The discovery of antibacterial agents using diversity-oriented synthesis

Warren R. J. D. Galloway,[a] Andreas Bender,[b] Martin Welch[c] and David R. Spring*[a]

Received (in Cambridge, UK) 25th September 2008, Accepted 24th February 2009
First published as an Advance Article on the web 1st April 2009
DOI: 10.1039/b816852k

The emergence and increasing prevalence of multi-drug resistance bacterial strains represents a clear and present danger to the standard of human healthcare as we know it. The systematic study of modulating biological systems using small molecules (so-called chemical genetics) offers a potentially fruitful means of discovering critically needed new antibacterial agents. Crucial to the success of this approach is the ready availability of functionally diverse small molecule collections. In this feature article we focus upon the use of a diversity-oriented synthesis (DOS) approach for the efficient generation of such compound collections, and discuss the utility of DOS for the discovery of new antibacterial agents.

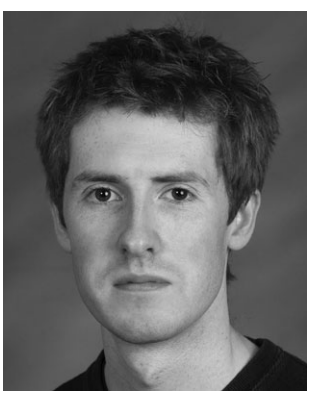
Introduction

The development of antibacterial agents is one of the greatest successes of 20th century medicine.[1] Since the serendipitous discovery of penicillin by Alexander Fleming in 1928, an arsenal of antibacterial agents have been developed and found widespread clinical application. However, bacteria have quickly become resistant to commonly prescribed antibiotics.[2,3] Combined with the lack of fundamental antibiotic research carried out by pharmaceutical companies over recent decades we are left with a legacy of relatively few efficacious drugs.[4–6] Indeed, there is little doubt amongst many involved in antibacterial research that the existing drugs we have in hand for the treatment of infectious diseases are insufficient to protect us in the long term.[3] Thus, bacterial infection, particularly from multi-drug resistant strains, remains a serious threat to human lives[3,4,7,8] and there is a clear and critical medical need for the discovery of novel antibacterial agents. In this context, small molecular mass compounds (so-called small molecules†) are certain to play a prominent role.

Small organic molecules have always been of interest in chemistry and biochemistry due to their ability to exert powerful effects on the functions of macromolecules that comprise living systems.[9,10] Indeed, the underlying approach of using small molecules to treat disease represents the basis for medicinal chemistry as we know it.[11] There are a several

[a] *Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, UK. E-mail: drspring@ch.cam.ac.uk; Fax: +44 (0) 1223-336362; Tel: +44 (0) 1223-336498*
[b] *Leiden/Amsterdam Center for Drug Research, Leiden University, Gorlaeus Laboratories, Leiden, 2300, The Netherlands*
[c] *Department of Biochemistry, University of Cambridge, Cambridge, UK. E-mail: mw240@mole.bio.cam.ac.uk; Tel: +44 (0) 1223-333640*

† The term 'small molecules' has no strict definition, but it is usually used to describe carbon-based compounds whose molecular weight is generally under 2000 g mol^{-1} and always less than that of macromolecules such as DNA, RNA and proteins.[128,129]



Warren Galloway

Warren Galloway was born in Dundee, Scotland in 1981 and attended the University of Cambridge for his undergraduate chemistry degree, graduating in 2004. He stayed at Cambridge for his PhD studies which were funded by the EPSRC and Pfizer. Under the supervision of Dr David Spring, his doctoral research focused upon the development of strategies for diversity-oriented synthesis. Warren received his PhD in 2008 and is currently a postdoctoral research assistant for Professor Robert Glen and Dr David Spring at Cambridge, working on the design and synthesis of novel modulators of 5-HT receptors.

Andreas Bender

Andreas Bender is an Assistant Professor for Cheminformatics with the Leiden/Amsterdam Center for Drug Research (LACDR). Before joining the LACDR, he was a Presidential Postdoctoral Fellow with the Novartis Institutes for Biomedical Research in Cambridge/MA, working on in silico *chemogenomics methods to understand small molecule bioactivity space. He earned his PhD from the University of Cambridge working with Prof Robert Glen on small-molecule data analysis. Andreas' work is centered on cheminformatics topics, dealing with small molecule activity models, the analysis of high-content and cellular screening data and the prediction of physicochemical properties of molecules.*

benefits associated with the use of small molecules as therapeutic agents including improved stability over peptides in oral administration (where peptide bonds are easily cleaved by proteases), synthetic accessibility and, perhaps most notably, optimisation of compound bioactivity is considerably easier for a small molecule than for complex macromolecules.

Methods which utilise small molecules as 'chemical probes' to modulate biological systems can be described by the umbrella term[12] *chemical genetics*.[13–15] Whereas traditional genetics uses gene knockouts (or knock-ins) on the level of the DNA, chemical genetics uses biologically active small molecules to directly attenuate the corresponding biological macromolecular product and thus affect a biological response, for example, the inhibition of bacterial growth. The first step of a chemical genetics experiment involves the identification of a small molecule which induces a desired phenotype (forward chemical genetics) or modulates the function of a specific protein of interest (reverse chemical genetics). Thus, in the former case, investigations proceed from phenotype to protein, whereas in the latter case investigations progress from protein to phenotype. Small molecules that exhibit biological effects can be discovered by both forward and reverse chemical genetics approaches through the screening of collections (or 'libraries') of small molecules to identify those with the desired characteristics (so-called 'hits').

This feature article discusses various aspects associated with the use of a forward chemical genetics approach for the discovery of new antibacterial agents. In particular, we focus on the crucial importance of 'high-quality' rather than 'high-quantity' small molecule libraries, and the use of diversity-oriented synthesis (DOS) to generate such compound collections in a highly efficient manner. We conclude with some representative examples from our own group which highlight the utility of a DOS approach for the discovery of new antibacterial agents.

Chemical genetics and antibacterial discovery

The ultimate aim of any biological screening project is to identify biologically active small molecules. In particular, there is a demand for *novel* biologically active molecules that exert their biological effect through a unique mechanism. This is perhaps most pertinent in the field of antibacterial discovery, where it is widely recognised that dealing with the problem of bacterial resistance is not simply a case of synthesising new variants of existing drugs (see Fig. 1 for some examples of existing antibiotics).‡

Established antibiotics are based around a limited number of structural classes which interfere with a small number of biological targets and operate by a limited array of mechanisms.[17–20] Resistance to one antibiotic of a class frequently leads to resistance to its' whole class;[21,22] although new derivatives of existing agents can improve efficacy temporarily, the underlying resistance mechanisms are still present in the environment.[5] Consequently, there is an urgent need for new antibacterial agents that have new modes of action or interfere with novel targets[23] and, as such, are not subject to the resistance mechanisms that have evolved to combat previous generations of antibacterials.[24]

In the quest to discover novel antibacterial agents the use of forward chemical genetics is particularly powerful; due to the inherent lack of biological target-bias associated with the approach, a single forward chemical genetics experiment has the potential to identify numerous bioactive small molecules which could operate *via* several different modes of action. A forward chemical genetics programme generally consists of three defined stages: the synthesis of a 'suitably designed' small molecule library, a phenotypic screen to identify bioactive compounds and a biological target identification method.[12] Work has have been carried out within the author's research

‡ An inherent problem associated with the 'antibiotic approach' is that by inhibiting bacterial cellular processes that are essential for microbial survival an intense selection pressure for drug-resistant mutations is created. Consequently the development of resistance is inevitably the result of antibiotic use and ultimately limits the efficacy and life span of every antibiotic.[17] That is, bacterial resistance to antibiotics can never be overcome, it is simply a case of managing it as well as possible.

Martin Welch

Martin Welch grew up near Newmarket and went on to study Biochemistry at Oxford University. His PhD was with Prof Michael Eisenbach (Weizmann Institute), where he worked on the mechanism of switching in the bacterial flagellar motor. Next, he joined the laboratory of Prof Jean-Pierre Samama (Toulouse), where he solved the X-ray crystal structure of one of the protein complexes involved in bacterial chemotaxis. Later he joined the lab of Prof George Salmond, establishing an interest in quorum sensing. Presently, he is a lecturer in Biochemistry in Cambridge. His research interests lie in understanding how antibiotics affect cell function.

David Spring

David Spring is currently an EPSRC Advanced Fellow at the University of Cambridge Chemistry Department. Previous to this appointment he spent two and a half years as a Wellcome Trust Postdoctoral Fellow and Fulbright Scholar at Harvard University with Professor Stuart L. Schreiber. He gained his DPhil for work on the proposed biosynthesis of the manzamine alkaloids at Oxford University under the supervision of Professor Sir Jack E. Baldwin FRS. David's research programme is focused on diversity-oriented synthesis, synthetic methodology and chemical genetics.

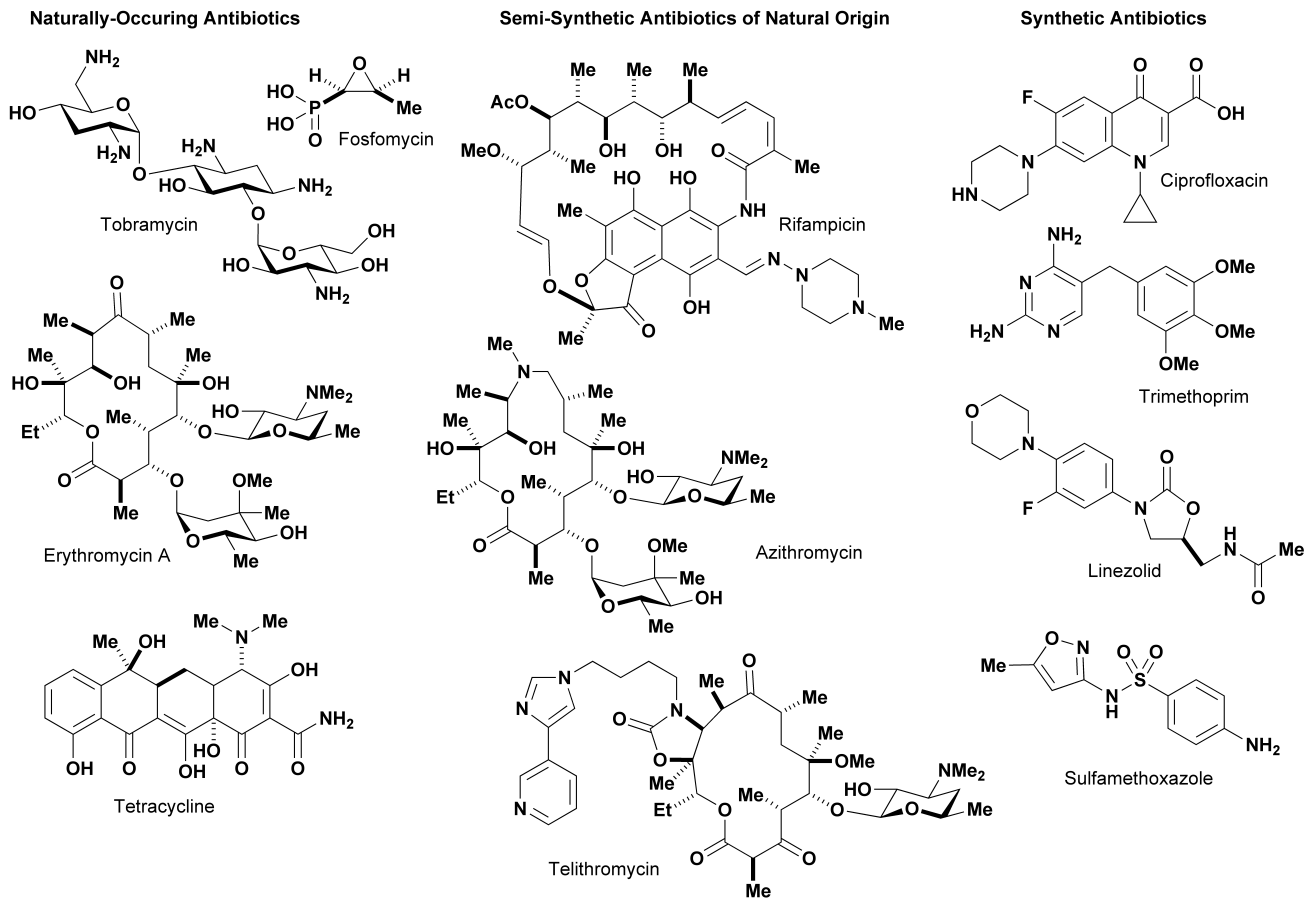


Fig. 1 Some selected examples of antibiotics. Most of the major classes of antibiotics in therapeutic use are natural products or semi-synthetic derivatives thereof.[16] There are three main classes of synthetic antibiotics in clinical use: sulfa drugs (*e.g.* sulfamethoxazole), the quinolones (*e.g.* ciprofloxacin) and oxazolidinones (*e.g.* linezoild).[16]

groups to address all of these elements within the context of the discovery of novel antibacterial agents.

Library design in the context of antibacterial discovery: biological and structural diversity

The ready availability of small molecules is of vital importance in biological screening experiments.[25] However, a fundamental consideration is what compounds should be synthesised and employed in these screening processes.[26] Ultimately this is determined by the desired outcome of the experiment and the requirements this places upon the compounds.[27] For example, if the goal of a screening process is to identify a small molecule for application as an orally bioavailable drug, then several observations have been made as to molecular characteristics that are desirable (most notably Lipinski's 'rule-of-five' criteria).[27–29] However, since antibacterial compounds are generally more hydrophilic and have a greater molecular weight than other drugs in general (Fig. 2), the guidelines that govern the discovery of orally bioavailable drugs are unlikely to be useful for the identification of new antibiotic compounds.[5]

An important consideration is the nature of the biological target of interest. When a specific biological molecule or family of molecules is targeted the compounds used in the screening process are usually carefully designed based on knowledge of the targets structure or the structure of known natural ligands.[27,31] However, if the screening process is unbiased, such as in a forward chemical genetics experiment where the precise nature of the biological target is unknown, the selection criteria for small molecules is dramatically complicated and cannot be defined *a priori*.[26,32]

The identification of novel biologically active small molecules may be aided by screening functionally (biologically) diverse compound libraries, since it has been argued that a greater sample of the bioactive chemical universe (*i.e.* of all bioactive molecules) increases the chance of identifying a compound with the desired properties (for example, a small molecule with antibacterial properties that operates *via* a novel mode of action).[33,34] The overall functional (biological) diversity of a small molecule library has been shown to be directly correlated with its overall structural diversity, which in turn is related to the amount of *chemical space* the library occupies.[33,35]

Structural diversity and chemical space

Chemical compounds can be characterized by a wide variety of physiochemical and topological 'descriptors' which contain information about either the bulk properties of the compound, such as molecular mass and lipophilicity, or its topological features such as degree of branching.[36–39] 'Chemical space' is a

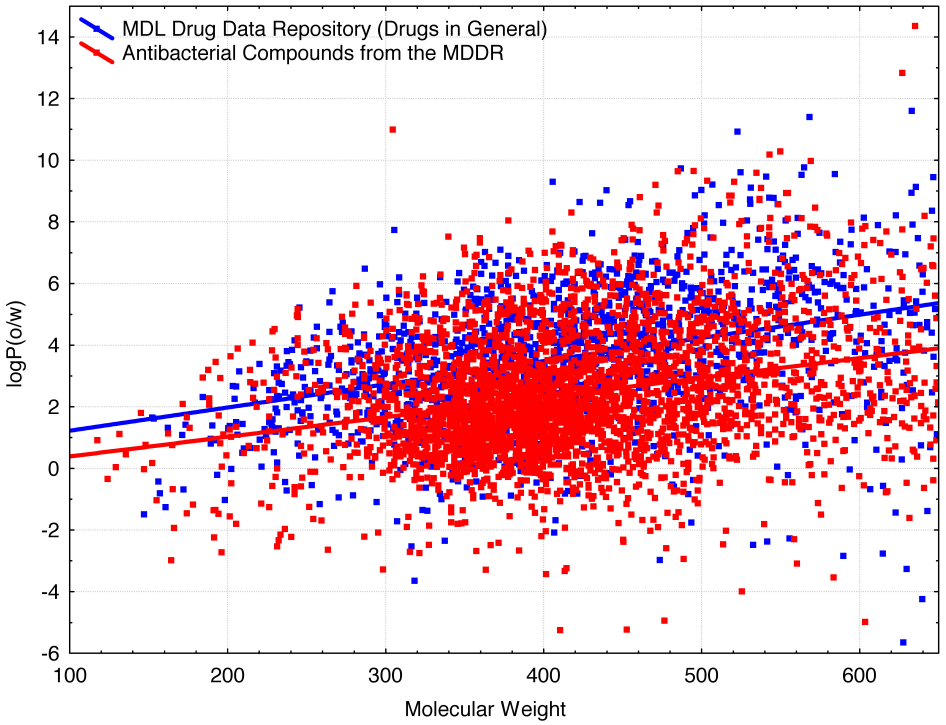


Fig. 2 A plot of calculated log *P*(o/w) (partition coefficient between octanol and water)[30] *vs.* molecular mass of antibacterial compounds (red squares) compared with compounds from the MDL Drug Data Repository (MDDR) database (blue circles). Antibacterial molecules are, on average, more hydrophilic and typically have a greater molecular weight, which is in agreement with previous reports.[5]

term often used in place of multidimensional descriptor space; it is a region defined by a particular choice of descriptors and the limits placed upon them.[40] In the context of small molecule libraries, chemical space can be defined as the total descriptor space that encompasses all the small carbon-based molecules that could, in principle, be created.[40]

Each specific molecule will have a unique combination of molecular properties and thus descriptor values, and will therefore reside at a discrete point in chemical space. Consequently, the structural features present within a collection of molecules will influence the overall distribution of the collection in chemical space. Thus it follows that the more structurally diverse the library, the more chemical space it interrogates. As molecular shape is intrinsically linked to biological activity, the greater the structural diversity and thus chemical space coverage of a library, the greater its overall functional (biological) diversity and consequently the better the odds of identifying small molecule modulators for a broad range of biological targets.[35] This viewpoint is supported by the similar property principle[41] which states that structurally similar molecules will exhibit similar physiochemical and biological properties.[42]

Biologically relevant chemical space

Biologically relevant chemical space has been defined as those parts of chemical space in which biologically active compounds reside.[40] The degree of overlap between biologically relevant chemical space and total chemical space is somewhat of a contentious issue and subject to much debate in the literature.[9,26,27,43] The limits of biologically relevant chemical space are defined by the specific binding interactions that must occur between small molecules and the three-dimensional (3D) molecular recognition patterns on biological molecules such as proteins.[29] What is uncertain is the size of this region in comparison to total chemical space, *i.e.* whether the biologically relevant region is 'small' and most of the chemical universe is 'empty' (containing no therapeutically interesting compounds).[29] That is, are the regions of chemical space defined by natural products and known drugs the most fertile regions for discovering biologically active small molecules,[9] or is there scope for discovering useful bioactive molecules, particularly those with novel modes of action, from 'un-tapped' areas of chemical space?[27,29] What is clear is that if we do not try and access such regions, we will never know! In particular, since antibacterials are rather 'un-drug-like' structures in the sense that they are, on average, larger and more hydrophilic than other drug compounds (Fig. 2), it is likely that chemical space with these properties has not yet been sampled sufficiently. In spite of this controversy, structural diversity (and thus total chemical space coverage), though rarely the 'end-game' in a chemical genetics experiment, is generally perceived to be an important consideration in small molecule library synthesis, particularly when the precise nature of the biological target molecule is unknown or the identification of a novel biologically active molecule is desired.[25,44,45]

Maximising structural diversity and the importance of structural complexity

It is widely accepted that it is not synthetically feasible to produce all theoretically stable, small carbon-based

molecules.[27,29,32] Therefore the generation of a collection of structurally diverse small molecules offers a unique challenge to the synthetic chemist.[26,32] Making molecules costs, both in terms of time and money; therefore *selectivity* in synthesis is an important consideration. This issue has spurred the development of several different approaches to efficiently interrogate wide regions of chemical space simultaneously or to identify and access areas of chemical space which have an enhanced probability of containing bioactive compounds.[26]

Though the term 'diversity' is to some degree a subjective one, there are four main components of structural diversity that have been consistently identified in the literature:[32,46]

1. *Appendage diversity* (or building-block diversity)—variation in different structural moieties around a common skeleton;

2. *Functional group diversity*—variation in the functional groups present;

3. *Stereochemical diversity*—variation in the orientation of potential macromolecule-interacting elements;[9]

4. *Skeletal diversity*—presence of many distinct molecular skeletons (or frameworks/scaffolds).§

Increasing the skeletal diversity in a small molecule library is widely regarded as one of the most effective ways of increasing the overall structural diversity of the library.[25,35,47–49] Furthermore, computational analyses have been carried out to support the notion that small multiple scaffold libraries are superior to large single-scaffold libraries in terms of bio-relevant diversity.[34] Libraries which are based around a single scaffold, regardless of their size, are restricted to a limited number of molecular shapes, as opposed to smaller libraries designed around multiple scaffolds.[34,35] Consequently, the compounds in libraries which are based around a common molecular skeleton tend to display chemical information similarly in 3D space, thus limiting the range of potential binding partners to a small set of macromolecules with a complementary 3D binding surface.[9,47]

In addition to structural diversity, *structural complexity* is another characteristic that is important in small molecule libraries. Although there is some debate in the literature, it has been argued that molecules that are structurally complex are more likely to interact with biological macromolecules in a selective and specific manner.[25,27,29,50]

Sources of small molecules

In the context of the discovery of novel biologically active small molecules, including antibacterial agents, what is required is a small molecule library that efficiently interrogates large areas of chemical space simultaneously and whose members have the potential to act as selective modulators of biological molecules. Thus what is needed are molecules that are *structurally diverse* (and consequently functionally diverse) and *structurally complex*. A collection of such compounds has the potential to provide hits against a panel of biological targets, allowing the discovery of small molecules with previously unknown (and potentially novel) biological effects.[51]

There are a number of potential sources of small molecules for use in biological screens. Nature has traditionally served as a rich source of biologically active molecules[46,52,53] which exhibit enormous structural diversity[54] and usually a high degree of specificity for their biological target. A multitude of natural products have proven to be useful as drugs or leads[55,56] and are still a major source of innovative therapeutic agents for infectious diseases.[17,52] Indeed, most of the major classes of antibiotics in therapeutic use are natural products or semi-synthetic derivatives thereof (Fig. 1).[16] In addition, due the evolution of natural products to interact with target proteins, hit rates in high-throughput screens are usually several times higher for natural products compared to small molecule libraries from conventional synthetic sources (*vide infra*). Unfortunately there are several problems associated with using natural product compounds in screening experiments including difficulties with purification, bioactive component identification, structural assignment and chemical modification and analogue synthesis (processes which are pertinent in the drug development process).[25] Of particular importance for the discovery of novel antibacterial agents is the fact that natural products only occupy a small proportion of total chemical space[29,40] which runs the risk of omitting a vast number of possibly biologically valuable small molecules from any screening process.[9]

The problems associated with the use of natural product compounds in biological screening experiments has spurred the development of several different *synthetic* approaches for the *de novo* creation of small molecule collections. It is possible to divide these approaches into two broad categories based upon the nature of the molecular skeletons present in the small molecules that are generated:

1. *Biased approaches*: the synthetic route is designed with a pre-encoded skeletal bias such that all compounds are based around a similar molecular skeleton;

2. *Non-biased approaches:* the synthetic route is designed with no pre-encoded skeletal bias such that a range of different molecular skeletons are present in the final small molecule collection.

It should be noted that this distinction is somewhat arbitrary and most approaches towards small molecule library synthesis lie somewhere along a continuum between these two extremes. Nevertheless, this division does provide a useful framework for the following discussion.

Biased approaches: combinatorial synthesis and synthesis around privileged structures

Pharmaceutical proprietary compound collections are traditionally very important sources of small molecules.[35,54] At the beginning of the 1990s a typical pharmaceutical company would have possessed numerous collections of structurally related compounds which had been synthesised during traditional medicinal chemistry-lead optimisation campaigns.[57] Such compound collections would have been

§ The term 'molecular skeleton' has no strict definition. Within the context of this report, the description recently outlined by Schreiber is appropriate; the term skeleton loosely denotes rigidifying elements in small molecules; these can be atom connectivities that yield either linked, fused, bridged or spiro rings, or acyclic conformational elements that provide substantial rigidification by avoiding non-bonding interactions.[96]

panned for possible antibacterials (and many other drug types) over the course of many years and it was acknowledged at this time that the 'low hanging fruit' from the antibiotic tree had probably already been picked.[5] Companies therefore sought new sources of compounds to augment their screening collections.[5,57] Combinatorial chemistry emerged as a key strategy in this regard.

In a very broad sense combinatorial chemistry may be defined as the rapid synthesis and screening of libraries of varied compounds to identify agents with desired functional properties.[58] Combined with established high-throughput screening (HTS) techniques, the development of combinatorial chemistry strategies in the early 1990s enabled the synthesis and testing of libraries of hundreds of thousands of different compounds at comparably low cost.[54] The approach was quickly embraced by the pharmaceutical industry with the hope that new drug leads, including novel antibiotic classes, would be produced by sheer weight of numbers. However, the expected surge in productivity, both in terms of antibiotic discovery and drug discovery in general, has not materialised. Indeed, as of the end of the calendar year 2007 there was only one reported *de novo* new chemical entity resulting from this method of chemical discovery which had been approved for drug use (as an antitumour agent).[56]

This 'innovation deficit' is generally attributed to defects in the nature of the libraries produced.[59] Early combinatorial libraries have been described as being intrinsically useless for drug discovery[60] because the compounds were too similar to each other, having limited structural diversity[54] and consequently offering only a narrow slice of chemical space.[60] This limited degree of overall structural diversity can be primarily attributed to a lack of skeletal diversity. Traditionally, combinatorial libraries contained large numbers of molecules which were each based around a specific molecular skeleton (a 'one-synthesis/one skeleton approach') which resulted in a high degree of appendage and functional group diversity (and possibly stereochemical diversity) but little variety in the nature of the core molecular scaffold.[9,25,35] Recent years have witnessed the development of more considered synthetic strategies in an attempt to increase the structural diversity exhibited by combinatorial libraries.[35] Nevertheless even these approaches are generally limited to known biologically active frameworks and, as such, have met with limited success in identifying *novel* biologically active small molecules.

In the context of antibacterial discovery it has been argued that pharmaceutical companies should place a greater emphasis on generating libraries which access the unique physiochemical property space spanned by known antibacterial agents;[61] existing antibacterials do not generally follow Lipinski's 'rule-of-five' criteria whereas corporate compound collections have always been heavily biased towards compounds that do (*vide supra*).[5] Therefore, to some researchers it is unsurprising that such collections have been found to be inadequate for the identification of new antibacterial agents.[5,61] Indeed, in a recent review on antibiotic discovery, researchers from GlaxoSmithKline commented that a new paradigm for library design in the context of antibacterial discovery is needed, with an emphasis upon greater molecular diversity and a better understanding of which physical chemical properties are important for antibacterials.[5] However, it is important to note that new antibacterial agents with novel modes of action may exist outside of known 'antibiotic physiochemical space'; consequently, small molecules libraries that simultaneously cover *both* known 'antibiotic chemical space' and unexplored regions of chemical space may prove particularly valuable in this context.

Although molecules based around a specific core skeleton are usually limited to a smaller range of biological partners than molecules based around a wide variety of skeletons, it has emerged that there is a sub-set of molecular skeletons whose presence in molecules confers upon them more flexible binding properties. These skeletons are known as 'privileged' frameworks (or substructures[62]) which have been defined as molecular scaffolds with versatile binding properties, such that a single scaffold is able to provide potent and selective ligands for a range of different biological targets (including multiple, unrelated classes of protein receptors[63]) through modification of functional groups.[64,65] Privileged scaffolds are common to known biologically active molecules (usually natural products) and thus have proven biological relevance.[26,43,64]

Recent years have witnessed an increase in the synthesis of small molecule libraries which are based around such privileged structures.[66–71] The rationale behind using such an approach is based on two main hypotheses. Firstly, that millions of years of evolutionary pressure has 'pre-validated' natural products, and thus compounds that are structurally similar, to be able to modulate protein function.[26,43,72–74] Secondly, that the chemical space explored by natural products and protein structure during evolution is strongly limited in size and highly conserved.[26] That is, there is a concept of 'evolutionary convergence of structures' in that natural products have evolved to interact with multiple proteins.[73,74] Synthesis around a privileged scaffold has been described as being distinct from the process of (focused) combinatorial library synthesis because the ultimate goal is the identification of compounds with novel biological properties distinct from those of the original privileged compound.[72] The screening of compound collections containing various privileged scaffold motifs has successfully identified numerous novel compounds with antibacterial activity (Fig. 3, though at the time of writing none have been successfully developed into marketed drugs).

Critics of the privileged scaffold approach have argued that many privileged structures have limited utility due to their promiscuous nature (that is, a lack of target specificity).[62] Typical privileged structure analyses have attempted to identify *minimal* substructures that are frequently found in ligands associated with a particular biological target family (*e.g.* G-coupled protein receptors, serine proteases, protein kinases and ligand-gated ion channels).[78] However, this process can easily detect structures that are merely drug-like and/or promiscuous protein binders and therefore confer bioactivity across a wide range of target families.[78] In recent years the definition of privileged structure has been modified by some researchers to describe commonly occurring fragments within ligands associated with a *particular* target family; that is, the term has become identified with those substructures found to be promiscuous within a given target

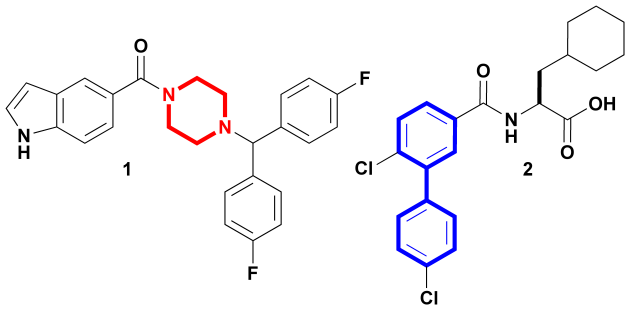


Fig. 3 Some examples of compounds with antibacterial activity based around privileged scaffolds. Compound **1** was found to act as an inhibitor of the *Mycobacterium tuberculosis* enoyl acyl carrier protein reductase InhA.[75] **1** is based around a piperazine scaffold (highlighted in red) which is frequently found in biologically active compounds across a number of different therapeutic areas, including antifungals, antidepressants and antivirals.[75,76] Compound **2** was found to exhibit antibacterial activity against a range of Gram-positive bacteria.[77] **2** is based around a biaryl scaffold (highlighted in blue) which has shown activity against across a wide range of therapeutic classes.[62]

family and carries that implication that these substructures are *specific* to that target family.[78] However a library based around a 'promiscuous privileged scaffold' may be advantageous if the compounds are to be screened against a wide variety of targets.[78] It has been argued that selectivity could possibly be introduced at a later date through appropriate substitution of the molecular framework.[62,78]

Related approaches include 'natural-product guided' methods such as 'biologically oriented synthesis'[79,80] and 'biology inspired synthesis'[26] which involve the generation of compound libraries based around the core structures of natural product templates with proven biological relevance (*i.e.* in some sense 'traditional' combinatorial diversity around a natural product scaffold or fragment thereof, with the aim of providing screening collections with high biological relevance).[81] These methods have had success in the identification of novel bioactive molecules, including antibacterials.[81–83] However, conservation of a significant portion of the core structure of a parent antibacterial compound may be a potential drawback in the context of antibiotic development; the antibacterial 'hits' from such a collection are likely to share a similar mode of action as the parent compound and thus be more susceptible to existing resistance mechanisms (*vide supra*). That is, this approach may be less likely to lead to *novel* antibacterials that exert their biological effect through a unique mechanism.

Clearly, such 'biased' synthesis methods are particularly relevant when a specific biological target is being considered. When a less focused approach is required, for example the biological target is unknown or we are hoping to find novel biologically active molecules, such as antibacterial agents, the use of non-biased synthetic approaches, which aim to access a wider range of chemical descriptor space, may be more useful.

Non-biased approaches: diversity-oriented synthesis as a source of skeletally diverse small molecules

Diversity-oriented synthesis (DOS) has recently emerged as a new synthetic approach towards library synthesis which aims to meet the challenge of synthesising structurally diverse small molecule collections.[9,32,47,84]

DOS has been defined as the deliberate, simultaneous and efficient synthesis of more than one target compound in a diversity-driven approach.[32] The aim of a DOS is to efficiently interrogate wide areas of chemical space simultaneously; this may include known bioactive regions of chemical space *and* unexplored regions of chemical space.[9,31,32,85] The hope is that by sampling a greater total area of chemical space the functional (biological) diversity of the library will be greater, thus increasing the chances of identifying a compound with the desired properties, for example antibacterial activity through a novel mode of action.[33–35]

A comparison of the synthetic strategies used in a DOS and a 'traditional' combinatorial approach is given in Fig. 4. A DOS pathway is analysed in the *forward* sense; simple starting materials (in this case, a single compound) are converted into a collection of structurally diverse small molecules in no more then five synthetic steps (in order to maximise synthetic efficiency).[9,32,86] The overall aim is the broad coverage of chemical space, which can be contrasted with the outcomes of more traditional combinatorial syntheses.

A successful DOS must address the four main types of structural diversity mentioned previously; that is, appendage, functional group, stereochemical and skeletal.[9,25,26,86,87] The most challenging facet of DOS, and of central importance to its success, is the ability to incorporate skeletal diversity into a compound collection.[9,47,85,88] Making molecules costs, both in terms of time and money. Therefore the ideal synthesis of a structurally diverse small molecule collection is one in which this diversity is achieved in the most efficient manner. The efficient generation of multiple molecular scaffolds is regarded as one of the most effective methods of increasing the overall structural diversity of a collection of molecules and has been reported to increase the odds of addressing a broad range of biological targets[34] (relative to a single-scaffold library).[25,35,47,48]

There is a clear distinction between DOS and traditional combinatorial methods; DOS libraries are usually smaller but consist of molecules that are typically structurally more complex, have a greater variety of core structures (skeletons) and possess richer stereochemical variation.[58] However, the boundary between modern, more considered combinatorial methods and DOS is less clear-cut, and the terms DOS and combinatorial chemistry are often used interchangeably in the literature. Indeed, since DOS incorporates many of the principles of combinatorial chemistry, it is probably best to consider DOS as a more evolved version of traditional combinatorial methodologies.¶ Recently our group has introduced the concept of the '*molecular diversity spectrum*' as a useful qualitative means for comparing the structural diversity associated with a particular molecular collection (Fig. 5). It should therefore be the goal of a DOS to synthesize, in a

¶ It is widely accepted in the literature that the use of a traditional combinatorial approach (diversity around a single scaffold) as a means for structural optimisation *once a biological active molecular skeleton has been identified* is without par; the principle benefit of DOS is in the initial discovery of (potentially novel) biologically active skeletons.[56]

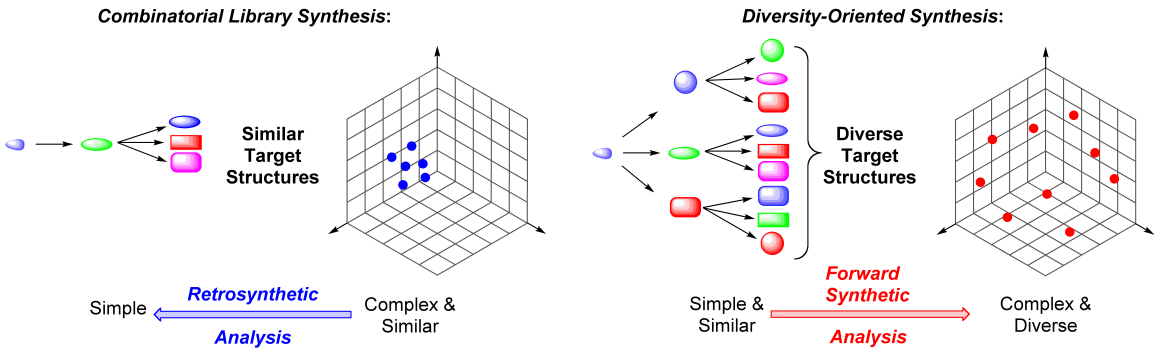


Fig. 4 A comparison of the planning strategies used in a traditional combinatorial synthesis and a DOS together with a visual representation of the chemical space coverage achieved in both cases (*i.e.* focused around a specific point or diverse coverage).

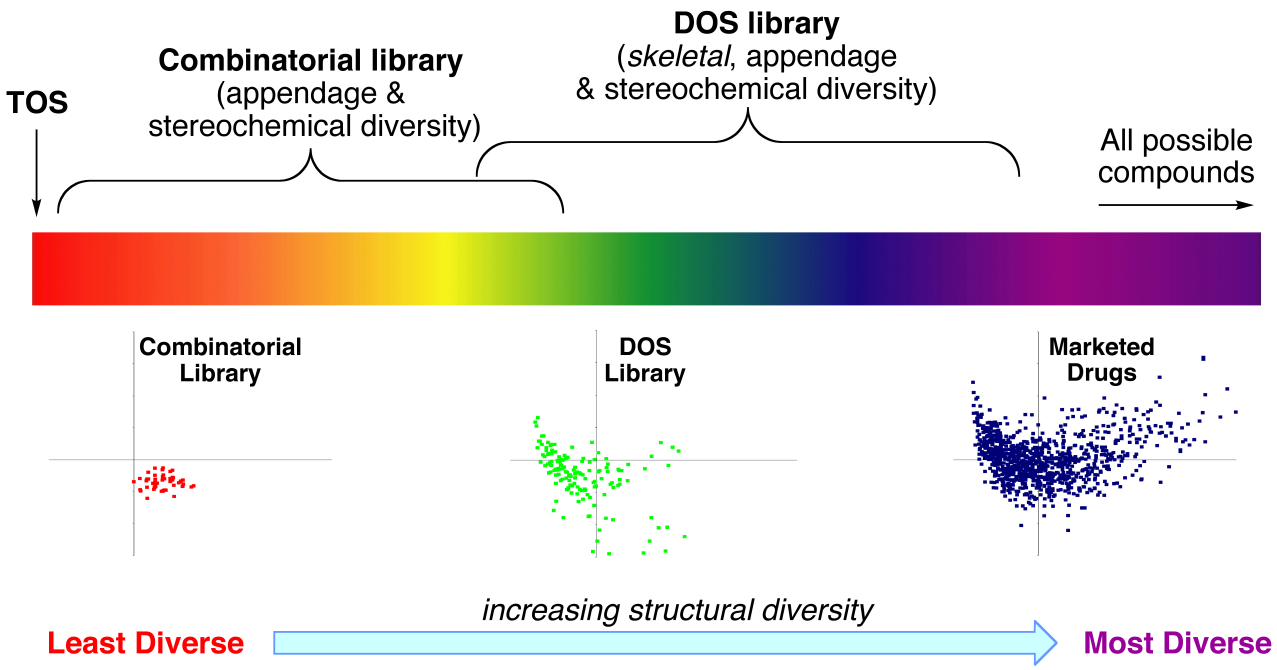


Fig. 5 The '*molecular diversity spectrum*'. 'Diversity' can be viewed as a spectrum ranging from a target-oriented synthesis (TOS) of a specific molecule to the synthesis of all possible compounds (*i.e.* total chemical space coverage); a traditional combinatorial approach and a DOS produce compound collections that sit between these two extremes. Below the diversity spectrum are two diagrams, created using principal component analysis, which show the distribution of representative chemical collections produced by a traditional combinatorial synthesis and a DOS in chemical descriptor space. Also included is a diagram showing the distribution of all known drug molecules in chemical descriptor space. These diagrams illustrate the ability of a DOS library to cover a wider area of chemical space than a traditional combinatorial library. The DOS library also achieves a better coverage of known bioactive chemical space. The nature of these libraries is discussed in more detail later in the text (see sections on 'Achieving skeletal diversity in a DOS', and 'Assessing the diversity of a small molecule library'). More detail regarding the use of principal component analysis is given in Fig. 8.

qualitative sense, collections of small molecules which are as near as possible to the right hand side of this spectrum.[25]

In is important to note that although a DOS aims to access a wide area of chemical space, there is a still a certain degree of bias associated with the compounds produced. This is because the goal of a DOS is to generate compounds that are capable of interacting with biological macromolecules; that is, a DOS is *designed* to access biologically relevant space. Though the boundaries of biological relevant chemical space are not known, and thus there is undoubted value in generating novel molecules from unexplored regions of chemical space, there are nevertheless certain chemical features that are required in a molecule in order for it to be able to interact with a biological system in a useful manner. In the context of antibacterial discovery using DOS, it is important to note that such synthesis design criteria do not include limiting the library to compounds that span the physiochemical property space occupied by know antibiotic compounds (*vide supra*). There are, however, molecular features that are desirable in terms of any further development of the molecule into a therapeutic agent (for example, favourable ADME (absorption, distribution, metabolism and elimination) properties as well as low

toxicity). Consequently, a DOS is *not* simply a random generation of compounds, but a carefully considered process that aims to generate small molecules with diverse molecular structures that are natural product-like and drug-like in terms of their capability to interact with, and modulate, biological systems.

DOS and antibacterial discovery

Since Schreiber's seminal works in the field[9,47,84,89,90] the use of DOS as a method for the generation of small molecule libraries has increased dramatically. Many different general strategies have been developed which aim to address all of these aspects of diversity and the reader is directed to recent review articles prepared by our group and others for a more detailed discussion of this topic.[25,46,86,91] There are numerous examples of novel, biologically useful small molecules which have been discovered through the screening of DOS libraries.[66,72,90,92–94] However DOS is arguably an under-exploited tool in the field of antibacterial discovery and there are very few reports in the literature pertaining to the screening of DOS collections (which incorporate skeletal diversity) against bacterial targets. Therefore for the purposes of this discussion we will highlight some examples from our own laboratories on the use of DOS to discover new antibacterial agents.

The overall aim of a DOS is the generation of a small molecule collection that spans as much bioactive chemical space as possible as this should increase the probability of discovering small molecules with the desired biological properties (in this case, antibacterial activity). As outlined previously, this functional diversity is most efficiently achieved through the efficient incorporation of skeletal (scaffold) diversity in a small molecule library.

Achieving skeletal diversity in a DOS

There are two main methods that have been employed in a DOS context for the generation of skeletal diversity, based upon either the *reagent* (the reagent-based approach) or the *substrate* (the substrate-based approach).[9,95,96]

The substrate-based approach to skeletal diversity

The *substrate-based approach* to skeletal diversity is based around a folding process. It involves the conversion of a collection of substrates, which contain appendages with suitable 'pre-encoded' skeletal information (so-called σ elements), into products have distinct molecular skeletons using a common set of conditions (Fig. 6).[9,85,95] In practice, such methods are usually based around intramolecular reactions that 'pair' strategically positioned functional groups in the substrates, resulting in compounds with diverse skeletons.[96] This approach is exemplified in a recent DOS pathway developed by Morton and co-workers which resulted in the generation of a total of 84 distinct molecular skeletons, representing a significant landmark in the degree of skeletal diversity incorporated in a synthetically-derived small molecule library.[97]

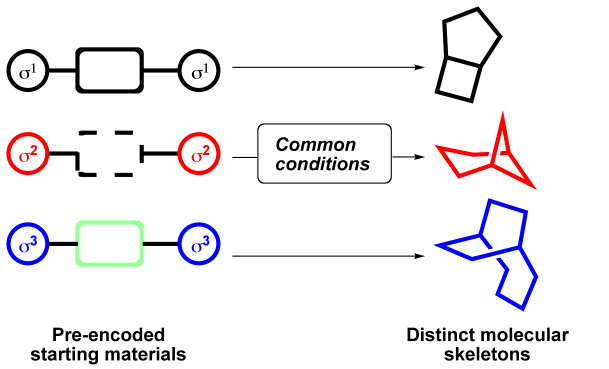


Fig. 6 An illustration of the substrate-based approach to skeletal diversity. A collection of substrates which are each based around a similar core skeleton but which carry different appendages (so-called σ-elements, labelled σ^1–σ^3) is shown. Under a common set of reaction conditions, each substrate is converted into a product having a different molecular skeleton, with the skeletal outcome dependent upon the nature of the σ-elements present in the substrate; *e.g.* 'red' σ^2 elements react in such a fashion so as to generate the molecular skeleton highlighted in red.

The reagent-based approach to skeletal diversity

The *reagent-based approach* to skeletal diversity is a branching synthetic strategy which involves a short series of divergent, complexity generating reactions from a common starting material to generate a collection of compounds with distinct molecular skeletons (Fig. 7).[9]

In practise, reagent-based skeletal diversity is achieved *via* two main methods:[9,25]

1. The use of a *densely functionalised molecule* where different functionalities in the same molecule are transformed by different reagents;

2. The use of a *pluripotent functional group* (*i.e.* one that can participate in a number of different reactions) where exposure of a given molecule to different reagents results in different reactions occurring at the same part (functional group) of the molecule.

The use of a pluripotent functional group strategy

Within the Spring group, we have developed several reagent-based DOS approaches that utilise starting materials with a pluripotent functionality. Critical to the success of a pluripotent DOS is the choice of a synthetically versatile starting material that is capable of undergoing a wide variety of

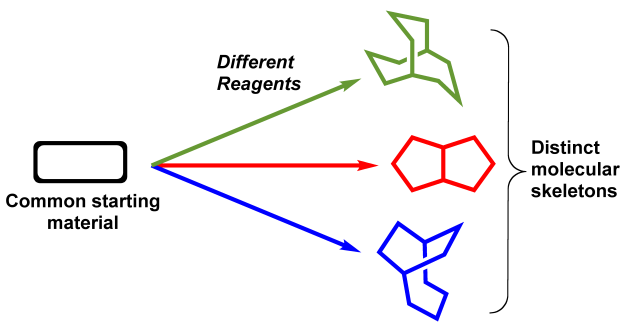


Fig. 7 The reagent-based approach to generating skeletal diversity.

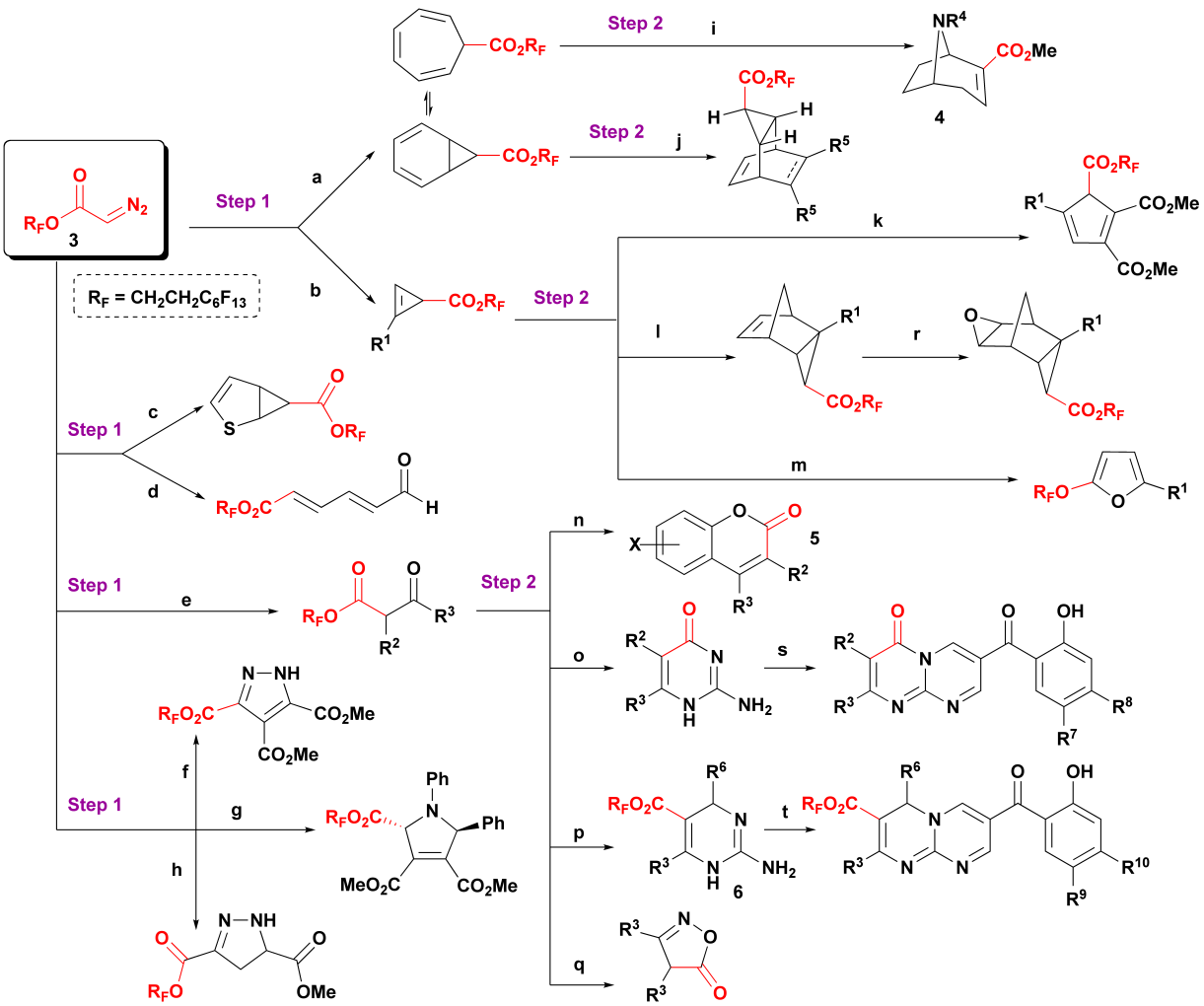
different chemical transformations and thus has the potential to be converted into several products with different molecular skeletons through the variation of reagents alone. These products should themselves be suitable for further diversification, preferably in further complexity-generating and branching reaction sequences; this provides a means to augment the skeletal diversity of the library further and ideally offer a route for the introduction of stereochemical diversity. Variation in the substrates used in these reactions provides scope for the introduction of building block and functional group diversity into the final library.

These considerations are illustrated by two representative case studies from our group. In both cases the overall aim was the generation of a small molecule collection that spanned as much bioactive chemical space as possible in order to increase the probability of discovering small molecules with antibacterial activity. Towards this end, we wanted the libraries to cover known regions of bioactive space, as by definition these are fruitful areas for the discovery of biologically active molecules. Additionally, we wanted to access regions of chemical space that are not covered by known natural products or pharmacologically active compounds, as bioactive molecules from such regions may be more likely to exert their biological activity through a novel mode of action. However, the two DOS approaches differed in terms of the relative degree of emphasis that was placed on exploring these different regions of chemical space.

The first of these methods involved the use of fluorous tagged diazaoacetate **3** as a starting unit in a branching DOS strategy (Scheme 1).[48] In this DOS the main focus was on the efficient generation of compounds that, while structurally novel, were also based around a variety of natural product-like scaffolds; that is, we aimed to explore areas of chemical space 'close' to those occupied by known bioactive small molecules through the incorporation of compounds based known bioactive frameworks. Diazoacetate **3** was identified as an attractive starting unit for two main reasons; (i) diazoacetate compounds exhibit enormous synthetic versatility, permitting the use of a wide variety of different synthetic transformations and (ii) polyfluorocarbon tag technology allowed standard solution phase parallel synthesis methods to be coupled with the benefits of fluorous-based purification protocols,[98,99] thus simplifying the isolation and purification of the library compounds. The main pathways are summarised in Scheme 1. In the first stage of the DOS the reactive diazoacetate functionality was exploited in four main branching reactions (Step 1, Scheme 1, reactions a through to g). The second stage of the DOS involved a series of complexity-generating reactions to diversify these molecular frameworks further, increasing the skeletal diversity of the library (Step 2, Scheme 1, reactions i through to q). In selected cases, a third stage of reactions was carried out (reactions r, s and t) to introduce additional complexity and diversity. Variation in the substrates used in these branching reactions (*i.e.* variation in R^{1-10}) introduced appendage and functional group diversity into the library. In addition, further appendage and functional group diversity was incorporated into the products from these pathways *via* the use of different ester cleavage mechanisms (*i.e* ester hydrolysis, transesterification, transamidation and ester reduction; R replaced by R^{11}, R^{12} *etc.*). Various substrates were used in the branching synthetic routes outlined in Scheme 1 to synthesize a library of 223 compounds based on 30 different core molecular skeletons in no more then four linear synthetic steps.

Several of these molecular skeletons have proven biological relevance. Consequently, the inclusion of molecules based around such scaffolds should allow the DOS library to span regions of (and close to) known bioactive chemical space (*vide infra*). For example, **4** is an ecgonine analogue; ecgonine itself is a member of the tropane alkaloid family of natural products and known bioactive molecules based on similar scaffolds include cocaine and the molecules β-CIT and β-CPT (radiolabelled forms of which have been used to help investigate dopamine uptake *in vivo*).[100,101] Compounds **5** are based around a coumarin skeleton. This privileged scaffold is found in a wide variety of synthetic and naturally occurring compounds that display a range of different biological activities, for example the anti-coagulant warfarin[102] and the rodenticide brodifacoum.[103] Furthermore, the dihydropyrimidone framework present in **6** is a well-established privileged sub-structure; compounds with this core scaffold have been reported as anti-hypertensives and anti-cancer leads and this class of compounds also exhibits anti-viral and anti-inflammatory activity.[62] Compounds based around these privileged scaffolds are desirable to include in a structurally diverse compound collection for use in chemical genetics screens.

The second example of a branching DOS strategy from our own group involves the generation of a library of small molecules starting from a solid-supported phosphonate **7** (Scheme 2).[104] In this DOS, an emphasis was placed upon the exploration of uncharted regions of chemical space through the development of routes to unnatural-product-like (*i.e.* completely novel) molecular scaffolds. In the first step of the DOS (Step 1, Scheme 2), **7** was reacted with a variety of aldehyde building blocks in an *E*-selective Horner–Wadsworth–Emmons reaction (building block diversity) to deliver twelve α,β-unsaturated acyl-imidazolidinones **8**. The pluripotent nature of **8** was exploited in the second step of the DOS (Step 2, Scheme 2) which involved three catalytic enantioselective divergent reaction pathways (stereochemical diversity): (i) [2 + 3] cycloaddition (reaction b); (ii) dihydroxylation (reaction c); and (iii) [4 + 2] cycloaddition (reaction d) to yield a collection of molecules based on three molecular frameworks (skeletal diversity). The next step of the DOS (Step 3, Scheme 2) involved a series of branching reactions to diversify these key branch-point substrates further. For example, the norbornene derivatives **9** (formed in step d) served as versatile intermediates for a series of branching reactions (reactions l to o) to generate five different molecular scaffolds (skeletal diversity). Of particular note was an interesting tandem ring-closing–opening–closing metathesis reaction (reaction o) which created skeletally diverse tricyclic products **10a** (7-5-7) and **10b** (7-5-8). In some cases, a fourth stage of reactions was carried out (Step 4, Scheme 2) to introduce additional complexity and diversity. The imidazolidinone portion of **7** (R group in Scheme 2) allowed



Scheme 1 DOS of a library of small molecules from a simple diazoacetate starting material **3**. Step 1 refers to the first step of the DOS, Step 2 refers to the second step of the DOS. *Reagents and conditions*: (a) C_6H_6, $Rh_2(OCOCF_3)_4$; (b) R^1CCH, $Rh_2(OAc)_4$, CH_2Cl_2; (c) thiophene, $Rh_2(OAc)_4$; (d) furan, $Rh_2(OAc)_4$ then I_2; (e) LDA −78 °C, then R^2COR^3, THF then $Rh_2(OAc)_4$, CH_2Cl_2; (f) DMAD; (g) PhCHO, $PhNH_2$ then DMAD, $Rh_2(OAc)_4$ or PhMe·$[Cu(OTf)]_2$, CH_2Cl_2; (h) methyl acrylate; (i) R^4NH_2, NaOH, H_2O, 180 °C then MeOH, H_2SO_4, 60 °C; (j) dienophile, toluene, reflux; (k) DMAD, toluene, 100 °C; (l) cyclopentadiene, CH_2Cl_2, 0 °C to rt; (m) Grubbs' second-generation catalyst, toluene, ethylene, reflux; (n) phenol derivative, conc. H_2SO_4; (o) guanidine, EtOH, reflux; (p) guanidine, R^6CHO, DMF, 75 °C; (q) NH_2OH, THF, reflux; (r) *m*CPBA, CH_2Cl_2, rt; (s) substituted 3-formyl chromone, EtOH, reflux; (t) substituted 3-formyl chromone, EtOH, reflux. DMAD = dimethyl acetylenedicarboxylate, rt = room temperature.

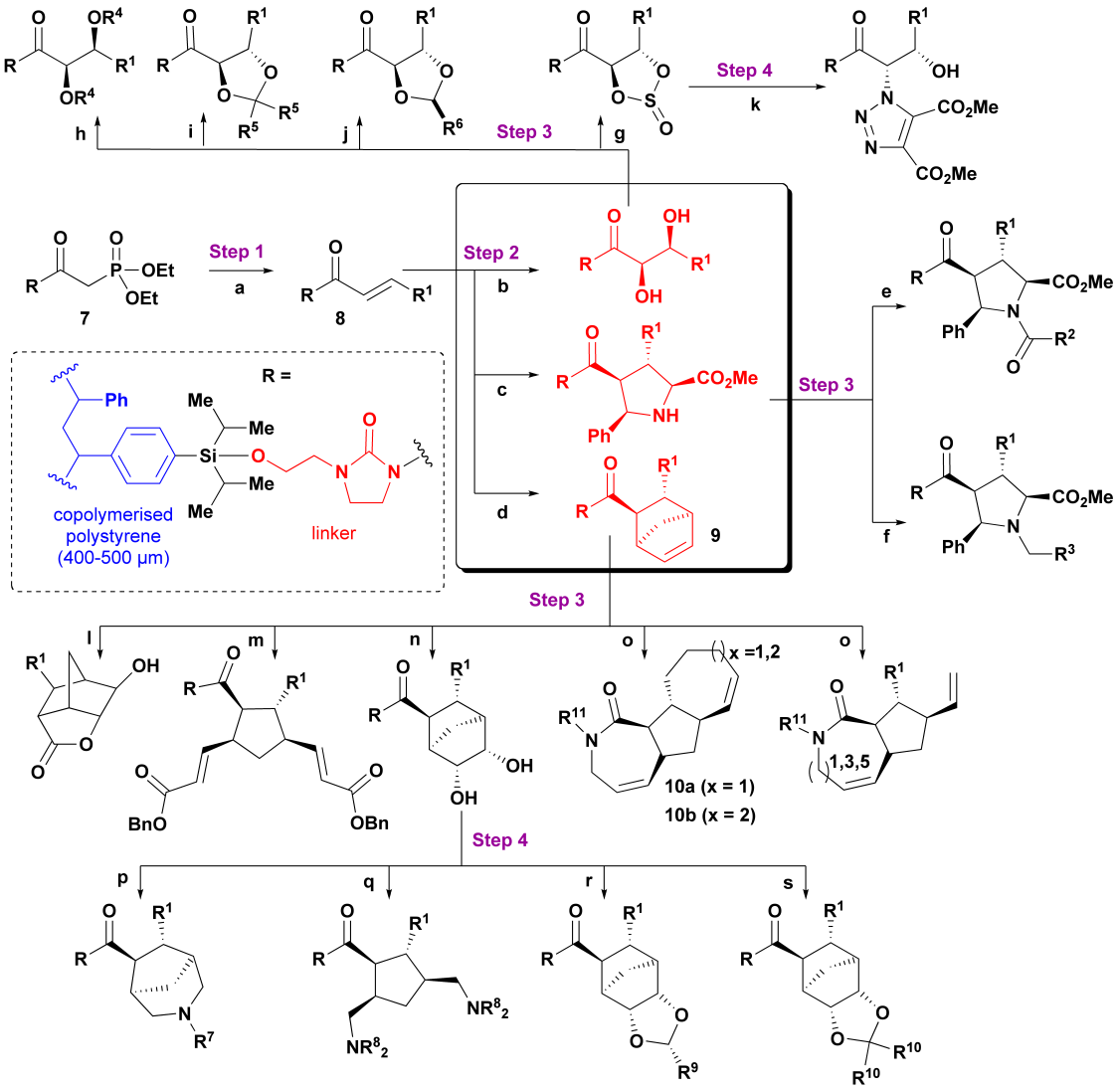
the attachment of compounds at each stage of the synthesis to a novel silyl-polystyrene solid support resin developed previously within the group,[105] which greatly simplified purification during library synthesis. In the final step of the DOS (not shown), the compounds were cleaved off the solid support using a variety of reagents which provided a means for the introduction of further appendage diversity.

Using the chemistry shown in Scheme 2 and a limited number of structurally diverse building blocks, a DOS of 242 small molecules was achieved, which have 18 molecular frameworks among other unique structural features. Many of these frameworks have no known representation in nature, highlighting the capability of this DOS approach to generate products that populate new, unexplored regions of chemical space.

Assessing the diversity of a small molecule library

There is a clear justification for the production of structurally diverse small molcule libraries as a means to improve coverage of bioactive chemical space. However, a fundamental issue when attempting such a synthesis is the subjective nature of diversity itself; that is, how can one compare the overall diversity present in different small molecule collections?[37] Recent years have witnessed significant progress in the development of computational methods that allow an assessment of the diversity present in different chemical collections in a more quantitative fashion.[37,106–108] The goal of these methods is not to provide an absolute measure of diversity, but a relative measure that also agrees to a good extent with chemical intuition.

Within our group we have utilised a computational process for diversity assessment based around the calculation of



Scheme 2 DOS of 242 compounds based of 18 discrete molecular frameworks. *Reagents and conditions*: (a) LiBr, 1,8-diazabicyclo[5.4.0]undec-7-ene, R^1CHO, MeCN; (b) AD-mix α, THF–H_2O (1 : 1); (c) (*R*)-QUINAP, AgOAc, *i*Pr_2NEt, α-imino-ester, THF, −78 to 25 °C; (d) chiral bis(oxazoline), $Cu(OTf)_2$, 3 Å MS, CH_2Cl_2, C_5H_6; (e) R^2COCl, DMAP, pyridine, CH_2Cl_2; (f) R^3CHO, BH_3·pyridine, MeOH; (g) $SOCl_2$, pyridine, CH_2Cl_2, 40 °C; (h) R^4Br, Ag_2O, CH_2Cl_2, 40 °C; (i) $R^5C(O)R^5$, TsOH, DMF, 65 °C; (j) R^6CHO, TsOH, DMF, 65 °C; (k) NaN_3, DMF, 100 °C then DMAD, toluene, 65 °C; (l) *m*CPBA, CH_2Cl_2 then MeOH, 65 °C; (m) $CH_2{=}CHCO_2Bn$, Grubbs' second-generation catalyst, ethylene, toluene, 120 °C; (n) OsO_4, NMO, $CH_3C(O)CH_3$–H_2O (10 : 1); (o) RNH_2, Me_2AlCl, toluene, 120 °C; then NaH, $R^{11}X$, DMF, THF; then toluene, 120 °C, Grubbs' second-generation catalyst, ethylene; (p) $NaIO_4$, THF–H_2O (1 : 1); then R^7NH_2, $NaB(OAc)_3H$, CH_2Cl_2; (q) $NaIO_4$, THF–H_2O (1 : 1); then R^8NHR^8, $NaB(OAc)_3H$, CH_2Cl_2; (r) R^9CHO, DMF, TsOH, 60 °C; (s) $R^{10}C(O)R^{10}$, DMF, TsOH, 60 °C. DMAD = dimethyl acetylenedicarboxylate.

molecular descriptor values followed by principal component analysis (PCA).[109] Though a detailed discussion of this procedure is beyond the scope of this review, the significant features are illustrated in the following case studies taken from our own laboratories. For each compound in the DOS libraries synthesised by Wyatt *et al.* and Thomas *et al.* (Scheme 1 and Scheme 2), the values of 184 different physiochemical and topological chemical descriptor properties (*e.g.* molecular weight, degree of branching, pK_a, charges) were calculated. The data sets produced for each compound were analysed on a per molecule basis using PCA. This method allowed a two-dimensional (2D) visual representation of the distribution of the compounds of the DOS libraries in chemical descriptor space to be derived (Fig. 8). The distribution in chemical space of three 'benchmark' molecule collections was also computed and included in these diagrams. The first of these collections was a focused library which was produced *via* a more traditional combinatorial approach (whereby a common scaffold is decorated with different appendages).[110] The second of these 'benchmark' collections was a sample of all known drug molecules with a similar weight range to the compounds present in the DOS libraries (molecular weight cut-off 650), taken from the MDL Drug Data Repository (MDDR) database.[111] Finally, the chemical space coverage achieved by the 3762 compounds marked as 'antibacterial' in the MDDR database is also included. By this

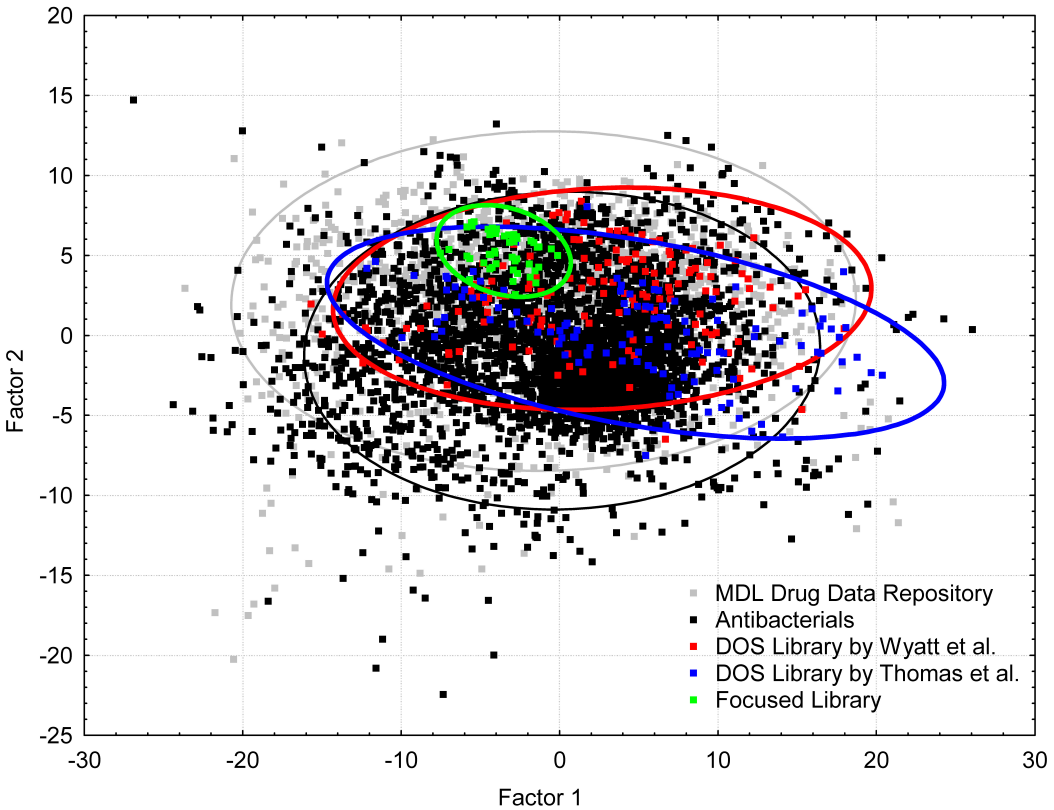


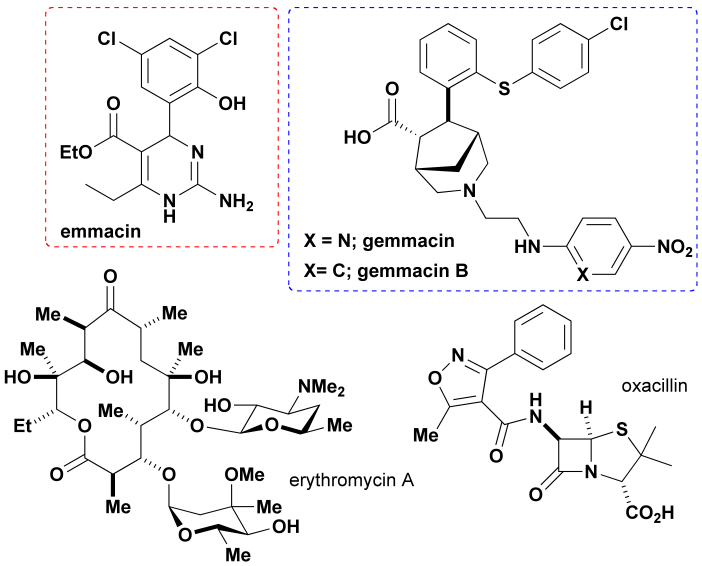
Fig. 8 A 2D visual representation of the distribution of different chemical collections in physicochemical and topological space derived using molecular operating environment (MOE) descriptors followed by principal component analysis (PCA). The DOS libraries synthesized are depicted by red dots (Wyatt *et al.*) and blue dots (Thomas *et al.*). For comparison, a focused library (green dots), the MDL Drug Data Repository (MDDR; grey dots), and antibacterial drugs (black dots) are depicted. PCA is a mathematical algorithm that reduces the dimensionality of a data set while retaining most of the variation in the data set.[109] In the original data sets each compound is described by the values of 184 different variables (molecular descriptors). PCA discovers new variables called principal components which account for the majority of the variation in the data sets. Each compound can then be described by the values of a relatively small number of these principal components. Thus each compound can be represented by a relatively few numbers instead of by values for 184 variables. This allows visually accessible 2D representations of the distribution of the data in physiochemical and topological space to be plotted.[109] Factor 1 and Factor are the first two principal components, which explain the majority of the variation in the data sets.

analysis, the two DOS libraries (high skeletal diversity) were shown to span a larger region of chemical space then that occupied by the 'focused' library (low skeletal diversity). This helps to validate the concept that maximising skeletal diversity in a small molecule library is crucial in terms of maximising overall structural diversity and thus chemical space coverage. The largest coverage of chemical space is achieved by the MDDR sample, which contains a very diverse compilation of compounds derived from a vast number of different sources, both natural and synthetic. However, the DOS libraries achieved a level of chemical space coverage (*i.e.* overall structural diversity) very close to that observed in the MDDR sample. This is despite the fact that the DOS libraries were each derived from a single synthetic pathway, and each library contains significantly fewer compounds than are present in the MDDR sample. This illustrates the value of our DOS approaches to generate structurally diverse products that span a wide area of chemical space in an efficient manner. In addition, the DOS libraries span a large region of chemical space covered by currently known antibacterial compounds, which is an important factor in the context of biological screening.

Biological screening

Computational analyses can be performed to determine if a high degree of structural diversity has been achieved in a small molecule library. However, it is important to emphasise that the ultimate success of any small molecule library is determined by the biological relevance of the compounds it contains; if the small molecule library does not yield hits in a chosen biological screening experiment, it will be deemed to be unsuccessful, no matter how structurally diverse it is. In the context of the DOS examples outlined in Schemes 1 and 2, we were concerned with identifying compounds with antibacterial activity.

Towards this end the DOS libraries of Wyatt *et al.* and Thomas *et al.* were screened in inhibition of proliferation phenotypic assays against three strains of *Staphylococcus aureus*: a methicillin susceptible *S. aureus* (MSSA) and two UK epidemic methicillin-resistant strains (EMRSA-15 and EMRSA-16).[112] These MRSA strains are responsible for the majority of infections with MRSA in the UK and both are resistant to penicillins and erythromycin.[113] This process identified a number of structurally novel compounds that



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

Fig. 9 Structures and activities (MIC_{50}) of emmacin, gemmacin, gemmacin B, erythromycin and oxacillin, which display growth inhibitory activity against methicillin-susceptible and -resistant strains of *S. aureus*. MSSA = methicillin-susceptible *S. aureus*. MIC_{50} = minimum inhibitory concentration required to inhibit the growth of 50% of organisms.

modulated bacterial growth (ranging from inhibition of proliferation to complete abolition of bacterial growth over the time course of the experiment) over a range of concentrations (100 to 10 μM). The most active compound identified in these preliminary screens was named gemmacin after Gemma Thomas, the researcher who first synthesised it (Fig. 9). Further phenotypic screening indicated that gemmacin has a broad range of activity against Gram-positive bacteria. The unusual molecular skeleton and unique structural features, together with the positive *in vitro* results against a variety of bacterial pathogens, highlight that gemmacin provides a new structure for the discovery of critically needed antibiotics.

These screening results demonstrate the utility of a DOS approach for the discovery of new antibacterial agents. These are compounds that are completely artificial, with novel structures never before accessed by nature and which display *useful* biological properties. This clearly illustrates the power of DOS to explore uncharted regions of chemical space that are 'biologically profitable' but have yet so far been ignored by the process of millions of years of natural evolution. In addition, this work helps to validate the usefulness of DOS as a tool in the drug-discovery process.

Though biologically active molecules (so-called 'hits') can be identified through the screening of structurally diverse small molecule collections, optimisation is usually required to transform these 'hits' into 'leads' that possess the right portfolio of on- and off-target activities as well as physiochemical and ADME/toxicity properties. Usually this is achieved by the synthesis of a focused library of analogues of the original bioactive molecule, through variation in appendage, functional group and stereochemical diversity elements. Biological screening of these compounds then allows structure activity relationships (SAR) to be investigated. Recently, Thomas and co-workers have reported such an SAR investigation of the antibacterial compound gemmacin.[114] These studies resulted in the identification of a new compound, gemmacin B, which demonstrated higher levels of antibacterial activity against EMRSA-16 (Fig. 9). Interestingly, these SAR studies indicated that little chemical diversification from the original gemmacin architecture was possible if antibacterial activity was to be retained.

Analogue syntheses and SAR analyses were also carried out on the most active compounds identified in the DOS library of Wyatt *et al.*[115] This work culminated in the discovery of an antibacterial agent named emmacin after Emma Wyatt, the researcher who first synthesised it (Fig. 9). Emmacin is a structurally novel antibacterial agent with an activity against EMRSA-15 and EMRSA-16 comparable to that of gemmacin and superior to that of two other clinically important antibacterial agents (Fig. 9). Both emmacin and gemmacin were found not to exhibit cytotoxic properties in a variety of mammalian-surrogate systems, and also lacked fungicidal and herbicidal properties, demonstrating the capability of the compounds to act as bacteria-selective agents. Such selectivity is of vital importance for any possible further development of these agents.

Target identification

Small molecules that exhibit antibacterial activity can be identified through the phenotypic screening of structurally diverse small molecule collections (*i.e.* forward chemical genetics). However, identification of the biological target(s) that physically interact with the compound of interest[12] (so-called target identification) is notoriously difficult. Such information provides knowledge regarding the underlying mode of action of the antibacterial agent, which is crucial for directed compound optimization and any pre-clinical and clinical development.[116]

Recent years have witnessed the emergence of several different methods for overcoming the problem of target identification, and the reader is directed towards some recently

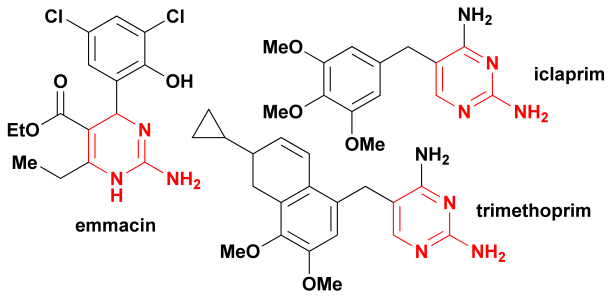


Fig. 10 Comparison of the structures of some known DHFR inhibitors (trimethoprim,[121] iclaprim[122]) with antibacterial activity with the structure of emmacin. The common elements of nitrogen-based heterocyclic frameworks present in each structure are highlighted.

published articles for a more detailed review of this area.[117–120] Unfortunately, no systematic target-identification methodologies of broad utility presently exist. Indeed, target identification has been described as the 'missing link' in chemical genetics, a technical hurdle that limits the methods huge potential.[117] Despite these difficulties, we have had some success in identifying the biological targets of hits from DOS libraries produced within the group, albeit on an *ad hoc* basis.

Target identification of gemmacin was attempted by screening the compound in a range of biological assays designed to identify common antimicrobial modes of action (such as modulation of kinase activity, protein synthesis and ATP synthesis uncoupling).[104] Gemmacin was found to be inactive in all of these assays, but did show activity in an assay designed to detect the generation of reactive oxygen species, which suggested that the compound may act as a cell-membrane disruptor. A membrane disruptor assay carried out with gemmacin did indeed demonstrate its capability to act as a selective disruptor of bacterial cell membranes, suggesting that this is at least one of the modes of action by which the compound exerts its antibacterial effects. Interestingly, membrane disruption is the primary mode of action of antimicrobial peptides such as defensins, gramicidin S, type A lantibiotics, and telavancin, which all have molecular weights greater than 1700 g mol^{-1}.[16] It is intriguing that the significantly smaller molecule gemmacin (539 g mol^{-1}) could have a similar mode of action.

Emmacin was subjected to a battery of biological assays, but was found to be inactive in all of them. A comparison of emmacin with the general structures of some known classes of antibacterials revealed structural similarities with a number of compounds known to act as inhibitors of the enzyme dihydrofolate reductase (DHFR) (Fig. 10).

DHFR is a ubiquitous enzyme in eukaryotic and prokaryotic cells which catalyses the reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate using NADPH as a cofactor.[123,124] Tetrahydrofolate is involved in the biosynthesis of nucleotide bases of DNA. Therefore inhibition of the DHFR enzyme blocks DNA synthesis, thereby arresting cell growth.[125] A DHFR inhibition assay indicated that emmacin was a prokaryote-selective, uncompetitive and reversible inhibitor of the EMRSA-16 DHFR enzyme $DfrB_{EMRSA16}$. Despite recent research into the discovery of new bacterial DHFR inhibitors,[121,122,125–127] the enzyme is still viewed as an underexploited target in the antibacterial field.[121] We could find no examples of substituted *dihydro*pyrimidine compounds, of the type exemplified in emmacin, having been applied to this therapeutic mode of action. Therefore, to the best of our knowledge, emmacin represents the first member of a new structural sub-class of bacterial-selective DHFR inhibitors.

Conclusions

Small molecule libraries produced *via* a DOS, which aim to display a high degree of structural and thus functional (biological) diversity, have proven to be valuable in chemical genetics experiments for the discovery of molecules that can modulate the activities of biological systems in a useful fashion. However, only recently has the power of this approach been directed towards the discovery of critically needed new antibacterial agents. Work in this area has already yielded positive results and offers great potential for future success, particularly in the discovery of antibacterial agents that operate *via* novel modes of action and are thus not susceptible to currently existing resistance mechanisms. However, continued improvements in each of the three elements of forward chemical genetics (namely library design and synthesis, biological screening and target identification) will be required if the goal of using this approach to regularly discover effective small molecule antibacterial agents which are suitable for further development is to be achieved.

As a final point, it may come as a surprise to some readers that we have made no attempt to patent *any* compounds generated in our DOS libraries, including those with the highest levels of antibacterial activity. This was a very deliberate decision. We openly encourage the distribution of this data in the hope that our bioactive hits may serve as inspirational starting points for subsequent development into more potent and selective antibacterial agents by others. We firmly believe that it is only through the intense collaboration of researchers from many disciplines, both industrial and academic, that the tide of the battle against infectious diseases will start to turn in our favour.

Acknowledgements

We gratefully acknowledge financial support from the EPSRC, BBSRC, MRC and Newman Trust. Also, we thank the referees for useful comments and suggestions to improve this article.

Notes and references

1 C. Walsh and G. Wright, *Chem. Rev.*, 2005, **105**, 391–394.
2 P. Fernandes, *Nat. Biotechnol.*, 2006, **24**, 1497–1503.
3 E. D. Brown and G. D. Wright, *Chem. Rev.*, 2005, **105**, 759–774.
4 R. E. W. Hancock, *Nat. Rev. Drug Discovery*, 2007, **6**, 28.
5 D. J. Payne, M. N. Gwynn, D. J. Holmes and D. L. Pompliano, *Nat. Rev. Drug Discovery*, 2007, **6**, 29–40.
6 M. L. Katz, L. V. Mueller, M. Polyakov and S. F. Weinstock, *Nat. Biotechnol.*, 2006, **24**, 1529–1531.
7 Editorial, *Nat. Biotechnol.*, 2006, **24**, 1489.
8 D. Hopwood, *Nat. Rev. Drug Discovery*, 2007, **6**, 8–12.
9 M. D. Burke and S. L. Schreiber, *Angew. Chem., Int. Ed.*, 2004, **43**, 46–58.
10 S. L. Schreiber, *Chem. Eng. News*, 2003, **81**(9), 51–61.

11 R. Pathania and E. D. Brown, *Biochem. Cell Biol.*, 2008, **86**, 111–115.
12 T. A. Walsh, *Pestic. Manage. Sci.*, 2007, **63**, 1165–1171.
13 D. P. Walsh and Y. T. Chang, *Chem. Rev.*, 2006, **106**, 2476–2530.
14 D. R. Spring, *Chem. Soc. Rev.*, 2005, **34**, 472–482.
15 S. L. Schreiber, *Bioorg. Med. Chem.*, 1998, **6**, 1127–1152.
16 C. T. Walsh, *Antibiotics: Actions, Origins, Resistance*, ASM Press, Washington, 2003.
17 F. von Nussbaum, M. Brands, B. Hinzen, S. Weigand and D. Habich, *Angew. Chem., Int. Ed.*, 2006, **45**, 5072–5129.
18 D. Habich and F. von Nussbaum, *ChemMedChem*, 2006, **1**, 951.
19 K. F. Blount and R. R. Breaker, *Nat. Biotechnol.*, 2006, **24**, 1558–1564.
20 U. Theuretzbacher and J. H. Toney, *Curr. Opin. Invest. Drugs*, 2006, **7**, 158–166.
21 J. Prieto, A. Calvo and M. L. Gomez-Lus, *J. Antimicrob. Chemother.*, 2002, **50**, 7–12.
22 L. B. Boyd, R. L. Atmar, G. L. Randall, R. J. Hamill, D. Steffen and L. Zechiedrich, *BMC Infect. Dis.*, 2008, **8**.
23 F. A. A. Amer, E. M. El-Behedy and H. A. Mohtady, *Biotechnol. Mol. Biol. Rev.*, 2008, **3**, 46–57.
24 L. L. Silver, *Nat. Rev. Drug Discovery*, 2007, **6**, 41–55.
25 R. J. Spandl and D. R. Spring, *Org. Biomol. Chem.*, 2008, **6**, 1149–1158.
26 M. Kaiser, S. Wetzel, K. Kumar and H. Waldmann, *Cell. Mol. Life Sci.*, 2008, **65**, 1186–1201.
27 A. Bender, S. Fergus, W. R. J. D. Galloway, F. G. Glansdorp, D. M. Marsden, R. L. Nicholson, R. J. Spandl, G. L. Thomas, E. E. Wyatt, R. C. Glen and D. R. Spring, *Diversity Oriented Synthesis: A Challenge for Synthetic Chemists*. Book chapter in *Chemical Genomics Small Molecule Probes to Study Cellular Function, Ernst Schering Research Foundation Workshop 58*, Springer, New York, 2006, pp. 47–60.
28 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 2001, **46**, 3–26.
29 C. Lipinski and A. Hopkins, *Nature*, 2004, **432**, 855–861.
30 A. Leo, C. Hansch and D. Elkins, *Chem. Rev.*, 1971, **71**, 525.
31 D. S. Tan, *Nat. Chem. Biol.*, 2005, **1**, 74–84.
32 D. R. Spring, *Org. Biomol. Chem.*, 2003, **1**, 3867–3870.
33 S. J. Haggarty, *Curr. Opin. Chem. Biol.*, 2005, **9**, 296–303.
34 M. K. Schawrz and W. H. B. Sauer, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 987–1003.
35 J. P. Kennedy, L. Williams, T. M. Bridges, R. N. Daniels, D. Weaver and C. W. Lindsley, *J. Comb. Chem.*, 2008, **10**, 345–354.
36 E. Estrada and E. Uriarte, *Curr. Med. Chem.*, 2001, **8**, 1573–1588.
37 S. Fergus, A. Bender and D. R. Spring, *Curr. Opin. Chem. Biol.*, 2005, **9**, 304–309.
38 T. I. Oprea and J. Gottfries, *J. Comb. Chem.*, 2001, **3**, 157–166.
39 T. I. Oprea, *Curr. Opin. Chem. Biol.*, 2002, **6**, 384–389.
40 C. M. Dobson, *Nature*, 2004, **432**, 824–828.
41 M. Johnson and G. M. Maggiora, *Concepts and Applications of Molecular Similarity*, Wiley, New York, 1990.
42 R. D. Brown and Y. Martin, *C. J. Chem. Inf. Comput. Sci.*, 1996, **36**, 572–584.
43 M. A. Koch and H. Waldmann, *Drug Discovery Today*, 2005, **10**, 471–483.
44 J. M. Mitchell and J. T. Shaw, *Angew. Chem., Int. Ed.*, 2006, **45**, 1722–1726.
45 L. V. Adriaenssens, C. A. Austin, M. Gibson, D. Smith and R. C. Hartley, *Eur. J. Org. Chem.*, 2006, 4998–5001.
46 G. L. Thomas, E. E. Wyatt and D. R. Spring, *Curr. Opin. Drug Discovery Dev.*, 2006, **9**, 700–712.
47 M. D. Burke, E. M. Berger and S. L. Schreiber, *Science*, 2003, **302**, 613–618.
48 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.
49 A. A. Shelat and R. K. Guy, *Nat. Chem. Biol.*, 2007, **3**, 442–446.
50 A. L. Hopkins, J. S. Mason and J. P. Overington, *Curr. Opin. Struct. Biol.*, 2006, **16**, 127–136.
51 C. Cordier, D. Morton, S. Murrison, A. Nelson and C. O'Leary-Steele, *Nat. Prod. Rep.*, 2008, **25**, 719–737.
52 J. Clardy and C. Walsh, *Nature*, 2004, **432**, 829–837.
53 M. Pucheault, *Org. Biomol. Chem.*, 2008, **6**, 424–432.
54 G. Schneider and K. Grabowski, *Curr. Chem. Biol.*, 2007, **1**, 115–127.
55 M. S. Butler, *J. Nat. Prod.*, 2004, **67**, 2141–2153.
56 D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2007, **70**, 461–477.
57 C. L. Waller, *J. Comput.-Aided Mol. Des.*, 2002, **16**, 299–300.
58 S. Borman, *Chem. Eng. News: Sci. Technol.*, 2004, 32–40.
59 G. Wess, *Drug Discovery Today*, 2002, **7**, 533–535.
60 A. M. Rouhi, *Chem. Eng. News*, 2003, **81**.
61 R. O'Shea and H. E. Moser, *J. Med. Chem.*, 2008, **51**, 2871–2878.
62 D. A. Horton, G. T. Bourne and M. L. Smythe, *Chem. Rev.*, 2003, **103**, 893–930.
63 K. C. Nicolaou, J. A. Pfefferkorn, A. J. Roecker, G. Q. Cao, S. Barluenga and H. J. Mitchell, *J. Am. Chem. Soc.*, 2000, **122**, 9939–9953.
64 R. W. DeSimone, K. S. Currie, S. A. Mitchell, J. W. Darrow and D. A. Pippin, *Comb. Chem. High Throughput Screening*, 2004, **7**, 473–493.
65 B. E. Evans, K. E. Rittle, M. G. Bock, R. M. Dipardo, R. M. Freidinger, W. L. Whitter, G. F. Lundell, D. F. Veber, P. S. Anderson, R. S. L. Chang, V. J. Lotti, D. J. Cerino, T. B. Chen, P. J. Kling, K. A. Kunkel, J. P. Springer and J. Hirshfield, *J. Med. Chem.*, 1988, **31**, 2235–2246.
66 S. K. Ko, H. J. Jang, E. Kim and S. B. Park, *Chem. Commun.*, 2006, 2962–2964.
67 B. T. Gregg, D. O. Tymoshenko, D. A. Razzano and M. R. Johnson, *J. Comb. Chem.*, 2007, **9**, 507–512.
68 J. S. Potuzak, S. B. Moilanen and D. S. Tan, *J. Am. Chem. Soc.*, 2005, **127**, 13796–13797.
69 S. B. Moilanen, J. S. Potuzak and D. S. Tan, *J. Am. Chem. Soc.*, 2006, **128**, 1792–1793.
70 S. Y. Shang, H. Iwadare, D. E. Macks, L. M. Ambrosini and D. S. Tan, *Org. Lett.*, 2007, **9**, 1895–1898.
71 J. F. Liu, C. J. Wilson, P. Ye, K. Sprague, K. Sargent, Y. Si, G. Beletsky, D. Yohannes and S. C. Ng, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 686–690.
72 B. C. Goess, R. N. Hannoush, L. K. Chan, T. Kirchhausen and M. D. Shair, *J. Am. Chem. Soc.*, 2006, **128**, 5391–5403.
73 R. Balamurugan, F. J. Dekker and H. Waldmann, *Mol. Biosyst.*, 2005, **1**, 36–45.
74 R. Breinbauer, I. R. Vetter and H. Waldmann, *Angew. Chem., Int. Ed.*, 2002, **41**, 2879–2890.
75 X. Hea, A. Alian and P. R. O. de Montellano, *Bioorg. Med. Chem.*, 2007, **15**, 6649–6658.
76 M. Berkheij, L. van der Sluis, C. Sewing, D. J. den Boer, J. W. Terpstra, H. Hiemstra, W. I. I. Bakker, A. van den Hoogenband and J. H. van Maarseveen, *Tetrahedron Lett.*, 2005, **46**, 2369–2371.
77 G. C. Look, C. Vacin, T. M. Dias, S. Ho, T. H. Tran, L. L. Lee, C. Wiesner, F. Fang, A. Marra, D. Westmacott, A. E. Hromockyj, M. M. Murphy and J. R. Schullek, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 1423–1426.
78 D. M. Schnur, M. A. Hermsmeier and A. J. Tebben, *J. Med. Chem.*, 2006, **49**, 2000–2009.
79 I. R. Correa, A. Noren-Muller, H. D. Ambrosi, S. Jakupovic, K. Saxena, H. Schwalbe, M. Kaiser and H. Waldmann, *Chem.–Asian J.*, 2007, **2**, 1109–1126.
80 A. Noren-Muller, I. Reis-Correa, 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.
81 O. Schwarz, S. Jakupovic, H. D. Ambrosi, L. O. Haustedt, C. Mang and L. Muller-Kuhrt, *J. Comb. Chem.*, 2007, **9**, 1104–1113.
82 O. Barun, K. Kumar, S. Sommer, A. Langerak, T. U. Mayer, O. Muller and H. Waldmann, *Eur. J. Org. Chem.*, 2005, 4773–4788.
83 K. C. Nicolaou, R. Hughes, J. A. Pfefferkorn, S. Barluenga and A. J. Roecker, *Chem.–Eur. J.*, 2001, **7**, 4280–4295.
84 S. L. Schreiber, *Science*, 2000, **287**, 1964–1969.
85 M. D. Burke, E. M. Berger and S. L. Schreiber, *J. Am. Chem. Soc.*, 2004, **126**, 14095–14104.
86 M. D. Burke and G. Lalic, *Chem. Biol.*, 2002, **9**, 535–541.
87 J. K. Mishra and G. Panda, *J. Comb. Chem.*, 2007, **9**, 321–338.
88 S. J. Taylor, A. M. Taylor and S. L. Schreiber, *Angew. Chem., Int. Ed.*, 2004, **43**, 1681–1685.

89 D. R. Spring, S. Krishnan and S. L. Schreiber, *J. Am. Chem. Soc.*, 2000, **112**, 5656–5657.
90 D. R. Spring, S. Krishnan, H. E. Blackwell and S. L. Schreiber, *J. Am. Chem. Soc.*, 2002, **124**, 1354–1363.
91 R. J. Spandl, M. Diaz-Gavilan, K. M. G. O'Connell, G. L. Thomas and D. R. Spring, *Chem. Rec.*, 2008, **8**, 129–142.
92 H. E. Pelish, N. J. Westwood, Y. Feng, T. Kirchhausen and M. D. Shair, *J. Am. Chem. Soc.*, 2001, **123**, 6740–6741.
93 G. D. Geske, R. J. Wezeman, A. P. Siegel and H. E. Blackwell, *J. Am. Chem. Soc.*, 2005, **127**, 12762–12763.
94 A. N. Koehler, A. F. Shamji and S. L. Schreiber, *J. Am. Chem. Soc.*, 2003, **125**, 8420–8421.
95 H. Oguri and S. L. Schreiber, *Org. Lett.*, 2005, **7**, 47–50.
96 T. E. Nielsen and S. L. Schreiber, *Angew. Chem., Int. Ed.*, 2008, **47**, 48–56.
97 (*a*) D. Morton, S. Leach, C. Cordier, S. Warriner and A. Nelson, *Angew. Chem., Int. Ed.*, 2009, **48**, 104–109; (*b*) W. R. J. D. Galloway, M. Diaz-Gavilan, A. Isidro-Llobet and D. R. Spring, *Angew. Chem., Int. Ed.*, 2009, **48**, 1194–1196.
98 D. Curran and Z. Y. Luo, *Green Chem.*, 2001, **3**, G39–G39.
99 W. Zhang, *Tetrahedron*, 2003, **59**, 4475–4489.
100 E. K. Shaya, U. Scheffel, R. F. Dannals, G. A. Ricaurte, F. I. Carroll, H. N. Wagner, Jr., M. J. Kuhar and D. F. Wong, *Synapse (N. Y.)*, 1992, **10**, 169–172.
101 U. Scheffel, J. W. Boja and M. J. Kuhar, *Synapse (N. Y.)*, 1989, **4**, 390–392.
102 A. M. Holbrook, J. A. Pereira, R. Labiris, H. McDonald, J. D. Douketis, M. Crowther and P. S. Wells, *Arch. Int. Med.*, 2005, **165**, 1095–1106.
103 P. S. van Heerden, B. C. B. Bezuidenhoudt and D. Ferreira, *J. Chem. Soc., Perkin Trans. 1*, 1997, 1141–1146.
104 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.
105 G. L. Thomas, M. Ladlow and D. R. Spring, *Org. Biomol. Chem.*, 2004, **2**, 1679–1681.
106 J. J. Perez, *Chem. Soc. Rev.*, 2005, **34**, 143–152.
107 A. Bender and R. C. Glen, *Org. Biomol. Chem.*, 2004, **2**, 3204–3218.
108 A. H. Lipkus, Q. Yuan, K. A. Lucas, S. A. Funk, W. F. Bartelt, R. J. Schenck and A. J. Trippe, *J. Org. Chem.*, 2008, **73**, 4443–4451.
109 M. Ringner, *Nat. Biotechnol.*, 2008, **26**, 303–304.
110 R. Faghih, W. Dwight, J. B. Pan, G. B. Fox, K. M. Krueger, T. A. Esbenshade, J. M. McVey, K. Marsh, Y. L. Bennani and A. A. Hancock, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 1325–1328.
111 MDL Drug Data Report, http://www.mdli.com.
112 P. C. Moore and J. A. Lindsay, *J. Med. Microbiol.*, 2002, **51**, 516–521.
113 A. P. Johnson, H. M. Aucken, S. Cavendish, M. Ganner, M. C. J. Wale, M. Warner and D. M. Livermore, *J. Antimicrob. Chemother.*, 2001, **48**, 141–156.
114 A. Robinson, G. L. Thomas, R. J. Spandl, M. Welch and D. R. Spring, *Org. Biomol. Chem.*, 2008, **6**, 2978–2981.
115 E. E. Wyatt, W. R. J. D. Galloway, G. L. Thomas, M. Welch, O. Loiseleur, A. T. Plowright and D. R. Spring, *Chem. Commun.*, 2008, 4962–4964.
116 C. P. Hart, *Drug Discovery Today*, 2005, **10**, 513–519.
117 L. Burdine and T. Kodadek, *Chem. Biol.*, 2004, **11**, 593–597.
118 Y. H. Ahn and Y. T. Chang, *Acc. Chem. Res.*, 2007, **40**, 1025–1033.
119 Y. K. Kim and Y. T. Chang, *Mol. Biosyst.*, 2007, **3**, 392–397.
120 C. C. Wong, K. W. Cheng, Q. Y. He and F. Chen, *Proteomics Clin. Appl.*, 2008, **2**, 338–354.
121 P. C. Wyss, P. Gerber, P. G. Hartman, C. Hubschwerlen, H. Locher, H. P. Marty and M. Stahl, *J. Med. Chem.*, 2003, **46**, 2304–2312.
122 P. Schneider, S. Hawser and K. Islam, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 4217–4221.
123 N. V. Kovalevskaya, Y. D. Smurnyy, B. Birdsall, J. Feeney and V. I. Polshakov, *Pharm. Chem. J.*, 2007, **41**, 350–353.
124 J. M. Blaney, C. Hansch, C. Silipo and A. Vittoria, *Chem. Rev.*, 1984, **84**, 333–407.
125 S. Hawser, S. Lociuro and K. Islam, *Biochem. Pharmacol.*, 2006, **71**, 941–948.
126 L. A. Sorbera, J. Castaner and X. Rabasseda, *Drugs Future*, 2004, **29**, 220–225.
127 H. H. Locher, H. Schlunegger, P. G. Hartman, P. Angehrn and R. L. Then, *Antimicrob. Agents Chemother.*, 1996, **40**, 1376–1381.
128 B. R. Stockwell, *Nat. Rev. Genet.*, 2000, **1**, 116–125.
129 S. L. Schreiber, *Nat. Chem. Biol.*, 2005, **1**, 64–66.

