Access to High Value Natural and Unnatural Products through Hyphenating Chemical Synthesis and Biosynthesis

Kevin P. P. Mahoney, Duncan R. M. Smith, Emma J. A. Bogosyan, Rebecca J. M. Goss*

School of Chemistry and BSRC, University of St Andrews, St Andrews, KY16 9ST, UK
Fax +44(1334)463808; E-mail: rjmg@st-andrews.ac.uk

Received: 08.04.2014; Accepted after revision: 04.06.2014

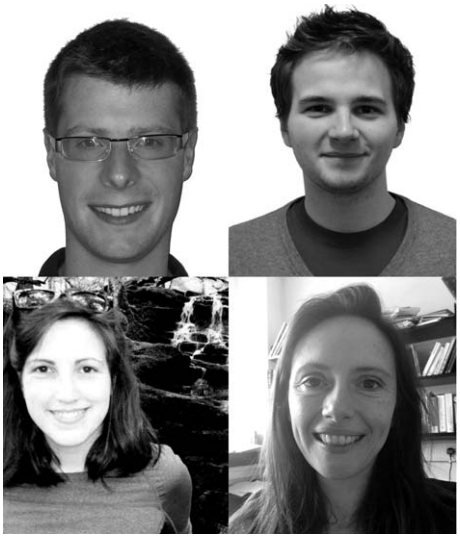
Abstract: Access to natural products and their analogues is crucial. Such compounds have, for many years, played a central role in the area of drug discovery as well as in providing tools for chemical biology. The ability to quickly and inexpensively acquire genome sequences has accelerated the field of natural product research. Access to genomic data coupled with new technologies for the engineering of organisms is resulting in the identification of large numbers of previously undiscovered natural products as well as an increased understanding of how the biosynthetic pathways responsible for the biogenesis of these compounds may be manipulated. This short review summarizes and reflects upon approaches to accessing natural products and has a particular focus on approaches combining molecular biology and synthetic chemistry.

1 Introduction
2 CHEM: Total Synthesis
3 BIO-CHEM: Semi-Synthesis
4 CHEM-BIO: Precursor Directed Biosynthesis
5 BIO-CHEM-BIO: Mutasynthesis
6 BIO-BIO: Combinatorial Biosynthesis
7 BIO-BIO-CHEM: Genochemetics: Gene Expression Enabling Synthetic Diversification
8 Conclusions and Future Directions in Natural Product Analogue Generation

Key words: alkyl halides, antibiotics, antifungal agents, antitumor agents, biosynthesis, chemoselectivity, cross-coupling

1 Introduction

Natural products provide an unparalleled starting point for drug discovery – over 60% of anticancer agents and over 70% of antibiotics entering clinical trials in the last three decades were based on such compounds.[1] In order to gain a full understanding of how a drug works and to generate compounds with improved biological activity and physicochemical properties, ready synthetic access to series of analogues is essential. In recent years, pharmaceutical industries have shied away from natural products due to the perceived synthetic intractability of such analogues and the misconception that it is not possible to carry out thorough structure activity relationship (SAR) assessment on these compounds. Libraries of natural products are often not readily accessible using conventional synthetic organic chemistry alone; this has caused the pharmaceutical industry to pursue combinatorial libraries of non-natural moieties. Whilst these combinatorial libraries may be easily made, they have so far delivered little of pharmaceutical interest. Christopher Lipinski, a key figure and authority in medicinal chemistry, urged in a C&EN cover story for scientists to return to and focus on natural products for drug discovery. As a result of largely abandoning natural products, industry's drug discovery pipelines are beginning to run dry.[2]



Kevin Mahoney (top left) completed his BSc in Medicinal Chemistry at Queen's University Belfast in 2009. He then worked in the pharmaceutical industry as a process and development chemist for two years before joining the Goss group in 2011 to study towards a PhD in elucidating and manipulating natural product biosynthesis, with a particular focus upon glycosylation.
Duncan Smith (top right) completed his MChem degree at the University of East Anglia in 2012, where he undertook a Masters research project within the Goss group. He then joined the Goss group as a PhD student and moved with the group to the University of St Andrews, where he is currently working on designing genochemetic systems that use epoxides as functionalizable handles.
Emma Bogosyan (bottom left) completed her BSc degree in chemistry at the University of East Anglia where she received the Katritzky Prize and Medal for outstanding performance in the Final Assessment. She is currently studying towards a PhD with Dr Rebecca Goss developing genochemetic systems.
Rebecca Goss (bottom right) is a Reader (Associate Professor) in Organic Chemistry and Biomolecular Chemistry in the School of Chemistry and BSRC at the University of St Andrews Scotland. Her research focuses on natural products of medicinal value, elucidating biosynthetic pathways, and in combining organic synthesis and biosynthesis to enable rapid access to series of natural product analogues. She is the recent recipient of an ERC Consolidator grant.

SYNTHESIS 2014, 46, 2122–2132
Advanced online publication: 09.07.2014
DOI: 10.1055/s-0034-1378359; Art ID: ss-2014-m0231-sr

Natural products are privileged molecules; being generated by proteins, they are predisposed to interact with proteins. Rather than dismiss such a treasure trove of useful drug leads, new approaches to address access to synthetically challenging series of natural product analogues are required. This short review highlights the approaches available for accessing new to nature compounds, examines their pros and cons and provides case studies exemplifying these routes to these designer molecules. If analogue access is the true goal, the philosophy must be to utilize the most expeditious combinations of synthetic organic chemistry and harnessed biosynthesis, as dictated by the natural product target and the desired analogues, to yield sufficient of each analogue for downstream studies.

To best summarize the different approaches that are available to access natural product analogues we defer to the simple classification as defined by Kirschning and Hahn,[3] which we have previously utilized.[4] In this system, two abbreviations are used: CHEM, referring to total or partial chemical synthesis, and BIO, describing a series of biosynthetic steps. For example, precursor-directed biosynthesis, where a biosynthetic pathway is coerced to accept and incorporate a synthetic substrate into a natural product, is termed CHEM-BIO; the biosynthetic substrate analogue is first synthesized (CHEM), then administered to an organism and incorporated by the organism's biosynthetic machinery (BIO) to generate an analogue of the natural product.

2 CHEM: Total Synthesis – The Traditional Approach

The two traditional methods of accessing natural product analogues are total synthesis (CHEM) of which a series of highlights are reviewed in 'Classics in Total Synthesis',[5] and semi-synthesis (BIO-CHEM, section 3). For certain natural products, total synthesis may currently provide the only means of access as isolation from a rare natural source may only have been achieved once. In other cases, the organism may be difficult to cultivate, intractable to genetic or chemical manipulation, or produce only low levels of the natural product or the analogues. For simple natural products, total synthesis may represent the most economical and expedient approach to accessing series of analogues or even the parent natural product itself.

A multitude of pharmaceutically attractive compounds are found in the marine environment.[6] Bioactive compounds which are isolated from tunicates and sponges may often prove difficult to culture in laboratories and access to more than a few milligrams of such compounds is uncommon leaving total synthesis providing the sole option. Eribulin mesylate (Figure 1), a polyketide macrocycle (marketed by Eisai Ltd. as Halaven) has potent anticancer activity and is used in the treatment of breast cancer. This analogue of halichondrin B (Figure 1), produced by the *Halichondria* genus of sponge, is structurally complex and is accessed through an impressive total synthesis.[7,8] Halaven represents a step-change for the pharmaceutical industry, being considerably more complex than other marketed pharmaceuticals that have been accessed through total synthesis and demonstrating the enhanced skill and determination of this industry.

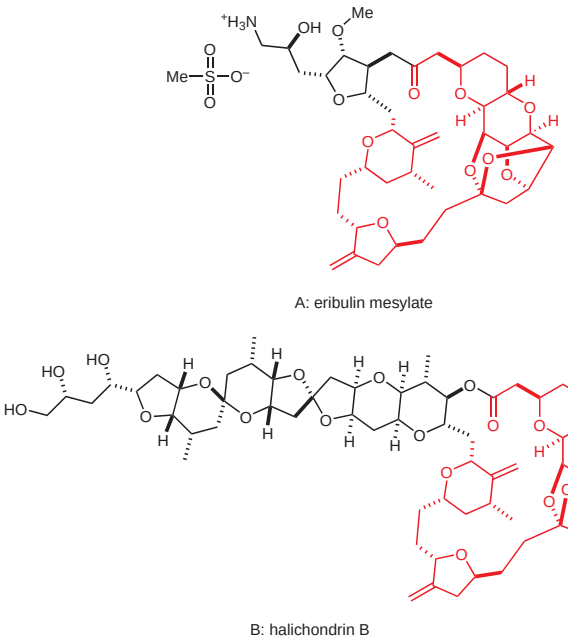
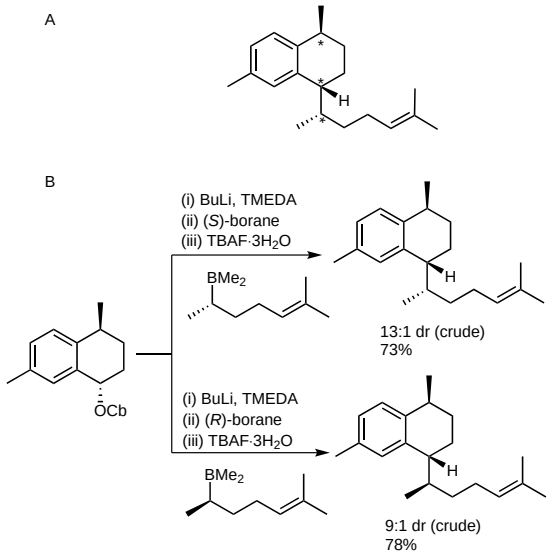


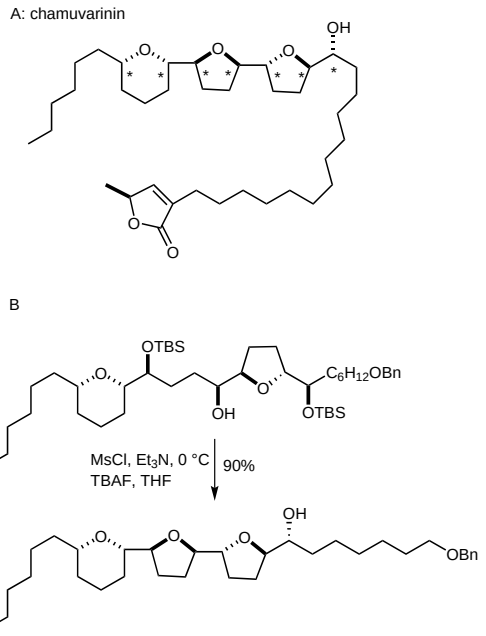
Figure 1 Eribulin mesylate (**A**) with 19 stereogenic atoms is a pharmaceutically optimized analogue of halichondrin (**B**) and provides a challenging case for total synthesis, which is the only option for non-innate production due to a lack of natural intermediates necessary for semi-synthetic processes.[8]

Total synthesis, as well as being essential to enable access to certain bioactive natural products, can also be invaluable for full structural and stereochemical characterization of newly isolated compounds or providing insight into a natural product's biogenesis. The exquisite structures of many of these molecules has provided inspiration for synthetic chemists to develop novel methods, of particular note is Aggarwal's elegant assembly of polyketides and polyketide like molecules using lithiation-borylation chemistry (Scheme 1).[9]

It is not always possible for isolation chemists to fully structurally characterize the natural products that they discover by spectroscopic methods alone and in certain challenging cases total synthesis is essential to confirm structural characterization.[10–13] Examples include the synthesis of chamuvarinin (Scheme 2), an acetogenin from the roots of the West African plant *Uvaria chamae*, enabling the relative and absolute stereochemistry to be assigned across the rare motif at the center of the structure in which a tetrahydropyran is linked to a bis-tetrahydrofuran;[14] synthesis of the C1–C12 fragment of tedanolide C (Scheme 3);[15] and synthesis of the stereochemically controversial C′D′E′F′ domain of vast marine polyketide maitotoxin (Scheme 4).[16]

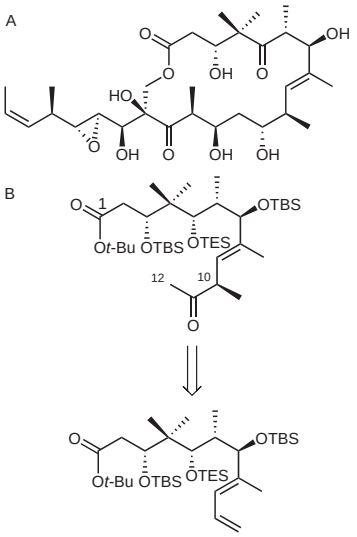


Scheme 1 The novel lithiation-borylation methodology has been showcased in the total synthesis of marine diterpene natural product (+)-erogorgiaene (**A**), which enabled sterechemical control at 3 stereocenters (*) by utilizing mixed boranes (**B**).[9]

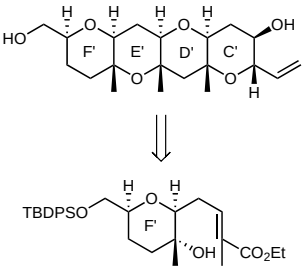


Scheme 2 Relative configuration of (+)-chamuvarinin (**A**), proposed by the Florence group, with 7 stereocenters (*) where relative and absolute configuration was determined through synthesis of the bis-tetrahydrofuran adjacent to a tetrahydropyran **B**.[14]

The total synthesis of marinomycin by the Evans group provides interesting insight into the compound's potential biogenesis (Figure 2). It is revealed that macrolactonization is shut down due to deactivation of the carbonyl if the hydroxyl of the salicylate moiety is available to hydrogen-bond to this motif. This is potentially a means of tightly controlling the macrolatonization of the two polyketide components of the dimer.[17] This insight is likely to apply to all salicylate lactones and esters.



Scheme 3 Tedanolide C (**A**). **B**: The C1–C12 domain constructed from a catalytically controlled asymmetric hydroformulation reaction.[15]



Scheme 4 Retrosynthetic analysis of the C′D′E′F′ domain of maitotoxin to obtain the desired 9 stereogenic centers with 3 consecutive methyl groups all axially placed on rings D′ and E′.[16]

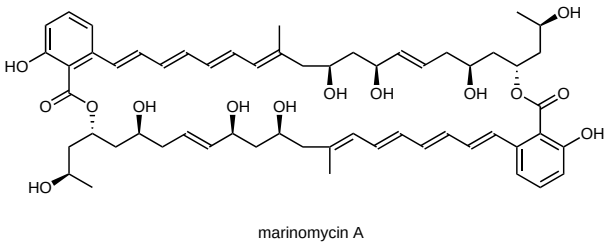


Figure 2 Marinomycin A[17]

Whilst total synthesis may currently be the only way of accessing a particular natural product or its analogues, advances in organism cultivation, genetic manipulation, and even synthetic biology mean that for many microbial and some plant-derived natural products an approach that harnesses biosynthesis may be more efficient.

3 BIO-CHEM: Semi-Synthesis

The biosynthesis and purification of natural products followed by diversification via chemical synthesis is collectively referred to as semi-synthesis. Semi-synthesis

involves the selective modification of a natural product's existing functional groups and relies upon chemical orthogonality present in the molecule. Just a few years after penicillin was used clinically, widespread resistance to this drug due to β-lactamase activity was apparent; to overcome this semi-synthetic analogues were developed.[18,19] Vancomycin has been used against Gram positive bacterial infections including methicillin resistant *Staphylococcus aureus* for 6 decades; however, in recent times it has been observed that its potency has been reduced and vancomycin-resistant *Enterococci* and *S. aureus* strains have emerged.[20] In analogy to the penicillin story, synthetic and semi-synthetic analogues of vancomycin are being sought to combat resistant strains[21,22] Peptidoglycan is a key component in bacterial cell walls, providing much of the strength and rigidity. It is composed of a β-1,4-linked glycan backbone of alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) units. Pendant pentapeptide chains are attached to the 3-position of each muramic acid with the terminal residues being D-Ala-D-Ala; neighboring this motif are either L-lysine or *meso*-diaminopimelic acid (DAP) residues. Crosslinking of these residues by amide bond formation serves to rigidify the polymer. Vancomycin (Figure 3) prevents cross linking by forming 5 hydrogen bonds to D-Ala-D-Ala, to overcome this resistant strains have evolved in which replacement of D-Ala-D-Ala with D-Ala-D-Lac reduces hydrogen bonding to vancomycin.[23] Analogues capable of killing vancomycin resistant strains are being sought. Recent work in this area includes the Boger lab's work on vancomycin aglycone converting the chloride to a boronic acid giving rise to a variety of functional groups via cross-coupling modification (Figure 3).[23]

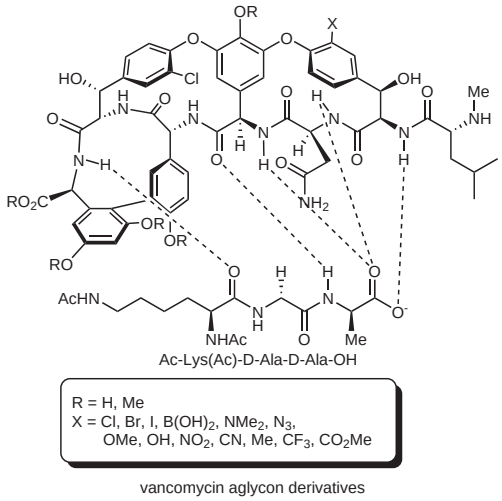
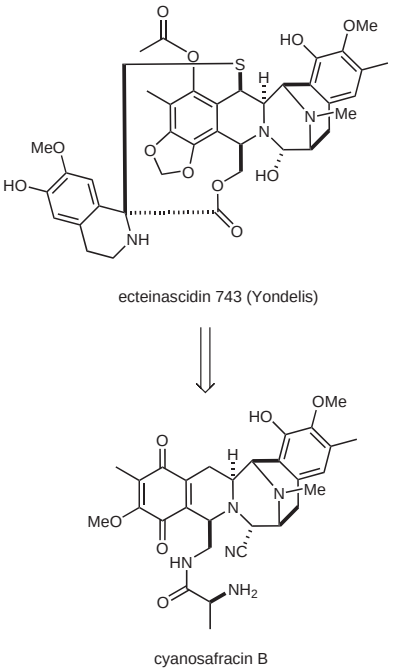


Figure 3 Derivatives of the vancomycin aglycone E-ring aryl chloride to determine its effect on D-Ala-D-Ala binding and antimicrobial activity.[23] Scheme adapted from Pathak.[24]

Complementary to this research, Miller has recently been taking semi-synthesis in an exciting new direction utilizing peptides, as catalysts, to control the site-selective halogenation of vancomycin and the related glycopeptide antibiotic teicoplanin;[24,25] basing the design of the peptide agent on the vancomycin binding site, regioselective halogenation has been achieved.[24] Using screened peptide libraries the Miller group have also been able to effect the semi-synthetic regiocontrolled halogenation of benzamides[26] and epoxidation of various isoprenols.[27]

Ecteinascdin 743 (Scheme 5), brand name Yondelis, is a structurally complex antitumor agent, of marine origin, generated by semi-synthesis.[28] Yondelis, initially discovered in a marine tunicate, is marketed by PharmaMar for the treatment of advanced or metastatic soft tissue sarcoma as well as for relapsed platinum-sensitive ovarian cancer. Due to low levels of production in the sea squirt, total synthesis was originally sought for this molecule and whilst many successful attempts have been made to obtain it on a large scale.[29,30] The drug is currently supplied via semi-synthesis from cyanosafracin B (Scheme 5).[28]

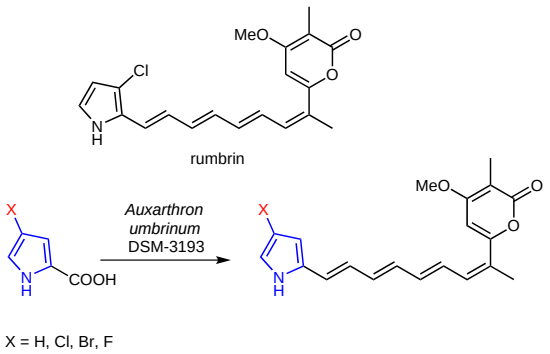


Scheme 5 Ecteinascidin 743 is produced on multi-gram scale via semi-synthesis using readily available cyanosafracin B[28]

4 CHEM-BIO Precursor Directed Biosynthesis

In precursor directed biosynthesis, the producing organism is fermented in the presence of a chemically synthesized analogue of a biosynthetic intermediate – a process referred to as 'analogue feeding'. The intention is that the precursor analogue is taken up by the producing organism and incorporated into the biosynthetic pathway resulting in the in vivo generation of the unnatural product. For precursor directed biosynthesis to be successful, the intermediate precursor must be synthetically accessible, must be taken up into the cell of the producing organism, and must be accepted as a substrate by downstream biosynthetic enzymes.

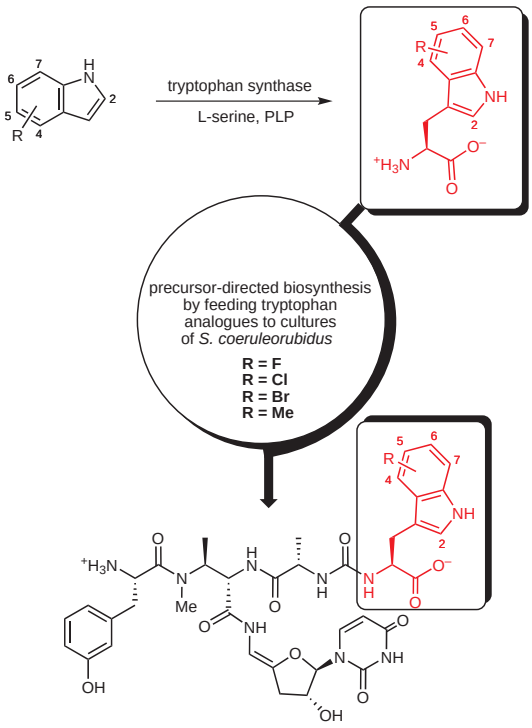
As a general rule, smaller modifications are more likely to successfully result in analogue production. Additionally, there is generally a greater degree of tolerance towards modified starter units in polyketide Synthase/Non-Ribosomal Peptide Synthetase (PKS/NRPS) systems, as seen in the generation of analogues of the PKS-derived cytoprotectant rumbrin (Scheme 6).[31] In the biosynthesis of rumbrin, the pyrrol-2-carboxylate starter unit is derived from proline before tethering to the PKS machinery.[32] By feeding analogues of the pyrrol-2-carboxylate (naturally chlorinated at the 2 position in rumbrin) Murphy et al. were able to demonstrate that the 3-chloropyrrole moiety was not essential for cytoprotectant activity, and observed improvements in activity against various cancer cell lines – most notably IC_{50} values of 35 nM and 28 nM for 3-bromoisorumbrin against HeLa (cervical) and A549 (lung) cancer cell lines respectively.[31]



Scheme 6 Feeding analogues of pyrrol-2-carboxylate to the rumbrin producer resulted in the generation of rumbrin analogues, which showed improved activity against various cancer cell lines[31]

MacMillan's access to a diverse suite of biologically active ammosamides, using precursor-directed biosynthesis without prior need for strain engineering, demonstrates how substrate flexible some biosynthetic pathways may be. The ammosamides are amidine analogues, which belong to the alkaloid ammosamide family and are isolated from the marine microbe *Streptomyces variabilis*. Supplementation of the fermentation media of *Streptomyces variabilis* with different aryl and alkyl amines gave access to a library of precursor-directed ammosamide analogues with a number of analogues demonstrating sub μM IC_{50} (0.4 to 0.8 μM) in vivo activities against a panel of non-small cell lung cancer (NSCLC) cell-lines.[33]

Analogue access, using precursor directed biosynthesis, may often be limited by specificity towards both position and type of substituent present within the analogue, as seen by the Goss group in the precursor-directed biosynthesis of derivatives of the antibiotic pacidamycin 4 upon feeding of tryptophan analogues (Scheme 7).[34] Good incorporation of the analogue was observed for halogen and methyl substituents at the 7 position and a methyl substituent at the 2 position of the tryptophan aromatic ring, but this incorporation was not as successful for substituents at other positions.[34] Previous work by this group had shown



Scheme 7 The Goss group have demonstrated pacidamycin 4 derivatives can be accessed using precursor directed biosynthesis by feeding tryptophan derivatives, which have been generated using tryptophan synthase to the pacidamycin producer *S. coeruleorubidus*[34]

that there was some flexibility in the terminal amino acids of pacidamycin, and using precursor directed biosynthesis several fluorophenylalanines and 2-chlorophenylalanine were incorporated into the natural product at either the C-terminus or both C- and N-termini. Strikingly, the titre of di-2-fluorophenylalanine derivative was over four times that of the corresponding unmodified natural product, and several other analogues also saw improved production.[35]

As the biosynthetic pathway of the natural product is left intact, in precursor directed biosynthesis, there will be production of the unmodified natural product alongside any production of an analogue. The relative levels of modified and unmodified natural product can vary depending on a number of factors, namely the ease of cellular uptake of the compound by the producing organism, levels of supply of the natural analogue, and the compatibility of the modified intermediate and the biosynthetic pathway. If an analogue is not a good substrate for the biosynthetic enzymes, this can lead to an undesirable accumulation of an intermediate and limited production of the natural product analogue. The synthesis and feeding of an intermediate downstream of a biosynthetic bottleneck can alleviate this problem, and if the natural intermediate at this stage is available only in limiting quantities, supplementation with an analogue can boost overall production of the natural product by bypassing a slow part of the biosynthesis.

The O'Connor group presented an elegant solution to such problems that were faced in the in planta generation of halogenated indole alkaloid natural products in *Catharanthus roseus*. They found that after transforming the producer with a tryptophan-7-halogenase, the halotryptophan intermediate was successfully generated, but proved to be a poor substrate for the next biosynthetic enzyme tryptophan decarboxylase. This led to accumulation of the 7-chlorotryptophan intermediate rather than the generation of the desired halogenated metabolites. By engineering the halogenase to act directly on the tryptophan decarboxylase product, tryptamine, the halogenation event was moved one step along the biosynthetic pathway, enabling the biosynthesis of these new to nature metabolites.[36]

In recent years, there have been efforts to develop the technique of precursor directed biosynthesis by engineering hosts. Well-known microbes such as *Escherischia coli* and *Saccharomyces cerevisiae* are attractive vehicles for heterologous natural product generation, but are not predisposed to the generation of the wide range of plant polyketide natural products as they lack several of the required metabolic pathways found in plants. By the feeding of a range of precursors that are difficult to access in the bacterial hosts, a wide range of plant natural products can be heterologously produced. However, as the pathways are foreign to the microbial hosts there is typically little regulation of production of toxic intermediates, which can negatively impact the growth of the bacteria.[37] *E. coli* has also been engineered to better yield erythromycin analogues from synthetic precursors, allowing access to an alkynated erythromycin analogue that enables further semi-synthetic diversification through click chemistry.[38] This shows that host engineering can facilitate precursor directed biosynthesis in otherwise untenable systems.

5 BIO-CHEM-BIO: Mutasynthesis

Mutasynthesis involves the genetic manipulation of a biosynthetic pathway prior to carrying out feeding of precursor analogues. There are two approaches to a mutasynthetic strategy:

i) Disrupting the biogenesis of the natural biosynthetic precursor:

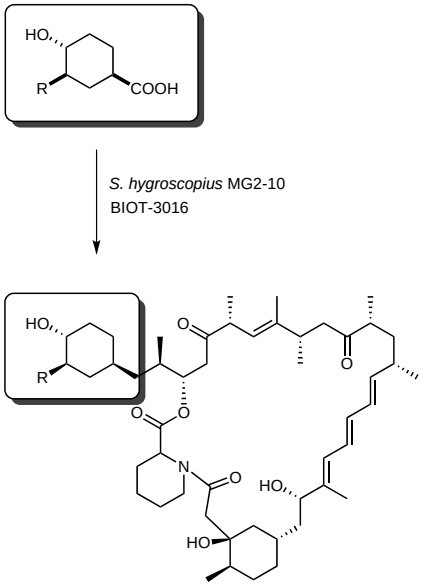
Without the natural precursors for biosynthesis the organism is unable to make the natural product. Natural product analogues are generated only if a suitable replacement for the natural precursor is administered to the organism. An advantage of this approach is that such knock-outs provide a blank canvas for analogue generation; analogues are generated as the sole products and there is no need to carry out oftentimes difficult purifications to separate natural products from their analogues.

ii) Modification of an enzyme involved in the biosynthesis of a natural product:

Mutagenesis can allow greater substrate flexibility, which can allow a broad and diverse range of analogues to be obtained. Alternatively, the protein can be exchanged with another protein. Using this approach, it is possible to achieve the introduction of completely new functionalities into the natural product. These new functionalities can in turn be used as 'handles' to allow further chemical modifications enabling a much more comprehensive library of analogues to be accessed.

5 i) Gene Disruption Approaches

Mutasynthesis has been successfully employed to enable access to a large series of rapamycins (Scheme 8). By deletion of the *rapK* gene, required for the biosynthesis of the 4*R*,5*R*-4,5-dihydroxycyclohex-1-enecarboxylic acid (DHCHC) precursor to the starter acid of rapamycin, a strain of *Streptomyces hygroscopicus* (MG2-10, BIOT-3016) is engineered that is incapable of making rapamycins unless supplemented with a suitable synthetic DHCHC replacements.[39] Using this approach, a number of cyclohexanecarboxylic and cycloheptanecarboxylic acid rapalogues were prepared. More recently, a very similar approach has been used in the preparation of tacrolimus analogues following a deletion of *fkbO* (a *rapK* homologue) from a producing organism *Streptomyces sp.* GT110507 by Maeng and co-workers.[40]

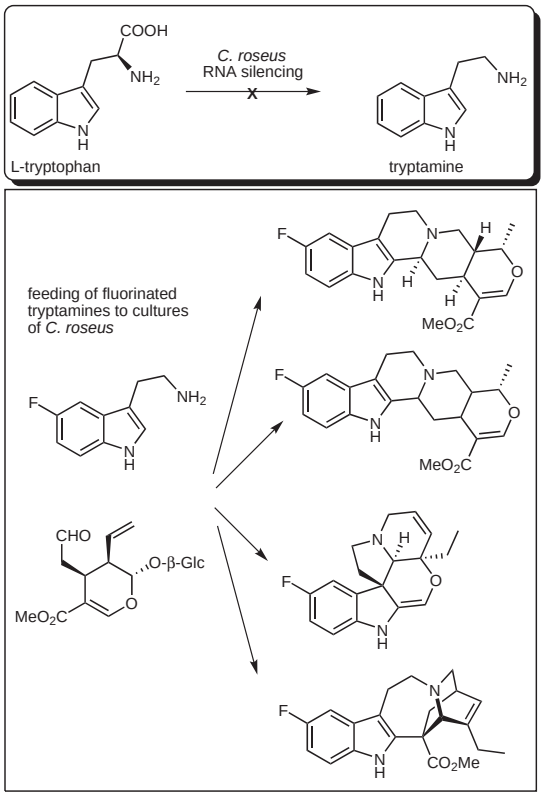


R = F, fluoro prerapamycin
(39-desmethoxy-39-fluoro-9-deoxo-16-*O*-desmethyl-27-desmethoxyrapamycin)

R = OH, prerapamycin

Scheme 8 Generation of rapamycin analogues by the deletion of *rapK*, responsible for the starter acid synthesis and feeding of analogues of DHCHC[39]

Other prominent examples of the gene disruption approach have been demonstrated by Micklefield and coworkers in the generation of analogues of calcium dependent antibiotic (CDA) by feeding 5-hydroxy- and 5-fluorotryptophans to an auxotrophic tryptophan *S. coelicolor*



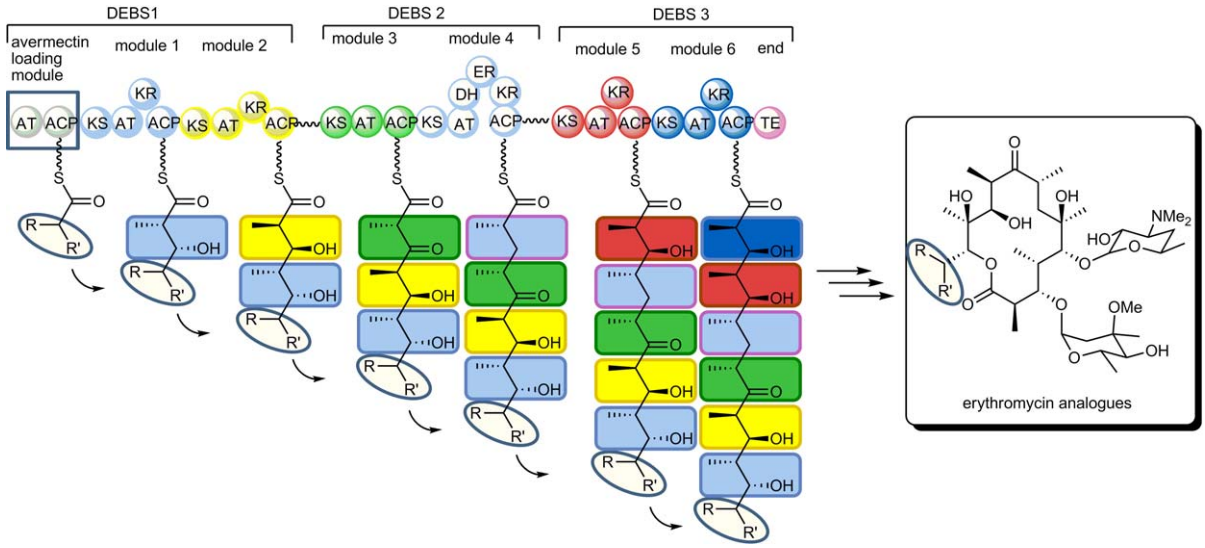
Scheme 9 BIO-CHEM-BIO approach to generate fluorinated monoterpene indole alkaloid analogues by feeding of fluorinated tryptamines to cultures of *C. roseus*[44]

strain.[41] Kirschning, Floss, and co-workers[42] elaborated on the work of Hong, Lee, and co-workers[43] to generate highly active geldanamycin derivatives by feeding 3-amino-5-hydroxybenzoic acid (AHBA) derivatives to a *napK* (AHBA synthase) deletion mutant.

The O'Connor group elaborated on their earlier work discussed above (Section 4), in which a tryptophan halogenase was introduced into the producing organism resulting in halogenated tryptophans being generated, which served as unnatural substrates in the biosynthesis of tryptamine. In their earlier work, they had sought to modify the biosynthesis of tryptamine in order to generate halogenated tryptamine analogues, which could serve as halogenated biosynthetic precursors to the monoterpene indole alkaloids. More recently, they have used RNA silencing of tryptophan decarboxylase to supress the biosynthesis of tryptamine in *C. roseus* altogether, opening up the potential of the biosynthetic pathway to incorporate synthetically-accessed fluorinated tryptamines to cultures of *C. roseus* (Scheme 9).[44] Through this approach, a number of fluorinated monoterpene indole alkaloids could be accessed. A similar approach was also adopted to generate a number of chlorinated and brominated analogues which could be further functionalized as part of a BIO-BIO-CHEM approach (Section 7).[45]

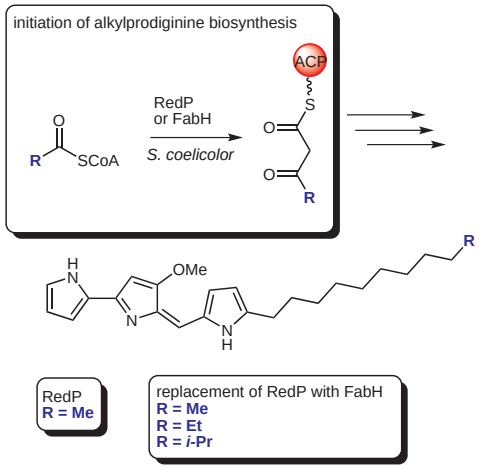
5 ii) Mutasynthesis by Enzyme Exchange or Modification

Pioneering work in the mutasynthesis of erythromycin analogues was carried out by the Staunton and Leadlay group. The erythromycin producing PKS [6-deoxyerythronolide B synthase (DEBS)] loading module in this case was exchanged with a wide-specificity loading module from the avermectin gene cluster of *S. avermitilis* (Scheme 10).[46] 2-Methylbutyric acid and isobutyric acid are the natural starter units for avermectin biosynthesis but a broad variety of over 40 alternative starter acids have been shown to also be tolerated by the loading module. The replacement strain of *S. erythrea* generated a diverse range of erythromycin analogues by supplementation of starter acid analogues.



Scheme 10 Introduction of the avermectin PKS loading module into the erythromycin gene cluster to enable greater chain starter unit substrate flexibility. Scheme adapted from Goss et al.[4]

A similar approach was employed in order to increase substrate flexibility in the biosynthesis of the prodiginines (Scheme 11). The enzyme, which is responsible for initiating the biosynthesis of the undecylpyrolle component of undecylprodiginine, RedP, a 3-Ketoacyl ACP synthase III initiation enzyme, was swapped with a homologue, FabH, allowing branched chain analogues to be produced.[47]



Scheme 11 Replacement of RedP in alkylprodiginine biosynthesis by FabH enables branched chain analogues to be produced by increasing the substrate flexibility in the initiation of alkylprodiginine biosynthesis[47]

Mupirocin, a clinically used antibiotic isolated from *Pseudomonas fluorescens* has a mode of action, which involves the inhibition of isoleucyl-transfer RNA synthetase in *Staphylococcus aureus* and *Streptococcus*. Simpson and co-workers successfully generated a number of mupirocin analogues, some with greater antibiotic activities by inactivating the polyketide synthase of *Pseudoalteromonas* spp SANK73390 and exogeneously feeding pseudomonic acids producing a series of epoxidated and nonepoxidated thiomarinol derivatives.[48]

Further work involved the supplementation of anhydroornithine and anhydrolysine to a non-ribosomal peptide synthetase deletion mutant, which yielded analogues with activity against methicillin-resistant *Staphylococcus aureus*.[48]

6 BIO-BIO: Combinatorial Biosynthesis

Combinatorial biosynthesis describes the ambition to be able to utilize enzymatic components from various different biosynthetic pathways in a combinatorial fashion, mixing and matching genetic elements encoding enzymes and creating designer biosynthetic pathways at will. Before the dream of combinatorial biosynthesis in its fullest sense can be realized, there is a long way to go in terms of understanding multifunctional enzyme structures, how these enzymes interact and passage intermediates, and how their substrate specificity of individual domains as well as entire systems may be reprogrammed. So far combinatorial biosynthesis has been used to achieve ring contraction[49–51] or ring expansion,[52] modification of the stereochemistry of moieties within polyketide synthases,[53] to alter glycosylation patterns,[54,55] and to generate new to nature halogenated metabolites,[56–58] with a recent example in this area being the impressive work on the enzymatic generation of fluoromalonyl-CoA and the processing of this substrate by the DEBSMod6+TE system, a step toward utilizing true combinatorial biosynthesis to incorporate fluorine into polyketide backbones.[59]

One of the earliest examples of the combining of gene clusters or the introduction of a gene with a desired function into another cluster in order to transform a natural product in a desired way was the engineering of a bacterial strain by Omura et al.[60] This engineered strain produced the novel polyketide antibiotic mederrhodin, a hybrid antibiotic produced by introducing the actinorhodin gene cluster heterologously into the medermycin producer (Figure 4).[60] As part of this work, a plasmid containing part of or the entire actinorhodin gene cluster was introduced into the medermycin producing *Streptomyces* sp. *strain AM7161* and ex-conjugants were screened for the production of novel secondary metabolites. Two new secondary metabolites were isolated, mederrhodin A and mederrhodin B. Mederrhodin A showed similar gram negative antibacterial activity to medermycin but mederrho-

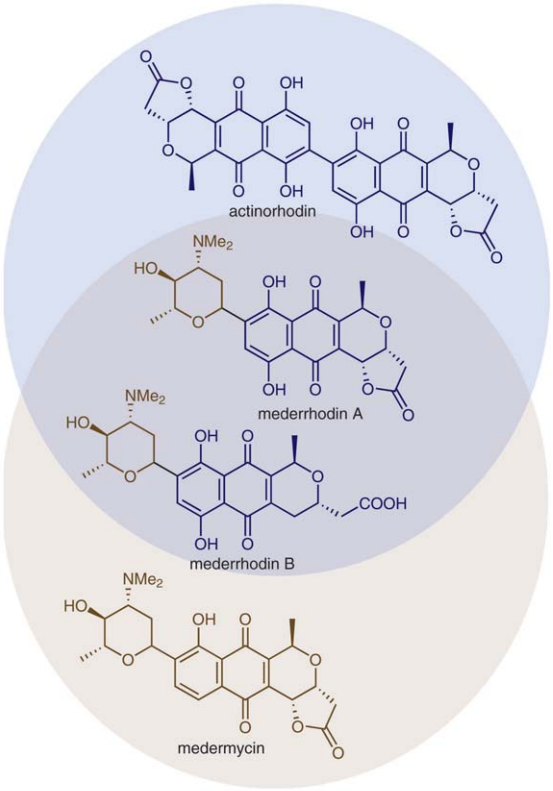


Figure 4 Introduction of the actinorhodin gene cluster into the medermycin producer in a combinatorial biosynthesis approach yields the new antibiotics mederrhodin A and mederrhodin B[60]

din B showed no activity against all gram positive and gram negative bacteria tested.

Daptomycin is a cyclolipopeptide used clinically for the treatment of skin infections caused by Gram-positive pathogens, bacteremia, and endocarditis. Baltz and co-workers have generated analogues of daptomycin by the exchange of single or multiple modules of the DptBC subunit of the NRPS and inactivation of a tailoring enzyme glutamic acid 3-methyltransferase from the producing organism *Streptomyces roseosporus*. Using this approach Baltz and co-workers were able to generate analogues with similar or better activities.[61]

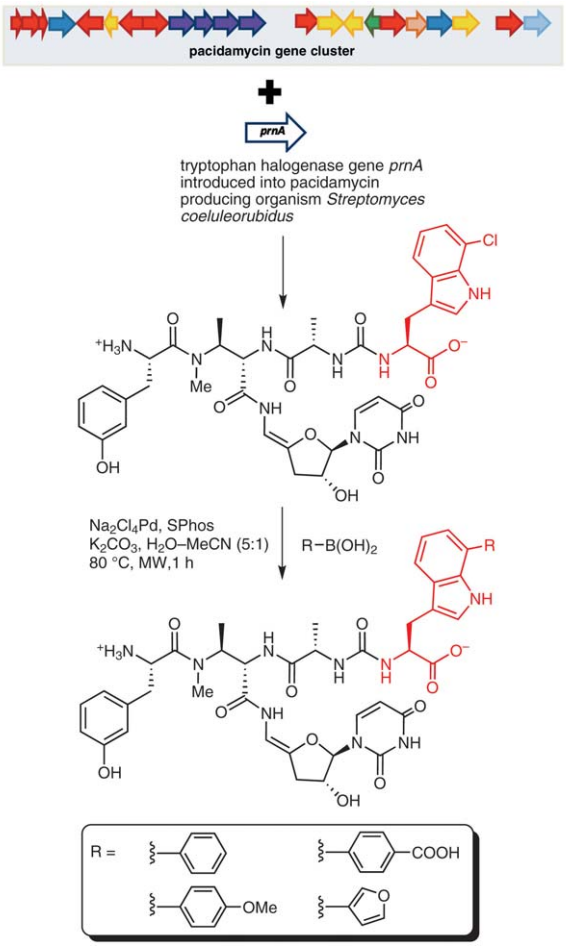
7 BIO-BIO-CHEM: Genochemetics: Gene Expression Enabling Synthetic Diversification

Genochemetics is a novel approach to the identification, investigation, and diversification of natural products. Described as 'gene expression enabling synthetic diversification,[62] genochemetics combines aspects of molecular biology and organic synthetic chemistry to construct a powerful toolkit for natural product research and analogue development.

Genochemetics can be considered as a union of combinatorial biosynthesis and semi-synthesis. In genochemetics, an enzyme from an unrelated biosynthetic pathway that is capable of introducing a reactive, chemically orthogonal, and selectively functionalizable handle into a natural product of interest is genetically introduced into the producing organism. The introduced enzyme acts in concert with the native biosynthetic machinery to generate a modified, unnatural product – sometimes alongside the unmodified natural product. The presence of the reactive and chemically orthogonal handle enables subsequent chemical derivatization to be performed selectively on the unnatural product, ideally as a component of the crude extracts and without need for prior purification or the use of protecting group chemistry.

To facilitate the genochemetic approach to natural product diversification, handles must be enzymatically installable and should enable facile and selective chemistry (such as cross-coupling or click chemistries). As a consequence of this first requirement, the handles must be found in an analogous structural context in a known (and ideally biosynthetically characterized) natural product.

The first example of genochemetic access to natural product analogues was demonstrated by the Goss group in 2010 (Scheme 12)[62] using the uridyl peptide antibiotic pacidamycin as the test bed.[1] The aim was to biosynthetically introduce a carbon–halogen bond, which could subsequently be modified using cross-coupling chemistries. The halotryptophan motif is found in several natural products, such as rebeccamycin and pyrrolnitrin (Figure 5), and tryptophan halogenases have been widely studied.[63] PrnA is responsible for the chlorination of tryptophan at the 7-position, which is the first step in pyrrolnitrin bio-



Scheme 12 Integration of the tryptophan halogenase gene *prnA* into the pacidamycin producer. *Streptomyces coeluleorubidus* gave access to halogenated pacidamycin analogues which could be further functionalized as part of a BIO-BIO-CHEM approach using Suzuki–Miyaura cross-coupling chemistries, in aqueous solvents and under mild conditions.[62,65]

synthesis.[64] The *prnA* gene was incorporated into the genome of the pacidamycin producer *Streptomyces coeruleorubidus* under the control of the constitutive promoter ermE*. This allowed production of the halogenase

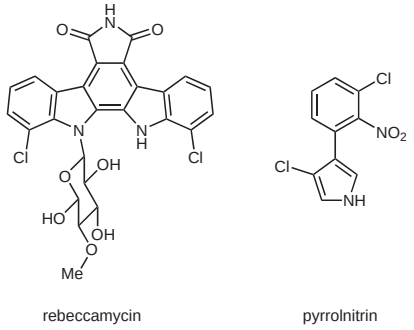


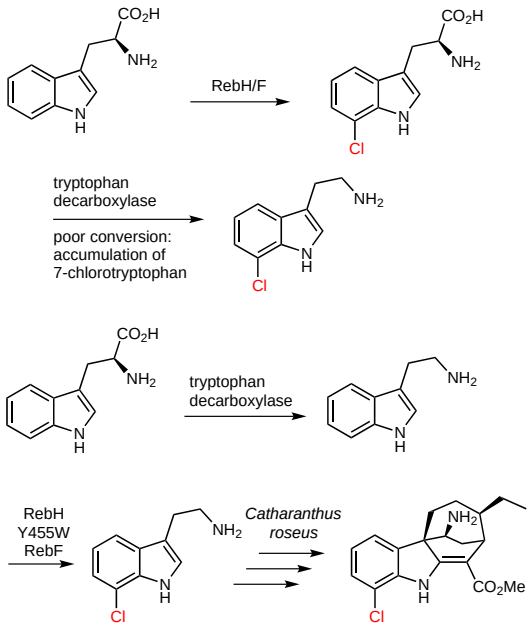
Figure 5 Antibiotics such as rebeccamycin and pyrrolnitrin contain halogens derived from the flavin-dependent tryptophan halogenases *rebH* and *prnA*. These enzymes can be introduced into other biosynthetic systems to generate halotryptophans in vivo.

alongside the pacidamycin biosynthetic machinery, and resulted in in vivo generation of chlorinated pacidamycin (Scheme 12).[62]

The group developed this example of combinatorial biosynthesis into a full genochemetic system (Scheme 12) by exploiting the selective reactivity of the newly introduced halogen through aqueous, palladium-catalyzed Suzuki–Miyaura cross-coupling chemistry under microwave irradiation, utilizing a sulfonated form of the SPhos ligand developed by Buchwald et al.[65]

The conversion could be performed on the crude extracts of fermentation cultures without need for purification beforehand. Using a series of 5 boronic acids, pacidamycin analogues were produced and identified by LC-MS/MS fragmentation patterns, in yields ranging from 33% to in excess of 95%. Two of these analogues, the phenyl- and the *p*-methoxyphenyl-substituted derivatives, were isolated in 52% and 67%, respectively.[62]

Recent work from the O'Connor group (Scheme 13) has presented the second example of genochemetic natural product analogue generation, this time in *C. roseus* as a development of their previous studies of the indole alkaloid metabolites of this plant (vide supra). The group employed cross-coupling chemistry similar to that previously used by the Goss group in order to functionalize halogenated metabolites generated both by precursor-directed biosynthesis and by combinatorial biosynthesis, with the latter giving the first eukaryotic genochemetic system.[45,62]



Scheme 13 Genochemetic natural product analogue generation resulting in the production of chlorinated indole alkaloid metabolites enabling post biosynthetic functionalization through cross-coupling.[45]

There are several examples of mutasynthesis in which the modification made to the natural product incorporates a chemically orthogonal handle that enables post-biosynthetic chemical modification. However, one of the key features of genochemetics is that the initial modification to the natural product is effected enzymatically and in vivo through modification of the biosynthetic pathway, unlike in mutasynthesis where it is still necessary to synthesize the modified biosynthetic precursor of interest.

8 Conclusions and Future Directions in Natural Product Analogue Generation

Natural products have historically been the leading molecules in the field of medicine; generated by biomolecules they are naturally predisposed to interact with biomolecules. With major recent advances in sequencing and analytical technologies fueling the increased rate of discovery of natural products these compounds are poised to become even more important to medicine.[66] Previously libraries of these oftentimes highly complex compounds, were considered intractable; by appropriately combining synthesis and biosynthesis it will become possible to access such libraries at will.

References

(1) Cragg, G. M.; Newman, D. J.; Snader, K. M. *J. Nat. Prod.* **1997**, *60*, 52.
(2) Rouhi, A. M. *Chem. Eng. News* **2003**, *81 (41)*, 77.
(3) Kirschning, A.; Hahn, F. *Angew. Chem. Int. Ed.* **2012**, *51*, 4012.
(4) Goss, R. J. M.; Shankar, S.; Fayad, A. A. *Nat. Prod. Rep.* **2012**, *29*, 870.
(5) (a) Nicolaou, K. C.; Sorensen, E. J. *Classics in Total Synthesis: Targets, Strategies, Methods*; Wiley: New York, **1996**. (b) Nicolaou, K. C.; Snyder, S. *Classics in Total Synthesis II: More Targets, Strategies, Methods*; Vol. 2; Wiley: New York, **2003**. (c) Nicolaou, K. C.; Chen, J. S. *Classics in Total Synthesis III*; Wiley: New York, **2011**.
(6) Vinothkumar, S.; Parameswaran, P. S. *Biotechnol. Adv.* **2013**, *31*, 1826.
(7) Towle, M. J.; Salvato, K. A.; Budrow, J.; Wels, B. F.; Kuznetsov, G.; Aalfs, K. K.; Welsh, S.; Zheng, W.; Seletsky, B. M.; Palme, M. H.; Habgood, G. J.; Singer, L. A.; Dipietro, L. V.; Wang, Y.; Chen, J. J.; Quincy, D. A.; Davis, A.; Yoshimatsu, K.; Kishi, Y.; Yu, M. J.; Littlefield, B. A. *Cancer Res.* **2001**, *61*, 1013.
(8) Yu, M. J.; Zheng, W.; Seletsky, B. M. *Nat. Prod. Rep.* **2013**, *30*, 1158.
(9) Elford, T. G.; Nave, S.; Sonawane, R. P.; Aggarwal, V. K. *J. Am. Chem. Soc.* **2011**, *133*, 16798.
(10) Nuhant, P.; Roush, W. R. *J. Am. Chem. Soc.* **2013**, *135*, 5340.
(11) Hu, D. X.; Clift, M. D.; Lazarski, K. E.; Thomson, R. J. *J. Am. Chem. Soc.* **2011**, *133*, 1799.
(12) Takada, A.; Hashimoto, Y.; Takikawa, H.; Hikita, K.; Suzuki, K. *Angew. Chem. Int. Ed.* **2011**, *50*, 2297.
(13) Custar, D. W.; Zabawa, T. P.; Hines, J.; Crews, C. M.; Scheidt, K. A. *J. Am. Chem. Soc.* **2009**, *131*, 12406.
(14) Florence, G. J.; Morris, J. C.; Murray, R. G.; Osler, J. D.; Reddy, V. R.; Smith, T. K. *Org. Lett.* **2011**, *13*, 514.
(15) Smith, T. E.; Fink, S. J.; Levine, Z. G.; McClelland, K. A.; Zackheim, A. A.; Daub, M. E. *Org. Lett.* **2012**, *14*, 1452.

(16) Nicolaou, K. C.; Seo, J. H.; Nakamura, T.; Aversa, R. J. *J. Am. Chem. Soc.* **2011**, *133*, 214.
(17) Evans, P. A.; Huang, M.; Lawler, M. J.; Maroto, S. *Nat. Chem.* **2012**, *4*, 680.
(18) Sheehan, J. C. *Ann. N. Y. Acad. Sci.* **1967**, *145*, 216.
(19) Kardos, N.; Demain, A. L. *Appl. Microbiol. Biotechnol.* **2011**, *92*, 677.
(20) Courvalin, P. *Clin. Infect. Dis.* **2006**, *42*, S25.
(21) Nakama, Y.; Yoshida, O.; Yoda, M.; Araki, K.; Sawada, Y.; Nakamura, J.; Xu, S.; Miura, K.; Maki, H.; Arimoto, H. *J. Med. Chem.* **2010**, *53*, 2528.
(22) Xie, J.; Pierce, J. G.; James, R. C.; Okano, A.; Boger, D. L. *J. Am. Chem. Soc.* **2011**, *133*, 13946.
(23) Pinchman, J. R.; Boger, D. L. *J. Med. Chem.* **2013**, *56*, 4116.
(24) Pathak, T. P.; Miller, S. J. *J. Am. Chem. Soc.* **2012**, *134*, 6120.
(25) Pathak, T. P.; Miller, S. J. *J. Am. Chem. Soc.* **2013**, *135*, 8415.
(26) Barrett, K. T.; Miller, S. J. *J. Am. Chem. Soc.* **2013**, *135*, 2963.
(27) Lichtor, P. A.; Miller, S. J. *Nat. Chem.* **2012**, *4*, 990.
(28) Cuevas, C.; Pérez, M.; Martín, M. J.; Chicharro, J. L.; Fernández-Rivas, C.; Flores, M.; Francesch, P.; Gallego, M.; Zarzuelo, F.; de la Calle, J.; García, C.; Polanco, I.; Rodríguez, I.; Manzanares, A. *Org. Lett.* **2000**, *2*, 2545.
(29) Kawagishi, F.; Toma, T.; Inui, T.; Yokoshima Fukuyama, T. *J. Am. Chem. Soc.* **2013**, *135*, 13684.
(30) Chen, J.; Chen, X.; Bois-Choussy, Zhu. J. *J. Am. Chem. Soc.* **2006**, *128*, 87.
(31) Clark, B. R.; O'Connor, S. E.; Fox, D.; Leroy, J.; Murphy, C. D. *Org. Biomol. Chem.* **2011**, *9*, 6306.
(32) Clark, B. R.; Murphy, C. D. *Org. Biomol. Chem.* **2009**, *7*, 111.
(33) Pan, E.; Oswald, N. W.; Legako, A. G.; Life, J. M.; Posner, B. A.; MacMillan, J. B. *Chem. Sci.* **2013**, *4*, 482.
(34) Grüschow, S.; Rackham, E. J.; Elkins, B.; Newill, P. L. A.; Hill, L. M.; Goss, R. J. M. *ChemBioChem* **2009**, *10*, 355.
(35) Ragab, A. E.; Grüschow, S.; Rackham, E. J.; Goss, R. J. M. *Org. Biomol. Chem.* **2010**, *8*, 3128.
(36) Glenn, W. S.; Nims, E.; O'Connor, S. E. *J. Am. Chem. Soc.* **2011**, *133*, 19346.
(37) Lussier, F.; Colatriano, D.; Wiltshire, Z.; Page, J. E.; Martin, V. J. J. *Comput. Struct. Biotechnol. J.* **2012**, *3*, e201210020.
(38) Lee, H. Y.; Harvey, C. J. B.; Cane, D. E.; Khosla, C. *J. Antibiot.* **2011**, *64*, 59.
(39) Goss, R. J. M.; Lanceron, S.; Deb Roy, A.; Sprague, S.; Nure-Alam, M.; Hughes, D. L.; Wilkinson, B.; Moss, S. J. *ChemBioChem* **2010**, *11*, 698.
(40) Kim, D. H.; Ryu, J. H.; Lee, K. S.; Lee, B. M.; Lee, M. O.; Lim, S.; Maeng, P. J. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 5881.
(41) Amir-Heidari, B.; Thirlway, J.; Micklefield, J. *Org. Biomol. Chem.* **2008**, *6*, 975.
(42) Eichner, S.; Floss, H. G.; Sasse, F.; Kirschning, A. *ChemBioChem* **2009**, *10*, 1801.
(43) Kim, W.; Lee, J. S.; Lee, D.; Cai, X. F.; Shin, J. C.; Lee, K.; Lee, C.; Ryu, S.; Paik, S.; Lee, J. J.; Hong, Y. *ChemBioChem* **2007**, *8*, 1491.
(44) Runguphan, W.; Maresh, J. J.; O'Connor, S. E. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13673.
(45) Runguphan, W.; O'Connor, S. E. *Org. Lett.* **2013**, *15*, 2850.
(46) Squire, C. M.; Goss, R. J. M.; Hong, H.; Leadlay, P. F.; Staunton, J. *ChemBioChem* **2003**, *4*, 1225.
(47) Mo, S.; Kim, B. S.; Reynolds, K. A. *Chem. Biol.* **2005**, *12*, 191.
(48) Murphy, A. C.; Fukuda, D.; Song, Z.; Hothersall, J.; Cox, R. J.; Willis, C. L.; Thomas, C. M.; Simpson, T. J. *Angew. Chem. Int. Ed.* **2011**, *50*, 3271.
(49) Thomas, I.; Martin, C. J.; Wilkinson, C. J.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **2002**, *9*, 781.
(50) Wu, N.; Cane, D. E.; Khosla, C. *Biochemistry* **2002**, *41*, 5056.
(51) Wu, N.; Kudo, F.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **2000**, *122*, 4847.
(52) Kao, C. M.; McPherson, M.; McDaniel, R. N.; Fu, H.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1997**, *119*, 11339.
(53) Caffrey, P. *Chem. Biol.* **2005**, *12*, 1060.
(54) Shepherd, M. D.; Liu, T.; Mendez, C.; Salas, J. A.; Rohr, J. *Appl. Environ. Microbiol.* **2011**, *77*, 435.
(55) Pérez, M.; Baig, I.; Braña, A. F.; Salas, J. A.; Rohr, J.; Méndez, C. *ChemBioChem* **2008**, *9*, 2295.
(56) Heide, L.; Westrich, L.; Anderle, C.; Gust, B.; Kammerer, B.; Piel, J. *ChemBioChem* **2008**, *9*, 1992.
(57) O'Connor, S. E. *Meth. Enzymol.* **2012**, *515*, 189.
(58) Roy, A. D.; Grüschow, S.; Cairns, N.; Goss, R. J. M. *J. Am. Chem. Soc.* **2010**, *132*, 12243.
(59) Walker, M. C.; Thuronyi, B. W.; Charkoudian, L. K.; Lowry, B.; Khosla, C.; Chang, M. C. Y. *Science* **2013**, *341*, 1089.
(60) Omura, S.; Ikeda, H.; Malpartida, F.; Kieser, H. M.; Hopwood, D. A. *Antimicrob. Agents Chemother.* **1986**, *29*, 13.
(61) Nguyen, K. T.; Ritz, D.; Gu, J. Q.; Alexander, D.; Chu, M.; Miao, V.; Brian, P.; Baltz, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17462.
(62) Roy, A. D.; Grüschow, S.; Cairns, N.; Goss, R. J. M. *J. Am. Chem. Soc.* **2010**, *132*, 12243.
(63) Smith, D. R.; Grüschow, S.; Goss, R. J. *Curr. Opin. Chem. Biol.* **2013**, *17*, 276.
(64) Dong, C.; Flecks, S.; Unversucht, S.; Haupt, C.; van Pee, K. H.; Naismith, J. H. *Science* **2005**, *309*, 2216.
(65) Barder, T. E.; Walker, S. D.; Martinelli, J. R.; Buchwald, S. L. *J. Am. Chem. Soc.* **2005**, *127*, 4685.
(66) Osbourn, A.; Goss, R. J. M.; Carter, G. T. *Natural Products: Discourse, Diversity, and Design*; Wiley: New York, **2014**.

