CHAPTER 4

Diversity-Oriented Synthesis

Warren R. J. D. Galloway
Richard J. Spandl
Andreas Bender
Gemma L. Thomas
Monica Diaz-Gavilan
Kieron M. G. O'Connell
David R. Spring

1.1. INTRODUCTION

Chemical genetics describes the use of molecules as "chemical probes" to investigate biological systems [1–3]. In contrast with traditional genetics, in which gene knockouts on the level of the DNA are used, chemical genetics uses biologically active small molecules to directly attenuate the corresponding biological macromolecular (usually protein) product. Thus, the ready availability of bioactive small molecules is of crucial importance in chemical genetics studies. Such small molecules can be identified by screening compound collections (libraries) in suitably designed assays. This chapter describes the use of *diversity-oriented synthesis* (DOS) to prepare structurally diverse small molecule libraries. Structurally diverse libraries show a greater variety in not only their physiochemical properties but also, and of most relevance here, in their biological activities. Herein we describe some of the most effective strategies that have been used in DOS library design and preparation.

1.2. SMALL MOLECULES, CHEMICAL GENETICS, AND CHEMICAL GENOMICS

Chemical genetics experiments can be performed in either a forward or a reverse sense (Figure 4.1). The first step of both approaches requires the identification of a small molecule that either 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.

Chemical genomics has been defined as the search for selective small molecule modulators of each function of all gene products (e.g., proteins) – that is, the application of reverse chemical genetics on a genome-wide scale [2]. Possession of such molecules would allow the systematic exploration and perturbation of biological systems with obvious potential for the development of improved chemotherapeutic treatments [2]. The challenges presented by a chemical genomics approach are daunting. It has been estimated that roughly 10% of the human genome (approximately thirty thousand genes) encode proteins that can bind to "drug-like" compounds [4]. However, small molecule partners for only approximately one thousand members of this total "druggable" proteome have been identified [5]. Fewer still can be considered as specific in their interaction. Clearly, there exists a significant need for the discovery of many more biologically active small molecules that are capable of selectively modulating gene product function.

One of the fundamental considerations of chemical biology is what type of compounds should be synthesized and employed in biochemical or biological screenings [6]. Ultimately, this is determined by the requirements placed on the compounds. 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 regarding molecular characteristics that are desirable, such as molecular size, shape, and tolerable functional groups [5, 7]. However, these demands are considerably different from those placed on a compound required for cell-based or in vitro assays [5, 8].

Biologically active small molecules can be identified by screening libraries of compounds in either phenotypic or

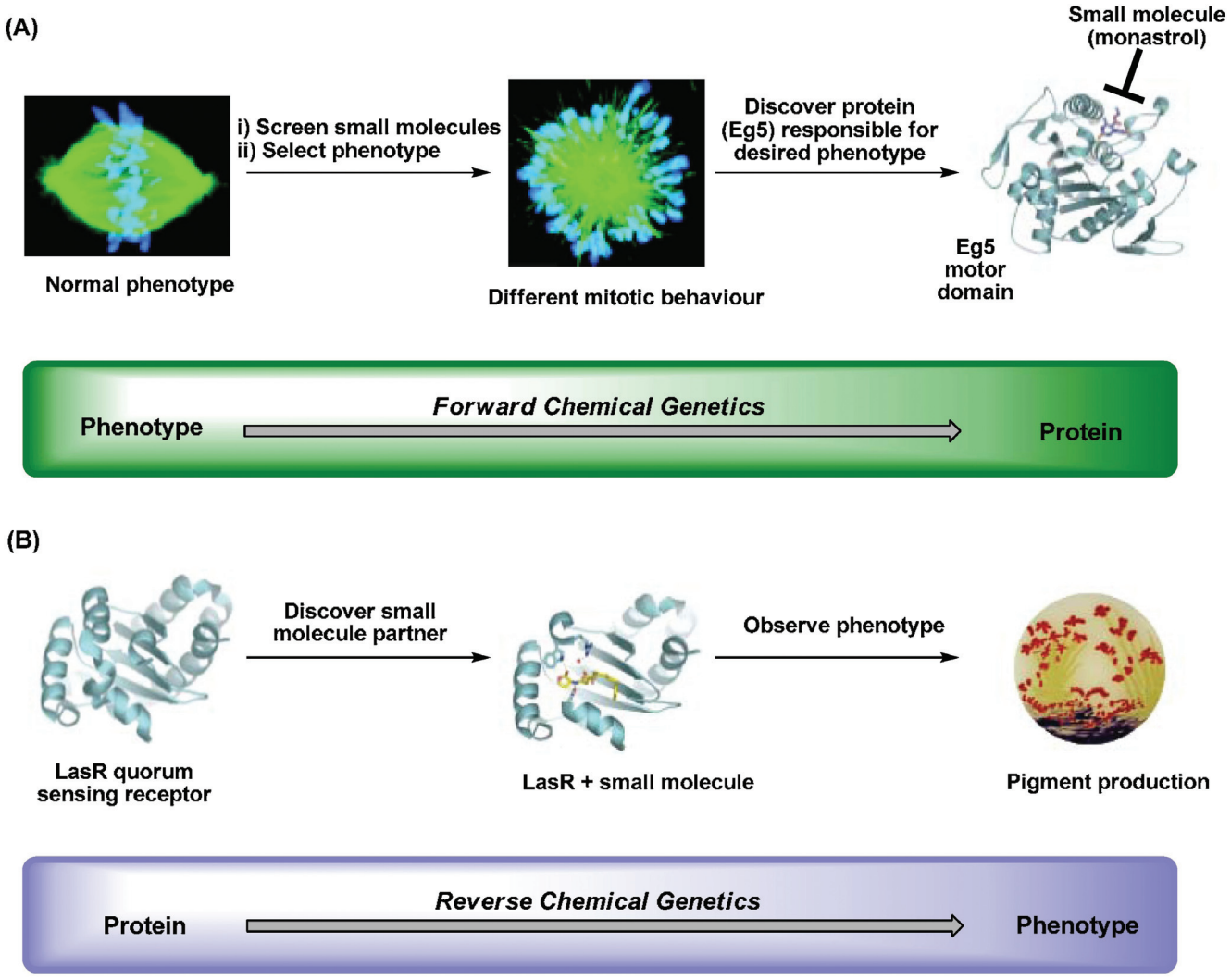


FIGURE 4.1: ***(A)*** *In a forward chemical genetics approach, a library of small molecules is screened to identify those that that induce a desired phenotypic effect (e.g., different mitotic behavior) on the biological system under investigation. Once a suitable small molecule has been identified, further investigation allows the gene product (i.e., protein) with which the molecule interacts to be discovered.* ***(B)*** *Reverse chemical genetics involves the use of small molecules against a known protein of interest (e.g., the LasR quorum sensing receptor protein). A small molecule that binds to the protein is identified, and the phenotypic effect induced by the action of this small molecule helps to define the role of its protein partner.*

protein-binding assays. In the latter case, where the target protein is known (e.g., in a reverse chemical genetics experiment), libraries containing compounds synthesized in a "hypothesis-driven" fashion – that is, designed to interact with a specific target – are more appropriate. Such compounds are usually selected/designed based on knowledge of the target's structure or the structure of known natural ligands [9]. In phenotypic assays, in which the precise nature of the eventual biological target is unknown, the selection criterion for small molecules is complicated dramatically. In these situations, the successful identification of "useful" biologically active small molecules may be aided by screening functionally (biologically) diverse compound libraries, because 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 [10, 11]. Such an approach may also be advantageous in the case of a reverse chemical genetic experiment in which the precise structural features of the biological target of interest are unknown (e.g., the structure of the target protein has not yet been fully characterized), or known natural ligands are unselective and interact with more than one protein target, meaning that closely related analogs may not be suitable for use.

The functional (biological) diversity of a library of small molecules has been shown to be directly correlated with its structural diversity, which in turn is related to the amount of *chemical space* the library occupies [11, 12].

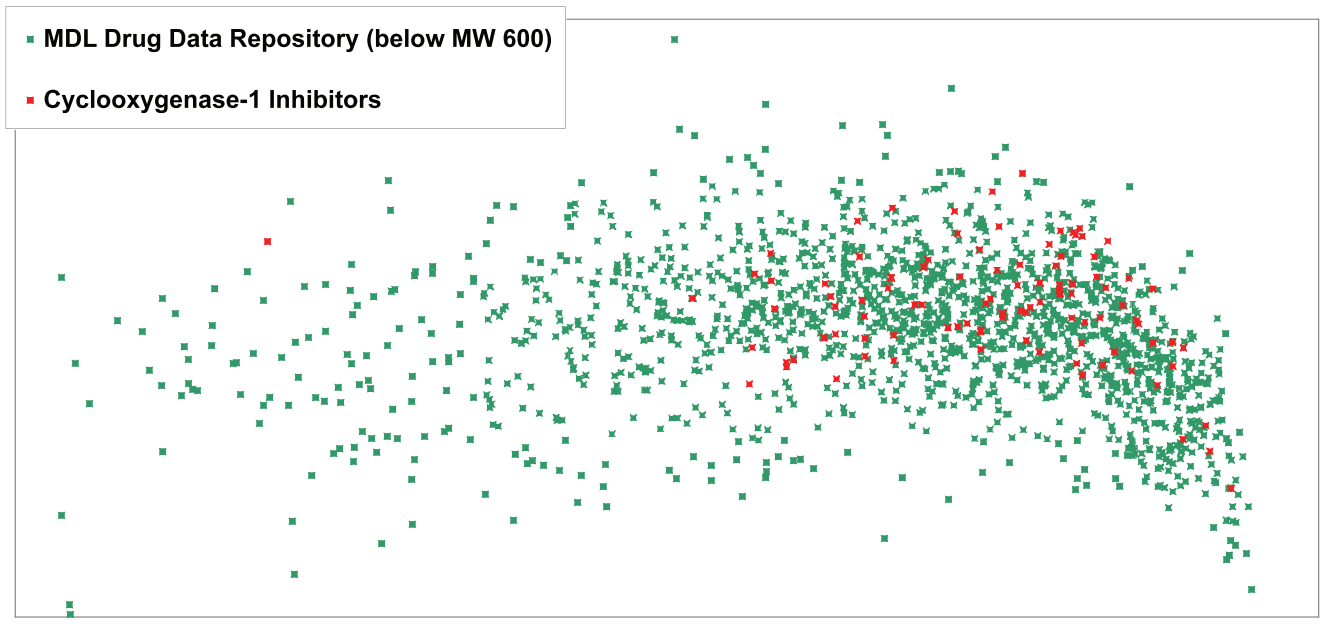


***FIGURE 4.2:** The distribution in chemical space of cyclooxygenase-1 inhibitors (red squares) and a subset of compounds from the MDL Drug Data Repository (MDDR) database (green squares). Once the chemical descriptors have been defined and calculated for each compound, this information can be condensed using a mathematical process known as principle component analysis (PCA). This allows for the construction of two-dimensional (2D) or three-dimensional (3D) displays that are accessible to human interpretation. This 2D visual representation shows that the cyclooxygenase-1 inhibitors populate a broad region of chemical space on the background of the MDDR compounds. In this diagram, each compound is plotted at a discrete point in chemical space. MW = molecular weight.*

1.3. STRUCTURAL DIVERSITY AND THE CONCEPT OF CHEMICAL SPACE

Chemicals can be characterized by a wide range of physiochemical and topological *descriptors* that 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 [13–16]. *Chemical space* is a 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 [17]. 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" [17].

A specific molecule will reside at a discrete point in chemical space because of its unique combination of molecular descriptor values. The structural features of a collection of molecules will therefore influence the distribution of the molecules in chemical space (Figure 4.2). Therefore, it follows that the more structurally diverse the library, the more chemical space it interrogates. Maximizing the structural diversity and thus chemical space coverage of a library should, in turn, increase its overall functional (biological) diversity; as "molecular shape is intrinsically linked to biological activity, the greater the structural diversity in a library, the better the odds of identifying ligands for a broad range of targets" [12].

1.4. BIOLOGICALLY RELEVANT CHEMICAL SPACE

The degree of overlap between total chemical space and biologically relevant chemical space is somewhat of a contentious issue and subject to much debate in the literature [5, 6, 18]. 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 [5]. What it not known is the size of this region in comparison to total chemical space – that is, whether the biologically relevant region is "small" and most of the chemical universe is "empty" (containing no therapeutically interesting compounds) [5]. In other words, are the regions of chemical space defined by natural products and known drugs the best or most fertile regions for discovering small molecules, or is their scope for discovering biologically useful molecules, particularly those with novel modes of action, from "untapped" areas of chemical space? [5, 18]

In spite of this controversy, structural diversity (and thus total chemical space coverage), though rarely the "end game" in a synthesis project, 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 [19]. Indeed, the

generation of libraries of pure, structurally diverse compounds is thought by some to be the "key to the discovery of new medicines and to the elucidation of biological pathways through chemical genetics" [20].

1.5. STRUCTURAL DIVERSITY AND THE IMPORTANCE OF STRUCTURAL COMPLEXITY

The synthesis of a collection of structurally diverse small molecules offers a unique challenge to the synthetic chemist [6, 21]. It is widely accepted that it is not synthetically feasible to produce all theoretically stable, small carbon-based molecules [5, 21]. Thus, selectivity in synthesis is an important consideration. This issue has spurred the development of a variety of different approaches that aim to efficiently interrogate wide regions of chemical space simultaneously or to identify regions of chemical space that have an enhanced probability of containing biologically active compounds [6]. Before evaluating the relative merits of some of the most commonly used methods, it is useful to consider what is meant by the term *structural diversity* in the context of library synthesis.

Though the word *diversity* is, to some degree, a subjective one, there are four principle components of structural diversity that have been consistently identified in the literature: [19, 21, 22]

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; and
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 [12, 24, 25]. 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 biorelevant diversity [11]. Libraries 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 [11, 12]. Libraries of "compounds that have a common molecular skeleton display chemical information similarly in 3D space, thus limiting the pool of potential binding partners to only those macromolecules with a complementary 3D binding surface" [18, 24]. Thus, variation in 3D structure, rather than the nature of the peripheral substituents, is key for interaction with a broad range of molecular targets [26]. Conversely, conservation of 3D structure (i.e., the nature of the core molecular skeleton) in a small molecule collection generally means that the molecules will bind a narrower range of molecular targets.

Despite the acknowledged correlation between skeletal diversity and overall structural and functional (biological) diversity, the need to incorporate skeletal diversity in small molecule libraries is a somewhat contentious issue and is very much application dependent. In instances where a specific protein is being targeted, small molecule libraries are generally based around a single molecular skeleton – for example, the skeleton of a known natural ligand that has demonstrated the ability to bind to the desired protein. Furthermore, the synthesis of small molecule libraries around scaffolds present in known biologically active compounds has been cited as a possible means of identifying small molecules with novel biological properties. The synthesis of such "biased" small molecule libraries is discussed later in this chapter.

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 likelier to interact with biology in a selective and specific manner [5, 27].

1.6. SOURCES OF SMALL MOLECULES

There are a number of potential sources of small molecules for use in biological screens, each of which addresses various aspects of the structural diversity and structural complexity criteria in a different manner and with varying degrees of success.

1.6.1. Natural Products

Traditionally, nature has served as a rich source of biologically active molecules [22, 28, 29]. Natural products exhibit enormous structural diversity [30], and though they may vary in terms of structural complexity, many exhibit a high degree of specificity for their biological target. Numerous natural products have proven to be useful as drugs or leads [31, 32] and are still a major source of innovative therapeutic agents for infectious diseases [28].

Unfortunately, there are several problems associated with using natural product compounds in screening experiments. These include difficulties with purification and bioactive component(s) identification. Additionally, chemical modification and analog synthesis, processes that are

* The term *molecular skeleton* has no strict definition. Within the context of this discussion, the description recently outlined by Schreiber is appropriate; "we use the term skeleton loosely to denote 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" [23].

particularly pertinent in the drug development process, may be extremely challenging because of the highly complex nature of most natural products [19]. Furthermore, natural products occupy only a small proportion of chemical space [5, 17], which runs the risk of omitting a vast number of possibly biologically valuable small molecules from any screening process [18].

1.6.2. The Synthesis of Small Molecule Libraries

The crucial need for high-quality compound libraries in biological screening experiments has spurred the development of several *synthetic* approaches toward small molecule library synthesis. It is possible to analyze these approaches in terms of the various design concepts they employ. One of the most fundamental of these is the control in the nature of the molecular skeletons present in the small molecules. Based on this consideration, synthetic approaches to structurally diverse small molecule libraries can be divided into two distinct, broad groups:

1) *Biased approaches*: The synthetic route is designed with a pre-encoded structural bias, such that all of the resultant molecules are based around a similar molecular skeleton (i.e., a biased library).
2) *Nonbiased approaches*: The synthetic route is designed with no pre-encoded structural bias, such that there are a variety of different molecular skeletons present in the final products (i.e., a nonbiased library).

1.6.2.1. Biased Approaches: Combinatorial Chemistry

Commercially available combinatorial libraries and pharmaceutical proprietary compound collections are traditionally very important sources of small molecules [12, 30]. Combinatorial chemistry may be defined, in a very broad sense, as the rapid synthesis and screening of libraries of varied compounds to identify agents with desired functional properties [33]. Combined with established high-throughput screening (HTS) techniques, the development of combinatorial chemistry strategies in the early 1990s enabled the generation and testing of libraries of hundreds of thousands of different compounds at comparably low cost [30]. The method was quickly embraced by the pharmaceutical industry, with the hope that drug leads would be produced by sheer weight of numbers. However, the expected surge in productivity has not materialized. Indeed, as of the end of 2007 there was only one reported de novo new chemical entity **(1)** resulting from this method of chemical discovery that had been approved for drug use (Figure 4.3) [32].

This disappointing degree of productivity is generally attributed to defects in the nature of the libraries produced. Early combinatorial libraries have been described as being "intrinsically useless for drug discovery" [34] because the compounds were too similar to each other, having limited

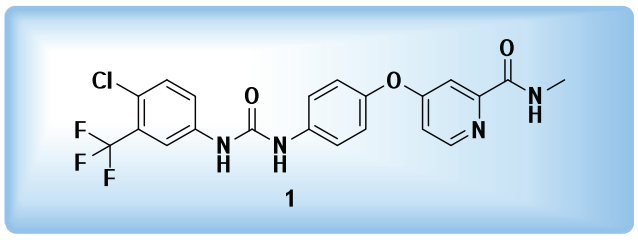


***FIGURE 4.3:** The antitumor compound 1 known as sorafenib (Nexavar) from Bayer, approved by the U.S. Food and Drug Administration (FDA) in 2005. It is a multikinase inhibitor and is in multiple clinical trials as both a combination and a single-agent therapy.*

structural diversity [30] and consequently offering "only a narrow slice" of chemical space [34]. This limited degree of overall structural diversity may be chiefly attributed to a lack of skeletal diversity. Traditionally, combinatorial libraries were based on a "one synthesis/one skeleton" approach, which resulted in a high degree of appendage diversity (and possibly stereochemical diversity) but little variety in the nature of the core molecular scaffold [12, 18, 19].

In an attempt to combat this problem, a more considered approach has been taken to combinatorial chemistry in recent years to try to increase the structural diversity exhibited by combinatorial libraries [12]. 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.*

It is widely recognized that molecules that have a common molecular skeleton tend to display chemical information similarly in 3D space and, as a result, are predisposed to bind to certain molecular targets that possess a complementary 3D binding surface [11, 18, 24]. In the case of a reverse chemical genetics experiment, in which small molecule modulators of a specific protein of interest are desired, it therefore can be advantageous to have a library of compounds based around a specific 3D structure (molecular skeleton) that has demonstrated affinity for the protein. The presence of specific structural features (such as a certain molecular skeleton) in a small molecule should predispose it toward binding the specific protein of interest. These desired/prerequisite binding structural features may be deduced from knowledge of the 3D structural features of the desired protein itself (e.g., a crystal structure) or of a known natural ligand for the protein. In the context of a reverse chemical genetics, it could be argued that it is inherently wasteful to synthesize (and screen) a library of skeletally diverse molecules that are designed to bind to a

* In the context of this discussion, a *novel* biologically active molecule can be defined as a molecule with a previously unknown biological activity that exerts this biological effect through a unique mechanism/mode of action.

wide range of proteins when only a specific protein is being targeted.

Although the 3D structure of a small molecule – that is, the nature of its core molecular skeleton – is of central importance in determining the range of molecular targets it can interact with, variation in the functional groups (structure and stereochemistry) present in the molecule is often important in terms of the strength and specificity of these interactions. Thus, reverse chemical genetics experiments may benefit from the screening of small molecule collections that display appendage, stereochemical, and functional group diversity around a common molecular skeleton, as this may facilitate the identification of ligands that bind with increased selectivity (and/or affinity) for the protein of interest.

1.6.2.2. Biased Approaches: Synthesis around a Privileged Molecular Skeleton Although molecules based around a specific core skeleton are generally limited to a smaller range of biological partners than molecules based around a diverse range of skeletons, it has emerged that there is a subset of molecular skeletons whose presence in molecules confers upon them more flexible binding properties. The frameworks in question are so-called privileged frameworks that 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 through modification of functional groups" [35]. These are molecular scaffolds that are common to known bioactive molecules (usually natural products) and thus have proven biological relevance [6, 36].

In recent years, there has been an increase in the synthesis of small molecule libraries that are based around such privileged structures. The rationale behind using such an approach is based on two main hypotheses. The first is that evolutionary pressure over millions of years has "prevalidated" natural products, and thus compounds that are structurally similar, to be able to modulate protein function. Second, the chemical space explored by natural products and protein structure during evolution is strongly limited in size and highly conserved. That is, there is a concept of *evolutionary convergence of structures* in that natural products have evolved to interact with multiple proteins [37, 38]. Synthesis around a privileged scaffold, which has also been referred to as natural product–guided synthesis, 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 [39]. It has been argued that privileged structure–based synthesis may permit such compounds to be found with enhanced probability and quality.

For example, Park et al. have recently reported the construction of a library of small molecules composed of twenty-two discrete and novel core skeletons embedded with a privileged benzopyran motif **(2)** through a branching DOS strategy [40]. The outline of the synthetic routes used is given in Scheme 1.

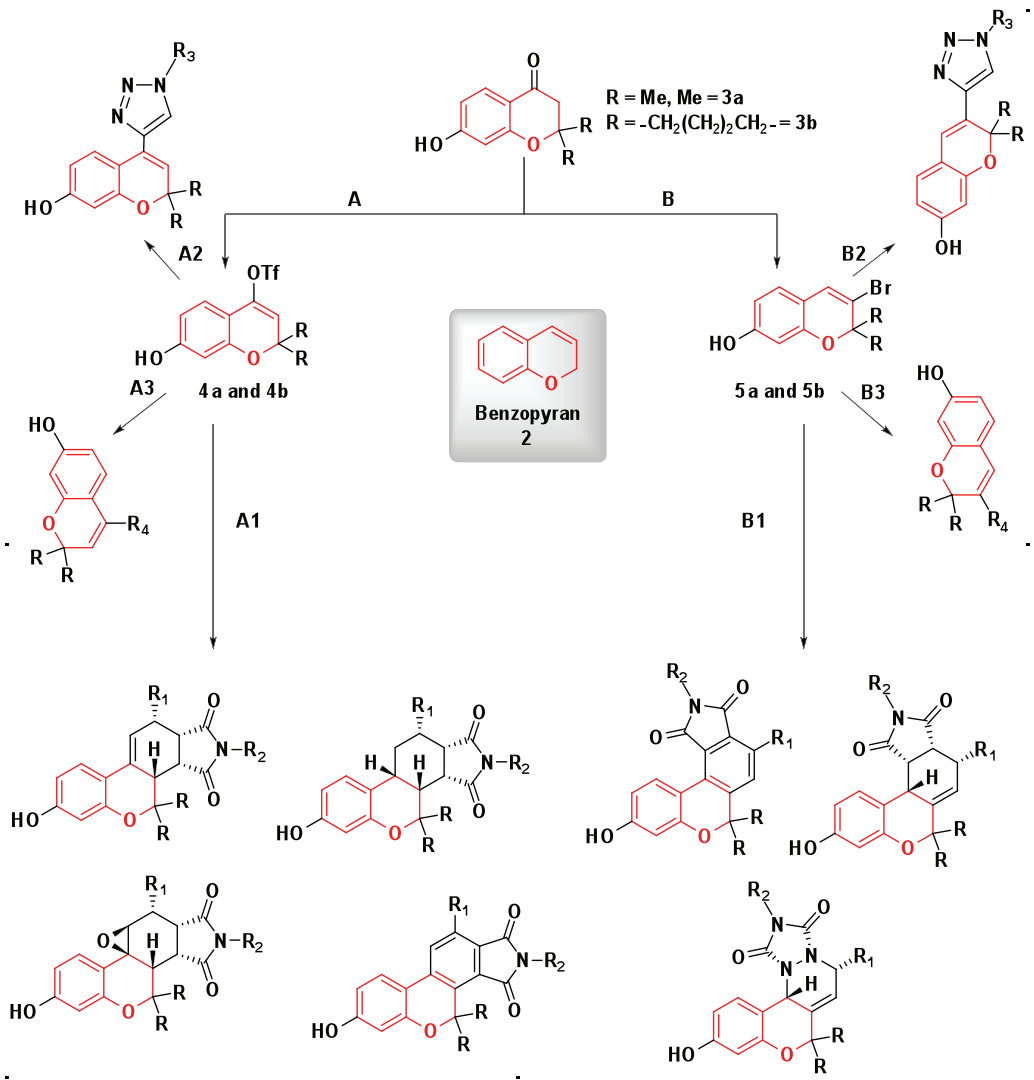
Starting from compounds **3a** and **3b**, intermediates **4** and **5** could be generated by reaction pathways A and B, respectively. These served as branch points for various chemical transformations such as Diels-Alder reactions (paths A1 and B1), click chemistry (A2 and B2), and palladium-mediated cross-coupling (A3 and B3). Using this strategy, Park and colleagues were able to synthesize twenty-two novel molecular architectures, each one containing the benzopyran substructure. The biological diversity of the library was demonstrated by the dramatic differences in the biological activities (IC_{50} values against a human cancer cell line) of compounds sharing the same appendices but having different 3D structures; that is, the variation in biological activity was reported to be a function of the core molecular skeleton rather than of the appendages present.

The use of a synthesis around a privileged structure approach again highlights the controversial issue of the degree of overlap between total chemical space and biologically relevant chemical space – that is, whether there is any point in exploring seemingly uncharted regions of chemical space (which nature may have indeed "sampled" through the process of evolution over the course of millions of years but ultimately ignored as a source of biologically useful molecules) when chemical space occupied by known natural products and medicinal compounds is enriched with bioactive structures. What is clear is that if we do not try to access such regions, we will never know! Clearly, library synthesis around a privileged structure is particularly relevant when a specific protein target is being considered. When a less focused approach is required – for example, if the protein target is unknown or we are hoping to find novel biologically active molecules – the use of nonbiased synthetic approaches that aim to access a wider range of chemical descriptor space may be more useful.

1.6.3. Nonbiased Approaches: DOS

The aim of a nonbiased approach toward library synthesis is to create a structurally diverse (including skeletal diversity) and functionally diverse small molecule collection with 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 [41]. DOS has recently emerged as a new synthetic approach to achieving this objective.

The goal 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, 18, 21, 42]. The hope is that by sampling a greater total area of chemical space, the functional



***SCHEME 1:** Outline of the synthetic strategy employed by Park and co-workers to generate a library based around twenty-two distinct core molecular skeletons embedded with a privileged benzopyran motif. The benzopyran motif common to all the members is highlighted in red. Tf = trifluoromethanesulfonate.*

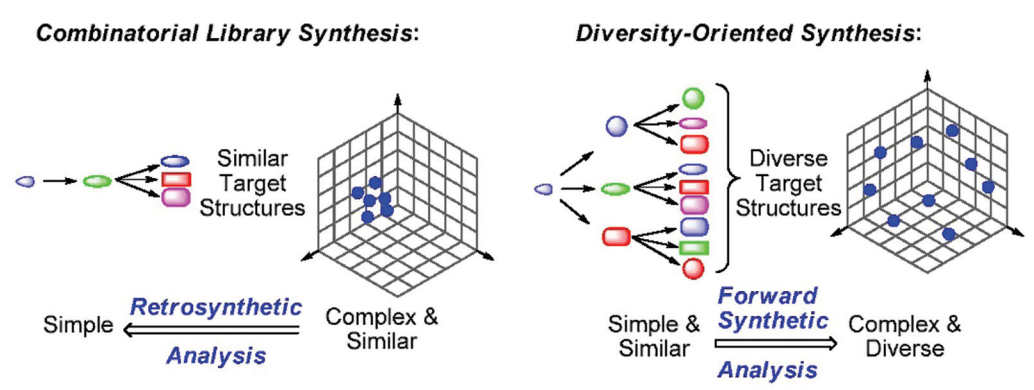


***FIGURE 4.4:** A comparison of the synthetic planning strategies used in a traditional combinatorial synthesis with a DOS, together with a visual representation of the chemical space coverage achieved in both cases (i.e., focused around a specific point or diverse coverage).*

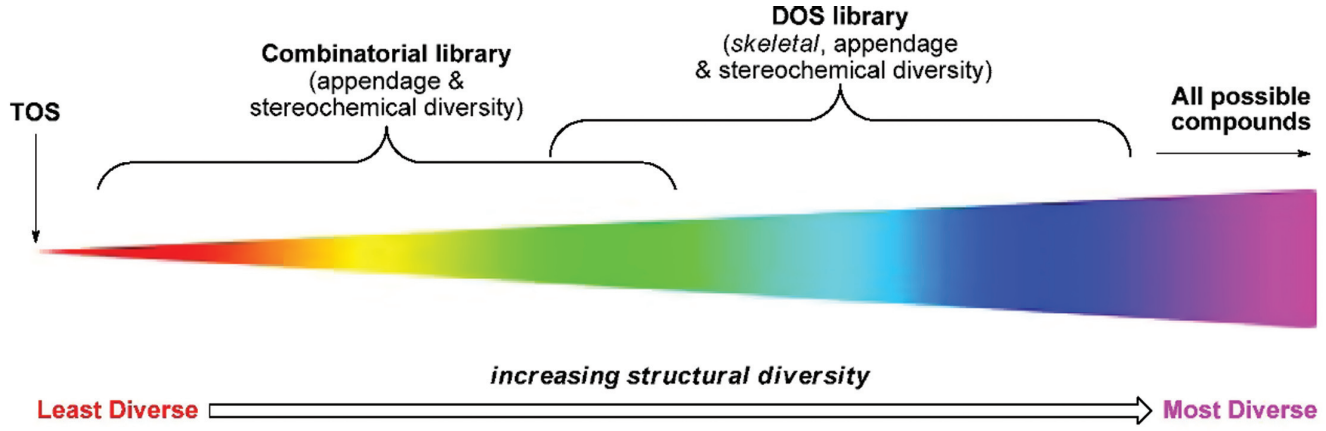


FIGURE 4.5: *The "molecular diversity spectrum." "Diversity" can be viewed as a spectrum ranging from a target-oriented synthesis (TOS) of a specific molecule (i.e., a total synthesis) to the synthesis of all possible compounds (i.e., total chemical space coverage); compound collections arising from a traditional combinatorial approach and those arising from a DOS sit between these two extremes [19].*

(biological) diversity of the library will be greater, increasing the chances of identifying a compound with the desired properties [10–12].

An illustration of the synthetic strategy used in a DOS is given in Figure 4.4. A DOS synthetic pathway is analyzed in the *forward* sense; a single, simple starting material is converted into a collection of structurally diverse small molecules in no more then five synthetic steps (to maximize synthetic efficiency) [18, 21, 43]. The overall aim is the broad, nonfocused coverage of chemical space, which can be contrasted with the outcomes of more traditional combinatorial syntheses.

A successful DOS must address the four principle types of structural diversity mentioned previously (Section 1.5), that is, appendage, functional group, stereochemical, and skeletal* [6, 18, 19, 43, 44]. It is the ability of a DOS to incorporate skeletal diversity into a compound collection that is the most challenging facet of this method and of central importance to its success [18, 24, 42, 45]. 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" [11] (relative to a single-scaffold library) [12, 19, 24, 46].

There is a clear distinction between DOS and traditional combinatorial methods; DOS libraries are generally smaller but consist of molecules that are structurally more complex, have a greater variety of core structures (skeletons), and possess richer stereochemical variation [33]. However, the boundary between modern, more considered combinatorial approaches and DOS is less clear-cut, and the terms DOS and combinatorial chemistry are often used interchangeably in the literature. Indeed, because many of the principles of combinatorial chemistry are used in DOS, it is probably best to consider DOS as a more evolved version of traditional combinatorial methodologies.† Recently, the concept of the "molecular diversity spectrum" has been introduced as a useful qualitative means for comparing the structural diversity associated with a particular molecular collection (Figure 4.5). It should therefore "be the goal of a DOS to synthesize, in a qualitative sense, collections of small molecules which are as near as possible to the right hand side of this spectrum" [19].

It should be noted that reverse chemical genetics experiments may benefit from the use of nonbiased (i.e., skeletally as well as functionally diverse) small molecule collections. Access to a range of different molecular architectures may allow the discovery of a small molecule that is completely unrelated structurally to any natural ligand(s) for the protein of interest but is capable of modulating the activity of this protein in a much more useful manner (e.g., is more selective, binds with a higher affinity, is easier to access).

* The process of varying functional group, appendage, and stereochemical diversity around privileged scaffolds is occasionally referred to in the literature as *DOS around a privileged scaffold*. However, the true ethos of DOS is based around a diverse, nonfocused coverage of chemical space, which is most efficiently achieved though variation in all aspects of diversity, including skeletal. In this context, we believe that the term *DOS around a privileged scaffold* is somewhat of a contradiction in terms and that other descriptions are more appropriate for the process of library generation around a privileged scaffold, e.g., *natural-product-inspired synthesis*.

† 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 optimization once a biologically 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 [32].

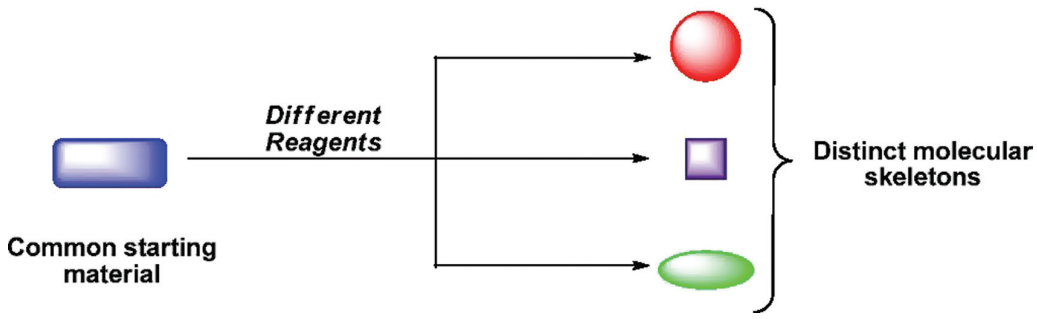


FIGURE 4.6: *Achieving skeletal diversity using the reagent-based approach.*

1.6.3.1. Achieving Skeletal Diversity via a DOS from Simple Starting Materials The efficient synthesis of a functionally, stereochemically, and *skeletally* diverse collection of small molecules from a common, simple starting material was the process described and developed by Schreiber in much of his pioneering work in the field of DOS [18, 24, 47]. The remainder of this chapter focuses on the general methods that have been developed to maximize skeletal diversity in small molecule libraries generated through a DOS approach.

There are two main approaches that have been developed to generate skeletal diversity in DOS libraries, based either upon the *reagent* (the reagent-based approach) or the *substrate* (the substrate-based approach) [18, 23, 48].

1.6.3.1.1. The reagent-based approach. The reagent-based approach is a branching synthetic strategy that involves a short series of divergent, complexity-generating reactions from a common starting material to generate a collection of compounds with distinct molecular skeletons (Figure 4.6) [18].

Critical to the success of this strategy is the choice of a synthetically versatile starting substrate that has the potential to be converted into two or more products with different molecular skeletons through the variation of reagents alone. These products in turn should be suitable for further diversification, preferably in further branching reactions (i.e., conversion into two or more skeletally diverse products again through choice of reagents).

In practice, reagent-based skeletal diversity is achieved via two main methods: [18, 19]

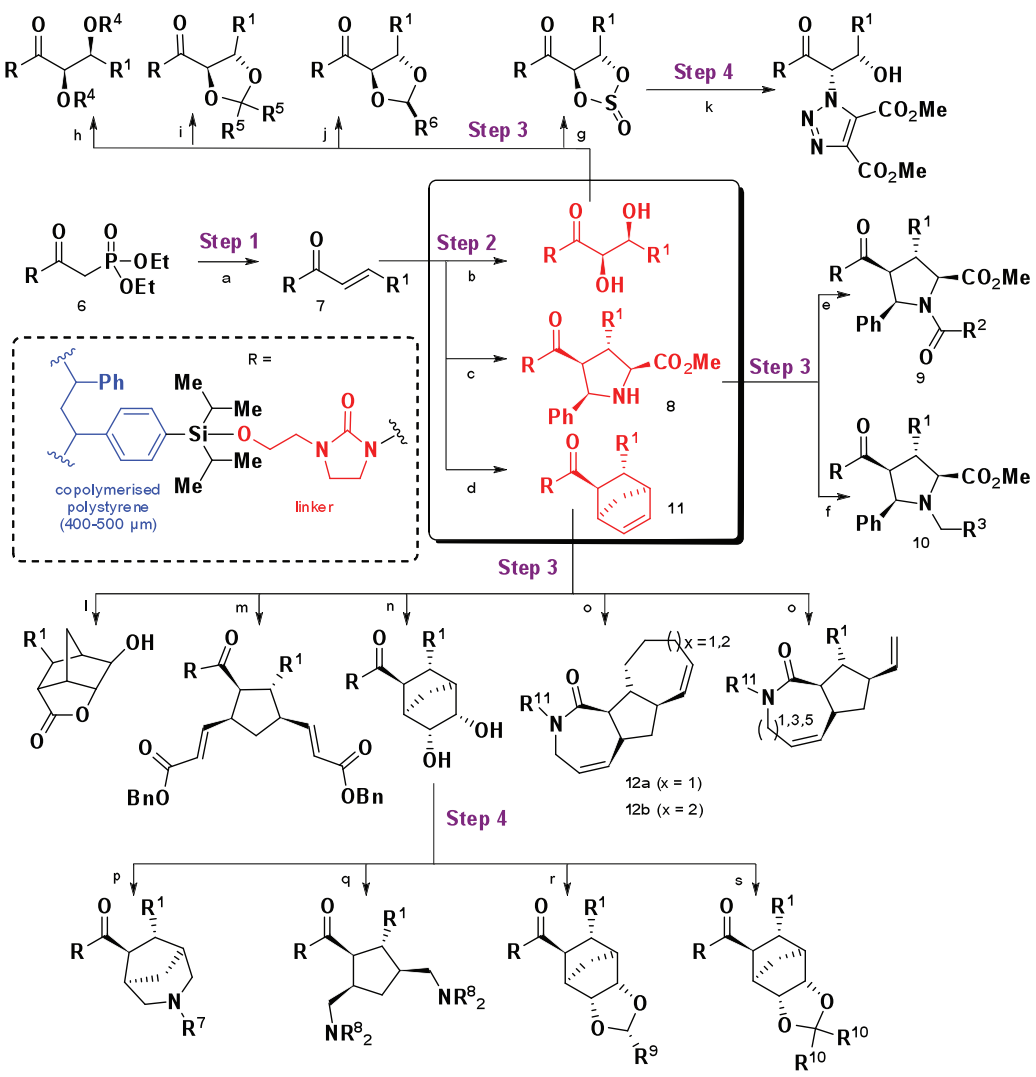
1) The use of a *pluripotent functionality* whereby exposure of a given molecule to different reagents results in different reactions occurring at the same part (functional group) of the molecule; and
2) The use of a *densely functionalized molecule* whereby different functionalities in the same molecule are transformed by different reagents.

1.6.3.1.2. The use of a pluripotent functionality. Thomas et al. have recently reported an example of the use of a pluripotent functional group strategy to generate a skeletally diverse compound collection (Scheme 2) [49]. This work employed the solid-supported phosphonate **6** as the starting unit. The immobilization of **6** on a silyl-polystyrene support greatly simplified product purification during library synthesis.

In the first step of the DOS, **6** was reacted with a variety of aldehyde building blocks (building block diversity) to deliver twelve α,β-unsaturated acyl-imidazolidinones **7**. The second step of the DOS involved three catalytic enantioselective divergent reaction pathways (stereochemical diversity): 1) dihydroxylation (reaction b); 2) [2 + 3] cycloaddition (reaction c); and 3) [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) involved a series of branching reactions to diversify these key branch-point substrates further. For example, the pyrrolidine products **8** could be acylated or alkylated (reactions e and f, respectively) to yield **9** and **10** (appendage diversity). The norbornene derivatives **11** (formed in step d) served as suitable intermediates for a series of branching reactions (reactions l to o) to generate five different molecular scaffolds (skeletal diversity). For example, a tandem ring-closing-opening-closing metathesis reaction was carried out (reaction o) to give skeletally diverse tricyclic products **12a** (7-5-7) and **12b** (7-5-8).* Of particular note was the generation of the *cis-trans*-fused 7-5-7 scaffold of **12a**, which has no known representation in nature, highlighting the capability of this DOS approach to generate products that populate new, unexplored regions of chemical space. In the final step of the DOS (step 4), the compounds were cleaved off the solid support using a variety of reagents (appendage diversity).

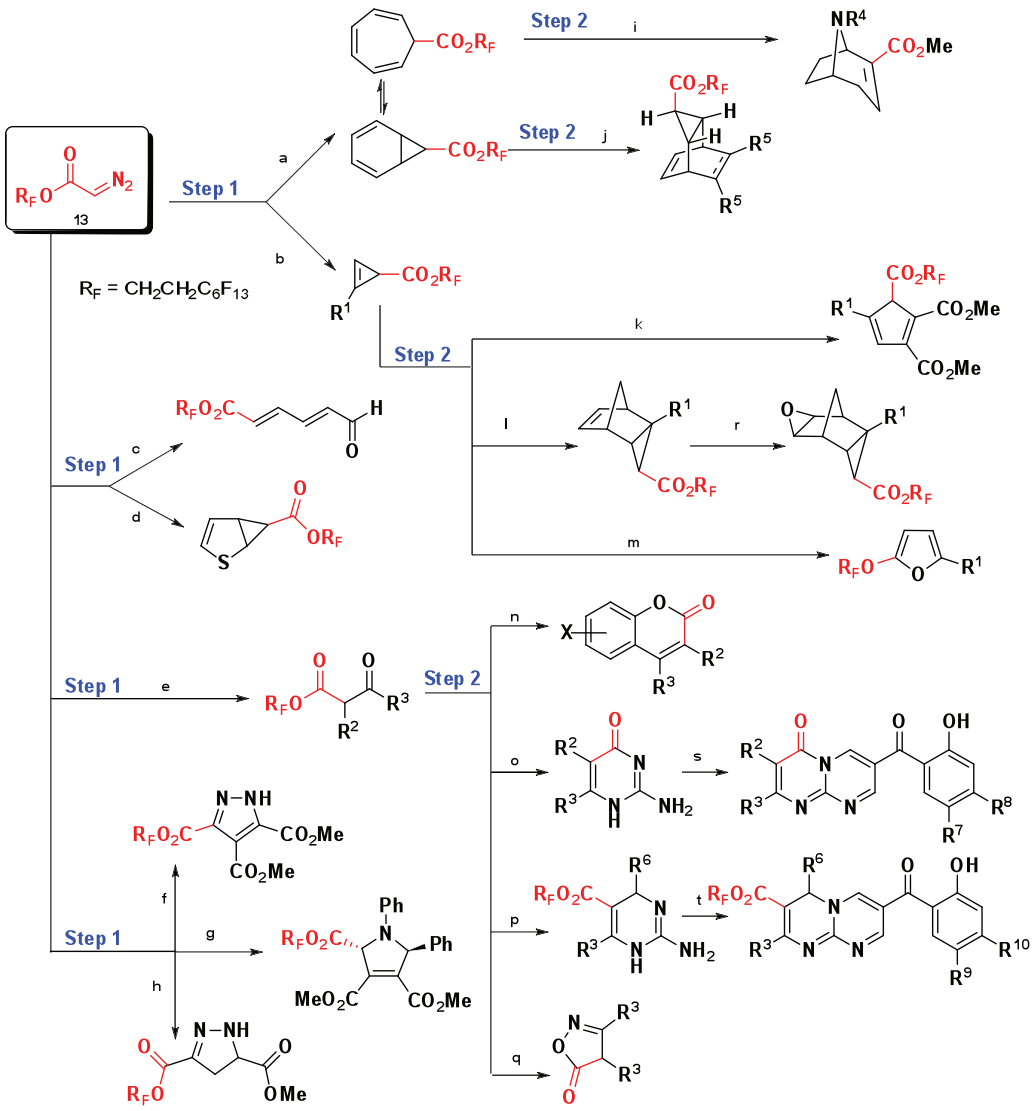
Using the chemistry shown in Scheme 2 and a limited number of structurally diverse building blocks, a DOS of 242 small molecules that have eighteen molecular frameworks, among other unique structural features, was achieved.

* Spandl and co-workers have recently developed a related tandem metathesis process that allows the generation of complex polycyclic molecular architectures from substituted norbornene derivatives in a highly efficient and atom-economical manner [50].



SCHEME 2: *DOS of 242 compounds based of 18 discrete molecular frameworks. Conditions: (a) LiBr, 1,8-diazabicyclo[5.4.0]undec-7-ene, R^1CHO, MeCN; (b) AD-mix β , THF:H2O (1:1); (c) (R)-QUINAP, AgOAc, iPr2NEt, α -imino-ester, THF, − 78°C to 25°C; (d) chiral bis(oxazoline), Cu(OTf)2, 3Å MS, CH2Cl2, C5H6; (e) R^2COCl, DMAP, pyridine, CH2Cl2; (f) R^3CHO, BH3-pyridine, MeOH; (g) SOCl2, pyridine, CH2Cl2, 40°C; (h) R^4Br, Ag2O, CH2Cl2, 40°C; (i) R^5C(O)R^5, TsOH, DMF, 65°C; (j) R^6CHO, TsOH, DMF, 65°C; (k) NaN3, DMF, 100°C then DMAD, toluene, 65°C; (l) mCPBA, CH2Cl2 then MeOH, 65°C; (m) CH2=CHCO2Bn, Grubbs' second-generation catalyst, ethylene, toluene, 120°C; (n) OsO4, NMO, CH3C(O)CH3:H2O (10:1); (o) RNH2, Me2AlCl, toluene 120°C; then NaH, R^{11}X, DMF, THF; then toluene, 120°C, Grubbs' second-generation catalyst, ethylene; (p) NaIO4, THF:H2O (1:1); then R^7NH2, NaB(OAc)3H, CH2Cl2; (q) NaIO4, THF:H2O (1:1); then R^8NHR^8, NaB(OAc)3H, CH2Cl2; (r) R^9CHO, DMF, TsOH, 60°C; (s) R^{10}C(O)R^{10}, DMF, TsOH, 60°C. (DHQD)PHAL = hydroquinidine 1,4-phthalazinediyl diether; DMAD = dimethyl acetylenedicarboxylate; DMAP = N,N-dimethylaminopyridine; DMF = N,N-dimethylformamide; Grubbs II = 1,3-(bis(mesityl)-2-imidazolidinylidene) dichloro (phenylmethylene) (tricyclohexylphosphine) ruthenium; mCPBA =* meta*-chloroperbenzoic acid; NMO = 4-methylmorpholine-N-oxide;THF = tetrahydrofuran; Ts =* para*-toluenesulfonyl.*

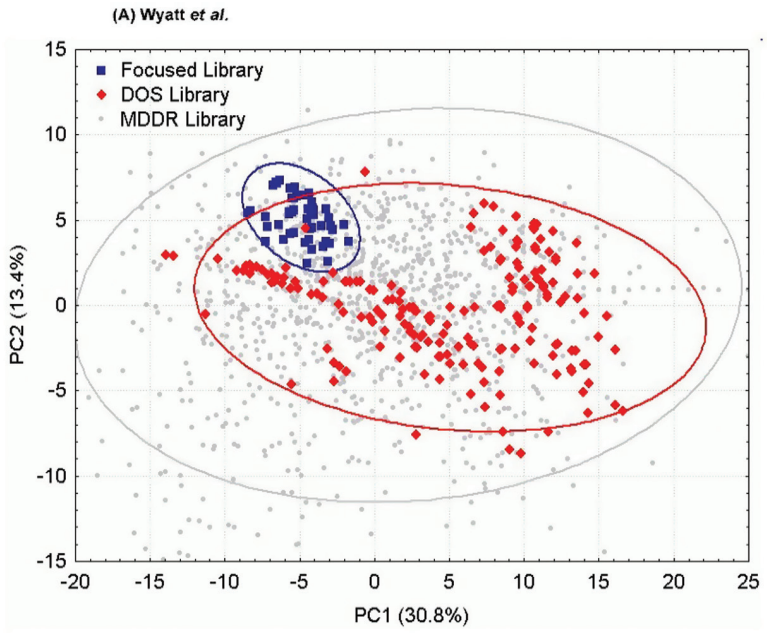
A branching DOS strategy was also utilized by Wyatt et al. for the synthesis of a skeletally diverse small molecule library (Scheme 3) [46]. The fluorous tagged diazoacetate compound **13** was identified as an attractive starting unit for two main reasons: 1) diazoacetate compounds exhibit enormous synthetic versatility, allowing a wide variety of different synthetic transformations to be carried out on the starting material; and 2) polyfluorocarbon tag technology allowed standard solution-phase parallel synthesis methods to be coupled with the benefits of fluorous-based purification protocols [51, 52], thus simplifying the isolation of the library compounds.



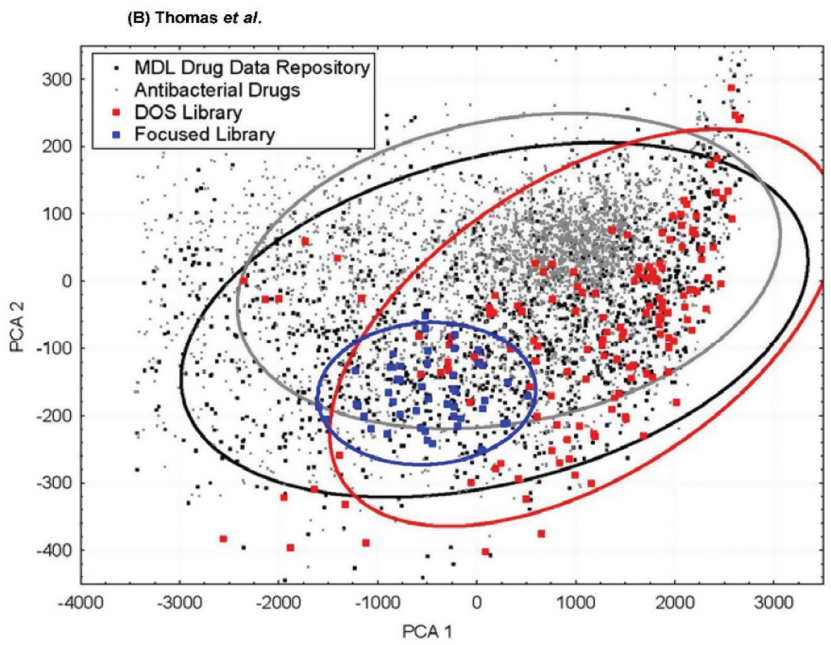
SCHEME 3: *The synthetic plan for the DOS of a library of small molecules from a simple diazoacetate starting material 13. Step 1 refers to the first step of the DOS; Step 2 refers to the second step of the DOS. (a) C6H6, Rh2(OCOCF3)4; (b) R^1CCH, Rh2(OAc)4, CH2Cl2; (c) Furan, Rh2(OAc)4 then I2; (d) Thiophene, Rh2(OAc)4; (e) LDA – 78° C, then R^2COR^3, THF then Rh2(OAc)4, CH2Cl2; (f) DMAD; (g) PhCHO, PhNH2 then DMAD, Rh2(OAc)4 or toluene. [Cu(OTf)]2, CH2Cl2; (h) methyl acrylate; (i) R^4NH2, NaOH, H2O, 180° C then MeOH, H2SO4, 60° C; (j) dienophile, toluene, reflux; (k) DMAD, toluene, 100° C; (l) cyclopentadiene, CH2Cl2, 0° C to rt; (m) Grubbs' II, toluene, ethylene, reflux; (n) phenol derivative, conc. H2SO4; (o) guanidine, EtOH, reflux; (p) guanidine, R^6CHO, DMF, 75° C; (q) NH2OH, THF, reflux; (r) mCPBA, CH2Cl2, rt; (s) substituted 3-formyl chromone, EtOH, reflux; (t) substituted 3-formyl chromone, EtOH, reflux. DMAD = dimethyl acetylenedicarboxylate; Grubbs' II = 1,3-(bis(mesityl)-2-imidazolidinylidene) dichloro (phenylmethylene) (tricyclohexylphosphine) ruthenium; LDA = lithium diisopropylamide; mCPBA =* meta*-chloroperbenzoic acid; Tf = trifluoromethanesulfonate; THF = tetrahydrofuran.*

The main pathways are summarized in Scheme 3. In the first stage of the DOS, the reactive diazoacetate functionality was exploited in four main branching reactions. Each branching reaction was chosen so as to generate at least one unique molecular skeleton (Step 1, Scheme 3):

1) Cycloaddition with benzene and alkynes (reactions a and b, respectively);
2) Cycloaddition with the heteroaromatic compounds furan and thiophene (reactions c and d, respectively);
3) 1,3-Dipolar cycloadditions with dimethyl acetylenedicarboxylate (DMAD) and methyl acrylate and a three-component ylide-mediated cycloaddition (reactions f, g, and h, respectively); and
4) α-Deprotonation followed by trapping of the resultant anion with an electrophile and subsequent



Library	Average 'chemical space' occupied per compound
Focused	0.70
MDDR	28.28
DOS	11.25

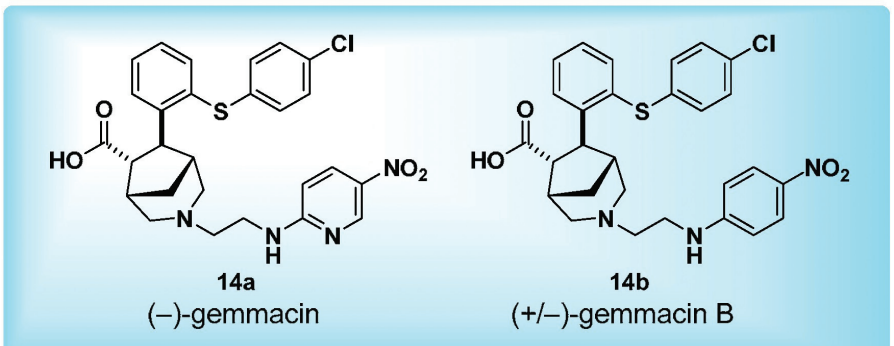


Library	Average 'chemical space' occupied per compound
Focused	0.64
MDDR	18.86
DOS	22.32
Antibacterial	12.79

FIGURE 4.7: 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 squares (A: Wyatt et al. B: Thomas et al.). For comparison, a focused library (blue squares), the MDL Drug Data Repository (MDDR; black dots), and antibacterial drugs (gray dots) (B only) are depicted. Data for the average "chemical space" occupied per compound calculated in the context of each of the DOS libraries are shown in the tables on the right of the corresponding chemical space diagrams.

metal-catalyzed hydrogen or carbon migration to form β-dicarbonyl compounds (reaction e).

The second stage of the DOS involved a series of complexity-generating reactions to diversify these molecular frameworks further, leading to the generation of more skeletal diversity in the library (Step 2, Scheme 3, reactions i through q). In some cases, a third stage of reactions was carried out (reactions r, s, and t) to introduce additional complexity and diversity. Appendage and functional group diversity were introduced into the library through variation in the substrates used in these branching reactions (i.e., variation in $R^{1–10}$). In addition, additional appendage and functional group diversity were incorporated into the



	MIC_{50} (μg ml^{-1})	
Compound	**EMRSA 15**	**EMRSA 16**
(-)-gemmacin	8	16
(+/-)-gemmacin B	8	8
erythromycin	>64	>64
oxacillin	>32	>32

FIGURE 4.8: ***Structure and activity of gemmacin (14a) and gemmacin b (14b) with growth inhibitory activity (MIC_{50}) against two strains of methicillin-resistant* Staphylococcus aureus*, EMRSA 15 and EMRSA 16. For comparison, the MIC_{50} values for erythromycin and oxacillin are also shown. ND = not determined. MSSA = methicillin-susceptible* S. aureus*. MIC_{50} = minimum inhibitory concentration required to inhibit the growth of 50% of organisms.***

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 (building blocks) were used in the branching synthetic routes outlined in Scheme 3 to synthesize a library of 223 compounds based on thirty different core molecular skeletons in no more than four linear synthetic steps from a simple diazoacetate starter unit.

A fundamental problem when attempting the synthesis of a diverse small molecule library is the subjective nature of diversity itself; that is, how does one determine how successful one library synthesis is compared with another in terms of diversity generation? Recent years have witnessed significant progress toward the development of computational methods that allow an assessment of the relative diversity present in different chemical collections in a more quantitative fashion. For example, for each compound in the DOS libraries synthesized by Wyatt et al. and Thomas et al., the values of 184 different physiochemical and topological chemical descriptor properties (e.g., molecular weight, degree of branching, pKa, charges) were calculated. The data sets produced for each compound were analyzed using principal component analysis (PCA) to generate a unique set of coordinates for each compound in chemical space.

Using this method, a two-dimensional (2D) visual representation of the distribution of the compounds of the DOS libraries in chemical space was derived (Figure 4.7A for the library of Wyatt et al., Figure 4.7B for the library of

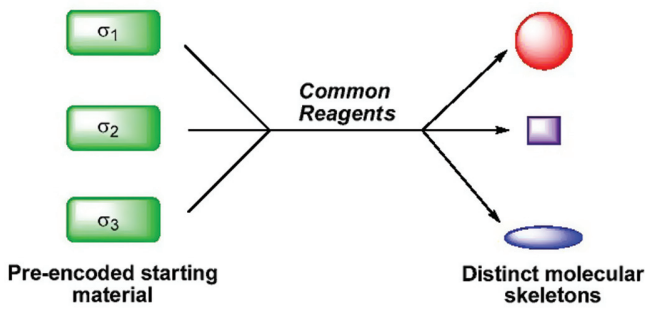


FIGURE 4.9: ***The substrate-based approach (folding process) to generating skeletal diversity.***

Thomas et al.). For comparison, the distribution in chemical space of two "benchmark" molecule collections was also computed and included in these diagrams. The first of these collections was a focused library (indicated by blue squares) that was produced via a more traditional combinatorial approach (whereby a common scaffold is decorated with different appendages) [53]. 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 cutoff of 650) taken from the MDL Drug Data Repository (MDDR) database (small gray dots) [54]. Finally, in the case of the library synthesized by Thomas et al. (Figure 4.7B), the chemical space coverage achieved by the 3,762 compounds marked as "antibacterial" in the MDDR database is also included. A quantitative estimation of the diversity achieved on a per-compound basis for each of these three chemical collections was also made (this can be interpreted as a measure of the average chemical space occupied per compound for each compound collection).

In both cases, the DOS libraries (high skeletal diversity) span a larger region of chemical space than that occupied by the focused library (low skeletal diversity). This supports the premise that maximizing skeletal diversity in a small molecule library is critical in terms of maximizing overall structural diversity and thus chemical space coverage. In the case of the library synthesized by Wyatt et al., the largest coverage of chemical space is achieved by the MDDR sample. However, the DOS library, despite having significantly fewer compounds, seems to occupy a relatively large region of chemical space, illustrating the value of this DOS approach to generate structurally diverse products that span a wide area of chemical space in an efficient manner. Perhaps of greater significance was the fact that the compound collection produced by Thomas et al. was shown to be even more diverse than the MDDR library in terms of (relative) diversity units – that is, 22 for the DOS library, 19 for MDDR, 13 for the antibacterials, and 0.6 for the focused library.

Computational analyses such as those described previously (Figure 4.7) can be performed to determine if a library synthesis has been successful in terms of achieving a high degree of diversity. However, it is important to remember 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 unsuccessful, no matter how structurally diverse it is. Phenotypic screening experiments performed using the DOS libraries of Wyatt et al. and Thomas et al. identified a number of structurally novel compounds that displayed antibacterial activity against pathogenic strains of methicillin-resistant *Staphylococcus aureus* (EMRSA 15 and EMRSA 16, the strains responsible for the majority of MRSA infections in the United Kingdom), demonstrating the utility of the DOS approach for the discovery of new antibacterial agents. The most active compound identified was **14a**, named gemmacin, which displayed a broad range of activity against Gram-positive bacteria and is believed to operate as a cell-membrane disruptor (Figure 4.8).

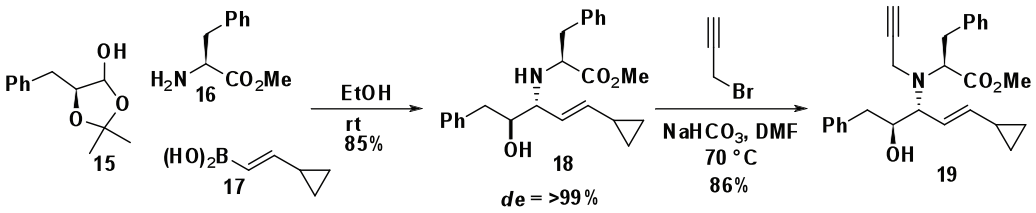
These examples clearly illustrate that biologically active molecules (so-called hits) can be identified through the screening of structurally diverse small molecule collections. However, optimization is usually required to transform these "hits" into "leads" that possess more useful properties (e.g., increased potency or specificity). Such optimization is usually achieved by the synthesis of a focused library around the original bioactive molecule through variation in appendage, functional group, and stereochemical diversity elements. Subsequent screening of these compounds allows structure-activity relationships (SARs) to be investigated. Recently, Thomas and coworkers have reported such a SAR investigation of the antibacterial compound gemmacin that was identified in the DOS campaign discussed previously (Scheme 2)[55]. Their studies identified compound **14b**, named gemmacin B, which demonstrated higher levels of bioactivity against EMRSA 16 (Figure 4.8). Interestingly, these SAR studies suggested that antibacterial activity was very dependent on the original structural features displayed by gemmacin; the authors surmised that "the gemmacin architecture appears to be situated on an 'isolated island of bioactivity' where little manoeuvrability (i.e., chemical diversification) is possible if antibacterial activity is to be retained."

Further examples of the use of a pluripotent functionality approach to DOS can be found in some reviews and recently published articles [18, 43, 45, 56].

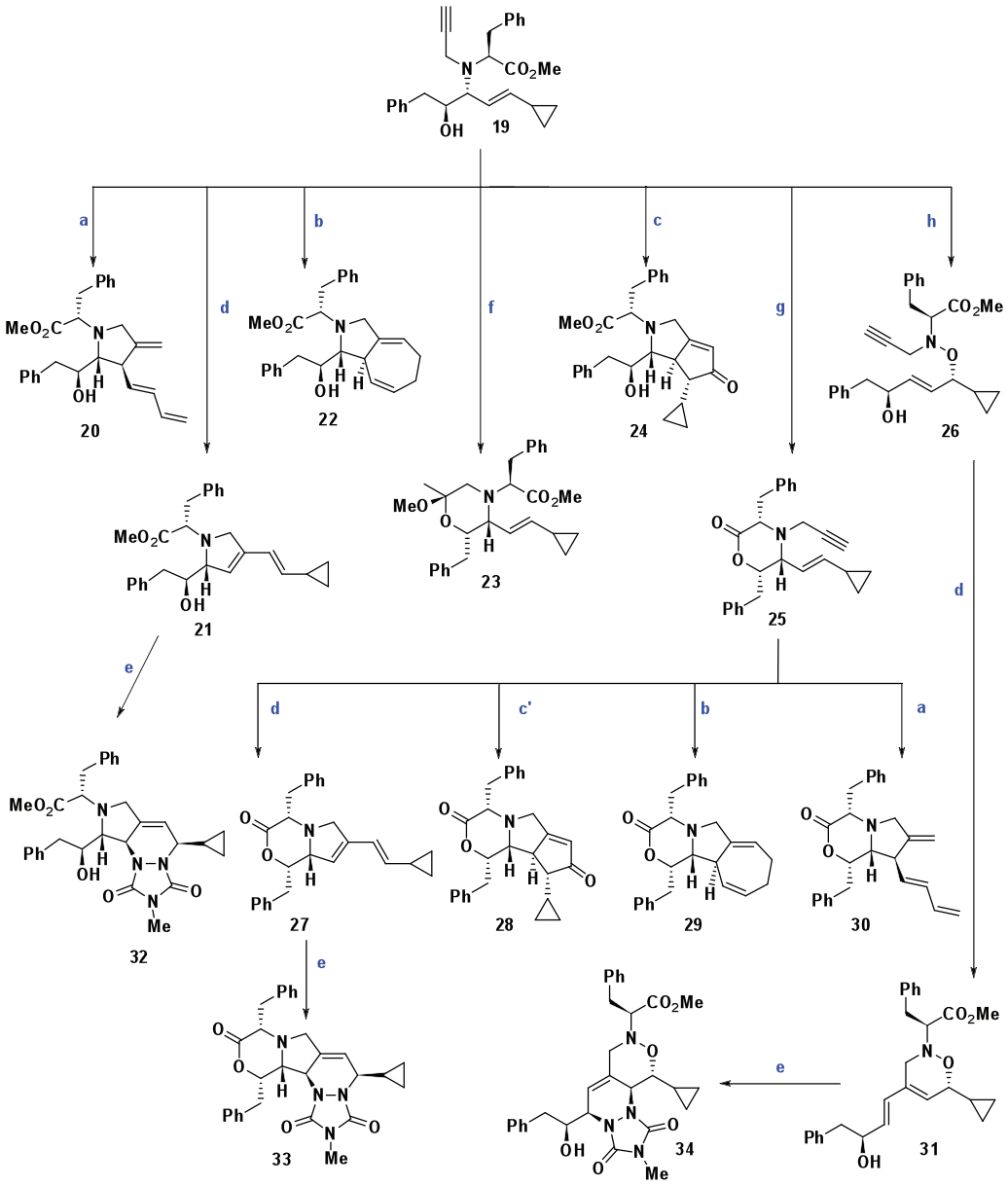
1.6.3.1.3. The use of a densely functionalized molecule. Schreiber et al. have reported the synthesis of a skeletally and stereochemically diverse small molecule collection via a DOS approach based around the varied reactivity of densely functionalized β-amino alcohol derivatives [57].

A Petasis three-component coupling of **15**, **16**, and **17** generated compound **18**; subsequent amine propargylation furnished the highly functionalized β-amino alcohol derivative **19** (Scheme 4). The stereochemical outcome of the Petasis reaction was found to be controlled by the lactol **15**, suggesting that all four possible anti–amino alcohol stereoisomers could (in principle) be generated (stereochemical diversity).

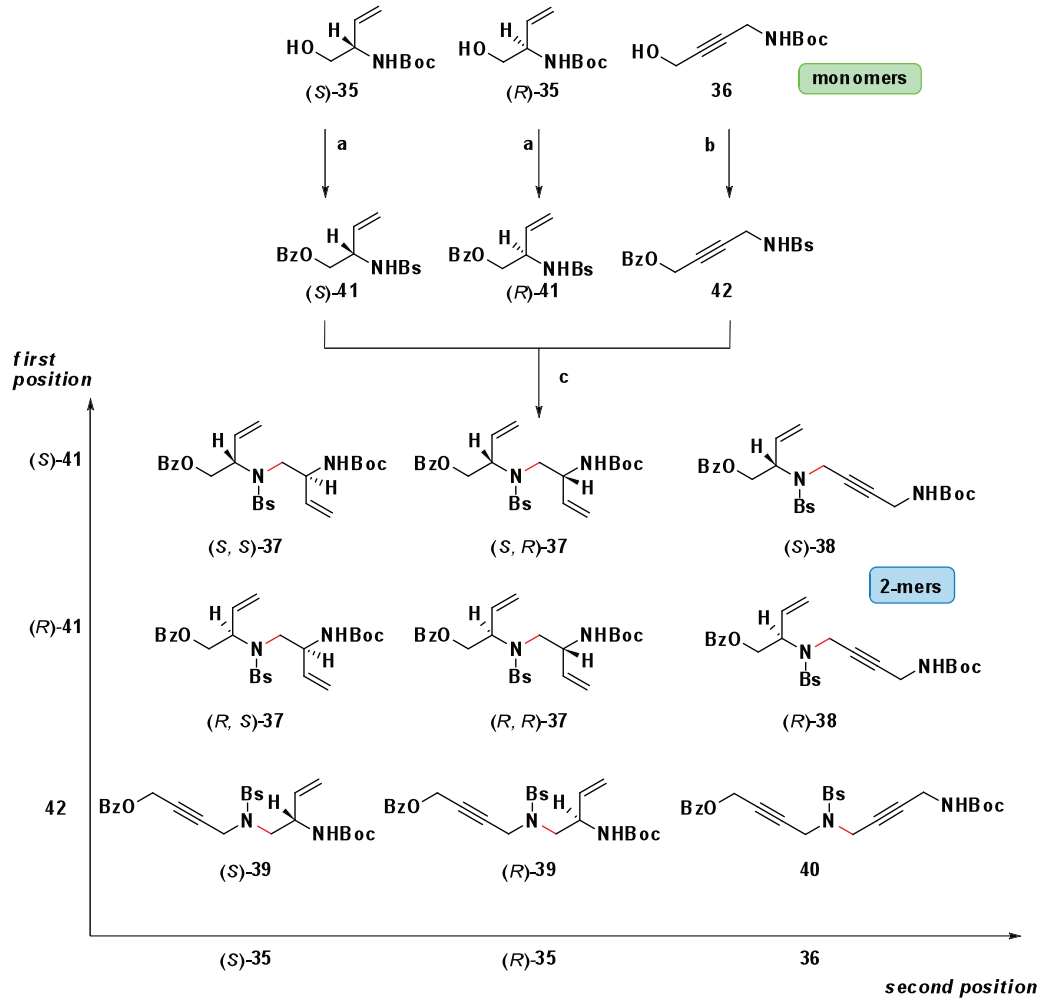
Initially, a series of seven skeletal diversification reactions was carried out on **19**, yielding products **20** through **26** (Scheme 5). These reactions were based around reactivity at four of the different functionalities present in **19** – that is, the hydroxyl group, the alkene, the alkyne, and the cyclopropane moieties: 1) palladium-catalyzed cycloisomerization (route a); 2) ruthenium-catalyzed cycloisomerization (route b); 3) cobalt-mediated Pauson-Khand



SCHEME 4: ***Synthesis of β -amino alcohol 19. DMF = N,N-dimethylformamide; rt = room temperature.***



SCHEME 5: ***Conditions: (a) [Pd(PPh3)2(OAc)2] (10 mol%), benzene, 80°C; (b) [CpRu(CH3CN)3PF6] (10 mol%), acetone, rt; (c) [Co2(CO)8], trimethylamine N-oxide, NH4Cl, benzene, rt; (c′) [Co2(CO)8], trimethylamine N-oxide, benzene, rt; (d) Hoveyda-Grubbs second-generation catalyst (10 mol%), CH2Cl2, reflux; (e) 4-methyl-1,2,4-triazoline-3,5-dione, CH2Cl2, rt; (f) NaAuCl4 (10 mol%), MeOH, rt; (g) NaH, toluene, rt; (h) mCPBA, THF, −78 to 0°C. rt = room temperature; mCPBA =*** meta-***chloroperbenzoic acid.***



SCHEME 6: Synthesis of 2-mer combinations from monomer units; (a) (i) TFA then BsCl/NaHCO3, EtOAc, (ii) BzCl, pyridine; (b) (i) BzCl, pyridine, (ii) TFA then BsCl, NEt3, CH2Cl2; (c) 2-mers were synthesized from the corresponding alcohol and brosylate using PPh3, DEAD, THF, 0°C to rt. Newly formed bonds are shown in red. Boc = tert-butoxycarbonyl; Bz = benzoyl; Bs = brosylate (BrC6H4SO2); Ac = acetyl; TFA = trifluoroacetic acid; DEAD = diethyl azodicarboxylate; THF = tetrahydrofuran.

reaction (route c); 4) enyne metathesis (route d); 5) gold-catalyzed intramolecular cyclization (route f); 6) lactone formation (route g); and 7) *m*CPBA-mediated rearrangement (route h).

Subjecting lactone **25** to a similar set of conditions used in the diversification of **19** led to products **27** through **30**. Further diversification of compounds containing a 1,3-diene moiety (**21**, **27**, and **31**) was achieved via a Diels-Alder reaction with 4-methyl-1,2,4-triazolin-3,5-dione to yield products **32** through **34** (route e). The whole DOS pathway was repeated using alternative amine building blocks in the Petasis reaction (substitutional diversity). The net result was that this DOS strategy enabled the synthesis, in only three to five steps, of a diverse collection of single-isomer small molecules whose members displayed substitutional, stereochemical, *and* skeletal diversity, with fifteen different molecular skeletons being present among the molecules produced.

1.6.3.2. The Substrate-Based Approach The substrate-based approach to skeletal diversity is based around a *folding process*. It involves the conversion of a collection of substrates that contain appendages with suitable "pre-encoded" skeletal information (so-called σ-elements) into products have distinct molecular skeletons using a common set of conditions (Figure 4.9) [18, 42, 48].

Spiegel and co-workers have reported the synthesis of a library of skeletally diverse small molecules using such an approach, based around reactive "oligomers" that could be subjected to a chemical transformation that causes them to fold up into distinct 3D shapes (Figure 4.10) [58].

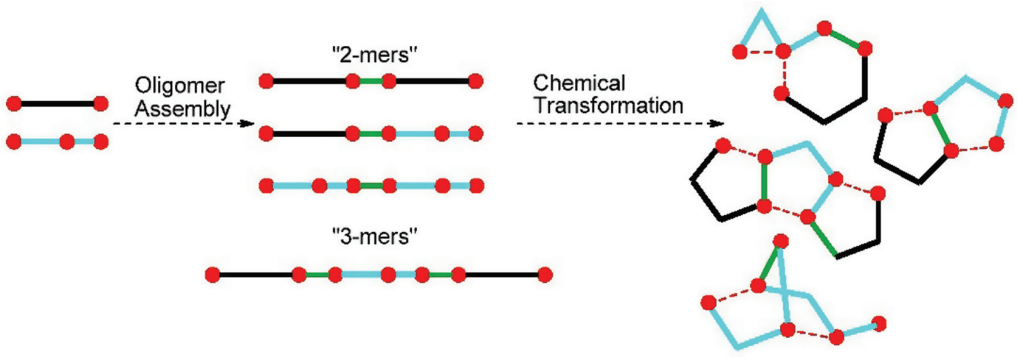
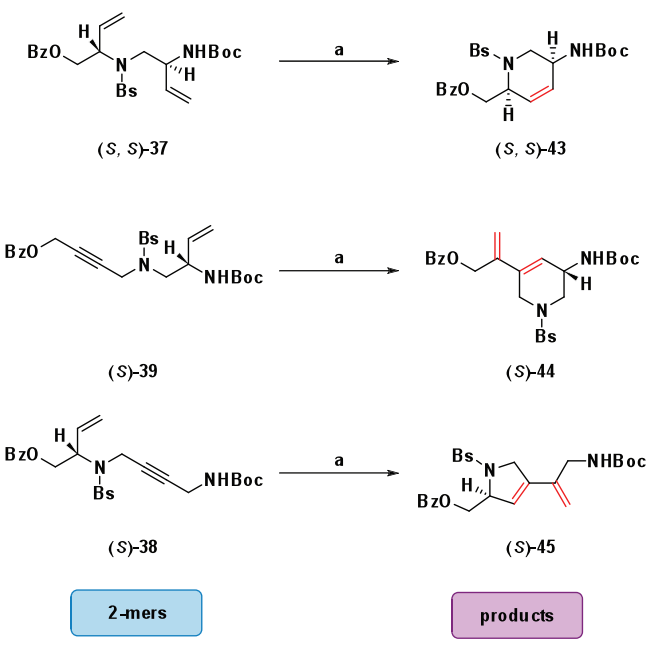


FIGURE 4.10: Schematic depiction of the oligomer-based approach used by Spiegel et al. [58] Red circles indicate reactive groups, green lines indicate monomer attachment sites, and dashed red lines indicate newly formed covalent bonds.

The library synthesis was based around three monomer units, (*S*)-**35**, (*R*)-**35** and **36**, which were converted into all nine possible 2-mers (a combination of 2 monomers) **37** to **40** via intermediate sulfonamides **41** and **42** (Scheme 6).

Treatment of all stereoisomeric variants **37** to **40** with Grubbs' first-generation catalyst provided eight different products having three types of skeletons: disubstituted tetrahydopyridine **43**, vinyl tetrahydropyridine **44**, and dihydropyrrole **45** (three selected examples are illustrated in Scheme 7).

Various 3-mers (a combination of 3 monomers) (*S*, *R*)-**46**, (*S*, *S*)-**46**, and **47** were then accessed from the 2-mers



SCHEME 7: Three examples of the generation of skeletally diverse products by treatment of three different 2-mers with a common reagent; (a) Grubbs' I (5 mol%), toluene, ethylene, reflux, 16–30 h. Newly formed bonds are shown in red. Boc =* tert*-butoxycarbonyl; Bs = brosylate (BrC6H4SO2); Bz = benzoyl; Grubbs' I = benzylidene-bis(tricyclohexylphosphine)dichlororuthenium.

(*S*)-**38** and **40** using a Fukuyama-Mitsunobu coupling reaction (Scheme 8). Under the common metathesis conditions employed for the reaction of the 2-mers, polycyclic compounds (*S*, *R*)-**48**, (*S*, *S*)-**48**, and **49** were formed.

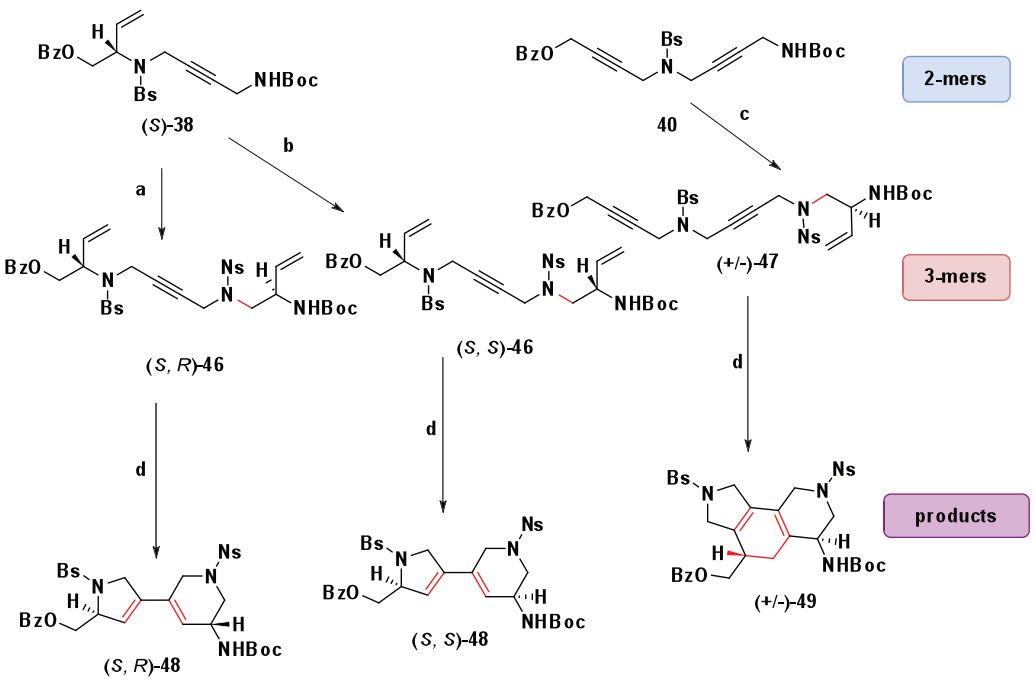
By the methods outlined above, a library of small molecules based on twelve distinct molecular skeletons was produced. The 1,3-diene products of the oligomerization-skeletalization sequence also served as substrates in a second skeletal diversification step using Diels-Alder reactions (not shown). Overall, starting from only three simple monomer units, a library of small molecules based around seventeen unique molecular skeletons was produced. Further examples of the use of a substrate-based approach to DOS can be found in some reviews and recently published articles [19, 23, 24, 48].

1.6.3.3. The Build-Couple-Pair Strategy Recent work by Schreiber has identified a common strategic feature present in many DOS pathways. This is the so-called build/couple/pair (B/C/P) three-phase strategy [23] (Figure 4.11).

The three phases involve can be defined in the following fashion: [23]

1) *Build:* Asymmetric syntheses of chiral building blocks;
2) *Couple:* Intermolecular coupling reactions that join the building blocks are performed; this process provides the basis for stereochemical diversity; and
3) *Pair*: Intramolecular coupling reactions that join pairwise combinations of functional groups incorporated in the "build" phase are performed; this process provides the basis for skeletal diversity.

The characteristics of a B/C/P approach toward skeletal diversity construction can be identified in both reagent- and substrate-based DOS pathways. For example, the DOS pathway outlined in Scheme 4 and Scheme 5 can be analyzed in terms of a B/C/P strategy; the "build" phase involved assembly of fragments **15**, **16**, and **17**; the Petasis



SCHEME 8: Formation of 3-mers and their subsequent diversification. (a) (i) TFA then NsCl, NaHCO3, EtOAc, (ii)(R)-29, PPh3, DEAD, THF, 0°C to rt; (b) (i) TFA then NsCl, NaHCO3, EtOAc, (ii)(S)-29, PPh3, DEAD, THF, 0°C to rt; (c) TFA then NsCl, NaHCO3, EtOAc, (ii)(rac)-29, PPh3, DEAD, THF, 0°C to rt; (d) Grubbs' I (5 mol%), toluene, ethylene, reflux, 16–30 hr. Newly formed bonds from each step are shown in red. Ac = acetyl; Boc = tert-butoxycarbonyl; Bs = brosylate (BrC6H4SO2); Bz = benzoyl; DEAD = diethyl azodicarboxylate; Grubbs' I = benzylidene-bis(tricyclohexylphosphine)dichlororuthenium; Ns = para-nitrophenylsulphonyl; TFA = trifluoroacetic acid; THF = tetrahydrofuran.

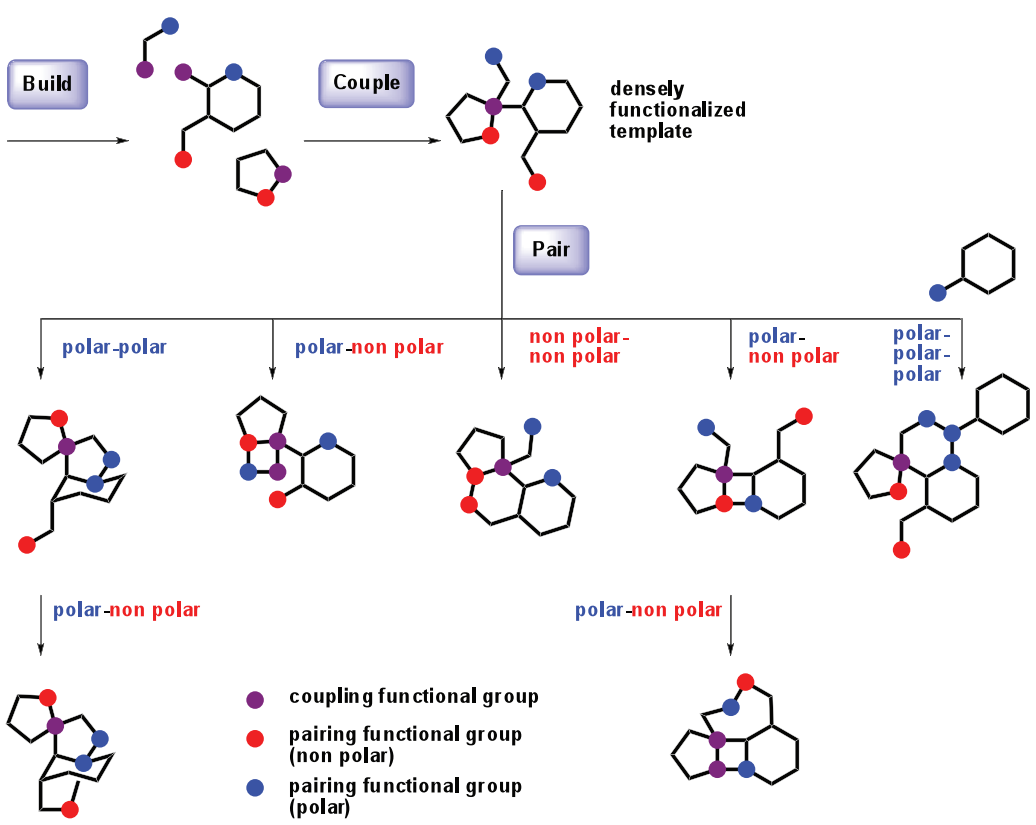


FIGURE 4.11: Generation of skeletal diversity with the B/C/P strategy: the pair phase consists of chemoselective and intramolecular joining of strategically placed polar (blue) and nonpolar (red) functional groups to afford diverse skeletons [23].

reaction followed by alkylation served as "couple" phases; and the subsequent reagent-controlled skeletal diversification reactions served as "pair" phases in which different combinations of the moieties of **19**, both polar and nonpolar, were "paired" in functional group specific reactions. For example, the palladium- and ruthenium-based catalysts selectively paired the nonpolar alkene, alkyne, and cyclopropane groups of **19**, enabling the cycloisomerization reactions leading to compounds **20** to **22**, whereas sodium hydride–mediated lactonization selectively paired the polar functional groups to form **25**.

The DOS pathway of Spiegel et al. outlined in Schemes 6, 7, and 8 can also be analyzed in a similar fashion. The "build" phase involved the synthesis of non-racemic monomers **35** and **36**. The combination of these monomers by reaction of their polar groups to form 2-mers and 3-mers comprised the "couple" phase, and the "pair" phase focused on joining the nonpolar groups of these 2-mers and 3-mers through the use of ruthenium-catalyzed metathesis reactions [23].

1.6.4. Conclusions

Small molecule libraries that display a high degree of structural and thus functional (biological) diversity have proven to be valuable in the discovery of molecules that can modulate the activities of biological macromolecules in a useful fashion. The synthesis of such diverse small molecule collections presents many challenges to the synthetic chemist. In recent years, many innovative DOS strategies have been developed in an attempt to increase the overall structural diversity of small molecule libraries in an efficient manner. The issue of maximizing skeletal diversity is widely recognized as the key to achieving this goal; "that the diversity of the (small molecule) library is defined by the diversity of the scaffolds (molecular skeletons) that make up the library is becoming an axiom" [25]. Many compound collections synthesized using a DOS approach have been successfully exploited in chemical genetics to identify useful modulators for biological systems [59, 60]. Nonetheless, there still remains a need for the development of new sources of diverse small molecules that can be exploited as potential therapeutic agents and research tools in biological systems. However, it is once again worth emphasizing that the ultimate success of any small molecule library is determined by the biological relevance of the compounds it contains. If the library does not yield hits in a chosen biological screening experiment, it will be deemed unsuccessful, no matter how structurally diverse it is.

REFERENCES

1. Walsh, D. P., and Chang, Y. T. (2006). Chemical genetics. *Chem Rev 106*, 2476–2530.
2. Spring, D. R. (2005). Chemical genetics to chemical genomics: small molecules offer big insights. *Chem Soc Rev 34*, 472–482.
3. Schreiber, S. L. (1998). Chemical genetics resulting from a passion for synthetic organic chemistry. *Bioorg Med Chem 6*, 1127–1152.
4. Hopkins, A. L., and Groom, C. R. (2002). The druggable genome. *Nat Rev Drug Discov 1*, 727–730.
5. Lipinski, C., and Hopkins, A. (2004). Navigating chemical space for biology and medicine. *Nature 432*, 855–861.
6. Kaiser, M., Wetzel, S., Kumar, K., and Waldmann, H. (2008). Biology-inspired synthesis of compound libraries. *Cell Mol Life Sci 65*, 1186–1201.
7. Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev 46*, 3–26.
8. Di, L., and Kerns, E. H. (2003). Profiling drug-like properties in discovery research. *Curr Opin Chem Biol* 7, 402–408.
9. Tan, D. S. (2005). Diversity-oriented synthesis: exploring the intersections between chemistry and biology. *Nat Chem Biol 1*, 74–84.
10. Haggarty, S. J. (2005). The principle of complementarity: chemical versus biological space. *Curr Opin Chem Biol 9*, 296–303.
11. Sauer, W. H. B., and Schwarz, M. K. (2003). Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. *J Chem Inf Comput Sci 43*, 987–1003.
12. Kennedy, J. P., Williams, L., Bridges, T. M., Daniels, R. N., Weaver, D., and Lindsley, C. W. (2008). Application of combinatorial chemistry science on modern drug discovery. *J Comb Chem 10*, 345–354.
13. Estrada, E., and Uriarte, E. (2001). Recent advances on the role of topological indices in drug discovery research. *Curr Med Chem 8*, 1573–1588.
14. Fergus, S., Bender, A., and Spring, D. R. (2005). Assessment of structural diversity in combinatorial synthesis. *Curr Opin Chem Biol 9*, 304–309.
15. Oprea, T. I., and Gottfries, J. (2001). Chemography: the art of navigating in chemical space. *J Comb Chem 3*, 157–166.
16. Oprea, T. I. (2002). Chemical space navigation in lead discovery. *Curr Opin Chem Biol 6*, 384–389.
17. Dobson, C. M. (2004). Chemical space and biology. *Nature 432*, 824–828.
18. Burke, M. D., and Schreiber, S. L. (2004). A planning strategy for diversity-oriented synthesis. *Angew Chem Int Ed 43*, 46–58.
19. Spandl, R. J., and Spring, D. R. (2008). Diversity-oriented synthesis; a spectrum of approaches and results. *Org Biomol Chem 6*, 1149–1158.
20. Adriaenssens, L. V., Austin, C. A., Gibson, M., Smith, D., and Hartley, R. C. (2006). Stereodivergent diversity oriented synthesis of piperidine alkaloids. *Eur J Org Chem*, 4998–5001.
21. Spring, D. R. (2003). Diversity-oriented synthesis; a challenge for synthetic chemists. *Org Biomol Chem 1*, 3867–3870.
22. Thomas, G. L., Wyatt, E. E., and Spring, D. R. (2006). Enriching chemical space with diversity-oriented synthesis. *Curr Opin Drug Discov Devel 9*, 700–712.

23. Nielsen, T. E., and Schreiber, S. L. (2008). Diversity-oriented synthesis – towards the optimal screening collection: a synthesis strategy. *Angew Chem Int Ed 47*, 48–56.
24. Burke, M. D., Berger, E. M., and Schreiber, S. L. (2003). Generating diverse skeletons of small molecules combinatorially. *Science 302*, 613–618.
25. Shelat, A. A., and Guy, R. K. (2007). Scaffold composition and biological relevance of screening libraries. *Nat Chem Biol 3*, 442–446.
26. An, H., Eum, S., Koh, M., Lee, S., and Par, S. (2008). Diversity-oriented synthesis of privileged benzopyranyl heterocycles from *s-cis*-enones. *J Org Chem 73*, 1752–1761.
27. Hopkins, A. L., Mason, J. S., and Overington, J. P. (2006). Can we rationally design promiscuous drugs? *Curr Opin Struct Biol 16*, 127–136.
28. Clardy, J., and Walsh, C. (2004). Lessons from natural molecules. *Nature 432*, 829–837.
29. Pucheault, M. (2008). Natural products: chemical instruments to apprehend biological symphony. *Org Biomol Chem 6*, 424–432.
30. Schneider, G., and Grabowski, K. (2007). Properties and architecture of drugs and natural products revisited. *Curr Chemical Biol 1*, 115–127.
31. Butler, M. S. (2004). The role of natural product chemistry in drug discovery. *J Nat Prod 67*, 2141–2153.
32. Newman, D. J., and Cragg, G. M. (2007). Natural products as sources of new drugs over the last 25 years. *J Nat Prod 70*, 461–477.
33. Borman, S. (2004). Rescuing Combichem. *Chem Eng News: Sci Technol 82*, 32–40.
34. Rouhi, A. M. (2003). Rediscovering natural products. *Chem Eng News 81*, 77–78, 82–83, 86, 88–91.
35. DeSimone, R. W., Currie, K. S., Mitchell, S. A., Darrow, J. W., and Pippin, D. A. (2004). Privileged structures: applications in drug discovery. *Comb Chem High Throughput Screening 7*, 473–493.
36. Koch, M. A., and Waldmann, H. (2005). Protein structure similarity clustering and natural product structure as guiding principles in drug discovery. *Drug Discov Today 10*, 471–483.
37. Balamurugan, R., Dekker, F. J., and Waldmann, H. (2005). Design of compound libraries based on natural product scaffolds and protein structure similarity clustering (PSSC). *Mol Biosyst 1*, 36–45.
38. Breinbauer, R., Vetter, I. R., and Waldmann, H. (2002). From protein domains to drug candidates – natural products as guiding principles in the design and synthesis of compound libraries. *Angew Chem Int Ed 41*, 2879–2890.
39. Goess, B. C., Hannoush, R. N., Chan, L. K., Kirchhausen, T., and Shair, M. D. (2006). Synthesis of a 10,000-membered library of molecules resembling carpanone and discovery of vesicular traffic inhibitors. *J Am Chem Soc 128*, 5391–5403.
40. Ko, S. K., Jang, H. J., Kim, E., and Park, S. B. (2006). Concise and diversity-oriented synthesis of novel scaffolds embedded with privileged benzopyran motif. *Chem Commun*, 2962–2964.
41. Cordier, C., Morton, D., Murrison, S., Nelson, A., and O'Leary-Steele, C. (2008). Natural products as an inspiration in the diversity-oriented synthesis of bioactive compound libraries. *Nat Prod Rep 25*, 719–737.
42. Burke, M. D., Berger, E. M., and Schreiber, S. L. (2004). A synthesis strategy yielding skeletally diverse small molecules combinatorially. *J Am Chem Soc 126*, 14095–14104.
43. Burke, M. D., and Lalic, G. (2002). Teaching target-oriented and diversity-oriented organic synthesis at Harvard University. *Chem. Biol. 9*, 535–541.
44. Mishra, J. K., and Panda, G. (2007). Diversity-oriented synthetic approach to naturally abundant S-amino acid based benzannulated enantiomerically pure medium ring heterocyclic scaffolds employing inter- and intramolecular Mitsunobu reactions. *J Comb Chem 9*, 321–338.
45. Taylor, S. J., Taylor, A. M., and Schreiber, S. L. (2004). Synthetic strategy toward skeletal diversity via solid-supported, Otherwise unstable reactive intermediates. *Angew Chem Int Ed 43*, 1681–1685.
46. Wyatt, E. E., Fergus, S., Galloway, W. R., Bender, A., Fox, D. J., Plowright, A. T., Jessiman, A. S., Welch, M., and Spring, D. R. (2006). Skeletal diversity construction via a branching synthetic strategy. *Chem Commun*, 3296–3298.
47. Schreiber, S. L. (2000). Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science 287*, 1964–1969.
48. Oguri, H., and Schreiber, S. L. (2005). Skeletal diversity via a folding pathway: synthesis of indole alkaloid-like skeletons. *Org Lett 7*, 47–50.
49. Thomas, G. L., Spandl, R. J., Glansdorp, F. G., Welch, M., Bender, A., Cockfield, J., Lindsay, J. A., Bryant, C., Brown, D. F. J., Loiseleur, O., Rudyk, H., Ladlow, M., and Spring, D. R. (2008). Anti-MRSA agent discovery using diversity-oriented synthesis. *Angew Chem Int Ed 47*, 2808–2812.
50. Spandl, R. J., Rudyk, H., and Spring, D. R. (2008). Exploiting domino enyne metathesis mechanisms for skeletal diversity generation. *Chem Commun*, 3001–3003.
51. Curran, D., and Luo, Z. Y. (2001). Fluorous techniques for the synthesis and separation of organic molecules. *Green Chem 3*, G3–G7.
52. Zhang, W. (2003). Fluorous technologies for solution-phase high-throughput organic synthesis. *Tetrahedron 59*, 4475–4489.
53. Faghih, R., Dwight, W., Pan, J. B., Fox, G. B., Krueger, K. M., Esbenshade, T. A., McVey, J. M., Marsh, K., Bennani, Y. L., and Hancock, A. A. (2003). Synthesis and SAR of aminoalkoxy-biaryl-4-carboxamides: novel and selective histamine H3 receptor antagonists. *Bioorg Med Chem Lett 13*, 1325–1328.
54. MDL Drug Data Report, http://www.symyx.com/products/pdfs/mddr_ds.pdf.
55. Robinson, A., Thomas, G. L., Spandl, R. J., Welch, M., and Spring, D. R. (2008). Gemmacin B: bringing diversity back into focus. *Org Biomol Chem 6*, 2978–2981.
56. Kumar, N., Kiuchi, M., Tallarico, J. A., and Schreiber, S. L. (2005). Small-molecule diversity using a skeletal transformation strategy. *Org Lett 7*, 2535–2538.
57. Kumagai, N., Muncipinto, G., and Schreiber, S. L. (2006). Short synthesis of skeletally and stereochemically diverse small molecules by coupling petasis condensation reactions to cyclization reactions. *Angew Chem Int Ed 45*, 3635–3638.

58. Spiegel, D. A., Schroeder, F. C., Duvall, J. R., and Schreiber, S. L. (2006). An oligomer-based approach to skeletal diversity in small-molecule synthesis. *J Am Chem Soc 128*, 14766–14767.
59. Koehler, A. N., Shamji, A. F., and Schreiber, S. L. (2003). Discovery of an inhibitor of a transcription factor using small molecule microarrays and diversity-oriented synthesis. *J Am Chem Soc 125*, 8420–8421.
60. Spring, D. R., Krishnan, S., Blackwell, H. E., and Schreiber, S. L. (2002). Diversity-oriented synthesis of biaryl-containing medium rings using a one bead/one stock solution platform. *J Am Chem Soc 124*, 1354–1363.