

Chemical probes and drug leads from advances in synthetic planning and methodology

Christopher J. Gerry^{1,2} and Stuart L. Schreiber^{1,2,3*}

Abstract | Screening of small-molecule libraries is a productive method for identifying both chemical probes of disease-related targets and potential starting points for drug discovery. In this article, we focus on strategies such as diversity-oriented synthesis that aim to explore novel areas of chemical space efficiently by populating small-molecule libraries with compounds containing structural features that are typically under-represented in commercially available screening collections. Drawing from more than a decade's worth of examples, we highlight how the design and synthesis of such libraries have been enabled by modern synthetic chemistry, and we illustrate the impact of the resultant chemical probes and drug leads in a wide range of diseases.

Organic small molecules with novel mechanisms of action (nMoAs) are needed to help address today's most challenging biomedical problems. In particular, potential disease-related molecular targets indicated by the study of human genetics are often poorly characterized and cannot be studied or modulated with existing chemical tools¹. Drugs with nMoAs are also urgently needed to address resistance to existing drugs for cancer and infectious diseases²⁻⁴.

The search for new chemical probes and drugs can begin in many places. Natural products and their derivatives, for example, have played important roles in the study and treatment of disease for thousands of years⁵⁻⁹. Alternatively, when the therapeutic target is known, advances in synthetic organic chemistry, structural biology, computational modelling and screening have enabled researchers to construct, identify and optimize bioactive compounds via structure-based design^{10,11}. While these techniques and others have long histories of success in drug discovery¹²⁻¹⁴, this article primarily discusses high-throughput screening of small-molecule libraries, in which researchers sift through collections of compounds (10^3 to $>10^6$ members) to find 'hits' that exhibit a desired biological activity.

The use of small-molecule screening in early-stage drug discovery expanded dramatically in the 1990s with the introduction of new screening technologies and more efficient methods to synthesize large numbers of compounds with combinatorial chemistry. Given the vastness of 'chemical space', all the molecules synthesized to date represent only a tiny fraction of all possible compounds that have properties similar to those of existing chemical

probes and drugs (BOX 1). This limitation has spurred chemists to want to populate the compound libraries being used in screening experiments with a more optimal set of molecules. Early and insightful thoughts into valuable features of compound libraries were offered by Paul Bartlett and colleagues, who suggested that libraries can be classified in one of two ways: focused libraries or prospecting libraries¹⁵. Focused libraries — also known as targeted or biased libraries — typically comprise analogues of a known bioactive compound to generate structure-activity relationships (SARs) that can inform optimization efforts. Conversely, prospecting libraries eschew a specific molecular architecture in favour of combinations of available starting materials that maximize the structural novelty and diversity of their products. This strategy is particularly appealing when no known bioactive compounds exist or an nMoA compound is desired.

Instead, initial library construction efforts were dominated by the acquisitions of increasingly large collections from vendors because larger collections are in general thought to increase the probability of finding good starting points for drug discovery¹⁶. Although this approach has been a productive source of drugs and probes, commercial vendors tend to increase library size by prioritizing ease of synthesis and availability of starting materials (for example, by using reactions that link compounds containing planar, aromatic rings). These reaction products typically have a higher content of sp^2 -hybridized carbons and a lower content of sp^3 -hybridized carbons, leading to an unequal distribution of structural features in small-molecule libraries. Therefore, most areas of synthetically accessible chemical space remain underexplored¹⁷.

¹Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA.

²The Broad Institute of Harvard & MIT, Cambridge, MA, USA.

³Howard Hughes Medical Institute, Cambridge, MA, USA.

*e-mail: schreiber@broadinstitute.org

doi:10.1038/nrd.2018.53
Published online 13 Apr 2018

Not all compounds containing under-represented structures, however, are equally worthy of attention in bioactive compound discovery. Taking a cheminformatics-based approach, Shelat and Guy¹⁸ have proposed that compounds that are structurally distinct from the contents of commercial screening collections are more likely to act via nMoAs. Natural products are one such class of compounds^{19,20}, and they have been a productive source of chemical tools and drugs, as noted above. However, billions of years of selection pressure commonly endow natural products with exquisite specificity for their targets, which are most often essential proteins (for example, tubulin and actin) whose inhibition will typically lead to cell death and in some cases overt toxicity²¹. In one study, the molecular targets of natural products were

enriched in proteins constituting the most connected nodes of protein–protein interaction networks, consistent with natural selection yielding toxins used by organisms to gain advantage over competitors²². Furthermore, given their often high level of structural complexity, it can be challenging to synthesize many analogues of natural products to explore SARs.

Intuitively, the high structural complexity of natural products supports the principle that the 3D topography of binding sites on macromolecular targets would benefit from similarly 3D compounds. Developments in synthetic methodology over the past three decades have facilitated the construction of molecules that exhibit the intricate structural architectures and other chemical features that are common in natural products but

Box 1 | Physicochemical properties of chemical libraries and ‘drug-likeness’

Increases in screening throughput have emphasized the need for compounds that can be translated easily into drug leads and clinical candidates. Even experienced medicinal chemists, however, have difficulty identifying such molecules by manual inspection of chemical structure alone²⁴⁴. In response, and pioneered originally by the late Corwin Hansch at Pomona College²⁴⁵, frameworks that connect ‘drug-like’ performance properties (for example, solubility and oral bioavailability) to small-molecule physicochemical properties (for example, molecular mass and lipophilicity) have emerged over the past decades. These metrics have been designed — often retrospectively — to guide chemists towards molecules with greater chances of success in both preclinical and clinical testing, but they should not be seen as strict limits or cut-offs²⁴⁶.

Properties correlated with success in drug discovery can also benefit the development of chemical probes²⁴⁷. Compounds containing structural motifs known to cause molecules to misbehave in high-throughput assays (for example, binding proteins non-selectively, generating spurious activity readings or disrupting cellular membranes) are called pan-assay interference compounds (PAINS)²⁴⁸. Identifying and de-prioritizing these compounds can reduce rates of false positives. Validating potency alone, however, does not guarantee the success of a probe discovery or drug discovery campaign. Compounds with intracellular targets must be sufficiently lipophilic to traverse at least one biological membrane en route to the appropriate cellular compartment. Simultaneously, they must be sufficiently hydrophilic to be soluble in aqueous assay buffers. Lipinski’s ‘rule of five’, which was designed to predict human oral bioavailability, recommends that chemists consider factors such as the calculated logarithm of the partition coefficient between *n*-octanol and water (cLogP) and molecular mass to achieve a suitable balance between solubility, membrane permeability and other drug-like properties¹⁷¹. Similarly, Veber and colleagues have proposed that a compound’s polar surface area and number of rotatable bonds constitute an orthogonal set of criteria for predicting oral bioavailability²⁴⁹. Groups have also applied computational tools such as machine learning algorithms to model the relationship between chemical structure and drug-likeness²⁵⁰.

Screening experiments with the 100,000-member diversity-oriented synthesis (DOS) collection at the Broad Institute provide an illustrative set of laboratory-based case studies. Because the library was not designed to address a specific disease or class of targets, it comprises hundreds of smaller sub-libraries with distinct scaffolds to maximize its biological performance diversity²⁵¹. Each sub-library was populated with compounds of intermediate complexity (as determined by factors such as *sp*³ content and quantity of stereogenic elements) containing rigidifying ring systems. While the library has afforded probes for targets throughout the proteome, important lessons can be gleaned from its shortcomings. In particular, some hits have exhibited poor aqueous solubility, although medicinal chemistry efforts have been able to improve the properties of many of these compounds^{79,93,117,181,205}. Next-generation libraries are currently being designed that de-emphasize the large, lipophilic structures that result from many of combinatorial chemistry’s most widely used transformations^{243,252}.

The compounds described in this article contain chemical features that are under-represented in commercially available drug-like libraries. To begin to assess the compatibility of these types of compounds with current standards of drug-likeness, we can compute cheminformatic descriptors (via the [Molinspiration molecular properties calculator](#)) for the 11 molecules highlighted in this article, as shown in the table. These values are broadly consistent with various traditional measures of drug-likeness^{171,249,253}. This simple analysis suggests that under-represented chemical features (for example, chirality and high *sp*³ content) are compatible with computed chemical properties that have been argued previously to be valuable in drug discovery.

Table | Values for cheminformatic descriptors for highlighted molecules

	Molecular mass (Da)	cLogP	TPSA (Å ²)	Rotatable bonds	Ro5 violations
Mean	476.71	4.04	84.3	4.9	0.8
Median	501.63	4.02	79.9	5	1

cLogP, calculated logarithm of the partition coefficient between *n*-octanol and water; Ro5, rule of five; TPSA, topological polar surface area.

under-represented in commercial screening collections. Some noteworthy examples include the establishment of transition-metal-mediated coupling and metathesis reactions, asymmetric catalysis and organocatalysis, all of which have been reviewed extensively^{23–25}. These transformations have enabled efficient syntheses of compounds that contain synthetically challenging structural features, such as medium-sized rings²⁶, non-peptidic macrocycles^{27,28} and spirocyclic, fused or bridged bicyclic and polycyclic ring systems²⁹, among many others^{30–32}.

The concept underlying diversity-oriented synthesis (DOS), which was introduced in 2000 (REF. 33), is to harness such advances to synthesize libraries of compounds that incorporate chemical features common to natural products, including *sp*³-hybridized basic nitrogen atoms, stereogenic elements and novel skeletons. The resulting compounds are not analogues of specific classes of natural products, yet they have the overall appearance of natural products. In order to overcome the tendency of natural selection to select for the limited number of essential targets, this process purposefully breaks the link between natural selection and the generation of natural-product-like compounds. (We use the term ‘natural-product-like’ to refer to compounds that have chemical features found in ensembles of natural products, not necessarily just in specific ones.)

After a decade of developing these synthetic pathways, executing on them to make libraries of compounds that have under-represented chemical features and testing those compounds in biological assays, in this article we focus on the remarkable activities of the resulting probes and drug leads. These studies begin to address the question of whether novel chemistry is a productive method to generate compounds that act via novel mechanisms. It is not experimentally practical, or arguably feasible, to determine whether this approach is more or less effective than other means to identify nMoA compounds, such as fragment-based drug discovery or other small-molecule synthesis and screening techniques^{7,12,34}. And given the vast numbers of accessible chemical structures, a diversity of approaches to find nMoA compounds could be valuable. Our aim here is simply to illustrate the value of compounds that have been identified using DOS strategies and related approaches to produce compounds that are structurally dissimilar from those in commercially available screening collections.

So, after briefly summarizing the strategic considerations behind the use of modern asymmetric synthesis to construct small-molecule screening libraries, we will focus on the performances of the resulting chemical probes and drug leads by discussing examples in a wide range of therapeutic areas.

Strategies for library synthesis

The strategy and implementation of DOS have been reviewed previously^{33,35,36}, but briefly, short reaction pathways are designed such that libraries of skeletally and stereochemically diverse compounds can be synthesized efficiently. Initial efforts relied upon sequences of complexity-generating reactions (for example, multi-component reactions, cycloadditions and ring-opening

or ring-closing metathesis reactions) to construct intricate scaffolds from readily available precursors rapidly³⁷. Pathways were designed such that diversity could stem from large arrays of building blocks (appendage diversity)³⁸, enumeration of stereoisomers (stereochemical diversity)³⁹ and late-stage branching points that alter scaffold topography (skeletal diversity)^{40,41}.

An early DOS pathway that furnished a selective histone deacetylase 6 (HDAC6) inhibitor provides an illustrative case study. Inspired by the metal-binding features of the natural products trichostatin A and trapoxin A, a library of 7,200 1,3-dioxanes was synthesized⁴². Metal-chelating functional groups were installed to bias the library towards HDAC inhibition, and the novel 1,3-dioxane core and its appendages were conceived to impart specificity. One library member, tubacin, was found to be a domain-specific inhibitor of HDAC6 (effective concentration for half-maximum response (EC_{50}) = 2.9 μ M)^{43,44}. Tubacin proved to be a useful tool compound in cancer and neurodegenerative disease research^{45,46}, but the impracticality of its synthesis hindered its optimization and clinical potential. A more efficient synthetic route would have aided the elaboration of SARs and the optimization of pharmacokinetic properties for in vivo studies. Recognizing this deficiency led to a rethinking of synthetic strategies, leading to novel compound collections.

Modular pathways that facilitate analogue synthesis can enable chemists to meet the demands required of probe and drug molecules. These syntheses can be realized by mimicking natural product biosynthesis, in which building blocks are coupled intermolecularly and then cyclized intramolecularly — from penicillins to terpenes⁴⁷. This process introduces both topographic complexity and rigidifying structural elements into the reaction products⁴⁸, the latter of which are important for decreasing the entropic cost of binding macromolecules. In the context of DOS, this concept is known as the ‘build–couple–pair’ strategy (discussed further in examples below)⁴⁹.

Target-oriented syntheses of complex molecules, especially natural products, have also benefited tremendously from the pairing of convergent, modular synthesis with asymmetric chemistry. For example, the development of the oncology drug eribulin was made possible by the Kishi group’s efficient synthesis of the marine natural product halichondrin B and subsequent analysis of structural analogues^{50,51}. This strategy has been extended to the synthesis of focused libraries of natural product derivatives, which can be screened to find a molecule with improved physicochemical properties or activity against drug-resistant cells or pathogens⁵². Along these lines, the Myers group has developed platforms for fully synthetic tetracycline and macrolide antibiotics^{53,54}. The construction of such libraries would have been impractical without advances in reaction methodology and synthetic planning.

As suggested above, however, the search for nMoA compounds is probably better served by prospecting libraries synthesized via DOS. But because nothing is known a priori about the activities of library members,

the corresponding pathways should permit facile alterations throughout the molecule; ideally, each position should be amenable to functionalization. Strategic implementation of the build–couple–pair concept can fulfil this criterion.

The remainder of this article discusses the synthesis, optimization and biomedical impact of compounds containing under-represented chemical features. The broad range of their biological activities shows that such compounds can act as useful probes and potential drugs throughout medicine.

Probes for heritable disease targets

Advances in genotyping, sequencing and analysis of genetic data have facilitated the large-scale study of disease-related genetic variations^{55,56}. Despite the substantial challenges involved in translating genomic information into therapeutic insights and new medicines⁵⁷, genetics-based ‘experiments of nature’ can reveal a type of dose–response of gene activity that suggests the physiological consequences of modulating a target with a therapeutic agent⁵⁸. Analyses of human genomic data have revealed variants associated with type 2 diabetes⁵⁹, inflammatory bowel disease⁶⁰ and psychiatric disorders^{61,62}, and these variants have pointed to disease-relevant pathways not previously recognized⁶³.

Alongside genetic approaches, chemical probes capable of potently and selectively modulating protein activity are valuable tools for exploring therapeutic hypotheses before they are tested in the clinic^{64–66}. If the protein target of interest is known, various *in vitro* binding and functional assays can enable target-based screening campaigns to identify a corresponding small-molecule modulator. Ensuing optimization efforts are then driven by secondary assays and cell-based disease models. Alternatively, phenotypic screening can identify compounds that elicit a notable effect on disease-relevant physiology, most often in human cells. This technique is particularly useful when the phenotype of interest is not strongly associated with a causal protein target or when searching for compounds with nMoAs. Although an in-depth discussion of the relative benefits and challenges of these two strategies is beyond the scope of this article, the strategies can be complementary in some circumstances^{67–70}.

The examples shown in FIG. 1 illustrate how the interrogation of compound libraries synthesized using DOS and related strategies has afforded small-molecule bioactive compounds across a wide range of heritable diseases (defined here as diseases that may be influenced by heritable genetic characteristics)^{71–82}; notably absent are compounds for the study of genetic targets in cancer, which will be discussed in a separate section. Here, we highlight three vignettes that provide greater insight into the underlying synthetic pathways, compound libraries and chemical biology.

Robotnikinin — developmental disorders. The Hedgehog signalling pathway plays a role in the organization of developing vertebrates and insects^{83,84}. While also studied in the context of oncogenesis and tumour

maintenance⁸⁵, disruptions in Hedgehog signalling in humans have been linked to developmental disorders such as holoprosencephaly⁸⁶. Complementing genetic studies in model organisms, small-molecule tool compounds have greatly improved our understanding of both individual ligands and the signalling pathway as a whole⁸⁷.

Before 2009, however, no direct modulators of sonic hedgehog (SHH), one of the three mammalian Hedgehog orthologues, had been discovered. To address this gap, a target-based screen was performed against ShhN (the active portion of SHH) using small-molecule microarrays⁷⁷. The microarrays consisted of ~10,000 natural products, known bioactive compounds and products of various diversity synthesis pathways attached to an isocyanate-functionalized glass microscope slide⁸⁸.

Efforts to optimize the most promising initial hits resulted in the discovery of robotnikinin (FIG. 1), a 12-membered macrolactone that binds to purified ShhN (dissociation constant (K_d) = 3.1 μ M) and exhibits concentration-dependent inhibition of ShhN-mediated signalling in multiple cell lines, such as Shh-LIGHT2 and C3H10T1/2 cells (1.9–62.5 μ M)⁷⁷. No inhibition was observed in cell lines that lacked its transmembrane receptor, Patched. These data suggest that robotnikinin disrupts the interaction between ShhN and Patched, although the exact mechanism by which this occurs is not well understood. With regard to selectivity, later studies that revealed novel macrocyclic peptide binders also showed that robotnikinin elicits a 3–4-fold greater response in an enzyme-linked immunosorbent assay (ELISA)-based competition assay relative to the other two Hedgehog orthologues⁸⁹.

Robotnikinin was derived from a library of 12-membered to 14-membered chiral macrocycles⁹⁰. After an Evans asymmetric alkylation was used to set the first stereocentre, additional stereogenic elements were introduced via intermolecular couplings with chiral 1,2-aminoalcohols (FIG. 2a). The final step, a Ru-catalysed ring-closing metathesis, effected cyclization of the linear precursor 11. The modular synthetic pathway, in addition to providing an efficient route to under-represented non-peptidic macrocycles, facilitated systematic optimization of ring size, substituents and stereochemistry.

As the first direct binder of SHH, robotnikinin has been used to study the Hedgehog signalling pathway⁹¹. On the molecular level, it is capable of distinguishing between closely related proteins and disrupting a critical protein–protein interaction required for proper signalling.

BRD0476 — diabetes. A therapeutic agent that reverses or prevents the autoimmune destruction of the insulin-producing β -cells of the pancreas in patients with type 1 diabetes would address a substantial unmet medical need. With this goal in mind, a cell-based phenotypic screen was performed to identify small molecules that inhibit cytokine-mediated β -cell apoptosis⁷³. The screening library comprised 6,488 chiral [6,8]-fused bicyclic lactams generated via DOS.

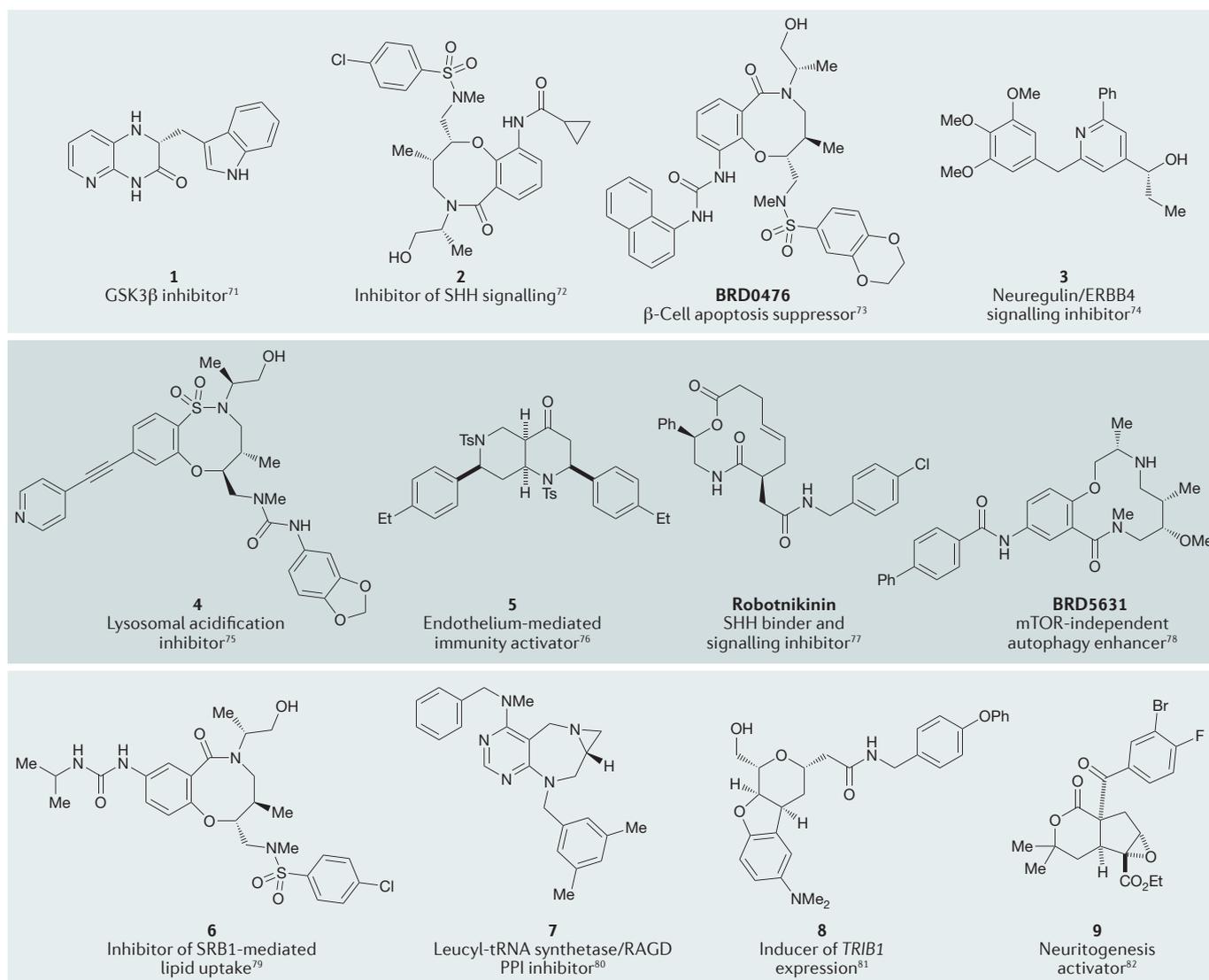


Figure 1 | A selection of compounds generated by diversity-oriented synthesis and related strategies as probes for a wide range of heritable diseases. Studies that use δ -lactone **9**, an activator of neurite growth in multiple classes of murine neurons⁸², may lead to insights into neurological disorders. The pyrimidodiazepine **7** disrupts the leucyl-tRNA synthetase–RAS-related GTP-binding protein D (RAGD) protein–protein interaction (PPI), which may improve our understanding of autophagy, ageing and immunosuppression via control of the mechanistic target of rapamycin complex 1 (mTORC1) signalling pathway^{80,260}. The benzannulated sultam **4** modulates lysosome acidification and thus

serves as a probe for V-ATPase function⁷⁵, which has been associated with loss of bone density and decreased kidney function, among other conditions^{261,262}. Other compounds interact with targets implicated in Alzheimer disease (**1**)⁷¹, holoprosencephaly (**2**)⁷², schizophrenia (**3**)⁷⁴, inflammatory diseases (**5**)⁷⁶, lipid metabolism (**6**)⁷⁹ and coronary artery disease (**8**)⁸¹. The remaining compounds shown are described in more detail in the vignettes in the main text. Properties relevant to their value as probes are discussed in BOX 2. GSK3 β , glycogen synthase kinase 3 β ; SHH, sonic hedgehog; SRB1, scavenger receptor class B member 1; *TRIB1*, tribbles homologue 1.

One of the most promising suppressors of apoptosis was optimized for potency via a medicinal chemistry campaign to afford BRD0476 (REF. ⁷³) (FIG. 1). In validation experiments, BRD0476 reduced caspase 3 activity and nitrite production in concentration-dependent manners, and it restored several aspects of normal β -cell function⁷³. Subsequent research involved the synthesis of analogues with improved aqueous solubility and the discovery that BRD0476 promotes β -cell survival by selectively inhibiting ubiquitin-specific peptidase 9X (USP9X)-dependent

Janus kinase–signal transducer and activator of transcription (JAK–STAT) signalling (significant activity at 2–5 μ M in independent primary samples)^{92,93}. Biochemical experiments revealed no significant inhibition of 96 human kinases, including JAK1–3 (<40% inhibition at 10 μ M) or 11 other members of the deubiquitinating enzyme (DUB) superfamily⁹². Not only do these results provide the first example of non-kinase-mediated inhibition of JAK–STAT signalling but they also suggest USP9X as a new potential therapeutic target for type 1 diabetes.

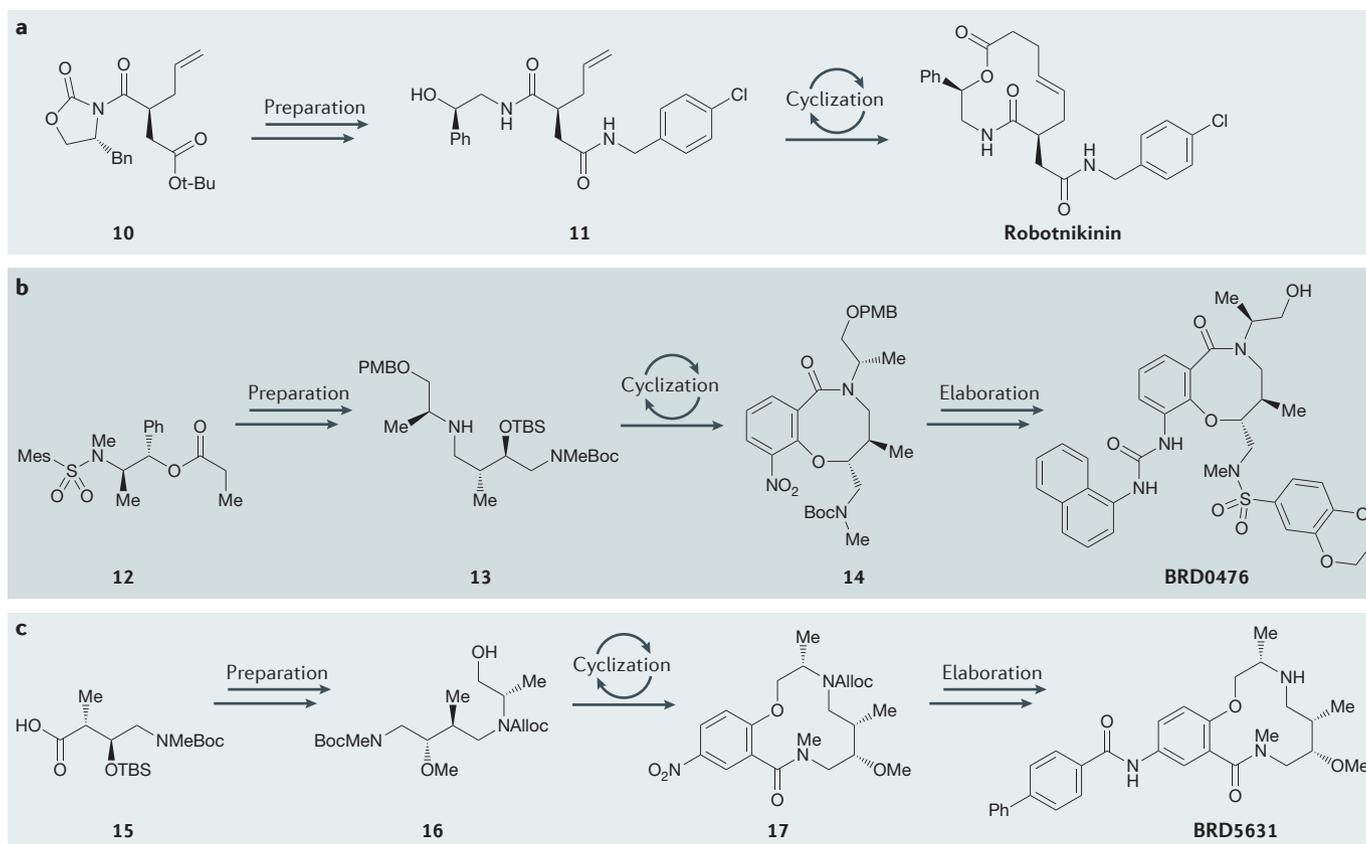


Figure 2 | Key transformations along the synthetic routes to selected compounds. The general strategy of building (or purchasing) chiral building blocks, coupling them together and cyclizing key linear intermediates via intramolecular transformations has generated a diverse array of chemical structures and biological activities in heritable disease and other areas of medicine. **a** | Chiral oxazolidinone **10** was functionalized strategically to afford intermediate **11**, which subsequently underwent esterification and ring-closing metathesis to afford the sonic hedgehog

(SHH) probe robotnikinin. **b** | Similarly, intermediate **13** was prepared from chiral ester **12**. S_NAr -based cyclization afforded the benzannulated 8-membered ring (**14**) at the core of BRD0476, a probe for type 1 diabetes and Janus kinase–signal transducer and activator of transcription (JAK–STAT) signalling. **c** | Intermediate **16**, which was prepared from chiral acid **15** and closely resembles **13**, was subjected to macrolactamization conditions to afford 12-membered macrocycle **17** en route to the autophagy probe BRD5631.

The synthesis of BRD0476 (FIG. 2b) begins with an auxiliary-mediated asymmetric aldol reaction and intermolecular coupling of a chiral 1,2-aminoalcohol⁹⁴. This key linear intermediate was then cyclized via S_NAr cycloetherification to construct the central 8-membered ring, a structural motif that is under-represented in current screening collections²⁶. With the core in place, appendage elaboration afforded BRD0476. The inclusion of all possible stereoisomers in the parent library aided the triage of screening data owing to the existence of built-in stereochemistry-based SARs⁷³.

As the development of BRD0476 has shown, the value of a probe is linked to the detailed molecular understanding of its MoA. These types of study can afford biological insights with implications for new therapeutics in the future.

BRD5631 — autophagy. Autophagy is the catabolic mechanism by which cells shuttle macromolecules, organelles and pathogens to the lysosome for recycling or degradation⁹⁵. Alterations in this critical maintenance process have been implicated in human genetic studies of

Crohn's disease, and additional studies have connected it to nonalcoholic fatty liver disease, Huntington's disease and other disorders^{96,97}. The development of both inhibitors and enhancers of autophagy would improve our understanding of multiple diseases and facilitate the discovery of new medicines.

One effort to search for small-molecule modulators of autophagy used a cell-based high-throughput screen of nearly 60,000 compounds from multiple DOS pathways prepared using asymmetric synthesis⁷⁸. Briefly, compound-treated HeLa cells were imaged via fluorescence microscopy to detect and quantify the presence of autophagosomes, the characteristic cellular structures of autophagy. BRD5631 (FIG. 1) was shown to increase autophagosome formation in a concentration-dependent manner ($EC_{50} = 3.1 \mu M$). Follow-up studies revealed that BRD5631 enhances autophagy via a mechanistic target of rapamycin (mTOR)-independent mechanism. While mTOR inhibitors are known to induce autophagy^{98,99}, their use in the clinic is limited by unwanted side effects¹⁰⁰. Therefore, BRD5631 represents an additional tool for studying autophagy.

The first few steps of the BRD5631 synthesis (FIG. 2c) closely mirror those of BRD0476 — an asymmetric aldol reaction followed by the intermolecular coupling of a chiral 1,2-aminoalcohol — but S_NAr chemistry was incapable of closing the 12-membered ring. Instead, a head-to-tail macrolactamization successfully generated all eight stereoisomers of the core scaffold, which were then functionalized to afford a 7,936-member library that included BRD5631 (REF. 101).

The structural similarity of intermediates **13** and **16** highlights the power of strategically designed compounds whose functional groups can be connected in different ways to generate different molecular skeletons. This example illustrates the build–couple–pair concept that mimics the way in which biosynthetic pathways construct natural products^{33,35,36,48,102}. As a key tenet of DOS, the ability to access multiple scaffolds from common precursors promotes skeletal diversity^{31,35,103}. Skeletally diverse libraries are becoming more valuable because they are thought to be more performance diverse¹⁰⁴, which would aid the search for nMoA compounds.

Probes for cancer targets

Historically, cancer drug discovery was based in part on phenotypic assays that identified compounds that killed cancer cells more effectively than they killed normal cells. Thus, many traditional (and still widely used) anticancer drugs affect processes important in all dividing cells, such as DNA synthesis, with their therapeutic window thought to be derived from the greater importance of such processes in rapidly dividing cancer cells^{105–107}. However, starting in the early 1980s, studies of the molecular basis of cancer have revealed that many cancers are driven by mutations in particular signalling proteins¹⁰⁸; targeting these mutated proteins — which are not present in normal cells — may enable selective cancer cell killing. Consequently, target-based screening has become an increasingly popular strategy in oncology over the past two decades¹⁰⁹. This era of molecularly targeted cancer therapy — enabled by technologies for investigating genetic differences between cancer cells and normal cells^{110,111} — has substantially improved the treatment of cancer. Many indications, however, remain without a targeted therapy, and resistance to new drugs continues to arise.

Cancer drug discovery has also been advanced by the development of potent and specific small-molecule probes that help characterize the protein targets that emerge from cancer genetics research. However, these proteins are often challenging to study with either small molecules or other modalities such as monoclonal antibodies. Therefore, if we are to exploit the findings of contemporary cancer research fully, the next generation of chemical probes should be capable of modulating the activities of historically challenging targets such as oncogenic transcription factors and GTPases^{112–114}. The screening of compounds that contain under-represented structural features, generated via DOS or other methods, appears to be a promising strategy, and it has already produced several probes for cancer

research^{115–127} (FIG. 3). The following vignettes discuss the discovery, optimization and use in cancer research of four notable tool compounds.

BRD7880 — aurora kinases B/C. Our first example does not discuss the targeting of a novel protein or protein class, but it does illustrate an unusual degree of selectivity in targeting members of a well-known class — protein kinases — with compounds resulting from DOS. Misregulation of the cell cycle, especially mitosis, is one of many hallmarks of cancer. Some of the key regulators of this process are the aurora kinases, a family of serine/threonine protein kinases that coordinate the complex cytoskeletal manoeuvres that occur during cell division¹²⁸. Alterations in aurora kinase activity have been linked to oncogenesis, and aurora kinases are often overexpressed in colorectal, breast and ovarian tumours, among others^{129,130}. Several aurora kinase inhibitors have since entered clinical trials^{131–133}, but none of them has won US Food and Drug Administration (FDA) approval. Achieving selective kinase inhibition can be challenging¹³⁴, so the lack of clinical success may be a result of insufficient specificity.

An unbiased phenotypic screen of 8,000 compounds from multiple DOS pathways identified BRD7880 (FIG. 3) as a highly selective inhibitor of aurora kinases B and C¹¹⁹. In this study, pools of DNA-barcoded cancer cell lines were used in a series of small-molecule screens. Cell lines were barcoded via viral transduction of DNA identifiers, pooled, treated with compounds and assessed for viability via amplification and sequencing of the DNA barcodes¹¹⁹. One of the hits from the screen that exhibited selective and concentration-dependent activity was BRD7880. After comparing its sensitivity profile — which cell lines were killed and which were not — to those of over 400 compounds with known MoAs in the Cancer Therapeutics Response Portal (CTRP), its close similarity to the sensitivity profile of the pan-aurora kinase inhibitor tozasertib suggested that BRD7880 is an aurora kinase inhibitor. Follow-up biochemical assays revealed not only that this hypothesis was correct but also that BRD7880 — obtained directly from the screening collection with no further optimization — was extremely potent and considerably more selective than tozasertib (FIG. 4), which had previously been tested in clinical trials¹³³. The only 2 kinases (out of 308) with >50% inhibition when treated with 30 nM BRD7880 were the aurora kinases B (half-maximal inhibitory concentration (IC_{50}) = 7 nM) and C (IC_{50} = 12 nM)¹¹⁹.

The synthesis of BRD7880 (FIG. 5a) closely resembles that of BRD0476, a structurally similar benzannulated lactam (FIG. 2b). However, an important difference is the relative stereochemistry of the methyl group and the silyl ether in linear intermediate **28**. An Evans asymmetric aldol reaction was performed to install the desired 1,2-*syn* stereodiad⁹⁴. Cyclization via S_NAr followed by functionalization of orthogonally protected lactam **29** furnished BRD7880.

BRD7880 is a particularly striking example of the impact of screening compounds that are accessible only via modern stereoselective synthesis. Because the aurora

kinases had already been the subject of multiple drug-discovery campaigns, it is remarkable that the most specific inhibitor to date was identified directly from a screen of 8,000 compounds. This result highlights the potential efficiency of exploring new regions of chemical

space made accessible by DOS strategies and advances in synthetic methodology. The rapid determination of the MoA of BRD7880 via high-dimensional assay data also provides a valuable precedent for future phenotypic screening campaigns. But the most direct impact of

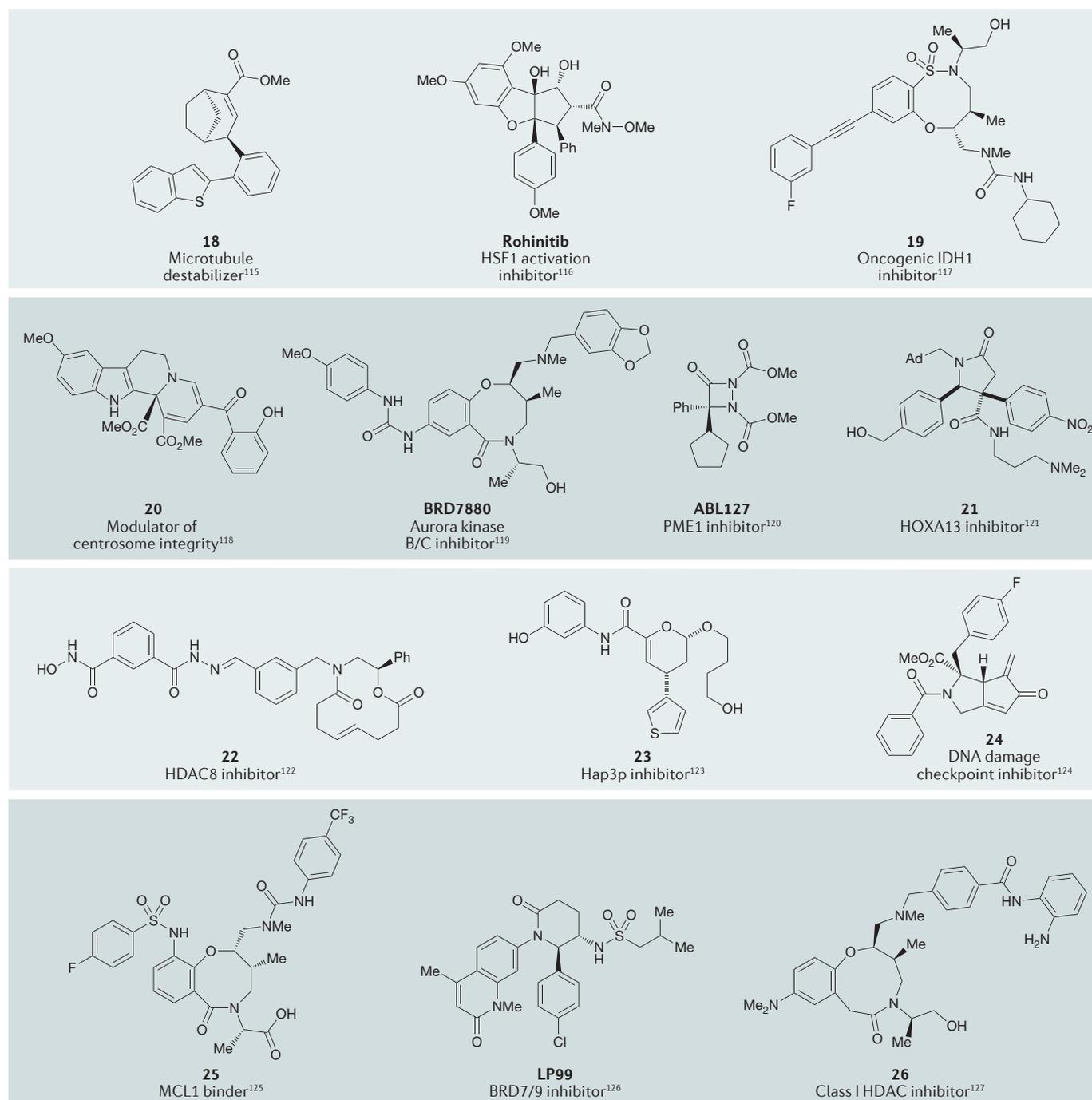


Figure 3 | Chemical probes for cancer targets. These probes include benzothiophene **18** and indoloquinolizine **20** as tools to study the cytoskeleton, a common cancer target^{115,118}; compounds such as hydrazide **22** (REF. ¹²²) and benzannulated lactam **26** (REF. ¹²⁷) that are highly specific and can discriminate between proteins within the same sub-family; probes discovered via target-based screening that exhibit significant stereochemistry-based structure–activity relationships, such as **19** (REF. ¹¹⁷) and **25** (REF. ¹²⁵); dienone **24**, which sensitizes p53-deficient cells to

DNA-damaging agents¹²⁴; and probes of ostensibly ‘undruggable’ transcription factors and protein–protein interactions, such as lactam **21** (REF. ¹²¹) and dihydropyran **23** (REF. ¹²³). Other compounds shown are described in more detail in the vignettes in the main text. Properties relevant to their value as probes are discussed in BOX 2. BRD7/9, bromodomain-containing protein 7/9; HDAC8, histone deacetylase 8; HOXA13, homeobox A13; HSF1, heat shock factor 1; IDH1, isocitrate dehydrogenase 1; MCL1, myeloid cell leukaemia 1; PME1, phosphatase methylesterase 1.

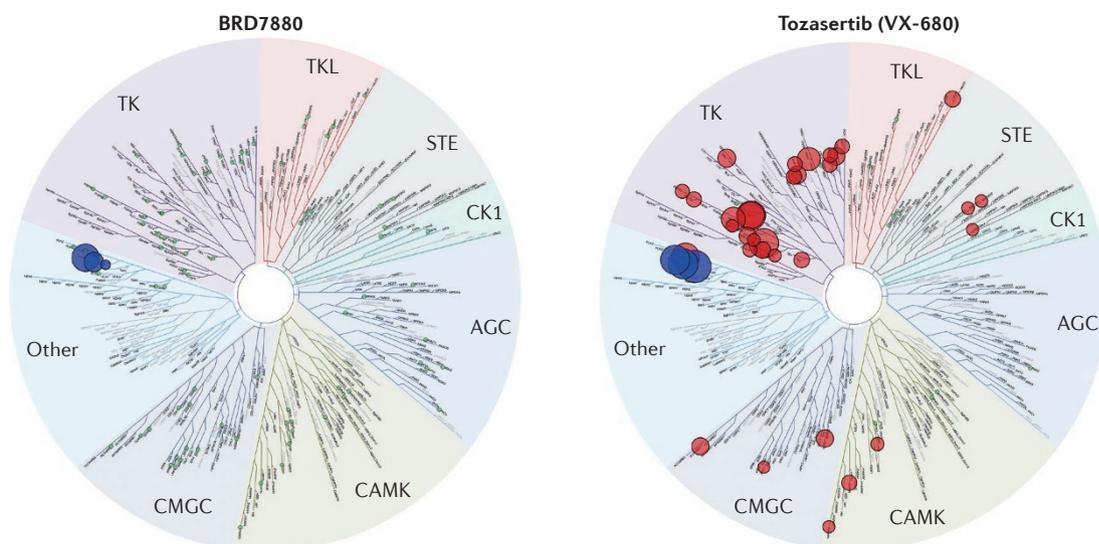


Figure 4 | Kinase affinity profiling reveals that BRD7880 is significantly more selective than tozasertib. The KinomeScan assay reveals kinases for which a given small molecule shows significant affinity (compound decreases binding of control by >75%). Aurora kinases A, B and C are shown in blue; other kinase binding partners are shown in red. CAMK, Ca²⁺/calmodulin-dependent protein kinase; CK1, casein kinase 1; STE, serine/threonine kinase; TK, tyrosine kinase; TKL, tyrosine kinase-like. Reproduced from REF. ¹¹⁹, Macmillan Publishers Limited.

BRD7880 may stem from its use in studying the therapeutic hypothesis of aurora kinase inhibition more rigorously than had been possible before with chemical tools.

Rohinitib — HSF1. Heat shock factor 1 (HSF1), a transcription factor that coordinates the heat shock response in human cells¹³⁵, is a regulator of ribosome biosynthesis that has also been associated with oncogenesis and poor outcomes in cancer patients^{136–138}. Exploring the connections between HSF1 activity, translational flux and maintenance of the tumorigenic cell state could reveal new therapeutic targets and strategies for cancer.

While studying the effects of varying ribosomal activity on transcription, Santagata and colleagues discovered a potent and selective inhibitor of HSF1 (REF. ¹¹⁶). They first observed that inhibition of translation elongation significantly decreased HSF1 activity, suggesting that HSF1 is involved in ribosome-mediated regulation of transcription. A screen of more than 300,000 compounds from the National Institutes of Health Molecular Libraries Probe Production Centers Network (NIH MLPCN) collection was then performed to identify small-molecule inhibitors of HSF1 activity. The resulting ~2,500 hits were then subjected to a dual reporter-based secondary screen to eliminate nonspecific compounds; remarkably, several ostensibly selective HSF1 inhibitors suppressed both reporters^{116,139}, which underscores the challenges associated with modulating transcription factor activity¹⁴⁰. The most potent and selective (>20-fold) hit, a natural product named rocaglamide A, was optimized to afford rohinitib (IC₅₀ = ~20 nM; FIG. 3). This unique chemical probe was then used to characterize extensively the link between protein translation and HSF1 activity in cancer cells. In addition to illuminating a number of fundamental insights into cancer biology, rohinitib suppressed tumour growth both in vitro and in vivo¹¹⁶.

The optimization of rocaglamide A involved both fully synthetic analogues and additional rocaglate natural products. All synthetic analogues, including rohinitib, were derived from the same 3-hydroxyflavone derivative, **30**. The fused tricyclic rocaglate core was constructed via a biomimetic photocycloaddition- α -ketol rearrangement-reduction sequence^{141,142} (FIG. 5b). The use of a chiral Brønsted acid can render the [3 + 2] cycloaddition enantioselective¹⁴³; alternatively, the desired enantiomer can be accessed via chiral resolution¹⁴⁴. This synthetic pathway demonstrates the utility of oft-underutilized photochemical transformations in generating natural-product-like scaffolds from flat, *sp*²-rich precursors¹⁴⁵.

LP99 — BRD7/9. Switch/sucrose non-fermentable (SWI/SNF) chromatin remodelling complexes are multi-protein complexes that regulate transcriptional processes involved in proliferation, DNA repair and other critical cellular functions. Loss of overall SWI/SNF function via inactivation of specific subunits has been linked to tumorigenesis, and SWI/SNF subunit mutations are found in roughly 20% of human cancers^{146,147}. One such subunit is bromodomain-containing protein 7 (BRD7), which is an essential cofactor for p53-mediated tumour suppression¹⁴⁸. Conversely, a subunit known as BRD9 is upregulated in some cancers and promotes the growth of acute myeloid leukaemia cell lines¹⁴⁹. Because the human proteome contains dozens of bromodomain-containing proteins, selective inhibitors are needed to assess the viability of BRD7 and BRD9 as therapeutic targets.

Looking to develop the first selective BRD7/9 inhibitor, Clark and colleagues¹²⁶ found that 1-methylquinolone could be elaborated into a potent and selective probe via asymmetric synthesis. 1-Methylquinolone had previously been identified via a crystallographic fragment

screen against a BRD9-related protein called ATPase family AAA domain-containing protein 2 (ATAD2) (REF. 150). From this promising starting point, additional structure-based and biophysical data informed the design of analogues that bind BRD9 with sub-micromolar affinity. The most potent compound, LP99 ($K_D = 99$ nM; FIG. 3), stabilized only 2 of the 48 expressible human bromodomain-containing proteins at $10 \mu\text{M}$ as measured via differential scanning fluorimetry (BRD7/9, change in transition midpoint for thermal unfolding

(ΔT_m) $> 4^\circ\text{C}$; all others, $\Delta T_m < 1^\circ\text{C}$)¹²⁶. LP99 was shown to disrupt the BRD7/9–chromatin interaction in cells without causing indiscriminate cell death. Preliminary experiments using LP99 as a tool compound found that BRD7 and BRD9 regulate inflammatory cytokine production, suggesting a novel anti-inflammatory therapeutic strategy.

While early analogues were synthesized as racemates and resolved via preparative-scale chiral chromatography, the active enantiomers of the most potent

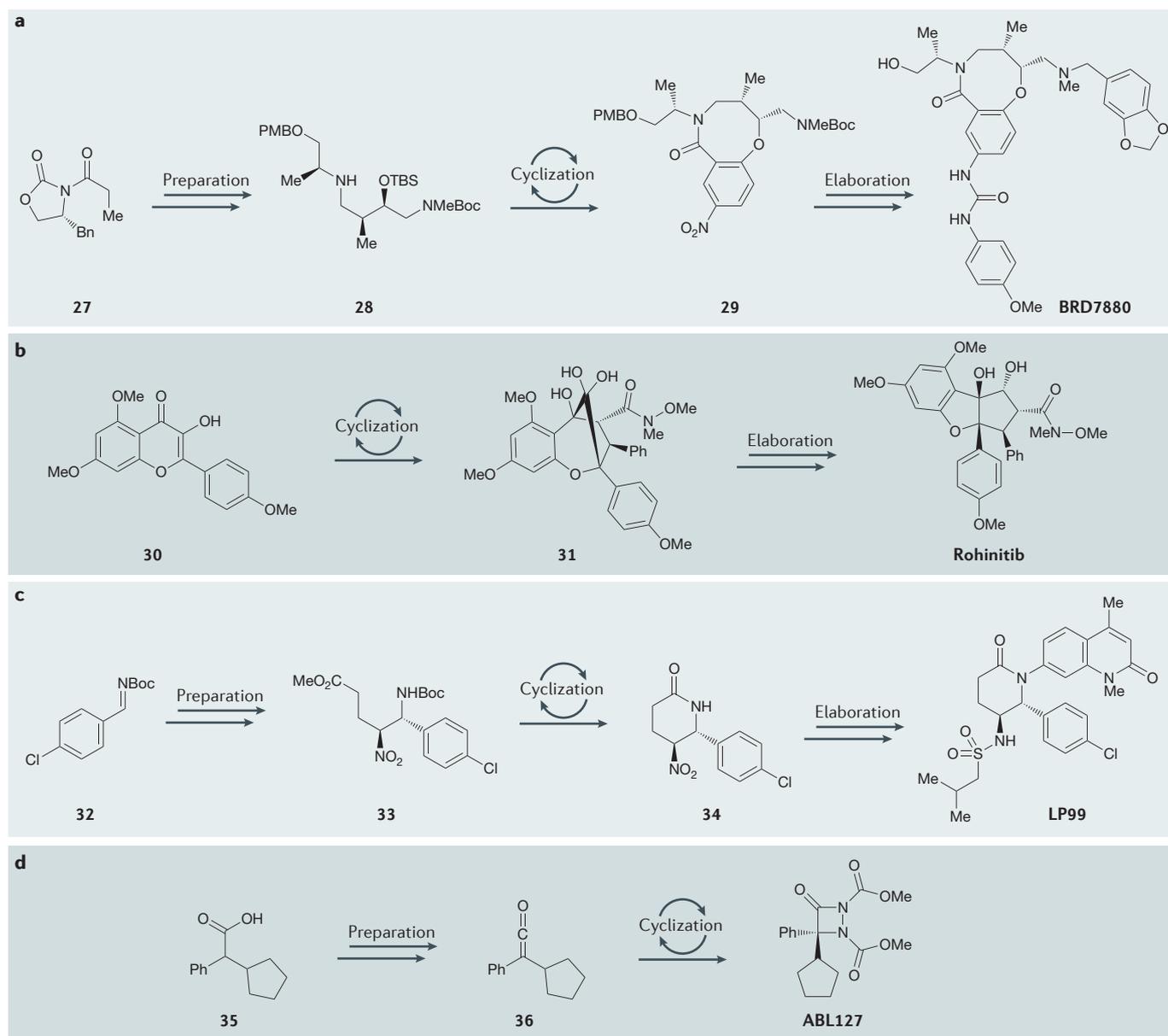


Figure 5 | Key transformations along the synthetic routes to selected compounds. **a** | In a synthesis that resembles that of BRD0476, chiral oxazolidinone **27** was functionalized strategically to afford intermediate **28**, which underwent $S_N\text{Ar}$ -based cyclization and further elaboration to provide BRD7880, a remarkably selective aurora kinase inhibitor. **b** | Fully synthetic analogues of rocaglate natural products originated from hydroxyflavone **30**. Rohinitib, a probe for heat shock factor 1 (HSF1) activity, and other rocaglate derivatives were synthesized via a biomimetic approach, proceeding through key intermediate **31**.

c | Starting from imine **32**, linear intermediate **33** was generated via a diastereoselective and enantioselective nitro-Mannich reaction. After cyclization to afford lactam **34**, appendage decoration furnished the bromodomain-containing protein 7/9 (BRD7/9) inhibitor LP99 as a single stereoisomer. **d** | The asymmetric synthesis of ABL127, a potent and selective probe of phosphatase methyltransferase 1 (PME1), was achieved via generation of ketene **36** from carboxylic acid **35**, followed by a [2 + 2] cycloaddition that was rendered enantioselective via a chiral nucleophilic catalyst.

compounds, including LP99, were prepared using asymmetric synthesis. The key transformation in the pathway (FIG. 5c) is achieved via an organocatalyzed nitro-Mannich reaction that furnishes the desired intermediate **33** with high diastereoselectivity and enantioselectivity (d.r. = 7:1, e.r. = 19:1)¹²⁶. Furthermore, this reaction can be performed on a gram scale, which facilitates the generation of additional analogues and the stockpiling of material for biological studies. Installation and decoration of the lactam core enabled the full synthesis of LP99 to be realized in just eight steps from commercially available reagents.

LP99 is a potent and selective chemical probe that has already uncovered previously unknown — and therapeutically promising — biology. Its development leveraged the interplay of many techniques, such as asymmetric catalysis and X-ray crystallography, and demonstrates that fragment-based drug design¹² is also capable of generating compounds with under-represented chemical features.

ABL127 — PME1. The widespread development of kinase inhibitors to study and treat cancer demonstrates the importance of the phosphoproteome in oncology¹³⁴. Various protein phosphatases regulate cellular dynamics by catalysing the dephosphorylation of serine, threonine and tyrosine residues. But despite the well-characterized link between aberrant phosphatase activity and cancer^{151,152}, many fewer chemical probes exist for protein phosphatases than for their kinase counterparts, contributing to the belief that phosphatases are largely ‘undruggable’ (REFS^{153,154}). Therefore, compounds that selectively modulate protein phosphatase activity, either directly or indirectly, would be useful small-molecule probes.

Protein phosphatase 2A (PP2A), a potent tumour suppressor that is responsible for the majority of serine/threonine dephosphorylation in eukaryotic cells, is inactivated in a variety of cancers^{155,156}. Its negative regulation is mediated in part by protein phosphatase methyltransferase 1 (PME1)¹⁵⁷, a serine hydrolase whose overexpression has been linked to increased cellular proliferation and other disease phenotypes¹⁵⁸. PME1 inhibition, therefore, is a potential anticancer therapeutic strategy.

A collaboration between the Cravatt and Fu laboratories afforded the first small-molecule inhibitor of PME1 (REF.¹²⁰). After a high-throughput fluorescence polarization-activity-based protein profiling assay of PME1 activity was developed, a screen of more than 315,000 members of the NIH MLPCN collection, a small fraction of which included compounds submitted by academic laboratories practising asymmetric synthesis (for example, BRD7880 above), was conducted. This experiment revealed a set of four aza- β -lactams, including ABL127 (IC₅₀ = 4.2 nM; FIG. 3), that inhibited PME1. Twenty-two other aza- β -lactams (including the enantiomers of three of the four hits) were significantly less active, which suggested that the observed PME1 inhibition resulted from a specific interaction. Follow-up studies showed that ABL127 inhibits PME1 in both cells (IC₅₀ = 6.4–11.1 nM) and mice (50 mg/kg) and that this inhibition decreases demethylated PP2A levels. Additional activity-based protein profiling and ‘click’ chemistry-based experiments

revealed that ABL127 exhibits selectivity across the entire proteome¹²⁰, further increasing its value as an nMoA chemical tool to study both PME1 and PP2A.

Enantioselective nucleophilic catalysis enables a three-step asymmetric synthesis of ABL127 (FIG. 5d). After generation of ketene **36**, a [2 + 2] cycloaddition with dimethyl azodicarboxylate affords ABL127. The cyclization is rendered enantioselective via the inclusion of a planar-chiral 4-pyrrolidinopyridine catalyst, and the transformation is compatible with a broad range of ketene and azo coupling partners¹⁵⁹. An extension of this methodology was used to synthesize oxo- β -lactams and develop a general platform for crafting specific serine hydrolase inhibitors¹⁶⁰.

Probes for infectious disease targets

Because agents of infectious disease are distinct from the host organism, the search for antimicrobials presents unique sets of opportunities and challenges. For the most common strategy of targeting the pathogen directly — rather than host-based mechanisms — on-target toxicity can be less of a concern because the drug target in the pathogen may not be found or may be significantly altered in the human patient. On the other hand, many pathogens can rapidly acquire drug resistance due to high rates of reproduction and DNA mutation^{161,162}. Accordingly, high-throughput phenotypic screening for compounds that kill pathogens but not human cells is commonly used to find potential drug leads for infectious diseases¹⁶³. Once a lead has emerged from screening, initial clues regarding the MoA of the compound can be gained via genetic methods. One approach involves incubating pathogens with sub-lethal concentrations of the compound to select for resistant strains that have altered sequences or copy numbers of the compound’s target (as revealed by genome sequencing)^{164,165}. Alternatively, systematically treating panels of pathogens in which a single gene is either deleted or overexpressed in each clone can point to relevant biology¹⁶⁶.

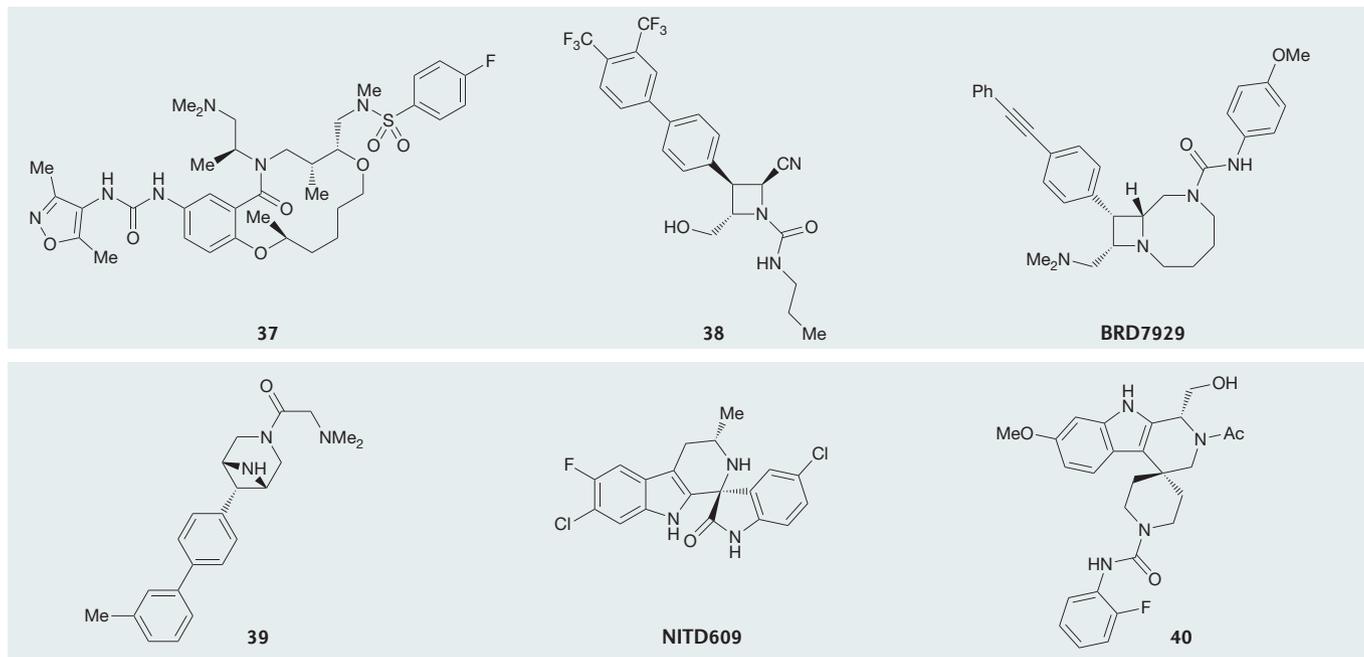
Reflecting their often evolutionary origin in competition between microorganisms, natural products have been a prolific source of leads and drugs against infectious diseases, but the emergence of resistance and the exhaustion of ‘low-hanging fruit’ have led this strategy to become less productive in recent decades^{7,8,167,168}. For some diseases, such as HIV/AIDS and hepatitis C, target-based drug discovery (often aided by knowledge of pathogen genetics) has made major contributions to the development of new medicines^{169,170}. But in others, particularly in the search for novel antibiotics, such strategies have been much less fruitful. In these project areas, the compounds typically available in both commercial libraries and industry collections could be a major limiting factor in anti-infective drug discovery. For example, ‘rule of five’ compliance continues to shape the contents of many screening collections even though it has little bearing on the success of antimicrobials^{168,171}.

One strategy to address these issues is to synthesize derivatives of known natural-product-based antimicrobials^{53,54,172,173}. Alternatively, advances in synthetic methodology may be leveraged to access scaffolds that

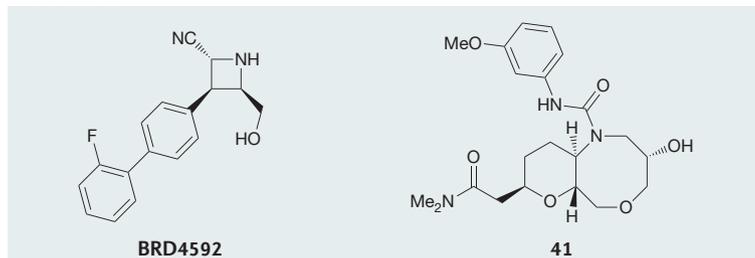
harbour chemical features frequently found in natural products without directly mimicking any specific class. Many promising next-generation antimicrobial agents contain such structural features^{174–185} (FIG. 6). The following four vignettes provide more detailed examples of the value of this approach in generating compounds that can be used to study and possibly treat infectious diseases.

NITD609 — malaria. Malaria is caused by several species of mosquito-borne *Plasmodium* parasites, with *Plasmodium falciparum* being responsible for the majority of fatalities¹⁸⁶. But most antimalarials treat only the asexual blood stage of *Plasmodium*¹⁸⁷, sparing the liver-stage and sexual blood-stage parasites that can sit dormant in the body and cause relapse at a later date¹⁸⁸.

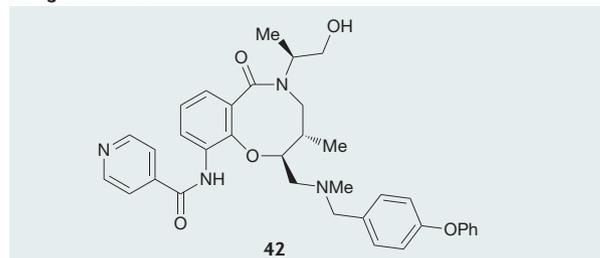
Malaria^{174–178}



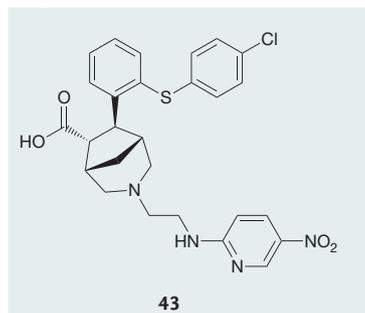
Tuberculosis^{179,180}



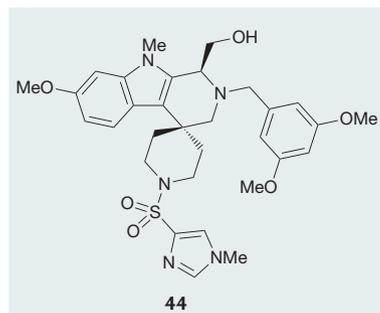
Chagas disease¹⁸¹



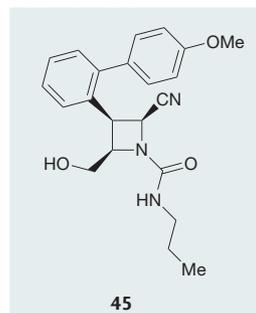
MRSA¹⁸²



RSV¹⁸³



Leishmaniasis¹⁸⁴



CDI¹⁸⁵

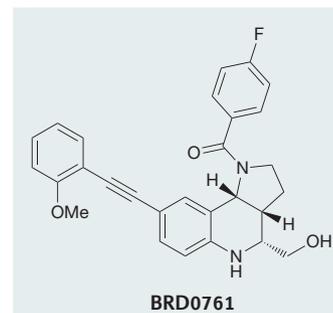


Figure 6 | Chemical probes and clinical candidates for infectious disease targets. A library of chiral azetidines originally designed to mimic central nervous system (CNS)-active compounds²⁰⁶ has afforded promising compounds against parasites such as *Leishmania donovani* (the causative agent of visceral leishmaniasis; **45**) and *Plasmodium falciparum* (the causative agent of malaria; **38,39** and BRD7929)^{175,176,178,184}. Macrocyclic **37** (malaria) and benzannulated lactam **42** (Chagas disease) also exhibit antiparasitic activity^{174,181}. Target-based and phenotypic

screening strategies were used to identify oxazocane **41** and bridged bicyclic **43**, which have bactericidal properties^{180,182}. The Pictet–Spengler-derived spirocycles **40** and **44** show antimalarial and antiviral activity, respectively^{178,183}. Other compounds shown are described in more detail in the vignettes in the main text. Properties relevant to their value as probes are discussed in BOX 2. CDI, *Clostridium difficile* infection, MRSA, methicillin-resistant *Staphylococcus aureus*; RSV, respiratory syncytial virus.

Furthermore, the risk of drug resistance^{189–191} necessitates easy-to-follow dosing regimens that maximize patient compliance — ideally a single-dose cure with prophylaxis and transmission-blocking activities¹⁸⁷. nMoA compounds that inhibit *Plasmodium* growth at multiple life stages are needed to address these challenges.

Prompted by the emergence of malaria-causing parasites resistant to artemisinin, an endoperoxide natural product^{192,193}, scientists at the Novartis Institute for Tropical Diseases and their collaborators performed phenotypic screens with *Plasmodium* to find nMoA antimalarial leads¹⁷⁷. After a screen of ~12,000 natural products and natural-product-like synthetic compounds, subsequent optimization yielded NITD609 (FIG. 6), a chiral spiroindolone that exhibits potent in vitro activity ($IC_{50} = \sim 0.5\text{--}10\text{ nM}$) against several strains of *P. falciparum* and *Plasmodium vivax* via an nMoA. Its molecular target, as first suggested by genomic sequencing of resistant parasites, was confirmed to be PfATP4, a cation-transporting P-type adenosine triphosphatase that maintains *Plasmodium* sodium homeostasis^{177,187,194,195}. Studies measuring toxicity (50% cytotoxicity concentration (CC_{50})/ $IC_{50} > 10^4$; human ether-a-go-go-related gene potassium channel (hERG) $IC_{50} > 30\text{ }\mu\text{M}$), pharmacokinetics (100% bioavailability) and in vivo efficacy (single-dose cure in a *Plasmodium berghei* mouse model at 100 mg/kg) suggested that NITD609 would be compatible with once-daily oral dosing¹⁷⁷, which was critical for its advancement into human trials as the first nMoA antimalarial clinical candidate in two decades¹⁸⁷. Since entering the clinic, not only has NITD609 exhibited impressive efficacy¹⁹⁶ but it has also been shown to prevent *P. falciparum* transmission from humans back to mosquitoes¹⁹⁷. Its success has prompted the development of additional antimalarials that target PfATP4 (REFS 198–200).

Advances in asymmetric catalysis have made the synthesis of NITD609 efficient. Starting from 6-chloro-5-fluoroindole, the original synthetic route afforded racemic NITD609 in eight steps; chiral chromatography was needed to isolate the active enantiomer¹⁷⁷. Due to the high cost of preparative-scale chiral chromatography, asymmetric routes would likely be more practical. The last step of the original synthesis, a Pictet–Spengler reaction, is highly diastereoselective but necessarily produces a racemate because the reaction lacks a chiral influence. Use of enantioenriched tryptamine derivatives or inclusion of a chiral phosphoric acid catalyst have been shown to furnish the appropriate enantiomer^{201,202}. Alternative routes that do not rely upon the Pictet–Spengler reaction have also been devised. One such pathway incorporates a Ni-catalysed enantioselective azadiels–Alder reaction to construct NITD609's key spirocyclic ring system (FIG. 7a), affording the final product in just three steps from indole **46** and ketimine **47** (REF. 203).

BRD7929 — malaria. Buoyed by the success of earlier pilot experiments^{174,204,205}, another effort to develop multistage antimalarial compounds began with a high-throughput screen of ~100,000 DOS-derived small molecules against a multidrug (MDR)-resistant strain of

blood-stage *P. falciparum*¹⁷⁸. These compounds represent hundreds of scaffolds — several of which were described above — not currently found outside of the Broad Institute. Hits from the primary screen were subjected to a panel of counter-screens to identify nMoA compounds that exhibited multistage activity. This analysis revealed clusters of structurally related bioactive compounds, most notably a series of bicyclic azetidines that showed nanomolar activity against all three *Plasmodium* life stages. For this series, genome sequencing of intentionally evolved resistant parasites showed mutations in the genetic locus predicted to encode phenylalanyl-tRNA synthetase (PfPheRS) — a novel antimalarial target. Follow-up experiments confirmed that an optimized analogue of the original hit, BRD7929 (FIG. 6), was not only a potent inhibitor of PfPheRS activity in vitro ($IC_{50} = 23\text{ nM}$) but also a single-dose cure, prophylactic and inhibitor of disease transmission in mouse models of malaria when dosed at 25 mg/kg (REF. 178).

The unusual bicyclic core of BRD7929 was constructed via two non-successive cyclizations (FIG. 7b). First, the azetidine was installed via treatment of chloride **50** with strong base to effect intramolecular ring closure, affording a 6:5 mixture of epimers that were separated by chromatography²⁰⁶. Then, after functionalization of azetidine **51**, ring-closing metathesis constructed the 8-membered diazocine. Subsequent Sonogashira coupling, isocyanate condensation and Mitsunobu-mediated amination afforded BRD7929 (REF. 178). While not currently applicable to BRD7929, recent advances in stereospecific C–H arylation methodology have improved the efficiency of synthetic pathways to bicyclic azetidines that lack functionalization at the 4-position²⁰⁷.

The development of NITD609 and BRD7929 highlights the power of combining asymmetric synthesis with phenotypic screening to discover new antimalarials. Multiple chemotypes showed potent activity against multidrug-resistant *Plasmodium* directly out of the screening deck; data surrounding many of these are freely available at the Malaria Therapeutics Response Portal (MTRP). Once optimized, both NITD609 and BRD7929 showed remarkable in vivo results that build upon an expanding pipeline of antimalarial compounds making their way through preclinical and clinical trials^{208,209}.

BRD0761 — Clostridium difficile infection. Narrow-spectrum nMoA antibiotics could have an important role in addressing the growing threat posed by the emergence of drug resistance in the Gram-positive bacterium *Clostridium difficile*, an opportunistic pathogen that can cause severe diarrhoea, sepsis and death^{210,211}. A high-throughput screen of the same ~100,000-member DOS library described above against eight bacterial organisms revealed potent and selective inhibitors of *C. difficile* growth¹⁸⁵. Two structurally distinct series of compounds, one of which is represented by BRD0761 (FIG. 6), were explored further. Compounds from both series were more selective than fidaxomicin, an FDA-approved antibiotic that was expressly approved for its selective activity against *C. difficile*²¹². Furthermore, BRD0761 exhibited greater potency against several *C. difficile* isolates (minimum

inhibitory concentration (MIC) = 0.06–1 µg/mL) than did vancomycin, which is used to treat severe cases of *C. difficile* infection^{210,213}. Genomic analysis of resistant mutants and in silico modelling suggested that the target of BRD0761 is glutamate racemase, an essential protein involved in cell wall biosynthesis. While this enzyme is a known target in *Helicobacter pylori*, *Mycobacterium tuberculosis* (Mtb) and other Gram-positive bacteria²¹⁴, it has not yet been validated in *C. difficile*¹⁸⁵.

The preparation of BRD0761 (FIG. 7c) demonstrates the value of combining complexity-generating transformations with asymmetric catalysis. An organocatalyzed Povarov reaction developed in the Jacobsen laboratory was identified as an efficient method to generate tetrahydroquinoline-containing compounds with three contiguous stereocentres. This reaction was used to synthesize tetrahydroquinoline **56** with high diastereoselectivity and enantioselectivity²¹⁵. Prepared in only

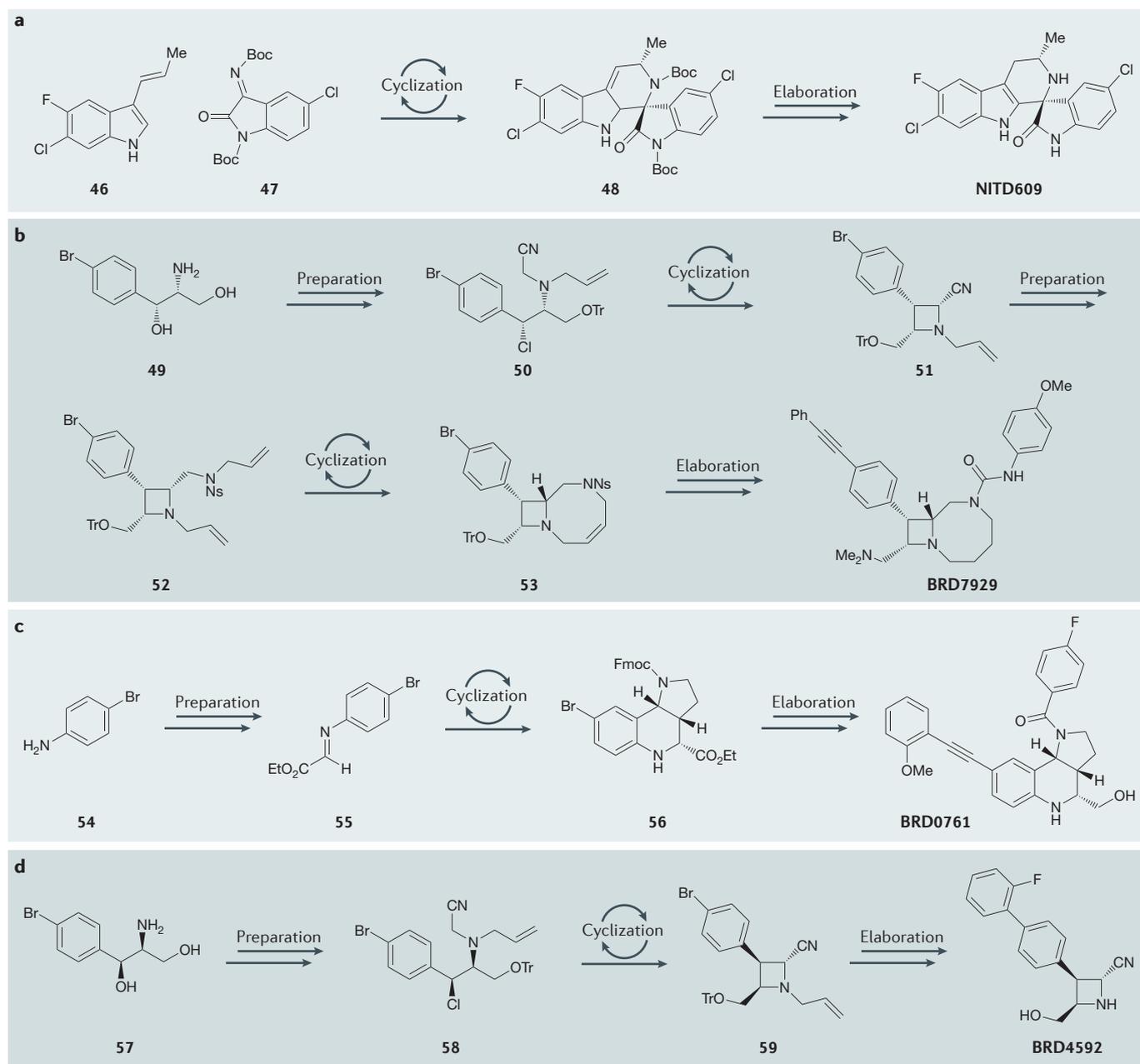


Figure 7 | Key transformations along the synthetic routes to selected compounds. a | One synthetic route to NITD609, a potent antimalarial with a novel mechanism of action (nMoA), involved an enantioselective aza-Diels-Alder reaction between indole **46** and ketimine **47** to afford spirocycle **48**, which is one step from NITD609. **b** | The synthesis of antimalarial BRD7929 contained two key cyclizations. First, treatment of chloride **50** with strong base supplied azetidine **51**. Subsequent functionalization furnished the

ring-closing metathesis substrate **52**, which was cyclized to diazocine **53** en route to BRD7929. **c** | From aniline **54**, imine **55** was an ideal substrate for a diastereoselective and enantioselective Povarov reaction. The resulting tetrahydroquinoline **56** was readily converted to BRD0761, which shows nMoA activity against *Clostridium difficile*. **d** | Similar to BRD7929, the tuberculosis probe BRD4592 was accessed via a chiral azetidine (**59**), which in turn was generated from chloride **58** and amino alcohol **57**.

two steps from commercially available starting materials, the resulting scaffold was readily elaborated into a 2,328-member library that contained BRD0761 (REF. 185).

Despite their structural and biological differences, as described above, BRD7929 and BRD0761 were derived from the same screening collection. These results highlight the ability of prospecting libraries to generate nMoA leads with diverse chemotypes across different diseases.

BRD4592 — tuberculosis. Recent years have seen the alarming rise of MDR, extensively drug-resistant (XDR) and totally drug-resistant (TDR) strains of *Mtb*^{216–218}, the causative agent of tuberculosis (TB). So, there is a substantial need for nMoA anti-TB compounds. Hoping to discover such compounds, a collaboration among researchers at the Broad Institute, the University of Chicago and Argonne National Laboratory screened roughly 83,000 compounds of the Broad DOS collection for activity against green fluorescent protein (GFP)-expressing *Mtb*¹⁷⁹. This screen identified BRD4592 (FIG. 6), a chiral azetidine with three contiguous stereocentres, as an inhibitor of *Mtb* growth ($MIC_{90} = 3 \mu M$). Notably, its other seven stereoisomers were not scored as active, which suggests that its bactericidal activity is due to specific target engagement. Genetic analysis of resistant mutants followed by extensive kinetic, thermodynamic and biochemical characterization revealed the molecular target to be both subunits of tryptophan synthase ($IC_{50}^{\alpha} = 71 \text{ nM}$; $IC_{50}^{\beta} = 23 \text{ nM}$), a previously untargeted essential metabolic enzyme^{179,219}. BRD4592 stabilizes multiple enzyme conformations via allosteric binding (FIG. 8), which probably contributes to its high specificity; by contrast, substrate mimetics that target enzyme active sites are likely to exhibit substantial off-target activity²²⁰. BRD4592 was then shown to inhibit *Mycobacterium marinum* growth at 15 μM in zebrafish embryos, a common in vivo model of TB²²¹.

The versatile azetidine-based synthetic pathway that afforded BRD7929 can be redirected to generate BRD4592 (FIG. 7d). As before, cyclization of chloride 58 produced a near-equal mixture of epimers at the 2-position that

were easily separated via chromatography²⁰⁶. Suzuki cross-coupling and functional group deprotections then furnished BRD4592. BRD7929 and BRD4592 highlight just two of the many distinct skeletons that can be synthesized from azetidines 51 and 59 and their stereoisomers; the full complement of bridged, fused and spirocyclic ring systems that can be accessed via this chemistry is discussed elsewhere²⁰⁶.

The discovery of BRD4592 offers lessons for antimicrobial drug discovery. The use of phenotypic screening both yielded membrane-permeable hits (an important issue in finding inhibitors of *Mtb*^{164,222,223}) and enabled the discovery of a new therapeutic target outside of previously drugged pathways¹⁷⁹. In addition, although the similarities between the syntheses and chemical structures of BRD4592 and BRD7929 are striking, we hesitate to call the azetidine a ‘privileged scaffold’; for example, the corresponding pyrrolidines were not included in the library and are therefore not available for comparison. Rather, we propose that these two vignettes provide evidence that compounds that are structurally dissimilar from those in screening collections are likely to lead to novel biological MoAs and therapeutic insights — different chemistry yields different outcomes.

Conclusions and outlook

The ability to modulate biological systems reversibly, selectively and with temporal control is one of the most powerful weapons in the chemical biologist’s arsenal. By advancing our understanding of disease biology, high-quality small-molecule probes have catalysed new research efforts to test therapeutic hypotheses and develop new medicines⁶⁶. However, we must be wary of compounds whose activities have not been sufficiently or accurately defined^{65,224} (BOX 2). Furthermore, much of our currently available chemical matter is both structure-redundant and performance-redundant²²⁵, which hampers our ability to develop nMoA compounds that can exploit insights from human genetics or combat drug-resistant pathogens.

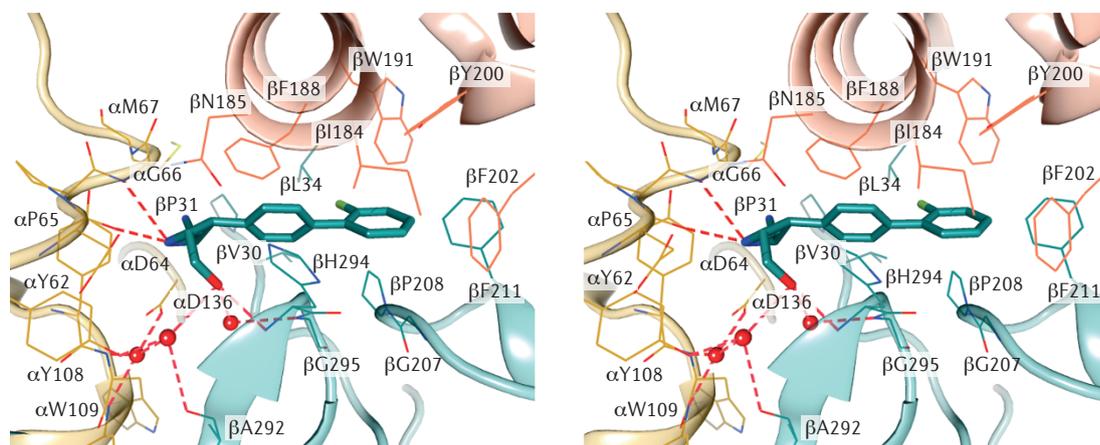


Figure 8 | Stereo view of BRD4592 (cyan) bound at the interface of the α and β subunits of tryptophan synthase. X-ray crystallography was used to solve the co-crystal structure of BRD4592 and tryptophan synthase. Hydrogen bonds and water molecules are shown as dashed lines and red spheres, respectively. Reproduced from REF. 179, Macmillan Publishers Limited.

Box 2 | Chemical probes: promises and pitfalls

Chemical probes can contribute to the development of drug leads by gauging the chemical tractability and physiological relevance of a potential therapeutic target²⁵⁴. But these two classes of molecule serve fundamentally different purposes. Unlike drugs, probes need not be optimized for safety and therapeutic efficacy in human patients — they are designed to answer a specific biological question in models of disease. Therefore, structural properties related to traditional measures of drug-likeness such as oral bioavailability and metabolic stability are not essential. For their corresponding biological results to be compelling, however, probes must be potent and selective and act via known mechanisms of action, among many other factors²²⁴. Failure to understand fully the biological consequences of probe treatment could result in the misassignment of molecular cause to phenotypic effect. Such mistakes have probably contributed to the reproducibility crisis in biomedical research²⁵⁵.

The growing popularity and importance of small-molecule probes has prompted systematic reviews of probe quality. Established mechanisms to correct the scientific literature are often not sufficiently nimble or comprehensive to curb errors that can propagate at an exponential rate. In response, groups such as the Structural Genomics Consortium (SGC) have established guidelines for describing valuable, high-quality tool compounds. Originally founded in 2008, the SGC is a public–private partnership that curates collections of probes related primarily to epigenetic signalling. To be included in the SGC database, probes must demonstrate potency both in vitro (half-maximal inhibitory concentration (IC_{50}) or dissociation constant (K_d) < 100 nM) and in cells (on-target activity at 1 μ M) and greater than 30-fold selectivity over other protein family members. As an alternative to objective but strict guidelines, Paul Workman and colleagues have proposed a ‘fit-for-purpose’ model that emphasizes a series of principles that are more practical to implement^{256,257}. These factors fall into a handful of broad categories related to chemistry (for example, solubility, stability and membrane permeability), potency (for example, in biochemical, cellular or in vivo assays) and other properties. Probes that do not fulfil all of these criteria can illuminate remarkable biology²⁵⁶, but the conclusions they suggest must always be placed in the appropriate context.

To complement the guidelines established by the SGC, Workman and others, various web-based resources have recently been established. These databases allow researchers to sift through experimental data and recommendations from the scientific community to identify the best compound for their particular research project. One such example is the [Chemical Probes Portal](#)⁶⁵, which features reviews of several hundred probes from members of its scientific advisory board. The [Probes & Drugs Portal](#), on the other hand, provides experimental data relating to ~30,000 bioactive compounds (updated on a monthly basis)²⁵⁸. Finally, [Probe Miner](#) is an ambitious effort to centralize data from the vast medicinal chemistry literature²⁵⁹. Analysis of >3.9 million experiments on >1.8 million compounds led to in-depth scoring of hundreds of thousands of probes against 2,220 human protein targets. These and similar resources should be used early and often in biomedical research projects that hope to draw biological conclusions from small-molecule-based experiments.

As this article has attempted to illustrate, advances in synthesis have enabled the construction of libraries of natural-product-like compounds. These novel chemistries have afforded nMoA compounds with potential therapeutic value across a broad range of diseases. While novel does not necessarily mean better, probe discovery and drug discovery campaigns are profoundly influenced by the quality of their starting points (‘front loading’)²²⁶. Earlier analyses, such as those conducted by Michael Hann and colleagues^{227,228}, have suggested that libraries enriched in these complex compounds are more likely to be inefficient in high-throughput screening experiments. However, the probes and drug leads discussed in this article were all derived from screens of 8,000 to ~315,000 compounds, suggesting that such libraries can generate useful hits and leads at least as efficiently as the large libraries used in the pharmaceutical industry²²⁹.

Biomedical research would also benefit from general, high-throughput methods for discovering small-molecule binders to the novel therapeutic targets uncovered by human and microbial genetics, molecular biology, cellular biology and chemical biology. Even if they do not alter protein activity, binders can be used to study how biological systems respond to various perturbations^{230,231}. For example, they can be applied to discover compounds that promote new protein–protein interactions or stabilize or destabilize

their cellular targets. In addition, the binders can be equipped with additional chemical motifs that promote the cellular mis-localization or degradation of targeted proteins^{232,233}. Unfortunately, methods for discovering binders are often time consuming and resource intensive^{234,235}. Affinity enrichment-based screening of libraries of DNA-barcoded compounds, however, is a promising technique that is gaining in popularity due to advances in DNA sequencing, among other factors²³⁶. Current DNA-encoded libraries are largely populated by peptidomimetics and *sp*²-rich compounds²³⁷, similar to commercial libraries of ‘conventional’ small molecules. Perhaps, then, DNA-encoded libraries could also benefit from the inclusion of compounds that contain under-represented structural features. This concept poses a future challenge for synthetic organic synthesis because modern chemistries are needed that are compatible with the conditions required for concomitant DNA barcoding.

Looking ahead, the building of performance-diverse small-molecule libraries would be more efficient if a rapid feedback mechanism were to exist between synthetic organic chemistry and biology. The current multi-year retrospective and time-consuming method for assessing small-molecule bioactivity precludes biological performance from influencing library design²²⁵, and structural diversity alone has been shown to be a poor predictor of biological performance diversity²³⁸.

On the other hand, high-dimensional (multiplexed) biological assays that track changes in gene expression²³⁹, cell morphology²⁴⁰ and other cellular features²⁴¹ can provide chemists with near-immediate feedback on their synthetic decisions. By associating each compound with thousands of biological changes, these assays constitute the backbone of a means for real-time biological annotation of reaction products^{242,243}. Such a feedback loop would indicate during library construction — rather than years later — the reactions, scaffolds and synthetic pathways worth pursuing for maximal performance diversity.

They also enable the hand-picking of small clusters of compounds, with each cluster having either known or novel MoAs, and thus collectively provide an empirical means for the construction of performance-diverse compound collections for probe and drug development. Chemists have neither the time nor the resources to synthesize every possible compound, so a prudent approach to bioactive compound discovery would incorporate the principles of the scientific method — hypotheses, experiments, observations and conclusions — whenever possible.

- Plenge, R. M. Disciplined approach to drug discovery and early development. *Sci. Transl. Med.* **8**, 349ps15 (2016).
- Holohan, C., Van Schaebroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer* **13**, 714–726 (2013).
- Spellberg, B., Bartlett, J. G. & Gilbert, D. N. The future of antibiotics and resistance. *N. Engl. J. Med.* **368**, 299–302 (2013).
- Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. V. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* **13**, 42–51 (2015).
- Koehn, F. E. & Carter, G. T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.* **4**, 206–220 (2005).
- Gerwick, W. H. & Moore, B. S. Lessons from the past and charting the future of marine natural products drug discovery and chemical biology. *Chem. Biol.* **19**, 85–98 (2012).
- Cragg, G. M. & Newman, D. J. Natural products: a continuing source of novel drug leads. *Biochim. Biophys. Acta* **1830**, 3670–3695 (2013).
- Ganesan, A. The impact of natural products upon modern drug discovery. *Curr. Opin. Chem. Biol.* **12**, 306–317 (2008).
- Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661 (2016).
- Ferreira, L. G., Dos Santos, R. N., Oliva, G. & Andricopulo, A. D. Molecular docking and structure-based drug design strategies. *Molecules* **20**, 13384–13421 (2015).
- Moitessier, N. et al. Medicinal chemistry projects requiring imaginative structure-based drug design methods. *Acc. Chem. Res.* **49**, 1646–1657 (2016).
- Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. & Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **15**, 605–619 (2016).
- Lionta, E., Spyrou, G., Vassiliadis, D. K. & Cournia, Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr. Top. Med. Chem.* **14**, 1923–1938 (2014).
- Sliwoski, G., Koithiwale, S., Meiler, J. & Lowe, E. W. Computational methods in drug discovery. *Pharmacol. Rev.* **66**, 334–395 (2013).
- Spaller, M. R., Burger, M. T., Fardis, M. & Bartlett, P. A. Synthetic strategies in combinatorial chemistry. *Curr. Opin. Chem. Biol.* **1**, 47–53 (1997).
- Dandapani, S. & Marcaurelle, L. A. Accessing new chemical space for 'undruggable' targets. *Nat. Chem. Biol.* **6**, 861–863 (2010).
- Reymond, J. L. The chemical space project. *Acc. Chem. Res.* **48**, 722–730 (2015).
- Shelat, A. A. & Guy, R. K. Scaffold composition and biological relevance of screening libraries. *Nat. Chem. Biol.* **3**, 442–446 (2007).
- 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).
- 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).
- Salvador-Reyes, L. A. & Luesch, H. Biological targets and mechanisms of action of natural products from marine cyanobacteria. *Nat. Prod. Rep.* **32**, 478–503 (2015).
- Dančik, V., Seiler, K. P., Young, D. W., Schreiber, S. L. & Clemons, P. A. Distinct biological network properties between the targets of natural products and disease genes. *J. Am. Chem. Soc.* **132**, 9259–9261 (2010).
- Dalko, P. I. & Moisan, L. Enantioselective organocatalysis. *Angew. Chem. Int. Ed. Engl.* **40**, 3726–3748 (2001).
- Taylor, M. S. & Jacobsen, E. N. Asymmetric catalysis by chiral hydrogen-bond donors. *Angew. Chem. Int. Ed. Engl.* **45**, 1520–1543 (2006).
- Fürstner, A. Olefin metathesis and beyond. *Angew. Chem. Int. Ed. Engl.* **39**, 3012–3043 (2000).
- Bauer, R. A., Wenderski, T. A. & Tan, D. S. Biomimetic diversity-oriented synthesis of benzannulated medium rings via ring expansion. *Nat. Chem. Biol.* **9**, 21–29 (2012).
- Kopp, F., Stratton, C. F., Akella, L. B. & Tan, D. S. A diversity-oriented synthesis approach to macrocycles via oxidative ring expansion. *Nat. Chem. Biol.* **8**, 358–365 (2012).
- Nie, F. et al. A multidimensional diversity-oriented synthesis strategy for structurally diverse and complex macrocycles. *Angew. Chem. Int. Ed. Engl.* **55**, 11139–11143 (2016).
- Antonchick, A. P. et al. Highly enantioselective synthesis and cellular evaluation of spirooxindoles inspired by natural products. *Nat. Chem.* **2**, 735–740 (2010).
- Moura-Letts, G., Diblasi, C. M., Bauer, R. A. & Tan, D. S. Solid-phase synthesis and chemical space analysis of a 190-membered alkaloid/terpenoid-like library. *Proc. Natl Acad. Sci. USA* **108**, 6745–6750 (2011).
- Morton, D., Leach, S., Cordier, C., Warriner, S. & Nelson, A. Synthesis of natural-product-like molecules with over eighty distinct scaffolds. *Angew. Chem. Int. Ed. Engl.* **48**, 104–109 (2009).
- Siau, W. Y. & Bode, J. W. One-step synthesis of saturated spirocyclic N-heterocycles with stannyl amine protocol (SnAP) reagents and ketones. *J. Am. Chem. Soc.* **136**, 17726–17729 (2014).
- Schreiber, S. L. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* **287**, 1964–1969 (2000).
- Wetzel, S., Bon, R. S., Kumar, K. & Waldmann, H. Biology-oriented synthesis. *Angew. Chem. Int. Ed. Engl.* **50**, 10800–10826 (2011).
- Nielsen, T. E. & Schreiber, S. L. Towards the optimal screening collection: a synthesis strategy. *Angew. Chem. Int. Ed. Engl.* **47**, 48–56 (2008).
- Burke, M. D. & Schreiber, S. L. A. Planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed. Engl.* **43**, 46–58 (2004).
- Lee, D., Sello, J. K. & Schreiber, S. L. Pairwise use of complexity-generating reactions in diversity-oriented organic synthesis. *Org. Lett.* **2**, 709–712 (2000).
- Tan, D. S., Foley, M. A., Stockwell, B. R., Shair, M. D. & Schreiber, S. L. Synthesis and preliminary evaluation of a library of polycyclic small molecules for use in chemical genetic assays. *J. Am. Chem. Soc.* **121**, 9073–9087 (1999).
- Lee, D., Sello, J. K. & Schreiber, S. L. A. Strategy for macrocyclic ring closure and functionalization aimed toward split-pool syntheses. *J. Am. Chem. Soc.* **121**, 10648–10649 (1999).
- Oguri, H. & Schreiber, S. L. Skeletal diversity via a folding pathway: synthesis of indole alkaloid-like skeletons. *Org. Lett.* **7**, 47–50 (2005).
- Kumagai, N., Muncipinto, G. & Schreiber, S. L. Short synthesis of skeletally and stereochemically diverse small molecules by coupling petasis condensation reactions to cyclization reactions. *Angew. Chem. Int. Ed. Engl.* **45**, 3635–3638 (2006).
- Sternson, S. M., Wong, J. C., Grozinger, C. M. & Schreiber, S. L. Synthesis of 7200 small molecules based on a substructural analysis of the histone deacetylase inhibitors trichostatin and tapoxin. *Org. Lett.* **3**, 4239–4242 (2001).
- Haggarty, S. J., Koeller, K. M., Wong, J. C., Butcher, R. A. & Schreiber, S. L. Multidimensional chemical genetic analysis of diversity-oriented synthesis-derived deacetylase inhibitors using cell-based assays. *Chem. Biol.* **10**, 383–396 (2003).
- Haggarty, S. J., Koeller, K. M., Wong, J. C., Grozinger, C. M. & Schreiber, S. L. Domain-selective small-molecule inhibitor of histone deacetylase 6 (HDAC6)-mediated tubulin deacetylation. *Proc. Natl Acad. Sci. USA* **100**, 4389–4394 (2003).
- Hideshima, T. et al. Small-molecule inhibition of proteasome and aggresome function induces synergistic antitumor activity in multiple myeloma. *Proc. Natl Acad. Sci. USA* **102**, 8567–8572 (2005).
- Tan, D. S. Diversity-oriented synthesis: exploring the intersections between chemistry and biology. *Nat. Chem. Biol.* **1**, 74–84 (2005).
- Schreiber, S. L. Organic chemistry: molecular diversity by design. *Nature* **457**, 153–154 (2009).
- Schreiber, S. L. Organic synthesis toward small-molecule probes and drugs. *Proc. Natl Acad. Sci. USA* **108**, 6699–6702 (2011).
- Dandapani, S. & Marcaurelle, L. A. Current strategies for diversity-oriented synthesis. *Curr. Opin. Chem. Biol.* **14**, 362–370 (2010).
- Aicher, T. D. et al. Total synthesis of halichondrin B and norhalichondrin B. *J. Am. Chem. Soc.* **114**, 3162–3164 (1992).
- Towle, M. J. et al. In vitro and in vivo anticancer activities of synthetic macrocyclic ketone analogues of halichondrin B. *Cancer Res.* **61**, 1013–1021 (2001).
- Paterson, I. & Anderson, E. A. The renaissance of natural products as drug candidates. *Science* **310**, 451–453 (2005).
- Charest, M. G., Lerner, C. D., Brubaker, J. D., Siegel, D. R. & Myers, A. G. A convergent enantioselective route to structurally diverse 6-deoxytetracycline antibiotics. *Science* **308**, 395–398 (2005).
- Seiple, I. B. et al. A platform for the discovery of new macrocyclic antibiotics. *Nature* **533**, 338–345 (2016).
- Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
- Altshuler, D., Daly, M. J. & Lander, E. S. Genetic mapping in human disease. *Science* **322**, 881–888 (2008).
- Dugger, S. A., Platt, A. & Goldstein, D. B. Drug development in the era of precision medicine. *Nat. Rev. Drug Discov.* **17**, 183–196 (2017).
- Plenge, R. M., Scolnick, E. M. & Altshuler, D. Validating therapeutic targets through human genetics. *Nat. Rev. Drug Discov.* **12**, 581–594 (2013).
- The SIGMA Type 2 Diabetes Consortium et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* **506**, 97–101 (2013).
- Momozawa, Y. et al. Resequencing of positional candidates identifies low frequency IL23R coding variants protecting against inflammatory bowel disease. *Nat. Genet.* **43**, 43–47 (2011).
- Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381**, 1371–1379 (2013).

62. Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.* **43**, 977–983 (2011).
63. Sekar, A. et al. Schizophrenia risk from complex variation of complement component 4. *Nature* **530**, 177–183 (2016).
64. Bunnage, M. E., Piatnitski Chekler, E. L. & Jones, L. H. Target validation using chemical probes. *Nat. Chem. Biol.* **9**, 195–199 (2013).
65. Arrowsmith, C. H. et al. The promise and peril of chemical probes. *Nat. Chem. Biol.* **11**, 536–541 (2015).
66. Schreiber, S. L. et al. Advancing biological understanding and therapeutics discovery with small-molecule probes. *Cell* **161**, 1252–1265 (2015).
67. Swinney, D. C. Phenotypic versus target-based drug discovery for first-in-class medicines. *Clin. Pharmacol. Ther.* **93**, 299–301 (2013).
68. Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. & Prunotto, M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat. Rev. Drug Discov.* **16**, 531–543 (2017).
69. Gilbert, I. H. Drug discovery for neglected diseases: molecular target-based and phenotypic approaches. *J. Med. Chem.* **56**, 7719–7726 (2013).
70. Zheng, W., Thorne, N. & McKew, J. C. Phenotypic screens as a renewed approach for drug discovery. *Drug Discov. Today* **18**, 1067–1073 (2013).
71. Wang, Y. et al. Diversity-oriented synthesis as a strategy for fragment evolution against GSK3 β . *ACS Med. Chem. Lett.* **7**, 852–856 (2016).
72. Schaefer, G. I. et al. Discovery of small-molecule modulators of the sonic hedgehog pathway. *J. Am. Chem. Soc.* **135**, 9675–9680 (2013).
73. Chou, D. H.-C. et al. Synthesis of a novel suppressor of β -cell apoptosis via diversity-oriented synthesis. *ACS Med. Chem. Lett.* **2**, 698–702 (2011).
74. Gray, B. L., Wang, X., Brown, W. C., Kuai, L. & Schreiber, S. L. Diversity synthesis of complex pyridines yields a probe of a neurotrophic signaling pathway. *Org. Lett.* **10**, 2621–2624 (2008).
75. Aldrich, L. N. et al. Discovery of a small-molecule probe for V-ATPase function. *J. Am. Chem. Soc.* **137**, 5563–5568 (2015).
76. Cruz, D., Wang, Z., Kibbie, J., Modlin, R. & Kwon, O. Diversity through phosphine catalysis identifies octahydro-1,6-naphthyridin-4-ones as activators of endothelium-driven immunity. *Proc. Natl Acad. Sci. USA* **108**, 6769–6774 (2011).
77. Stanton, B. Z. et al. A small molecule that binds hedgehog and blocks its signaling in human cells. *Nat. Chem. Biol.* **5**, 154–156 (2009).
78. Kuo, S.-Y. et al. Small-molecule enhancers of autophagy modulate cellular disease phenotypes suggested by human genetics. *Proc. Natl Acad. Sci. USA* **112**, E4281–E4287 (2015).
79. Dockendorff, C. et al. Benzo-fused lactams from a diversity-oriented synthesis (DOS) library as inhibitors of scavenger receptor BI (SR-BI)-mediated lipid uptake. *Bioorg. Med. Chem. Lett.* **25**, 2100–2105 (2015).
80. Kim, J. et al. Diversity-oriented synthetic strategy for developing a chemical modulator of protein–protein interaction. *Nat. Commun.* **7**, 13196 (2016).
81. Nagiec, M. M. et al. Modulators of hepatic lipoprotein metabolism identified in a search for small-molecule inducers of tribbles pseudokinase 1 expression. *PLoS ONE* **10**, 1–26 (2015).
82. Dakas, P. Y. et al. Discovery of neurotogenic compound classes inspired by natural products. *Angew. Chem. Int. Ed. Engl.* **52**, 9576–9581 (2013).
83. Chiang, C. et al. Cyclopia and defective axial patterning in mice lacking sonic hedgehog gene function. *Nature* **383**, 407–413 (1996).
84. Ingham, P. W. & McMahon, A. P. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.* **15**, 3059–3087 (2001).
85. Thayer, S. P. et al. Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. *Nature* **425**, 851–856 (2003).
86. Roessler, E. et al. Mutations in the human sonic hedgehog gene cause holoprosencephaly. *Nat. Genet.* **14**, 357–360 (1996).
87. Stanton, B. Z. & Peng, L. F. Small-molecule modulators of the sonic hedgehog signaling pathway. *Mol. Biosyst.* **6**, 44–54 (2010).
88. Bradner, J. E. et al. A robust small-molecule microarray platform for screening cell lysates. *Chem. Biol.* **13**, 493–504 (2006).
89. Owens, A. E. et al. Design and evolution of a macrocyclic peptide inhibitor of the sonic hedgehog/patched interaction. *J. Am. Chem. Soc.* **139**, 12559–12568 (2017).
90. Peng, L. F., Stanton, B. Z., Maloof, N., Wang, X. & Schreiber, S. L. Syntheses of aminoalcohol-derived macrocycles leading to a small-molecule binder to and inhibitor of sonic hedgehog. *Bioorg. Med. Chem. Lett.* **19**, 6319–6325 (2009).
91. Justilien, V. & Fields, A. P. Molecular pathways: novel approaches for improved therapeutic targeting of hedgehog signaling in cancer stem cells. *Clin. Cancer Res.* **21**, 505–513 (2015).
92. Chou, D. H. C. et al. Kinase-independent small-molecule inhibition of JAK–STAT signaling. *J. Am. Chem. Soc.* **137**, 7929–7934 (2015).
93. Scully, S. S. et al. Small-molecule inhibitors of cytokine-mediated STAT1 signal transduction in beta-cells with improved aqueous solubility. *J. Med. Chem.* **56**, 4125–4129 (2013).
94. Marcaurrelle, L. A. et al. An aldol-based build/couple/pair strategy for the synthesis of medium- and large-sized rings: discovery of macrocyclic histone deacetylase inhibitors. *J. Am. Chem. Soc.* **132**, 16962–16976 (2010).
95. Feng, Y., He, D., Yao, Z. & Klionsky, D. J. The machinery of macroautophagy. *Cell Res.* **24**, 24–41 (2014).
96. Hidvegi, T. et al. An autophagy-enhancing drug promotes degradation of mutant α -1-Antitrypsin Z and reduces hepatic fibrosis. *Science* **329**, 229–232 (2010).
97. Rubinsztein, D. C., Codogno, P. & Levine, B. Autophagy modulation as a potential therapeutic target for diverse diseases. *Nat. Rev. Drug Discov.* **11**, 709–730 (2012).
98. Kim, J., Kundu, M., Viollet, B. & Guan, K.-L. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nat. Cell Biol.* **13**, 132–141 (2011).
99. Ravikumar, B. et al. Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nat. Genet.* **36**, 585–595 (2004).
100. Martins, F. et al. A review of oral cytotoxicity associated with mTOR inhibitor therapy in cancer patients. *Oral Oncol.* **49**, 293–298 (2013).
101. Fitzgerald, M. E. et al. Build/couple/pair strategy for the synthesis of stereochemically diverse macrolactams via head-to-tail cyclization. *ACS Comb. Sci.* **14**, 89–96 (2012).
102. Galloway, W. R. J. D., Isidro-Llobet, A. & Spring, D. R. Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. *Nat. Commun.* **1**, 1–13 (2010).
103. Burke, M. D., Berger, E. M. & Schreiber, S. L. A selective strategy yielding skeletally diverse small molecules combinatorially. *J. Am. Chem. Soc.* **126**, 14095–14104 (2004).
104. Sauer, W. H. B. & Schwarz, M. K. Size doesn't matter: scaffold diversity, shape diversity and biological activity of combinatorial libraries. *Chimia* **57**, 276–283 (2003).
105. Moffat, J. G., Rudolph, J. & Bailey, D. Phenotypic screening in cancer drug discovery – past, present and future. *Nat. Rev. Drug Discov.* **13**, 588–602 (2014).
106. Fesik, S. W. Promoting apoptosis as a strategy for cancer drug discovery. *Nat. Rev. Cancer* **5**, 876–885 (2005).
107. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
108. Fruman, D. A. & Rommel, C. PI3K and cancer: lessons, challenges and opportunities. *Nat. Rev. Drug Discov.* **13**, 140–156 (2014).
109. Sawyers, C. Targeted cancer therapy. *Nature* **432**, 294–297 (2004).
110. Bredel, M. & Jacoby, E. Chemogenomics: an emerging strategy for rapid target and drug discovery. *Nat. Rev. Genet.* **5**, 262–275 (2004).
111. Jones, S. et al. Personalized genomic analyses for cancer mutation discovery and interpretation. *Sci. Transl. Med.* **7**, 283ra53 (2015).
112. Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J. & Der, C. J. Drugging the undruggable RAS: mission possible? *Nat. Rev. Drug Discov.* **13**, 828–851 (2014).
113. Verdine, G. L. & Walensky, L. D. The challenge of drugging undruggable targets in cancer: lessons learned from targeting BCL-2 family members. *Clin. Cancer Res.* **13**, 7264–7270 (2007).
114. Yan, C. & Higgins, P. J. Drugging the undruggable: transcription therapy for cancer. *Biochim. Biophys. Acta* **1835**, 76–85 (2013).
115. Ibbeson, B. M. et al. Diversity-oriented synthesis as a tool for identifying new modulators of mitosis. *Nat. Commun.* **5**, 1–8 (2014).
116. Santagata, S. et al. Tight coordination of protein translation and HSF1 activation supports the anabolic malignant state. *Science* **341**, 1258303 (2013).
117. Law, J. M. et al. Discovery of 8-membered ring sulfonamides as inhibitors of oncogenic mutant isocitrate dehydrogenase 1. *ACS Med. Chem. Lett.* **7**, 944–949 (2016).
118. Dückert, H. et al. Natural product-inspired cascade synthesis yields modulators of centrosome integrity. *Nat. Chem. Biol.* **8**, 179–184 (2011).
119. Yu, C. et al. High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. *Nat. Biotechnol.* **34**, 419–423 (2016).
120. Bachovchin, D. A. et al. Academic cross-fertilization by public screening yields a remarkable class of protein phosphatase methyltransferase-1 inhibitors. *Proc. Natl Acad. Sci. USA* **108**, 6811–6816 (2011).
121. Ng, P. Y., Tang, Y., Knosp, W. M., Stadler, H. S. & Shaw, J. T. Synthesis of diverse lactam carboxamides leading to the discovery of a new transcription-factor inhibitor. *Angew. Chem. Int. Ed. Engl.* **46**, 5352–5355 (2007).
122. Tang, W., Luo, T., Greenberg, E. F., Bradner, J. E. & Schreiber, S. L. Discovery of histone deacetylase 8 selective inhibitors. *Bioorg. Med. Chem. Lett.* **21**, 2601–2605 (2011).
123. Koehler, A. N., Shamji, A. F. & Schreiber, S. L. Discovery of an inhibitor of a transcription factor using small molecule microarrays and diversity-oriented synthesis. *J. Am. Chem. Soc.* **125**, 8420–8421 (2003).
124. Huryn, D. M. et al. Chemical methodology as a source of small-molecule checkpoint inhibitors and heat shock protein 70 (Hsp70) modulators. *Proc. Natl Acad. Sci. USA* **108**, 6757–6762 (2011).
125. Fang, C. et al. Single diastereomer of a macrolactam core binds specifically to myeloid cell leukemia 1 (MCL1). *ACS Med. Chem. Lett.* **5**, 1308–1312 (2014).
126. Clark, P. G. K. et al. LP99: discovery and synthesis of the first selective BRD7/9 bromodomain inhibitor. *Angew. Chem. Int. Ed. Engl.* **54**, 6217–6221 (2015).
127. Frumm, S. M. et al. Selective HDAC1/HDAC2 inhibitors induce neuroblastoma differentiation. *Chem. Biol.* **20**, 715–725 (2013).
128. Carmona, M. & Earnshaw, W. C. The cellular geography of aurora kinases. *Nat. Rev. Mol. Cell Biol.* **4**, 842–854 (2003).
129. Bischoff, J. R. et al. A homologue of *Drosophila* aurora kinase is oncogenic and amplified in human colorectal cancers. *EMBO J.* **17**, 3052–3065 (1998).
130. Harrington, E. A. et al. VX-680, a potent and selective small-molecule inhibitor of the aurora kinases, suppresses tumor growth in vivo. *Nat. Med.* **10**, 262–267 (2004).
131. Bavetsias, V. & Linardopoulos, S. Aurora kinase inhibitors: current status and outlook. *Front. Oncol.* **5**, 278 (2015).
132. Melichar, B. et al. Safety and activity of alisertib, an investigational aurora kinase A inhibitor, in patients with breast cancer, small-cell lung cancer, non-small-cell lung cancer, head and neck squamous-cell carcinoma, and gastro-oesophageal adenocarcinoma: a five-arm phase. *Lancet Oncol.* **16**, 395–405 (2015).
133. Dar, A. A., Goff, L. W., Majid, S., Berlin, J. & El-Rifai, W. Aurora kinase inhibitors — rising stars in cancer therapeutics? *Mol. Cancer Ther.* **9**, 268–278 (2010).
134. Zhang, J., Yang, P. L. & Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* **9**, 28–39 (2009).
135. Sonna, L. A., Fujita, J., Gaffin, S. L. & Lilly, C. M. Invited review: effects of heat and cold stress on mammalian gene expression. *J. Appl. Physiol.* **92**, 1725–1742 (2002).
136. Santagata, S. et al. High levels of nuclear heat-shock factor 1 (HSF1) are associated with poor prognosis in breast cancer. *Proc. Natl Acad. Sci. USA* **108**, 18378–18383 (2011).
137. Dai, C., Whitesell, L., Rogers, A. B. & Lindquist, S. Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. *Cell* **130**, 1005–1018 (2007).
138. Meng, L., Gabai, V. L. & Sherman, M. Y. Heat-shock transcription factor HSF1 has a critical role in human epidermal growth factor receptor-2-induced cellular transformation and tumorigenesis. *Oncogene* **29**, 5204–5213 (2010).

139. Whitesell, L. & Lindquist, S. Inhibiting the transcription factor HSF1 as an anticancer strategy. *Expert Opin. Ther. Targets* **13**, 469–478 (2009).
140. Koehler, A. N. A complex task? Direct modulation of transcription factors with small molecules. *Curr. Opin. Chem. Biol.* **14**, 331–340 (2010).
141. Gerard, B., Jones, G. & Porco, J. A. A Biomimetic approach to the rocaglamides employing photogeneration of oxidopyryliums derived from 3-hydroxyflavones. *J. Am. Chem. Soc.* **126**, 13620–13621 (2004).
142. Roche, S. P., Cencic, R., Pelletier, J. & Porco, J. A. Biomimetic photocycloaddition of 3-hydroxyflavones: synthesis and evaluation of rocaglate derivatives as inhibitors of eukaryotic translation. *Angew. Chem. Int. Ed. Engl.* **49**, 6533–6538 (2010).
143. Gerard, B., Sangji, S., O'Leary, D. J. & Porco, J. A. Enantioselective photocycloaddition mediated by chiral brønsted acids: asymmetric synthesis of the rocaglamides. *J. Am. Chem. Soc.* **128**, 7754–7755 (2006).
144. Rodrigo, C. M., Cencic, R., Roche, S. P., Pelletier, J. & Porco, J. A. Synthesis of rocaglamide hydroxamates and related compounds as eukaryotic translation inhibitors: synthetic and biological studies. *J. Med. Chem.* **55**, 558–562 (2012).
145. Kárkás, M. D. et al. Photochemical approaches to complex chemotypes: applications in natural product synthesis. *Chem. Rev.* **116**, 9683–9747 (2016).
146. Wilson, B. G. & Roberts, C. W. M. SWI/SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer* **11**, 481–492 (2011).
147. Reisman, D., Glaros, S. & Thompson, E. A. The SWI/SNF complex and cancer. *Oncogene* **28**, 1653–1668 (2009).
148. Drost, J. et al. BRD7 is a candidate tumour suppressor gene required for p53 function. *Nat. Cell Biol.* **12**, 380–389 (2010).
149. Hohmann, A. F. et al. Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition. *Nat. Chem. Biol.* **12**, 672–679 (2016).
150. Chaikuad, A., Petros, A. M., Fedorov, O., Xu, J. & Knapp, S. Structure-based approaches towards identification of fragments for the low-druggability ATAD2 bromodomain. *Med. Chem. Commun.* **5**, 1843–1848 (2014).
151. Saha, S. et al. A Phosphatase associated with metastasis of colorectal cancer. *Science* **294**, 1343–1346 (2001).
152. Li, J. et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* **275**, 1943–1947 (1997).
153. Tonks, N. K. Protein tyrosine phosphatases — from housekeeping enzymes to master regulators of signal transduction. *FEBS J.* **280**, 346–378 (2013).
154. Lazo, J. S. & Sharlow, E. R. Drugging undruggable molecular cancer targets. *Annu. Rev. Pharmacol. Toxicol.* **56**, 23–40 (2016).
155. Khanna, A., Pimanda, J. E. & Westermarck, J. Cancerous inhibitor of protein phosphatase 2A, an emerging human oncoprotein and a potential cancer therapy target. *Cancer Res.* **73**, 6548–6553 (2013).
156. Ciccone, M., Calin, G. A. & Perrotti, D. From the biology of PP2A to the PADs for therapy of hematologic malignancies. *Front. Oncol.* **5**, 21 (2015).
157. Xing, Y. et al. Structural mechanism of demethylation and inactivation of protein phosphatase 2A. *Cell* **133**, 154–163 (2008).
158. Wandzioch, E. et al. PME-1 modulates protein phosphatase 2A activity to promote the malignant phenotype of endometrial cancer cells. *Cancer Res.* **74**, 4295–4305 (2014).
159. Berlin, J. M. & Fu, G. C. Enantioselective nucleophilic catalysis: the synthesis of aza- β -lactams through [2 + 2] cycloadditions of ketenes with azo compounds. *Angew. Chem. Int. Ed. Engl.* **47**, 7048–7050 (2008).
160. Zuhl, A. M. et al. Competitive activity-based protein profiling identifies Aza- β -lactams as a versatile chemotype for serine hydrolase inhibition. *J. Am. Chem. Soc.* **134**, 5068–5071 (2012).
161. Little, S. J. et al. Antiretroviral-drug resistance among patients recently infected with HIV. *N. Engl. J. Med.* **347**, 385–394 (2002).
162. Cohen, M. L. Epidemiology of drug resistance: implications for a post-antimicrobial era. *Science* **257**, 1050–1055 (1992).
163. Katsuno, K. et al. Hit and lead criteria in drug discovery for infectious diseases of the developing world. *Nat. Rev. Drug Discov.* **14**, 751–758 (2015).
164. Lewis, K. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **12**, 371–387 (2013).
165. Davies, J. & Davies, D. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* **74**, 417–433 (2010).
166. Farha, M. A. & Brown, E. D. Strategies for target identification of antimicrobial natural products. *Nat. Prod. Rep.* **33**, 668–680 (2016).
167. Chu, J. et al. Discovery of MRSA active antibiotics using primary sequence from the human microbiome. *Nat. Chem. Biol.* **12**, 1004–1006 (2016).
168. Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **6**, 29–40 (2007).
169. Wensing, A. M. J., van Maarseveen, N. M. & Nijhuis, M. Fifteen years of HIV protease inhibitors: raising the barrier to resistance. *Antiviral Res.* **85**, 59–74 (2010).
170. Scheel, T. K. H. & Rice, C. M. Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nat. Med.* **19**, 837–849 (2013).
171. 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).
172. Fischbach, M. A. & Walsh, C. T. Antibiotics for emerging pathogens. *Science* **325**, 1089–1093 (2009).
173. Wright, P. M., Seiple, I. B. & Myers, A. G. The evolving role of chemical synthesis in antibacterial drug discovery. *Angew. Chem. Int. Ed. Engl.* **53**, 8840–8869 (2014).
174. Lukens, A. K. et al. Diversity-oriented synthesis probe targets *Plasmodium falciparum* cytochrome *b* ubiquinone reduction site and synergizes with oxidation site inhibitors. *J. Infect. Dis.* **211**, 1097–1103 (2015).
175. Maetani, M. et al. Discovery of antimalarial azetidine-2-carbonitriles that inhibit *P. falciparum* dihydroorotate dehydrogenase. *ACS Med. Chem. Lett.* **8**, 438–442 (2017).
176. Plouffe, D. M. et al. High-throughput assay and discovery of small molecules that interrupt malaria transmission. *Cell Host Microbe* **19**, 114–126 (2016).
177. Rottmann, M. et al. Spiroindolones, a potent compound class for the treatment of malaria. *Science* **329**, 1175–1180 (2010).
178. Kato, N. et al. Diversity-oriented synthesis yields novel multistage antimalarial inhibitors. *Nature* **538**, 344–349 (2016).
179. Wellington, S. et al. A small-molecule allosteric inhibitor of *Mycobacterium tuberculosis* tryptophan synthase. *Nat. Chem. Biol.* **13**, 943–950 (2017).
180. Park, S. W. et al. Target-based identification of whole-cell active inhibitors of biotin biosynthesis in *Mycobacterium tuberculosis*. *Chem. Biol.* **22**, 76–86 (2015).
181. Dandapani, S. et al. Diversity-oriented synthesis yields a new drug lead for treatment of Chagas disease. *ACS Med. Chem. Lett.* **5**, 149–153 (2014).
182. Thomas, G. L. et al. Anti-MRSA agent discovery using diversity-oriented synthesis. *Angew. Chem. Int. Ed. Engl.* **47**, 2808–2812 (2008).
183. Duvall, J. R. et al. Novel diversity-oriented synthesis-derived respiratory syncytial virus inhibitors identified via a high throughput replicon-based screen. *Antiviral Res.* **131**, 19–25 (2016).
184. Nuhs, A. et al. Development and validation of a novel *Leishmania donovani* screening cascade for high-throughput screening using a novel axenic assay with high predictivity of leishmanicidal intracellular activity. *PLoS Negl. Trop. Dis.* **9**, 1–17 (2015).
185. Duvall, J. R. et al. Identification of highly specific diversity-oriented synthesis-derived inhibitors of *Clostridium difficile*. *ACS Infect. Dis.* **3**, 349–359 (2017).
186. Baniecki, M. L., Wirth, D. F. & Clardy, J. High-throughput *Plasmodium falciparum* growth assay for malaria drug discovery. *Antimicrob. Agents Chemother.* **51**, 716–723 (2007).
187. Flannery, E. L., Chatterjee, A. K. & Winzler, E. A. Antimalarial drug discovery — approaches and progress towards new medicines. *Nat. Rev. Microbiol.* **11**, 849–862 (2013).
188. Campo, B., Vandal, O., Wesche, D. L. & Burrows, J. N. Killing the hypnozoite — drug discovery approaches to prevent relapse in *Plasmodium vivax*. *Pathog. Glob. Health* **109**, 107–122 (2015).
189. International Artemisinin Study Group. Artesunate combinations for malaria. *Lancet* **365**, 9–17 (2004).
190. White, N. J. Antimalarial drug resistance. *J. Clin. Invest.* **113**, 1084–1092 (2004).
191. Winzler, E. A. & Manary, M. J. Drug resistance genomics of the antimalarial drug artemisinin. *Genome Biol.* **15**, 544 (2014).
192. Klayman, D. L. Qinghaosu (artemisinin): an antimalarial drug from China. *Science* **228**, 1049–1055 (1985).
193. Dondorp, A. M. et al. Artemisinin resistance in *Plasmodium falciparum* malaria. *N. Engl. J. Med.* **361**, 455–467 (2009).
194. Dyer, M., Jackson, M., McWhinney, C., Zhao, G. & Mikkelsen, R. Analysis of a cation-transporting ATPase of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **78**, 1–12 (1996).
195. Spillman, N. J. et al. Na⁺ regulation in the malaria parasite *Plasmodium falciparum* involves the cation ATPase PfATP4 and is a target of the spiroindolone antimalarials. *Cell Host Microbe* **13**, 227–237 (2013).
196. White, N. J. et al. Spiroindolone KAE609 for falciparum and vivax malaria. *N. Engl. J. Med.* **371**, 403–410 (2014).
197. Van Pelt-Koops, J. C. et al. The spiroindolone drug candidate NITD609 potently inhibits gametocytogenesis and blocks *Plasmodium falciparum* transmission to *Anopheles* mosquito vector. *Antimicrob. Agents Chemother.* **56**, 3544–3548 (2012).
198. Flannery, E. L. et al. Mutations in the P-type cation-transporter PfATP4, PfATP4, mediate resistance to both aminopyrazole and spiroindolone antimalarials. *ACS Chem. Biol.* **10**, 413–420 (2015).
199. Vaidya, A. B. et al. Pyrazoleamide compounds are potent antimalarials that target Na(+) homeostasis in intraerythrocytic *Plasmodium falciparum*. *Nat. Commun.* **5**, 5521 (2014).
200. Jiménez-Díaz, M. B. et al. (+)-SJ733, a clinical candidate for malaria that acts through ATP4 to induce rapid host-mediated clearance of *Plasmodium*. *Proc. Natl Acad. Sci. USA* **111**, E5455–E5462 (2014).
201. Ang, S. H. et al. Spiro-indole derivatives for the treatment of parasitic diseases. Int. Patent EP2009/053902 (2009).
202. Badillo, J. J., Silva-García, A., Shupe, B. H., Fetting, J. C. & Franz, A. K. Enantioselective plectrospengler reactions of isatins for the synthesis of spiroindolones. *Tetrahedron Lett.* **52**, 5550–5553 (2011).
203. Zheng, H. et al. Regio- and enantioselective aza-diels-alder reactions of 3-vinylindoles: a concise synthesis of the antimalarial spiroindolone NITD609. *Angew. Chem. Int. Ed. Engl.* **54**, 10958–10962 (2015).
204. Heidebrecht Jr, R. W. et al. Diversity-oriented synthesis yields a novel lead for the treatment of malaria. *ACS Med. Chem. Lett.* **3**, 112–117 (2012).
205. Comer, E. et al. Diversity-oriented synthesis-facilitated medicinal chemistry: toward the development of novel antimalarial agents. *J. Med. Chem.* **57**, 8496–8502 (2014).
206. Lowe, J. T. et al. Synthesis and profiling of a diverse collection of azetidine-based scaffolds for the development of CNS-focused lead-like libraries. *J. Org. Chem.* **77**, 7187–7211 (2012).
207. Maetani, M. et al. Synthesis of a bicyclic azetidine in vivo antimalarial activity enabled by stereospecific, directed C(sp³)-H arylation. *J. Am. Chem. Soc.* **139**, 11300–11306 (2017).
208. Younis, Y. et al. 3,5-Diaryl-2-aminopyridines as a novel class of orally active antimalarials demonstrating single dose cure in mice and clinical candidate potential. *J. Med. Chem.* **55**, 3479–3487 (2012).
209. Baragaña, B. et al. A novel multiple-stage antimalarial agent that inhibits protein synthesis. *Nature* **522**, 315–320 (2015).
210. Cohen, S. H. et al. Clinical practice guidelines for *Clostridium difficile* infection in adults: 2010 update by the Society for Healthcare Epidemiology of America (SHEA) and the Infectious Diseases Society of America (IDSA). *Infect. Control Hosp. Epidemiol.* **31**, 431–455 (2010).
211. Kelly, C. P., Pothoulakis, C. & LaMont, J. T. *Clostridium difficile* colitis. *N. Engl. J. Med.* **330**, 257–262 (1994).
212. Louie, T. J., Emery, J., Krulicki, W., Byrne, B. & Mah, M. OPT-80 eliminates *Clostridium difficile* and is sparing of bacteroides species during treatment of *C. difficile* infection. *Antimicrob. Agents Chemother.* **53**, 261–263 (2009).

213. Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H. & Kuijper, E. J. *Clostridium difficile* infection. *Nat. Rev. Dis. Primers* **2**, 16020 (2016).
214. Fisher, S. L. Glutamate racemase as a target for drug discovery. *Microb. Biotechnol.* **1**, 345–360 (2008).
215. Gerard, B. et al. Application of a catalytic asymmetric povarov reaction using chiral ureas to the synthesis of a tetrahydroquinoline library. *ACS Comb. Sci.* **14**, 621–630 (2012).
216. World Health Organization. *Global tuberculosis report 2016* (WHO, 2016).
217. Dheda, K., Barry, C. E. & Maartens, G. Tuberculosis. *Lancet* **387**, 1211–1226 (2016).
218. Munro, S. A. et al. Patient adherence to tuberculosis treatment: a systematic review of qualitative research. *PLoS Med.* **4**, e238 (2007).
219. Dunn, M. F. Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Arch. Biochem. Biophys.* **519**, 154–166 (2012).
220. Nussinov, R. & Tsai, C.-J. Allostery in disease and in drug discovery. *Cell* **153**, 293–305 (2013).
221. Meijer, A. H. Protection and pathology in TB: learning from the zebrafish model. *Semin. Immunopathol.* **38**, 261–273 (2016).
222. Stanley, S. A. et al. Identification of novel inhibitors of *M. tuberculosis* growth using whole cell based high-throughput screening. *ACS Chem. Biol.* **7**, 1377–1384 (2012).
223. Silhavy, J., T., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* **2**, 1–17 (2010).
224. Frye, S. V. The art of the chemical probe. *Nat. Chem. Biol.* **6**, 159–161 (2010).
225. Wawer, M. J. et al. Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling. *Proc. Natl Acad. Sci. USA* **111**, 10911–10916 (2014).
226. Drews, J. Drug discovery: a historical perspective. *Science* **287**, 1960–1964 (2000).
227. Hann, M. M., Leach, A. R. & Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **41**, 856–864 (2001).
228. Hann, M. M. & Oprea, T. I. Pursuing the leadlikeness concept in pharmaceutical research. *Curr. Opin. Chem. Biol.* **8**, 255–263 (2004).
229. Kogej, T. et al. Big pharma screening collections: more of the same or unique libraries? The AstraZeneca–Bayer Pharma AG case. *Drug Discov. Today* **18**, 1014–1024 (2013).
230. Terai, T. & Nagano, T. Small-molecule fluorophores and fluorescent probes for bioimaging. *Pflügers Arch.* **465**, 347–359 (2013).
231. Sakamoto, K. M. et al. Protacs: chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. *Proc. Natl Acad. Sci. USA* **98**, 8554–8559 (2001).
232. Winter, G. E. et al. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348**, 1376–1381 (2015).
233. Bondeson, D. P. et al. Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat. Chem. Biol.* **11**, 611–617 (2015).
234. Meyer, B. & Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed. Engl.* **42**, 864–890 (2003).
235. Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **3**, 301–317 (2004).
236. Salamon, H., Klika Skopic, M., Jung, K., Bugain, O. & Brunschweiler, A. Chemical biology probes from advanced DNA-encoded libraries. *ACS Chem. Biol.* **11**, 296–307 (2016).
237. Franzini, R. M. & Randolph, C. Chemical space of DNA-encoded libraries. *J. Med. Chem.* **59**, 6629–6644 (2016).
238. Petrone, P. M. et al. Biodiversity of small molecules — a new perspective in screening set selection. *Drug Discov. Today* **18**, 674–680 (2013).
239. Subramanian, A. et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**, 1437–1452 (2017).
240. Gustafsdottir, S. M. et al. Multiplex cytological profiling assay to measure diverse cellular states. *PLoS ONE* **8**, e80999 (2013).
241. Kang, J. et al. Improving drug discovery with high-content phenotypic screens by systematic selection of reporter cell lines. *Nat. Biotechnol.* **34**, 70–77 (2016).
242. Gerry, C. J. et al. Real-time biological annotation of synthetic compounds. *J. Am. Chem. Soc.* **138**, 8920–8927 (2016).
243. Nelson Jr., S. D., Wawer, M. J. & Schreiber, S. L. Divergent synthesis and real-time biological annotation of optically active tetrahydrocyclopenta[c]pyranone derivatives. *Org. Lett.* **18**, 6280–6283 (2016).
244. Lajiness, M. S., Maggiora, G. M. & Shanmugasundaram, V. Assessment of the consistency of medicinal chemists in reviewing sets of compounds. *J. Med. Chem.* **47**, 4891–4896 (2004).
245. Hansch, C. The advent and evolution of QSAR at Pomona College. *J. Comput. Aided Mol. Des.* **25**, 495–507 (2011).
246. Lipinski, C. Chris Lipinski. Interview by Peter Kirkpatrick. *Nat. Rev. Drug Discov.* **11**, 900–901 (2012).
247. Lipinski, C. & Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **432**, 855–861 (2004).
248. Baell, J. & Walters, M. A. Chemistry: chemical con artists foil drug discovery. *Nature* **513**, 481–483 (2014).
249. Veber, D. F. et al. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **45**, 2615–2623 (2002).
250. Tian, S. et al. The application of in silico drug-likeness predictions in pharmaceutical research. *Adv. Drug Deliv. Rev.* **86**, 2–10 (2015).
251. Blake, J. F. Integrating cheminformatic analysis in combinatorial chemistry. *Curr. Opin. Chem. Biol.* **8**, 407–411 (2004).
252. Haftchenary, S. et al. Efficient routes to a diverse array of amino alcohol-derived chiral fragments. *ACS Comb. Sci.* **18**, 569–574 (2016).
253. Ghose, A. K., Viswanadhan, V. N. & Wendoloski, J. J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative quantitative characterization of known drug databases. *J. Comb. Chem.* **1**, 55–68 (1999).
254. Garbaccio, R. M. & Parmee, E. R. The impact of chemical probes in drug discovery: a pharmaceutical industry perspective. *Cell Chem. Biol.* **23**, 10–17 (2016).
255. Frye, S. V. et al. Tackling reproducibility in academic preclinical drug discovery. *Nat. Rev. Drug Discov.* **14**, 733–734 (2015).
256. Workman, P. & Collins, I. Probing the probes: fitness factors for small molecule tools. *Chem. Biol.* **17**, 561–577 (2010).
257. Blagg, J. & Workman, P. Choose and use your chemical probe wisely to explore cancer biology. *Cancer Cell* **32**, 9–25 (2017).
258. Skuta, C. et al. Probes and drugs portal: an interactive, open data resource for chemical biology. *Nat. Methods* **14**, 759–760 (2017).
259. Antolin, A. A. et al. Objective, quantitative, data-driven assessment of chemical probes. *Cell Chem. Biol.* **25**, 194–205 (2017).
260. Cornu, M., Albert, V. & Hall, M. N. mTOR in aging, metabolism, and cancer. *Curr. Opin. Genet. Dev.* **23**, 53–62 (2013).
261. Batlle, D. & Haque, S. K. Genetic causes and mechanisms of distal renal tubular acidosis. *Nephrol. Dial. Transplant.* **27**, 3691–3704 (2012).
262. Bhargava, A. et al. Osteopetrosis mutation R444L causes endoplasmic reticulum retention and misprocessing of vacuolar H⁺-ATPase α 3 subunit. *J. Biol. Chem.* **287**, 26829–26839 (2012).

Acknowledgements

The authors thank B. Melillo, T. Sharifnia, and B. Wagner for providing critical feedback on the manuscript. C.J.G. is supported by a National Science Foundation Graduate Research Fellowship (DGE1144152). S.L.S. is an investigator at the Howard Hughes Medical Institute. Research in his laboratory in the area covered in this Review has been generously supported for three decades by the National Institute of General Medical Sciences (GM038627).

Competing interests

The authors declare no competing interests.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

RELATED LINKS

Chemical Probes Portal: <http://www.chemicalprobes.org/>
 Molinspiration molecular properties calculator: <http://www.molinspiration.com/cgi-bin/properties>
 Probe Miner: <http://probeminer.ic.ac.uk/>
 The Probes & Drugs Portal: <https://www.probes-drugs.org/home/>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF