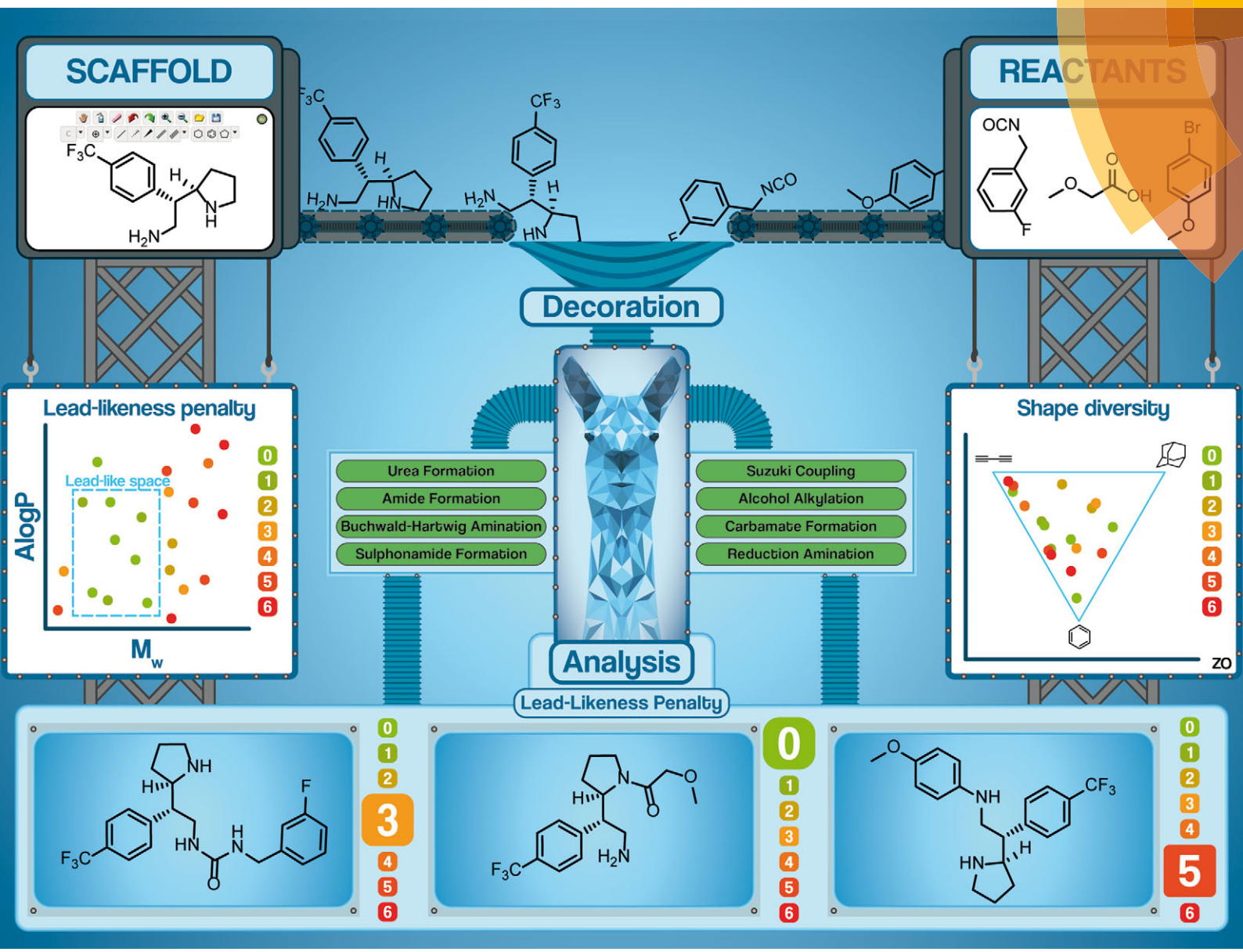
Volume 52 | Number 45 | 7 June 2016 | Pages 7175–7316

ChemComm

Chemical Communications

www.rsc.org/chemcomm



ISSN 1359-7345

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

Stephen P. Marsden, Adam Nelson *et al.*

A divergent synthetic approach to diverse molecular scaffolds: assessment of lead-likeness using LLAMA, an open-access computational tool

175 YEARS

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2016, 52, 7209

Received 18th April 2016,
Accepted 29th April 2016

DOI: 10.1039/c6cc03244c

www.rsc.org/chemcomm

A divergent synthetic approach to diverse molecular scaffolds: assessment of lead-likeness using LLAMA, an open-access computational tool†

Ignacio Colomer,[a] Christopher J. Empson,[ab] Philip Craven,[ab] Zachary Owen,[a] Richard G. Doveston,[ab] Ian Churcher,[c] Stephen P. Marsden*[a] and Adam Nelson*[ab]

Complementary cyclisation reactions of hex-2-ene-1,6-diamine derivatives were exploited in the synthesis of alternative molecular scaffolds. The value of the synthetic approach was analysed using LLAMA, an open-access computational tool for assessing the lead-likeness and novelty of molecular scaffolds.

Controlling molecular properties is crucial in drug discovery because the probability of successful progression is influenced by parameters including lipophilicity, molecular weight, the number of aromatic rings and the fraction of sp^3-hybridised carbons (Fsp^3).[1] As a result, guidelines, such as Lipinski's rule-of-five (concerning oral bioavailability),[2] have been formulated to help chemists to target drug-relevant chemical space.[3]

In turn, controlling the molecular properties of lead compounds is advisable since optimisation generally increases lipophilicity, molecular weight and complexity.[4] As a result, lead-like chemical space can be described in terms of both molecular properties (*e.g.*[5] $-1 < c\log P < 3$; $14 \leq$ heavy atoms ≤ 26) and structural features.[5] Unfortunately, most commercially-available compounds[5,6]‡ and new synthetic methods[5] do not target lead-like chemical space. The problem is exacerbated when diversity is considered since chemical space has been explored very unevenly and unsystematically.[7]

The challenge[5,8] of exploring novel lead-like chemical space has prompted us[6b,9] and others[10] to develop lead-oriented synthetic approaches. Demonstrating the value of such approaches requires tools for virtual library enumeration and evaluation that are not commonly available within academia. We have therefore developed LLAMA (Lead-Likeness and Molecular Analysis),§ an open-access tool that enables decoration[11] and assessment of the lead-likeness of small molecule scaffolds (Fig. 1). Each product is assigned a "lead-likeness penalty" (LLP) which penalises

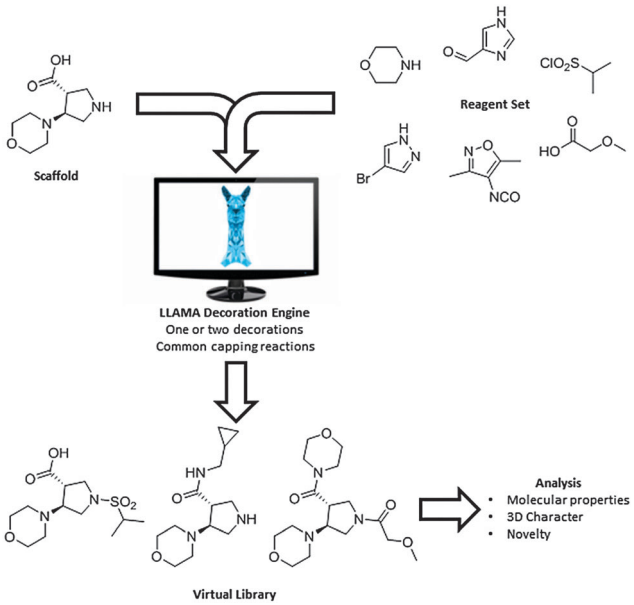


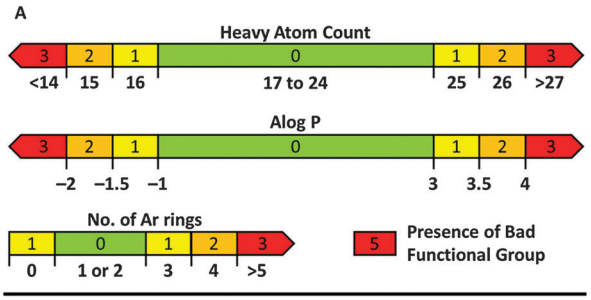
Fig. 1 Graphical representation of the LLAMA workflow. A virtual library is enumerated by decoration of entered scaffolds once or twice using definable capping reagents and reactions. The lead-likeness, three-dimensionality and novelty of the virtual library may be analysed.

properties and features that are not lead-like (Fig. 2; ESI†). Rather than applying strict filters, the penalty increases with deviation from lead-like space. Scaffold novelty is assessed by comparing the corresponding Murcko assemblies[12] (with and without alpha atoms) with those of, or embedded in, a random 2% of the "available now" set of the ZINC database[13] of commercially-available compounds. Finally, to capture the shape diversity of the compounds, the principal moments of inertia[14a] (PMI) and mean deviation from the plane of best fit[14b] of low-lying conformers are determined. To demonstrate LLAMA's utility, we analysed the lead-likeness of some scaffolds prepared using an approach that we have developed.

We envisaged a divergent synthetic approach in which unsymmetrical unsaturated diamine derivatives **1** would be converted

[a] *School of Chemistry, University of Leeds, Leeds, LS2 9JT, UK. E-mail: s.p.marsden@leeds.ac.uk, a.s.nelson@leeds.ac.uk*
[b] *Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, LS2 9JT, UK*
[c] *GlaxoSmithKline Medicines Research Centre, Stevenage, SG1 2NY, UK*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cc03244c

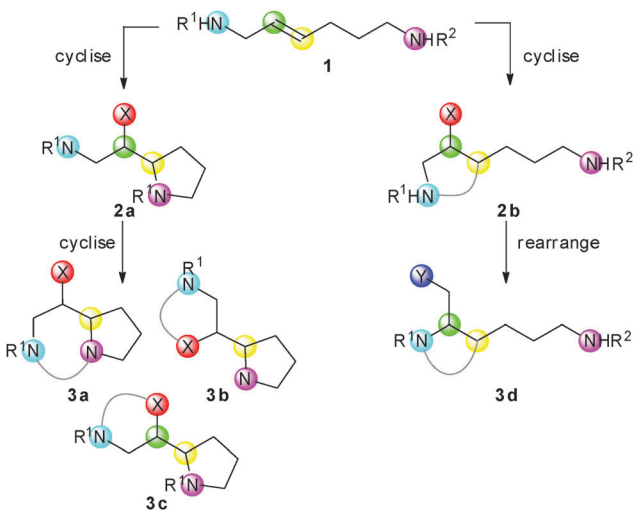


B

Property	Value	Penalty
Heavy atoms	30	3
Alog P	1.35	0
No. of Ar rings	3	1
Function Group Filter	0	0
Total		4

Property	Value	Penalty
Heavy atoms	22	0
Alog P	-0.53	0
No. of Ar rings	0	1
Function Group Filter	0	0
Total		1

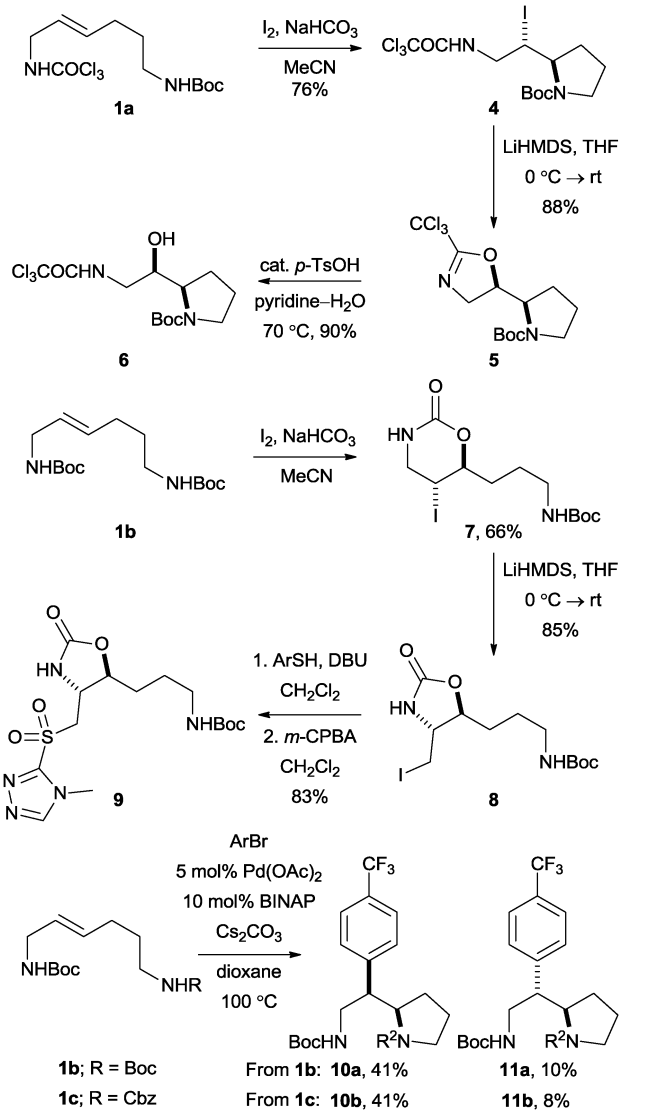
Fig. 2 Lead-likeness penalty. Panel A: Graphical representation of contributions to the penalty. Panel B: Penalties for two exemplar compounds.



Scheme 1 Synthetic approach for the conversion of hex-2-ene-1,6-diamine derivatives **1** into diverse molecular scaffolds. Alternative reactive sites are indicated as coloured circles. Complementary cyclisations would yield scaffolds **2** which might be further cyclised or rearranged to yield additional scaffolds **3**.

into alternative scaffolds (Scheme 1). Thus, complementary cyclisations would yield alternative heterocyclic intermediates **2**; further cyclisation or rearrangement would then yield additional molecular scaffolds **3**. The approach would be reminiscent of a branching pathway approach which enabled a single substrate to be converted into twelve natural product-like scaffolds.[15]

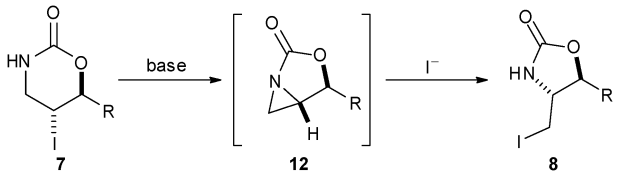
The fate of the iodocyclisation reactions of hex-2-ene-diamine derivatives **1** was dependent on the specific protecting groups



Scheme 2 Complementary cyclisations onto the central alkene of the hex-2-ene-1,6-diamine derivatives 1. The specific Ar groups used are shown in the products.

used (Scheme 2). Thus, treatment of the differentially-protected **1a** with iodine and sodium bicarbonate in acetonitrile resulted in cyclisation of the remote Boc-protected amine to yield the pyrrolidine **4** in 76% yield.[16] In stark contrast, under the same conditions, the doubly Boc-protected amine **1b** cyclised through the allylic carbamate to give a 95 : 5 mixture of the oxazinone **7** and the corresponding oxazolidinone (66% and 4% yield respectively).[17] In each case, regioselectivity was determined using the diagnostic upfield ^{13}C NMR chemical shift of the iodine-substituted carbons.

Treatment of the trichloroacetamide **4** with LiHMDS induced cyclisation to yield the oxazoline **5** which could be hydrolysed to give the differentially-protected diamine **6**. In contrast, treatment of the oxazinone **7** with LiHMDS triggered an unexpected rearrangement to give the isomeric oxazolidinone **8**. Presumably, deprotonation of **7** results in participation to yield the corresponding aziridine **12** which is then ring-opened by iodide (Scheme 3).

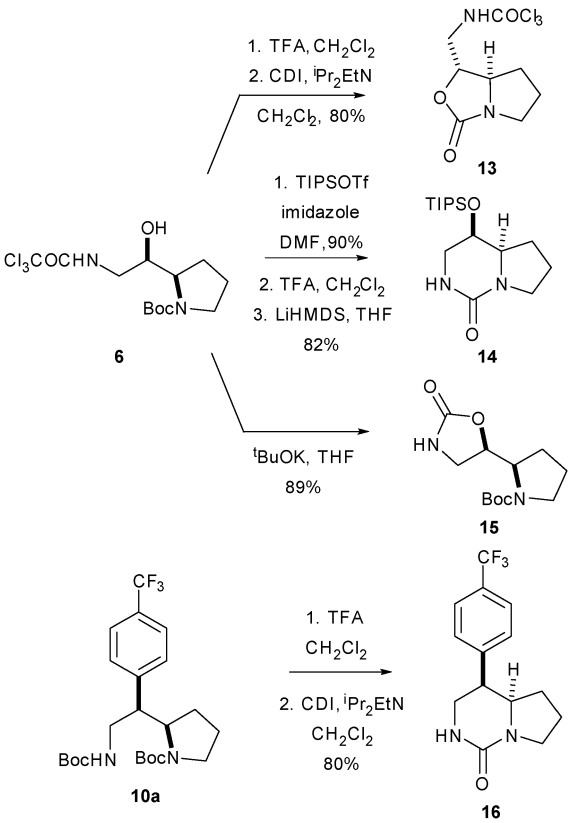


Scheme 3 Possible mechanism for the rearrangement of oxazin-2-one **7**.

The *trans* configuration of the oxazolidinone **8** was assigned by the characteristic[18] coupling constant (4.5 Hz) between the ring protons. Although we are unaware of related rearrangements of oxazinones, aziridines related to the intermediate **12** have been prepared[19] and ring-opened with nucleophiles.[20]

Palladium-catalysed aminoarylation also enabled cyclisation onto the central alkene of **1b** and the differentially-protected **1c.** With 5 mol% $Pd(OAc)_2$, 10 mol% BINAP, Cs_2CO_3 in dioxane at 100 °C,[21] cyclisation occurred as expected exclusively through the distal nitrogen to afford the corresponding pyrrolidines **10a–b** and **11a–b** with ~85 : 15 diastereoselectivity.

A range of bicyclic scaffolds was prepared from the cyclisation products **6** and **10a** (Scheme 4). Thus, Boc-deprotection of **6,** and reaction with CDI, yielded the bicyclic oxazolidinone **13.** Similarly, **10a** was converted into the related bicyclic scaffold **16.** Alternatively, capping of the hydroxyl group of **6** by silylation, followed by Boc deprotection and LiHMDS-mediated cyclisation yielded the alternative bicyclic scaffold **14.** Finally, ^{t}BuOK in THF triggered cyclisation of the hydroxy group of **6** (with displacement of trichloromethyl anion)



Scheme 4 Synthesis of bicyclic molecular scaffolds.

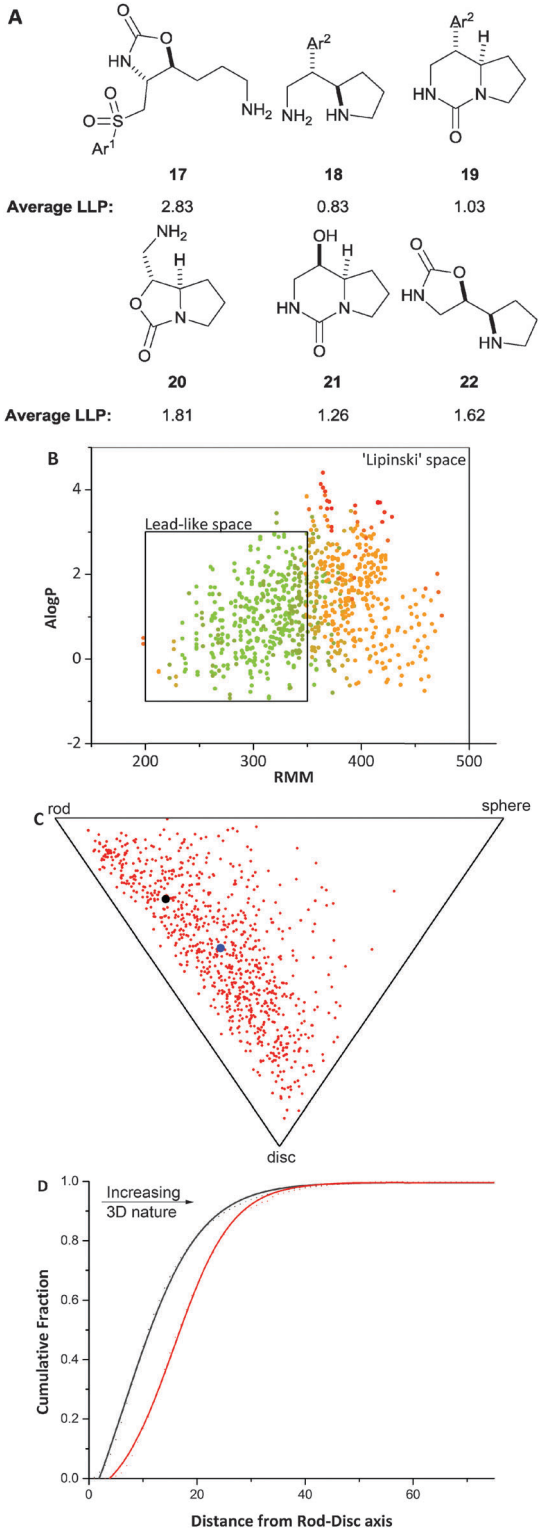


Fig. 3 Evaluation of the molecular scaffolds. Panel A: Molecular scaffolds uploaded to LLAMA (Ar^1 = 2-imidazole, 2-pyridyl, 2-pyrimidyl or 2-*N*-methyl-triazole; Ar^2 = 5-pyrimidyl, 4-trifluoromethylphenyl, 3-pyridyl or 3-cyanophenyl). Panel B: Molecular properties of the enumerated compounds (coloured according to LLP: 0, green; 3, orange; 6+, red). Panel C: PMI analysis of the enumerated compounds (red) and centre of gravities for the virtual library (large blue circle) and 2% of ZINC database (large black circle). Panel D: Cumulative percentage of molecules within a defined distance of the rod-disc axis for the virtual library (red) and 2% of ZINC database (black) (Panel D).

to yield the oxazolidinone **15**; similar cyclisations[22] have been previously reported.

Overall, six diverse scaffolds were therefore prepared from the differentially-protected diamines **1**. LLAMA was used to assess the value of the scaffolds **17–22** (Fig. 3). In each case, the enumerated library comprised compounds that had been decorated once or twice with exemplar medicinal chemistry capping groups.¶

Our approach yielded scaffolds that allow significant lead-like chemical space to be explored [average lead-likeness penalty: 1.57 (σ = 1.44) for our virtual library *cf.* 4.17 (σ = 3.17) for compounds from the ZINC database] (Panel B). In addition, the enumerated compounds are significantly more three-dimensional than a random 2% of the ZINC database (Panels C and D). Finally, the novelty assessment showed that the Murcko assembly (without alpha-atoms) of only one scaffolds was known (**18** with Ar^2 = 4-trifluoromethylphenyl) in the 2% selection of the ZINC database; however, the Murcko assembly with alpha atoms was not known for this scaffold, indicating that its substitution pattern is novel. We note that LLAMA may also be used prospectively to design scaffolds (*e.g.* specific Ar^2 groups in **19**) that are both novel and lead-like.

In conclusion, our synthetic approach exploited complementary cyclisations of hex-2-ene-1,6-diamine derivatives to yield a range of novel, lead-like small molecule scaffolds. The computational tool LLAMA can support the development of lead-oriented synthetic approaches by enabling assessment of the lead-likeness of alternative product scaffolds.

We acknowledge funding from EPSRC (EP/J00894X/1) and GSK for the development of LLAMA. We acknowledge 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) and EFPIA companies' in kind contribution for scaffold synthesis. IC (Ignacio Colomer), SPM and AN contributed to the scaffold syntheses; CJE, PC, ZO, RD, IC (Ian Churcher), SPM and AN contributed to the development of LLAMA.

Notes and references

‡ The proportion has been estimated to 2.6% (ref. 5), 32% (ref. 6*a*) and 23% (ref. 6*b*) depending on the specific criteria and reference set used.
§ LLAMA is available at: https://llama.leeds.ac.uk.
¶ The scaffolds **17–19** already contain a variable group (Ar^1 or Ar^2) and were subjected to only one further decoration.

1 (*a*) M. J. Waring, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 2844; (*b*) P. D. Leeson and B. Springthorpe, *Nat. Rev. Drug Discovery*, 2007, **6**, 881; (*c*) T. J. Ritchie and S. J. F. Macdonald, *Drug Discovery Today*, 2009, **14**, 1011; (*d*) M. C. Wenlock, R. P. Austin, P. Barton, A. M. Davis and P. D. Leeson, *J. Med. Chem.*, 2003, **46**, 1250; (*e*) F. Lovering, J. Bikker and C. Humblet, *J. Med. Chem.*, 2009, **52**, 6752.
2 (*a*) C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 1997, **23**, 3; (*b*) C. A. Lipinski, *Drug Discovery Today*, 2004, **1**, 337.
3 A single metric has been developed to estimate drug-likeness by integrating a number of relevant parameters: G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan and A. L. Hopkins, *Nat. Chem.*, 2012, **4**, 90.
4 (*a*) T. I. Oprea, A. M. Davis, S. J. Teague and P. D. Leeson, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 1308; (*b*) E. Perola, *J. Med. Chem.*, 2010, **53**, 2986; (*c*) M. M. Hann, *Med. Chem. Commun.*, 2011, **2**, 349; (*d*) G. M. Keserü and G. M. Makara, *Nat. Rev. Drug Discovery*, 2009, **8**, 203.
5 A. Nadin, C. Hattotuwagama and I. Churcher, *Angew. Chem., Int. Ed.*, 2012, **51**, 1114.
6 (*a*) A. Chuprina, O. Lukin, R. Demoiseaux, A. Buzko and A. Shivanyuk, *J. Chem. Inf. Model.*, 2010, **50**, 470; (*b*) R. G. Doveston, P. Tosatti, M. Dow, D. J. Foley, H. Y. Li, A. J. Campbell, D. House, I. Churcher, S. P. Marsden and A. Nelson, *Org. Biomol. Chem.*, 2015, **13**, 859.
7 A. H. Lipkus, Q. Yuan, K. A. Lucas, S. A. Funk, W. F. Bartelt III, R. J. Schenck and A. J. Trippe, *J. Org. Chem.*, 2008, **73**, 4443.
8 (*a*) P. MacLellan and A. Nelson, *Chem. Commun.*, 2013, **49**, 2383; (*b*) R. Doveston, S. P. Marsden and A. Nelson, *Drug Discovery Today*, 2014, **19**, 813.
9 (*a*) D. J. Foley, R. G. Doveston, I. Churcher, A. Nelson and S. P. Marsden, *Chem. Commun.*, 2015, **51**, 11174; (*b*) T. James, I. Simpson, J. A. Grant, V. Sridharan and A. Nelson, *Org. Lett.*, 2013, **15**, 6094; (*c*) T. James, P. MacLellan, G. M. Burslem, I. Simpson, J. A. Grant, S. Warriner, V. Sridharan and A. Nelson, *Org. Biomol. Chem.*, 2014, **12**, 2584.
10 (*a*) M. Lüthy, M. C. Wheldon, C. Haji-Cheteh, M. Atobe, P. S. Bond, P. O'Brien, R. E. Hubbard and I. J. Fairlamb, *Bioorg. Med. Chem.*, 2015, **23**, 2680; (*b*) S. V. Ryabukhin, D. M. Panov, D. S. Granat, E. N. Ostapchuk, D. V. Kryvoruchko and O. O. Grygorenko, *ACS Comb. Sci.*, 2014, **16**, 146; (*c*) A. Borisov, V. Voloshchuk, M. Nechayev and O. Grygorenko, *Synthesis*, 2013, 2413.
11 The default set of capping groups is based on those that we have used in previous analyses (ref. 6*b*). See ESI†.
12 G. W. Bemis and M. A. Murcko, *J. Med. Chem.*, 1996, **39**, 2887.
13 J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad and R. G. Coleman, *J. Chem. Inf. Model.*, 2012, **52**, 1757.
14 (*a*) W. H. B. Sauer and M. K. Schwarz, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 987; (*b*) N. C. Firth, N. Brown and J. Blagg, *J. Chem. Inf. Model.*, 2012, **52**, 2516.
15 D. Robbins, A. F. Newton, C. Gignoux, J.-C. Legeay, A. Sinclair, M. Rejzek, C. A. Laxon, S. K. Yalamanchili, W. Lewis, M. A. O'Connell and R. A. Stockman, *Chem. Sci.*, 2011, **2**, 2232.
16 For examples of iodocyclisations of *N*-Boc pentenamines, see: (*a*) K. Kamiyama, Y. Urano, S. Kobayashi and M. Ohno, *Tetrahedron Lett.*, 1987, **28**, 3123; (*b*) S. G. Davies, G. Darren, M. Smyth and A. Chippindale, *J. Chem. Soc., Perkin Trans. 1*, 1999, 3089.
17 J. M. Jordá-Gregori, M. E. González-Rosende, P. Cava-Montesinos, J. Sepúlveda-Arques, R. Galeazzi and M. Orena, *Tetrahedron: Asymmetry*, 2000, **11**, 3769.
18 A. G. Wee and D. D. McLeod, *J. Org. Chem.*, 2003, **68**, 6268.
19 For similar rearrangements of γ-lactams and cyclic ureas, see: (*a*) A. Sudau, W. Münch, J.-W. Bats and U. Nubbemeyer, *Eur. J. Org. Chem.*, 2002, 3304; (*b*) C. Agami, F. Amiot, F. Couty and L. Dechoux, *Tetrahedron Lett.*, 1998, **39**, 5373.
20 For examples for similar aziridines being ring-opening by a range of nucleophiles, see: (*a*) A. Padwa, A. C. Flick, C. A. Leverett and T. Stengel, *J. Org. Chem.*, 2004, **69**, 6377; (*b*) S. Knapp and Y. Yu, *Org. Lett.*, 2007, **9**, 1359.
21 M. B. Bertrand, J. D. Neukom and J. P. Wolfe, *J. Org. Chem.*, 2008, **73**, 8851.
22 (*a*) R. Link, G. Subharekha, M. Gallant, S. J. Danishefsky, T. C. Chou and L. M. Ballas, *J. Am. Chem. Soc.*, 1996, **118**, 2825; (*b*) K. Stankov, M. Martinkov, J. Gonda, M. Bago, M. Piltov and G. Gönciov, *Tetrahedron: Asymmetry*, 2015, **26**, 1394.

