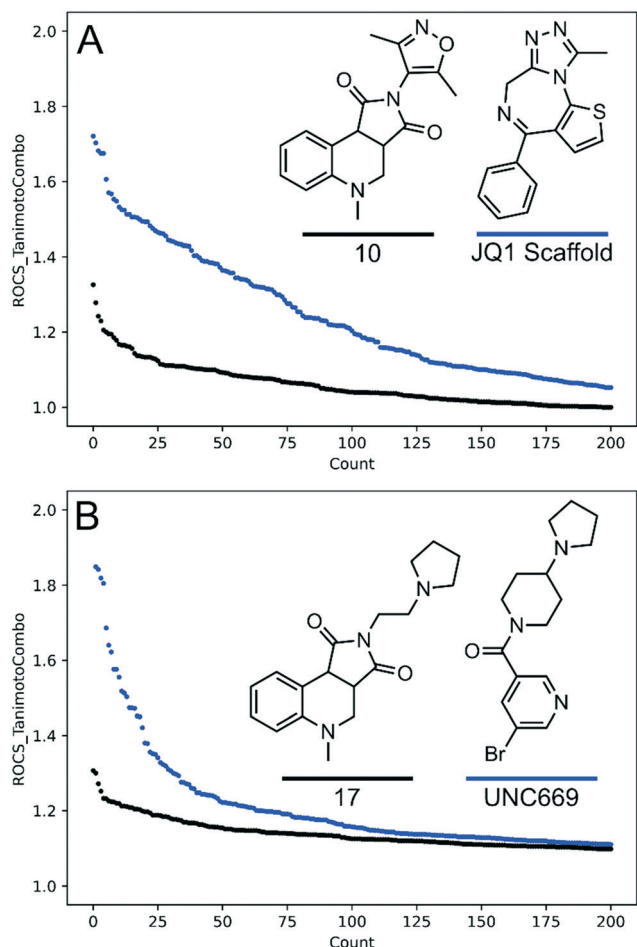
Fig. 4 Analysis of the top 200 ROCS ranked BRD4 or L3MBTL1 ligands in the ChEMBL database compared with exemplar compounds bearing privileged epigenetic chemotypes. Count refers to the ROCS ranking where **1** is the best ranked compound (most similar). A: Compound **10** (black) or the scaffold of JQ1 (blue) *vs.* BRD4 ligands deposited to ChEMBL (2168 molecules). B: Compound **17** (black) or UNC669 (blue) *vs.* L3MBTL1 ligands deposited to ChEMBL (10 847 molecules, UNC669 was removed from the test set).

scores greater than 1.4 indicate molecular shapes with significant similarity between the query molecule and the test compound.[32,33] Analysis for compound **10** reveals that there are no shape similar compounds within the BRD4 ligand set, and only six ligands were identified with ROCS scores ≥1.2 (Fig. 4A). Identical analysis of compound **17** also showed that no ligands with scores ≥1.4 were contained within the L3MBTL1 set, and 19 ligands had scores ≥1.2 (Fig. 4B). This demonstrates that these new compounds are distinct in shape from previously utilized chemical matter for exemplar drug targets containing acetyl- or methyl- lysine reader domains and may form new interactions with proteins, thus augmenting current screening collections.

To test whether shape similar compounds could be identified from these test sets we repeated the ROCS analysis using either the molecular scaffold of JQ1,[34] a potent BRD4 ligand, or the L3MBTL1 ligand UNC669[25] (Fig. 4). When using the JQ1 scaffold, 41 BRD4 ligands were identified with

ROCS scores ≥1.4 and 101 ligands had scores ≥1.2. Similarly, 22 L3MBTL1 ligands were identified with scores ≥1.4 and 66 with scores ≥1.2. This analysis supports the finding that there are essentially no ligands that are similar to the compounds contained within our focused library despite employing privileged chemotypes that are enriched within the BRD4 or L3MBTL1 ChEMBL sets. This demonstrates that it is possible to construct screening collections that are enriched with bioisosteres or privileged chemotypes for epigenetic targets without compromising the uniqueness of the compounds contained within the library. This design strategy therefore enables access to new highly relevant chemical space that has not been otherwise explored.

Conclusions

In summary, the strategy for the design and synthesis of a focused compound collection described here has the potential to address the low hit rates frequently observed in drug discovery campaigns against epigenetic targets and can be applied to a range of chemistries. Here we utilized a convergent synthetic route to generate a small exemplar library of compounds with a unique combination of scaffold and anchoring moiety. The ideal molecular properties of these compounds, and their location in underpopulated regions of chemical space, may facilitate the identification and optimization of ligands for previously undrugged epigenetic targets. Future development of the library will focus on the inclusion of new scaffolds, additional isosteres and evaluation of the hit rate in various assays to further validate this design approach. It may also be possible to facilitate the discovery of dual inhibitors of epigenetic targets using this approach through the inclusion of multiple isosteres in a single bifunctional molecule further broadening the utility of hits from this library.[35–37] We envisage that the chemo-informatically driven development of synthetic routes to novel scaffolds decorated with privileged isosteres will alleviate some of the issues with identifying epigenetic inhibitors.

Conflicts of interest

There is no conflict of interest to declare.

Acknowledgements

Academic License to support our research. We thank members of the Burslem lab for useful discussions.

References

1 J. P. Hughes, S. Rees, S. B. Kalindjian and K. L. Philpott, *Br. J. Pharmacol.*, 2011, **162**, 1239–1249.

2 D. M. Volochnyuk, S. V. Ryabukhin, Y. S. Moroz, O. Savych, A. Chuprina, D. Horvath, Y. Zabolotna, A. Varnek and D. B. Judd, *Drug Discovery Today*, 2019, **24**, 390–402.

3 D. G. Brown and J. Boström, *J. Med. Chem.*, 2018, **61**, 9442–9468.

4 A. I. Green and G. M. Burslem, *J. Med. Chem.*, 2021, **64**, 7231–7240.

5 J. J. Irwin, G. Gaskins, T. Sterling, M. M. Mysinger and M. J. Keiser, *J. Chem. Inf. Model.*, 2018, **58**, 148–164.

6 S. K. Verma, X. Tian, L. V. LaFrance, C. Duquenne, D. P. Suarez, K. A. Newlander, S. P. Romeril, J. L. Burgess, S. W. Grant, J. A. Brackley, A. P. Graves, D. A. Scherzer, A. Shu, C. Thompson, H. M. Ott, G. S. V. Aller, C. A. Machutta, E. Diaz, Y. Jiang, N. W. Johnson, S. D. Knight, R. G. Kruger, M. T. McCabe, D. Dhanak, P. J. Tummino, C. L. Creasy and W. H. Miller, *ACS Med. Chem. Lett.*, 2012, **3**, 1091–1096.

7 Z. Sessions, N. Sánchez-Cruz, F. D. Prieto-Martínez, V. M. Alves, H. P. Santos, E. Muratov, A. Tropsha and J. L. Medina-Franco, *Drug Discovery Today*, 2020, **25**, 2268–2276.

8 P. Czodrowski, G. Hölzemann, G. Barnickel, H. Greiner and D. Musil, *J. Med. Chem.*, 2015, **58**, 457–465.

9 M. Schiedel, M. Moroglu, D. M. H. Ascough, A. E. R. Chamberlain, J. J. A. G. Kamps, A. R. Sekirnik and S. J. Conway, *Angew. Chem., Int. Ed.*, 2019, **58**, 17930–17952.

10 A. Ganesan, P. B. Arimondo, M. G. Rots, C. Jeronimo and M. Berdasco, *Clin. Epigenet.*, 2019, **11**, 174.

11 J. N. Beyer, N. R. Raniszewski and G. M. Burslem, *ChemBioChem*, 2021, **22**, 17–42.

12 D. S. Hewings, M. Wang, M. Philpott, O. Fedorov, S. Uttarkar, P. Filippakopoulos, S. Picaud, C. Vuppusetty, B. Marsden, S. Knapp, S. J. Conway and T. D. Heightman, *J. Med. Chem.*, 2011, **54**, 6761–6770.

13 O. Fedorov, H. Lingard, C. Wells, O. P. Monteiro, S. Picaud, T. Keates, C. Yapp, M. Philpott, S. J. Martin, I. Felletar, B. D. Marsden, P. Filippakopoulos, S. Müller, S. Knapp and P. E. Brennan, *J. Med. Chem.*, 2014, **57**, 462–476.

14 J. M. Herold, L. I. James, V. K. Korboukh, C. Gao, K. E. Coil, D. J. Bua, J. L. Norris, D. B. Kireev, P. J. Brown, J. Jin, W. P. Janzen, O. Gozani and S. V. Frye, *MedChemComm*, 2012, **3**, 45–51.

15 D. C. Blakemore, L. Castro, I. Churcher, D. C. Rees, A. W. Thomas, D. M. Wilson and A. Wood, *Nat. Chem.*, 2018, **10**, 383–394.

16 G. Perumal, M. Kandasamy, B. Ganesan, K. Govindan, H. Sathya, M.-Y. Hung, G. Chandru Senadi, Y.-C. Wu and W.-Y. Lin, *Tetrahedron*, 2021, **80**, 131891.

17 K. Itoh, S. Nagao, K. Tokunaga, S. Hirayama, F. Karaki, T. Mizuguchi, K. Nagai, N. Sato, M. Suzuki, M. Hashimoto and H. Fujii, *Chem. – Eur. J.*, 2021, **27**, 5171–5179.

18 N. F. Nikitas, M. A. Theodoropoulou and C. G. Kokotos, *Eur. J. Org. Chem.*, 2021, **2021**, 1168–1173.

19 D. W. Manley, A. Mills, C. O'Rourke, A. M. Z. Slawin and J. C. Walton, *Chem. – Eur. J.*, 2014, **20**, 5492–5500.

20 ChEMBL, https://www.ebi.ac.uk/chembl/, (accessed 30 March 2021).

21 C.-W. Hsu and H. Sundén, *Org. Lett.*, 2018, **20**, 2051–2054.

22 G. E. M. Crisenza, D. Mazzarella and P. Melchiorre, *J. Am. Chem. Soc.*, 2020, **142**, 5461–5476.

23 J. Y. W. Mak, K.-C. Wu, P. K. Gupta, S. Barbero, M. G. McLaughlin, A. J. Lucke, J. Tng, J. Lim, Z. Loh, M. J. Sweet, R. C. Reid, L. Liu and D. P. Fairlie, *J. Med. Chem.*, 2021, **64**, 2186–2204.

24 O. Gilan, I. Rioja, K. Knezevic, M. J. Bell, M. M. Yeung, N. R. Harker, E. Y. N. Lam, C. Chung, P. Bamborough, M. Petretich, M. Urh, S. J. Atkinson, A. K. Bassil, E. J. Roberts, D. Vassiliadis, M. L. Burr, A. G. S. Preston, C. Wellaway, T. Werner, J. R. Gray, A.-M. Michon, T. Gobbetti, V. Kumar, P. E. Soden, A. Haynes, J. Vappiani, D. F. Tough, S. Taylor, S.-J. Dawson, M. Bantscheff, M. Lindon, G. Drewes, E. H. Demont, D. L. Daniels, P. Grandi, R. K. Prinjha and M. A. Dawson, *Science*, 2020, **368**, 387–394.

25 J. M. Herold, T. J. Wigle, J. L. Norris, R. Lam, V. K. Korboukh, C. Gao, L. A. Ingerman, D. B. Kireev, G. Senisterra, M. Vedadi, A. Tripathy, P. J. Brown, C. H. Arrowsmith, J. Jin, W. P. Janzen and S. V. Frye, *J. Med. Chem.*, 2011, **54**, 2504–2511.

26 P. J. Slavish, L. Chi, M.-K. Yun, L. Tsurkan, N. E. Martinez, B. Jonchere, S. C. Chai, M. Connelly, M. B. Waddell, S. Das, G. Neale, Z. Li, W. R. Shadrick, R. R. Olsen, K. W. Freeman, J. A. Low, J. E. Price, B. M. Young, N. Bharatham, V. A. Boyd, J. Yang, R. E. Lee, M. Morfouace, M. F. Roussel, T. Chen, D. Savic, R. K. Guy, S. W. White, A. A. Shelat and P. M. Potter, *Cancer Res.*, 2020, **80**, 3507–3518.

27 R. Kikumoto, Y. Tamao, T. Tezuka, S. Tonomura, H. Hara, K. Ninomiya, A. Hijikata and S. Okamoto, *Biochemistry*, 1984, **23**, 85–90.

28 P. Holzer, K. Masuya, P. Furet, J. Kallen, T. Valat-Stachyra, S. Ferretti, J. Berghausen, M. Bouisset-Leonard, N. Buschmann, C. Pissot-Soldermann, C. Rynn, S. Ruetz, S. Stutz, P. Chène, S. Jeay and F. Gessier, *J. Med. Chem.*, 2015, **58**, 6348–6358.

29 R. C. Clevenger and K. D. Turnbull, *Synth. Commun.*, 2000, **30**, 1379–1388.

30 M. Brand, A. M. Measures, B. G. Wilson, W. A. Cortopassi, R. Alexander, M. Höss, D. S. Hewings, T. P. C. Rooney, R. S. Paton and S. J. Conway, *ACS Chem. Biol.*, 2015, **10**, 22–39.

31 A. Nadin, C. Hattotuwagama and I. Churcher, *Angew. Chem., Int. Ed.*, 2012, **51**, 1114–1122.

32 G. Maggiora, M. Vogt, D. Stumpfe and J. Bajorath, *J. Med. Chem.*, 2014, **57**, 3186–3204.

33 Y. C. Martin, J. L. Kofron and L. M. Traphagen, *J. Med. Chem.*, 2002, **45**, 4350–4358.

34 P. Filippakopoulos, J. Qi, S. Picaud, Y. Shen, W. B. Smith, O. Fedorov, E. M. Morse, T. Keates, T. T. Hickman, I. Felletar,

M. Philpott, S. Munro, M. R. McKeown, Y. Wang, A. L. Christie, N. West, M. J. Cameron, B. Schwartz, T. D. Heightman, N. La Thangue, C. A. French, O. Wiest, A. L. Kung, S. Knapp and J. E. Bradner, *Nature*, 2010, **468**, 1067–1073.

35 J. P. Smalley, S. M. Cowley and J. T. Hodgkinson, *Molecules*, 2020, **25**, 4394.

36 J. Chen, Y. Li, J. Zhang, M. Zhang, A. Wei, H. Liu, Z. Xie, W. Ren, W. Duan, Z. Zhang, A. Shen and Y. Hu, *Eur. J. Med. Chem.*, 2021, **209**, 112868.

37 S. J. Atkinson, P. E. Soden, D. C. Angell, M. Bantscheff, C. Chung, K. A. Giblin, N. Smithers, R. C. Furze, L. Gordon, G. Drewes, I. Rioja, J. Witherington, N. J. Parr and R. K. Prinjha, *MedChemComm*, 2014, **5**, 342–351.