Journal Name

Dynamic Article Links ►

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

A Conceptual Framework for Analysing and Planning Synthetic Approaches to Diverse Lead-like Scaffolds

Paul Maclellan[a,b] and Adam Nelson*[a,b]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

Historically, chemists' exploration of chemical space has been exceptionally uneven and unsystematic. This feature article outlines a comprehensive conceptual framework that may be used to capture, analyse and plan synthetic approaches that may address this historically uneven exploration. Illustrative examples of synthetic approaches that target, or have potential to target, broad tracts of lead-like chemical space are presented within the context of this conceptual framework. Particular emphasis is placed on synthetic approaches that enable the combinatorial variation of molecular scaffold, particularly within the boundaries of lead-like chemical space.

Introduction

The discovery of bioactive small molecules has been shaped, in large part, by the historic exploration of chemical space by chemical synthesis (and biosynthesis). The scaffolds of known bioactive small molecules, in particular, play a key role in guiding chemists' navigation of biologically-relevant chemical space.[1,2] Vast sources of historic structure–activity relationship data have allowed the chemical space defined by known bioactive ligands to be mapped.[3] The field of biology-oriented synthesis (BIOS),[4] for example, seeks to focus on biologically-relevant chemical space by using biologically-validated scaffolds[5,6,7] to inspire library design.

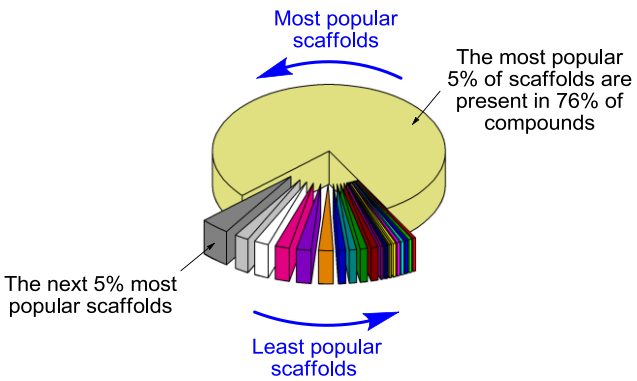


Figure Scaffold diversity of the organic chemistry universe.[8] The 24 282 284 cyclic compounds in the CAS registry in 2008 are grouped from the most popular 5% to the least popular 5% of scaffolds. Around half of the known compounds are based on just 0.25% of known molecular scaffolds!

Worryingly, however, chemists' historical exploration of chemical space has been exceptionally uneven and unsystematic. Around half of all known compounds are based on just 0.25% of the known molecular scaffolds (Figure)![8,‡] This uneven exploration is also reflected in small molecule screening collections:[6,9] consequently, the biological properties of small molecules have not been annotated systematically, with huge emphasis on a small number of synthetically accessible scaffolds. To what extent, then, are our collective views about the biological relevance of chemical space skewed by our uneven exploration by synthesis?

In this article, we describe synthetic approaches that have emerged recently to allow chemical space to be explored more systematically. We introduce a conceptual framework that allows these valuable approaches to be analysed, compared and extended. We have emphasised approaches that allow significant variation of the molecular scaffold prepared. In contrast to our previous review of diversity-oriented synthetic approaches,[10] we focus here on approaches that facilitate exploration within the boundaries of lead-like chemical space. Thus in each case, the approaches highlighted may yield products that could be decorated to give large numbers of lead-like small molecules.

Lead-like chemical space

There is a strong link between the molecular properties of clinical candidates, and the probability of their successful negotiation of the development process to yield marketed drugs.[11] Molecular properties that have been shown to correlate with success in the development process include molecular size and lipophilicity (clogP),[12] the number of aromatic rings[13] (nAr) and the fraction of sp^3-hybridised carbons (Fsp^3).[14] The optimisation process almost inevitably increases both molecular weight and lipophilicity, making it essential to control the properties of initial leads.[15,16] The concept of lead-oriented synthesis has recently been introduced to highlight specific challenges associates with preparing lead-like small molecules i.e. molecules that would be good starting points for lead optimisation (see Table 1).[17]

Table 1 Preferred molecular properties and features that have been proposed for lead-like small molecules[17]

Molecular property/feature	Preferred values
Lipophilicity	-1 < clogP < 3
Molecular size	14 ≤ heavy atoms ≤ 26[a]
Aromatic rings	nAr ≤ 3
Shape	More 3D shape[b]
Sub-structures	Absence of chemically-reactive, electrophilic or redox-active groups

[a]Molecular weight, M_R, ~200-350. [b]The fraction of sp^3 carbons (Fsp^3) can be a useful parameter for assessing three dimensionality.[14]

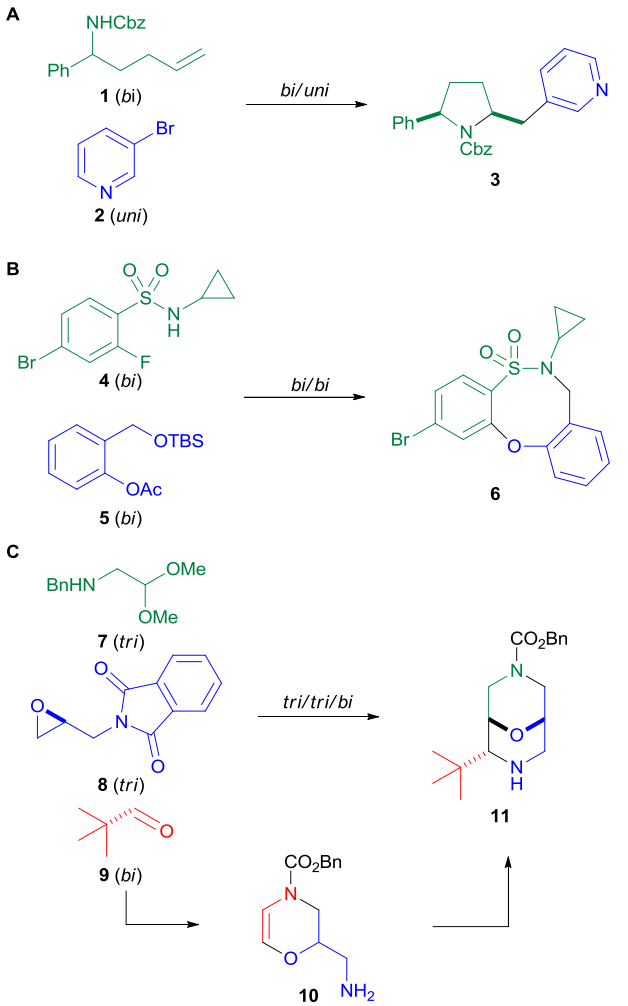
A recent survey[17] showed that the vast majority (>99%) of commercially-available screening compounds do not have lead-like properties.[¶] Furthermore, most compounds (~98%) reported in recent synthetic methodology papers are also not lead-like.[‖] This problem may stem, in part, from the rather narrow toolkit of reactions that is widely used to support drug discovery.[18,19] Although diversity-oriented approaches[10,20,21,22] have emerged to allow the variation of molecular scaffold, these approaches have tended to yield small molecules that lie well outside lead-like chemical space. The development of robust methodology for preparing large numbers of lead-like molecules for screening thus remains a significant and largely unmet challenge.

A comprehensive conceptual framework

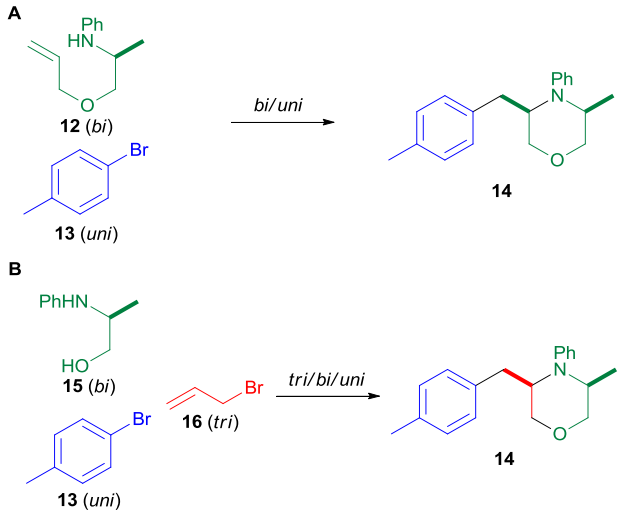
Some frameworks have been developed to describe alternative synthetic approaches to diverse molecular scaffolds. For example, the "build–couple–pair"[20,**] framework is broad, and has been used to describe synthetic approaches in which building blocks are combined to yield alternative molecular scaffolds. Ambiphile pairing reactions[23] (in which pairs of bifunctional building blocks are combined to yield scaffolds) may be considered to be a subset of reactions that exemplify the "build-couple-pair" approach.

Here, we describe a comprehensive conceptual framework that classifies synthetic approaches that convert building blocks into diverse small molecule scaffolds (often in more than one step). The framework is hierarchical, and therefore captures the relative power of different approaches. Synthetic approaches to molecular scaffolds are more powerful if more bonds are formed to individual building blocks, or if more building blocks are used.

Synthetic approaches are classified according to the number of new bonds formed to each of the building blocks in the conversion into molecular scaffolds; new bonds between functional groups *within* a building block are counted only once. Illustrative examples of three approaches to molecular scaffolds are shown in Scheme 1.[††] For example, the aminoarylation reaction to give the pyrrolidine **3** involves cyclisation of the (bi-connective) pentenamine derivative **1** with concomitant arylation (with the uni-connective aryl bromide **2**) (Panel A);[24] the aminoarylation reaction **1** + **2** → **3** is thus classified as a *bi/uni* process. An ambiphile pairing reaction (**4** + **5**) leading to the formation of a cyclic sultam (**6**) is shown in Panel B:[23] here, both of the building blocks are bi-connective, and so the overall approach is classified as a *bi/bi* process. The formation of the bridged bicyclic scaffold **11** is even more powerful because three building blocks are used, and three bonds are formed to two of these building blocks;[25] the approach exploits two tri-connective (**7** and **8**) and one bi-connective (**9**) building block, and thus receives the *tri/tri/bi* classification.



Scheme 1 Overview of our conceptual framework for classifying synthetic approaches to molecular scaffolds. The framework classifies the number of new bonds to each of the building block in the conversion into a molecular scaffold. Illustrative examples of three approaches to molecular scaffolds are provided: the application of an aminoarylation in the synthesis of pyrrolidines (Panel A), an ambiphile pairing reaction in the synthesis of cyclic sultams (Panel B), and a reaction sequence leading to bridged bicyclic scaffolds (Panel C). In each case, colour is used to show the correspondence between building block and substructures in intermediates and products; new connections between and within building blocks are shown in black.



Scheme 2 The classification describes the overall conversion of building blocks into molecular scaffolds. The classification of synthetic approaches thus depends on the building blocks that are exploited (compare Panels A and B). In this review, we consider that it should be possible to vary independently all of the building blocks used in a synthetic approach; thus, unless variation of the alkenyl building block in the illustrated aminoarylation reaction is possible (shown explicitly in Panel B), it would be appropriate to classify the reaction as a *bi/uni* process (Panel A).

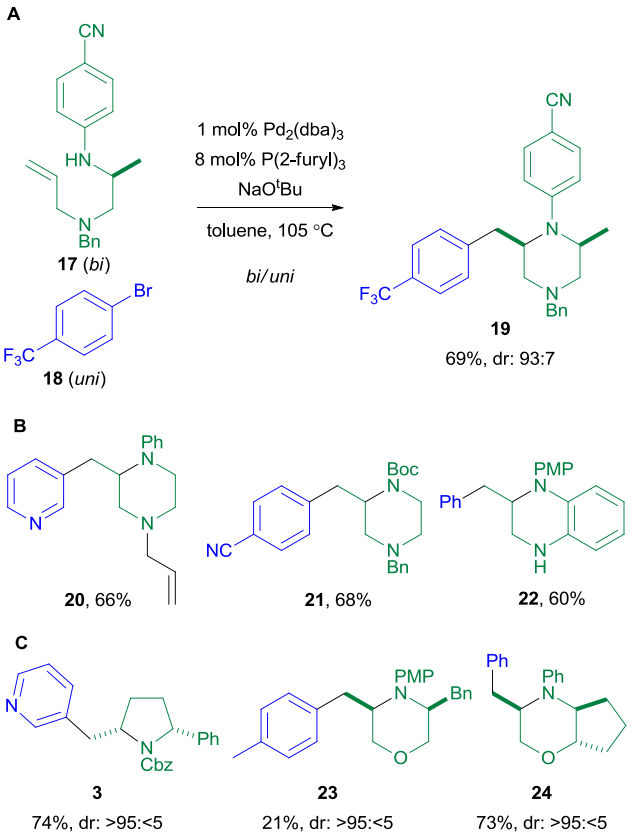
It is important to note that the classification describes the overall conversion of building blocks into molecular scaffolds. Thus, the classification depends on the specific building blocks that are considered to be starting materials (and not just the product scaffold). This classification thus captures the increased molecular complexity of the product scaffold relative to the starting building blocks. In this review, we consider that it should be possible to vary independently all of the building blocks. For this reason, the synthesis of the morpholine **14** (Scheme 2) is probably best classified as a *bi/uni* process (**12** + **13** → **14**, Panel A) because the component **16**, allyl bromide, (Panel B) was not varied. Ultimately, the value of specific synthetic approaches will depend on the accessibility of the starting materials, the increased molecular complexity of the product scaffolds, the diversity of the product scaffolds, and the molecular properties of the derivatives that may be prepared.

Approaches to the synthesis of lead-like scaffolds

This article focuses on synthetic approaches that may allow lead-like chemical space to be explored more systematically. A central challenge in this area is to develop synthetic approaches that allow variation of molecular scaffold within the boundaries of lead-like chemical space. Many reactions are systematically less successful with more polar building blocks,[17] and few reactions have been retooled to allow lead-like chemical space to be targeted more effectively.[26] This article focuses, therefore, on illustrative approaches that have emerged that allow (or have potential to allow) control over the scaffold of lead-like small molecules; it is not a comprehensive review of these approaches. The approaches are organised according to the conceptual framework that we have developed to classify, capture and extend approaches to the synthesis of diverse molecular scaffolds.

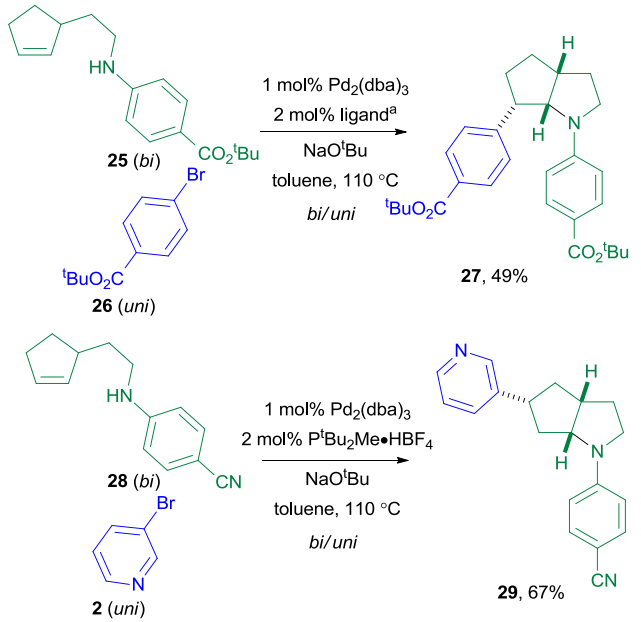
Approaches involving a bi- and a uni-connective building block

The simplest class of transformation highlighted in this article exploits reactions between a bi- and a uni-connective building block. We have deliberately not focused on cyclisations of a single uni-connective building block (which cannot allow combinatorial variation of the product scaffold) or reactions between two uni-connective building blocks (which are simply cross-coupling reactions).



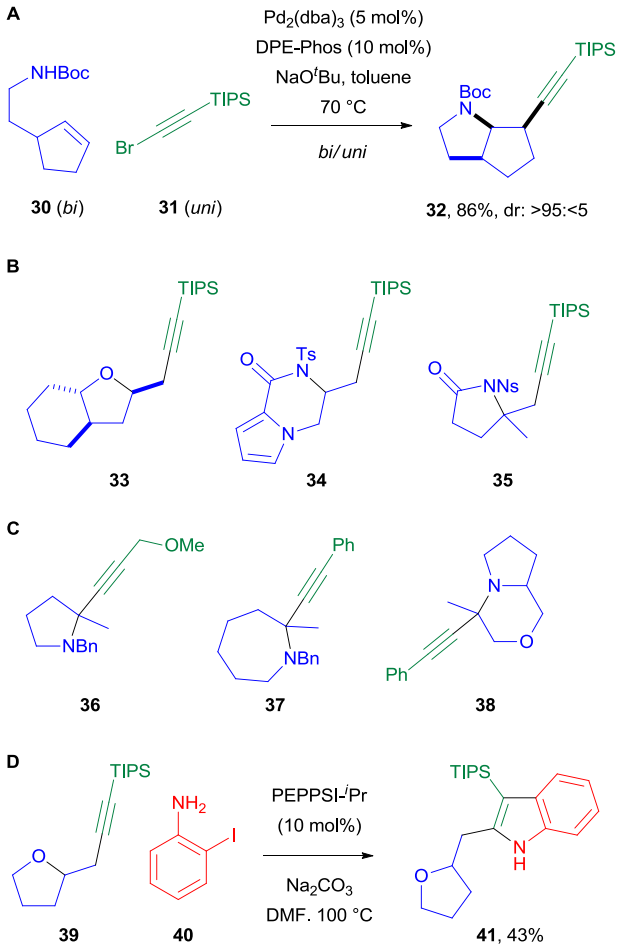
Scheme 3 Application of aminoarylation reactions in the synthesis of heterocyclic scaffolds. Panel A: Illustrative example of the approach in the synthesis of a piperazine. Panel B: Examples of other piperazines that may be prepared using the approach described in Panel A. Panel C: Examples of other heterocycles to which the general approach has been applied.

Aminoarylation reactions have been exploited in the synthesis of a wide range of heterocyclic scaffolds (Scheme 3). For example, Pd-catalysed cyclisation of (bi-connective) unsaturated amine derivatives (e.g. **17**), with concomitant arylation, enables the synthesis of piperazines (e.g. **19**-**22**), generally with high diastereoselectivity (Panels A and B).[27] Under related reaction conditions, the approach may be exploited in the synthesis of alternative heterocycles including pyrrolidines[24,28] (e.g. **3**) and morpholines[29] (e.g. **23** and **24**) (Panel C). The approach is likely to be most valuable when at least one nitrogen atom can be subsequently decorated (either directly, or after removal of a protecting group e.g. PMP or Boc).



Scheme 4 Control of the position of arylation in the formation of octahydrocyclopenta[*b*]pyrroles. [a]ligand = 2-diphenylphosphino-2'-(*N*,*N*-dimethylamino)biphenyl.

Remarkably, the regioselectivity of arylation could be controlled in cyclisations of substrates such as **25** and **28** (Scheme 4).[30] Crucially, careful choice of ligand controlled the position of the introduced substituent (e.g. → **27** or **29**) for many different aryl groups. In the case of **29**, migration of the palladium must occur before reductive elimination to give the product.



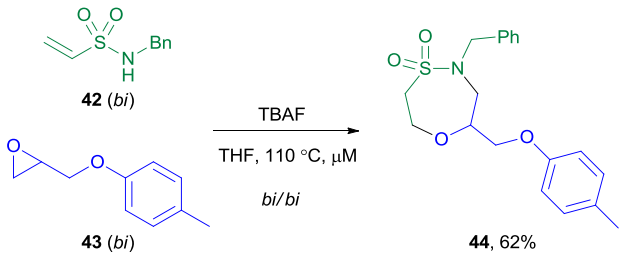
Scheme 5 Application of aminoalkynylation and related reactions in the synthesis of heterocyclic scaffolds. Panel A: Illustrative example of the approach. Panel B: Examples of other heterocycles that may be prepared using the approach. Panel C: Illustrative examples of products of a related Cu-catalysed cyclisation of alkynyl amines. Panel D: Exploitation of the alkyne functionality in a product scaffold as a handle for further functionalisation.

A valuable cyclisation reaction has been exploited in the synthesis of a range of heterocyclic scaffolds (Scheme 5).[31] Pd-catalysed cyclisation of unsaturated amine derivatives (e.g. **30**) could be followed by reaction with a terminal alkyne building block (e.g. → **32**) (Panel A). The approach could be extended to yield related scaffolds through cyclisation of alcohol (→ **33**) or *N*-sulfonyl amide (→ **34** or **35**) nucleophiles (Panel B). In addition, a related Cu-catalysed cyclisation reaction of alkynyl amines has been exploited in the synthesis of related scaffolds (e.g. **36**-**38**) (Panel C).[32] The alkyne substituent in the product has been shown to be a useful handle for further functionalisation (e.g. **39** + **40** → **41**; Panel D). The exploitation of the approach in the synthesis of lead-like scaffolds would require the judicious selection of appropriately polar and functionalised building blocks.

A range of other reactions between bi- and uni-connective building blocks have been developed that may be exploited in the synthesis of lead-like scaffolds: for example, Pd-catalysed aminoacetoxylation.[33] In many cases, variation of both building blocks has not been exemplified.

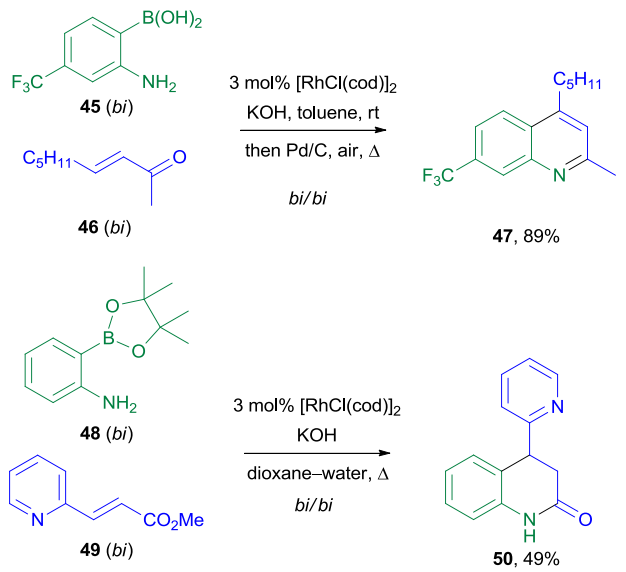
Approaches involving two bi-connective building blocks

Reactions between pairs of bifunctional building blocks can be extremely valuable in lead-oriented synthesis, particularly if variation of molecular scaffold is possible. Such transformations have previously been dubbed ambiphile pairing reactions.[23] For example, ring-opening of epoxides (e.g. **43**) with unsaturated sulfonamides (e.g. **42**) could be followed by oxa-Michael cyclisation to give the corresponding seven membered sultams (e.g. **44**) (Scheme 6).



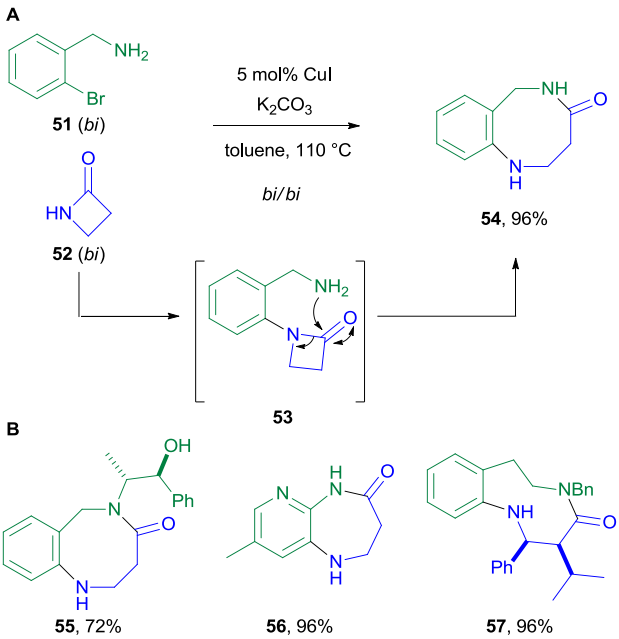
Scheme 6 Reaction between two bi-connective building blocks to yield a cyclic sultam scaffold.

Rh-catalysed Michael additions of *o*-substituted boronic acid derivatives have been exploited in the synthesis of a range of heterocycles (Scheme 7). For example, 2-aminophenyl boronic acids (e.g. **45**) were reacted with a range of α,β-unsaturated ketones (e.g. **46**); subsequent cyclisation, and oxidation, yielded the corresponding quinolines (e.g. **47**).[34] It was possible to vary the scaffold prepared by varying the chemistry exploited in the cyclisation step: for example, Rh-catalysed addition of the 2-aminophenyl boronate ester **48** to an α,β-unsaturated ester **49** yielded the corresponding tetrahydroquinolinone **50**.[35] It is envisaged that the approach may be extended to the synthesis of other heterocyclic scaffolds through further variation of the electrophile and the *o*-substituted phenylboronic acid derivative used.



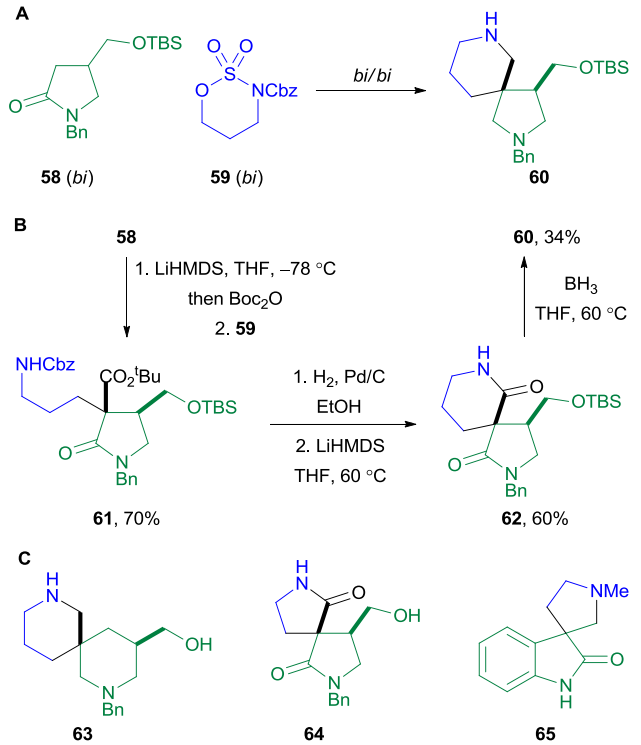
Scheme 7 Exploitation of Rh-catalysed conjugate addition reactions in the synthesis of heterocyclic scaffolds.

A range of benzo-fused medium lactam scaffolds have been prepared by Cu-catalysed amination, followed by ring expansion (Scheme 8).[36] For example, Cu-catalysed coupling of the aryl bromide **51** and the β-lactam **52** (→ **53**) was followed by ring expansion to give the medium ring lactam **54** (Panel A). Variation of the building blocks enabled the synthesis of related scaffolds, often with minor modification of the reaction conditions used.



Scheme 8 Synthesis of medium lactams by Cu-catalysed amination, and transamidation. Panel A: Illustrative example of the approach. Panel B: Examples of other lactams that have been prepared using a broadly similar approach.

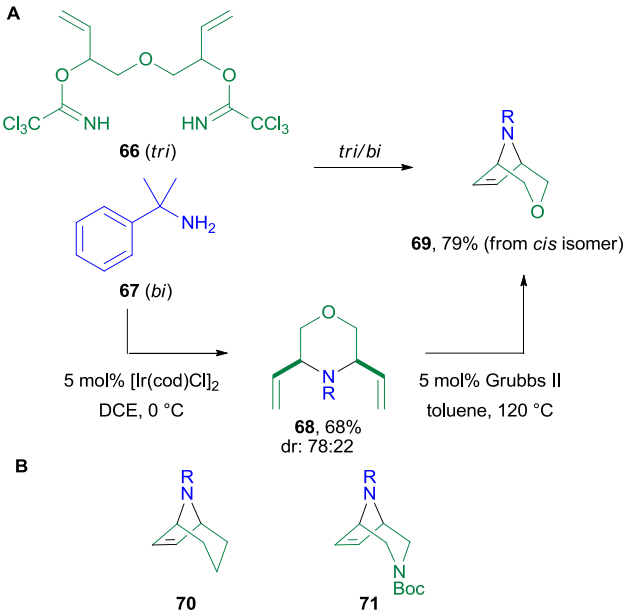
The reactions of cyclic sulfamidates with enolates have enabled the synthesis of a range of spirocyclic scaffolds (Scheme 9).[37] For example, enolisation of **58**, by treatment with LiHMDS, and reaction with Boc_2O yielded a stabilised enolate that was subsequently reacted with the cyclic sulfamidate **59** (→ **61**) (Panels A and B). In this example, we have not considered Boc_2O to be a building block because it cannot be varied to yield alternative scaffolds. Removal of the Cbz group triggered spirocyclisation (→ **62**); reduction gave the scaffold **60**. The approach was highly flexible, and a range of scaffolds could be prepared by varying the building blocks used: for example, it was possible to vary the ring size of both rings in the spirocycle (e.g. **63** and **64**), and to prepare benzo-fused scaffolds (e.g. **65**) (Panel C). In addition, cyclic sulfamidates are generally easily prepared from amino alcohols, which may allow extension to substituted analogues of these scaffolds.



Scheme 9 Synthesis of spirocyclic scaffolds by ring-opening of cyclic sulfamidates with enolates, and subsequent cyclisation. Panel A: Overview of the approach illustrated by a specific example. Panel B: Illustrative synthesis of a spirocyclic scaffold. Panel C: Examples of other spirocyclic scaffolds that have been prepared using a broadly similar approach.

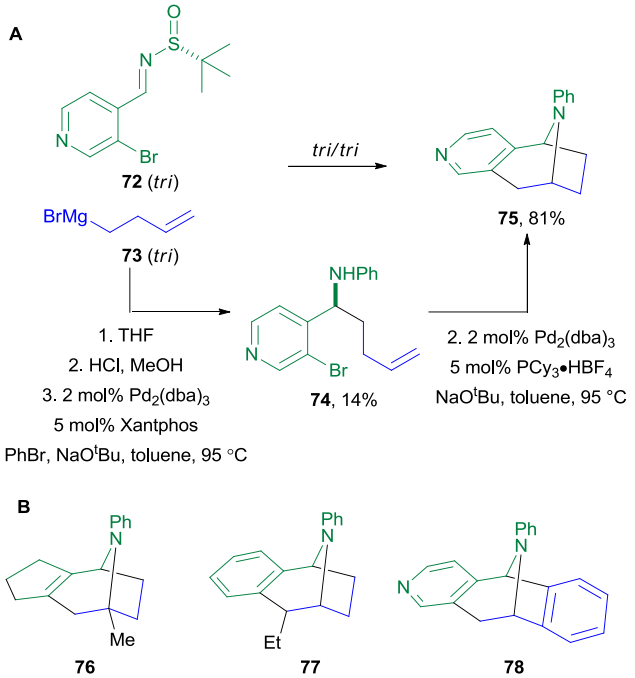
More complex approaches exploiting pairs of building blocks

Scaffolds with more complex topologies may be prepared by exploiting tri-connective building blocks. For example, the Ir-catalysed amination reaction between cumyl amine (**67**) and the bis-allylic trichloroacetimidate **66** gave the corresponding morpholine **68** with moderate diastereoselectivity in favour of the *cis* diasteroisomer (Panel A, Scheme 10).[38] The *cis*-disubstituted substrate **68** underwent a remarkable ring-closing metathesis reaction to give the bridged bicyclic scaffold **69**. The related scaffolds **70** and **71** have been prepared using an analogous approach (Panel B). In this reaction sequence, the exploitation of a tri-connective building block (**66**) enabled the synthesis of a bridged scaffold from a pair of building blocks.

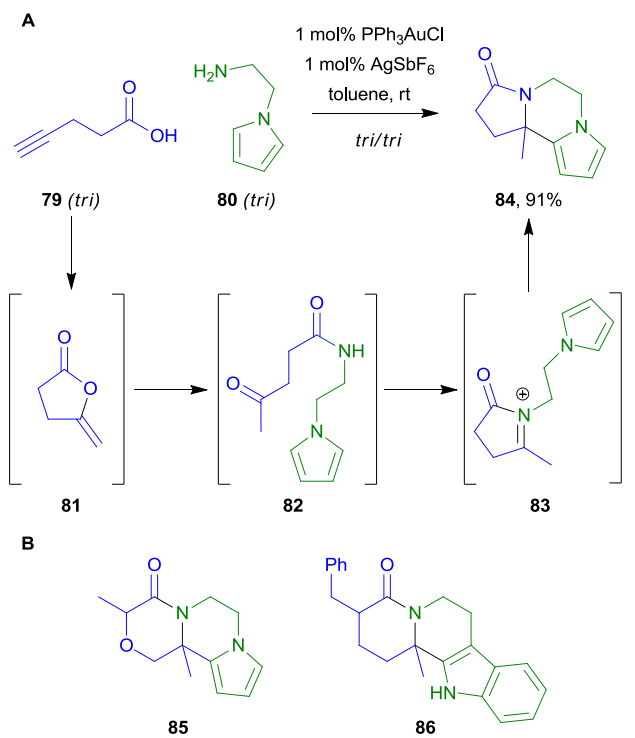


Scheme 10 Synthesis of bridged scaffolds by Ir-catalysed allylic amination, and ring-closing metathesis. Panel A: Illustrative example of the approach. Panel B: Examples of other scaffolds that have been prepared using an analogous approach. R = cumyl.

A Pd-catalysed aminoarylation reaction was exploited in the synthesis of bridged scaffolds such as **75** (Scheme 11).[39] The approach exploited two tri-connective building blocks – **72** and **73** – which were reacted together to give pentenamine derivatives such as **74** (Panel A). Subsequent Pd-catalysed aminoarylation (in which the amine, the alkene and the heteroaryl bromide are all in the same molecule) gave the bridged scaffold **75**. The approach was also exploited in the synthesis of alternative bridged scaffolds (including **76**, **77** and **78**) (Panel B). The overall approach illustrates the value of tri-connective building blocks in the synthesis of scaffolds with more complex topologies.



Scheme 11 Synthesis of bridged scaffolds by Pd-catalysed aminoarylation. Panel A: Illustrative example of the approach. Panel B: Examples of other scaffolds that have been prepared.

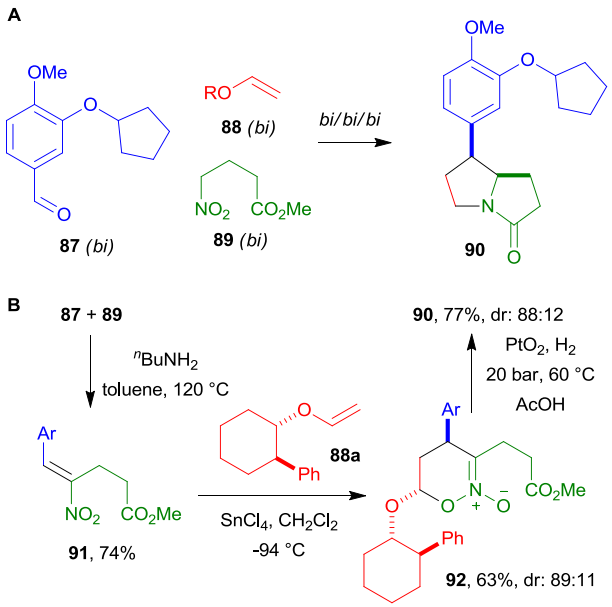


Scheme 12 Synthesis of tri- and tetracyclic scaffolds using a Au-catalysed cascade reaction. Panel A: Illustrative example of the approach. Panel B: Examples of related scaffolds that have also been prepared.

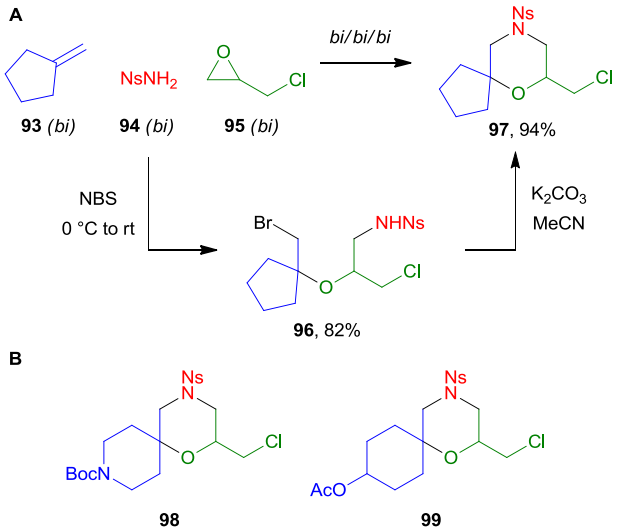
A one-pot Au-catalysed cascade has been exploited in the preparation of skeletally-diverse alkaloid-like small molecules (Scheme 12).[40] Initially, Au-catalysed cyclisation of alkynyl carboxylic acids (e.g. **79**) gave cyclic enol ethers (e.g. **81**) which underwent attack by amine nucleophiles (e.g. **80**) to give ketoamides (e.g. **82**) (Panel A). Under the same reaction conditions, the ketoamide intermediate (e.g. **82**) was then converted into an *N*-acyl iminium ion (e.g. **83**), which was trapped by a tethered nucleophile (e.g. → **84**). The approach has been exploited in the synthesis of a range of alkaloid-like scaffolds (e.g. **85** and **86**) by varying the combination of alkynyl carboxylic acid and heterocycle-tethered amine used (Panel B).

Extension to three or more building blocks

The power of approaches to small molecule scaffolds is greatly increased when more than two building blocks are used. For example, three bi-connective building blocks – the aldehyde **87**, the enol ether **88**, and the methoxycarbonyl-substituted nitro compound **89** – were exploited in the synthesis of the phosphodiesterase (PDE) IVb inhibitor **90** (Scheme 13).[41] Initially, the aldehyde **87** and the nitro compound **89** were condensed to yield the α,β-unsaturated nitro compound **91**. The nitroalkene **91** underwent efficient cycloaddition with the chiral enol ether **88a** to give the nitronate **92** with good diastereoselectivity. Finally, hydrogenation of **92** triggered reductive amination and amide formation to yield the bicyclic lactam **90**. It is likely that this synthetic approach could be extended to a range of related scaffolds with alternative substitution patterns.



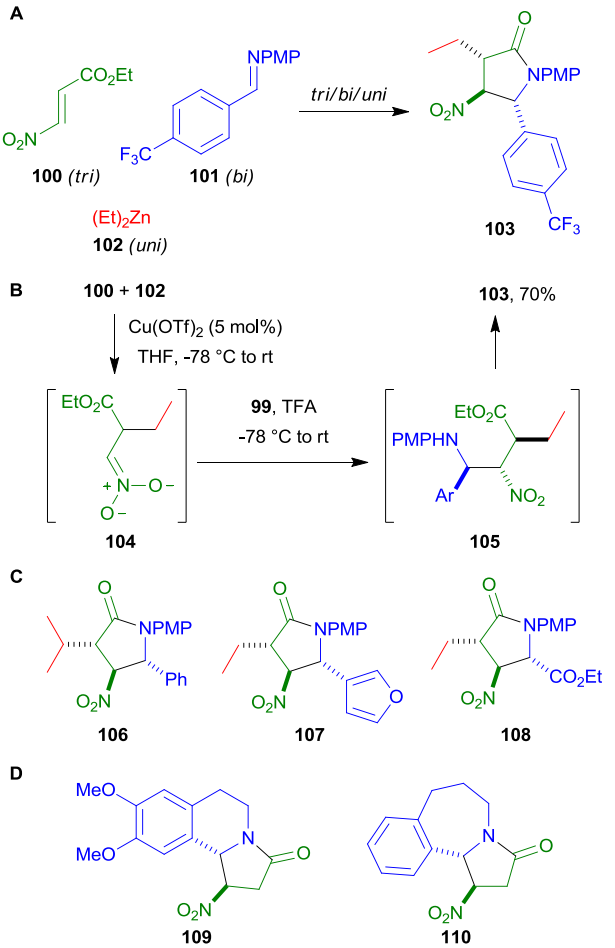
Scheme 13 Synthesis of the PDE IVb inhibitor **90**. Panel A: Overview of the synthesis of **90**. Panel B: Specific reactions exploited in the synthesis of **90**.



Scheme 14 A multicomponent approach to morpholine derivatives. Panel A: Illustration of the approach in the synthesis of the spirocyclic scaffold **97**. Panel B: Examples of related scaffolds that have also been prepared using the approach.

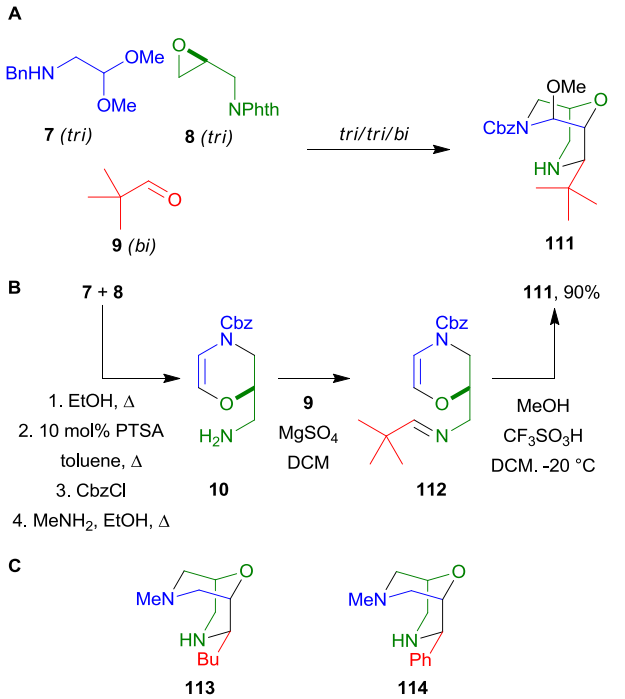
A multicomponent approach to morpholine derivatives that exploits three bi-connective building blocks is shown in Scheme 14.[42] Initially, treatment with NBS triggered the reaction between an alkene (e.g. **93**), an epoxide (e.g. **95**) and a sulfonamide (e.g. **94**) (→ **96**) (Panel A). Subsequent base-catalysed cyclisation yielded a morpholine derivative (e.g. **97**). The approach has also been exploited in the synthesis of more lead-like scaffolds in which the spiro-fused ring is heterocyclic (as in **98**) or substituted (as in **99**). In each case, the scaffolds are ripe for functionalisation, for example by substitution of chloride and/or after removal of the nitrophenylsulfonyl (Ns) protecting group.

A three-component reaction has been exploited in the synthesis of densely-functionalised pyrrolidinones (Scheme 15):[43] in this reaction, the three components were a (tri-connective) nitro acrylate (e.g. **100**), a (bi-connective) imine (e.g. **101**) and a (uni-connective) dialkylzinc reagent (e.g. **102**) (Panel A). Initially, conjugate addition of the dialkylzinc reagent yielded a nitro anion (e.g. **104**) which underwent a nitro-Mannich reaction with the imine (e.g. **101**) (→ **105**); finally, lactamisation yielded a pyrrolidinone scaffold (e.g. **103**). The approach has been exploited in the synthesis of a range of alternatively-substituted pyrrolidinones **106**-**108** (Panel C), and a related two-component process has yielded benzo-fused tricyclic scaffolds (e.g. **109** and **110**) (Panel D).[44]



Scheme 15 Synthesis of pyrrolidinones by a conjugate addition–nitro-Mannich–lactamisation cascade. Panel A: Overview of the approach. Panel B: Illustrative synthesis of a specific scaffold. Panel C: Examples of analogues that have been prepared using the approach. Panel D: Examples of scaffolds that may be prepared using a related two-component process exploiting cyclic imines.

The power of approaches to small molecule scaffolds further increases with the connectivity of the building blocks used. A synthesis of substituted oxabispidine scaffolds (Scheme 16) exploits two tri-connective building blocks – an amine-substituted epoxide (e.g. **8**) and an amino-substituted acetal (e.g. **7**) – and one bi-connective building block – an aldehyde (e.g. **9**) (Panel A).[25] Initially, the two tri-connective building blocks were condensed to give, after protecting group manipulation, an oxazine (e.g. **10**); subsequent condensation with an aldehyde (e.g. **9**) enabled formation of a bridged bicyclic scaffold which could be trapped either an α-methoxy amine (as in **111**) or as a benzotriazole adduct (not shown) (Panel B). The approach was exploited in the synthesis of a range of substituted bispidines such as **113** and **114** (Panel C). It is possible that additional substituents might be introduced by substitution (rather than simple reduction) of the intermediate α-methoxy amine (e.g. **111**) or benzotriazole adduct.

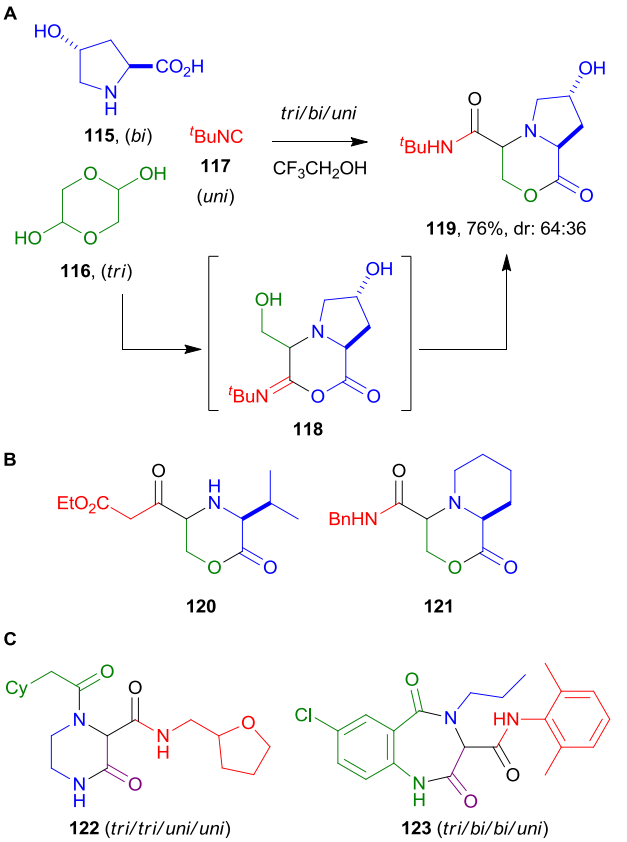


Scheme 16 Synthesis of bridged tricyclic oxabispidine scaffolds. Panel A: Overview of the approach. Panel B: Illustrative synthesis of a specific scaffold. Panel C: Examples of related scaffolds that have been prepared using the approach.

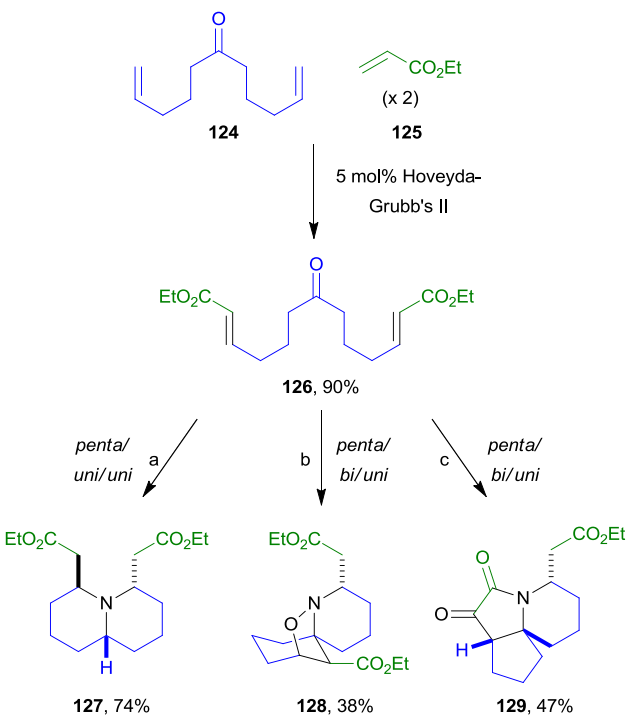
Approaches in which the connectivity of building blocks may be varied

The approaches that have thus far been highlighted in this article have not allowed wide variation of the connectivity of the building blocks used. However, approaches that do allow the connectivity of building blocks to be varied – crucially, whilst maintaining a common approach – are particularly valuable in the synthesis of many diverse molecular scaffolds.

Multicomponent reactions can enable highly efficient synthesis of diverse molecular scaffolds.[45,46] A three-component reaction between an amino acid (e.g. **115**), an aldehyde derivative (e.g. **116**) and an isocyanide (e.g. **117**) has been exploited in the synthesis of heterocyclic scaffolds (such as **119**) (Scheme 17).[47,48] Iminium ion formation, and attack of the isocyanide yielded a nitrilium ion which was intercepted (intramolecularly) by the carboxylic acid (→ **118**); subsequent rearrangement yielded the scaffold **119** (Panel A). By varying the amino acid used – i.e. using either an acyclic or a different cyclic amino acid – alternative scaffolds (e.g. **120** and **121**) could be prepared (Panel B). The overall approach is highly general because many alternative scaffolds are accessible through variation of the tethering strategy used. The scaffolds **119**-**121** were accessible by tethering the amine and the carboxylic acid (i.e. by using an amino acid building block). However, by tethering different nucleophiles to the amine component (e.g. → **122**) or the carboxylic acid component (e.g. → **123**), alternative scaffolds are accessible (Panel C).[49,50]

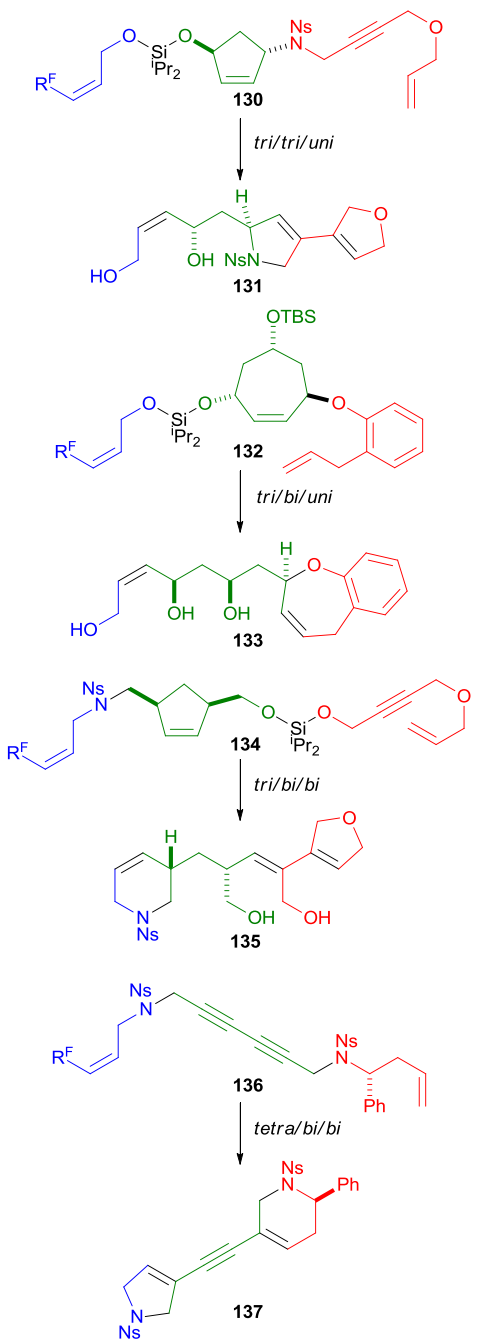


Scheme 17 Application of multicomponent reactions in the synthesis of diverse scaffolds. Panel A: Overview of a three component reaction involving an amino acid, an aldehyde derivative and an isocyanide. Panel B: Examples of scaffolds that have also been prepared using the approach. Panel C: Scaffolds that have been prepared using related four component reactions.



Scheme 18 Illustrative examples of diverse scaffolds accessible using a branching pathway. Reagents: (a) (i) NH_3, $NaBH_4$, $Ti(OEt)_4$, EtOH; (ii)

AcOH; (b) (i) $NH_2OH{\cdot}HCl$, NaOAc, MeCN; (ii) toluene, 140 °C; (c) (i) $NH_2OH{\cdot}HCl$, NaOAc, MeOH–MeCN, 60 °C; (ii) NaOEt, EtOH.



Scheme 19 Illustrative syntheses of natural product-like molecules with unprecedented scaffold diversity using metathesis cascade chemistry. The power of the approach lay in the ability to vary the connectivity of the building blocks exploited. R^F is a fluorous-tagged substituent.

Two-directional synthetic approaches have been exploited in the synthesis of diverse molecular scaffolds. The ketone **126**, prepared by cross-metathesis of the unsaturated ketone **124** and ethyl acrylate, has been exploited as the key intermediate in a branching pathway leading to twelve natural product-like scaffolds (Scheme 18).[51] Thus, reductive amination could be followed by double conjugate addition to give the alkaloid-like scaffold **127**; in this reaction, **124** functions as a penta-connective building block. The approach may be extended to related bicyclic scaffolds by using unsymmetrical substrates.[52] Alternatively, oxime formation from **126**, and conjugate addition, enabled the formation of a nitrone which could undergo intramolecular dipolar cycloaddition; alternative scaffolds could be prepared by exploiting either kinetic control (not shown) or thermodynamic control (**128**). Furthermore, fragmentation of one of these products enabled the preparation of the tricyclic scaffold **129**.

An approach to natural product-like molecules of unprecedented scaffold diversity (over 80 distinct scaffolds) has been developed (Scheme 19).[53] The approach relied on the iterative attachment of building blocks (shown in green and red) to a fluorous-tagged linker (shown in blue) using either permanent or temporary linkages to yield metathesis substrates (e.g. **130**, **132**, **134** and **136**). Metathesis cascade reactions were then used to 'reprogramme' the molecular scaffolds to give, after deprotection, skeletally-diverse products (e.g. **131**, **133**, **135** and **137**). The connectivity of the building blocks could be varied by exploiting both permanent and temporary linkages, and, by exploiting some building blocks for which intra-building block metathesis reaction was possible.

Conclusions

In this Feature Article, we have described a comprehensive conceptual framework that may be used to capture, analyse and plan synthetic approaches to diverse molecular scaffolds. A wide range of synthetic approaches that target, or have potential to target, lead-like chemical space were organised using the conceptual framework. Particular emphasis was placed on synthetic approaches that enable the combinatorial variation of molecular scaffold, particularly within lead-like chemical space.

Table 2 Relative power of synthetic approaches to molecular scaffolds

Building blocks		Power of approach	
Number	Connectivity	Generality	Increase in complexity
2	*bi/uni* or *bi/bi*	Often high	Low
2	≥1 building block with *tri* or higher connectivity	Usually low	High
≥3	Many combinations of building block connectivities possible	High in some cases	High

The framework is hierarchical, and can thus allow easy comparison of the relative power of dissimilar synthetic approaches. Specifically, the ability of approaches to increase molecular complexity – from starting building blocks to product scaffolds – is captured (Table 2).

There are many examples of synthetic approaches that exploit a pair of building blocks with low connectivity (*bi/uni* or *bi/bi* processes). In many cases, wide variation of the product scaffold is possible, for example by using either cyclic or acyclic building blocks, or by varying the distance between the reactive functional groups. However, the increase in molecular complexity – from building blocks to scaffolds – is inherently relatively low.

More powerful synthetic approaches are possible if the connectivity of the building blocks, or the number of building blocks exploited, is increased. For example, complex scaffolds are accessible if at least one of a pair of building blocks has higher (i.e. *tri* or higher) connectivity. Such approaches can allow more complex scaffolds to be prepared (such as bridged scaffolds), but the preparation of many different scaffolds using a single approach has rarely been exemplified (for a rare example, see Scheme 11).

Synthetic approaches that exploit more than two building blocks can sometimes be general – yielding many different scaffolds – and also increase molecular complexity greatly: for example, an approach exploiting three bi-connective building blocks has yielded a wide range of macrocyclic and medium ring scaffolds.[54] However, as the connectivity of the building blocks is increased, generality is usually compromised.

The connectivity of building blocks can sometimes be varied using a common synthetic approach; such approaches are remarkably powerful and general. The Ugi reaction, for example, has been exploited in the synthesis of an extremely wide range of scaffolds by tethering different reacting functional groups (for some examples, see Scheme 17). A synthesis of natural products of unprecedented scaffold diversity (>80 scaffolds) was possible because variation of the linkages between building blocks, and the connectivity of the building blocks, was possible (see Scheme 19).

Syntheses of diverse molecular scaffolds are still, however, dominated by a small number of highly general reactions, most notably alkene metathesis and the Ugi reaction. A major challenge for synthetic chemists is, therefore, to develop powerful, yet general, synthetic approaches to lead-like molecular scaffolds. Such approaches will help to address the uneven exploration of chemical space by synthesis, and may yield valuable starting points for drug discovery programmes.

Acknowledgements

We thank EPSRC, AstraZeneca and GSK for funding our research into the synthesis of diverse lead-like scaffolds. We thank members of the Nelson group and Ian Churcher, Andrew Grant, Ed Griffen, Steve Marsden, Iain Simpson and Stuart Warriner and for helpful discussions.

Notes and references

[a] *School of Chemistry, University of Leeds, Leeds, LS2 9JT, UK. Fax: +44 113 343 6401; Tel: +44 113 343 6502; E-mail: a.s.nelson@leeds.ac.uk*
[b] *Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, LS2 9JT, UK.*

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

‡ These data refer to the analysis of scaffold diversity at the graph-node level i.e. capturing connectivity and atom types (but not bond types) within scaffolds (see ref. 8).

¶ >99% of 4.9M commercially-available screening compounds failed at least one of the molecular size, lipophilicity or structural filters described in the Table.

‖98.2% of compounds reported in a selection of key synthetic chemistry journals in 2009 failed at least one of the molecular size, lipophilicity or structural filters described in the Table.

** In this approach, building blocks are prepared ("built"), linked ("coupled") and cyclised ("paired") to yield molecular scaffolds.

†† Throughput this article, colour is used to illustrate the correspondance between building blocks and substructures in intermediates and products. New connections between and within building blocks are shown in black.

1. (a) S. Wetzel, K. Klein, S. Renner, D. Rauh, T. I. Oprea, P. Mutzel, and H. Waldmann, *Nature Chem. Biol.* 2009, **5**, 581; (b) S. Renner, W. A. L. van Otterlo, M. Dominguez Seoane, S. Möcklinghoff, B. Hofmann, S. Wetzel, A. Schuffenhauer, P. Ertl, T. I. Oprea, D. Steinhilber, L. Brunsveld, D. Rauh, and H. Waldmann, *Nature Chem. Biol*. 2009, **5**, 585.
2. Y. Hu, A. M. Wassermann, E. Lounkine, and J. Bajorath, *J. Med. Chem.* 2010, **53**, 752.
3. G. V. Paolini, R. H. B. Shapland, W. P. van Hoorn, J. S. Mason, and A. L. Hopkins, *Nature Biotechnol.* 2006, **24**, 805.
4. S. Wetzel, R. S. Bon, K. Kumar, and H. Waldmann, *Angew. Chem. Int. Ed.* 2011, **50**, 10800.
5. 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. Med. Chem.*, 1988, **31**, 2235.
6. A. A. Shelat and R. K. Guy, *Nature Chem. Biol.* 2007, **3**, 442.
7. Y. Hu, D. Stumpfe, and J. Bajorath, *J. Chem. Inf. Model.* 2011, **51**, 1742.
8. A. H. Lipkus, Q. Yuan, K. A. Lucas, S. A. Funk, W. F. Bartelt, R. J. Schenck, and A. J. Trippe, *J. Org. Chem.* 2008, **73**, 4443.
9. M. Krier, G. Bret, and D. Rognan, *J. Chem. Inf. Model.* 2006, **46**, 512.
10. M. Dow, M. Fisher, T. James, F. Marchetti and A. Nelson, *Org. Biomol. Chem.* 2012, **10**, 17.
11. M. C. Wenlock, R. P. Austin, P. Barton, A. M. Davis, and P. D. Leeson, *J. Med. Chem.* 2003, **46**, 1250.
12. P. D. Leeson and B. Springthorpe, *Nat. Rev. Drug. Discov.* 2007, **6**, 881.
13. T. J. Ritchie and S. J. F. Macdonald, *Drug Discov. Today*2009, **14**, 1011.
14. F. Lovering, J. Bikker, and C. Humblet, *J. Med. Chem.* 2009, **52**, 6752.
15. T. I. Oprea, A. M. Davis, S. J. Teague, and P. D. Leeson, *J. Chem. Inf. Comput. Sci.* 2001, **41**, 1308.
16. G. M. Keserü and G. M. Makara, *Nat. Rev. Drug. Discov.* 2009, **8**, 203.
17. A. Nadin, C. Hattotuwagama, and I. Churcher, *Angew. Chem. Int. Ed.* 2012, **51**, 1114.
18. S. D. Roughley and A. M. Jordan, *J. Med. Chem.* 2011, **54**, 3451.
19. T. W. J. Cooper, I. B. Campbell, and S. J. F. Macdonald, *Angew. Chem. Int. Ed.* 2010, **49**, 8082.
20. T. E. Nielsen and S. L. Schreiber, *Angew. Chem. Int. Ed.* 2008, **47**, 48.
21. M. D. Burke and S. L. Schreiber, *Angew. Chem. Int. Ed.* 2004, **43**, 46.
22. W. R. J. D. Galloway, A. Isidro-Llobet, and D. R. Spring, *Nature Commun.* 2010, **1**, 80.
23. A. Rolfe, T. B. Samarakoon, and P. R. Hanson, *Org. Lett.* 2010, **12**, 1216.

24. M. B. Bertrand, M. L. Leathen, and J. P. Wolfe, *Org. Lett.* 2007, **9**, 457.
25. H. Brice, D. M. Gill, L. Goldie, P. S. Keegan, W. J. Kerr, and P. H. Svensson, *Chem. Commun.* 2012, **48**, 4836.
26. For an example, see: P. Tosatti, J. Horn, A. J. Campbell, D. House, A. Nelson, and S. P. Marsden, *Adv. Synth. Cat.* 2010, **352**, 3153.
27. J. S. Nakhla, D. M. Schultz, and J. P. Wolfe, *Tetrahedron* 2009, **65**, 6549.
28. M. B. Bertrand, J. D. Neukom, and J. P. Wolfe, *J. Org. Chem.* 2008, **73**, 8851.
29. M. L. Leathen, B. R. Rosen, and J. P. Wolfe, *J. Org. Chem.* 2009, **74**, 5107.
30. J. E. Ney and J. P. Wolfe, *J. Am. Chem. Soc.* 2005, **127**, 8644.
31. S. Nicolai and J. Waser, *Org. Lett.* 2011, **13**, 6324.
32. J. Han, B. Xu, and G. B. Hammond, *J. Am. Chem. Soc.* 2010, **132**, 916.
33. E. J. Alexanian, C. Lee, and E. J. Sorensen, *J. Am. Chem. Soc.* 2005, **127**, 7690.
34. J. Horn, S. P. Marsden, A. Nelson, D. House, and G. G. Weingarten, *Org. Lett.* 2008, **10**, 4117.
35. J. Horn, H. Y. Li, S. P. Marsden, A. Nelson, R. J. Shearer, A. J. Campbell, D. House, and G. G. Weingarten, *Tetrahedron* 2009, **65**, 9002.
36. A. Klapars, S. Parris, K. W. Anderson, and S. L. Buchwald, *J. Am. Chem. Soc.* 2004, **126**, 3529.
37. C. Hirschh, J. S. Parker, M. W. D. Perry, M. F. Haddow, and T. Gallagher, *Org. Lett.* 2012, **14**, 4846.
38. R. A. Brawn, C. R. W. Guimarães, K. F. McClure, and S. Liras, *Org. Lett.* 2012, **14**, 4802.
39. D. M. Schultz and J. P. Wolfe, *Org. Lett.* 2011, **13**, 2962.
40. T. Yang, L. Campbell, and D. J. Dixon, *J. Am. Chem. Soc.* 2007, **129**, 12070.
41. A. Y. Sukhorukov, Y. D. Boyko, Y. V. Nelyubina, S. Gerard, S. L. Ioffe, and V. A. Tartakovsky, *J. Org. Chem.* 2012, **77**, 5465.
42. J. Zhou, L. Zhou, and Y.-Y. Yeung, *Org. Lett.* 2012, **14**, 5250.
43. J. C. Anderson, L. R. Horsfall, A. S. Kalogirou, M. R. Mills, G. J. Stepney, and G. J. Tizzard, *J. Org. Chem.* 2012, **77**, 6186.
44. S. M.-C. Pelletier, P. C. Ray, and D. J. Dixon, *Org. Lett.* 2009, **11**, 4512.
45. E. Ruijter, R. Scheffelaar, and R. V. A. Orru, *Angew. Chem. Int. Ed.* 2011, **50**, 6234.
46. I. Akritopoulou-Zanze, *Curr. Opin. Chem. Biol.* 2008, **12**, 324.
47. I. W. Ku, S. Cho, M. R. Doddareddy, M. S. Jang, G. Keum, J.-H. Lee, B. Y. Chung, Y. Kim, H. Rhim, and S. B. Kang, *Bioorg. Med. Chem. Lett.* 2006, **16**, 5244.
48. Y. B. Kim, E. H. Choi, G. Keum, S. B. Kang, D. H. Lee, H. Y. Koh, and Y. Kim, *Org. Lett.* 2001, **3**, 4149.
49. C. Hulme, M. M. Morrissette, F. A. Volz, and C. J. Burns, *Tetrahedron Lett.* 1998, **39**, 1113.
50. C. Hulme and M.-P. Cherrier, *Tetrahedron Lett.* 1999, **40**, 5295.
51. D. Robbins, A. F. Newton, C. Gignoux, J.-C. Legeay, A. Sinclair, M. Rejzek, C. A. Laxon, S. K. Yalamanchili, W. Lewis, M. a. O'Connell, and R. A. Stockman, *Chem. Sci.* 2011, **2**, 2232.
52. K. M. G. O'Connell, M. Díaz-Gavilán, W. R. J. D. Galloway, and D. R. Spring, *Beilstein J. Org. Chem.* 2012, **8**, 850.
53. D. Morton, S. Leach, C. Cordier, S. Warriner, and A. Nelson, *Angew. Chem. Int. Ed.* 2009, **48**, 104.
54. L. A. Marcaurelle, E. Comer, S. Dandapani, J. R. Duvall, B. Gerard, S. Kesavan, M. D. Lee, H. Liu, J. T. Lowe, J.-C. Marie, C. A. Mulrooney, B. A. Pandya, A. Rowley, T. D. Ryba, B.-C. Suh, J. Wei, D. W. Young, L. B. Akella, N. T. Ross, Y.-L. Zhang, D. M. Fass, S. A. Reis, W.-N. Zhao, S. J. Haggarty, M. Palmer, and M. A. Foley, *J. Am. Chem. Soc.* 2010, **132**, 16962.