26
Diversity-Oriented Synthesis: Developing New Chemical Tools to Probe and Modulate Biological Systems

Warren R. J. D. Galloway, David Wilcke, Feilin Nie, Kathy Hadje-Georgiou, Luca Laraia, and David R. Spring

26.1
Introduction

Small molecular mass chemical entities (the so-called small molecules) are capable of interacting with biological macromolecules and exerting profound effects upon their function [1]. The use of small molecules to selectively perturb biological systems underpins the field of chemical biology and forms the basis of modern medicine [1, 2]. Put simply, humanity has a significant dependence on biologically active small molecules [3]. It is therefore unsurprising that significant effort is directed toward the identification of new small molecules with specific biological activity.

26.2
The Biological Problem

26.2.1
How to Discover New Chemical Modulators of Biological Function?

In situations where the biological target is structurally well defined, it is often possible to use this information to rationally design or select small-molecule-binding partners. Similarly, if the structure of the natural ligand is known, this can be used as a "template" to guide compound synthesis or selection [1, 4]. However, in cases where the precise nature of the biological target is unknown (e.g., a phenotypic screen), natural ligands are unidentified, or a novel mode of binding to a particular target is desired, a "rational" design, or selection process is clearly not possible. In such situations, the discovery of bioactive small molecules relies upon the biological assessment (screening) of collections (or "libraries") of small molecules to identify those with the desired properties (the so-called "hits"). The success of any such screening endeavor will clearly be inherently dependent upon the molecular composition of the library [1, 4].

Concepts and Case Studies in Chemical Biology, First Edition. Edited by Herbert Waldmann and Petra Janning.

26.2.2 Sources of Small Molecules for Screening

The molecules comprising small-molecule screening collections may be obtained from natural (natural products) or nonnatural (chemical synthesis) sources [4].

26.2.2.1 Natural Products

Natural products show a wide range of biological activities and have been used medicinally throughout the course of human history [4]. However, there are problems associated with using natural products in screening experiments; they can be difficult to source, and the identification, purification, and chemical modification (to improve properties, e.g., potency) of the bioactive components can be very challenging [1]. In addition, there may be areas of chemical space (Box 26.1) that nature has ignored, which nonetheless contain compounds with biologically interesting properties; such compounds would not be detected if screening was limited to natural products only. Thus, it is not always realistic, or indeed desirable, to generate, and screen libraries consisting solely of natural products [4].

Box 26.1 Chemical Space

Chemical compounds can be characterized by a wide variety of molecular "descriptors" such as physiochemical properties (e.g., their lipophilicity) and topological features (e.g., degree of branching) [5, 6]. The term *chemical space* is commonly used in place of "multi-dimensional descriptor space": it is a region defined by a particular selection of molecular descriptors and the upper and lower values (limits) placed upon them [5, 6]. In the context of small-molecule collections, chemical space is generally defined as the total descriptor space that encompasses all small carbon-based molecules that could, in principle, be created [5, 6]. Each chemically unique small molecule will have a unique set of molecular properties and thus molecular descriptor values, and will therefore reside at a discrete point in chemical space [5].

26.2.2.2 Chemical Synthesis and the Need for Structural Diversity

Deliberate chemical synthesis represents an important alternative means of obtaining small-molecule libraries for biological screening [4]. But what molecules should be synthesized? Between the late 1980s and mid-1990s, a "brute force" approach was adopted; libraries of large numbers of compounds (literally millions in some cases) could be efficiently produced by combinatorial-type methods and these were routinely screened using high-throughput methods [7]. However, libraries of this sort have had limited success in the discovery of novel biologically active small molecules [4]. This has been largely attributed to the lack of structural variation between the compounds within such collections. It is now widely acknowledged that the success of any screening process (in terms of the hit frequency) is inherently dependent upon the structural diversity of the

library used; the size of the library is not everything [1, 8, 9]. There is a direct correlation between the overall structural diversity of a small-molecule library and its functional diversity (i.e., the range of biological activities displayed by the library). High functional diversity is clearly valuable in screens where the precise nature of the biological target is unknown or ill defined (e.g., a phenotypic screen) [1]. The presence of multiple structural classes within a molecular collection being screened against a single, specific target also increases the probability of discovering a molecule capable of binding in a novel manner [4].

Why should the functional diversity of a small-molecule library be related to its overall structural diversity? Biological macromolecules interact with each other in a three-dimensional (3D) manner [2]. On a molecular scale, biomolecules can be thought of as large 3D environments with certain defined potential binding regions. Consequently, they will only interact with small molecules that display a complementary 3D binding surface [4]. That is, a given biomolecule imposes a degree of shape selection for binding partners [1, 5, 8, 9]. Thus, the 3D shape of a small molecule is the most important factor controlling its biological effects [1]. Molecular shape is dictated by molecular structure [1, 7]. Structurally diverse libraries should therefore contain compounds with a diverse range of distinct molecular shapes; consequently, the library as a whole would cover a broad range of potential biological binding partners [4, 10].

There are four principal components of structural diversity that are typically identified [1]:

1) **Appendage diversity** – variation in structural moieties around a common scaffold
2) **Functional group diversity** – variation in the functional groups present
3) **Stereochemical diversity** – variation in the orientation of potential macromolecule-interacting elements
4) **Scaffold diversity** – presence of many distinct molecular scaffolds.

Scaffold diversity is the most crucial of these aspects in terms of the functional diversity of a library [10]. The shape–space coverage of any compound set (and thus its functional diversity) stems mainly from the nature and 3D geometries of the central scaffolds, with the peripheral substituents being of considerably less importance in this regard [7, 10]. Traditional combinatorial libraries typically possessed low levels of scaffold diversity; the molecules in such collections were broadly similar structures, with structural variation restricted to the presence of different appendages around a common scaffold. This explains their poor performance in many biological screens, especially those where the precise nature of the biological target was poorly defined or unknown [7].

Many commercially available and proprietary compound collections are synthesized in a combinatorial-type manner and so suffer from a lack of structural (principally scaffold) diversity [4, 11]. Another drawback of such collections stems from the nature of medicinal chemistry research over the course of the past few decades, which has focused upon a limited set of biological targets [7]. As a result, commercially available and proprietary compound libraries are

often heavily biased toward compounds that satisfy predefined criteria for the modulation of such "traditional" targets (e.g., the Lipinski "rule of 5" criteria for oral bioavailability [1, 7, 12]). Consequently, these libraries are intrinsically biased toward *known* bioactive chemical space (the chemical space spanned by known biologically active molecules), leaving potentially large swathes of biologically relevant chemical space underexplored. There is a widespread acknowledgement that the targets of the current pharmacopeia represent only a small fraction of potential targets that could impact on health [7, 11]. There are many other human-disease-related targets (such as protein–protein interactions), which have traditionally been termed *undruggable* as they have proved difficult, if not impossible, to address with small molecules [7, 11, 13]. However, it is becoming increasingly evident that these targets are indeed tractable to small molecule modulation; it is simply that they have traditionally been challenged with the wrong types of molecules [2, 11]. The molecules comprising typical commercially available and proprietary compound collections seem to be well suited to modulating "traditional" medicinal chemistry targets, but lack the structural features necessary to affect other processes [7, 11, 13].

26.3 The Chemical Approach

26.3.1 Diversity-Oriented Synthesis

Diversity-oriented synthesis (DOS) was developed over the past decade in order to address the need for new small-molecule collections with higher levels of structural, and thus functional, diversity [2, 14, 15]. DOS libraries aim to efficiently interrogate large areas of chemical space simultaneously. This includes known biologically relevant chemical space (by definition, a fruitful region for the discovery of useful small-molecule modulators of biological function) *and* under-explored (and, indeed, completely novel) regions of chemical space, which may contain molecules with unusual or exciting biological properties (e.g., the capability to modulate classically "undruggable" targets) [1, 7]. The screening of such libraries should provide hits against a broad range of biological targets with increased frequency and decreased cost relative to less diverse libraries, facilitating the discovery of new agents for therapeutic intervention and novel probes for biological research [1, 11].

26.3.1.1 DOS and Scaffold Diversity

Ideally, a DOS should address all four of the principal types of structural diversity mentioned previously. However, the ideal synthesis of a structurally diverse small molecule collection is one in which the diversity is achieved in the most efficient manner possible [1]. As alluded to previously, it is the scaffold diversity of the library that is the key parameter in this regard. It is generally acknowledged

that increasing the scaffold diversity in a small-molecule library is one of the most effective ways of increasing its overall structural diversity (and, consequently, its shape, and thus functional diversity) [1, 10]. Thus, the efficient incorporation of multiple molecular scaffolds in a single library is of central importance to the success of a DOS. This is undoubtedly the most challenging facet of any DOS program [1, 2].

There are two basic strategies for generating scaffold diversity in a DOS context (Figure 26.1). The first is a "branching" approach, where divergent reactions are carried out on a substrate to produce compounds with distinct molecular scaffolds. The second is a "folding" approach, where intramolecular reactions are used to "pair up" strategically positioned functional groups. This could involve either the use of different starting materials and common reaction conditions, such that each starting material yields a product containing a different molecular scaffold, or a densely functionalized molecule where different functional groups can be reacted together under distinct reaction conditions and so create a number of different scaffolds. These strategies are not orthogonal to each other and many

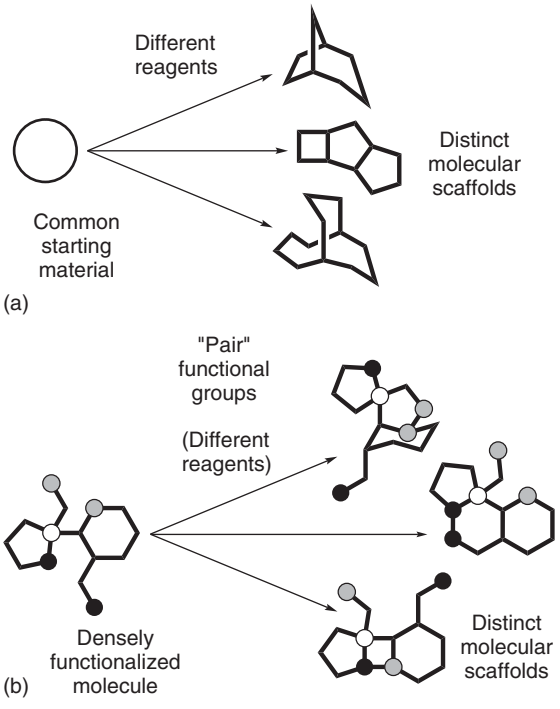


Figure 26.1 *Examples of strategies for generating scaffold diversity in DOS.* (a) An example of the "branching" approach. Here, the exposure of a given starting material to different reagents results in the generation of different molecular scaffolds. (b) An example of the "folding" approach. Here, different "complementary" functional groups (indicated by different colored circles) of a densely functionalized molecule are reacted together (the "pair" process) in functional group-specific intramolecular reactions, to yield different scaffolds [8].

DOS programmers will utilize both [1, 4]. The resulting products should ideally contain synthetic handles for further transformations, thereby providing scope for additional diversification [5]. Variation in starting materials and/or reagents allows for the introduction of appendage, functional group, and stereochemical diversity (the latter of which may also be incorporated through the use of stereoselective reactions) [1].

26.4 Chemical Biology Research

26.4.1 DOS as a Tool for Identifying New Modulators of Mitosis

Antimitotic compounds are used clinically for the treatment of cancer, and this target class is widely regarded to still hold great promise for anticancer therapy. How can new, structurally distinct antimitotic agents be identified? Recently, Spring and coworkers described the discovery of new small-molecule modulators of mitosis using DOS, illustrating the utility of this synthetic approach for the identification of new biologically relevant chemical entities [16].

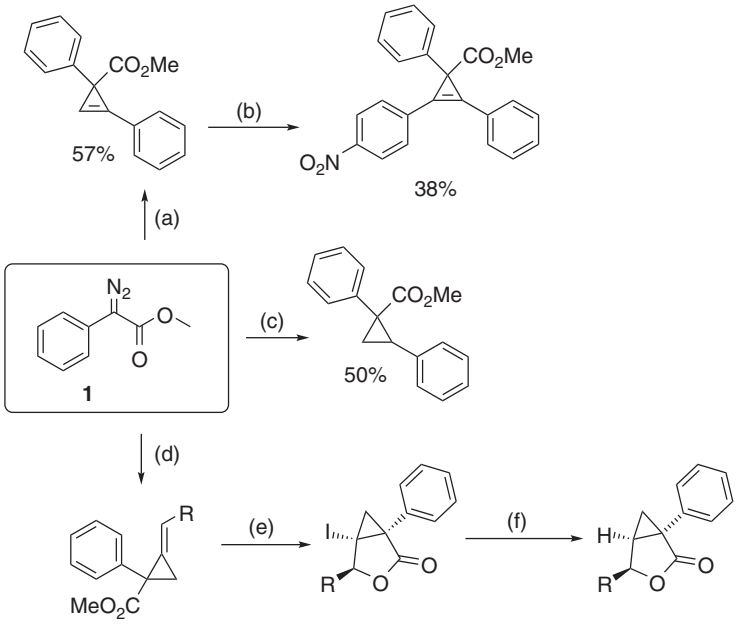
26.4.1.1 DOS Library Synthesis

Diazoacetates represent attractive starting units for the branching DOS pathways [5, 17]. The diazoacetate functionality exhibits enormous synthetic versatility, permitting the use of a wide variety of different synthetic transformations. Thus, diazoacetate compounds have the potential to be converted into several products with different scaffolds, which should themselves be suitable for further diversification [5].

Spring and coworkers recently reported the use of two different, readily accessible, phenyldiazoacetate compounds (**1** and **2**) as starter units for two different branching DOS pathways (Schemes 26.1 and 26.2) [16]. The second pathway (Scheme 26.2) utilized the highly functionalized derivative **3** as a key branch-point intermediate. The presence of both an electron deficient and an electron neutral olefin, coupled with the proximal aryl bromide and a carboxylic ester, afforded the opportunity to regioselectively modify the scaffold of **3** in a multidirectional approach. Overall, these two DOS pathways combined generated a library totaling 35 small molecules, with 10 distinct molecular scaffolds, comprising complex fused ring systems of varying sizes and a multiplicity of stereocenters present. Cheminformatic analysis of the library indicated that it accessed biologically relevant areas of chemical space and had a good level of shape diversity.

26.4.1.2 Biological Studies: Phenotypic Screening for Antimitotic Effects

The DOS library compounds were screened for their ability to induce mitotic arrest in human osteosarcoma cells (U2OS line) [18, 19]. Cells were incubated with compounds at a range of concentrations and then stained for the mitotic

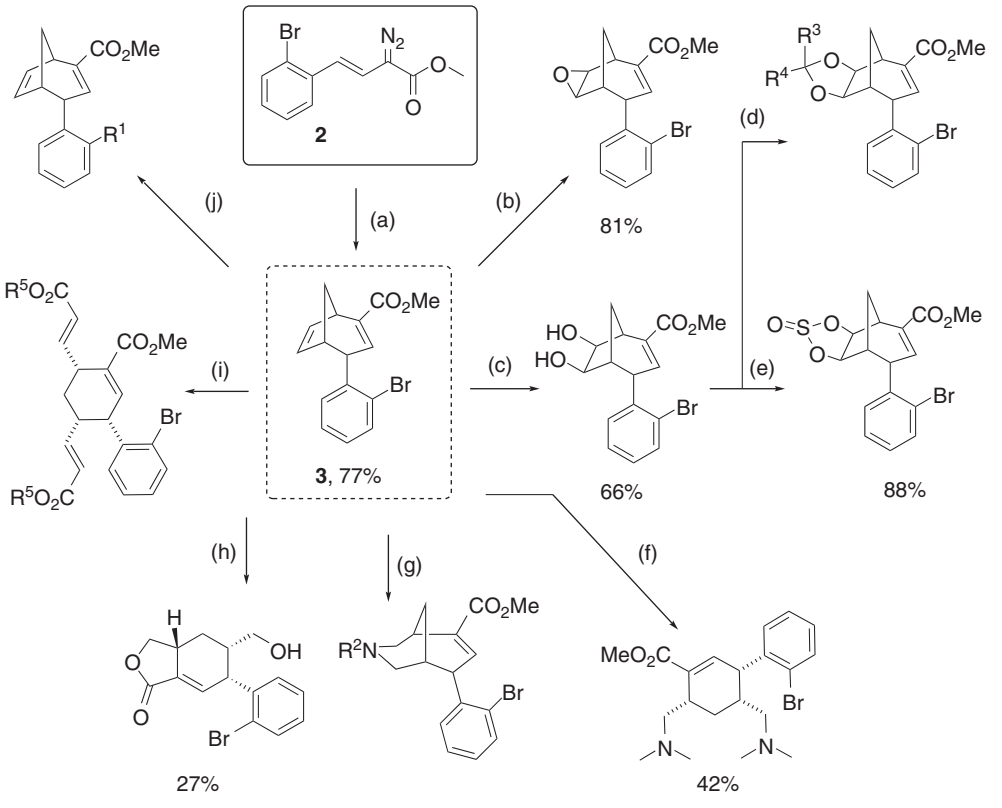


Scheme 26.1 Synthesis of subset of DOS library from compound **1**; (a) phenylacetylene, $Rh_2(OAc)_4$ (1 mol%), CH_2Cl_2; (b) *p*-nitroiodobenzene, $Pd(OAc)_2$ (10 mol%), K_2CO_3, DMF; (c) styrene, $Rh_2(OAc)_4$ (1 mol%), CH_2Cl_2; (d) allene, $Rh_2(OAc)_4$ (1 mol%), CH_2Cl_2; (e) *N*-iodosuccinimide, MeCN-H_2O (2 : 1), 50 °C; and (f) Bu_3SnH, AIBN, PhH, 80 °C. DMF: *N,N*-dimethylformamide, AIBN: 2,2′-azobis(2-methylpropionitrile).

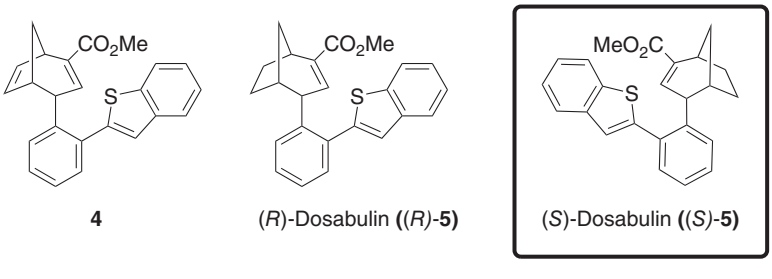
marker phosphohistone H3 and imaged on a Cellomics Arrayscan high-content microscope. The percentage of cells arrested in mitosis following compound treatment was then calculated by image analysis. The most potent compound (**4**) gave a large (35–40%) mitotic arrest (Figure 26.2 for structure, Table 26.1 for screening data). On the basis of this result, the partially saturated analog of **4**, compound **5**, was prepared in a racemic form (Figure 26.2). Compound **5**, subsequently termed *dosabulin*, was also found to give a mitotic arrest in U2OS cells, with a twofold increase in potency compared to **4**. Treatment with dosabulin also resulted in growth inhibition in the low micromolar range over a period of 72 h (Figure 26.3, Table 26.1). Separation of both enantiomers of dosabulin by preparative chiral high-performance liquid chromatography (HPLC) and subsequent retesting, revealed that all the activity resided in the *(S)*-enantiomer (Figure 26.3, Table 26.1). It was found that *(S)*-dosabulin treated cells died through apoptosis while cells treated with *(R)*-dosabulin did not.

26.4.1.3 **Biological Studies: Target Identification**

While phenotypic screening allows for the rapid identification of biologically active molecules from a library, subsequent target identification (identification of the biological target(s) that interact with a compound of interest) is notoriously



Scheme 26.2 Synthesis of subset of DOS library from compound **2**; (a) cyclopentadiene, $Rh_2(OAc)_4$ (1 mol%), CH_2Cl_2; (b) *m*-CPBA, CH_2Cl_2; (c) OsO_4 (2.5 mol%), NMO, acetone-H_2O (9 : 1); (d) aldehyde/ketone, CSA (10 mol%), 3 Å molecular sieves, CH_2Cl_2; (e) $SOCl_2$, CH_2Cl_2; (f) 2,6-lutidine, NMO, OsO_4 (2.5 mol%), $PhI(OAc)_2$, acetone-H_2O (10 : 1), then dimethylamine, $NaBH(OAc)_3$, CH_2Cl_2; (g) 2,6-lutidine, NMO, OsO_4 (2.5 mol%), $PhI(OAc)_2$, acetone-H_2O (10 : 1), then primary amine, $NaBH(OAc)_3$, CH_2Cl_2; (h) 2,6-lutidine, NMO, OsO_4 (2.5 mol%), $PhI(OAc)_2$, acetone-H_2O (10 : 1), then $NaBH_4$, MeOH; (i) alkene, Hoveyda-Grubbs (II) catalyst (10 mol%), ethylene, PhMe, 100 °C; and (j) $Pd(OAc)_2$ (10 mol%), boronic acid, PPh_3 (15 mol%), 2N K_2CO_3, PhMe, 90 °C. NMO: *N*-methylmorpholine-*N*-oxide, CSA: camphorsulfonic acid.



4 (*R*)-Dosabulin (**(*R*)-5**) (*S*)-Dosabulin (**(*S*)-5**)

Figure 26.2 Structures of some antimitotics from the DOS library.

Table 26.1 Mitotic arrest (EC_{50}) and growth inhibition (IC_{50}) values for selected compounds from the DOS library.

Compound	Mitotic index (MI) EC_{50} (μM)	Growth inhibition $(GI)_{50}$ (μM)
4	6.25 ± 0.91	3.70 ± 0.71
(*rac*)-Dosabulin	3.13 ± 0.32	1.47 ± 0.03
(*R*)-Dosabulin	N/A	N/A
(*S*)-Dosabulin	1.23 ± 0.10	0.81 ± 0.37

EC_{50}, effective concentration 50; IC_{50}, inhibitor concentration 50.
Growth inhibition assessed by sulforhodamine B colorimetric assay for cytotoxic effects. All values are mean ± standard deviation. N/A = not active.

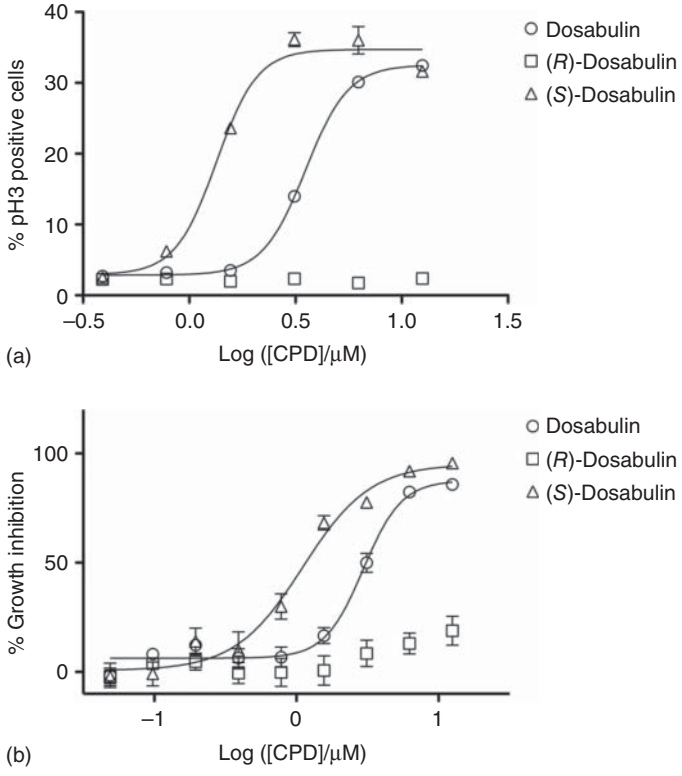


Figure 26.3 (a) Representative mitotic index assay data for racemic dosabulin and its purified enantiomers. Data points are mean ± standard error in the mean of an experiment conducted in triplicate. "% PH3 positive cells" refers to the proportion of cells stained with an antibody against phosphor-histone H3. "CPD" = compound under investigation. (b) Growth inhibition curves assessed by sulforhodamine B assay for racemic dosabulin and its purified enantiomers. Data points are mean ± standard error in the mean of an experiment conducted in triplicate. "CPD" = compound under investigation.

difficult [5, 19]. However, careful observation of the phenotype may sometimes offer clues [20]. Spring and coworkers used confocal microscopy to look at the key mitotic protein, tubulin; it was found that the tubulin network was heavily disrupted upon treatment with *(S)*-dosabulin. This led to the hypothesis that *(S)*-dosabulin was targeting tubulin itself. This phenotype was recapitulated by nocodazole, a known tubulin depolymerizer, providing indirect evidence for this assertion. Existing agents targeting tubulin suffer from administration and resistance problems; thus, new antimitotics-targeting tubulin are of significant therapeutic interest [21]. An *in vitro* tubulin polymerization assay established that *(S)*-dosabulin acts as a tubulin depolymerizing agent. Several small-molecule-binding sites are known to exist in the tubulin polymer [22]. For example, vinblastine binds the β-tubulin subunit, while colchicine binds at the α/β interface. Further work demonstrated that *(S)*-dosabulin partially inhibits the binding of colchicine to tubulin, suggesting that it may bind in a site vicinal or allosteric to colchicine.

26.5 Conclusion

Over the course of the past decade, DOS has established itself as a powerful tool for the efficient *de novo* creation of structurally, and thus functionally, diverse small molecule collections. Many ingenious DOS strategies have been reported, which have enabled the efficient synthesis of libraries based on tens of different molecular scaffolds; and the screening of DOS libraries has led to the identification of numerous bioactive small molecules (including modulators of a range of undruggable targets) [7]. For example, Spring and coworkers used DOS to discover dosabulin, a novel small molecule that causes mitotic arrest and cancer cell death by apoptosis at submicromolar concentrations.

A key challenge for future DOS campaigns is to improve the balance between broad chemical space coverage and biological relevance [1]. A DOS should aim to specifically and efficiently access *both* known and unknown *biologically relevant* chemical space, rather than regions that are not going to provide biologically useful small molecules [1]. Toward this end, future years may witness the emergence of more "constrained" DOS campaigns that seek to generate maximum structural diversity within preselected limits, such that a better balance between structural (scaffold) diversity (which is valuable for broad bioactive chemical space coverage) and target relevance and/or drug likeness is achieved.

References

1. Galloway, W.R.J.D., Isidro-Llobet, A., and Spring, D.R. (2010) Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. *Nat. Commun.*, **1**, 80.
2. O'Connor, C.J., Beckmann, H.S.G., and Spring, D.R. (2012) Diversity-oriented synthesis: producing chemical tools for dissecting biology. *Chem. Soc. Rev.*, **41**, 4444–4456.
3. O'Connor, C.J., Laraia, L., and Spring, D.R. (2011) Chemical genetics. *Chem. Soc. Rev.*, **40**, 4332–4345.
4. O'Connell, K.M.G., Galloway, W.R.J.D., Ibbeson, B.M., Isidro-Llobet, A., O'Connor, C.J., and Spring, D.R.

(2012) in *Solid-Phase Organic Synthesis:Conepts, Strategies and Applications* (eds P.H. Toy and Y. Lam), John Wiley & Sons, Inc., Hoboken, NJ, pp. 131–150.
5. Galloway, W.R.J.D., Bender, A., Welch, M., and Spring, D.R. (2009) The discovery of antibacterial agents using diversity-oriented synthesis. *Chem. Commun.*, 2446–2462.
6. Dobson, C.M. (2004) Chemical space and biology. *Nature*, **432**, 824–828.
7. Galloway, W.R.J.D. and Spring, D.R. (2013) Towards drugging the undruggable: enhancing the scaffold diversity of synthetic small molecule screening collections using diversity-oriented synthesis. *Divers. Oriented Synth.*, **1**, 21–28.
8. Burke, M.D. and Schreiber, S.L. (2004) A planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed.*, **43**, 46–58.
9. Burke, M.D., Berger, E.M., and Schreiber, S.L. (2003) Generating diverse skeletons of small molecules combinatorially. *Science*, **302**, 613–618.
10. Sauer, W.H. and Schwarz, M.K. (2003) Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. *J. Chem. Inf. Comput. Sci.*, **43**, 987–1003.
11. Schreiber, S.L. (2009) Organic chemistry: molecular diversity by design. *Nature*, **457**, 153–154.
12. Lipinski, C. and Hopkins, A. (2004) Navigating chemical space for biology and medicine. *Nature*, **432**, 855–861.
13. Altmann, K.H., Buchner, J., Kessler, H., Diederich, F., Krautler, B., Lippard, S., Liskamp, R., Muller, K., Nolan, E.M., Samori, B., Schneider, G., Schreiber, S.L., Schwalbe, H., Toniolo, C., van Boeckel, C.A., Waldmann, H., and Walsh, C.T. (2009) The state of the art of chemical biology. *ChemBioChem*, **10**, 16–29.
14. Lee, D., Sello, J.K., and Schreiber, S.L. (2000) Pairwise use of complexity-generating reactions in diversity-oriented organic synthesis. *Org. Lett.*, **2**, 709–712.
15. Schreiber, S.L. (2000) Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science*, **287**, 1964–1969.
16. Ibbeson, B.M., Laraia, L., Alza, E., O'Connor, C.J., Tan, Y., Davies, H.M.L., McKenzie, G., Venkitaraman, A.R., and Spring, D.R. (2014) Diversity-oriented synthesis as a tool for identifying new modulators of mitosis. *Nat. Commun.*, **5**, 3155.
17. Wyatt, E.E., Fergus, S., Galloway, W.R.J.D., 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.
18. Young, D.W., Bender, A., Hoyt, J., McWhinnie, E., Chirn, G.W., Tao, C.Y., Tallarico, J.A., Labow, M., Jenkins, J.L., Mitchison, T.J., and Feng, Y. (2008) Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nat. Chem. Biol.*, **4**, 59–68.
19. Ziegler, S., Pries, V., Hedberg, C., and Waldmann, H. (2013) Target identification for small bioactive molecules: finding the needle in the haystack. *Angew. Chem. Int. Ed.*, **52**, 2744–2792.
20. Neumann, B., Walter, T., Heriche, J.K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U., Cetin, C., Sieckmann, F., Pau, G., Kabbe, R., Wunsche, A., Satagopam, V., Schmitz, M.H.A., Chapuis, C., Gerlich, D.W., Schneider, R., Eils, R., Huber, W., Peters, J.M., Hyman, A.A., Durbin, R., Pepperkok, R., and Ellenberg, J. (2010) Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature*, **464**, 721–727.
21. Dumontet, C. and Jordan, M.A. (2010) Microtubule-binding agents: a dynamic field of cancer therapeutics. *Nat. Rev. Drug Discovery*, **9**, 790–803.
22. Gigant, B., Wang, C., Ravelli, R.B.G., Roussi, F., Steinmetz, M.O., Curmi, P.A., Sobel, A., and Knossow, M. (2005) Structural basis for the regulation of tubulin by vinblastine. *Nature*, **435**, 519–522.