nature chemistry

ARTICLES

PUBLISHED ONLINE: 19 JANUARY 2014 | DOI: 10.1038/NCHEM.1844

Probing chemical space with alkaloid-inspired libraries

Michael C. McLeod, Gurpreet Singh, James N. Plampin III, Digamber Rane, Jenna L. Wang, Victor W. Day and Jeffrey Aubé*

Screening of small-molecule libraries is an important aspect of probe and drug discovery science. Numerous authors have suggested that bioactive natural products are attractive starting points for such libraries because of their structural complexity and sp^3-rich character. Here, we describe the construction of a screening library based on representative members of four families of biologically active alkaloids (Stemonaceae, the structurally related cyclindricine and lepadiformine families, lupin and Amaryllidaceae). In each case, scaffolds were based on structures of the naturally occurring compounds or a close derivative. Scaffold preparation was pursued following the development of appropriate enabling chemical methods. Diversification provided 686 new compounds suitable for screening. The libraries thus prepared had structural characteristics, including sp^3 content, comparable to a basis set of representative natural products and were highly rule-of-five compliant.

Natural products have played a central role in medicine for as long as humans have sought to cure and ameliorate disease[1,2]. Many have been fine-tuned by evolution for purposes that bear a mechanistic relationship to a given therapeutic need[3]. Natural products are often potent, selective and able to cross biological membranes, although many do not adhere to common paradigms for oral absorption[4] (it is worth recalling that natural products were specifically excluded from Lipinski's guidelines of drug-like properties[5]). For these reasons, natural products continue to inspire creativity in both medicinal chemistry and chemical synthesis.

As screening of small-molecule libraries remains an important aspect of early-stage drug and probe discovery, there has been interest in increasing the representation of natural products and related structures in libraries[6–10]. Approaches include the straightforward approach of collecting natural products or natural-product extracts from their natural sources, which requires access to libraries obtained from bioprospecting and, for extracts, a downstream deconvolution step. To supplement such sources, synthetic chemists have co-opted natural product structures for the construction of natural product-like libraries. More often than not, these efforts provide purpose-built libraries for biological indications closely related to known bioactivities of the natural products itself[11–13]. Diversity-oriented synthesis (DOS) has also been used to create natural product-inspired libraries[9,14,15]. Some authors have suggested that the higher sp^3/sp^2 content and rich stereochemistry typical of natural products and, by extension, libraries derived from them is correlated with suitability as drug candidates[16–20]. In all of these approaches, the complexity of natural products presents synthetic challenges that must be surmounted to provide screening libraries that contain chemotypes that can be modified in the case of attractive hits[21,22].

We sought an approach to natural product-like screening libraries that would balance the likelihood of finding molecules useful in the pursuit of new biology with synthetic tractability. We chose to focus on selected families of alkaloids, preferring those with established biological activity at multiple targets, hypothesizing that such families might embody a 'privileged structure'[23–28] that could be optimized for new biological properties following suitable modification. We therefore created a nested set of synthetically derived cores that represented salient structural features of the natural-product starting point. These were further modified to produce 'secondary scaffolds' that differ more substantially from the original structure but retain attractive elements of scaffold design. In previous work, we used these tenets to create a library based on Stemonaceae alkaloids that ultimately led to potent Sigma-1 ligands, an activity not known to be associated with this family of natural products[29].

Here, we generalize this concept to structurally diverse alkaloids of the cylindricine, Amaryllidaceae and lupin families. We sought to address the synthetic challenges presented by these families by repurposing a suite of thematically related chemical reactions to library construction, most of which were developed in the context of total synthesis. Additional method development ultimately allowed us to obtain diversifiable scaffolds unavailable directly from natural-product starting materials. Overall, we synthesized a total of 686 new compounds, of which >90% were prepared in >20 mg quantities and all in >90% purity.

Figure 1 depicts scaffolds inspired by the architecture of four biologically active alkaloid families: (1) Stemonaceae alkaloids (exemplified by neostenine)[30,31]; (2) the structurally related cylindricine, lepadiformine and fascicularine families of marine alkaloids (here, collectively called the cylindricine series)[32]; (3) the Amaryllidaceae alkaloid mesembrine[33]; and (4) sparteine, a lupin alkaloid[34,35]. Structurally, each starting alkaloid contains at least one fused pair of rings, but one spiro and one bridged ring system are also represented. Biologically, these classes represent a variety of reported activities, ranging from those described in traditional medicine for the natural-product source to pharmacologically verified and clinically used agents (Supplementary Table 1). According to the approach outlined above, each polycyclic alkaloid was simplified to the primary and secondary scaffolds indicated (Fig. 1). The secondary scaffolds contain the same number of rings as the central scaffold, but with different ring sizes and/or connectivities.

Department of Medicinal Chemistry, Delbert M. Shankel Structural Biology Center, University of Kansas, 2034 Becker Drive, Lawrence, Kansas 66047, USA. *e-mail: jaube@ku.edu

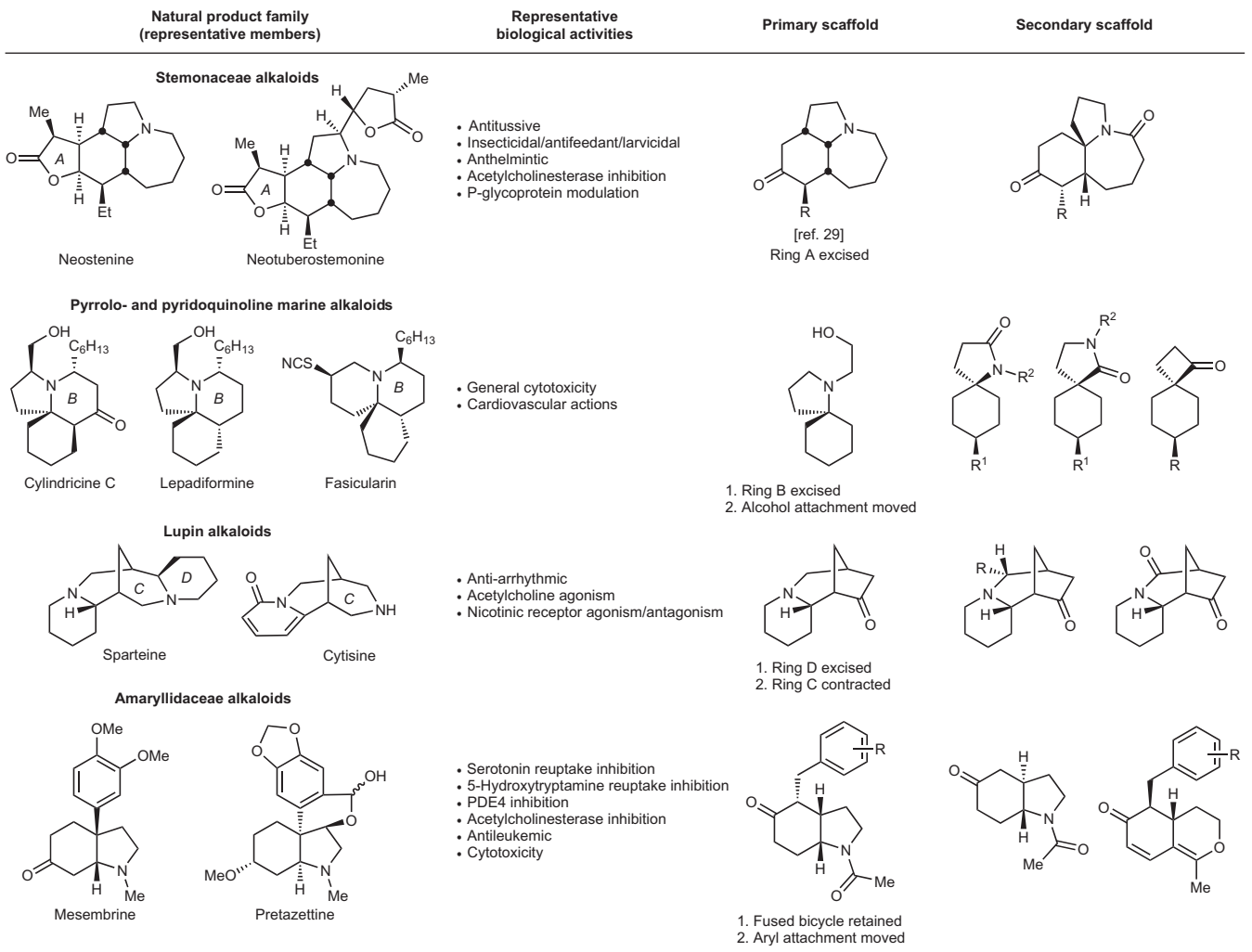

Figure 1 | Strategic overview of natural-product families selected for library expansion and corresponding scaffold selection. Each family of natural products is exemplified by two to three members and the biological activities cited are representative of each family as a whole. Reviews[30–35] provide overviews of the biological landscape, with additional references provided in Supplementary Table 1. In each case, a primary scaffold embodies the minimal structural aspects of the natural-product family that were pursued. This involved removal of some features to enhance both versatility and synthetic accessibility. Secondary scaffolds (far right) represent more substantially modified variants of the primary scaffold, either through functional group or substitution changes or modification of ring structures.

Results

Synthesis of scaffolds. To construct libraries containing 40−320 members each, it was necessary to develop enabling chemistries that would permit practical access to the desired chemotypes (Fig. 2). This was pursued using the azido-Schmidt conversion of keto azides to lactams as the primary unifying technology[36]. Four variations of this reaction were used to create the desired scaffolds: (1) the combined Diels–Alder/Schmidt reaction between a silyloxy diene and an azide-containing dienophile[37] (Stemonaceae alkaloid series and one scaffold for the mesembrine series (Fig. 2a,d); (2) the reaction of a ketone with a hydroxyalkyl azide[38] (cylindricine series; Fig. 2b); (3) the intramolecular Schmidt reaction of a ketone tethered to an azide[39] (sparteine series; Fig. 2c); and (4) another combined Diels–Alder/Schmidt reaction, but this time one in which the azide is attached to the diene (the second mesembrine scaffold; Fig. 2d).

The Diels–Alder/Schmidt sequence is particularly powerful, as it extends the applicability of the ubiquitous Diels–Alder reaction by tying it to the *in situ* conversion of the new ring into a heterocycle. The other sequences rely in one case (Fig. 2b) on a spiroannulation step to afford a cyclobutanone (itself pressed into service as a scaffold; see below) that can be converted into the desired spirocyclic intermediates. Finally, the route in Fig. 2c parlays an advanced total synthesis intermediate into the scaffold **16**. To use the Diels–Alder/Schmidt chemistry in Fig. 2a and d, we needed to explore new variations using highly substituted dienes. Schmidt reactions related to those shown in Fig. 2b and c were first developed during total synthesis efforts toward sparteine, and the lepadiformine and cylindricine alkaloids[39,40]. The azido alcohol-mediated Schmidt reaction has been previously used for building a library of γ-turn mimetics[41]. We prepared most of the scaffolds in racemic form, as there was no reason to favour a particular enantiomer for broad screening, often in straightforward biochemical assays (in a few cases, we did make scaffolds from L-configured amino-acid derivatives).

The Lewis acid-promoted reaction of unsubstituted diene **1a** with **2** has been reported previously[37], but dienes **1b–d**, which contain an additional element of diversity and are readily prepared in three or fewer steps from commercially available starting material, were unknown to engage in Diels–Alder/Schmidt chemistry before this work. All four dienes underwent the desired conversions, which were reproducible and scalable to provide up to 12 g of lactam with no loss of efficiency. Mechanistically, such reactions are believed to proceed via a Diels–Alder reaction, from which the product stereochemistry arises, followed by an intramolecular

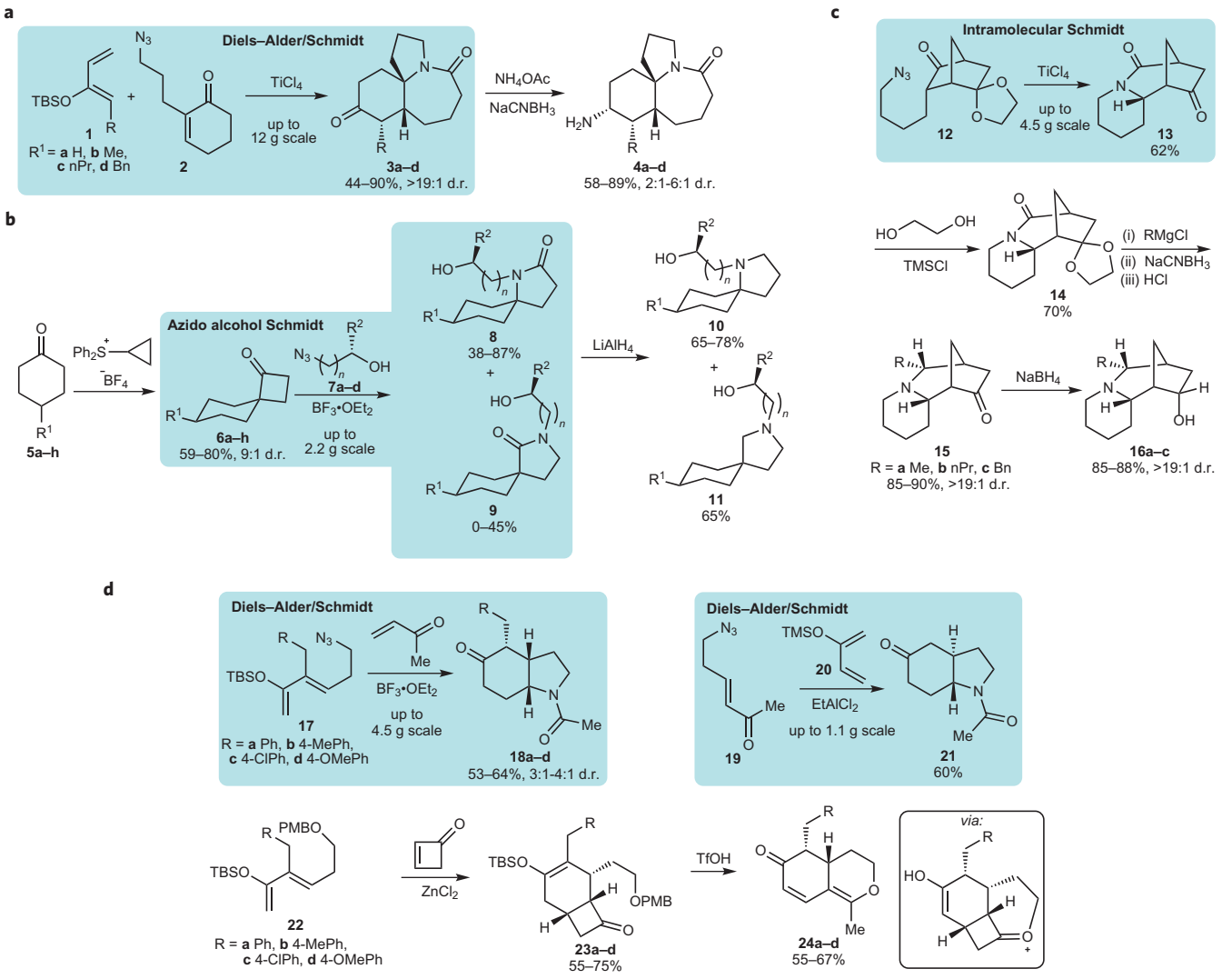

Figure 2 | Construction of primary and secondary scaffolds. In each section, coloured boxes indicate enabling chemistry developed for scaffold syntheses. **a**, Synthesis of Stemonaceae alkaloid scaffolds using a Diels–Alder/Schmidt reaction. **b**, Cylindricine scaffolds were prepared by sulfur ylide-mediated spiroannulation followed by ring expansion using an azido alcohol variant of the Schmidt reaction. **c**, Azide **12** was prepared from a previously reported C2-symmetrical diketone[39] using an improved method (see Supplementary Section II) and converted to the tricyclic sparteine scaffolds using the intramolecular Schmidt reaction. **d**, Two different variations of bicyclic analogues of mesembrine were prepared using a Diels–Alder/Schmidt reaction. In addition, reaction of **22** with the interesting dienophile cyclobutenone[43] followed by rearrangement afforded a non-nitrogenous scaffold **24**. Full synthetic schemes, including experimental details and characterization data of representative compounds, are provided in the Supplementary Sections II and V.

Schmidt reaction[37]. The relative stereochemistries of lactams **3b–d** were confirmed by X-ray crystallography. Reductive amination with ammonium acetate provided amines **4a–d** in good yield in 2:1 to 6:1 d.r. The diastereomers of amines **4b–d** were separated by reverse-phase chromatography (the isomers of **4a** were inseparable and not used for library synthesis). The relative stereochemistry of **4b** and **4c** was confirmed by X-ray crystallography of sulfonamide derivatives (see Supplementary Section VI), and amine **4d** was assigned by analogy.

Trost spiroannulation was used to create spirocyclic ketones **6a–h** (Fig. 2b)[42]. Once in hand, these ketones were used as the basis of one sublibrary, but our main objective was to convert them to the cylindricine-inspired substructures **8** and **9**. This was accomplished by previously unexamined Schmidt reactions of **6a–h** with hydroxyalkyl azides **7a–d**. This led to a mixture of readily separable lactam regioisomers **8** and **9**, thereby affording a pair of useful scaffolds from ketones **6**. The use of a two-carbon hydroxyalkyl azide led predominantly to the formation of scaffolds **8**, whereas three-carbon homologues afforded ~1:1 mixtures of constitutional isomers. Previous work has shown the selectivity of such reactions to be highly substrate-dependent, and a discussion of the relevant topics has been published[38]. Additionally, lactams **8** and **9** were reduced to afford tertiary amine cores **10** and **11**.

The tricyclic lactam core of the sparteine-inspired scaffolds (Fig. 2c) was prepared by intramolecular Schmidt reaction of azide **12**, an intermediate in the total synthesis of sparteine[39]. Scale-up to 4.5 g quantities was possible by optimizing the previous route to compound **12**. The multistep reduction shown in the scheme afforded amines **15a–c** as single diastereomers, and further reduction with $NaBH_4$ similarly led to isomerically pure alcohols **16a–c**. The stereochemistry of the endo-selective Grignard addition and ketone reduction was confirmed by X-ray crystallography of a derived carbamate (see Supplementary Section VI).

The primary scaffold of the mesembrine series **18** was prepared in moderate yield and 3:1 to 4:1 d.r. by reacting previously unknown

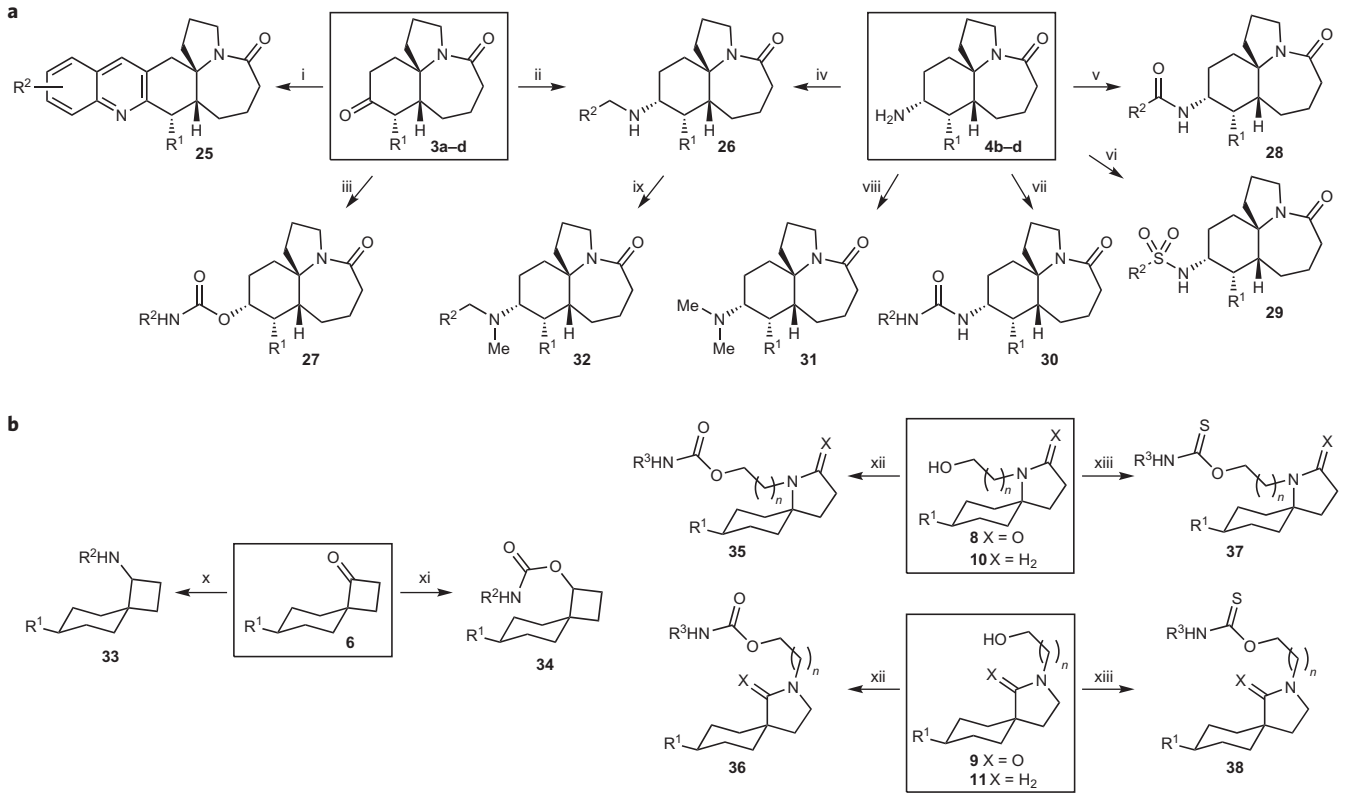

Figure 3 | Library construction from Stemonaceae and cylindricine alkaloid-inspired scaffolds. a, Ketone scaffolds **3a–d** and amines **4b–d** were converted into libraries of quinolines, amines, amides, sulfonamides, ureas and carbamates. Key scaffolds are indicated by boxes. **b**, Ketone scaffolds **6**, lactams **8** and **9**, and amines **10** and **11** were converted into a series of libraries of amines, carbamates and thiocarbamates. A total of 499 unique structures, each obtained in >90% purity (high-performance liquid chromatography (HPLC), ultraviolet detector at 214 nm) and >20 mg quantities were obtained. Scaffold (number of final products obtained): **3a–d** (131), **4c–d** (191), **6** (112), **8** (19), **9** (12), **10** (32), **11** (2). (i) 2-Nitrobenzaldehyde, Fe^0, 0.1 M aq. HCl, EtOH, 85 °C; then **3a–d**, KOH, 85 °C; (ii) amine, AcOH, $Na(OAc)_3BH$, CH_2Cl_2, rt; (iii) (a) L-Selectride, THF, −78 °C to rt, 90–95%, 9:1 to >19:1 d.r., (b) 4-nitrophenyl-chloroformate, pyridine, rt, 55−76%, (c) R^2NH_2, CH_2Cl_2, rt; (iv) R^2CHO, AcOH, $Na(OAc)_3BH$, CH_2Cl_2, rt; (v) R^2CO_2H, EDC, DMAP, CH_2Cl_2, rt; (vi) R^2SO_2Cl, Et_3N, CH_2Cl_2, rt; (vii) $R^2N{=}C{=}O$, PhMe, rt; (viii) $H_2C{=}O$, HCO_2H, 95 °C; (ix) $H_2C{=}O$, AcOH, $Na(OAc)_3BH$, CH_2Cl_2, rt; (x) R^2NH_2, DCE, microwave, 150 °C; (xi) (a) $NaBH_4$, THF, MeOH, rt, 94−97%, (b) $R^2N{=}C{=}O$, Et_3N, THF, rt; (xii) $R^2N{=}C{=}O$, Et_3N, THF, rt; (xiii) $R^2N{=}C{=}S$, NaH, THF, rt.

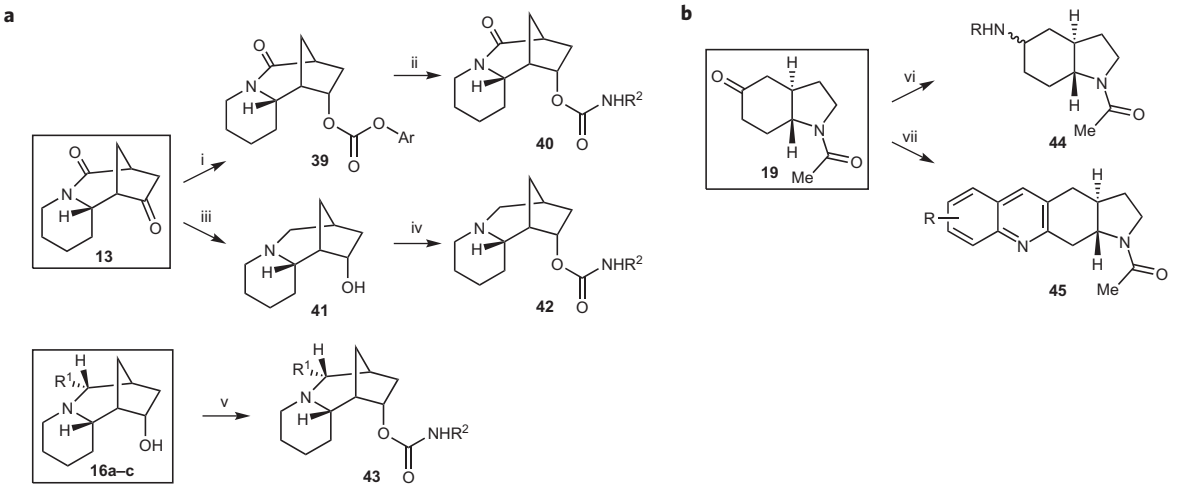

Figure 4 | Library construction from sparteine and mesembrine-inspired scaffolds. a, Library construction from sparteine-inspired scaffolds led to carbamates generated both from lactam **13** directly or by first converting it to the amine-containing scaffold **41**. Similar chemistry could be used on the additional alkyl-group-containing scaffolds **16a–c**. **b**, Amide scaffold **19** was converted into amine and quinoline libraries. A total of 132 unique structures, each obtained in >90% purity (HPLC, ultraviolet detector at 214 nm) and >10 mg quantities, were obtained. Scaffold (number of final products obtained): **13** (44), **16a–c** (52), **19** (36). (i) (a) $NaBH_4$, MeOH, rt, 88%, >19:1 d.r., (b) 4-nitrophenyl chloroformate, pyridine, THF, rt, 95%; (ii) R^2NH_2, DCE, rt; (iii) $LiAlH_4$, THF, reflux, 80%; (iv) $R^2N{=}C{=}O$, MeCN, microwave, 110 °C; (v) $R^2N{=}C{=}O$, THF, rt; (vi) RNH_2, AcOH, $Na(OAc)_3BH$, THF; (vii) 2-nitrobenzaldehyde, Fe^0, 0.1 M HCl, EtOH, 85 °C; then **19**, KOH, 85 °C.

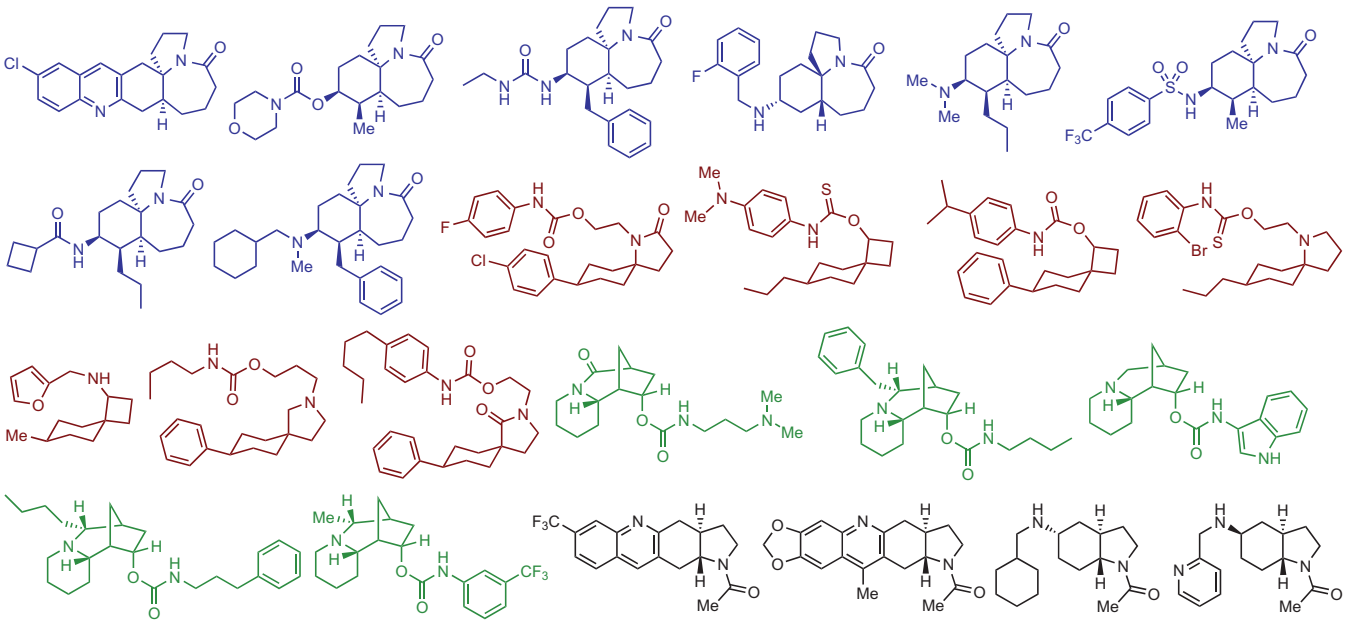

Figure 5 | Representative selection of library compounds used in cheminformatic analyses. For a full list of scaffold and library compounds used in this analysis, see Supplementary Figs 3 and 4.

dienes **17** with methylvinylketone. These isomers proved inseparable and were not subjected to additional diversification. We similarly prepared secondary scaffold **21** by a modification of the previously reported Diels–Alder/Schmidt reaction of diene **20** and azide-tethered enone **19** (Fig. 2d)[37]. Attaching the azide to either the diene or dienophile allows access to either the *cis* or *trans* fused scaffold, (cf. routes to **18** versus **21**). Hypothesizing that greater selectivity might be obtained with a more biased ring system, we reacted **17** with the unusual dienophile cyclobutenone (introduced to the Diels–Alder reaction by Danishefsky[43]), but a complex mixture of products was obtained, with no desired lactam evident. For comparison, we prepared azide-less silyloxydiene **22** and submitted it to the Diels–Alder reaction conditions with cyclobutenone. In this case, the cycloaddition proceeded smoothly to afford bicyclic ketone **23,** but, surprisingly, treatment with trifluoromethanesulfonic acid resulted in the formation of the previously unknown isochromenones **24** as single diastereomers. This reaction is previously unknown and could proceed via a retro-Michael reaction from the intermediate shown (an alternative enolization/electrocyclic ring-opening sequence is also possible, but available evidence does not currently allow us to differentiate between the possibilities).

Library construction. We took three main tacks toward diversifying our scaffolds: direct conversion of ketones to additionally fused heterocycles or amines, conversion of alcohols to carbamates, or decoration of amines as amides, sulfonamides or ureas. These methods were chosen because they increase structural diversity, provide various degrees of hydrogen-bonding capabilities, and yield functional groups consistent with probe or drug development.

For the Stemonaceae-derived library, scaffold **3a** gave quinolines **25** by a modified Friedländer reaction and secondary amines **26** by reductive amination with various benzylamines (Fig. 3a). Analogous reactions with scaffolds **3b–d** resulted in an inseparable mixture of diastereomers, which were not pursued. Further diversification was achieved by reducing ketones **3a–d** with L-Selectride to generate the corresponding alcohols with excellent diastereoselectivity (relative stereochemistry of the major product was determined by X-ray crystallography). These were used to prepare a library of carbamates **27**. Direct reaction of the alcohol with isocyanates resulted in the formation of an unidentified by-product that co-eluted with the carbamate product upon purification. However, by first converting the alcohol to the corresponding *para*-nitrobenzyl carbonate, subsequent reaction with a wide range of amines resulted in a much cleaner and more effective preparation of library members. Amine scaffolds **4b–d** were used to synthesize sublibraries of amides **28**, sulfonamides **29**, ureas **30** and secondary amines **26**, while tertiary amine libraries were obtained either by reductive amination of secondary amines **26** with formaldehyde, or by Eschweiler–Clarke reaction of primary amines **4b–d**.

Spirocyclic ketone scaffolds **6** were converted into libraries of amines **33** by microwave-assisted reductive amination, and libraries of carbamates **34** by reduction with $NaBH_4$ followed by reaction with isocyanates (Fig. 3b). Similarly, the alcohol moieties of lactams **8** and **9** and cyclic amines **10** and **11** were reacted with isocyanates and isothiocyanates to produce libraries of carbamates **35** and **36** and thiocarbamates **37** and **38**, respectively.

The ketone of scaffold **13** was stereoselectively reduced to afford the corresponding endo-alcohol (Fig. 4a), the structure of which was determined by X-ray of the corresponding *para*-bromobenzoyl ester (see Supplementary Section VI). The alcohol was converted into a library of carbamates **40** by carbonate formation and subsequent reaction with a range of amines. The corresponding library containing a basic amine in the ring system was prepared by double reduction with $LiAlH_4$ in THF, giving **41**, which was further diversified into carbamates **42** by microwave-promoted reaction with isocyanates. Similar treatment of **16a–c** led to libraries exemplified by **43**.

Reductive amination of mesembrine-inspired scaffold **19** provided the corresponding secondary amines **44** in 2:1 to 1:1 d.r. (Fig. 4b). We ascribe the poor stereoselectivity to the relatively flat nature of the *trans* ring-fused system inherent in **19**. In this case, the amine diastereomers were separable by reverse-phase chromatography, which allowed us to incorporate a modest number of amines containing this scaffold into our library (relative stereochemistries were determined by two-dimensional NMR). The use of anilines in the reductive amination resulted in inseparable mixtures of diastereomeric products, which were unsuitable for our screening requirements. A library of quinolines **45** was also synthesized by reacting scaffold **19** with either substituted 2-aminobenzaldehydes or 2-aminobenzophenones or acetophenones.

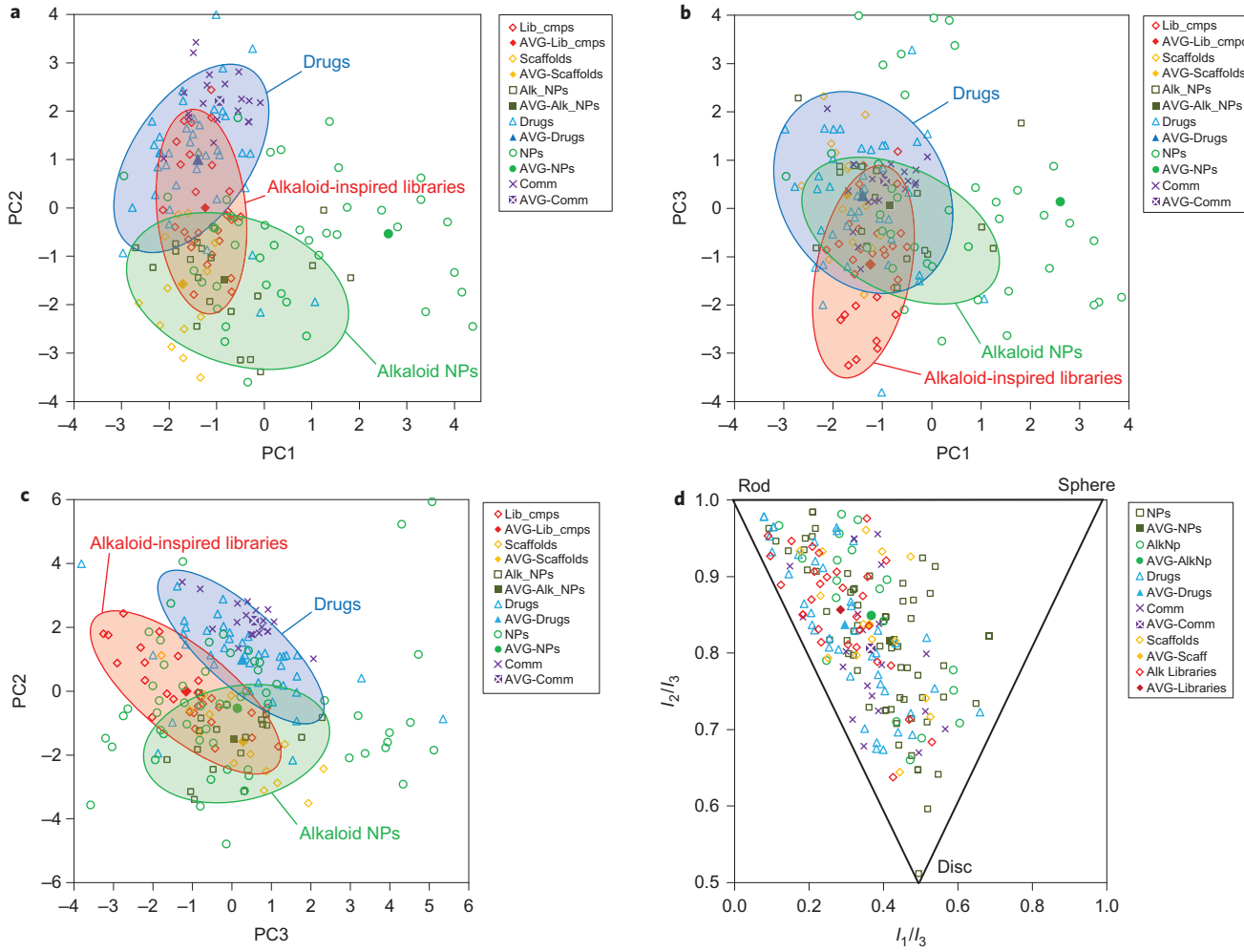

Figure 6 | Cheminformatic analysis of alkaloid-inspired scaffolds and library members. Structural and physiochemical properties of a representative selection of synthesized scaffolds (14 compounds, yellow diamonds) and library members (29 compounds, red diamonds) were compared with those of alkaloid natural products (NPs, 20 compounds, green squares) and an established reference set[48] of drugs (40 compounds, blue triangles), commercially available drug-like molecules (20 compounds, purple crosses) and natural products (60 compounds, green circles) using PCA and PMI analysis. The hypothetical average (mean) structure for each series is also plotted (AVG-). **a**, PCA plot of PC1 versus PC2. **b**, PCA plot of PC1 versus PC3. **c**, PCA plot of PC2 versus PC3. **d**, PMI plot showing the three-dimensional shape of the lowest-energy conformer of each compound. The shaded red, green and blue areas outline the regions of the plot where the majority of our alkaloid inspired libraries, alkaloid natural products and drugs, respectively, are located.

Cheminformatic analysis. The chemical properties of our alkaloid-inspired libraries were analysed using principal component analysis (PCA), a statistical tool to condense multidimensional chemical properties (for example, molecular weight, log P, ring complexity) into single dimensional numerical values (principal components), allowing greater ease of comparison with different sets of compounds[44]. This analysis was performed on a representative sample of scaffolds and library members, as summarized in Fig. 5 (for full list see Supplementary Figs 3 and 4), which compared them with a selection of alkaloid natural products and a reference set of drugs, natural products and commercial drug-like compounds using the protocols employed by Tan[45]. The parameters that have the greatest influence on principal component 1 (PC1) are molecular weight, number of oxygen atoms, number of hydrogen-bond acceptors, and topological polar surface area (tPSA). Together, these parameters have the effect of moving compounds to the right in plots PC1 versus PC2 and PC1 versus PC3. The descriptors with the largest loading on PC2 are the number of nitrogen atoms, number of aromatic rings and the number of ring systems, which shift compounds upward in plots PC1 versus PC2 and PC2 versus PC3. In contrast, the nStMW parameter (defined as the number of *R*–*S* stereocentres; this may be viewed as a rough descriptor of stereochemical complexity), shifts compounds downward in these plots. Finally, PC3 is affected to the greatest degree by X log P (calculated octanol/water partition coefficient), number of rings, and ALOGPs (an alternative log P calculation), which together shift compounds in a negative direction along the PC3 axis in plots PC1 versus PC3 and PC2 versus PC3, and ALOGpS (calculated aqueous solubility), which shifts compounds in a positive direction in these plots.

In two of the three variations, the data show significant overlap between our alkaloid-inspired library compounds with both drug and natural-product regions (Fig. 6a,b). In the PC2 versus PC3 plot, there is less overlap between our library compounds and commercial drug space, but the compounds still overlap considerably with natural alkaloid space and abut drug space (Fig. 6c). This might be expected to arise naturally from the combination of two different modes of synthetic chemistry: the natural-product inspired routes that led to the key scaffolds versus the diversity-generating steps that followed in the library expansion phase.

Principal moment of inertia (PMI) analysis[46] was also used to compare the three-dimensional shapes of the lowest-energy

conformations of our scaffolds and library members with the above reference sets[45]. Our compounds were found to lie along the rod-disc side of the triangle, with a preference for the rod vertex (Fig. 6d). We note that drugs and the commercial screening libraries represented in the reference sets also reside in this region of the plot[46].

These analyses suggest that our scaffolds and library compounds are similar to natural products, particularly alkaloid natural products, while also sharing attractive properties of drugs believed to be compatible with bioavailability. We note that the fraction of carbon atoms with an sp^3 centre (Fsp^3) in our scaffolds (0.77) and library compounds (0.66) is very similar to the values obtained for our reference natural products (0.64) and alkaloids (0.65), and substantially higher than that for drugs (0.41) (Supplementary Table 2). Alkaloids (average M_w of 319, rotatable bonds 2.8) are generally smaller and more rigid than non-alkaloid natural products (629, 9.7), and more closely resemble our library compounds (355, 4.1). Although we did not particularly set out to adhere to any filters for drug-likeness in our design, we point out that 72% of our library compounds satisfied each of Lipinski's rules of five, with 100% meeting three out of four of the criteria. Moreover, all library members fulfilled Veber's requirements for good oral bioavailability ($\leq$10 rotatable bonds; $\leq$140 Å^2 total polar surface area)[47].

Discussion

We have described one approach to balancing two concerns that arise in the development of libraries for biological screening: (1) the desire to provide compound collections that are different enough from existing libraries to inform interesting new biology, and (2) dealing with the sheer vastness of chemical space, which makes it hard to create useful new bioactive molecules that are strikingly different from typical drug-like libraries. Using four alkaloids as inspiration agents, we used straightforward reaction sequences and moderately advanced synthetic intermediates to generate scaffold structures that were easily converted to libraries using parallel synthesis technology. The particular routes combine azide-mediated methods for the incorporation of nitrogen into organic frameworks with established ketone syntheses. Regarding the latter, we used one very well known reaction for the generation of fused ring systems (Diels–Alder) and a powerful but somewhat underutilized method for generating spirocycles (Trost spiroannulation). Overall, a total of 686 previously unknown structures were synthesized, comprising 55 separate scaffolds and 631 analogues. Of these, 266 (39%) of the compounds contained a basic nitrogen atom (ranging from 21% for the Stemonaceae alkaloid libraries to 68% of those derived from mesembrine). Overall, more than 90% of the library members were obtained in the 20 mg quantities and 90% purities that we had initially targeted (with the remainder achieving the purity goal but only being obtained in 10–20 mg quantities), with all of the compounds being >90% diastereomerically pure. The synthetic routes developed are amenable for use with both hit re-synthesis and downstream structure–activity relationship studies, both critical for real-world applications of small-molecule libraries.

The computational assessment of our libraries shows that the new compounds have many of the attributes that some authors have proposed to arise from using natural products in the first place—notably high sp^3 counts relative to commercial libraries and comparable to those in a previously used natural-product set[48]. This feature naturally arose from the selection of alkaloid-based scaffolds and reflects the choice of targets in the first place. In fact, the slight drop in average sp^3 content in the libraries made versus the scaffolds themselves can be attributed to our selection of common 'medicinal chemistry' subunits (for example, aromatics) for our specific diversification efforts. On the other hand, the observation that the libraries were highly Lipinski- and Veber-compliant may be viewed as a surprise, given the conventional viewpoint that natural products have very different chemical properties from synthetically derived drug scaffolds. Overall, the PCA and PMI analyses support the fact that the primary goal of this project was achieved insofar as we created libraries different enough from highly occupied drug discovery space to be interesting (and thus addressing our mission of 'probing chemical space'), but not so different as to lack potential in screens. The ultimate determination of this potential by screening against a variety of targets (for example, by submission into the Small Molecule Repository of the US National Institutes of Health) and the extension of the concept to other libraries is currently being pursued.

Methods

General procedures for the key Schmidt reactions used in the syntheses of scaffolds **3**, **8**, **9**, **13**, **18** and **21** are described below. All reactions were performed using flame-dried glassware under an argon atmosphere. Additional experimental details and analytical data for scaffold synthesis and library preparation are provided in the Supplementary Methods, together with full details of the cheminformatic analysis. CAUTION: The authors remind all experimentalists contemplating use of these methods to follow established safety protocols in the use of alkyl azides and their precursors.

General procedure for Diels–Alder/Schmidt reaction for preparation of scaffolds 3 (Stemonaceae alkaloid series). Titanium tetrachloride (1 M solution in dichloromethane, 2.5 equiv.) was added dropwise to a solution of azide **2** (1 equiv.) and silyloxydiene **1a**–**d** (2.5 equiv.) in anhydrous dichloromethane (0.06 M w.r.t. azide) at 0 °C under argon. The resulting red/brown solution was stirred at 0 °C for 2 h, then allowed to warm slowly to room temperature overnight. The reaction mixture was then quenched with water and stirred at room temperature for 1 h. The organic layer was removed, and the aqueous extracted with dichloromethane (×3). The combined organics were dried (Na_2SO_4) and concentrated to afford a brown oil. The crude product was purified by chromatography (silica gel, 95:5 ethyl acetate: methanol) to afford lactams **3a**–**d**.

General procedure for azido-alcohol Schmidt reaction for preparation of scaffolds 8 and 9 (cylindricine series). Boron trifluoride diethyl etherate (5 equiv.) was added dropwise to a solution of spiro[3.5]nonan-1-one **6a**–**h** (1 equiv.) in anhydrous dichloromethane (0.15 M w.r.t. spirononanone) at −78 °C under argon. The reaction mixture was stirred for 30 min at −78 °C, then a solution of hydroxyalkyl azide (3 equiv.) in dichloromethane (1.5 M) was added. The reaction mixture was stirred at −78 °C for 3 h, then allowed to warm slowly to room temperature overnight. The reaction mixture was then concentrated under reduced pressure and the resulting oil was dissolved in 15% aqueous potassium hydroxide solution and stirred for 30 min. The reaction mixture was then extracted with dichloromethane (×3). The combined organics were washed with water, dried ($MgSO_4$) and concentrated. The crude product was purified by automated chromatography (silica gel 0−100% ethyl acetate in hexanes) to afford lactams **8** and **9**.

Alkyl azide Schmidt reaction for preparation of scaffold 13 (sparteine series). Titanium tetrachloride (20.7 ml, 188.5 mmol, 5 equiv.) was added dropwise to a solution of azide **12** (10 g, 37.7 mmol, 1.0 equiv.) in anhydrous dichloromethane (350 ml) at 0 °C under argon. A yellow precipitate was formed. The reaction mixture was allowed to warm to room temperature, stirred for 24 h and then quenched with water. The aqueous layer was extracted with dichloromethane. The organic layer was dried (Na_2SO_4) and concentrated to give an oil. The crude product was purified by chromatography (silica gel, 100% ethyl acetate) to afford lactam **13** (4.5 g, 62%) as a colourless solid.

General procedure for Diels–Alder/Schmidt reaction for preparation of scaffolds 18 (mesembrine series). Methyl vinyl ketone (1 equiv.) was added to a solution of silyloxydiene **20** (1.5 equiv.) in anhydrous dichloromethane (0.4 M w.r.t. diene) and the reaction mixture cooled to −78 °C. Boron trifluoride diethyl etherate (1.5 equiv.) was then added and the reaction mixture stirred at −78 °C for 4 h then warmed to room temperature. A further aliquot of boron trifluoride diethyl etherate (2 equiv.) was then added and the reaction mixture stirred at room temperature for an additional 16 h. The reaction mixture was diluted with dichloromethane and quenched with saturated aqueous $NaHCO_3$. The organic layer was washed with saturated aqueous NH_4Cl, water and brine, dried ($MgSO_4$) and concentrated. The crude product was purified by automated column chromatography (silica gel 0 to 100% ethyl acetate in hexanes, then 90:10 dichloromethane:methanol) to afford lactams **18**.

Diels–Alder/Schmidt reaction for preparation of scaffold 21 (mesembrine series). Ethylaluminium dichloride (1 M solution in hexane, 25 ml, 1 equiv.) was added dropwise to a solution of trimethylsilyloxy butadiene **20** (3.5 g, 25 mmol, 2.5 equiv.) and azide **19** (1.4 g, 10 mmol, 1 equiv.) in anhydrous dichloromethane (15 ml) under argon at −78 °C. After the addition was complete, the reaction

mixture was stirred at −78 °C for 2 h then gradually warmed to room temperature and stirred for a further 16 h, then quenched with water and diluted with dichloromethane (100 ml). The combined organics were washed with saturated aqueous $NaHCO_3$ (50 ml) and brine (50 ml), dried (Na_2SO_4) and concentrated. The crude product was purified by chromatography (silica gel 100% ethyl acetate) to afford lactam **19** (800 mg, 45%) as a light yellow oil.

Received 22 July 2013; accepted 5 December 2013; published online 19 January 2014

References

1. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* **70,** 461–477 (2007).
2. Butler, M. S. Natural products to drugs: natural product-derived compounds in clinical trials. *Nat. Prod. Rep.* **25,** 475–516 (2008).
3. Paterson, I. & Anderson, E. A. Chemistry. The renaissance of natural products as drug candidates. *Science* **310,** 451–453 (2005).
4. Ganesan, A. The impact of natural products upon modern drug discovery. *Curr. Opin. Chem. Biol.* **12,** 306–317 (2008).
5. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **23,** 3–25 (1997).
6. Boldi, A. M. Libraries from natural product-like scaffolds. *Curr. Opin. Chem. Biol.* **8,** 281–286 (2004).
7. Thomas, G. L. & Johannes, C. W. Natural product-like synthetic libraries. *Curr. Opin. Chem. Biol.* **15,** 516–522 (2011).
8. Camp, D., Davis, R. A., Evans-Illidge, E. A. & Quinn, R. J. Guiding principles for natural product drug discovery. *Future Med. Chem.* **4,** 1067–1084 (2012).
9. Lachance, H., Wetzel, S., Kumar, K. & Waldmann, H. Charting, navigating, and populating natural product chemical space for drug discovery. *J. Med. Chem.* **55,** 5989–6001 (2012).
10. Zuegg, J. & Cooper, M. A. Drug-likeness and increased hydrophobicity of commercially available compound libraries for drug screening. *Curr. Top. Med. Chem.* **12,** 1500–1513 (2012).
11. Njardarson, J. T., Gaul, C., Shan, D., Huang, X-Y. & Danishefsky, S. J. Discovery of potent cell migration inhibitors through total synthesis: lessons from structure–activity studies of (+)-migrastatin. *J. Am. Chem. Soc.* **126,** 1038–1040 (2004).
12. Szpilman, A. M. & Carreira, E. M. Probing the biology of natural products. Molecular editing by diverted total synthesis. *Angew. Chem. Int. Ed.* **49,** 9592–9628 (2010).
13. Wetzel, S., Bon, R. S., Kumar, K. & Waldmann, H. Biology-oriented synthesis. *Angew. Chem. Int. Ed.* **50,** 10800–10826 (2011).
14. Burke, M. D. & Schreiber, S. L. A planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed.* **43,** 46–58 (2004).
15. Huigens, R. W. *et al.* A ring-distortion strategy to construct stereochemically complex and structurally diverse compounds from natural products. *Nature Chem.* **5,** 195–202 (2013).
16. Kim, Y. K. *et al.* Relationship of stereochemical and skeletal diversity of small molecules to cellular measurement space. *J. Am. Chem. Soc.* **126,** 14740–14745 (2004).
17. Shelat, A. A. & Guy, R. K. Scaffold composition and biological relevance of screening libraries. *Nature Chem. Biol.* **3,** 442–446 (2007).
18. Lovering, F., Bikker, J. & Humblet, C. Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* **52,** 6752–6756 (2009).
19. Clemons, P. A. *et al.* Quantifying structure and performance diversity for sets of small molecules comprising small-molecule screening collections. *Proc. Natl Acad. Sci. USA* **108,** 6817–6822 (2011).
20. Lovering, F. Escape from flatland 2: complexity and promiscuity. *MedChemComm* **4,** 515–519 (2013).
21. Nandy, J. P. *et al.* Advances in solution- and solid-phase synthesis toward the generation of natural product-like libraries. *Chem. Rev.* **109,** 1999–2060 (2009).
22. MacLellan, P. & Nelson, A. A conceptual framework for analysing and planning synthetic approaches to diverse lead-like scaffolds. *Chem. Commun.* **49,** 2383–2393 (2013).
23. Evans, B. E. *et al.* Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* **31,** 2235–2246 (1988).
24. DeSimone, R. W., Currie, K. S., Mitchell, S. A., Darrow, J. W. & Pippin, D. A. Privileged structures: applications in drug discovery. *Comb. Chem. High T. Scr.* **7,** 473–493 (2004).
25. Costantino, L. & Barlocco, D. Privileged structures as leads in medicinal chemistry. *Curr. Med. Chem.* **13,** 65–85 (2006).
26. Duarte, C. D., Barreiro, E. J. & Fraga, C. A. M. Privileged structures: a useful concept for the rational design of new lead drug candidates. *Mini-Rev. Med. Chem.* **7,** 1108–1119 (2007).
27. Costantino, L. & Barlocco, D. Privileged structures as leads in medicinal chemistry. *Front. Med. Chem.* **5,** 381–422 (2010).
28. Verma, A. *et al.* Nitrogen-containing privileged structures and their solid phase combinatorial synthesis. *Comb. Chem. High T. Scr.* **16,** 345–393 (2013).
29. Frankowski, K. J. *et al.* Synthesis and receptor profiling of *Stemona* alkaloid analogues reveal a potent class of sigma ligands. *Proc. Natl Acad. Sci. USA* **108,** 6727–6732 (2011).
30. Pilli, R. A., Rosso, G. B. & de Oliveira, M. d. C. F. The *Stemona* alkaloids. *Alkaloids* **62,** 77–173 (2005).
31. Pilli, R. A., Rosso, G. B. & de Oliveira, M. d. C. F. The chemistry of *Stemona* alkaloids: an update. *Nat. Prod. Rep.* **27,** 1908–1937 (2010).
32. Weinreb, S. M. Studies on total synthesis of the cylindricine/fasicularin/lepadiformine family of tricyclic marine alkaloids. *Chem. Rev.* **106,** 2531–2549 (2006).
33. Gericke, N. & Viljoen, A. M. Sceletium—a review update. *J. Ethnopharm.* **119,** 653–663 (2008).
34. Perez, E. G., Mendez-Galvez, C. & Cassels, B. K. Cytisine: a natural product lead for the development of drugs acting at nicotinic acetylcholine receptors. *Nat. Prod. Rep.* **29,** 555–567 (2012).
35. Daly, J. W. Nicotinic agonists, antagonists, and modulators from natural sources. *Cell. Mol. Neurobiol.* **25,** 513–552 (2005).
36. Wrobleski, A., Coombs, T. C., Huh, C. W., Li, S-W. & Aubé, J. The Schmidt reaction. *Org. React.* **78,** 1–320 (2012).
37. Zeng, Y., Reddy, D. S., Hirt, E. & Aube, J. Domino reactions that combine an azido–Schmidt ring expansion with the Diels–Alder reaction. *Org. Lett.* **6,** 4993–4995 (2004).
38. Gracias, V., Frank, K. E., Milligan, G. L. & Aubé, J. Ring expansion by *in situ* tethering of hydroxy azides to ketones: the Boyer reaction. *Tetrahedron* **53,** 16241–16252 (1997).
39. Smith, B. T., Wendt, J. A. & Aube, J. First asymmetric total synthesis of (+)-sparteine. *Org. Lett.* **4,** 2577–2579 (2002).
40. Meyer, A. M., Katz, C. E., Li, S-W., Vander Velde, D. & Aube?, J. A tandem Prins/Schmidt reaction approach to marine alkaloids: formal and total syntheses of lepadiformines A and C. *Org. Lett.* **12,** 1244–1247 (2010).
41. Fenster, E. *et al.* Three-component synthesis of 1,4-diazepin-5-ones and the construction of γ-turn-like peptidomimetic libraries. *J. Comb. Chem.* **10,** 230–234 (2008).
42. Trost, B. M. & Bogdanowicz, M. J. New synthetic reactions. X. Versatile cyclobutanone (spiroannelation) and γ-butyrolactone (lactone annelation) synthesis. *J. Am. Chem. Soc.* **95,** 5321–5334 (1973).
43. Li, X. & Danishefsky, S. J. Cyclobutenone as a highly reactive dienophile: expanding upon Diels–Alder paradigms. *J. Am. Chem. Soc.* **132,** 11004–11005 (2010).
44. Xue, L., Stahura, F. & Bajorath, J. in *Chemoinformatics* Vol. 275 (ed. Jürgen Bajorath) Ch. 9, 279–289 (Humana, 2004).
45. Kopp, F., Stratton, C. F., Akella, L. B. & Tan, D. S. A diversity-oriented synthesis approach to macrocycles via oxidative ring expansion. *Nature Chem. Biol.* **8,** 358–365 (2012).
46. Sauer, W. H. B. & Schwarz, M. K. Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. *J. Chem. Inf. Comput. Sci.* **43,** 987–1003 (2003).
47. Veber, D. F. *et al.* Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **45,** 2615–2623 (2002).
48. Bauer, R. A., Wurst, J. M. & Tan, D. S. Expanding the range of 'druggable' targets with natural product-based libraries: an academic perspective. *Curr. Opin. Chem. Biol.* **14,** 308–314 (2010).

Acknowledgements

The authors thank P. Porubsky and B. Neuenswander for purification of library compounds. The authors also acknowledge financial support from the US Institute of General Medical Sciences (P41 GM089164, Pilot-scale Libraries Initiative; 5P50GM069663, KU Chemical Methodologies and Library Development Center) and an NSF-MRI grant (CHE-0923449) for the purchase of an X-ray diffractometer.

Author contributions

M.C.M., J.N.P., D.R., G.S. and J.A. designed the experiments and analysed the data. M.C.M., J.N.P., D.R. and G.S. performed the synthesis and characterization. J.L.W. and M.C.M. performed the cheminformatic analysis and V.W.D. performed and analysed the X-ray structures. M.C.M. and J.A. wrote the manuscript.

Additional information

Supplementary information and chemical compound information are available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.A.

Competing financial interests

The authors declare no competing financial interests.