Cell Subset Targeting of Glycosylated Polyketides Revealed by Multiplexed Phenotypic Screening of Natural Product Fraction Libraries and Bioengineered Analogs

By

David Earl

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Chemical and Physical Biology

May 11, 2018

Nashville, Tennessee

Approved:

Gary A. Sulikowski, Ph.D.

Brian O. Bachmann, Ph.D.

Jonathan M. Irish, Ph.D.

Eric P. Skaar, Ph.D.

Kevin L. Schey, Ph.D.

ACKNOWLEDGMENTS

First and foremost, this work would not have been possible without the support of my advisor, Dr. Brian Bachmann. I am deeply grateful for the encouragement and flexibility he has provided me in exploring my scientific curiosities and allowing me to pursue new research avenues outside of the typical focus of the lab.

Similarly, I must thank Dr. Jonathan Irish for literally giving me the keys to his lab and offering the training and resources necessary to develop the cytometry assays for this project. I also appreciate the time and advice given by his lab members especially the patience shown by Dr. Nalin Leelatian in teaching me how to properly perform barcode staining.

Many thanks are due to Dr. Brent Ferrell for giving me crash courses in Hematology and Immunology, for his help in planning and analyzing experiments, and for generously providing patient samples.

I am also grateful for the assistance given by our collaborators in the Sulikowski lab in preparing analogs, probe compounds, and substrates for feeding studies.

I would also like to thank my committee members for their advice and guidance in developing my research and goals.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF SPECTRA

Spectrum **Page**

CHAPTER 1

Introduction

A Historical Perspective on Natural Product Discovery

From their initial incorporation as part of shamanistic rituals by tribal healers to modern use by physicians, mankind has relied upon natural products to treat their ailments. The evolution from herbal remedies and accompanying superstitions to the current use of purified bioactive constituents of natural product extracts as pharmaceuticals is directly tied to the technological and philosophical advances of humanity. The earliest written record of formulations for treating disease is the Ebers Papyrus compiled by ancient Egyptian priests[1]. As the main theory of disease was demonic possession, these remedies where used as diuretics and to induce emesis, as the priests believed physical expulsion would affect spiritual expulsion.

The Greeks continued the study and use of medicinal plants and while they rejected the demonic theory of disease, they instead relied on astrological events to inform prescriptions. The physician Galen began testing formulations of multiple plants believing that each plant possessed a unique potency and was among the first to advocate empirical testing of plant extracts. He also compiled the medicinal knowledge of the Greeks in his work, *On the Art of Healing*, which shaped the practice of medicine for the next 1500 years.

The next major advance in the development of therapeutics was made by the Swiss physician Paracelsus, who borrowing from the thinking of his contemporary alchemists, advocated the finding of the singular healing *arcanum* within each herbal remedy[2].

The growing influence of empiricism in the 1700s lead most physicians to largely disregard traditional medicines and begin evaluating the potential of heliotherapy, hydrotherapy, and electrotherapy instead, albeit with limited success. However, the founding of the Societe de Pharmacie in Paris in 1803 marked a return to the study of plants as a source of therapeutics. The founding director, Nicolas Vauquelin, pushed his students and faculty to apply analytical chemistry to plant materials to identify the active components. Within its first year, Jean-Francois Derosne reported the isolation of a crystalline salt with alkaline properties while working with opium. However due to the prevailing theory that organic acids were likely to be the responsible for bioactivity, he attributed it as a contaminant from potash.

Concurrently, an Austrian pharmacist, Friedrich Wilhelm Serturner, isolated meconic acid from opium and tested the compound on dogs but found no narcotic activity. He then tested an alkali substance from the opium extract and found that it was active. He refined his methods and published them in 1817 in a paper entitled *Morphium, a Salt-like Base and Meconic Acid as the Chief Constituents of Opium.* Included in the report was a description of the effects upon the author and several friends after swallowing 100 mg each of the newly discovered salt[3].

Upon reading Serturner's work, the editor of the *Annales de Chimie*, Joseph Gay-Lussac translated the paper and republished it with an accompanying editorial predicting the discovery of many more plant alkali salts and proposed using the suffix '-ine' for these

compounds. This prediction proved correct as Joseph Pelletier reported the isolation of emetine from the root plant *Cephaelis ipecacuanha* later that year[4].

Pelletier next set out to test the biologist Carl Linnaeus's hypothesis that plants of related genus would have similar pharmacological properties and with his student Joseph Caventou isolated strychnine from three different species of the *Strychnos* family. However, their next project proved to be of even greater significance as they were able to isolate quinine from cinchona bark. Importantly, in their 1820 paper reporting its isolation, they advocated for physicians to begin studying the effects of administering purified compounds. This work sparked the genesis of the modern pharmaceutical industry and by 1826, 150,000 kg of cinchona bark were being processed annually to yield 3600 kg of quinine[4].

Bioactive constituents from plant sources began to be discovered with increasing frequency such as isolation of cocaine in 1860 by Albert Niemann[5]. The success of this avenue of research lead other scientists to begin exploring other organisms as sources of natural products. Epinephrine was isolated by John Jacob Abel in 1897 from sheep adrenal glands[6] and the development of bacterial culture by Robert Kock in the 1880s established the groundwork that culminated in the discovery penicillin by Fleming and allowed the ushering in the 'golden age' of antibiotic discovery from the 1940s to 1960s[7].

Today the effort to identify and isolate chemical compounds produced by organisms that exert a biological effect upon other organisms or upon itself continues to be a central focus of the field of natural product chemistry. Obtaining the purified component of interest allows for reproducible dosing, structural determination, synthetic investigation, and chemical biology assay development for determining mode of action.

Overlap of Natural Products with Biological Space

In addition to their importance to the development of modern medicine, the study of natural product chemistry and biochemistry is motivated by the desire to address several basic science questions. The observation that drugs may be found in nature leads to the fundamental question of why compounds that were naturally selected for in a non-disease context are useful as therapeutics. Natural products have long been viewed as a privileged class of compounds and with the advent of large compound databases and collections, chemoinformatic approaches have offered several intriguing hypotheses to address this question.

In 1999 Reichel compared databases of natural products and synthetic compounds and found that natural products typically contain three or more stereocenters, more oxygen, and less nitrogen atoms then drugs and synthetics. They also observed that natural products contained more pharmacophoric features, which are arrangements of functional groups conducive to interacting with macromolecules[8].

This work was extended by Lee and Schneider who compared chemical properties developed by Lipinski known as the rule-of-five[9]. The rule-of-five states that a drug candidate should have a molecular weight less than 500 Daltons, a log P of less than 5, and contain no more than 5 hydrogen donors. They concluded that while natural products are more lipophilic in general, they violated the rule-of-five at similar frequency compared to synthetic compounds, which is in contrast to the commonly held industry view that natural products are especially disadvantaged as therapeutic candidates[10].

These types of approaches led to a growing interest in mapping chemical space where each dimension corresponds to a molecular descriptor. An estimated 10^{60} carbon based small molecules could populate chemical space, whereas most organisms produce no more than a few thousand small molecules[11]. Therefore, biological chemical space is a vastly small subset of total chemical space. Feher and Schmidt were among the first to use principal component analysis to visualize the overlap of synthetic compounds and natural products with bioactive molecules in their 2003 publication[12].

One theory on the relatively small size of chemical biological space is that there are relativity few known protein folds given the infinite number of potential peptide sequences, with several highly conserved superfamilies occurring in most genomes. Waldmann et al. argue that this macromolecular constraint is the reason for limited size of bioactive chemical space[13]. Quinn et al. further hypothesized that shared protein fold topology in biosynthetic enzymes and in druggable human proteins accounts for the privilege of natural product scaffolds, as natural products must have reasonable affinity to the enzymes that are synthesizing them[14].

Therapeutic Importance of Natural Products

The overlap of chemical and biological space allows bioactive natural products to selectively interact with discrete biochemical targets, engendering remarkable phenotypic changes in the biology of treated organisms. Correspondingly, natural products have played a central role for drug discovery programs and as probes of biological function leading to new potential therapeutic targets. For example, the hedgehog-signaling pathway was discovered using the alkaloid cyclopamine, leading to the discovery of the smoothened (SMO) target[15]. This discovery provided the basis for the development of selective SMO target inhibitors such as vismodegib for treatment of basal cell carcinoma[16]. Newman and Cragg have written extensive reviews on the clinical importance of natural products and note that from 1980 - 2010, 79% of small molecule drugs entering the clinic were either natural products, natural product derivatives, or natural product mimics. Recent examples include the antibiotics biapenem, ceftobiprole, and telavancin, the antiviral peramivir, and the anticancer compounds cabazitaxel and pralatrexate[17].

Often natural products with promising bioactivity are abandoned if they do not possess ideal pharmacological properties or have other undesirable properties such as inefficient synthetic supply or tractability for modifications. However the recent examples of daptomycin[18] and bryostatin 1[19] show how development of new formulations or combinations of biosynthetic and synthetic approaches can overcome these issues.

Statement of Dissertation

Despite the historical and continued therapeutic importance of natural products, efficient methods for exploring and defining structural activity relationships of natural products are still lacking. Mapping structure activity relationships identifies sites on the molecule that are amenable to modification without loss of potency which illuminates strategies to improve pharmacokinetics or develop analogs that are useful as chemical probes in mechanism of action studies. However, the structural complexity of natural products limits the success of traditional medicinal chemistry efforts and hinders their development as therapeutics and tool compounds.

Another remaining challenge is that the isolation and structural elucidation of novel natural products remains laborious and time intensive. This problem is further compounded by the fact that a single organism often has multiple secondary metabolite gene clusters and within each cluster a suite of analogs may be produced. This abundance of candidate molecules for isolation requires a means of prioritization for resource allocation.

The following chapters of this thesis present research that addresses these two challenges. The first area covered is expanding access to analogs of structurally complex natural products that are intractable to traditional medicinal chemistry approaches. In particular the work focuses on the apoptolidins and ammocidins, a family of glycosylated polyketides. Experimental results on the characterization and manipulation of the biosynthetic pathways are reported as well as methods for obtaining new analogs from mutant strains of the producing organisms in Chapter 2. Chapter 3 presents experimental results on the biological activity of these analogs and their development as chemical probes.

The second area covered in this thesis concerns the development of methods that extend traditional methods of natural product discovery via multiplexing bioactivity guided fractionation assays. This topic is addressed in Chapter 4 and experimental results are presented on the development of the platform and several applications thereof. Importantly this method provides a preliminary annotation of secondary metabolites present in a crude extract and prioritizes compounds for isolation that are most probable to be bioactive based on correlation analysis of time dependent chemoinformatic and bioassay data.

References

1 Ebbell, B. *The Papyrus Ebers: The Greatest Egyptian Medical Document* (Levin and Munksgaard-Ejnar Munksgaard, 1937).

2 Conrad, L. I. *The Western Medical Tradition: 800 BC to AD 1800* (Cambridge University Press, 1995).

3 Atanasov, A. G. *et al.* Discovery and resupply of pharmacologically active plant-derived natural products: A review *Biotechnology Advances* **33**, 1582-1614, (2015).

4 Sneader, W. *Drug Discovery* (John Wiley & Sons, Ltd, 2006).

5 Niemann, A. Ueber eine neue organische Base in den Cocablättern. *Archiv der Pharmazie* **153**, 129-155, (1860).

6 Abel, J. J. On the blood-pressure-raising constituent of the suprarenal capsule. *Bulletin of Johns Hopkins Hospital* **8**, 151-157, (1897).

7 Blevins, S. M. & Bronze, M. S. Robert Koch and the 'golden age' of bacteriology. *International Journal of Infectious Diseases* **14**, e744-e751, (2010).

8 Henkel, T., Brunne, R. M., Müller, H. & Reichel, F. Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angewandte Chemie International Edition* **38**, 643-647, (1999).

9 Lipinski, C. A. Lead and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today* **1**, 337-341, (2004).

10 Lee, M. L. & Schneider, G. Scaffold architecture and pharmacophoric properties of natural products and trade drugs: application in the design of natural product-based combinatorial libraries. *Journal of Combinatorial Chemistry* **3**, 284-289, (2001).

11 Dobson, C. M. Chemical space and biology. *Nature* **432**, 824-828, (2004).

12 Feher, M. & Schmidt, J. M. Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *Journal of Chemical Information and Computer Sciences* **43**, 218-227, (2003).

13 Breinbauer, R., Vetter, I. R. & Waldmann, H. From protein domains to drug candidates-natural products as guiding principles in the design and synthesis of compound libraries. *Angewandte Chemie International Edition* **41**, 2879-2890, (2002).

14 Kellenberger, E., Hofmann, A. & Quinn, R. J. Similar interactions of natural products with biosynthetic enzymes and therapeutic targets could explain why nature produces such a large proportion of existing drugs. *Natural Product Reports* **28**, 1483-1492, (2011).

15 Taipale, J. *et al.* Effects of oncogenic mutations in Smoothened and Patched can be reversed by cyclopamine. *Nature* **406**, 1005, (2000).

16 Sekulic, A. *et al.* Efficacy and safety of vismodegib in advanced basal-cell carcinoma. *The New England Journal of Medicine* **366**, 2171-2179, (2012).

17 Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of Natural Products* **75**, 311-335, (2012).

18 Fowler , V. G. J. *et al.* Daptomycin versus standard therapy for bacteremia and endocarditis caused by Staphylococcus aureus. *New England Journal of Medicine* **355**, 653-665, (2006).

19 Wender, P. A. & Hardman, C. T. Scalable synthesis of bryostatin 1 and analogs, adjuvant leads against latent HIV. *Science* **358**, 218-223, (2017).

CHAPTER 2

Biosynthetic Engineering of Glycosylated Macrolides

Discovery of the Apoptolidins and Ammocidins

Macrolides comprise a structurally and pharmacologically diverse class of natural products, selectively addressing an equally diverse array of cellular targets, such as immunosuppressive signaling (FK-506, FKB12 calcineurin)[1], splicing factors (SF3b, pladienolide)[2], and ion channels (avermectin, glutamate gated chloride channel)[3]. Moreover, variations within structural families can possess entirely different targeting properties, which has motivated significant activity towards generating modified macrolides, via both chemical synthesis and biosynthetic pathway engineering[4]. Correspondingly, the impact of macrolides has been realized in both the clinic and as chemical biological tools for uncovering new insights into cell biology.

Of interest to this work are a family of macrolides generated by actinomycetes represented by the apoptolidins and ammocidins, which are 20/21-membered glycosylated macrolides possessing potent and selective cell targeting properties.

Isolation of the apoptolidins

Apoptolidin A was originally discovered in a screen for selective inducers of apoptosis in E1A oncogene transformed cell lines and was isolated from *Nocardiopsis* sp. FU40 by Seto

et al[5]. Further evaluation of the compound against the NCI-60 collection revealed that activity was greatest in cell lines that do not exhibit the Warburg effect instead relying on oxidative phosphorylation[6].

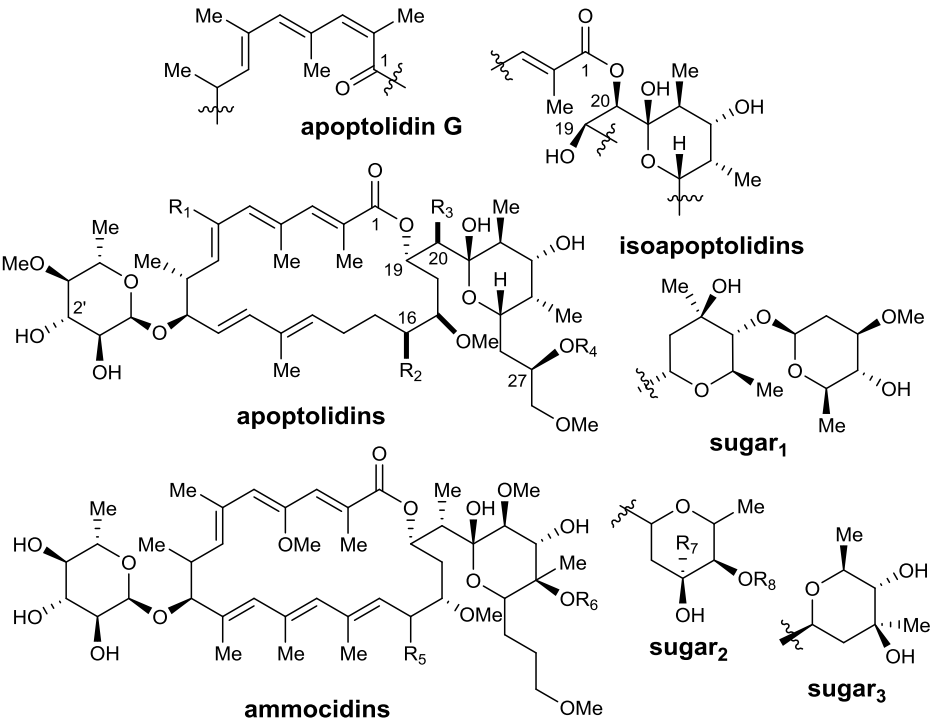
Since this initial report, several more apoptolidin natural products have been isolated from *Nocardiopsis* sp. FU40 as minor metabolites, now designated apoptolidins B–G. Apoptolidin B and C both lack hydroxyl groups at C16 and both C16 and C18, respectively, but are otherwise identical to the primary metabolite, apoptolidin A[7]. Apoptolidin D lacks a C6 methyl group in contrast to other apoptolidins[8]. Apoptolidins E and F are deoxygenated at C16 and C18 but differ in the type of deoxy sugar at C9 and C27[9].

Also, apoptolidin has been reported to undergo a ring expansion to give a 21 membered isomeric macrolactone named isoapoptolidin A[10]. Similar isomeric forms have subsequently been identified for apoptolidins B and D as well. A second type of isomerization is also possible as the C2-C3 double bond geometry can by inverted by light exposure resulting in the configurational isomer apoptolidin G[11]. Finally, it should be noted that Mahmud group recently isolated 2'O and 3'O succinylated forms of the apoptolidins from an *Amycolatopsis* species and proposed that this modification is useful for secretion of the compounds as they are readily labile to hydrolysis[12].

Isolation of the ammocidins

The structurally related natural product, ammocidin A, was isolated from *Saccharothrix* sp. AJ9571 by Hayakawa et al. in a screen against Ras oncogene transformed cell lines[13]. The same group has since reported the isolation of several additional related minor metabolites, the ammocidins B-D[14]. Structures of the apoptolidins and ammocidins

summarized in **Figure 2.1**.



apoptolidin A (**1**, R_1 = Me; R_2 = R_3 = OH; 2' = β-OH; R_4 = $sugar_1$)
apoptolidin B (**2**, R_1 = Me; R_2 = H; R_3 = OH; 2' = β-OH; R_4 = $sugar_1$)
apoptolidin C (**3**, R_1 = Me; R_2 = R_3 = H; 2' = β-OH; R_4 = $sugar_1$)
apoptolidin D (**4**, R_1 = H; R_2 = R_3 = OH; 2' = β-OH; R_4 = $sugar_1$)
apoptolidin E (**5**, R_1 = Me; R_2 = R_3 = H; 2' = α-OH; R_4 = $sugar_1$)
apoptolidin F (**6**, R_1 = Me; R_2 = R_3 = H; 2' = α-OH; R_4 = H)
apoptolidin G (**7**, R_1 = Me; R_2 = R_3 = OH; 2' = β-OH; R_4 = $sugar_1$)
apoptolidin H (**8**, R_1 = Me; R_2 = R_3 = OH; 2' = β-OH; R_4 = H)
isoapoptolidin A (**9**, R_1 = Me; R_2 = R_3 = OH; R_4 = $sugar_1$)
isoapoptolidin B (**10**, R_1 = Me; R_2 = H R_3 = OH; R_4 = $sugar_1$)
isoapoptolidin D (**11**, R_1 = H; R_2 = R_3 = OH; R_4 = $sugar_1$)
ammocidin A (**12**, R_5 = OH; R_6 = $sugar_2$; R_7 = H; R_8 = $sugar_3$)
ammocidin B (**13**, R_5 = OH; R_6 = $sugar_2$; R_7 = Me; R_8 = $sugar_3$)
ammocidin C (**14**, R_5 = H; R_6 = $sugar_2$; R_7 = Me; R_8 = H)
ammocidin D (**15**, R_5 = OH; R_6 = H)

Figure 2.1: Structures of the apoptolidins and ammocidins

Synthetic Investigations of the Apoptolidins and Ammocidins

Due to the striking selectivity exhibited by apoptolidin, several groups undertook its total synthesis, thereby generating synthetic and semisynthetic analogs. The first total synthesis of apoptolidin A was reported by Nicolaou and co-workers in 2003[15]. Utilizing late stage glycosylations and macrolactonization in their strategy, they were also able to prepare three novel apoptolidin analogues including C27-hydroxy apoptolidin A and two macrolactones missing the pyran fragment.

The Sulikowski laboratory has developed routes to multiple apoptolidin agylcones[16,17] and in collaboration with the Bachman lab have demonstrated that synthetically prepared aglycones can be glycosylated upon incubation with *Nocardiopsis* strains in which the polyketide synthase has been knocked down[18]. Intriguingly, glycosylation of apoptolidinone D was realized at C27 only suggesting that C9 glycosylation may precede cyclization. The Sulikowski group has also prepared hexaacetylated apoptolidin A, which allowed for deglycosylation at C27 by treatment with methanolic hydrochloric acid to provide acylated apoptolidin monosaccharide and are currently finalizing a synthetic route to apoptolidinone C. Significant progress has also been made on synthetic routes to access the ammocidins[19].

The Wender group has also generated many analogs via semi-synthesis to explore structure activity relationships. Notable examples include a truncated apoptolidin by acid hydrolysis of the C27 disaccharide accompanied by elimination of water in the pyran moiety; functionalizing the various hydroxyl groups of apoptolidin using careful protecting group manipulation or peptide catalysis; and production of a few analogues of the macrocycle

and δ-lactone apoptolidin fragment by way of oxidative cleavage of apoptolidin A[20,21].

Biosynthetic Investigations of the Apoptolidins and Ammocidins

Sequencing of the apoptolidin and ammocidin producers

We have reported the sequencing and initial characterization of the apoptolidin biosynthetic gene cluster which contains a modular type I polyketide synthase, a biosynthetic cassette for a unique (R)-2-methoxymalonyl-ACP starter unit, genes necessary for the biosynthesis and attachment of the C9 monosaccharide and the C27 disaccharide, as well as tailoring oxidases and methyltransferases[22]. The organization of the gene cluster is shown in **Figure 2.2**

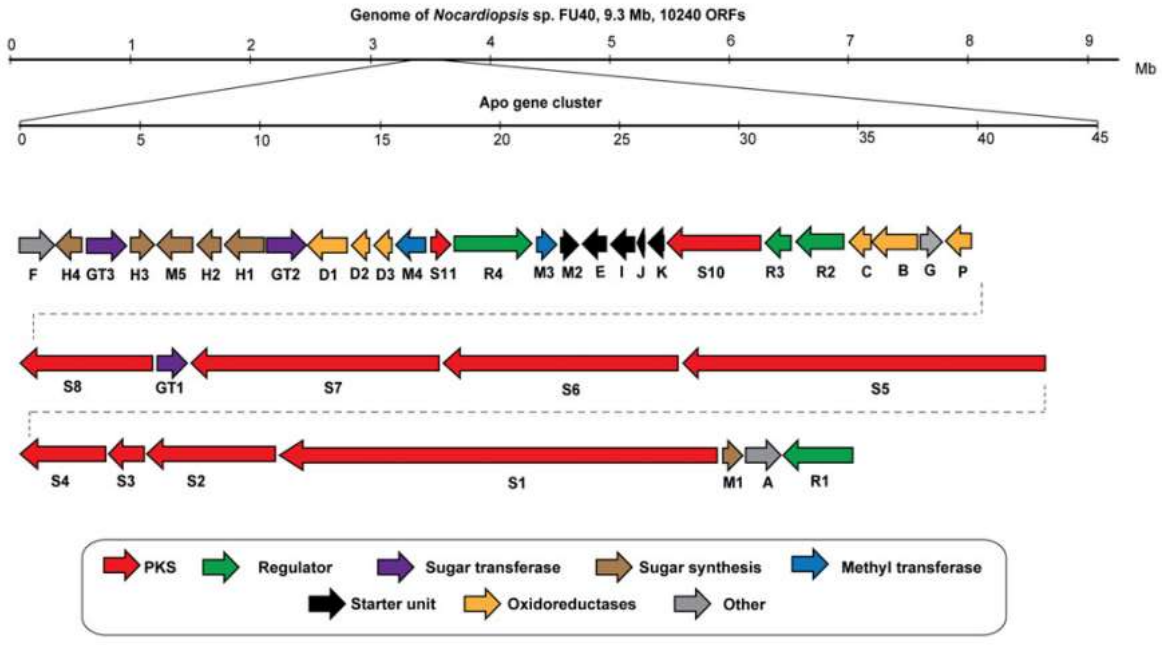


Figure 2.2: The apoptolidin biosynthetic gene cluster

Similarly, we have sequenced the genome of *Saccharothrix* sp. AJ9571, the ammocidin producer using Illumina MiSeq. Due to the repetitive sequence of PKS modules, assembly of the entire cluster on a single contig was not successful given the short read lengths from MiSeq. However, we could identify non-overlapping contigs with portions of the ammocidin gene cluster consistent with expected domain sequence. High overall sequence similarity to the apoptolidin gene cluster was observed including conserved arrangement of polyketide synthase modules, genes for methoxymalonate starter unit biosynthesis, and similar tailoring enzymes including oxidases and glycosyltransferases. Sequence similarity for genes approaches 90% which, along with segments of conserved gene organization, suggests the possibility of a common evolutionary origin. While the ammocidin and apoptolidin polyketide synthases are superficially similar, many predicted structural and biosynthetic differences are expected. A representation of the predicted domain sequences for apoptolidin and ammocidin *seco* acid biosynthesis is shown in **Figure 2.3** with expected differences highlighted in red.

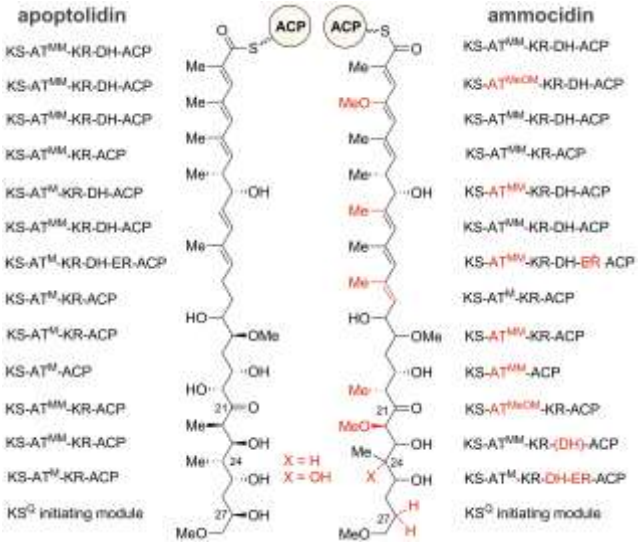


Figure 2.3: Comparison of the apoptolidin and predicted ammocidin functional *seco* acid PKS biosynthesis. Predicted differences in domain organization highlighted in red.

Organization of the apoptolidin polyketide synthase

Apoptolidin is the product of a type I polyketide synthase (PKS). Type 1 PKS's are multi-domain megasynthases that construct the core of many polyketide natural products[23]. Chain elongation is achieved through a non-iterative series of decarboxylative Claisen condensations and β-reductions. A typical catalytic cycle consists of the following steps, first acyltransferase (AT) domains transfer the elongation units (various C2 substituted malonate derivatives) from a molecule of coenzyme A (CoA) to the phosphopantetheine group of an acyl carrier protein domain (ACP) for activation. The activated malonate unit is transferred to a cysteine residue by the ketosynthase (KS) domain which catalyzes the decarboxylative Claisen condensation onto the downstream malonyl ACP. The oxidation state of the resulting β-ketoacyl moiety is then determined by the action of ketoreductase (KR), dehydratase (DH) and enoylreductase (ER) domains to the appropriate oxidation state dependent upon the total number of present active reductive domains. When chain elongation is complete, a thioesterase (TE) domain then catalyzes the cleavage and/or cyclization to the final polyketide product.

In *Nocardiopsis* sp. FU40, the apoptolidin polyketide core is assembled by thirteen modules across eight PKS proteins, ApoS1 to ApoS8. The proposed organization of the modules and domains responsible for the biosynthesis of the apoptolidin polyketide *seco* acid is shown in **Figure 2.4**. Polyketide biosynthesis is often initiated by a module containing a KS domain with a cysteine to glutamate codon modification mutating the essential active site cysteine (C193Q) involved in the transthioesterification reaction that precedes KS mediated condensation. ApoS1 was identified as possessing the likely initiating module as it contains the characteristic sequence of this type of decarboxylative

KSQ loading module.

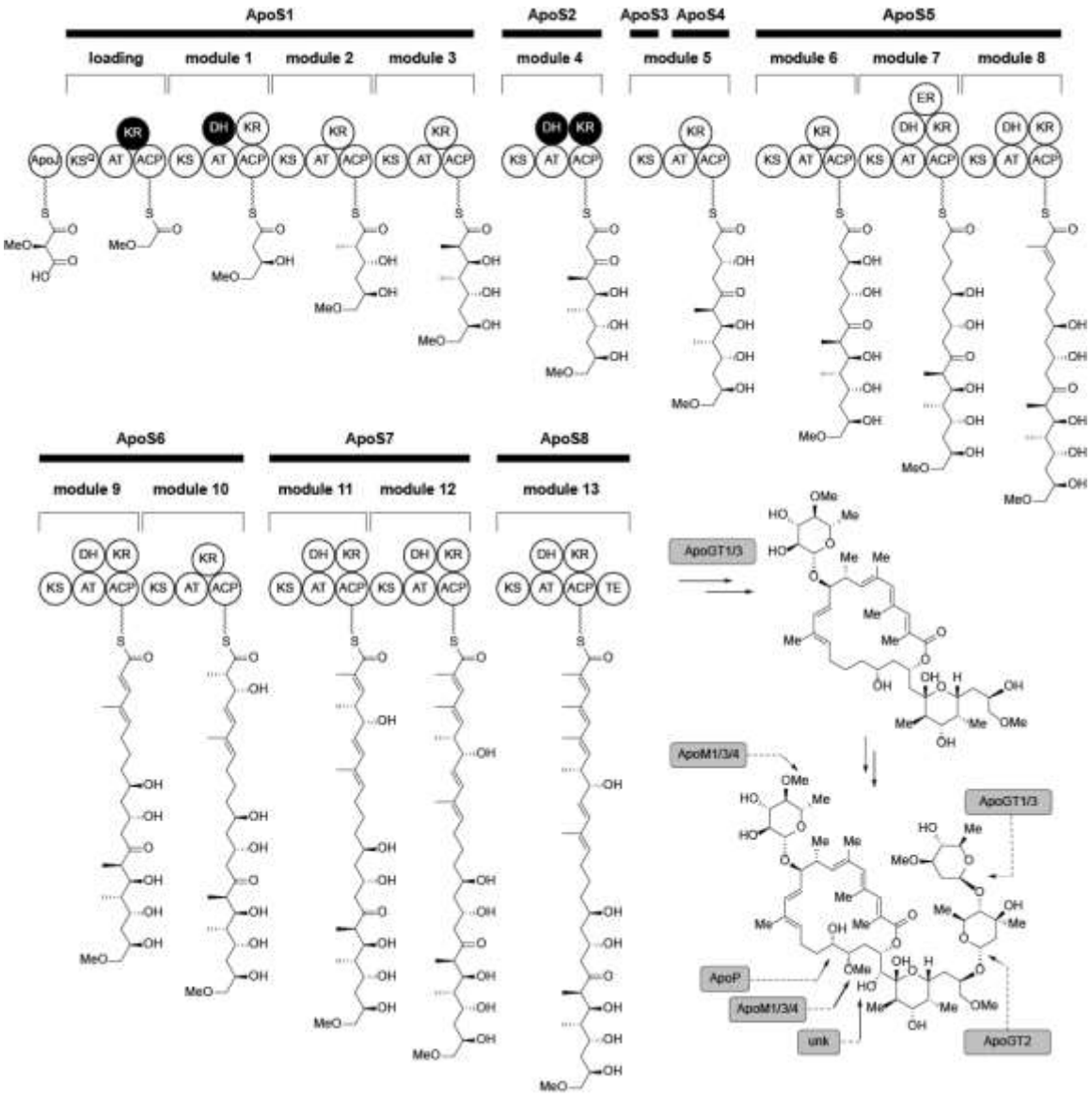


Figure 2.4: The apoptolidin biosynthetic pathway. The polyketide synthase consists of a loading module and 12 *seco* acid extension modules. Inactive domains are in black. Upon cleavage by the terminal thioesterase (TE) domain several tailoring reactions occur including glycosylations, oxidations, and methylations.

The next 12 extension modules are proposed to be contained on ApoS1 to S7 and are ordered according to the predicted arrangement of required domains necessary for the apoptolidin polyketide backbone. ApoS8 is the likely terminating module as it contains a TE domain, required for *seco* acid release from the megasynthase machinery. Two additional open reading frames where identified with sequence homology to polyketide synthase domains. One contains an incomplete module sequence 'KS-AT-KR' while the other encodes a free-standing thioesterase protein that may be important in hydrolytic release and/or macrocyclization.

Description of deoxysugar biosynthesis genes

The apoptolidins are decorated by three sugar residues. The C9 hydroxyl is appended with 6-deoxy-4-O-L-methyl glucose or alternatively with 4-O-L-methyl-L-rhamnose dependent on growth conditions. The sugars have been demonstrated to be important for selective cytotoxicity[24]. The necessary genes for sugar biosynthesis are contained in a single sub-cluster along with genes for two glycosyl transferases at the beginning portion of the whole apoptolidin cluster. A third glycosyl transferase is located within the putative PKS encoding sub-cluster. BLAST analysis of the sugar biosynthetic genes revealed five genes encoding putative enzymes suitable for the biosynthesis of the mono and disaccharide sugars. The biosynthesis of these sugars starts with a common intermediate, NDP-L-4-keto-3-deoxyglucose, which is converted into NDP-D-oleandrose by the action of Apo1,2,3/4 or TDP-L-olivomycose by ApoH1/2 ApoM2, a 3,5-epimerase, and Apo3/4. The 3,5-epimerase is not contained within the cluster but a suitable open reading frame designated *nsf5842* encoding a 3,5-epimerase was found in the *Nocardiopsis* genome.

Finally, NDP-L-4-keto-3-deoxyglucose is likely converted into NDP-D-6-deoxy-4-O-methyl-D-glucose by ApoH3 or ApoH4.

Description of tailoring O-methyltransferases

The apoptolidin gene cluster contains two putative O-methyltransferases encoded by *apoM1* and *apoM3*. ApoM1 is most closely related to sugar O-methyltransferases, indicating a likely role in the methylation of the 4' position of the C9 monosaccharide, leaving ApoM3 as the likely candidate for methylation of the C17 hydroxyl group.

Description of tailoring oxidases

Analysis of the pattern of oxidation in the polyketide backbone of apoptolidin and the isolation of apoptolidins B and C suggest that C16 and C20 hydroxyl groups are the result of post-PKS oxygenation reactions. The two proteins thought to be responsible for these C-H bond oxidative reactions are ApoP, a P450 with 38/56% identity/similarity to EryK from the erythromycin biosynthetic gene cluster of *Saccharopolyspora erythraea*[25] and 43/61% similarity to TylI of tylosin biosynthesis in *Streptomyces fradiae*[26]. The second proposed oxidase is ApoD1-3, an apparent three component Rieske non-heme iron oxygenase similar to oxygenases involved in the oxidation of aryl and biaryl functional groups.

Genetic Manipulation of the Apoptolidin Gene Cluster

apoS8

Containing a *seco* acid releasing TE domain, ApoS8 is predicted to contain the last extension module of the apoptolidin PKS. To confirm the identification of the apoptolidin cluster a site directed disruption mutant, FU40-*apoS8*::*aac*(3)IV was generated. This strain was cultured in fermentation medium, extracted with ethyl acetate and analyzed via HPLC-MS for the production of known apoptolidins. Notably no apoptolidin A or any known apoptolidin analogs were produced by this strain confirming the identification of the cluster. However, pulse feeding of an apoptolidin monosaccharide largely restored production of the apoptolidin A which validated the targeted gene disruption. Importantly this result also demonstrates that transcriptionally downstream genes involved in biosynthesis and glycosylation of apoptolidin monosaccharide remain intact[22].

apoP

ApoP was identified as a cytochrome P450 monooxygenase based on sequence analysis. A selective *apoP* gene replacement mutant, FU40-*apoP*::*aac*(3)IV was generated and the fermentation yielded apoptolidin analogs 18 Da less than parent apoptolidins according to HPLC-MS analysis, indicating that the ApoP is a hydroxylating P450 and that targeted replacement of the corresponding gene did not generate polar effects downstream[22].

apoGT2

The apoptolidin gene cluster contains three genes encoding for glycosyl transferases (*apoGT1-3*). To date only a selective knockout of *apoGT2* has been produced. The analysis of the crude extracts from the resulting mutant revealed that production of apoptolidins A-

G had been eliminated. Instead, the organism produced two compounds with the signature UV absorbance of apoptolidin (λmax= 232, 333), with a *m/z* of 858.5[27].

apoJ and apoK

The translated sequences of the five gene (*R*)-2-methoxymalonyl-acyl carrier protein (MeOM-ACP) contiguous gene cassette (*apoK-M2*) has broad sequence identity to translated *fkbG-K* which encode the biosynthesis of this extender unit from the FK520 gene cluster in *S. hygroscopicus*[28], with *ApoJ* displaying 53% identity to the acyl carrier protein for the extender unit. However, while MeOM-ACP has been reported to function as an extender unit by intercepting polyketide synthase in *trans*, it has never been reported as a chain initiator. To confirm priming of the apoptolidin polyketide synthase with MeOM-ACP, we endeavored to disrupt *apoJ* in this cluster. We employed two step PCR-targeting replacement, in which genes were first replaced by antibiotic resistance markers in cosmids containing the apo gene cluster and subsequently transferred into *Nocardiopsis* to select for double crossover events.

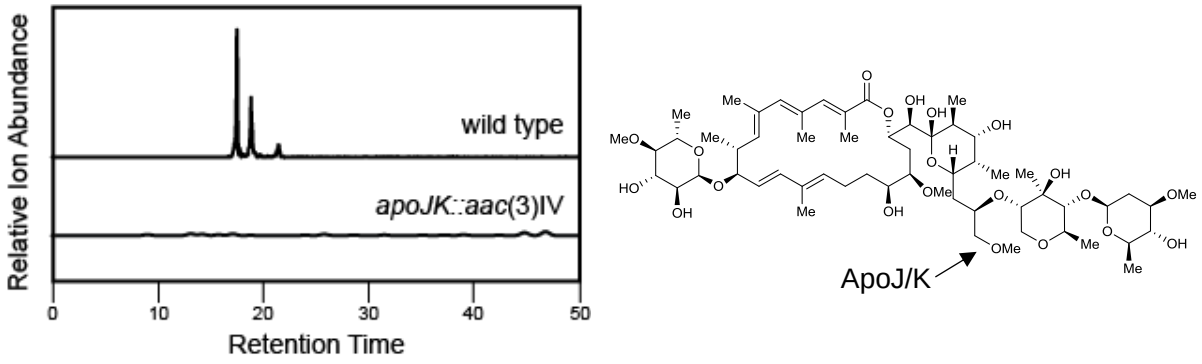


Figure 2.5: Selected ion trace (*m/z* 1146.5) for targeted disruption mutant of *apoJK*

Attempts to replace the 291 bp *apoJ* did not yield recombinant clones, however double gene replacement of *apoJ* and *apoK* was successful, resulting in a mutant strain *Nocardiopsis* sp. FU40-*apoJK*::*aac*(3)IV. The translated *apoK* gene possesses 65% identity with FkbK, an oxidase responsible for dehydrogenation of glyceryl-ACP en route to hydroxymalonate, and its deletion should have no effect of down-stream apoptolidin biosynthesis. HPLC-MS analysis of extracts of fermentation cultures of this strain demonstrated a complete abolishment of production of all apoptolidins, supporting the hypothesis of (*R*)-2-methoxymalonyl-ACP biosynthetic initiation (**Figure 2.5**).

apoD1 and apoD2

The fully elaborated apoptolidins require two aglycone C-H bond oxidations to generate the hydroxyl groups at C16 and C20. We annotated two potential enzymatic systems for oxidation within the apo gene cluster. In addition to *apoP*, putatively encoding a translated cytochrome p450, preliminary annotation of the gene cluster also revealed the presence of a potential Rieske non-heme dioxygenase, ApoD1-3. This family of enzymes consist of three components, a two-component heterodimer (apoD1/2) with an α–subunit containing the active site, a β–subunit, which serves a purely structural role, and an interacting redox accessory ferredoxin subunit (apoD3). The putative apoD1/D2 system is most similar to arene dioxygenase enzymes (e.g. napthalene dioxygenase, 42% identity) that process arene substrates to cis-dihydrodiol products. Notably, these systems can function as monooxygenases hydroxylating aliphatic C-H bonds via a proposed HO-Fe(V)=O intermediate[29]. Given the absence of other oxidative enzymes in to apo gene cluster, we hypothesized that the dioxygenase like system is recruited in one of the tailoring oxidations.

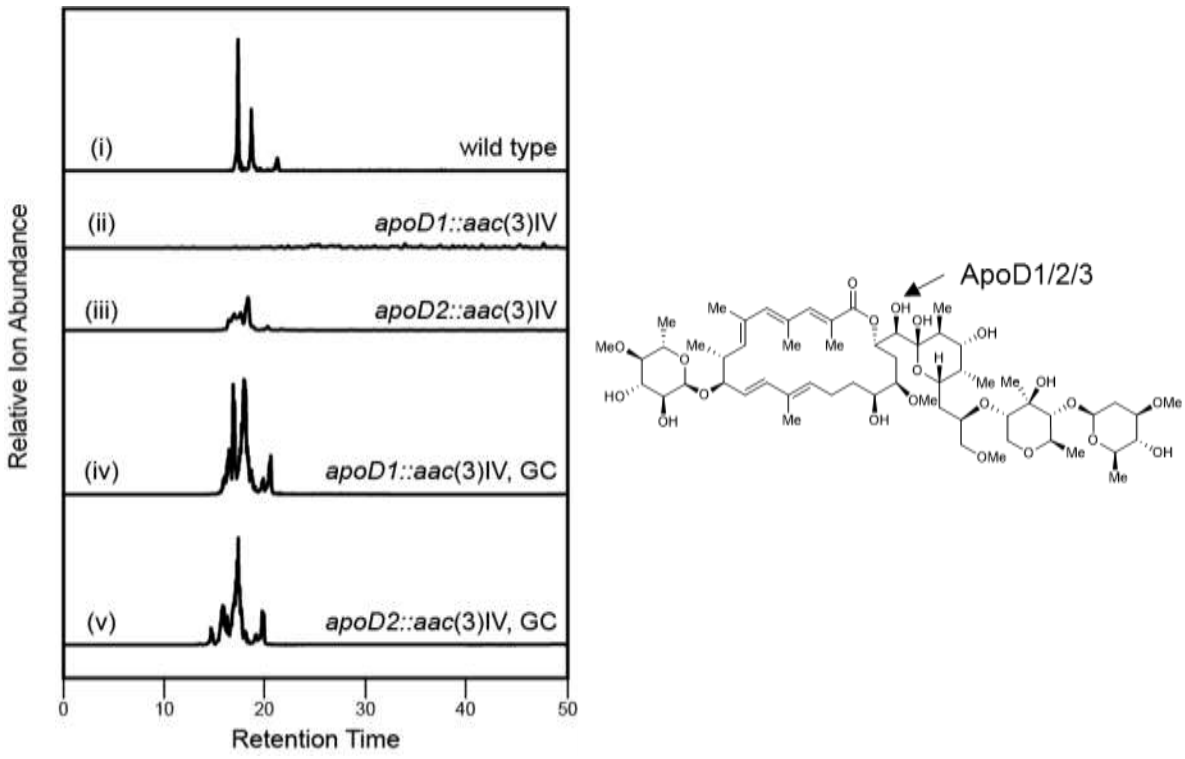


Figure 2.6: Selected ion trace (*m/z* 1146.5) for targeted disruption mutants of *apoD1/D2* (α and β subunits of Rieske oxidase) normalized to wild type counts ($4e^6$). (i) wild type (ii) *apoD1*disruption mutant (iii) *apoD2* disruption mutant (iv) *apoD1* mutant genetically complemented with *apoD1* (v) *apoD2* mutant genetically complemented with *apoD2*

To test the involvement of apoD1/D2 in the biosynthesis of apoptolidins, *apoD1* or *apoD2* was replaced in cosmids with *aac*(3)IV. Modified cosmids were then conjugatively transformed into *Nocardiopsis* sp. FU40 and double cross-over mutants were selected in apramycin containing medium. **Figure 2.6** shows the selected ion trace for *m/z* 1146.5, the $[M + NH_4]$ adduct of apoptolidin A, from crude extracts from the wild type producer and disrupted mutant strains. Analysis of the HPLC-MS chromatograms for the *apoD1* mutant (**Figure 2.6**, trace ii) revealed that the production of apoptolidin A and isomers was abolished. In contrast, the *apoD2* mutant (**Figure 2.6**, trace iii) was still able to produce apoptolidin A, albeit at a drastically reduced level compared to the wildtype producer. Taken together these results confirm that the dioxygenase is a critical and necessary

component of the apoptolidin gene cluster.

Genetic conformation of targeted disruptions

To confirm that the *apoJK* and *apoD1* were double crossover mutants Southern blot analysis was performed (**Figure 2.7**). In both cases the expected bands were observed confirming the genetic location of the disruptions.

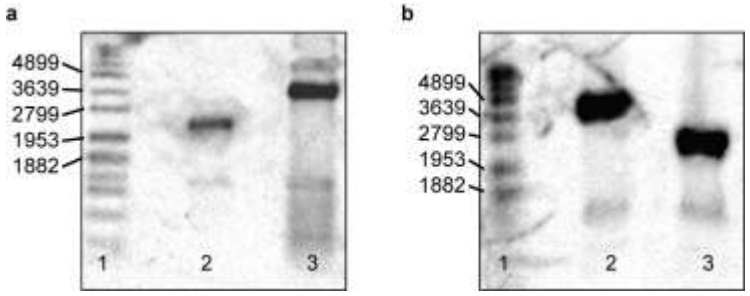


Figure 2.7: Southern blot analysis for (a) *apoJK* and (b) *apoD1*. Lane 1 is DNA ladder, Lane 2 is the disrupted DNA fragment and lane 3 is the wildtype DNA fragment. Expected size of *apoJK* wildtype and disrupted fragments are 2.6 Kb and 3.6 Kb respectively. Expected size of *apoD1* wildtype and disrupted fragments are 4.0 Kb and 2.7 Kb respectively.

Additionally, in order to confirm that *apoD1/D2* results were due to gene specific inactivation and not polar effects, a modified pSET152 plasmid was constructed containing a hygromycin resistance gene and a copy of either *apoD1* or *apoD2* and transformed into the respective mutants via conjugation. In both cases HPLC-MS analysis of fermentation cultures confirmed production of apoptolidin was restored (**Figure 2.6**, trace iv and v). Since genetic complementation and Southern analysis (**Figure 2.7b**) indicate a clean functional gene replacement, it is possible apoD1/D2 are essential for polyketide

maturation and that this oxidase may act on the *seco* acid, or an intermediate, prior to aglycone formation and release. Finally the *apoJ/K* mutant was successfully chemically complemented and results are shown in **Section 2.9**

Isolation of Analogs from Mutant Strains

Isolation of parent compounds

Shake flask fermentation cultures of *Nocardiopsis* sp. FU40 yields 60-80 mg per liter of apoptolidin A. Purified compound is readily obtained by extraction of culture supernatant with ethyl acetate, concentration and then running a series of LH-20 and HPLC columns. Similarly, 20-30 mg per liter of ammocidin A can be obtained from fermentation cultures of *Saccharothrix* sp. AJ9571**.** These relatively high yields provide an excellent and renewable source of material for probe development and biological assays.

Isolation of apoptolidin H

As noted above a new *m/z* of 858.5 was observed to be produced by the *apoGT2* mutant strain. Isolation and NMR of this metabolite confirmed that the ion corresponds to the ammonium adduct of the apoptolidin polyketide core, glycosylated at C9 only and confirms the role of ApoGT2 in the installation of the disaccharide moiety at the C27 hydroxyl group. This new analog was assigned the name apoptolidin H and is obtained in yields of approximately 40 mg per liter of fermentation cultures.

Isolation of 16-deoxyapoptolidin

As noted previously, the *apoP* mutant strain produced a metabolite with the apoptolidin

chromophore and with *m/z* of 1116.5 which is consistent with the loss of a hydroxyl group. To confirm the location where the hydroxyl group was lost, the analog was purified from 1 liter fermentation cultures. During isolation isomerization occurred with a chromophore consistent with ring expansion via acyl migration between C19 and C20 making C16 the likely candidate for hydroxyl loss. The purified compound was characterized by 1D and 2D NMR (**Table 2.2**). Initial analysis of the 1H NMR spectrum showed similarities with the apoptolidin B spectrum and preliminary proton shift assignments for C1-14, C21-C28, and Sugar Protons were made by comparison with reported shifts. HSQC experiments were used to establish 1H-13C connections and revealed the presence of an additional methine, consistent with the loss of a hydroxyl group. O-methylations were established by HMBC experiments. COSY experiments were used to identify spin systems and to 'walk in' from C12 to C16 and from C20 to C16 using 1H-1H correlations. Additional TOCSY experiments confirmed the loss of the hydroxyl at C16.

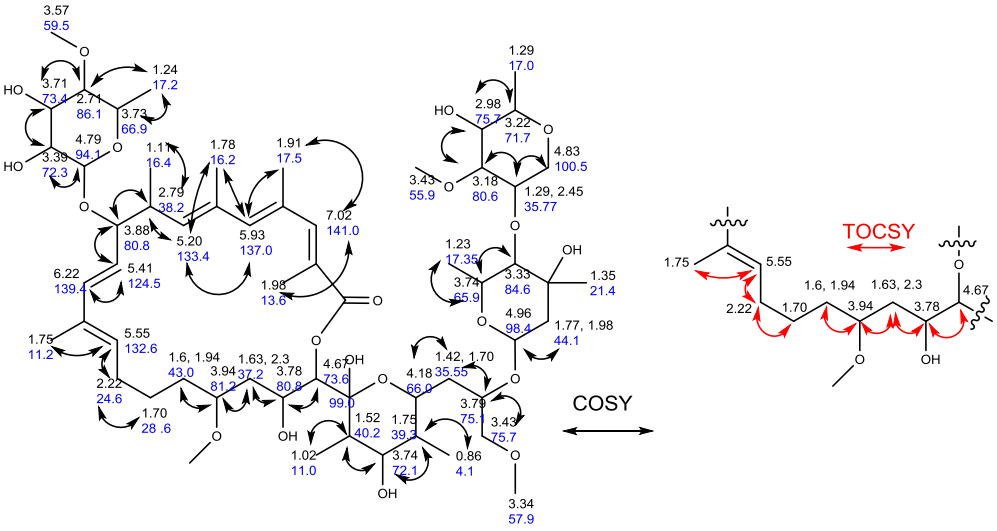


Figure 2.8. 16-deoxyapoptolidin shift assignments and correlations. ^{1}H shifts in black, ^{13}C shifts in blue.

Position	1H NMR δH	13C NMR δC	1H-13C HSQC NMR δC	1H-13C HMBC NMR δC	1H-1H TOCSY NMR δH
1		174			
2		128.8			
3	7.02		141	13.5, 17.3, 128.8, 136.9, 174	1.98, 5.93
4		132.2			
5	5.93		137	16.2, 17.4, 131.9, 133.4, 141.0	1.78, 1.91, 1.98, 5.2, 7.02
6		136.5			
7	5.2		133.4	16.2, 38.2, 80.8, 137	1.78, 2.79, 5.93
8	2.79		38.2	16.2, 80.8, 133.1	1.11, 3.88, 5.2
9	3.88		80.8	16.2, 38.2, 94.2, 133.4, 139.4	2.79, 5.41
10	5.41		124.5	38.2, 80.8, 133.2	3.88, 6.22
11	6.22		139.4	1.11, 80.8, 132.6	3.88, 5.41, 5.55
12					
13	5.55		132.6	11.1, 24.6, 28.6, 139.4	1.76, 2.2
14	~2.22		24.6	28.6, 81.2, 132.6	1.7, 5.55
15	1.68, 1.72		28.6		3.94
16	1.6, 1.94		43		2.3, 3.78, 4.66
17	3.94		81.2		
18	1.63, 2.3		37.2		
19	3.78		81		
20	4.66		73.6		
21		99			
22	1.52		40.2		1.02, 3.74
23	3.73		72.1		0.86, 1.02, 1.52
24	1.75		39.4		3.73, 4.18
25	4.18		66		0.86, 1.42, 1.69, 3.43, 3.79,
26	1.42, 1.69		35.5		3.43, 3.79, 4.18
27	3.79		75.1		1.42, 1.69, 3.79
28	3.43		75.7		
2-Me	1.98		13.6	128.8, 141, 174	5.93, 7.02
4-Me	1.91		17.4	132.2, 136.5, 137.0, 141.0	5.93, 7.02
6-Me	1.78		16.2	132.2, 136.5, 137	1.11, 5.2, 5.93
8-Me	1.11		16.3	38.2, 80.8, 133.4	2.79, 3.88, 5.2
12-Me	1.76		11.2	139.4	5.55
17-OMe	3.32		55.8	81	
22-Me	1.02		11	40.2, 72.1, 99	
24-Me	0.86		4.2	39.4, 66, 72.1	
28-OMe	3.35		57.9	75.7	
1'	4.79		94.2	66.6, 73.4, 80.8	3.39
2'	3.39		72.3	73.4	3.71, 4.79
3'	3.71		73.4		2.71, 3.39
4'	2.71		86.1	59.5, 66.6,	3.73
5'	3.74		66.6	72.3, 86.1	1.24
6'	1.24		17.2	66.6, 86.1	2.71, 3.74
4'-OMe	3.57		59.5	86.1	
1"	4.96		98.4	65.9, 71.6	1.77, 1.98
2"	1.77, 1.98		44.1	98.4	4.96
3"		71.6			
4"	3.33		84.6	17.35, 21.3, 65.9, 100.5	3.73
5"	3.73		65.9		1.23
6"	1.23		17.35	21.3, 65.9	3.73
3"-Me	1.35		21.3	44.1, 71.6, 84.6	
1'"	4.83		100.5	35.8, 84.6	1.29, 2.45
2'"	1.29, 2.45		35.8	80.5, 100.5	3.18, 4.83
3'"	3.18		80.5	75.7	1.29, 2.45, 2.98
4'"	2.98		75.7	17, 75.7, 80.5	3.18, 3.22
5'"	3.22		71.7	17, 100.5	1.29, 2.98
6'"	1.29		17	71.7, 75.7, 100.5	2.98, 3.22
3'"-OMe	3.43		55.9	35.7, 80.5	

Table 2.1: Shift assignments for 16-deoxyapoptolidin analog

Precursor Directed Biosynthetic Studies

Synthetic thioesters of polyketide chain initiating and extension building blocks and intermediates have been shown to load KS active site cysteine thiols domains in vitro[30], and have also been successfully employed in chemical complementation studies of blocked polyketide biosynthetic pathways[31]. Chemical rescue of the *apoJK* knockout strain was performed with the *N*-acetylcysteamine, (SNAC)[32] thioester of (*R*)-2-methoxymalonate (MeOMe-SNAC). Initial studies of synthesized MeOMe-SNAC added to early stage *Nocardiopsis* growth cultures substantially restored apoptolidin A biosynthesis. Optimal incorporation efficiency was determined by evaluating pulsed dosing schedules in which 60 mg were fed in equal portions over the seven day fermentation. It was determined that pulsed supplementation of MeOMe-SNAC with 50 μL aliquots of 8 mg/mL in DMSO starting on the 2nd day of seed culture yielded the best results with production of apoptolidin A restored to near wildtype levels (**Figure 2.10**, trace iii). These results are consistent with the site of acylation of the first polyketide synthase protein ApoS1 being the active site cysteine in the second KS domain. The efficient biochemical bypass of MeOM-ACP with its SNAC-activated analog motivated us to survey the ability of the polyketide synthase loading module to accept synthetic precursor analogs. A series of synthetic starter units (**Figure 2.9**) were generated by EDCI coupling of corresponding carboxylic acids and N-acetylcysteamine: acetic acid **3**, 3-hydroxy-4-methoxybutanioc acid **4**, 3-azido-4-methoxybutanioc acid **5**, butanoic acid **6**, pentanoic acid **7**, and but-3-ynoic acid **8**. Synthetic starters units were pulse fed into cultures beginning on the second day of seed culture and through the fifth day of production cultures. Extracts were collected and analyzed by HPLC-MS. Incorporation of **4**, which embodies the nascent *seco* acid

after second module elongation, was also successful albeit with lower efficiency (**Figure 2.10**, trace iv). However no incorporation of SNAC thioesters was observed for starter units that contained unnatural stereochemistry, chain lengths, or otherwise unnatural functional group modifications.

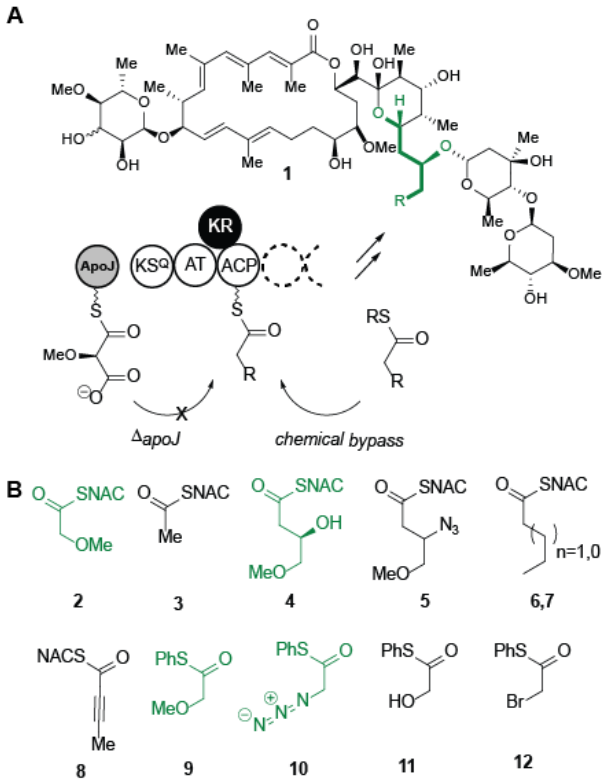


Figure 2.9: Proposed chemical bypass strategy and structures of tested synthetic starter units. (A). Scheme for ApoJ/K analysis and biochemical bypass. (B) Structures of starter unit surrogates synthesized and incubated with *apoJK* null strain. HPLC-MS analysis of targeted replacement of *apoJK* and chemical complementation with natural loading units, advanced diketide, and starter unit analogs. Successfully incorporated units are highlighted in green

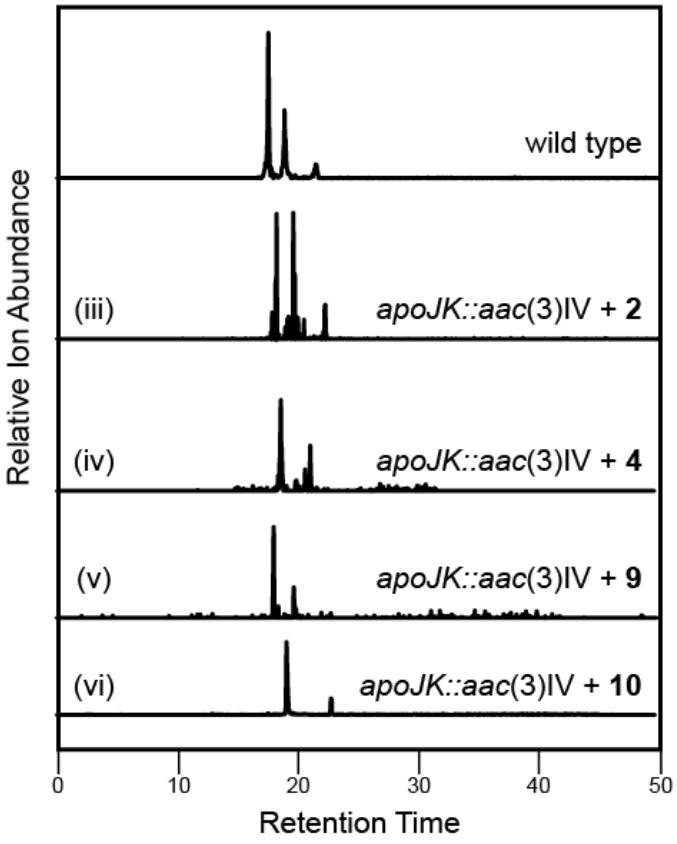


Figure 2.10: Selected ion traces from chemical complementation experiments**.** HPLC-MS analysis of targeted deletion of *apoJK* and chemical complementation with natural loading units, advanced diketide, and starter unit analogs.

Recently, thiophenol thioesters were demonstrated to load the KS domain pikromycin pathway[34]. Thiophenol esters of methoxyacetic acid **9**, 2-azidoacetic acid **10**, 2-hydroxyacetic acid **11** and 2-bromoacetic acid **12** were prepared and supplemented into *Nocardiopsis* sp. FU40-*apoJK*::*aac*(3)IV cultures to evaluate their efficacy of biochemical rescue. HPLC-MS analysis of extracts revealed that the thiophenol ester of methoxyacetate successfully complemented the *apoJK* deletion (**Figure 2.10**, trace v) and that 2-azidoacetic acid bypass resulted in a new metabolite of *m/z* =1139 with a UV max of 290 and 330 nm consistent with the properties of other apoptolidins suggesting successful incorporation of the azide at C28 (**Figure 2.10**, trace vi). To confirm the identity of this

newly observed compound, we performed collision induced dissociation studies on the putative azide containing analog. The three sugars of the apoptolidins provide diagnostic fragments for identification. Loss of the C27 sugar with dehydration yielded fragments with *m/z* 306 and 835. Subsequent loss of the C9 sugar resulted in fragments with *m/z* of 163 and 675. Finally, observation of fragmentation across the C22-C23 bond and dehydration (*m/z* of 457) confirmed incorporation of the azide at the terminal end of apoptolidin (**Spectrum 2.5**).

Conclusions

In summary, genetic manipulation of the apoptolidin gene cluster has produced six mutant strains to date. Of these two stains (FU40-*apoP*::*aac*(3)IV and FU40-*apoGT2*::*aac*(3)IV) are beneficial for generating analogs via fermentation, one has found use as a biotransformation reagent (FU40-*apoS8*::*aac*(3)IV), and one has been used in precursor directed biosynthesis studies (FU40-*apoJK*::*aac*(3)IV). Studies on the biological activity of apoptolidin and ammocidin analogs are discussed in chapters 3 and 4.

Successful chemical complementation of the FU40-*apoJK*::*aac*(3)IV strain with a synthetically prepare SNAC activated starting unit resorted production of apoptolidin A and motivated us to generate and test a small library of synthetic precursors for incorporation by the rest of the intact polyketide synthase. Overall little substrate flexibility was observed, however, using a thiophenol activated azide starter unit was successful in generating an azido-apoptolidin analog. The observation that thiophenol activated starter units may be more efficiently incorporated by the polyketide synthase

presents the potential for a broader range of structurally distinct starters units to be incorporated than initial experiments indicated.

Our study of the biosynthesis and evaluation of biological activity of the ammocidins remains comparably immature leaving several intriguing avenues of research open. The development of accurate long read sequencing techniques, such as those by Pacific Biosciences, affords an opportunity for compete closure of the *Saccharothrix* sp AJ9571 genome and unambiguous analysis of the ammocidin gene cluster which would aid in the design of targeted mutagenesis experiments. While initial attempts at transformation via electroporation or conjugation with selection via nalidixic acid of the ammocidin producer were unsuccessful, the development of high efficiency transformation of via 'conjugational lagoons' in the Bachmann lab warrant further explorations in this area.

Experimental Methods

Chemicals, strains, media, and general DNA techniques

Oligonucleotide primers were synthesized by Sigma Aldrich, plasmid DNA was purified by Qiagen miniprep kit (Catalog #27106), genomic DNA was purified using a Promega Wizard Kit (Catalog #A1120), Agarose gel purified DNA fragments were isolated using Qiagen gel extraction kit (Catalog #28706. *E. coli* BW25113 was used to maintain cosmids and plasmids for lambda Red targeted gene replacement. *E. coli* S17-1 was used for conjugative transformation of *Nocardiopsis*. Strains were grown on LB agar or in LB medium supplemented with antibiotics necessary for plasmid maintenance. Optical densities of cultures were determined absorbance at 600 nM. All chemical were purchased

from Sigma Aldrich (St, Louis MO) unless otherwise specified.

ApoS8-F:	5’ CGGGTGCGGGCCGCCGCCGGGCTGGCCGCGCTGGACGTGATTCCGGGGATCCGTCGAC 3’
ApoS8-R:	5’ GCTCACGCAGGACTCCTTTCAACGTTTTTCAAAACCTTATGTAGGCTGGAGCTGCTTC 3’
ApoP-F:	5’ GCGTAAGGTTTTGAAAAACGTTGAAAGGAGTCCTGCGTGATTCCGGGGATCCGTCGACC 3’
ApoP-R:	5’ACGGCGGCCGGAAGCGAACCTCCCGGCCGCCGTTCCTCATGTAGGCTGGAGCTGCTTC 3’
ApoGT2-F:	5’-ATTTTCTCCCGATCCGATGGCGAAAGGCTCACGCGCGTGATTCCGGGGATCCGTCGACC-3’
ApoGT2-R:	5’-GAGCTACCCCCCCTGTGGCGGCTCCGGCCCGAAACCCTTATGTAGGCTGGAGCTGCTTC-3’
ApoJK-F:	5’ AGAAGGTCCGCCAGCAGCTGAAGCTGGCCCGGATCATGATTCCGGGGATCCGTCGACC 3’
ApoJK-R:	5’ TCCAGGAAGACGACCAGTTCCATCGCGAACAGCGAGGACTGTAGGCTGGAGCTGCTTC 3’
ApoD1-F:	5’ CATCCGAACCTCGACCCGCGACGGGAGCCCGCGAAGATGTTCCGGGGATCCGTCGACC 3’
ApoD1-R:	5’ ATCGCACGACAGGCCCGGCCTCGGTGGAGGGCACCGCCGTGTAGGCTGGAGCTGCTTC 3’
ApoD2-F:	5’ CGTGGCCCCGCGCCCCTGAACCTGAGGGGGTTTCGCATGATTCCGGGGATCCGTCGACC 3’
ApoD2-R:	5’ ACTGGCGCGCTTAGGTAGGCTTGGAGCTGGGCGCTGCCCTGTAGGCTGGAGCTGCTTC 3’
Del-up:	5’ GGTCGACGGATCCCCGGAAT 3’
Del-down:	5’ GAAGCAGCTCCAGCCTACA 3’
GIB-D1GS:	5’ GTCGATCAGATCGGCGTACATGGATCCTACCAACCGGCAC 3’
GIB-D1GE:	5’ GCCAACGGCCGGGGCTAGAGCGCATATGCTCGAGAAG 3’
GIB-D1VUP:	5’ GTGCCGGTTGGTAGGATCCATGTACGCCGATCTGATCGAC 3’
GIB-D1VDN:	5’ CTTCTCGAGCATATGCGCTCTAGCCCCGGCCGTTGGC 3’
GIB-D2GS:	5’ GACGGCCCGGTGAATGCGGTCATGGATCCTACCAACCGGCA 3’
GIB-D2GE:	5’ CTCGGCCTGTTCTTCTGAAGCGCATATGCTCGAGAAG 3’
GIB-D2VUP:	5’ GTGCCGGTTGGTAGGATCCATGACCGCATTCACCGGGCCCGTC 3’
GIB-D2VDN:	5’ CTTCTCGAGCATATGCGCTTCAGAAGAACAGGCCGAG 3’
Hygbcheck-F:	5’ GAAGGCGTTGAGATGCAGTT 3’
Hygbcheck-R:	5’ GATTCGGATGATTCCTACGC 3’

Table 2.1: Primers used for genetic manipulation and of the apoptolidin gene cluster

Fermentation of Nocardiopsis and Saccharothrix sp.

50 μl from a glycerol stock was streaked on Bennet's agar (yeast extract 1 g, beef extract 1 g, NZ amine A 2 g, glucose 10 g, and agar 20 g per 1 L H_2O, pH 7.2) and grown for 3-5 days at 30°C until sporulation occurred. A loop full of aerial mycelia was used to inoculate 5 mL seed cultures (soluble starch 10 g, molasses 10 g, peptone 10 g, and beef extract 10 g per 1 L of H_2O, pH 7.2) in 50 mL flacon tubes and grown for 3 days at 30 °C. Seed cultures were then transferred to 50 mL of fermentation media (glycerol 20 g, molasses 10 g, Casamino acids 5 g, peptone 1 g, $CaCO_3$ 4 g per 1 L H_2O, pH 7.2) in 250 Erlenmeyer flasks and grown for 7 days at 30 °C with shaking.

Extraction of fermentation cultures

Fermentation cultures were extracted as previously described[22].

PCR based gene-targeting

Cosmids containing the relevant biosynthetic genes were disrupted using the previously described lambda red methodology[22,35].

Conjugative transformation of Nocardiopsis

E. coli strain S17-1 containing the modified cosmids was grown to an OD of 0.2 and 250 μL were mixed with approximately $10e^8$ spores of *Nocardiopsis* sp. FU40 that had been heat shocked for 10 minutes at 50 °C. The mixed cells were then plated on dried MS agar plates and incubated for 30 min at 37 °C and then for 16 hrs at 30 °C at which point the plates were overlaid with 12.5 μg/mL of nalidixic acid and 80 μg/mL of apramycin. After 1 week individual resistant colonies were picked. Fermentation cultures of resistant colonies were carried out and crude extracts were analyzed by HPLC-MS.

PCR and Southern analysis of mutant strains

Disruption cassettes were prepared as described previously for *apoS8*, *apoP*, and *apoGT2*[22]. Primers ApoJK-F and ApoJK-R, ApoD1-F and ApoD1-R, and ApoD2-F and ApoD2-R (**Table 2.1**) were used to amplify the apramycin resistance cassette from pIJ773 for each respective disrupted biosynthetic gene. Putative disrupted cosmids and transformed strains were initially confirmed using primers Del-up and Del-down (**Table 2.1**) to amplify the *acc*(3)IV gene.

Southern analysis, digoxigenin probe labeling, hybridization, and detection was performed according to the suppliers protocols (Roche DIG High prime DNA Labeling and Detection kit, Sigma)

Genetic complementation of gene disrupted mutants

Primers GIB-D1GS and GIB-D1GE were used to amplify *apoD1* from cosmid DNA isolated from *E. coli*. Primers GIB-D1VUP and GIB-D1VDN were used to amplify linear plasmid DNA from a derivative of pSET152 in which the apramycin resistance was replaced with hygromycin resistance and *ermE** was inserted before the multiple cloning site (**Table 2.1**). PCR products were gel purified and isolated and then joined using a Gibson assembly cloning kit (New England Biolabs, Catalog #E5510S) per manufactures protocol. The resulting plasmids phygD1 and phygD2 were transformed into electrocompetent *E. coli* S17-1. Plasmids were isolated from hygromycin resistant cultures and checked by amplification of the *aph*(7'')-la gene using primers Hygbcheck-F and Hygbcheck-R. Plasmids were transformed into *Nocardiopsis* using conjugation.

Chemical complementation and precursor directed biosynthesis of gene replacement mutants

Synthetic starters units were then dissolved in DMSO at 8 mg/mL and then daily pulse fed into cultures at in 50 μL aliquots beginning on the second day of seed culture and through the fifth day of production cultures resulting in a total amount of ~30 mg of starting unit being added

Isolation of apoptolidin A and H and ammocidin A

Compounds were isolated according to previously reported procedures[22].

Isolation and structural analysis of 16-dexoyapoptolidin

A total of 1 L of fermentation culture was centrifuged at 3700 g for 30 min. The supernatant was extracted with 3 volumes of ethyl acetate, combined, and concentrated to 200 mg total solids per mL of methanol by rotary evaporation. The concentrated extract was then fractionated with Sephadex LH-20 Resin (GE Healthcare Bio-Sciences) with methanol as the eluent at 4 °C in the dark. Fractions containing *m/z* 1116 were combined and further purified by preparative HPLC (Waters, XBridge C18 Prep, 5 uM) (10 mL/min, 0 min – 1 min: 75% solution A, 25% solution B, 70 min: 100% solution B) (Solution A = 95:5, H_2O:MeCN, 10 mM NH_4OAc; Solution B: 5:95 H_2O:MeCN, 10mM NH_4OAc). Fractions containing 16-dexoyapoptolidin were evaporated using a Genevac HT-2 system at 30 °C and reconstituted in methanol-d_4 for structural characterization by nuclear magnetic resonance.

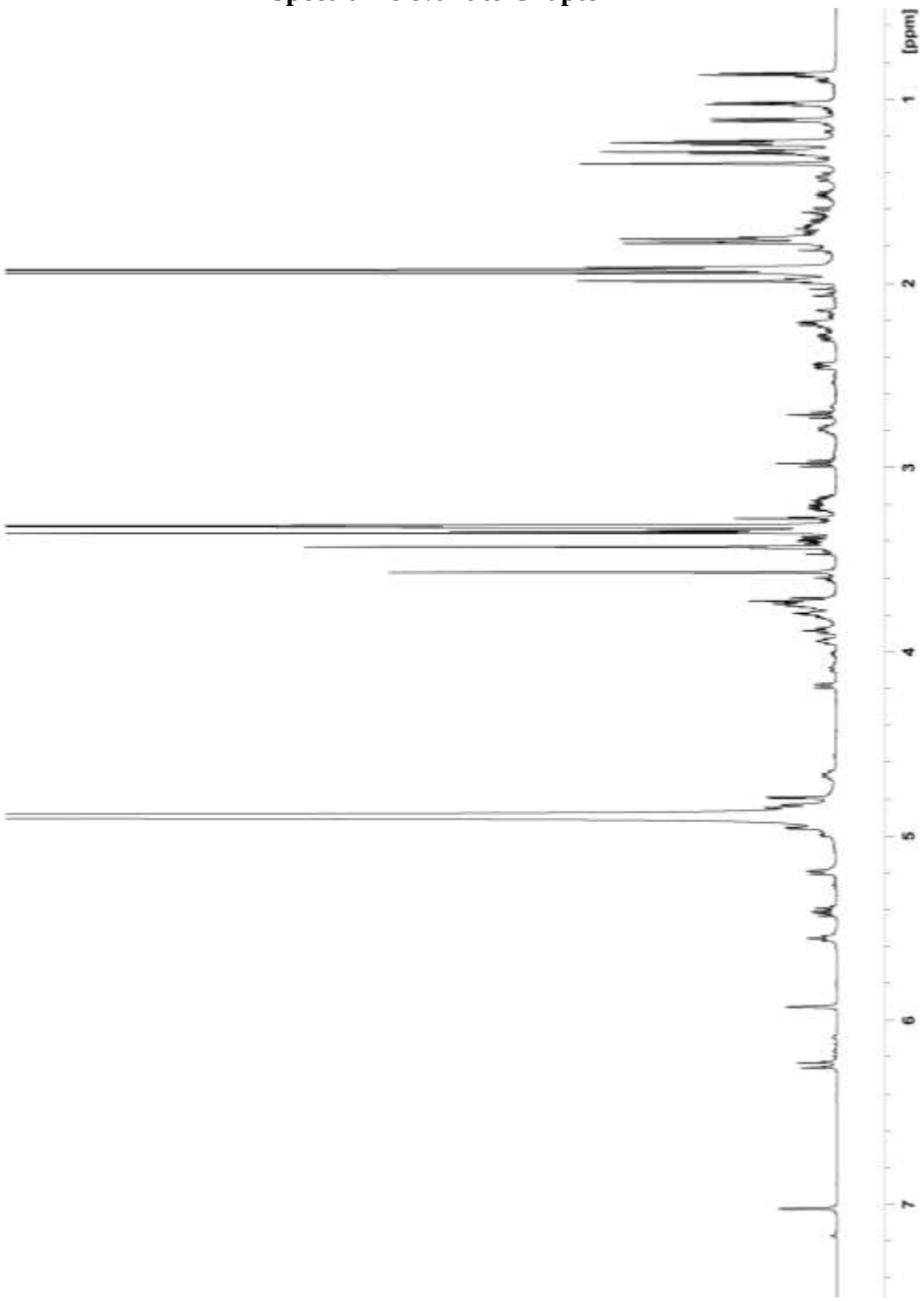
NMR and MS parameters

Proton nuclear magnetic resonance (^{1}H NMR) spectra and carbon-13 (^{13}C NMR) spectra were recorded on a 600 MHz spectrometer at ambient temperature. ^{1}H NMR data are reported as δ values relative to residual non-deuterated solvent δ 3.31 ppm from methanol-d_4. For ^{13}C spectra, chemical shifts are reported relative to the δ 49.00 ppm resonance of methanol-d_4.

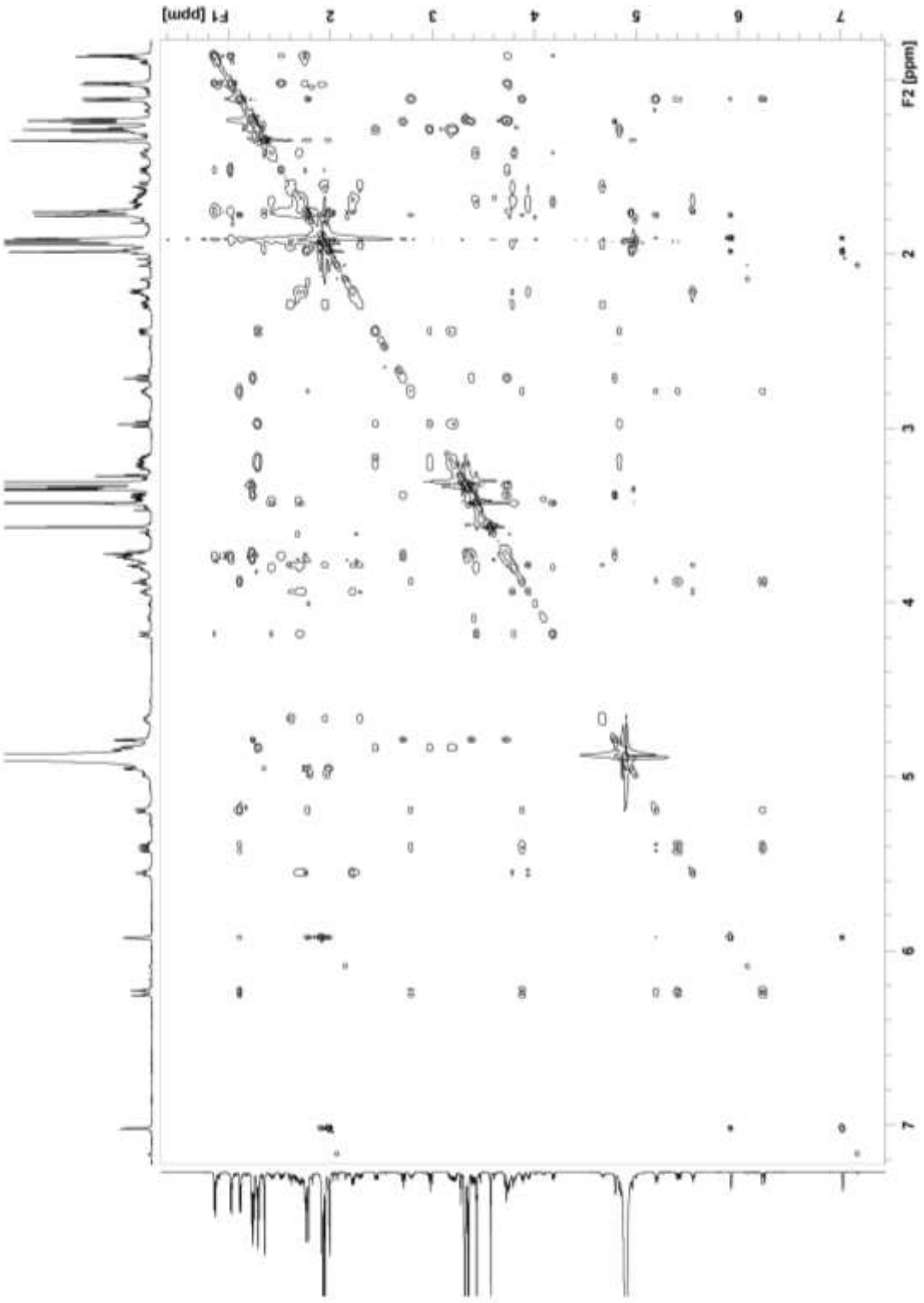
Mass spectrometry was performed by using a TSQ Triple Quadrapole mass spectrometer equipped with an electrospray ionization source and Surveyor PDA Plus detector. For positive ion mode, the following settings were used: capillary temperature was 270 °C; spray voltage 4.2 kV; spray current 30 mA; capillary voltage 35 V; tube lens 119 V;

skimmer offset 15 V. For negative ion mode, capillary temperature 270 °C; spray voltage 30 kV; spray current 20 mA; capillary voltage 35 V; tube lens 119 V; skimmer offset 15 V. For MS/MS fragmentation the collision gas pressure was 1 mTorr with collision energy of 20 V and energy ramp of 0 eV.

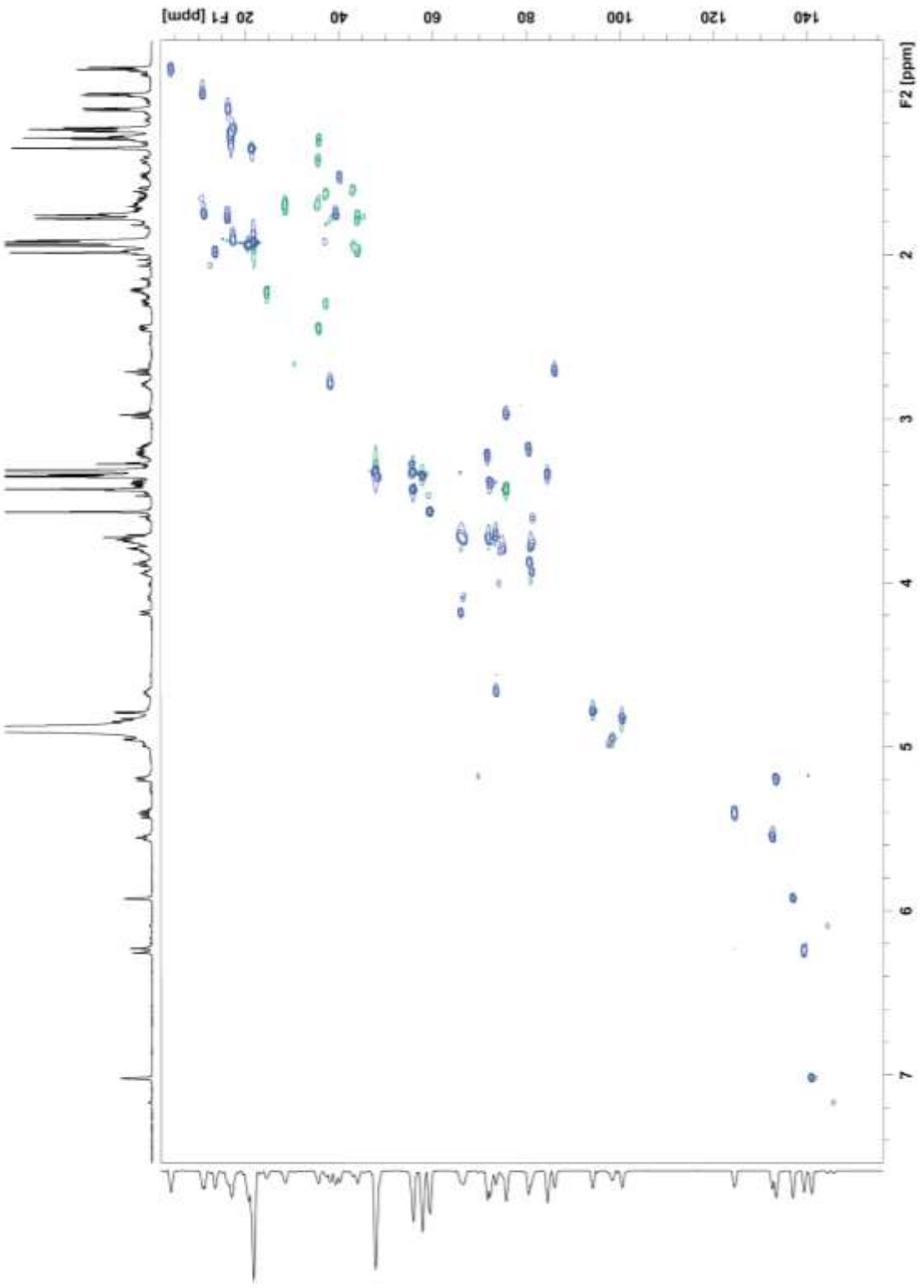
Spectra Relevant to Chapter 2



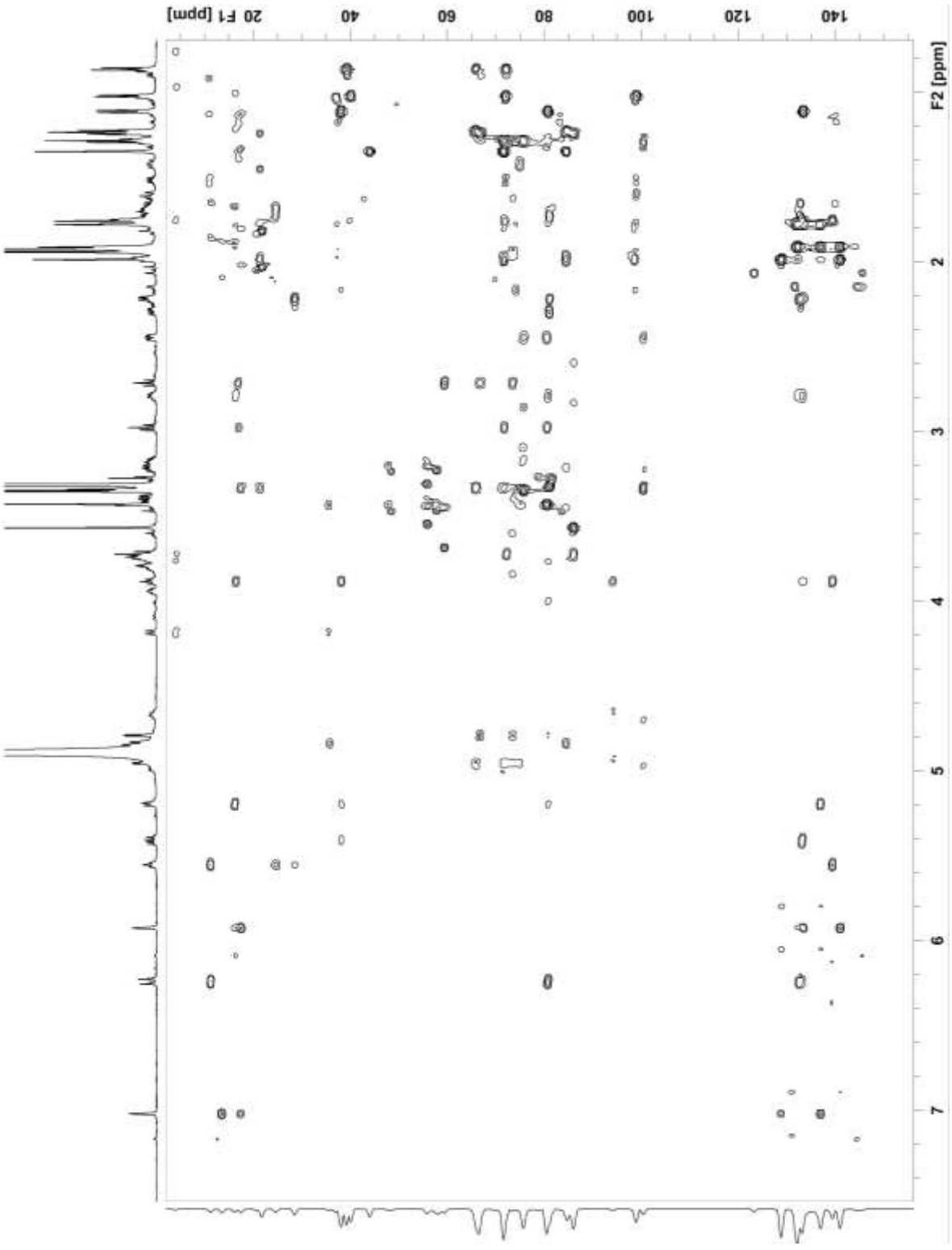
Spectrum 2.1: ^{1}H NMR, 600 MHz, 16-deoxyapoptolidin analog in methanol-d_4



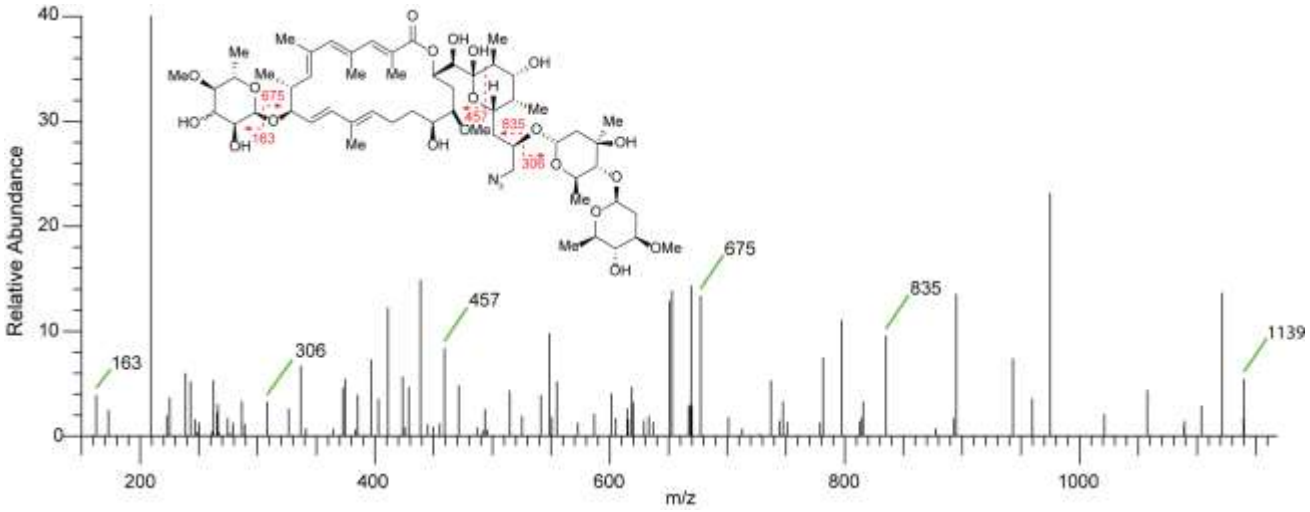
Spectrum 2.2:2D TOCSY NMR, 600 MHz, 16-deoxyapoptolidin analog in methanol-d_4



Spectrum 2.3:2D HSQC NMR, 600 MHz, 16-deoxyapoptolidin analog in methanol-d_4



Spectrum 2.4: 2D HMBC NMR, 600 MHz, 16-deoxyapoptolidin analog in methanol-d_4



Spectrum 2.5: MS/MS fragmentation of azido-apoptolidin analog. CID fragmentation spectrum of parent *m/z* 1139 in positive ion mode.

References

1 Liu, J. *et al.* Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66**, 807-815, (1991).

2 Kotake, Y. *et al.* Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nature Chemical Biology.* **3**, 570-575, (2007).

3 Pong, S.-S. & Wang, C. C. Avermectin B1 modulation of γ-aminobutyric acid receptors in rat brain membranes. *Journal of Neurochemistry.* **38**, 375-379, (1982).

4 Woerly, E. M., Roy, J. & Burke, M. D. Synthesis of most polyene natural product motifs using just 12 building blocks and one coupling reaction. *Nature Chemistry* **6**, 484, (2014).

5 Kim, J. W., Adachi, H., Shin-ya, K., Hayakawa, Y. & Seto, H. Apoptolidin, a new apoptosis inducer in transformed cells from Nocardiopsis sp. *Journal of Antibiotics.* **50**, 628-630, (1997).

6 Salomon, A. R., Voehringer, D. W., Herzenberg, L. A. & Khosla, C. Understanding and exploiting the mechanistic basis for selectivity of polyketide inhibitors of F0F1-ATPase. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14766-14771, (2000).

7 Wender, P. A., Sukopp, M. & Longcore, K. Apoptolidins B and C: isolation, structure determination, and biological activity. *Organic Letters* **7**, 3025-3028, (2005).

8 Wender, P. A. & Longcore, K. E. Isolation, structure determination, and anti-cancer activity of apoptolidin D. *Organic Letters* **9**, 691-694, (2007).

9 Wender, P. A. & Longcore, K. E. Apoptolidins E and F, new glycosylated macrolactones isolated from Nocardiopsis sp. *Organic Letters* **11**, 5474-5477, (2009).

10 Wender, P. A., Gulledge, A. V., Jankowski, O. D. & Seto, H. Isoapoptolidin: structure and activity of the ring-expanded isomer of apoptolidin. *Organic Letters* **4**, 3819-3822, (2002).

11 Bachmann, B. O. *et al.* Light-induced isomerization of apoptolidin a leads to inversion of C2-C3 double bond geometry. *Organic Letters* **12**, 2944-2947, (2010).

12 Sheng, Y. *et al.* Succinylated Apoptolidins from Amycolatopsis sp. ICBB 8242. *Organic Letters* **17**, 2526-2529, (2015).

13 Murakami, R. *et al.* Ammocidin, a new apoptosis inducer in Ras-dependent cells from Saccharothrix sp. I. Production, isolation and biological activity. *Journal of Antibiotics.* **54**, 710-713, (2001).

14 Murakami, R. *et al.* Ammocidins B, C and D, new cytotoxic 20-membered macrolides from Saccharothrix sp. AJ9571. *The Journal of Antibiotics* **62**, 123-127, (2009).

15 Nicolaou, K. C. *et al.* Total synthesis of apoptolidin: completion of the synthesis and analogue synthesis and evaluation. *Journal of the American Chemical Society* **125**, 15443-15454, (2003).

16 Wu, B., Liu, Q. & Sulikowski, G. A. Total synthesis of apoptolidinone. *Angewandte Chemie International Edition* **43**, 6673-6675, (2004).

17 Pennington, J. D., Williams, H. J., Salomon, A. R. & Sulikowski, G. A. Toward a stable apoptolidin derivative: identification of isoapoptolidin and selective deglycosylation of apoptolidin. *Organic Letters* **4**, 3823-3825, (2002).

18 Ghidu, V. P. *et al.* Combined chemical and biosynthetic route to access a new apoptolidin congener. *Organic Letters* **11**, 3032-3034, (2009).

19 Chau, S. T., Hayakawa, Y. & Sulikowski, G. A. 18O assisted analysis of a gamma,delta-epoxyketone cyclization: synthesis of the C16-C28 fragment of ammocidin D. *Organic Letters* **13**, 756-759, (2011).

20 Wender, P. A., Jankowski, O. D., Tabet, E. A. & Seto, H. Toward a structure-activity relationship for apoptolidin: selective functionalization of the hydroxyl group array. *Organic Letters* **5**, 487-490, (2003).

21 Lewis, C. A., Longcore, K. E., Miller, S. J. & Wender, P. A. An approach to the site-selective diversification of apoptolidin A with peptide-based catalysts. *Journal of Natural Products* **72**, 1864-1869, (2009).

22 Du, Y. *et al.* Biosynthesis of the apoptolidins in Nocardiopsis sp. FU40. *Tetrahedron* **67**, 6568-6575, (2011).

23 Staunton, J. & Weissman, K. J. Polyketide biosynthesis: a millennium review. *Natural Product Reports* **18**, 380-416, (2001).

24 Salomon, A. R., Voehringer, D. W., Herzenberg, L. A. & Khosla, C. Apoptolidin, a selective cytotoxic agent, is an inhibitor of F0F1-ATPase. *Chemisty & Biology.* **8**, 71-80, (2001).

25 Summers, R. G. *et al.* Sequencing and mutagenesis of genes from the erythromycin biosynthetic gene cluster of Saccharopolyspora erythraea that are involved in L-mycarose and D-desosamine production. *Microbiology* **143 (Pt 10)**, 3251-3262, (1997).

26 Merson-Davies, L. A. & Cundliffe, E. Analysis of five tylosin biosynthetic genes from the tyllBA region of the Streptomyces fradiae genome. *Molecular Microbiology* **13**, 349-355, (1994).

27 DeGuire, S. M. *et al.* Fluorescent probes of the apoptolidins and their utility in cellular localization studies. *Angewandte Chemie International Edition.* **54**, 961-964, (2015).

28 Wu, K., Chung, L., Revill, W. P., Katz, L. & Reeves, C. D. The FK520 gene cluster of Streptomyces hygroscopicus var. ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* **251**, 81-90, (2000).

29 Bassan, A., Blomberg, M. R. A. & Siegbahn, P. E. M. A theoretical study of the cis-dihydroxylation mechanism in naphthalene 1,2-dioxygenase. *Journal of Biological Inorganic Chemistry* **9**, 439-452, (2004).

30 Tsukamoto, N. *et al.* 6-Deoxyerythronolide B synthase 1 is specifically acylated by a diketide intermediate at the β-ketoacyl-acyl carrier protein synthase domain of module 2. *Biochemistry* **35**, 15244-15248, (1996).

31 Franke, J. & Hertweck, C. Biomimetic thioesters as probes for enzymatic assembly lines: Synthesis, aApplications, and challenges. *Cell Chemical Biology* **23**, 1179-1192, (2016).

32 Jacobsen, J. R., Hutchinson, C. R., Cane, D. E. & Khosla, C. Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase. *Science* **277**, 367-369, (1997).

33 Kubota, T. *et al.* Determination of the cryptic stereochemistry of the first PKS chain-extension step in ansamitocin biosynthesis by Actinosynnema pretiosum. *Chembiochem:* **7**, 1221-1225, (2006).

34 Hansen, D. A., Koch, A. A. & Sherman, D. H. Substrate controlled divergence in polyketide synthase catalysis. *Journal of the American Chemical Society* **137**, 3735-3738, (2015).

35 Gust, B. *et al.* Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces. *Advances in Applied Microbiology* **54**, 107-128, (2004).

CHAPTER 3

Biological Evaluation of Apoptolidin and Ammocidin Analogs*

Cytotoxicity of Apoptolidins and Ammocidins

Apoptolidin A is readily obtained by fermentation of the actinomycete *Nocardiopsis* sp. FU40. As described in Chapter 2 identification and manipulation of the apoptolidin gene cluster provides an opportunity to access glycovariants of apoptolidin A by targeted gene disruption. Mutation of *apoGT2* led to production of a glycovariant of apoptolidin A. In this case fermentation provided a new apoptolidin lacking the C27 disaccharide termed apoptolidin H. Ammocidin A is readily obtained from fermentation of *Saccharothrix* sp. AJ9571. Access to these major analogs motivated us to compare the cytotoxicity of this class of compounds with particular interest in the structure activity relationships of glycosylation.

In our first attempt of a standard cell viability assay using H292 human lung cancer cells, apoptolidin A induced cell growth arrest without any indication of cell death. In this experiment, cells at ~20% confluency were treated with apoptolidin A and after 48 hours assayed for cell viability with no significant cell death detected. Even treatment of cells with apoptolidin A for as long as 5 days resulted in only an antiproliferative effect but no loss of cell integrity. The assay was repeated using cells grown to high confluency (~70%)

* Results presented in this chapter have been previously published[11, 12] and portions and figures have been reproduced or adapted.

prior to apoptolidin A treatment. This attempt resulted in >95% cell death after 4 days with a calculated EC_{50} of 20-30 nM.

This observation of confluency dependence motivated us to quantify the effect in a standardized fashion. Cells were systematically plated in a 96 well format (10, 15, 20 and 25 thousand cells per well) 16 hours prior to treatment with apoptolidin A and then assayed for viability after 4 days. As shown in **Figure 3.1**, 25,000 cells per well resulted in a EC_{50} of 13 nM matching reported literature values. The results summarized in **Figure 3.1** also illustrate the antiproliferative activity of apoptolidin A (EC_{50} <100 nM) against lower confluency cells (10-20K cells). We also performed this type of assay using a titration of ammocidin A and a different adherent cell line (A2058 melanoma cells) and observed the same relationship between confluency and cytotoxicity shown in **Figure 3.2**.

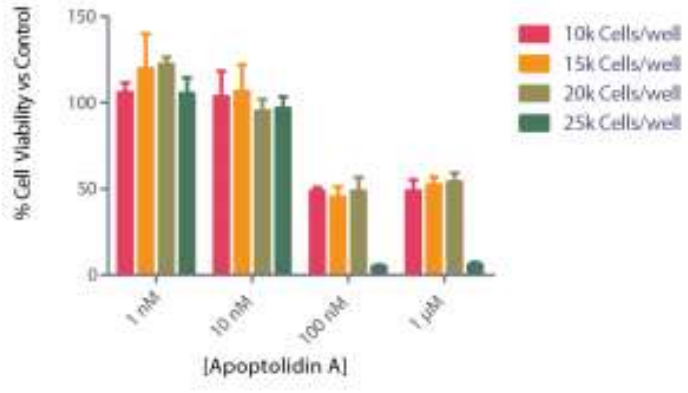


Figure 3.1: Cell density dependence cytotoxicity of apoptolidin A against H292 cells

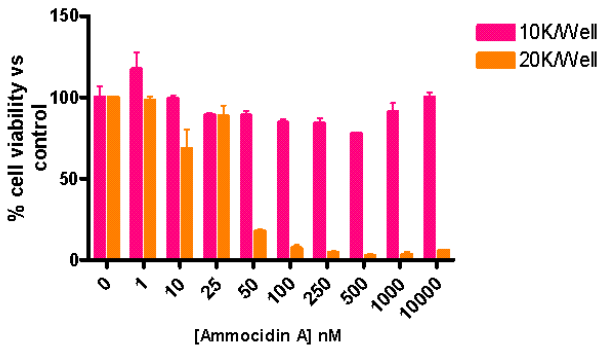


Figure 3.2: Cell density dependence cytotoxicity of ammocidin A against A2058 cells

In separate experiments, we have observed the cytotoxicity of apoptolidin A is potentiated by using cell culture media formulations of increasingly reduced glucose. Notably, such nutrient starvation conditions have been proposed to mimic poorly vascularized cells seen in solid tumors. We hypothesize that high and low confluency cells differ in metabolic flux with low confluency cells primarily utilizing the Embden-Meyerhof glycolytic pathway and high-density cells using the more energetic oxidative phosphorylation (OXPHOS) metabolism[1]. These results are in agreement with Salomon's results[2] as they demonstrated glycolytic cells previously insensitive to apoptolidin were sensitized to apoptolidin by the addition of 2-deoxyglucose or oxamate, small molecules known to channel carbon flux from the Embden-Meyerhof to the OXPHOS pathway.

Using the high confluency condition, we determined the EC_{50} values for a titration series of apoptolidin A and H and ammocidin A against H292 cells using a MTT assay to measure cell viability. Apoptolidin A was the most potent with an EC_{50} of 13 nM, followed by ammocidin A (50 nM) and apoptolidin H was almost 45 times less potent than apoptolidin A (600 nM). The EC_{50} curves are shown in **Figure 3.3**. The apoptolidinones A and D

(both fully deglycosylated) were also evaluated but where not cytotoxic up to concentrations of 10 μM. In contrast, the apoptolidin D disaccharide had an EC_{50} of 200 nM.

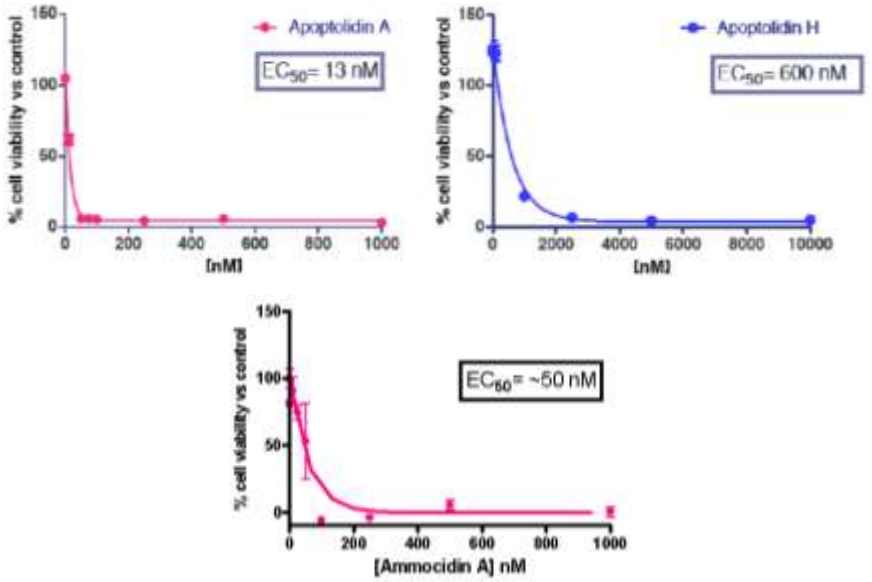


Figure 3.3: EC_{50} curves for apoptolidin A and H and ammocidin A. Cytotoxicities of apoptolidin A and H and ammocidin A against H292 cells

Inhibition of Mitochondrial ATPase

The F_O/F_1 ATPase was proposed as the molecular target by Salomon and coworkers who reported apoptolidin A to be a low micromolar inhibitor of the ATPase[3]. The fact that there is a difference of several orders of magnitude between the reported cytotoxic and enzymatic potency motivated us to try to reproduce their results using a yeast mitochondrial derived F_O/F_1 ATPase assay. Total mitochondrial protein was determined with a BCA protein assay kit and adjusted to a suitable range to measure inhibition of ATPase activity monitoring the rate of oxidation of NADH by following the decrease in adsorption at 350 nM over 10 minutes. Apoptolidin A and H showed modest and comparable inhibition with K_i values of 4.9 and 13.7 μM respectively, matching the value reported by Solomon. As glycosylation of apoptolidin had no apparent effect on enzymatic inhibition this result suggested a potential role in cellular localization by the sugar residues instead.

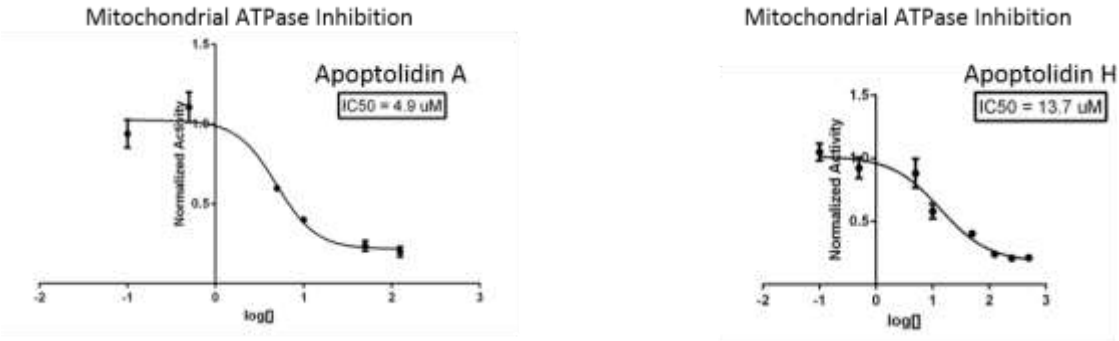


Figure 3.4: Inhibition of mitochondrial F_O/F_1 ATPase

Probe Development

In order to initiate chemical probe studies for cellular localization we required the introduction of an azido functional group within the apoptolidin core to enable conjugation to either fluorescent or affinity tags using click chemistry. To this end we took advantage of a report by the Wender group describing the selective benzoylation of the C2' hydroxyl group of the C9 sugar[4]. This approach proved successful as selective acylation of the C2' hydroxyl group of apoptolidins A and H was obtained using 5-azidopentanoic acid to afford azido derivatives of A and H in 30-40% yield. Importantly, when evaluated in the cell viability assay (**Figure 3.5**), azido analogs maintained activity comparable to their parent substrates (EC_{50} 12 and 350 nM).

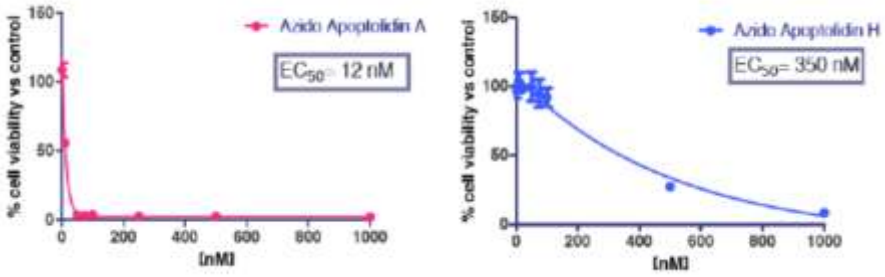


Figure 3.5: EC_{50} values of azido apoptolidin A and H against H292 cells

As partner alkyne tags we selected the cyanine dye Cy3 and biotin PEG tethered to click ready bicycle[6.1.0]nonynes (BNE). Coupling of BNE-Cy3 with azido apoptolidins A and H proceeded smoothly in methanol at room temperature over 4 hours to give fluorescently labeled apoptolidins A and H in 39% and 32% yield, respectively. In addition, biotin-BNE was reacted with apoptolidin A under identical conditions to give biotinylated apoptolidin

A in 29% yield. Cy3 conjugates maintained activity relative to their parent macrolides (EC_{50} 20 nM and EC_{50} 600 nM) when evaluated in the H292 cell assay.

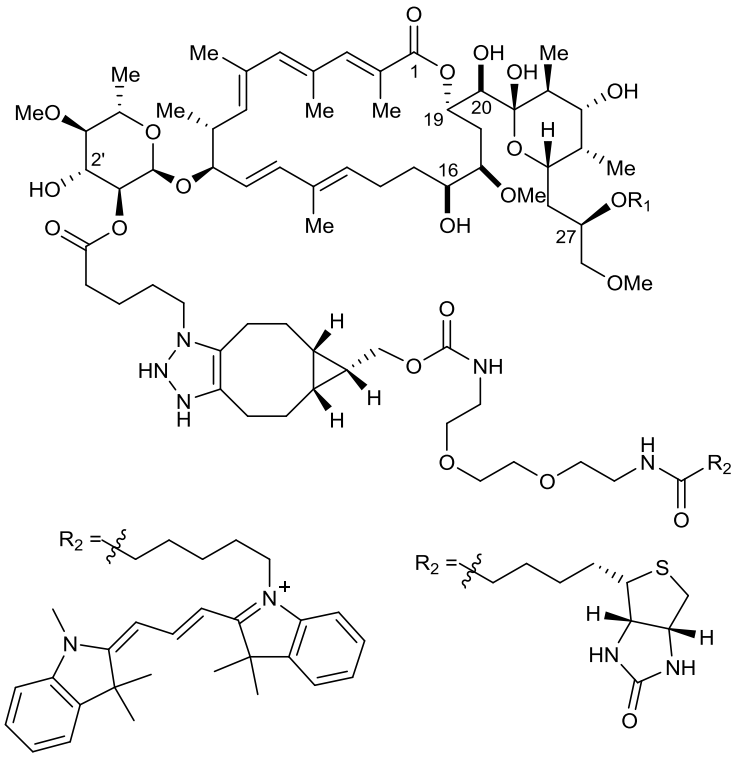


Figure 3.6: Structures of apoptolidin probe compounds

Fluorescent Microscopy Studies

Imaging of localization of apoptolidin probe compounds by H292 cells

Confocal microscopy studies were conducted with Cy3 apoptolidin A and H conjugates at concentrations of 200 nM with H292 human lung cancer cells. In these experiments compound treatment for 15 min was followed by a 60 min washout in order to remove nonspecific binding. Cellular images of experiments using Cy3 apoptolidin A are shown in **Figure 3.7**. Staining of washed out cells with Mitotracker Green FM (**Figure 3.7A**) was conducted in order to evaluate whether Cy3 apoptolidin A (**Figure 3.7B**) localized in the mitochondria. Inspection of the merged image (**Figure 3.7D**) confirmed co-localization with Mitotracker stain. Colocalization was further quantified by Costes' analysis which showed excellent overlap with a Pearson's coefficient of 0.89. An identical set of experiments using Cy3 apoptolidin H demonstrated similar localization in the mitochondria. This observation is in agreement with earlier reports describing apoptolidin A as a ATP synthase inhibitor.

These results should not be considered conclusive as cationic dyes such as Cy3 tend to localize in the mitochondria[5]. To more effectively judge whether the bioactivity enabled by glycosylation of the apoptolidins is due to directing localization within the mitochondria we are now in the process of examining non-cationic dyes conjugated to apoptolidins and the non-toxic aglycone (apoptolidinone) in microscopy experiments.

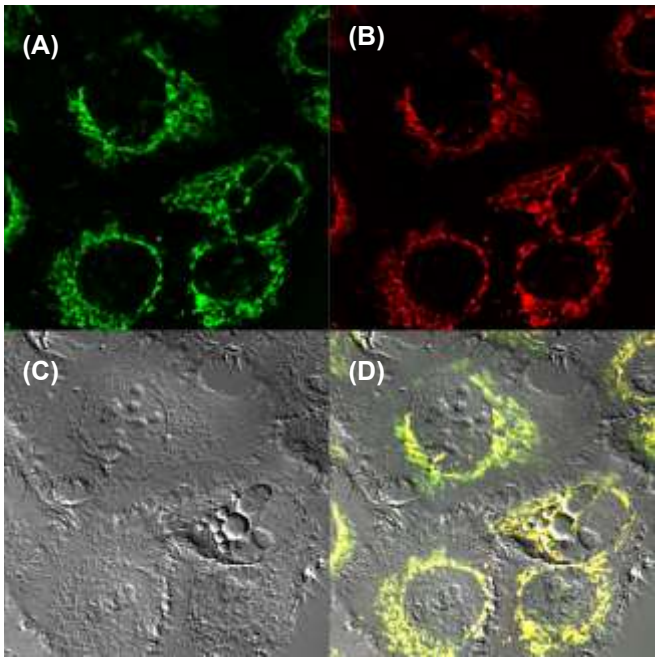


Figure 3.7: Mitochondrial localization of Cy3 Apoptolidin. Flourescent images of (A) Mitotraker (B) Cy3 apoptolidin, (C) differential interference contrast image, and (D) overlay of all images.

Imaging of uptake of apoptolidin probe compounds by PBMCs, A549, and U87 cells

We also used confocal fluorescent microscopy to characterize the uptake of Cy3 apoptolidin A and H in healthy peripheral blood mononuclear cells (PBMCs) and human lung adenocarcinoma (A549) and human glioblastoma (U87) tumor cells after a 1-hour treatment (**Figure 3.8)**. The confocal images revealed minimal uptake of Cy3 apoptolidins by healthy PBMCs, but higher uptake of Cy3 apoptolidins by A549 and U87 tumor cells.

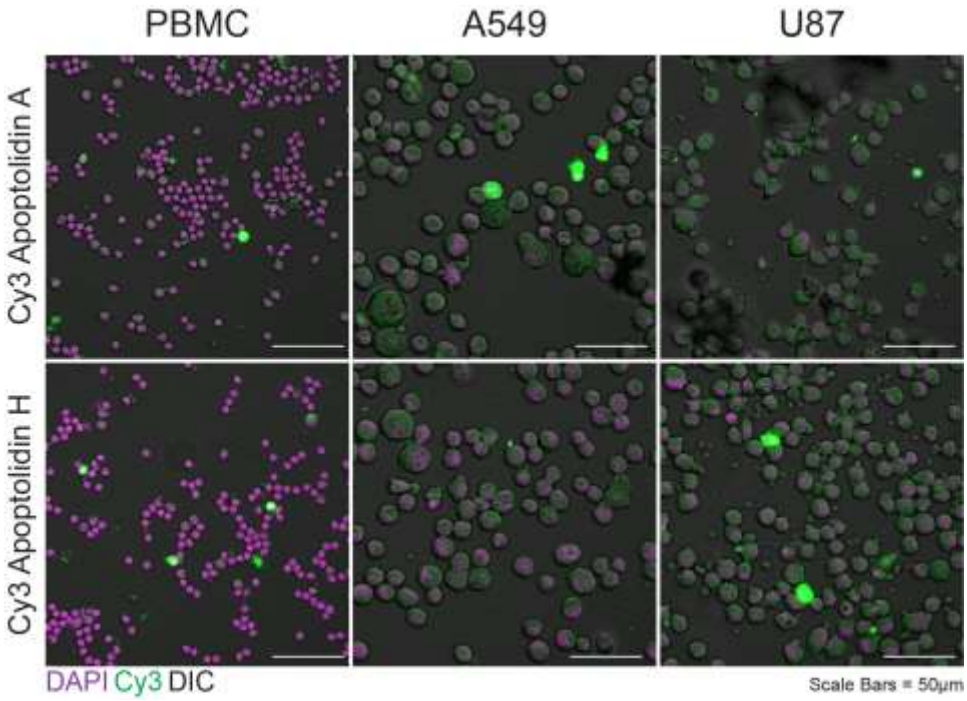


Figure 3.8: Cellular uptake of Cy3 apoptolidins in PBMCs, A549, and U87 cells.

Preliminary Flow Cytometry Results

Analysis of apoptolidins against H292 cells

Apoptolidin A's selective toxicity for tumor cells suggests that it is a promising lead for the treatment of cancer. However, to harness the potential of this natural product requires a complete understanding of its cellular target and its mechanism of action. At the end of their publication on the mechanism of action of apoptolidin, Salomon and coworkers concluded that apoptolidin A induces apoptosis in LYas mouse lymphoma cells on the basis of staining for annexin V and propidium iodide (PI).

We likewise performed flow cytometry analysis on apoptolidin A treated H292 cells and observed only live cells after 48 hours or alternatively, complete cell death (PI positive cells) after 2-3 days with minimal annexin V positive staining (**Figure 3.9**). Subsequent attempts to capture apoptosis at other time points was also unsuccessful.

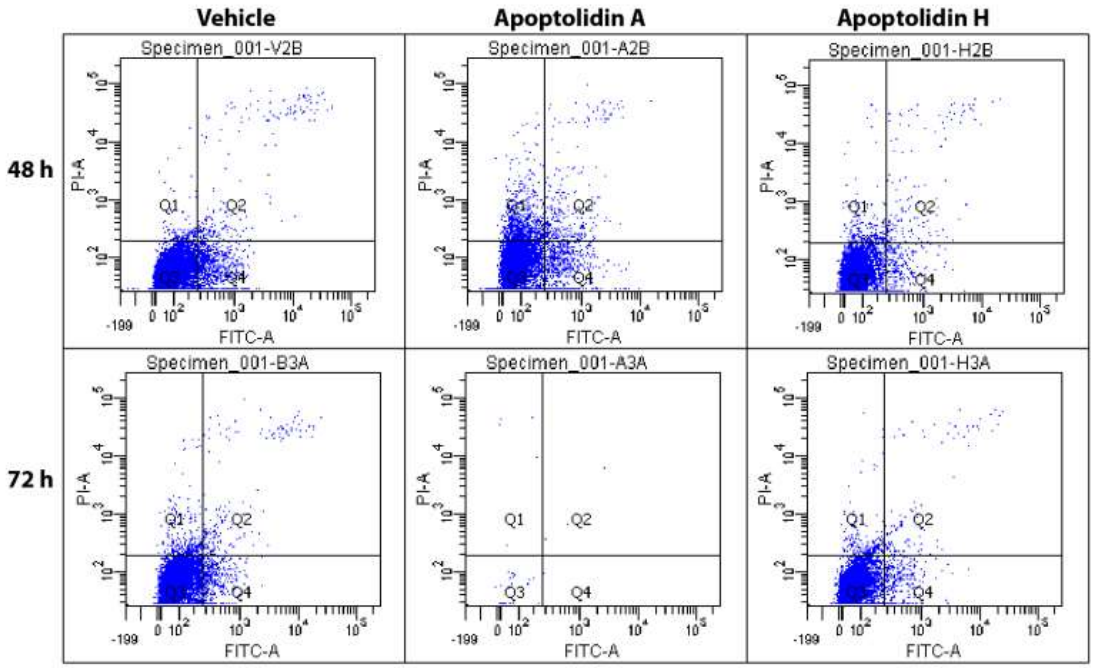


Figure 3.9: Annexin V assay. Biaxial plots of PI (viability) and FITC (Annexin V). The majority of events occur in quadrant 3 (negative for both markers) indicating most cells remain viable.

The initial data suggest that the killing of H292 cells by apoptolidin might not be occurring by apoptosis, but rather by necrosis, as cells are completely ruptured, and membrane integrity is lost. Alternatively, we hypothesize that apoptolidin A may cause cell death by apoptosis or necrosis, depending on cell type and environmental conditions.

FACS analysis of apoptolidins against PBMCs and sensitive and insensitive cell lines

In the original isolation paper and later when evaluated against the National Cancer Institute 60 (NCI-60) human cancer cell line panel, apoptolidin A was described as a selective inhibitor of cell growth[2,6]. One rational for cell type selective activity is selective cellular uptake. Hecht and co-workers have demonstrated cyanine tagged bleomycin is

selectively taken up in most cancer cell lines in comparison to "normal" cell counterparts in cell culture[7]. To test this hypothesis, we decided to characterize the uptake of apoptolidin A and H by different human cell types, as well as their signaling responses to treatment.

Fluorescent phospho-specific flow cytometry (phospho-flow) employs fluorescently tagged antibodies to dissect activation of cell signaling pathways in single cells in response to treatment with small molecules including natural products[8]. As fluorophores with different emission wavelengths can be monitored on different channels the uptake of fluorescent small molecules can be monitored as well as cell response. For example, the cellular uptake of fluorescent anticancer agents such as daunomycin as well as fluorescent nanoparticles has been monitored by traditional flow cytometry[9].

Using phospho-flow, we monitored cellular uptake of Cy3 apoptolidins and phosphorylation of acetyl-CoA carboxylase (ACC) after short-term (1 hour) treatment with vehicle (DMSO) or apoptolidins. We tested the response in PBMCs, the apoptolidin sensitive glioblastoma cell lines LN229 and U87 (reported to undergo autophagy by way of AMPK activation, as indicated by increased phosphorylation of AMPK (Thr 172), ACC (Ser79) and ULK1 (Ser555)[10], the apoptolidin sensitive colon cancer cell line SW620 and in A549 cells which appeared to be apoptolidin insensitive when evaluated in the NCI-60 cell line screen.

After 1 hour of treatment, healthy PBMCs and all four cancer cell lines showed almost complete (>98%) uptake of Cy3 apoptolidins A and H. Cancer cells showed higher Cy3 signal compared to healthy PBMCs, suggesting greater uptake of Cy3 apoptolidin A and

H, corresponding to imaging data in **Figure 3.8**.

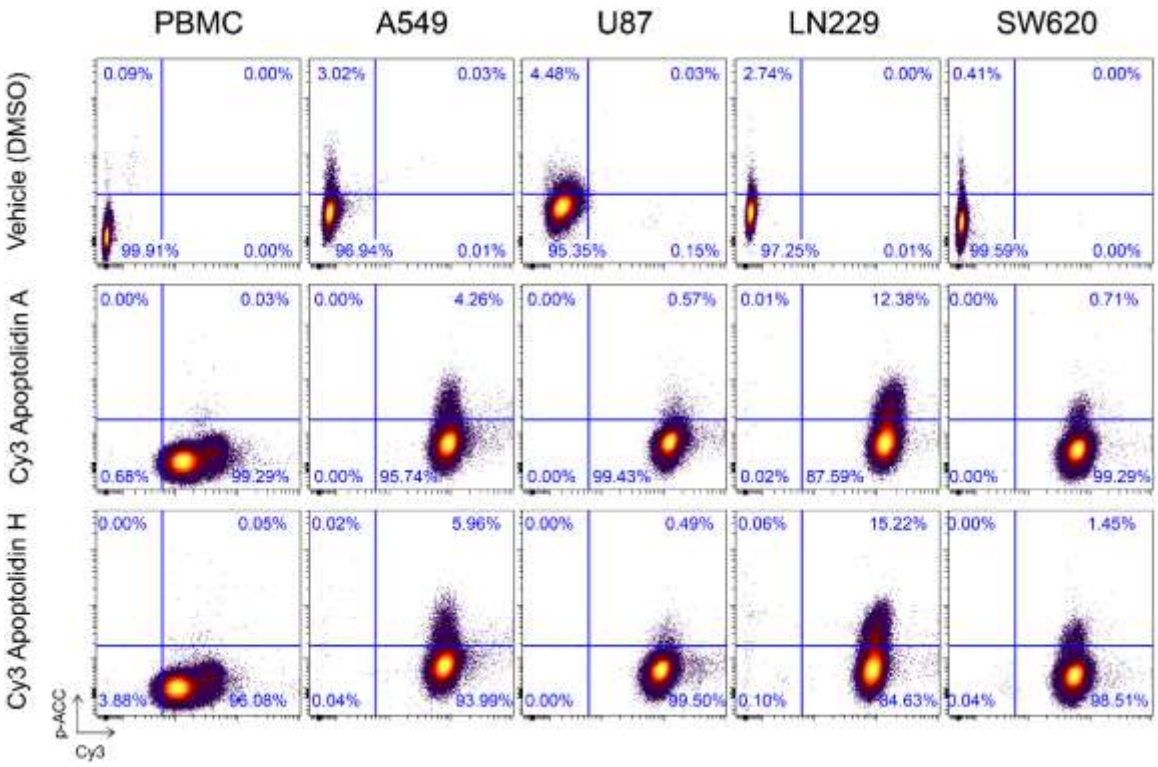


Figure 3.10: Biaxial plots of p-ACC vs Cy3-apoptolidin uptake by cell type

We also measured phosphorylation-specific ACC (p-ACC; y-axis), a marker indicative of autophagy (**Figure 3.10**). In all cancer cell lines, we observed a proportion of cell subset that showed high p-ACC signal at baseline (DMSO treatment). LN229 glioblastoma cells showed the highest increase in the abundance of this subset, from 2.74% (DMSO) to 12.39% and 15.28% after Cy3 apoptolidin A and H treatments, respectively. The majority (>99%) of p-ACC expressing LN229 cells after apoptolidin treatments were among the cells that had Cy3 apoptolidin-A and H uptake. In contrast, healthy human PBMCs did not show an increase in p-ACC expression in response to apoptolidin treatments. A549 and

SW620 cells showed only a minimal increase in abundance of cells expressing p-ACC after treatment with Cy3 apoptolidin-A and H, suggesting that A549 and SW620 cells were relatively insensitive to apoptolidins compared to LN229 cells at 1 hour. However, U87 cells showed no increase in p-ACC activity after short-term treatment with apoptolidins A or H.

Conclusions

Since the initial reports of their discovery and cell type specific cytotoxicity, the apoptolidins and ammocidins have been the focus of significant research efforts by several laboratories. Development of synthetic and biosynthetic approaches to access novel congeners has generated a ready supply of parent compounds and analogs. Testing of these compounds first in cytotoxicity assays with adherent cell lines indicated that the state of glycosylation comprises a significant structure activity relationship as exampled by the difference in potency between apoptolidin A and H (15 nM and 600 nM, respectively). Evaluating the enzymatic inhibition of the F_O/F_1 ATPase by apoptolidin A and H indicated the sugar residues do not effect this activity as both compounds had a similar Ki (5 μM and 14 μM, respectively).

To investigate potential causes of differences in activity, a series of probe compounds have been developed via selective acylation of the 2' hydroxyl of the C9 sugar. Fortuitously modifications at this site have minor impact on the activity of analogs. Fermentation of the producing organism and a mutant strain in which the glycosyltransferase apoGT2 is inactivated provides an economic and abundant supply of the needed starting materials for probe development.

Applications of fluorescent probes have shown similar rates of uptake and localization in the mitochondria independent of the presence or absence of the C27 disaccharide. However, these results are confounded by the observation that the positive charge of the fluorophore is sufficient to cause accumulation in the mitochondrial membrane. Seeking to understand if the disaccharide might be influencing the rate of uptake by cells or if cell type dependent uptake might serve as a cause of activity difference, several experiments were performed utilizing both fluorescent microscopy and flow cytometry to measure uptake in apoptolidin sensitive and resistant cell lines and PBMCs. All cell types tested showed more than 98% uptake of apoptolidins after 1 hour of treatment. Additionally LN229 cells responded with a marked increase in p-ACC expressing cells, suggesting their sensitivity to apoptolidins. Even though the responses were not as striking, A549 and SW620 cells also showed minimal increase in p-ACC after 1-hour of apoptolidin treatment, whereas healthy human PBMCs and U87 cells did not change p-ACC expression.

Experimental Methods

General procedures

All non-aqueous reactions were performed in flame-dried or oven dried round-bottomed flasks under an atmosphere of argon. Stainless steel syringes or cannula were used to transfer air- and moisture-sensitive liquids. Reaction temperatures were controlled using a thermocouple thermometer and analog hotplate stirrer. Reactions were conducted at room temperature (rt, approximately 23 °C) unless otherwise noted. Flash column chromatography was conducted using silica gel 230-400 mesh. Analytical thin-layer chromatography (TLC) was performed on E. Merck silica gel 60 F254 plates and visualized using UV, and potassium permanganate stain. Yields were reported as isolated, spectroscopically pure compounds.

Materials

Solvents were obtained from either an MBraun MB-SPS solvent system or freshly distilled (tetrahydrofuran was distilled from sodium-benzophenone; toluene was distilled from calcium hydride and used immediately; dimethyl sulfoxide was distilled from calcium hydride and stored over 4 Å molecular sieves). Commercial reagents were used as received. The molarity of n-butyllithium solutions was determined by titration using diphenylacetic acid as an indicator (average of three determinations).

Instrumentation

Semi-preparative reverse phase HPLC was conducted on a Waters HPLC system using a Phenomenex Luna 5 μm C18(2) 100A Axia 250 x 10.00 mm column or preparative reverse phase HPLC (Gilson) using a Phenomenex Luna column (100 Å, 50 x 21.20mm, 5 μm

C18) with UV/Vis detection. Infrared spectra were obtained as thin films on NaCl plates using a Thermo Electron IR100 series instrument and are reported in terms of frequency of absorption (cm-1). LC/MS was conducted and recorded on an Agilent Technologies 6130 Quadrupole instrument. High-resolution mass spectra were obtained from the Department of Chemistry and Biochemistry, University of Notre Dame using either a JEOL AX505HA or JEOL LMS-GC mate mass spectrometer or by the Vanderbilt University Center for Neuroscience Drug Discovery (VCNDD) on a Micromass Q-Tof API-US mass spectrometer.

Production and chemical synthesis of apoptolidins and fluorescent derivatives

Apoptolidins A and H were produced by fermentation of the apoptolidin producer FU 40 and a mutant strain (ApoGT2 knockout). Cyanine-3 derivatives of apoptolidin A and H were prepared by semi synthesis as described previously[11].

MTT cell viability cell density experiments

Low passage (P#<25) H292 human lung carcinoma cells (obtained from the American Type Culture Collection, ATCC) were plated at 5, 10, 15, 20 or 25 thousand cells per well in 96-well plates in 100 μL of RMPI 1640 medium containing 10% fetal bovine serum and 100 IU penicillin and 100 mg/mL streptomycin and incubated for 16 hours to attach. Apoptolidin A was dissolved in DMSO at1 μM, 10 μM, 100μM and 1 mM. The resulting DMSO stock solutions were diluted in complete RPMI medium1000:1 to yield medium solutions containing 1 nM, 10 nM, 100 nM, and 1 μM apoptolidin A or DMSO to give a final DMSO concentration of 0.1%. Media was removed from each well by aspiration and replaced with media containing DMSO vehicle or apoptolidin A for a total of n=4 wells

per cell density and concentration. Cells were incubated for four days (96 hours). The media from each well was then aspirated and replaced with 100 μL of complete RPMI medium containing 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) at 0.5 mg/mL and returned to the incubator at 37 °C for two hours. Media from each well was aspirated and replaced with 100 μL DMSO. Absorbance was measured at 560 nM using a GloMax Multiplate reader (Promega, Madison, WI, USA). Blank absorbance from the average of 8 wells treated with MTT in cell-free medium was subtracted from each value. The percent cell viability of each well was calculated as the fraction of the average absorbance of DMSO control treated cells at each condition.

MTT cell viability EC_{50} experiments

Low passage (P#<25) H292 or A5409 cells were plated at 25,000 cells per well in 96-well plates in 100 μL of RMPI 1640 medium containing 10% fetal bovine serum and 100 IU penicillin and 100 mg/mL streptomycin and incubated for 16 hours to attach. Compounds used for testing were dissolved in DMSO at various concentrations from 1 μM to 10 mM. The resulting DMSO stock solutions were diluted in complete RPMI medium 1000:1 to yield medium solutions containing 1 nM to 10 μM of the respective compound to be assayed and a uniform DMSO concentration of 0.1%. Media was removed from each well by aspiration and replaced with media containing DMSO vehicle or compound for a total of n=8 wells per concentration. Cells were incubated for four days (96 hours). The media from each well was then aspirated and replaced with 100 μL of complete RPMI medium containing MTT at 0.5 mg/mL and returned to the incubator at 37 °C for two hours. Media from each well was aspirated and replaced with 100 μL DMSO. Absorbance was measured at 560 nM using a GloMax Multiplate reader (Promega, Madison, WI, USA). Blank

absorbance from the average of 8 wells treated with MTT in cell-free medium was subtracted from each value. The percent cell viability of each well was calculated as the fraction of the average absorbance of DMSO control treated cells. Data was plotted GraphPad Prism 5 and fitted with a non-linear regression curve. Effective concentration 50 (EC50) was estimated graphically as the concentration at which 50 % of control formazan product absorbance was detected.

Effect of culture media on toxicity of apoptolidin A

H292 cells were plated at a density of 500 per well in 96 well plates and treated for 7 days with apoptolidin A, from 3-30 nM. Cell viability was detected by loading the cells with Calcein-AM reagent for 30 min at 37 °C, followed by measurement of Calcein Fluorescence (485 nm excitation, 520 nm emission). Glucose concentrations of media types are shown in parenthesis MEM (1.0 g/L), DMEM (4.5 g/L), and RPMI 1640 (2.0 g/L).

F_O/F_1-ATPase inhibition assay

A single colony of DBY7286 (mat A, ura-/-) was inoculated into a pre-culture (50 mL in a 250 mL shake flask) of semisynthetic media (3 g yeast extract,0.5 g glucose, 0.5 g $CaCl_2 \cdot H_2O$, 0.5 g NaCl, 0.6 g $MgCl_2 \cdot H_2O$, 0.1 g KH_2PO_4, 0.1 g NH_4Cl, 22mL 90% DL-lactic acid, 8 g NaOH, and 1L ddI water, pH = 5.5), and incubated with orbital shaking for 15 h in flasks at 30 °C. Semisynthetic media (5 x 1 L of in 3 L Fernbach Flasks)were inoculated with 1% of preculture and incubated at 30 °C for 16 hrs to an OD600 of 3. Cells were collected by centrifugation at 2000 x g for 15 min. Supernatant was removed and the combined cell pellets were resuspended in 150 mL of distilled H2O. The suspension was

then transferred to pre-weighed centrifuge bottles and centrifuged at 2000 x g for 5 minutes. After decanting the supernatant, the wet weight of the cell pellet was determined. The 3 to 4 g pellet was then resuspended in 25 mL of 0.1 M Tris, 10 mM dithiothreitol, pH 9.4, and incubated for 15 minutes in a 30 °C water bath. The cells were then centrifuged at 2000 x g, resuspended in 20 mL of buffer A (1.2 M sorbitol, 20 mM KH_2PO_4, pH 7.4), and converted to spheroplasts by incubation with Zymolyase (2.5 mg/g cell pellet) for 30 min at 30 °C. Spheroplasts were collected by centrifugation at 4000 x g, washed, and resuspended twice in 20 mL cold buffer A. Washed spheroplasts were resuspended in 50 mL cold buffer B (0.6 M sorbitol, 20 mM K+MES, 0.5 mM PMSF, pH 6.0), homogenized using a Dounce homogenizer, diluted to 125mL with buffer B, and centrifuged at 1500 x g for 5 minutes. The supernatants were retained and the pellets resuspended in 50 mL buffer B, homogenized and centrifuged at 1500 x g for 5minutes. The pellets were then discarded, the supernatant suspensions were pooled, and centrifuged at 12000 x g for 10 minutes. The resulting supernatant was removed and the pellets were resuspended in 60 mL buffer B without PMSF with a homogenizer and centrifuged at 1500 x g for 5 minutes. The supernatant suspensions were then centrifuged at 12000 x g for10 minutes. The mitochondria containing pellets were resuspended in 1 mL buffer B and total mitochondrial protein was determined with a BCA protein assay kit (Thermo, Inc). Samples were adjusted to desired protein concentration with buffer C (0.6 M sorbitol, 20 mM K+HEPES, pH 7.4) To measure inhibition of ATPase activity 20 µg of mitochondrial protein was added to 100 µL of a solution of 50 mM Tris (pH 8.0), 3.3 mM $MgCl_2$, 2 µg/mL Antimycin, 5 u/mL lactate dehydrogenase, 3 u/mL pyruvate kinase, and 0.3 mM NADH, at 25 °C. The ATP consumption assay was initiated by addition of 1 mM ATP and 1 mM phosphoenol

pyruvate, and rate of oxidation of NADH was monitored by following the decrease in adsorption at 350 nM over 10 minutes.

Confocal microscopy localization experiments

Low passage (P < 25) H292 human lung carcinoma cells were plated in MaTek dishes at 15% confluence in 2.0 mL of RMPI 1640 medium containing 10%fetal bovine serum and allowed to attach and grow for 40 hours. 200 μM stock solutions of Cy3 apoptolin A, Cy3 apoptolidin H, or BME-Cy3 in DMSO was prepared. Each dish was treated by the following protocol. Media was removed by aspiration and replaced with 2.0mL serum free RPMI 1640 media. 2.0 μL of the appropriate DMSO stock solution was added to the dish and cells were returned to the incubator for 15 min. Media containing fluorophores was removed by aspiration followed by a wash with serum free RPMI 1640 media (3 x 2.0 mL) followed by a final addition of 2.0 mL of serum free media. After incubation for 30 min, media was removed and replaced by a freshly sonicated (important for mitotracker solubility) 30 nM solution of Mitotracker Green FM. After an additional incubation of 30 min in the incubator, media was removed by aspiration, (PBS, 2 x 2.0 mL) and replaced with 2.0 mL PBS. Each dish was then imaged at ten random fields by confocal microscopy. Confocal microscopy was performed on a LSM780 (Zeiss) using a c- Apochromat 40x 1.2 W Corr M27 oil immersion objective. Cy3 fluorescence was excited using 488 nm laser (2%) and emission was measured with a 492-542 nm bandpass. Mitotracker Green FM fluorescence was excited with a 488 nm laser (2%) and emission as measured with a bandpass of 552-683 nm. All images were acquired using 512x512, 0.14 μm diameter pixels, a 12.6 μs pixel dwell time, 12-bit gray levels and a 2.4μm optical section. Each compound was tested in 3 dishes of cells. Pearson's coefficients were calculated using the

JACoP plugin7 for ImageJ8 1.46r software for each field from each dish and are reported as the average of the 30 calculations.

Uptake of apoptolidins A and H in various cell types

Human cancer cell lines and peripheral blood mononuclear cells (PBMCs) were used to characterize uptake of apoptolidin A and apoptolidin H. The following cell lines were included: SW620 (colon cancer), U87-MG (glioblastoma), LN229 (glioblastoma), and A549 (lung adenocarcinoma). Cell lines were cultured under ATCC recommended protocols. Cells were detached using Trypsin and resuspended in recommended culture media at $1x10^6$ cells/mL prior to drug treatment. Human PBMCs were collected from a healthy donor following protocols approved by Vanderbilt University Medical Center Institutional Review Board, processed by standard Ficoll preparation protocol, and cryopreserved in liquid nitrogen. PBMCs were thawed and resuspended in warm RPMI 1640 media containing 10% FBS at $1x10^6$ cells/mL prior to drug treatment. Cells were treated with either vehicle (DMSO), 1 μM of Cy3 apoptolidin A, or 1 μM of Cy3 apoptolidin H for 1 hour at 37 °C. Cells were washed twice in PBS and fixed with 1.6% paraformaldehyde for 10 minutes at room temperature, and were permeabilized with ice-cold methanol for 30 minutes.

Fluorescent flow cytometry

After methanol permeabilization, cells were stained with 1:250 anti p-ACC antibody (Cell Signaling) for 30 minutes in the dark at room temperature. Cells were then stained with 1:1000 Donkey anti-Rabbit Ax647 (Life Technologies) for 30 minutes in the dark at room temperature, and were washed and resuspended in PBS for analysis on 5-laser BD LSRII

(BD Biosciences, San Jose, CA) at the Vanderbilt Flow Cytometry Shared Resource and evaluated using Cytobank software.

Confocal microscopy uptake experiments

The stained cell suspensions described above were placed on glass slides for imaging on LSM 710 META inverted (Zeiss) at the Vanderbilt Cell Imaging Shared Resource. Data were analyzed using Zen 2011 software.

References

1 Bereiter-Hahn, J., Munnich, A. & Woiteneck, P. Dependence of energy metabolism on the density of cells in culture. *Cell Structure and Function* **23**, 85-93, (1998).

2 Salomon, A. R., Voehringer, D. W., Herzenberg, L. A. & Khosla, C. Understanding and exploiting the mechanistic basis for selectivity of polyketide inhibitors of F0F1-ATPase. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14766-14771, (2000).

3 Salomon, A. R., Voehringer, D. W., Herzenberg, L. A. & Khosla, C. Apoptolidin, a selective cytotoxic agent, is an inhibitor of F0F1-ATPase. *Chemistry and Biology* **8**, 71-80, (2001).

4 Wender, P. A., Jankowski, O. D., Tabet, E. A. & Seto, H. Toward a structure-activity relationship for apoptolidin: selective functionalization of the hydroxyl group array. *Organic Letters* **5**, 487-490, (2003).

5 Kim, Y. K. *et al.* Control of muscle differentiation by a mitochondria-targeted fluorophore. *Journal of the American Chemical Society* **132**, 576-579, (2010).

6 Kim, J. W., Adachi, H., Shin-ya, K., Hayakawa, Y. & Seto, H. Apoptolidin, a new apoptosis inducer in transformed cells from Nocardiopsis sp. *Journal of Antibiotics* **50**, 628-630, (1997).

7 Schroeder, B. R. *et al.* The disaccharide moiety of bleomycin facilitates uptake by cancer cells. *Journal of the American Chemical Society* **136**, 13641-13656, (2014).

8 Krutzik, P. O., Trejo, A., Schulz, K. R. & Nolan, G. P. Phospho flow cytometry methods for the analysis of kinase signaling in cell lines and primary human blood samples. *Methods in Molecular Biology* **699**, 179-202, (2011).

9 Dordal, M. S. *et al.* Flow cytometric assessment of the cellular pharmacokinetics of fluorescent drugs. *Cytometry* **20**, 307-314, (1995).

10 Serrill, J. D. *et al.* Apoptolidins A and C activate AMPK in metabolically sensitive cell types and are mechanistically distinct from oligomycin A. *Biochemical Pharmacology* **93**, 251-265, (2015).

11 DeGuire, S. M. *et al.* Fluorescent probes of the apoptolidins and their utility in cellular localization studies. *Angewandte Chemie International Edition* **54**, 961-964, (2015).

12 Chong, K.M., *et al.* The use of fluorescently-tagged apoptolidins in cellular uptake and response studies. *Journal of Antibiotics* **69**, 327-30, (2016).

CHAPTER 4

Development of Multiplexed Activity Metabolomics for Phenotypic Discovery[‡]

Design and Validation of a Multiplexed Activity Metabolomics Platform

Generation of natural product fraction libraries and cheminformatic annotation

Despite the centrality of metabolite functional analysis, the development of a generalizable 'omics-scale solution for uncovering the functional roles of secondary metabolites within disease relevant cellular contexts remains a substantial challenge.[1] It is now possible to convert biological extracts (*e.g.*, of microbial culture, plant/tissue origin) into highly characterized chromatographic microtiter arrays by split-flow liquid chromatographic mass spectrometry.[2] The biological characterization of such untargeted metabolomic arrays results in the generation of 'bioactivity chromatograms', and correlation analysis to matched extracted ion current (EIC) mass chromatograms identifies candidate metabolites linked to measured bioassay targets. However, per-well single assay modalities greatly limit the efficiency of this approach, and targeted biochemical assays or phenotypic assays against cell lines reveal only a fraction of significant roles of metabolites in arrays.

[‡] Results presented in this chapter have been previously published[26] and portions and figures have been reproduced or adapted.

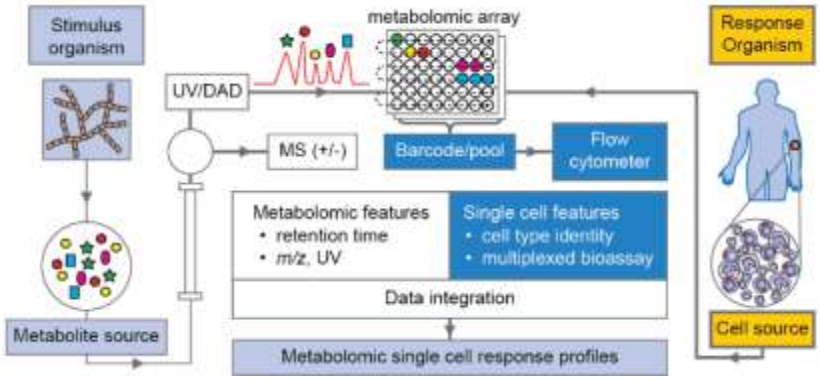


Figure 4.1: Schematic for assaying natural product libraries. High data content metabolomic arrays are generated in replicate from a 'stimulus' organism via split flow polarity switching chromatography mass spectrometry. A suspension of disaggregated tissue cells from a 'response' organism (human) is added to the metabolomic array.

The Multiplexed Activity Metabolomics (MAM) workflow first generates a metabolomic array in microtiter plate format via reversed phase liquid chromatographic separation of a crude biological extract produced by a 'stimulus' organism. A portion of the effluent is diverted to a polarity-switching electrospray mass spectrometric analyzer (ESI-MS) and the remainder of the effluent to a microtiter plate fraction collector after passing through a UV/VIS diode array detector. Following evaporation and resuspension of collected fractions, cell preparations from a 'response' organism (e.g. humans, represented by tissue cells) are added to the microtiter wells for incubation with the metabolomic fractions to induce cellular responses (**Figure 4.1**).

Multiplexed cytometric analysis utilizing fluorescent cell barcoding

Cells within wells are then stained for viability, fixed and permeabilized, and fluorescent cell barcoding (FCB, **Figure 4.2**) is used to label the well contents via differential staining of cells with *N*-hydroxy succinimide (NHS) ester-functionalized fluorescent dyes.[3] Thusly

'barcoded', cells in the microtiter wells are then pooled and stained with multiple fluorescent antibodies to quantitate cell status and targeted cell type-specific responses to metabolites. Critically, flow cytometric gating based on the barcoding fluorophores facilitates the assignment of cells to their original coordinates on the microtiter plate metabolite array (i.e. 'deconvolutes' treatment conditions for each cell), yielding simultaneous bioassay marker quantitation per well for each targeted antibody-fluorophore conjugate. Barcoding enables high throughput antibody assays by using a fraction of antibody reagents compared to a microtiter format and pooling also ensures the uniformity of antibody staining of cells across all wells, decreasing experimental variation. The result is a multiplexed series of well coordinate-linked immunoassay profiles running through the metabolomic fraction array.

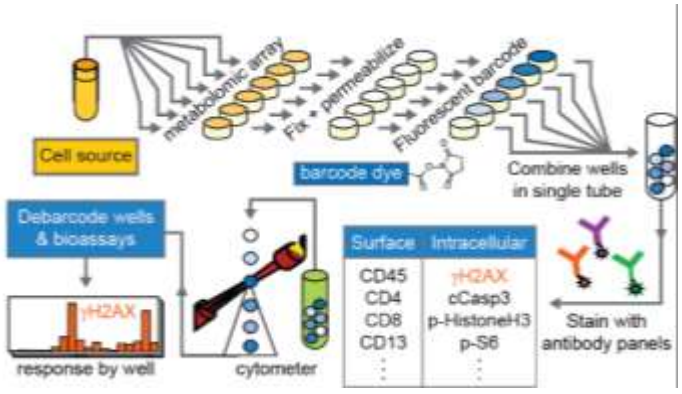


Figure 4.2: Multiplexing assays with FCB. Flow cytometric cell barcoding and multiplexed immunoassays are used to identify multiple cell type/sub-type specific biological responses to metabolites in the array. Correlation analysis of the resulting bioactivity and UV/ESI/MS(+)(-) data generate putative functional activities for metabolites.

To maximize available fluorescence channels for multiparameter flow-cytometry, FCB was adapted to barcode 48 wells with two fluorescent NHS-activated ester dye gradients of NHS-Pacific Orange and NHS-Pacific Blue. After two dimensional barcoding, wells were pooled into a single tube and stained with fluorescently tagged antibodies.

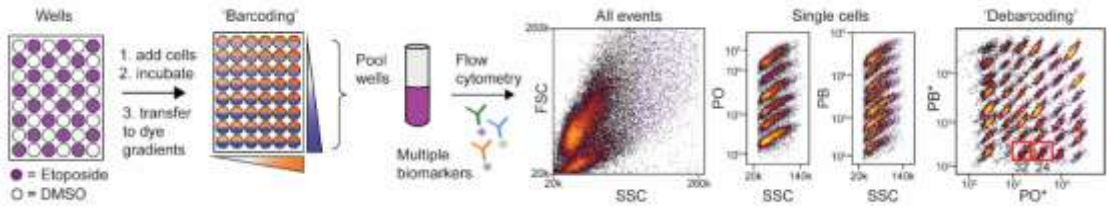


Figure 4.3: Design of a checkerboard validation experiment using Kasumi cells and a DNA active natural product. Overview of 48-well fluorescent cell barcoding and debarcoding validation. Compounds and vehicle are added in a checkerboard pattern to 48 wells and cells were added and incubated prior to being barcoded using dye gradients of N-hydroxysuccinimide functional Pacific Orange and Pacific Blue. Cells are stained with Ax700, fixed, permeabilized and pooled prior to immunoassay with antibodies tagged with non-overlapping fluorescent dyes. Cells are analyzed by flow cytometry and gated selecting (i) intact, single cells, (ii) Pacific Orange (PO) to reveal columns, and (iii) Pacific Blue to reveal rows and generate populations for each well.

Checkerboard validation experiment with etoposide

To test the robustness of the FCB assay, Kasumi-1 cells were incubated in 48 wells in a checkerboard fashion with vehicle dimethylsulfoxide (DMSO) or one of two benchmark natural products: the podophyllotoxin derivative etoposide[4], a potent topoisomerase inhibitor and inducer of double-strand DNA breaks, or the bisindole alkaloid staurosporine[5], a classical inducer of apoptosis. After treatment, cells were stained with a permeability/viability indicator, Alexafluor 700 (Ax700)[6], barcoded, combined into a single tube, and then stained with fluorescently labeled antibodies specific to either

cleaved caspase-3 (cCasp3), a protein activated in apoptosis[7], or γH2AX, a histone phosphorylated during genomic damage[8,9]. A representative workflow and data for etoposide is shown in **Figure 4.3**. Analysis of single cell events revealed two populations for each readout, and biaxial plots of Pacific Orange versus Pacific Blue yielded 48 distinct populations.

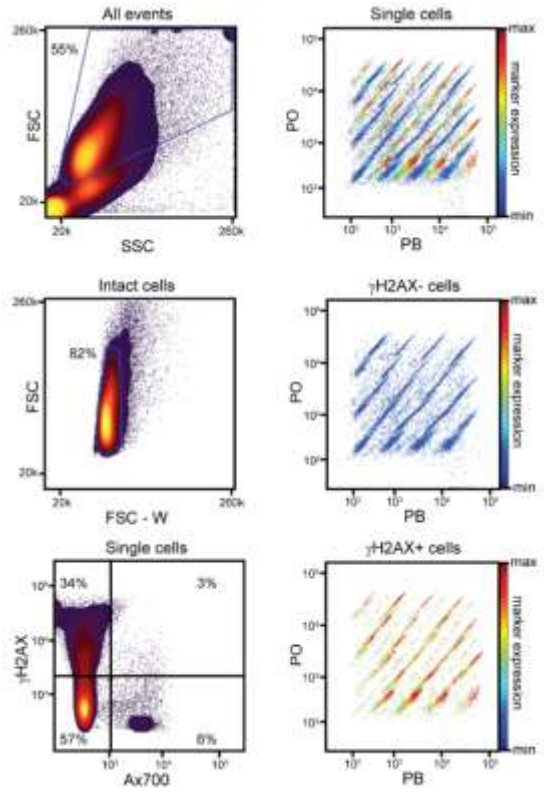


Figure 4.4: Gating strategy for etoposide checkerboard. All collected events from staurosporine checkerboard experiment were gated for intact cells (FSC vs SSC), then single cells (FSC vs FSC-W) and finally for γH2AX expression.

Recovery of the well coordinates and determination of antibody binding in debarcoded populations was accomplished using Cytobank, a cloud-based cytometric analysis platform, to confirm compound-specific effects (**Figure 4.4**).

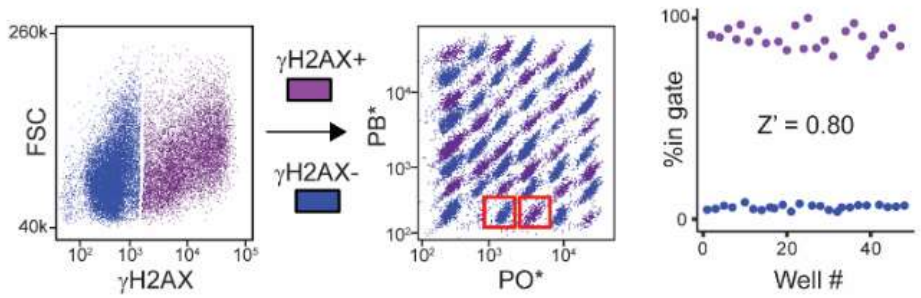


Figure 4.5: Z-score analysis of etoposide validation checkerboard. Biaxial plots of single cells (PO vs PB) colored by γH2AX expression visually reflect the checkboard pattern. Plots of only γH2AX - cells or γH2AX+ cells show populations whose FCB coordinates match assay wells with vehicle or compound respectively.

In the case of etoposide, gating for γH2AX and then debarcoding illustrated the bifurcated response within the checkerboard (**Figure 4.5**), and biaxially gated percent changes reflect bioassay results.

Comparable results were obtained for staurosporine (**Figure 4.6 and 4.7**), and these results were used to calculate the standard deviation of each assay plate, which conformed to levels in standard practice in high-throughput screening analysis (Z-factor > 0.77).

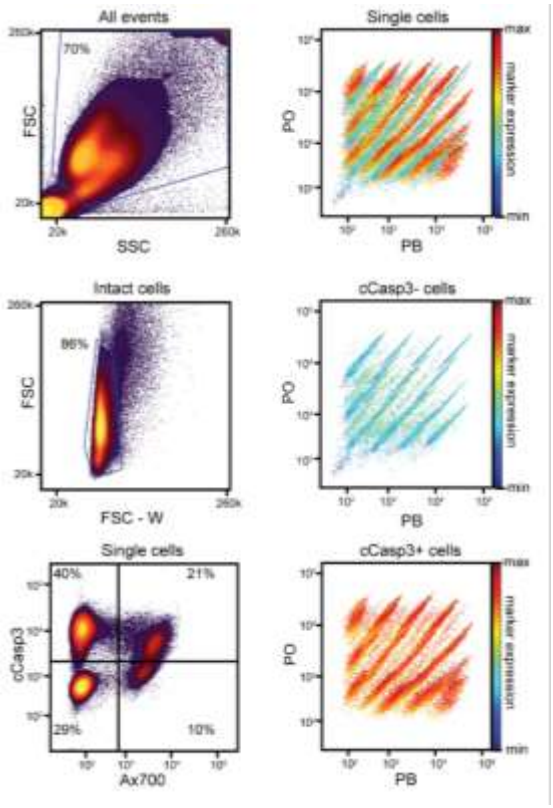


Figure 4.6: Gating strategy for staurosporine checkerboard. All collected events from staurosporine checkerboard experiment were gated for intact cells (FSC vs SSC), then single cells (FSC vs FSC-W) and finally for cCasp3 expression.

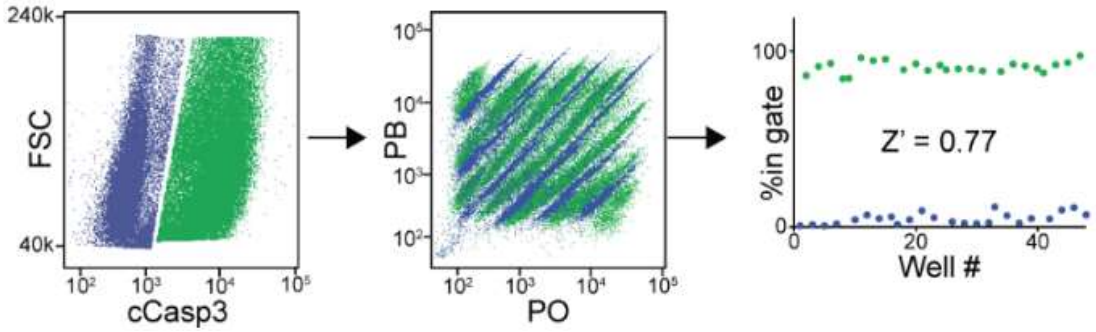


Figure 4.7: Z-score analysis of staurosporine validation checkerboard. Biaxial plots of single cells (PO vs PB) colored by cCasp3 expression visually reflect the checkboard pattern. Plots of only cCasp3- cells or cCasp3+ cells show populations whose FCB coordinates match assay wells with vehicle or compound respectively.

As an additional evaluation of barcoding, cCasp3 and γH2AX expression induced by staurosporine and etoposide, respectively, were demonstrated to be dose-dependent within assay conditions by separate cytometric barcoding of concentration response curves and quantitating antibody binding (**Figure 4.8**)

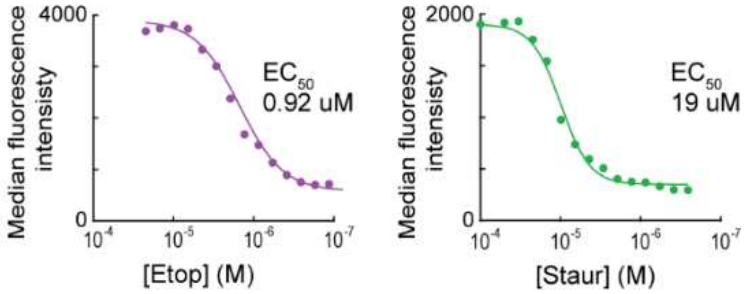


Figure 4.8: Dose Response Curves from etoposide and staurosporine titrations. A titration series of etoposide and staurosporine was prepared on a microtiter plate, incubated with KG1 cells, barcoded and stained. EC_{50} values were calculated using the median fluorescent intensity of cell populations from each well.

Validation with mixture of known compounds

The integrated analysis of an HPLC-MS-generated chromatographic array in conjunction with FCB and cellular response data (MAM) was validated using a chemically defined mixture of bioactive compounds. A mixture of six structurally and mechanistically diverse cytotoxic small molecules was chromatographically arrayed and assayed against a human myeloid leukemia-derived cell line (KG1) using the MAM platform. EIC chromatograms for the six compounds can be readily compared to bioactivity chromatograms and demonstrated specific and mechanistically expected responses to multiplexed immunoassays (**Figure 4.9**). For instance, the EIC peak for the known apoptosis-inducing

secondary metabolite staurosporine (*m/z* = 467.5), was the highest correlating peak in the well 25 bioactivity bin for cCasp3. Similarly, the largest response for γH2AX occurred in well 20, matching the retention time of the potent topoisomerase inhibitor etoposide (*m/z* = 606.5). Of note, in this experiment of modest complexity, a single cytometric flow run generates an aggregate of 240 individual raw immunoassays, which may be further combined into additional function assays that can be compared to arrayed compound elution profiles. Importantly, MAM successfully identified and differentiated compounds in a mixture based on their elution profile and differential response to a multiplexed antibody panel.

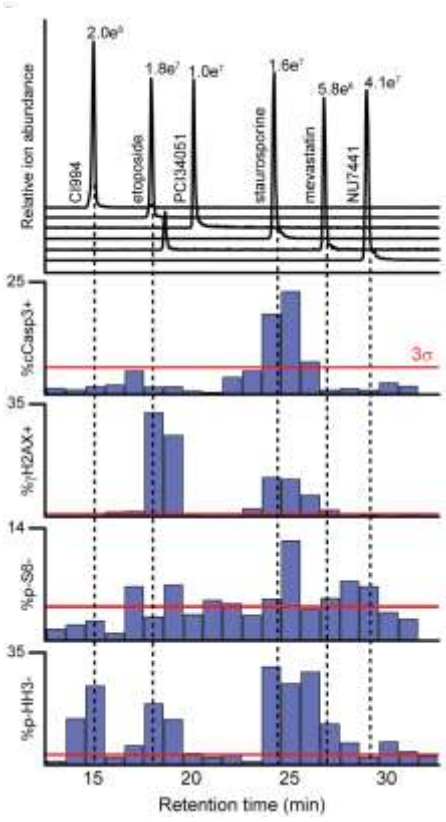


Figure 4.9. Validation with a mixture of pure compounds. Integration and validation of chromatographic arrays and FCB. Chromatographic arraying is performed using split flow HPLC/UV/MS with polarity switching mass scanning resulting in an array of highly characterized fractions. A mixture of six bioactive small molecules was arrayed onto a microtiter plate via split flow HPLC/MS fractionation and solvent was evaporated. Subsequently KG1 cells were added to the wells of the plate for incubation with the various toxicants. Cells were stained with Alexa-700 dye to indicate cell viability, fixed, permeabilized, barcoded, pooled and then immuno-stained with antibody-dye conjugates for DNA damage and apoptosis using anti-γH2AX and anti-cCasp3, respectively, and additional conjugates directed against phosphorylated Histone H3 (p-HH3), and phosphorylated ribosomal protein S6K (p-S6). The sample was analyzed via flow cytometry, and reconstructed bioactivity chromatograms were generated by gating on viable and marker positive (γH2AX and cCasp3) or marker negative (p-Histone H3 and p-S6) cells. Selective ion traces are aligned with bioactivity chromatograms.

Validation with a crude extract

The identification of bioactive molecules within complex cellular (e.g., microbial) metabolomes using MAM requires that barcoding and bioassay cytometric measurements be stable to potential interferences present under typical secondary metabolite-producing conditions, such as soluble extractable cellular metabolites, cell wall components, and spent growth medium species. The robustness of MAM was therefore tested by fractionating and analyzing a concentrated methanolic microbial extract generated from a *Streptomyces* strain grown in complex media and spiked with etoposide and staurosporine prior to chromatography. Prior to spiking, the extract possessed no measurable bioactivity. After fractionation, wells were evaporated, and KG1 cells were added to the plate and incubated for 16 hrs. Subsequent to fixation and permeabilization, cells were barcoded, pooled, and assayed using antibodies against γH2AX and cCasp3. Bioactivity chromatograms for these markers were generated from the de-barcoded data set and formatted for correlation analysis. Cells were effectively assigned as distinct populations to the 48 wells according to the dye-gradient selection, demonstrating no cytometric interference with FCB from extract components. Moreover, as shown in **Figure 4.10a**, plots of median fluorescence intensity for cCasp3, and γH2AX expression per debarcoded population generated bioactivity chromatograms for correlation analysis.

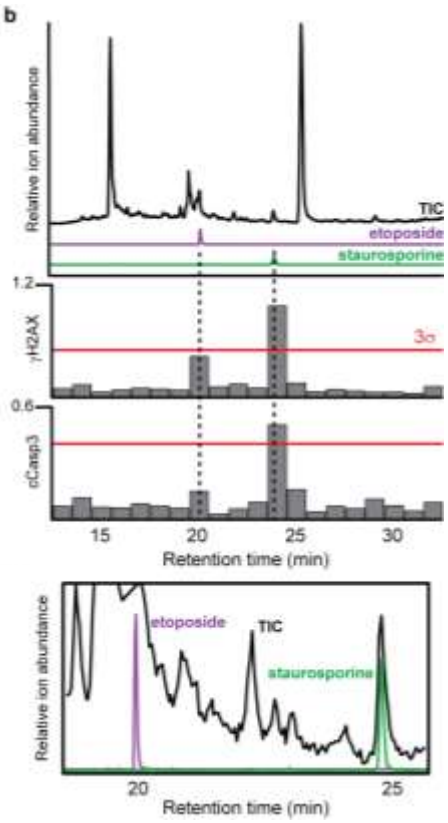


Figure 4.10: Validation of MAM using know compounds in a crude extract. An inactive extract was spiked with etoposide and staurosporine prior to fractionation. Bioactivity chromatograms were constructed using the arcsinh transformed median of all cells per well. the bottom panel shows an expansion of TIC, and EIC for etoposide and staurosporine. Red line denotes threshold for signal greater than 3 standard deviations of the readout from 4 blank control wells.

Importantly, although the EIC abundance of staurosporine and etoposide was below the threshold of the average intensity of the TIC (**Figure 4.10b**), both were the highest Pearson-correlating components in the bioactive fractions[10]. Despite the presence of high abundance products of cellular metabolism and media components, no additional potent cCasp3 or γH2AX-modulating activities were observed in the test metabolome.

Applications of Multiplexed Activity Metabolomics

MAM of apoptolidins and ammocidins

The MAM approach is also readily adaptable to titration and time point studies of purified compounds in microtiter plate formats. Concentration series of apoptolidin A, ammocidin A, 16-deoxyapoptolidin, and apoptolidin H were added to the first 4 columns of a 96 well plate alongside control wells containing DMSO. Approximately 200,000 Jurkat cells in 200 μL of media were added to each well and incubated for 16 hrs. Subsequently wells were viability stained (Alexa 700), fixed with paraformaldehyde, and permeabilized with methanol to allow for intracellular staining and biomarker detection by immunohistochemistry. Thusly prepared well contents were transferred to a second plate containing a gradient of two fluorescent dyes, Pacific Orange (PO) in decreasing concentration across rows, and Pacific Blue (PB) in decreasing concentration down columns yielding a distinct 2 color barcode for each well. After barcode staining and washing all wells were combined in a single tube and stained with antibodies for apoptosis (cCasp3), DNA damage response (γH2AX)22, cell cycle/chromatin status (p-Histone H3) and proliferative signaling (p-S6). Samples were then analyzed via fluorescent flow cytometry. Single cell measurements were assigned to individual experiment conditions by gating on biaxial plots of FSC vs PO fluorescent intensity (5 populations corresponding to columns 1-5) and then biaxial plots of FSC vs PB for each PO population (7 populations corresponding to rows A-G).

To survey changes in general cell status upon exposure to these macrolides, the percent of viable and marker+ cells was determined for each of the 32 populations (**Figure 4.11**).

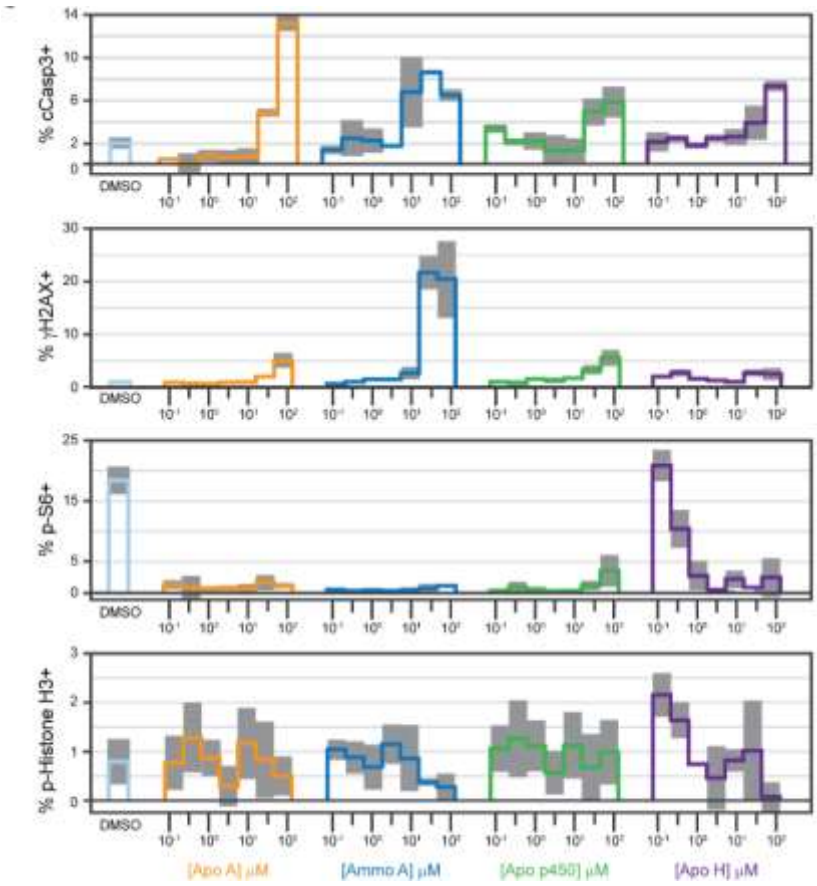


Figure 4.11: MAM with titrated macrolides. Structure activity relationships of 20/21 membered glycosylated macrolides using flow cytometric barcoding and multiparametric antibody analysis of central cell status checkpoints. Concentration response plots of debarcoded cell data in response to purified compounds (100 nM to 100 µM). Step graphs are based on the average of 2 replicate experiments and grey shaded regions are s.d.

The greatest apoptotic response was observed for apoptolidin A followed by ammocidin A. Strikingly, only ammocidin A elicited a strong DNA damage response, suggesting that despite structural similarity glycosylated macrolides may have disparate targets or unique polypharmacological effects. All analogs decreased p-S6 activity in agreement with previous growth inhibition experiments. Notably treatment with 1 µM apoptolidin H, which is ~50 times less cytotoxic compared to apoptolidin A, was sufficient to suppress

almost all S6 phosphorylation. None of the analogs had a significant effect on the number of cells progressing through M-phase as measured by p-Histone H3.

MAM of S. specus: finding metabolites within metabolomes with anti-cancer activity in human tissue

In addition to quantitating intracellular events and cell status immuno-markers, single cell characterization via cytometry facilitates the differentiation of cell types within heterogeneous mixtures based on cell size, shape, complexity, (via differential light scattering), and the detection of cell type-selective surface markers[11,12]. This enables characterization of the ways in which the components of metabolomic arrays effect molecular phenotypic changes in mixtures of cells, including primary cell preparations that more closely approximate a native cancerous microenvironment than pure immortalized cell lines. Acute myeloid leukemia (AML) patient bone marrow samples were selected as an advantageous system for MAM due to their beneficial cytometric properties and clinical significance. AML remains a deadly adult cancer, and treatments have not greatly improved the five-year overall survival rate, which is 21.3% overall and remains under 5% for patients who are 65 and older[13]. Bone marrow biopsies that are routinely obtained from patients being treated for AML contain a complex mixture of multiple cancer and normal cell types. These tissues are fully 'suspended', require minimal processing (e.g., disaggregation) for cytometric analysis, and contain a mixture of cell types representative of *in vivo* therapeutic contexts. Cytometric characterization of AML via immunophenotyping is widespread in the diagnosis and management of this disease, providing a strong basis for biomarker selection and analysis.

To test the ability of MAM to assess the effects of a bioactive metabolomic arrays against a heterogeneous cell mixture, microbial metabolomic arrays were incubated with cell preparations derived from AML biopsy samples from two separate patients. The patient samples used in this experiment represent two common underlying genetic mutational profiles occurring in AML. Patient 001 was a 23 year old female with a gene translocation (MLL-MLLT3) correlated to intermediate prognosis but without other tested common molecular mutations. Patient 015 was a 68 year old male with the FLT3 internal tandem duplication (FLT3-ITD) strongly associated with poor prognosis[14], but with otherwise normal cytogenetics. Subsequent to aspiration from bone marrow, red blood cells and platelets were removed from the patient samples via density gradient separation, resulting in bone marrow mononuclear cells containing a mixture of heterogeneous AML blasts and non-malignant myeloid and lymphoid cells and their progenitors[15]. These heterogeneous mixtures served as the response organism system for multiplexed cellular and biochemical analysis. For the microbial metabolomic array source organism, we selected an actinomycete strain designated *Streptomyces specus* that we had isolated from Blue Springs cave in Sparta, Tennessee. *S. specus* was of particular relevance, as it had been observed via dereplication analysis of HPLC/MS data to produce a family of anthracycline natural products related to the clinically employed AML drug daunorubicin, including baumycins, unusual natural acetal functionalized congeners,[11] and related anthracycline functional metabolites with apparent masses not previously reported.

After incubation of the two patient-derived samples with the metabolomic array, cell mixtures were barcoded using the two-color gradient as previously described and pooled. The cellular effects of the array were then analyzed via a 6-marker panel: Ax700 (viability),

cCasp3 (apoptosis), γH2AX (DNA damage), p-S6 (protein synthesis, growth, and mTOR-mediated metabolism)[16-18], p-Histone H3 (M phase/proliferation)[19,20], and CD45 (leukocyte cell surface marker)[21]. After initially gating for intact single cells, bone marrow mononuclear cells were gated by CD45 expression and side scatter (SSC) to distinguish lymphocytes, myeloid, and leukemia blast cells. The cells in each sample were debarcoded, and markers were quantitated resulting in 48 debarcoded well populations for blasts, myeloid, and leukocyte cell types with 6 readouts for each. Thus, the MAM platform generated bioactivity chromatograms for each cell status marker for each cell type, representing at least 864 unique raw bioassays in a single flow cytometric run, typically acquired in a few minutes. Combined marker analysis via biaxial gating yielded additional phenotypic assays. For instance, combining viability via +/- Ax700 and biomarker expression increased the effective number of distinct phenotypic assays per extract up to 36 assays per well, or 1728 per array.

In the case of the *S. specus* metabolomic array interacting with AML biopsy samples, and gating for the three major cell types, the strongest bioactivity was observed in the viable (Ax700-) and γH2AX+ and cCasp3+ subsets in both patients. The sample derived from Patient 015 contained the three readily discernable subpopulations of leukemia blasts, myeloid cells, and lymphocytes, which could be separately debarcoded to yield defined cell type response profiles in each well. Patient 001's sample was comprised of predominantly leukemia blasts and lymphocytes. Individual biaxial plots (**Figure 4.12 and 4.13**) were used to generate well bioactivity profiles based on set positivity thresholds.

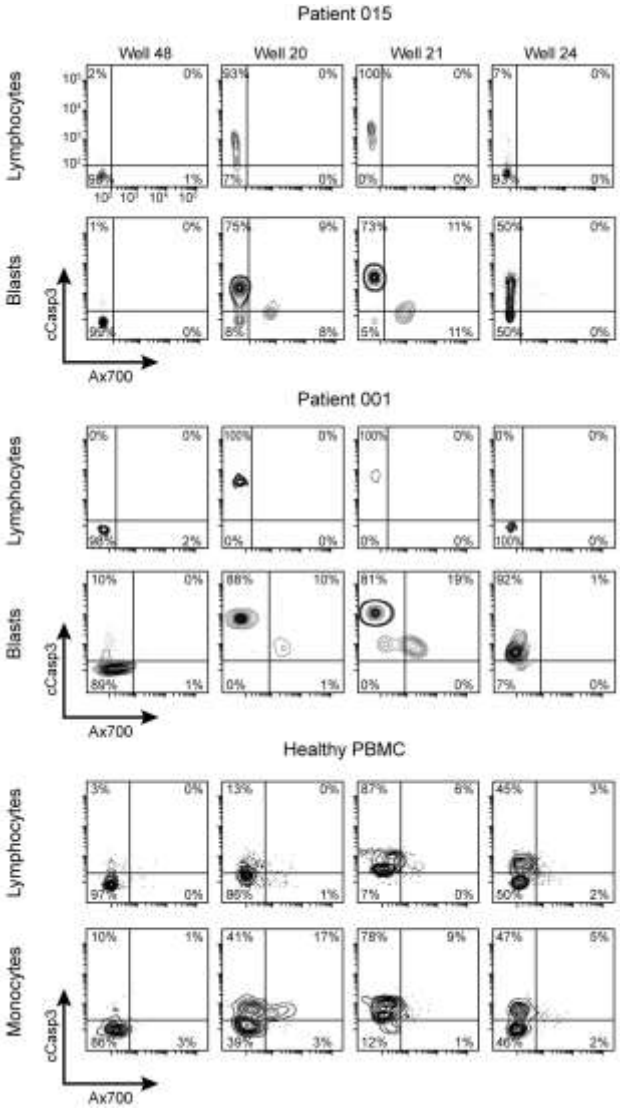


Figure 4.12. Biaxial plots of cCasp3 vs Ax700 from fraction wells containing specumycins. The predominant analogs of the specumycins elute from 20 to 21 minutes with well 21 containing the highest amount of specumycin A. Also shown are the biaxial plots from well 48 which contained only elution buffer and well 24 containing the *m/z* 1054.

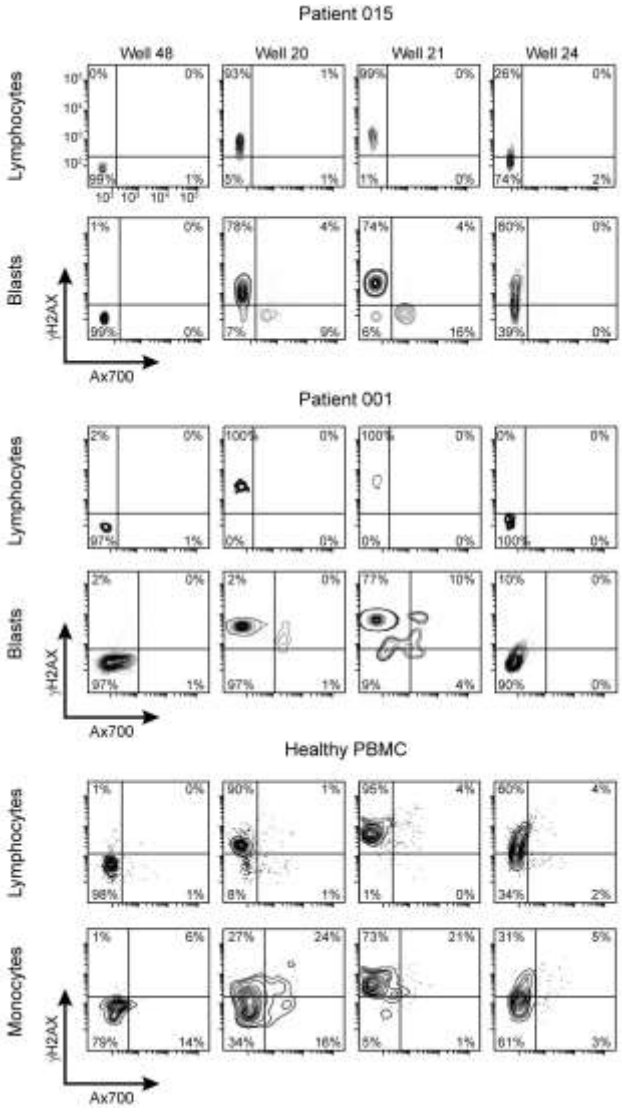


Figure 4.13. Biaxial plots of γH2AX vs Ax700 from fraction wells containing specumycins. Same as **Figure 4.12**. except corresponding plots for γH2AX are shown instead. Bioactivity profiles represent averages of thousands of single cell measurements and are highly reproducible between biological replicates (**Figure 4.14**).

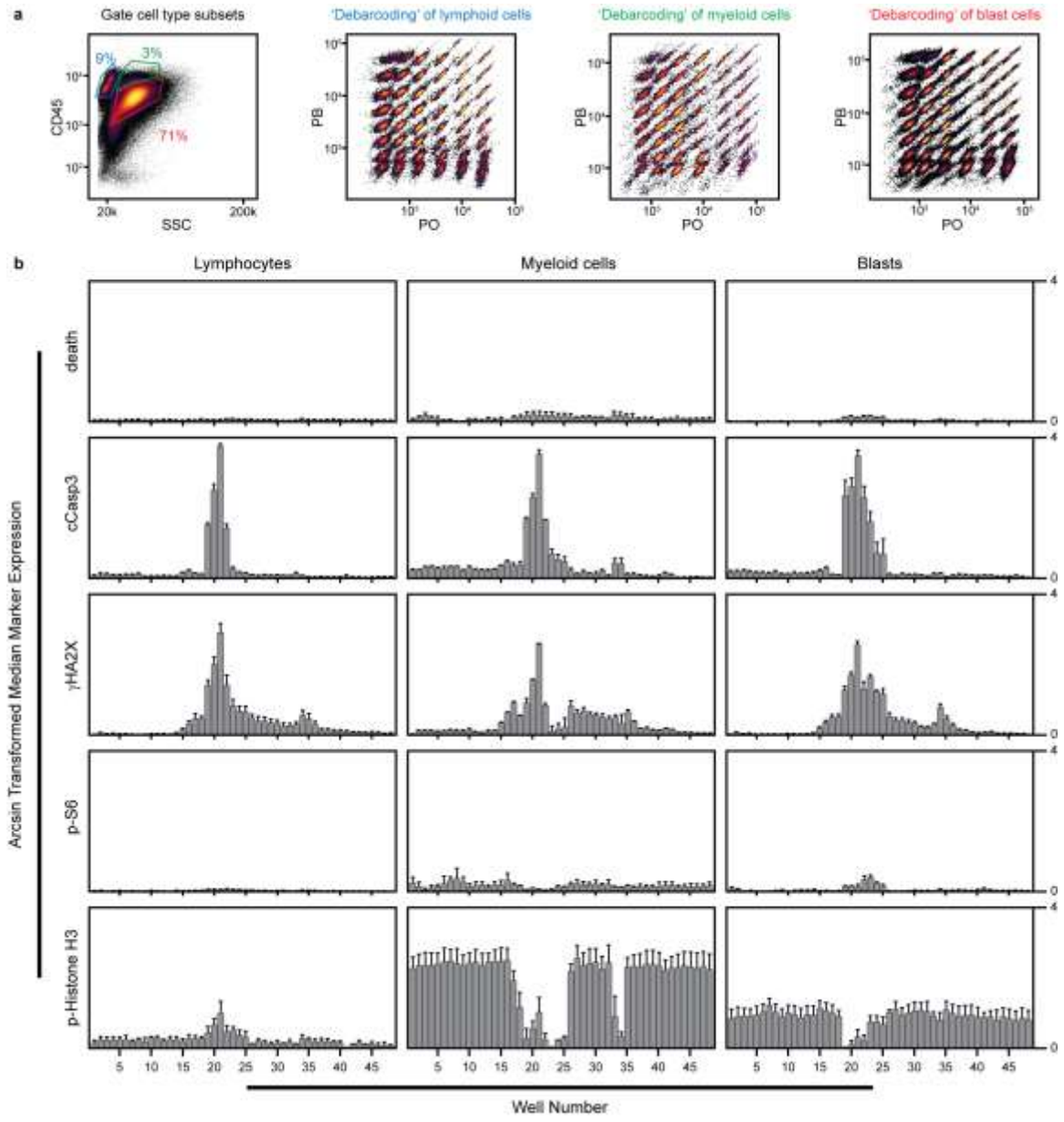


Figure 4.14. Comparison of replicates of MAM with primary patient samples. Samples from patient 015 were incubated with two replica plates of the fractionated *S. specus* extract. **a.** Gating and debarcoding of 3 cell subsets. **b.** Per well response for each marker. Bar graphs show the average of the arcsin transformed medians for each marker for the two experiments. Error bars are standard deviations.

A subset of these bioactivity chromatograms are shown in **Figure 4.15**, and indicate the presence of bioactive molecules in the *S. specus* extract, which can be preliminarily identified via comparison of EIC to bioactivity profiles.

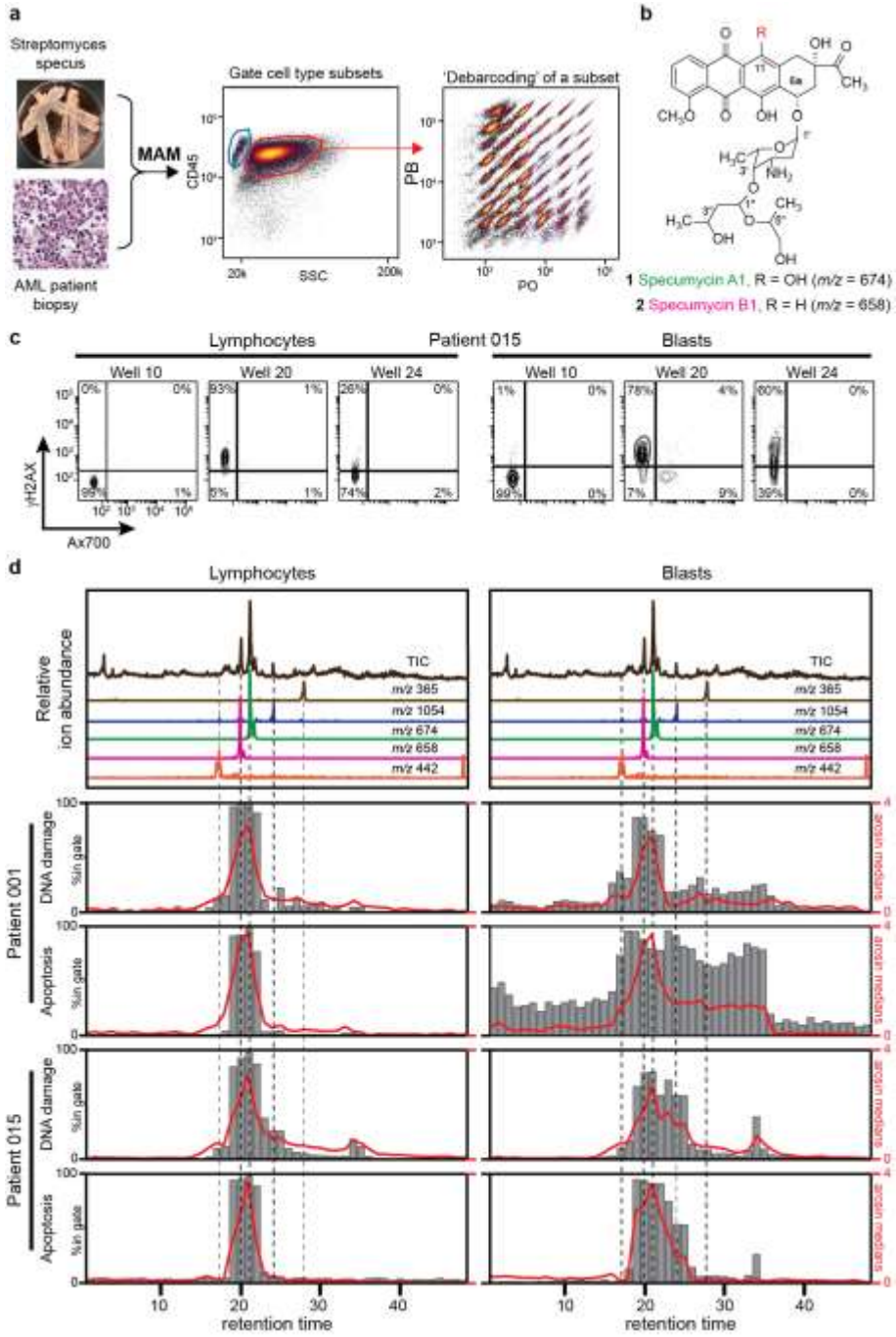


Figure 4.15: Bioactivity chromatograms from MAM of S. Specus. A structurally novel acetal-functional anthracycline selectively targets leukemic blast cells and not non-malignant lymphocytes within a human bone marrow biopsy. **a.** Chromatographic arraying of *Streptomyces specus* was performed using split flow HPLC/UV/MS with polarity switching mass scanning resulting in an array of highly characterized fractions from a crude extract of the baumycin producer *S. specus*. Primary cell preparations were prepared from AML patient biopsy. The metabolite array was incubated with heterogeneous cell samples, and the cells were then viability stained, fixed, barcoded, and stained with antibodies for biomarker and surface maker expression. Standard gating was performed using biaxial plots of CD45 expression vs SSC was used to determine cell type subsets (blue gate: lymphocytes, red gate: blasts) that were each individually analyzed for γH2AX and cCasp3 expression. **b**. Structures of Specumycin A1 and B1 **c**. Biaxial plots of selected wells gated for lymphocytes and leukemia blasts. **d**. Total ion current and selected extracted ion currents of metabolites within the metabolome of *S. specus* correlating to bioactive wells from assays against 2 patient samples. Bar graphs of percent of cells in upper left quadrant of marker/viability gate (marker positive and viable cells). Solid red line is the arcsinh transformed median of the marker.

Two observations result from this data set. First, apparent cell type selectivity was demonstrated for several features in the metabolomic array. For instance, a 2.5 and 47-fold increase in selective blast-targeting bioactivity vs leukocytes was observed eluting in wells 17, and 24, respectively. Second, patient-specific activities were evident in bioactivity profiles. Examination of HPLC/MS and bioactivity profiles of well 17, containing the most abundant m/z = 442, revealed a 30-fold increase in apoptosis and 5-fold increase in DNA damage in patient 001 versus patient 015. A similar trend was observed in later eluting wells containing anthracycline chromophores. Notably, patient 015 possessed the FLT3(ITD) phenotype, which is an internal tandem deletion in kinase encoding gene *FLT3* demonstrated to confer resistance to anthracyclines[14]. Therefore, these observations are consistent with substantially enhanced resistance to anthracyclines in patient 015.

Isolation of specumycins

To validate bioactive features putatively identified using MAM, most abundant correlating mass features were isolated. The most potent bioactivity peak was observed in well 21 and correlated to the most abundant eluting anthracycline (*m/z* = 674), termed specumycin A1. Specumycin A1 was isolated in scale up fermentations and its structure determined by multidimensional nuclear magnetic resonance (NMR) experiments (**Table 4.1**). The planar structure of specumycin A1 is identical to the structures of baumycin A1/2, which contain an unusual acetal appending the 3'-O-methyl on the daunosamine sugar[22]. The next most abundant feature, and the primary feature in well 20, was specumycin B1 (**Table 4.2**), a previously unreported an 11-deoxy congener. Specumycin B1 was observed to be as active to A1 under assay conditions but 3-fold less abundant, suggesting a potentially more potent congener.

Comprehensive isolation of low abundance bioactive species is beyond the scope of this study. However, cell type and patient-specific responses identified by MAM, such as bioactive metabolites demonstrating enhanced activity against a FLT3(ITD) AML sample and selective activity for leukemia cells (e.g. well 24, *m/z* = 1054, **Figure 4.15**), demonstrate the potential of this platform for performing preliminary analysis and prioritization of activity differences within a natural product family for common AML subclasses.

Position	^{1}H NMR δ_H *(J in Hz)*	^{13}C NMR δ_C	^{1}H - ^{1}H COSY NMR δ_H	^{1}H - ^{13}C HMBC NMR δ_C
1	7.89 d (5.9)	120.1	7.71	187.2, 136.1, 121.6, 118.9
2	7.71 t (7.8)	136.0	7.89, 7.32	161.8, 136.1
3	7.32 d (7.8)	118.9	7.71	187.2, 161.8, 120.1
7	5.17	69.8	2.27, 2.08	135.3, 77.6
8	2.27 d (14.6), 2.08 dd (14.8, 4.2)	35.4	5.17, 3.16	77.6, 69.8, 33.8
10	3.16 d(18.6), 2.89 d (18.6)	33.8	2.27	212.6, 156.3, 135.1, 77.6, 35.4
14	2.40 s	25.4		212.6, 77.6
1'	5.5	100.5	3.82, 1.93	67.8, 46.8
2'	1.93-1.88 m	32.7	5.50, 3.35	46.8
3'	3.35	46.8	3.82, 1.93	74.2
4'	3.82	74.2	5.50, 4.15, 3.35, 1.93	100.5, 46.8, 32.7, 17.6
5'	4.15 q (6.4)	67.8	3.82, 1.29	100.5, 74.2, 46.8, 17.6
6'	1.29 d (6.7)	17.6	4.15	74.2, 67.8
1"	4.84	101.2	1.86	
2"	1.86-1.81 m	42.4	4.84, 4.23	101.2, 64.4
3"	4.23 m	64.4	1.86, 1.20	
4"	1.20 d (6.2)	24.2	4.23	64.4, 42.4
5"	3.76	73.0	3.54, 3.50, 1.13	73.0, 16.6
6"	3.54 dd (11.8, 2.1) 3.50 dd (11.8, 7.1)	66.5	3.76	
7"	1.13 d (6.3)	16.6	3.76	73.0, 66.5
4		161.8	7.71, 7.32	
5				
6				
9		77.6	5.17, 3.16, 2.89, 2.40	
11		156.3	3.16, 2.89	
12		187.2	7.89, 7.32	
13		212.6	3.16, 2.89, 2.40	
4a		121.6	7.89, 7.32	
5a				
6a		135.3	5.17	
10a		135.1	3.18, 2.89	
11a				
12a		136.1	7.89, 7.71	
4-OMe	4.01	57.2	7.32	161.8

Table 4.1 NMR Shift Assignments for specumycin A1

Position	^{1}H NMR δ_H *(J in Hz)*	^{13}C NMR δ_C	^{1}H - ^{1}H COSY NMR δ_H	^{1}H - ^{13}C HMBC NMR δ_C
1	7.88 d (7.6)	120.5	7.71, 7.32	182.8, 120.9, 118.8
2	7.71 t 8.0)	136.1	7.88, 7.32	161.3
3	7.32 d (8.4)	118.8	7.88, 7.71	188.7, 161.3, 120.5
7	5.14	69.8	2.25, 2.09	130.0, 77.8
8	2.25 d (14.6), 2.09 dd (14.8, 4.1)	36.1	5.14, 3.00	130.0, 77.8, 69.8, 39.7
10	3.15 d (17.4), 3.00 d (17.4)	39.7	7.41, 2.25	212.6, 143.1, 130.0, 120.0, 77.8, 36.1
14	2.37 s	25.4		212.6, 77.8
1'	5.48	99.8	3.85, 2.06, 1.99	69.8, 67.1, 46.8
2'	2.06, 1.99	30.2	5.48, 3.56	
3'	3.56	46.8	3.85, 2.06, 1.99	
4'	3.85	72.6	4.17, 3.56	100.7, 46.8, 30.2
5'	4.17 q (5.9)	67.1	1.27	99.8, 72.6, 17.5
6'	1.27 d (6.6)	17.5	4.17	72.6, 67.1, 30.2
1"	4.78	100.7	1.87, 1.80	
2"	1.87, 1.80	41.4	4.78, 4.25	100.7, 64.0
3"	4.25	64	1.87, 1.80, 1.18	
4"	1.18 d (6.1)	24.4	4.25	64.0, 41.4
5"	3.7	72.6	3.62, 3.51, 1.14	100.7, 66.1
6"	3.62 dd (18.7, 8.0), 3.51 d (10.8)	66.1	3.7	72.6
7"	1.14 d (6.0)	16.2	3.7	72.6, 66.1
4		161.3	7.71, 7.32, 4.00	
5		188.7	7.32	
6				
9		77.8	5.14, 3.15, 3.00, 2.37, 2.25, 2.09	
11	7.41 s	120	3.15, 3.00	188.7, 182.8, 130.0, 39.7
12		182.8	7.88	
13		212.6	3.15, 3.00, 2.37	
4a				
5a				
6a		130	5.14, 3.15, 3.00, 2.25, 2.09	
10a		143.1	3.15, 3.00	
12a		120.9	7.88	
4-OMe	4.00 s	57.1	7.32	161.3, 118.8

Table 4.2. NMR Shift Assignments for specumycin B1

MAM of Nocardiopsis. sp. FU40

The cell targeting potential of secondary metabolites within metabolomic arrays in the anthracycline-resistant phenotype sample (015), was further explored employing the soil actinobacterium *Nocardiopsis* sp. FU40 as a source organism. This strain produces a family of bioactive compounds called apoptolidins (A – H)[23], which are cytotoxic glycosylated macrolides, and a pair of cytotoxic glycosylated polyene macrolactams, ciromicins A – B[24]. Thus, a metabolome array was generated from *Nocardiopsis* sp. FU40, and AML Patient 015-derived anthracycline resistant cell preparations were incubated with the array and subjected the samples to MAM analysis.

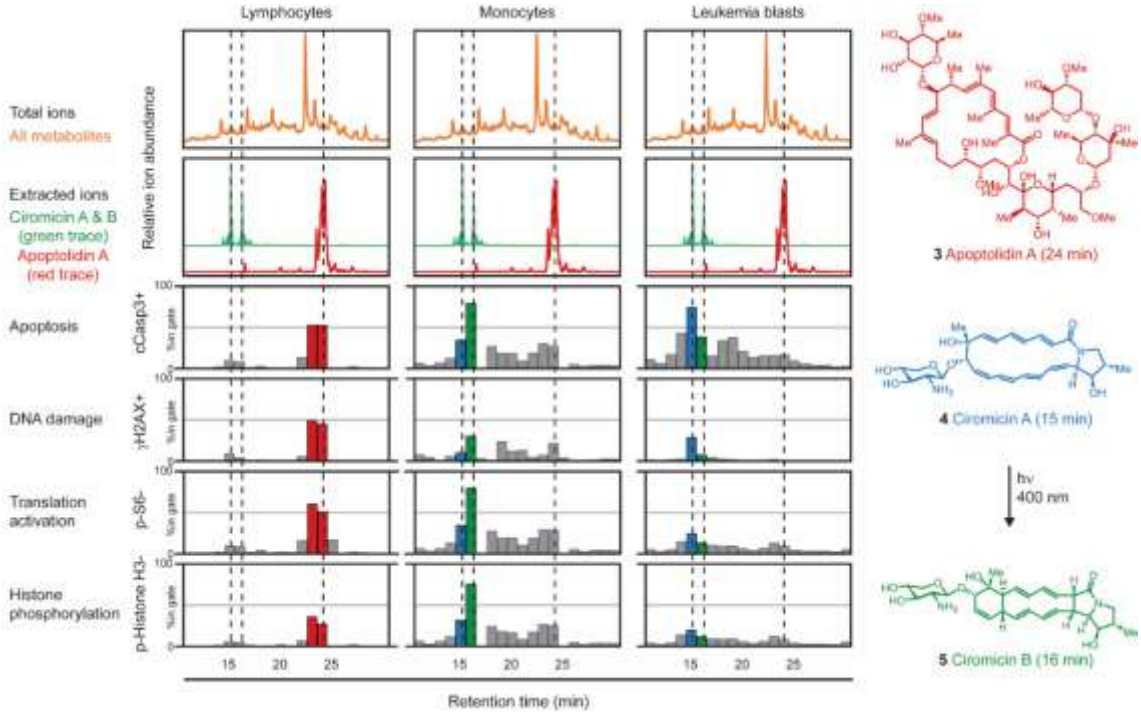


Figure 4.16. Bioactivity chromatograms from MAM of N. FU40. An optochemical cell selectivity switching natural product in *Nocardiopsis* revealed by primary cell MAM. Top metabolome row (orange) shows total ion current and extracted ions for ciromicins A and B (*m/z* = 515, 15 and 16 min) and aptoptolidin (*m/z* = 1129.5, 23.4 min) with their elution times shown in dotted lines. The next four rows show selected bioactivity chromatograms from a single flow experiment, which were generated by adding an aspirated preparation of bone marrow mononuclear cells from an AML patient, barcoding, immunostaining, and de-barcoding. The immunostaining panel contained, CD45 (leukocyte-common antigen), cCasp3 (cleaved caspase), γH2AX (DNA damage), p-Histone H3 (cell cycle marker upregulated during mitosis), and p-S6 (marker for active translation). Histograms for each marker for highlighted wells are shown in **Figure 4.17**.

Shown in **Fig. 4.16** is a selection of bioactivity profiles generated from this single data set indicating how the arrayed metabolome obtained from *Nocardiopsis* sp. FU40 can be mined for bioeffectors that have selective activity against different cell types present in an AML patient. For instance, apoptolidins selectively induced caspase-dependent apoptosis in lymphocytes, whereas ciromicins induced apoptosis most prominently in leukemic cells. Similarly, apoptolidin A induced γH2AX, apoptosis and decreased p-Histone H3 signaling

selectively in lymphocytes, and ciromicins induced more DNA damage in monocytes and blast cells. Taken together, these data demonstrate the identification of differential cell targeting of secondary metabolites against primary cell mixtures in the background of an extracted microbial metabolome. In contrast to the specumycins, ciromicins demonstrate potent selectivity for blasts in comparison to lymphocytes in the anthracycline resistant phenotype.

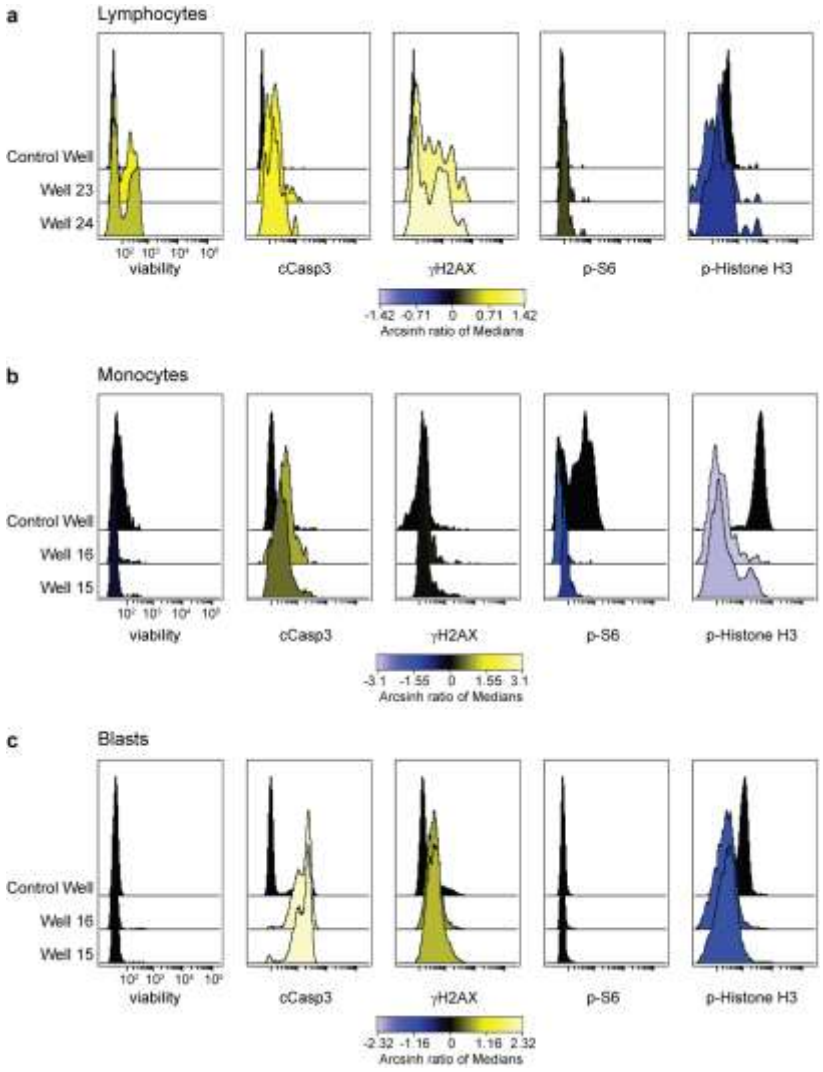


Figure 4.17. Histogram plots of active wells from MAM of N. FU40. Histogram plots of each marker in control wells and highlighted wells in **Figure 4.16**. **a** Marker distribution in wells 23 and 24 which had the maximal response in lymphocytes and contained the apoptolidins. **b** and **c** Marker distribution in wells 15 and 16 which contained the ciromicins which induced the largest response in monocytes and blasts

Validation of Cell Subset Targeting

Expansion of the bioactivity chromatogram in the region of ciromicin elution revealed that the isobaric metabolites ciromicin A, and ciromicin B were resolved into separate wells with strikingly distinct biological phenotypes. Specifically, ciromicin A displayed maximal apoptosis markers in leukemia blast cells whereas its photo-isomerization product ciromicin B stimulated monocyte apoptosis. We recently reported the discovery of ciromicins in an apoptolidin polyketide synthase knockout strain of *Nocardiopsis* sp. FU40, and demonstrated that ciromicin B is the product of an unexpected visible light-initiated 12-π electron photo-isomerization of ciromicin A58. The identification of ciromicins here in the wild-type monoculture of *Nocardiopsis* sp. FU40 was surprising because it is produced in low levels in the wild-type strain. Thus, the sensitivity of the MAM platform using primary cells was capable of effectively identifying the bioactivity of this low abundance secondary metabolite family, and the modest resolution of 48-well binning of fractions was sufficient to resolve bioactivities of closely eluting species.

The discovery of putative cell type-specific cellular responses to ciromicin isomers using MAM may be considered as a primary screening 'hit', describing a multidimensional response of a metabolomics fraction with associated correlation coefficient-ranked metabolite features. To validate the unusual photochemically triggered modulation of primary cell selectivity, pure ciromicins A and B were isolated from scaled up cultures and assayed them against the same biopsy sample using an enhanced panel of 29 cell surface markers that classifies all myeloid cell populations when paired with unsupervised machine learning tools. The viSNE algorithm, which allows robust identification of both non-malignant and leukemia cell subsets[15], was applied to datasets collected by mass cytometry

after a 48 h treatment of patient-derived bone marrow mononuclear cells (BMMC) with ciromicin A, ciromicin B, or DMSO. Shown in **Figure 4.18** are viSNE maps showing the overall changes in the cellular landscape of the primary BMMC after this treatment. Proximity in viSNE space corresponds to similarity in cell type identity while differences in immunophenotype drive separation of cells (dots) on a viSNE map. To quantify the overall shifts in cellular subsets, gates were drawn on the viSNE map corresponding to prominent populations based on abundance (**Figure 4.18a**). The relative abundance of phenotypically distinct cell subsets present in different treatment conditions and the enriched features of these populations were characterized Per cell marker expression and median values allowed assignment of cellular identity to populations (**Figure 4.19**). Changes in the relative abundance of each population demonstrated that photoisomerization polarizes overall cellular immunophenotype within leukemia cells. Based on subset gating within viSNE, ciromicin B reduced the relative abundance of leukemia stem cells and hematopoietic stem cells and smaller blast subsets, while ciromicin A reduced the largest blast subset (subset 9, **Figure 4.18a**). Both isomers had comparatively little impact on lymphoid cells, which were comprised largely of CD4+ and CD8+ T cells (subsets 2 and 4, **Figure 4.18a**). Next, marker enrichment modeling (MEM) was used to characterize feature enrichment in comparison to a population of CD34+CD38-Lin- cells within the leukemia sample. MEM identified changes in populations between ciromicin A and ciromicin B, including greater enrichment for CD13, a marker of myeloid differentiation, and CD43, or sialophorin, which commonly expressed on more mature myeloid cells such as granulocytes and monocytes, after treatment with ciromicin A within a major leukemia population (subset 9, **Figure 4.18b**). Overall, the pattern of cell type

selectivity expanded the depth of the initial screen, validating MAM's ability to identify selectively bioactive compounds.

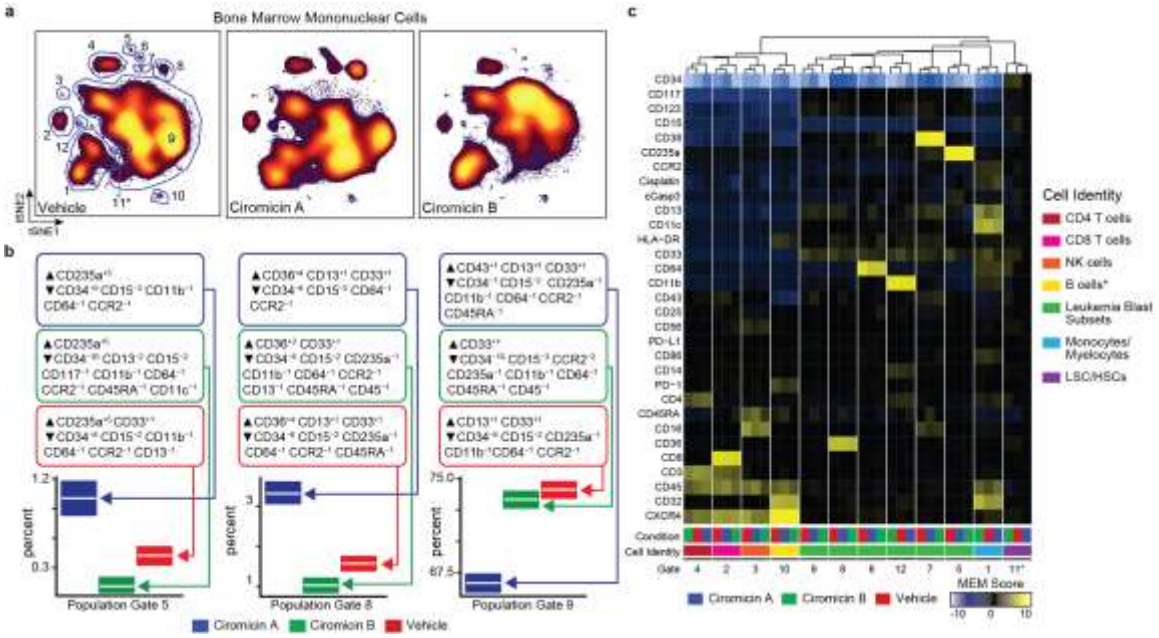


Figure 4.18. In depth profiling of ciromicins by mass cytometry and viSNE. Photochemical isomers ciromicins A and B selectively target different cell subsets within the heterogeneous mixture of patient biopsy cells. Mass cytometry uses DOTA-chelated metals detected by ICP-MS to eliminate spectral overlap, expanding the feature range to 29 antibody-quantified features per cell. **a.** viSNE maps of 20,000 individual cells from each condition are organized according to differences in their surface marker profiles for each treatment condition. **b.** MEM labels for 3 blasts subsets and plots of population prevalence with observed prevalence in white and 95% binomial confidence interval represented by box. **c.** Marker enrichment modeling (MEM) was used to characterize major populations within the samples and highlight differences in marker expression relative to a gated population of phenotypic hematopoietic stem cells (gate 11). Heat maps of hierarchal clustered MEM labels reveal subsets specific differences and cellular identification. Heat maps of median marker expression in patient sample 015 are shown in **Figure 4.19**.

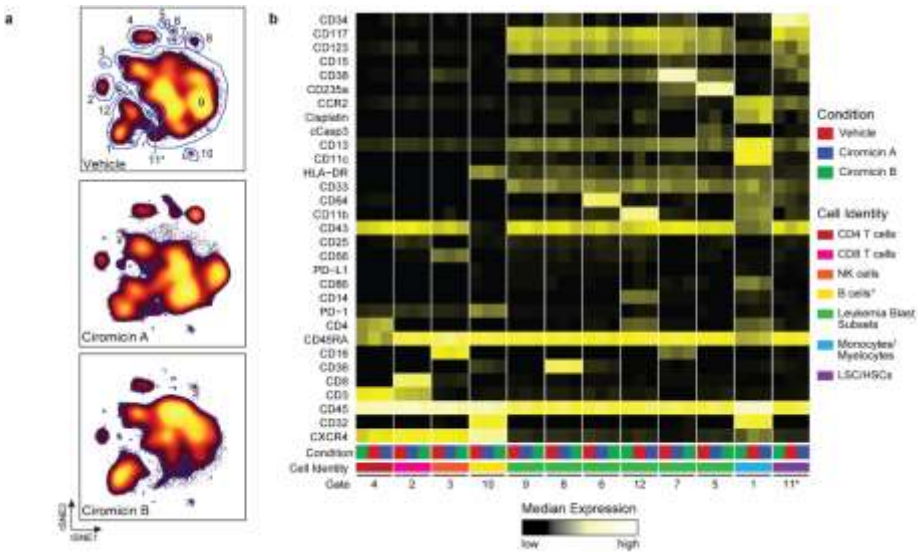


Figure 4.19. Median marker expression of viSNE populations. 12 major populations were identified after viSNE analysis and gated. **b** Median marker expression for each gated population after treatment with vehicle, ciromicin A or B.

Finally, in order to provide a comparison to a healthy control and demonstrate reproducibility, we produced fresh fermentation cultures of the ciromicin producing organism, re-isolated and purified ciromicin A and B, and performed concentration response experiments with both compounds against new aliquots of both patient 015 and healthy PBMC samples. The results are shown in a **Figure 4.20**.

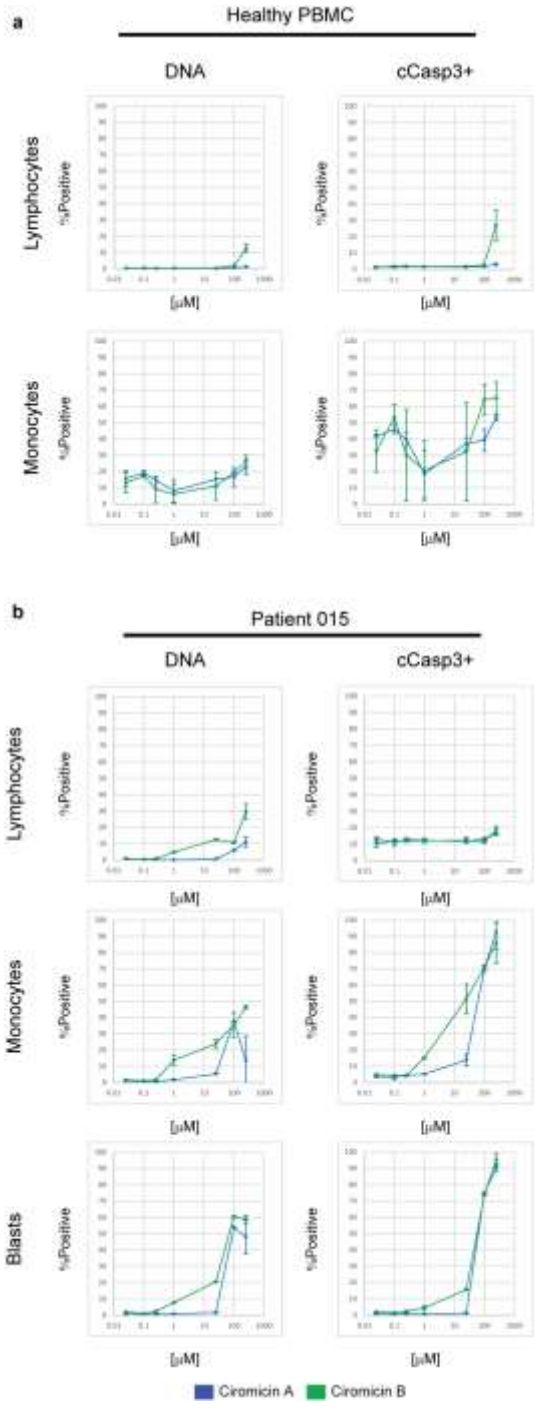


Figure 4.20. Titration of ciromicins against primary AML and PBMCs. 24 hour titration of ciromicins assayed against **a** PBMCs from a healthy donor and **b** a AML patient sample.

Conclusions

Cancer is challenging to study and to therapeutically manipulate, due in part to the complexity of cell signaling processes affected by pharmacological interactions, and system heterogeneity as seen in the polyclonal nature of cancer cells, the complexity of the supporting stroma, and the infiltrating immune cells[1,2]. MAM as implemented here provides a generalizable system to link metabolomic feature data from one organism or system to functional targets or their causally related networks within another heterogeneous cellular environment. A key feature of the cytometric strategy underpinning MAM is its ability to analyze heterogeneous mixtures of cells, which more closely approximate a native cellular milieu than immortalized cell lines, using multiplexed markers of cell status and type. Specifically, MAM was employed here to study the interkingdom interactions of metabolomes of two secondary metabolite-producing soil bacteria with primary cell preparations from two phenotypically distinct patient-derived AML cell samples. The combination of metabolomics, single cell biology, and cheminformatics used here identified biologically active secondary metabolites produced at low levels that mediate apoptosis, DNA damage, and cell signaling in a cohort of cells present in AML patient's bone marrow samples. In this primary cytological screen, differential activities were identified for secondary metabolites present within complex and concentrated microbial extracts.

Minor structural analogs are often dismissed as uninteresting and abandoned before a thorough evaluation of biological activity is undertaken leading to the loss of potentially useful tool compounds for chemical biology and therapeutic development. While structure activity relationships are typically used to improve binding affinity to a single target or

pharmacological properties, structural changes can have off targets effects leading to potential polypharmacological effects. Our assays of the apoptolidins and ammocidins utilizing activity metabolomics reveals relatively minor structural changes that profoundly affect biological activity beyond changes in potency of overall cytotoxicity in a family of glycosylate macrolides previously thought to share a common mechanism of action.

The cave-derived bacterium *Streptomyces specus* is a producer of multiple compounds that share the anthracycline core of daunorubicin, which is used in combination therapy with nucleoside analog cytarabine as the standard of care in the treatment of AML, but differ in decorating glycosides[3]. Specumycins described here are daunorubicin variants similar to baumycins, appended with an unusual acid labile-acetal moiety on the 3´-hydroxyl of daunosamine[4], and are reported to demonstrate broad cytotoxicity comparable to that of daunorubicin[5]. Despite its clinical significance, daunorubicin is actually a low abundance biosynthetic intermediate *en route* to baumycin in most daunorubicin producers, and is typically isolated by acid catalyzed degradation of baumycin glycosides[4]. However, though being the major product of most daunorubicin biosynthetic pathways, the potential role of the baumycin acetal moiety in cytotoxicity and cell targeting has remained untested prior to this study. Applying MAM to the metabolomic array of *Streptomyces specus* revealed activities of a spectrum of specumycin polyketides and related metabolites against divergent primary cell phenotypes. Along with the discovery of the previously unreported and more potent compound specumycin B1, these data suggest previously unnoticed potential for the 3'-acetal functional in AML anthracycline therapy. Validating the observed bioactivity trends, the two most abundant features demonstrated potent activity against leukemia blasts and leukocyte cells, and were isolated and structurally elucidated.

Numerous less abundant species displayed remarkable differential cell-type targeting between patients suggesting an untapped potential for discovery of more selective pharmacological agents within the anthracycline family in biosynthetically competent actinomycetes strains.

Nocardiopsis sp. FU40 was selected as a subject for the MAM platform as its metabolomic array is complex, both in terms of the sheer number of apoptolidin analogs it produces (denoted A - H), and in its capacity to simultaneously produce polyene macrolactams and aromatic polyketides that were previously reported to possess moderate to potent cytotoxicity against cell lines. Applying MAM to test *Nocardiopsis* arrays against AML primary cell preparations successfully deconvoluted apoptolidins from ciromicins and revealed distinct cell-targeting phenotypes. Apoptolidin A and its isobaric analogs present in the extract (isoapoptolidins A and G) correlated to the most potent lymphocyte-targeting activity across all markers. The induction of cCasp3 in lymphocytes is consistent to prior studies of this compound performed in cell lines, which also present evidence in support of mitochondrial FoF1-ATPase-targeting within this family[6]. Other apoptolidin congeners are generally chromatographically dispersed from apoptolidin A, but did not display this degree of activity. Notably, the apoptolidins only nominally affected marker expression in leukemia blasts and monocytes, demonstrating how MAM readily identifies first pass cell-targeting activity in primary tissue samples. The distinctive blast/myeloid cell type targeting observed for ciromicins A and B was notable, and careful examination of their elution region revealed a remarkable switch of cell specificity between blast and non-blast myeloid lineages for the two compounds.

As a follow up to using MAM as a primary assay for lead discovery, mass cytometry was used as a secondary validation and deep cell profiling assay. A 29-marker mass cytometry panel was used to classify the cellular effects of purified ciromicins A and B on subsets in primary cell preparations. Mass cytometry revealed changes in differentiated immunophenotypic subsets and demonstrated that visible light induced photoisomerization of ciromicin A to B induces wholesale shifts in cell type targeting and indicating the importance of aglycone structure and geometry to the mechanism of action of this family of macrolactams. For instance, bicyclic Michael-acceptor containing ciromicin A exerted its greatest influence on the largest subset of AML cells, whereas tetracyclic potentially less electrophillic ciromicin B targeted stem-like myeloid progenitors, a subset that may be beneficial to address in therapy.[7,8] Mass cytometry also revealed that ciromicins target leukemia blasts in a patient with an anthracycline-resistant leukemia phenotype and, unlike anthracyclines in the previous study, have little negative effect on lymphoid cells. Finally, mass cytometric findings, performed in concert with MAM using patient samples, validated and provided a deeper profiling of bioactive compounds discovered. Overall, the multiplexed single cell approaches used here represent a paradigm shift in comparison to typical discovery efforts using monoclonal immortalized cell lines or other research models that do not accurately reflect the cell diversity and composition of primary human tumors and leukemic tissues. That ciromicins A and B represent photo-switching natural products with distinct cell subtype-targeting phenotypes provides potential tools for investigating the pharmacology of this family and the effects of targeting cell sub-types. Notably, molecular photo-pharmacological switches currently find broad application toward

understanding cellular function by leveraging the spatiotemporal control afforded by such compounds[9].

In summary, a general method is demonstrated for searching preliminary structure activity relationships in secondary metabolite families in producing organisms, without the need for compound isolation, and provide insight into how bioactive lead compounds affect diseased and normal cell types in major patient phenotypes using clinical samples. Given that there are a limited number of distinct clinical subsets, automated cytometric analysis of untargeted metabolomic inventories against sets of relevant patent phenotypes provides a process for 'personalized' natural product discovery. This is a proof of principle study of a viable drug discovery platform. In a full scale implementation, cells derived from multiple patients, including cells derived from healthy individuals, would be necessary to realize the full scope of lead-compound preclinical assessment. While applied here for the case of identifying bioactive secondary metabolites within metabolomes, the MAM platform enables the discovery of cellular responses to molecular inventories, regardless of sources. Given the importance all chemical communications in mediating life processes within and between organisms, a generalizable method for identifying functional roles for metabolites has significant potential in applications spanning a broad range of applications in cellular chemical biology.

Future direction: FCB of bacteria

We have been exploring the potential application of fluorescent cell barcoding to bacterial cells to allow for MAM based assays with bacteria serving as the response organism. Preliminary results (**Figure 4.21**) indicate the feasibility of this approach. 12 liquid cultures of *Staphylococcus aureus* were alternatively treated with antibiotic or DMSO and

then stained with a 3x4 barcode. However reduced staining efficiency due potentially to small cellular size or membrane interference, has limited the total number of barcode levels on a single channel. New barcoding strategies and staining protocols are being developed to overcome this issue.

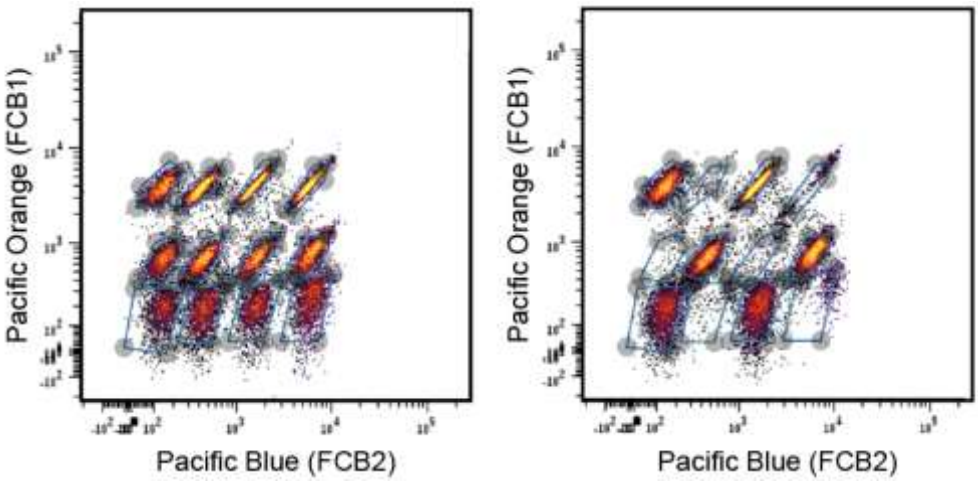


Figure 4.21: Barcoding of *S. aureus*. 2 dye barcode of *S. aureus* recovers antibiotic checkerboard. The left panel shows a barcode of 12 untreated samples (three levels of Pacific Orange and four levels of Pacific Blue). The right panel shows the barcode of a checkerboard antibiotic treatment.

Future direction: MAM screening of cave organisms

Currently the Bachmann lab collection of cave organisms is being evaluated using a two phased MAM screening approach. Crude extracts are first evaluated for dose dependent responses for viability, apoptosis or DNA stress. Extracts which show a favorable response are then fully fractionated and evaluated with broader panels.

Future direction: Automated data analysis pipeline

Given the multiple combinations of biomarkers and the potential extension of FCB to more than two dimensions, we have been prompted to develop an automated means for data

analysis of MAM experiments. One of the most time-consuming steps is the manual gating of each barcode population to debarcode the samples. Additionally, samples contain cells of varying size and type which are labeled with the barcode with varying efficiency which can cause overlap between populations making it difficult to determine where to draw gates to separate the populations. To address this problem and to automate the process of debarcoding the data we are developing DebarcodeR, an R package that allows users to correct for cellular variability in their barcoded data using multiple regression and then classify each barcoded cell to an experimental sample using mixture modeling. DebarcodeR currently can be run as a Shiny app using local files or using the Cytobank API.

The first module (**Figure 4.22**) lets the user specify how their data is barcoded and the cellular populations they want to debarcode. Corresponding data points are then passed to a regression module. Uptake of FCB dyes has been shown to be related to cell size which

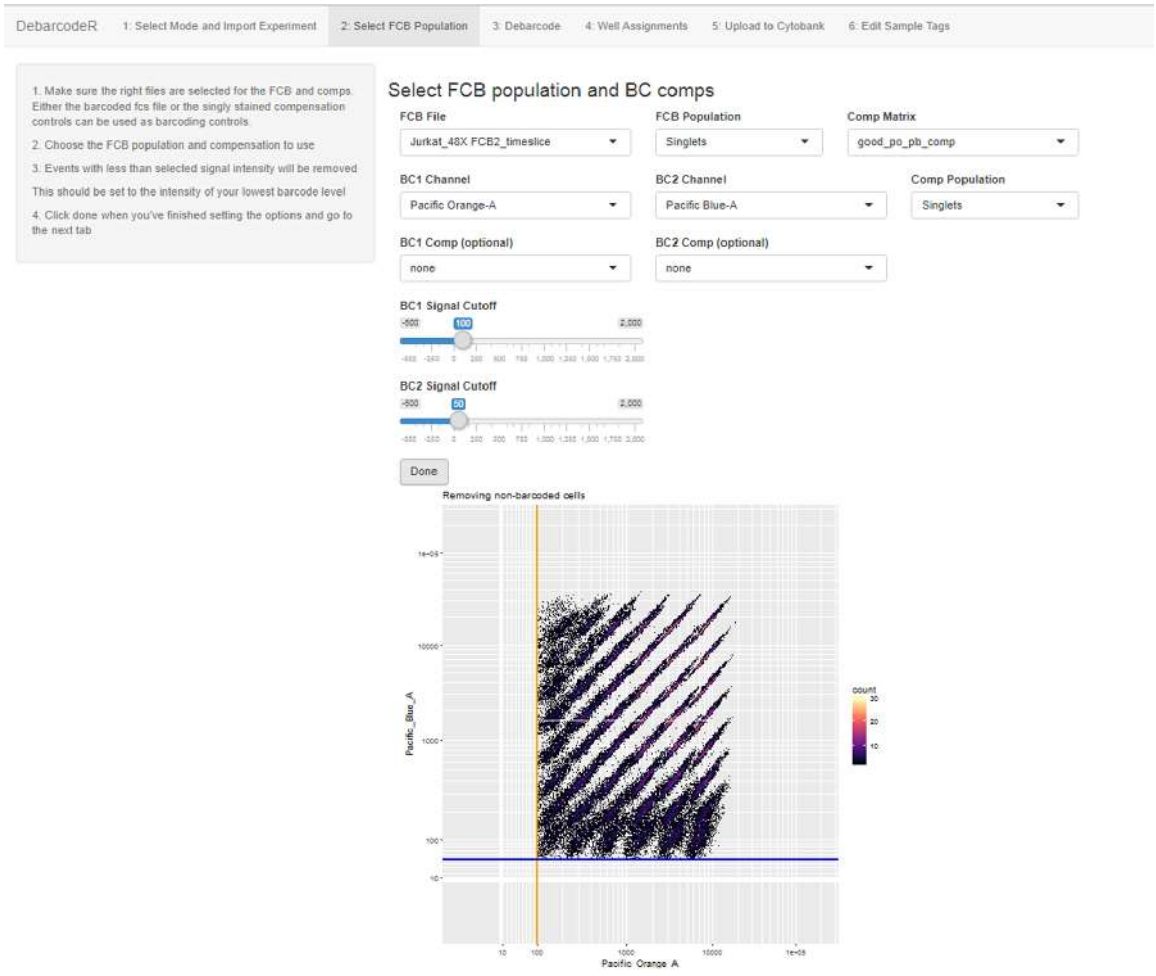


Figure 4.22: Screen shot of DebarcodeR setup module.

can be corrected for by fitting a regression model on FSC and SSC to resolve apparent overlap between adjacent barcoding levels[10]. Data sets are next passed to the debarcoding module where each barcoding channel is fit to a mixture of skew normal distributions according to the number of specified levels. Cells are then assigned to the modeled distribution under which they most likely fall. In the case of very large data sets a subset of data points can be used to fit the model for improved performance. After the modeling is finished for the first barcoding channel a histogram plot of the fit is returned (**Figure 4.23**)

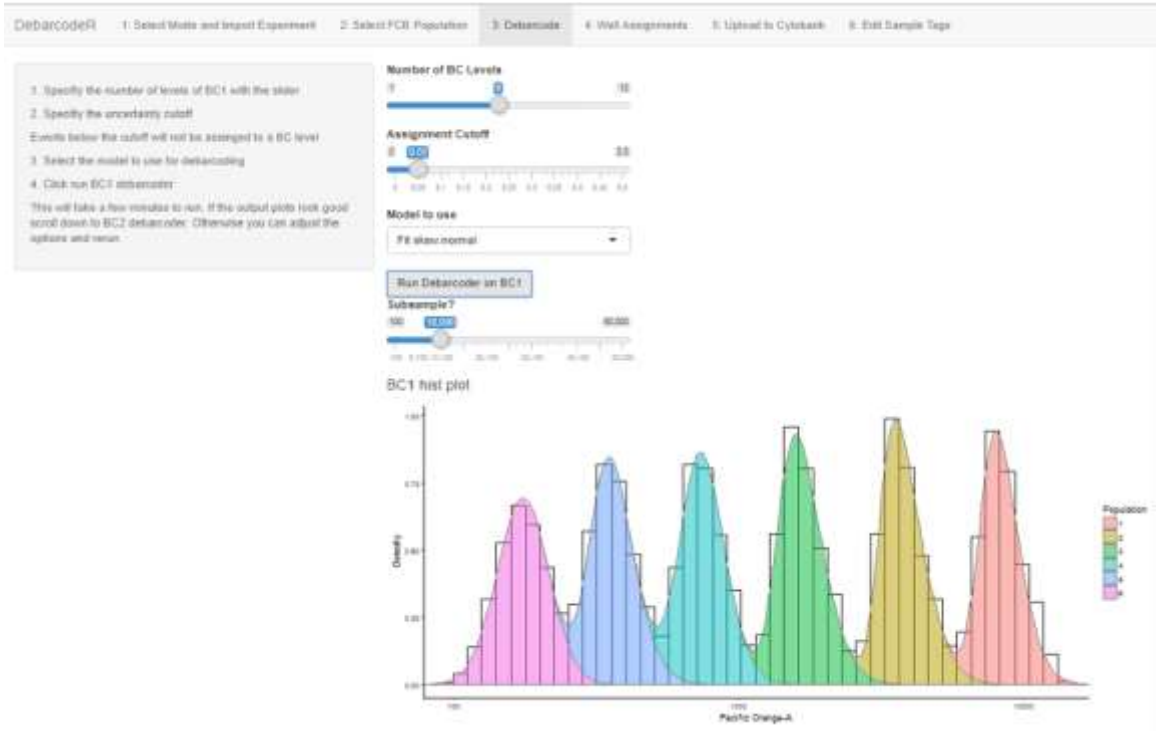


Figure 4.23: 6 levels Pacific Orange fit with mixture modeling.

The user can then specify the number of levels that should be present in BC2 (8 levels of Pacific Blue in the example data) and the debarcoding module runs on BC2. Plots for each BC1 level are returned again with cellular data points colored by the assigned BC2 level (**Figure 4.24**). Points between barcoded sub-populations that did not meet the uncertainty cutoff are classified as level 0.

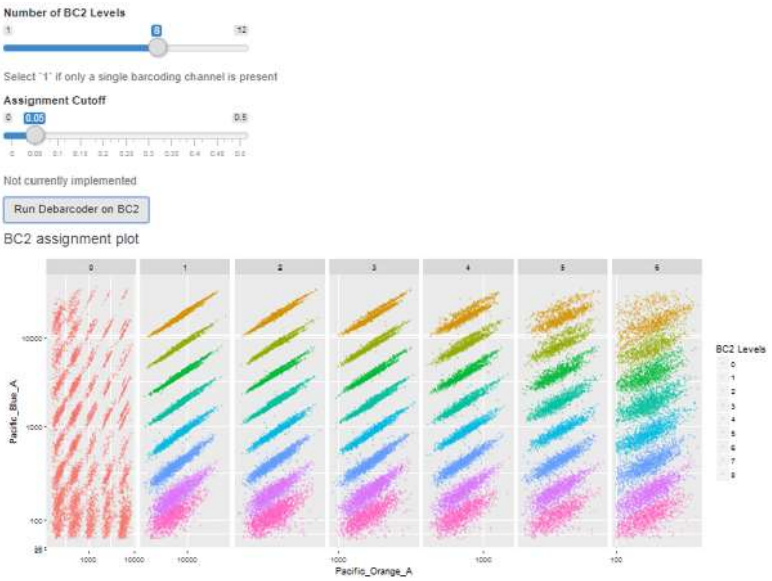


Figure 4.24: Assignment of cells to 48 barcoded populations by automated debarcoding

Experimental Methods

Preparation of microbial crude extracts

Streptomyces strains were maintained on ISP2 agar (yeast extract 4 g/L, malt extract 10 g/L, glucose 4 g/L, agar 20 g/L, pH 7.2). Loops of mycelia were used to inoculate 5 mL seed cultures in ISP2 medium (yeast extract 4 g/L, malt extract 10 g/L, glucose 4 g/L, pH 7.2) for *Streptomyces* strains, incubating them for 3 days at 30 °C. Seed cultures were then transferred to 250 mL Erlenmeyer flasks containing 25 mL of BA medium (soybean powder 15 g/L, glucose 10 g/L, soluable starch 10 g/L, NaCl 3 g/L, $MgSO_4$ 1 g/L, K_2HPO_4 1 g/L and trace elemental solution 1 mL/L, pH 7.2) and grown for 7 days at 30 °C with shaking. Aqueous fermentation broth was extracted by shaking with Diaion® HP20 synthetic absorbent resin (Alfa Aesar) (125 mL HP20 bead/H_2O slurry per 500 mL aqueous broth) for 2 h. Fermentation broth was then centrifuged (3700 x g, 30 min) and the supernatant decanted. Metabolites were eluted from absorbent resin and cells with methanol (250 mL methanol/ 125 mL HP20 bead/H_2O slurry) by shaking for 1.5 h, followed by centrifugation (3700 x g, 30 min) and decanting of the methanol extract. Further extraction was performed with acetone (250 mL acetone / 125 mL HP20 bead/H_2O slurry) by shaking for 1.5 h, followed by centrifugation (3700 x g, 30 min) and decanting of the acetone extract. *Nocardiopsis* strains were cultured and extracted as previously described[23]. Purified ciromicins A and B were isolated from co-cultures as previously described[24]. Kasumi-1 and KG-1 cell lines were obtained from ATCC and identified using mass cytometry analysis of 35 myeloid proteins as reported previously[34].

Specumycin A1 and B1 isolation

Crude acetone extract was concentrated and fractionated with Sephadex LH-20 resin (GE Healthcare Bio-Sciences) with methanol as the eluent. Fractions were analyzed by analytical HPLC/MS, and fractions containing the compound(s) of interest were pooled and further purified by preparative HPLC (Waters, XBridge C18 Prep, 5 uM) (10 mL/min, 0 min – 1 min: 100% solution A, 5 min: 85% solution A; 15% solution B, 65 min: 15% solution A; 85% solution B, 70 min: 100% solution B) (Solution A = 95:5, H_2O:MeCN, 10 mM NH_4OAc; Solution B: 5:95 H_2O:MeCN, 10 mM NH_4OAc). In order to obtain analytical purity, fractions containing the compound of interest (34 - 35 min) were pooled and purified by flash column chromatography (98:2 CH_2Cl_2:MeOH to 95:5 CH_2Cl_2:MeOH). The structure of specumycins A1 and B1 were elucidated using a combination of mass spectrometry and two-dimensional nuclear magnetic resonsance spectroscopy data. Mass spectrometry data produced with electrospray ionization and collected in both positive and negative modes provided the molecular weight of specumycins A1 and B1. Correlated nuclear magnetic spectroscopy (COSY) allowed for the assignment of the spin systems present in the aglycone, amino sugar and acetal moieties of specumycins A1 and B1. Multiplicity-edited heteronuclear single quantum coherence spectroscopy (HSQC) allowed for assigned 1H shifts to be correlated to their corresponding 13C shifts, as well as for the assignment of shifts as corresponding to methylenes or methines. Full structure elucidation was completed with heteronuclear multiple bond correlation spectroscopy (HMBC), which allowed for the assignment of remaining shifts based upon their proximity to assigned shifts.

Generation of metabolomic arrays

Mass spectrometry was performed by using a TSQ Triple quadrapole mass spectrometer equipped with an electrospray ionization source and Surveyor PDA Plus detector. For positive ion mode, the following settings were used :capillary temperature was 270 ºC; spray voltage 4.2 kV; spray current 30 mA; capillary voltage 35 V; tube lens 119 V; skimmer offset 15 V. For negative ion mode, capillary temperature 270 ºC; spray voltage 30 kV; spray current 20 mA; capillary voltage 35 V; tube lens 119 V; skimmer offset 15 V. Fraction plates were prepared by injecting 20 µL of purified compounds in methanol or concentrated extract via a Thermo PAL auto injector onto a phenomenex luna 5 µm C18(2) reverse phase HPLC column. The sample was fractionated using a gradient of 100% Buffer A (95% H_2O, 5% acetonitrile) to 100% Buffer B (5% acetonitrile, 95% H_2O) over 48 min at a flow rate of 1 mL/min a fixed splitter with a 3:1 ratio with 3 parts going to the photodiode array detector and fraction collector and 1 part going to the MS. Fractions were collected in 1 min intervals in a 96 deep well plate. 150 µL of eluent from each well was transferred to 4 replica plates and dried in vacuuo using a Genevac HT-2 system at 30 °C.

Fluorescent cell barcoding of cell seeded metabolomic arrays

Eight serial 1:2.14 dilutions of Pacific Blue were prepared, covering a concentration range from 0.038-7.67 µg/mL. Six serial 1:2.5 dilutions of Pacific Orange were prepared, covering a concentration range from 0.22 -21 µg/mL. Each dilution of Pacific blue was added to all wells in a single row of a 96-well plate (10 µL/well), so that the dye concentration in each row decreased from the top to the bottom of the plate. Similarly, each dilution of Pacific Orange was added to all wells in a column of the same 96-well plate (10 µL/well), so that the concentration in each column decreased from columns 1-6 and from

columns 7-12. This procedure yielded two sets of 48 barcoded wells per plate. Approximately 200,000 cells (180 μL suspended in phosphate-buffered saline (PBS) were added to each well and incubated in the dark at room temperature for 30 min. Staining was then quenched by addition of 75 μL of 1% BSA (Sigma) in PBS.

Antibody staining

Cells were stained with antibodies in 100 μL staining medium for 30 min in the dark, unless otherwise noted. Individual antibodies were added in accordance with manufacturer's instructions. Staining was quenched with 1% BSA in PBS, and stained cells were washed with PBS prior to analysis.

Validation checkerboard

Kasumi-1 cells were incubated with either 20 μM etoposide or 1 μM staurosporine and/or DMSO in a checkerboard pattern overnight. After treatment, cells were stained with Alexa 700, fixed, permeabilized, barcoded, pooled, and then stained with anti-γH2AX-PerCP-Cy5.5 (clone N1-431, BD) or anti-cleaved caspase-3-PE (clone C92-605, BD). Subsequent to staining, samples were run on a 5 laser BD Fortessa flow cytometer. Upon gating single cell events, wells were debarcoded, and the percent of positive cells for each respective marker was determined for each of the 48 populations. Z scores were calculated according to the formula $Z = 1 - 3(\sigma_p + \sigma_n)/|\mu_p - \mu_n|$.

Dose response curves

Serial dilutions of etoposide and staurosporine were prepared from DMSO stocks (10 mM) covering a range from 100 μM to 100 nM. 1 μL of each dilution point was added to a well. Each compound titration was handled individually on a separate plate. KG1 Cells [150,000

in 199 µL of culture medium (RPMI1640 + 20% FBS + 1% penicillin/streptomycin)] were added to each well and mixed by pipetting. After incubation for 16 h, cells were stained with Alexa 700, fixed with 1.6% paraformaldehyde, and permeabilized in methanol for 20 min at -20 °C. Wells were then barcoded as described above, combined, and then stained with antibodies specific for anti-cleaved caspase-3-PE (clone C92-605, BD), or anti-γH2AX-PerCP-Cy5.5 (clone N1-431, BD).

MAM protocol with 6 pure compounds

DMSO stocks (10 mM) of etoposide, staurosporine, CL994, PF04708671, PCI34051, mevastatin, and NU7441 were added (1 µL each) to 43 µL of methanol and fractionated as described above. Before addition of cells, compounds in the wells were suspended by addition of 1 µL of DMSO and mixing by Vortex. KG1 Cells [150,000 in 199 µL of culture medium (RPMI1640 + 20% fetal bovine serum (FBS) + 1% penicillin/streptomycin)] were added to each well and mixed by pipetting. After incubation for 16 h, cells were stained with Alexa 700, fixed with 1.6% paraformaldehyde, and permeabilized in methanol for 20 min at -20 °C. Wells were then barcoded as described above, combined, and then stained with the following antibodies: anti-cleaved caspase-3-PE (clone C92-605, BD), anti-p-Histone H3-PE-Cy7 (clone HTA28, BioLegend), anti-γH2AX-PerCP-Cy5.5 (clone N1-431, BD), and anti-p-S6-Ax647 (clone D57.22E, CST).

MAM using crude extract with internal standards

DMSO stocks (10 mM) of etoposide and staurosporine were added (1 µL each) to 48 µL of a crude extract (200 mg/mL in 50% methanol/water) and fractionated as described above. Extract was generated from a *Streptomyces* cave strain grown in BA medium, and

extracted with 50% methanol prior to evaporation in vacuo. Before treatment, compounds in the wells were suspended by addition of 1 µL of DMSO and mixing by vortexing. KG1 Cells [150,000 in 199 µL of culture medium (RPMI1640 + 20% fetal bovine serum (FBS) + 1% penicillin/streptomycin)] were added to each well and mixed. After incubation for 16 h, cells were stained with Ax700, fixed with 1.6% paraformaldehyde, and permeabilized in methanol for 20 min at -20 °C. Wells were then barcoded as described above, combined, and then stained with anti-cleaved caspase-3-PE (clone C92-605, BD) and anti-γH2AX-PerCP-Cy5.5 (clone N1-431, BD).

AML patient samples

All specimens were obtained in accordance with the Declaration of Helsinki following protocols approved by the Vanderbilt University Medical Center Institutional Review Board. Details of patients and sample acquisition were previously published[15]. Briefly, consent was obtained via an approved written consent form, and eligibility criteria included >=18 years of age with suspected acute myeloid leukemia undergoing clinical evaluation at Vanderbilt. Samples analyzed here were collected from bone marrow prior to any treatment. Once obtained, samples underwent immediate (within <30 min) density gradient separation of mononuclear cells using a BD Vacutainer® CPT™ Cell Preparation Tube with Sodium Heparin (BD Biosciences, Franklin Lakes, NJ). The separated mononuclear cells were then pelleted with low speed centrifugation (200 x *g*) and aliquoted into multiple cryotubes in an 88% FBS + 12% DMSO solution. Samples were stored at -80 ^{o}C for 24-72 h prior to long-term storage in liquid nitrogen. Patient 1305001 (001) was found to have the MLL-MLLT3 t(9;11)(p22;q23) translocation in all cells by karyotype analysis and was without other tested common molecular mutations (which included FLT3 internal tandem

duplication (ITD), NPM1, CEPBA, & c-KIT)[25]. Patient 1305015 (015) had a normal karyotype, but was found to have both a FLT3-ITD and an NPM1 mutation.

Fluorescent flow cytometry barcoding and bioactivity analysis (macrolides)

To make barcoding plates seven serial 1:2.14 dilutions of Pacific Blue were prepared, covering a concentration range from 0.083-7.67 μg/mL. Five serial 1:2.5 dilutions of Pacific Orange were prepared, covering a concentration range from 0.53-21 μg/mL. Each dilution of Pacific blue was added to all wells in a single row of a 96-well plate (10 μL/well), so that the dye concentration in each row decreased from the top to the bottom of the plate. Similarly, each dilution of Pacific Orange was added to all wells in a column of the same 96-well plate (10 μL/well), so that the concentration in each column decreased from columns 1-5. Concentration series were prepared from DMSO stocks at 200X the final concentrations (100 nM to 100 uM) and each treatment condition was added to wells in 1 μL aliquots. Jurkat cells [200,000 in 199 μL of culture medium (RPMI1640 + 10% fetal bovine serum (FBS) + 1% penicillin/streptomycin)] were added to each well and mixed by pipetting. After incubation for 16 h, cells were stained with Alexa 700, fixed with 1.6% paraformaldehyde, and permeabilized in methanol for 20 min at -20 °C. Cells were then centrifuged and washed once with PBS, resuspended in 180 μL PBS, transferred to the barcoding plate well and incubated in the dark at room temperature for 30 min. Staining was then quenched by addition of 75 μL of 1% BSA (Sigma) in PBS. Wells were then combined and stained with the following antibodies: anti-cleaved caspase-3-PE (clone C92-605, BD), anti-p-Histone H3-PE-Cy7 (clone HTA28, BioLegend), anti-γH2AX-PerCP-Cy5.5 (clone N1-431, BD), and anti-p-S6-Ax647 (clone D57.22E, CST). Cells were stained with antibodies in 100 μL staining medium for 30 min in the dark. Individual

antibodies were added in accordance with manufacturer's instructions. Staining was quenched with 1% BSA in PBS, and stained cells were washed with PBS prior to analysis.

MAM of Streptomyces Specus and Nocardiopsis sp. FU40 against primary cell preparations

Primary cell preparations were thawed, and 200,000 cells were added to each well of a fraction plate containing a metabolite array that was generated from a crude extract from *S. specus* or *Nocardiopsis* sp. FU40. After a 16 h incubation, cells were stained for viability, fixed, permeabilized, barcoded, and stained with the following antibodies: anti-Human CD45-Ax488 (clone HI30, BioLegend), anti-cleaved caspase-3-PE (clone C92-605, BD), anti-p-Histone H3-PE-Cy7 (clone HTA28, BioLegend), anti-γH2AX-PerCP-Cy5.5 (clone N1-431, BD), and anti-p-S6-Ax647 (clone D57.22E, CST). SSC and CD45 expression were used to define lymphocyte, monocyte, and blast populations. Each population was then debarcoded, and readouts were determined for the 48 wells per cell type.

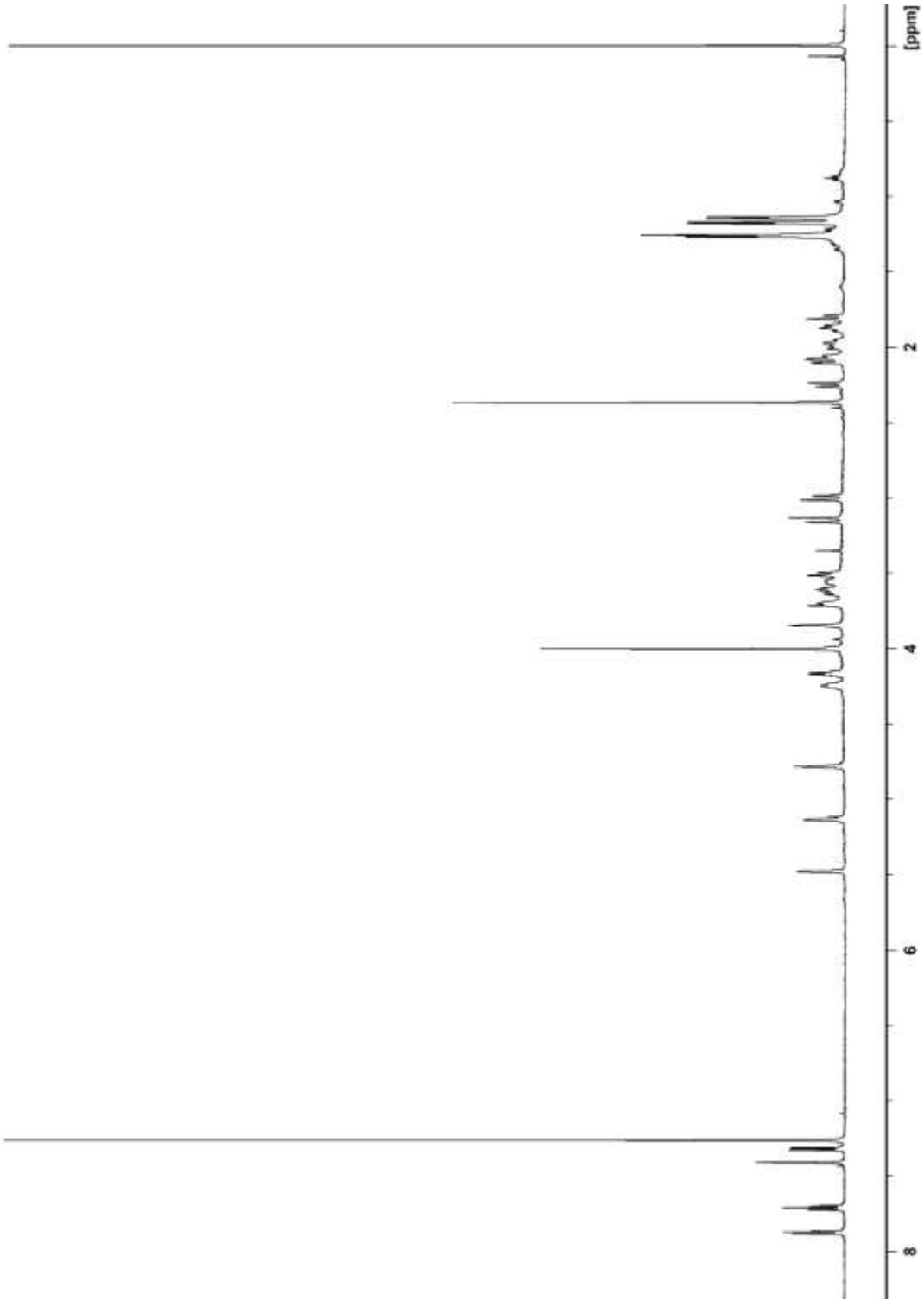
Deep profiling of ciromicins A/B against primary cell preparations

Ciromicins were purified by separation on a Water 600 HPLC system with a reverse phase column using a linear gradient of water/acetonitrile containing 0.1% formic acid. Fractions with UV absorbance indicative of ciromicins were then combined and applied on a size exclusion Sephadex LH-20 column for a gravity elution in methanol. Ciromicin A and B were then separated by a secondary HPLC purification. Approximately 6 million cells (2 million per condition) from a thawed primary AML sample were incubated in culture medium [80% RPMI 1640 (Mediatech, Inc., Manassas, VA) + 20% FBS (Gibco® standard FBS, life technologies, Grand Island, NY) with 10 μM ciromicin A, ciromicin B, or

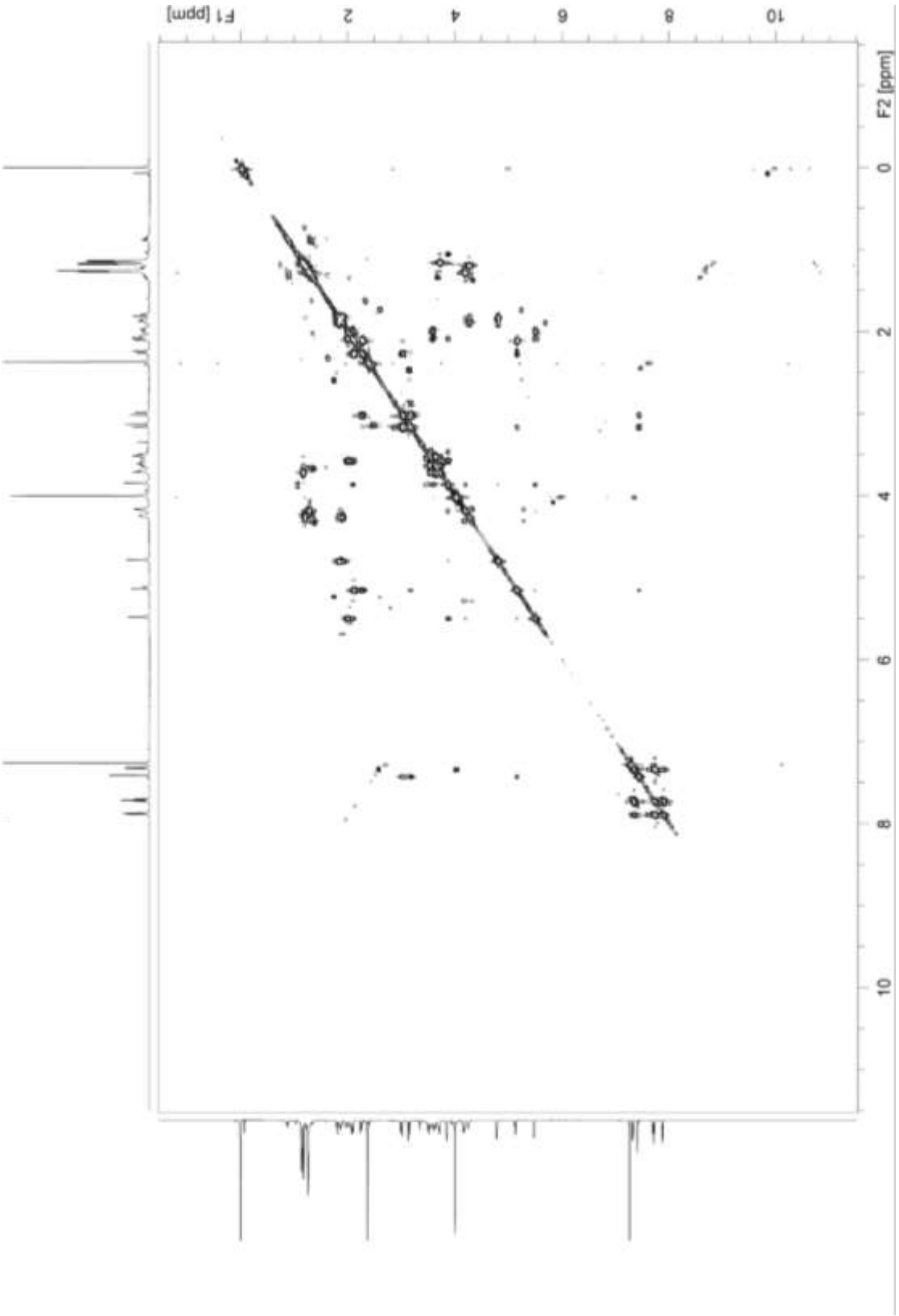
DMSO] for 48 h. Mass cytometry experiments were performed as previously described[15]. Briefly, after incubation with Ciromicin A, B, or vehicle, samples were pelleted by centrifugation at 200 x *g*, resuspended, and washed with PBS (HyClone®, HyClone Laboratories, Logan, UT), pelleted, and resuspended in PBS. They were then stained with Cell-ID™ Cisplatin (Fluidigm, South San Francisco, CA) per the manufacturer's recommended protocol. The cells were washed and resuspended in staining medium [CSM: PBS + 1% BSA (Fisher Scientific, Fair Lawn, NJ)]. Cells were then stained with a mass cytometry antibody panel of 29 extracellular antibodies designed to characterize AML blasts and most non-AML peripheral blood mononuclear cells consisting of anti-human CD235a-141 (clone HIR2, Fluidigm), anti-human CD117-143 (clone 104D2, Fluidigm), anti-human CD11b-144 (clone ICRF44, Fluidigm), anti-human CD4-145 (clone RPAT4, Fluidigm), anti-human CD64-146 (clone 10.1, Fluidigm), anti-human CD36-147 (clone 5-271, BioLegend), anti-human CD34-148 (clone 581, Fluidigm), anti-human CCR2-149 (clone K036C2, BioLegend), anti-human CD43-150 (clone 84-3C1, Fluidigm), anti-human CD123-151 (clone 6H6, Fluidigm), anti-human CD13-152 (clone WM15, Fluidigm), anti-human CD45RA-153 (clone HI100, Fluidigm), anti-human CD45-154 (clone HI30, Fluidigm), anti-human CD86-156 (clone IT2.2, Fluidigm), anti-human CD33-158 (clone WM53, Fluidigm), anti-human CD11c-159 (clone BU15, Fluidigm), anti-human CD14-160 (clone M5E2, Fluidigm), anti-human CD32-161 (clone FUN-2, BioLegend), anti-human CDHLA-DR-163 (clone L243, BioLegend), anti-human CD15-164 (clone W6D3, Fluidigm), anti-human CD16-165 (clone 3G8, Fluidigm), anti-human CD38-167 (clone HIT2, Fluidigm), anti-human CD8-168 (clone SK1, Fluidigm), anti-human CD25-169 (clone 2A3, Fluidigm), anti-human CD3-170 (clone UCHT1, Fluidigm), anti-human

CD184-173 (clone 12G5, Fluidigm), anti-human PD1-174 (clone EH12.2H7, Fluidigm), anti-human PD-L1-175 (clone 29E.2A3, Fluidigm), and anti-human CD56-176 (clone CMSSB, Fluidigm). A master mix of these antibodies was added to each sample to give a final staining volume of 50 µL and incubated at room temperature for 30 min. Cells were then washed twice, first with CSM and then with PBS and then permeabilized in -20 °C 100% methanol for 20 min. Following permeabilization, cells were washed, stained with 250 nM iridium intercalator (Fluidigm, San Francisco, CA) for 30 min at 4 °C, washed twice in PBS, and then re-suspended in 500 µL ddH_2O for CyTOF analysis. Samples were analyzed using a CyTOF 1.0 cytometer (Fluidigm, San Francisco, CA)

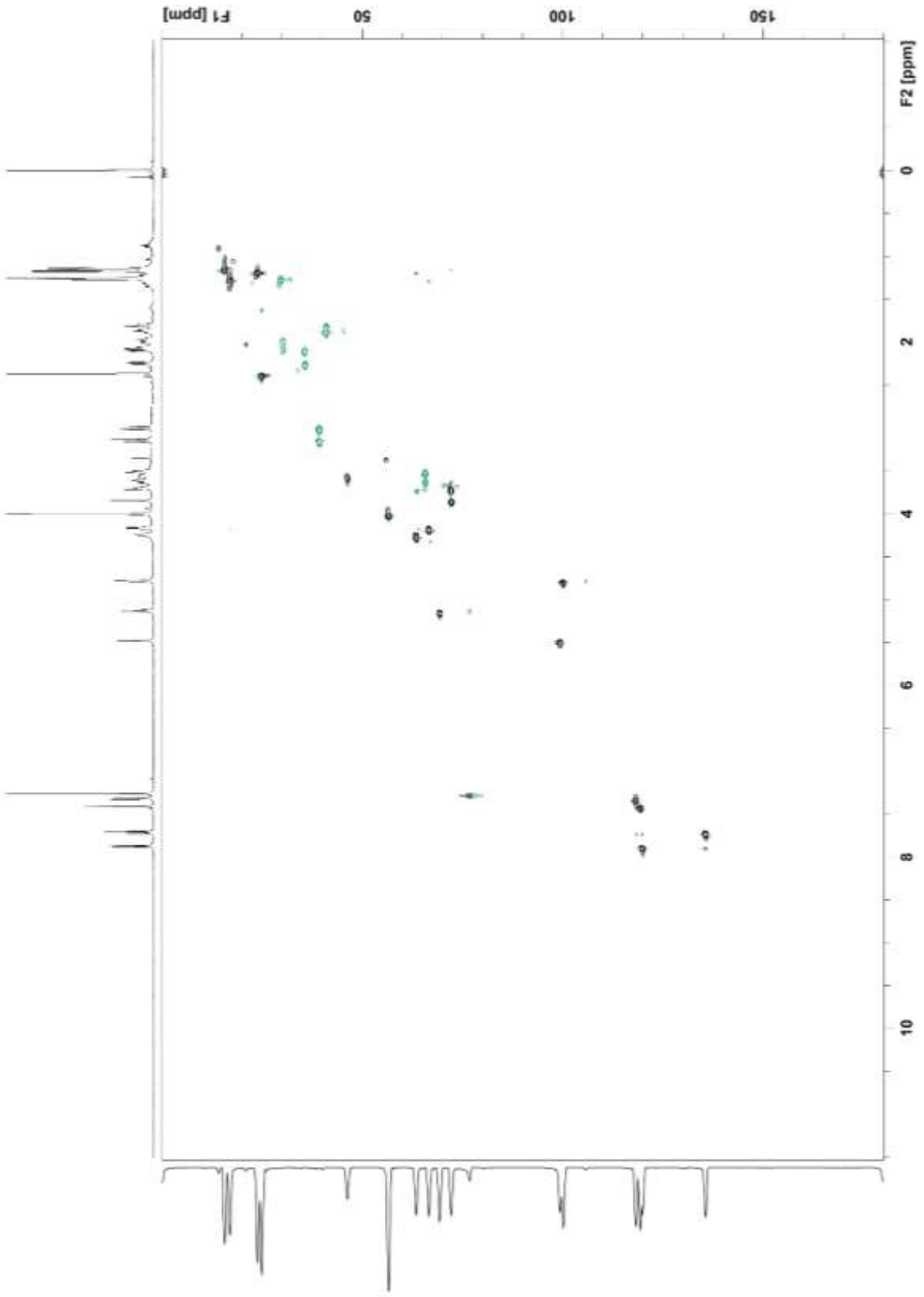
Spectra Relevant to Chapter



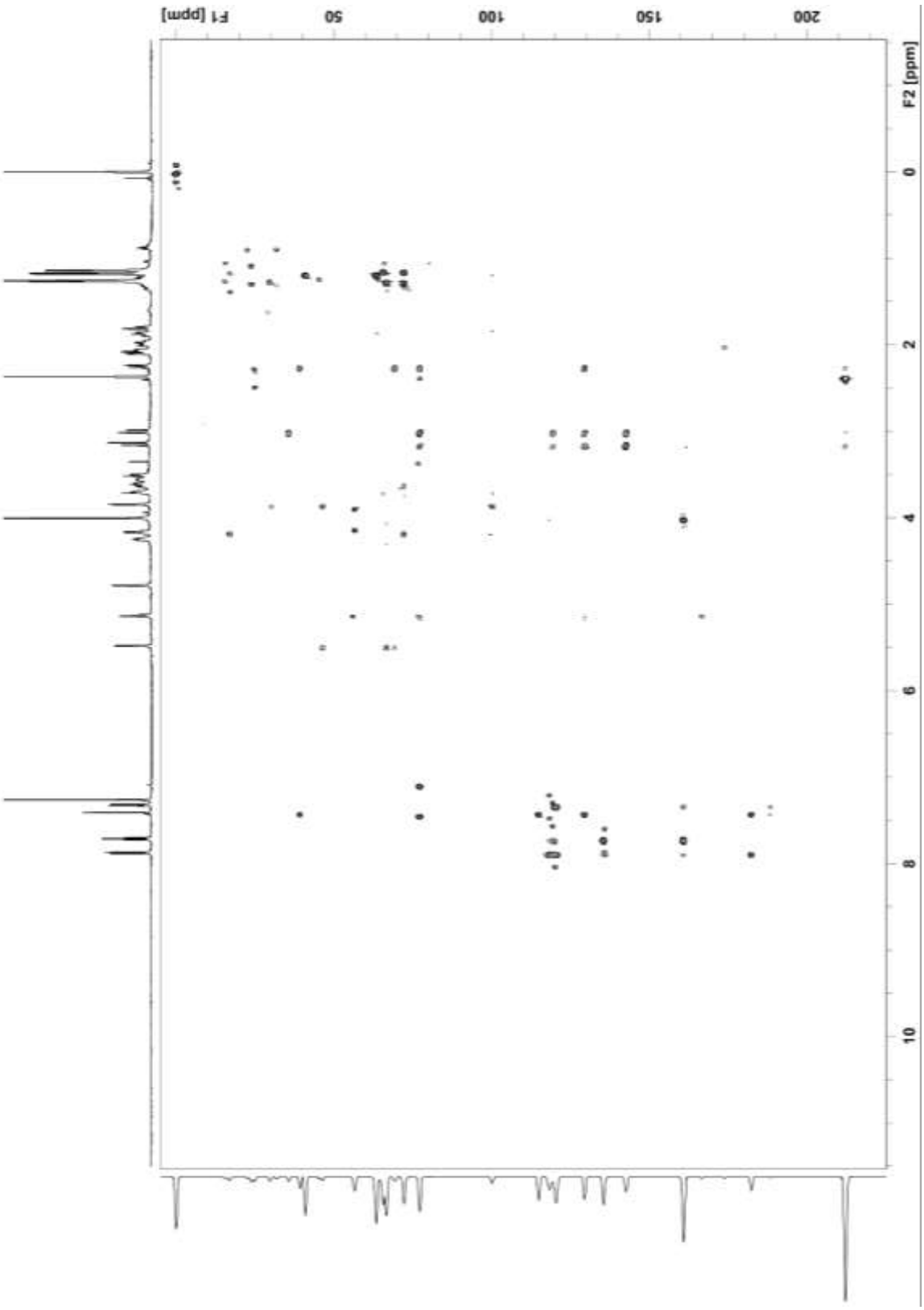
Spectrum 4.1: ^{1}H NMR, 600 MHz, of specumycin B1 in $CDCl_3$



Spectrum 4.2: COSY NMR, 600 MHz, of specumycin B1 in $CDCl_3$



Spectrum 4.3: HSQC NMR, 600 MHz, of specumycin B1 in $CDCl_3$



Spectrum 4.4: HMBC NMR, 600 MHz, of specumycin B1 in $CDCl_3$

References

1 Patti, G. J., Yanes, O. & Siuzdak, G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology* **13**, 263-269, (2012).

2 Johnson, T. A. *et al.* Natural product libraries to accelerate the high-throughput discovery of therapeutic leads. *Journal of Natural Products* **74**, 2545-2555, (2011).

3 Krutzik, P. O., Clutter, M. R., Trejo, A. & Nolan, G. P. Fluorescent cell barcoding for multiplex flow cytometry. *Current Protocols in Cytometry* **55**, 6.31.31-36.31.15, (2011).

4 Chen, G. L. *et al.* Nonintercalative antitumor drugs interfere with the breakage-reunion reaction of mammalian DNA topoisomerase II. *Journal of Biological. Chemistry.* **259**, 13560-13566, (1984).

5 Belmokhtar, C. A., Hillion, J. & Segal-Bendirdjian, E. Staurosporine induces apoptosis through both caspase-dependent and caspase-independent mechanisms. *Oncogene* **20**, 3354-3362, (2001).

6 Perfetto, S. P. *et al.* Amine reactive dyes: an effective tool to discriminate live and dead cells in polychromatic flow cytometry. *Journal of Immunological Methods* **313**, 199-208, (2006).

7 Vermeulen, K., Van Bockstaele, D. R. & Berneman, Z. N. Apoptosis: mechanisms and relevance in cancer. *Annals of Hematolology.* **84**, 627-639, (2005).

8 Kuo, L. J. & Yang, L. X. Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo* **22**, 305-309, (2008).

9 Tu, W. Z. *et al.* gammaH2AX foci formation in the absence of DNA damage: mitotic H2AX phosphorylation is mediated by the DNA-PKcs/CHK2 pathway. *FEBS Letters.* **587**, 3437-3443, (2013).

10 R: A Language and Environment for Statistical Computing (2016).

11 Borowitz, M. J., Guenther, K. L., Shults, K. E. & Stelzer, G. T. Immunophenotyping of acute-leukemia by flow cytometric analysis - use of CD45 and right-angle light scatter to gate on leukemic blasts in 3-color analysis. *American Journal of Clinical Pathololgy.* **100**, 534-540, (1993).

12 Rothe, G. *et al.* Consensus protocol for the flow cytometric immunophenotyping of hematopoietic malignancies. *Leukemia* **10**, 877-895, (1996).

13 Lewis, C. A., Longcore, K. E., Miller, S. J. & Wender, P. A. An approach to the site-selective diversification of apoptolidin A with peptide-based catalysts. *Journal of Natural Products* **72**, 1864-1869, (2009).

14 Estey, E. H. Acute myeloid leukemia: 2012 update on diagnosis, risk stratification, and management. *American Journal of Hematology.* **87**, 89-99, (2012).

15 Ferrell, P. B., Jr. *et al.* High-dimensional analysis of acute myeloid seukemia reveals phenotypic changes in persistent cells during induction therapy. *PLoS One* **11**, e0153207, (2016).

16 Dufner, A. & Thomas, G. Ribosomal S6 kinase signaling and the control of translation. *Experiments Cell Research* **253**, 100-109, (1999).

17 Magnuson, B., Ekim, B. & Fingar, D. C. Regulation and function of ribosomal protein S6 kinase (S6K) within mTOR signalling networks. *Biochemical Journal* **441**, 1-21, (2012).

18 Tang, H. *et al.* Amino acid-induced translation of TOP mRNAs is fully dependent on phosphatidylinositol 3-kinase-mediated signaling, is partially inhibited by rapamycin, and is independent of S6K1 and rpS6 phosphorylation. *Molecular Cellular Biology* **21**, 8671-8683, (2001).

19 Hans, F. & Dimitrov, S. Histone H3 phosphorylation and cell division. *Oncogene* **20**, 3021-3027, (2001).

20 Sugiyama, K. *et al.* Aurora-B associated protein phosphatases as negative regulators of kinase activation. *Oncogene* **21**, 3103-3111, (2002).

21 Huntington, N. D. & Tarlinton, D. M. CD45: direct and indirect government of immune regulation. *Immunology Letters* **94**, 167-174, (2004).

22 Takahashi, Y. *et al.* The structure of baumycins A1, A2, B1, B2, C1 and C2. *Journal of Antibiotics* **30**, 622-624, (1977).

23 Du, Y. *et al.* Biosynthesis of the apoptolidins in Nocardiopsis sp. FU 40. *Tetrahedron* **67**, 6568-6575, (2011).

24 Derewacz, D. K., Covington, B. C., McLean, J. A. & Bachmann, B. O. Mapping microbial response metabolomes for induced natural product discovery. *ACS Chemical Biology* **10**, 1998-2006, (2015).

25 Marcucci, G., Haferlach, T. & Dohner, H. Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. *Journal of Clinical Oncology* **29**, 475-486, (2011).

26 Earl, D. C., Ferrell, P. B., Leelatian, N., Froese, J. T., Reisman, B. J., Irish, J. M., Bachmann, B. O. Discovery of human cell selective effector molecules using single cell multiplexed activity metabolomics. *Nature Communications*. **9**, 39, (2018).