CHAPTER 2

The Application of Diversity-oriented Synthesis in Chemical Biology

NATALIA MATEU, SARAH L. KIDD, THOMAS J. OSBERGER, HANNAH L. STEWART, SEAN BARTLETT, WARREN R. J. D. GALLOWAY, ANDREW J. P. NORTH, HANNAH F. SORE AND DAVID R. SPRING*

University of Cambridge, Department of Chemistry, Lensfield Road, Cambridge, CB2 1EW, UK
*Email: spring@ch.cam.ac.uk

2.1 Introduction

Communicable and non-communicable diseases are prevalent worldwide. Whilst treatments and cures exist for several of these, for many more they are either ineffective or non-existent. Cancer, malaria and antibiotic resistance are some representative examples of conditions that cause huge burdens on worldwide healthcare. However, despite research and development investment being higher than ever, the biomedical community struggles to provide effective solutions for such devastating diseases. This can be attributed to the inability to identify appropriate targets to modulate the disease, the lack of suitable compounds to interact with identified targets, or the failure of compounds to pass through clinical trials. In such a situation, the question remains as to where the problem fundamentally lies and how to address it. This chapter will analyse the importance of screening

Chemical Biology No. 10
Chemical and Biological Synthesis: Enabling Approaches for Understanding Biology
Edited by Nick J. Westwood and Adam Nelson

Published by the Royal Society of Chemistry, www.rsc.org

high-quality small-molecule collections to facilitate the identification of novel biologically-active small molecules and how different diversity-oriented synthesis (DOS) approaches have efficiently addressed this issue.

2.1.1 Small Molecule Screening Collections

Investigations into the function of different biological systems and the role of specific biological targets in a given disease can provide vital information for both chemical genetics and drug discovery. For both of these approaches, the identification of suitable small molecules that can modulate different pathways selectively is highly desirable. Within the field of drug discovery, failures resulting from low observed efficacy of drug candidates in clinical trials are often attributed to the inadequate prediction of the physiological response that may result from modulating a given biological target. Initial investigations into biological targets for their therapeutic potential are often first conducted *via* genetic association studies and knockout models using DNA mutations.[1] However, these predicted effects do not always translate into those observed when a target is modulated using small molecules. In contrast, the use of small molecules to selectively perturb the biological function of macromolecules, such as proteins, has been validated as an effective means to study biological systems and thus may provide more reliable tools for predicting biological responses.[2] Nevertheless, the identification of effective small molecules (*probes* or *'tool compounds'*) capable of modulating a target of interest is significantly challenging. These tool compounds not only provide new understanding of biological targets and pathways related to a given disease (*target identification/validation*), but can also form the first step towards the development of a new medicine (*hit identification*) (Figure 2.1). In cases where the biological target is well defined, rational design of potential modulators is often possible, particularly when the structure of a natural substrate is known. However, this is not possible in the absence of structural information, and hence the discovery of these chemical probes relies on the screening of small-molecule libraries. High-throughput screening (HTS) of large small-molecule libraries is one of the most commonly employed techniques for this purpose.[3] In these assays, the quality of the compound collection, in terms of composition and size, ultimately determines the chances of finding good hits.[4] It is therefore vital to identify the key chemical features that constitute a high-quality screening collection in order to facilitate the discovery of new small-molecule modulators for challenging, and currently underexplored, biological targets.

2.1.1.1 Structural Diversity

An enlightening analysis carried out by GlaxoSmithKline (GSK) after their failure to identify new tractable antibiotic candidates from HTS[5] highlighted an underlying problem within pharmaceutical screening collections. Screening libraries have grown massively in size, but there has not been a

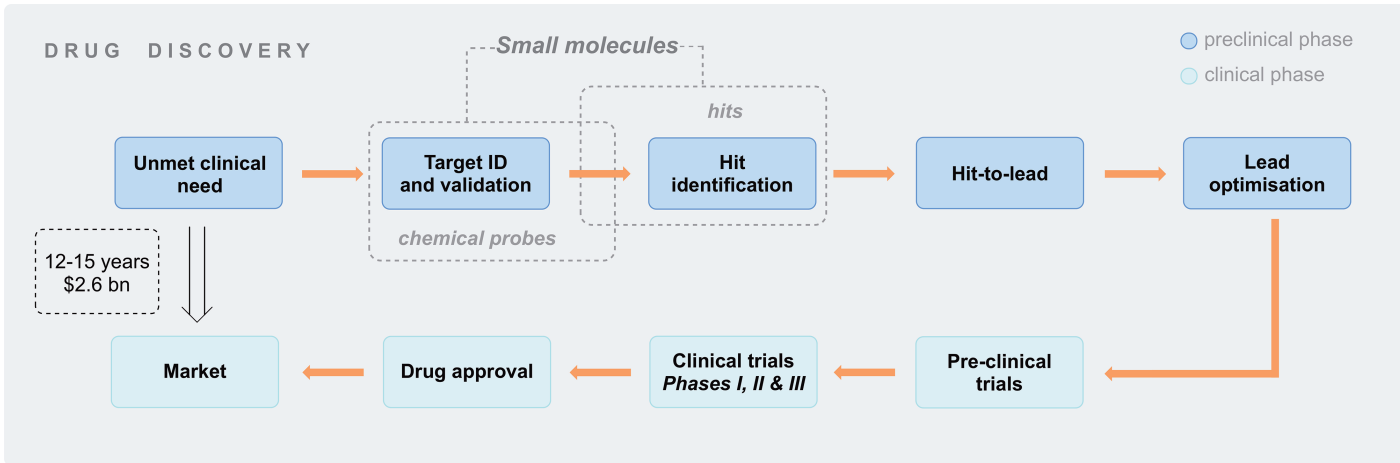


Figure 2.1 The involvement of small molecules in the early stages of drug discovery.

corresponding increase in the number of new structural classes or scaffolds within these collections.[4] Thus, there is a lack of *structural diversity* within these screening collections. It has been proposed that this is a result of the way the libraries are constructed. Combinatorial-type libraries are typically generated by combining a number of building blocks in different ways using the same synthetic methods, resulting in similar scaffolds with varied substituents; hence these libraries possess low levels of structural diversity.[6,7] Analysis supports the suggestion that increasing the library size or number of peripheral substituents does not significantly increase library diversity; instead, the molecular shape distribution of the library as a whole tends to be determined by the nature of the central scaffolds.[8]

An additional factor which has contributed to the structural bias of existing libraries towards similar structures is the fact that medicinal chemistry research over the past few decades has typically focused on a limited set of biological targets.[9] Approximately 29% of all drugs are enzyme modulators and nearly 36% act upon G-protein-coupled receptors.[7,10,11] In addition, even within major target classes, proteins have been studied highly unevenly.[12] As such, most libraries are directed towards identifying modulators of these and other ''traditional'' targets, including a strong preference for molecules that fulfil Lipinski's 'Rule of five' for oral bioavailability.[13–15] To add to the problem, a significant overlap between synthetic screening libraries across the pharmaceutical industry will be present, since the same chemical methods and vendors have often been used.[5]

2.1.1.2 Structural Complexity

In the early 1990s, within drug discovery, drug attrition predominately occurred in early clinical trials as a result of adverse pharmacokinetics and bioavailability.[16] Since the pharmaceutical industry has responded to these issues by improving the physicochemical properties (so called drug-like properties) of the lead candidates, there has been a shift in attrition towards late-stage clinical trials. The primary causes of late-stage attrition currently include efficacy and safety issues, such as toxicology and clinical safety.[16–18]

Studies exploring these causes of late-stage attrition[15–30] revealed that an increase in the *molecular complexity* of a lead is associated with increased potency and decreased toxicity. This was rationalised through the understanding that toxicity is often due to off-target effects.[31] Given that biological macromolecules can be considered as three-dimensional (3D) environments with defined binding regions,[6,7] a given biological macromolecule will only interact with molecules that have a complementary 3D binding surface.[6,7,32,33] Crucially, it is the molecular structure which determines a molecule's 3D shape.[7,33] Thus, by increasing their molecular structural complexity, both drug-candidates or chemical probes should have a better complementarity for a specific target and a decreased chance of binding off-targets, aiding further investigations.[22,34]

Molecular complexity is an intuitive principle that has been quantitatively defined in a number of ways. However, it is most often judged using the fraction of sp^3 carbons (Fsp^3) and the number of chiral centres; increased saturation and stereochemical complexity corresponds to improved selectivity and decreased toxicity in clinical candidates.[19,22,23,34,35] Therefore, going forwards, high-quality screening libraries will need to incorporate these features.

2.1.1.3 Chemical Space Coverage

Chemical space is a chemoinformatic concept that refers to all possible molecules that could be, in principle, synthesised as defined by a set of preselected physicochemical and topological descriptors. The chemical space of 'drug-like' molecules (or potential hits) has been estimated to be in the order of 10^{62} achievable compounds.[36] From a biomedicinal point of view, the relationship between biology and chemistry has been described as biologically-relevant chemical space.[37] This defines the regions of chemical space that are populated by biologically active molecules; although some of these regions have been already explored with known bioactive molecules, this space remains largely ''unexplored''. Accordingly, in the context of hit identification, the ''ideal'' screening collection would encompass all ''drug-like'' chemical space and thus access all the biologically relevant regions, but this is physically unattainable. Computational investigations have revealed that the structural diversity and complexity of a screening library indeed correlates to the capability of the collection to achieve broad chemical space coverage and a broad range of distinct biological activities.[19,22,38] In this manner, screening libraries should ideally feature structurally *diverse* and *complex* small molecules in order to maximise the coverage of chemical space and therefore biological space to aid future investigations.

2.1.2 Sources of Complex and Diverse Libraries

To address these issues, there is a need for new screening libraries of *complex* and *diverse* small molecules. However, creating these is no simple task and the question then arises of how to populate a library. There are two main sources of small molecules: natural products and synthetic molecules.

2.1.2.1 Natural Products

Natural products are a rich source of complex and diverse bioactive molecules.[39,40] Since natural products result from coevolution with proteins, it has been suggested that their scaffolds are pre-encoded to bind to evolutionarily conserved protein binding sites and feature pre-validated biological activities (see Chapter 3).[41–43] It is not surprising that natural products, characterised by their enhanced molecular complexity, high degree of saturation and chirality content, provide highly specific biological

activities and have been widely used as templates for the successful development of drugs that mimic their actions.[40,44] Indeed, these molecules have played a pivotal role in medicine: 33% of small-molecule drugs developed between 1981 and 2014 are derived from natural products.[45]

Despite the fruitful prospects of natural products for the discovery of novel bioactive molecules, difficulties in isolation and purification of these compounds often lead to the screening of mixtures, which can hinder the identification of the active species.[46] Additionally, in order to screen natural products, kilograms of the natural source are often required to isolate only milligrams of the active compound.[47] As such there has been a sustained interest within the synthetic community in designing synthetic routes to construct these complex molecules.[48] Although this has aided our understanding of natural products these synthetic routes are often incredibly challenging and time-consuming, due to the sheer complexity of these molecules. Moreover, analogue generation can be challenging, something that is usually essential during hit optimisation.

2.1.2.2 Synthetic Small Molecules

A vital alternative to natural product libraries is the chemical synthesis of small molecules. The over-reliance on combinatorial-type and split-pool syntheses has resulted in vast screening libraries with an inherent lack of structural diversity (see above). These routes usually generate a limited range of molecular scaffolds, predominantly based on achiral or 'flat' architectures.[35] Thus, the development of novel strategies for the generation of screening libraries which feature the two attributes of structural complexity and structural diversity has become a major drive for synthetic chemists. In this manner, the development of novel synthetic strategies that take inspiration from the structural complexity of natural products, but feature improved synthetic tractability and modular aspects has attracted increasing interest. Within the last two decades several different methodologies have emerged, including biology-oriented synthesis (Chapter 3),[41,49–51] DNA-encoded/-templated libraries (Chapter 7)[52–55] and diversity-oriented synthesis (this chapter).[7,10,56]

The subject of structural diversity is a subjective concept. However, four common principles are commonly identified within the literature to guide the development of these libraries:

(1) Skeletal diversity: variation in the molecular architecture/ring structure of a molecule.
(2) Appendage diversity: variation in the structural moieties around a common scaffold.
(3) Stereochemical diversity: variation in the orientation of different elements from the scaffold.
(4) Functional group diversity: variation of the functional groups present which have the potential to interact with a given biological macromolecule.

Accordingly, the ideal library would, therefore, feature molecules that exhibit variation of all four types of diversity. Among the strategies developed, diversity-oriented synthesis has proven to be particularly fruitful for the discovery of novel chemical probes and hit compounds. In the remainder of this section we will discuss the underlying synthetic principles of DOS.

2.1.3 Synthetic Strategies for the Construction of Complex and Diverse Libraries: Diversity-oriented Synthesis

The diversity-oriented synthesis concept, 'DOS', was conceived and developed in the early 2000s.[57,58] DOS describes the deliberate and efficient divergent synthesis of complex small-molecule collections that interrogate large areas of chemical space. This contrasts with traditional retrosynthetic and combinatorial synthetic strategies, which commonly follow a linear process to generate focused libraries around a specific molecular framework occupying a defined area of chemical space. Instead, DOS adopts a forward synthetic analysis approach (Figure 2.2), whereby simple and common starting substrates are converted, through a wide range of complexity-generating transformations, into a collection of different scaffolds that incorporate all four principles of diversity. In particular, the incorporation of skeletal diversity is crucial for ensuring the functional diversity of a DOS library.[7,10]

Principal moment of inertia (PMI) plots provide a useful tool for the visual assessment of the chemical space coverage displayed by compound collection. In this approach, the normalised ratio of principal moments of inertia from the lowest energy conformation of each component of the library is calculated and plotted on a triangular graph. Each vertex of the graph represents one of three representative molecular shapes: rod-like, disk-like, or sphere-like. The broader the distribution within the three corners the compound collection has, the more diverse and complex it is (Figure 2.2). In comparison with combinatorial library synthesis, DOS approaches have been shown to deliver libraries with increased structural diversity.

2.1.3.1 Synthetic Principles in DOS

Broadly speaking, two approaches for the generation of scaffold diversity have been employed in DOS campaigns to date: reagent- and substrate-based DOS (Figure 2.3). In the *reagent-based approach,* branching pathways are carried out on a pluripotent substrate to yield different compounds with distinct molecular scaffolds. In this case, it is the choice of reagents and co-substrates that dictate the stereochemical and skeletal diversity within the final scaffolds. Alternatively, in the *substrate-based approach,* strategically positioned functional groups are reacted together (''paired'') so as to ''fold'' a

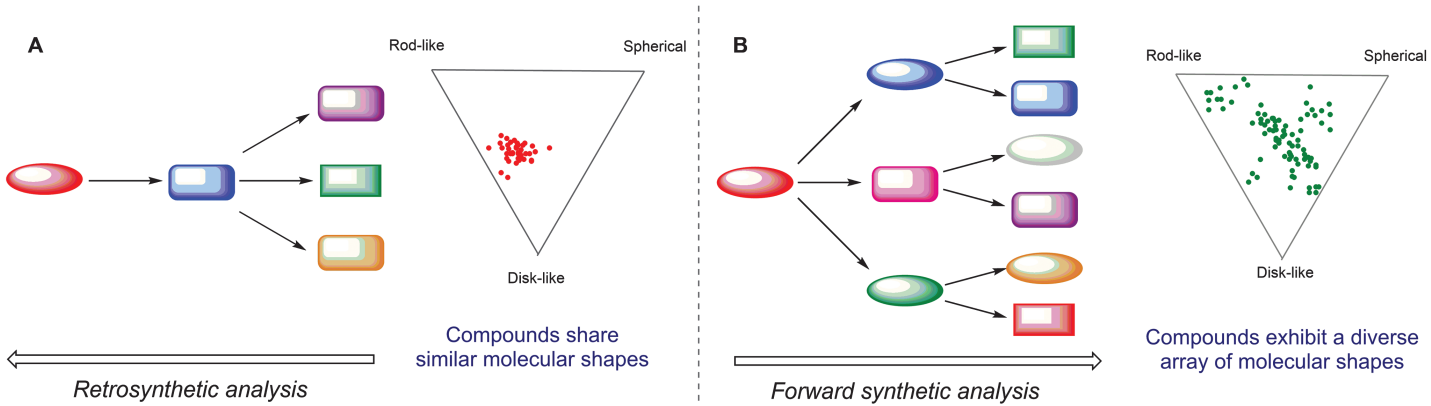


Figure 2.2 Comparison of Combinatorial (A) and DOS (B) strategies for library synthesis and the respective chemical space coverage of each as determined by principal moment of inertia plots.

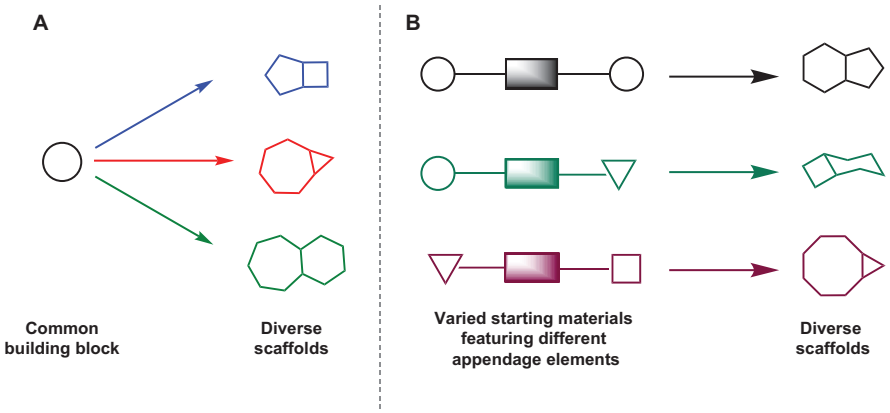


Figure 2.3 DOS strategies to synthesise distinct molecular scaffolds. (A) Reagent-based. (B) Substrate-based.

substrate into distinct molecular scaffolds. In this case, the pre-encoded functionalities within the substrate determine the skeletal composition in one of two ways: (A) starting from a densely-functionalised molecule such that different reaction conditions yield multiple scaffolds from one substrate; or (B) using different starting materials and common reaction conditions such that each starting material will furnish a product based around a different scaffold (Figure 2.3). For both approaches, the challenge is centred on the careful selection of the starting substrates.

These two approaches are not orthogonal and most modern DOS strategies incorporate aspects of both.[6,10,33] Other features of diversity (appendage, functional group and stereochemical) can be introduced into the compound libraries through variation in the starting materials and/or reagents used.[33,59]

Independent of the approach taken, the build–couple–pair (B–C–P) algorithm (Figure 2.4) is a powerful method in DOS.[7,60] This three-phase strategy begins with the *build phase*, involving the synthesis of building blocks containing orthogonal and ideally some chiral functionalities. These building blocks can then be linked together or coupled with other substrates intermolecularly to produce complex and densely functionalised molecules. This step is termed the *couple phase* and provides the basis for the introduction of stereochemical diversity. Finally, the *pair phase* involves the intramolecular coupling of different moieties within the intermediates to produce rigidified scaffolds. This stage results in the skeletal diversity of the library.

In the remainder of this chapter, we will describe several state-of-the-art examples of DOS and how these libraries have been applied to aid the identification of novel biologically active molecules for use as both hits (hit identification) and chemical probes (for chemical genetics or target identification/validation).

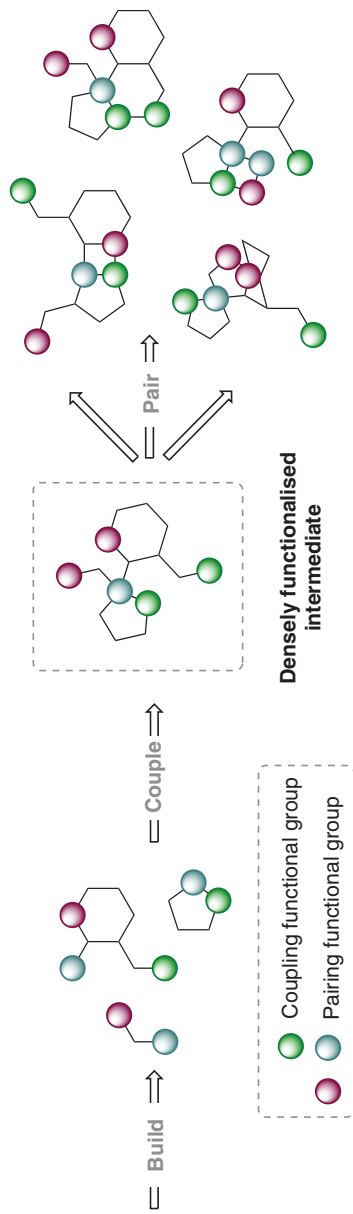


Figure 2.4 Build–Couple–Pair in DOS.

2.2 Application of Diversity-oriented Synthesis for the Identification of Small Molecule Modulators

2.2.1 Structural Diversity and Phenotypic Screening

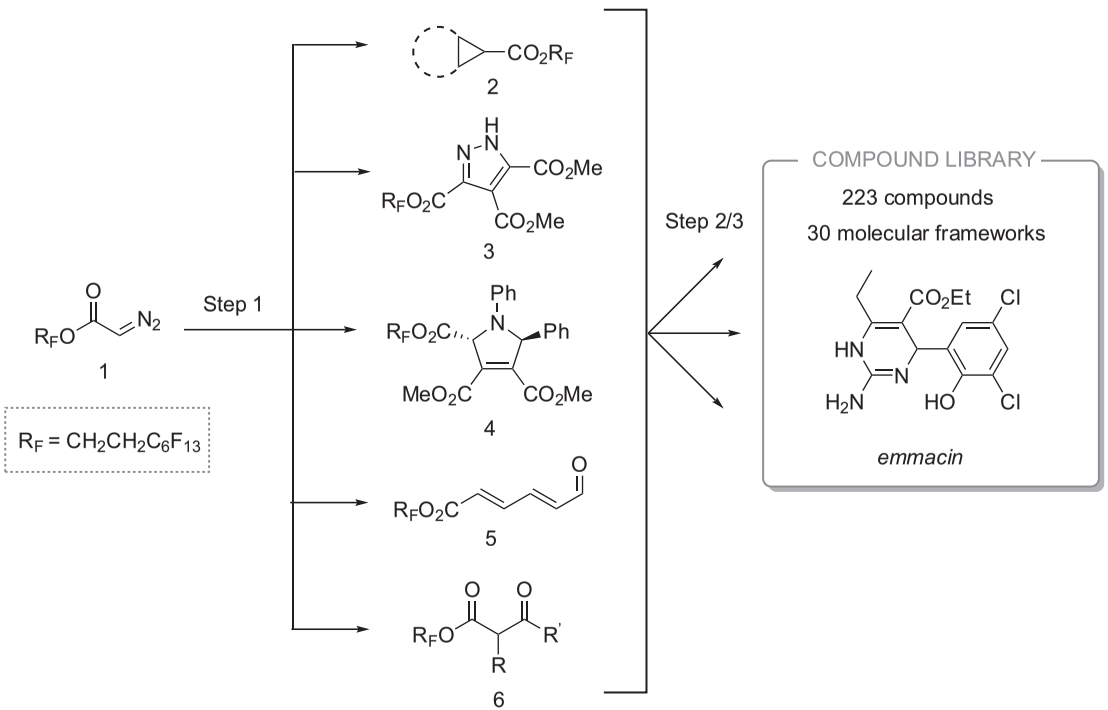
The identification of small-molecule modulators through the screening of compound collections comprises different *in vitro* or *in vivo* strategies. Among these is phenotypic screening: testing small molecules and observation of phenotypic change in a biological system, which has been successfully used in the pharmaceutical industry for many years.[61–63] Indeed, 56% of ''first-in-class'' medicines approved by the FDA between 1999 and 2008 have been identified in this manner.[64] In a phenotypic screening approach the molecular mechanism of action and the biological target protein can remain unknown, however, predictive studies can be undertaken by comparing multi-parametric phenotypic profiles or ''phenoprints'' of the new hits with reference compounds with an established mechanism of action.

Phenotypic screening involves testing small-molecule compound collections in cellular or *in vivo* models directly without any previous target knowledge. In these unbiased processes, where the precise nature of the biological target is unknown, the selection criteria for the ''right'' small molecule within chemical space are dramatically complicated due to the lack of any structural information.[65] The question then arises, how can the ideal small-molecule collection be selected for these assays without any structural guidelines? As previously discussed, small-molecule collections with a high degree of structural diversity and complexity are predicted to display a higher hit-rate and broader scope of biological activity.[22,66]

There are many examples in the literature where the combination of structurally diverse and complex screening collections, often produced from DOS, with phenotypic screening has resulted in the identification of new biologically active molecules.[67–72] The selected examples discussed later in this chapter (Section 2.2.1.1) are focused on studies that have provided access to new hits related to challenging therapeutic areas, such as antibiotic resistance and cancer.

2.2.1.1 Efficient Identification of Novel Antibiotics from Diverse Collections

In the context of DOS, the power of this strategy as a tool to identify new and much-needed antibacterials through phenotypic screening has been validated.[59,73] One of the first case studies was reported by Wyatt *et al.* in 2006 and involved the exploitation of the different reactivity of a two-carbon fluorous-tagged diazoacetate **1** following a reagent-based DOS strategy (Scheme 2.1).[74] In the first set of branching reactions, three key synthetic transformations were applied: (i) three-membered ring formation, (ii) 1,3-dipolar cycloadditions



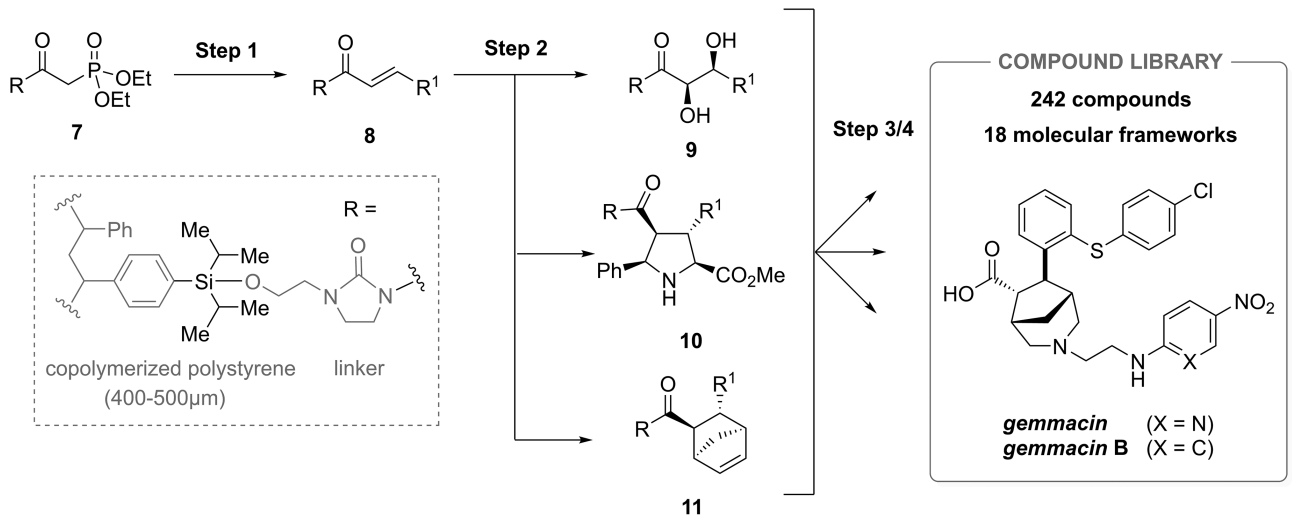
Scheme 2.1 New sub-structural class of bacterial-selective DHFR inhibitor *emmacin* discovered from a DOS library of 223 compounds based on 30 molecular frameworks.

and (iii) α-deprotonation followed by quenching with an electrophile and carbenoid formation. These structurally diverse intermediates **2**–**6** were subjected to further complexity-generating reactions to diversify the molecular frameworks and thus increase the skeletal diversity of the library (Scheme 2.1, Step 2/3). As a result, a collection of 223 small molecules based around 30 distinct molecular skeletons was efficiently generated in no more than four synthetic steps from the common diazoacetate unit.

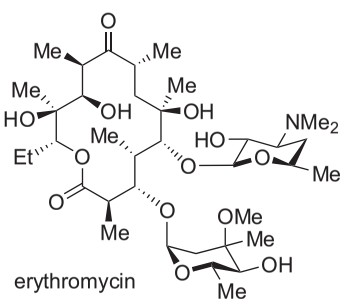
Preliminary phenotypic screens revealed that 64 structurally diverse compounds of the DOS library were able to modulate the growth of strains of methicillin-resistant *Staphylococcus aureus* (MRSA).[67] Further investigation revealed that the compound named *emmacin* (shown in Scheme 2.1) was a potent inhibitor of methicillin-susceptible strains of *S. aureus* (MSSA) and two UK epidemic strains of MRSA (EMRSA-15 and EMRSA-16[75,76]). In addition, this compound exhibited no cytotoxic properties in a variety of mammalian-surrogate systems. Target identification revealed that *emmacin* is a prokaryote-selective, uncompetitive and reversible inhibitor of dihydrofolate reductase of EMRSA-16 ($DfrB_{EMRSA16}$).[77,78]

Following a similar strategy, Thomas *et al.*[68] successfully combined the synthesis of a DOS library and further phenotypic screening to identify a new antibiotic with similar potency to those used clinically. The generation of the small-molecule library was achieved starting from a solid-supported phosphonate **7** using a reagent-based DOS strategy (Scheme 2.2). In this case, the reacted phosphonate functionality allowed the *E*-selective formation of highly functionalized α,β-unsaturated acyl imidazolidinones **8** (Step 1) through a Horner–Wadsworth–Emmons reaction with different aldehydes. Following a divergent scheme, **8** was reacted in three catalytic and enantioselective pathways: (i) [2 + 3] cycloaddition; (ii) dihydroxylation, and (iii) [4 + 2] cycloaddition to yield a small set of molecules based on three different molecular frameworks (Step 2). The next steps (Step 3/4) of the synthetic strategy involved a series of branching reactions to further diversify these key branch-point substrates **9**–**11**. In this manner, a collection of 242 natural-product-like small molecules based on 18 molecular scaffolds, with high levels of skeletal diversity, was synthesized.

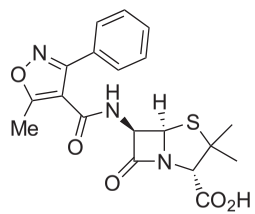
Similarly to the previous case study, the resulting DOS collection was screened to study the effect on the growth of three strains of *Staphylococcus aureus* (MSSA, EMRSA 15 and EMRSA 16). Three structurally novel compounds inhibited bacterial growth, but one of them, named *gemmacin*, showed good activity against EMRSA 15 and 16 (Figure 2.5). Stereoselective synthesis and further SAR studies permitted the identification of (−)-*gemmacin*-B, which demonstrated higher levels of antibacterial activity against EMRSA 16.[69] Target identification was achieved after the observation that *gemmacin* generated reactive oxygen species, which indicated that the compound may act as a cell-membrane disruptor. This supposition was confirmed using a membrane disruptor assay where *gemmacin* acted as a selective disruptor of bacterial cell membranes.



Scheme 2.2 Novel antibacterial *gemmacin* discovered from a DOS library of 242 compounds based on 18 molecular frameworks.



erythromycin



oxacilin

	MIC_{50} (μg ml^{-1})		
	MSSA	EMRSA 15	EMRSA 16
emmacin	**2**	**9**	**9**
(-)-*gemmacin*	**2**	**8**	**16**
(+/-)-*gemmacin* B	**–**	**8**	**8**
erythromycin	0.5	>64	>64
oxacilin	0.5	>32	>32

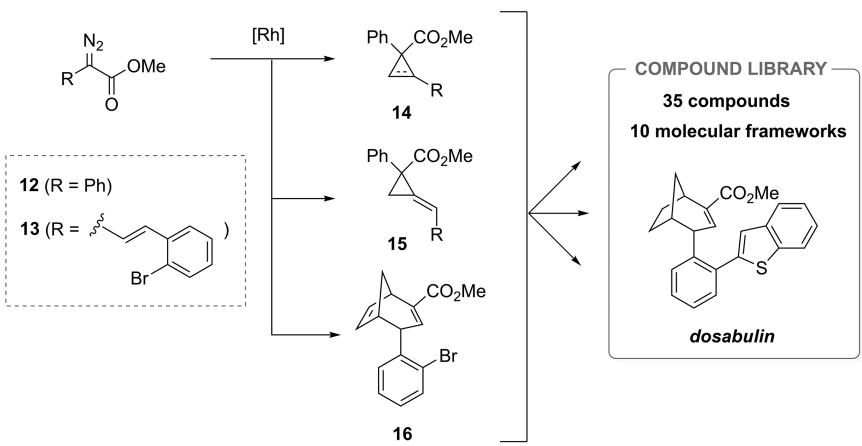
Figure 2.5 MIC_{50} values for *emmacin, gemmancin* and *gemmacin B* against MSSA, EMRSA 15 and EMRSA 16 are shown. For comparison, the MIC_{50} values for erythromycin and oxacillin are also shown. (MIC_{50} = minimum inhibitory concentration required to inhibit the growth of 50% of organisms).

2.2.1.2 DOS as a Tool for the Identification of New Anticancer Small Molecules

Among the many existing techniques developed to understand cancer genetics, the identification of small molecules capable of modifying cancer phenotypes is particularly effective for feeding early drug discovery programs.[79–82] In particular, the screening of structurally diverse compound collections has been demonstrated to be an excellent strategy to identify new modulators of cancer cells.[83] In this context, some examples are discussed below where the combination of DOS collections and phenotypic screening using cancer cell lines has delivered new anticancer hit compounds.

In one of these case studies, rhodium carbenoid chemistry was used as a key step for the reagent-based DOS of a structurally diverse small-molecule collection starting from phenyldiazo ester compounds.[84] Following a divergent scheme, the α-diazo ester **12** was reacted with terminal alkynes, alkenes and allenes *via* rhodium-catalysed cyclopropanation reaction to give rise to a small set of different substituted three-membered rings (Scheme 2.3, **14–15**). Encouraged by the scope of this rhodium(II)-catalysed reaction, the styryl diazo ester derivative **13** was synthesized and treated under similar reaction conditions. Accordingly, after a cyclopropanation-Cope reaction with cyclopentadiene, intermediate **16** was stereoselectively generated. These three restricted and highly functionalized scaffolds **14–16** were considered privileged starting points for further diversification through a wide range of chemical transformations. This strategy delivered a library of 35 three-dimensional and structurally diverse compounds with high sp^3 content and broad coverage of biologically relevant chemical space.

Screening of the resulting compound collection for antimitotic activity in human U2OS osteosarcoma cells[85] was performed, and two compounds

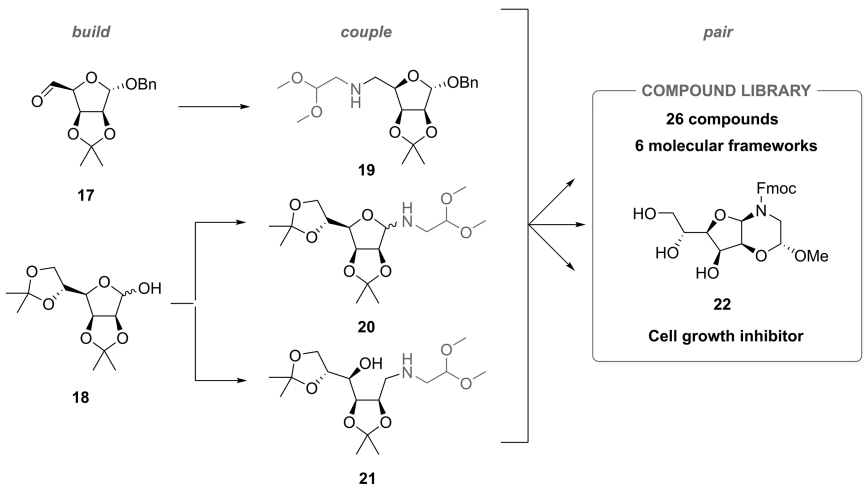


Scheme 2.3 Rhodium carbenoid methodology used for DOS leading to identification of *dosabulin* as an inhibitor of tubulin polymerization.

from the DOS library were observed to cause mitotic arrest. Further chemical modifications carried out on these initial active molecules resulted in the identification of the (*S*)-enantiomer of *dosabulin* as a more potent inhibitor of mitosis [concentration giving 50% effect (EC_{50}) 1.2 μM] causing cell death by apoptosis. Target identification was investigated by competition studies observing the displacement of colchicine or vinblastine from tubulin by (*S*)-*dosabulin* using confocal microscopy. These experiments indicated that *dosabulin* was not binding to the vinblastine site on tubulin, but that the binding site was in the vicinity of, or allosteric to, the colchicine site.

Recently, the use of carbohydrates has attracted interest in the field of DOS[86–89] due to their immense potential for generating stereochemical and structural diversity as well as their biomedical applications.[90–92] In notable recent work in this area, Trabocchi and colleagues[93] described how building blocks derived from D-mannose and glycine could be used to create the structural complexity following a substrate-based build–couple/pair strategy (Scheme 2.4). The protected sugar derivatives **17** and **18**, resulting from D-mannose modifications (*build* phase), were combined in the *couple* step with amino-acetaldehyde dimethyl acetal through three reducing-based reactions. The presence of polyhydroxylated species and the protected carbonyl function coming from the amino acid derivative in intermediates **19**–**21** allowed intramolecular *trans*-acetalisations as the key pairing strategy. Thus, after the *pair* phase, a collection of 26 polyhydroxylated nitrogen-containing scaffolds based on six molecular frameworks with high levels of structural diversity (demonstrated by computational analysis) was generated.

Further investigations were focused on testing the resulting DOS library in a phenotypic screen using MDA-MB-231 cell lines for the identification



Scheme 2.4 Cell growth inhibitor **22** identified from a DOS library of polyhydroxylated nitrogen-containing scaffolds.

of modulators of the breast carcinoma cell cycle mechanism.[70] These investigations were based on the preliminary evidence regarding the capability of iminosugars to inhibit the growth of breast cancer cells.[94,95] The MDA-MB-231 cell line is a simple model system for the study of triple-negative breast cancer, which shows a major tendency towards early metastasis, not responding to hormonal chemotherapy and accounts for 15% of all breast carcinomas. In this experiment, after 48 hours of incubating members of the DOS library with MDA-MB-231 cells, compound **22** showed the best range of inhibition of cell proliferation (exhibiting 40% inhibition). Despite subsequent synthesis of analogues of **22**, none of the newly synthesised compounds displayed better inhibition values, validating compound **22** as a new anticancer modulator based on a polyhydroxylated scaffold. Further studies towards the identification of the molecular mechanism of action to identify the target related with the phenotypic effect are ongoing.

2.2.2 The Role of DOS in Target Validation through the Discovery of New Chemical Probes

A complementary method for target validation in a drug discovery context is the use of small-molecule chemical probes, which can modulate biological systems in order to predict therapeutic potential.[96] Importantly, as well as small-molecule inhibitors this includes the discovery of probes that can activate signalling pathways, since this can shed light on the workings of complex biological systems.[97] In this manner, the discovery of novel, high-quality small molecules capable of inducing functional pharmacology and proving phenotype perturbation can play a vital role in facilitating target validation. DOS represents an ideal synthetic strategy to deliver novel structurally diverse molecules for use as chemical probes, since the limiting factor in the application of small molecules for this process is often the availability of appropriate chemical modulators. Since DOS aims to populate new areas of chemical space, this provides the potential to widen our 'chemical probe toolbox', enabling us to modulate and interrogate the functional behaviour of challenging, poorly characterized and even novel biological systems to seed new drug-discovery programs.

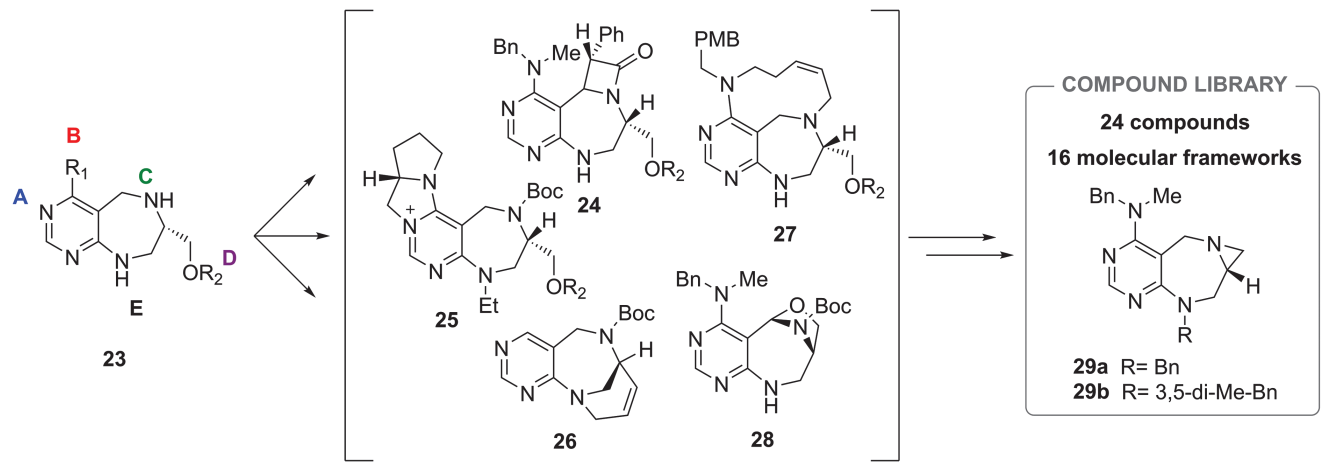
2.2.2.1 DOS Yields a Novel PPI Modulator

Protein–protein interactions (PPIs) play a critical regulatory role within a range of biological functions within the body and have been implicated in a vast range of disease states, including cancer.[98] Despite the fruitful prospects of PPI inhibition for drug discovery, the challenging nature of inhibiting these interactions using 'traditional' small-molecule screening collections has led to the perception of PPIs as 'undruggable'. Recent work by Kim *et al.*, however, has demonstrated the application of DOS in the

identification of a modulator of a key PPI interaction implicated in the amino-acid-dependent activation of the mechanistic target of rapamycin complex 1 (mTORC1), a regulator of cellular growth, proliferation and autophagy with oncogenic implications.[99]

In this work, the authors utilised a privileged-DOS (pDOS) strategy[100] to generate 'biological navigators' using pyrimidodiazepine and pyrimidine cores, considered 'privileged' scaffolds due to their presence within various bioactive molecules and marketed drugs. Additionally, they incorporated a diazepine motif within these fused core scaffolds, to increase the 3D character of the library *via* higher sp^3 content and conformational flexibility. Following a B–C–P algorithm five reactive sites were installed upon the pyrimidodiazepine core, that by utilizing a reagent-based approach were selectively reacted in a pairwise fashion. This produced several tri-cyclic and tetra-cyclic scaffolds in a synthetically efficient manner from **23**. Five different reactive site combinations were then employed (Scheme 2.5)—A–B (**25**), B–C (**27**), C-pair (**24**), C–D (**28**) and D–E (**26**)—to generate 24 compounds representing 16 distinct frameworks. A variety of functional group pairings was employed including intramolecular nucleophilic substitution, ring-closing metathesis (RCM) of pre-installed unsaturated moieties and exploitation of the both the nucleophilic and electrophilic nature of an imine at C. Additionally, the use of a rhodium-catalysed [2 + 2] cycloaddition (to give **24**), a [2,3]-sigmatropic ring expansion (not shown) and RCM (to give (**26**) allowed the formation of challenging ring systems such as a β-lactam, a 6,10-benzoxazecine and a bridge-head [4,3,1] structure, respectively. A comprehensive collection of 3–10 membered carbocycles and hetereocycles was installed upon the key pyrimidine core. The results of chemoinformatic analysis indicated that the resulting library displayed a broad shape distribution, similar to that of 71 bioactive natural products, but an improved distribution compared with 15 pyrimidine-containing FDA approved drugs.[99]

The leucyl-tRNA synthetase (LRS)–Ras-related GTP-binding protein D (RagD) PPI plays an important role in the amino-acid-dependent activation of mTORC1 *via* leucine sensing and signalling to mTORC1.[101] Initial screening for inhibition of this PPI was conducted using ELISA-based HTS using LRS and glutathione-*S*-transferase tagged RagD, leading to the identification of **29a** and **29b** as dose-dependent inhibitors. Subsequent biological experiments focused on target validation of LRS–RagD PPI inhibition through investigations into the effects on mTORC1 activity and cell proliferation. Western blot experiments revealed that **29a** and **29b** suppressed the phosphorylation of p70 ribosomal protein S6 kinase 1 (S6K1)—a known kinase substrate for mTORC1—whilst **29b** proved to selectively downregulate two further mTORC1 substrates but not mTORC2 and 5′ AMP-activated protein kinase (AMPK) substrates in cancer cell lines. This indicated selectivity for this signalling pathway; however, notable differences in these phosphorylations were observed compared with the known mTORC1 inhibitor rapamycin, which the authors suggested was a result of



Scheme 2.5 The pDOS strategy utilised by Kim *et al.*[99] using pyrimidodiazepine and pyrimidine core scaffolds. Five reactive sites were added to the cores, allowing functional group pairings to yield 16 distinct compounds (selected examples shown)—two of which proved to inhibit the LRS–RagD PPI.

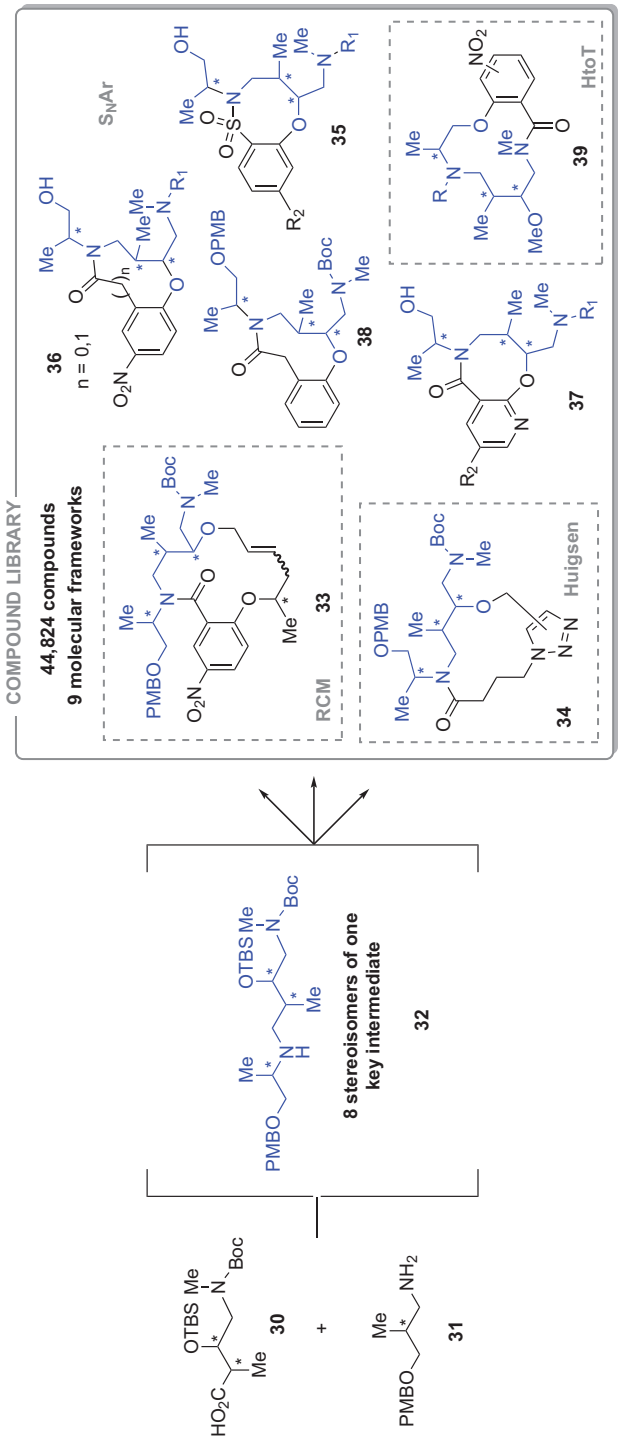
29a acting *via* an alternative mode of action. Finally western blot, live-cell imaging and cell proliferation assays were used to confirm that **29a** is capable of inducing autophagy and a reduction in proliferation, even in the presence of leucine, as a result of inhibition of the LRS–RagD PPI.[99]

2.2.2.2 *A Single DOS Library Yields Multiple Chemical Tools for Multiple Biological Systems*

Recent efforts by researchers at the Broad Institute have begun to investigate the application of a single DOS library for the identification of numerous chemical probes. Strikingly, this library has proven to be particularly fruitful for discovery, with several publications detailing the identification of biological hits from this collection in the seven years since the library's publication, including antiparasitic molecules[102] and probes which modulate autophagy.

A seminal report by Marcaurelle *et al.* in 2010[103] first described the methodology used to construct this complex library, describing an elegant strategy for the construction of a collection of medium-to-large sized rings with diverse stereochemical and skeletal features. Following a B–C–P algorithm, the build phase consisted of stereodivergent *syn*- and *anti*-aldol reactions, producing all four stereoisomers of a Boc-protected γ-amino acid (**30**), as well as the separate production of both stereoisomers of a protected alaninol derivative (**31,** Scheme 2.6). Joining of the two blocks *via* amide formation followed by reduction yielded the intermediate **32**. Importantly, the modular nature of the synthetic route facilitated access to all eight diastereomers of **32** and therefore of each resultant scaffold, enabling future stereochemical structure–activity relationship (SSAR) data to be generated.

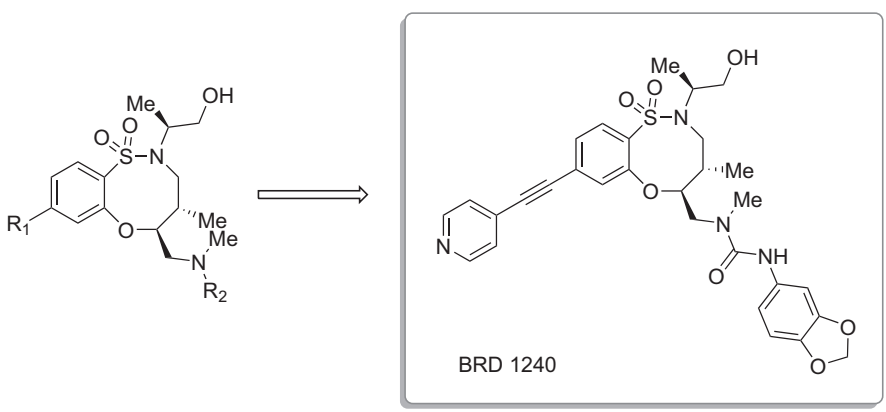
Finally, a reagent-based strategy was applied in the pair phase, using nucleophilic aromatic substitution (S_NAr), Huisgen triazole formation and RCM chemistries to construct 24 8–14-membered-ring scaffolds stemming from five distinct and complex frameworks. Three further papers[104–106] from the Marcaurelle laboratory described additional extensions of this methodology, using the same stereochemically-rich intermediate and four alternative pairing partners in S_NAr and head to tail (H to T) cyclisations to generate a further five scaffolds using **32**. The innovative incorporation of suitable branching points, such as aromatic halides, aliphatic and aromatic amino functionalities (*via* nitro reduction), across the whole library resulted in the presence of multiple potential exit vectors within each scaffold and allowed facile library expansion. Importantly, the robust nature of the chemistry allowed multi-gram quantities of key intermediates and final scaffolds to be isolated. A combination of these factors enabled significant expansion of the library *via* solid-phase synthesis using combinatorial-type modifications to yield over 44 000 compounds as stereoisomers and analogues of the initial nine frameworks in a matrix-like fashion.



Scheme 2.6 The synthetic strategy employed by Macaurelle *et al.*,[103] Gerard *et al.*,[104] Fitzgerald *et al.*[105] and Chou *et al.*[106] for the formation of nine molecular frameworks, with subsequent library expansion yielding over 44 000 compounds.

2.2.2.2.1 Identification of a Novel Probe for VATPase Function. Aldrich *et al.*[107] investigated the application of the Marcaurelle DOS library for the identification of chemical probes for autophagy. This led to the identification of inhibitors of lysosomal acidification *via* high-content screening followed by target identification studies. BRD1240 (Scheme 2.7) showed low micromolar activity in an initial phenotypic screen detecting modulation of the number of autophagosomes within HeLa cells, with key differences in activity between all possible stereoisomers of the scaffold being noted; only two of the eight proved to be active. Further experiments investigated autophagosomal turnover, where it was found that BRD1240 and Bafilomycin A1 (a known vacuolar-type H^+–ATPase (V-ATPase) inhibitor) inhibited the turnover, indicating a possible mechanism of action. Additionally, BRD1240 was found to modulate lysosomal function through inhibiting lysosomal acidification and protease activity. Structure–activity relationship (SAR) data indicated that the 4-position nitrogen within the pyridine was vital for activity, whilst an electron-rich urea moiety was also required for optimal activity. Using cancer cell line sensitivity profiling and a comparison of BRD1240 and BafA1, the authors concluded that BRD1240 perturbs V-ATPase function. This hypothesis was validated by the suppression of V-ATPase function by BRD1240 in biochemical assays. Surprisingly, the kinetics of these biochemical assays indicated that BRD1240 may act *via* a novel mode of action compared with the known BafA1. Thus, BRD1240 serves as a novel probe for the investigation of lysosomal acidification *via* V-ATPase modulation.

2.2.2.2.2 Identification of a Small-molecule Modulator of Autophagy Independent of mTOR and Lysosomal Function. Compounds in the Marcaurelle library were also shown to yield novel hits that modulated autophagy; however, importantly, the results of studies into their



Scheme 2.7 The development of a V-ATPase probe. From initial hits, BRD1240 was chosen as the lead compound for further investigation.

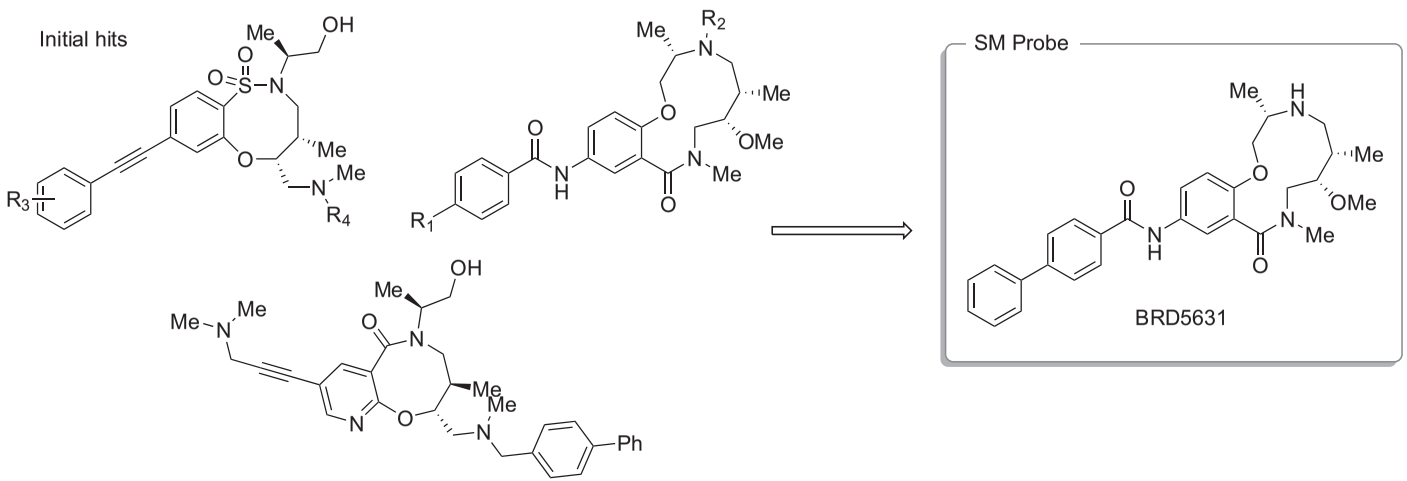
mechanism of action indicated that these compounds in fact promoted autophagy and did not perturb mTOR signalling or lysosomal function. Thus, these small molecules can serve as orthogonal probes for autophagic processes. Kuo *et al.*[108] screened 59 541 DOS compounds for modulation of autophagosome number in HeLa cells. Five hits, which were shown to increase autophagesome number, were selected for further investigations where they were found to promote autophagy, whilst not disrupting mTOR signalling pathways or lysosomal function determined by western blot and protein phosphorylation experiments. Furthermore, the lead hit BRD5631 (Scheme 2.8) was also shown to modulate disease-associated phenotypes, including protein aggregation, cell survival, bacterial replication and inflammatory cytokine production as a result of autophagy activation. Whilst investigations into the precise mode of action and target of BRD5631 are ongoing, this molecule will continue to be useful for illuminating the biological relevance and therapeutic potential of promoting autophagy.

2.2.2.3 Application of DOS for Discovery of Novel Antimalarial Compounds

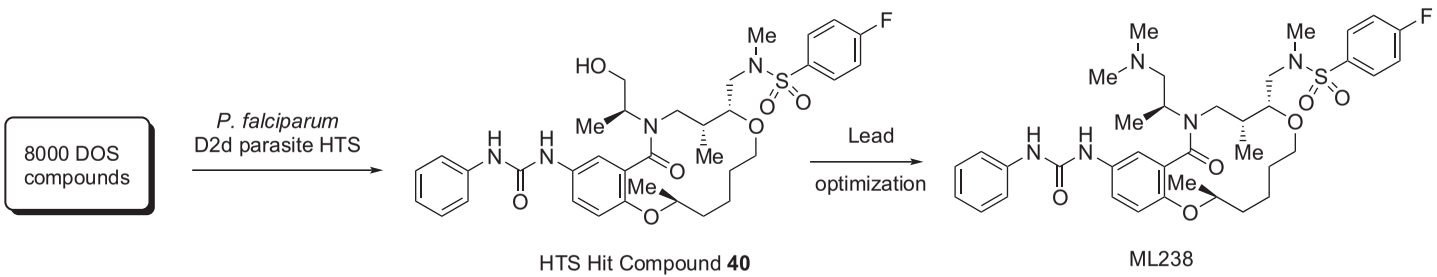
As part of a pilot investigation into the application of DOS molecules as antimalarials, a phenotypic screen of a DOS-derived library containing a subset of compounds from the Marcaurelle library against multidrug-resistant D2d *Plasmodium falciparum* asexual blood-stage parasites was undertaken.[109] Of the 8000 molecules screened, 560 displayed over 90% D2d growth inhibition. From the initial 560 promising results, 26 molecules exhibited particularly potent inhibition (over 50% at 280 nM), and 20 of these were macrolactam scaffolds synthesized in the Marcaurelle DOS campaign using RCM. Ultimately, compound **40** was identified as the most active in the screen (Scheme 2.9).

Due to the stereochemically comprehensive design and construction of the original DOS library, the stereochemical SAR (SSAR) of all 16 stereoisomers of **40** was easily determined, with the (2*S*,5*R*,6*R*,12*S*) stereoisomer being found to be most potent. Additional SAR of the peripheral substituents identified that switching the hydroxyl to a dimethylamino group increased solubility. These initial efforts led to the designation of ML238 as an antimalarial probe, which inhibits two strains of *P. falciparum* (D2d and 3D7) and exhibits low off-target cytotoxicity, as well as high stability in human plasma.

Taking ML238 as a lead compound, further medicinal chemistry optimization was undertaken to remove unfavourable properties such as potential cardiotoxicity [Human ether-a-go-go-related gene (hERG) product binding], poor stability in microsomes of a mouse model and low phosphate buffered saline (PBS) solubility, factors which could lead to low bioavailability or very high doses in any future therapeutic application. The short,



Scheme 2.8 From five prioritised hit compounds, BRD5631 was chosen as the lead compound for further investigation.



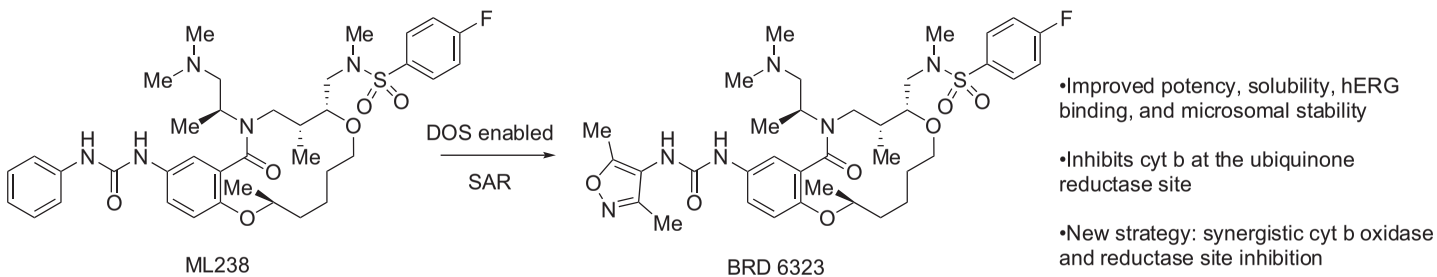
Scheme 2.9 HTS of DOS library yields novel antimalarial probe ML238.

modular synthesis of this class of macrolactams enabled further facile optimization, due to the flexibility of the DOS strategy to modify and evaluate nearly every position of the core ring, including ring size and effects of heteroatom exchange, simply by exchanging building blocks used at the outset of the route.[110] The application of appendage and skeletal SAR ultimately yielded an improved analogue, BRD6323, which was studied *in vivo* and further investigated for its mechanism of action (Scheme 2.10).

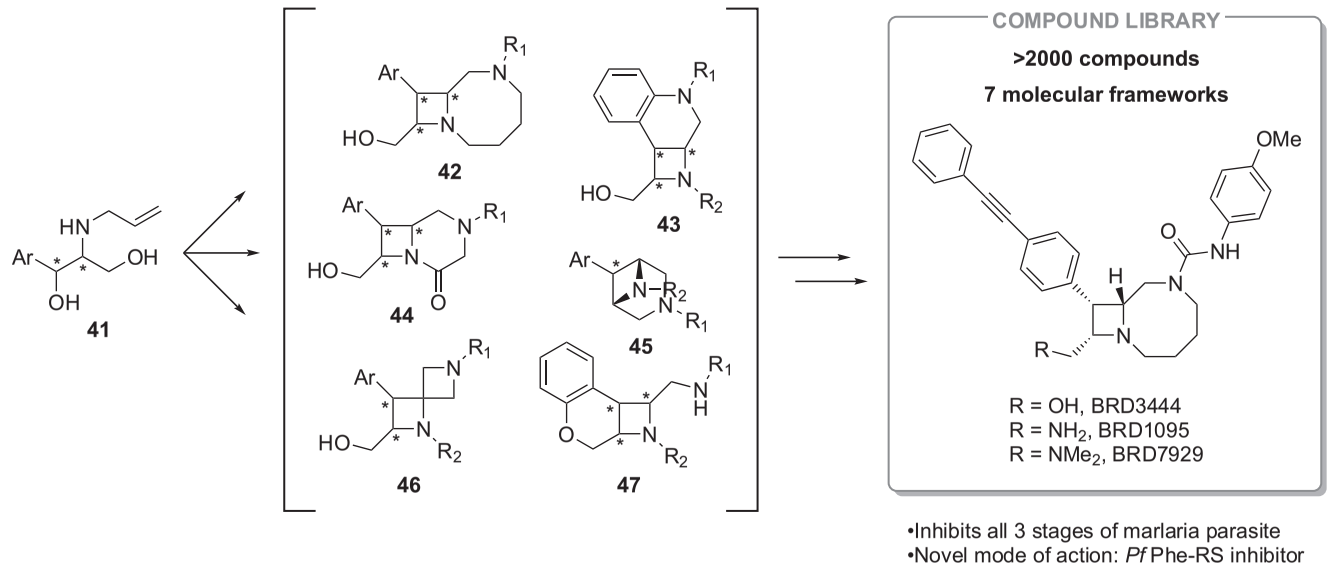
In initial experiments ML238 and BRD6323 proved effective against parasite isolates with drug-resistant genotypes, indicating a different mode of action to current treatments. To determine the cellular target of these molecules, resistance selection and whole-genome sequencing were performed on resistant strains, ultimately revealing mutations occurring solely in the *cytb* locus over a number of BRD6323-resistant lines.[111] It was further determined that BRD6323 and ML238 inhibit the ubiquinone reductase site (Q_i) of cytochrome b in *P. falciparum*. When used in concert with known cytochrome b ubiquinol oxidase site (Q_o) inhibitors like atovaquone, a synergistic effect was observed. These observations illuminated a new potential combination therapy strategy due to dual inhibition of an enzyme at two sites, which may prove useful in the fight against malarial infection in the future.

2.2.2.3.1 DOS for the Discovery of Azetidine-based Antimalarials and Identification of Novel Cellular Targets. Encouraged by the results of the pilot-scale HTS of a DOS library yielding new leads with important biological activity and underexplored modes of action, an expanded antimalarial screen of the complete Broad Institute DOS library, comprising nearly 100 000 molecules, was undertaken.[112] Included in this library were a collection of highly substituted azetidine molecules. The azetidine library was constructed based on an ephedrine-like scaffold.[113] All possible stereoisomers of an aryl-containing *N*-allyl amino diol were constructed (Scheme 2.11). Subsequent transformations led to the diastereoselective formation of azetidines, which were elaborated *via* a reagent-based approach to generate a series of bicyclic compounds. This methodology yielded several distinct scaffold classes of bicyclic azetidine molecules, including bridged bicyclic, monoketopiperazine, azocane and azaspirocyclic scaffolds, representing seven molecular scaffolds and over 2000 compounds after a combinatorial-type solid-phase synthesis effort. The features of this library included comprehensive stereochemical information for each scaffold and, as a result, high three-dimensionality of the molecules, as well as high scaffold diversity of the library.

Similar to the pilot-scale study, these and other molecules were screened against *P. falciparum* D2d cells in a phenotypic blood-stage parasite growth-inhibition assay. In this study, counter-screens against drug-resistant clones and parasite isolates were performed to determine if the compounds acted according to a known mechanism of action, in order to prioritize compounds displaying novel modes of action. Additional assays against



Scheme 2.10 DOS-enabled SAR yields improved antimalarial probe BRD6323.



Scheme 2.11 DOS of azetidine compounds leads to discovery of a novel multistage inhibitor of malaria.

liver-stage and transmission-stage parasites identified four compound series that exhibited multistage inhibition. Of these four series, three represented new scaffolds against known targets, and one (BRD3444, Scheme 2.11) was found to inhibit *P. falciparum* through an unknown target. Medicinal chemistry optimization of BRD3444 using the modular synthetic approach developed in the original DOS led to analogues with improved solubility, bioavailability and potency, resulting in molecules BRD7929, with a dimethylamino group replacing a hydroxymethyl group in the original structure, and BRD1095, with a methylamino group replacement.

Using these molecules, resistance selection and whole-genome sequencing of the resistant clones were performed, and analysis predicted the inhibition of cytosolic phenylalanine-tRNA synthetase (*Pf* Phe-RS). This target was confirmed by experiments with purified recombinant protein. Inhibition of this target results in the elimination of asexual blood-, liver- and transmission-stage parasites, a unique effect among antimalarial drugs, with demonstrated *in vivo* efficacy in mouse malaria models. When applied to mouse models, it was demonstrated that treatment with BRD7929 prevents transmission, ensures prophylaxis and provides single-dose cures for malaria infections.

2.3 Conclusions and Outlook

The case studies presented herein demonstrate the ability of DOS-derived molecules to accelerate discovery by generating hits and facilitating derivatisation, to serve as probes to enable new biological insights and to generate lead compounds to address some of the most pressing problems in medicine. The molecules generated using DOS possess complexity and diversity that is reminiscent of natural products whilst maintaining the synthetic tractability of simpler, drug-like compounds. These features have enabled the rapid construction of diverse libraries which, when combined with modern methods for high-throughput screening, present tremendous opportunities for new discoveries in chemistry, biology and medicine. Future efforts towards these goals will be aided by developments in synthetic methodology and strategy, enabling access to new areas of chemical space and increases in synthetic efficiency. As we have shown, DOS libraries access vast, often untapped, potential for finding novel bioactive molecules, often beyond the uses for which the molecules were originally envisioned. Many of the examples discussed herein demonstrate the importance of follow-up studies beyond initial chemical probe discovery, and can lead to novel areas of biological inquiry or reveal new targets or modes of action. Therefore, the continued screening and biological evaluation of existing and new DOS libraries will considerably increase our biological understanding and therapeutic application of these molecules.

Acknowledgements

The Spring lab acknowledges support from AstraZeneca, the European Union, the Engineering and Physical Sciences Research Council (EP/P020291/1) and the Royal Society.

References

1. S. L. Schreiber, J. D. Kotz, M. Li, J. Aubé, C. P. Austin, J. C. Reed, H. Rosen, E. L. White, L. A. Sklar, C. W. Lindsley, B. R. Alexander, J. A. Bittker, P. A. Clemons, A. De Souza, M. A. Foley, M. Palmer, A. F. Shamji, M. J. Wawer, O. McManus, M. Wu, B. Zou, H. Yu, J. E. Golden, F. J. Schoenen, A. Simeonov, A. Jadhav, M. R. Jackson, A. B. Pinkerton, T. D. Y. Chung, P. R. Griffin, B. F. Cravatt, P. S. Hodder, W. R. Roush, E. Roberts, D. H. Chung, C. B. Jonsson, J. W. Noah, W. E. Severson, S. Ananthan, B. Edwards, T. I. Oprea, P. J. Conn, C. R. Hopkins, M. R. Wood, S. R. Stauffer and K. A. Emmitte, *Cell*, 2015, **161**, 1252–1265.
2. C. J. O' Connor, L. Laraia and D. R. Spring, *Chem. Soc. Rev.*, 2011, **40**, 4332.
3. S. Fox, S. Farr-Jones, L. Sopchak, A. Boggs, H. W. Nicely, R. Khoury and M. Biros, *J. Biomol. Screening*, 2006, **11**, 864–869.
4. K. H. Bleicher, H.-J. Böhm, K. Müller and A. I. Alanine, *Nat. Rev. Drug Discovery*, 2003, **2**, 369–378.
5. D. J. Payne, M. N. Gwynn, D. J. Holmes and D. L. Pompliano, *Nat. Rev. Drug Discovery*, 2007, **6**, 29–40.
6. K. M. G. O'connell, W. R. J. D. Galloway and D. R. Spring, in *Diversity-Oriented Synthesis:Basics and Applications in Organic Synthesis, Drug Discovery and Chemical Biology*, ed. A. Trabocchi, John Wiley & Sons, Inc, Hoboken, NJ, 2013, pp. 253–287.
7. W. R. J. D. Galloway, J. E. Stokes and D. Spring, in *Small Molecule Medicinal Chemistry: Strategies and Technologies*, ed. W. Czechtizky and P. Hamley, John Wiley & Sons, Inc., 2016, pp. 79–101.
8. W. H. B. Sauer and M. K. Schwarz, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 987–1003.
9. W. R. J. D. Galloway and D. R. Spring, *Diversity-Oriented Synth.*, 2013, **1**, 21–28.
10. C. J. O' Connor, H. S. G. Beckmann and D. R. Spring, *Chem. Soc. Rev.*, 2012, **41**, 4444–4456.
11. M. Rask-Andersen, M. S. Almén and H. B. Schiöth, *Nat. Rev. Drug Discovery*, 2011, **10**, 579–590.
12. A. M. Edwards, R. Isserlin, G. D. Bader, S. V. Frye, T. M. Willson and F. H. Yu, *Nature*, 2011, **470**, 163–165.
13. C. Lipinski and A. Hopkins, *Nature*, 2004, **432**, 855–861.
14. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 2012, **64**, 4–17.

15. P. D. Leeson and B. Springthorpe, *Nat. Rev. Drug Discovery*, 2007, **6**, 881–890.
16. I. Kola and J. Landis, *Nat. Rev. Drug Discovery*, 2004, **3**, 711–716.
17. M. J. Waring, J. Arrowsmith, A. R. Leach, P. D. Leeson, S. Mandrell, R. M. Owen, G. Pairaudeau, W. D. Pennie, S. D. Pickett, J. Wang, O. Wallace and A. Weir, *Nat. Rev. Drug Discovery*, 2015, **14**, 475–486.
18. D. Cook, D. Brown, R. Alexander, R. March, P. Morgan, G. Satterthwaite and M. N. Pangalos, *Nat. Rev. Drug Discovery*, 2014, **13**, 419–431.
19. F. Lovering, J. Bikker and C. Humblet, *J. Med. Chem.*, 2009, **52**, 6752–6756.
20. F. Lovering, *Med. Chem. Commun.*, 2013, **4**, 515–519.
21. W. P. Walters, J. Green, J. R. Weiss and M. A. Murcko, *J. Med. Chem.*, 2011, **54**, 6405–6416.
22. M. M. Hann, A. R. Leach and G. Harper, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 856–864.
23. A. R. Leach and M. M. Hann, *Curr. Opin. Chem. Biol.*, 2011, **15**, 489–496.
24. F. López-Vallejo, M. A. Giulianotti, R. A. Houghten and J. L. Medina-Franco, *Drug Discovery Today*, 2012, **17**, 718–726.
25. T. J. Ritchie and S. J. F. Macdonald, *Drug Discovery Today*, 2009, **14**, 1011–1020.
26. T. J. Ritchie, S. J. F. MacDonald, R. J. Young and S. D. Pickett, *Drug Discovery Today*, 2011, **16**, 164–171.
27. T. J. Ritchie, S. J. F. Macdonald, S. Peace, S. D. Pickett and C. N. Luscombe, *Med. Chem. Commun.*, 2012, **3**, 1062.
28. T. J. Ritchie, S. J. F. Macdonald, S. Peace, S. D. Pickett and C. N. Luscombe, *MedChemComm*, 2013, **4**, 673.
29. P. D. Leeson and A. M. Davis, *J. Med. Chem.*, 2004, **47**, 6338–6348.
30. A. A Shelat and R. K. Guy, *Nat. Chem. Biol.*, 2007, **3**, 442–446.
31. J. D. Hughes, J. Blagg, D. A. Price, S. Bailey, G. A. DeCrescenzo, R. V. Devraj, E. Ellsworth, Y. M. Fobian, M. E. Gibbs, R. W. Gilles, N. Greene, E. Huang, T. Krieger-Burke, J. Loesel, T. Wager, L. Whiteley and Y. Zhang, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4872–4875.
32. C. J. Gerry and S. L Schreiber, *Nat. Rev. Drug Discovery*, 2018, **17**, 333–352.
33. W. R. J. D. Galloway, A. Isidro-Llobet and D. R. Spring, *Nat. Commun.*, 2010, **1**, 80.
34. O. Méndez-Lucio and J. L. Medina-Franco, *Drug Discovery Today*, 2017, **22**, 120–126.
35. P. A. Clemons, N. E. Bodycombe, H. A. Carrinski, J. A. Wilson, A. F. Shamji, B. K. Wagner, A. N. Koehler and S. L. Schreiber, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 18787–18792.
36. R. Bohacek and C. McMartin, *Med. Res.*, 1996, **16**, 3–50.
37. Z. L. Deng, C. X. Du, X. Li, B. Hu, Z. K. Kuang, R. Wang, S. Y. Feng, H. Y. Zhang and D. X. Kong, *J. Chem. Inf. Model.*, 2013, **53**, 2820–2828.

38. A. Schuffenhauer, N. Brown, P. Selzer, P. Ertl and E. Jacoby, *J. Chem. Inf. Model.*, 2006, **46**, 525–535.
39. A. L. Harvey, *Drug Discovery Today*, 2008, **13**, 894–901.
40. T. Rodrigues, D. Reker, P. Schneider and G. Schneider, *Nat. Chem.*, 2016, **8**, 531–541.
41. S. Wetzel, R. S. Bon, K. Kumar and H. Waldmann, *Angew. Chem. - Int. Ed.*, 2011, **50**, 10800–10826.
42. R. S. Bon and H. Waldmann, *Acc. Chem. Res.*, 2010, **43**, 1103–1114.
43. H. Van Hattum and H. Waldmann, *J. Am. Chem. Soc.*, 2014, **136**, 11853–11859.
44. K. Grabowski and G. Schneider, *Curr. Chem. Biol.*, 2007, **1**, 115–127.
45. D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2016, **79**, 629–661.
46. R. J. Spandl, A. Bender and D. R. Spring, *Org. Biomol. Chem.*, 2008, **6**, 1149–1158.
47. I. S. Young and M. A. Kerr, *J. Am. Chem. Soc.*, 2007, **129**, 1465–1469.
48. K. C. Nicolaou, C. R. H. Hale, C. Nilewski and H. A. Ioannidou, *Chem. Soc. Rev.*, 2012, **41**, 5185.
49. S. Basu, B. Ellinger, S. Rizzo, C. Deraeve, M. Schürmann, H. Preut, H.-D. Arndt and H. Waldmann, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 6805–6810.
50. H. van Hattum and H. Waldmann, *J. Am. Chem. Soc.*, 2014, **136**, 11853–11859.
51. A. Nören-Müller, I. Reis-Corrêa, H. Prinz, C. Rosenbaum, K. Saxena, H. J. Schwalbe, D. Vestweber, G. Cagna, S. Schunk, O. Schwarz, H. Schiewe and H. Waldmann, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 10606–10611.
52. A. I. Chan, L. M. McGregor and D. R. Liu, *Curr. Opin. Chem. Biol.*, 2015, **26**, 55–61.
53. R. M. Franzini, D. Neri and J. Scheuermann, *Acc. Chem. Res.*, 2014, **47**, 1247–1255.
54. R. E. Kleiner, C. E. Dumelin, G. C. Tiu, K. Sakurai and D. R. Liu, *J. Am. Chem. Soc.*, 2010, **132**, 11779–11791.
55. Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder and D. R. Liu, *Science*, 2004, **305**, 1601–1605.
56. D. Morton, S. Leach, C. Cordier, S. Warriner and A. Nelson, *Angew. Chem. Ind. Ed.*, 2009, **48**, 104–109.
57. D. Lee, J. K. Sello and S. L. Schreiber, *Org. Lett.*, 2000, **2**, 709–712.
58. S. L. Schreiber, *Science*, 2000, **287**, 1964–1969.
59. W. R. J. D. Galloway, A. Bender, M. Welch and D. R. Spring, *Chem. Commun.*, 2009, 2446–2462.
60. T. E. Nielsen and S. L. Schreiber, *Angew. Chem. - Int. Ed.*, 2008, **47**, 48–56.
61. M. Schirle and J. L. Jenkins, *Drug Discovery Today*, 2016, **21**, 82–89.
62. W. Zheng, N. Thorne and J. C. McKew, *Drug Discovery Today*, 2013, **18**, 1067–1073.
63. J. A. Lee, M. T. Uhlik, C. M. Moxham, D. Tomandl and D. J. Sall, *J. Med. Chem.*, 2012, **55**, 4527–4538.

64. D. C. Swinney and J. Anthony, *Nat. Rev. Drug Discovery*, 2011, **10**, 507–519.
65. D. R. Spring, *Chem. Soc. Rev.*, 2005, **34**, 472–482.
66. M. Hansson, J. Pemberton, O. Engkvist, I. Feierberg, L. Brive, P. Jarvis, L. Zander-Balderud and H. Chen, *J. Biomol. Screening*, 2013, **19**, 727–737.
67. E. Wyatt, W. Galloway, G. Thomas, M. Welch, O. Loiseleur, A. Plowright and D. Spring, *Chem. Commun.*, 2008, **40**, 4962–4964.
68. G. L. Thomas, R. J. Spandl, F. G. Glansdorp, M. Welch, A. Bender, J. Cockfield, J. A. Lindsay, C. Bryant, D. F. J. Brown, O. Loiseleur, H. Rudyk, M. Ladlow and D. R. Spring, *Angew. Chem. - Int. Ed.*, 2008, **47**, 2808–2812.
69. A. Robinson, G. L. Thomas, R. J. Spandl, M. Welch and D. R. Spring, *Org. Biomol. Chem.*, 2008, **6**, 2978–2981.
70. E. Lenci, R. Innocenti, A. Biagioni, G. Menchi, F. Bianchini and A. Trabocchi, *Molecules*, 2016, **21**, 1405.
71. L. Li, Z.-L. Li, F.-L. Wang, Z. Guo, Y.-F. Cheng, N. Wang, X.-W. Dong, C. Fang, J. Liu, C. Hou, B. Tan and X.-Y. Liu, *Nat. Commun.*, 2016, **7**, 13852.
72. S. Oh, H. J. Nam, J. Park, S. H. Beak and S. B. Park, *ChemMedChem*, 2010, **5**, 529–533.
73. A. M. Socha, N. Y. Tan, K. L. Laplante and J. K. Sello, *Bioorg. Med. Chem.*, 2010, **18**, 7193–7202.
74. E. E. Wyatt, S. Fergus, W. R. J. D. Galloway, A. Bender, D. J. Fox, A. T. Plowright, A. S. Jessiman, M. Welch and D. R. Spring, *Chem. Commun.*, 2006, 3296–3298.
75. P. C. L. Moore and J. A. Lindsay, *J. Med. Microbiol.*, 2002, **51**, 516–521.
76. M. R. Brown and A. W. Smith, *J. Antimicrob. Chemother.*, 2001, **48**, 141–142.
77. N. V. Kovalevskaya, E. D. Smurnyi, B. Birdsall, J. Feeney and V. I. Polshakov, *Pharm. Chem. J.*, 2007, **41**, 350–353.
78. K. Bowden, N. V. Harris and C. A. Watson, *J Chemother.*, 1993, **5**, 377–388.
79. S. Hoelder, P. A. Clarke and P. Workman, *Mol. Oncol.*, 2012, **6**, 155–176.
80. V. Lavanya, A. A. Mohamed Adil, N. Ahmed, A. K. Rishi and S. Jamal, *Integr. Cancer Sci. Ther.*, 2014, **1**, 39–46.
81. J. Zhang, P. Yang and N. Gray, *Nat. Rev. Cancer*, 2009, **9**, 28–39.
82. J. L. Adams, J. Smothers, R. Srinivasan and A. Hoos, *Nat. Rev. Drug Discovery*, 2015, **14**, 603–622.
83. I. Collins and A. M. Jones, *Molecules*, 2014, 17221–17255.
84. B. M. Ibbeson, L. Laraia, E. Alza, C. J. O' Connor, Y. S. Tan, H. M. L. Davies, G. McKenzie, A. R. , Venkitaraman and D. R. Spring, *Nat. Commun.*, 2014, **5**, 3155.
85. K. N. Niforou, A. K. Anagnostopoulos, K. Vougas, C. Kittas, V. G. Gorgoulis and G. T. Tsangaris, *Cancer Genomics Proteomics*, 2008, **5**, 63–77.

86. L. D. S. Yadav, V. P. Srivastava, V. K. Rai and R. Patel, *Tetrahedron*, 2008, **64**, 4246–4253.
87. H. Kubota, J. Lim, K. M. Depew and S. L. Schreiber, *Chem. Biol.*, 2002, **9**, 265–276.
88. A. Aravind, P. S. Kumar, M. G. Sankar and S. Baskaran, *Eur. J. Org. Chem.*, 2011, 6980–6988.
89. S. Senthilkumar, S. S. Prasad, P. S. Kumar and S. Baskaran, *Chem. Commun.*, 2014, **50**, 1549–1551.
90. N. Asano, R. J. Nash, R. J. Molyneux and G. W. J. Fleet, *Tetrahedron: Asymmetry*, 2000, **11**, 1645–1680.
91. N. Asano, T. Yamauchi, K. Kagamifuchi, N. Shimizu, S. Takahashi, H. Takatsuka, K. Ikeda, H. Kizu, W. Chuakul, A. Kettawan and T. Okamoto, *J. Nat. Prod.*, 2005, **68**, 1238–1242.
92. Y. Hu, K. Wang and J. B. MacMillan, *Org. Lett.*, 2013, **15**, 390–393.
93. E. Lenci, G. Menchi, A. Guarna and A. Trabocchi, *J. Org. Chem.*, 2015, **80**, 2182–2191.
94. E. M. Sánchez-Fernández, R. Gonçalves-Pereira, R. Rísquez-Cuadro, G. B. Plata, J. M. Padrón, J. M. García Fernández and C. Ortiz Mellet, *Carbohydr. Res.*, 2016, **429**, 113–122.
95. E. M. Sánchez-Fernández, R. Rísquez-Cuadro, M. Chasseraud, A. Ahidouch, C. Ortiz Mellet, H. Ouadid-Ahidouch and J. M. García Fernández, *Chem. Commun.*, 2010, **46**, 5328–5330.
96. C. H. Arrowsmith, J. E. Audia, C. Austin, J. Baell, J. Bennett, J. Blagg, C. Bountra, P. E. Brennan, P. J. Brown, M. E. Bunnage, C. Buser-Doepner, R. M. Campbell, A. J. Carter, P. Cohen, R. A. Copeland, B. Cravatt, J. L. Dahlin, D. Dhanak, A. M. Edwards, S. V. Frye, N. Gray, C. E. Grimshaw, D. Hepworth, T. Howe, K. V. M. Huber, J. Jin, S. Knapp, J. D. Kotz, R. G. Kruger, D. Lowe, M. M. Mader, B. Marsden, A. Mueller-Fahrnow, S. Müller, R. C. O'Hagan, J. P. Overington, D. R. Owen, S. H. Rosenberg, B. Roth, R. Ross, M. Schapira, S. L. Schreiber, B. Shoichet, M. Sundström, G. Superti-Furga, J. Taunton, L. Toledo-Sherman, C. Walpole, M. A. Walters, T. M. Willson, P. Workman, R. N. Young and W. J. Zuercher, *Nat. Chem. Biol.*, 2015, **11**, 536–541.
97. J. A. Zorn and J. A. Wells, *Nat. Chem. Biol.*, 2010, **6**, 179–188.
98. D. E. Scott, A. R. Bayly, C. Abell and J. Skidmore, *Nat. Rev. Drug Discovery*, 2016, **15**, 533–550.
99. J. Kim, J. Jung, J. Koo, W. Cho, W. S. Lee, C. Kim, W. Park and S. B. Park, *Nat. Commun.*, 2016, **7**, 13196.
100. J. Kim, H. Kim and S. B. Park, *J. Am. Chem. Soc.*, 2014, **136**, 14629–14638.
101. J. M. Han, S. J. Jeong, M. C. Park, G. Kim, N. H. Kwon, H. K. Kim, S. H. Ha, S. H. Ryu and S. Kim, *Cell*, 2012, **149**, 410–424.
102. S. Dandapani, A. R. Germain, I. Jewett, S. le Quement, J.-C. Marie, G. Muncipinto, J. R. Duvall, L. C. Carmody, J. R. Perez, J. C. Engel, J. Gut, D. Kellar, J. Lage Siqueira-Neto, J. H. McKerrow, M. Kaiser,

A. Rodriguez, M. A. Palmer, M. Foley, S. L. Schreiber and B. Munoz, *ACS Med. Chem. Lett.*, 2014, **5**, 149–153.
103. L. A. Marcaurelle, E. Comer, S. Dandapani, J. R. Duvall, B. Gerard, S. Kesavan, M. D. Lee, H. Liu, J. T. Lowe, J. C. Marie, C. A. Mulrooney, B. A. Pandya, A. Rowley, T. D. Ryba, B. C. Suh, J. Wei, D. W. Young, L. B. Akella, N. T. Ross, Y. L. Zhang, D. M. Fass, S. A. Reis, W. N. Zhao, S. J. Haggarty, M. Palmer and M. A. Foley, *J. Am. Chem. Soc.*, 2010, **132**, 16962–16976.
104. B. Gerard, J. R. Duvall, J. T. Lowe, T. Murillo, J. Wei, L. B. Akella and L. A. Marcaurelle, *ACS Comb. Sci.*, 2011, **13**, 365–374.
105. M. E. Fitzgerald, C. A. Mulrooney, J. R. Duvall, J. Wei, B. C. Suh, L. B. Akella, A. Vrcic and L. A. Marcaurelle, *ACS Comb. Sci.*, 2012, **14**, 89–96.
106. D. H. C. Chou, J. R. Duvall, B. Gerard, H. Liu, B. A. Pandya, B. C. Suh, E. M. Forbeck, P. Faloon, B. K. Wagner and L. A. Marcaurelle, *ACS Med. Chem. Lett.*, 2011, **2**, 698–702.
107. L. N. Aldrich, S. Y. Kuo, A. B. Castoreno, G. Goel, P. Kuballa, M. G. Rees, B. A. Seashore-Ludlow, J. H. Cheah, I. J. Latorre, S. L. Schreiber, A. F. Shamji and R. J. Xavier, *J. Am. Chem. Soc.*, 2015, **137**, 5563–5568.
108. S.-Y. Kuo, A. B. Castoreno, L. N. Aldrich, K. G. Lassen, G. Goel, V. Dančík, P. Kuballa, I. Latorre, K. L. Conway, S. Sarkar, D. Maetzel, R. Jaenisch, P. A. Clemons, S. L. Schreiber, A. F. Shamji, R. J. Xavier, M. A. Fischbach and E. Latz, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E4281–E4287.
109. R. W. Heidebrecht, C. Mulrooney, C. P. Austin, R. H. Barker, J. A. Beaudoin, K. C. Cheng, E. Comer, S. Dandapani, J. Dick, J. R. Duvall, E. H. Ekland, D. A. Fidock, M. E. Fitzgerald, M. Foley, R. Guha, P. Hinkson, M. Kramer, A. K. Lukens, D. Masi, L. A. Marcaurelle, X. Su, C. J. Thomas, M. We, R. C. Wiegand, D. Wirth, M. Xia, J. Yuan, J. Zhao, M. Palmer, B. Munoz and S. Schreiber, *ACS Med. Chem. Lett.*, 2012, **3**, 112–117.
110. E. Comer, J. A. Beaudoin, N. Kato, M. E. Fitzgerald, R. W. Heidebrecht, M. D. Lee, D. Masi, M. Mercier, C. Mulrooney, G. Muncipinto, A. Rowley, K. Crespo-Llado, A. E. Serrano, A. K. Lukens, R. C. Wiegand, D. F. Wirth, M. A. Palmer, M. A. Foley, B. Munoz, C. A. Scherer, J. R. Duvall and S. L. Schreiber, *J. Med. Chem.*, 2014, **57**, 8496–8502.
111. A. K. Lukens, R. W. Heidebrecht, C. Mulrooney, J. A. Beaudoin, E. Comer, J. R. Duvall, M. E. Fitzgerald, D. Masi, K. Galinsky, C. A. Scherer, M. Palmer, B. Munoz, M. Foley, S. L. Schreiber, R. C. Wiegand and D. F. Wirth, *J. Infect. Dis.*, 2015, **211**, 1097–1103.
112. N. Kato, E. Comer, T. Sakata-Kato, A. Sharma, M. Sharma, M. Maetani, J. Bastien, N. M. Brancucci, J. A. Bittker, V. Corey, D. Clarke, E. R. Derbyshire, G. Dornan, S. Duffy, S. Eckley, M. A. Itoe, K. M. J. Koolen, T. A. Lewis, P. S. Lui, A. K. Lukens, E. Lund, S. March, E. Meibalan, B. C. Meier, J. McPhail, B. Mitasev, E. L. Moss, M. Sayes,

Y. VanGessel, M. J. Wawer, T. Yoshinaga, A.-M. Zeeman, V. M. Avery, S. N. Bhatia, J. E. Burke, F. Catteruccia, J. C. Clardy, P. A. Clemons, K. J. Dechering, J. R. Duvall, M. A. Foley, F. Gusovsky, C. H. M. Kocken, M. Marti, M. L. Morningstar, B. Munoz, D. E. Neafsey, A. Sharma, E. A. Winzeler, D. F. Wirth, C. A. Scherer and S. L. Schreiber, *Nature*, 2016, **538**, 344–349.

113. J. T. Lowe, M. D. Lee, L. B. Akella, E. Davoine, E. J. Donckele, L. Durak, J. R. Duvall, B. Gerard, E. B. Holson, A. Joliton, S. Kesavan, B. C. Lemercier, H. Liu, J. C. Marié, C. A. Mulrooney, G. Muncipinto, M. Welzel-O'Shea, L. M. Panko, A. Rowley, B. C. Suh, M. Thomas, F. F. Wagner, J. Wei, M. A. Foley and L. A. Marcaurelle, *J. Org. Chem.*, 2012, **77**, 7187–7211.