nature chemistry

ARTICLES

PUBLISHED ONLINE: 20 JANUARY 2013 | DOI: 10.1038/NCHEM.1549

A ring-distortion strategy to construct stereochemically complex and structurally diverse compounds from natural products

Robert W. Huigens III, Karen C. Morrison, Robert W. Hicklin, Timothy A. Flood Jr, Michelle F. Richter and Paul J. Hergenrother*

High-throughput screening is the dominant method used to identify lead compounds in drug discovery. As such, the makeup of screening libraries largely dictates the biological targets that can be modulated and the therapeutics that can be developed. Unfortunately, most compound-screening collections consist principally of planar molecules with little structural or stereochemical complexity, compounds that do not offer the arrangement of chemical functionality necessary for the modulation of many drug targets. Here we describe a novel, general and facile strategy for the creation of diverse compounds with high structural and stereochemical complexity using readily available natural products as synthetic starting points. We show through the evaluation of chemical properties (which include fraction of sp^3 carbons, ClogP and the number of stereogenic centres) that these compounds are significantly more complex and diverse than those in standard screening collections, and we give guidelines for the application of this strategy to any suitable natural product.

Collections of small molecules are used routinely in high-throughput screens to find new drug leads. In fact, from 1999 to 2008, 45 of the 50 Food and Drug Administration approvals for first-in-class small-molecule new molecular entities originated from a screen[1]. The composition of compound-screening collections, therefore, has a significant impact on the types of drugs that come to market and the efficiency by which next-generation therapeutics are developed.

Many high-throughput screening (HTS) success stories involve biological targets that can be modulated by low molecular weight and relatively planar organic compounds of high sp^2 character and low, if any, stereochemical complexity. For example, kinases are outstanding drug targets whose enzymatic activity is typically inhibited at the ATP binding site by organic compounds with no stereogenic centres and high aromatic content[2,3]. However, for more complex biological targets, tremendous challenges in lead identification still exist. For example, disruptors of protein–protein interactions[4] and inhibitors of transcription factors[5] are rarely small planar compounds with little stereochemical complexity. Thus, most HTS campaigns versus these targets using compounds present in standard screening collections will fail. In addition, compounds active in certain therapeutic areas (for example, antibacterials) tend to be larger and more complex than the average screening compound[6]. For these reasons, and many others, there is a pressing need for the creation of compounds that are structurally complex and diverse[7].

Recognizing this need, creative strategies to rapidly generate collections of complex molecules have appeared. One approach is diversity-oriented synthesis, in which simple starting materials are coupled to make diverse structures that are more like natural products in terms of size, percentage of sp^3 carbons and number of stereogenic centres[8–12]. Other methods include the synthesis of scaffolds inspired by natural products that can be decorated efficiently and differentially[13,14], skeletal diversifications[15–17], use of building blocks derived from natural products for combinatorial synthesis[18], biology-oriented synthesis[19] and the synthesis of chiral and conformationally constrained oligomers[20].

Here we report a new approach for the rapid creation of complex and diverse small molecules. In this process structurally complex natural products are converted, in an average of three chemical steps, into markedly different core scaffolds that are distinct from each other and from the parent natural product. Using chemoselective reactions, the core ring structures of readily available natural products are altered systematically via reactions that distort the ring system (for example, ring cleavage, ring expansion, ring fusion, ring rearrangements and combinations thereof (Fig. 1)). Importantly, this method stands in contrast to traditional optimization campaigns with goals to enhance the inherent biological activity or improve drug-like properties of a natural product (for example, erythromycin to azithromycin, penicillin to amoxicillin

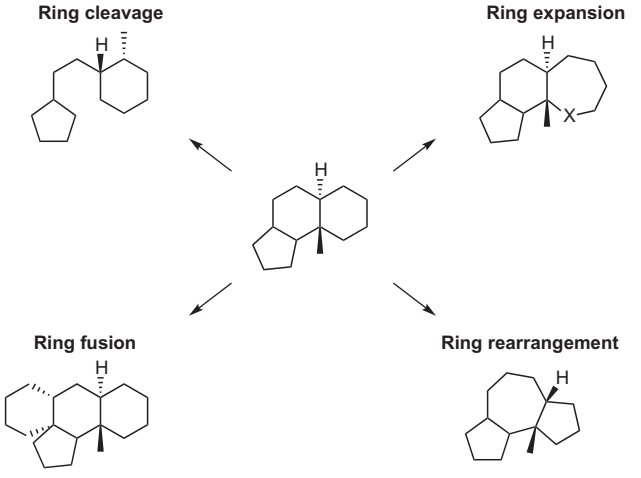

Figure 1 | Ring-distortion reactions can be used readily to convert natural products into complex and diverse scaffolds.

University of Illinois, Department of Chemistry, Roger Adams Laboratory, 600 South Mathews, Urbana, Illinois 61801, USA. *e-mail: hergenro@illinois.edu

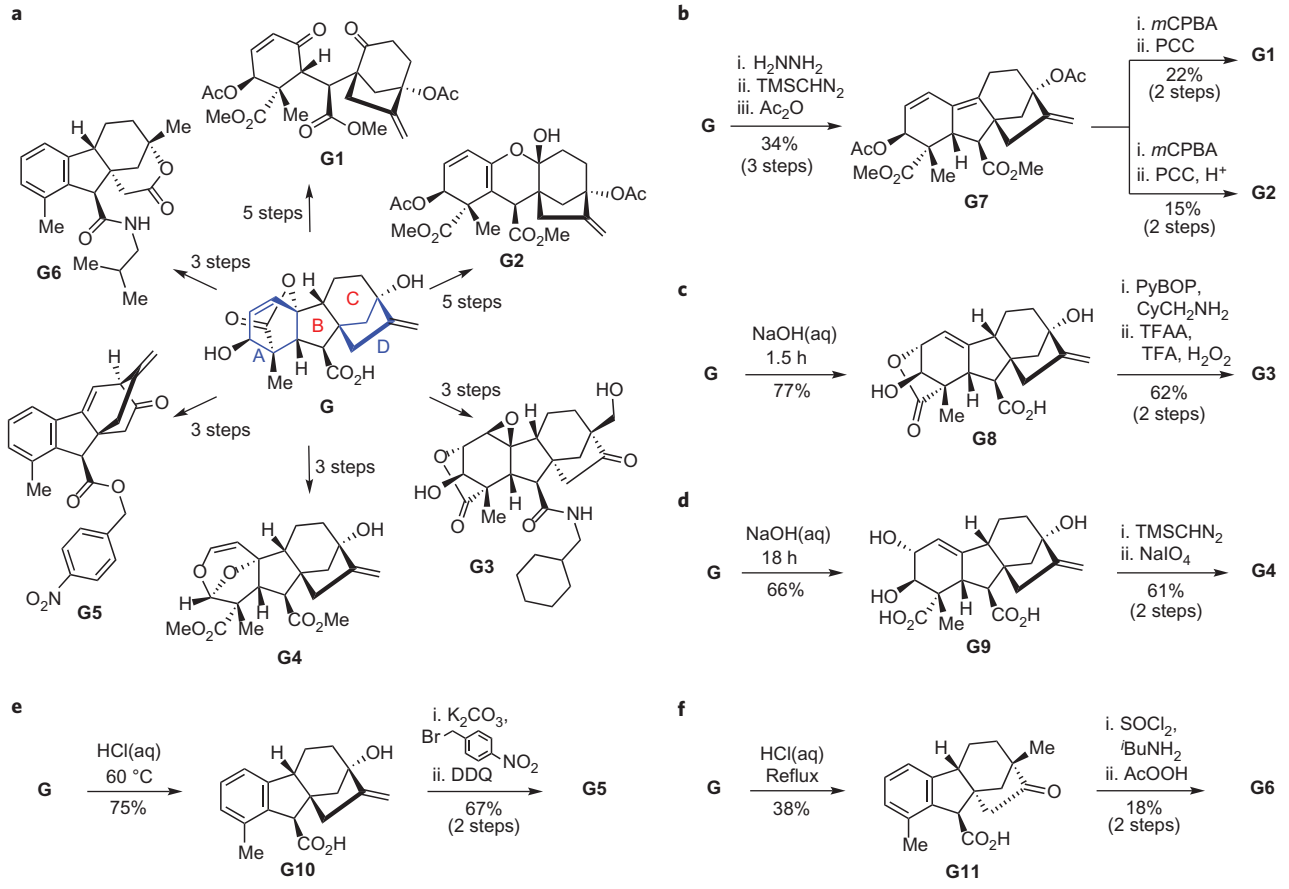

Figure 2 | Application of ring-distortion reactions in the synthesis of complex and diverse small molecules from G. a, Sequences of between three and five steps can be used to convert gibberellic acid into the six structures G1–G6. **b**, Synthetic route to G1 and G2 from **G** using an oxidative cleavage of the B ring. **c**, Synthetic route to G3 from **G** using a base-catalysed lactone rearrangement. **d**, Synthetic route to G4 from **G** using a base-catalysed lactone rearrangement and opening. **e**, Synthetic route to G5 from **G** using an acid-catalysed aromatization and an oxidative ring rearrangement. **f**, Synthetic route to G6 from **G** using an acid-catalysed aromatization and ring rearrangement, followed by oxidative ring expansion. Cy = cyclohexyl, PyBOP = benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate, TFAA = trifluoroacetic anhydride, TFA = trifluoroacetic acid, TMS = trimethylsilyl.

and so on). To demonstrate this ring-distortion strategy, we selected three readily available and well-studied natural products from different structural classes: the diterpene gibberellic acid (**G**), the steroid adrenosterone (**A**) and the alkaloid quinine (**Q**) (Figs 2– 4, respectively). However, dozens of readily available natural products could be converted into diverse and complex molecules using this strategy. This method takes inspiration from the manner in which nature creates certain complex natural products using a common intermediate to generate scores of compounds that are very different from one another[21].

Results and discussion

Diversifying gibberellic acid. G (Fig. 2) is a plant hormone isolated from *Gibberella fujikuroi*[22] and produced industrially[23] on the tonne scale. **G** contains a tetracyclic diterpene core with a fused lactone, two allylic alcohols, an exocyclic olefin and a carboxylic acid, which enables the selective and independent functionalization of each ring of the core structure via a variety of reactions that distort the ring system. We exploited these structural features in concert with known degradation reactions of **G** (refs 24–27) in the construction of complex and diverse scaffolds in three to five steps from **G** (Fig. 2, G1–G6).

Hydrazine-promoted elimination of the lactone on **G** (ref. 27) followed by methylation and acetylation affords triene G7. Treatment with *meta*-chloroperbenzoic acid (*m*CPBA) yields an intermediate epoxide with complete selectivity for the tetrasubstituted olefin which, when subjected to oxidative cleavage conditions using pyridinium chlorochromate (PCC), produces diketone G1. With exposure of G1 to silica or acid, the A-ring ketone tautomerizes and collapses onto the C-ring ketone to form ketal G2; this can be achieved directly from the epoxide precursor using PCC and an acidic workup.

Exposure of **G** to basic conditions leads to a lactone rearrangement and the generation of alkene G8 (ref. 26). Amidation of G8, followed by treatment with trifluoroperacetic acid (generated *in situ*), provides G3 via epoxidation of both alkenes and a Wagner–Meerwein rearrangement to afford the primary alcohol. Prolonged exposure of **G** to base leads to the cleavage of the lactone ring to provide diol G9 (ref. 28). Methylation of the carboxylic acids followed by oxidative cleavage of the diol with sodium periodate and intramolecular [4 + 2] cycloaddition provides acetal G4.

Treatment of **G** with dilute hydrochloric acid results in the elimination of the lactone and decarboxylation to aromatize the A ring, which enables the isolation of *allo*-gibberic acid (G10)[24]. Esterification followed by oxidative rearrangement with 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ)[29] gives G5. Exposure of **G** to refluxing hydrochloric acid results in aromatization and a Wagner–Meerwein rearrangement to form gibberic acid (G11)[24]. Amidation through an intermediate acyl chloride followed by Baeyer–Villiger oxidation produces lactone G6.

Diversifying adrenosterone. A (Fig. 3) is a steroid hormone produced in the adrenal cortex of mammals[30]. **A**'s structurally complex steroidal framework contains five contiguous stereogenic

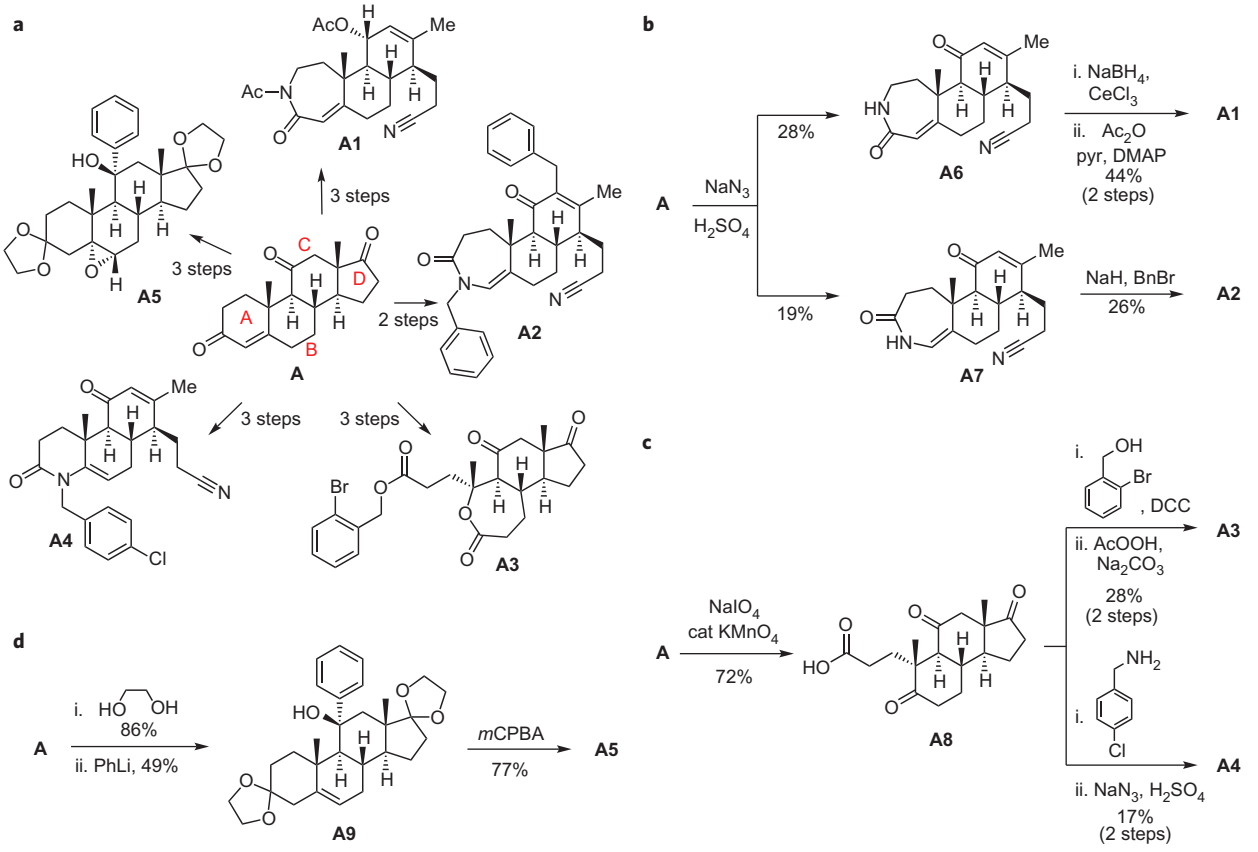

Figure 3 | Application of ring-distortion reactions in the synthesis of complex and diverse small molecules from A. a, Sequences of between two and three steps can be used to convert adrenosterone into the five structures A1–A5. **b**, Novel Schmidt reaction of **A** and synthetic elaboration for the conversion of **A** into A1 and A2. **c**, Oxidative cleavage of the A-ring of **A** and further elaboration using the Baeyer–Villiger or Schmidt reaction pathways for the rapid conversion of **A** into A3 and A4. **d**, Ring-fusion reaction pathway that enables the conversion of **A** into A5. cat = catalytic, pyr = pyridine.

centres; in addition, each of the four individual carbocyclic rings of **A** is functionalized with an enone or ketone. Although embedded in the A ring, the enone is also connected to the B ring as an exocyclic double bond, and the C and D rings are each functionalized with a ketone. These key functional groups provide synthetic handles for known chemical reactivity and can be manipulated strategically to synthesize novel, diverse and complex chemical scaffolds in three or fewer steps (Fig. 3, A1–A5).

During these synthetic investigations of **A**, a novel substrate-dependent Schmidt reaction was discovered that effected both ring expansion and ring cleavage in a single synthetic transformation. Subjecting **A** to Schmidt conditions for one hour gives two constitutional isomers (A6 and A7 in Fig. 3) that result from a tandem D-ring cleavage and A-ring expansion. Final dehydration of the subsequent primary amide in concentrated sulfuric acid results in the observed cyano groups in A6 and A7.

Enamide A6 and lactam A7 were each elaborated to give novel complex molecular scaffolds. Lactam A6 undergoes a stereoselective Luche reduction of the C-ring enone, which after treatment with acetic anhydride in pyridine with catalytic 4-dimethylaminopyridine (DMAP) affords A1. Treatment of enamide A7 with sodium hydride followed by benzyl bromide results in enone A2.

Although not known specifically for **A**, oxidative cleavage of certain steroidal A-ring enones can be effected using $NaIO_4$ and catalytic $KMnO_4$ (ref. 31). The oxidative cleavage of **A**'s A ring with $NaIO_4$ and $KMnO_4$ gives acid A8, which on treatment with 2-bromobenzyl alcohol and *N*,*N*′-dicyclohexylcarbodiimide (DCC) provides the corresponding ester. A selective ring expansion at the B-ring ketone using a Baeyer–Villiger reaction with peracetic acid provides lactone A3. Acid A8 condenses with 4-chlorobenzylamine on heating in ethanol to provide the corresponding A-ring substituted enamide. A final ring-cleavage reaction of the D ring using our Schmidt protocol yields A4.

A undergoes a double-ring fusion reaction (at the A and D rings) on treatment with ethylene glycol and catalytic *p*-toluenesulfonic acid[32]. The resulting ketone reacts with phenyllithium to give **A9** (refs 33,34). Final treatment of A9 with *m*CPBA results in epoxide-ring fusion at the B ring to yield A5.

Diversifying quinine. Q (Fig. 4), an alkaloid isolated from the bark of the genus *Cinchona*, is available in high purity and low cost through its use as an antimalarial therapeutic, food additive and catalyst scaffold[35]. Unlike other natural products employed herein, **Q** is composed of two discrete ring systems; however, the stereochemical complexity and diverse functionality (a tertiary amine, a secondary alcohol, an olefin and a quinoline) of **Q** make it amenable to selective ring-system distortion to create diverse molecular scaffolds (Fig. 4, Q1–Q5).

In the course of these investigations we discovered an unprecedented tandem ring cleavage/ring fusion of **Q** effected by treatment with thionochloroformate. Ring cleavage of the quinuclidine by *O*-phenyl thionochloroformate occurs selectively at N1–C2. In addition to the expected ring cleavage and chloride addition, this reaction also leads to a diastereoselective rearrangement of the free alcohol and thiocarbamate to form the *S*-alkyl thiocarbamate Q1 as a single diastereomer, as confirmed by X-ray crystallography.

In contrast to the selectivity observed with *O*-phenyl thionochloroformate, an acid-catalysed Hofmann-type elimination of **Q** is known to occur exclusively at N1–C8 (ref. 36), and the addition of benzyl chloroformate to the crude degradation product results

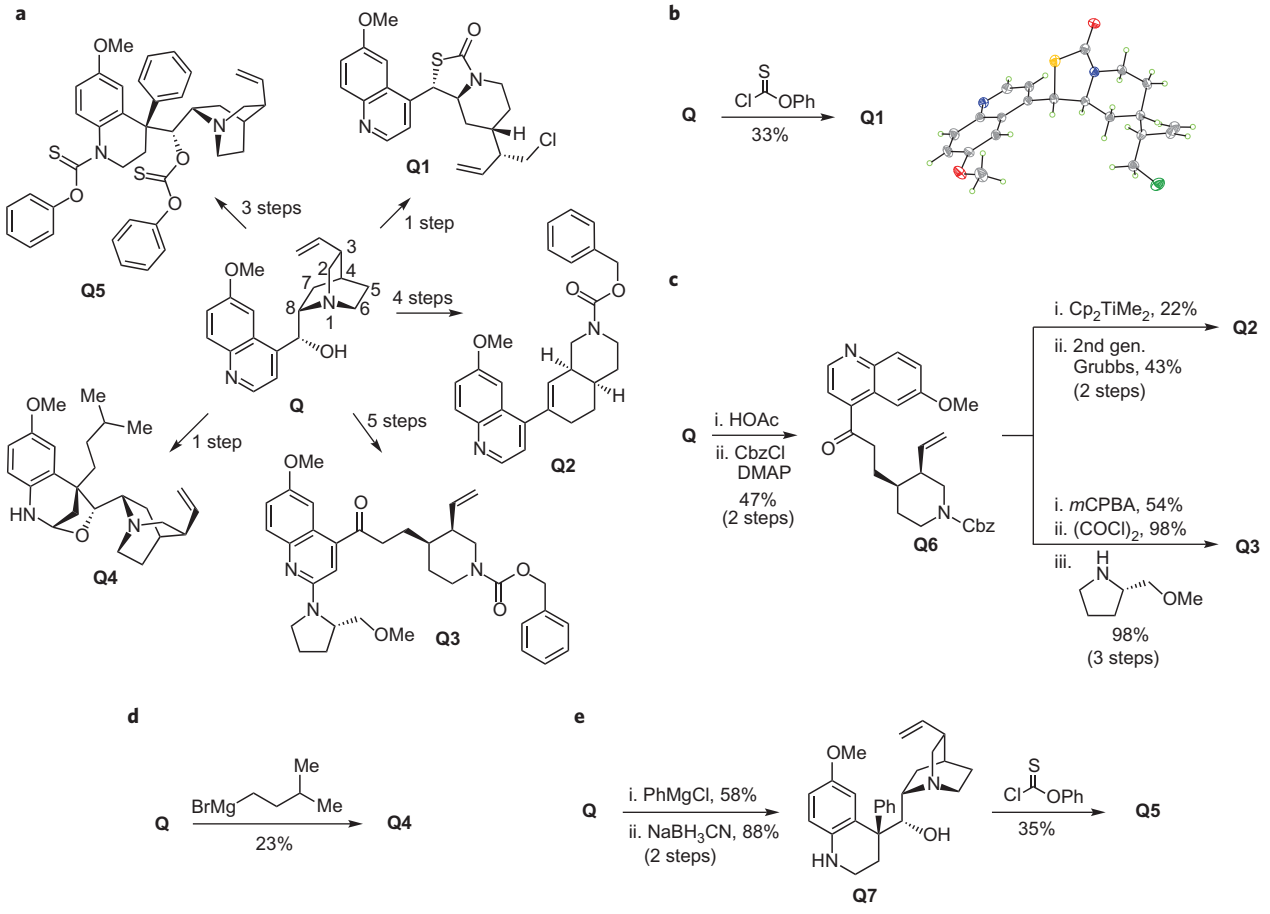

Figure 4 | Application of ring-distortion reactions in the synthesis of complex and diverse small molecules from Q. a, Sequences of between one and five steps can be used to convert quinine into the five structures Q1–Q5. **b**, Cleavage of the quinuclidine ring and rearrangement forms Q1 (the crystal structure is shown). **c**, Acid-catalysed ring opening of the quinuclidine ring produces ketone Q6. Olefination using the Petasis reagent followed by ring-closing metathesis leads to Q2. Alternatively, *N*-oxidation of the Q6 quinoline followed by rearrangement, chlorination and substitution with an amine leads to Q3. **d**, Addition of isoamylmagnesium bromide leads to alkylation and hemiaminal ether formation to provide Q4. **e**, Grignard addition and reductive ring opening of the resulting hemiaminal ether produces tetrahydroquinoline Q7. Exposure of Q7 to *O*-phenyl thionochloroformate results in the formation of Q5. Cbz = benzyloxycarbonyl, Cp = cyclopentadienyl.

in ketone Q6, which was elaborated to form two unique structures (Q2 and Q3). Petasis methylenation of ketone Q6 followed by 1,2-ring fusion via ring-closing metathesis using second-generation Grubbs catalyst forms [4.4.0]-bicycle Q2. Quinoline *N*-oxidation of Q6 using *m*CPBA, chlorination with oxalyl chloride and nucleophilic displacement of the chloride by (*S*)-2-(methoxymethyl)pyrrolidine provides amine Q3.

Further ring-system distortions of the quinoline ring were accomplished through the addition of Grignard reagents[37]. Exposure of **Q** to isoamylmagnesium bromide in toluene results in nucleophilic addition to the quinoline ring followed by hemiaminal ether formation to provide Q4 as a single diastereomer. Alternatively, reduction of the hemiaminal ether formed through the addition of phenylmagnesium chloride to **Q** with sodium cyanoborohydride provides tetrahydroquinoline Q7. Treatment of Q7 with *O*-phenyl thionochloroformate results in bis-acylation to form Q5 as the major product with no observed chlorine incorporation.

Compound analysis. In contrast to most standard small-molecule library constructions in which simple starting materials are built up into more complex products, an important consequence of starting with natural products is that all the intermediates are complex structurally and worthy of inclusion in the final library in their own right. For example, in the course of synthesizing the compounds depicted in Figs 2–4 a total of 19, 18 and 12 complex structures were produced, respectively (all the structures are shown in Supplementary Fig. S1). Thus, detailed in Figs 2–4 is the synthesis of 49 structurally and stereochemically complex small molecules from three readily available natural products. Most of these compounds are created in good yield and, importantly, all the compounds are created on a multimilligram scale, which allows a full structural characterization (see Supplementary Information) and multiple biological screens. In addition, each of these compounds possesses sites for diversification, which allows for the facile and rapid creation of dozens of complex compounds (as described below).

Advances in chemoinformatics enable the evaluation of massive chemical and biological data sets, which allows for a rough determination of the structural features of small molecules that correlate with biological activity. It is apparent that many compounds in screening collections have non-trivial liabilities, including non-specific reactivity and a propensity to aggregate, which lead to false positives and complicate the development of a hit into a drug[38]. When specific disease areas are examined, the problem is more acute. For example, analysis of compounds that kill Gram-negative bacterial pathogens shows an average ClogP of −0.1 (ref. 39), a realm occupied by vanishingly few compounds in commercial screening collections.

In an attempt to quantify the structural complexity and diversity of the novel compounds created through this paradigm, structural features known to track with biological activity were analysed. A

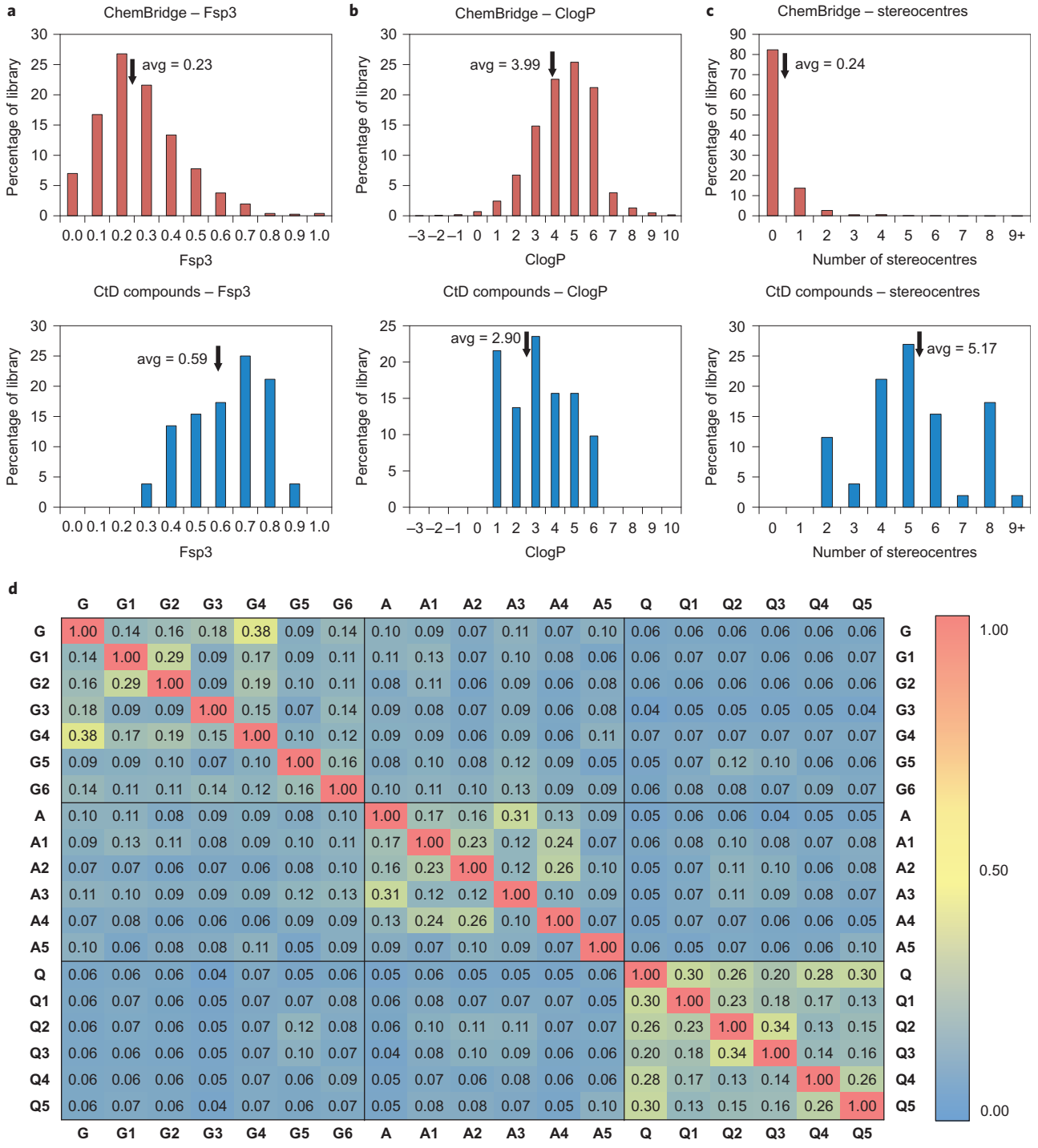

Figure 5 | Comparison of compounds created by the CtD method with commercial screening collections. a–c, For this analysis, a 150,000-compound collection from ChemBridge Corporation was compared to the 49 CtD compounds synthesized (Figs 2–4) for Fsp3 (**a**), ClogP (**b**) and number of stereogenic centres per compound (**c**). **d**, Tanimoto similarity coefficients for CtD compounds (from Figs 2a, 3a and 4a) relative to the three natural products and to each other (1.0 represents perfect similarity). For a full Tanimoto-matrix analysis on the 49-compound set, see Supplementary Fig. S3. Thus, compounds created through CtD have markedly different properties from those in commercial screening collections, and are structurally diverse from one another and from the parent natural products. avg = average.

recent study examined eight structural parameters (molecular weight, ClogP, polar surface area, rotatable bonds, hydrogen-bond donors and acceptors, and complexity and fraction of sp^3-hybridized carbons (Fsp3)) of compounds synthesized by medicinal chemists over the past 50 years, and then compared them to marketed drugs[40]. One of the most important conclusions from this analysis is that medicinal chemists are creating compounds with lower-than-ideal Fsp3 values, and with ClogP values that are higher than ideal. We calculated Fsp3 and ClogP of the compounds described herein and compared them to those of compounds in large screening collections. For this analysis we used a 150,000-member compound collection from the ChemBridge MicroFormat Library (http://chembridge.com), a standard commercial screening collection and one used by others in comparison analyses[41].

Fsp3 is the number of sp^3-hybridized carbon atoms in a compound divided by the sum of the sp^3- and sp^2-hybridized carbon atoms[42]. Studies show the benefits of a higher Fsp3, which include lower melting points and enhanced aqueous solubility[43], and reveal that 'discovery' compounds have lower Fsp3 values than those of actual drugs (0.36 versus 0.47) (ref. 43). An analysis of medicinal compounds synthesized over the past 50 years shows that the average Fsp3 is declining[40,44], a result attributed in part to the increasing ease of sp^2–sp^2 coupling reactions[43]. ClogP is often used as a rough measure of lipophilicity; among other things, compounds with higher ClogP values tend to have non-ideal solubility, promiscuity and off-target toxicity[45]. Analysis shows that the average ClogP for all medicinal compounds synthesized since 1985 increased significantly, and is higher than the average for marketed drugs[40,45]. Indeed, a survey of 18 pharmaceutical companies from 2000 to 2010 shows the majority are still synthesizing compounds with a mean ClogP >4 (ref. 44). As shown by the histograms in Fig. 5a, the new compounds described herein have an average Fsp3 of 0.59, which is considerably higher than that in the commercial collection (0.23). In a similar vein, the average ClogP for these compounds is 1.1 log units lower than that in the commercial screening set (2.90 versus 3.99, Fig. 5b), which corresponds to a 12-fold reduction in hydrophobicity.

The presence of stereogenic centres in a compound can also be used as a surrogate for molecular complexity. Compounds with stereogenic centres may interact more specifically with their chiral receptors, and compounds with low or no stereogenic centres are more prone to attrition during the various stages of drug discovery[43]. Commercial screening collections are dominated by achiral compounds; for example, of the 150,000 compounds in the ChemBridge collection, 82% have no stereogenic centres and 14% have a single stereocentre, which leaves only 4% of these compounds with multiple stereogenic centres (Fig. 5c). Obviously, the synthesis of complex and diverse compounds using natural products as input materials offers a tremendous advantage in this regard. Of the 49 compounds disclosed herein, all have two or more stereogenic centres, with the median number being five (Fig. 5c).

Although visual inspection of the structures in Figs 2–4 readily reveals considerable structural diversity, we applied a similarity metric to make more quantitative comparisons. For this, Tanimoto coefficients[46] were generated in Discovery Studio (Accelrys) using ECFP_6 molecular fingerprints[47]. Final structures in Figs 2–4 were used as the reference input for every other compound in the set, and a similarity score was obtained for each pair on a scale from 0 to 1, with 1 representing perfect similarity. As shown by the data in Fig. 5d, the final compounds depicted in Figs 2–4 are very different, both from one another and from the parent natural products. The **G**, **A** and **Q** compound sets are expected to be quite different from one another; however, even within the sets the compounds show low Tanimoto coefficients (average for the **G** set is 0.15, that for the **A** set is 0.15 and that for the **Q** set is 0.22), which indicates considerable structural diversity. For calibration purposes, this analysis was also performed on structures that represent simple modifications to the parent compounds. As expected, these minor structural changes afford higher similarity scores (average of 0.67, see Supplementary Fig. S2), consistent with the work of others using Tanimoto coefficients[48]. For the similarity matrix of the full 49-compound set, see Supplementary Fig. S3.

To demonstrate that traditional derivatization strategies can be applied even to these highly complex compounds that contain an array of chemical moieties, small libraries were synthesized based on 12 of the 49 compounds. As shown in Supplementary Fig. S4, small collections of imides, *N*-benzylated amides, aryl amides, amides, lactones, secondary and tertiary alcohols, epoxides, triazoles, ureas and sulfonamides were created readily from these 12 small molecules, and in this manner an additional 119 highly complex compounds were synthesized.

Guidelines. This new complexity-to-diversity (CtD) approach is demonstrated here for three complex natural products; however, the same logic and methods can be applied to a multitude of other natural products. In general, the compounds most amenable to diversification through this method are available in suitable quantities (either from commercial sources or through isolation) and possess orthogonal functional groups that allow for ring-system distortion and diversification through chemoselective reactions. As exemplified with **G**, **A** and **Q**, certain common ring-distortion strategies facilitate the rapid diversification (in ≤5 synthetic steps) of complex natural products:

- Ring-cleavage reactions enable dramatic structural changes in one chemical step, and they are the most utilized tactic in our CtD approach. A benefit of ring-cleavage reactions is that typically they provide new functional groups that can be further diversified. Examples include the base-promoted hydrolysis of the lactone on isogibberellic acid (G9), oxidative cleavage on **A** (A8) and N–C cleavage on **Q** (Q1 and Q6).
- Ring-expansion reactions are useful in forming novel ring skeletons or as a prelude to ring-cleavage reactions. There are several options for chemical reactions that induce ring expansion, with the Baeyer–Villiger and Schmidt reactions being powerful methods to target ketone and enone functionalities for ring expansion. Ring expansion, as a tactic for ring-system distortion, was applied successfully in the synthesis of target structures G6, A3, A6 and A7.
- Ring-fusion reactions can provide further diversification by connecting disparate structural elements in the pre-existing ring system or by the simple addition of a new constrained ring to the ring system. Various modes of ring fusion were used to demonstrate this tactic on each natural product. For example, the ring-closing metathesis product Q2 (1,2-ring fusion), formation of the [4 + 2] cycloaddition product G4 and formation of the bis-ketal A9 (1,1-ring fusion) result from ring-fusion reactions. Ring-substitution reactions can also be an example of ring fusion. In the example of ring substitution, the composition of the ring changes without altering the ring size, as exemplified by the creation of the A-ring enamide in A4.
- Ring-rearrangement reactions that dramatically reorganize the core structure are dictated by the natural product and are thus applicable on a case-by-case basis. This tactic is illustrated with **G** in the Wagner–Meerwein rearrangements of G3 and G11, or the DDQ oxidation of G10. These transformations are facilitated by the propensity of the tertiary alcohol in **G**'s C ring to form a ketone on carbon migration, which alters the molecular topology.

Conclusion

Historically, natural products or their close analogues were considered as end points in the drug-discovery process. Indeed, this thinking is quite fruitful; for example, 41% of anticancer drugs and 65% of antibacterial drugs are natural products or very close derivatives thereof[49]. The features that make natural products different from most synthetic compounds (for example, high Fsp3, low ClogP and the presence of stereogenic centres) give these compounds a propensity to bind to their macromolecular target with a high affinity and specificity, and yet still retain the solubility and cell permeability needed for a therapeutic agent. The approach described herein uses natural products not as the end point, but as the initial point for the discovery process that starts with compounds inherently biased for biological success and systematically

transforms them into diverse compounds of equal complexity. Beyond the three demonstrations herein, Supplementary Fig. S5 presents ten other complex and readily available natural products that are highly suited to manipulation from complexity to diversity, and this strategy is generalizable to scores of additional natural products.

Certain chemical properties, such as molecular complexity and multiple stereogenic centres, are extremely difficult to include when large collections of compounds are produced for HTS. The systematic application of ring-distortion reactions on appropriate natural product starting materials offers a convenient approach to the rapid generation of large numbers of complex and diverse small molecules. These compounds possess a high degree of molecular complexity, as shown by examination of the Fsp3 and number of stereogenic centres, and are diverse structurally, as indicated by a Tanimoto similarity analysis. Depending on the exact application, specific structural features (for example, molecular mass, ClogP, hydrogen-bond donors/acceptors and so on) can be programmed in by careful selection of diversification reactions and building blocks. This method to construct complex and diverse small molecules can rapidly provide compounds with properties suitable for a wide variety of biological and medicinal applications.

Methods

Full experimental details and characterization data for all new compounds are included in the Supplementary Information.

Received 12 October 2012; accepted 6 December 2012; published online 20 January 2013

References

1. Swinney, D. C. & Anthony, J. How were new medicines discovered? *Nature Rev. Drug Discov.* **10,** 507–519 (2011).
2. Flaherty, K. T., Yasothan, U. & Kirkpatrick, P. Vemurafenib. *Nature Rev. Drug Discov.* **10,** 811–812 (2011).
3. Tsai, J. *et al.* Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Natl Acad. Sci. USA* **105,** 3041–3046 (2008).
4. Domling, A. Small molecular weight protein–protein interaction antagonists: an insurmountable challenge? *Curr. Opin. Chem. Biol.* **12,** 281–291 (2008).
5. Grivas, P. D., Kiaris, H. & Papavassiliou, A. G. Tackling transcription factors: challenges in antitumor therapy. *Trends Mol. Med.* **17,** 537–538 (2011).
6. Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature Rev. Drug Discov.* **6,** 29–40 (2007).
7. Kodadek, T. The rise, fall and reinvention of combinatorial chemistry. *Chem. Commun.* **47,** 9757–9763 (2011).
8. Schreiber, S. L. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* **287,** 1964–1969 (2000).
9. Galloway, W. R., Isidro-Llobet, A. & Spring, D. R. Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. *Nature Commun.* **1,** 80 (2010).
10. Spaller, M. R., Burger, M. T., Fardis, M. & Bartlett, P. A. Synthetic strategies in combinatorial chemistry. *Curr. Opin. Chem. Biol.* **1,** 47–53 (1997).
11. Clemons, P. A. *et al.* Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *Proc. Natl Acad. Sci. USA* **107,** 18787–18792 (2010).
12. Cui, J. *et al.* Creation and manipulation of common functional groups *en route* to a skeletally diverse chemical library. *Proc. Natl Acad. Sci. USA* **108,** 6763–6768 (2011).
13. Pelish, H. E., Westwood, N. J., Feng, Y., Kirchhausen, T. & Shair, M. D. Use of biomimetic diversity-oriented synthesis to discover galanthamine-like molecules with biological properties beyond those of the natural product. *J. Am. Chem. Soc.* **123,** 6740–6741 (2001).
14. Goess, B. C., Hannoush, R. N., Chan, L. K., Kirchhausen, T. & Shair, M. D. Synthesis of a 10,000-membered library of molecules resembling carpanone and discovery of vesicular traffic inhibitors. *J. Am. Chem. Soc.* **128,** 5391–5403 (2006).
15. Kumar, N., Kiuchi, M., Tallarico, J. A. & Schreiber, S. L. Small-molecule diversity using a skeletal transformation strategy. *Org. Lett.* **7,** 2535–2538 (2005).
16. Balthaser, B. R., Maloney, M. C., Beeler, A. B., Porco, J. A. & Snyder, J. K. Remodelling of the natural product fumagillol employing a reaction discovery approach. *Nature Chem.* **3,** 969–973 (2011).
17. 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).
18. Niggemann, J., Michaelis, K., Frank, R., Zander, N. & Hofle, G. Natural product-derived building blocks for combinatorial synthesis. Part 1. Fragmentation of natural products from myxobacteria. *J. Chem. Soc. Perkin Trans.* **1,** 2490–2503 (2002).
19. 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).
20. Aquino, C. *et al.* A biomimetic polyketide-inspired approach to small-molecule ligand discovery. *Nature Chem.* **4,** 99–104 (2012).
21. O'Connor, S. E. & Maresh, J. J. Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat. Prod. Rep.* **23,** 532–547 (2006).
22. Curtis, P. J. & Cross, B. E. Gibberellic acid. A new metabolite from the culture of filtrates of *Gibberella fujikuroi*. *Chem. Ind.* 1066 (1954).
23. Rodrigues, C., Vandenberghe, L. P., de Oliveira, J. & Soccol, C. R. New perspectives of gibberellic acid production: a review. *Crit. Rev. Biotechnol.* **32,** 263–273 (2012).
24. Cross, B. E. Gibberellic acid. Part I. *J. Chem. Soc.* 4670–4676 (1954).
25. Mulholland, T. P. C. Gibberellic acid. Part 9. The structure of allogibberic acid. *J. Chem. Soc.* 2693–2701 (1958).
26. Henderson, J. H. & Graham, H. D. A possible mechanism for biological and chemical activity of gibberellic acid. *Nature* **193,** 1055–1056 (1962).
27. Grove, J. F. & Mulholland, T. P. C. Gibberellic acid. Part 12. The stereochemistry of allogibberic acid. *J. Chem. Soc.* 3007–3022 (1960).
28. Cross, B. E., Grove, J. F. & Morrison, A. Gibberellic acid. 18. Some rearrangements of ring A. *J. Chem. Soc.* 2498–2515 (1961).
29. Cross, B. E. & Markwell, R. E. Rearrangements of the gibbane skeleton during reactions with 2,3-dichloro-5,6-dicyanobenzoquinone. *J. Chem. Soc. Perkin Trans. I* 1476–1487 (1973).
30. Idler, D. R., Schmidt, P. J. & Bitners, I. Isolation and identification of adrenosterone in salmon (*Oncorhynchus nerka*) plasma. *Can. J. Biochem. Physiol.* **39,** 1653–1654 (1961).
31. Borthakur, M. & Boruah, R. C. A microwave promoted and Lewis acid catalysed solventless approach to 4-azasteroids. *Steroids* **73,** 637–641 (2008).
32. Bernstein, S., Littell, R. & Williams, J. H. Steroidal cyclic ketals. IV. The conversion of 11-keto- to 11α-hydroxysteroids. The preparation of 11-epi-hydrocortisone, and Δ4-androstene-11α-ol-3,17-dione. *J. Am. Chem. Soc.* **75,** 1481–1482 (1953).
33. Lecomte, V., Stephan, E., LeBideau, F. & Jaouen, G. Improved addition of organolithium reagents to hindered and/or enolisable ketones. *Tetrahedron* **59,** 2169–2176 (2003).
34. Stephan, E., Brossat, M., Lecomte, V. & Bouit, P. A. Synthesis of the 11 beta-hydroxymethyl-androst-4-en-3,17-dione. *Tetrahedron* **62,** 3052–3055 (2006).
35. Song, C. E. *Cinchona Alkaloids in Synthesis and Catalysis* 1–10 (Wiley, 2009).
36. Smith, A. C. & Williams, R. M. Rabe rest in peace: confirmation of the rabe-kindler conversion of D-quinotoxine into quinine: experimental affirmation of the Woodward–Doering formal total synthesis of quinine. *Angew. Chem. Int. Ed.* **47,** 1736–1740 (2008).
37. Hintermann, L., Schmitz, M. & Englert, U. Nucleophilic addition of organometallic reagents to cinchona alkaloids: simple access to diverse architectures. *Angew. Chem. Int. Ed.* **46,** 5164–5167 (2007).
38. Baell, J. B. & Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **53,** 2719–2740 (2010).
39. O'Shea, R. & Moser, H. E. Physicochemical properties of antibacterial compounds: implications for drug discovery. *J. Med. Chem.* **51,** 2871–2878 (2008).
40. Walters, W. P., Green, J., Weiss, J. R. & Murcko, M. A. What do medicinal chemists actually make? A 50-year retrospective. *J. Med. Chem.* **54,** 6405–6416 (2011).
41. Feher, M. & Schmidt, J. M. Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* **43,** 218–227 (2003).
42. Yan, A. & Gasteiger, J. Prediction of aqueous solubility of organic compounds by topological descriptors. *QSAR Comb. Sci.* **22,** 821–829 (2003).
43. 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).
44. Leeson, P. D. & St-Gallay, S. A. The influence of the 'organizational factor' on compound quality in drug discovery. *Nature Rev. Drug Discov.* **10,** 749–765 (2011).
45. Leeson, P. D. & Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Rev. Drug Discov.* **6,** 881–890 (2007).
46. Rogers, D. J. & Tanimoto, T. T. A computer program for classifying plants. *Science* **132,** 1115–1118 (1960).

47. Rogers, D. & Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **50,** 742–754 (2010).
48. Huggins, D. J., Venkitaraman, A. R. & Spring, D. R. Rational methods for the selection of diverse screening compounds. *ACS Chem. Biol.* **6,** 208–217 (2011).
49. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Natural Prod.* **75,** 311–335 (2012).

Acknowledgements

We thank the Office of Naval Research (N00014-09-1-0249) and the University of Illinois for support of this work. R.W.H. III is an American Cancer Society postdoctoral fellow. K.C.M. is a National Science Foundation predoctoral fellow and a Robert C. and Carolyn J. Springborn graduate fellow. M.F.R. is a member of the National Institutes of Health Chemistry–Biology Interface (Training Grant NRSA 1-T-32-GM070421). We thank D. Gray for the X-ray analysis of Q1, M. Brichacek for technical assistance with selected two-dimensional NMR experiments in addition to many helpful discussions and R. J. Rafferty for technical assistance. We also thank M. Ahmed and A. Al-khouja for the synthesis of various intermediates during these investigations.

Author contributions

P.J.H. and R.W.H. III conceived the study. R.W.H. III designed and executed the synthesis of the **A** compound set and constructed libraries based on compounds A1, A2, A3, A5, A9, A12 and A15. K.C.M. designed and executed the synthesis of the **G** compound set and constructed libraries based on G6, G10 and G19. R.W.H. designed and executed the synthesis of the **Q** compound set and constructed the library based on Q1. T.A.F. performed computational analyses of all the compound sets discussed in this Article. M.F.R. constructed libraries of compounds A3, A12, G10 and G16. P.J.H. supervised this research and wrote this manuscript with the assistance of R.W.H. III, K.C.M., R.W.H. and T.A.F.

Additional information

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

Competing financial interests

The authors declare no competing financial interests.