



Universiteit  
Leiden  
The Netherlands

## **Discovery of novel antibiotics from actinomycetes by integrated metabolomics & genomics approaches**

Wu, Changsheng

### **Citation**

Wu, C. (2016, October 26). *Discovery of novel antibiotics from actinomycetes by integrated metabolomics & genomics approaches*. Retrieved from <https://hdl.handle.net/1887/43768>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/43768>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/43768> holds various files of this Leiden University dissertation

**Author:** Wu, Shangsheng

**Title:** Discovery of novel antibiotics from actinomycetes by integrated metabolomics & genomics approaches

**Issue Date:** 2016-10-26

## Chapter 3

### Metabolic profiling as a tool for prioritizing antimicrobial compounds

Changsheng Wu <sup>1,2</sup>, Young Hae Choi <sup>2</sup>, and Gilles P. van Wezel <sup>1,3</sup>

<sup>1</sup> *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands;*

<sup>2</sup> *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands;*

<sup>3</sup> *Department of Microbial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, The Netherlands;*

*J. Ind. Microbiol. Biotechnol.* **2016**, 43, 299–312

**ABSTRACT**

Metabolomics is an analytical technique that allows scientists to globally profile low molecular weight metabolites between samples in a medium- or high-throughput environment. Different biological samples are statistically analysed and correlated to a bioactivity of interest, highlighting differentially produced compounds as potential biomarkers. Here, we review NMR- and MS-based metabolomics as technologies to facilitate the identification of novel antimicrobial natural products from microbial sources. Approaches to elicit the production of poorly expressed (cryptic) molecules are thereby key to allow statistical analysis of samples to identify bioactive markers, while connection of compounds to their biosynthetic gene cluster is a determining step in elucidating the biosynthetic pathway and allows downstream process optimization and upscaling. The review focuses on approaches built around NMR-based metabolomics, which enables efficient dereplication and guided fractionation of (antimicrobial) compounds.

## 1. INTRODUCTION

The discovery and development of antibiotics to treat bacterial infections is one of the greatest triumphs of modern medicine. However, the exponential increase of antimicrobial resistance means that bacterial infections now once more pose a major threat to human health.<sup>1</sup> The high frequency of re-discovery of known molecules thereby frustrates screening efforts, and this necessitates new approaches to bolster the antibiotic pipelines.<sup>2-5</sup> Filamentous fungi and bacteria of the order of *Actinomycetales* are the major producers of biomedical and agricultural natural products, and these microorganisms are responsible for producing the vast majority of the known antibiotics.<sup>6, 7</sup> Some two thirds of all antibiotics are produced by actinomycetes, the majority of which are sourced by members of the genus *Streptomyces*. However, this is likely to represent only a tiny portion of the repertoire of total chemical space of bioactive compounds that these microorganisms may produce.<sup>6, 7</sup> Firstly, while the number of filamentous microbes in soil and marine environments is enormous, the bulk of them resist cultivation in the laboratory and thus escape screening and exploitation.<sup>8</sup> Recent work showed that enabling growth of these ‘uncultivable’ microorganisms might open up a new area of the chemical space of natural products.<sup>9, 10</sup> Secondly, many natural product biosynthetic gene clusters (BGCs) of cultivable microbes remain silent during standard laboratory cultivation.<sup>11, 12</sup> As an example, genome sequencing of actinomycetes revealed that the producing capacity of even the best-studied model organisms has been grossly underestimated.<sup>13-16</sup> And thirdly, those gene clusters that are expressed, in particular if expression is relatively low as compared to other bioactive compounds produced by the same organism, often escape discovery due to the very demanding chemical analysis or chromatographic purification that is required.<sup>17</sup> New approaches are therefore required to exploit the huge unexplored reservoir of bioactive natural products and rejuvenate the drug-discovery pipelines.

The preferred method of identifying compounds in complex mixtures is metabolomics, which is a high-throughput analytical technique that offers a global analysis of the products of all cellular metabolic reactions, regardless of the reactions that lead to their production.<sup>18</sup> Spectroscopic techniques, and in particular nuclear magnetic resonance (NMR) and mass spectrometry (MS), are applied to chemically profile the metabolites that are produced by microorganisms. In connection with a wide range of chemometric methods, the metabolic differences among experimental groups are subsequently unraveled by comparison.<sup>19, 20</sup> In this review, we discuss ways to enforce fluctuations in the production of bioactive compounds, and the application of metabolomics to facilitate uncovering those compounds, whereby we zoom in on NMR-based metabolic profiling methods, which are less commonly applied in microbial drug discovery.

## 2. APPROACHES TO ELICIT ANTIBIOTIC PRODUCTION

Genome mining of actinomycetes on average reveals the presence of 30-40 secondary metabolite biosynthetic gene clusters (BGCs) per species.<sup>21</sup> Only a small number of these gene clusters has been matched to a natural product. A major problem is that many BGCs remain silent or poorly expressed under routine screening conditions. If we are to even start attempting to visualize these compounds by metabolomics approaches, their activation is the first essential step. Generic approaches to activate silent antibiotics to uncover the hidden chemical diversity include: (i) perturbed growth media with additives,<sup>22, 23</sup> (ii) inducing

antibiotic resistance, *e.g.* to streptomycin<sup>24</sup> and rifampicin,<sup>25-27</sup> (iii) the addition of chemical elicitors, such as *N*-acetylglucosamine,<sup>28</sup>  $\gamma$ -butyrolactones,<sup>29, 30</sup> or histone deacetylase inhibitors;<sup>31</sup> and (iv) microbial cocultivation.<sup>32-34</sup> Instead of inducing the production of an antibiotic, an interesting alternative may be offered by screening for inducible resistance, in particular associated with classes of antibiotics that are less frequently identified in classical screens. Growth on media containing vancomycin was used to enrich a library of actinomycetes for producers of glycopeptide antibiotics, and subsequent phylogeny-based screening according to specific fingerprints of biosynthetic genes, lead to the identification of the novel glycopeptide antibiotic pekiskomycin.<sup>35</sup>

In terms of the activation of antibiotic production, a concept study was recently done using a collection of some 800 strains that were grown under 40 different growth conditions and then assayed for changes in the production of antimicrobial activity against the so-called ESKAPE pathogens.<sup>23</sup> This showed that many strains that fail to produce antibiotics under routine growth conditions, produce bioactivities that inhibits the growth of MDR pathogens under specific nutritional conditions or after addition of elicitors.<sup>23</sup>

## 2.1. Harnessing the regulatory mechanisms that control antibiotic biosynthesis

To develop new directed approaches, it is imperative that we better understand the underlying regulatory networks. The regulation of antibiotic production involves multiple regulatory cascades and networks. Here, we will guide the reader with some general principles gleaned from the model organism *Streptomyces coelicolor*; for more extensive reviews on the control of antibiotic production we refer the reader elsewhere.<sup>36-39</sup>

Antibiotic BGCs specify biosynthesis and modification, resistance, transport and pathway-specific control. The best-studied example of cluster-encoded, pathway-specific activators is that of the *Streptomyces* antibiotic regulatory proteins (SARPs). SARPs typically bind to direct repeats in promoter regions of biosynthetic genes, thereby activating transcription of some or all of the antibiotic biosynthetic genes in the cluster.<sup>40</sup> In *S. coelicolor*, the pathway-specific regulatory genes *actII-ORF4*, *cdaR* and *redD* encode the SARPs for actinorhodin (ACT), calcium-dependent antibiotic (CDA) and undecylprodigiosin (RED) production, respectively. Of these, *actII-ORF4* and *redD* have been extensively studied. Deletion of these genes abolishes the production of actinorhodin or undecylprodigiosin, respectively,<sup>41</sup> and their transcription is activated in a growth phase-dependent manner.<sup>42, 43</sup> Interestingly, positioning *redD* behind a developmental or nitrogen-responsive promoter is sufficient to produce undecylprodigiosin in aerial hyphae or bring it under nitrogen control, respectively, which suggests a lack of downstream control.<sup>44</sup> Hence, one effective approach to activate gene clusters is overexpression of their activator genes, but this involves genetic manipulation which is not amenable to high-throughput approaches.

Recent years have underlined the importance of global regulators as a higher order regulatory network, and their possible application for the activation of antibiotic production. Many pleiotropic regulators characterized so far are required only under specific environmental conditions.<sup>45, 46</sup> The nutrient sensory GntR-family regulator DasR controls among others aminosugar metabolism and transport, the chitinolytic system and antibiotic production.<sup>28, 47-51</sup> DasR is a highly pleiotropic regulator, as demonstrated by recent systems biology analysis of chitin- or *N*-acetylglucosamine-induced cultures of *S. coelicolor*.<sup>52, 53</sup>

DasR directly controls the transcription of *actII-ORF4* and *redD* as well as *redZ*, a response regulator required for the activation of *redD* and thus of undecylprodigiosin production.<sup>28</sup> Systems-wide DNA binding experiments using ChIP-chip analysis revealed that in *S. coelicolor* DasR likely controls all pathway-specific activator genes,<sup>53</sup> but this appears to be more an exception than a rule. The activity of DasR is modulated by metabolic derivatives of *N*-acetylglucosamine, and addition of this amino sugar to the culture media activates antibiotic production in several actinomycetes, including cryptic antibiotics.<sup>28</sup>

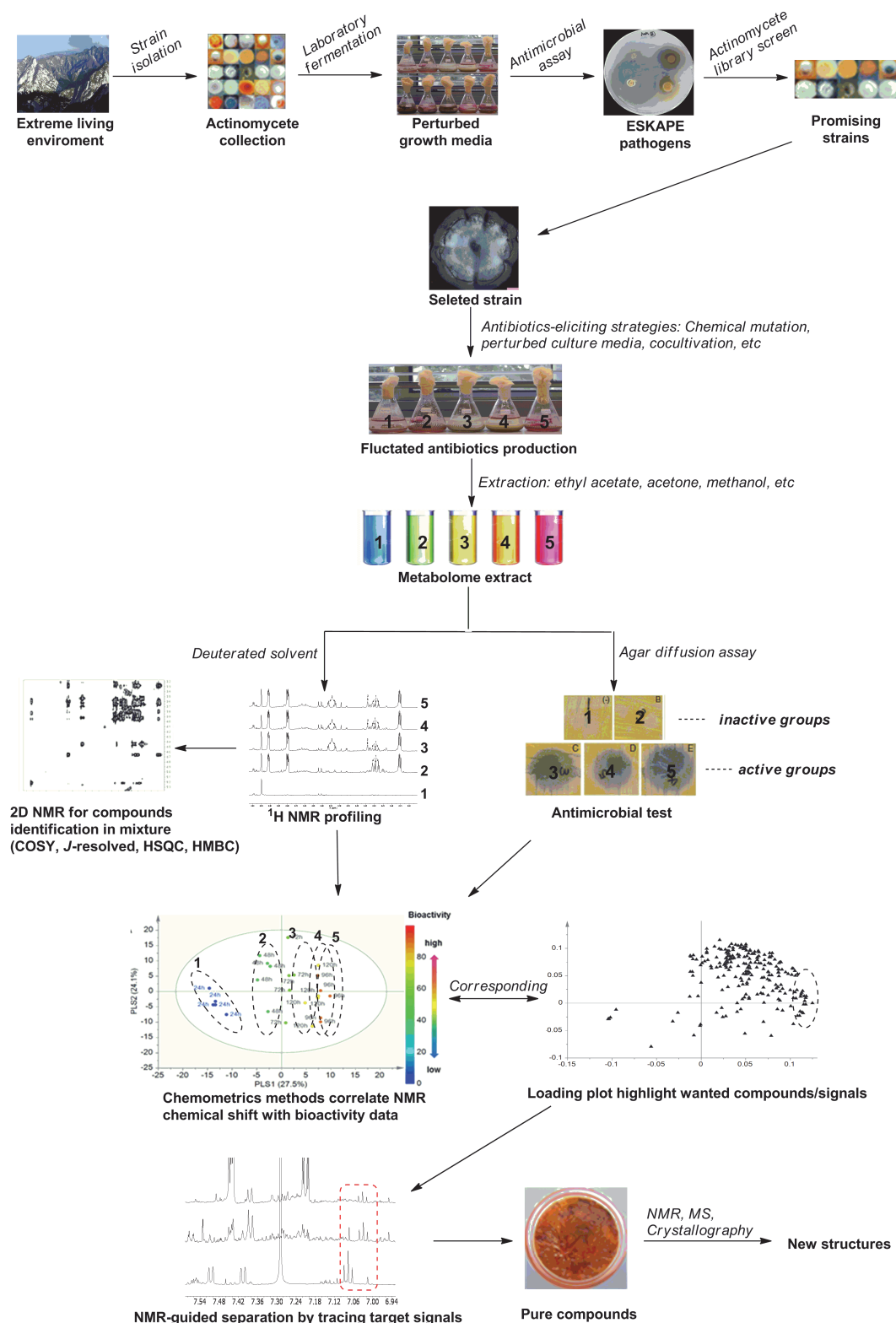
Interesting cross-talk is seen for DasR with other higher order antibiotic regulators. One is the TetR-family regulatory protein AtrA that in *S. coelicolor* is required for the transcription of *actII-ORF4*<sup>54</sup> and in *S. griseus* of *strR*, the pathway-specific activator gene of streptomycin production.<sup>55</sup> The antibiotic activating role of AtrA may be widespread, as suggested by recent data that AtrA also controls biosynthesis of the important lipoglycopeptide antibiotic daptomycin in *Streptomyces roseosporus*.<sup>56</sup> AtrA is controlled by the level of phosphate and repressed by the PhoRP system.<sup>57</sup> AtrA activates development<sup>58</sup> as well as *N*-acetylglucosamine import,<sup>59</sup> and thus antagonizes the repressing activities of DasR, whereby the metabolic balance likely plays a deciding role on the net outcome. Further complexity is offered by Rok7B7, a member of the ROK family of proteins, which are predominantly sugar regulatory proteins and sugar kinases, including glucose kinase.<sup>60, 61</sup> Rok7B7 also pleiotropically affects primary and secondary metabolism, and is like AtrA required for actinorhodin production.<sup>62, 63</sup> Recent evidence suggests that Rok7B7 may be activated by a derivative of the C5 sugar xylose.<sup>63</sup>

Other pleiotropic antibiotic regulators involved in the control of *actII-ORF4* are the large SARP-family regulator AfsR and the phosphate regulator PhoP. AfsR derives its name from a putative relationship to synthesis of the hormone-like signaling molecule A-factor, belonging to the  $\gamma$ -butyrolactones. The A-factor-responsive AdpA controls the onset of development and antibiotic production in *S. griseus*,<sup>64, 65</sup> but the precise relationship with AfsR (if any) so far remains unresolved. In *S. coelicolor*, AfsR is conditionally required for Act and Red production<sup>41</sup> and its control is somehow transmitted through control of the small downstream gene *afsS*, thereby activating its transcription.<sup>66</sup> While the precise function of *afsS* is unclear, similarly to AtrA, its pleiotropic and positive effect on antibiotic production makes it an attractive target for the activation of antibiotic production. PhoP represses actinorhodin production in response to phosphate,<sup>67</sup> and it may also relay its control via *afsS*.<sup>57</sup> These are examples to highlight the complexity of the control of antibiotic production, whereby an astonishing number of around 15 regulatory proteins have been shown to control *actII-ORF4* alone.<sup>46</sup> Improved understanding of these multiple and intertwined regulatory networks will allow scientists to design new approaches to activate poorly expressed BGCs.

## 2.2. Connecting eliciting approaches to metabolomics-driven lead discovery

In the framework of this review, the activation of gene expression is important to activate silent gene clusters and particularly to achieve significant fluctuation of the production of (cryptic) antibiotics and thus differential bioactivity between cultures; this serves as enabling technology for statistical correlation between a bioactivity of interest and the responsible (sought-after) compounds. Having established such differential expression, the next challenge is to efficiently identify the induced antimicrobial compounds against a background of an

inevitably highly complex metabolome matrix.



**Figure 1. A pipeline for the discovery of antibiotics produced by actinomycetes based on NMR-based metabolomics.** Strains are subjected to eliciting approaches to activate the production of poorly expressed antimicrobials, followed by NMR-based metabolomics to identify the bioactive molecules. <sup>1</sup>H NMR profiling of active and inactive groups, aided by 2D NMR, allows dereplication of known molecules and chemometric methods then correlate proton signals in the active fractions to the bioactive compound(s).



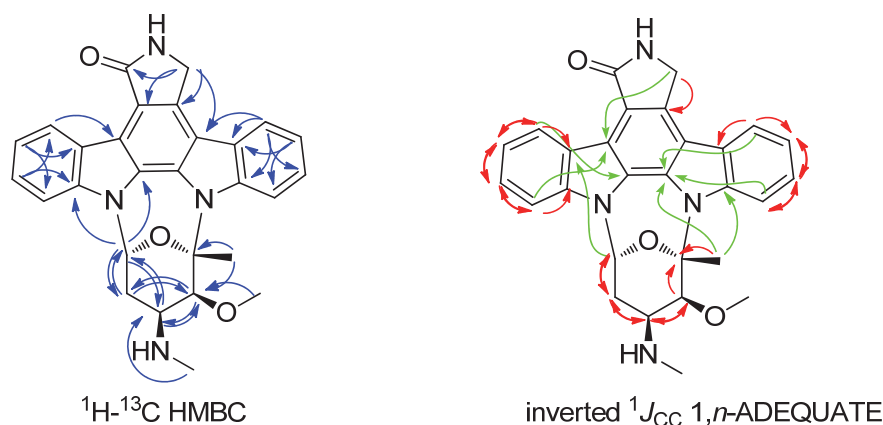
Metabolomics is an effective tool for facilitating the discovery of new antibiotics, because it allows the multivariate comparison of active and inactive metabolic samples, highlighting particularly the differentially produced compounds which serve as potential biomarkers, thereby avoiding chemical redundancy in the very early stage.<sup>68-70</sup> Integrating metabolomics approaches with new eliciting strategies should allow scientists to streamline their drug-discovery pipeline, so as to prioritize novel molecules over repeated (and frustrating) isolation of known molecules. A schematic representation of a feasible antibiotics-discovery pipeline based on NMR-based metabolomics is presented in Figure 1. Compounds can be readily identified in complex biological matrices without time-consuming chromatographic separation, further aided by 2D NMR experiments. In the next section, we detail some of the NMR techniques that are most appropriate for metabolomics approaches.

### 3. METABOLOMICS OF NATURAL PRODUCTS BASED ON NMR SPECTROSCOPY

NMR techniques that are most frequently used in NMR metabolomics are  $^1J$ -resolved, COSY, TOCSY, HSQC, and HMBC.<sup>70-73</sup>  $^1\text{H}$  NMR signals corresponding to molecules that exist in both the active and the inactive groups are discarded based on lack of statistical relevance.<sup>74</sup> Chemometric analysis is then used for further data mining of the  $^1\text{H}$  NMR data, extracting information that is intractable to visual detection. In principle, the methodology of NMR measurements, spectral data preprocessing, and multivariate data analysis basically follow a protocol that was originally designed for analysis of plant metabolites.<sup>71</sup> The  $^1\text{H}$  NMR signals correlated with the better bioactivity and separation of the active from the inactive groups are used as probing tools for NMR-guided isolation of the sought-after bioactivity. Advantages of NMR-based over MS-based (LC-MS or GC-MS) metabolomics are: (i) the chemical shifts and splitting patterns of proton resonance data provide partial structural information that allows early assessment of the type of compound causing the bioactivity, thus aiding in avoiding redundancy; (ii) the analysis of crude mixtures can be done in very short time (in the order of minutes), with high signal robustness (*i.e.* minimum technical variation), so as to guide scientists in a medium- to high-throughput environment. These advantages should be weighed against the lack of true high-throughput application which is a key selling point of MS-based approaches. As soon as the major interfering signals are removed from the NMR spectra, the material (even if not yet pure) is likely of sufficient quality for full structure elucidation. In some cases, structural elucidation of unknown compounds is challenging, especially for those entities with unprecedented molecular scaffold. Two-dimensional NMR (2D NMR) techniques have formed the foundation of most contemporary approaches for structure elucidation,<sup>75</sup> allowing assembly of the molecular architecture by defining atom-to-atom connectivity. In general, NMR-based metabolomics uses a combination of 2D NMR techniques for deconvoluting metabolites in congested mixtures, namely  $J$ -resolved,  $^1\text{H}$ -based 2D (COSY and TOCSY),  $^{13}\text{C}$ -based 2D NMR (HSQC and HMBC).<sup>75</sup>  $J$ -resolved solves the signal purity of  $^1\text{H}$  NMR by showing splitting pattern with coupling constants in F2 axis which requires for chemical structures of corresponding signals. COSY and TOCSY give information of  $^1\text{H}$  connectivity within the same molecule. Along with the connectivity, the  $^1\text{H}$ -based COSY and TOCSY also help signal identification in overlapped regions. Many signals are overlapped which obscures some of the peaks in the mixture, but these can be deconvoluted by the correlations in the 2D NMR spectra.  $^{13}\text{C}$ -related spectra such as HSQC

and HMBC give more detailed structural information, particularly for the carbon skeleton. Although important for identifying the carbon skeleton of organic compounds,  $^{13}\text{C}$ -NMR is rarely applied in NMR-based metabolomics because of the low abundance of  $^{13}\text{C}$  in nature and its long relaxation time, and also because  $^{13}\text{C}$  broad-band decoupling cannot be used quantitatively. HSQC and HMBC are therefore primarily applied to support the  $^1\text{H}$  NMR data.

Deciphering the structure of organic molecules by using conventional routine suites, based on for example  $J$ -resolved-COSY (or TOCSY), HSQC and HMBC<sup>76</sup> often gives rise to misassignment. This partly arises from the inherent challenges of differentiating two-bond heteronuclear ( $\text{H} \rightarrow \text{C}$ ,  $^2J_{\text{CH}}$ ) correlations from three-bond correlations ( $^3J_{\text{CH}}$ ) in the HMBC spectrum. In addition, as HMBC normally just provides  $^2J_{\text{CH}}$  and  $^3J_{\text{CH}}$  but rarely  $^4J_{\text{CH}}$  correlations, structure elucidation becomes progressively challenging the more molecules become proton-deficient.<sup>77</sup> To address these problems, many other approaches have been tried such as combined HSQC-TOCSY.<sup>78</sup> Researchers from Merck proposed that adequate sensitivity double-quantum (1,  $n$ -ADEQUATE) NMR spectroscopy should be reconsidered as a complement for HMBC in structure elucidation.<sup>79</sup> 1,1-ADEQUATE exclusively presents  $^1J_{\text{CC}}$  correlations and can help distinguish  $^2J_{\text{CH}}$  from  $^3J_{\text{CH}}$  correlations in the HMBC spectrum, which would be a powerful method for determining the proton-rich small molecules by using a combination of 1,1-ADEQUATE + HMBC.<sup>80, 81</sup> 1,  $n$ -ADEQUATE experiments predominantly yield homonuclear  $^3J_{\text{CC}}$  correlations (equal to  $^4J_{\text{CH}}$ ) that are difficult to obtain with HMBC NMR. This feature makes 1, $n$ -ADEQUATE a complement for HMBC in defining proton-deficient structures, such as a large group of aromatic polyketides produced by actinomycetes.<sup>80, 82</sup> Furthermore, newly developed inverted  $^1J_{\text{CC}}$  1, $n$ -ADEQUATE spectroscopy allows the integration of 1,1-ADEQUATE and 1, $n$ -ADEQUATE information in a single spectrum, while they remain distinguished by plotting the output in different colors (Figure 2).<sup>83</sup>



**Figure 2. Diagram of HMBC (  $\rightarrow$  ), and inverted  $^1J_{\text{CC}}$  1, $n$ -ADEQUATE correlations by using the proton-deficient alkaloid staurosporine as an example.** The skeleton of this model compound can be unequivocally assigned by HMBC + inverted  $^1J_{\text{CC}}$  1, $n$ -ADEQUATE strategy. In inverted  $^1J_{\text{CC}}$  1, $n$ -ADEQUATE,  $^1J_{\text{CC}}$  correlations are shown as  $\rightarrow$ , while important  $^3J_{\text{CC}}$  correlations are shown as  $\rightarrow$ .

Though NMR spectroscopy is undoubtedly powerful in the structure elucidation, it works better on pure organic compounds, as otherwise signal-overlapping issues come into play. To address this, chromatography-hyphenated NMR (HPLC-NMR) has been introduced into metabolic analysis.<sup>84</sup> The major advantages of on-flow HPLC-NMR lie in the real-time

access to structure information of chromatographic peaks by using NMR spectrometers as detector.<sup>85</sup> To balance the inherent lower sensitivity of NMR spectra and the limited injection volume onto the HPLC unit, solid-phase-extraction (SPE) is coupled to trap the analytes onto an identical SPE cartridge when repeated HPLC analysis, followed by final elution with a deuterated solvent into miniaturization tube for NMR analysis.<sup>86</sup> By using this LC-SPE-NMR approach,<sup>87, 88</sup> comprehensive structural information about chemical composition of crude extracts can be obtained online, thus enabling early and prospective assessment of the valuable components of an extract or fraction.<sup>89</sup> Furthermore, when integrating MS into the system, like LC-DAD-SPE-MS-NMR instrumentation,<sup>90, 91</sup> the architecture of novel compounds can be elucidated *de novo* and online without laborious purification of the analyte(s) from its highly complex matrix.

#### 4. CONNECTING GENOTYPE TO CHEMOTYPE: INTEGRATING GENOMICS WITH METABOLOMICS

Before looking at applications of metabolomics to identify novel compounds, in this section we will look into how metabolomics may aid in connecting a natural product of interest to its BGC. This is a crucial step as it facilitates molecular approaches such as directed mutagenesis, overexpression of pathway-specific regulators and heterologous expression. Furthermore, identification of the gene cluster is often required for determining the precise biosynthetic pathway.

Next-generation sequencing (NGS) technologies have uncovered the genetic architecture of thousands of BGCs, and this has revolutionized the drug-discovery approaches.<sup>92</sup> Exploiting this rapidly increasing source of information, a recent network analysis of the phylogeny and distribution of BGCs in microbial genomes revealed a vast number of yet underexplored genetic resources, belonging to many hundreds of gene cluster families.<sup>93</sup>

Gene clusters such as those for polyketides (PKS) or nonribosomal peptides (NRPS) are readily identified using bioinformatics, and to some extent the domain structures of the biosynthetic proteins allows prediction of the molecule that they specify.<sup>94, 95</sup> In recent years, excellent bioinformatics packages have been developed for the identification of BGCs. These include antiSMASH (ANTIbiotics & Secondary Metabolite Analysis SHell)<sup>96</sup> and SMURF (Secondary Metabolite Unknown Regions Finder), as well as packages that enable the identification of biosynthesis genes for specific subclasses of NPs (reviewed in<sup>97</sup>). Examples are SBSPKS (Structure Based Sequence Analysis of Polyketide Synthases) for polyketide BGCs,<sup>98</sup> NRPSPredictor for nonribosomal peptide gene clusters<sup>99</sup> and BAGEL for the identification of biosynthetic clusters for bacteriocins and lantibiotics.<sup>100</sup>

##### 4.1. From genotype to chemotype: bioinformatics-guided natural product discovery

How can we determine whether the compound that is derived from a given BGC is novel (*genotype to chemotype*) or, conversely, link a natural product of interest to its gene cluster (*chemotype to genotype*)? The *genotype-to-chemotype* roadmap (Figure 3A)<sup>101, 102</sup> provides a glimpse of ‘hidden treasures’, *e.g.* aided by the exploitation of pathway-specific regulators<sup>103, 104</sup> or via heterologous gene expression.<sup>105, 106</sup> Metagenomic libraries from environmental samples thereby allow scientists access to the ‘dark matter’, namely the vast biosynthetic potential of uncultivable microbes.<sup>107, 108</sup> However, an intrinsic disadvantage of genome-based

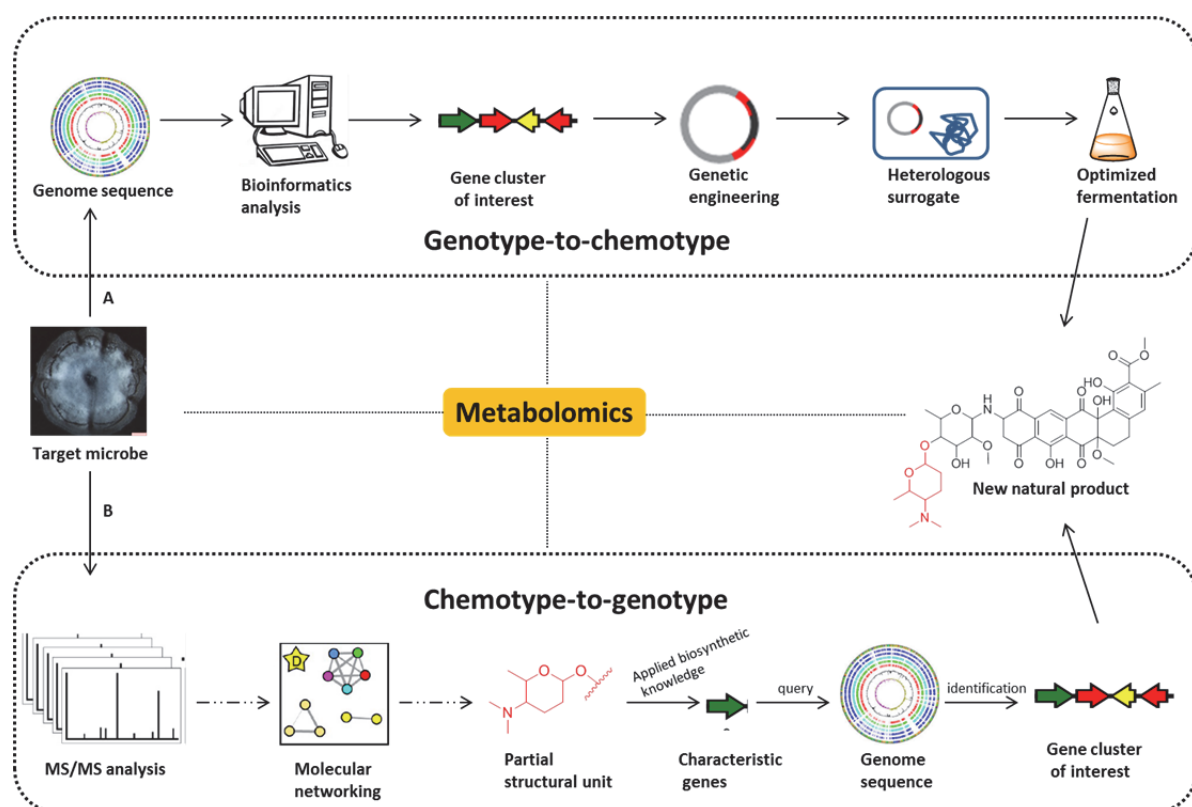
bioprospecting is that current technologies target a single BGC at the time, which is time-consuming, while the discovery efficiency is low as it is very difficult to prioritize ‘promising’ gene clusters. After all, if a gene cluster appears novel, it does not mean that the compound has not already been identified in HT screening regimes in the pre-genomic era. We predict, however, that synthetic biology and heterologous expression of gene clusters will soon become part of a more automated environment, which would make these approaches much more feasible. In addition, approaches directed at the (heterologous) expression of gene clusters depend on the functional analysis of biosynthetic enzymes and/or detection of intact clusters, while biosynthetic reactions may be catalyzed by enzymes that are not encoded by the gene cluster itself and/or by enzymes with low specificity,<sup>109</sup> or are catalyzed non-enzymatically.<sup>110, 111</sup> Indeed, metabolites are the end products of many cellular processes, so that the bioactive end product is not necessarily specified by individual genes or clusters.<sup>18</sup>

In the case of intertwined metabolic and biosynthetic pathways giving rise to a single compound, targeted genomic approaches lose their effectiveness.<sup>111-113</sup> In such cases, metabolomic approaches are required to bridge the gap between bioinformatics-driven gene cluster analysis and NP discovery.<sup>114, 115</sup> The fundamental principle is based on the fact that structurally related NPs are typically characterized by similar MS/MS fragmentation patterns. The MS/MS structural relatedness among molecules can be detected in an automated manner, and can subsequently generate a molecular network wherein analogues cluster together. In this way, known structures and “known unknowns” can be identified in extracts, offering an effective dereplication strategy and a way of prioritizing novel compounds.<sup>116</sup>

#### 4.2. From chemotype to genotype: metabolomics to identify biosynthetic gene clusters

MS/MS analysis allows the detection of “structural unit tags” or building blocks whose biosynthesis needs specific enzymes bearing substrate specificity. Candidate genes can be used as query against the genome databases to locate the target gene cluster for the compound(s) of interest. MS-based metabolomics in conjunction with molecular networking builds up a *chemotype-to-genotype* roadmap (Figure 3B)<sup>101, 102, 115, 117</sup> for NP discovery, directly linking secondary metabolites to their corresponding gene clusters. Examples for this concept include peptidogenomics,<sup>101, 118</sup> glycogenomics<sup>102</sup> and natural product proteomining.<sup>119</sup> Peptidogenomics is an MS-based genome-mining method that links peptide natural products to their BGCs through iteratively matching tandem MS (MS<sup>n</sup>) sequence tagging with genomics-derived peptide structures,<sup>101</sup> supported by software automation packages such as the recently developed Pep2Path.<sup>120</sup> For an excellent review of peptidogenomics techniques and their application we refer the reader elsewhere.<sup>115</sup> Glycogenomics is a *chemotype-to-genotype* approach for fast characterization of glycosylated natural products (GNPs) and their biosynthetic pathways. Tandem mass spectrometry (MS<sup>n</sup>) allows rapid identification of the structural information of sugar substituents of GNPs in microbial metabolic extracts.<sup>121, 122</sup> The biosynthesis of certain sugars need specific enzymes; as an example, genes for forosamine production encode 2,3- and 3,4-dehydratases, 3-ketoreductase, aminotransferase and *N,N*-dimethyltransferase. Applying the bacterial sugar biosynthetic knowledge, MS<sup>n</sup>-characterized glycosyl groups can be matched to corresponding glycosylation genes. In microbial genomes, the glycosylation-associated biosynthetic genes are usually co-clustered with those for a wide range of aglycones. Programs such as

antiSMASH are specifically designed with the identification of such signature BGCs in mind. The relevant aglycone biosynthetic genes are then identified, which in turn guides further isolation and structural identification of glycoconjugates by applying knowledge of the aglycone-specific BGC. As aglycones could be any type of structural scaffolds, for sugar-modified NPs, glycogenomics has an advantage over peptidogenomics. This technique therefore expands the applicability of MS-guided genome mining for different classes of NPs, such as nonribosomal and ribosomal peptides, polyketides, terpenes, and alkaloids.<sup>122</sup> Furthermore, since glycosylated products are produced via the decoration of an aglycone, which is often bioactive by itself, we envisage that glycogenomics can be extended to discover even non-glycosylated NPs that serve as intermediates for final GNPs.



**Figure 3. General workflows for genotype-to-chemotype (A) and chemotype-to-genotype (B) genome mining for natural product discovery.**

Natural product proteomining is based on the concept that gene cluster expression profiles correlate directly to the level of the corresponding natural product.<sup>119</sup> When growth conditions are chosen such that a bioactivity of interest is differentially produced, *e.g.* high under some conditions and intermediate or not at all in others, statistical correlation between bioactivity (metabolomics) and gene expression profiles (proteomics, RNAseq) will allow prioritizing those gene clusters with the optimal match. Since the metabolomics will also provide clues on the type of molecule, conclusive evidence of the BGC responsible for the product of interest can thus be obtained, depending on the number of candidate BGCs present in the genome and the number of samples analyzed.<sup>119</sup> The concept of metabolomics aiding genomics in finding new NPs is also exemplified by the case of myxoprincomide from the *Myxococcus xanthus*.<sup>123</sup> Initially, 13 BGCs in *M. xanthus* DK1622 were found for which the

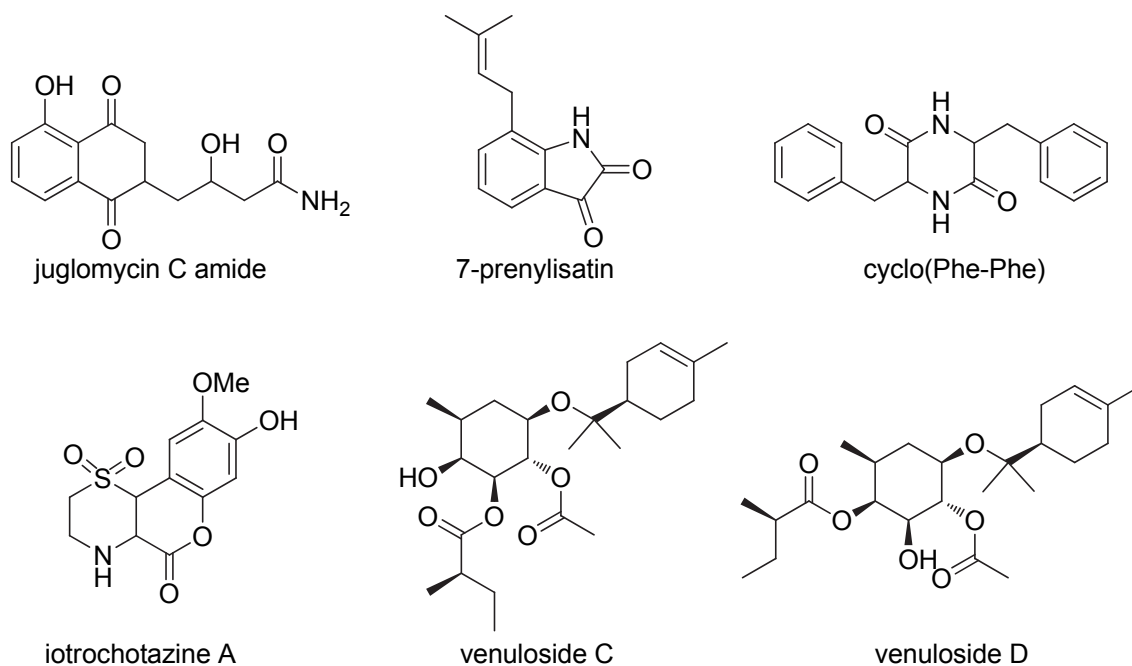
metabolites remained elusive. After targeted inactivation of these cryptic gene clusters, MS-based metabolomics was applied to pairwise compare extracts of mutants and the wild-type strain, highlighting the subtle metabolic differences by statistical filtering. Thus, a metabolite with molecular formula  $C_{45}H_{76}N_{10}O_{16}$  was assigned to a NRPS/PKS hybrid gene cluster and identified as a molecule with the unprecedented structural skeleton of myxoprincomide. Gene knockout/untargeted metabolomics strategy could also facilitate discovery of intermediates and/or new side products of biosynthesis pathway, as exemplified by characterization of 2-alkyl-4,5-dihydrothiazole-4-carboxylates from *Pseudomonas aeruginosa*.<sup>124,125</sup> Comparison of the extracellular metabolome of *P. aeruginosa pch* mutants with that of the wild-type strain identified 198 secondary metabolites regulated by these pathway-specific *pch* genes. Therefore, perturbation of the expression of global or pathway-specific regulators followed by metabolomics is a promising strategy to rapidly connect molecules to genes or gene clusters.

## 5. EXPERIMENTAL APPROACHES BASED ON NMR METABOLOMICS

As explained above, metabolomics approaches allow elucidating the nature of poorly bioactive molecules in complex mixtures, and for this achieving differential production of the bioactivity of interest is a key step. Such fluctuation may be achieved by among others different culturing conditions, inducing drug-resistance or by co-culturing.<sup>19, 20</sup> NMR-based metabolomics is very common for identifying plant NPs, but has yet seen relatively little application in microbial drug discovery. A general scheme is presented in Figure 1. We recently applied NMR-based metabolomics for the mining of novel soil isolates with promising antibiotic-producing potential, of which we provide a few examples here for illustration purposes. To identify a bioactivity produced by the soil isolate *Streptomyces* sp. MBT70, the strain was grown in media supplemented with different additives followed by  $^1H$  NMR spectroscopy and statistical comparison of the extracts by partial least square modeling-discriminant analysis (PLS-DA). A main discriminator between bioactive and inactive cultures was a signature H-5 residue belonging to a naphthoquinone. Further HPLC fractionation targeting naphthoquinones then led to the isolation and identification of juglomycin C amide (Figure 4).<sup>119</sup> The generation of streptomycin resistant ( $Str^R$ ) mutants in soil isolate *Streptomyces* sp. MBT28 resulted in the specific activation of a bioactive compound, which was subsequently identified by comparing the active ( $Str^R$ -induced) and non-active fractions by  $^1H$  NMR and statistical analysis by projection to latent structures (PLS); this allowed correlation of the enhanced bioactivity to a set of distinctive aromatic signals. NM-guided separation by tracking characteristic  $^1H$  NMR signals as probes, resulted in the characterization of the isatin-type antibiotic 7-prenylisatin (our unpublished data). As a third example, co-cultivation of the filamentous model microbes *S. coelicolor* and *Aspergillus niger* substantially influenced their NPs profiles. NMR spectroscopy of extracts obtained from the separate strains and the coculture combined with multivariate data analysis revealed diketopiperazine-type compounds like cyclo(Phe-Phe) (Figure 4), which were exclusively produced by *A. niger* when grown in co-culture with *S. coelicolor*.<sup>126</sup>

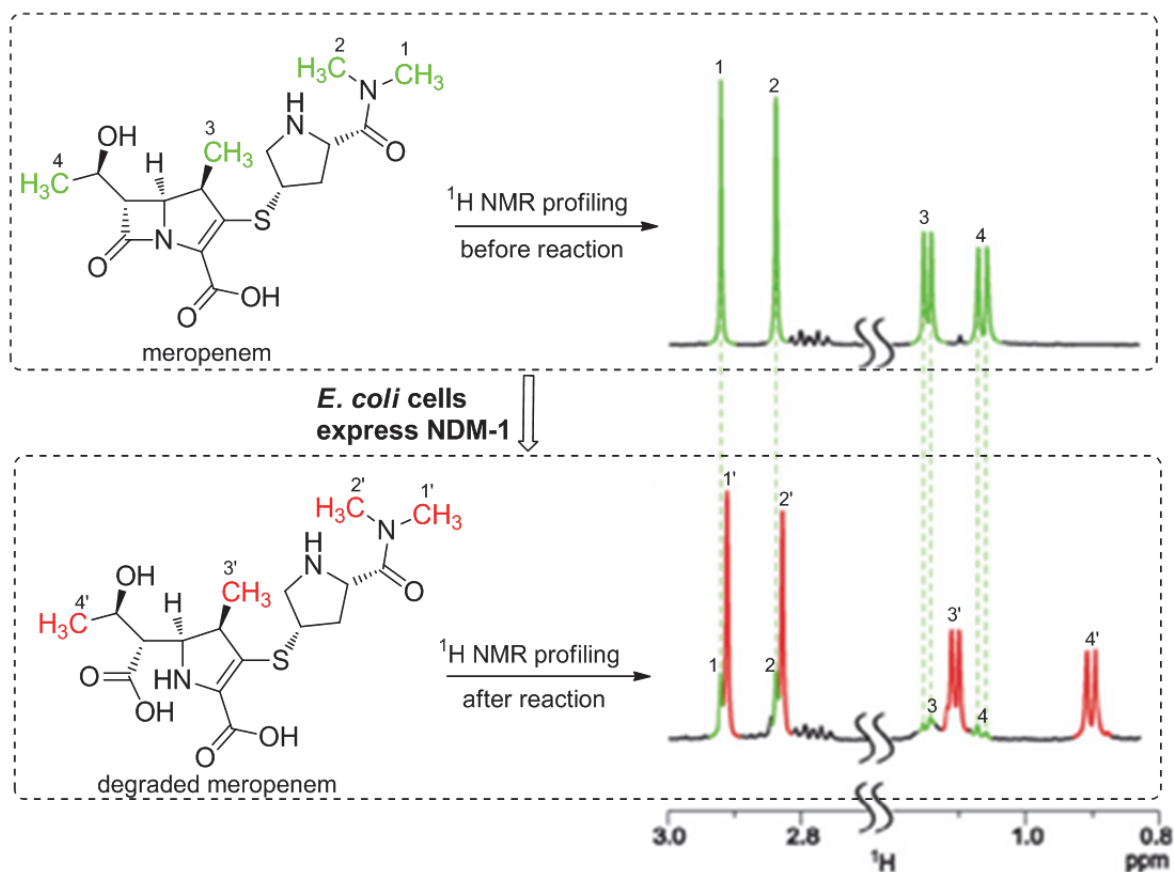
$^1H$  NMR based profiling also finds its application in lead discovery. Quinn and coworkers integrated solid phase extraction (SPE) with NMR fingerprinting. An extract library was prepared featuring lead-like enhanced (LLE) fractions selected on the basis of favorable physicochemical properties<sup>127</sup>, and this LLE library was subsequently profiled by

$^1\text{H}$  NMR spectroscopy to present the structural information of the small molecules contained in the library. In this way, the complicated and overlapping  $^1\text{H}$  NMR spectra typical of crude extracts were avoided by prefractionation, facilitating the detection of minor compounds. Using the proton signals as tracking signals for NMR-guided isolation, an unprecedented scaffold of iotrochotazine A was identified in LLE-fractionated extracts from sponges<sup>128</sup> and low-abundant venulosides C and D (Figure 4) in extracts from plants.<sup>129</sup>



**Figure 4.** Chemical structures of the compounds mentioned in the final section of the review (“Experimental approaches based on NMR metabolomics”).

An important class of NPs in the fight against antibiotic resistance is that of resistance inhibitors. Penicillins are still the most administered antibiotics,<sup>130, 131</sup> but resistance caused by  $\beta$ -lactamases was discovered almost as soon as the antibiotic itself<sup>132</sup> and has become a major problem for their clinical application.<sup>1</sup> The combined use of  $\beta$ -lactam antibiotics together with  $\beta$ -lactamase inhibitors is exemplified by the combination of amoxicillin and clavulanic acid, marketed as Augmentin®.<sup>133, 134</sup> A target-based in-cell NMR approach was recently proposed for use in high-throughput screening approaches to identify resistance inhibitors with activity against a specific  $\beta$ -lactamase.<sup>135</sup> The New Delhi metallo- $\beta$ -lactamase (NDM-1), which opens the cyclic amide ring in  $\beta$ -lactams antibiotics,<sup>136</sup> was monitored in real time in *E. coli* cells by  $^1\text{H}$  NMR spectroscopy.<sup>137</sup> The approach is summarized in Figure 5. The chemical shifts of four characteristic methyl groups are different between the substrate meropenem and its NDM-1-degraded product, which was used as a reporter of the lytic reaction. Inhibition of NMD-1 enzymatic activity was readily detected since the characteristic methyl groups of meropenem remain unshifted in the  $^1\text{H}$  NMR spectrum. NMR-guided fractionation was then applied to trace the responsible antibiotic adjuvants in the active fractions.



**Figure 5.** *In vivo*  $^1\text{H}$  NMR spectroscopy monitors NDM-1 activity in living *Escherichia coli* cells. Cyclic amide ring in  $\beta$ -lactam antibiotic meropenem is hydrolyzed under catalysis of enzyme NDM-1. The  $^1\text{H}$  NMR chemical shifts of four characteristic methyls (green) in the substrate accordingly move upfield in the hydrolytic product (red). This reaction can be monitored in real-time without disturbing the experimental system.

## 6. CONCLUSION

In a time where dereplication is one of the major challenges in screening efforts, metabolomics offers an effective strategy to prioritize novel molecules produced by microorganisms. Combination of eliciting strategies with comprehensive chemometric comparison of secondary metabolomes and genomics is an excellent alternative over traditional HT approaches. Metabolomics thereby complements (meta-)genomics and synthetic biology approaches in the mining of producing organisms, in particular for poorly expressed (cryptic) molecules. In addition, connecting chemotype to genotype and *vice versa* is important to facilitate molecular approaches and to accelerate unraveling of the biosynthetic pathway, and thus enable metabolic engineering and upscaling. Also taking the huge promise of the unculturable microorganisms into account, it is clear that natural product discovery is far from an echo from the past. Instead, as noted by many scientists worldwide, we may well be witnessing a new era of drug discovery, where many different disciplines such as genomics, metabolomics, bioinformatics, synthetic biology, ecology, chemical biology and industrial screening come together to meet the drug-discovery challenges and deliver new solutions to treat infectious diseases associated with drug resistance.



## REFERENCES

- (1) WHO, *Antimicrobial Resistance: Global Report on Surveillance*. Geneva, Switzerland, 2014.
- (2) Baltz, R. H. *Curr Opin Pharmacol* **2008**, *8*, 557-563.
- (3) Cooper, M. A.; Shlaes, D. *Nature* **2011**, *472*, 32.
- (4) Genilloud, O.; Gonzalez, I.; Salazar, O.; Martin, J.; Tormo, J. R.; Vicente, F. *J Ind Microbiol Biotechnol* **2011**, *38*, 375-389.
- (5) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat Rev Drug Discov* **2007**, *6*, 29-40.
- (6) Bérdy, J. *J Antibiot (Tokyo)* **2005**, *58*, 1-26.
- (7) Newman, D. J.; Cragg, G. M. *J Nat Prod* **2007**, *70*, 461-477.
- (8) Kaeberlein, T.; Lewis, K.; Epstein, S. S. *Science* **2002**, *296*, 1127-1129.
- (9) Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Schaberle, T. F.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. A.; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K. *Nature* **2015**, *517*, 455-459.
- (10) Wilson, M. C.; Mori, T.; Ruckert, C.; Uriá, A. R.; Helf, M. J.; Takada, K.; Gernert, C.; Steffens, U. A.; Heycke, N.; Schmitt, S.; Rinke, C.; Helfrich, E. J.; Brachmann, A. O.; Gurgui, C.; Wakimoto, T.; Kracht, M.; Crusemann, M.; Hentschel, U.; Abe, I.; Matsunaga, S.; Kalinowski, J.; Takeyama, H.; Piel, J. *Nature* **2014**, *506*, 58-62.
- (11) Bergmann, S.; Schumann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. *Nat Chem Biol* **2007**, *3*, 213-217.
- (12) Zhu, H.; Sandiford, S. K.; van Wezel, G. P. *J Ind Microbiol Biotechnol* **2014**, *41*, 371-386.
- (13) Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. *Nature* **2002**, *417*, 141-147.
- (14) Cruz-Morales, P.; Vijgenboom, E.; Iruegas-Bocardo, F.; Girard, G.; Yanez-Guerra, L. A.; Ramos-Aboites, H. E.; Pernodet, J. L.; Anne, J.; van Wezel, G. P.; Barona-Gomez, F. *Genome Biol Evol* **2013**, *5*, 1165-1175.
- (15) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat Biotechnol* **2003**, *21*, 526-531.
- (16) van Keulen, G.; Dyson, P. J. *Adv Appl Microbiol* **2014**, *89*, 217-266.
- (17) Pauli, G. F.; Chen, S. N.; Friesen, J. B.; McAlpine, J. B.; Jaki, B. U. *J Nat Prod* **2012**, *75*, 1243-1255.
- (18) Perić-Concha, N.; Long, P. F. *Drug Disc Today* **2003**, *8*, 1078-1084.
- (19) Gaudêncio, S. P.; Pereira, F. *Nat Prod Rep* **2015**, *32*, 779-810.
- (20) Wu, C.; Kim, H. K.; van Wezel, G. P.; Choi, Y. H. *Drug Disc Today: Technol* **2015**, *13*,

- 11–17.
- (21) Weber, T.; Charusanti, P.; Musiol-Kroll, E. M.; Jiang, X.; Tong, Y.; Kim, H. U.; Lee, S. Y. *Trends Biotechnol* **2015**, *33*, 15-26.
  - (22) Bode, H. B. r.; Bethe, B.; H<sup>^</sup>fs, R.; Zeeck, A. *ChemBioChem* **2002**, *3*, 619-627.
  - (23) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, *160*, 1714-1725.
  - (24) Ohnishi, Y.; Ishikawa, J.; Hara, H.; Suzuki, H.; Ikenoya, M.; Ikeda, H.; Yamashita, A.; Hattori, M.; Horinouchi, S. *J Bacteriol* **2008**, *190*, 4050-4060.
  - (25) Fu, P.; Jamison, M.; La, S.; Macmillan, J. B. *Org Lett* **2014**, *16*, 5656-5659.
  - (26) Hosaka, T.; Ohnishi-Kameyama, M.; Muramatsu, H.; Murakami, K.; Tsurumi, Y.; Kodani, S.; Yoshida, M.; Fujie, A.; Ochi, K. *Nat Biotechnol* **2009**, *27*, 462-464.
  - (27) Ochi, K.; Tanaka, Y.; Tojo, S. *J Ind Microbiol Biotechnol* **2014**, *41*, 403-414.
  - (28) Rigali, S.; Titgemeyer, F.; Barends, S.; Mulder, S.; Thomae, A. W.; Hopwood, D. A.; van Wezel, G. P. *EMBO Rep* **2008**, *9*, 670-675.
  - (29) Willey, J. M.; Gaskell, A. A. *Chem Rev* **2011**, *111*, 174-187.
  - (30) Hsiao, N. H.; Gottelt, M.; Takano, E. *Methods Enzymol* **2009**, *458*, 143-157.
  - (31) Albright, J. C.; Henke, M. T.; Soukup, A. A.; McClure, R. A.; Thomson, R. J.; Keller, N. P.; Kelleher, N. L. *ACS Chem Biol* **2015**, In press.
  - (32) Bertrand, S.; Bohni, N.; Schnee, S.; Schumpp, O.; Gindro, K.; Wolfender, J.-L. *Biotechnol Adv* **2014**, *32*, 1180-1204.
  - (33) Hoshino, S.; Wakimoto, T.; Onaka, H.; Abe, I. *Org Lett* **2015**, *17*, 1501-1504.
  - (34) Sugiyama, R.; Nishimura, S.; Ozaki, T.; Asamizu, S.; Onaka, H.; Kakeya, H. *Org Lett* **2015**, *17*, 1918-1921.
  - (35) Thaker, M. N.; Wang, W.; Spanogiannopoulos, P.; Waglechner, N.; King, A. M.; Medina, R.; Wright, G. D. *Nat Biotechnol* **2013**, *31*, 922-927.
  - (36) Bibb, M. J. *Curr Opin Microbiol* **2005**, *8*, 208-215.
  - (37) Liu, G.; Chater, K. F.; Chandra, G.; Niu, G.; Tan, H. *Microbiol Mol Biol Rev* **2013**, *77*, 112-143.
  - (38) Sanchez, S.; Chavez, A.; Forero, A.; Garcia-Huante, Y.; Romero, A.; Sanchez, M.; Rocha, D.; Sanchez, B.; Avalos, M.; Guzman-Trampe, S.; Rodriguez-Sanoja, R.; Langley, E.; Ruiz, B. *J Antibiot (Tokyo)* **2010**, *63*, 442-459.
  - (39) van Wezel, G. P.; McDowall, K. J. *Nat Prod Rep* **2011**, *28*, 1311-1333.
  - (40) Wietzorrek, A.; Bibb, M.; Chakraborty, R. *Mol Microbiol* **1997**, *25*, 1181-1184.
  - (41) Floriano, B.; Bibb, M. *Mol Microbiol* **1996**, *21*, 385-396.
  - (42) Arias, P.; Fernandez-Moreno, M. A.; Malpartida, F. *J Bacteriol* **1999**, *181*, 6958-6968.
  - (43) Takano, E.; Gramajo, H. C.; Strauch, E.; Andres, N.; White, J.; Bibb, M. J. *Mol Microbiol* **1992**, *6*, 2797-2804.
  - (44) van Wezel, G. P.; White, J.; Hoogvliet, G.; Bibb, M. J. *J Mol Microbiol Biotechnol* **2000**, *2*, 551-556.
  - (45) Bibb, M.; Hesketh, A. *Methods Enzymol* **2009**, *458*, 93-116.
  - (46) van Wezel, G. P.; McKenzie, N. L.; Nodwell, J. R. *Methods Enzymol* **2009**, *458*, 117-141.
  - (47) Colson, S.; Stephan, J.; Hertrich, T.; Saito, A.; van Wezel, G. P.; Titgemeyer, F.; Rigali, S. *J Mol Microbiol Biotechnol* **2007**, *12*, 60-66.

- (48) Colson, S.; van Wezel, G. P.; Craig, M.; Noens, E. E.; Nothaft, H.; Mommaas, A. M.; Titgemeyer, F.; Joris, B.; Rigali, S. *Microbiology* **2008**, *154*, 373-382.
- (49) Craig, M.; Lambert, S.; Jourdan, S.; Tenconi, E.; Colson, S.; Maciejewska, M.; Ongena, M.; Martin, J. F.; van Wezel, G.; Rigali, S. *Environ Microbiol Rep* **2012**, *4*, 512-521.
- (50) Rigali, S.; Nothaft, H.; Noens, E. E.; Schlicht, M.; Colson, S.; Muller, M.; Joris, B.; Koerten, H. K.; Hopwood, D. A.; Titgemeyer, F.; van Wezel, G. P. *Mol Microbiol* **2006**, *61*, 1237-1251.
- (51) Swiatek, M. A.; Tenconi, E.; Rigali, S.; van Wezel, G. P. *J Bacteriol* **2012**, *194*, 1136-1144.
- (52) Nazari, B.; Kobayashi, M.; Saito, A.; Hassaninasab, A.; Miyashita, K.; Fujii, T. *Appl Environ Microbiol* **2012**, *79*, 707-713.
- (53) Swiatek-Polatynska, M. A.; Bucca, G.; Laing, E.; Gubbens, J.; Titgemeyer, F.; Smith, C. P.; Rigali, S.; van Wezel, G. P. *PLoS One* **2015**, *10*, e0122479.
- (54) Uguru, G. C.; Stephens, K. E.; Stead, J. A.; Towle, J. E.; Baumberg, S.; McDowall, K. J. *Mol Microbiol* **2005**, *58*, 131-150.
- (55) Hong, B.; Phornphisutthimas, S.; Tilley, E.; Baumberg, S.; McDowall, K. J. *Biotechnol Lett* **2007**, *29*, 57-64.
- (56) Mao, X. M.; Luo, S.; Zhou, R. C.; Wang, F.; Yu, P.; Sun, N.; Chen, X. X.; Tang, Y.; Li, Y. Q. *J Biol Chem* **2015**, *290*, 7992-8001.
- (57) Santos-Beneit, F.; Rodriguez-Garcia, A.; Sola-Landa, A.; Martin, J. F. *Mol Microbiol* **2009**, *72*, 53-68.
- (58) Kim, S. H.; Traag, B. A.; Hasan, A. H.; McDowall, K. J.; Kim, B. G.; van Wezel, G. P. *Antonie Van Leeuwenhoek* **2015**, *108*, 201-213.
- (59) Nothaft, H.; Rigali, S.; Boomsma, B.; Swiatek, M.; McDowall, K. J.; van Wezel, G. P.; Titgemeyer, F. *Mol Microbiol* **2010**, *75*, 1133-1144.
- (60) Mahr, K.; van Wezel, G. P.; Svensson, C.; Krengel, U.; Bibb, M. J.; Titgemeyer, F. *Antonie Van Leeuwenhoek* **2000**, *78*, 253-261.
- (61) Titgemeyer, F.; Reizer, J.; Reizer, A.; Saier, M. H., Jr. *Microbiology* **1994**, *140*, 2349-2354.
- (62) Martinez, A.; Kolvek, S. J.; Hopke, J.; Yip, C. L.; Osburne, M. S. *Appl Environ Microbiol* **2005**, *71*, 1638-1641.
- (63) Swiatek, M. A.; Gubbens, J.; Bucca, G.; Song, E.; Yang, Y. H.; Laing, E.; Kim, B. G.; Smith, C. P.; van Wezel, G. P. *J Bacteriol* **2013**, *195*, 1236-1248.
- (64) Tomono, A.; Tsai, Y.; Yamazaki, H.; Ohnishi, Y.; Horinouchi, S. *J Bacteriol* **2005**, *187*, 5595-5604.
- (65) Higo, A.; Hara, H.; Horinouchi, S.; Ohnishi, Y. *DNA Res* **2012**, *19*, 259-273.
- (66) Lee, P. C.; Umeyama, T.; Horinouchi, S. *Mol Microbiol* **2002**, *43*, 1413-1430.
- (67) Sola-Landa, A.; Moura, R. S.; Martin, J. F. *Proc Natl Acad Sci U S A* **2003**, *100*, 6133-6138.
- (68) Patti, G. J.; Yanes, O.; Siuzdak, G. *Nat Rev Mol Cell Biol* **2012**, *13*, 263-269.
- (69) Kuehnbaum, N. L.; Britz-Mckibbin, P. *Chem Rev* **2013**, *113*, 2437-2468.
- (70) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Trends Biotechnol* **2011**, *29*, 267-275.
- (71) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat Protoc* **2010**, *5*, 536-549.
- (72) Mahrous, E. A.; Farag, M. A. *J Adv Res* **2015**, *6*, 3-15.

- (73) Novoa-Carballal, R.; Fernandez-Megia, E.; Jimenez, C.; Riguera, R. *Nat Prod Rep* **2011**, *28*, 78-98.
- (74) Sidebottom, A. M.; Johnson, A. R.; Karty, J. A.; Trader, D. J.; Carlson, E. E. *ACS Chem Biol* **2013**, *8*, 2009-2016.
- (75) Breton, R. C.; Reynolds, W. F. *Nat Prod Rep* **2013**, *30*, 501-524.
- (76) Nicolaou, K. C.; Snyder, S. A. *Angew Chem Intl Ed* **2005**, *44*, 1012-1044.
- (77) Senior, M. M.; Williamson, R. T.; Martin, G. E. *J Nat Prod* **2013**, *76*, 2088-2093.
- (78) Kövér, K. E.; Hruby, V. J.; Uhrin, D. *J Magn Res* **1997**, *129*, 125-129.
- (79) Martin, G. E., Using 1,1- and 1,*n*-ADEQUATE 2D NMR data in structure elucidation protocols. In *Annual Rep NMR Spectr*, Elsevier: London, 2011; pp 215-291.
- (80) Meyer, S. W.; Köck, M. *J Nat Prod* **2008**, *71*, 1524-1529.
- (81) Cheatham, S. F.; Kline, M.; Sasaki, R. R.; Blinov, K. A.; Elyashberg, M. E.; Molodtsov, S. G. *Magn Res Chem* **2010**, *48*, 571-574.
- (82) Buevich, A. V.; Williamson, R. T.; Martin, G. E. *J Nat Prod* **2014**, *77*, 1942-1947.
- (83) Martin, G. E.; Williamson, R. T.; Dormer, P. G.; Bermel, W. *Magn Res Chem* **2012**, *50*, 563-568.
- (84) Wolfender, J.-L.; Ndjoko, K.; Hostettmann, K. *Phytochem Anal* **2001**, *12*, 2-22.
- (85) Månsson, M.; Phipps, R. K.; Gram, L.; Munro, M. H. G.; Larsen, T. O.; Nielsen, K. F. *J Nat Prod* **2010**, *73*, 1126-1132.
- (86) Seger, C.; Sturm, S.; Stuppner, H. *Nat Prod Rep* **2013**, *30*, 970-987.
- (87) Sprogøe, K.; Stærk, D.; Jäger, A. K.; Adsersen, A.; Hansen, S. H.; Witt, M.; Landbo, A. K.; Meyer, A. S.; Jaroszewski, J. W. *J Nat Prod* **2007**, *70*, 1472-1477.
- (88) Lambert, M.; Stærk, D.; Hansen, S. H.; Sairafianpour, M.; Jaroszewski, J. W. *J Nat Prod* **2005**, *68*, 1500-1509.
- (89) Johansen, K. T.; Wubshet, S. G.; Nyberg, N. T.; Jaroszewski, J. W. *J Nat Prod* **2011**, *74*, 2454-2461.
- (90) Liu, H.; Zheng, A.; Liu, H.; Yu, H.; Wu, X.; Xiao, C.; Dai, H.; Hao, F.; Zhang, L.; Wang, Y.; Tang, H. *J Agr Food Chem* **2012**, *60*, 129-135.
- (91) Exarchou, V.; Godejohann, M.; Beek, T. A. V.; Gerothanassis, I. P.; Vervoort, J. *Anal Chem* **2003**, *75*, 6288-6294.
- (92) Corre, C.; Challis, G. L. *Nat Prod Rep* **2009**, *26*, 977-986.
- (93) Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland Brown, L. C.; Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J.; Birren, B. W.; Takano, E.; Sali, A.; Linington, R. G.; Fischbach, M. A. *Cell* **2014**, *158*, 412-421.
- (94) Bachmann, B. O.; Ravel, J. *Methods Enzymol* **2009**, *458*, 181-217.
- (95) Lautru, S.; Challis, G. L. *Microbiology* **2004**, *150*, 1629-1636.
- (96) Medema, M.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. *Nucleic Acids Res* **2011**, *39*, W339-W346.
- (97) Fedorova, N. D.; Moktali, V.; Medema, M. H. *Methods Mol Biol* **2012**, *944*, 23-45.
- (98) Anand, S.; Prasad, M. V.; Yadav, G.; Kumar, N.; Shehara, J.; Ansari, M. Z.; Mohanty, D. *Nucleic Acids Res* **2010**, *38*, W487-496.
- (99) Rottig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. *Nucleic Acids Res* **2011**, *39*, W362-367.
- (100) de Jong, A.; van Heel, A. J.; Kok, J.; Kuipers, O. P. *Nucleic Acids Res* **2010**, *38*,

- W647-651.
- (101) Kersten, R. D.; Yang, Y. L.; Xu, Y.; Cimermancic, P.; Nam, S. J.; Fenical, W.; Fischbach, M. A.; Moore, B. S.; Dorrestein, P. C. *Nat Chem Biol* **2011**, *7*, 794-802.
  - (102) Kersten, R. D.; Ziemert, N.; Gonzalez, D. J.; Duggan, B. M.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. *Proc Natl Acad Sci U S A* **2013**, *110*, 4407-4416.
  - (103) Laureti, L.; Song, L.; Huang, S.; Corre, C.; Leblond, P.; Challis, G. L.; Aigle, B. *Proc Natl Acad Sci U S A* **2011**, *108*, 6258-6263.
  - (104) Gottelt, M.; Kol, S.; Gomez-Escribano, J. P.; Bibb, M.; Takano, E. *Microbiology* **2010**, *156*, 2343-2353.
  - (105) Baltz, R. H. *J Ind Microbiol Biotechnol* **2010**, *37*, 759-772.
  - (106) Ongley, S. E.; Bian, X.; Neilan, B. A.; Müller, R. *Nat Prod Rep* **2013**, *30*, 1121-1138.
  - (107) Brady, S. F.; Simmons, L.; Kim, J. H.; Schmidt, E. W. *Nat Prod Rep* **2009**, *26*, 1488-1503.
  - (108) Kang, H. S.; Brady, S. F. *ACS Chem Biol* **2014**, *9*, 1267-1272.
  - (109) Schwab, W. *Phytochemistry* **2003**, *62*, 837-849.
  - (110) Mayer, M.; Thiericke, R. *J Antibiot (Tokyo)* **1993**, *46*, 1372-1380.
  - (111) Wu, C.; van Wezel, G. P.; Choi, Y. H. *J Antibiot (Tokyo)* **2015**, In press.
  - (112) Vingadassalon, A.; Lorieux, F.; Juguet, M.; Gof, G. r. L.; Gerbaud, C.; Pernodet, J.-L.; Lautru, S. *ACS Chem Biol* **2015**, *10*, 601-610.
  - (113) Tsukioka, Y.; Yamashita, Y.; Nakano, Y.; Oho, T.; Koga, T. *J Bacteriol* **1997**, *179*, 4411-4414.
  - (114) Doroghazi, J. R.; Albright, J. C.; Goering, A. W.; Ju, K.-S.; Haines, R. R.; Tchalukov, K. A.; Labeda, D. P.; Kelleher, N. L.; Metcalf, W. W. *Nat Chem Biol* **2014**, *10*, 963-970.
  - (115) Krug, D.; Müller, R. *Nat Prod Rep* **2014**, *31*, 768-783.
  - (116) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; Wong, W. R.; Linington, R. G.; Zhang, L.; Debonisi, H. M.; Gerwick, W. H.; Dorrestein, P. C. *J Nat Prod* **2013**, *76*, 1686-1699.
  - (117) Helfrich, E. J. N.; Reiter, S.; Piel, J. *Curr Opin Biotechnol* **2014**, *29*, 107-115.
  - (118) Nguyen, D. D.; Wu, C. H.; Moree, W. J.; Lamsa, A.; Medema, M. H.; Zhao, X.; Gavilan, R. G.; Aparicio, M.; Atencio, L.; Jackson, C.; Ballesteros, J.; Sanchez, J.; Watrous, J. D.; Phelan, V. V.; van de Wiel, C.; Kersten, R. D.; Mehnaz, S.; De Mot, R.; Shank, E. A.; Charusanti, P.; Nagarajan, H.; Duggan, B. M.; Moore, B. S.; Bandeira, N.; Palsson, B. O.; Pogliano, K.; Gutierrez, M.; Dorrestein, P. C. *Proc Natl Acad Sci U S A* **2013**, *110*, E2611-2620.
  - (119) Gubbens, J.; Zhu, H.; Girard, G.; Song, L.; Florea, B. I.; Aston, P.; Ichinose, K.; Filippov, D. V.; Choi, Y. H.; Overkleeft, H. S.; Challis, G. L.; van Wezel, G. P. *Chem Biol* **2014**, *21*, 707-718.
  - (120) Medema, M. H.; Paalvast, Y.; Nguyen, D. D.; Melnik, A.; Dorrestein, P. C.; Takano, E.; Breitling, R. *PLoS Comput Biol* **2014**, *10*, e1003822.
  - (121) Herget, S.; Toukach, P. V.; Ranzinger, R.; Hull, W. E.; Knirel, Y. A.; von der Lieth, C.-W. *BMC Struct Biol* **2008**, *8*, 35.
  - (122) Elshahawi, S. I.; Shaaban, K. A.; Kharel, M. K.; Thorson, J. S. *Chem Soc Rev* **2015**, In press.

- 
- (123) Cortina, N. S.; Krug, D.; Plaza, A.; Revermann, O.; Müller, R. *Angew Chem Intl Ed* **2012**, *51*, 811-816.
- (124) Kersten, R. D.; Dorrestein, P. C. *ACS Chem Biol* **2009**, *4*, 599-601.
- (125) Vinayavekhin, N.; Saghatelian, A. *ACS Chem Biol* **2009**, *4*, 617-623.
- (126) Wu, C.; Zacchetti, B.; Ram, A. F. J.; van Wezel, G. P.; Claessen, D.; Choi, Y. H. *Sci Rep* **2015**, *5*, 10868.
- (127) Camp, D.; Davis, R. A.; Campitelli, M.; Ebdon, J.; Quinn, R. J. *J Nat Prod* **2012**, *75*, 72-81.
- (128) Grkovic, T.; Pouwer, R. H.; Vial, M. L.; Gambini, L.; Noël, A.; Hooper, J. N. A.; Wood, S. A.; Mellick, G. D.; Quinn, R. J. *Angew Chem Intl Ed* **2014**, *53*, 6070-6074.
- (129) Grkovic, T.; Pouwer, R. H.; Wang, Q.; Guymer, G. P.; Holst, J.; Quinn, R. J. *J Nat Prod* **2015**, In press.
- (130) Demain, A. L. *Proc Natl Sci Counc Repub China B* **1991**, *15*, 251-265.
- (131) Southgate, R.; Elson, S. *Fortschr Chem Org Naturst* **1985**, *47*, 1-106.
- (132) Abraham, E. P.; Chain, E. *Nature* **1940**, *46*, 837.
- (133) Drawz, S. M.; Bonomo, R. A. *Clin Microbiol Rev* **2010**, *23*, 160-201.
- (134) Fisher, J. F.; Meroueh, S. O.; Mobashery, S. *Chem Rev* **2005**, *105*, 395-424.
- (135) Ma, J.; Cao, Q.; Mcleod, S. M.; Ferguson, K.; Gao, N.; Breeze, A. L. *Angew Chem Intl Ed* **2015**, *54*, 4764-4767.
- (136) Yong, D.; Toleman, M. A.; Giske, C. G.; Cho, H. S.; Sundman, K.; Lee, K.; Walsh, T. R. *Antimicrob Agents Chemother* **2009**, *53*, 5046-5054.
- (137) Ma, J.; McLeod, S.; MacCormack, K.; Sriram, S.; Gao, N.; Breeze, A. L.; Hu, J. *Angew Chem Intl Ed* **2014**, *53*, 2130-2133.