COMMUNICATION

A biosynthesis-inspired approach to over twenty diverse natural product-like scaffolds†

James D. Firth,[a] Philip G. E. Craven,[a] Matthew Lilburn,[a] Axel Pahl,[b] Stephen P. Marsden*[a] and Adam Nelson*[ac]

A synthetic approach to diverse scaffolds was developed that was inspired by diterpene biosynthesis. Initial scaffolds, generated using Diels–Alder reactions of furyl-functionalised amines, were transformed into alternative scaffolds using cleavage, ring expansion, annulation and rearrangement reactions. In total, 25 diverse scaffolds were prepared that were shown to have high natural product-likeness.

The exploration of biologically-relevant chemical space is an enduring challenge in both medicinal chemistry and chemical biology. Natural products arise through the evolution of biosynthetic pathways driven by functional benefit to the host organism,[1] In biology-oriented synthesis,[2] this pre-validated relevance[3] is exploited in the design of productive small molecules informed by the frameworks of natural products.[4] The structural features of natural products are highly distinctive, and may be captured in a natural product-likeness score.[5] In particular, a high fraction of sp3-hybridised carbons[6] (Fsp3) is an attractive feature because it correlates strongly with the successful translation of clinical candidates.[7] Indeed, about a third of small molecule drugs approved between 1981–2010 were inspired by natural products.[8]

We envisaged a synthetic approach to diverse natural product-like scaffolds that was broadly inspired by diterpene biosynthesis (Scheme 1).[9] The bi- and polycyclic scaffolds of labdanes and clerodanes are formed in two steps by diterpene synthases (diTCs). Class II diTCs catalyse the initial cyclisation of geranylgeranyl diphosphate (GGPP) to give bicyclic diphosphate intermediates

(Panel A); subsequent enzyme- (*e.g.* class I diTC) catalysed reactions then yield the final diterpene scaffolds. Inspired by this two-stage synthesis of natural product scaffolds,[10] we envisaged that a range of initial scaffolds might be analogously prepared[11] by reacting furyl-substituted amines with maleic anhydride (Panel B). The initial scaffolds would then be transformed into product scaffolds using a suite of cleavage, ring expansion, annulation and rearrangement reactions. Thus, analogously to diterpene biosynthesis, sp2-rich starting materials would be transformed into diverse, sp3-rich polycyclic scaffolds.

To start with, a range of substituted furyl-substituted amines (ESI†) was reacted with maleic anhydride to give the corresponding cycloadducts **2a–k** in moderate to excellent yield (Scheme 2). For example, DMB-protected (2-furyl)-amine **1** yielded the tricyclic adduct **2a** in 90% yield as a single diastereomer. Whilst syntheses of similar γ-lactams are known,[11] we were delighted that the synthesis of the related[12] δ-lactams **2b** and **2c** was also efficient. Substitution of the 4-[11c] or 5-[11b] position of the furan ring was well tolerated (→ **2d** and **2e** respectively). Substitution α to nitrogen[11d,e] was also tolerated but the diastereoselectivity was poor: however, after esterification, the diastereomeric cycloadducts (**2f**/*epi*-**2f** and **2g**/*epi*-**2f**) could be readily separated.

The fusion of addition rings to the furyl-substituted amine starting material was possible in several contexts. For example, the tetracyclic scaffolds **2h** and **2i** were obtained in 90% and 70% yield respectively (the latter notably as a single diastereomer). In the case of 7-(2-furyl)-1,4-diazepane, esterification and hydrogenation (to prevent retro-cycloaddition) gave the diastereomeric tetracyclic scaffolds **2j** and *epi*-**2j** in 44% and 17% yield respectively. The reaction of a 2-(2-furyl)-pyrrolidine was, however, more problematic: reaction with maleic anhydride gave an open-chain adduct which underwent (reversible) cycloaddition only after esterification. After subsequent hydrogenation of the thermodynamic mixture of products, the scaffold **2k** was obtained in 48% overall yield. The relative configuration of the cycloadducts **2f**, *epi*-**2f**, *epi*-**2g 2j**, *epi*-**2j** and **2k** was determined by X-ray crystallography (Fig. 2 and ESI†).

*a School of Chemistry, University of Leeds, Leeds LS2 9JT, UK.*
  E-mail: a.s.nelson@leeds.ac.uk, s.p.marsden@leeds.ac.uk
b Max Planck Institute of Molecular Physiology, Otto-Hahn Strasse 11, 44227
 *Dortmund, Germany*
c Astbury Centre for Structural Molecular Biology, University of Leeds,
  Leeds LS2 9JT, UK
† Electronic supplementary information (ESI) available: Experimental procedures, characterisation data, computational analyses and crystallographic data in CIF or other electronic forms. CCDC 1478957–1478963. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c6cc04662b
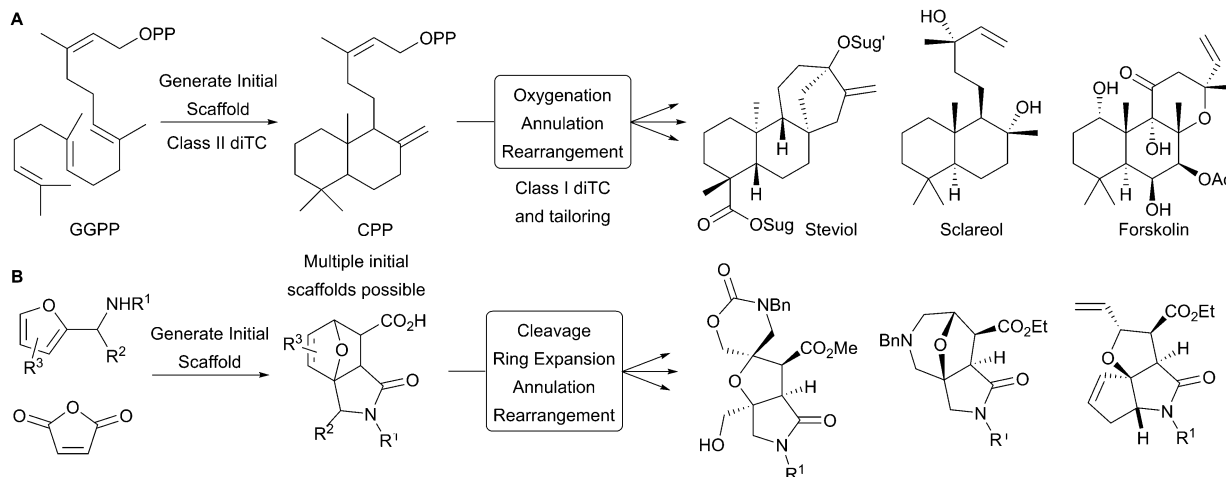
Scheme 1 Pathways to natural products and natural product-like scaffolds. Panel A: class II diTCs catalyse the cyclisation of GGPP into alternative initial scaffolds (e.g. CPP); diterpene scaffolds are formed in subsequent enzyme- (e.g. class I diTC) catalysed reactions. Panel B: envisaged synthetic strategy in which the generation of alternative initial scaffolds would be followed by formation of natural product-like scaffolds. CPP, copalyl diphosphate. Sug and Sug' denote glycosyl groups.
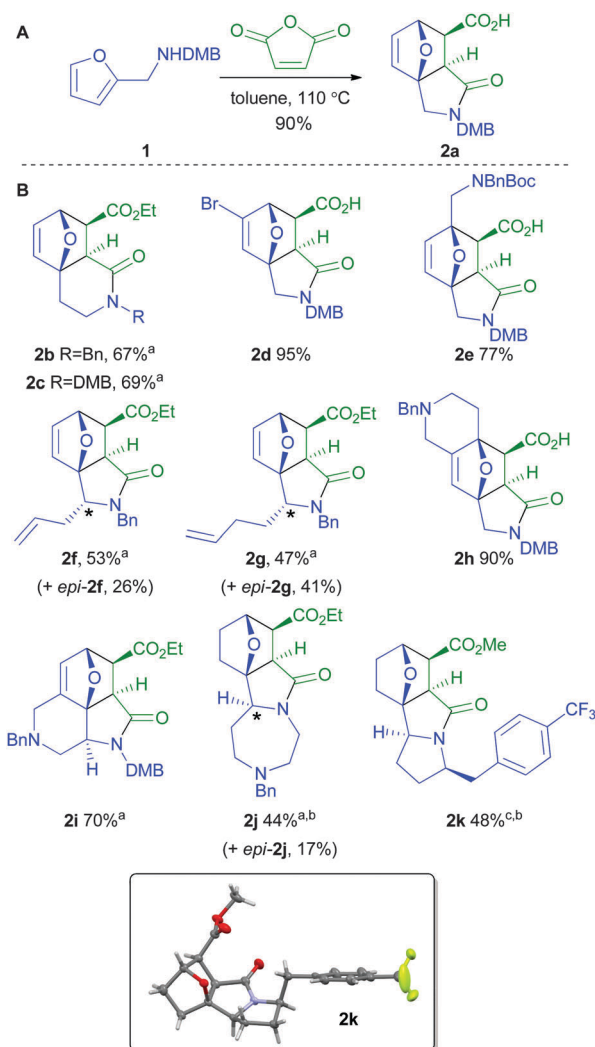
Scheme 2 Synthesis of alternative initial scaffolds. Panel A: synthesis of an exemplar initial scaffold. Panel B: additional initial scaffolds prepared. Deviations from standard conditions: [a] then conc. HCl$_{(aq)}$, EtOH, 78 °C; [b] then H$_2$, 10% Pd/C, EtOH; [c] MeCN, rt then HCl, MeOH.

We established a suite of methods to transform the cycloadducts into a wide range of molecular scaffolds (Scheme 3). Initially, we focused on reactions that would cleave (and hence simplify) the initial scaffold (Panel A). Cleavage of the dihydrofuran ring of **3** (from **2a**; ESI†) was possible by dihydroxylation of the alkene and treatment with periodate: subsequent reaction with sodium borohydride gave the fused bicyclic scaffold **4** in 75% yield. Crucially, the approach was effective with other initial scaffolds, to give the related bi- and tricyclic scaffolds **5–8**; remarkably, the scaffold **6** was obtained as a single diastereomer, albeit in only 20% yield due to a difficult purification.

Next, we focused on formal ring expansions of the initial scaffolds (Panel B).[11f] After dihydroxylation and oxidative cleavage of **3** and **2c**, it was possible to intercept the resulting dialdehydes: thus, treatment with benzylamine and sodium triacetoxyboro-hydride gave the ring-expanded analogues **9** and **10** in 82% and 62% yield respectively.

Annulation allowed access to more structurally complex scaffolds (Panel C). For example, Boc-deprotection of **2e**, followed by TBTU-mediated γ-lactamisation, gave the tetracyclic scaffold **11** in 87% yield. In addition, Boc-deprotection of **8**, and concomitant esterification, was followed by treatment with CDI to give the spirocyclic scaffold **12** in 97% yield.

We harnessed alkene metathesis to induce rearrangement of initial scaffolds (Panel D).[13] Treatment of the diene **2f** with 5 mol% Hoveyda–Grubbs second generation catalyst in TBME gave the tricyclic scaffold **13** in 31% yield. In a similar vein, the dienes **2g** and *epi*-**2g** were converted into the related tricycles **14** and *epi*-**14**. The relative configuration of *epi*-**14** was determined by X-ray crystallography.

Finally, we employed functional group interconversions to modify the substitution of two scaffolds (Scheme 4). These transformations may be considered analogous to steps (e.g. hydroxylation; esterification) that tailor the substitution of diterpene scaffolds. Esterification of **2a**, followed by hydrogenation and LiAlH$_4$ reduction, gave the hydroxymethyl-substituted scaffold **15**. Mitsunobu reaction of
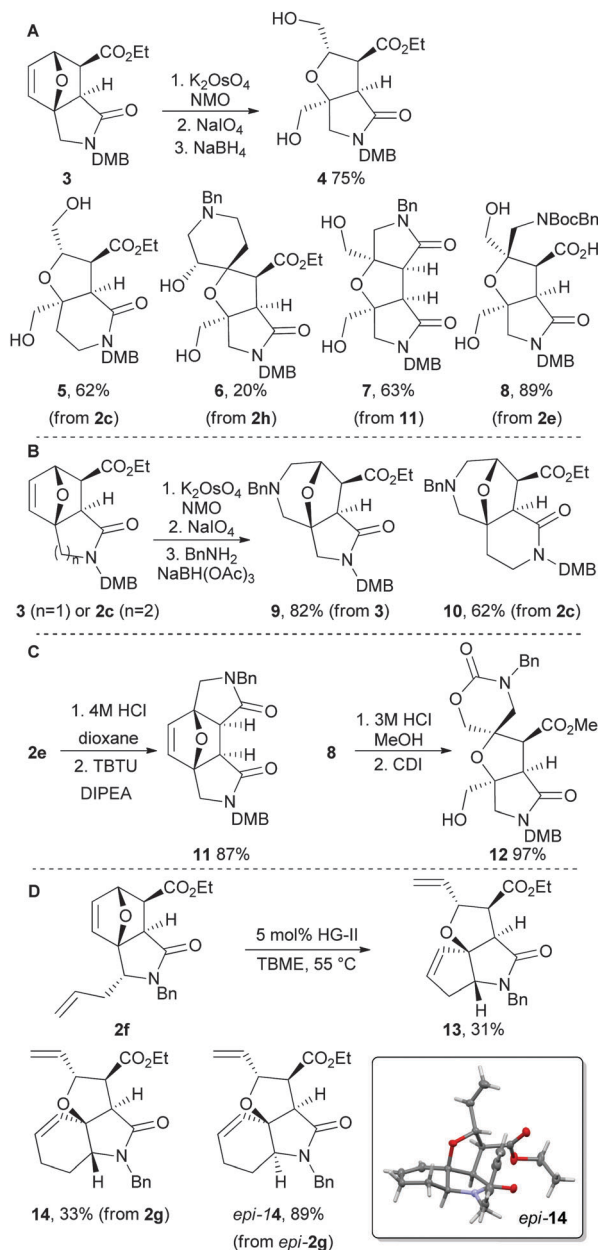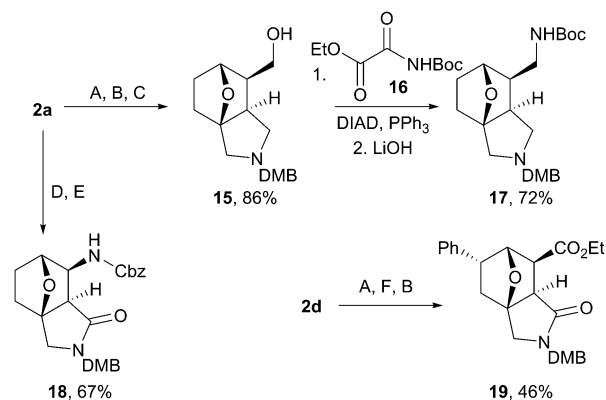
Scheme 3 Generation of diverse natural product-like scaffolds. Panel A: by ring Cleavage. Panel B: by ring expansion. Panel C: by annulation. Panel D: by rearrangement. HG-II, Hoveyda–Grubbs second generation catalyst.

Scheme 4 Tailoring of natural product-like scaffolds. Methods: (A) HCl, EtOH; (B) H₂, Pd/C; (C) LiAlH₄; (D) NH₄ HCO₂, Pd/C; (E) DPPA, BnOH; (F) PhB(OH)₂, K₃PO₄, 5 mol% Pd(PPh₃)₄.

The diversity of the scaffolds was assessed by constructing a hierarchical scaffold tree (Fig. 1).[17] Twenty one different frameworks are represented at the graph-node-bond level, which are related hierarchically to seven parent (monocyclic) frameworks. There is significant scaffold diversity at each level of hierarchy, meaning that the scaffolds are not simply closely related derivatives.

We determined natural product-likeness scores[5] for both the 25 (deprotected) scaffolds prepared, and for a range of their derivatives (Fig. 2). A virtual library of 11 223 decorated compounds was enumerated using our open-access tool, LLAMA,[18] by combining the 25 (deprotected) scaffolds with up to two typical medicinal chemistry capping groups (ESI†). To assess complementarity to existing collections, natural product-likeness scores were compared with those of 278 365 compounds from

15 with *N*-Boc ethyl oxamate (**16**) as nucleophile, followed by hydrolytic work up,[14] gave the orthogonally-protected diamine **17** in 72% yield. Alternatively, transfer hydrogenation of **2a**, followed by Curtius rearrangement in the presence of benzyl alcohol, gave the related Cbz-protected scaffold **18** in 67% yield. Finally, esterification of **2d**, Suzuki reaction with phenyl boronic acid, and hydrogenation gave the scaffold **19** as a single diastereoisomer (46% yield over 3 steps).
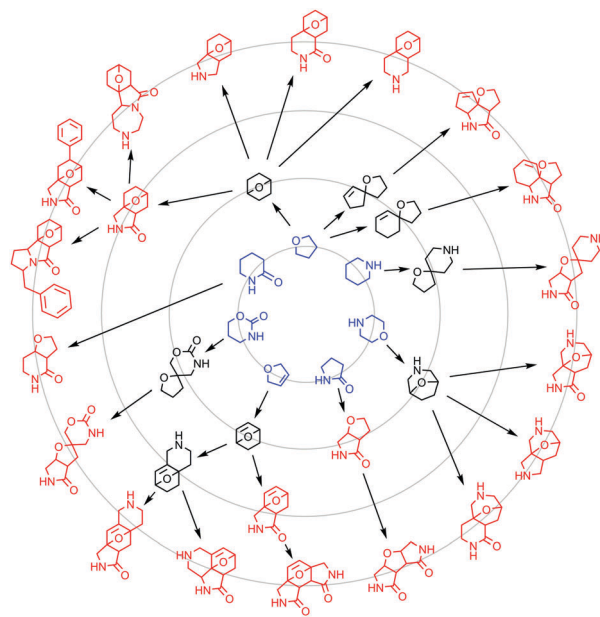
In total, 25 scaffolds were prepared using our biosynthesis-inspired approach (ESI†). The scaffolds were highly novel: the Murcko framework[15] of only two deprotected scaffolds were substructures in a 2% sample of the ZINC database.[16]

Fig. 1 Hierarchical scaffold tree. The twenty one frameworks (red) at the graph-node-bond level are related hierarchically to bicyclic (black) and parent (monocyclic; blue) frameworks.
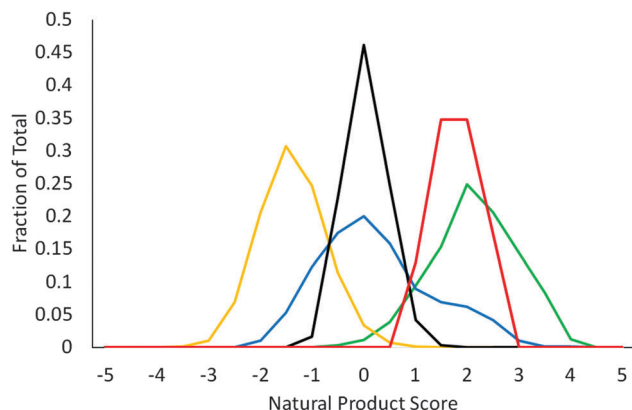
Fig. 2 Distribution of natural product-likeness scores for the 25 deprotected scaffolds (red); 11 445 decorated analogues of these scaffolds (black); 4460 compounds from a natural product screening library (green); 1822 FDA-approved drugs (blue); and 278 365 compounds from a commercial screening collection (yellow).

a commercial screening collection ($\mu = -1.63$); 1822 FDA-approved drugs ($\mu = -0.01$); and 4460 compounds from a natural product screening library ($\mu = 1.8$) (ESI†). The 25 deprotected scaffolds are highly natural product-like, with a mean natural product-likeness score of 1.58. Even after decoration with typical medicinal chemistry capping groups ($\mu = -0.21$), natural product-likeness is comparable to that of FDA-approved drugs and much higher than that of the commercial screening collection.

The high natural product-likeness of the scaffolds makes them highly attractive for exploitation in the synthesis of distinctive screening compounds. Like diterpenes biosynthesised from GGPP, the scaffolds are much more three-dimensional (*e.g.* higher Fsp^3) than the corresponding starting materials. Crucially, decoration can yield small molecules with good drug-like properties (ESI†). To exploit these favourable properties, >1300 screening compounds have since been prepared from the scaffolds **9**, **15** and **17** for addition to the Joint European Compound Library[19] (JECL) of the European Lead Factory.

Notes and references

1 R. D. Firn and C. G. Jones, *Nat. Prod. Rep.*, 2003, **20**, 382.
2 (*a*) S. Wetzel, R. S. Bon, K. Kumar and H. Waldmann, *Angew. Chem., Int. Ed.*, 2011, **50**, 10800; (*b*) H. van Hattum and H. Waldmann, *J. Am. Chem. Soc.*, 2014, **126**, 11853.
3 A. L. Harvey, R. Edrada-Ebel and R. J. Quinn, *Nat. Rev. Drug Discovery*, 2015, **14**, 111.
4 For some complementary synthetic approaches to libraries inspired by natural products, see ref. 2 and: (*a*) D. Morton, S. Leach, C. Cordier, S. Warriner and A. Nelson, *Angew. Chem., Int. Ed.*, 2009, **48**, 104; (*b*) K. C. Morrison and P. J. Hergenrother, *Nat. Prod. Rep.*, 2014, **31**, 6; (*c*) A. Masarwa, M. Weber and R. Sarpong, *J. Am. Chem. Soc.*, 2015, **137**, 6327; (*d*) T. Rodrigues, D. Reker, P. Schneider and G. Schneider, *Nat. Chem.*, 2016, **8**, 531.
5 P. Ertl, S. Roggo and A. Schuffenhauer, *J. Chem. Inf. Model.*, 2008, **48**, 68. Natural product likeness scores were calculated using the implementation in RDKit v2015.09.2 (Greg Landrum; Open Source Cheminformatics; http://www.rdkit.org; last accessed 27-Apr-2016).
6 (*a*) H. Lachance, S. Wetzel, K. Kumar and H. Waldmann, *J. Med. Chem.*, 2012, **55**, 5989; (*b*) D. H. Drewry and R. Macarron, *Curr. Opin. Chem. Biol.*, 2010, **14**, 289.
7 F. Lovering, J. Bikker and C. Humblet, *J. Med. Chem.*, 2009, **52**, 6752.
8 D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2012, **75**, 311.
9 (*a*) P. Zerbe and J. Bohlmann, *Trends Biotechnol.*, 2015, **33**, 419; (*b*) R. J. Peters, *Nat. Prod. Rep.*, 2010, **27**, 1521; (*c*) P. M. Dewick, *Nat. Prod. Rep.*, 2002, **19**, 181.
10 See also: M. Garcia-Castro, L. Kremer, C. D. Reinkemeier, C. Unkelbach, C. Strohmann, S. Ziegler, C. Ostermann and K. Kumar, *Nat. Commun.*, 2015, **6**, 6516.
11 (*a*) A. V. Varlamov, E. V. Boltukhina, F. I. Zubkov, N. V. Sidorenko, A. I. Chernyshev and D. G. Grudinin, *Chem. Heterocycl. Compd.*, 2004, **40**, 22; (*b*) F. I. Zubkov, V. P. Zaytsev, E. V. Nikitina, V. N. Khrustalev, S. V. Gozun, E. V. Boltukhina and A. V. Varlamov, *Tetrahedron*, 2011, **67**, 9148; (*c*) R. Murali, H. Surya Prakash Rao and H. W. Scheeren, *Tetrahedron*, 2001, **57**, 3165; (*d*) E. V. Boltukhina, F. I. Zubkov, E. V. Nikitina and A. V. Varlamov, *Synthesis*, 2005, 1859; (*e*) F. I. Zubkov, E. V. Boltukhina, K. F. Turchin and A. V. Varlamov, *Tetrahedron*, 2004, **60**, 8455; (*f*) T. Flagstad, G. Min, K. Bonnet, R. Morgentin, D. Roche, M. H. Clausen and T. E. Nielsen, *Org. Biomol. Chem.*, 2016, **14**, 4942.
12 (*a*) R. A. Tromp, J. Brussee and A. van der Gen, *Org. Biomol. Chem.*, 2003, **1**, 3592; (*b*) M. Treus, L. M. Harwood, J. C. Estévez, C. Salas, M. G. B. Drew and R. J. Estévez, *Synlett*, 2013, 2221.
13 For related rearrangements, see: (*a*) J. K. Lam, Y. Schmidt and C. D. Vanderwal, *Org. Lett.*, 2012, **14**, 5566; (*b*) J. K. Lam, H. V. Pham, K. N. Houk and C. D. Vanderwal, *J. Am. Chem. Soc.*, 2013, **135**, 17585.
14 F. Berrée, G. Michelot and M. Le Corre, *Tetrahedron Lett.*, 1998, **39**, 8275.
15 G. W. Bemis and A. M. Murcko, *J. Med. Chem.*, 1996, **39**, 2887.
16 J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad and R. G. Coleman, *J. Chem. Inf. Model.*, 2012, **52**, 1757.
17 A. Schuffenhauer, P. Ertl, S. Roggo, S. Wetzel, M. A. Koch and H. Waldmann, *J. Chem. Inf. Model.*, 2007, **47**, 47.
18 I. Colomer, C. J. Empson, P. Craven, Z. Owen, R. G. Doveston, I. Churcher, S. P. Marsden and A. Nelson, *Chem. Commun.*, 2016, **52**, 7209; LLAMA is available at https://llama.leeds.ac.uk.
19 (*a*) J. Besnard, P. S. Jones, A. L. Hopkins and A. D. Pannifer, *Drug Discovery Today*, 2015, **20**, 181; (*b*) A. Karawajczyk, F. Giordanetto, J. Benningshof, D. Hamza, T. Kalliokoski, K. Pouwer, R. Morgentin, A. Nelson, G. Müller, A. Piechot and D. Tzalis, *Drug Discovery Today*, 2015, **20**, 1310.