Molecular Diversity

Synthesis and Demonstration of the Biological Relevance of sp$^3$-rich Scaffolds Distantly Related to Natural Product Frameworks

Daniel J. Foley,[a, b] Philip G. E. Craven,[a, b] Patrick M. Collins,[c] Richard G. Doveston,[a, b] Anthony Aimon,[a, b] Romain Talon,[d] Ian Churcher,[e] Frank von Delft,[c, d] Stephen P. Marsden,[b] and Adam Nelson*[a, b]

Abstract: The productive exploration of chemical space is an enduring challenge in chemical biology and medicinal chemistry. Natural products are biologically relevant, and their frameworks have facilitated chemical tool and drug discovery. A “top-down” synthetic approach is described that enabled a range of complex bridged intermediates to be converted with high step efficiency into 26 diverse sp$^3$-rich scaffolds. The scaffolds have local natural product-like features, but are only distantly related to specific natural product frameworks. To assess biological relevance, a set of 52 fragments was prepared, and screened by high-throughput crystallography against three targets from two protein families (ATAD2, BRD1 and JMJ2D). In each case, 3D fragment hits were identified that would serve as distinctive starting points for ligand discovery. This demonstrates that frameworks that are distantly related to natural products can facilitate discovery of new biologically relevant regions within chemical space.

Introduction

Small molecules continue to dominate Man’s ability to treat disease, and can transform our understanding of fundamental biology. Yet historically, the exploration of chemical space has been highly uneven,[1, 2] in part because a narrow toolkit of reliable reactions has underpinned molecular discovery.[3] Natural products can facilitate the identification of biologically relevant chemical space[4] since they have arisen through the function-driven evolution of biosynthetic pathways.[5] Indeed, around a third of recent small molecule drugs have been inspired by natural products.[6] In biology-oriented synthesis,[7] natural product frameworks[8] inform the design of productive small molecule screening collections and fragment sets that can be exploited in the discovery of ligands for unrelated protein targets.[8, 9] In addition, synthetic approaches have been developed to convert specific natural products into alternative complex frameworks.[10] Natural product-inspired compounds can provide highly distinctive starting points for discovery that contrast starkly with most synthetic screening compounds.[11] In particular, the typically high fraction of sp$^3$-hybridised carbons (Fsp$^3$) is attractive since it correlates with the successful translation of clinical candidates.[12]

We envisaged a “top-down” synthetic approach in which alternative complex, yet readily accessible, intermediates would be converted into many natural product-like scaffolds (Scheme 1). Initially, bridged scaffolds of general structure 2 would be prepared using intramolecular [5+2] cycloaddition reactions[13, 14] (e.g. 1 → 2; ring cleavage (red); for example, 2 → 3), ring expansion (magenta; for example, 2 → 5), annulation (blue; for example, 2 → 5), or functional group modification (green; for example, 2 → 6) would then yield diverse scaffolds. The approach contrasts with diversity-oriented strategies[15, 16] in which building blocks are prepared (“built”) and linked (“coupled”) to yield intermediates which are then cyclised (“paired”) to yield alternative scaffolds. Although diversity-oriented approaches to sp$^3$-rich fragments have been developed,[17] their biological relevance has only rarely[17b] been demonstrated. To undertake a preliminary assessment of biological relevance, we have therefore prepared and screened a...
The complex bridged intermediates 2 were transformed into diverse molecular scaffolds (Scheme 3 and Supporting Information). Cleavage of specific bonds enabled scaffold simplification. Specifically, we exploited 1,2-diol cleavage (e.g.—the fused bicyclic scaffold 7); alkene ozonolysis (e.g.—the spirofused scaffold 3); and reductive allylic ether cleavage (→ the fused bicyclic scaffold 12). Interception of dialdehydes formed by 1,2-diol cleavage enabled overall ring expansion. Thus, cleavage of the regioisomeric 1,2-diols formed from 2a and 2c, followed by cyclative double reductive amination, yielded the tricyclic scaffolds 4 and 15, respectively.

Annulation enabled more complex scaffolds to be prepared. For example, fusion of alternative heterocycles was possible either by condensation–aromatisation of the masked diketone of 2a, 2b or 2c (e.g.—the pyrrole 9, the quinoxalines 10 and 14 or the imidazole 11); or regioselective [3+2] annulation to the enone 2d (→ the pyrrole 17). Alternatively, treatment of the regioisomeric intermediates 2a and 2c with MeLi effected 1,2-addition; reductive amination and oxazolidinone formation then gave the regioisomeric tetracycles 8 and 16. Furthermore, intramolecular reductive Heck reaction of 2e gave the complex tetracycle 19. Finally, substituted analogues were prepared in which the framework of the initial cycloadduct had been retained (e.g. 13). Overall, the 26 scaffolds were prepared from commercially available starting materials in a total of just 64 steps (processes conducted in a single reaction vessel).

An hierarchical tree was constructed to capture the relationship between the scaffolds (Figure 1). Here, scaffolds were systematically simplified using an established protocol by removal of rings until a parent monocyclic ring system was ultimately obtained.[21] Twenty two different graph-node-bond frameworks (capturing atom and bond type) were represented, which was then simplified to give nine parent monocycles. The scaffolds are based on a wide range of parent ring systems and there is thus significant diversity at each level of hierarchy of the scaffold tree. The exploitation of several different complex intermediates 2 was critical to the diversity that was possible, for example by enabling variation of regiochemistry (e.g. 4/15, 8/16 and 10/14) and heteroatom position and identity (e.g. 11/17). Such an approach would unlikely be possible by modification of natural products, since several related starting materials would be required in multi-gram quantities.

To compare with other screening sets, we determined natural product likeness scores[11] for the deprotected scaffolds, a natural product screening library (4,460 compounds) and a commercial screening collection (278,365 largely synthetic compounds) (Figure 2, Panel A). The distribution of the scores for the scaffolds was broadly similar to that of the natural product screening library but highly distinctive from that of the large screening collection. The local structural features of our scaffolds are thus reminiscent of those found in natural products. Despite the high natural product likeness, however, only one of the 22 graph-node-bond frameworks is actually a substructure in the roughly 281,000 compounds in the Dictionary of Natural Products[20] (Figure 1 and Supporting Information). Indeed, significant simplification to mono- or bicyclic

---

**Results and Discussion**

### Synthesis of natural product-like scaffolds

The bridged scaffolds 2a–f were prepared using intramolecular [5+2] cycloaddition chemistry (Scheme 2). The oxygen-bridged cycloadducts were prepared by silyl transfer-induced [5+2] cycloaddition of 3-oxidopyrrolidinones (→ 2a–c).[14] In a similar vein, the nitrogen-bridged cycloadducts were prepared by intramolecular cycloaddition of 3-oxidopyrrolidinyl ylides (→ 2d–f).[16]
frameworks is required before sub-structures of natural products are found. Our scaffolds thus have high natural product likeness, but their frameworks lie in branches that augment the scaffold trees of natural product frameworks.

**Synthesis of a fragment set and high-throughput crystallographic screens**

To enable preliminary assessment of biological relevance, we prepared 52 racemic fragments based on 23 of the scaffolds. Here, a fragment-based approach was exploited to enable efficient exploration of chemical space accessible using our synthetic approach. The set was designed to have high shaped diversity, and to comprise fragments with controlled size (13 to 19 heavy atoms) and lipophilicity ($\log P < 3$) (Supporting Information). The fragment set was significantly more three-dimensional (Supporting Information), and more natural product-like (Figure 2, Panel B), than commercially available fragments with the same heavy atom range.

The fragment set was screened against three protein targets from two different mechanistic classes involved in epigenetic biology: the ATAD2 and BRD1 (also known as BRPF2) bromodomains, and the histone demethylase JMJD2D (also known as KDM4D). Here, the objective was to perform a preliminary assessment of biological relevance rather than to provide specific starting points for discovery. High expression levels of ATAD2[25] and JMJD2 family members[26] correlate with poor outcomes in several cancers, whilst BRD1 is a member of the BRPF family of scaffolding proteins whose role in acute myeloid leukemia is now emerging.[27] The two bromodomains are contrasting targets: ATAD2 has a shallow N-acetyl lysine binding site and has been suggested to have particularly low druggability.[28] The three target proteins were all amenable to high-throughput protein crystallography. Protein crystals were soaked with the 52 individual racemic fragments, picked and then subjected to automated X-ray diffraction. Fragment hits were identified through detection of additional electron density, and inspection for polar interactions with the protein. Fragment hits were successfully identified for each of the target proteins: two hits for JMJD2D, eight hits for the BRD1 bromodomain and seven hits for the ATAD2 bromodomain (Figure 3 and Supporting Information).

The fragment screen against JMJD2D revealed two hits, both of which targeted a peripheral binding site (Panel B1, Figure 3 and Supporting Information). The hits complement those found in previous fragment screens against JMJD2D,[26,31] specifically, X-ray crystallography had revealed 23 fragments all target the enzyme active site. The functional importance of this peripheral binding site could now be investigated, for ex-

---

**Scheme 3.** Representative syntheses of natural product-like scaffolds. Scaffolds were prepared from cycloadducts by ring cleavage (red), ring expansion (magenta), ring formation (blue) or substitution (green). Typical conditions (see Supporting Information for full details): (a) NaBH₄ then CSA (from 2a: 77%; from 2c: 34%); (b) NaIO₄ then NaBH₄, 44%; (c) NaIO₄ then BN₃H₆, NaBH(OAc)₃, 32%; (d) MeLi (from 2a: 91%; from 2c: 53%); (e) NH₃, Ti(OiPr)₄, NaBH₄, 77%; (f) MeOCOCI, Et₂N, 63%; (g) TBAF then NaH, 29%; (h) ethylene diamine, AcOH, 180 °C, 40%; (i) 1,2-diaminoobenzene, AcOH, heat (from 2b: 21% from 2c: 32%); (j) NH₃·OAc, paraformaldehyde, 60 °C, 38%; (k) NaBH₄ then 2,2-dimethoxypropane, p-TsOH, 60 °C, 71%; (l) O₃ then Me,S then NaBH₄, 44%; (m) LiAlH₄, THF, Δ, 75%; (n) NH₃·Ti(OiPr)₄, NaBH₄, 22%; (o) NaIO₄ then DMBNH₂, NaBH(OAc)₃, 30%; (p) MeOCOCI, Et₂N then TBAF then NaH, 34%; (q) EN₃O₂, PPh₃, NET₃, thenq DDQ, (from 2d: 37% from 2f: 26%); (r) NaBH₄, CeCl₃, 7H₂O, −78 °C, 87%; (s) H₂, Pd(OH)₂/C, HCl, 89%; (t) 20 mol% Pd(OAc)₂, 40 mol% PPh₃, NET₃, 11%; (u) PhB(OH)₂, NET₃, 1 mol% [Rh(cod)Cl]₂, then H₂, Pd(OH)₂/C, HCl, 46%; (v) NaBH₄, 19% (plus 19% epimer).
ample to reveal opportunities for allosteric modulation of the enzyme.

For the BRD1 bromodomain, the eight fragment hits target-
ed the N-acetyl lysine binding site (Panel B2, Figure 3 and Sup-
porting Information; see Ref. [27] and references therein for
known ligands). The fragment hits contained several different
N-acetyl mimetics that interacted with both N110 and, either
directly or via a bridging water, Y67.

For the ATAD2 bromodomain, the fragment screen yielded
seven hits (based on six distinct frameworks) that targeted the
N-acetyl lysine binding site (Panel B3, Figure 3 and Supporting
Information). Six of the hits mirrored the binding mode of the
N-acetyl lysine side chain, interacting directly with N1064
and, via a bridging water, with Y1021. Four hits were in
common with BRD1 although, in three cases, the interaction
networks were more extensive: for example ent-23 made
water-mediated interactions with V1008 and D1014 (Panel B3,
top, Figure 3) and ent-25 interacted directly with E1017
(Figure 4).

To compare directly with a more conventional fragment set,
we also screened 700 commercially available fragments against
ATAD2 by high-throughput crystallography and identified nine
hits that targeted N-acetyl lysine binding site (Figure 4 and
Supporting Information). As a group, the interactions of these
fragments parallel those of other sp²-rich fragments: they inter-
act directly with N1064 and/or, via a bridging water, with
Y1021 but make few additional polar contacts. For screens against many targets, flatter fragments have been ob-
served to have higher hit rates. With ATAD2, however, a sig-
nificantly higher hit rate was observed with our shape-diverse
natural product-like fragments (7/52) than with more conven-
tional flatter fragment sets (9/700). This outcome is consis-
tent with the low hit rate observed in a previous fragment
screen by NMR (65/13800, subsequently triaged to yield 12
hits with $K_d < 1 \text{ mm}$).
We have developed a "top-down" synthetic approach in which alternative complex, yet readily accessible, intermediates were converted into many diverse scaffolds. These scaffolds have local natural product-like features, but are only distantly related to specific natural product frameworks. As a set of 52 fragments based on 23 of the scaffolds was screened against three epigenetic targets from two distinct protein families. In each case, hits were obtained that may provide distinctive opportunities for subsequent fragment growth. We have therefore demonstrated that frameworks that are distantly related to natural products can facilitate identification of novel regions of biologically relevant chemical space. Synthetic approaches to such frameworks may thus help identify fertile chemical space for bioactive small molecule discovery that is inaccessible to existing compound collections and biosynthetic pathways.

Acknowledgements

We thank EPSRC for a Doctoral Prize Fellowship (to DJF) and grants (EP/J00894X; EP/N025652/1); GSK for funding; Aleksandra Szykowska and Michael Fairhead for protein production and purification; and support from the Innovative Medicines Initiative Joint Undertaking under grant number 115489, resources of which are composed of financial contribution from the European Union’s Seventh Framework Programme (F97/2008–2013), EFPIA companies’ in kind contribution. The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFP) [ULTRA-DD grant no. 115766], Janssen, Merck & Co., Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome Trust [092809/Z/10/Z].

Conflict of interest

The authors declare no conflict of interest.
[8] B. Over, S. Wetzl, C. Grütter, Y. Nakai, S. Renner, D. Rauh, H. Waldmann, Nat. Chem. 2013, 5, 21 – 28.

[9] T. Rodríguez, D. Reker, J. Kunze, P. Schneider, G. Schneider, Angew. Chem. Int. Ed. 2015, 54, 10516 – 10520; Angew. Chem. 2015, 127, 10662 – 10666.

[10] a) R. W. Huigens III, K. C. Morrison, R. W. Hicklin, T. A. Flood Jr., M. F. Richter, P. J. Hergenrother, Nat. Chem. 2013, 5, 195 – 202; b) R. J. Rafferty, R. W. Hicklin, K. A. Maloof, P. J. Hergenrother, Angew. Chem. Int. Ed. 2014, 53, 220 – 224; Angew. Chem. 2014, 126, 224 – 228; c) A. Masarwa, M. Weber, R. Sarpong, J. Am. Chem. Soc. 2015, 137, 6327 – 6334; d) M. Weber, K. Owens, A. Masarwa, R. Sarpong, Org. Lett. 2015, 17, 5432 – 5435.

[11] P. Ertl, S. Roggo, A. Schuffenhauer, J. Chem. Inf. Model. 2008, 48, 68 – 74. Implemented in RDKit v2015.09.2 (Greg Landrum; Open Source Cheminformatics; http://www.rdkit.org; last accessed 27-Apr-2016).

[12] F. Lovering, J. Bikker, C. Humblet, J. Med. Chem. 2009, 52, 6752 – 6756.

[13] J. L. Mascareñas, in Advances in Cycloadditions, Volume 6, pp. 1 – 54 (Ed. Harmata, M.) JAI Press Inc., Stamford CT, 1999.

[14] J. L. Mascareñas, A. Rumbo, L. A. Castedo, J. Org. Chem. 1997, 62, 8620 – 8621.

[15] T. Nielsen, S. L. Schreiber, Angew. Chem. Int. Ed. 2008, 47, 48 – 56; Angew. Chem. 2008, 120, 52 – 61.

[16] For key reviews, see: a) Diversity-Oriented Synthesis: Basics and Applications in Organic Synthesis Discovery and Chemical Biology (Ed. Trabocchi, A.) Wiley, Hoboken NJ, 2013; b) W. R. J. D. Galloway, A. Isidro-Llobet, D. R. Spring, Nat. Commun. 2010, 1, 80; c) M. D. Burke, S. L. Schreiber, Angew. Chem. Int. Ed. 2004, 43, 46 – 58; Angew. Chem. 2004, 116, 48 – 60.

[17] a) A. W. Hung, A. Ramek, Y. Wang, T. Kaya, J. A. Wilson, P. A. Clemons, D. W. Young, Proc. Natl. Acad. Sci. USA 2011, 108, 6799 – 6804; b) Y. Wang, ACS Med. Chem. Lett. 2016, 7, 852 – 856; c) D. G. Twigg, N. Kondo, S. L. Mitchell, W. R. J. D. Galloway, H. F. Sore, A. Madin, D. R. Spring, Angew. Chem. Int. Ed. 2016, 55, 12479 – 12483; Angew. Chem. 2016, 128, 12667 – 12671.

[18] A. Rumbo, A. Mourinho, L. Castedo, J. L. Mascareñas, J. Org. Chem. 1996, 61, 6114 – 6120.

[19] A. Amorese, A. Arcadi, E. Bernocchi, S. Cacchi, S. Cerrini, W. Fedeli, G. Ortar, Tetrahedron 1989, 45, 813 – 828.

[20] Pairs of diastereomers were prepared for two scaffolds.

[21] a) S. Wetzl, K. Klein, S. Renner, D. Rauh, T. I. Oprea, P. Mutzel, H. Waldmann, Nat. Chem. Biol. 2009, 5, 581 – 583; b) S. Renner, Nat. Chem. Biol. 2009, 5, 585 – 592.

[22] Dictionary of Natural Products Online, http://dnp.chemnetbase.com/ intro/ (Accessed: 2016).

[23] M. Congreve, R. Carr, C. Muray, H. A. Jhoti, Drug Discovery Today 2003, 8, 876 – 877.

[24] I. Colomer, C. J. Empson, P. Craven, Z. Owen, R. G. Doveston, I. Churcher, S. P. Marsden, A. Nelson, Chem. Commun. 2016, 52, 7209 – 7212. LLAMA is available at https://llama.leeds.ac.uk.

[25] P. Bamborough, Angew. Chem. Int. Ed. 2016, 55, 11382 – 11386; Angew. Chem. 2016, 128, 11554 – 11558.

[26] M. Kocznyska, J. Med. Chem. 2016, 59, 1580 – 1598.

[27] N. Igoe, J. Med. Chem. 2017, 60, 668 – 680.

[28] a) A. Chakikad, A. M. Petros, O. Fedorov, J. Xu, S. Knapp, MedChemComm 2014, 5, 1843 – 1848; b) M. J. Harner, B. A. Chauder, J. Phan, S. W. Fesik, J. Med. Chem. 2014, 57, 9687 – 9692.

[29] P. M. Collins, J. T. Ng, R. Talon, K. Nekrosiute, T. Krojer, A. Douangamath, J. Brandao-Neto, N. Wright, N. M. Pearce, F. von Delft, Acta Crystallogr. Sect. D 2017, 73, 246 – 255.

[30] a) N. M. Pearce, A. R. Bradley, T. Krojer, B. D. Marsden, C. M. Deane, F. von Delft, Struct. Dyn. 2017, 4, 032104; b) N. M. Pearce, Nat. Commun. 2017, 8, 15123.

[31] a) V. Bavetsias, J. Med. Chem. 2016, 59, 1388 – 1409; b) S. M. Westaway, J. Med. Chem. 2016, 59, 1357 – 1369.

[32] a) E. H. Demont, J. Med. Chem. 2015, 58, 5649 – 5673; b) G. Poncet-Montange, Biochem. J. 2015, 466, 337 – 346.

[33] R. J. Hall, P. N. Mortenson, C. W. Murray, Prog. Biophys. Mol. Biol. 2014, 116, 82 – 91.

[34] A. D. Morley, Drug Discovery Today 2013, 18, 1221 – 1227.

Manuscript received: September 6, 2017
Version of record online: October 6, 2017