# An Introduction to Diversity-Oriented Synthesis

**Richard J. Spandl,[a] Gemma L. Thomas,[a] Monica Diaz-Gavilan,[a] Kieron M. G. O'Connell[a] and David R. Spring[a]***

[a] *Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK.*
*\* To whom correspondence should be addressed.*

## Table of Contents

## 1.1 Introduction

*Diversity-oriented synthesis* (DOS) aims to synthesize collections of molecules that represent the variety of charges, polarities, bonding interactions, and architectures that can potentially be recognised by nature's biomolecules.[1-3] The structural variety present in DOS libraries confers both physicochemical and biological diversity to the compound collection.[4] DOS libraries can be used in screening experiments to identify novel biologically active small molecule modulators.[5-8] Although there are other potential applications, it is in this context that DOS will be discussed.

Since the publication of the completed human genome sequence in 2004,[9] computational efforts towards annotation have been less than trivial.[10] Although the long established techniques of traditional genetics have a part to play, the more recently developed technique of *chemical genetics* is also predicted to be of use.[11-14] Chemical genetics, instead of using gene knock-outs directly on the level of the DNA, employs chemical methods to perturb the corresponding gene products (proteins). Small molecules can therefore be used to have a modulating effect on proteins and, consequently, the biological system under investigation can be dissected. As with traditional genetic approaches, chemical genetics can be performed in either a forward sense (*i.e.* induce a phenotype

and identify the protein target) or reverse sense (*i.e.* perturb a protein and observe the phenotype) (Fig. 1).

**[Fig. 1 Best Inserted Here]**

Although chemical genetics has been exploited successfully on an *ad hoc* basis,[13] its use is restricted by a lack of generality. To illustrate this point, of the 25,000 human genes proposed to encode proteins which will bind drug-like molecules (the "druggable genome" *i.e.* 10% of the total),[15] only about 1,000 have known small molecule partners. Furthermore, of these interactions, fewer can be considered as specific with regards to their effects on other proteins. Promiscuity of this nature will be detrimental to a chemical genetics experiment and is more likely when a protein with large families of homologues (*e.g.* proteases or kinases) is being investigated. With this in mind, it has been proposed that using complex small molecules aids interaction specificity;[16] there is however debate in the literature about this point.[17, 18]

The data above demonstrates that the lack of specific small molecule modulators is a major limiting factor in increasing the generality of chemical genetic approaches. Access to skeletally diverse libraries would therefore (potentially) be advantageous, especially if the most ambitious aim of chemical genetics is to be realised; *i.e.* the identification of small molecule partners for every known protein ('*chemical genomics*').[19]

Although there are a number of potential sources of small molecules (discussed below), *de novo* library synthesis using DOS may be an alternative. DOS describes the deliberate, simultaneous, and efficient synthesis of multiple targets that are not only diverse in the appendages they display, but also in their molecular architectures.[2] Thus, the aim of DOS is to synthesize skeletally diverse collections of small molecules that interrogate areas of *chemical space*, in the natural product-like/drug-like region (in a broad sense), that have not previously been explored.

## 1.2 Exploring Chemical Space

With the aim of many screening projects being to identify novel biologically active small molecules, success may be aided by screening biologically (functionally) diverse compound libraries. It is the skeletal diversity of a library that has been shown to correlate to its biological diversity.[4] As a method of assessing library quality, library diversity has been calculated using chemical space analysis.[20, 21]

Chemical space encompasses all possible chemical entities and can be defined and analysed computationally.[22] The location of a molecule in chemical space is a function of the abstract representation created by an analysis of the compound's associated *chemical descriptors*. Chemical descriptors can be used to describe not only the bulk properties of a molecule,[23] but also its 3D arrangement in space (*i.e.* topological features).[24] Once the chemical descriptors (of which there will be many) have been defined and calculated for each library member, this information can be condensed using principle component analysis (PCA). This allows for the construction of visually gratifying 2D or 3D displays that are accessible to human interpretation. In these representations each molecule is plotted at a discrete point in chemical space (more correctly called multidimensional descriptor space) and the relative proportion of chemical space covered by different compound collections can be compared. As a result library diversity can be analysed.[4, 20, 22, 25]

As an example of chemical space representation, cyclooxygenase-1 inhibitors, defined and analysed as described above, were plotted on a background of pharmacologically active compounds (MDL Drug Data Repository) (Fig. 2). Examination of these results demonstrates that these inhibitors are diverse in structure and not clustered in a tightly defined region.

**[Fig. 2 Best Inserted Here]**

Since a variety of chemical descriptors can be used in the construction of chemical space representations, there may be ambiguity in these assessments of diversity. For example, in our experience, using certain chemical descriptors, a library of amides synthesized (hypothetically) from a wide variety carboxylic acids and amines appeared extremely diverse on analysis. Since this diversity is the result of the building blocks used and not the bond forming reaction itself, this highlights the problem of programming human intuition (*i.e.* diversity assessment) into computers. Although diversity assessment may be useful, it can also be misleading. Like-with-like comparisons are often thwarted as it is difficult to say exactly what constitutes diversity and also to know how it should be calculated.

When the target molecule in a screening programme is unknown, *i.e.* in forward chemical genetics, screening skeletally diverse collections of compounds may be useful. Although achieving skeletal diversity *per se* is rarely the 'end-game', its incorporation may increase hit identification.

## 1.3 Sources of Skeletally Diverse Small Molecules

Although DOS can be used to prepare skeletally diverse small molecules libraries, there are alternatives to this recently developed, and sometimes challenging, approach. For example, natural products, combinatorial libraries, and proprietary compound collections may provide small molecule collections that offer both complexity and diversity. There are however potential drawbacks.

Natural products are diverse in structure and have specific modulating effects on biomolecules. Their isolation, purification, and characterisation can, however, be a very complicated process.[26, 27] Furthermore, the chemical derivatisation of natural products, as a result of the complex multi-step syntheses often required, can be extremely challenging. Such chemical derivatisation is important in *focused-library synthesis*; this technique gives access to novel compounds based around the original natural product scaffold and can be used in lead-optimisation or structure-activity-relationship (SAR) studies.

Combinatorial libraries, synthesized using *traditional combinatorial chemistry*, although they offer complexity and quantity, tend to contain mainly flat molecules with fewer chiral centres than natural products or drugs. Analysis of the mean number of chiral centres, calculated from the analysis of various databases in terms of average per molecule, demonstrates this; natural products, drugs, and combinatorial library members were found to contain 6.2, 3.3, and 0.4 chiral centres respectively.[28] Likewise, a comparison of the average molecular weights (414:340:393) and number of rings (4.1:2.6:3.2) also highlights both differences and similarities between these compound types.

Combinatorial libraries tend to be synthesized using a 'one-synthesis/one-skeleton'[29] approach and, as a result, potentially exhibit limited skeletal diversity. This potential limitation can be offset by employing many chemists to perform many syntheses; the proprietary compound collections of pharmaceutical companies are, therefore, diverse in

structure. As a result of being biased by previous drug discovery programmes[30] or by meeting pre-defined criteria, *e.g.* the Lipinski rule of 5 (RO5),[31] these compound collections may, however, have their drawbacks. In relation to the latter point, there has been debate about the value of restricting chemists to synthesising compound that meet the RO5 criteria.[32, 33]

## 1.4 Enriching Chemical Space Using DOS

Since natural products and known drugs occupy only a small proportion of bioactive chemical space,[16, 22] exploring previously uncharted regions, as is the aim in DOS, may be advantageous. When bioactive molecules are sought, although DOS aims to interrogate 'novel' regions of chemical space, the compounds prepared should still be 'natural product-like' or 'drug-like' in terms of structure. The term 'diversity' must therefore be taken in context.

A comparison with *target-oriented synthesis* (TOS) and traditional combinatorial chemistry (focused-library synthesis) serves to highlight the DOS approach to populating chemical space. TOS and focused library synthesis attempt to synthesize compounds at a discrete point in, or clustered within a certain region of, chemical space. In these approaches *retro-synthetic analysis* is used to theoretically deconstruct complex target molecules into simple starting materials. In DOS, a '*forward synthetic*' strategy must be envisaged to facilitate the transformation of simple and similar starting materials into an array of complex and diverse products (Fig. 3).

**[Fig. 3 Best Inserted Here]**

To be an effective strategy for library synthesis, DOS, as is the case in TOS and focused library synthesis, requires highly efficient reactions in terms of both yield and stereoselectivity.

## 1.5 The Subjective Nature of 'Diversity'

Libraries synthesized using DOS are not 'truly' diverse, nor are they designed to be. It may be better therefore, instead of describing DOS libraries as 'non-focused', to describe them as '*soft-focused*'. Although the aim of DOS is to populate diverse regions of chemical space, the libraries constituents are, nevertheless, required to interact with biomolecules (in the case of a biological modulator or probe).

As an extension of this, the terms 'diversity' and DOS can be confusing since both are used freely in the literature. For example, the racemic synthesis of a chiral target molecule could be classified as a DOS. Also, when any compound collection is synthesized, since the constituents are not identical, there is some degree of diversity incorporated. Although these are extreme cases, they highlight the subjectivity of diversity.

To address these issues we envisaged the '*molecular diversity spectrum*' (Fig. 4).[1] In one extreme of the spectrum would be where maximal chemical space coverage has been achieved and, in the other extreme, would be a TOS.[1] Although quantification of this spectrum would be difficult (see above), it should be the goal of a DOS to synthesize, in a qualitative sense, collections as near to the right hand side of this spectrum as possible.

**[Fig. 4 Best Inserted Here]**

The four principle types of diversity that can be incorporated into a compound collection are highlighted on the molecular diversity spectrum, these are: 1) appendage (building block) diversity, 2) functional group diversity, 3) stereochemical diversity, and 4) skeletal diversity.[29, 34] Whereas the first three types of diversity can be introduced by using reagent controlled (stereoselective) transformations, the most challenging facet of DOS is the incorporation of skeletal diversity. It is skeletal diversity which is of critical importance in a DOS project. Thus, in contrast to traditional combinatorial chemistry where diversity primarily arises by building block variation, DOS places emphasis on accessing different molecular skeletons.

## 1.6 Differing Strategies Towards Similar Goals

A successful DOS should allow not only for combinatorial variation in the building blocks used, but should also allow for these appendages to be displayed on different three dimensional scaffolds. In a general sense, most of the DOS strategies reported in the literature are either based on privileged scaffolds or designed from simple starting materials. These strategies, *i.e.* 'DOS based on privileged scaffolds' and 'DOS from simple starting materials', are different in their approaches and, to some extent, their end-goals.

### 1.6.1 DOS based on privileged scaffolds

As a result of the evolutionary pressure for specific ligand-biomolecule interactions, compounds bearing privileged scaffolds are predisposed to exhibit bioactivity. Privileged structures are defined as those which display high affinity binding to multiple protein classes; the term was first used to describe the benzodiazepines and the benzazepines by Evans and co-workers.[35] In some instances, basing libraries on privileged scaffolds may be advantageous.[36-40] In contrast to combinatorial natural-product like libraries, where limited diversification is explored,[41] DOS libraries of this nature aim to incorporate privileged motifs into structurally diverse (skeletally and stereochemically) architectures. The aim of these investigations is to identify compounds with novel bioactivities that are distinct from the natural product (or class of natural products) in which the privileged motifs are commonly found.[42]

This approach was exploited by Park and co-workers in their synthesis of a natural product-like DOS library based around the benzopyran motif **1**.[38] The benzopyran motif is common in many natural products (4636 of the compounds from the '*dictionary of natural products*' are benzyopyraniod)[43] and the synthesis of bioactive benzopyran containing compounds has been extensively studied. Using a branching strategy (also referred to as a library-from-library approach), Park and co-workers were able to synthesize 22 novel molecular architectures, each one containing the benzopyran sub-structure. From the starting materials **2** and **3**, the scaffolds **4**-**14** could be accessed using one of two major branching pathways (pathway A or B) (Scheme 1).

## [Scheme 1 Best Inserted Here]

The reactions used in library generation were atom economical and proceeded in good yields and with good diastereoselectivities. To investigate the effect on bioactivity of using unbiased and diverse natural product-like libraries, *in vitro* cytotoxicity studies were performed against a human cancer cell line. The compounds screened showed a wide range of $IC_{50}$ values and these variations were shown to correlate to differences in the

molecular skeletons and not in the appendages attached. Thus, skeletal diversity led to functional (biological) diversity.

An alternative but related approach to discovering novel biologically active small molecules, known as '*biologically-oriented synthesis*' (BIOS), is noteworthy. This approach, pioneered by Waldmann and co-workers, "builds on the inspiration given by nature through natural products".[44] Since nature uses only a very small fraction of chemical and biological space, this allows small molecule modulators and functional proteins to be classified and grouped. For example, in the case of small molecules, the 'structural classification of natural products' (SCONP) is used to examine the relationships between the different biologically relevant scaffolds and also to group these skeletons in a hierarchical tree.[45] Likewise, using the technique of 'protein structure similarity clustering' (PSSC), proteins can be classified on the basis of their fold topology and also their inhibitory profiles.[46]

By merging these concepts, biologically pre-validated scaffolds can be chosen by identifying links with the target protein through architectural commonalties with other known inhibitors.[44]  This process, which is the essence of BIOS and has only been discussed here briefly, allows small focused libraries to be synthesized. Indeed, BIOS has shown some initial success.[47]

Although DOS libraries based on privileged scaffolds may offer advantages when the target protein is known, there is a need to explore the chemical space occupied by neither natural products nor drugs.[1, 4] As DOS libraries from simple starting materials contain no pre-encoded bias (*i.e.* no pre-selected motifs), these can give access to more diverse regions of chemical space than the approaches discussed above.

1.6.2 DOS from simple starting materials

The process described by Schreiber in much of his pioneering work in DOS, *i.e.* generating skeletally diverse libraries from simple starting materials, will be the subject of the remainder of this chapter.[48, 49] As a general strategy to incorporate diversity, the use of branching pathways to access distinct molecular scaffolds will be discussed.[3]

Branching pathways commonly make use of reactions that increase structural complexity, *i.e.* complexity generating reactions. Although the complexity of a compound collection bears no relation to its diversity, it has been reported to confer specificity (see above).[16, 50] Tandem processes, where the product of one complexity generating reaction is the substrate for the next, are of immense value in DOS as both structural complexity and diversity are increased efficiently.[29] This allows the complex 3D scaffolds required to be generated. Although structural complexity is relatively straightforward to achieve, accessing skeletal diversity efficiently is a more challenging goal.

**1.7 Generating Skeletal Diversity**

When synthesising a library, skeletal diversity can be incorporated using either: branching pathways, where a common starting material is transformed into distinct skeletons using different reagents (the 'reagent based approach'); or, folding pathways, where different starting materials, containing pre-encoded skeletal information, are subjected to a common set of conditions and converted to different scaffolds (the 'substrate based approach') (Fig. 5).[34]

Successful DOS processes utilise these approaches to achieve skeletal diversity in a number of ways; a review of the literature suggests three general strategies. These are: 1) the use of a pluripotent functional group where the same part of a molecule is subjected to different transformations induced by different reagents; 2) the use of a densely functionalised molecule where different functionalities in the same molecule are transformed by different reagents; or, 3) the use of a folding process where different structurally encoding elements, contained in different substrates, are subjected to the same reaction conditions.

Schreiber and co-workers have recently described a 'build/couple/pair' strategy which combines approaches 2 and 3.[51] After firstly 'building' the required chiral starting units, possibly from chiral pool reagents, 'coupling' reactions afford a densely functionalised molecule; it is common for multi-component coupling reactions (MCRs) to be used. Different parts of the densely functionalised molecule, in functional group specific reactions, are then 'paired' to generate different molecular skeletons.

### 1.7.1 Strategy 1: Pluripotent Functional Groups

After a suitable functional group, which can participate in a rich variety of chemical transformations, has been chosen, other factors must also be considered for this approach to be successful. The synthesis of the starting material should allow for building block diversity to be incorporated and, more importantly, the novel scaffolds generated in the initial branching pathways should have the potential to be diversified further (preferably in subsequent branching reactions). To this end, the high reactivity and mechanistic flexibility of the diazoacetate moiety was exploited in the generation of a small molecule library by Wyatt *et al.*[52]

The flourous-tagged diazoacetate **15** was chosen as it could participate in a variety of complexity generating C-C bond forming reactions (the products of which could be diversified further). Also, the polyfluorocarbon tag present in **15** allowed for generic purification of the library compounds by fluorous solid-phase extraction (SPE), reverse fluorous SPE, or liquid-liquid extraction. The starting material **15** therefore not only gave access to multiple branching pathways, but facilitated high-throughput purification and hence increased the efficiency of the DOS process.

From the starting unit **15**, three major branching pathways were initially used (step 1): the three-membered ring forming reaction, yielding the cyclopropanes **16** and **17**; the α-deprotonation/electrophilic quenching reaction, yielding the 1,3-keto esters **19** and **20**; and, the 1,3-dipolar cycloaddition reaction, yielding the heterocycles **21** and **22** (Scheme 2). These newly formed branch point products **16**, **17**, **19**, and **20** were then converted to a second generation of compounds **18**, **23**-**29** in sequential complexity generating reactions (step 2). These tandem processes thus served to increase complexity and diversity; a collection of 223 small molecules, based around 20 discrete molecular frameworks, was synthesized.

A skeletally diverse small molecule library was synthesized, also using a similar approach to that above, by Thomas *et al.* (Scheme 3).[53] The solid supported phosphonate **31** was chosen as the starting unit as it could be readily synthesized from **30** (in an *E*-selective

Horner Wadsworth Emmons reaction, step 1) and as it allowed access to three divergent reaction pathways.

The reactions from **31**, which were catalytic and enantioselective thus allowing stereochemical diversity to be incorporated, were: the 1,3-dipolar cycloaddition branching pathway to give **32** (performed using the AgOAc/(*S*)-QUINAP system);[54] the dihydroxylation branching pathway to give **33** (performed using a modified Sharpless asymmetric dihydroxylation (AD) reaction protocol);[55] and, the Diels Alder branching pathway to give **34** (asymmetric catalysis was possible as the nature of **31** allowed two-point catalyst binding *i.e.* using Evans methodology)[56] (step 2). Subsequent transformations enabled a second generation of compounds (step 3) and a third generation of compounds (step 4) to be synthesized; these processes allowed both the appendage and the skeletal diversity to be increased and afforded 242 compound based around 18 distinct scaffolds.

Similar to the flourous-tagged technology used by Wyatt *et al.*, the solid supported nature of these compounds (Scheme 3) allowed for high-throughput synthesis. The imidazolidinone portion allowed the appropriate chemical entity to be attached to a novel solid support resin.[57] Simple protocols (such as amide, acid, and alcohol synthesis) could be used to cleave the target compounds, at the exocyclic carbonyl of the imidazolidinone, from the immobilised portion of the molecule.

**[Scheme 3 Best Inserted Here]**

The diversities of the libraries synthesized by Wyatt *et al.* and Thomas *et al.* were assessed using selected chemical descriptors and PCA (Fig. 6A and 6B respectively). The diversities of these libraries were then compared to: 1) MDDR compounds (molecular weight cut-off 650); 2) two focused libraries (synthesized using traditional combinatorial chemistry); and (in the case of Fig. 6B only), the 3762 compounds marked as 'antibacterial' in the MDDR database. Both DOS libraries, as expected, were found to occupy a greater area of chemical space than the focused libraries. More significant however, using the datasets chosen, was 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; *i.e.* 22 for the DOS library, 19 for MDDR, 13 for the antibacterials, and 0.6 for the focused library (Fig. 6B).

**[Fig. 6 Best Inserted Here]**

The more diverse compound collection synthesized by Thomas *et al.* was screened for antibacterial activity. Screens were performed against two UK epidemic strains of methicillin-resistant *Staphylococcus aureus* (EMRSA 15 and EMRSA 16); these strains are responsible for the majority of MRSA infections in the UK.[58] The most active compound **35**, which was called gemmacin, showed a broad range of activity against Gram-positive bacteria (Table 1). Target identification suggested that gemmacin **35** acted as a cell membrane disruptor. Thus, a unique scaffold, which could potentially be useful in the antibacterial development, had been identified using DOS.

**[Table 1 Best Inserted Here]**

Multi-component coupling reactions (MCRs) feature regularly in DOS libraries as they serve to increase structural complexity; the products are, however, of identical molecular architecture. To exploit MCRs most successfully in DOS, two approaches can be used,

either: use the MCR to produce a densely functionalised molecule that can then be diversified further (see Strategy 2);[59] or, by incorporating a 'folding process' into the MCR (see Strategy 3).[60]


1.7.2 Strategy 2: Pluripotent (Densely Functionalised) Molecules

Using the Petasis three-component coupling of the lactol **36**, the amino acid **37**, and the boronic acid **38**, Schreiber and co-workers synthesized the β-amino alcohol **40** in good yield and diastereoselectivity (Scheme 4) *via* amine propargylation of the intermediate compound **39**. The densely functionalised compound **40** was used as the starting point for a DOS.[59]

**[Scheme 4 Best Inserted Here]**


A DOS strategy was exploited whereby the different combinations of the moieties of **40**, both polar and non-polar, were 'paired' in functional group specific reactions. This allowed diverse molecular skeletons to be generated using either: cycloisomeriation reactions (**41** and **43**); an enyne metathesis reaction (**42**); a gold mediated alkyne addition reaction (**44**); a Pauson-Khand reaction (**45**); or, a lactonisation reaction (**46**) (Scheme 5). In addition to this first generation of compounds, the lactone **46**, where the unsaturated functionalities remained unchanged, could be converted to a second generation of compounds **47**-**50** using identical pairing reactions to those used previously. The 1,3-dienes generated *via* the enyne metathesis reaction (*i.e.* **42** and **47**) were diversified further in a tandem Diels-Alder process to give the adducts **51** and **52**.[59]

**[Scheme 5 Best Inserted Here]**

The stereochemical outcome of the Petasis reaction was controlled by the lactol, and both diastereoisomers of **40** could be produced; stereochemical diversity could therefore be incorporated into any library synthesized using this approach. Although a small molecule library was not synthesized, different building blocks (*i.e.* different amino acids) were investigated in the initial MCR and subsequent pairing reactions. This is an elegant example of how both structural diversity and complexity can be achieved using a densely functionalised starter unit.

A similar pairing strategy has also been reported more recently by Proco and co-workers.[61] Using an enantioselective 1,4-addition to a nitro compound, the densely functionalised starter **53** (containing a combination of alkyne, alkene, nitro, and ester functional groups) could be synthesized readily. Different pairing reactions of the functional groups present in **53** led to the formation of different scaffolds: pairing the alkyne and alkene groups (route a) gave **54**; pairing the nitro and alkyne groups (route b) gave **55**; and, pairing the ester and nitro groups (route c) gave **56** (Scheme 6).

**[Scheme 6 Best Inserted Here]**

Polycyclic scaffolds could be accessed by stepwise pairing reactions. For example, the starting material **57** was first converted to the 1,3-diene **58** *via* an enyne metathesis reaction. The diene **58** was transformed, in the same pot, to the Diels Alder adduct **59** which was then converted, in a pairing reaction of the nitro and ester group, to the polycyclic compound **60** (Scheme 7).

**[Scheme 7 Best Inserted Here]**


1.7.3 Strategy 3: Folding Pathways

Garcia-Tellado and co-workers reported the use of an ABB' three-component coupling reaction in the synthesis of diverse scaffolds.[60] Pre-encoded information in the starting materials **61** and **64** (*i.e.* their comparative acidities and electrophilicities), in addition to the properties of the catalyst **62**, resulted in chemodifferentiation. The starting alkyne **61** and the α-dicarbonyl compounds **64** (which were either acidic or non-acidic α-keto esters or α-keto amides) were converted to the architectures **65**-**67** (Scheme 8).

**[Scheme 8 Best Inserted Here]**

The use of alternative ABB' MCRs has also been reviewed.[62] It is more common, however, for MCRs to feature as discrete steps in DOS pathways and be used to generate starter units. These starter units are often designed to participate in folding pathways.[48, 63, 64]


An excellent example of this was reported by Oguri and Schreiber who synthesized the starting scaffolds **68** with various combinations of key moieties at the three reactive sites A, B, and C. Attached to these sites was either a silyl either linker **69**, an α-diazo ketocarbonyl group **70**, or an indole moiety **71**. The relative arrangements of these groups in **68** encoded the skeletal information required to synthesize the distinct indole-like scaffolds **73**-**75** (Scheme 9).[63]

The folding pathway involved a Rh(II) induced cyclisation in which the α-diazo ketocarbonyl group of **68** reacted to form the carbonyl ylide **72**. The newly formed ylide **72** then participated in an intramolecular 1,3-dipolar cycloaddition reaction with the 2-3 double bond of the pendant indole group to give **73**-**75** (Scheme 9).

**[Scheme 9 Best Inserted Here]**

**1.8 DOS and Solid-Phase Organic Synthesis**

The ability to prepare compound libraries using traditional combinatorial chemistry has been aided by the use of solid phase organic synthesis (SPOS). Since the pioneering work by Merrfield in relation to peptide synthesis,[65] a wide range of solution phase reactions can now be performed on the solid support.[66] In addition to being amenable to automation, SPOS allows simple purification protocols which frequently involve only a filtration to be exploited (Scheme 10). As a result, compound collections have been generated in a high-throughput fashion therefore increasing the efficiency of library synthesis. Indeed, as stressed above, the problem is not the quantity of the small molecules that can be generated but the quality of the library members in terms of structural diversity.

Whereas SPOS has frequently been exploited in traditional combinatorial chemistry, the two technologies are not explicitly linked. More recently, as a result of its ever increasing applicability and generality, SPOS has been applied to the synthesis of structurally diverse compound libraries (*e.g.* Scheme 3).[7, 50, 53] Whereas the development of novel reactions on solid support is an important area of research in its own right, the major focus of this section is the use 'diversity-linkers' in the generation of DOS libraries.

**[Scheme 10 Best Inserted Here]**

1.8.1 An Overview of Linkage Cleavage Strategies

To mirror the diverse array of conditions employed in organic transformations,[67] over the last 20 years, more than 200 linkers have been developed.[68] To be useful a linker must: 1) contain a functional group through which the substrate to be attached; 2) be 'long' enough to prevent any unwanted resin-substrate interactions; and, 3) be stable (especially the linker-substrate bond) to the reaction conditions employed. This latter property confers selective cleavability.

Linkers are broadly classified depending on the functionality which remains at the site of cleavage (Scheme 10). Whereas with classic/traditional linkers the functional group used to attach the substrate to the resin is still present after cleavage, with traceless linkers only a hydrogen atom remains.[67, 68] In contrast to these approaches, diversity linkers allow further structural variation to be incorporated and can be of value in DOS.

Although the focus of this section is not classic and traceless linkers, some considerations are noteworthy. Since classic linkers are usually acid or base labile (*e.g.* **76** to **77**, Scheme 11), which ensures the target compound can be removed under mild conditions, this can be problematic as cleavage may occur before it is required. To overcome this problem so-called safety-catch linkers were developed; these require specific activation (*i.e.* **78** to **79**) before cleavage can occur (*i.e.* **79** to **80**, Scheme 11). Furthermore, in the context of a chemical genetics study where this cleavage is method exploited, all the of the compounds screened would contain a common polar functionality. Not only would this reduce the diversity of the collection, but the functional group itself may adversely effect bioactivity. Although this second point can be offset by the use of traceless linkers (*i.e.* **81** to **82**, Scheme 11), further chemical manipulation at this site is then problematic.

**[Scheme 11 Best Inserted Here]**

Scott and Steel have identified three sub-classes of diversity linkers in an excellent review: 1) those where diversity is incorporated by the nucleophilic component; 2) those where diversity is incorporated by an electrophilic component; and, 3) those which do not fit neatly into either of these classifications.[68] Herein we will highlight examples to demonstrate key principles as a way of introducing this topic.

1.8.2 Diversity Linkers; A Summary Of The Approaches Used

Although the linker used by Thomas *et al.* (Scheme 3) could be removed to yield acids, esters, and a variety of amides by varying the nucleophilic component in the cleavage reaction, examples where electophilic intermediate is initially generated allow greater diversity to be incorporated. For examples the triazene linker **83** has been used to initially generate an elecrophilic aryldiazonium salt **84**; this was intercepted by a diverse range of nucleophiles to generate compounds **85**-**90** (Scheme 12).[69, 70]

**[Scheme 12 Best Inserted Here]**

In a similar vein, diversity can be incorporated using an electrophilic component. Generally these linkers contain a suitably labile bond or are converted to reactive intermediates that are subsequently intercepted by a diverse range of electrophiles. As an alternative to these 'nucleophilic' and 'electrophilic' approaches,  Diels Alder reactions, ring closing

metathesis reactions, radical cleavage reactions, and modified Friedlander reactions have also been exploited. Again these have allowed diversity to be incorporated in the cleavage step in a library synthesis.

Although many diversity linkers provide a useful method of incorporating appendage diversity, approaches such as 'cyclo-release cleavage', where different skeletons can be produced under different conditions, may be more useful still.

## 1.9 Conclusion

Although a recently developed concept, over the last few years, novel and imaginative strategies have been used to prepare DOS libraries. Many of these compound collections have been successfully exploited in chemical genetics to identify modulators for biological systems.[5-8] DOS, to populate diverse regions of chemical space, still represents a potentially rewarding challenge for chemists.

## 1.9 References

[1]  R. J. Spandl, D. R. Spring, A. Bender, *Org. Biomol. Chem.* **2008**, DOI: 10.1039/B719372F.

[2]  D. R. Spring, *Org. Biomol. Chem.* **2003**, *1*, 3867.

[3]  G. L. Thomas, E. E. Wyatt, D. R. Spring, *Curr. Opin. Drug Discovery Dev.* **2006**, *9*, 700.

[4]  S. J. Haggarty, *Curr. Opin. Chem. Biol.* **2005**, *9*, 296.

[5]  Y. K. Kim, M. A. Arai, T. Arai, J. O. Lamenzo, E. F. Dean, N. Patterson, P. A. Clemons, S. L. Schreiber, *J. Am. Chem. Soc.* **2004**, *126*, 14740.

[6]  A. N. Koehler, A. F. Shamji, S. L. Schreiber, *J. Am. Chem. Soc.* **2003**, *125*, 8420.

[7]  D. R. Spring, S. Krishnan, H. E. Blackwell, S. L. Schreiber, *J. Am. Chem. Soc.* **2002**, *124*, 1354.

[8]  F. G. Kuruvilla, A. F. Shamji, S. M. Sternson, P. J. Hergenrother, S. L. Schreiber, *Nature* **2002**, *416*, 653.

[9]  F. S. Collins, E. S. Lander, J. Rogers, R. H. Waterston, *Nature* **2004**, *431*, 931.

[10]  M. R. Brent, *Nat. Rev. Genet.* **2008**, *9*, 62.

[11]  D. R. Spring, *Chem. Soc. Rev.* **2005**, *34*, 472.

[12]  S. L. Schreiber, *Bioorg. Med. Chem.* **1998**, *6*, 1127.

[13]  S. L. Schreiber, *Chem. Eng. News* **2003**, *81*, 51.

[14]  D. P. Walsh, Y. T. Chang, *Chem. Rev.* **2006**, *106*, 2476.

[15]  A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727.

[16]  C. Lipinski, A. Hopkins, *Nature* **2004**, *432*, 855.

[17]  M. M. Hann, A. R. Leach, G. Harper, *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856.

[18]  A. Schuffenhauer, N. Brown, P. Selzer, P. Ertl, E. Jacoby, *J. Chem. Inf. Mod.* **2006**, *46*, 525.

[19]  G. MacBeath, *Genome Biol.* **2001**, *2*, 2005.1.

[20]  S. Fergus, A. Bender, D. R. Spring, *Curr. Opin. Chem. Biol.* **2005**, *9*, 304.

[21]  J. J. Perez, *Chem. Soc. Rev.* **2005**, *34*, 143.

[22]  C. M. Dobson, *Nature* **2004**, *432*, 824.

[23]  G. M. Downs, P. Willett, W. Fisanick, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1094.

[24]  E. Estrada, E. Uriarte, *Curr. Med. Chem.* **2001**, *8*, 1573.

[25]  S. H. Fitzgerald, M. Sabat, M. Geysen, *J. Comb. Chem.* **2007**, *9*, 724.

[26]  M. S. Butler, *Nat. Prod. Rep.* **2008**, *22*, 162.

[27]  K. S. Lam, *Trends Microbiol.* **2007**, *15*, 279.

[28]  M. Feher, J. M. Schmidt, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 218.

[29]  M. D. Burke, S. L. Schreiber, *Angew. Chem. Int. Ed.* **2004**, *43*, 46.

[30]  M. J. Valler, D. Green, *Drug Discov. Today* **2000**, *5*, 286.

[31]  C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* **1997**, *23*, 3.

[32]  T. H. Keller, A. Pichota, Z. Yin, *Curr. Opin. Chem. Biol.* **2006**, *10*, 357.

[33]  C. Abad-Zapatero, *Drug Discov. Today* **2007**, *12*, 995.

[34]  M. D. Burke, E. M. Berger, S. L. Schreiber, *J. Am. Chem. Soc.* **2004**, *126*, 14095.

[35]  B. E. Evans, K. E. Rittle, M. G. Bock, R. M. Dipardo, R. M. Freidinger, W. L. Whitter, G. F. Lundell, D. F. Veber, P. S. Anderson, R. S. L. Chang, V. J. Lotti, D. J. Cerino, T. B. Chen, P. J. Kling, K. A. Kunkel, J. P. Springer, J. Hirshfield, *J. Med. Chem.* **1988**, *31*, 2235.

[36]  A. Reayi, P. Arya, *Curr. Opin. Chem. Biol.* **2005**, *9*, 240.

[37]  J. Clardy, C. Walsh, *Nature* **2004**, *432*, 829.

[38]  S. K. Ko, H. J. Jang, E. Kim, S. B. Park, *Chem. Commun.* **2006**, 2962.

[39]  C. X. Zhou, A. V. Dubrovsky, R. C. Larock, *J. Org. Chem.* **2006**, *71*, 1626.

[40]  R. Messer, C. A. Fuhrer, R. H. Iner, R. Haner, *Curr. Opin. Chem. Biol.* **2005**, *9*, 259.

[41]   A. M. Boldi, D. R. Dragoli,  (Ed.: A. M. Boldi), Taylor and Francis/CRC Press, London, **2006**, pp. 1.

[42]   B. C. Goess, R. N. Hannoush, L. K. Chan, T. Kirchhausen, M. D. Shair, *J. Am. Chem. Soc.* **2006**, *128*, 5391.

[43]   J. Buckingham, *Dictionary of Natural Products (CD ROM)*, Taylor and Francis/CRC Press, London, **2005**

[44]   M. Kaiser, S. Wetzel, K. Kumar, H. Waldmann, *Cell Mol. Life Sci* **2008**, 1.

[45]   M. A. Koch, A. Schuffenhauer, M. Scheck, S. Wetzel, M. Casaulta, A. Odermatt, P. Ertl, H. Waldmann, *Proc. Nat. Acad. Sci. U.S.A.* **2005**, *102*, 17272.

[46]   F. J. Dekker, M. A. Koch, H. Waldmann, *Curr. Opin. Chem. Biol.* **2005**, *9*, 232.

[47]   A. Noren-Muller, I. Reis-Correa, H. Prinz, C. Rosenbaum, K. Saxena, H. J. Schwalbe, D. Vestweber, G. Cagna, S. Schunk, O. Schwarz, H. Schiewe, H. Waldmann, *Proc. Nat. Acad. Sci. U.S.A.* **2006**, *103*, 10606.

[48]   M. D. Burke, E. M. Berger, S. L. Schreiber, *Science* **2003**, *302*, 613.

[49]   S. L. Schreiber, *Science* **2000**, *287*, 1964.

[50]   N. Kumar, M. Kiuchi, J. A. Tallarico, S. L. Schreiber, *Org. Lett.* **2005**, *7*, 2535.

[51]   T. E. Nielsen, S. L. Schreiber, *Angew. Chem. Int. Ed.* **2008**, *47*, 48.

[52]   E. E. Wyatt, S. Fergus, W. Galloway, A. Bender, D. J. Fox, A. T. Plowright, A. S. Jessiman, M. Welch, D. R. Spring, *Chem. Commun.* **2006**, 3296.

[53]   G. L. Thomas, R. J. Spandl, F. G. Glansdorp, M. Welch, A. Bender, J. Cockfield, J. A. Lindsey, C. Bryant, D. F. J. Brown, O. Loiseleur, H. Rudyk, M. Ladlow, D. R. Spring, *Angew. Chem. Int. Ed.* **2008**, DOI: 10.1002/anie.200705415.

[54]   C. Chen, X. D. Li, S. L. Schreiber, *J. Am. Chem. Soc.* **2003**, *125*, 10174.

[55]   H. C. Kolb, M. S. Vannieuwenhze, K. B. Sharpless, *Chem. Rev.* **1994**, *94*, 2483.

[56]   J. S. Johnson, D. A. Evans, *Acc. Chem. Res.* **2000**, *33*, 325.

[57]   G. L. Thomas, M. Ladlow, D. R. Spring, *Org. Biomol. Chem.* **2004**, *2*, 1679.

[58]   A. P. Johnson, H. M. Aucken, S. Cavendish, M. Ganner, M. C. Wale, M. Warner, D. M. Livermore, B. D. Cookson, *J. Antimicrob. Chemother.* **2001**, *48*, 143.

[59]   N. Kumagai, G. Muncipinto, S. L. Schreiber, *Angew. Chem. Int. Ed.* **2006**, *45*, 3635.

[60]   D. Tejedor, A. Santos-Exposito, F. Garcia-Tellado, *Chem. Eur. J.* **2007**, *13*, 1201.

[61]   E. Comer, E. Rohan, L. Deng, J. A. Porco, *Org. Lett.* **2007**, *9*, 2123.

[62]   D. Tejedor, F. Garcia-Tellado, *Chem. Soc. Rev.* **2007**, *36*, 484.

[63]   H. Oguri, S. L. Schreiber, *Org. Lett.* **2005**, *7*, 47.

[64]   J. K. Sello, P. R. Andreana, D. S. Lee, S. L. Schreiber, *Org. Lett.* **2003**, *5*, 4125.

[65]   R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149.

[66]   R. E. Dolle, *J. Comb. Chem.* **2005**, *7*, 739.

[67]   N. Jung, M. Weihn, S. Brase, *Top. Curr. Chem.* **2007**, *278*, 1.

[68]   P. J. H. Scott, P. G. Steel, *Eur. J. Org. Chem.* **2006**, 2251.

[69]   S. Brase, M. Schroen, *Angew. Chem. Int. Ed.* **1999**, *38*, 1071.

[70]   S. Brase, *Acc. Chem. Res.* **2004**, *37*, 805.
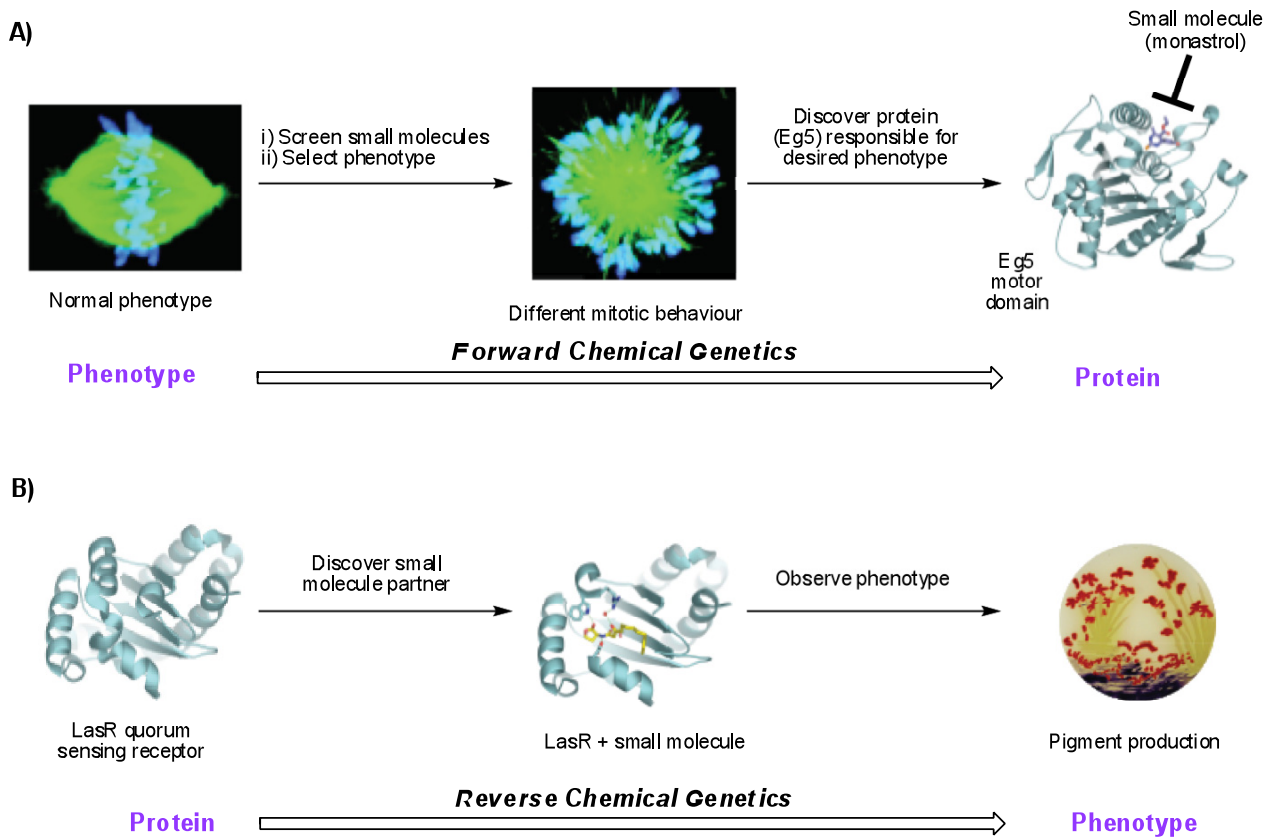
# Figures and Schemes

## Figure 1

**A)**



i) Screen small molecules
ii) Select phenotype

Normal phenotype

Discover protein (Eg5) responsible for desired phenotype

Different mitotic behaviour

Small molecule (monastrol)

Eg5 motor domain

**Phenotype** → *Forward Chemical Genetics* → **Protein**

**B)**



Discover small molecule partner

Observe phenotype

LasR quorum sensing receptor

LasR + small molecule

Pigment production

**Protein** → *Reverse Chemical Genetics* → **Phenotype**

**Fig. 1: A)** In forward chemical genetics a library is screened and small molecules that induce a desired phenotype (*e.g.* different mitotic behaviour) are identified. Further investigation allows the protein responsible for this change (*i.e.* the protein partner of the small molecule identified from the initial screen) to be discovered. **B)** In reverse chemical genetics the small molecules partner of the protein under investigation (*e.g.* LasR) is first discovered. The phenotype induced by the action of this pre-selected small molecule is then observed.

**Figure 2**
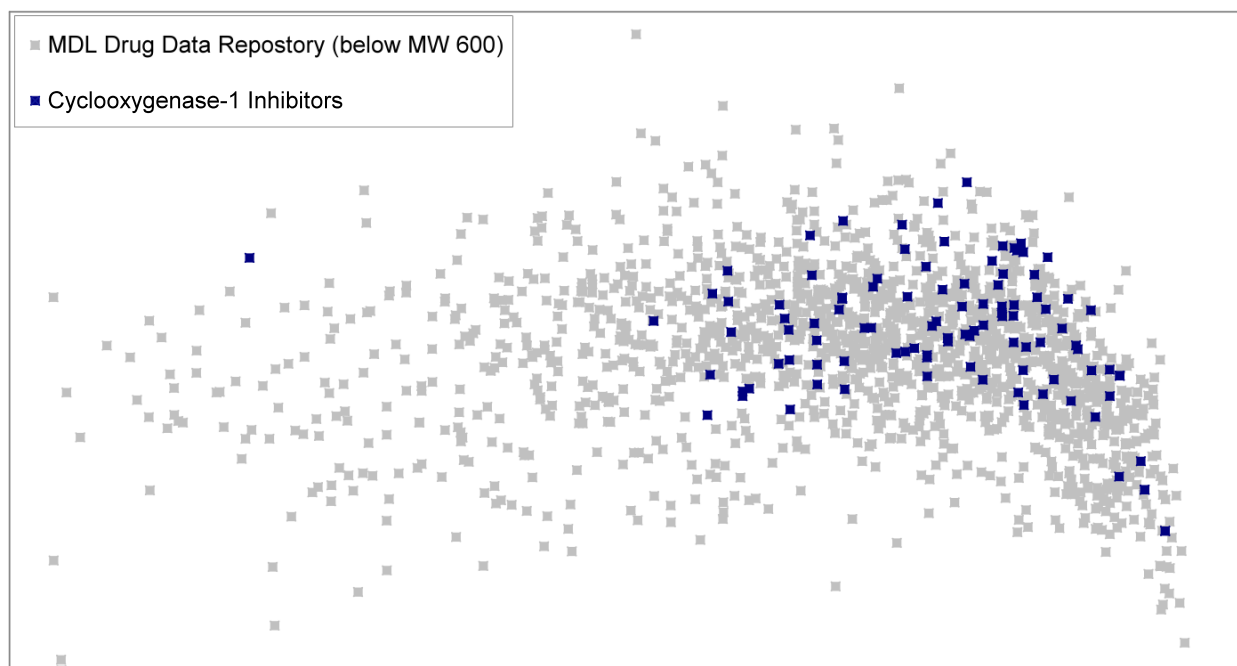


**Fig. 2: A chemical space analysis of cyclooxygenase-1 inhibitors (blue squares) and MDDR compounds (grey squares). This visual representation shows that the cyclooxygenase-1 inhibitors populate a diffuse region of chemical space on the background of the MDDR compounds.**

**Figure 3**
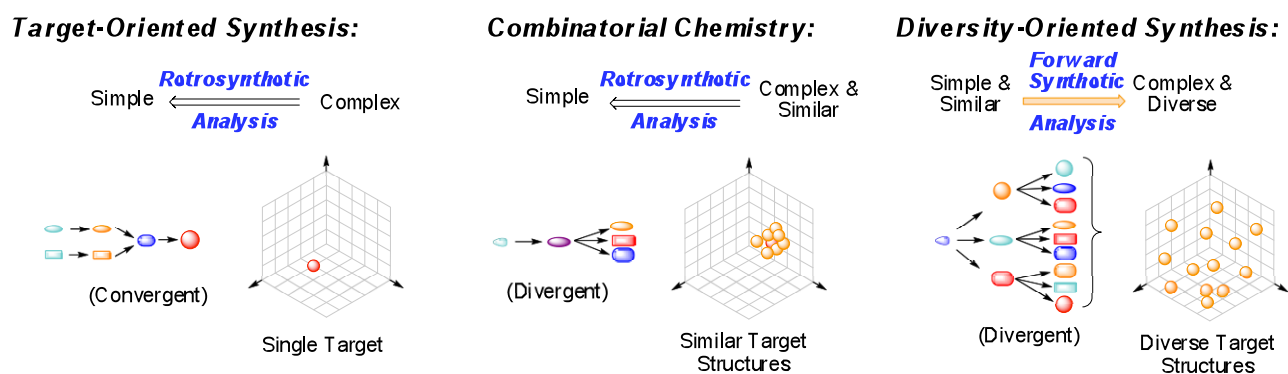


**Fig. 3: In target-oriented synthesis and traditional combinatorial chemistry, retro-synthetic analysis is used; the target molecule or molecules occupy a discrete point in, or more densely populate a region of, chemical space. In contrast, diversity-oriented synthesis uses forward synthetic analysis to plan, and subsequently produce, compound collections that interrogate a diffuse region of chemical space.**

16

# Figure 4
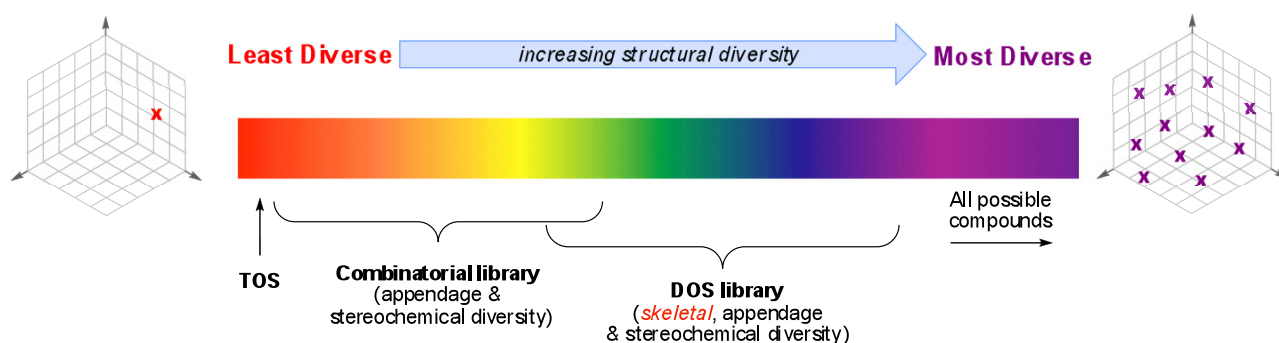


Fig. 4: The '*molecular diversity spectrum*'; a qualitative description of TOS, combinatorial chemistry, and DOS in terms of structural diversity. The left-hand side of the spectrum represents minimal skeletal variety within a compound collection whereas the right-hand side represent the maximal diversity (theoretically) achievable. Combinatorial chemistry libraries and DOS libraries are placed on this axis going from left (least diverse) to right (most diverse).[1]

# Scheme 1
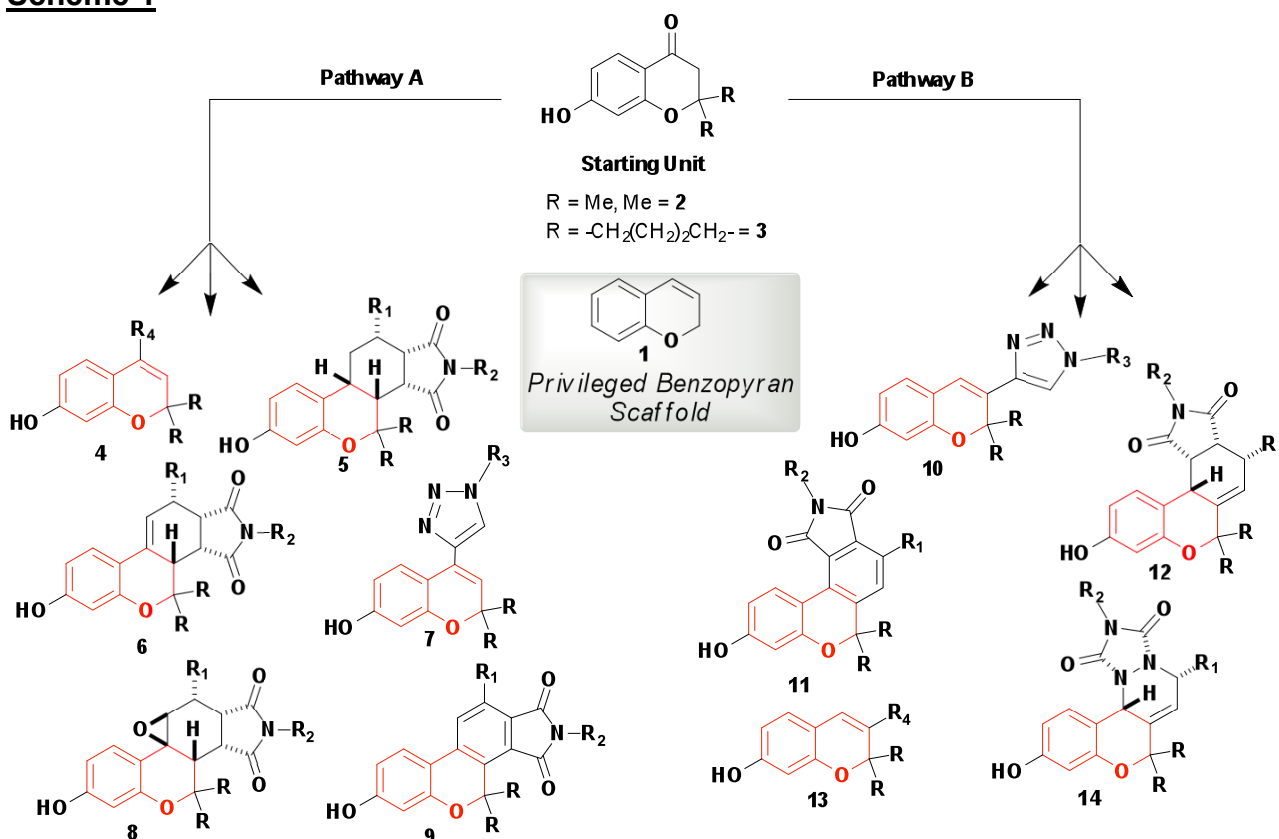


Scheme 1: The scaffolds of a DOS library based around the privileged benzopyran scaffold. A range of IC$_{50}$ values were reported (biological diversity) when this library was screened against a human cancer cell line.[38]

# Figure 5

## A) Branching Pathway



Common starting material → Different Reagents → Distinct skeletons

reagent–based approach

## B) Folding Process

$\sigma_1$, $\sigma_2$, $\sigma_3$ → Common Reagents → Distinct skeletons

Pre-encoded starting material

substrate–based approach
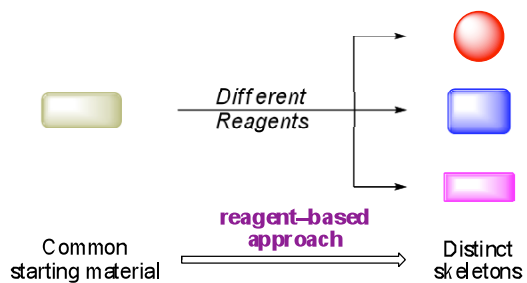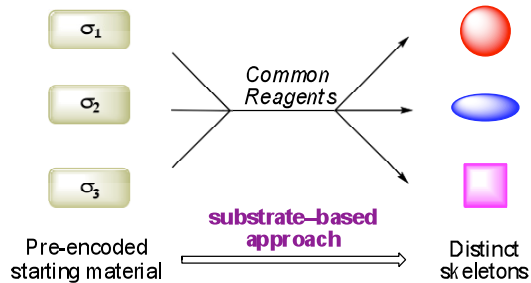
Fig. 5: The two general strategies to accessing skeletal diversity in a compound library; either branching pathways (A) or folding pathways (B).

## Scheme 2
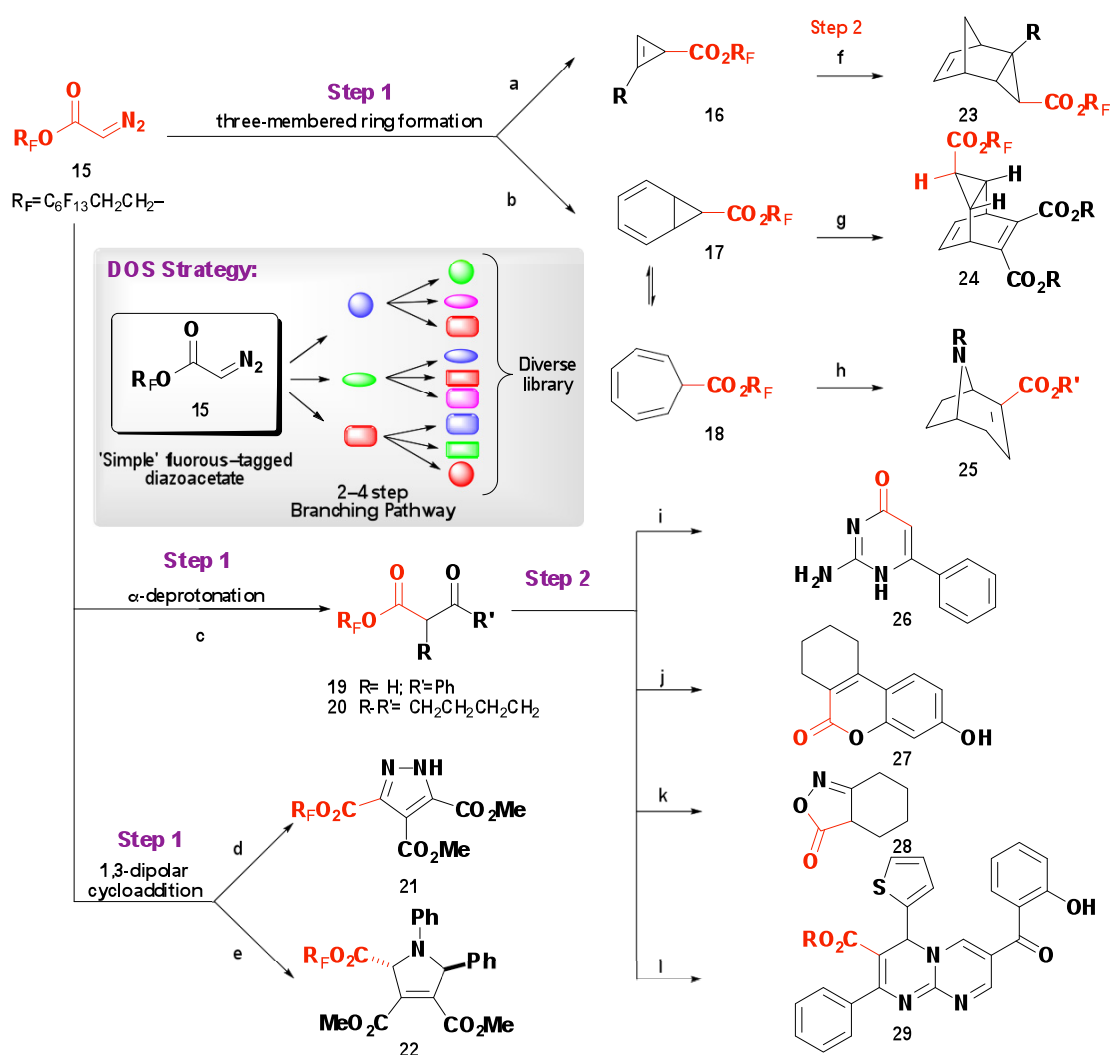


**Scheme 2: Diversity-oriented synthesis of 223 small molecules based on 30 discrete frameworks.** Steps 1: (a) RCCH, $Rh_2(OAc)_4$, [BuCCH, 57%]; (b) $C_6H_6$, $Rh_2(O_2CCF_3)_4$, 70%; (c) LDA, RCOR' then $Rh_2(OAc)_4$; 36: 49% (90%); 37: 68% (97%); (d) DMAD, 84% (88%); (e) PhCHO, PhNH$_2$, then DMAD, $Rh_2(OAc)_4$, *d.r.* = 20 : 1, 51% (80%). Steps 2: (f) $C_5H_6$, 92%; (g) dienophile [dimethyl acetylenedicarboxylate, 59%]; (h) $RNH_2$, NaOH then MeOH, $H_2SO_4$, [MeNH2, 35%]; (i) Guanidine carbonate 62% (96%); (j) Resorcinol, $H_2SO_4$, 74% (95%); (k) $NH_2OH$, 77% (89%); (l) Thiophene-2-carboxaldehyde, guanidine carbonate, then 3-formylchromone, 43% (98%). Yields and purity (in brackets) of the product example following generic purification using (reverse) fluorous solid phase extraction or precipitation shown. Purity determined by HPLC, LCMS or ^{1}H NMR. DMAD = dimethyl acetylenedicarboxylate.[52]

19
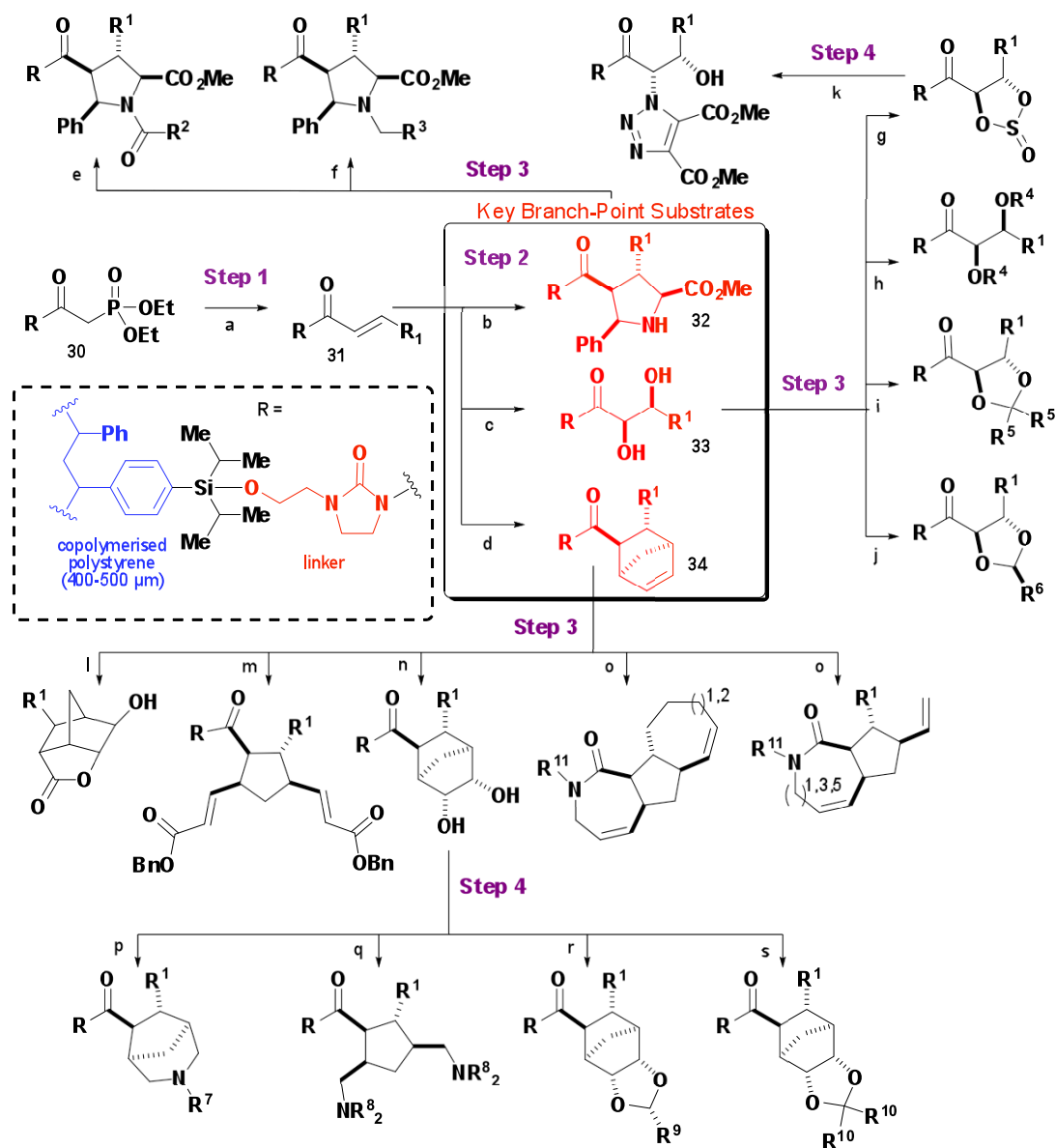
# Scheme 3



Scheme 3: Diversity-oriented synthesis of 242 compounds based of 18 discrete molecular frameworks. Conditions: a) LiBr, 1,8-diazabicyclo[5.4.0]undec-7-ene, $R^1$CHO, MeCN; b) (*R*)-QUINAP, AgOAc, i-Pr$_2$NEt, THF, -78 °C → 25 °C; c) AD-mix, (DHQD)PHAL, THF:H2O (1:1); d) chiral bis(oxazoline), Cu(OTf)$_2$, 3Å MS, CH$_2$Cl$_2$, C$_5$H$_6$; e) $R^2$COCl, DMAP, pyridine, CH$_2$Cl$_2$; f) $R^3$CHO, BH$_3$• pyridine, MeOH; g) SOCl$_2$, pyridine, CH$_2$Cl$_2$, 40 °C; h) $R^4$Br, Ag$_2$O, CH$_2$Cl$_2$, 40 °C; i) $R^5$C(O)$R^5$, TsOH, DMF, 65 °C; j) $R^6$CHO, TsOH, DMF, 65 °C; k) NaN$_3$, DMF, 100 °C then dimethyl acetylenedicarboxylate, PhMe, 65 °C; l) *m*CPBA, CH$_2$Cl$_2$ then MeOH, 65 °C; m) CH$_2$=CHCO$_2$Bn, PhMe, 120 °C, Grubbs II, CH$_2$=CH$_2$; n) OsO$_4$, NMO, CH$_3$C(O)CH$_3$:H$_2$O (10:1); o) RNH$_2$, Me$_2$AlCl, PhMe 120 °C; then NaH, $R^{11}$X, DMF, THF; then PhMe, 120 °C, Grubbs II, CH2=CH2; p) NalO$_4$, THF:H$_2$O (1:1); then $R^7$NH$_2$, NaB(OAc)$_3$H, CH$_2$Cl$_2$; q) NalO$_4$, THF:H$_2$O (1:1); then $R^8$NHR$^8$, NaB(OAc)$_3$H, CH$_2$Cl$_2$; r) $R^9$CHO, DMF, TsOH, 60 °C; s) $R^{10}$C(O)R$^{10}$, DMF, TsOH, 60 °C. DMF = *N,N*-dimethylformamide; THF = tetrahydrofuran; DMAP = *N,N*-dimethylaminopyridine; (DHQD)PHAL = hydroquinidine 1,4-phthalazinediyl diether; *m*CPBA = meta-chloroperbenzoic acid; Ts = para-toluenesulfonyl; Grubbs II = 1,3-(bis(mesityl)-2-imidazolidinylidene) dichloro (phenylmethylene) (tricyclohexylphosphine) ruthenium; NMO = 4-methylmorpholine-N-oxide.[53]

# Figure 6
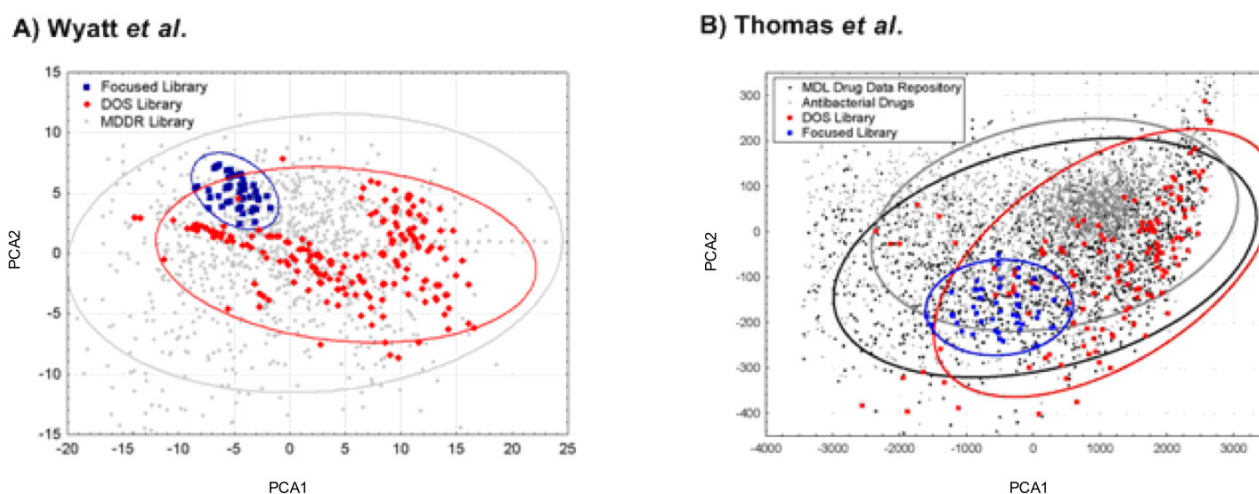


| Library | Wyatt *et al.*[52] | Thomas *et al.*[53] |
|---|---|---|
| | *Average 'chemical space' occupied per compound* | *Average 'chemical space' occupied per compound* |
| Focused | 0.70 | 0.64 |
| MDDR | 28.28 | 18.86 |
| DOS | 11.25 | 22.32 |
| Antibacterial | Not calculated | 12.79 |

**Fig. 6: Visual representation of the diversity of different chemical collections in physicochemical and topological space 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 (grey dots) (B only) are depicted. Analysis of the average 'chemical space' occupied per compound is shown in the table below.[52, 53]**

# Table 1



**35**
**(−)-gemmacin**

| | MIC$_{50}$ (µg ml$^{-1}$) | | |
|---|---|---|---|
| | MSSA | EMRSA 15 | EMRSA 16 |
| (±)-gemmacin | 2 | 16 | 32 |
| **(−)-gemmacin** | ND | **8** | **16** |
| (+)-gemmacin | ND | 16 | 32 |
| erythromycin | 0.5 | >64 | >64 |
| oxacillin | 0.5 | >32 | >32 |

**Table 1: Structure and activity of gemmacin 35 with growth inhibitory activity (MIC$_{50}$) against three strains of *S. aureus*. For comparison the MCI$_{50}$ values for erythromycin and oxacillin are also shown. ND = not determined. MSSA = methicillin-susceptible *S. aureus*.[53]**
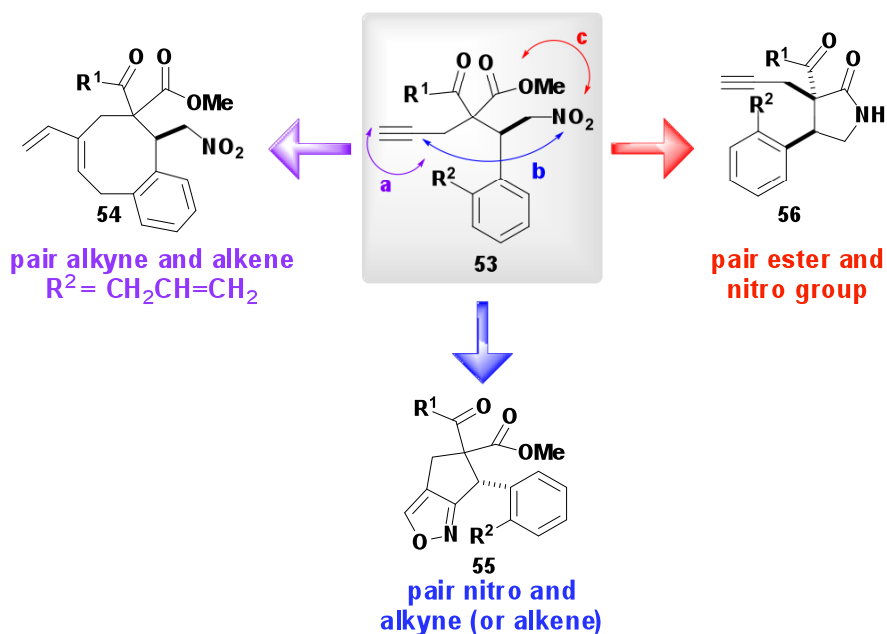
## Scheme 4
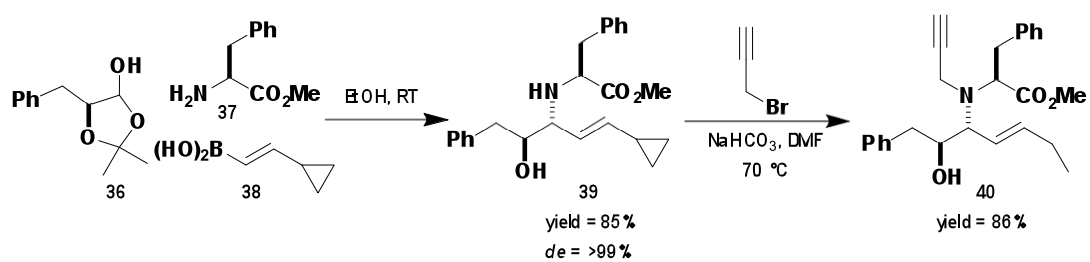


**Scheme 4: The synthesis of the β-amino alcohol 40. A Petasis three-component coupling reaction was used to synthesize the intermediate 39 which was then converted to 40; this compound was used as the starting point for the DOS.[59]**
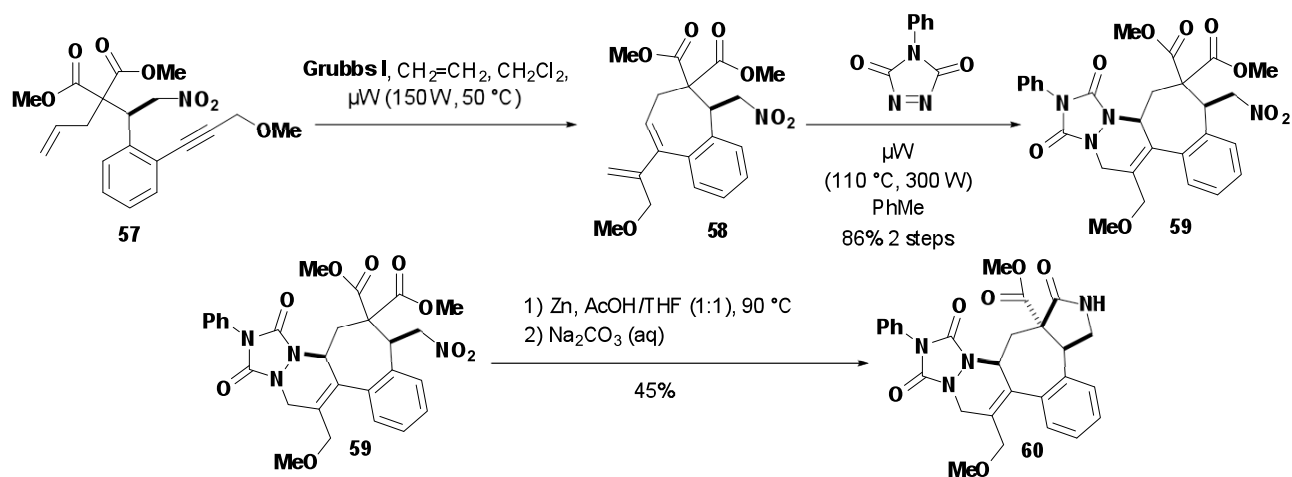
# Scheme 5
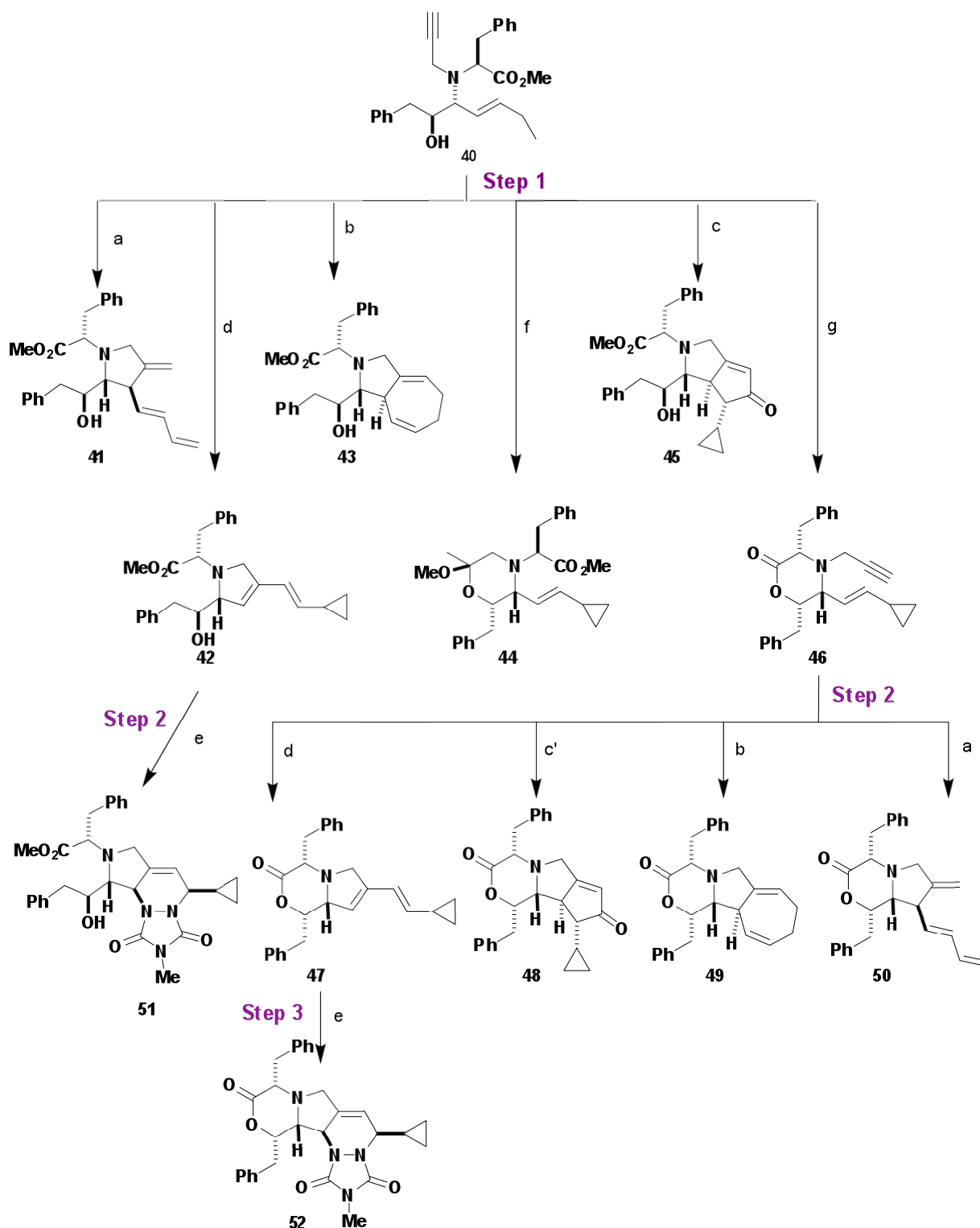


**Scheme 5: An example of the use of densely functionalised molecules in a DOS. Conditions: a) [Pd(PPh₃)₂(OAc)₂] (10 mol%), benzene, 80 °C; b) [CpRu(CH₃CN)₃PF₆] (10 mol%), acetone, RT; c) [Co₂(CO)₈], trimethylamine *N*-oxide, NH₄Cl, benzene, RT; c') [Co₂(CO)₈], trimethylamine *N*-oxide, benzene, RT d) 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.[59]**

## Scheme 6



**Scheme 6: After the synthesis of 53, the different scaffolds 54-56 were formed by pairing the functionalities present in the starting material.**[61]
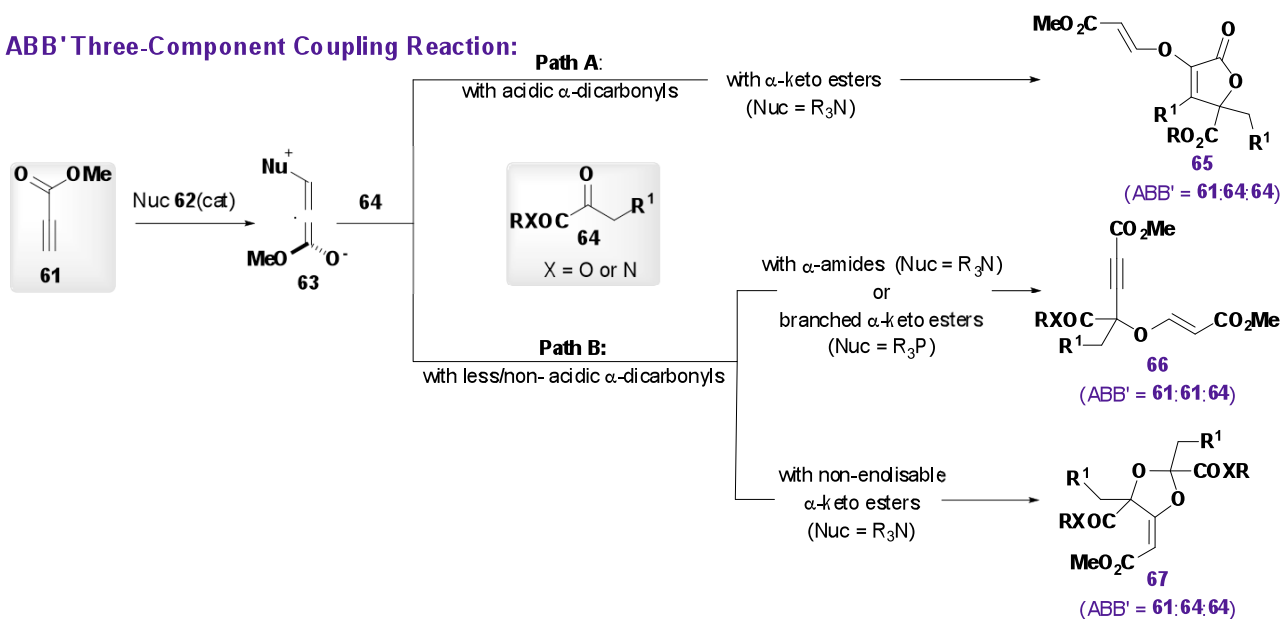
## Scheme 7



**Scheme 7: Polycyclic scaffolds such as 60 could be synthesized using stepwise coupling reactions.**[61]

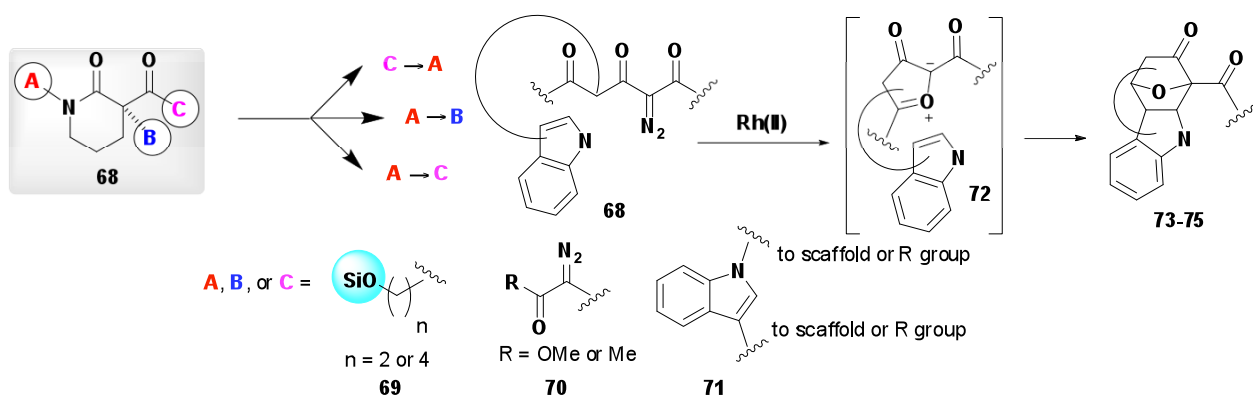## Scheme 8

**ABB' Three-Component Coupling Reaction:**



**Path A:** with acidic α-dicarbonyls

with α-keto esters (Nuc = R₃N) → **65** (ABB' = **61**:**64**:**64**)

**Path B:** with less/non-acidic α-dicarbonyls

with α-amides (Nuc = R₃N) or branched α-keto esters (Nuc = R₃P) → **66** (ABB' = **61**:**61**:**64**)

with non-enolisable α-keto esters (Nuc = R₃N) → **67** (ABB' = **61**:**64**:**64**)

**Scheme 8:** This ABB' MCR allowed the synthesis of scaffolds 65-67 from the alkyne 61 and the α-dicarbonyl 64 in the presence of the catalyst 62. The chemodifferentiation, which allowed this folding process to occur, was a result of the natures of the starting materials and the catalyst.[60]
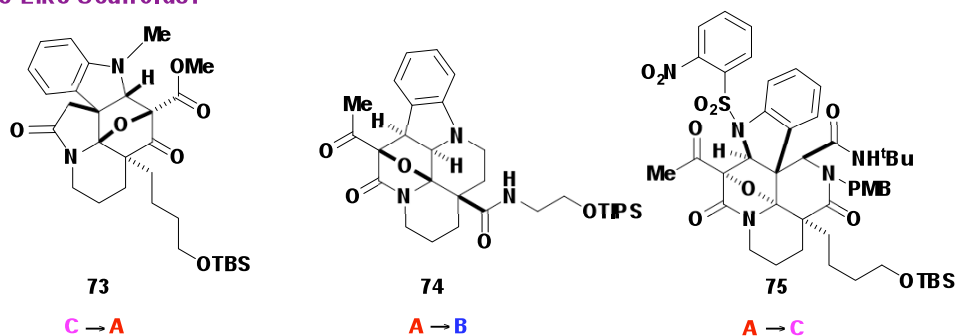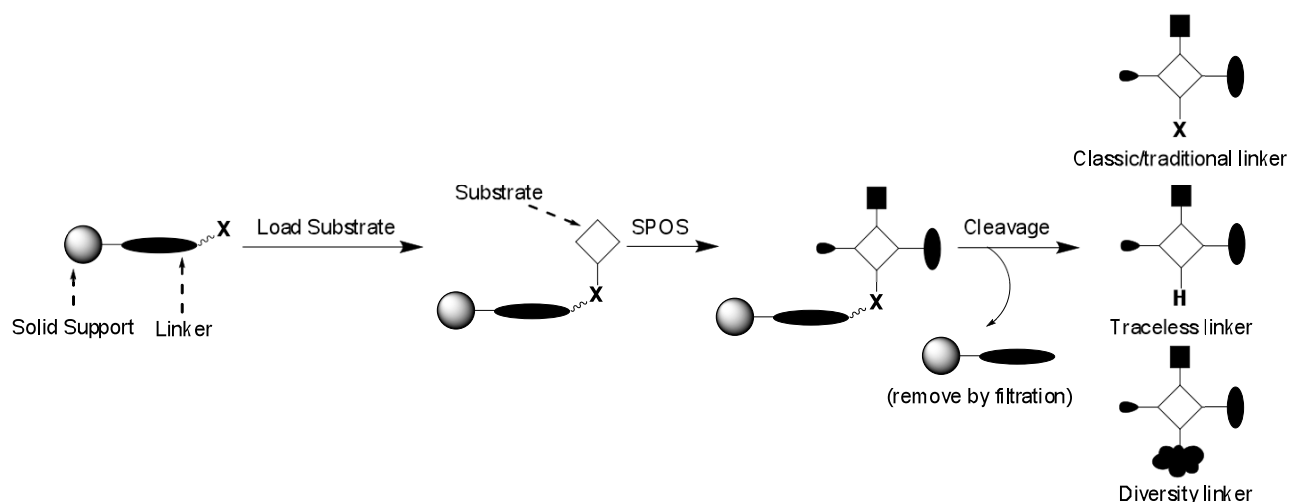
## Scheme 9

**Common Substrate:** / **Rh(II) Induced Folding Reaction:**



**Distinct Indole-Like Scaffolds:**

**73** C → A   **74** A → B   **75** A → C

**Scheme 9: By elaborating on the substrate scaffold 67, with various combination of 68-70, a Rh(II) induced cyclisation allowed the formation of indole-like scaffolds 68-70 in this folding pathway.[63]**



**Scheme 10: An overview of SPOS; a substrate molecule is attached, *via* a selectively cleavable linker, to a solid support. The substrate is carried through sequential rounds of chemical transformations (SPOS) before being cleaved from the resin. Linkers are classified as either: classic/traditional linkers; traceless linkers; or, diversity linkers. Although the first two types are generally cleaved using a common 'cleavage cocktail', the latter can be used to incorporate structural diversity into the product.[68]**
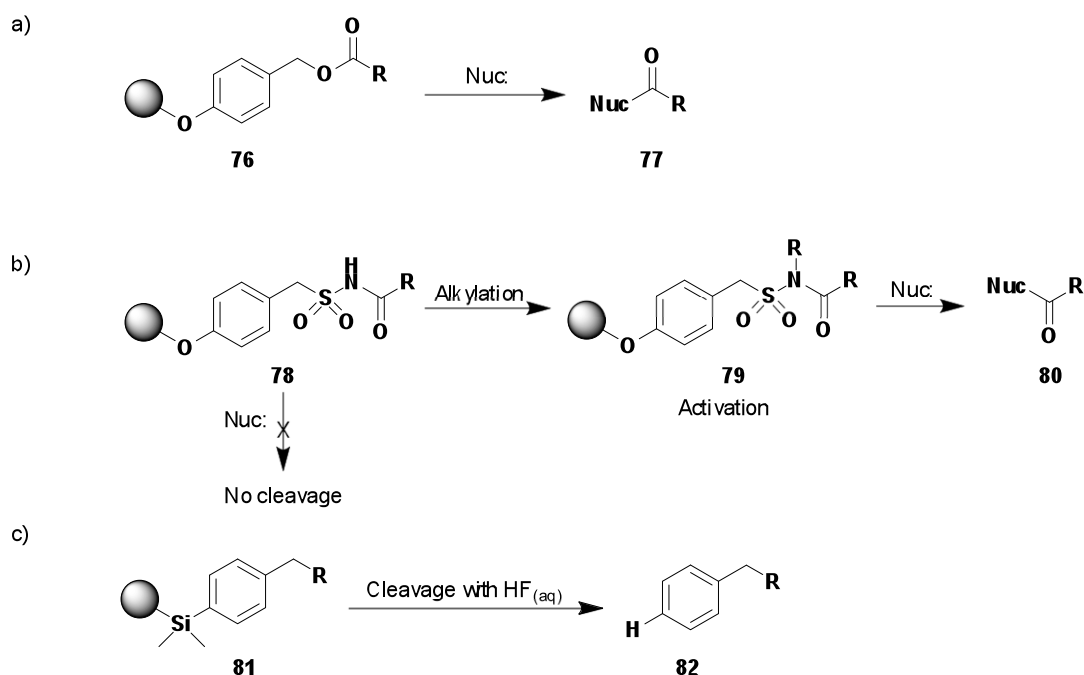


**Scheme 11: a) An example of the use of a classic linker. Cleavage in base allows the removal of the product from the resin; the polar functional group remains. b) An extension of this strategy is the use of a safety catch linker, in this example the nitrogen must be alkylated before base induced cleavage can occur. c) An example of the use of a traceless linker. Using the silicon-based linker, after acid induced cleavage, only a hydrogen remains at the initial site of attachment.**
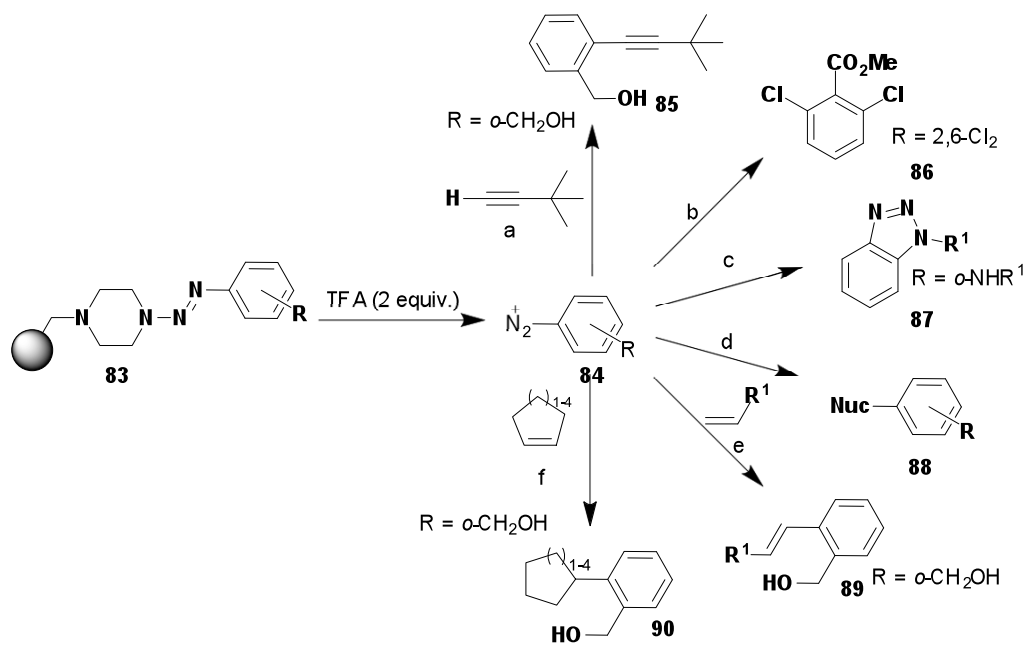
**Scheme 12: The generation of a highly reactive electrophilic intermediate that is then intercepted by a range of nucleophiles.**[69, 70]