

CHAPTER I

INTRODUCTION

Natural products and mass spectrometry

Natural products (also known as secondary metabolites) have always been a significant source of new lead compounds in pharmaceutical industries. About half of the drugs currently in clinical use are natural products or synthetic molecules based on natural product scaffolds.¹ Between 1981 and 2006, 34% of all small molecule new chemical entities were natural products (NP) or natural product derivatives (NPD) and 17% were synthetic molecules whose pharmacophore was from a natural product.¹ In 1999, nine out of the 20 best-selling non-protein drugs were either NPD or developed from natural product leads with combined annual sales greater than \$ 16 billion.² NP/NPD have various medicinal uses (Figure 1-01) such as antibiotics (e.g., amoxicillin and tigecycline), immunosuppressive agents (e.g., cyclosporine and rapamycin), hypocholesterolemic agents (e.g., lovastatin and pravastatin), antimigraine agents (e.g., sumatriptan and eliptriptan), antihypertensive agents (e.g., captopril and enalapril) and anticancer agents (e.g., taxol and calicheamicin).^{1, 2}

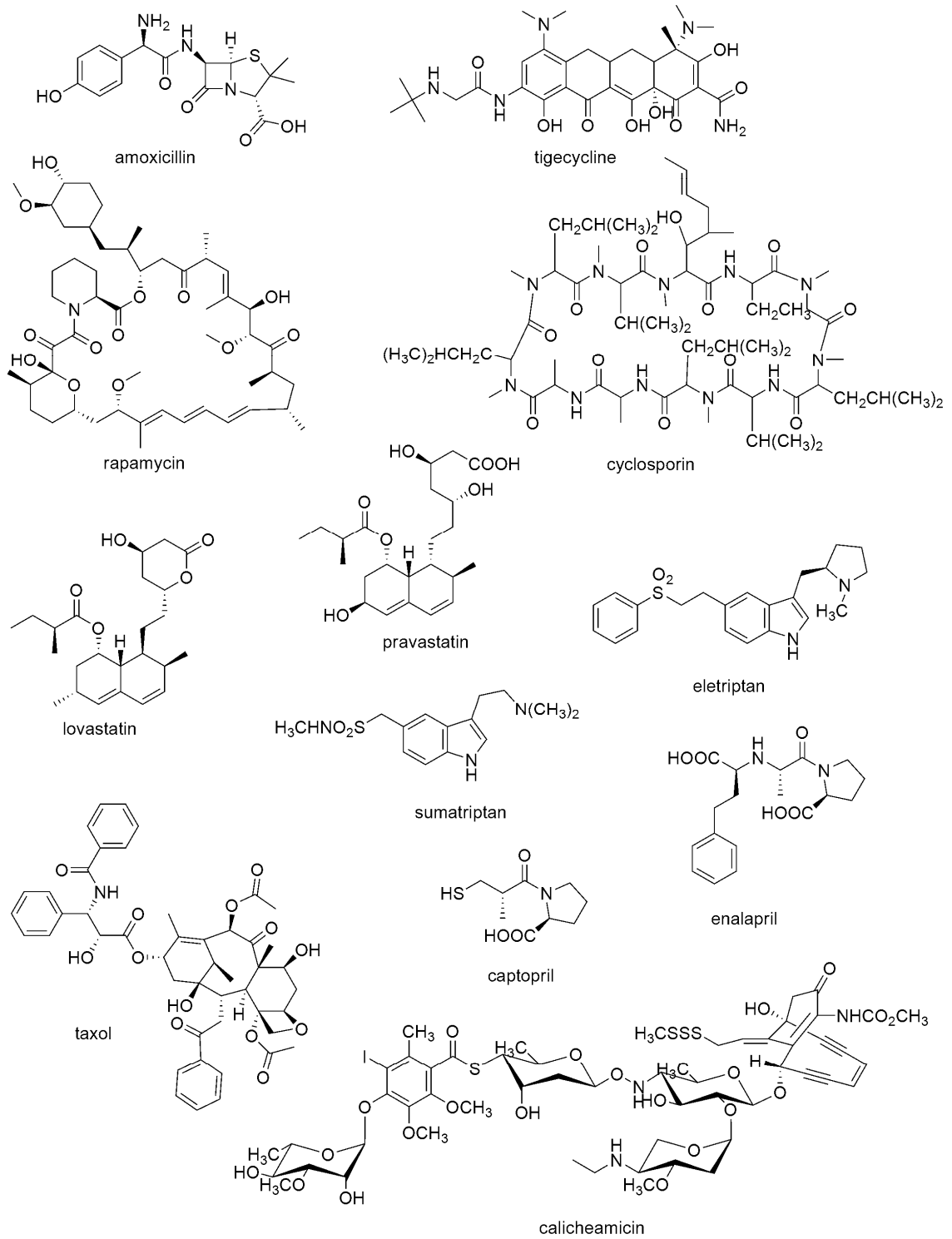


Figure 1-01. Structures of drugs currently in clinical use derived from NP.

Several spectroscopic techniques are used in the identification and characterization of new NP, such as mass spectrometry (MS), nuclear magnetic resonance (NMR), ultraviolet spectroscopy (UV), infrared spectroscopy (IR), and X-ray crystallography. Recent advancements in MS have enabled MS to become an invaluable tool in the study of natural products.

MS has long been used for identifying and quantifying compounds. MS is based on the production of gaseous, positively or negatively charged ions that are subsequently separated according to their mass-to-charge (m/z) ratio and detected.³ A typical mass spectrometer consists of an ion source, which generates ions, a mass analyzer, which separates ions based on their m/z , and a detector, which measures the separated ions^{3, 4} (Figure1-02a).

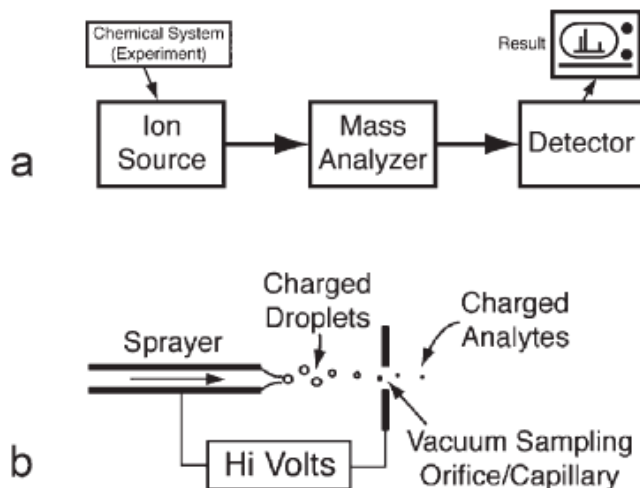


Figure 1-02. Schematic diagrams of mass spectrometry. (a) A generic MS system consisting of an ion source, a mass analyzer and a detector. (b) A schematic diagram of electrospray ionization (ESI).⁴

In traditional MS, purified samples were subjected to high-energy ionization under ultra-high vacuum conditions (“hard” ionization; e.g., electron impact [EI]).³ The development of electrospray ionization (ESI) MS has marked a milestone in the analysis of natural products. In contrast to earlier ionization techniques such as EI, which were applicable only to thermally stable, low molecular weight volatile compounds, virtually any ion (ranging from inorganic salts to large macromolecules such as proteins) can be analyzed by ESI-MS.³ Another advantage of ESI-MS over other ionization techniques is that it can be directly coupled to high performance liquid chromatography (HPLC).³ The sample is sprayed into the ion source as a solution. In the source, the solvent is evaporated under atmospheric pressure in the presence of an electric field, which generates charged ions to be separated by the mass analyzer^{3, 4} (Figure 1-02b). In this way, the mass analyzer has become a unique kind of chromatographic detector, able to give information about the molecular mass, the chemical formula and structure of several components of complex biological samples. As a result, the interface of HPLC with ESI-MS has provided an excellent method in the identification and isolation of new secondary metabolites from complex extracts.

Typically the first and most important step in the identification of a compound is the determination of its molecular mass. ESI is quite efficient at providing this information. In contrast to “hard” ionization techniques, ESI rarely generates fragments. Molecules are ionized by protonation, cationization, or deprotonation to form pseudomolecular ions.³ During protonation and

cationization, positively charged ions are formed such as $[M+H]^+$ and $[M+Na]^+$, respectively. During deprotonation, a proton is removed from the molecule resulting in negatively charged ions $[M-H]^-$. Performing both positive ion and negative ion ESI-MS experiments can give reliable evidence for the molecular weight of a compound. In addition, depending on the mass analyzer used, high accuracy in the determination of the molecular weight can be achieved. From this accurately determined molecular mass (± 0.001 mass units), the chemical formula and the number of double bonds, rings or heteroatoms can be inferred. Figure 1-03 illustrates the high resolution and accurate mass measurement of C36-keto-meridamycin in both positive and negative ion mode.⁵ The masses and formulae from both ionization modes are in agreement with the expected mass and formula for C36-keto-meridamycin ($C_{45}H_{72}NO_{12}$, theoretical mass 819.51328). C36-keto-meridamycin is an analogue of the natural product meridamycin produced through genetic manipulation of the meridamycin gene cluster.

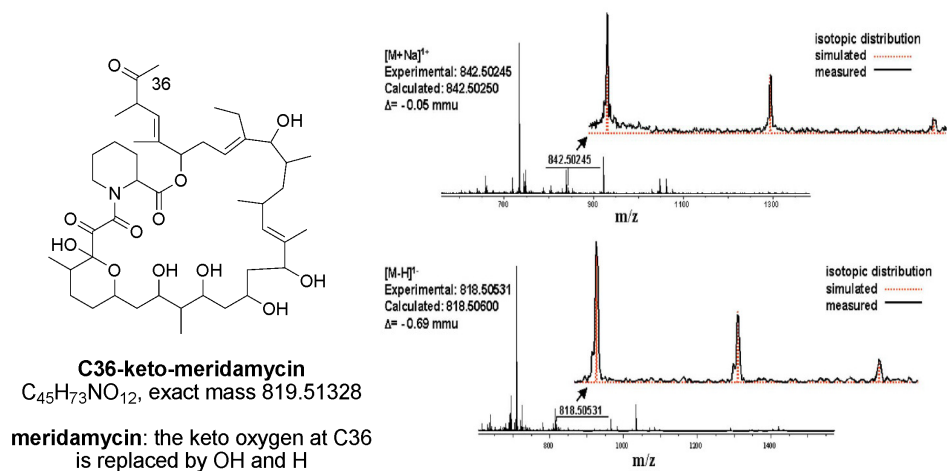


Figure 1-03. Mass spectra of C36-keto-meridamycin in both positive (top) and negative (bottom) ion modes.⁵

Once the molecular weight has been determined, further structural information may be obtained by tandem MS. ESI, being a soft ionization technique, does not produce a significant amount of fragment ions. Fragmentation is achieved by collision-induced dissociation (CID). For example, using a triple quadrupole mass analyzer, a single ion is selected using the first quadrupole, Q1. Subsequently, only the selected ion (precursor ion) enters the second quadrupole (Q2, collision chamber) where it collides with argon atoms and fragments are generated. These fragments (product ions) can then be analyzed by the third quadrupole (Q3).³ Since the fragments generated are characteristic of the chemical structure of the parent ion, structural information about the parent ion can be obtained by analyzing the product ions.

An excellent example of the power of tandem MS in elucidating the structures of complex natural products is everninomicin (EV), a family of oligosaccharide antibiotics produced by *Micromonospora carbonacea*.⁶ These antibiotics are highly active against gram-positive bacteria including methicillin-resistant *staphylococci* and vancomycin-resistant *enterococci*. This family of compounds is characterized by an eight-sugar backbone with two ortho ester functionalities, a nitrosugar, a completely substituted aromatic ester containing two chlorines, and a methylene dioxy group as illustrated by SCH 27899, a member of this family (Figure 1-04).

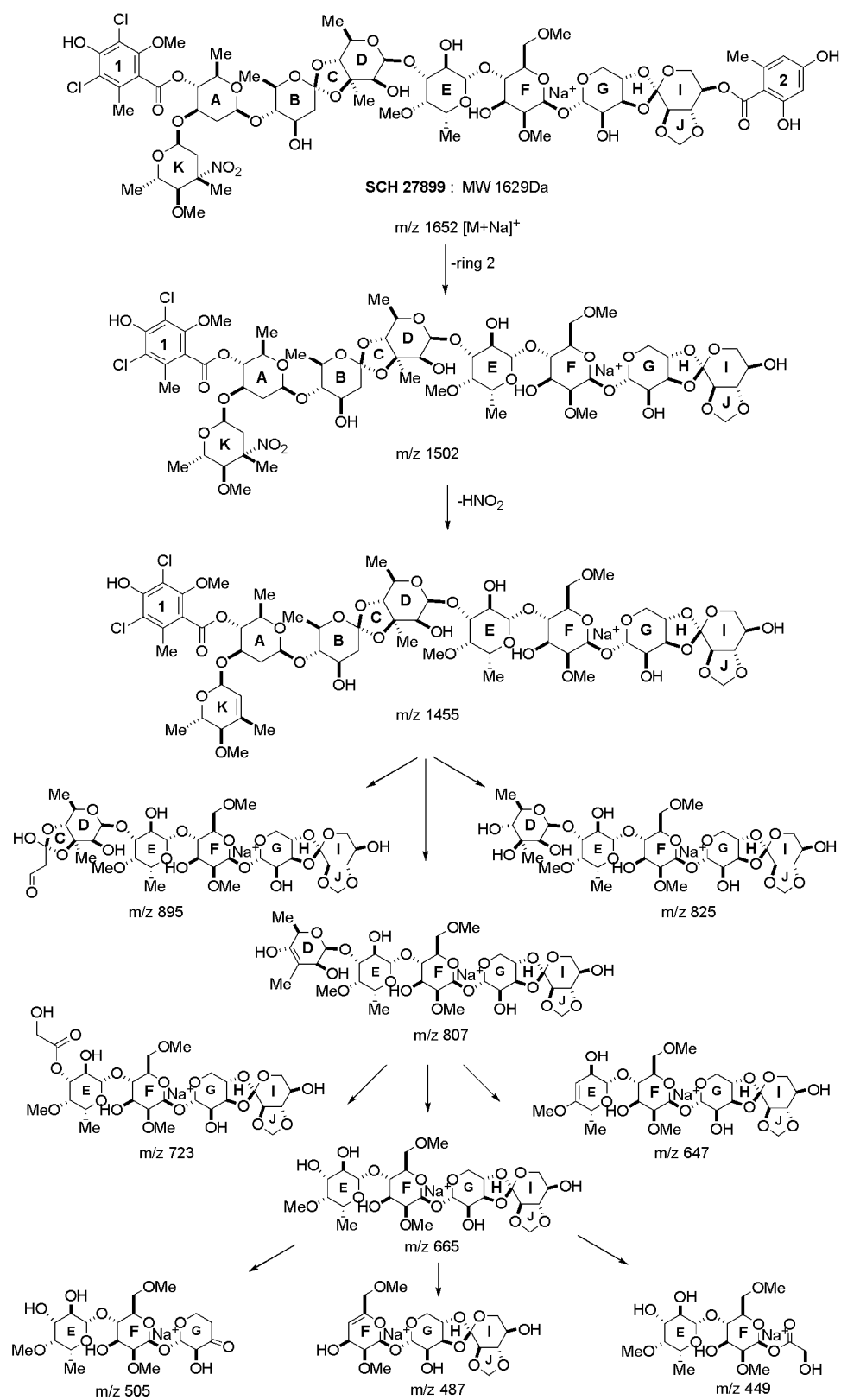


Figure 1-04. Fragmentation pattern of sodiated SCH 27899, a representative of the everninomicin family.⁶

Early structural characterization work on the everninomicin family was based on chemical degradation and fast-atom bombardment (FAB) MS.⁷ However, since then an ESI tandem MS method has been developed for the structural identification of everninomicins.⁶ New members of this class can be discovered based on their isotopic and fragmentation pattern. This is based on the notion that compounds that are structurally similar will have a similar fragmentation pattern. In the case of everninomicins, several characteristic fragment ions are associated with changes on certain functional groups of the molecule (Figure 1-04). For example, a loss of HNO₂ indicates the presence of a nitro group. Also, the loss of ring **2** provides evidence for its presence on sugar ring **I**. In addition, the characteristic isotopic pattern consistent with the presence of two chlorine atoms in the molecule can aid in identifying this molecules. This very specific fragmentation pattern of everninomicins is extremely valuable in the identification and structural elucidation of new members of this family.

MS has found broad application not only in the molecular weight determination and structural elucidation of natural products but also in probing the biosynthetic pathways that lead to their production. Understanding the biosynthetic pathways that form natural products is of immense significance since this information can be used to improve production yields and to synthesize structural analogues via manipulation of the pathway. Feeding studies using isotopic labeled precursors followed by inspection of the labeling pattern in the final product have very frequently been the first step in the investigation of a biosynthetic system. MS has proven to be especially powerful in biosynthetic

studies with stable isotopic labeled compounds. MS can distinguish compounds differing by less than 1 Da in mass with high sensitivity. In addition, the location of the labeled atoms can be deduced by performing tandem MS and analyzing the resulting fragment ions.

Stable isotopic labeling studies of naturally occurring antibiotic dioxapyrrolomycin clearly demonstrate the application of MS and tandem MS in the elucidation of biosynthetic pathways. Dioxapyrrolomycin (**1.1**) was isolated from growing cultures of *Streptomyces fumanus* and is characterized by a nitrated pyrrole moiety and high halogenation of both aromatic rings (Figure 1-05).⁸ Fermentation extracts of this organism also contain several pyrrolomycins (Figure 1-05). To investigate the order of biosynthetic transformations leading to dioxapyrrolomycin and the precursor relationship of the various pyrrolomycins to it, several ²H, ¹⁵N, and ¹³C labeled compounds were introduced to growing cultures of the producing organism.

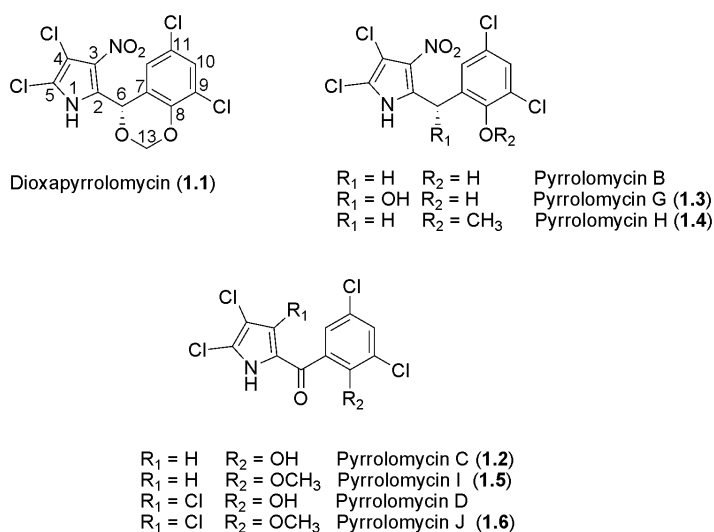


Figure 1-05. Dioxapyrrolomycin and related pyrrolomycins

Cultures of *S. fumanus* were grown with Na¹⁵NO₃ as the sole nitrogen source. Culture extracts were analyzed by MS and were found to contain compounds **1.1-1.4**, in which each nitrogen atom was highly enriched (>80%) in ¹⁵N. Compounds **1.2-1.4** (in their ¹⁵N-labeled form) were reintroduced to growing cultures of *S. fumanus* to assess their intermediacy in dioxapyrrolomycin biosynthesis. When ¹⁵N-labeled pyrrolomycins C was fed, **1.1** isolated showed 8-18% incorporation of ¹⁵N at N-1. Tandem MS experiments unambiguously established the regiospecificity of ¹⁵N labeling, as the NO₂⁻ fragment ion generated was entirely at *m/z* 46, corresponding to the ¹⁴N isotope. The conversion of **1.2** to **1.1** demonstrated that **1.2** is a true intermediate and not just a shunt metabolite of the biosynthesis. In addition, when cultures of *S. fumanus* were supplemented with ¹⁵N₂-labeled **1.3**, ¹⁵N₂-labeled **1.4** was isolated with 86% ¹⁵N enrichment. Cultures of *S. fumanus* grown in the presence of ¹⁵N₂-labeled **1.4** produced ¹⁵N₂-labeled **1.1** with 79% ¹⁵N enrichment, establishing the precursor relationship of **1.4** to **1.1**.

Feeding studies using isotopic labeled methionine were used to study the formation of the methylenedioxy bridge in **1.1** in greater detail. Specifically, in the presence of L-[*methyl*-¹³C]-methionine, **1.1**, **1.4**, **1.5**, and **1.6** isolated from cultures of *S. fumanus* exhibited >85% ¹³C enrichment in C-13. These results clearly indicated that the C-13 methyl group of **1.4**, **1.5**, and **1.6** and the C-13 methylenedioxy group of **1.1** are derived from methionine. Furthermore, feeding of L-[*methyl*-¹³CD₃]-methionine yielded more information into the formation of the methylenedioxy bridge. The dioxapyrrolomycin (**1.1**) isolated had molecular ions

shifted three mass units higher than unlabeled **1.1**. Isolated labeled **1.4** was reintroduced to growing cultures of *S. fumanus* yielding **1.1** enriched in ^{13}C at the C-13 by 57%. These last two results clearly showed that **1.4** is the ultimate precursor of **1.1** and the transformation of the C-13 methyl group of **1.4** to the methylenedioxy bridge of **1.1** occurs without significant hydrogen-deuterium exchange. These feeding studies enabled the establishment of the biosynthetic precursors of dioxapyrrolomycin and the timing of transformations (nitration, *O*-methylation, methylenedioxy bridge formation) leading to **1.1**, as outlined in Figure 1-06.⁸

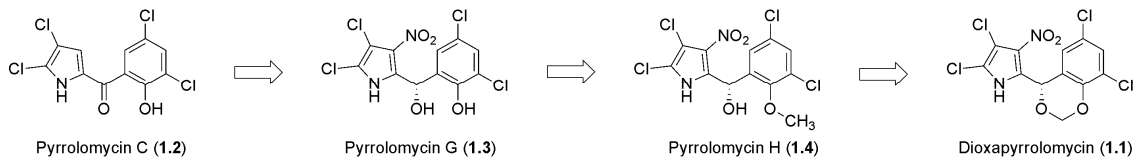


Figure 1-06. Biosynthetic pathway of dioxapyrrolomycin as established by stable isotopic feeding studies.⁸

Once the intermediates of a natural product's biosynthesis have been established, they can be used in conjunction with MS to identify and isolate the enzymes involved in the transformations leading to the compound of interest. MS has long been used in the field of enzymology. Unlike other analytical techniques that require the presence of a chromophore or a radioactive label to study an enzymatic reaction, the only requirement for an MS-based method is that the reaction produces a mass change. Most enzymatic reactions, except for racemates and isomerization reactions, do result in a change in molecular mass.

Thus, MS can unambiguously identify substrates, intermediates, and products by their molecular mass and unique fragmentation pattern.

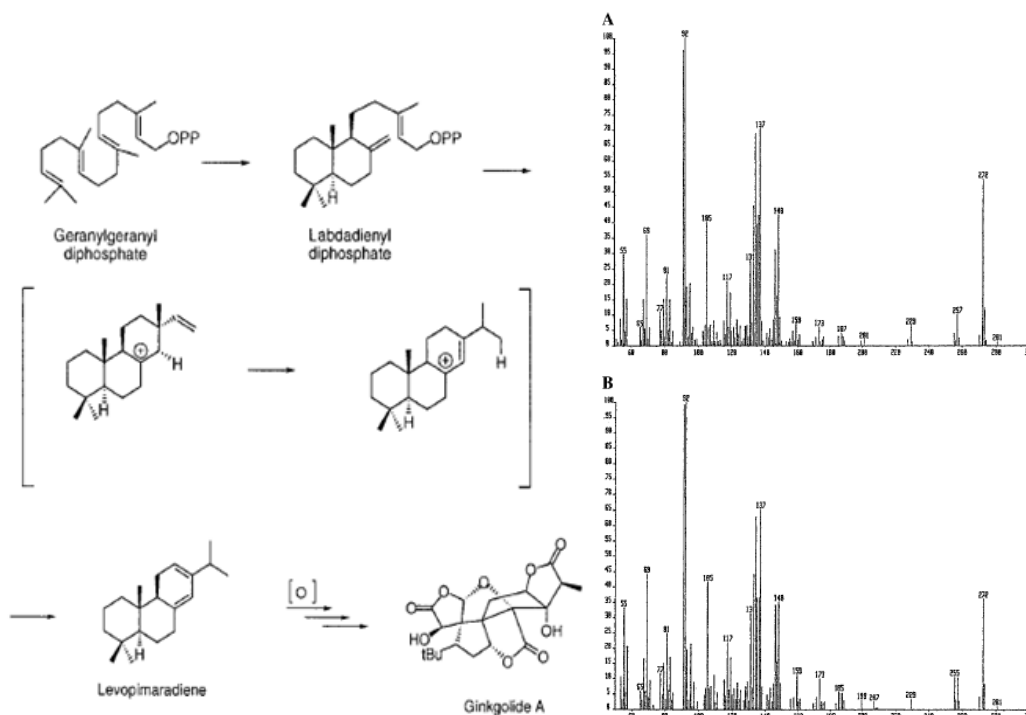


Figure 1-07. Left: Cyclization of GGDP to levopimaradiene en route to ginkgolide A. Right: GCMS of biosynthetic (A) and synthetic (B) levopimaradiene.⁹

For instance, levopimaradiene synthase, a key enzyme in ginkgolide biosynthesis was cloned and functionally characterized via GC/MS.⁹ Ginkgolides are a family of diterpenoids isolated from *Ginkgo biloba* leaf extracts, which are highly specific platelet-activating factor receptor antagonists. Levopimaradiene synthase catalyzes the initial cyclization step of geranylgeranyl diphosphate (GGDP) to levopimaradiene, a key intermediate in ginkgolide biosynthesis (Figure 1-07). Matsuda and coworkers⁹ expressed *G. biloba* cDNA encoding this

synthase into *Escherichia coli* and incubated resultant cell lysates with GGDP. Low yields of levopimaradiene required a sensitive detection method to assess the enzyme's activity. Thus, all reactions were extracted with hexane, concentrated, and analyzed by GC/MS. Identical elution time and fragmentation pattern of biosynthesized levopimaradiene to a synthetic standard verified that the enzyme product was indeed levopimaradiene (Figure 1-07).⁹

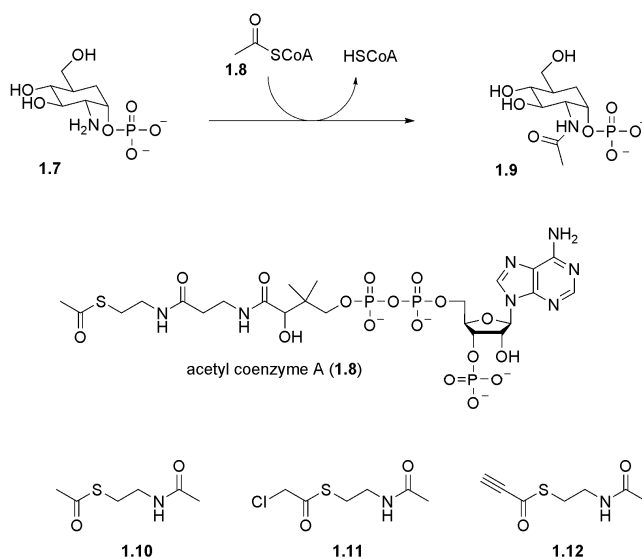


Figure 1-08. Acyltransferase activity as reported by Pohl *et al*¹⁰

When the function of an enzyme has been established, further studies involving enzyme mechanism, kinetics and its substrate specificity or flexibility can be undertaken using MS. For example, Pohl and coworkers studied the kinetics and substrate specificity of the acyltransferase activity of a bifunctional enzyme using an MS based assay.¹⁰ In short, the cloned enzyme was incubated in the presence of glucosamine-1-phosphate (1.7) and acetyl coenzyme A (1.8) (Figure 1-08). Over time, peaks corresponding to 1.7 ($m/z = 258$) disappeared as

new peaks corresponding to *N*-acetylglucosamine-1-phosphate (**1.9**, $m/z = 300$) appeared. Peak quantification with internal standards and calibration curves allowed the calculation of kinetic constants (K_M , k_{cat} , and k_{cat}/K_M) for the reaction. In addition, the substrate flexibility of the enzyme was probed by using truncated analogues of acetyl coenzyme A (**1.10-1.12**). The structures of products formed were determined by the same MS method used for enzyme kinetics. The modifications pictured may not deviate vastly from the original structure of acetyl coenzyme A, but could allow the researchers to tag the product using highly specific reactions, such as alkyne/azide coupling.¹¹

Recent advances in MS instrumentation allow the direct investigation of covalent and non-covalent enzyme intermediates. Anderson and coworkers¹² reported the first direct identification of a hypothesized unstable hemiketal phosphate intermediate bound to the enzyme 3-deoxy-*D*-manno-2-octulosonate-8-phosphate (KDO8P) synthase in a noncovalent complex. KDO8P synthase is essential for Gram negative bacteria as it is involved in the lipopolysaccharide biosynthetic pathway. It catalyzes the condensation of arabinose-5-phosphate (A5P), H_2O , and phosphoenolpyruvate (PEP) to form KDO8P and inorganic phosphate (P_i) (Figure 1-09). It is thought that the mechanism involves labile intermediate **I**, but this intermediate had never been directly observed before. High resolution ESI-MS interfaced with a novel rapid-mixing technique allowed the detection of the enzyme bound intermediate (**E•I**) along with enzyme complexes with each substrate and each product at different ratios as the reaction progressed (Figure 1-10).

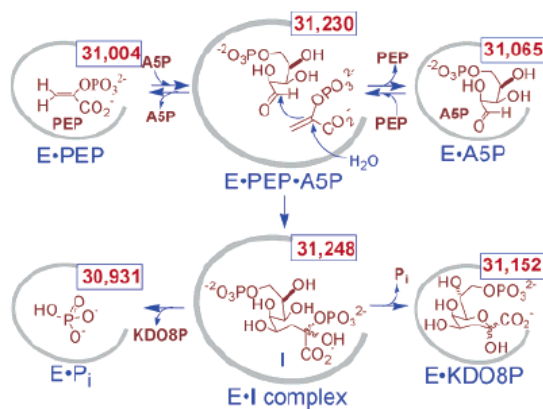


Figure 1-09. KDO8P synthase reaction pathway¹²

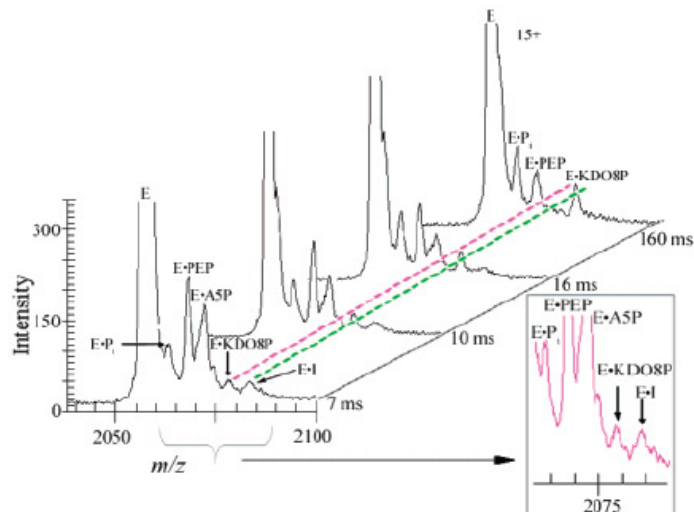


Figure 1-10. ESI-TOF mass spectra of KDO8P synthase during catalysis.¹²

An example of the use of MS to study enzyme intermediates has been the investigation of nonribosomal peptide synthetases (NRPS) and polyketide synthases (PKS).¹³ Both classes of enzymes biosynthesize their natural products via covalent intermediates on phosphopantetheine arms. The yersiniabactin biosynthetic pathway represents the most thoroughly studied NRPS/PKS hybrid

system.¹⁴ All covalent intermediates in the pathway have been identified by MS (Figure 1-11).

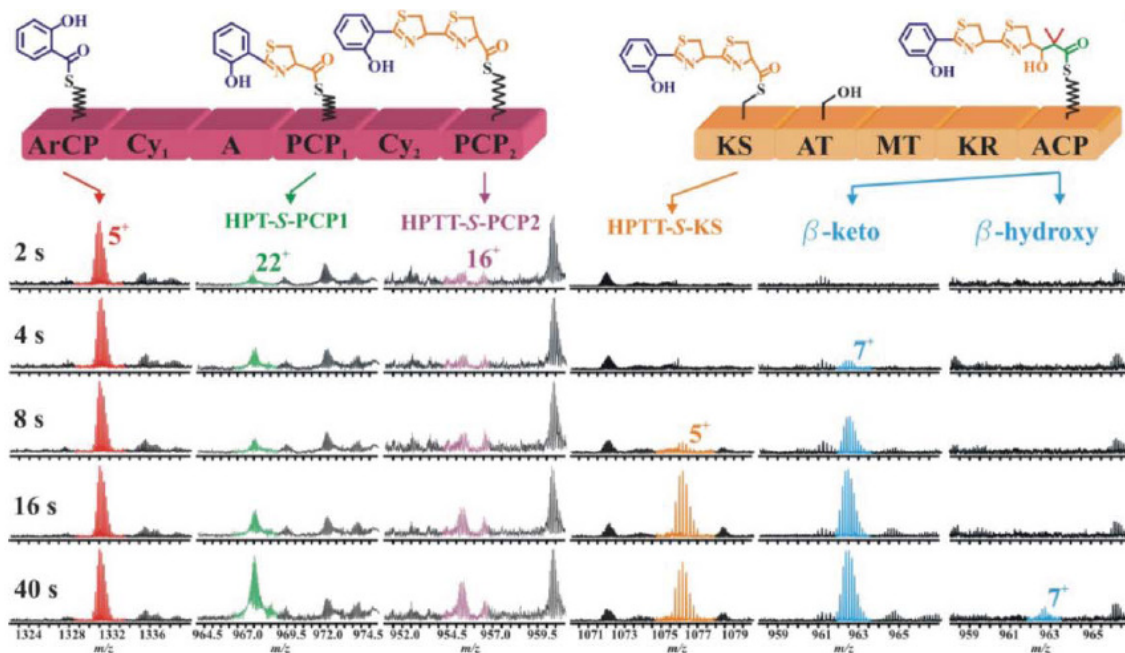


Figure 1-11. Detection of enzyme-bound intermediates in the yersiniabactin biosynthetic pathway.^{13, 14}

In summary, MS is a powerful technique that can be used at every stage of natural product studies, from discovery and structural characterization of new compounds to biosynthetic enzyme identification and manipulation. This dissertation illustrates the utility of MS in NP investigations by describing its use in two natural product studies: I) K-26, a hypotensive phosphonopeptide metabolite, and II) apoptolidin, an anticancer natural product of polyketide origin.

Case Study I: K-26, a hypotensive phosphonopeptide metabolite

Natural and synthetic products containing a carbon-phosphorus (C-P) bond have been found to possess potent biological activities and unique biological roles in primary and secondary metabolism (Table 1-1) and thus, they have been studied thoroughly over the last fifty years. Significant biological activities of C-P containing natural products include herbicidal, antibacterial, antiparasitic, and antihypertensive activities.^{15, 16} These activities can be attributed to the physical and structural similarity of phosphonic and phosphinic acids to the phosphate ester and carboxylic acid groups present in many biomolecules. In this way, phosphonates can act as substrate or transition state analogs and inhibit enzyme catalysis. In addition, the C-P bond is chemically and thermally highly inert making organophosphonates resistant to chemical hydrolysis, thermal decomposition and photolysis.

Table 1-1. Phosphonates in current use

Phosphonate	Class	Application
Cidofovir ¹	antiviral	cytomegalovirus, HIV
Adefovir ¹	antiviral	hepatitis B, HIV
Fosamax ¹	antiosteoclastic	osteoporosis
Fosfomycin ²	antibacterial	systemic
Fosmidomycin ²	antiparasitic	Malaria
Bialaphos ²	herbicidal	weed control

¹synthetic compounds ²natural products

The first C-P containing natural product to be identified and isolated from natural sources is 2-aminoethylphosphonic acid (AEP; **1.13**). It is one of the most abundant naturally occurring phosphonates. It is found in marine invertebrates and microorganisms as part of phosphonolipids and proteoglycans, although it has also been isolated from human brain.^{15, 16} Other C-P bond containing natural products include bialaphos (**1.16**; antifungal, antibiotic, herbicide), fosfomycin (**1.14**; antibiotic), fosmidomycin (**1.15**; antimalarial), and K-26 (**1.17**; ACE inhibitor).

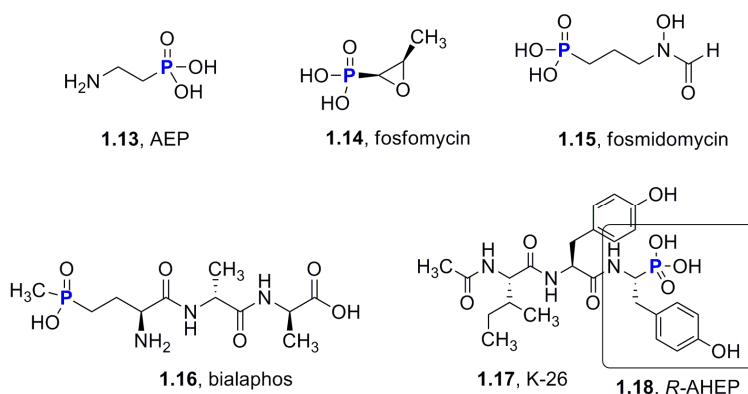


Figure 1-12. Bioactive C-P containing natural products.

Considering such diverse activities, it is surprising that only a small number of C-P containing natural products have been characterized biosynthetically. Bialaphos (**1.16**) is a member of the phosphinothricin family, a well-characterized subclass containing a phosphonic acid analog of glutamate, and is used as a natural herbicide. Members of this subclass have been studied extensively and the mechanism by which the C-P bond is formed in all of them seems to be unified. The key enzyme in the biosynthesis of bialaphos and all

studied C-P bond containing natural products to date is phosphoenolpyruvate mutase (PEP mutase).¹⁷⁻²⁰ PEP mutase catalyzes the intramolecular rearrangement of phosphoenolpyruvate to phosphonopyruvate (Ppyr). The decarboxylation of Ppyr produces phosphonoacetaldehyde, which is a biosynthetic precursor of AEP, fosfomycin, and bialaphos (Figure 1-13).

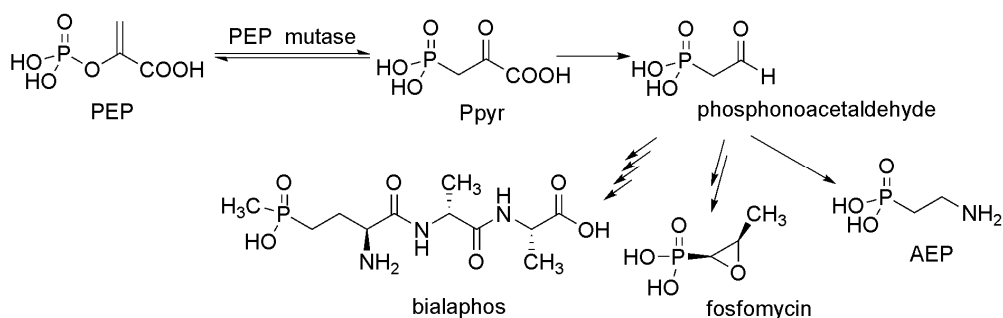


Figure 1-13. C-P bond formation by PEP mutase and resulting C-P containing natural products.

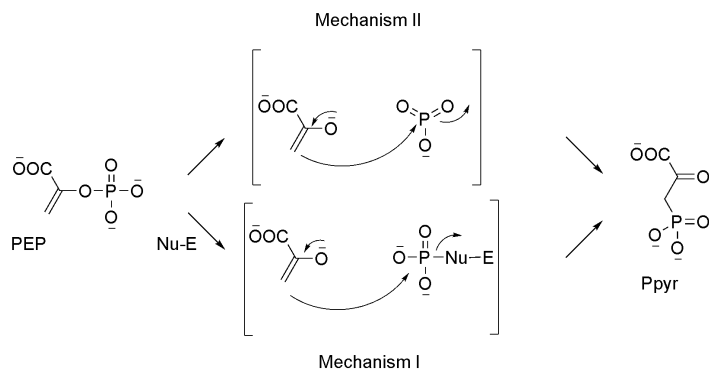
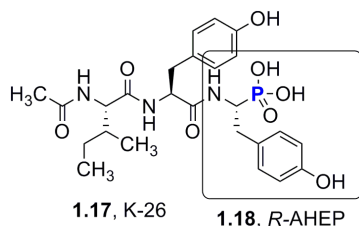


Figure 1-14. Possible mechanisms of PEP mutase.

PEP mutase has been studied extensively but the exact catalytic mechanism is still a matter of debate. It has been suggested that the phosphoryl transfer occurs intramolecularly and with retention of stereochemistry at

phosphorus. There are two possible mechanisms as indicated in Figure 1-14. In mechanism I the phosphoryl group is first transferred to an active site residue of the enzyme and then followed by addition to the C(3) of the pyruvate enol intermediate. Mechanism II involves dissociation of a PO_3^- group followed by addition to the C(3) of the pyruvate enol intermediate.¹⁷⁻²⁰

A second example of a C-P bond forming enzyme is carboxyphosphoenol pyruvate (CPEP) mutase.²¹⁻²³ CPEP mutase catalyzes the formation of the second C-P bond in bialaphos. The two enzymes have 25% amino acid sequence identity and 45% amino acid sequence similarity, suggesting they possess a similar structure and catalytic mechanism.



K-26 (**1.17**) from Actinomycete sp. K-26 (NRRL 12379)²⁴⁻²⁸ is representative of an uninvestigated class of natural phosphonates which incorporate a phosphonic acid analog of tyrosine. K-26 is reported to possess angiotensin converting enzyme inhibitory activity with an IC_{50} value of 12.5 nM, comparable to the widely prescribed antihypertensive drug Captopril. K-26 was initially discovered via ACE bioassay guided fractionation of extracts of Actinomycete sp. K-26. NMR, mass spectrometry, degradation and synthetic studies have demonstrated that K-26 is comprised of N-acetylated L-isoleucine,

tyrosine, and the non-proteinogenic amino acid, (*R*)-1-amino-2-(4-hydroxyphenyl)ethylphosphonic acid (AHEP, **1.18**). This phosphotyrosine functionality is shared among several related ACE inhibitors produced by *Streptosporangium*²⁹ and *Actinomadura*²⁶ species.

Despite the potent hypotensive activity of K-26 and related compounds, the biosynthetic pathways by which K-26 and more specifically AHEP are biosynthesized remain uncharacterized. The resemblance of AHEP to the amino acid tyrosine suggests that its origin lies in the shikimic acid pathway. Although PEP mutase has been found to be common in the biosynthesis of all C-P containing natural products to date, it is difficult to rationalize how a PEP mutase generated precursor or analogous mutase reaction could lead to the formation of AHEP due to the positioning of the aromatic ring in AHEP. Comparing the structure of bialaphos (**1.16**) to AHEP (**1.18**), it becomes apparent that the C-P bond formation in these two molecules may be substantially different. Our goal is to elucidate the mechanism of C-P bond formation of K-26, which appears to be distinct from previously investigated C-P containing natural products.

K-26 is a small peptide molecule containing a highly unusual amino acid. It belongs in a subclass of bioactive natural products, which contain not only the common twenty amino acids but also unique groups such as D-amino acids, heterocycles, and amino acid analogs. In contrast to proteins, small peptide natural products are frequently not synthesized by the ribosome in bacteria.³⁰ Instead they are constructed by nonribosomal peptide synthetases (NRPS). These synthetases are organized in modules. Each module is responsible for the

incorporation of one specific amino acid into the final product. It is further subdivided into domains responsible for amino acid recognition, activation, modification, condensation and release. The different modules may or may not be on the same protein.³⁰ We hypothesize that there is a NRPS system for the biosynthesis of K-26 and that AHEP is appended to the dipeptide precursor by a megasynthetase (Figure 1-15). Alternatively, K-26 may be assembled via an NRPS-independent (NIS) pathway³¹ or a ligase similar to the ones involved in glutathione biosynthesis.³²

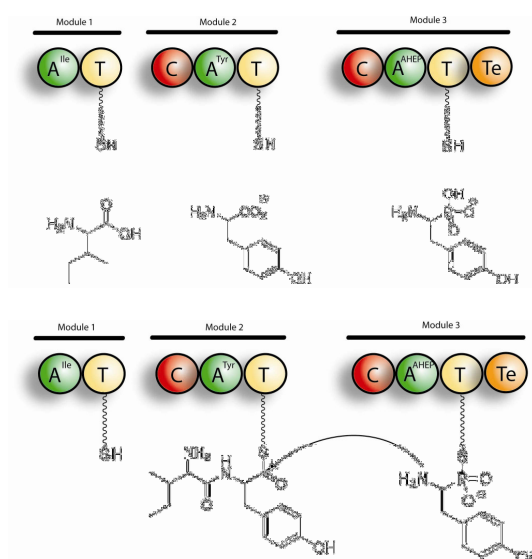


Figure 1-15. Biosynthetic hypothesis for K-26 formation by an NRPS.

Our goal is to understand the mechanism by which K-26 is biosynthesized, as a starting point for the production of new natural products containing the same functionality. Comprehending the biosynthetic system

responsible for the synthesis of AHEP, a proven pharmacophore, is an important step towards the design of more potent C-P containing natural products.

Case Study II: Apoptolidin, an anticancer natural product of polyketide origin

Apoptolidin A (**1.19**) is a glycosylated macrolactone that exhibits selective activity against cancer cells.³³ It was initially isolated by Seto and coworkers in 1997 from the fermentation of soil dwelling actinomycete *Nocardioopsis sp* using a screen designed to identify specific apoptosis inducers.³³ The molecular formula of apoptolidin A was determined as C₅₈H₉₆O₂₁ by high-resolution FAB-MS and extensive NMR studies were used to obtain its molecular structure.^{33, 34} Apoptolidin has a highly unsaturated 20-membered macrolide with a side chain containing a 6-membered hemiketal. A novel sugar, 6-deoxy-4-O-methyl-L-glucose, is attached to O9 and a disaccharide consisting of L-olivomycose and D-oleandrose is appended to O27. Originally, the isolation yield of apoptolidin A was reported as 109 mg/L of fermentation medium.³³ Wender and coworkers optimized the culture conditions resulting in an increased yield of 130mg/L of apoptolidin A, as well as the isolation of two apoptolidin A analogues, apoptolidin B (**1.20**) and apoptolidin C (**1.21**) at low milligram amounts.³⁵ Apoptolidin B and C only differ from apoptolidin A by the absence of the hydroxyl group on C16 (both B and C) and on C20 (only C). Recently, a new analogue was isolated and characterized, named apoptolidin D (**1.22**).³⁶ Its only difference with apoptolidin A is the absence of a methyl group on C6 (Figure 1-16).

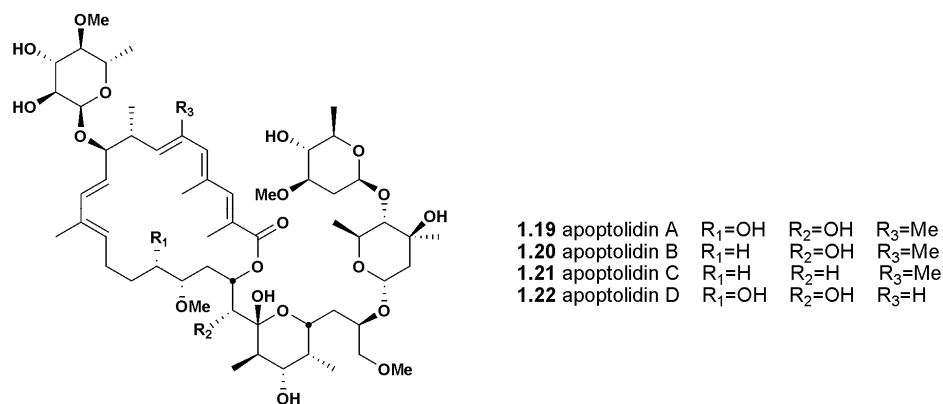


Figure 1-16. Structures of apoptolidins A-D

It has been reported that apoptolidin A can rearrange into its 21-membered isomer isoapoptolidin A (**1.23**) by acyl migration from O19 to O20.^{37, 38} The equilibrium constant of this rearrangement is about 0.6 in Dulbecco's phosphate-buffered saline (PBS) at 37°C.³⁸ Apoptolidin A has been found to be stable for up to three months at -20 °C when dissolved in chloroform or methylene chloride. Similarly, isoapoptolidin A is stable as a solution in methanol at -20 °C for up to six months. In contrast, a dilute aqueous solution of apoptolidin A at ambient temperature can be observed to convert to isoapoptolidin when monitored by HPLC.³⁸ Isomerization of apoptolidin B to isoapoptolidin B (**1.24**) and apoptolidin D to isoapoptolidin D (**1.25**) has also been observed.^{35, 36} Apoptolidin C is unable to undergo this rearrangement as it is missing the hydroxyl group on C20.

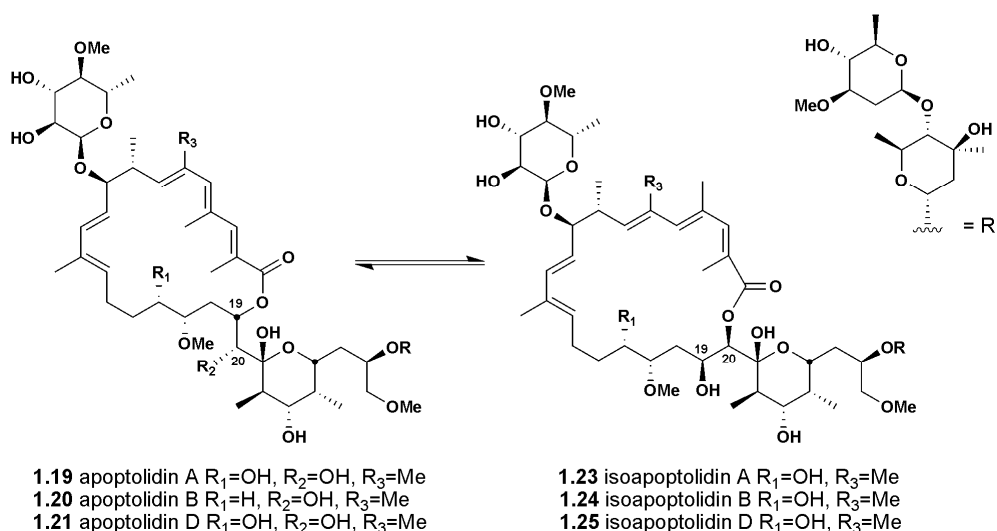


Figure 1-17. Isomerization of apoptolidins

Apoptolidin A was shown to induce apoptosis at concentrations as low as 10nM in oncogene-containing E1A-transformed glial cells.³³ In contrast, normal glial cells were insensitive to apoptolidin A even at micromolar concentrations. It was found that apoptolidin was among the top 0.1% most selective agents tested in the National Cancer Institute's 60-cell line screening panel.³⁹ This high selectivity of apoptolidin for cancer cells over normal cells makes it an attractive lead for the development of new anticancer agents with minimal side effects.

The mechanism of action of apoptolidin was investigated by Khosla and coworkers. Due to the resemblance of the apoptolidin A aglycone (**1.26**, apoptolidinone A) to oligomycin (**1.30**), a known mitochondrial F_0F_1 -ATPase inhibitor, apoptolidin A was tested against F_0F_1 -ATPase and was found to be a potent inhibitor ($IC_{50} = 0.7\mu M$).³⁹ These results provided evidence that inhibition of mitochondrial F_0F_1 -ATPase may be the basis of apoptolidin's biological activity. However, structure-activity relationship (SAR) studies undertaken by

Wender and coworkers proved that *in vitro* inhibition of mitochondrial F₀F₁-ATPase is not an accurate predictor of the *in vivo* activity of apoptolidin derivatives, the selective induction of apoptosis in cancer cells.⁴⁰ These findings suggest that there might be an alternate target of apoptolidin and its mode of action may be more complex than originally suggested.

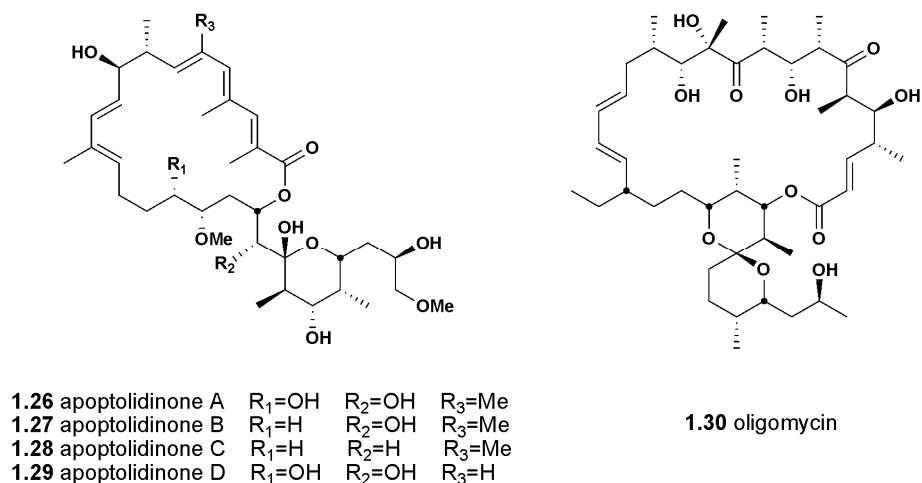


Figure 1-18. Structures of apoptolidinones and oligomycin

We intend to synthesize apoptolidin analogues by precursor directed biosynthesis and use these compounds to identify the cellular target of apoptolidin. Specifically, we plan to supplement the fermentation broth of *Nocardioopsis* sp. with synthetic aglycone analogues and use the glycosylating machinery of this bacterial organism to glycosylate the unnatural aglycones. In this way, we will be able to introduce an azido handle on the apoptolidin core in order to conjugate a “reporter” group consisting of a fluorophore or biotin. The resulting chemical probe will be used to screen the proteome of apoptolidin

sensitive cancer cells and after purification, the protein adduct may be identified by mass spectrometry.

Dissertation Goals

This dissertation exemplifies the significance of mass spectrometry as a tool in natural product studies. The versatility of this tool enables scientists to use MS to isolate new natural products, structurally characterize them, identify their biosynthetic intermediates and enzymes and even study their mechanism of action. In detail:

Chapter II describes structure-activity relationship studies undertaken in order to identify the ACE pharmacophore in phosphonopeptide K-26. A series of K-26 analogues were synthesized and their ACE inhibitory activity was evaluated in order to determine critical functional determinants.

Chapter III describes the identification of K-26 and AHEP biosynthetic precursors via heavy-atom isotope labeled substrate incorporation experiments. A highly sensitive selected reaction monitoring (SRM) based method for isotopic incorporation estimation was developed. The results suggest a new mechanism of phosphonate group biosynthesis in natural products.

Chapter IV describes our attempts in isolating the *N*-acetyltransferase (NAT) involved in the K-26 biosynthesis. Several chromatographic techniques were used to purify the NAT. An SRM based assay was developed for detecting acetylation and proteomics were employed for identification of the K-26 relevant NAT.

Chapter V describes the use of LC/MS in the production and isolation of apoptolidin analogues. Apoptolidinone analogues were introduced to growing cultures of *Nocardiosis* sp. and their glycosylation was detected by mass spectrometry.

References

1. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs over the last 25 years. *Journal of natural products* **2007**, 70, (3), 461-77.
2. Harvey, A., Strategies for discovering drugs from previously unexplored natural products. *Drug Discov Today* **2000**, 5, (7), 294-300.
3. Stroobant, V.; Hoffmann, E. d., *Mass spectrometry: principles and applications*. 2nd ed.; Wiley: New York, 2001.
4. Shipovskov, S.; Reimann, C. T., Electrospray ionization mass spectrometry in enzymology: uncovering the mechanisms of two-substrate reactions. *Analyst* **2007**, 132, (5), 397-402.
5. He, M.; Haltli, B.; Summers, M.; Feng, X.; Hucul, J., Isolation and characterization of meridamycin biosynthetic gene cluster from *Streptomyces* sp. NRRL 30748. *Gene* **2006**, 377, 109-18.
6. Chen, G.; Pramanik, B. N.; Bartner, P. L.; Saksena, A. K.; Gross, M. L., Multiple-stage mass spectrometric analysis of complex oligosaccharide antibiotics (everninomicins) in a quadrupole ion trap. *J Am Soc Mass Spectrom* **2002**, 13, (11), 1313-21.
7. Ganguly, A. K.; Pramanik, B. N.; Girijavallabhan, V. M.; Sarre, O.; Bartner, P. L., The use of fast atom bombardment mass spectrometry for the determination of structures of everninomicins. *The Journal of antibiotics* **1985**, 38, (6), 808-12.
8. Charan, R. D.; Schlingmann, G.; Bernan, V. S.; Feng, X.; Carter, G. T., Dioxapyrrolomycin biosynthesis in *Streptomyces fumanus*. *Journal of natural products* **2006**, 69, (1), 29-33.
9. Schepmann, H. G.; Pang, J.; Matsuda, S. P., Cloning and characterization of *Ginkgo biloba* levopimaradiene synthase which catalyzes the first committed step in ginkgolide biosynthesis. *Arch Biochem Biophys* **2001**, 392, (2), 263-9.
10. Mizanur, R. M.; Jaipuri, F. A.; Pohl, N. L., One-step synthesis of labeled sugar nucleotides for protein O-GlcNAc modification studies by chemical function analysis of an archaeal protein. *J Am Chem Soc* **2005**, 127, (3), 836-7.

11. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie (International ed)* **2002**, 41, (14), 2596-9.
12. Li, Z.; Sau, A. K.; Shen, S.; Whitehouse, C.; Baasov, T.; Anderson, K. S., A snapshot of enzyme catalysis using electrospray ionization mass spectrometry. *J Am Chem Soc* **2003**, 125, (33), 9938-9.
13. Dorrestein, P. C.; Kelleher, N. L., Dissecting non-ribosomal and polyketide biosynthetic machineries using electrospray ionization Fourier-Transform mass spectrometry. *Natural product reports* **2006**, 23, (6), 893-918.
14. McLoughlin, S. M.; Kelleher, N. L., Kinetic and regiospecific interrogation of covalent intermediates in the nonribosomal peptide synthesis of yersiniabactin. *J Am Chem Soc* **2004**, 126, (41), 13265-75.
15. Fields, S. C., Synthesis of natural products containing a C-P bond. *Tetrahedron* **1999**, 55, (42), 12237-12272.
16. Seto, H.; Kuzuyama, T., Bioactive natural products with carbon-phosphorus bonds and their biosynthesis. *Natural product reports* **1999**, 16, (5), 589-96.
17. Jia, Y.; Lu, Z.; Huang, K.; Herzberg, O.; Dunaway-Mariano, D., Insight into the mechanism of phosphoenolpyruvate mutase catalysis derived from site-directed mutagenesis studies of active site residues. *Biochemistry* **1999**, 38, (43), 14165-73.
18. Kim, J.; Dunaway-Mariano, D., Phosphoenolpyruvate mutase catalysis of phosphoryl transfer in phosphoenolpyruvate: kinetics and mechanism of phosphorus-carbon bond formation. *Biochemistry* **1996**, 35, (14), 4628-35.
19. Lee, S. H.; Hidaka, T.; Nakashita, H.; Seto, H., The carboxyphosphoenolpyruvate synthase-encoding gene from the bialaphos-producing organism *Streptomyces hygroscopicus*. *Gene* **1995**, 153, (1), 143-4.
20. Liu, S.; Lu, Z.; Jia, Y.; Dunaway-Mariano, D.; Herzberg, O., Dissociative phosphoryl transfer in PEP mutase catalysis: structure of the

enzyme/sulfoxyruvate complex and kinetic properties of mutants. *Biochemistry* **2002**, 41, (32), 10270-6.

21. Freeman, S.; Pollack, S. J.; Knowles, J. R., Synthesis of the unusual metabolite carboxyphosphoenolpyruvate. Cloning and expression of carboxyphosphoenolpyruvate mutase. *J Am Chem Soc* **1992**, 114, (1), 377-378.

22. Hidaka, T.; Imai, S.; Hara, O.; Anzai, H.; Murakami, T.; Nagaoka, K.; Seto, H., Carboxyphosphoenolpyruvate phosphonmutase, a novel enzyme catalyzing C-P bond formation. *Journal of bacteriology* **1990**, 172, (6), 3066-72.

23. Hidaka, T.; Imai, S.; Seto, H., Biosynthesis mechanism of carbon-phosphorus bond formation. Isolation of carboxyphosphoenolpyruvate and its conversion to phosphinopyruvate. *J Am Chem Soc* **1989**, 111, (20), 8012-8013.

24. Hirayama, N.; Kasai, M.; Shirahata, K., Structure and Conformation of a Novel Inhibitor of Angiotensin-I Converting Enzyme - a Tripeptide Containing Phosphonic Acid. *International Journal of Peptide and Protein Research* **1991**, 38, (1), 20-24.

25. Kasai, M.; Yoshida, N.; Hirayama, N.; Shirahata, K., Structure Elucidation of New Inhibitors of angiotensin I Converting Enzyme, K-26 and K-4. *Symposium Papers, The 27th Symposium on the Chemistry of Natural Products* **1985**, 577.

26. Koguchi, T.; Yamada, K.; Yamato, M.; Okachi, R.; Nakayama, K.; Kase, H., K-4, a Novel Inhibitor of Angiotensin-I Converting Enzyme Produced by *Actinomadura-Spiculosospora*. *Journal of Antibiotics* **1986**, 39, (3), 364-371.

27. Yamato, M.; Kase, H.; Kawamoto, I.; Kasai, M.; Shirahata, K.; Deguchi, T.; Shuto, K.; Okachi, R.; Nakayama, K. Novel physiologically active substances K-26, a process for preparation thereof and a pharmaceutical composition containing the same. US4,683,220, 1987.

28. Yamato, M.; Koguchi, T.; Okachi, R.; Yamada, K.; Nakayama, K.; Kase, H.; Karasawa, A.; Shuto, K., K-26, a Novel Inhibitor of Angiotensin-I Converting Enzyme Produced by an Actinomycete K-26. *Journal of Antibiotics* **1986**, 39, (1), 44-52.

29. Ohuchio, S.; Kurihara, K.; Shinohara, A.; Takei, T.; Yoshida, J.; Amano, S.; Miyadoh, S.; Matsushida, Y.; Somura, T.; Sezaki, M., Studies on New Angiotensin Converting Enzyme Inhibitors, SF2513A,B and C, Produced by *Streptosporangium nondiastaticum*. *Sci. Reports of Meiji Seika Kaisha* **1988**, *27*, 46 - 54.
30. Sieber, S. A.; Marahiel, M. A., Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem Rev* **2005**, *105*, (2), 715-38.
31. Challis, G. L., A widely distributed bacterial pathway for siderophore biosynthesis independent of nonribosomal peptide synthetases. *Chembiochem* **2005**, *6*, (4), 601-11.
32. Janowiak, B. E.; Griffith, O. W., Glutathione synthesis in *Streptococcus agalactiae*. One protein accounts for gamma-glutamylcysteine synthetase and glutathione synthetase activities. *The Journal of biological chemistry* **2005**, *280*, (12), 11829-39.
33. Kim, J. W.; Adachi, H.; Shin-ya, K.; Hayakawa, Y.; Seto, H., Apoptolidin, a new apoptosis inducer in transformed cells from *Nocardiosis* sp. *The Journal of antibiotics* **1997**, *50*, (7), 628-30.
34. Hayakawa, Y.; Kim, J. W.; Adachi, H.; Shin-ya, K.; Fujita, K. i.; Seto, H., Structure of Apoptolidin, a Specific Apoptosis Inducer in Transformed Cells. *J Am Chem Soc* **1998**, *120*, (14), 3524-3525.
35. Wender, P. A.; Sukopp, M.; Longcore, K., Apoptolidins B and C: isolation, structure determination, and biological activity. *Organic letters* **2005**, *7*, (14), 3025-8.
36. Wender, P. A.; Longcore, K. E., Isolation, structure determination, and anti-cancer activity of apoptolidin D. *Organic letters* **2007**, *9*, (4), 691-4.
37. 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* **2002**, *4*, (22), 3823-5.

38. Wender, P. A.; Gullledge, A. V.; Jankowski, O. D.; Seto, H., Isoapoptolidin: structure and activity of the ring-expanded isomer of apoptolidin. *Organic letters* **2002**, 4, (22), 3819-22.
39. Salomon, A. R.; Voehringer, D. W.; Herzenberg, L. A.; Khosla, C., Understanding and exploiting the mechanistic basis for selectivity of polyketide inhibitors of F(0)F(1)-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, (26), 14766-71.
40. Wender, P. A.; Jankowski, O. D.; Longcore, K.; Tabet, E. A.; Seto, H.; Tomikawa, T., Correlation of F0F1-ATPase inhibition and antiproliferative activity of apoptolidin analogues. *Organic letters* **2006**, 8, (4), 589-92.

CHAPTER II

IDENTIFICATION OF ACE PHARMACOPHORE IN K-26

Introduction

Phosphonic acids are important pharmacophores of significant relevance to human health. Under physiological conditions, the phosphonate moiety ($R-PO_3H_2$, $pK_a \sim 2.5, 8.0^1$) can serve as an isosteric replacement for phosphate or carboxylate functional groups, which are ubiquitous ligands in the active sites of many enzymes. The mode of action of phosphonates is often either through competitive interaction with substrate binding regions or as an analog of tetrahedral transition states. As a result, synthetic phosphonates find real world application in medicine as antibiotics, antivirals, antiosteoclastics and as environmentally benign herbicides.²

Naturally occurring phosphonopeptide K-26 (**1.17**) was initially discovered via ACE bioactivity guided fractionation of extracts of a soil dwelling prokaryote, "actinomycete strain K-26".³⁻⁵ It has been reported to possess ACE inhibitory activity comparable to the widely prescribed antihypertensive drug Captopril in both *in vitro* assays and in intravenously administered animal models. NMR, mass spectrometry, degradation and synthetic studies have demonstrated that K-26 is comprised of *N*-acetylated L-isoleucine, L-tyrosine, and the nonproteinogenic amino acid, (*R*)-1-amino-2-(4-hydroxyphenyl)ethylphosphonic acid (AHEP, **1.18**). This "phosphonotyrosine" functional group is shared among

several ACE inhibitor peptide analogues of K-26 produced by various *Streptosporangium*⁶ and *Actinomadura*⁷ species. The stereochemistry of AHEP has been established by x-ray analysis of the ethyl ester of K-26.

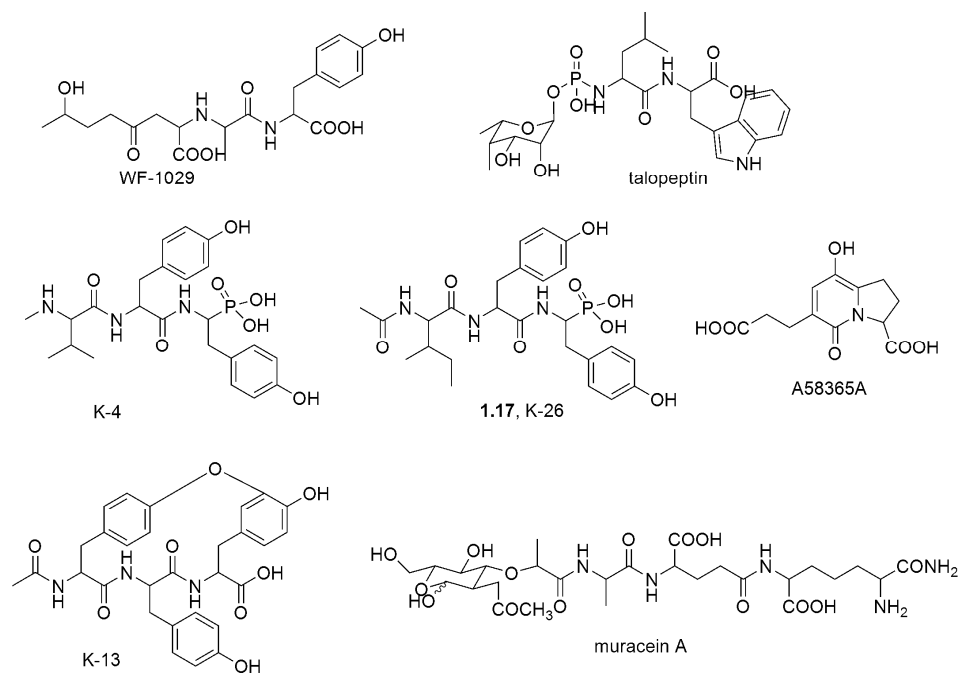


Figure 2-01. ACE inhibitors produced by bacteria

The isolation of a mammalian ACE modulating substance from a bacterium may appear somewhat surprising. However, there are several ACE inhibitors (Figure 2-01) derived from microbial sources. These include WF-1029 from *Doratomyces putredinis*,⁸ talopeptin from *Streptomyces mozunensis*,⁹ K-4 from *Actinomadura*,⁷ K-13 from *Micromonospora halophytica*,¹⁰ muracein A from *Nocardia orientalis*,¹¹ and A58365A from *Streptomyces chromofuscus*.¹² In addition, recently bacterial ACE homologues have been identified in *Xanthomonas*¹³, suggesting that the bacterial homologues of eukaryotic ACE

may play important roles in bacterial physiology. While the natural target of these terminal phosphonopeptides remains unidentified, the potent inhibitory activity of this entire class of compounds towards ACE suggests the possibility that these compounds may target analogous metalloproteases, either endogenously or metalloproteases of consequence in the ecological niche of the producing organisms.

ACE inhibitors are widely prescribed for cardiovascular diseases, including high blood pressure, heart failure, heart attack and kidney failure and have combined annual sales in excess of six billion dollars.¹⁴ ACE has a critical role in the regulation of blood pressure by catalyzing the hydrolytic cleavage of His-Leu dipeptide from decapeptide angiotensin I to yield angiotensin II, a potent vasoconstrictor. ACE also cleaves bradykinin resulting in antagonistic effects to angiotensin II. Human endothelial somatic ACE is a monomeric zinc metalloenzyme, heavily glycosylated, localized to the plasma membrane and characterized by the presence of two homologous domains, each containing a functional active site.¹⁴ Typical ACE inhibitors consist of peptide or peptidomimetic structures incorporating a functional group that interacts with the active site zinc, such as a thiol, a carboxylate, a phosphinic acid (Figure 2-02).

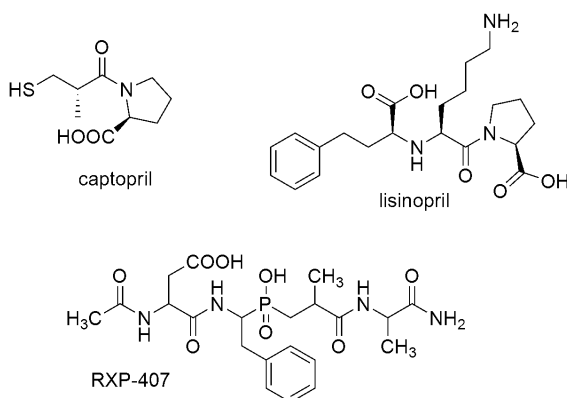


Figure 2-02. Representative ACE inhibitors

It has been previously suggested that the phosphonate moiety of K-26 may interact with the active site zinc atom and contribute to the potent inhibitory activity of this class of compounds.⁴ However, small proteinogenic di- and tripeptides have also been reported to possess significant ACE activity. For instance isoleucyl tyrosine and other proteinogenic analogs of K-26 have been reported to be low micromolar inhibitors of mammalian ACE¹⁵ and it has remained unclear how the ACE activity of K-26 is modulated by the substitution of carboxyl group functionality and *N*-acetylation. To identify the structural determinants of ACE activity in the class of chiral α -amino phosphonyl natural products we have synthesized K-26 and several peptide analogues.

Results and Discussion

Production levels of K-26 in the producing organism are low (est. < 10 $\mu\text{g/L}$ in our hands), necessitating the development of a synthetic route for K-26 and analogues. AHEP was synthesized as previously described¹⁶⁻²⁰ with minor modifications. Commercially available 4-benzyloxyphenylacetyl chloride (**2.1**)

was reacted with triethyl phosphite in an Arbuzov type reaction to yield the corresponding α -keto phosphonate (not isolated). Subsequent conversion to the oxime (**2.2**) with hydroxylamine and reduction with zinc/formic acid resulted in racemic AHEP diethyl ester (**2.3**). Standard peptide coupling conditions²¹ were employed in coupling the free base of AHEP diethyl ester to Ac-Ile-Tyr(Bzl)-COOH (**2.4**). In a two-step deprotection process, benzyl groups were removed from the resultant tripeptide by catalytic hydrogenation and ethyl groups were removed by reaction with iodotrimethylsilane. This synthetic route resulted in the preparation of a mixture of two major diastereomers of K-26 (**1.17a** and **1.17b**), which were readily separated via C18 chromatography using a gradient of 5-40% acetonitrile containing 0.1% trifluoroacetic acid. Stereochemistry was assigned by comparison of chemical shift values and NOESY cross-peak correlations to peptide analogue Ac-Ile-Tyr-Tyr (**2.7a**) and Ac-Ile-Tyr-D-Tyr (**2.7b**).

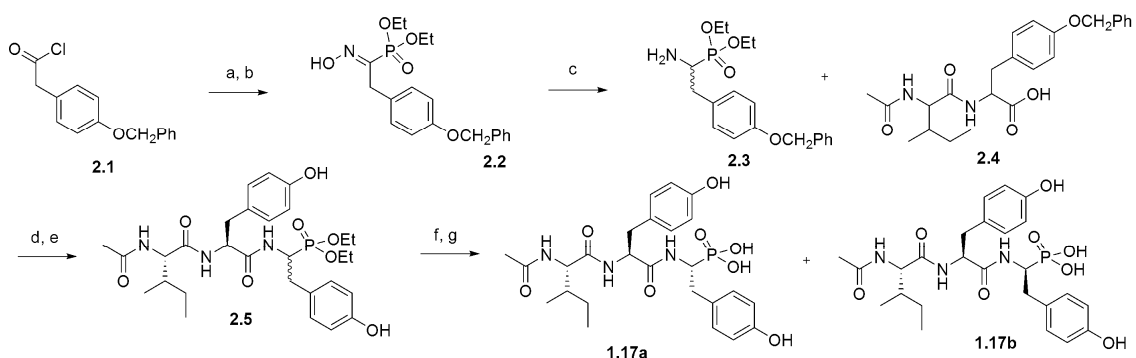


Figure 2-03. Synthesis of two diastereomers of K-26. Reagents and conditions: (a) $\text{P}(\text{OEt})_3$, THF, $0 \rightarrow 70^\circ\text{C}$; (b) $\text{NH}_2\text{OH}\cdot\text{HCl}$, pyridine, ethanol, 25°C ; (c) Zn/HCOOH , 25°C ; (d) $\text{EDC}\cdot\text{HCl}$, HOBT, 2,4,6-trimethylpyridine, DMF, 25°C ; (e) H_2 , Pd/C, 25°C ; (f) TMSI, thioanisole, MeCN, 0°C .

Des-acetyl K-26 (**2.6a,b**) was synthesized in an analogous fashion to K-26 with the exception that Cbz-Ile-Tyr(Bzl)-COOH was coupled to AHEP in place of Ac-Ile-Tyr(Bzl)-COOH. Stereochemical assignments of the two des-acetyl K-26 diastereomers (**2.6a** and **2.6b**) were also established based on chemical shift values and NOESY correlations in comparison to synthetic Ile-Tyr-Tyr (**2.8a**) and Ile-Tyr-D-Tyr (**2.8b**). Of note, while the structure of K-26 has been previously unambiguously determined, this work provides the first high-resolution NMR spectroscopic data for this series of compounds.

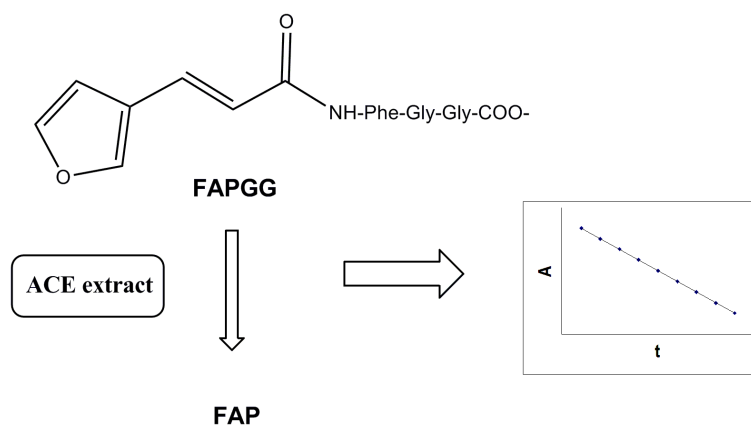


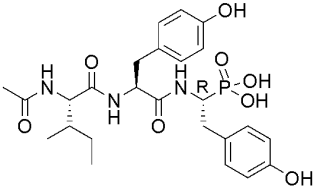
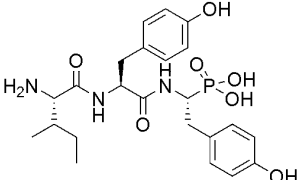
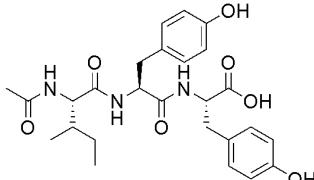
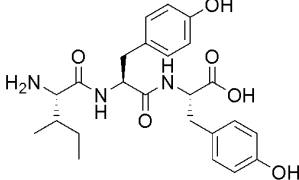
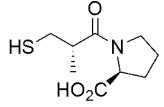
Figure 2-04. ACE assay

The method outlined by Holmquist and coworkers²² for measuring ACE activity was adapted for assay in 96-well format (see supplementary material). In short, ACE activity was extracted from rabbit lung acetone powder by soaking with 100 mM borate buffer followed by ultracentrifugation.²³ Furanacryloyl-L-phenylalanyl-glycylglycine (FAPGG) was used as the chromogenic substrate for measuring initial velocities. The progress of the reaction can be monitored in a

continuous fashion based on the hypsochromic shift of the absorption spectra that occurs upon hydrolysis of FAPGG to FAP and GG. In a typical reaction, ACE extract was preincubated with a range of inhibitor concentrations in the responsive concentration region for five minutes followed by the addition of a 1 mM solution of FAPGG. The rate of FAPGG hydrolysis was obtained by measuring the change in absorbance at 340 nm versus time. IC_{50} values were obtained by fitting triplicate measurements to a sigmoidal dose response curve. As a benchmark, the IC_{50} of Captopril was determined in parallel to ensure the accuracy of the assay. The IC_{50} values of the peptide inhibitors and Captopril are summarized in Table 2-1.

Measurement of ACE inhibition by the natural tripeptide analog of K-26, L-Ile-L-Tyr-L-Tyr (**2.8a**), confirmed the previously reported micromolar range activities of related small natural peptides. *N*-Acetylation of this compound to (**2.7a**) improved its activity by approximately ten fold. However, the potentiating effect of phosphonyl substitution in this compound class was more pronounced. K-26 (**1.17a**) was found to be 1500-fold more potent than Ac-Ile-Tyr-Tyr (**2.7a**). Indeed, all analogues containing AHEP were found to be markedly improved inhibitors relative to their carboxylate congeners.

Table 2-1. Measured IC₅₀ values of K-26 and analogues

compound	IC ₅₀ (nM)
	
1.17a , (K-26) Ac-L-Ile-L-Tyr-R-AHEP	14.4
1.17b , Ac-L-Ile-L-Tyr-S-AHEP	139
1.17c , Ac-L-Ile-L-Tyr-R-AHEP(OEt) ₂	3.83 x 10 ⁵
	
2.6a , L-Ile-L-Tyr-R-AHEP	234.3
2.6b , L-Ile-L-Tyr-S-AHEP	5.46 x 10 ⁴
	
2.7a , Ac-L-Ile-L-Tyr-L-Tyr	2.1 x 10 ⁴
2.7b , Ac-L-Ile-L-Tyr-D-Tyr	2.0 x 10 ⁶
	
2.8a , L-Ile-L-Tyr-L-Tyr	2.36 x 10 ⁵
2.8b , L-Ile-L-Tyr-D-Tyr	> 10 ⁶
	
captopril	7.7

Other general trends emerge from these data. Separation of the synthetic diastereomers permitted the evaluation of the effect of stereochemistry of AHEP

on ACE inhibition. We observed that the cost of inverting the stereochemistry from the natural product (*R*-AHEP) was an approximately ten-fold decrease in activity and similar trends for acetylated and des-acetyl analogues in all series were observed. Additionally, alkylation of the phosphonate group (ethyl esters) decreased activity over 25,000-fold, as demonstrated by the comparison of **1.17a** and **1.17c**. The influence of *N*-acetylation on ACE activity in K-26 was also consistent, with an across the board improvement of 10 - 20 fold for most acetylated analogues.

These data provide an additional basis for understanding the improvement in ACE inhibition by K-26 versus the carboxylate analogues. The ACE potentiating effects of phosphonate substitution may originate from a stronger ionic interaction of the phosphonate centered anion with the enzyme, either through side chain interactions or by direct interaction with Zn²⁺ ligand. Previous X-ray structural studies⁵ have also suggested that phosphonyl substitution may limit the conformational flexibility of the peptide backbone relative to the carboxyl analog. This notion is supported by the observation of weak long range dipolar cross-coupling between the γ -methyl of isoleucine (0.5 ppm) and the AHEP phenyl ring (6.7 ppm) in K-26 NOESY experiments. No such distal interactions were observed in Ac-Ile-Tyr-Tyr.

While the bonafide target of K-26 within the producing organism (or within the microbial ecosystem from which the producer was isolated) remains unknown, the specific functionalities that distinguish K-26 from its proteinogenic congener, *N*-acetylation and α -phosphonylation, are shown here to direct

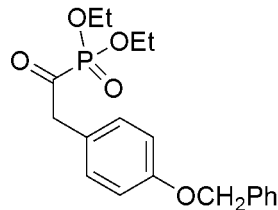
markedly improved activity against the zinc metalloprotease ACE. It is therefore possible that the natural target of K-26 is a catalytically similar metalloprotease and these results suggest that the nature's strategy of C-terminal phosphonyl substitution may find meritorious application in the discovery of new C-terminal phosphonate inhibitors. The demonstration of the potency of the phosphonate pharmacophore in this class of compounds underlines the utility of further studies aimed at identifying the biosynthetic gene(s) required for assembly of α -aminophosphonates.

Materials and Methods

Organic Synthesis

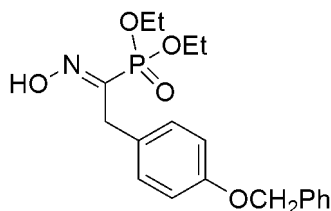
General. All reagents were commercially available and used without purification unless specified otherwise. Thin-layer chromatography was performed on precoated silica gel (60 F₂₅₄) plates obtained from EMD. Silica gel (35-70 μ m) from EM Science was used for column chromatography. All NMR spectra were acquired on a Bruker DRX-400 or a DRX-501 instrument unless otherwise noted. High resolution mass spectral data was acquired by Nonka Senova at Notre Dame University. Low resolution mass spectral data was obtained on a Thermofinnigan (San Jose, CA) TSQ® Quantum triple quadrupole mass spectrometer equipped with a standard electrospray ionization source. Amino acids and all peptide coupling reagents were purchased from EMD. 4-benzyloxyphenylacetyl chloride was obtained from TCI America.

Synthesis of diethyl (4-benzyloxy)phenyl acetyl phosphonate.^{16, 20}



4-benzyloxyphenylacetyl chloride (**2.1**, 1.0g, 3.8mmol) was dissolved in 5 mL THF and chilled to 0°C. Triethyl phosphite (0.87mL, 5mmol) was added dropwise under anhydrous conditions. When the addition was complete, the reaction was warmed to 70 °C and stirred for 15 minutes. The solvent was removed *in vacuo* and the product was not purified or characterized before using it in the next step.

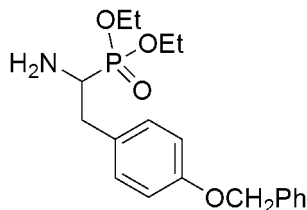
Synthesis of *O,O*-diethyl-1-oximato-2-(4-benzyloxyphenyl) ethylphosphonate (**2.2**).¹⁸



To a suspension of hydroxylammonium chloride (0.3433g, 4.94 mmol) in 0.5mL pyridine and 1mL ethanol was added 1.38g diethyl (4-benzyloxy)phenyl acetyl phosphonate dissolved in 2mL ethanol dropwise with stirring. After stirring for 12 hours, the clear solution was evaporated and the residue was quenched with 10 mL 2% HCl and extracted with dichloromethane. Organic fractions were combined, dried and concentrated to yield a yellow oil, which was purified by silica gel chromatography (3:2 ethyl acetate/hexane). The total yield for two steps was 0.717g (50%) of a clear oil. ¹H NMR (CDCl₃, 300MHz) δ 7.45-7.35 (m, 5H), 7.29-7.27 (d, *J* = 8.5Hz, 2H), 6.92-6.89 (d, *J* = 8.5Hz, 2H), 5.06 (s, 2H), 4.16-3.93

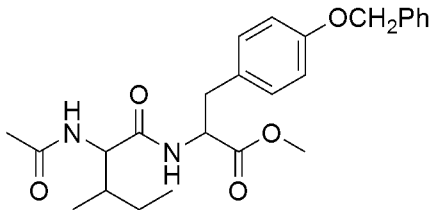
(m, 4H), 3.89-3.84 (d, $J = 14\text{Hz}$, 2H), 1.26-1.21 (t, $J = 7\text{Hz}$, 6H); ^{13}C NMR (CDCl_3 , 125MHz) δ 157.5, 153.4, 137.0, 131.0, 128.5, 127.8, 127.3, 114.7, 69.9, 62.8, 32.1, 16.0; ^{31}P NMR (CDCl_3 , 121MHz) δ 9.14; HRMS (FAB, M^+) Calcd. for $\text{C}_{19}\text{H}_{25}\text{NO}_5\text{P}$ 378.1470, found 378.1465.

*Synthesis of diethyl 1-amino-2-(4-benzyloxy phenyl)ethyl phosphonate (2.3).*¹⁷



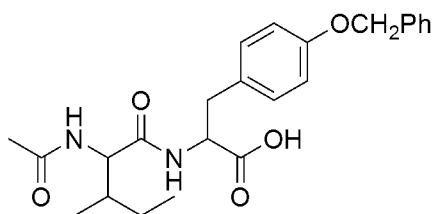
To a suspension of zinc (0.442 g, 6.76 mmol) in 4mL formic acid was added the oxime (**2.2**, 0.639 g, 1.69 mmol) under argon atmosphere. The reaction was allowed to proceed overnight after which the suspension was filtered and the filtrate evaporated to yield a clear oil. The oil was dissolved in ethyl acetate and extracted with saturated sodium bicarbonate and water. The organic fractions were dried over anhydrous MgSO_4 and concentrated *in vacuo* to yield a clear oil (0.356 g, 58%) ^1H NMR (CDCl_3 , 400MHz) δ 7.37-7.23 (m, 5H), 7.10-7.08 (d, $J = 8.5\text{Hz}$, 2H), 6.88-6.85 (d, $J = 8.5\text{Hz}$, 2H), 4.98 (s, 2H), 4.14-4.06 (m, 4H), 3.18-3.06 (m, 2H), 2.59-2.50 (m, 1H), 1.30-1.26 (t, $J = 7\text{Hz}$, 6H); ^{13}C NMR (CDCl_3 , 100MHz) δ 157.6, 137.0, 130.1, 128.5, 127.9, 127.4, 114.9, 70.0, 62.2, 51.1, 49.6, 36.9, 16.5; ^{31}P NMR (CDCl_3 , 162Hz) δ 29.36; HRMS (FAB, M^+) Calcd for $\text{C}_{19}\text{H}_{27}\text{NO}_4\text{P}$ 364.1678, found 364.1677.

*Synthesis of N-acetyl-L-isoleucyl-O-benzyl-L-tyrosine methyl ester.*²¹



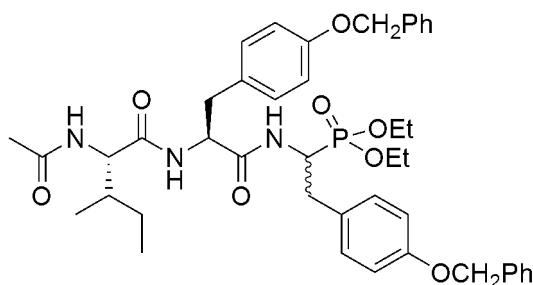
N-acetyl-L-isoleucine (112.6 mg, 0.65 mmol), *O*-benzyl-L-tyrosine methyl ester hydrochloride (200mg, 0.62mmol) and *N*-hydroxybenzotriazole (104.1 mg, 0.68 mmol) were dissolved in 6 mL anhydrous dimethylformamide and cooled to 0°C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl)carbodiimide hydrochloride (130.4 mg, 0.68 mmol) and 2,4,6-trimethylpyridine (172.3 μ L, 1.30mmol), stirred at 0 °C for one hour and at room temperature overnight. Ethyl acetate (150 mL) was added to the reaction which was extracted with 1N HCl (3 \times 30mL), 1N NaHCO₃ (3 \times 30mL) and brine (3 \times 30mL). The organic fraction was dried over MgSO₄ and evaporated to give white solid (236 mg, 86.5%). ¹H NMR (CDCl₃, 400MHz) δ 7.36-7.23 (m, 5H), 6.96-6.94 (d, *J* = 8.6Hz, 2H), 6.84-6.82 (d, *J* = 8.6Hz, 2H), 6.20-6.18 (d, *J* = 7.8Hz, 1H), 6.02-6.00 (d, *J* = 8.6Hz, 1H), 4.96 (s, 2H), 4.79-4.72 (m, 1H), 4.22-4.18 (m, 1H), 3.67 (s, 3H), 3.03-2.93 (m, 2H), 1.92 (s, 3H), 1.74-1.68 (m, 1H), 1.44-1.38 (m, 1H), 1.09-1.01 (m, 1H), 0.83-0.78 (m, 6H); ¹³C NMR (CDCl₃, 100MHz) δ 171.5, 170.7, 169.8, 157.9, 136.9, 130.2, 128.5, 127.9, 127.6, 127.4, 115.0, 69.9, 57.5, 53.1, 52.3, 37.4, 37.0, 24.9, 23.2, 15.1, 11.2. ESI-MS (M+H⁺): 441.2

*Synthesis of N-acetyl-L-isoleucyl-O-benzyl-L-tyrosine (2.4).*²⁴



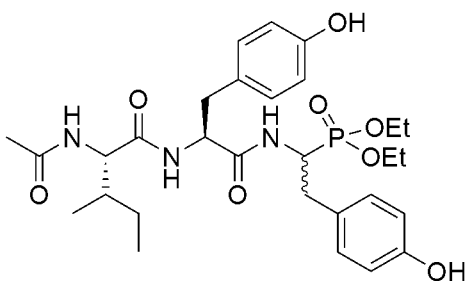
A mixture of *N*-acetyl-L-isoleucyl-O-benzyl-L-tyrosine methyl ester (190 mg, 0.43 mmol) in acetonitrile (25mL) and 3% Na₂CO₃ (38mL) was stirred overnight at room temperature. The reaction was extracted with hexane (3×50mL), acidified with 1N HCl to pH 3 and extracted with chloroform (3×60mL). The combined chloroform fractions were washed with brine (2×50mL), dried over MgSO₄ and evaporated to a white solid (147 mg, 80.2%). ¹H NMR (CDCl₃, 400MHz) δ 7.31-7.24 (m, 5H), 7.00-6.98 (d, *J* = 8.5Hz, 2H), 6.81-6.79 (d, *J* = 8.5Hz, 2H), 6.66-6.64 (d, *J* = 7.8Hz, 1H), 6.19-6.17 (d, *J* = 9.0Hz, 1H), 4.91 (s, 1H), 4.75-4.70 (q, *J* = 6.5Hz, 1H), 4.27-4.22 (t, *J* = 8, 8.6Hz, 1H), 3.09-3.04 (dd, *J* = 5.5, 14Hz, 1H), 2.89-2.83 (dd, *J* = 6.4, 14.1Hz, 1H), 1.92 (s, 3H), 1.68 (s, 1H), 1.41-1.35 (m, 1H), 1.03-0.98 (m, 1H), 0.80-0.76 (m, 6H); ¹³C NMR (DMSO, 100MHz) δ 173.2, 171.6, 169.3, 157.4, 137.6, 130.5, 130.0, 128.8, 128.1, 128.0, 127.9, 114.8, 69.5, 56.8, 54.0, 37.1, 36.2, 24.5, 22.8, 15.6, 11.3. ESI-MS (M+H⁺): 427.2

Synthesis of diethyl N-(N-acetyl-L-isoleucyl-O-benzyl - L-tyrosyl)-1-amino-2-(4-benzyloxyphenyl)ethylphosphonate.^{3, 21}

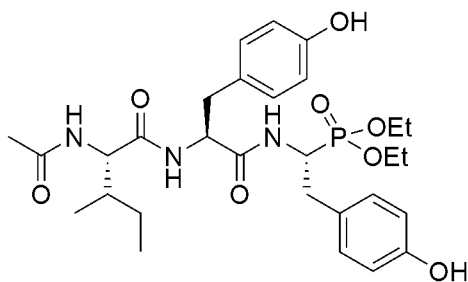


N-acetyl-L-isoleucyl-*O*-benzyl-L-tyrosine (183.4 mg, 0.43 mmol), diethyl 1-amino-2-(4-benzyloxyphenyl)ethylphosphonate (150 mg, 0.41 mmol) and *N*-hydroxybenzotriazole (68.9 mg, 0.45 mmol) were dissolved in 4 mL anhydrous dimethylformamide and cooled to 0 °C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl)carbodiimide hydrochloride (86.3 mg, 0.45 mmol) and 2,4,6-trimethylpyridine (59.6 µL, 0.45 mmol), stirred at 0 °C for one hour and at room temperature overnight. Ethyl acetate (100 mL) was added to the reaction which was extracted with 1N HCl (3×20mL), 1N NaHCO₃ (3×20mL) and brine (3×20mL). The organic fraction was dried over MgSO₄, evaporated and the sample was further purified via flash column chromatography (4% methanol in chloroform) to yield a white solid (215.2 mg, 68.0%). ¹H NMR (CDCl₃, 500MHz) δ 7.44-7.31 (m, 10H), 7.17-7.00 (m, 4H), 6.93-6.78 (m, 5H), 6.36-6.19 (m, 2H), 5.08-4.98 (m, 4H), 4.74-4.68 (m, 2H), 4.24-4.00 (m, 5H), 3.19-3.14 (m, 1H), 3.04-2.94 (m, 1H), 2.89-2.70 (m, 2H), 1.97 (s, 3H), 1.81-1.79 (m, 1H), 1.43-1.24 (m, 7H), 1.08-1.02 (m, 1H), 0.91-0.76 (m, 6H); ¹³C NMR (CDCl₃, 100MHz) δ 171.6, 171.3, 170.2, 157.5, 157.4, 137.0, 136.9, 130.4, 130.1, 130.0, 129.4, 129.3, 128.8, 128.7, 128.4, 127.8, 127.7, 127.4, 127.3, 127.2, 114.7, 114.6, 114.5, 114.4, 69.7, 69.5, 62.6, 62.4, 60.3, 57.6, 53.5, 46.1, 38.4, 38.1, 37.5, 34.8, 29.6, 25.9, 25.2, 22.8, 21.0, 16.6, 16.4, 15.3, 14.1, 11.3; ³¹P NMR (CDCl₃, 202MHz) δ 24.9, 24.7. ESI-MS (M+H⁺): 771.5

Synthesis of diethyl *N*-(*N*-acetyl-*L*-isoleucyl-*L*-tyrosyl)-1-amino-2-(4-hydroxyphenyl) ethylphosphonate (**2.5**).³

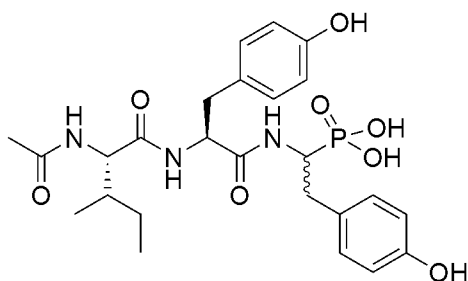


To a suspension of diethyl *N*-(*N*-acetyl-*L*-isoleucyl-*O*-benzyl-*L*-tyrosyl)-1-amino-2-(4-benzyloxyphenyl)ethyl phosphonate (215.2 mg, 0.28 mmol) in 5mL methanol, palladium on activated carbon (10%Pd, 22mg) was added and the mixture was allowed to stir overnight under a hydrogen atmosphere. The suspension was filtered through Celite and the filtrate was evaporated to yield an off-white solid (136 mg, 82.1%). The product was a mixture of two diastereomers that were separated via preparative HPLC using a linear water-acetonitrile gradient (95:5 to 60:40 H₂O:CH₃CN over 30 min) containing 0.1% trifluoroacetic acid. ¹H NMR (MeOD, 500MHz) δ 7.10-7.02 (m, 4H), 6.73-6.66 (m, 4H), 4.60-4.50 (m, 2H), 4.19-4.08 (m, 5H), 3.10-2.55 (m, 2H), 1.96 (s, 3H), 1.70-1.61 (m, 1H), 1.39-1.32 (m, 7H), 1.10-1.03 (m, 1H), 0.89-0.70 (m, 6H); ¹³C NMR (MeOD, 125MHz) δ 171.8, 171.7, 171.6, 155.8, 155.7, 129.9, 129.7, 129.5, 127.5, 127.4, 127.3, 114.8, 114.6, 114.5, 62.9, 62.7, 57.9, 54.0, 37.0, 36.1, 33.6, 24.3, 21.0, 15.3, 15.2, 14.3, 9.8; ³¹P NMR (MeOD, 202MHz) δ 25.5. ESI-MS (M+H⁺): 592.3



1c: ^1H NMR (MeOD, 500MHz) δ 7.06-7.05 (d, $J = 8.4\text{Hz}$, 2H), 7.03-7.02 (d, $J = 8.4\text{Hz}$, 2H), 6.70-6.66 (m, 4H), 4.60-4.53 (m, 2H), 4.16-4.08 (m, 5H), 3.15-3.09 (m, 1H), 2.98-2.94 (dd, $J = 5.4, 14\text{Hz}$, 1H), 2.84-2.77 (m, 1H), 2.71-2.66 (dd, $J = 8.9, 14\text{Hz}$, 1H), 1.96 (s, 3H), 1.70 (m, 1H), 1.36-1.33 (m, 7H), 1.10-1.06 (m, 1H), 0.87-0.84 (t, $J = 7.4\text{Hz}$, 3H), 0.73-0.72 (d, $J = 6.8\text{Hz}$, 3H); ^{13}C NMR (MeOD, 125MHz) δ 171.8, 171.7, 171.6, 155.8, 155.7, 129.9, 129.7, 129.5, 127.5, 127.4, 127.3, 114.8, 114.6, 114.5, 62.9, 62.7, 57.9, 54.0, 37.0, 36.1, 33.6, 24.3, 21.0, 15.3, 15.2, 14.3, 9.8; ^{31}P NMR (MeOD, 202MHz) δ 25.5. ESI-MS ($\text{M}+\text{H}^+$): 592.3

*Synthesis of N-(N-acetyl-L-isoleucyl-L-tyrosyl)-1-amino-2-(4-hydroxyphenyl)ethyl phosphonic acid (1.17, K-26).*²⁵



To a suspension of diethyl *N*-(*N*-acetyl-*L*-isoleucyl-*L*-tyrosyl)-1-amino-2-(4-hydroxyphenyl)ethylphosphonate (25.2 mg, 0.04 mmol) in 2mL acetonitrile and 0.2mL thioanisole at 0°C, iodotrimethylsilane (1 mL) was added dropwise and the reaction was allowed to proceed at 0°C for 3 hours. The solvent was removed under reduced pressure to give a brown oily residue of the phosphonate

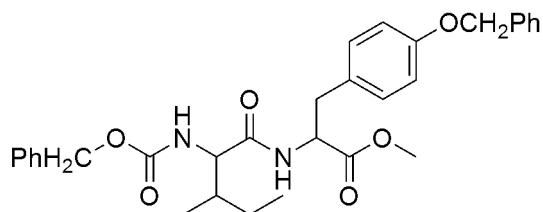
trimethylsilylester which was hydrolysed by treatment with 8:2 methanol/water (10mL). After evaporation of the solvents, the residue was dissolved in methanol and the two diastereomers of K-26 were separated via preparative HPLC using a linear water-acetonitrile gradient (95:5 to 60:40 H₂O:CH₃CN over 30min) containing 0.1% trifluoroacetic acid. Evaporation of the HPLC solvents yielded **1.17a** (3.3 mg, 15.4%) and **1.17b** (3.1 mg, 14.5%).

1.17a: ¹H NMR (D₂O, 500MHz) δ 7.06-7.05 (d, *J* = 8.4Hz, 2H), 7.01-6.99 (d, *J* = 8.4Hz, 2H), 6.70-6.69 (d, *J* = 8.4Hz, 2H), 6.69-6.68 (d, *J* = 8.4Hz, 2H), 4.52-4.49 (dd, *J* = 4.0, 10.4Hz, 1H), 4.17-4.12 (t, *J* = 12.5Hz, 1H), 3.86-3.84 (d, *J* = 8.55Hz, 1H), 3.12-3.09 (d, *J* = 14.3Hz, 1H), 2.97-2.94 (dd, *J* = 4.1, 14.1, 1H), 2.63-2.55 (m, 2H), 1.86 (s, 3H), 1.49 (s, 1H), 1.14-1.12 (m, 1H), 0.92-0.86 (m, 1H), 0.70-0.67 (t, *J* = 7.4Hz, 3H), 0.51-0.50 (d, *J* = 6.8Hz, 3H); ¹³C NMR (D₂O, 151MHz) δ 173.5, 172.3, 171.7, 154.0, 153.8, 130.6, 130.2, 129.9, 129.8, 128.4, 115.1, 58.1, 53.9, 50.3-49.4 (d, ¹*J*_{CP} = 147Hz), 36.8, 35.8, 34.6, 24.2, 21.5, 14.4, 9.9; ³¹P NMR (D₂O, 202MHz) δ 19.31. ESI-MS (M+H⁺): 536.2

1.17b: ¹H NMR (D₂O, 500MHz) δ 7.10-7.08 (d, *J* = 8.5Hz, 2H), 6.75-6.73 (d, *J* = 8.5Hz, 2H), 6.71-6.69 (d, *J* = 8.5Hz, 2H), 6.64-6.63 (d, *J* = 8.5Hz, 2H), 4.46-4.44 (m, 1H), 4.32-4.27 (m, 1H), 3.87-3.86 (d, *J* = 8Hz, 1H), 3.15-3.12 (m, 1H), 2.69-2.56 (m, 2H), 2.38-2.33 (dd, *J* = 9.5, 14Hz, 1H), 1.88 (s, 3H), 1.54 (m, 1H), 1.1-1.09 (m, 1H), 0.94-0.90 (m, 1H), 0.68-0.65 (t, *J* = 7Hz, 3H), 0.61-0.60 (d, *J* = 7Hz); ¹³C NMR (D₂O, 151MHz) δ 174.3, 173.6, 171.9, 154.3, 130.6, 130.5, 130.4, 130.2, 128.9, 115.3, 115.2, 58.7, 57.7, 53.9, 50.3-49.4 (d, ¹*J*_{CP} = 147Hz),

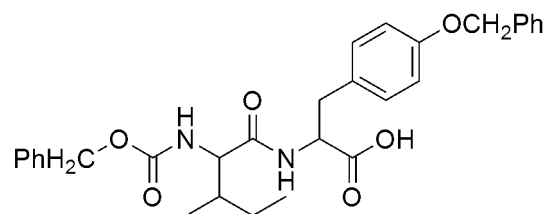
36.3, 35.6, 34.6, 25.3, 21.5, 13.5, 10.7 ; ^{31}P NMR (D_2O , 162MHz) δ 19.04. ESI-MS ($\text{M}+\text{H}^+$): 536.2

*Synthesis of N-benzyloxycarbonyl-L-isoleucyl-O-benzyl-L-tyrosine methyl ester.*²¹



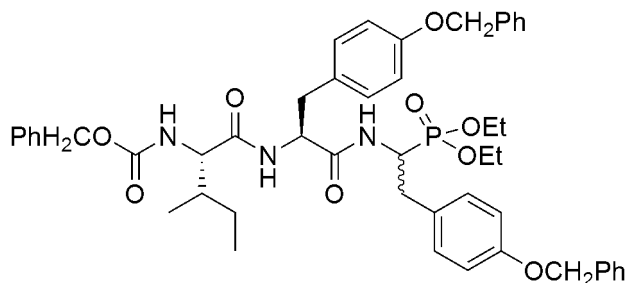
N-benzyloxycarbonyl-L-isoleucine (172.4 mg, 0.65 mmol), *O*-benzyl-L-tyrosine methyl ester hydrochloride (200mg, 0.62mmol) and *N*-hydroxybenzotriazole (104.1 mg, 0.68 mmol) were dissolved in 6 mL anhydrous dimethylformamide and cooled to 0 °C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl)carbodiimide hydrochloride (130.4 mg, 0.68 mmol) and 2,4,6-trimethylpyridine (172.3 μL , 1.30mmol), stirred at 0°C for one hour and at room temperature overnight. Ethyl acetate (150 mL) was added to the reaction which was extracted with 1N HCl (3 \times 30mL), 1N NaHCO_3 (3 \times 30mL) and brine (3 \times 30mL). The organic fraction was dried over MgSO_4 and evaporated to give white solid (296 mg, 89.6%). ^1H NMR (CDCl_3 , 300MHz) δ 7.35-7.24 (m, 10H), 5.95-6.92 (d, J = 8.4Hz, 2H), 6.82-6.80 (d, J = 8.4Hz, 2H), 6.13-6.11 (d, J = 7.8Hz, 1H), 5.21-5.19 (d, J = 8.7Hz, 1H), 5.04 (s, 2H), 4.94 (s, 2H), 4.79-4.74 (q, J = 5.8Hz, 1H), 3.96-3.92 (m, 1H), 3.65 (s, 3H), 3.05-2.91 (m, 2H), 1.79-1.73 (m, 1H), 1.36 (s, 1H), 1.06-1.02 (m, 1H), 0.84-0.79 (m, 6H); ^{13}C NMR (CDCl_3 , 100MHz) δ 171.7, 170.7, 158.0, 136.9, 136.2, 130.3, 128.6, 128.2, 128.0, 127.7, 127.5, 115, 70.0, 67.1, 59.6, 53.1, 52.3, 37.4, 37.1, 24.7, 15.3, 11.4. ESI-MS ($\text{M}+\text{H}^+$): 533.3

*Synthesis of N-benzyloxycarbonyl-L-isoleucyl-O-benzyl-L-tyrosine.*²⁴



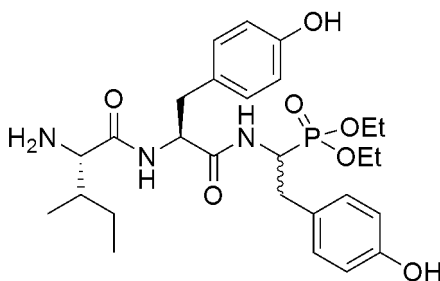
A mixture of *N*-benzyloxycarbonyl-L-isoleucyl-O-benzyl-L-tyrosine methyl ester (296 mg, 0.56 mmol) in acetonitrile (30mL) and 3% Na₂CO₃ (50mL) was stirred overnight at room temperature, then washed with hexane (3×50mL), acidified with 1N HCl to pH 3 and extracted with chloroform (3×60mL). The combined chloroform fractions were washed with brine (2×50mL), dried over MgSO₄ and evaporated to a white solid (234 mg, 81.3%). ¹H NMR (DMSO, 400MHz) δ 8.08-8.06 (d, *J* = 7.8Hz, 1H), 7.35-7.22 (m, 10H), 7.15-7.13 (d, *J* = 9.2Hz, 1H), 7.10-7.07 (d, *J* = 8.4Hz, 2H), 6.82-6.80 (d, *J* = 8.4Hz, 2H), 4.96 (s, 2H), 4.95 (s, 2H), 4.38-4.33 (q, *J* = 8.1Hz, 1H), 3.89-3.85 (t, *J* = 8.4Hz), 2.95-2.90 (m, 1H), 2.80-2.75 (m, 1H), 1.62-1.60 (m, 1H), 1.33-1.28 (m, 1H), 1.03-0.99 (m, 1H), 0.70-0.68 (m, 6H); ¹³C NMR (DMSO, 100MHz) δ 173.3, 171.6, 157.4, 156.3, 137.6, 137.4, 130.6, 130.0, 128.8, 128.7, 128.1, 128.0, 114.8, 69.5, 65.8, 59.5, 53.9, 36.9, 36.4, 24.6, 15.6, 11.2. ESI-MS (M+H⁺): 519.2

Synthesis of diethyl N-(N-benzyloxycarbonyl-L-isoleucyl-O-benzyl-L-tyrosyl)-1-amino-2-(4-benzyloxyphenyl)ethyl phosphonate.^{3, 21}



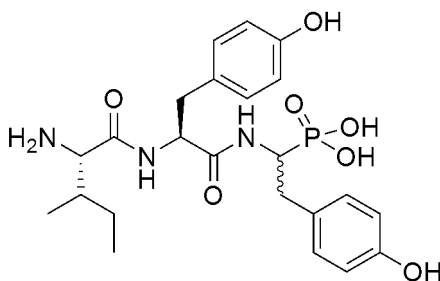
N-benzyloxycarbonyl-L-isoleucyl-*O*-benzyl-L-tyrosine (223.0 mg, 0.43 mmol), diethyl 1-amino-2-(4-benzyloxyphenyl)ethylphosphonate (150mg, 0.41mmol) and *N*-hydroxybenzotriazole (68.9 mg, 0.45 mmol) were dissolved in 4 mL anhydrous dimethylformamide and cooled to 0°C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl)carbodiimide hydrochloride (86.3 mg, 0.45 mmol) and 2,4,6-trimethylpyridine (59.6 μ L, 0.45 mmol), stirred at 0 °C for one hour and at room temperature overnight. Ethyl acetate (100 mL) was added to the reaction which was extracted with 1N HCl (3 \times 20mL), 1N NaHCO₃ (3 \times 20mL) and brine (3 \times 20mL). The organic solvent was dried over MgSO₄, evaporated and the sample was further purified via flash column chromatography (4%methanol in chloroform) to give white solid (255 mg, 72.0%). ¹H NMR (CDCl₃, 500MHz) δ 7.42-7.31 (m, 15H), 7.12-7.05 (m, 4H), 6.93-6.81 (m, 5H), 6.39-6.26 (m, 2H), 5.15-4.97 (m, 6H), 4.71-4.61 (m, 2H), 4.16-4.03 (m, 4H), 3.95-3.92 (m, 1H), 3.19-2.98 (m, 2H), 2.87-2.73 (m, 2H), 1.84 (s, 1H), 1.38-1.24 (m, 7H), 1.05-0.91 (m, 1H), 0.85-0.76 (m, 6H); ¹³C NMR (CDCl₃, 125MHz) δ 170.9, 170.1, 157.8, 157.7, 137.0, 136.9, 130.3, 130.2, 130.1, 128.6, 128.5, 128.4, 128.3, 128.2, 128.0, 127.9, 127.4, 114.9, 114.8, 114.7, 70.0, 69.9, 67.3, 67.2, 62.8, 62.7, 62.6, 60.1, 53.9, 47.1, 37.4, 36.8, 34.7, 24.5, 16.5, 16.4, 15.5, 15.4, 11.4; ³¹P NMR (CDCl₃, 202MHz) δ 24.7, 24.6. ESI-MS (M+H⁺): 864.4

*Synthesis of diethyl N-(L-isoleucyl-L-tyrosyl)-1-amino-2-(4-hydroxyphenyl)ethyl phosphonate.*³



To a suspension of diethyl *N*-(*N*-benzyloxycarbonyl-*L*-isoleucyl-*O*-benzyl-*L*-tyrosyl)-1-amino-2-(4-benzyloxyphenyl)ethyl phosphonate (250 mg, 0.29 mmol) in 5mL methanol, palladium on activated carbon (10% Pd, 25mg) was added and the mixture was allowed to stir overnight under a hydrogen atmosphere. The suspension was filtered through Celite and the filtrate was evaporated to give off-white solid (125 mg, 78.3%). The product was a mixture of two diastereomers that may be separated via preparative HPLC using a linear water-acetonitrile gradient (95:5 to 60:40 H₂O:CH₃CN over 30 min) containing 0.1% trifluoroacetic acid. ¹H NMR (D₂O, 400MHz) δ 6.98-6.93 (m, 3H), 6.69-6.59 (m, 5H), 4.48-4.24 (m, 2H), 4.02-3.90 (m, 4H), 3.60-3.46 (m, 1H), 3.02-2.96 (m, 2H), 2.38-2.36 (d, *J* = 6.7Hz, 1H), 1.70-1.57 (m, 1H), 1.22-1.10 (m, 7H), 0.86-0.76 (m, 1H), 0.74-0.47 (m, 6H) ; ¹³C NMR (D₂O, 101MHz) δ 171.6, 171.5, 168.4, 167.8, 154.4, 154.3, 154.1, 130.4, 130.3, 130.2, 130.1, 128.2, 128.1, 127.6, 127.5, 115.4, 115.3, 115.2, 64.3, 64.2, 64.1, 64.0, 63.9, 57.2, 54.0, 48.2-46.6 (d, ¹*J*_{CP} = 155Hz), 47.7-46.2 (d, ¹*J*_{CP} = 155Hz), 36.9, 36.2, 36.1, 33.1, 24.1, 23.9, 15.6, 15.5, 13.8, 13.6, 10.4, 10.1; ³¹P NMR (D₂O, 162MHz) δ 26.3, 25.9. ESI-MS (M+H⁺): 550.3

Synthesis of *N*-(*L*-isoleucyl-*L*-tyrosyl)-1-amino-2-(4-hydroxyphenyl)ethyl phosphonic acid (**2.6**).²⁵



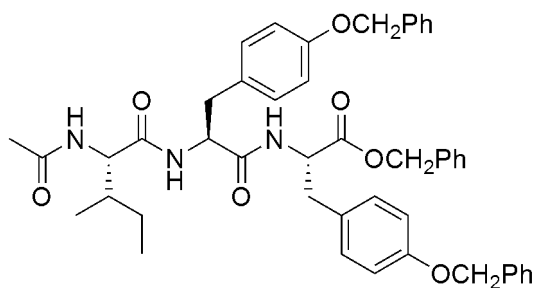
To a suspension of diethyl *N*-(*L*-isoleucyl-*L*-tyrosyl)-1-amino-2-(4-hydroxyphenyl)ethylphosphonate (10 mg, 0.018 mmol) in 2mL acetonitrile and 0.2mL thioanisole at 0°C, iodotrimethylsilane (1 mL) was added dropwise and the reaction was allowed to run at 0°C for 3 hours. The solvent was removed under reduced pressure to give a brown oily residue of the phosphonate trimethylsilylester which was hydrolysed by treatment with 8:2 methanol/water (10mL). After evaporation of the solvents, the residue was dissolved in methanol and the two diastereomers of (**2.6**) were purified via preparative HPLC using a linear water-acetonitrile gradient (ranging from 95:5 to 60:40 H₂O:CH₃CN over 30min) containing 0.1% trifluoroacetic acid. Evaporation of the HPLC solvents yielded **2.6a** (1.4 mg, 15.8%) and **2.6b** (1.3 mg, 14.6%).

2.6a: ¹H NMR (D₂O, 400MHz) δ 7.00-6.98 (d, *J* = 8.4Hz, 2H), 6.99-6.97 (d, *J* = 8.4Hz, 2H), 6.68-6.66 (d, *J* = 8.4Hz, 2H), 6.60-6.58 (d, *J* = 8.4Hz, 2H), 4.49-4.45 (m, 1H), 4.06-4.00 (t, *J* = 12.6Hz, 1H), 3.45-3.44 (d, *J* = 6.2Hz, 1H), 3.03-3.00 (d, *J* = 14Hz, 1H), 2.94-2.89 (dd, *J* = 4.1, 14.5Hz, 1H), 2.64-2.46 (m, 2H), 1.57 (m, 1H), 1.08 (m, 1H), 0.82-0.74 (m, 1H), 0.67-0.63 (t, *J* = 7.3Hz, 3H), 0.52-0.50 (d, *J* = 6.9Hz, 3H); ¹³C NMR (D₂O, 151MHz) δ 167.9, 163.0, 154.3, 153.8, 130.5,

130.3, 129.9, 128.4, 117.2, 115.3, 115.0, 57.3, 54.1, 50.6-49.6 (d, $^1J_{CP} = 147\text{Hz}$), 36.7, 36.2, 34.6, 24.0, 13.5, 10.2; ^{31}P NMR (D_2O , 162MHz) δ 19.2. ESI-MS ($\text{M}+\text{H}^+$): 494.2

2.6b: ^1H NMR (D_2O , 500MHz) δ 7.11-7.09 (d, $J = 8.5\text{Hz}$, 2H), 6.75-6.74 (d, $J = 8.5\text{Hz}$, 4H), 6.70-6.68 (d, $J = 8.5\text{Hz}$, 2H), 4.49-4.47 (m, 1H), 4.24-4.18 (m, 1H), 3.65-3.64 (d, $J = 5.6\text{Hz}$, 1H), 3.15-3.12 (m, 1H), 2.65-2.58 (m, 1H), 2.51-2.38 (m, 2H), 1.78 (m, 1H), 1.33-1.29 (m, 1H), 1.06-1.00 (m, 1H), 0.83-0.81 (d, $J = 7\text{Hz}$, 3H), 0.79-0.76 (t, $J = 7.4\text{Hz}$, 3H); ^{13}C NMR (D_2O , 151MHz) δ 171.6, 168.8, 154.3, 154.0, 130.4, 130.2, 127.9, 116.9, 115.3, 57.5, 54.5, 50.2-49.2 (d, $^1J_{CP} = 146\text{Hz}$), 37.8, 36.1, 34.6, 23.7, 13.9, 10.4; ^{31}P NMR (D_2O , 162MHz) δ 18.9. ESI-MS ($\text{M}+\text{H}^+$): 494.2

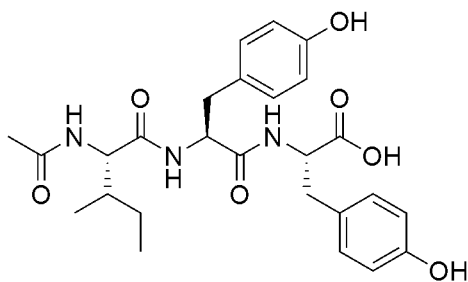
*Synthesis of N-acetyl-L-isoleucyl-O-benzyl-L-tyrosyl-O-benzyl-L-tyrosine benzyl ester.*²¹



N-acetyl-L-isoleucyl-O-benzyl-L-tyrosine (145 mg, 0.34 mmol), *O*-benzyl-L-tyrosine benzyl ester hydrochloride (127.3mg, 0.32mmol) and *N*-hydroxybenzotriazole (55.1 mg, 0.36 mmol) were dissolved in 3 mL anhydrous dimethylformamide and cooled to 0°C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl)carbodiimide hydrochloride (69.0 mg, 0.36 mmol) and 2,4,6-trimethylpyridine (88.8 μL , 0.67 mmol), stirred at 0°C for one hour and at

room temperature overnight. Ethyl acetate (75 mL) was added to the reaction and was extracted with 1N HCl (3×15mL), 1N NaHCO₃ (3×15mL) and brine (3×15mL). The organic solvent was dried over MgSO₄ and evaporated to give white solid (203 mg, 82.3%). ¹H NMR (CDCl₃, 400MHz) δ 7.32-7.21 (m, 15H), 7.03-6.99 (d, *J* = 8.4Hz, 2H), 6.81-6.79 (d, *J* = 8.4Hz, 2H), 6.74-6.69 (m, 4H), 6.39-6.37 (d, *J* = 7.6Hz, 1H), 6.10-6.08 (d, *J* = 7.2Hz, 1H), 5.88-5.86 (d, *J* = 8.4Hz, 1H), 5.06-5.00 (m, 2H), 4.94 (s, 2H), 4.92 (s, 2H), 4.70-4.66 (m, 1H), 4.57-4.50 (m, 1H), 4.16-4.12 (m, 1H), 2.98-2.82 (m, 4H), 1.90 (s, 3H), 1.73-1.68 (m, 1H), 1.39-1.33 (m, 1H), 1.04-0.96 (m, 1H), 0.81-0.75 (m, 6H); ¹³C NMR (CDCl₃, 101MHz) δ 170.8, 170.6, 169.9, 169.8, 157.9, 136.8, 135.0, 130.4, 130.2, 128.5, 128.4, 128.3, 127.9, 127.5, 127.4, 115.0, 114.8, 69.9, 67.1, 57.8, 54.2, 53.4, 37.3, 37.0, 36.9, 24.9, 23.1, 15.3, 11.2. ESI-MS (M+H⁺): 770.4

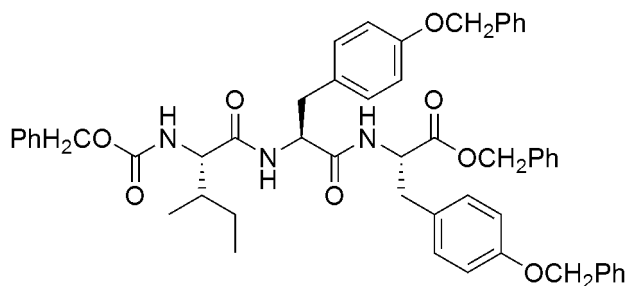
Synthesis of N-acetyl-L-isoleucyl-L-tyrosyl-L-tyrosine (2.7a).



To a suspension of *N*-acetyl-L-isoleucyl-O-benzyl-L-tyrosyl-O-benzyl-L-tyrosine benzyl ester (200 mg, 0.26 mmol) in 5mL methanol, palladium on activated carbon (10%Pd, 20mg) was added and the mixture was allowed to stir overnight under a hydrogen atmosphere. The suspension was filtered through Celite and the filtrate was evaporated to give off-white solid. The product was purified via preparative HPLC using a linear water-acetonitrile gradient (95:5 to 60:40

H₂O:CH₃CN over 30min) containing 0.1% trifluoroacetic acid. Evaporation of the HPLC solvents yielded off-white solid (106 mg, 86.8%). ¹H NMR (D₂O, 500MHz) δ 7.01-6.99 (m, 4H), 6.74-6.70 (m, 4H), 4.53-4.50 (m, 1H), 4.43-4.40 (m, 1H), 3.91-3.90 (d, *J* = 8.3Hz, 1H), 3.04-3.00 (dd, *J* = 5, 14Hz, 1H), 2.97-2.93 (dd, *J* = 5.8, 14Hz, 1H), 2.82-2.77 (dd, *J* = 8.6, 14Hz, 1H), 2.74-2.69 (dd, *J* = 9.2, 14Hz, 1H), 1.22(s, 3H), 1.56-1.55 (m, 1H), 1.22-1.17 (m, 1H), 0.98-0.92 (m, 1H), 0.72-0.69 (t, *J* = 7.4Hz, 3H), 0.58-0.57 (d, *J* = 6.7Hz, 3H); ¹³C NMR (MeOD, 101MHz) δ 176.0, 171.9, 171.8, 171.5, 155.7, 129.9, 129.8, 127.5, 114.7, 114.6, 58.0, 54.3, 36.7, 36.2, 24.3, 21.0, 14.4, 9.8. ESI-MS (M+H⁺): 500.2

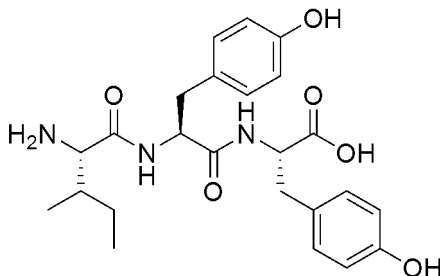
*Synthesis of N-benzyloxycarbonyl-L-isoleucyl-O-benzyl-L-tyrosyl-O-benzyl-L-tyrosine benzyl ester.*²¹



N-benzyloxycarbonyl-L-isoleucyl-O-benzyl-L-tyrosine (234 mg, 0.45 mmol), *O*-benzyl-L-tyrosine benzyl ester hydrochloride (171 mg, 0.43 mmol) and *N*-hydroxybenzotriazole (72 mg, 0.47 mmol) were dissolved in 4 mL anhydrous dimethylformamide and cooled to 0°C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl)carbodiimide hydrochloride (90 mg, 0.47 mmol) and 2,4,6-trimethylpyridine (119.7 μL, 0.90 mmol), stirred at 0°C for one hour and at room temperature overnight. Ethyl acetate (100 mL) was added to the reaction and was extracted with 1N HCl (3×20mL), 1N NaHCO₃ (3×20mL) and brine

(3×20mL). The organic solvent was dried over MgSO₄ and evaporated to give white solid (122.7 mg, 82.8%). ¹H NMR (CDCl₃, 500MHz) δ 7.47-7.29 (m, 4H), 7.11-7.10 (d, *J* = 8.4Hz, 2H), 6.88-6.87 (d, *J* = 8.4Hz, 2H), 6.81-6.80 (m, 4H), 6.59-6.58 (d, *J* = 7.4Hz, 1H), 6.32-6.31 (d, *J* = 6.8Hz, 1H), 5.33-5.32 (d, *J* = 8.1Hz, 1H), 5.17-5.00 (m, 8H), 4.78-4.77 (m, 1H), 4.66-4.65 (m, 1H), 4.06-4.03 (m, 1H), 3.05-2.93 (m, 4H), 1.87-1.85 (m, 1H), 1.31-1.28 (m, 1H), 1.09-1.03 (m, 1H), 0.88-0.85 (m, 6H); ¹³C NMR (CDCl₃, 125MHz) δ 170.9, 170.7, 170.0, 157.9, 136.9, 136.1, 135.0, 130.4, 130.2, 128.5, 128.2, 128.0, 127.9, 127.4, 114.9, 114.8, 69.9, 67.1, 59.5, 54.2, 53.5, 37.3, 37.1, 36.9, 24.8, 15.4, 11.3. ESI-MS (M+H⁺): 862.4

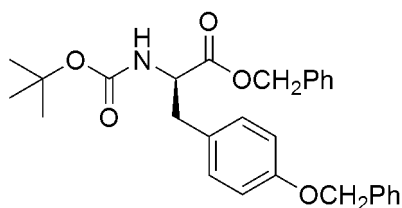
Synthesis of L-isoleucyl-L-tyrosyl-L-tyrosine (2.8a).



To a suspension of *N*-benzyloxycarbonyl-L-isoleucyl-O-benzyl-L-tyrosyl-O-benzyl-L-tyrosine benzyl ester (100 mg, 0.12 mmol) in 4mL methanol, palladium on activated carbon (10%Pd, 10mg) was added and the mixture was allowed to stir overnight under a hydrogen atmosphere. The suspension was filtered through Celite and the filtrate was evaporated to give off-white solid. The product was purified via preparative HPLC using a linear water-acetonitrile gradient (95:5 to 60:40 H₂O:CH₃CN over 30min) containing 0.1% trifluoroacetic acid. Evaporation of the HPLC solvents yielded off-white solid (48.6 mg, 88.6%). ¹H

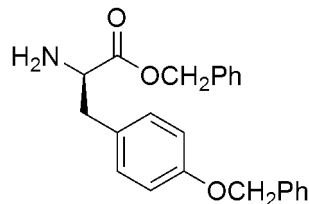
NMR (D₂O, 500MHz) δ 7.03-7.02 (d, J = 8.5Hz, 2H), 7.01-6.99 (d, J = 8.5Hz, 2H), 6.75-6.73 (d, J = 8.5Hz, 2H), 6.71-6.69 (d, J = 8.5Hz, 2H), 4.52-4.49 (t, J = 7.6Hz, 1H), 4.39-4.37 (m, 1H), 3.65-3.63 (d, J = 6.1Hz, 1H), 3.02-2.70 (m, 4H), 1.76-1.71 (m, 1H), 1.28-1.23 (m, 1H), 1.00-0.94 (m, 1H), 0.78-0.75 (t, J = 7.4Hz, 3H), 0.70-0.69 (d, J = 6.9Hz, 3H); ¹³C NMR (MeOD, 101MHz) δ 171.2, 167.8, 155.9, 155.7, 129.9, 129.8, 127.6, 127.3, 114.7, 114.6, 57.4, 54.9, 54.4, 36.7, 36.3, 23.6, 13.8, 10.2. ESI-MS (M+H⁺): 458.2

*Synthesis of N-(tert-butyloxycarbonyl)-O-benzyl-D-tyrosine benzyl ester.*²⁶



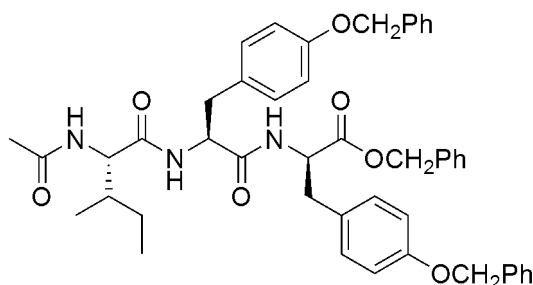
N-(*tert*-butyloxycarbonyl)-*D*-tyrosine (1.00 g, 3.55 mmol), potassium carbonate (1.2273 g, 8.88 mmol), benzyl bromide (1.05 mL, 8.88 mmol) and tetrabutylammonium iodide (162.5 mg, 0.44 mmol) were suspended in 15 mL dimethylformamide. The reaction was stirred at room temperature overnight and partitioned into ether and water. The aqueous layer was extracted with ether, and the combined ether extracts were washed with 1N HCl and brine, dried over MgSO₄, and concentrated *in vacuo* to give white solid (1.315 g, 80.3%). ¹H NMR (CDCl₃, 500MHz) δ 7.46-7.32 (m, 10H), 6.98-6.97 (d, J = 8.4Hz, 2H), 6.87-6.86 (d, J = 8.4Hz, 2H), 5.22-5.12 (m, 2H), 5.05 (s, 2H), 5.01-4.99 (m, 1H), 4.63-4.61 (m, 1H), 3.06-3.05 (m, 2H), 1.45 (s, 9H); ¹³C NMR (CDCl₃, 151MHz) δ 171.8, 157.8, 155.1, 137.0, 135.2, 130.3, 128.6, 128.4, 128.0, 127.9, 127.4, 114.8, 79.8, 69.9, 67.0, 54.5, 37.4, 28.3. ESI-MS (M+H⁺): 462.2

*Synthesis of O-benzyl-D-tyrosine benzyl ester.*²⁷



N-(*tert*-butyloxycarbonyl)-*O*-benzyl-D-tyrosine benzyl ester (1.315 g, 2.85 mmol) was stirred in a solution of trifluoroacetic acid (25 mL) and anisole (2.5 mL) at 0°C for 30 min. Trifluoroacetic acid was removed by evaporation and the resulting white solid was dissolved in ethyl acetate and extracted with saturated sodium bicarbonate and water. The organic layer was dried over MgSO₄ and concentrated *in vacuo* to white solid (0.519 g, 50.4%). ¹H NMR (CDCl₃, 500MHz) δ 7.46-7.33 (m, 10H), 7.09-7.07 (d, J = 8.4Hz, 2H), 6.91-6.89 (d, J=8.4Hz, 2H), 5.19-5.14 (m, 2H), 5.06 (s, 2H), 3.78-3.75 (m, 1H), 3.06-3.03 (m, 1H), 2.91-2.86 (m, 1H), 1.56 (s, 2H); ¹³C NMR (CDCl₃, 125MHz) δ 174.9, 157.7, 137.0, 135.5, 130.3, 129.2, 128.5, 128.4, 128.3, 127.9, 127.4, 114.8, 69.9, 66.6, 55.9, 40.1. ESI-MS (M+H⁺): 362.2

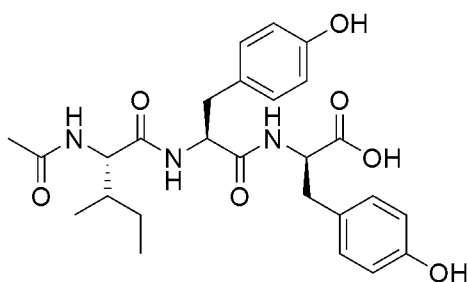
*Synthesis of N-acetyl-L-isoleucyl-O-benzyl-L-tyrosyl-O-benzyl-D-tyrosine benzyl ester.*²¹



N-acetyl-L-isoleucyl-*O*-benzyl-L-tyrosine (123.7 mg, 0.29 mmol), *O*-benzyl-D-tyrosine benzyl ester (100 mg, 0.28 mmol) and *N*-hydroxybenzotriazole (45.9 mg,

0.30 mmol) were dissolved in 3 mL anhydrous dimethylformamide and cooled to 0°C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl) carbodiimide hydrochloride (57.5 mg, 0.30 mmol) and 2,4,6-trimethylpyridine (39.8 μ L, 0.30 mmol), stirred at 0°C for one hour and at room temperature overnight. Ethyl acetate (75 mL) was added to the reaction and was extracted with 1N HCl (3 \times 15mL), 1N NaHCO₃ (3 \times 15mL) and brine (3 \times 15mL). The organic solvent was dried over MgSO₄ and evaporated to give white solid (167 mg, 79.8%). ¹H NMR (DMSO, 400MHz) δ 8.42-8.40 (m, 1H), 7.80-7.78 (m, 2H), 7.33-7.16 (m, 15H), 7.04-7.02 (d, 8.4Hz, 2H), 6.91-6.89 (d, 8.4Hz, 2H), 6.81-6.78 (d, 8.4Hz, 2H), 6.74-6.72 (d, 8.4Hz, 2H), 5.04-4.90 (m, 6H), 4.44-4.34 (m, 2H), 4.03-3.99 (m, 1H), 2.92-2.46 (m, 4H), 1.74 (s, 3H), 1.54-1.52 (m, 1H), 1.19-1.16 (m, 1H), 0.93-0.86 (m, 1H), 0.66-0.56 (m, 6H); ¹³C NMR (CDCl₃, 125MHz) δ 171.6, 171.4, 171.2, 169.6, 157.6, 157.3, 137.6, 137.5, 136.1, 130.6, 128.9, 128.8, 128.4, 128.2, 128.1, 127.9, 114.9, 114.6, 69.5, 66.4, 57.5, 54.3, 54.0, 36.7, 24.6, 22.9, 15.7, 11.3. ESI-MS (M+H⁺): 770.4

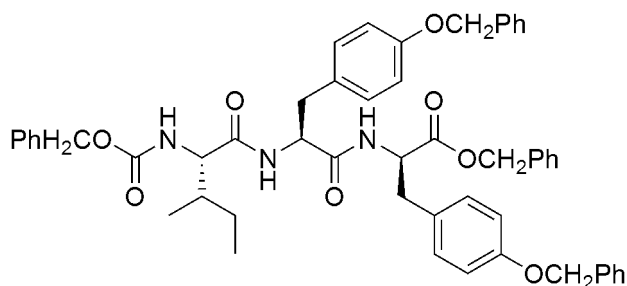
Synthesis of N-acetyl-L-isoleucyl-L-tyrosyl-D-tyrosine (2.7b).



To a suspension of *N*-acetyl-L-isoleucyl-O-benzyl-L-tyrosyl-O-benzyl-D-tyrosine benzyl ester (150 mg, 0.19 mmol) in 5mL methanol, palladium on activated carbon (10%Pd, 15mg) was added and the mixture was allowed to stir overnight

under a hydrogen atmosphere. The suspension was filtered through Celite and the filtrate was evaporated to give off-white solid. The product was purified via preparative HPLC using a linear water-acetonitrile gradient (95:5 to 60:40 H₂O:CH₃CN over 30min) containing 0.1% trifluoroacetic acid. Evaporation of the HPLC solvents yielded off-white solid (72.8 mg, 81.3%). ¹H NMR (D₂O, 500MHz) δ 6.85-6.83 (m, 4H), 6.69-6.67 (m, 4H), 4.52-4.50 (m, 1H), 4.44-4.41 (m, 1H), 4.03-3.99 (m, 1H), 2.92-2.46 (m, 4H), 1.81 (s, 3H), 1.78-1.76 (m, 1H), 1.30-1.25 (m, 1H), 1.03-0.97 (m, 1H), 0.81-0.72 (m, 6H); ¹³C NMR (D₂O, 125MHz) δ 174.3, 171.8, 170.9, 168.6, 162.8, 162.5, 154.4, 154.3, 130.3, 127.8, 127.5, 117.3, 115.3, 115.0, 57.4, 54.7, 53.8, 36.3, 36.1, 35.6, 23.8, 22.9, 13.8, 10.3. ESI-MS (M+H⁺): 500.2

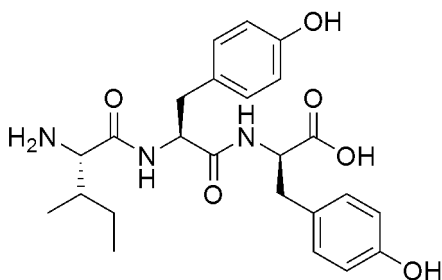
*Synthesis of N-benzyloxycarbonyl-L-iso-leucyl- O-benzyl- L-tyrosyl- O-benzyl- D-tyrosine benzyl ester.*²¹



N-benzyloxycarbonyl-L-iso-leucyl-*O*-benzyl-L-tyrosine (150.4 mg, 0.29 mmol), *O*-benzyl-D-tyrosine benzyl ester (100 mg, 0.28 mmol) and *N*-hydroxybenzotriazole (45.9 mg, 0.30 mmol) were dissolved in 3 mL anhydrous dimethylformamide and cooled to 0°C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl) carbodiimide hydrochloride (57.5 mg, 0.30 mmol) and 2,4,6-trimethylpyridine (39.8 μL, 0.30 mmol), stirred at 0°C for one hour and at room temperature

overnight. Ethyl acetate (75 mL) was added to the reaction and was extracted with 1N HCl (3×15mL), 1N NaHCO₃ (3×15mL) and brine (3×15mL). The organic solvent was dried over MgSO₄ and evaporated to give white solid (101.6 mg, 79.3%). ¹H NMR (DMSO, 400MHz) δ 8.46-8.44 (d, J = 7.9Hz, 1H), 7.74-7.72 (d, 8.7Hz, 1H), 7.31-7.16 (m, 20H), 7.04-7.02 (d, 8.4Hz, 2H), 6.90-6.88 (d, 8.4Hz, 2H), 6.82-6.79 (d, 8.4Hz, 2H), 6.73-6.71 (d, 8.4Hz, 2H), 5.01-4.91 (m, 8H), 4.49-4.37 (m, 2H), 3.78-3.74 (m, 1H), 2.93-2.47 (m, 4H), 1.51 (m, 1H), 1.16-1.15 (m, 1H), 0.94-0.87 (m, 1H), 0.65-0.62 (t, 7.4Hz, 3H), 0.56-0.54 (d, 6.8Hz, 3H); ¹³C NMR (DMSO, 101MHz) δ 171.7, 171.4, 171.1, 157.6, 157.3, 156.3, 137.6, 137.5, 136.1, 130.1, 130.0, 129.5, 128.8, 128.7, 128.4, 128.2, 128.1, 128.0, 127.9, 114.9, 114.6, 109.3, 106.9, 81.1, 69.5, 66.4, 65.8, 59.8, 54.2, 53.8, 37.5, 36.8, 36.5, 24.5, 15.6, 11.2. ESI-MS (M+H⁺): 862.4

Synthesis of L-isoleucyl-L-tyrosyl-D-tyrosine (2.8b).



To a suspension of *N*-benzyloxycarbonyl-L-isoleucyl-O-benzyl-L-tyrosyl-O-benzyl-D-tyrosine benzyl ester (100 mg, 0.12 mmol) in 4mL methanol, palladium on activated carbon (10%Pd, 10mg) was added and the mixture was allowed to stir overnight under a hydrogen atmosphere. The suspension was filtered through Celite and the filtrate was evaporated to give off-white solid. The product was purified via preparative HPLC using a linear water-acetonitrile gradient (95:5

to 60:40 H₂O:CH₃CN over 30min) containing 0.1% trifluoroacetic acid. Evaporation of the HPLC solvents yielded off-white solid (45.5 mg, 85.8%). ¹H NMR (D₂O, 500MHz) δ 6.85-6.83 (m, 4H), 6.69-6.67 (m, 4H), 4.52-4.50 (m, 1H), 4.44-4.41 (m, 1H), 3.71-3.70 (d, 5.5Hz, 1H), 2.91-2.87 (m, 1H), 2.72-2.68 (m, 3H), 1.81-1.76 (m, 1H), 1.30-1.25 (m, 1H), 1.03-0.97 (m, 1H), 0.81-0.75 (t, 7Hz, 3H), 0.74-0.72 (d, 7.4Hz, 3H); ¹³C NMR (D₂O, 125MHz) δ 174.3, 171.8, 168.6, 162.8, 162.5, 154.4, 154.3, 130.3, 127.8, 127.5, 117.3, 115.3, 115.0, 57.4, 54.7, 53.8, 36.3, 36.1, 35.6, 23.8, 13.8, 10.3. ESI-MS (M+H⁺): 458.2

ACE Inhibition

General. Captopril, FAPGG and rabbit lung acetone powder were purchased from Sigma. FAPGG reagent refers to a 1mM solution of FAPGG in 50mM Tris, 0.6M NaCl, pH 8.0. ACE extract was prepared according to FitzGerald and coworkers.²³ Specifically, 1g of rabbit lung acetone powder was suspended in 10mL 100mM sodium borate buffer containing 5%(v/v) glycerol at pH 8.3 with stirring at 4°C overnight. The suspension was then ultracentrifuged at 40,000×g for 30min. The supernatant exhibited ACE activity and was stored at 4°C for up to a week without observable loss of activity.

ACE inhibitor solutions. ACE inhibitory compounds were dissolved in 20mM Tris, pH 8 and serial dilutions were prepared to obtain solutions of concentrations ranging from millimolar to picomolar. DMSO (10%v/v) was added to all solution to facilitate dissolution of the compounds. The millimolar stock solutions were standardized by NMR. In detail, to 400μL of the stock solution, 100μL D₂O and 0.5μL dioxane were added. A standard proton NMR experiment

with an attenuated delay time ($d_1=15s$) was performed and from the integral values of the dioxane singlet and the inhibitor peaks we were able to calculate the accurate concentration of the inhibitor in the stock solution.

*ACE assay.*²² To each well of a 96-well plate were added 60 μ L ACE extract and 60 μ L inhibitor solution. After a five minute pre-incubation at room temperature, 100 μ L of FAPGG reagent was added. 20mM Tris pH 8 buffer was used for the background rate. Absorbance readings were made with a Spectramax UV/Vis spectrometer at 340 nm every 8 seconds over a period of 9 minutes. The experiment was performed in triplicate and the data was fitted on a sigmoidal curve using the Graphpad Prism 4 software. IC_{50} values were obtained from the equation of the sigmoidal curves illustrated in Appendix B.

References

1. Quin, L. D., *Organophosphorus Chemistry*. 1st ed.; Wiley-Interscience: New York, 2000.
2. Seto, H.; Kuzuyama, T., Bioactive natural products with carbon-phosphorus bonds and their biosynthesis. *Nat Prod Rep* **1999**, 16, (5), 589-96.
3. Kase, H.; Yamato, M.; Koguchi, T.; Okachi, R.; Kasai, M.; Shirahata, K.; Kawamoto, I.; Shuto, K.; Karasawa, A. Phosphorus-containing oligopeptides and a pharmaceutical composition containing them. EP 61172, 1982.
4. Yamato, M.; Koguchi, T.; Okachi, R.; Yamada, K.; Nakayama, K.; Kase, H.; Karasawa, A.; Shuto, K., K-26, a Novel Inhibitor of Angiotensin-I Converting Enzyme Produced by an Actinomycete K-26. *Journal of Antibiotics* **1986**, 39, (1), 44-52.
5. Hirayama, N.; Kasai, M.; Shirahata, K., Structure and Conformation of a Novel Inhibitor of Angiotensin-I Converting Enzyme - a Tripeptide Containing Phosphonic Acid. *International Journal of Peptide and Protein Research* **1991**, 38, (1), 20-24.
6. Ohuchio, S.; Kurihara, K.; Shinohara, A.; Takei, T.; Yoshida, J.; Amano, S.; Miyadoh, S.; Matsushida, Y.; Somura, T.; Sezaki, M., Studies on New Angiotensin Converting Enzyme Inhibitors, SF2513A,B and C, Produced by *Streptosporangium nondiastaticum*. *Sci. Reports of Meiji Seika Kaisha* **1988**, 27, 46 - 54.
7. Koguchi, T.; Yamada, K.; Yamato, M.; Okachi, R.; Nakayama, K.; Kase, H., K-4, a Novel Inhibitor of Angiotensin-I Converting Enzyme Produced by *Actinomadura-Spiculospora*. *Journal of Antibiotics* **1986**, 39, (3), 364-371.
8. Ando, T.; Okada, S.; Uchida, I.; Hemmi, K.; Nishikawa, M.; Tsurumi, Y.; Fujie, A.; Yoshida, K.; Okuhara, M., WF-10129, a novel angiotensin converting enzyme inhibitor produced by a fungus, *Doratomyces putredinis*. *J Antibiot (Tokyo)* **1987**, 40, (4), 468-75.
9. Fukuhara, K.-i.; Murao, S.; Nozawa, T.; Hatano, M., Structural elucidation of talopeptin (MK-I), a novel metallo proteinase inhibitor produced by *Streptomyces mozunensis* MK-23. *Tetrahedron Letters* **1982**, 23, (22), 2319-2322.

10. Kase, H.; Kaneko, M.; Yamada, K., K-13, a novel inhibitor of angiotensin I converting enzyme produced by *Micromonospora halophytica* subsp. *exilis*. I. Fermentation, isolation and biological properties. *J Antibiot (Tokyo)* **1987**, 40, (4), 450-4.
11. Bush, K.; Henry, P. R.; Slusarchyk, D. S., Muraceins--muramyl peptides produced by *Nocardia orientalis* as angiotensin-converting enzyme inhibitors. I. Taxonomy, fermentation and biological properties. *J Antibiot (Tokyo)* **1984**, 37, (4), 330-5.
12. Hunt, A. H.; Mynderse, J. S.; Samlaska, S. K.; Fukuda, D. S.; Maciak, G. M.; Kirst, H. A.; Occolowitz, J. L.; Swartzendruber, J. K.; Jones, N. D., Structure elucidation of A58365A and A58365B, angiotensin converting enzyme inhibitors produced by *Streptomyces chromofuscus*. *J Antibiot (Tokyo)* **1988**, 41, (6), 771-9.
13. Riviere, G.; Michaud, A.; Corradi, H.; Sturrock, E.; Acharya, K.; Cogez, V.; Bohin, J.; Vieau, D.; Corvol, P., Characterization of the first angiotensin-converting like enzyme in bacteria: Ancestor ACE is already active. *Gene* **2007**, 399, (1), 81-90.
14. Acharya, K. R.; Sturrock, E. D.; Riordan, J. F.; Ehlers, M. R., Ace revisited: a new target for structure-based drug design. *Nat Rev Drug Discov* **2003**, 2, (11), 891-902.
15. Wu, J.; Aluko, R. E.; Nakai, S., Structural requirements of Angiotensin I-converting enzyme inhibitory peptides: quantitative structure-activity relationship study of di- and tripeptides. *J Agric Food Chem* **2006**, 54, (3), 732-8.
16. Burk, M. J.; Stammers, T. A.; Straub, J. A., Enantioselective Synthesis of α -Hydroxy and α -Amino Phosphonates via Catalytic Asymmetric Hydrogenation. *Organic Letters* **1999**, 1, (3), 387-390.
17. Kowalik, J.; Kupczyk-Subotkowska, L.; Mastalerz, P., Preparation of dialkyl 1-aminoalkane phosphonates by reduction of dialkyl 1-hydroxy imino alkane phosphonates with zinc in formic acid. *Synthesis* **1981**, 57-8.
18. Maier, L.; Diel, P. J., Organic Phosphorus-Compounds .97. Synthesis and Properties of 1-Amino-2-Aryl- and 2-Pyridyl-Ethylphosphonic Acids and

Derivatives. *Phosphorus Sulfur and Silicon and the Related Elements* **1991**, 62, (1-4), 15-27.

19. Ntai, I.; Phelan, V. V.; Bachmann, B. O., Phosphonopeptide K-26 biosynthetic intermediates in *Astrosporangium hypotensionis*. *Chemical Communications* **2006**, (43), 4518-4520.

20. Tam, C. C.; Mattocks, K. L.; Tishler, M., Enol-keto tautomerism of α -ketophosphonates. *Proc Natl Acad Sci USA* **1981**, 78, (6), 3301-3304.

21. Carpino, L. A.; El-Faham, A.; Albericio, F., Efficiency in Peptide Coupling: 1-Hydroxy-7-azabenzotriazole vs 3,4-Dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine. *Journal of Organic Chemistry* **1995**, 60, (11), 3561-4.

22. Holmquist, B.; Bunning, P.; Riordan, J. F., Continuous Spectrophotometric Assay for Angiotensin Converting Enzyme. *Analytical Biochemistry* **1979**, 95, (2), 540-548.

23. Murray, B. A.; Walsh, D. J.; FitzGerald, R. J., Modification of the furanacryloyl-L-phenylalanyl-glycylglycine assay for determination of angiotensin-I-converting enzyme inhibitory activity. *Journal of Biochemical and Biophysical Methods* **2004**, 59, (2), 127-137.

24. Adamson, J. G.; Blaskovich, M. A.; Groenevelt, H.; Lajoie, G. A., Simple and convenient synthesis of tert-butyl ethers of Fmoc-serine, Fmoc-threonine, and Fmoc-tyrosine. *Journal of Organic Chemistry* **1991**, 56, (10), 3447-9.

25. Oishi, S.; Karki, R. G.; Kang, S.-U.; Wang, X.; Worthy, K. M.; Bindu, L. K.; Nicklaus, M. C.; Fisher, R. J.; Burke, T. R., Jr., Design and Synthesis of Conformationally Constrained Grb2 SH2 Domain Binding Peptides Employing α -Methylphenylalanyl Based Phosphotyrosyl Mimetics. *Journal of Medicinal Chemistry* **2005**, 48, (3), 764-772.

26. Chen, C.; Zhu, Y.-F.; Wilcoxon, K., An improved synthesis of selectively protected L-dopa derivatives from L-tyrosine. *Journal of Organic Chemistry* **2000**, 65, (8), 2574-2576.

27. Futaki, S.; Yagami, T.; Taike, T.; Akita, T.; Kitagawa, K., Sulfur trioxide/thiol: a novel system for the reduction of methionine sulfoxide. *Journal of*

the Chemical Society, Perkin Transactions 1: Organic and Bio-Organic Chemistry **1990**, (3), 653-8.

CHAPTER III

BIOSYNTHETIC INTERMEDIATES OF AHEP AND K-26

Introduction

The angiotensin converting enzyme inhibitor K-26 (**1.17**) is a naturally occurring *N*-acetylated tripeptide containing isoleucine, tyrosine and the nonproteinogenic amino acid, (*R*)-1-amino-2-(4-hydroxyphenyl)ethylphosphonic acid (AHEP, **1.18**).^{1, 2} The terminal phosphonate tyrosine analogue is reportedly present in a small group of tripeptide natural products produced by other members of the *Streptosporangiaceae* family.^{3, 4} This group of natural products possesses potent *in vitro* angiotensin converting enzyme inhibitory activities, which have been demonstrated to translate to effective hypotensive activity *in vivo*.^{1, 5} Despite the potent hypotensive activity of K-26 and related compounds, the biosynthetic pathways by which K-26 and more specifically AHEP are biosynthesized still remain uncharacterized.

In all C-P bond containing natural product pathways characterized to date, the C-P bond originates via phosphoenolpyruvate mutase (PEP mutase) which catalyzes the intramolecular rearrangement of phosphoenolpyruvate to phosphonopyruvate (Figure 3-01).^{6, 7} This endergonic phosphoryl transfer reaction is proposed to proceed through a dissociated metaphosphate intermediate and multiple subsequent elaboration steps are required to produce the final alkylphosphonate natural products. For example, at least 10 steps are

required to produce the C-P framework for the γ -phosphonyl amino acid phosphinothricin in the herbicidal C-P peptide bialaphos (**1.16**).⁸ Due to the positioning of the aromatic ring in AHEP, it is not obvious how the carbon-phosphorus backbone of α -amino phosphonate AHEP could be formed by this mode of phosphonyl migration. To probe the biosynthesis of K-26 and AHEP, heavy-atom isotope labelled precursor incorporation experiments were designed.

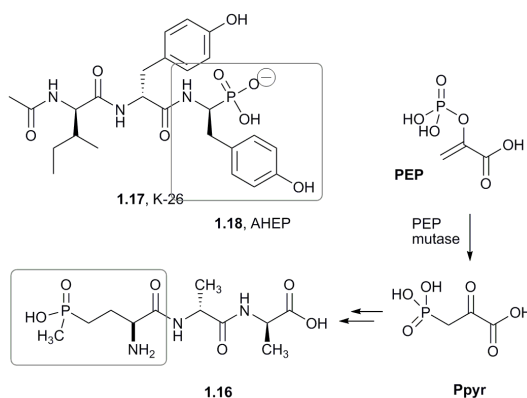


Figure 3-01. Comparing K-26 and bialaphos.

Results and Discussion

To investigate the biosynthetic origin of the atoms in the aromatic ring and side chain of AHEP a series of incorporation studies with isotopically-labeled tyrosines was undertaken. Extremely low production levels of K-26 precluded classical NMR based isotope estimation and required the development of a method with high sensitivity to accurately measure the incorporation of proposed primary metabolites into K-26. Less than 10 $\mu\text{g/L}$ of K-26 was detected in the fermentation broth, based on ACE inhibitory activity-based estimations. We developed tandem mass spectrometric methods using selected reaction

monitoring (SRM)^{9, 10} in which the [M-H] precursor ions of K-26 were fragmented by collision-induced dissociation. By judicious selection of precursor and product isotopomer masses in an experiment, quantitation of the isotopic enrichment in the charge-containing product ions and in the neutral loss fragments was possible. In effect, the easily selected precursor ion was used to indirectly observe the relative isotopic abundances in fragments. This application of SRM, is widely applicable to natural product biosynthetic studies, providing a highly sensitive and specific means of determining incorporation of monoisotopic or polyisotopic substitution in fragments by mass spectrometry.

Samples were prepared by pulse feeding growing 0.6 – 1.2 L cultures of Actinomycete sp. K-26 (1 mM/day for 5 days) with ring-*d*₄-tyrosine, 3,3-*d*₂-tyrosine and ¹⁵N-tyrosine. After six days of incubation, K-26 samples were isolated by solid phase extraction on HP-20 resin and enriched by centrifugal molecular weight filtration. K-26 was further separated from co-metabolites by reversed phase HPLC. Initial studies employing ACE bioassay guided fractionation by reported methodologies unambiguously identified K-26 by ESI-MS/MS and biological activity. Characteristic MS/MS fragmentation reactions, in comparison to authentic synthetic sample, verified the sequence of the unusual tripeptide and position of phosphonate moiety by loss of PO₃ from the C-terminal amino acid. These data enabled subsequent isotopic enrichment measurements which were performed on a triple quadrupole mass spectrometer using electrospray ionization (ESI), collision induced dissociation (CID) and selected reaction monitoring (SRM).

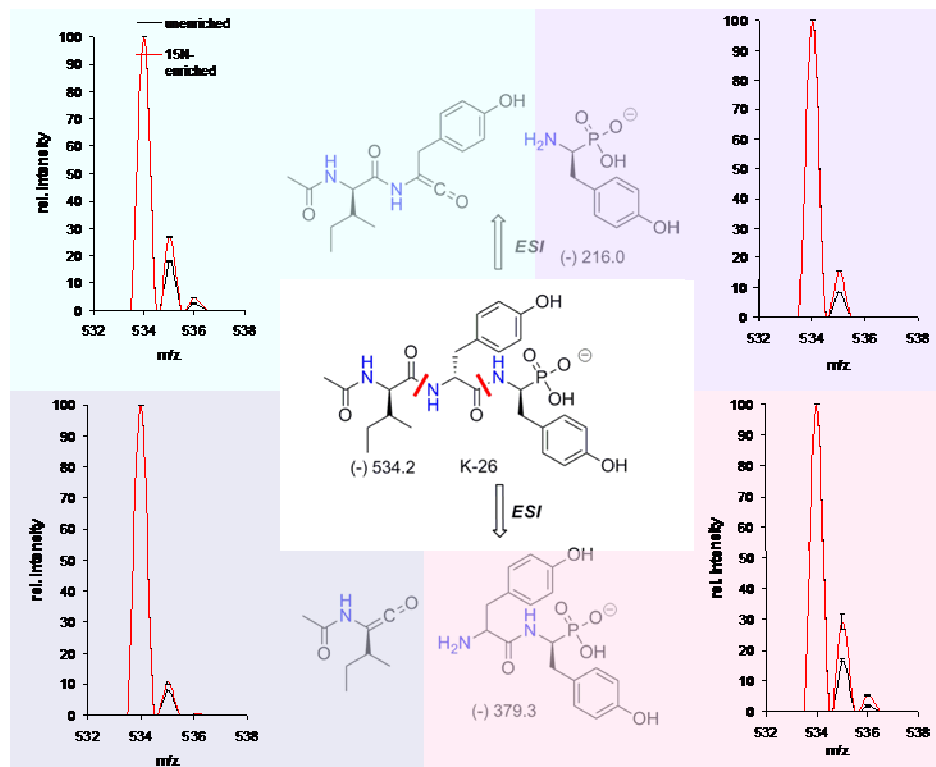


Figure 3-02. Simulated isotopic distributions based on ^{15}N SRM isotopic abundance calculations.

The incorporation of ^{15}N -tyrosine into K-26 amino acids was determined by numerical deconvolution using the theoretical mass isotopomer abundance of K-26 to compute the enrichment profiles in the tracer studies.¹¹ To accurately measure the isotopomer distribution in the individual amino acids, two CID reactions were monitored. First, we selected a specific precursor mass for K-26 (m/z 534, 535, 536) and monitored two characteristic product ions m/z 216 and m/z 217, which correspond to AHEP and AHEP+1, respectively. To determine if the ^{15}N was incorporated into the Ile or the central Tyr, the same precursor ions were chosen and we monitored product ions m/z 379, 380, and 381. SRM

experiments of these selected peptide fragmentations were used to accurately measure the isotope distribution ratios of enriched and unenriched samples using either neutral loss or charged residue fragments. Least squares fitting of labeled isotopomer data permitted the extraction of relative isotopic enrichments of each amino acid in the tripeptide K-26. The internal control of ^{15}N incorporation into isoleucine demonstrates the robustness of the applied isotopomer abundance back calculation technique and is presumably the result of transamination via α -keto-3-methyl valerate (Figure 3-03).

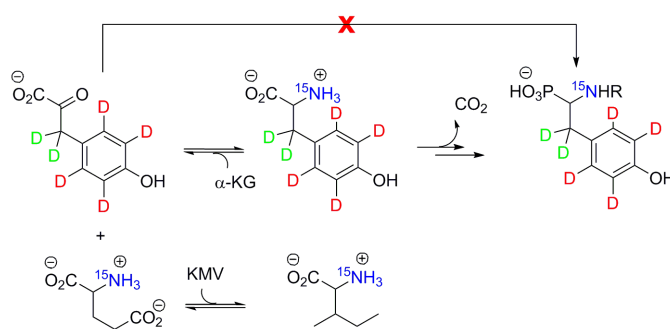
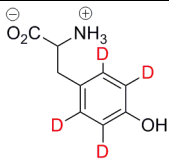

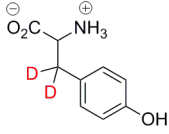


Figure 3-03. Summary of tyrosine isotopic enrichment experiments.

When ring- d_4 -tyrosine enriched samples were analyzed by MS, new pseudomolecular ion peaks appeared at the +4 and +8 position, indicating double incorporation of labeled tyrosine into K-26. SRM experiments were designed in order to decouple these data and determine the relative extent of incorporation into AHEP and tyrosine moieties. Transitions $536 \rightarrow 218$, $540 \rightarrow 218$, $540 \rightarrow 222$ were monitored in order to determine the relative amounts of unenriched K-26, incorporation in the central tyrosine, and incorporation in AHEP, respectively. Similarly, 3,3- d_2 -tyrosine incorporation was measured in

order to determine whether the biosynthesis of K-26 proceeds through an elimination intermediate. These data are summarized in Table 3-1.

Table 3-1. Incorporation of isotopically labeled tyrosine into K-26

Precursor	Ile	Tyr	AHEP
	N.D.	17.6±0.5	20.2±0.2
	2.6±0.1	6.2±0.6 ^a (5.4±1.4) ^b	7.1±0.4
	N.D.	16.1±0.6 ^b	18.7±0.6

^adetermined by subtraction of %Ile from % Ile-Tyr

^bdetermined by subtraction of %AHEP from % Tyr-AHEP

The trio of tyrosine incorporation studies clearly demonstrate that the non-proteinogenic amino acid AHEP is derived from tyrosine. Since de-aromatization of chorismate is highly unlikely, intact ring *d*₄-tyrosine incorporation suggests that AHEP biosynthesis branches from primary metabolism subsequent to chorismic acid. ¹⁵N-tyrosine incorporation data suggest that hydroxyphenylpyruvate is not an immediate progenitor for AHEP, as this would likely be evidenced by a washing out of ¹⁵N abundance in AHEP relative to tyrosine, which is not observed. The *d*₂-tyrosine incorporation data indicate that, in addition to intact

incorporation, the AHEP+1 peak was only weakly enhanced ($3.26\pm 0.1\%$) leading us to conclude that β -deuteriums are retained in the transformation of tyrosine to AHEP. In summary, these data unambiguously establish that enzymatic phosphorylation occurs subsequent to tyrosine biosynthesis and does not proceed via an eliminated intermediate. The substrate for phosphorylation may be tyrosine, a tyrosine metabolite or the tripeptide *N*-Ac-Ile-Tyr-Tyr. Any of these alternatives would indicate that the K-26 biosynthetic system is distinct from previously studied C-P containing natural products.

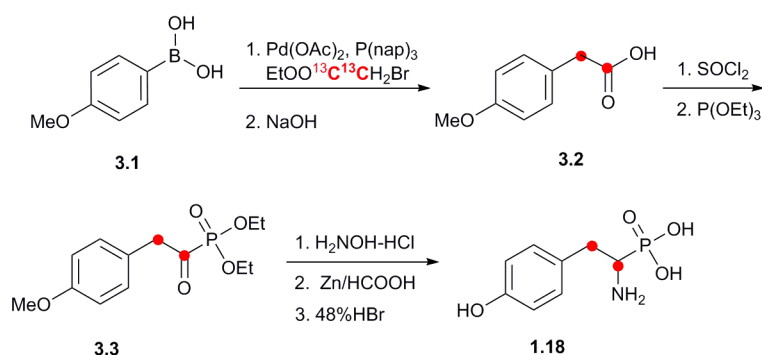


Figure 3-04. Synthesis of labelled AHEP

To determine the substrates and timing of the C-P bond forming step in the biosynthesis of K-26, $^{13}\text{C}_2$ -labeled AHEP (**1.18**) was synthesized for isotopic incorporation studies. Based on extant procedures,^{12, 13} we have developed an efficient synthetic route that facilitates the introduction of heavy atom isotopic labels in *rac*-AHEP (Figure 3-04). Suzuki coupling was used to couple 4-methoxyphenylboronic acid (**3.1**) with ethyl bromoacetate- $^{13}\text{C}_2$. In a four step one-pot process, the resulting ester was saponified, converted to the

corresponding acylchloride and reacted with triethylphosphite resulting in the α -keto phosphonate (**3.3**). This unstable compound was converted without isolation to the oxime with hydroxylamine, which was immediately reduced *in situ* in the presence of zinc/formic acid resulting in the respective amine. Global deprotection was achieved with hydrobromic acid and the final product *rac*-AHEP (**1.18**) was obtained in a 22% overall yield (based on ethyl bromoacetate- $^{13}\text{C}_2$) after purification by ion exchange filtration. AHEP was obtained in this two-pot process with only one chromatographic step required and >95% purity was confirmed by ^{31}P , ^{13}C , ^1H NMR and mass spectrometry.

Liquid cultures of the K-26 producing organism were separately supplemented with labeled and unlabeled synthetic racemic AHEP (0.3mM/day for 4 days). After six days of incubation, K-26 samples were isolated from fermentation supernatant by solid phase extraction with Diaion HP-20 polystyrene resin at pH = 3 and fractionated by 5K centrifugal molecular weight filtration. K-26 was further separated from co-metabolites by reverse phase HPLC/MS. Cultures that were supplemented with unlabeled AHEP were used as a reference and grew equally well as unsupplemented control cultures. The specific incorporation of $^{13}\text{C}_2$ -AHEP into K-26 amino acids was determined using the SRM method previously described.¹⁰ The isotopic mass distribution of K-26 from $^{13}\text{C}_2$ -AHEP fed cultures indicated the major isotopomer at $m/z=536$, a two Da shift relative to the unlabeled standard (Figure 3-05). Deconvolution of mass isotopomer data indicates that $^{13}\text{C}_2$ -AHEP was incorporated into K-26 at a level of 85% natural abundance.

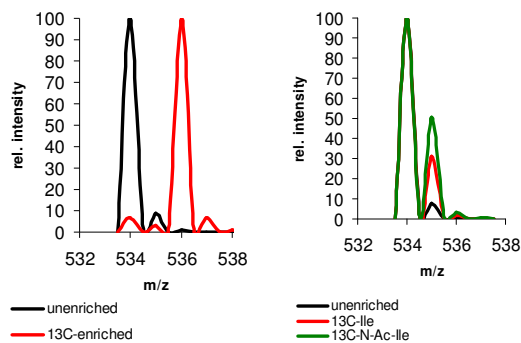


Figure 3-05. Calculated isotopic distributions based on SRM data. *Left:* calculated % $^{13}\text{C}_2$ -AHEP incorporation. *Right:* % ^{13}C -Ile and $^{13}\text{C}_2$ -*N*-acetyl $^{13}\text{C}_1$ -Ile incorporation.

The high level of incorporation of labeled AHEP, in combination with previous data suggests that indeed AHEP is a discrete precursor in the biosynthetic pathway of K-26 and the substrate of the C-P bond forming enzyme is most likely tyrosine or a closely related metabolite. Since decarboxylation of tyrosine is formally required for its conversion to AHEP we also investigated the precursor relationship of labeled tyramine to AHEP and K-26. Four separate incorporation studies (at 1.0 mM/day for 4 days) failed to demonstrate any detectable incorporation of d_4 -tyramine into K-26. While this result does not absolutely rule out the intermediacy of tyramine in the biosynthesis of K-26, it is most likely that tyrosine is a more direct metabolic precursor than tyramine in the biosynthetic pathway leading to AHEP. If tyrosine is indeed the direct precursor, this also suggests the possibility that decarboxylation is coupled to the C-P bond forming reaction in a single enzymatic reaction. In any event, these data demonstrate that C-P bond forming biochemistry other than PEP mutase is operative in the biosynthetic pathway of K-26.

A priori, it is expected that the K-26 gene cluster will encode an *N*-acetyltransferase (NAT) enzyme, which may acetylate isoleucine prior or subsequent to peptide coupling. In the biosynthesis of phosphinothricin, the γ -phosphonyl amino acid is acetylated prior to loading onto a nonribosomal peptide synthetase and is subsequently deacetylated at a later stage to produce the bioactive product.¹⁴ To determine if acetylation of isoleucine occurs prior to peptide coupling, as it does in the case of phosphinothricin, feeding studies with $^{13}\text{C}_1$ -labeled isoleucine and *N*-acetyl isoleucine were designed.

N-acetyl isoleucine was synthesized by reacting $^{13}\text{C}_2$ -labeled acetyl chloride and $^{13}\text{C}_1$ -isoleucine according to literature precedent.¹⁵ The ^{13}C -labeled isoleucine and *N*-acetyl isoleucine were separately introduced to growing cultures of the K-26 producer (0.4mM/day for 4 days). Samples were prepared as described above and incorporation levels were calculated in the same manner. Incorporation of $^{13}\text{C}_1$ -isoleucine and $^{13}\text{C}_2$ -*N*-acetyl $^{13}\text{C}_1$ -isoleucine into K-26 was calculated to be 24% and 43%, respectively (Figure 3-05). Interestingly, $^{13}\text{C}_2$ -*N*-acetyl $^{13}\text{C}_1$ -isoleucine was not incorporated intact. That *N*-acetyl isoleucine was incorporated only after being deacetylated was indicated by the enhancement of only the M+1 peak and not of the M+3. The higher level of enrichment of the M+1 peak in the case of $^{13}\text{C}_2$ -*N*-acetyl $^{13}\text{C}_1$ -isoleucine relative to isoleucine may be due to more favorable intracellular transport properties of this less charged compound. This isoleucine incorporation data demonstrate that acetylation must occur subsequent the first or both peptide couplings or perhaps acetylation of

isoleucine occurs while still covalently bound to an enzyme prior to condensation with tyrosine.

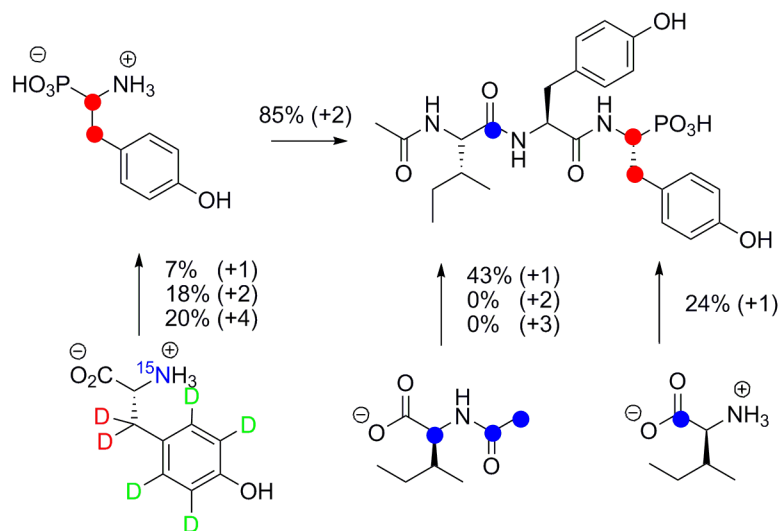


Figure 3-06. Summary of isotopic experiments

The combined data from all the feeding studies (Figure 3-06) delineate a minimal biosynthetic scheme for K-26. AHEP is a discrete intermediate in the biosynthesis of K-26 and tyrosine is a close precursor of AHEP. Tyrosine is not deaminated but it is decarboxylated *en route* to AHEP, but not via tyramine. Acetylation occurs subsequent to the formation of the first peptide bond and isoleucine is coupled to tyrosine prior to acetylation. The intermediacy of AHEP suggests that this nonproteinogenic amino acid is appended to a peptide precursor via a nonribosomal mechanism. In actinomycetes most small peptide natural products are synthesized by the canonical system of nonribosomal peptide synthetases (NRPS).¹⁶ Whether or not K-26 is synthesized by these systems or by an unrelated synthetase or ligase is unknown at this time.

However these results will inform ongoing studies focused on identifying the genes and enzymes involved in the K-26 biosynthesis, including those involved in peptide assembly and C-P bond formation.

Materials and Methods

Synthesis

General. All reagents were commercially available and used without purification unless specified otherwise. Thin-layer chromatography was carried out on precoated silica gel (60 F₂₅₄) plates obtained from EMD. Silica gel (35-70 μ m) from EM Science was used for column chromatography. All NMR spectra were acquired on a Bruker DPX-300 or a Bruker DRX-400 instrument unless otherwise noted. Mass spectral data was acquired by Nonka Sevova at Notre Dame University unless specified otherwise. Labeled ethyl bromoacetate, acetyl chloride and tyrosines were purchased from Cambridge Isotope Laboratories. Labeled isoleucine was obtained from Sigma-Aldrich. Labeled tyramine was synthesized by Vanessa Phelan.

Synthesis of (4-methoxy)phenyl [1,2-¹³C₂]acetic acid(3.2). To a dry flask was added palladium acetate (40.4 mg, 0.18 mmol), tri-1-naphthyl phosphine (218.6 mg, 0.53 mmol), K₃PO₄ (6.261 g, 29.5 mmol), ethyl bromoacetate (1 g, 5.9 mmol), 4-methoxyphenyl boronic acid (1.793 g, 11.8 mmol) and anhydrous tetrahydrofuran (25 mL) under argon atmosphere. The reaction was quenched with 45 mL water after 16hrs at room temperature and the mixture was extracted with dichloromethane (3×20mL). The combined organic layers were dried over

MgSO₄, the solvent was removed *in vacuo* and the product was purified via silica gel chromatography (19:1 hexane-ethyl acetate, R_f 0.25) to give an oil (0.658 g, 57.4%). ¹H NMR (CDCl₃) δ 1.20-1.25 (t, 3H, J=7.2Hz), 3.30-3.76 (dd, 2H, J=7.8Hz, ¹J_{CH}=129Hz), 3.78 (s, 3H), 4.08-4.16 (dq, 2H, J=7.2, ³J_{CH}=3Hz), 6.83-6.86 (d, 2H, J=11.7Hz), 7.16-7.20 (dd, 2H, J=11.7Hz, ³J_{CH}=3Hz) ¹³C NMR δ 13.9, 39.8-40.6 (d, J=57.6Hz), 55.0, 60.5, 113.6, 113.7, 129.9, 130.0, 158.3, 171.3-172.1 (d, J=57.6Hz) FAB-MS (M+H⁺): 196.1032.

Ethyl 2-(4-methoxy phenyl) acetate (1.891 g, 9.6 mmol) was stirred in 19.2mL 1N NaOH and 19.2mL methanol for 2 hours at room temperature. The solution was then extracted with ethyl acetate (3x10mL). The organic layers were combined, dried over MgSO₄, and the solvent was removed *in vacuo* to give a white solid (1.485 g, 92%). ¹H NMR (CDCl₃) δ 3.34-3.80 (dd, 2H, J=7.5Hz, ¹J_{CH}=129Hz), 3.78 (s, 3H), 6.83-6.86 (d, 2H, J=8.4Hz), 7.16-7.20 (dd, 2H, J=8.4Hz, ³J_{CH}=4.2Hz) ¹³C NMR δ 39.3-40.0 (d, J=55.4Hz), 55.0, 113.7, 113.8, 130.1, 158.6, 176.4-177.1 (d, J=55.4Hz) FAB-MS (M+H⁺): 168.0671.

Synthesis of diethyl (4-methoxy)phenyl [1,2- ¹³C₂]acetyl phosphonate (3.3). Thionyl chloride (0.7 mL, 9.7 mmol) was added to a suspension of (4-benzyloxy)phenyl [1,2- ¹³C₂]acetic acid (1.485 g, 8.8 mmol) in CHCl₃ (20 mL) containing DMF (0.1 mL). The mixture was stirred for 4 hours at reflux. The solvent and excess thionyl chloride were removed *in vacuo* yielding yellow oil. The product was used without isolation for the subsequent step.

The acyl chloride (1.642 g, 8.8 mmol) was dissolved in 5 mL THF and chilled to 0°C. Triethyl phosphite (1.53 mL, 8.8 mmol) was added dropwise under

anhydrous conditions. When the addition was complete, the reaction was warmed to 70°C and continued stirring for 15 minutes. The reaction was evaporated to obtain **(3.3)**, which was used without isolation in the following step.

Synthesis of (R)-1-amino-2-(4-hydroxyphenyl) [1,2-¹³C₂]ethylphosphonic acid (1.18). To a suspension of hydroxylammonium chloride (0.799g, 11.5mmol) in 1 mL pyridine and 2 mL ethanol was added slowly 2.536g **(3.3)** in 2 mL ethanol. After stirring for 12 hours the clear solution was evaporated and the oily residue treated with 10 mL 2% HCl. The mixture was extracted with dichloromethane. The organic layers were combined, dried and concentrated to yield a yellow oil (2.2735 g, 85.2%). The product was used in the next reaction without further purification or characterization.

To a suspension of zinc (1.9611 g, 30 mmol) in 7.5mL formic acid was added the oxime(2.2735 g, 7.5 mmol) under argon. The reaction was allowed to proceed overnight. The suspension was filtered and the filtrate was evaporated to give yellow oil (100%). ¹H NMR (CDCl₃) δ 1.20-1.39 (m, 6H), 2.80-2.92 (m, 1H), 3.10-3.21 (m, 1H), 3.48-3.52 (m, 1H), 3.78 (s, 3H), 4.09-4.20 (m, 4H), 6.84-6.86 (d, 2H, J=8.6Hz), 7.14-7.18 (dd, 2H, J=8.6Hz, ³J_{CH}=4.0Hz) ¹³C NMR δ 16.0, 34.1-34.5 (d, J=32.7Hz), 47.6-50.1 (dd, J=32.7Hz, ¹J_{CP}=155Hz), 55.0, 63.3, 63.4, 114.0, 130.1, 158.7 ³²P NMR δ 25.3-26.5 (d, ¹J_{CP}=155Hz) FAB-MS (M+Na⁺): 312.124574.

The amine (2.1696 g, 7.5 mmol) was dissolved in 100mL 48% HBr and refluxed for 2 hours. The solvent was evaporated and the resulting yellow residue was dissolved in 1mL 1N NaOH solution, run through a Dowex 50WX4-100

column (25mm ID × 100mm) and eluted with 100 mL water. Ninhydrin positive fractions were combined to give AHEP (940 mg, 57.2%). ^1H NMR (D_2O) δ 2.40-2.51 (m, 1H), 2.86-3.24 (m, 1H), 3.35-3.65 (m, 1H), 6.74-6.79 (d, 2H, $J=8.4\text{Hz}$), 7.08-7.12 (dd, 2H, $J=8.4\text{Hz}$, $^3J_{\text{CH}}=4\text{Hz}$) ^{13}C NMR δ 32.7-33.2 (d, $J=32.7\text{Hz}$), 49.5-51.8 (dd, $J=32.7\text{Hz}$, $^1J_{\text{CP}}=141\text{Hz}$), 115.5, 115.6, 130.2, 154.3 ^{32}P NMR δ 14.4-15.5 (d, $^1J_{\text{CP}}=141\text{Hz}$) FAB-MS ($\text{M}+\text{H}^+$): 220.0666.

Synthesis of [1,2- $^{13}\text{C}_2$]N-acetyl [1- $^{13}\text{C}_1$]isoleucine. Isoleucine (0.5 g, 3.8 mmol) was dissolved in 4N NaOH solution (2.85 mL, 11.4 mmol) and chilled to 0°C . Acetyl chloride (0.31 mL, 4.2 mmol) was added in 5 batches over 50 min. The reaction was acidified to pH 1.5 with concentrated HCl and extracted with ethyl acetate. The combined organic layers were dried over MgSO_4 and concentrated resulting in a white solid. The solid was recrystallized from ethyl acetate/hexane to give *N*-acetyl isoleucine (167 mg, 25%). ^1H NMR (DMSO) δ 0.80-0.84 (m, 6H), 1.10-1.23 (m, 1H), 1.31-1.42 (m, 1H), 1.66-1.72 (m, 1H), 1.61-2.06 (dd, 3H, $^1J_{\text{CH}}=127\text{Hz}$, $^2J_{\text{CH}}=6\text{Hz}$), 4.09-4.16 (m, 1H), 7.94-7.97 (m, 1H) ^{13}C NMR δ 11.3, 15.6, 22.0-22.7 (d, $J=50.8\text{Hz}$), 24.7, 36.4, 55.2, 169.0-169.6 (d, $J=50.8\text{Hz}$), 173.2 FAB-MS ($\text{M}+\text{H}^+$): 177.1216.

Stable isotopic incorporation experiments

Fermentation. The K-26 producing strain was obtained from the Agricultural Research Service (NRRL 12379). The production protocol suggested by Yamato *et al*¹ was followed with minor modifications. The seed medium consisted of dextrose, 0.1g/L; Difco soluble starch, 0.1g/L; Bacto beef extract, 0.05g/L; Bacto yeast extract, 0.05g/L; Bacto tryptone, 0.05g/L; and CaCO_3 ,

0.02g/L dissolved in distilled water and was adjusted to pH 7.2 before autoclaving. Fermentation was initiated by aseptically inoculating one loop of mycelia grown on an agar plate into a sterile 50-mL Falcon tube containing 10 mL seed medium (Phase I). The Falcon tube was incubated for 10-12 days at 28°C in a shaker incubator. In the second phase, 3 mL of the phase I seed culture were transferred into a 300-mL flask containing 30 mL of seed medium. The flask was incubated for 3-4 days at 28°C in a shaker incubator. In phase III, 30 mL of phase II seed culture were transferred into a 3000-mL Fernbach flask containing 300 mL production medium. The flask was incubated for 5-6 days at 28°C in a shaker incubator. The production medium contained Difco soluble starch, 0.4g/L; soy bean meal (Wild Oat), 0.3g/L; corn steep liquor (Sigma), 0.05g/L; K₂HPO₄, 5mg/L; MgSO₄, 2.4mg/mL; KCl, 3mg/L; and CaCO₃, 0.03g/L dissolved in distilled water and was adjusted to pH 7.8 before autoclaving.

Pulse feeding experiments. Millimolar amounts (based on the culture volume) of each precursor (0.3mM for AHEP, 0.4mM for isoleucine and *N*-acetylisoleucine, and 1mM for tyrosines and tyramine) were dissolved in 5 mL water and administered separately to phase III culture through a sterile syringe filter in every 24 hours for 4 days.

K-26 purification. The production culture was centrifuged and the supernatant was acidified to pH 3.0. To one liter of supernatant were added thirty grams of the hydrophobic resin Diaion HP-20 (activated by stirring with methanol and rinsing with water) and the suspension was stirred for 30 minutes. The resin was filtered and washed with water (2x30mL). The washed resin beads were

placed in a 50% methanol-50% water solution and stirred for 10 min. The suspension was then filtered, washed with 50% methanol (2x30 mL), and discarded. The filtrate was concentrated to 40 mL and neutralized. K-26 was further purified using Millipore Ultra 5000 MWCO centrifugal filter devices and centrifuging at 3450g for 30 minutes. The filtrate was concentrated to ca. one mL by rotary evaporation.

Mass spectrometry

General. Mass spectrometry was performed using ThermoFinnigan (San Jose, CA) TSQ® Quantum triple quadrupole mass spectrometer equipped with a standard electrospray ionization source outfitted with a 100- μ m I.D. deactivated fused Si capillary. Data acquisition and spectral analysis were conducted with Xcalibur™ Software, version 1.3, from ThermoFinnigan (San Jose, CA), on a Dell Optiplex GX270 computer running the Microsoft® Windows 2000 operating system. The source spray head was oriented at an angle of 90° to the ion-transfer-tube. Nitrogen was used for both the sheath and auxiliary gas. The sheath and auxiliary gases were set to 33 and 14 (arbitrary units) respectively. Samples were introduced by HPLC. A Surveyor® Autosampler and a Surveyor® MS Pump from ThermoFinnigan (San Jose, CA) were used. The injection volume was 10 μ L. K-26 was separated from co-metabolites using a Jupiter™ minibore 5 μ m C18 column (2.0mm \times 15cm) with a linear water-acetonitrile gradient (ranging from 95:5 to 5:95 H₂O:CH₃CN) containing 10mM ammonium acetate. The flow rate was 0.2mL/min.

Comparison of biogenic K-26 to synthetically prepared sample. Product ion spectra were generated from the precursor ion $m/z=534.1$. Only the precursor ion was allowed to pass through the first quadrupole, the ion was collisionally activated in the second quadrupole (20eV, argon @ 1.5mT). Product ion spectra were recorded for m/z 30.0-550.0. The mass-spectral resolution of both quads was set to a peak width of 0.7 u (full width at half maximum, FWHM).

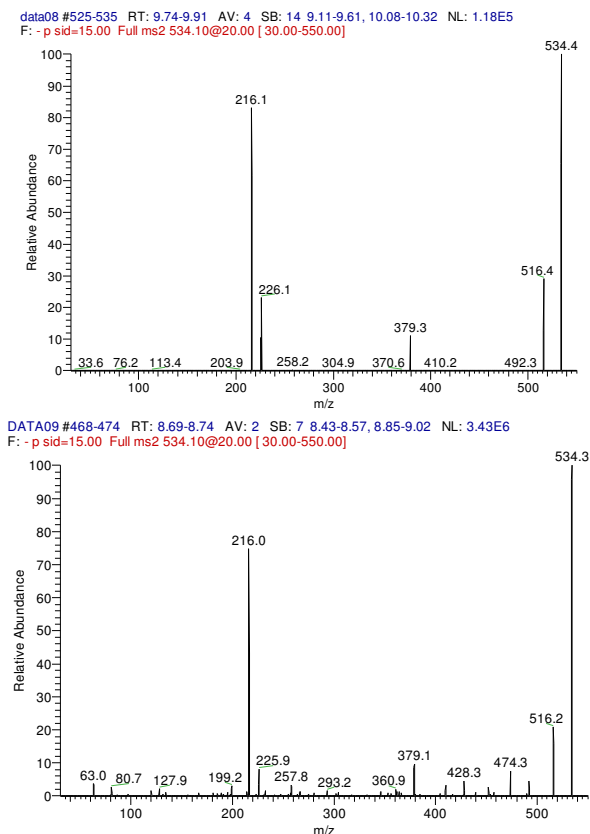


Figure 3-07. MS² of 534 ion of synthetic (above) versus biogenic (below) K-26.

Ring-d₄-tyrosine. The mass spectrometer was operated in the positive ion mode and the electrospray needle was maintained at 4200V. The ion transfer

tube was operated at 35V and 300°C. The tube lens voltage was set to 131V. Source CID (offset voltage between skimmer and the first ion guide, Q00) was used at 15V. The selected reaction monitoring (SRM) mode was used. Ions were collisionally activated with argon at an indicated pressure of 1.5mT. The mass-spectral resolution was set to a peak width (full width at half maximum, FWHM) of 0.70u and 0.70u for precursor and product ions respectively. Mass transitions at the specified collision energy (m/z 536→218; 20eV), (m/z 540→218; 20eV), and (m/z 540→222; 20eV) were monitored for unenriched K-26, enrichment in the central tyrosine and enrichment in AHEP, respectively. The scan width for product ions was 1.000u and the cycle time for each ion was 0.25 seconds. The electron multiplier gain was set to 2×10^6 . Data were acquired in profile mode. The resulting chromatogram is illustrated in Figure 3-08.

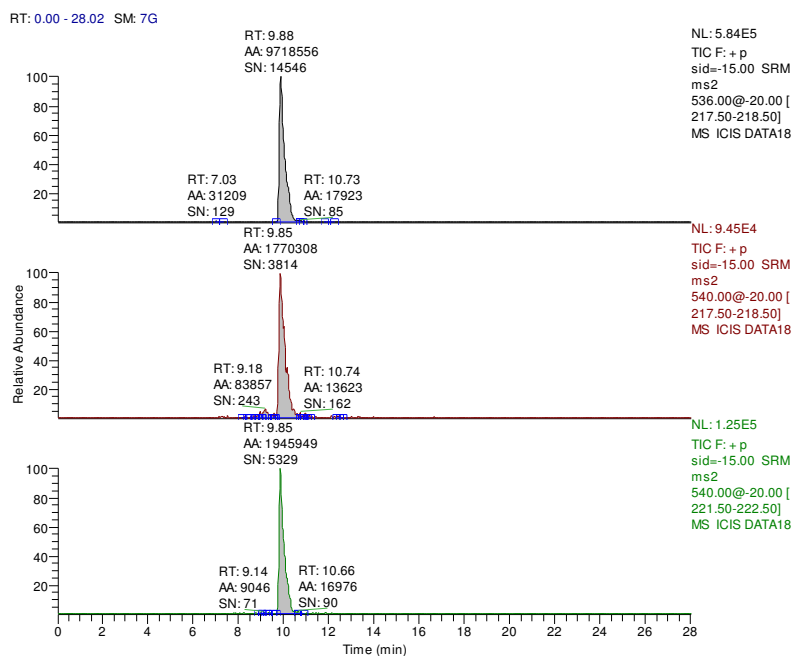


Figure 3-08. Chromatograms used for the calculation of ring- d_4 -tyrosine incorporation

Table 3-2. Calculation of ring-*d*₄-tyrosine incorporation

Transition	Area			Isotope Distribution				
	1	2	3	1	2	3	average	SD
536 --> 218	9701469	9674024	9736478	100	100	100	100	0
540 --> 218	1719261	1658059	1755359	17.72	17.13	18.03	17.63	0.45
540 --> 222	1983006	1954870	1962925	20.44	20.21	20.16	20.27	0.15

The isotope distribution was calculated by assigning the 536→218 peak an area of 100 and normalizing the other peaks accordingly.

¹⁵N-tyrosine. The mass spectrometer was operated in the negative ion mode and the electrospray needle was maintained at 4200V. The ion transfer tube was operated at -35V and 350°C. The tube lens voltage was set to -150V. Source CID (offset voltage between skimmer and the first ion guide, Q00) was used at 15V. The selected reaction monitoring (SRM) mode was used. Ions were collisionally activated with argon at an indicated pressure of 1.4mT. The mass-spectral resolution was set to a peak width (full width at half maximum, FWHM) of 0.50u and 0.50u for precursor and product ions respectively. Mass transitions at the specified collision energy (m/z 534→216; 35eV), and (m/z 535→217; 35eV) were monitored for AHEP and AHEP+1, respectively. Transitions (m/z 534→216; 35eV), (m/z 535→216; 35eV), and (m/z 536→216; 35eV) were monitored for *N*-Ac-Ile-Tyr, *N*-Ac-Ile-Tyr+1, and *N*-Ac-Ile-Tyr+2, respectively. Transitions (m/z 534→379; 35eV), (m/z 535→380; 35eV), and (m/z 536→381; 35eV) were monitored for Tyr-AHEP, Tyr-AHEP+1, and Tyr-AHEP+2, respectively. Finally, transitions (m/z 534→379; 35eV), (m/z 535→379; 35eV), and (m/z 536→379; 35eV) were monitored for *N*-Ac-Ile, *N*-Ac-Ile+1, and *N*-Ac-Ile+2, respectively. The scan width for product ions was 1.000u and the cycle

time for each ion was 0.15 seconds. The electron multiplier gain was set to 2×10^6 . Data were acquired in profile mode. The resulting chromatograms are illustrated in Figures 3-09, 3-10, and 3-11. Both a ^{15}N -tyrosine enriched sample and an unenriched sample were scanned. The unenriched sample was used to create theoretical curves of ratios of mass isotopomer abundance in K-26 to correct the isotopomer distribution of the enriched sample. Sample calculations can be found in Appendix C.

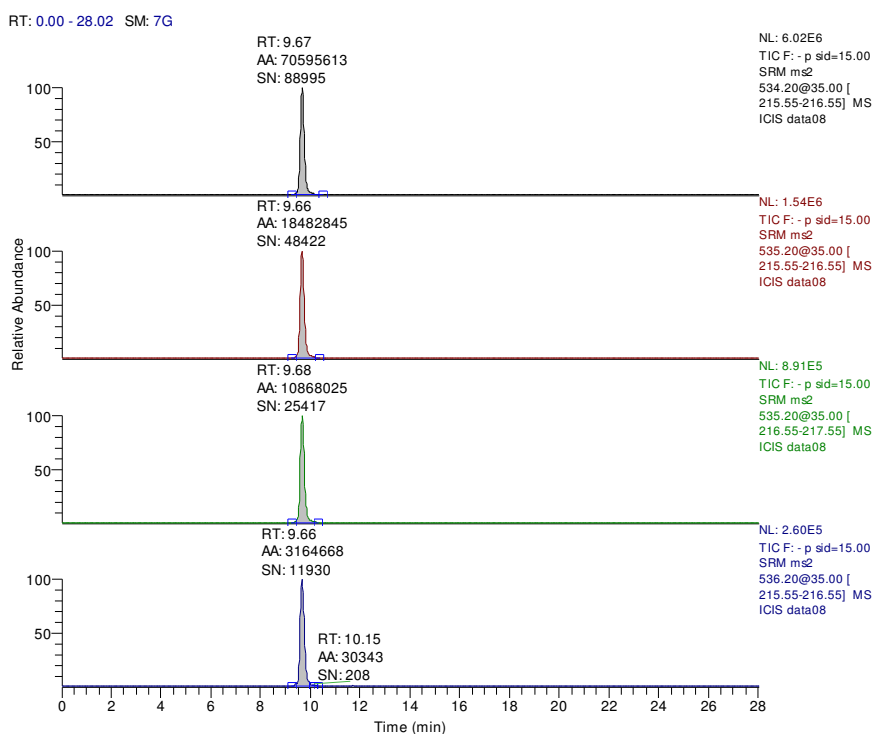


Figure 3-09. Chromatograms used in the calculation of ^{15}N -tyrosine incorporation in AHEP

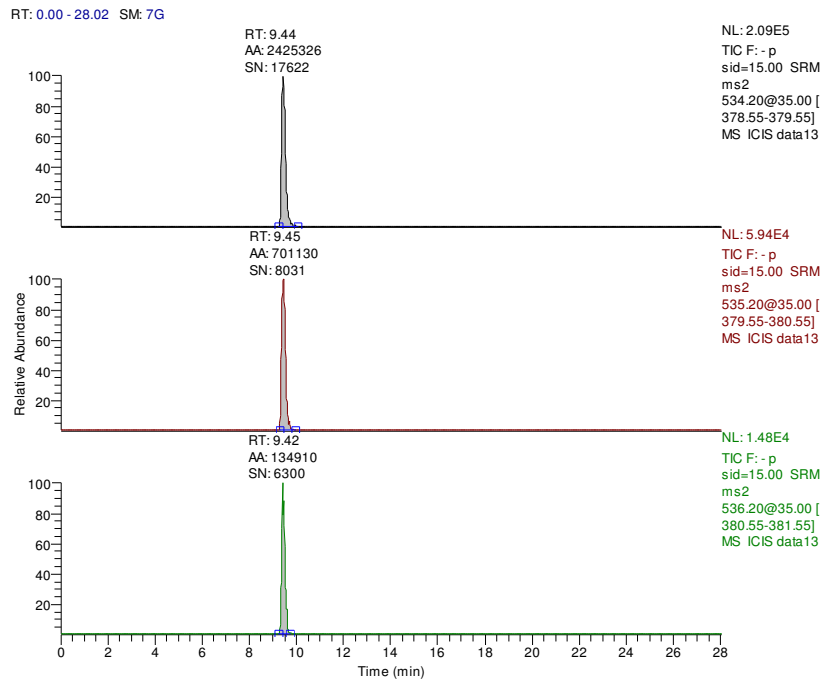


Figure 3-10. Chromatograms used in the calculation of ^{15}N -tyrosine incorporation in Tyr-AHEP

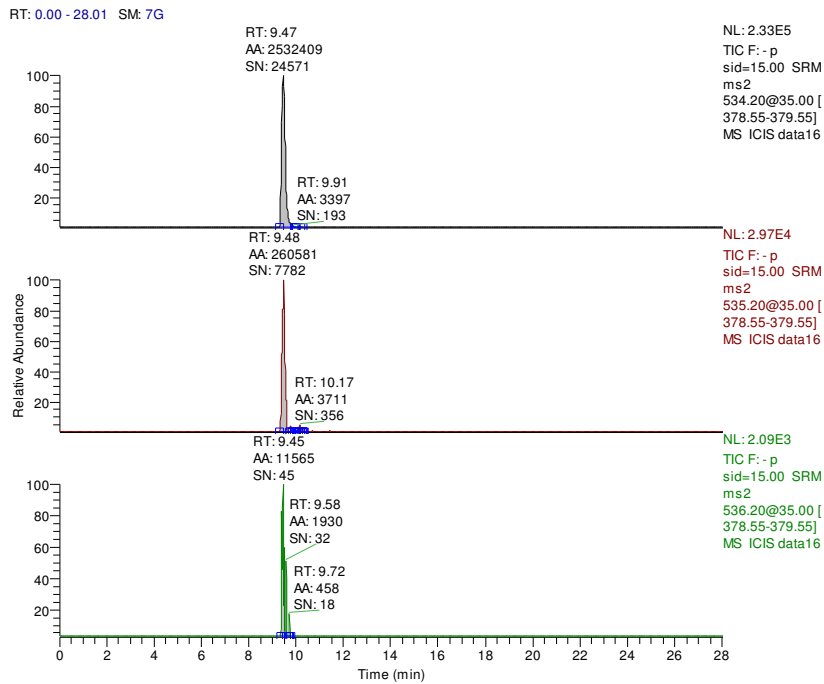


Figure 3-11. Chromatograms used in the calculation of ^{15}N -tyrosine incorporation in N-Ac-Ile-Tyr

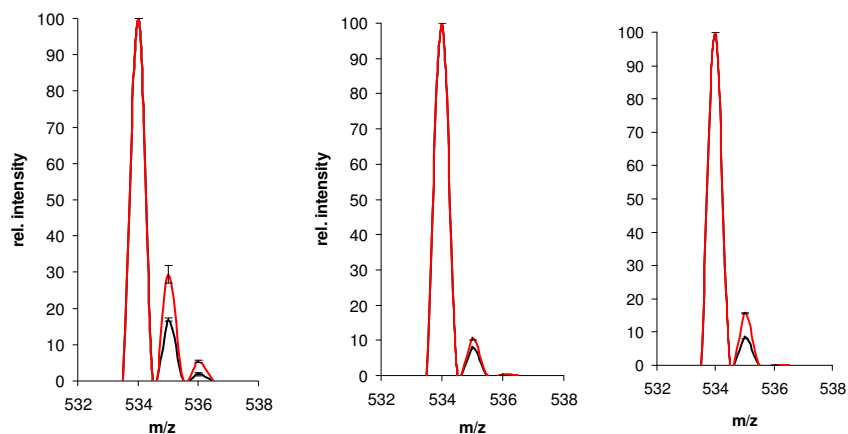


Figure 3-12. Simulated isotopic distributions based on least squares fitting of isotopomer data. Red is labeled black is unlabeled. *Left:* %¹⁵N enrichment in Tyr-AHEP fragment. *Middle:* %¹⁵N enrichment in N-Ac-Ile fragment. *Right:* %¹⁵N in AHEP alone.

*d*₂-tyrosine. The mass spectrometer was operated in the positive ion mode and the electrospray needle was maintained at 4200V. The ion transfer tube was operated at 35V and 300°C. The tube lens voltage was set to 131V. Source CID (offset voltage between skimmer and the first ion guide, Q00) was used at 15V. The selected reaction monitoring (SRM) mode was used. Ions were collisionally activated with argon at an indicated pressure of 1.5mT. The mass-spectral resolution was set to a peak width (full width at half maximum, FWHM) of 0.50u and 0.50u for precursor and product ions respectively. Mass transitions at the specified collision energy (m/z 536→218; 20eV), (m/z 537→219; 20eV), (m/z 538→220; 20eV), and (m/z 539→221; 20eV) were monitored for AHEP, AHEP+1, AHEP+2, and AHEP+3, respectively. Additionally, transitions at the specified collision energy (m/z 536→381; 20eV), (m/z 537→382; 20eV), (m/z

538→383; 20eV), (m/z 539→384; 20eV), and (m/z 540→385; 20eV) were monitored for Tyr+AHEP, Tyr+AHEP+1, Tyr+AHEP+2, Tyr+AHEP+3, and Tyr+AHEP+4, respectively. The scan width for product ions was 1.000u and the cycle time for each ion was 0.15 seconds. The electron multiplier gain was set to 2×10^6 . Data were acquired in profile mode. The resulting chromatograms are illustrated in Figures 3-13 and 3-14. Both a d_2 -tyrosine enriched sample and an unenriched sample were scanned. The unenriched sample was used to create a least squares curve for the data. The same method of calculation as in the ^{15}N -tyrosine experiment was used for determining the amount and position of incorporation. The enrichment level in the central tyrosine was calculated by subtracting the AHEP enrichment from the Tyr-AHEP enrichment. Simulated isotopic distributions based on least squares fitting of isotopomer data are shown in figure 3-15.

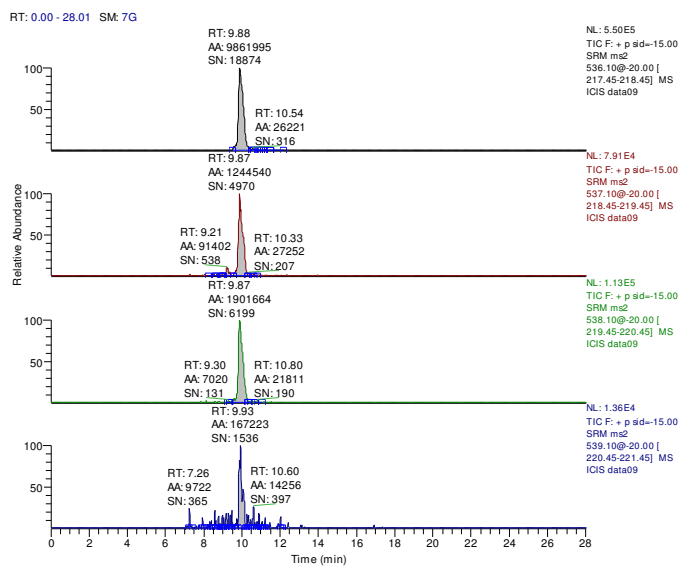


Figure 3-13. Chromatograms used in the calculation of 3,3- d_2 -tyrosine incorporation in AHEP fragment.

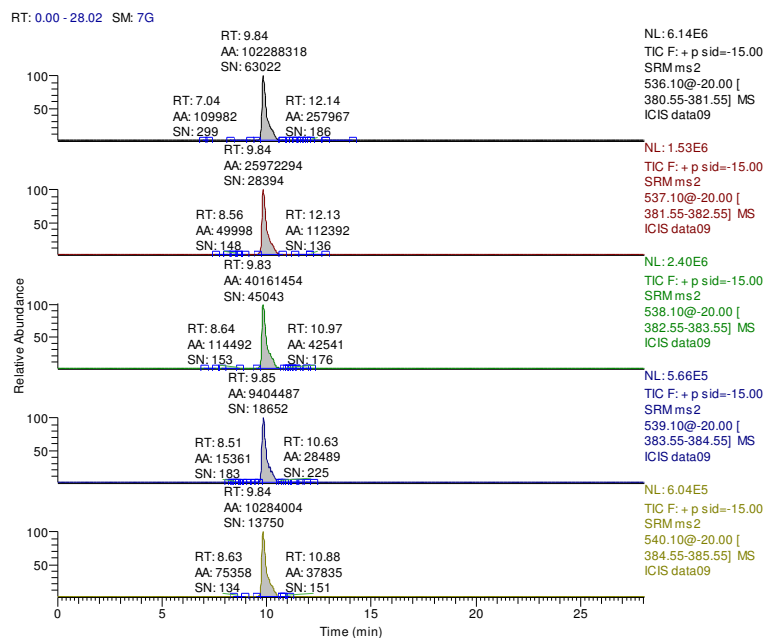


Figure 3-14. Chromatograms used in the calculation of 3,3- d_2 -tyrosine incorporation in Tyr-AHEP dipeptide fragment.

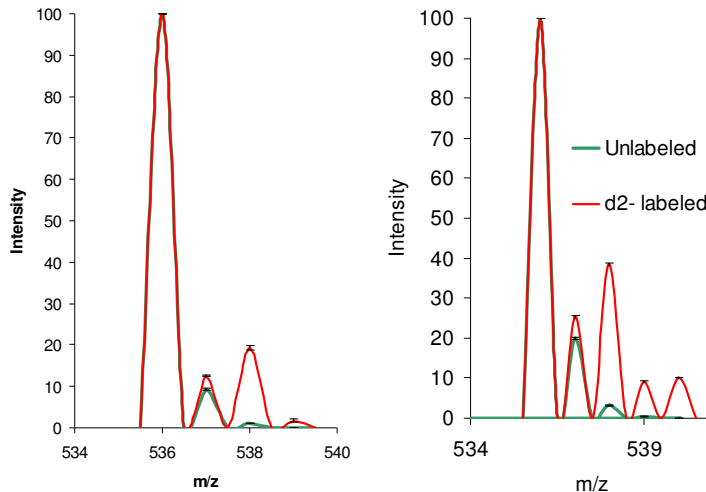


Figure 3-15. Simulated isotopic distributions based on least squares fitting of isotopomer data from chromatograms in figure 3-13 and 3-14. Red is labeled sample, green is unlabeled. *Left:* calculated % β - d_2 -Tyrosine in AHEP fragment. These data clearly show conservation of both deuteriums into AHEP and very little evidence of loss of single deuterium. *Right:* % β - d_2 -Tyrosine in Tyr-AHEP fragment confirming double incorporation.

¹³C₂-AHEP. The mass spectrometer was operated in the negative ion mode and the electrospray needle was maintained at 4200V. The ion transfer tube was operated at -35V and 350°C. The tube lens voltage was set to -150V. Source CID (offset voltage between skimmer and the first ion guide, Q00) was used at 15V. The selected reaction monitoring (SRM) mode was used. Ions were collisionally activated with argon at an indicated pressure of 1.4mT. The mass-spectral resolution was set to a peak width (full width at half maximum, FWHM) of 0.50u and 0.50u for precursor and product ions respectively. Mass transitions at the specified collision energy (m/z 534→216; 35eV), (m/z 536→218; 35eV) and (m/z 540→220; 35eV) were monitored for AHEP, AHEP+2, and AHEP+4, respectively. Transitions (m/z 534→379; 35eV), (m/z 536→381; 35eV), and (m/z 538→383; 35eV) were monitored for Tyr-AHEP, Tyr-AHEP+2, and Tyr-AHEP+4, respectively. The scan width for product ions was 1.000u and the cycle time for each ion was 0.15 seconds. The electron multiplier gain was set to 2 × 10⁶. Data were acquired in profile mode. Both a ¹³C₂-AHEP enriched sample and an unenriched sample were scanned. The unenriched sample was used to create theoretical curves of ratios of mass isotopomer abundance in K-26 to correct the isotopomer distribution of the enriched sample. Similar calculations as described earlier¹⁰ were used and the incorporation of ¹³C₂-AHEP in K-26 was calculated as 85%. No ¹³C enrichment was observed in the central tyrosine.

Ring-d₄-tyramine. The mass spectrometer was operated in the negative ion mode and the electrospray needle was maintained at 4200V. The ion transfer tube was operated at -35V and 350°C. The tube lens voltage was set to -150V.

Source CID (offset voltage between skimmer and the first ion guide, Q00) was used at 15V. The selected reaction monitoring (SRM) mode was used. Ions were collisionally activated with argon at an indicated pressure of 1.4mT. The mass-spectral resolution was set to a peak width (full width at half maximum, FWHM) of 0.50u and 0.50u for precursor and product ions respectively. Mass transitions at the specified collision energy (m/z 534→216; 35eV), (m/z 538→220; 35eV) and (m/z 538→220; 35eV) were monitored for unenriched K-26, enrichment in the central tyrosine and enrichment in AHEP, respectively. The scan width for product ions was 1.000u and the cycle time for each ion was 0.25 seconds. The electron multiplier gain was set to 2×10^6 . Data were acquired in profile mode. Both a ring-*d*₄-tyramine enriched sample and an unenriched sample were scanned. The mass spectral data for the two samples was almost identical, indicating no *d*₄ enrichment of K-26.

*¹³C*₃-*N*-acetyl isoleucine and *¹³C*₁-isoleucine. The mass spectrometer was operated in the negative ion mode and the electrospray needle was maintained at 4200V. The ion transfer tube was operated at -35V and 350°C. The tube lens voltage was set to -150V. Source CID (offset voltage between skimmer and the first ion guide, Q00) was used at 15V. The selected reaction monitoring (SRM) mode was used. Ions were collisionally activated with argon at an indicated pressure of 1.4mT. The mass-spectral resolution was set to a peak width (full width at half maximum, FWHM) of 0.50u and 0.50u for precursor and product ions respectively. Transitions (m/z 534→379; 35eV), (m/z 535→379; 35eV), (m/z 536→379; 35eV), and (m/z 537→379; 35eV) were monitored for *N*-Ac-Ile, *N*-Ac-

Ile+1, *N*-Ac-Ile+2, and *N*-Ac-Ile+3, respectively. The scan width for product ions was 1.000u and the cycle time for each ion was 0.15 seconds. The electron multiplier gain was set to 2×10^6 . Data were acquired in profile mode. An unenriched sample was used to create theoretical curves of ratios of mass isotopomer abundance in K-26 to correct the isotopomer distribution of the enriched sample. Similar calculations as described earlier¹⁰ were used and the incorporation of $^{13}\text{C}_3$ -*N*-acetyl isoleucine and $^{13}\text{C}_1$ -isoleucine in K-26 were calculated as 43% and 24%, respectively. In the case of *N*-acetyl isoleucine, there was only enrichment of the M+1 peak and not of the M+3, indicating incorporation of the amino acid occurred after deacetylation.

References

1. Yamato, M.; Koguchi, T.; Okachi, R.; Yamada, K.; Nakayama, K.; Kase, H.; Karasawa, A.; Shuto, K., K-26, a Novel Inhibitor of Angiotensin-I Converting Enzyme Produced by an Actinomycete K-26. *Journal of Antibiotics* **1986**, 39, (1), 44-52.
2. Kasai, M.; Yoshida, N.; Hirayama, N.; Shirahata, K., Structure Elucidation of New Inhibitors of angiotensin I Converting Enzyme, K-26 and K-4. *Symposium Papers, The 27th Symposium on the Chemistry of Natural Products* **1985**, 577.
3. Ohuchio, S.; Kurihara, K.; Shinohara, A.; Takei, T.; Yoshida, J.; Amano, S.; Miyadoh, S.; Matsushida, Y.; Somura, T.; Sezaki, M., Studies on New Angiotensin Converting Enzyme Inhibitors, SF2513A,B and C, Produced by *Streptosporangium nondiastaticum*. *Sci. Reports of Meiji Seika Kaisha* **1988**, 27, 46 - 54.
4. Koguchi, T.; Yamada, K.; Yamato, M.; Okachi, R.; Nakayama, K.; Kase, H., K-4, a Novel Inhibitor of Angiotensin-I Converting Enzyme Produced by *Actinomadura-Spiculosospora*. *Journal of Antibiotics* **1986**, 39, (3), 364-371.
5. Hirayama, N.; Kasai, M.; Shirahata, K., Structure and Conformation of a Novel Inhibitor of Angiotensin-I Converting Enzyme - a Tripeptide Containing Phosphonic Acid. *International Journal of Peptide and Protein Research* **1991**, 38, (1), 20-24.
6. Liu, S.; Lu, Z.; Jia, Y.; Dunaway-Mariano, D.; Herzberg, O., Dissociative phosphoryl transfer in PEP mutase catalysis: structure of the enzyme/sulfoxyruvate complex and kinetic properties of mutants. *Biochemistry* **2002**, 41, (32), 10270-6.
7. Liu, S. J.; Lu, Z. B.; Han, Y.; Jia, Y.; Howard, A.; Dunaway-Mariano, D.; Herzberg, O., Conformational flexibility of PEP mutase. *Biochemistry* **2004**, 43, (15), 4447-4453.
8. Blodgett, J. A.; Zhang, J. K.; Metcalf, W. W., Molecular cloning, sequence analysis, and heterologous expression of the phosphinothricin tripeptide biosynthetic gene cluster from *Streptomyces viridochromogenes* DSM 40736. *Antimicrob Agents Chemother* **2005**, 49, (1), 230-40.

9. Yost, R. A.; Boyd, R. K., Tandem Mass-Spectrometry - Quadrupole and Hybrid Instruments. *Methods in Enzymology* **1990**, 193, 154-200.
10. Ntai, I.; Manier, M. L.; Hachey, D. L.; Bachmann, B. O., Biosynthetic origins of C-P bond containing tripeptide K-26. *Organic Letters* **2005**, 7, (13), 2763-2765.
11. Culea, M.; Hachey, D. L., Determination of Multiply Labeled Serine and Glycine Isotopomers in Human Plasma by Isotope-Dilution Negative-Ion Chemical-Ionization Mass-Spectrometry. *Rapid Communications in Mass Spectrometry* **1995**, 9, (8), 655-659.
12. Drescher, M.; Li, Y. F.; Hammerschmidt, F., Enzymes in Organic-Chemistry .2. Lipase-Catalyzed Hydrolysis of 1-Acyloxy-2-Arylethylphosphonates and Synthesis of Phosphonic Acid Analogs of L-Phenylalanine and L-Tyrosine. *Tetrahedron* **1995**, 51, (17), 4933-4946.
13. Quin, L. D., *Organophosphorus Chemistry*. 1st ed.; Wiley-Interscience: New York, 2000.
14. Grammel, N.; Schwartz, D.; Wohlleben, W.; Keller, U., Phosphinothricin-tripeptide synthetases from *Streptomyces viridochromogenes*. *Biochemistry* **1998**, 37, (6), 1596-1603.
15. Chenault, H. K.; Dahmer, J.; Whitesides, G. M., Kinetic Resolution of Unnatural and Rarely Occurring Amino-Acids - Enantioselective Hydrolysis of N-Acyl Amino-Acids Catalyzed by Acylase-I. *Journal of the American Chemical Society* **1989**, 111, (16), 6354-6364.
16. Sieber, S. A.; Marahiel, M. A., Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem Rev* **2005**, 105, (2), 715-38.

Despite the potent hypotensive activity of K-26 and related compounds, the biosynthetic pathways by which K-26 and more specifically AHEP are biosynthesized still remain uncharacterized. Earlier studies in our lab^{5, 6} have established that the biosynthesis of K-26 is distinct from most other C-P bond containing metabolites^{7, 8} in that the precursor for AHEP is derived from tyrosine, instead of phosphoenolpyruvate. Labeling studies have established that AHEP is a discrete intermediate of K-26 and tyrosine is a close precursor of AHEP. The transformation of tyrosine to AHEP occurs with retention of nitrogen and stereochemical configuration at the α -carbon of tyrosine. The intermediacy of AHEP suggests that this nonproteinogenic amino acid is appended to a peptide precursor via a nonribosomal mechanism. In actinomycetes most small peptide natural products are synthesized by the canonical system of nonribosomal peptide synthetases (NRPS).⁹ Whether or not K-26 is synthesized by these systems or by an unrelated synthetase or ligase is unknown at this time. However these results underline the importance of further studies focused on identifying the genes and enzymes involved in the K-26 biosynthesis, including those involved in peptide assembly and C-P bond formation.

In an effort to identify the K-26 gene cluster, the sequencing of the entire genome of the K-26 producing organism has been completed. We have identified > 19 putative biosynthetic gene clusters and 16 of them encode nonribosomal peptide synthetase clusters. Rough annotation has provided structural classes and sub-structural clues for the metabolites encoded by these gene clusters. We

currently have a 3.5x draft of the genome and a 10,000 member fosmid library covering the entire genome.

Our original proposal was that K-26 was synthesized by a nonribosomal synthetase (NRPS) mechanism. If this were the case, we would expect domains for the adenylation (A), thioesterification (T), condensation (C) and thioesterase (Te) domains. In colinear fashion, the domain order for K-26 biosynthesis would be comprised of three modules [A-T]-[C-A-T]-[C-A-T-Te] with a possible acetyltransferase domain appended prior to the first C domain. Homology modeling of A-domains has proven to be an effective means of determining their amino-acid specificity.¹⁰ However, analysis of all the adenylation domains for the extant NRPS gene clusters in the producing organism has yet to reveal an obvious candidate gene cluster for the biosynthesis of K-26.

Due to the unusual nature of the K-26 biosynthetic pathway, sequence analysis alone may be insufficient for identifying the pathway. However, since the precursors of K-26 in the biosynthetic pathway have been identified, then reverse genetics may be employed to discover the K-26 gene cluster. Specifically, this chapter describes biochemical experiments focused in identifying the *N*-acetyltransferase involved in the K-26 biosynthesis. Biosynthetic genes for secondary metabolite production in most bacteria are generally found in clusters.¹¹ In that way, identification of gene sequence of a single biosynthetic enzyme, allows the identification of the entire biosynthetic gene cluster.

Previous studies⁶ have suggested that acetylation occurs after at least the first peptide coupling limiting the possible substrates of acetylation to only two: isoleucyl-tyrosine or des-acetyl K-26 (dAcK26, **2.6a**) as illustrated in Figure 4-02.

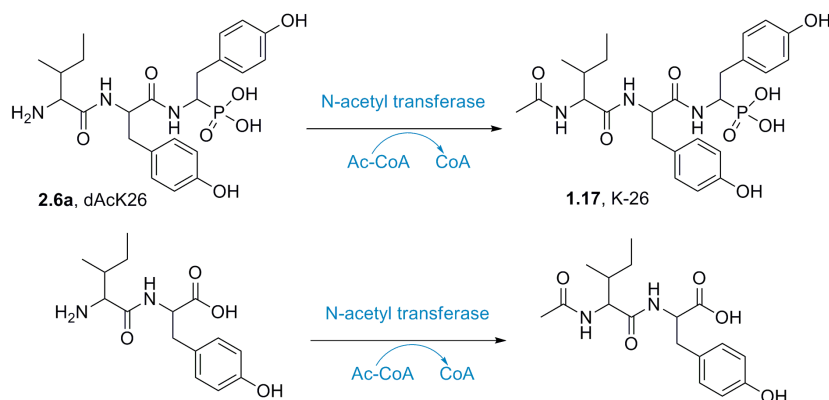


Figure 4-02. Possibilities for *N*-acetyltransferase activity in K-26 biosynthesis

To further investigate the *N*-acetyltransferase activity in K-26 biosynthesis, these substrates were synthesized and a highly sensitive and selective tandem MS method was developed to assay for *in vitro* acetylation reactions. Active protein fractions were analyzed using proteomics in order to identify the K-26 *N*-acetyltransferase.

In a typical proteomics experiment (Figure 4-03)¹², the proteins to be analyzed are further separated using one-dimensional gel electrophoresis and bands from the gel are cut and digested, typically with trypsin, resulting in smaller peptides. These peptides are separated by one or more steps of high-pressure liquid chromatography in very fine capillaries and eluted into an electrospray ion source. As they enter the mass spectrometer, a mass spectrum of the peptides eluting is taken. Each peptide ion is isolated and fragmented by energetic

collision with gas and its MS/MS spectrum is recorded. The MS and MS/MS spectra of each peptide are matched against protein sequence databases resulting in identifying the peptides and therefore the proteins present in the gel. In our case, since the genome of the organism has been sequenced, peptide sequences were compared to a database consisting of the organism's proteome and emphasis was placed on *N*-acetyltransferases.

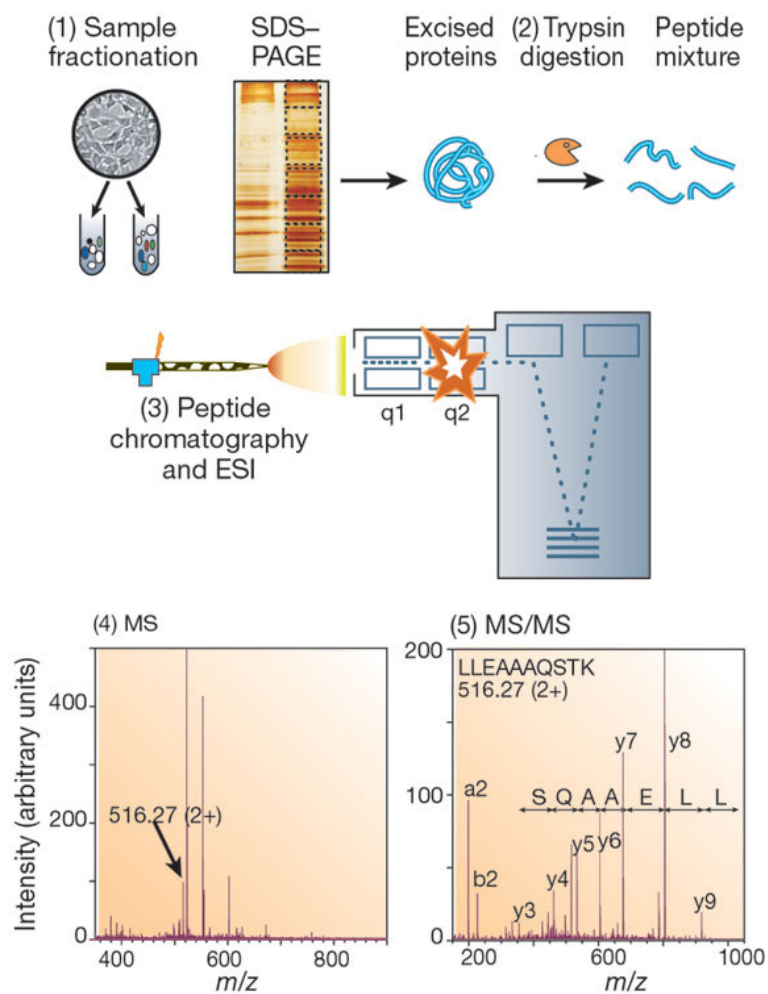


Figure 4-03. Typical proteomics experiment¹²

Results and Discussion

Low production levels of K-26 (<10 μ g/L) by fermentation required the development of a synthetic route for K-26 and des-acetyl K-26. These phosphonopeptides were synthesized as described earlier¹³ (Figure 2-03) without modifications. Isoleucyl-tyrosine and *N*-acetyl-isoleucyl tyrosine were synthesized using standard peptide coupling conditions.¹⁴

As a first step, cells were lysed and after removal of cellular debris by centrifugation, the resulting cell-free extract was incubated with these substrates in the presence of an excess of acetyl coenzyme A (AcCoA) for four hours at 30°C. *N*-acetyltransferase activity was determined using a colorimetric, a fluorometric and a mass spectrometric method. Both colorimetric and fluorometric assay detect a byproduct of the reaction, coenzyme A (CoA), while the mass spectrometric method detects disappearance of substrate and formation of product.

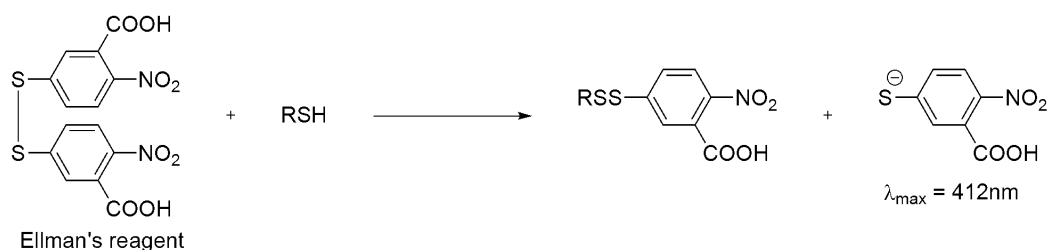


Figure 4-04. Detection of thiol RSH using Ellman's reagent

In detail, Ellman's reagent (colorimetric assay)¹⁵ reacts with thiols to form an anion that absorbs at 412nm (Figure 4-04). The fluorometric assay used employs coumarinylphenylmaleimide (CPM),¹⁶ which reacts with thiols to give a

highly fluorescent product (Figure 4-05). Typically, acetyltransferase activity is detected using a colorimetric, fluorometric or radiometric^{17, 18} assay. The drawback of those approaches is that they are not selective. As a result the background noise due to thiols present on proteins (cysteines) or thiols present in the buffer required for protein stability is extremely high. In addition, hydrolysis of AcCoA to form CoA may occur independently of acetylation of dAck26.

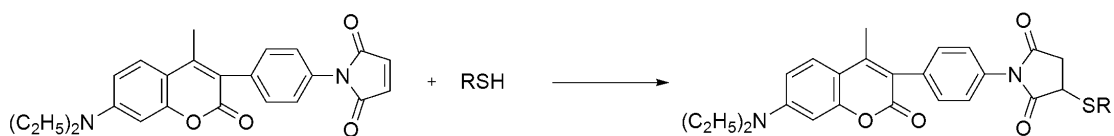


Figure 4-05. Detection of thiol RSH using CPM

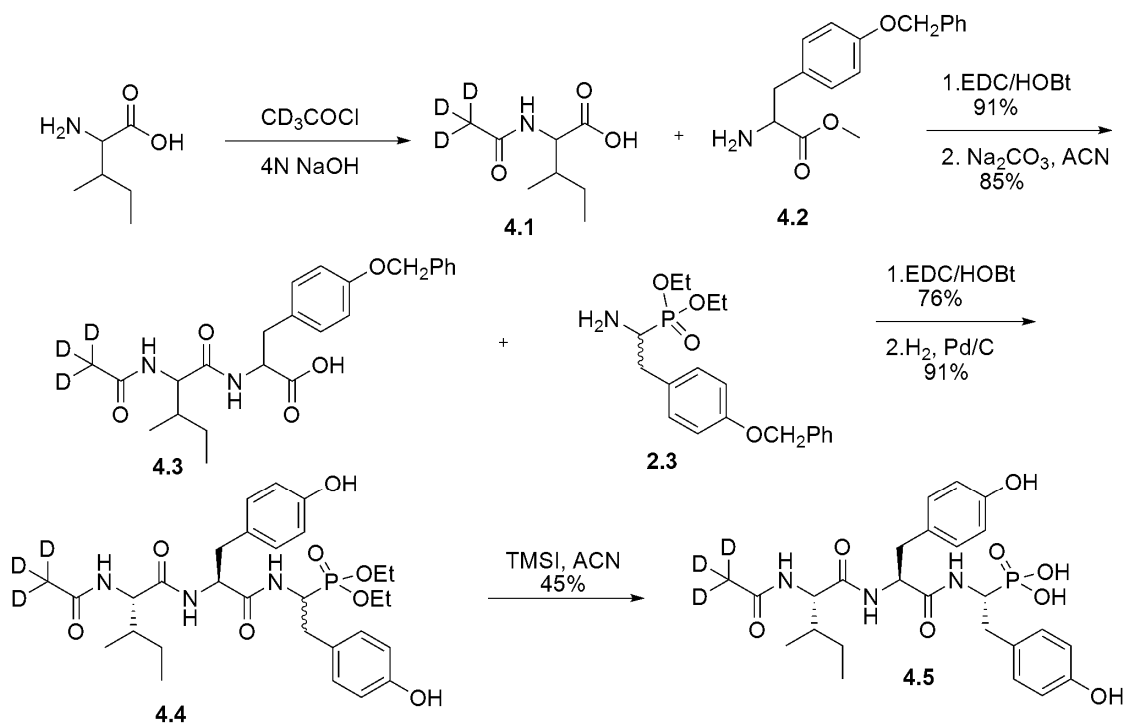


Figure 4-06. Synthesis of d_3 -K-26

A selected reaction monitoring (SRM) based detection method was developed that could accurately quantify the amounts of substrate and product simultaneously. To improve the robustness of the assay, deuterated K-26 (**4.5**) was also synthesized (Figure 4-06) using the same methodology as for K-26¹³ and used as an internal standard.

In short, *N*-acetyl isoleucine (**4.1**) was synthesized by reacting *d*₃-labeled acetyl chloride and unlabeled isoleucine according to literature precedent.¹⁹ Labeled *N*-acetyl isoleucine (**4.1**) was coupled to *O*-benzyl-L-tyrosine benzyl ester (**4.2**) and the resulting ester was saponified resulting in *d*₃-*N*-Ac-Ile-Tyr(Bzl)-COOH (**4.3**). Standard peptide coupling conditions¹⁴ were employed in coupling AHEP diethyl ester (**2.3**) to **4.3**. In a two-step deprotection process, benzyl groups were removed from the resultant tripeptide by catalytic hydrogenation and ethyl groups were removed by reaction with iodotrimethylsilane. This synthetic route resulted in a mixture of two major diastereomers of *d*₃-K-26. The diastereomer containing (R)-AHEP (same stereochemistry as biogenic K-26) was separated via C18 chromatography and used as an internal standard in our tandem MS assay.

Specifically, cell-free extract was incubated with des-acetyl K26 and AcCoA. After 4 hours, internal standard *d*₃-K-26 was added at 10nM final concentration and the protein was removed using 3000 MWCO centrifugal filter devices. To accurately measure the concentration of dAck26 and K-26 present, three CID reactions were monitored. Transitions 492→216, 534→216, and 537→216 gave information about the relative amounts of dAck26, K-26 and *d*₃-

K-26, respectively (Figure 4-07). The use of internal standard d_3 -K-26 provided accuracy, precision, and assay ruggedness. The presence of K-26 (and not an artifact) could be verified by co-elution with the internal standard. In addition, since the concentration of the internal standard was known and constant, quantitative data about the amount of K-26 formed and the level of N-acetyltransferase activity detected could be obtained.

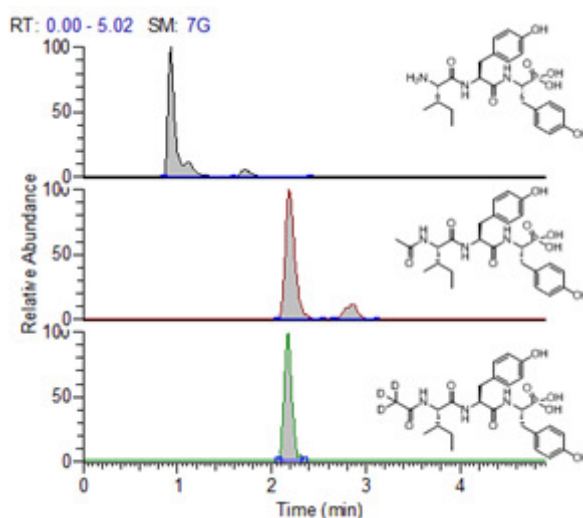


Figure 4-07. HPLC chromatogram for NAT activity detection

With a sensitive and accurate assay in hand, we proceeded with cell lysis and fractionation of the organism's proteome. Purification of NAT was initiated by preparing a cell-free extract of *Actinomyces K-26* in Tris buffer via sonication. Due to low production levels of K-26 and its biosynthetic enzymes, a large amount of cells (40-60g) had to be used. Nucleic acids were removed using streptomycin sulfate. Different ammonium sulfate cuts were tested for NAT activity and it was initially determined that the majority of NAT activity was

precipitated between 40 and 65% saturation. The pellet was resuspended in minimum amount of general enzyme buffer (GEB, see materials and methods for details) and salts were removed by dialysis. The resulting dialysate was further separated by anion exchange using Q-sepharose. Fractions were analyzed by monitoring absorbance at 280nm and activity was evaluated using the MS assay described above (Figure 4-08).

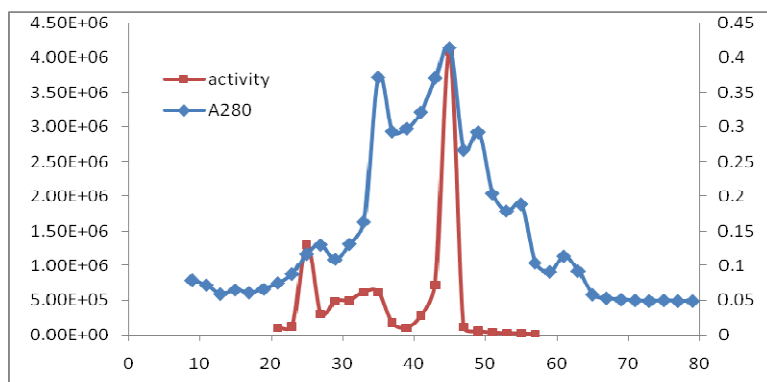


Figure 4-08. Anion exchange (Qsepharose) separation (blue: A280, red: NAT activity)

The most active fractions (43-46) were pooled and concentrated by bringing the combined fractions to 75% saturation with ammonium sulfate, recovering the proteins by centrifugation and resuspending the resulting pellet in buffer. This concentrated high salt protein solution was loaded on a hydrophobic interaction chromatography (HIC) column using an FPLC. Unfortunately, HIC was unsatisfactory as it suffered from poor resolution and low recovery (Figure 4-09). Unable to further purify the active fractions due to low protein and activity

levels, we resorted to mass spectrometry based proteomics analysis. The purification procedure is outlined in detail in Figure 4-10.

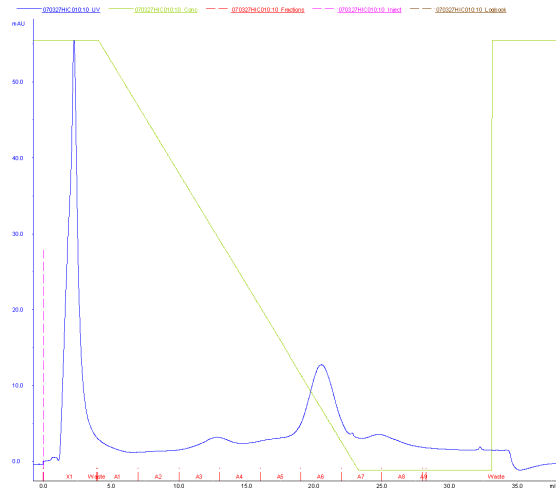


Figure 4-09. Hydrophobic interaction chromatogram

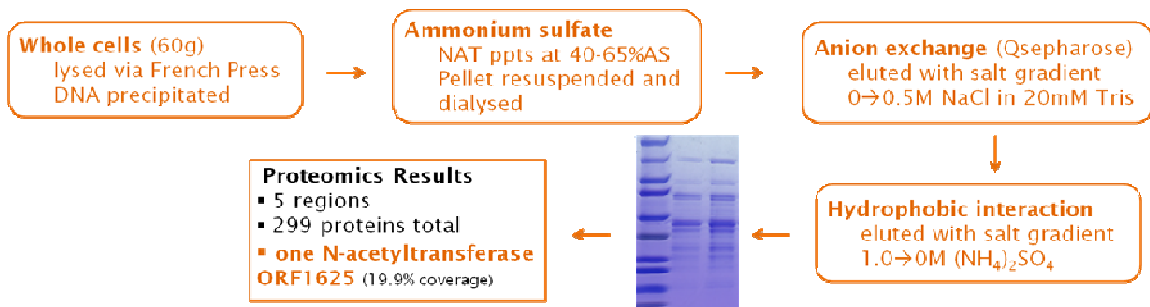


Figure 4-10. Purification outline (first attempt)

This first attempt to purify the K-26 NAT, resulted in only one hit, ORF1625 with 19.9% coverage. To validate this hit, purification was repeated with one change; instead of using HIC to further purify the active Q-sepharose fractions, size exclusion was employed. The most active fractions were analyzed by proteomics. ORF1625 was present in the sample with 33.3% coverage.

However, there were two more NATs, ORF0987 (39.1% coverage) and ORF0319 (42.0%). (Figure 4-11) These inconclusive results necessitated repeating the purification.

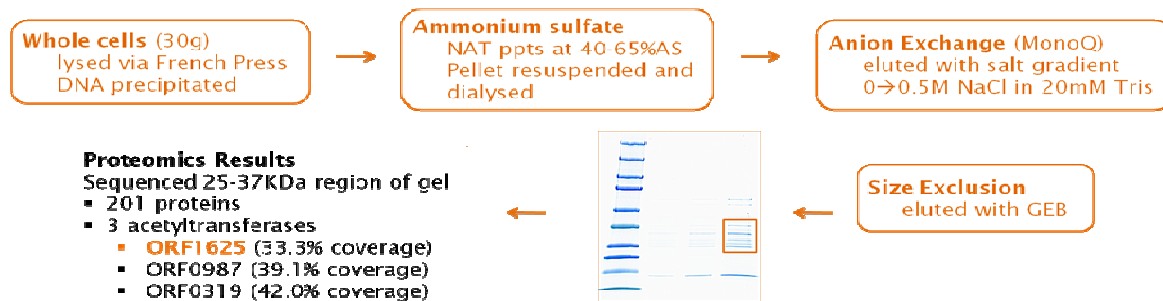


Figure 4-11. Purification outline (second attempt)

Once again, purification was initiated by preparing a cell-free extract of *Actinomyces K-26* in Tris buffer via sonication. Nucleic acids were removed using streptomycin sulfate. NAT activity was precipitated between 40 and 65% ammonium sulfate saturation. The pellet was resuspended in minimum amount of GEB and salts were removed by dialysis. The resulting dialysate was further separated by anion exchange using Q-sepharose at pH 8. The Q-sepharose chromatogram (Figure 4-12) was significantly different than before (Figure 4-08) but for consistency, fractions 43-46 (although not the most active) were purified further.

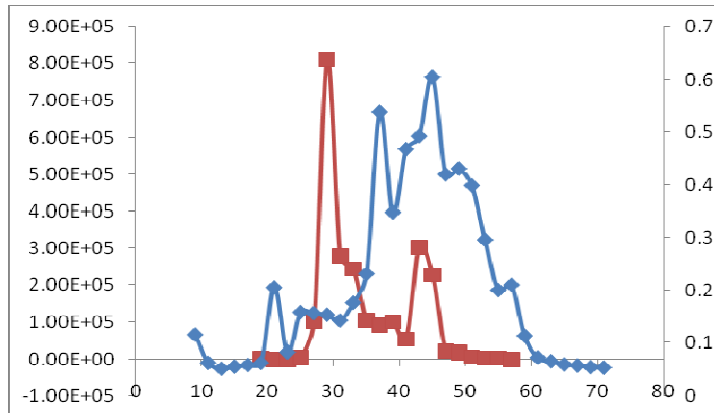


Figure 4-12. Anion exchange separation (blue: A280, red: NAT activity)

In detail, fractions 43-46 were pooled and brought to 75% saturation with ammonium sulfate. Proteins were recovered by centrifugation and resuspending the resulting pellet in piperazine buffer at pH 6. The resuspended proteins were dialyzed to remove salts and further purified by anion exchange on an FPLC. The elution buffer was at pH 6, a pH much lower than what had been used for anion exchange before. The most active fractions were pooled and concentrated as described above. The concentrated protein fractions were loaded on a HIC column and eluted with a buffer containing 5% glycerol for protein stability. The most active fraction was analyzed by proteomics and ORF1625 was the only NAT present in the sample (Figure 4-13).

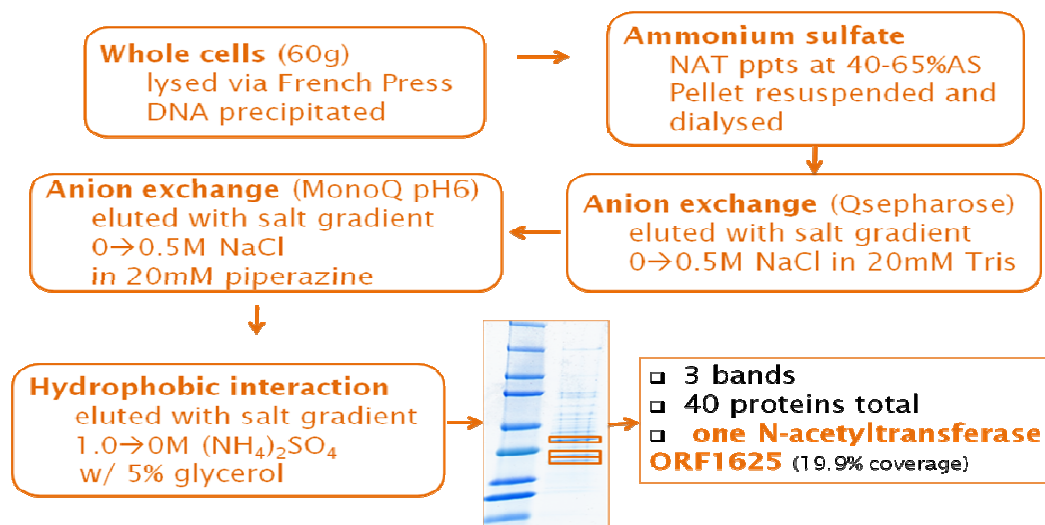


Figure 4-13. Purification outline (third attempt)

Fractions 29-31 (most active) from the Q-sepharose separation (Figure 4-12) were also combined and concentrated. The combined fractions were further purified by a second anion exchange step at pH 6. The active fractions of this step were loaded on a size exclusion column. The most active fraction of the size exclusion separation was analyzed by proteomics and the NAT's present were identified as ORF5577 (16.9% coverage), ORF6213 (16.6% coverage) and ORF5817 (62.3% coverage).

To further study ORF1625 and its NAT activity, it was amplified by PCR and cloned into pET28 for expression as N-terminal hexahistidine tagged protein in *E. coli* by Dr. Yu Du. The heterologously expressed NAT was purified by affinity chromatography and its acetyltransferase activity was evaluated using the MS based assay described above. It was found that the protein encoded by ORF1625 could only acetylate dAck26 (**2.6a**) with a 0.07% conversion yield and its diastereomer (**2.6b**) with a 0.01% yield. These results indicated that ORF1625

is not very efficient at acetylating dAck26, it may not be a dedicated NAT for K-26 biosynthesis or perhaps the presence of a hexahistidine tag was affecting its activity.

ORF5817 was also amplified by PCR and cloned into pET28 for expression as N-terminal hexahistidine tagged protein in *E. coli* by Dr. Yu Du, because of its high coverage. The heterologously expressed NAT was purified by affinity chromatography and its acetyltransferase activity was evaluated using the MS based assay described above. ORF5817 was able to acetylate dAck26 (**2.6a**) with a 0.01% yield but no acetylation of dAck6 diastereomer (**2.6b**) was observed. Low acetyltransferase activity for ORF5817 may be due to the presence of a hexahistidine tag. Perhaps, the conditions used to test for NAT activity are not optimal or ORF5817 is a general amino acid *N*-acetyltransferase that can acetylate a variety of peptide substrates but not just dAck26.

Another proteomics hit of interest was ORF6213. Although it was only present in an active protein sample once with low coverage (16.6%), inspection of its location in the K-26 producer genome revealed that it was part of a cluster coding for an NRPS (Table 4-1).

Table 4-1. ORF6193-6218, a candidate K-26 biosynthetic cluster

ORF06193	6386968	6386255	transcriptional regulator, GntR family
ORF06194	6387194	6388228	glucosamine-fructose-6-phosphate aminotransferase
ORF06195	6388283	6388519	ptsG-like product
ORF06196	6388532	6388978	putative phosphotransferase enzyme IIA component ypqE
ORF06197	6389069	6389305	HPr-like protein crh
ORF06198	6389313	6391067	phosphoenolpyruvate-protein phosphotransferase
ORF06200	6393733	6392375	peptide synthetase 1
ORF06202	6394249	6394770	surfactin synthetase subunit 3
ORF06203	6394770	6399860	peptide synthetase ScpsB
ORF06204	6399857	6400771	pyoverdine sidechain peptide nonribosomal III, L-Thr-L-Ser component
ORF06205	6400775	6404275	dimodular nonribosomal peptide nonribosomal
ORF06206	6404399	6405229	probable macrolide efflux protein, putative
ORF06207	6405317	6407008	putative membrane attached glycosyl hydrolase
ORF06208	6407078	6407770	haloacid dehalogenase-like hydrolase, putative
ORF06209	6407803	6408309	conserved hypothetical protein
ORF06210	6408397	6409035	haloacid dehalogenase-like hydrolase, putative
ORF06211	6409247	6409954	phosphoglycolate phosphatase
ORF06212	6411496	6410039	hydrolase, alpha/beta fold family domain protein
ORF06213	6412152	6411607	acetyltransferase, GNAT family family
ORF06214	6412326	6414236	acetoacetyl-CoA synthase
ORF06215	6415778	6414249	putative
ORF06217	6416202	6421010	Bacterial NAD-glutamate dehydrogenase superfamily
ORF06218	6422111	6421071	conserved hypothetical protein

ORF6213 was further investigated; it was amplified by PCR and cloned into pET28 for expression as N-terminal hexahistidine tagged protein in *E. coli*. The heterologously expressed NAT was purified by affinity chromatography and its acetyltransferase activity was evaluated using the MS based assay described above. When ORF6213 was incubated with either diastereomer of dAck26 (**2.6a** or **2.6b**), there was a 5% conversion to the corresponding diastereomer of K-26 (**1.17a** and **1.17b**, respectively). To determine if ORF6213 was a specific NAT for K-26 biosynthesis, it was also incubated with Ile-Tyr-Tyr and Ile-Tyr. Both

peptides were acetylated, the former with 55% yield and the latter with 11% yield. These results suggest that although ORF6213 is very efficient at acetylating dAck26 it is not specific. The fact that it can also acetylate Ile-Tyr-Tyr and Ile-Tyr indicates that it may be a general NAT, not dedicated to K-26 biosynthesis but able to recognize peptides containing N-terminal isoleucine.

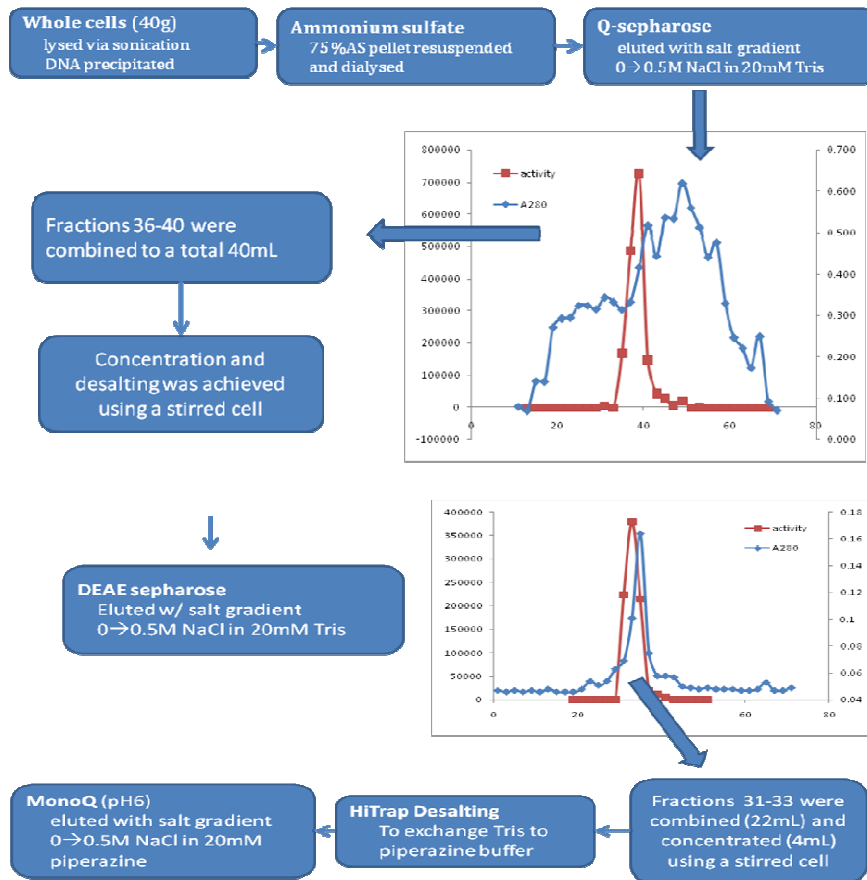


Figure 4-14. Purification outline (fourth attempt)

In a last attempt (Figure 4-14) to identify the NAT involved in K-26 biosynthesis, purification was initiated by preparing a cell-free extract of *Actinomyces K-26* in Tris buffer by sonication. Nucleic acids were removed

using streptomycin sulfate. NAT activity was precipitated with a single ammonium sulfate cut at 75% ammonium sulfate saturation. The pellet was resuspended in minimum amount of GEB and salts were removed by dialysis. The resulting dialysate was further separated by anion exchange using Q-sepharose at pH 8. The most active fractions from the Q-sepharose separation were combined, concentrated and desalted using a stirred cell. The desalted combined fractions were further purified by weak anion exchange on a DEAE-sepharose column at pH 8. The most active protein fractions were concentrated and desalted using a stirred cell and proteins were separated further using a third anion exchange step on a MonoQ at pH 6. The gradual purification of NAT as well as the decrease in protein concentration is illustrated in Figure 4-15. The most active fraction of the last anion exchange separation was analyzed by proteomics and the NAT's present were identified as ORF5577 and ORF5817.



Figure 4-15. SDS-PAGE analysis of NAT purification. (lane1: marker, lane2: streptomycin sulfate supernatant, lane3: 75%AS pellet resuspended and dialysed, lane4: Q-sepharose active fractions, lane5: DEAE-sepharose active fractions, lane6: monoQ active fraction)

In conclusion, using a combination of a NAT MS based assay and proteomics analysis, we have identified a few NAT's that are able to acetylate dAck26. After cloning and expressing these proteins in *E. coli*, we further investigated their activity. ORF6213 seemed to be the most efficient at acetylating dAck26. The location of ORF6213 in the K-26 producer's genome supports the hypothesis that it might be the NAT involved in K-26 biosynthesis. However, it is not specific. ORF6213 is highly efficient at acetylating Ile-Tyr-Tyr and Ile-Tyr, as well. These results are inconclusive. It is possible that the NAT is not part of the gene cluster and that the cells utilize a general peptide NAT to acetylate dAck26. In addition, we have recently discovered that the genome has several gaps. The gene coding for the NAT of interest may be in one of these gaps, in which case proteomics analysis could not identify it. Current efforts are focused on filling in the gaps of the genome and identifying the K-26 biosynthetic cluster by isolation of another enzyme, such as the C-P bond forming enzyme or the NRPS.

Materials and Methods

Organic Synthesis

General. All reagents were commercially available and used without purification unless specified otherwise. Thin-layer chromatography was carried out on precoated silica gel (60 F₂₅₄) plates obtained from EMD. Silica gel (35-70 *µm*) from EM Science was used for column chromatography. All NMR spectra

were acquired on a Bruker DPX-300 or a Bruker DRX-400 instrument unless otherwise noted. Low resolution mass spectral data was obtained on a Thermofinnigan (San Jose, C A) TSQ® Quantum triple quadrupole mass spectrometer equipped with a standard electrospray ionization source. Amino acids and all peptide coupling reagents were purchased from EMD. Labeled acetyl chloride was obtained from Sigma-Aldrich.

Synthesis of d_3 -N-acetyl isoleucine (4.1). Isoleucine (0.5 g, 3.8 mmol) was dissolved in 4N NaOH solution (2.85 mL, 11.4 mmol) and chilled to 0°C. d_3 -Acetyl chloride (0.30 mL, 4.2 mmol) was added in 5 batches over 50 min. The reaction was acidified to pH 1.5 with concentrated HCl and extracted with ethyl acetate. The combined organic layers were dried over MgSO₄ and concentrated resulting in a white solid. The solid was recrystallized from ethyl acetate/hexane to give *N*-acetyl isoleucine (270.8 mg, 40.3%). ¹H NMR (DMSO) δ 0.80-0.84 (m, 6H), 1.10-1.23 (m, 1H), 1.31-1.42 (m, 1H), 1.66-1.72 (m, 1H), 4.09-4.16 (m, 1H), 7.94-7.97 (m, 1H) ¹³C NMR δ 11.3, 15.6, 21.5 (m), 24.7, 36.4, 55.2, 169.4, 173.2 ESI-MS (M+H⁺): 177.1.

Synthesis of d_3 -N-acetyl-L-isoleucyl-O-benzyl-L-tyrosine methyl ester. *N*-acetyl-L-isoleucine (4.1, 270 mg, 1.53 mmol), *O*-benzyl-L-tyrosine methyl ester hydrochloride (4.2, 469.8mg, 1.46mmol) and *N*-hydroxybenzotriazole (246.5 mg, 1.61 mmol) were dissolved in 15 mL anhydrous dimethylformamide and cooled to 0°C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl) carbodiimide hydrochloride (308.6 mg, 1.61 mmol) and 2,4,6-trimethylpyridine (407 μL, 3.07mmol), stirred at 0 °C for one hour and at room temperature

overnight. Ethyl acetate (375 mL) was added to the reaction which was extracted with 1N HCl (3×75mL), 1N NaHCO₃ (3×75mL) and brine (3×75mL). The organic fraction was dried over MgSO₄ and evaporated to give white solid (595 mg, 91.9%). ¹H NMR (CDCl₃, 400MHz) δ 7.36-7.23 (m, 5H), 6.96-6.94 (d, *J* = 8.6Hz, 2H), 6.84-6.82 (d, *J* = 8.6Hz, 2H), 6.20-6.18 (d, *J* = 7.8Hz, 1H), 6.02-6.00 (d, *J* = 8.6Hz, 1H), 4.96 (s, 2H), 4.79-4.72 (m, 1H), 4.22-4.18 (m, 1H), 3.67 (s, 3H), 3.03-2.93 (m, 2H), 1.74-1.68 (m, 1H), 1.44-1.38 (m, 1H), 1.09-1.01 (m, 1H), 0.83-0.78 (m, 6H); ¹³C NMR (CDCl₃, 100MHz) δ 171.5, 170.7, 169.8, 157.9, 136.9, 130.2, 128.5, 127.9, 127.6, 127.4, 115.0, 69.9, 57.5, 53.1, 52.3, 37.4, 37.0, 24.9, 22.6 (m), 15.1, 11.2. ESI-MS (M+H⁺): 444.2

Synthesis of d₃-N-acetyl-L-isoleucyl-O-benzyl-L-tyrosine (4.3). A mixture of *d₃-N-acetyl-L-isoleucyl-O-benzyl-L-tyrosine methyl ester* (595 mg, 1.34 mmol) in acetonitrile (80mL) and 3% Na₂CO₃ (120mL) was stirred overnight at room temperature. The reaction was extracted with hexane (3×160mL), acidified with 1N HCl to pH 3 and extracted with chloroform (3×190mL). The combined chloroform fractions were washed with brine (2×160mL), dried over MgSO₄ and evaporated to a white solid (461 mg, 80.2%). ¹H NMR (CDCl₃, 400MHz) δ 7.31-7.24 (m, 5H), 7.00-6.98 (d, *J* = 8.5Hz, 2H), 6.81-6.79 (d, *J* = 8.5Hz, 2H), 6.66-6.64 (d, *J* = 7.8Hz, 1H), 6.19-6.17 (d, *J* = 9.0Hz, 1H), 4.91 (s, 1H), 4.75-4.70 (q, *J* = 6.5Hz, 1H), 4.27-4.22 (t, *J* = 8, 8.6Hz, 1H), 3.09-3.04 (dd, *J* = 5.5, 14Hz, 1H), 2.89-2.83 (dd, *J* = 6.4, 14.1Hz, 1H), 1.68 (s, 1H), 1.41-1.35 (m, 1H), 1.03-0.98 (m, 1H), 0.80-0.76 (m, 6H); ¹³C NMR (DMSO, 100MHz) δ 173.2, 171.6, 169.3,

157.4, 137.6, 130.5, 130.0, 128.8, 128.1, 128.0, 127.9, 114.8, 69.5, 56.8, 54.0, 37.1, 36.2, 24.5, 22.8 (m), 15.6, 11.3. ESI-MS (M+H⁺): 430.2

Synthesis of diethyl N-(d₃-N-acetyl-L-isoleucyl-O-benzyl - L-tyrosyl)-1-amino-2-(4-benzyloxyphenyl)ethylphosphonate. N-acetyl-L-isoleucyl-O-benzyl-L-tyrosine (**4.3**, 219 mg, 0.51 mmol), diethyl 1-amino-2-(4-benzyloxyphenyl)ethylphosphonate (177.6 mg, 0.49 mmol) and N-hydroxybenzotriazole (82.7 mg, 0.54 mmol) were dissolved in 5 mL anhydrous dimethylformamide and cooled to 0 °C. The reaction was treated with 1-ethyl-3(3'-dimethylaminopropyl)carbodiimide hydrochloride (103.5 mg, 0.54 mmol) and 2,4,6-trimethylpyridine (71.6 µL, 0.54 mmol), stirred at 0 °C for one hour and at room temperature overnight. Ethyl acetate (125 mL) was added to the reaction which was extracted with 1N HCl (3×25mL), 1N NaHCO₃ (3×25mL) and brine (3×25mL). The organic fraction was dried over MgSO₄, evaporated and the sample was further purified via flash column chromatography (4% methanol in chloroform) to yield a white solid (287.4 mg, 75.7%). ¹H NMR (CDCl₃, 500MHz) δ 7.44-7.31 (m, 10H), 7.17-7.00 (m, 4H), 6.93-6.78 (m, 5H), 6.36-6.19 (m, 2H), 5.08-4.98 (m, 4H), 4.74-4.68 (m, 2H), 4.24-4.00 (m, 5H), 3.19-3.14 (m, 1H), 3.04-2.94 (m, 1H), 2.89-2.70 (m, 2H), 1.81-1.79 (m, 1H), 1.43-1.24 (m, 7H), 1.08-1.02 (m, 1H), 0.91-0.76 (m, 6H); ¹³C NMR (CDCl₃, 100MHz) δ 171.6, 171.3, 170.2, 157.5, 157.4, 137.0, 136.9, 130.4, 130.1, 130.0, 129.4, 129.3, 128.8, 128.7, 128.4, 127.8, 127.7, 127.4, 127.3, 127.2, 114.7, 114.6, 114.5, 114.4, 69.7, 69.5, 62.6, 62.4, 60.3, 57.6, 53.5, 46.1, 38.4, 38.1, 37.5, 34.8, 29.6, 25.9, 25.2, 22.8 (m), 21.0, 16.6, 16.4, 15.3, 14.1, 11.3; ³¹P NMR (CDCl₃, 202MHz) δ 24.9, 24.7. ESI-MS (M+H⁺): 774.5

Synthesis of diethyl N-(N-acetyl-L-isooleucyl-L-tyrosyl)-1-amino-2-(4-hydroxy phenyl)ethylphosphonate (4.4). To a suspension of diethyl *N*-(*d*₃-*N*-acetyl - L - isooleucyl - O - benzyl - L -tyrosyl)-1-amino-2-(4-benzyloxyphenyl)ethyl phosphonate (287.4 mg, 0.37 mmol) in 5mL methanol, palladium on activated carbon (10%Pd, 29mg) was added and the mixture was allowed to stir overnight under a hydrogen atmosphere. The suspension was filtered through Celite and the filtrate was evaporated to yield an off-white solid (199.6 mg, 82.1%). ¹H NMR (MeOD, 500MHz) δ 7.10-7.02 (m, 4H), 6.73-6.66 (m, 4H), 4.60-4.50 (m, 2H), 4.19-4.08 (m, 5H), 3.10-2.55 (m, 2H), 1.70-1.61 (m, 1H), 1.39-1.32 (m, 7H), 1.10-1.03 (m, 1H), 0.89-0.70 (m, 6H); ¹³C NMR (MeOD, 125MHz) δ 171.8, 171.7, 171.6, 155.8, 155.7, 129.9, 129.7, 129.5, 127.5, 127.4, 127.3, 114.8, 114.6, 114.5, 62.9, 62.7, 57.9, 54.0, 37.0, 36.1, 33.6, 24.3, 21.0 (m), 15.3, 15.2, 14.3, 9.8; ³¹P NMR (MeOD, 202MHz) δ 25.5. ESI-MS (M+H⁺): 595.3.

Synthesis of N-(N-acetyl-L-isooleucyl-L-tyrosyl)-1-amino-2-(4-hydroxy phenyl)ethyl phosphonic acid (4.5, d₃-K-26). To a suspension of diethyl *N*-(*N*-acetyl-L-isooleucyl-L-tyrosyl)-1-amino-2-(4-hydroxyphenyl)ethylphosphonate (25.2 mg, 0.04 mmol) in 2mL acetonitrile and 0.2mL thioanisole at 0°C, iodotrimethylsilane (1 mL) was added dropwise and the reaction was allowed to proceed at 0°C for 3 hours. The solvent was removed under reduced pressure to give a brown oily residue of the phosphonate trimethylsilylester which was hydrolysed by treatment with 8:2 methanol/water (10mL). After evaporation of the solvents, the residue was dissolved in methanol and the two diastereomers of *d*₃-K-26 were separated via preparative HPLC using a linear water-acetonitrile

gradient (95:5 to 60:40 H₂O:CH₃CN over 30min) containing 0.1% trifluoroacetic acid. Evaporation of the HPLC solvents yielded **4.5** (5.4 mg, 25%) ¹H NMR (D₂O, 500MHz) δ 7.06-7.05 (d, *J* = 8.4Hz, 2H), 7.01-6.99 (d, *J* = 8.4Hz, 2H), 6.70-6.69 (d, *J* = 8.4Hz, 2H), 6.69-6.68 (d, *J* = 8.4Hz, 2H), 4.52-4.49 (dd, *J* = 4.0, 10.4Hz, 1H), 4.17-4.12 (t, *J* = 12.5Hz, 1H), 3.86-3.84 (d, *J* = 8.55Hz, 1H), 3.12-3.09 (d, *J* = 14.3Hz, 1H), 2.97-2.94 (dd, *J* = 4.1, 14.1, 1H), 2.63-2.55 (m, 2H), 1.49 (s, 1H), 1.14-1.12 (m, 1H), 0.92-0.86 (m, 1H), 0.70-0.67 (t, *J* = 7.4Hz, 3H), 0.51-0.50 (d, *J* = 6.8Hz, 3H); ¹³C NMR (D₂O, 151MHz) δ 173.5, 172.3, 171.7, 154.0, 153.8, 130.6, 130.2, 129.9, 129.8, 128.4, 115.1, 58.1, 53.9, 50.3-49.4 (d, ¹*J*_{CP} = 147Hz), 36.8, 35.8, 34.6, 24.2, 21.5 (m), 14.4, 9.9; ³¹P NMR (D₂O, 202MHz) δ 19.31. ESI-MS (M+H⁺): 539.2.

NAT purification

Production cultures of Actinomycete K-26 (prepared as described in Chapter III, p. 86) were harvested by centrifugation (3450 x g, 30min, 4°C). The cell paste was washed with 20% glycerol and the mycelia were frozen in liquid nitrogen for storage at -80°C. Purification was initiated by suspending frozen mycelia in GEB (general enzyme buffer: 50mM Tris, 10% glycerol, 1mM DTT, 0.2mM PMSF, pH 7) at 1g cells/mL buffer. Cells were ruptured at 0°C by ultrasonication at 50% duty cycle, power level 7 for 5 min and centrifuged (16300 x g, 30min, 4°C). Nucleic acids were precipitated by adding 5% streptomycin sulfate (at 1/5 volume), shaking at 0°C for 15min and centrifuging. Ammonium sulfate was added in portions over 10 min to the desired final saturation and

stirred for 1hr at 4°C and centrifuged. The protein pellets were redissolved in minimum volume of GEB and dialysed overnight versus 3 liters of GEB.

Q-sepharose. The dialysed solution from above was loaded on a Q-sepharose (Sigma-Aldrich) column equilibrated with 20mM Tris, pH 8 at 0.5mL/min using a peristaltic pump and eluted using a linear salt gradient 0→0.5M NaCl in 20mM Tris at pH 8. Active fractions were combined and concentrated by ammonium sulfate precipitation as above or using an Amicon stirred cell.

Hydrophobic Interaction chromatography. A high salt protein solution was loaded on a HiTrap Phenyl HP column (GE) equilibrated with 50mM Tris, 1M (NH₄)₂SO₄, 5% glycerol, pH 7 using an Akta FPLC at a 1mL/min flow rate. Proteins were eluted using a linear salt gradient 1.0→0M (NH₄)₂SO₄ in 50mM Tris containing 5% glycerol at pH 7.

Size exclusion chromatography. Protein solutions were loaded on a sephadex (GE) column equilibrated with GEB at 0.5mL/min and eluted with GEB.

MonoQ. Desalted protein solutions were loaded on a MonoQ (GE) column equilibrated with 20mM piperazine pH 6 buffer using an Akta FPLC at 1mL/min flow rate. Proteins were eluted using a linear gradient 0→0.5M NaCl in 20mM piperazine at pH 6.

DEAE-sepharose. Desalted protein solutions were loaded on DEAE-sepharose (Sigma-Aldrich) equilibrated with 20mM Tris, pH 8 buffer at 0.5mL/min using a peristaltic pump. Proteins were eluted using a linear gradient 0→0.5M NaCl in 20mM Tris at pH 8.

Enzyme assay. Protein fractions (30 μ L) were incubated with 1mM AcCoA and 300 μ M peptide substrate (dAcK26, IYY or IY) in 20mM Tris, 2mM MgCl₂, pH 8 at a total reaction volume of 90 μ L at 30°C for 4hours. Activity was evaluated using one of the methods below. Protein was removed using 3000 MWCO centrifugal filter devices, before analysis using the mass spectrometry based method.

Colorimetric assay for NAT detection. The method outlined by Brooke and coworkers¹⁵ was used with minor modifications. Specifically, a 100 μ L reaction was quenched with guanidine hydrochloride solution (6.4M guanidine hydrochloride, 0.1M Tris, pH 7.3, 25 μ L) containing 5mM DTNB and the absorbance at 405 nm was measured on a plate reader within 5min. Reactions without AcCoA, substrate or protein were used as controls. The amount of CoA produced was determined from a standard curve.

Fluorometric assay for NAT detection. The method outlined by Parvari and coworkers¹⁶ was used with minor modifications. Specifically, 20 μ L samples were removed from the reaction mixture and mixed with 20 μ L of 0.4mM CPM stock solution in isopropanol. One milliliter of 1% Triton X-100 was added to each sample. The excitation wavelength was 390nm, and the emission at 473nm was measured using a BIO-TEK Synergy HT plate reader.

Mass spectrometry based assay for NAT detection. Mass spectrometry was performed using ThermoFinnigan (San Jose, CA) TSQ® Quantum triple quadrupole mass spectrometer equipped with a standard electrospray ionization source outfitted with a 100- μ m I.D. deactivated fused Si capillary. Data

acquisition and spectral analysis were conducted with Xcalibur™ Software, version 1.3, from ThermoFinnigan (San Jose, CA), on a Dell Optiplex GX270 computer running the Microsoft® Windows 2000 operating system. The source spray head was oriented at an angle of 90° to the ion-transfer-tube. Nitrogen was used for both the sheath and auxiliary gas. The sheath and auxiliary gases were set to 33 and 14 (arbitrary units) respectively. Samples were introduced by HPLC. A Surveyor® Autosampler and a Surveyor® MS Pump from ThermoFinnigan (San Jose, CA) were used. The injection volume was 10µL. Peptides were separated using a Luna™ minibore 3µm C18(2) column (2.0mm × 5cm) with an isocratic mobile phase consisting of 90% water, 10% acetonitrile and 10mM ammonium acetate. The flow rate was 0.2mL/min.

The mass spectrometer was operated in the negative ion mode and the electrospray needle was maintained at 4200V. The ion transfer tube was operated at -35V and 350°C. The tube lens voltage was set to -150V. Source CID (offset voltage between skimmer and the first ion guide, Q00) was used at 15V. The selected reaction monitoring (SRM) mode was used. Ions were collisionally activated with argon at an indicated pressure of 1.4mT. The mass-spectral resolution was set to a peak width (full width at half maximum, FWHM) of 0.50u and 0.50u for precursor and product ions respectively. Mass transitions at the specified collision energy (m/z 492→216; 20eV), (m/z 534→216; 20eV) and (m/z 537→216; 35eV) were monitored for dACK26, K-26, and d₃-K-26, respectively. Transitions (m/z 293→180; 20eV), (m/z 335→180; 20eV), (m/z 456→180; 20eV), and (m/z 498→180; 20eV) were monitored for Ile-Tyr, Ac-Ile-

Tyr, Ile-Tyr-Tyr, and Ac-Ile-Tyr-Tyr, respectively. The scan width for product ions was 1.000u and the cycle time for each ion was 0.15 seconds. The electron multiplier gain was set to 2×10^6 . Data were acquired in profile mode.

Proteomics. Protein was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, Tris-HCl 10-20% gradients, Bio-Rad Laboratories). The gels were stained with Coomassie brilliant blue. Proteomics analysis was performed by Dr. Lisa Zimmermann in the Vanderbilt Proteomics Core Facility.

References

1. Yamato, M.; Koguchi, T.; Okachi, R.; Yamada, K.; Nakayama, K.; Kase, H.; Karasawa, A.; Shuto, K., K-26, a Novel Inhibitor of Angiotensin-I Converting Enzyme Produced by an Actinomycete K-26. *Journal of Antibiotics* **1986**, 39, (1), 44-52.
2. Hirayama, N.; Kasai, M.; Shirahata, K., Structure and Conformation of a Novel Inhibitor of Angiotensin-I Converting Enzyme - a Tripeptide Containing Phosphonic Acid. *International Journal of Peptide and Protein Research* **1991**, 38, (1), 20-24.
3. Ohuchio, S.; Kurihara, K.; Shinohara, A.; Takei, T.; Yoshida, J.; Amano, S.; Miyadoh, S.; Matsushida, Y.; Somura, T.; Sezaki, M., Studies on New Angiotensin Converting Enzyme Inhibitors, SF2513A,B and C, Produced by *Streptosporangium nondiastaticum*. *Sci. Reports of Meiji Seika Kaisha* **1988**, 27, 46 - 54.
4. Koguchi, T.; Yamada, K.; Yamato, M.; Okachi, R.; Nakayama, K.; Kase, H., K-4, a Novel Inhibitor of Angiotensin-I Converting Enzyme Produced by *Actinomadura-Spiculospora*. *Journal of Antibiotics* **1986**, 39, (3), 364-371.
5. Ntai, I.; Manier, M. L.; Hachey, D. L.; Bachmann, B. O., Biosynthetic origins of C-P bond containing tripeptide K-26. *Organic letters* **2005**, 7, (13), 2763-2765.
6. Ntai, I.; Phelan, V. V.; Bachmann, B. O., Phosphonopeptide K-26 biosynthetic intermediates in *Astrosporangium hypotensionis*. *Chemical Communications* **2006**, (43), 4518-4520.
7. Blodgett, J. A.; Thomas, P. M.; Li, G.; Velasquez, J. E.; van der Donk, W. A.; Kelleher, N. L.; Metcalf, W. W., Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide. *Nat Chem Biol* **2007**, 3, (8), 480-5.
8. Woodyer, R. D.; Shao, Z.; Thomas, P. M.; Kelleher, N. L.; Blodgett, J. A.; Metcalf, W. W.; van der Donk, W. A.; Zhao, H., Heterologous production of fosfomycin and identification of the minimal biosynthetic gene cluster. *Chem Biol* **2006**, 13, (11), 1171-82.

9. Sieber, S. A.; Marahiel, M. A., Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem Rev* **2005**, 105, (2), 715-38.
10. Challis, G. L.; Ravel, J.; Townsend, C. A., Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chemistry & Biology* **2000**, 7, (3), 211-224.
11. Clardy, J.; Walsh, C., Lessons from natural molecules. *Nature* **2004**, 432, (7019), 829-37.
12. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003**, 422, (6928), 198-207.
13. Ntai, I.; Bachmann, B. O., Identification of ACE pharmacophore in the phosphonopeptide metabolite K-26. *Bioorganic & medicinal chemistry letters* **2008**, 18, (10), 3068-3071.
14. Carpino, L. A.; El-Faham, A.; Albericio, F., Efficiency in Peptide Coupling: 1-Hydroxy-7-azabenzotriazole vs 3,4-Dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine. *Journal of Organic Chemistry* **1995**, 60, (11), 3561-4.
15. Brooke, E. W.; Davies, S. G.; Mulvaney, A. W.; Pompeo, F.; Sim, E.; Vickers, R. J., An approach to identifying novel substrates of bacterial arylamine N-acetyltransferases. *Bioorganic & medicinal chemistry* **2003**, 11, (7), 1227-34.
16. Parvari, R.; Pecht, I.; Soreq, H., A microfluorometric assay for cholinesterases, suitable for multiple kinetic determinations of picomoles of released thiocholine. *Analytical biochemistry* **1983**, 133, (2), 450-6.
17. Walker, K.; Ketchum, R. E.; Hezari, M.; Gatfield, D.; Goleniowski, M.; Barthol, A.; Croteau, R., Partial purification and characterization of acetyl coenzyme A: tax-4(20),11(12)-dien-5 α -ol O-acetyl transferase that catalyzes the first acylation step of taxol biosynthesis. *Archives of biochemistry and biophysics* **1999**, 364, (2), 273-9.
18. Fukuchi, J.; Kashiwagi, K.; Takio, K.; Igarashi, K., Properties and structure of spermidine acetyltransferase in *Escherichia coli*. *The Journal of biological chemistry* **1994**, 269, (36), 22581-5.

19. Chenault, H. K.; Dahmer, J.; Whitesides, G. M., Kinetic Resolution of Unnatural and Rarely Occurring Amino-Acids - Enantioselective Hydrolysis of N-Acyl Amino-Acids Catalyzed by Acylase-I. *Journal of the American Chemical Society* **1989**, 111, (16), 6354-6364.

CHAPTER V

SYNTHESIS OF APOPTOLIDIN ANALOGUES VIA PRECURSOR DIRECTED BIOSYNTHESIS

Introduction

Apoptolidin A (**1.19**) is a glycosylated macrolactone initially isolated by Hayakawa and coworkers in 1997 from the fermentation of soil dwelling actinomycete *Nocardioopsis sp* using a screen designed to identify specific apoptosis inducers.¹ The molecular formula of apoptolidin A was determined as C₅₈H₉₆O₂₁ by high-resolution FAB-MS and extensive NMR studies were used to obtain its molecular structure.^{1, 2} Apoptolidin has a highly unsaturated 20-membered macrolide with a side chain containing a 6-membered hemiketal. A novel sugar, 6-deoxy-4-*O*-methyl-L-glucose, is attached to O9 and a disaccharide consisting of L-olivomycose and D-oleandrose is appended via O27. Apoptolidin A was shown to induce apoptosis at concentrations as low as 10nM in oncogene-containing E1A-transformed glial cells.¹ In contrast, normal glial cells were insensitive to apoptolidin A even at micromolar concentrations. It was found that apoptolidin was among the top 0.1% most selective agents tested in the National Cancer Institute's 60-cell line screening panel.³ This high selectivity of apoptolidin for cancer cells over normal cells makes it an attractive lead for the development of new anticancer agents with minimal side effects.

Khosla and coworkers demonstrated that apoptolidin A is a mitochondrial F₀F₁-ATPase inhibitor (IC₅₀ = 0.7μM).³ These results provided evidence that

inhibition of mitochondrial F_0F_1 -ATPase may be the basis of apoptolidin's biological activity. However, structure-activity relationship (SAR) studies undertaken by Wender and coworkers proved that *in vitro* inhibition of mitochondrial F_0F_1 -ATPase is not an accurate predictor of the *in vivo* activity of apoptolidin derivatives, the selective induction of apoptosis in cancer cells.⁴ These findings suggest that there might be an alternate target of apoptolidin and its mode of action may be more complex than originally suggested.

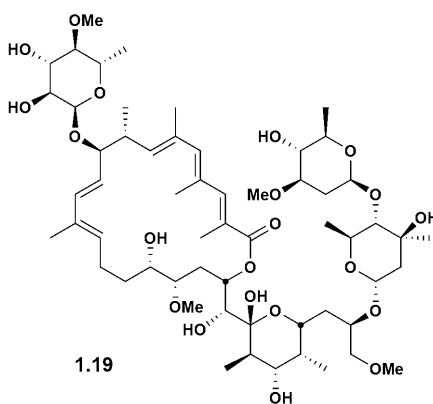


Figure 5-01. Structure of apoptolidin A

Our goal is to synthesize a chemical probe of apoptolidin to be used in identifying its cellular target(s) by a combination of techniques, including proteomics analysis. Although apoptolidin can be isolated from the fermentation broth of *Nocardioopsis sp.* in high yields (>100mg/L),^{1, 5} it is only readily modified within the sugar region. However, it has been shown that the sugars are critical to apoptolidin's activity^{6, 7} and any modification to introduce a tag could lead to loss of activity. While total syntheses of apoptolidin have been reported,^{6, 8} chemical glycosylation can be extremely challenging due to issues of protecting

groups and stereoselectivity. In contrast, enzymatic glycosylation does not require the use of protective groups and it is regio- and stereoselective. We plan to chemically synthesize azido apoptolidinone (**5.1**) and using the glycosylating machinery of the organism to prepare 16-azido apoptolidin (Figure 5-02). This analogue of apoptolidin can be readily conjugated to a biotin and/or a fluorophore by click chemistry⁹ to be used as a chemical probe to identify apoptolidin's cellular targets.

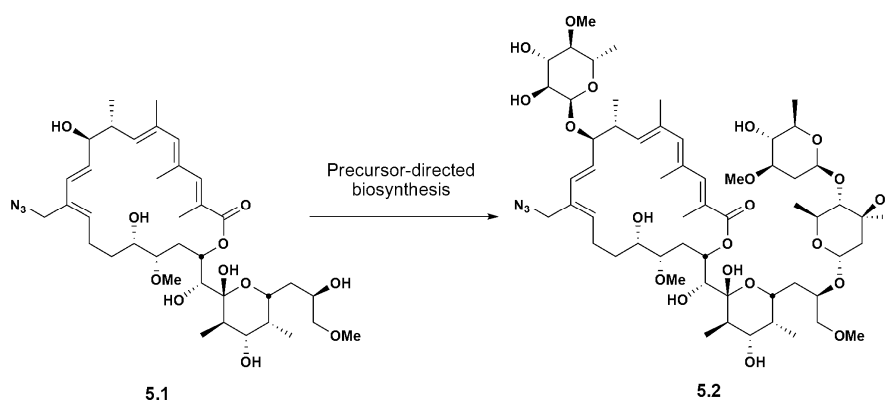


Figure 5-02. Enzymatic glycosylation of azido apoptolidinone

Instead of isolating the glycosyltransferases involved in apoptolidin biosynthesis, we have decided to use a whole cell approach, namely precursor-directed biosynthesis (PDB). PDB is defined as “the derivatization of a secondary metabolite by feeding biosynthetic precursor analogues to the fermentation broth of producing organisms”.¹⁰ While the biosynthesis of apoptolidin has not been established, inspection of its structure can result in a hypothetical biosynthetic pathway that follows the typical polyketide biosynthesis by a polyketide synthetase (PKS). In short, the PKS is responsible for activation of the starter

unit, followed by conjugation of extender malonyl or methylmalonyl units, modification by ketoreductase and dehydratase domain and finally macrolactonization and release of the polyketide core from the PKS.¹¹ Then, the sugar units are attached to the macrolactone by glycosyltransferases (Figure 5-03). However, there has been evidence to support that perhaps the linear precursor of the macrolactone (of a different natural product) may be the substrate of glycosylation.¹²

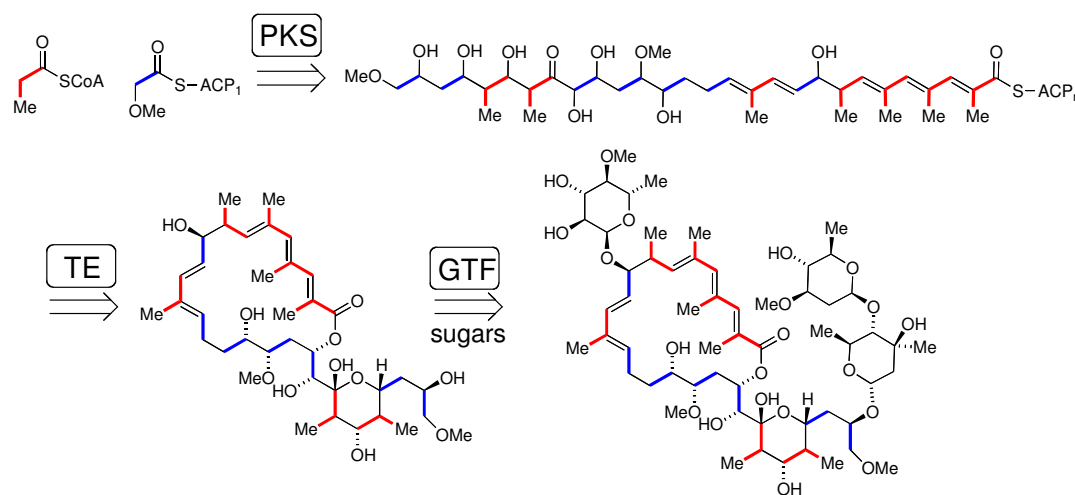


Figure 5-03. Hypothetical biosynthesis of apoptolidin

This chapter will focus on the development of a mass spectrometric detection method for apoptolidin and analogues and its use in our efforts to synthesize an apoptolidin chemical probe. We were able to use this MS method to assess the substrate flexibility of the glycosylating enzymes involved in apoptolidin biosynthesis and identify new derivatives of apoptolidin produced by precursor-directed biosynthesis.

Results and Discussion

The gene cluster coding for apoptolidin remains unknown. Thus, genetic manipulation to produce apoptolidin analogues is not a possibility. However, Omura and coworkers have developed a biotransformation strategy by which they were able to glycosylate unnatural aglycones using whole cells of macrolide producing organisms in the presence of cerulenin.¹³⁻¹⁵ Cerulenin is a known fatty acid biosynthesis and polyketide synthetase inhibitor. The effect of cerulenin can be described as a chemical “knockout” of the PKS as the megasynthetase’s function is inhibited. Our first step in evaluating the feasibility of this strategy to our system was to determine a concentration of cerulenin that would inhibit apoptolidin production without killing the producing organism.

The production protocol outlined by Seto and coworkers¹ was scaled down and cerulenin (0.2mM/day and 0.1mM/day for 5 days) was added to the production culture of *Nocardioopsis* sp. as a DMSO solution. To the control culture was added the same amount of DMSO. The cell mass was monitored on a daily basis as a measure of fatty acid biosynthesis inhibition. The cell mass seemed to be unaffected by the presence of cerulenin (Figure 5-04). The apoptolidin production inhibition was determined by LC/MS analysis of the culture’s extract on an ESI triple quadrupole mass spectrometer (Figure 5-05). It was determined that 0.2mM cerulenin per day based on culture volume was an appropriate concentration for reducing the production of apoptolidin to below 10% relative to the control.

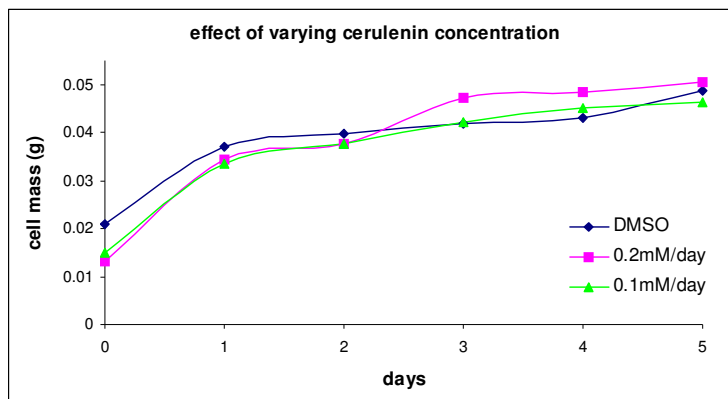


Figure 5-04. Effect of cerulenin on cell mass

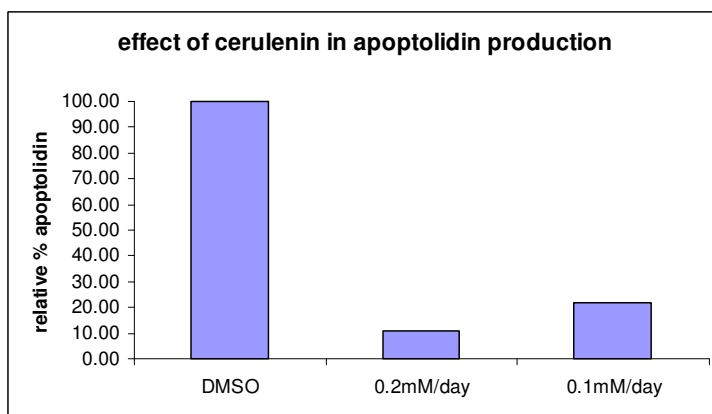


Figure 5-05. Effect of cerulenin in apoptolidin production

Before moving forward with the feeding of aglycone analogues, we wanted to verify that the addition of synthetic apoptolidinone would restore the production of apoptolidin. This would indicate that the glycosyltransferases can function independently of the PKS. To that effect, apoptolidinone A (**1.26**) was synthesized as previously described¹⁶ and added to growing cultures of *Nocardioopsis* sp. in combination with cerulenin. The apoptolidin production level

was compared to an uninhibited control culture and a culture supplemented with cerulenin (Figure 5-06).

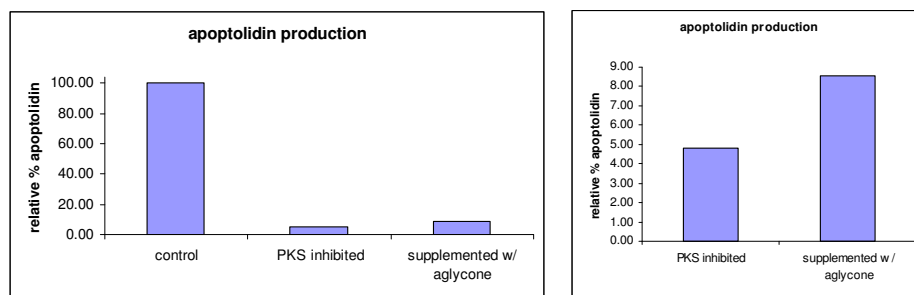
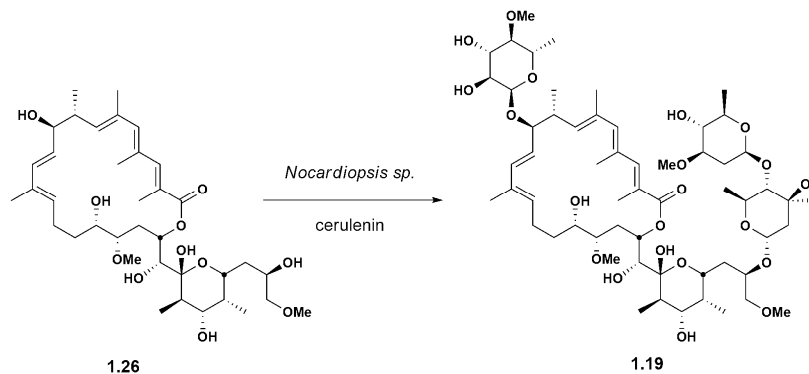


Figure 5-06. Enzymatic glycosylation of synthetic apoptolidinone A

Addition of synthetic apoptolidinone A in combination with cerulenin doubled the production of apoptolidin A in comparison to addition of cerulenin alone. However, the apoptolidin production levels in comparison to the control culture are extremely low (about 10%). After a closer look at the MS data from the feeding study, we have identified a new analogue of apoptolidin illustrated in figure 5-07.

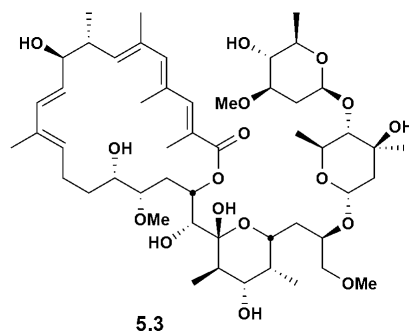


Figure 5-07. Structure of diglycosylated apoptolidinone A

Compound **5.3** differs from apoptolidin A only by the absence of sugar 6-deoxy-4-*O*-methyl-L-glucose at C9, while the disaccharide unit appended via O27 is still in place. Isolation of this compound explains the low production level of apoptolidin in the presence of synthetic apoptolidinone and cerulenin and also raises questions about the substrate of glycosylation. Specifically, these data may suggest that 6-deoxy-4-*O*-methyl-L-glucose is attached to the linear precursor of the macrolactone, while the disaccharide is added after release from the PKS. Diglycosylated apoptolidinone A (**5.3**) is currently being further purified in order to confirm its structure by NMR studies and evaluate its activity against cancer cells and F_0F_1 -ATPase, specifically.

To further validate this method of preparing apoptolidin analogues, we synthesized an unnatural aglycone, structurally very similar to apoptolidinone A (Figure 5-08). This macrolactone (**1.29**), initially named 6-normethyl apoptolidinone, differs from the natural aglycone by the absence of a methyl group on C6. The same synthetic methodology as described earlier¹⁶ was employed to obtain 15mg of **1.29**, which was introduced to growing cultures of *Nocardioopsis* sp. in the presence of cerulenin. The culture broth was extracted

with ethyl acetate and the concentrated extracts analyzed by LC/MS. A peak corresponding to the fully glycosylated aglycone (**1.22**, $m/z = 1132.6$) is observed. However, several peaks of that mass are present in the control culture (Figure 5-09).

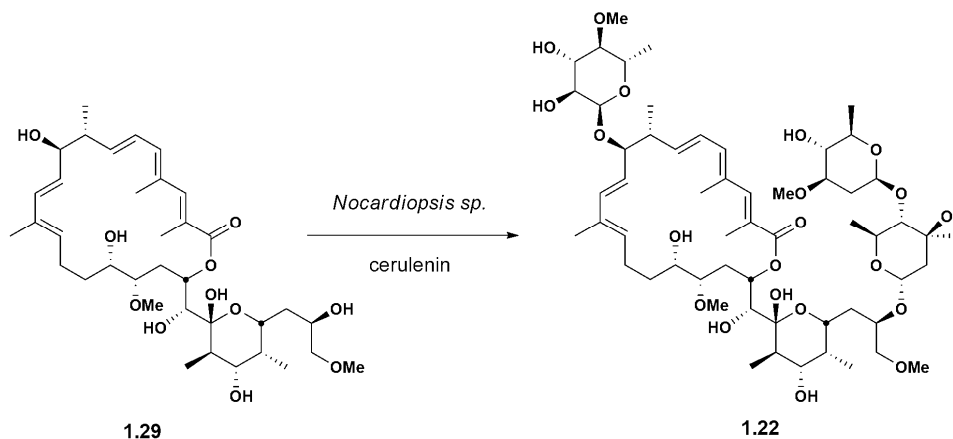


Figure 5-08. Glycosylation of 6-normethyl apoptolidinone (apoptolidinone D)

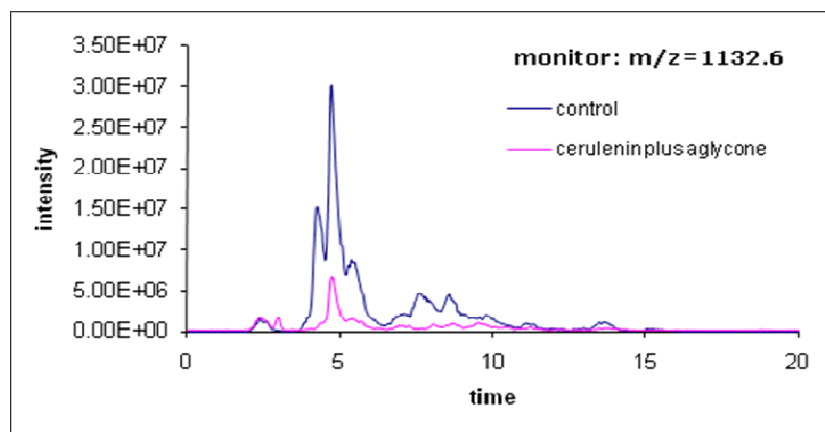


Figure 5-09. Chromatogram from 6-normethyl apoptolidinone feeding study

Initial speculations about these peaks suggested that they may correspond to naturally occurring analogues of apoptolidin that are missing one

of the methyl ethers. However, after completion of these experiments, Wender and coworkers reported the isolation of a new apoptolidin, apoptolidin D.¹⁷ Surprisingly, apoptolidin D (**1.29**) is missing a methyl group at the C6 position. Thus, the initially thought unnatural aglycone 6-normethyl apoptolidinone is a naturally occurring macrolactone, named apoptolidinone D by the Wender group. This aglycone is most likely biosynthesized by the same PKS as apoptolidinone A and their structural difference is due to incorporation of a malonyl extender unit instead of a methylmalonyl extender unit. This substrate flexibility of PKS is not uncommon and has been observed before.¹⁸

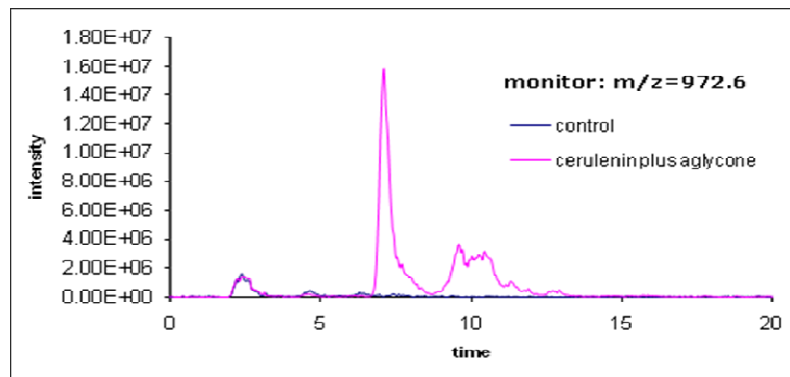


Figure 5-10. Detection of diglycosylated apoptolidinone D

LC/MS analysis (Figure 5-10) of the 6-normethyl apoptolidinone supplemented culture also clearly identified a new metabolite with $m/z = 972.6$ corresponding to the diglycosylated apoptolidinone D (**5.4**). Once again, as in the case of apoptolidinone A, the organism is able to attach the disaccharide at O27 but the sugar on C9 is missing (Figure 5-11).

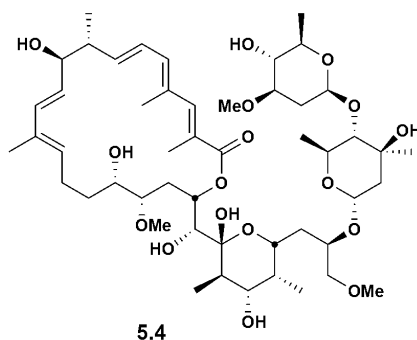


Figure 5-11. Structure of diglycosylated apoptolidinone D

Apoptolidin analogue **5.4** was further purified via normal phase HPLC by Dr. Victor Ghidu and extensive 1D and 2D NMR experiments (Figure 5-12) confirmed the structure hypothesized based on MS data.

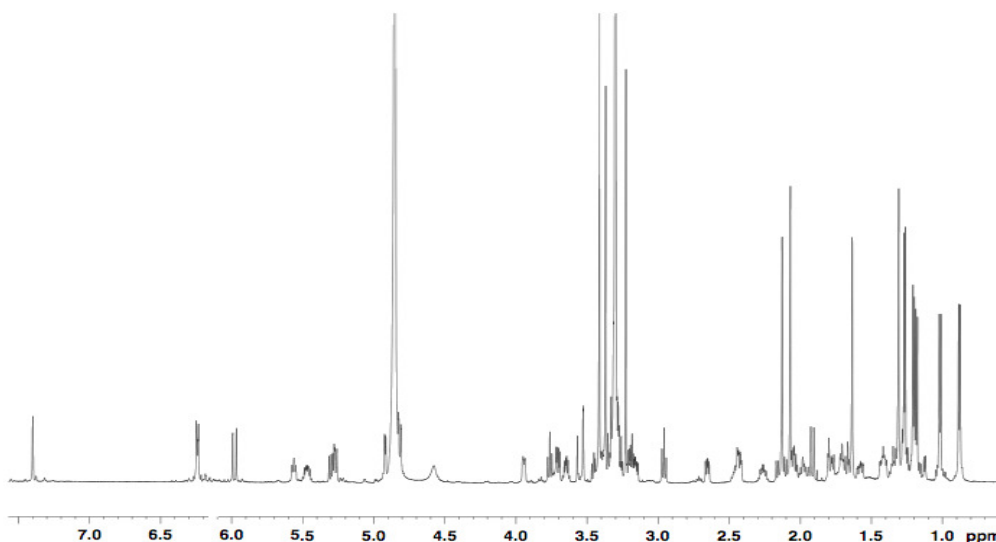


Figure 5-12. ^1H NMR of diglycosylated apoptolidinone D

The disaccharide decorated apoptolidinone D was tested against NCI H292 cells (human lung cancer cells) and its activity compared to apoptolidin A

by Dr. Aaron Jacobs. He was able to show that cells treated with either compound underwent apoptosis after 5-7 days. Apoptolidin A was determined to be more potent ($EC_{50} = 20\text{nM}$) than diglycosylated apoptolidinone D by a factor of 10 (Figure 5-13). However, it is noteworthy that the aglycone alone has $EC_{50} > 10\mu\text{M}$ in the same assay. While it is yet still under debate why only the disaccharide is added to the synthetic aglycones during our feeding studies, it appears that glycosylation at O27 is crucial for activity of the resulting apoptolidin analogues.

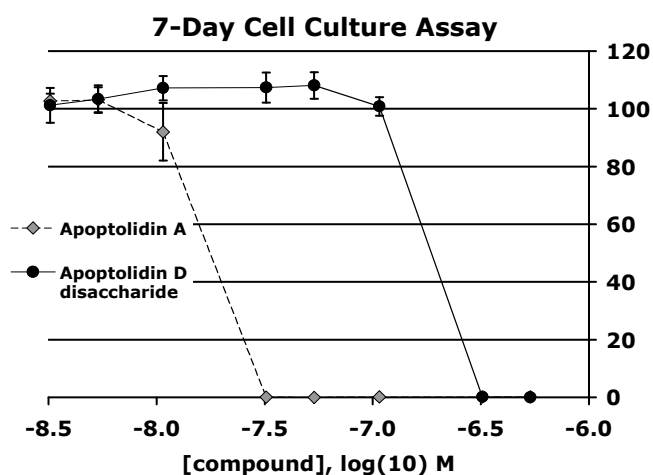


Figure 5-13. Apoptotic activity of diglycosylated apoptolidinone D

These preliminary results support our hypothesis that *Nocardioopsis* sp. can be used to afford apoptolidin analogues by glycosylation of synthetic aglycones. Synthesis of azido apoptolidinone (**5.1**) is currently being pursued by Jingqi Wang and upon completion, it will be added in combination with cerulenin to cultures of the apoptolidin producer. Based on the results from the aglycone

feeding studies described above, it is highly possible that complete glycosylation of the azido apoptolidinone may not be observed. However, we are confident that addition of the disaccharide on O27 will be sufficient to produce an apoptolidin analogue having nanomolar EC₅₀. The diglycosylated azido apoptolidinone (**5.5**, Figure 5-14) should be able to interact with the apoptolidin cellular targets despite missing the 6-deoxy-4-*O*-methyl-L-glucose.

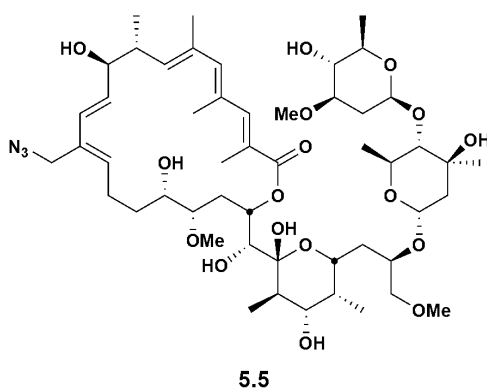


Figure 5-14. Structure of diglycosylated azido apoptolidinone

Azido apoptolidin (**5.2**), if isolated, and diglycosylated azido apoptolidinone can then be conjugated to a difunctional probe (Figure 5-15) using “click” azide-alkyne coupling chemistry as described by Cravatt and coworkers.^{9, 19, 20} The resulting trifunctional probe contains elements for target recognition (apoptolidin analogue), detection (e.g. rhodamine) and isolation (e.g. biotin).

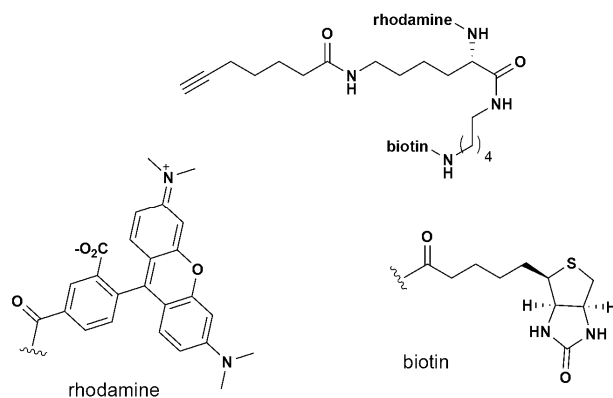


Figure 5-15. Difunctional probe containing terminal alkyne for conjugation via “click” chemistry⁹

Incubation of cancer cells with this trifunctional probe will allow isolation of target proteins via avidin affinity chromatography and visualization of the isolated protein by in-gel fluorescent scanning.⁹ MS analysis of the isolated proteins after tryptic digest and comparison of MS data to protein databases will identify the protein target of apoptolidin.

Materials and Methods

Fermentation. One hundred microliters of frozen glycerol stock of *Nocardiosis sp.* was plated onto a Petri dish containing Bennett’s medium and incubated at 30°C for 3 days. Bennett’s medium consisted of yeast extract, 1g/L; beef extract, 1g/L; NZ amine A (casein digest), 2g/L; glucose, 10g/L; and agar, 20g/L dissolved in distilled water and was adjusted to pH 7.0 before autoclaving. It was then poured into sterile Petri dishes (~25mL/dish).

The production protocol suggested by Hayakawa and coworkers¹ was followed. Specifically, fermentation was initiated by aseptically inoculating one loopful of mycelia grown on Bennett’s agar plates into a sterile 50-mL Falcon

tube containing 5mL seed medium. The seed cultures were incubated at 30°C for 3 days in a shaker incubator. The seed medium consisted of soluble starch, 10g/L; molasses, 10g/L; peptone, 10g/L; and beef extract, 10g/L dissolved in deionized water and was adjusted to pH 7.2 before autoclaving.

The 5 mL seed culture was transferred into a 2800-mL Fernbach flask containing 250mL production medium. The flask was incubated for 5 days at 30°C in a shaker incubator. The production medium consisted of glycerol, 20g/L; molasses, 10g/L; casamino acids, 5g/L; peptone, 1g/L; and calcium carbonate (CaCO₃), 4g/L dissolved in deionized water and was adjusted to pH 7.2 before autoclaving.

Feeding experiments. To determine the appropriate cerulenin concentration, 0.2mM (4mg/100mL production culture) and 0.1mM (2mg/100mL production culture) were dissolved in 1mL DMSO and administered separately to 100mL production culture through a sterile syringe filter every 24 hours starting at the time of inoculation for a total of 5 times. To the control culture 1mL DMSO (no cerulenin) was added in the same manner daily.

For the aglycone feeding studies, aglycones were added at the time of inoculation while cerulenin was pulse fed daily. Specifically, 7mg apoptolidinone A were dissolved in 1mL DMSO and added through a sterile syringe filter to a 100mL production medium at the time of inoculation while 0.2mM amount of cerulenin (4mg/100mL culture) was dissolved in 0.8mL DMSO and added through a sterile syringe filter every 24 hours for 5 days. To the control culture

1mL DMSO was added at the time of inoculation and 0.8mL DMSO in the same manner daily.

15mg apoptolidinone D were dissolved in 1mL DMSO and added through a sterile syringe filter to a 200mL production medium at the time of inoculation while 0.2mM amount of cerulenin (4mg/100mL culture) was dissolved in 1mL DMSO and added through a sterile syringe filter every 24 hours for 5 days. To the control culture 1mL DMSO was added at the time of inoculation and 0.8mL DMSO in the same manner daily.

Apoptolidin and apoptolidin analogues purification. The production culture was centrifuged at 4000rpm for 15 min. The supernatant was extracted with ethyl acetate (3 x 100mL) followed by TLC. The mycelia were stirred with acetone (100mL) for one hour and centrifuged at 4000rpm for 15 minutes. The acetone supernatant was extracted with ethyl acetate (3 x 50mL). The ethyl acetate extracts were combined and concentrated to ca. one milliliter by rotary evaporation.

Mass spectrometry. Mass spectrometry was performed using ThermoFinnigan (San Jose, CA) TSQ® Quantum triple quadrupole mass spectrometer equipped with a standard electrospray ionization source outfitted with a 100-µm I.D. deactivated fused Si capillary. Data acquisition and spectral analysis were conducted with Xcalibur™ Software, version 1.3, from ThermoFinnigan (San Jose, CA), on a Dell Optiplex GX240 computer running the Microsoft® Windows 2000 operating system. The source spray head was oriented at an angle of 90° to the ion-transfer tube. Nitrogen was used for both

the sheath and auxiliary gas. The sheath and auxiliary gases were set to 33 and 14 (arbitrary units) respectively.

Samples were introduced by HPLC. A Surveyor® Autosampler and a Surveyor® MS Pump from ThermoFinnigan (San Jose, CA) were used. The injection volume was 10µL. Crude apoptolidin extracts were separated using a Jupiter™ minibore 5µm C18 column (2.0mm × 15cm) with an isocratic mobile phase consisting of 65% water, 35% acetonitrile and 10mM ammonium acetate. The flow rate was 0.2mL/min.

The mass spectrometer was operated in the positive ion mode and the electrospray needle was maintained at 4200V. The ion transfer tube was operated at 35V and 342°C. The tube lens voltage was set to 85V. Source CID (offset voltage between skimmer and the first ion guide, Q00) was used at 15V.

The mass spectrometer was operated in full scan mode using Quad 1. The mass spectral resolution was set to a peak width of 0.70u (full width at half maximum, FWHM). Full scan spectra were acquired from m/z 600.0 to 1200.0 over 1.0 second. Data were acquired in profile mode. The electron multiplier gain was set to 3×10^5 .

Apoptolidin and analogues formed ammonium adducts and using the Xcalibur™ Software, data was digitally filtered for the m/z of interest (Figure 5-16).

data13 #790-832 RT: 15.37-16.18 AV: 43 SB: 190 16.97-18.72, 12.98-14.89 NL:
T: + p sid=-15.00 Q1MS [600.00-1500.00]

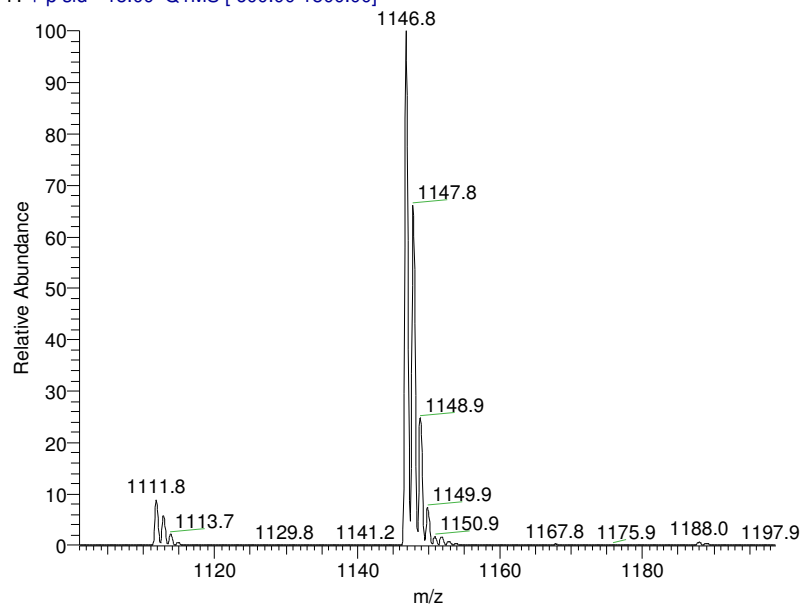


Figure 5-16. Mass spectrum of apoptolidin A

References

1. Kim, J. W.; Adachi, H.; Shin-ya, K.; Hayakawa, Y.; Seto, H., Apoptolidin, a new apoptosis inducer in transformed cells from *Nocardiosis sp.* *J Antibiot (Tokyo)* **1997**, 50, (7), 628-30.
2. Hayakawa, Y.; Kim, J. W.; Adachi, H.; Shin-ya, K.; Fujita, K. i.; Seto, H., Structure of Apoptolidin, a Specific Apoptosis Inducer in Transformed Cells. *J Am Chem Soc* **1998**, 120, (14), 3524-3525.
3. Salomon, A. R.; Voehringer, D. W.; Herzenberg, L. A.; Khosla, C., Understanding and exploiting the mechanistic basis for selectivity of polyketide inhibitors of F(0)F(1)-ATPase. *Proc Natl Acad Sci U S A* **2000**, 97, (26), 14766-71.
4. Wender, P. A.; Jankowski, O. D.; Longcore, K.; Tabet, E. A.; Seto, H.; Tomikawa, T., Correlation of F0F1-ATPase inhibition and antiproliferative activity of apoptolidin analogues. *Org Lett* **2006**, 8, (4), 589-92.
5. Wender, P. A.; Sukopp, M.; Longcore, K., Apoptolidins B and C: isolation, structure determination, and biological activity. *Org Lett* **2005**, 7, (14), 3025-8.
6. Nicolaou, K. C.; Li, Y.; Sugita, K.; Monenschein, H.; Guntupalli, P.; Mitchell, H. J.; Fylaktakidou, K. C.; Vourloumis, D.; Giannakakou, P.; O'Brate, A., Total synthesis of apoptolidin: completion of the synthesis and analogue synthesis and evaluation. *J Am Chem Soc* **2003**, 125, (50), 15443-54.
7. Salomon, A. R.; Zhang, Y.; Seto, H.; Khosla, C., Structure-activity relationships within a family of selectively cytotoxic macrolide natural products. *Org Lett* **2001**, 3, (1), 57-9.
8. Wehlan, H.; Dauber, M.; Mujica Feraud, M. T.; Schuppan, J.; Mahrwald, R.; Ziemer, B.; Juarez Garcia, M. E.; Koert, U., Total synthesis of apoptolidin. *Angew Chem Int Ed Engl* **2004**, 43, (35), 4597-601.
9. Adam, G. C.; Vanderwal, C. D.; Sorensen, E. J.; Cravatt, B. F., (-)-FR182877 is a potent and selective inhibitor of carboxylesterase-1. *Angew Chem Int Ed Engl* **2003**, 42, (44), 5480-4.

10. Thiericke, R.; Rohr, J., Biological variation of microbial metabolites by precursor-directed biosynthesis. *Nat Prod Rep* **1993**, 10, (3), 265-89.
11. Keating, T. A.; Walsh, C. T., Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr Opin Chem Biol* **1999**, 3, (5), 598-606.
12. Kao, C. L.; Borisova, S. A.; Kim, H. J.; Liu, H. W., Linear aglycones are the substrates for glycosyltransferase DesVII in methymycin biosynthesis: analysis and implications. *J Am Chem Soc* **2006**, 128, (17), 5606-7.
13. Omura, S.; Sadakane, N.; Tanaka, Y.; Matsubara, H., Chimeramycins: new macrolide antibiotics produced by hybrid biosynthesis. *J Antibiot (Tokyo)* **1983**, 36, (7), 927-30.
14. Sadakane, N.; Tanaka, Y.; Omura, S., Hybrid biosynthesis of derivatives of protylonolide and M-4365 by macrolide-producing microorganisms. *J Antibiot (Tokyo)* **1982**, 35, (6), 680-7.
15. Sadakane, N.; Tanaka, Y.; Omura, S., Hybrid biosynthesis of a new macrolide antibiotic by a daunomycin-producing microorganism. *J Antibiot (Tokyo)* **1983**, 36, (7), 921-2.
16. Wu, B.; Liu, Q.; Sulikowski, G. A., Total synthesis of apoptolidinone. *Angew Chem Int Ed Engl* **2004**, 43, (48), 6673-5.
17. Wender, P. A.; Longcore, K. E., Isolation, structure determination, and anti-cancer activity of apoptolidin D. *Org Lett* **2007**, 9, (4), 691-4.
18. Hardt, I. H.; Steinmetz, H.; Gerth, K.; Sasse, F.; Reichenbach, H.; Hofle, G., New natural epothilones from *Sorangium cellulosum*, strains So ce90/B2 and So ce90/D13: isolation, structure elucidation, and SAR studies. *J Nat Prod* **2001**, 64, (7), 847-56.
19. Adam, G. C.; Sorensen, E. J.; Cravatt, B. F., Chemical strategies for functional proteomics. *Mol Cell Proteomics* **2002**, 1, (10), 781-90.

20. Adam, G. C.; Sorensen, E. J.; Cravatt, B. F., Trifunctional chemical probes for the consolidated detection and identification of enzyme activities from complex proteomes. *Mol Cell Proteomics* **2002**, 1, (10), 828-35.

CHAPTER VI

SYNOPSIS AND CONCLUDING REMARKS

Natural products and mass spectrometry

Natural products have always played and will always play a significant role in the development of new therapeutic agents. Crude plant extracts have long been used as alternative folk medicines. In modern medicine, some of the best-selling drugs are natural products or natural product derivatives. About half of the drugs currently in clinical use are natural products or synthetic molecules based on natural product scaffolds. The drug discovery model in the pharmaceutical industry is focused on the high throughput screening of as many compounds as possible in order to find new structures and new activities for drug development. Recent advancements in MS, have made it an invaluable tool for high throughput screening of biological extracts. MS can give information not only about the molecular weight of a new natural product but also its structural features.

In addition, MS can be used to study the biosynthesis of a secondary metabolite. Probing the biosynthesis of an active natural product is of great importance because that information can be used to engineer new analogues with possibly more potent or even different activities. MS can be used not only to identify biosynthetic intermediates but also to study the enzymes involved. Enzyme activity, specificity/flexibility, mechanism and kinetics can all be studied

by MS. Finally, MS can also be used to study the mechanism of action of natural products, which is very important when trying to design a new drug.

The goal of this dissertation was to illustrate the utility of mass spectrometry in natural products studies through two examples: I) phosphonopeptide K-26 and II) macrolide apoptolidin.

Phosphonopeptide K-26

Despite the diverse and significant biological activities of carbon-phosphorus bond containing natural products, the biosynthetic pathways that lead to their formation are not well understood. Of interest to us is the C-P bond containing natural product, K-26. The potent angiotensin converting enzyme inhibitor K-26 is a naturally occurring *N*-acetylated tripeptide containing isoleucine, tyrosine and the non-proteinogenic amino acid (*R*)-1-amino-2-(4-hydroxyphenyl)ethyl phosphonic acid (AHEP). Extremely low production levels of K-26 (less than 10µg/L) required development of a highly sensitive method of detection. Isotopically labeled precursor incorporation experiments were designed to investigate the biosynthetic origin of AHEP and isotopic enrichment was calculated based on selected reaction monitoring (SRM) experiments. We have unambiguously identified tyrosine as an immediate precursor of AHEP and AHEP itself as a discrete precursor of K-26.

After establishing the basic building blocks of K-26 and AHEP, biochemical experiments were undertaken in order to identify the genes and enzymes responsible for K-26 biosynthesis. Specifically, attempts were made to

isolate the *N*-acetyltransferase involved in the biosynthesis of K-26. An SRM based method was developed for detecting acetylation of dAck26 and we proceeded with cell lysis and fractionation of the organism's proteome. Different chromatographic techniques were employed, such as ion exchange, hydrophobic interaction, and size exclusion. Unfortunately, due to low production levels of K-26 and its biosynthetic enzymes, a large amount of cells (40-60g) had to be used each time. In addition, activity was diminished after two or three chromatographic steps. For that reason, protein fractions were further separated using SDS-PAGE and regions on the gel were cut out, digested and sequenced. Sequences were compared to a database consisting of the organism's proteome and emphasis was placed on the hits for *N*-acetyltransferases. Hits that were repeated or were in interesting regions of the genome were cloned into pET28 and expressed as N-terminal hexahistidine tagged proteins. Their activity was evaluated using the established LC/MS assay.

ORF6213 seemed to be the most efficient at acetylating dAck26. The location of ORF6213 in the K-26 producer's genome supports the hypothesis that it might be the NAT involved in K-26 biosynthesis. However, it is not specific. It is possible that the NAT is not part of the gene cluster and that the cells utilize a general peptide NAT to acetylate dAck26. In addition, we have recently discovered that the genome has several gaps. The gene coding for the NAT of interest may be in one of these gaps, in which case proteomics analysis could not identify it. Current efforts are focused on filling in the gaps of the genome and

identifying the K-26 biosynthetic cluster by isolation of another enzyme, such as the C-P bond forming enzyme or the NRPS.

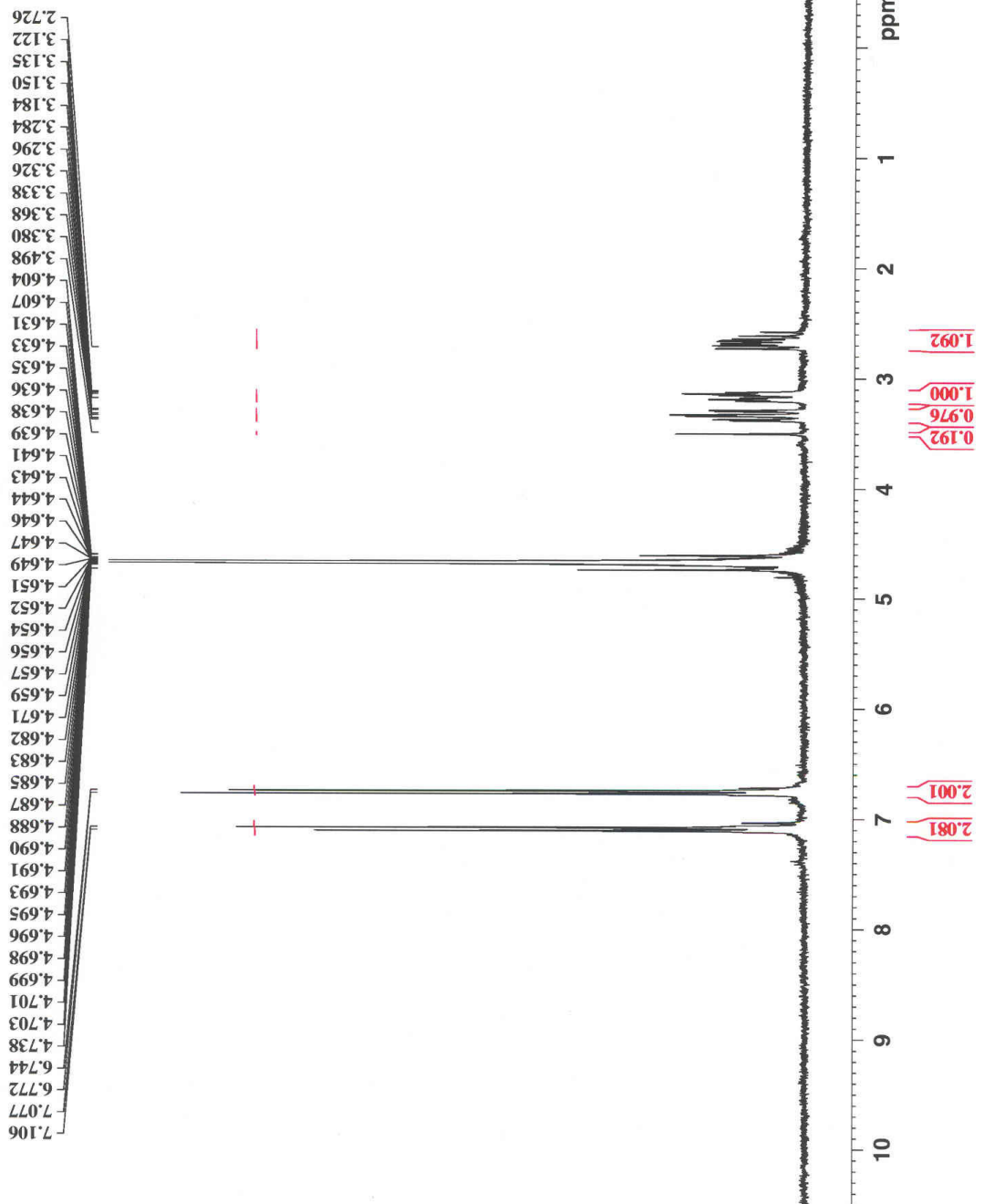
Macrolide apoptolidin

Apoptolidin is an anticancer agent produced by *Nocardioopsis sp* that induces cell death by a mitochondrial dependent apoptotic pathway. Several groups including Dr. Sulikowski's have attempted the total synthesis of this molecule. The sugars that are appended on the aglycone core seem to significantly increase the cytotoxicity of the molecule. Chemical glycosylation can be extremely challenging due to issues of protecting groups and stereoselectivity. In contrast, enzymatic glycosylation does not require the use of protective groups and it is regio- and stereoselective. The aglycone part of the molecule is synthesized by a polyketide synthetase (PKS) and after release it is glycosylated by a series of enzymes. We proposed using the glycosylating machinery of the organism to prepare analogues of apoptolidin by feeding synthetic aglycone analogues and detecting glycosylation by mass spectrometry. We plan to chemically synthesize azido apoptolidinone and using the glycosylating machinery of the organism to prepare 16-azido apoptolidin. This analogue of apoptolidin can be readily conjugated to a biotin and/or a fluorophore by click chemistry to be used as a chemical probe to identify apoptolidin's cellular targets.

APPENDIX A

NMR SPECTRA

IN-1-237
AHEP

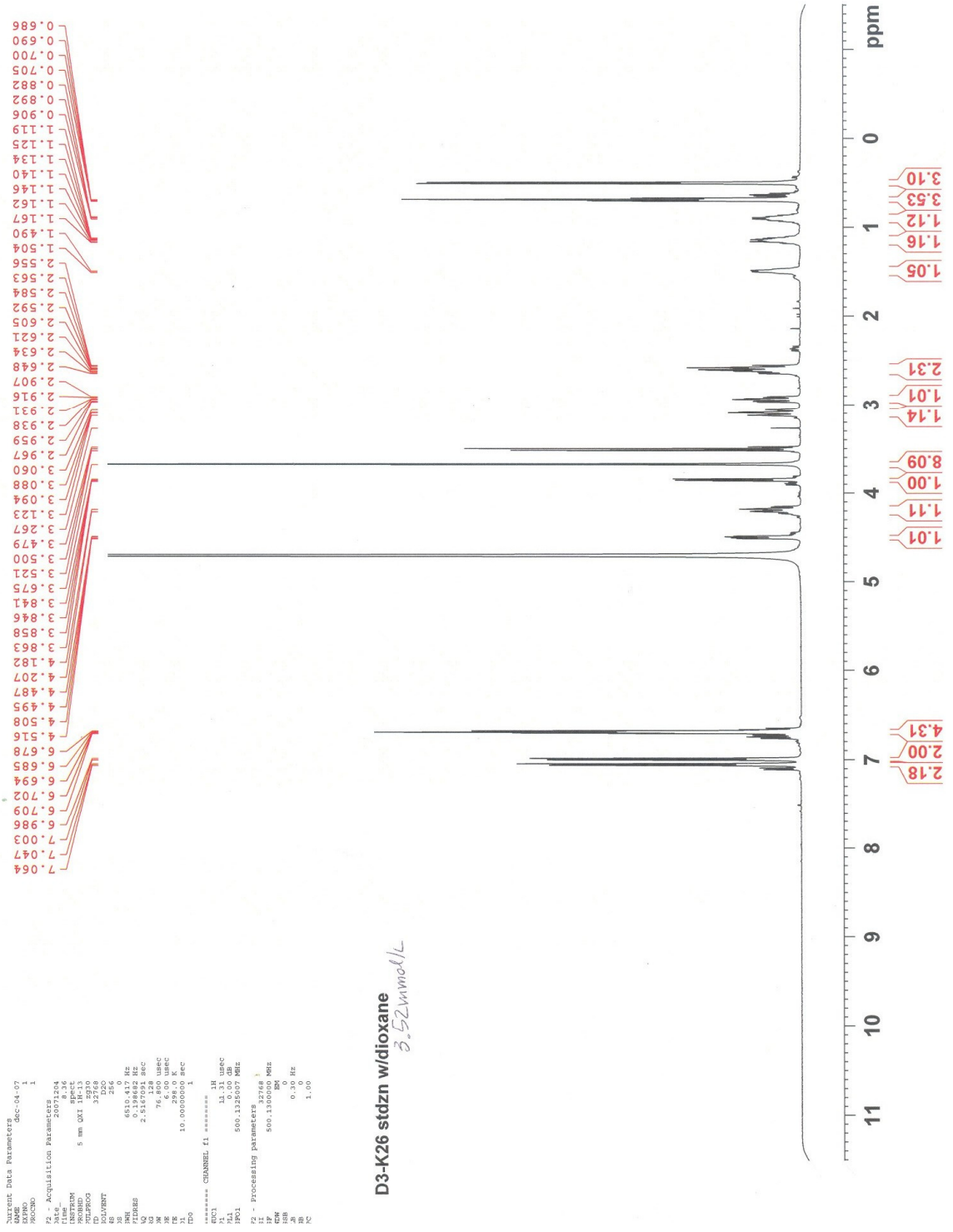


Current Data Parameters
 NAME: ap-21-05
 EXPNO: 1
 PROCNO: 1

F2 - Acquisition Parameters
 Date_: 20050421
 Time: 12.27
 Operator: M
 PROBRD: 5 mm QNP 1H/1
 PULPROG: zg30
 TD: 16384
 SOLVENT: D2O
 DS: 0
 SWH: 3906.250 Hz
 FIDRES: 0.238419 Hz
 AQ: 2.0972021 sec
 RG: 327.5
 DW: 128.000 usec
 DE: 6.00 usec
 TE: 0.0 K
 DCREST: 2.0000000 sec
 MCREST: 0.0150000 sec
 MCWRR: 0.0150000 sec

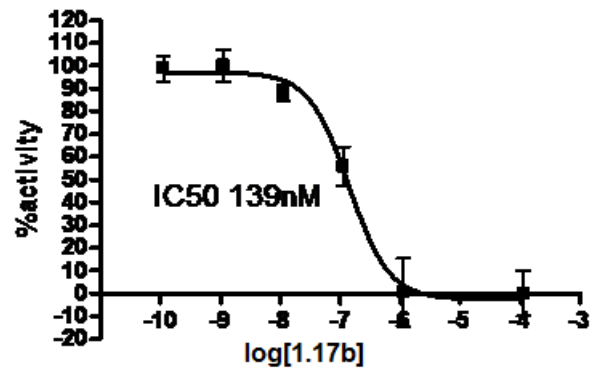
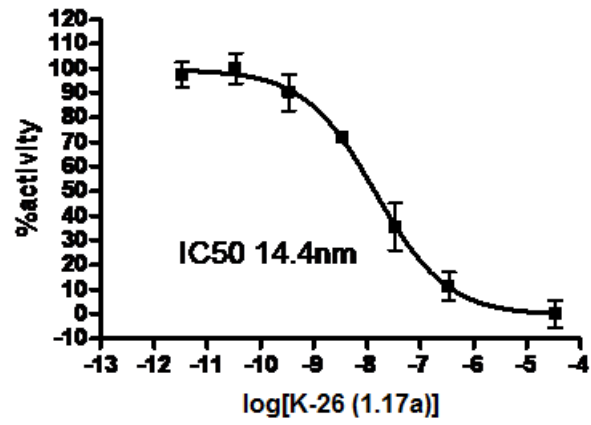
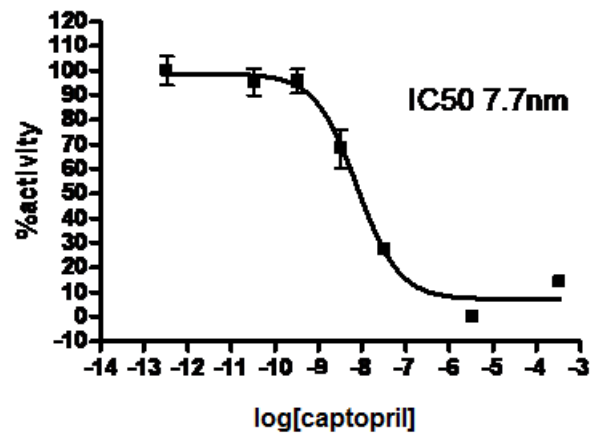
===== CHANNEL f1 =====
 NUCL1: 1H
 P1: 9.10 usec
 PL1: 0.00 dB
 SFO1: 300.1316000 MHz

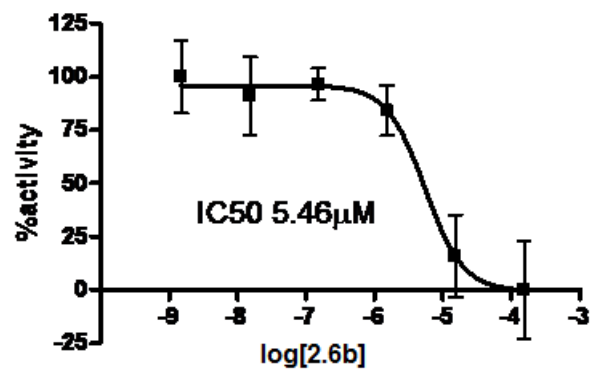
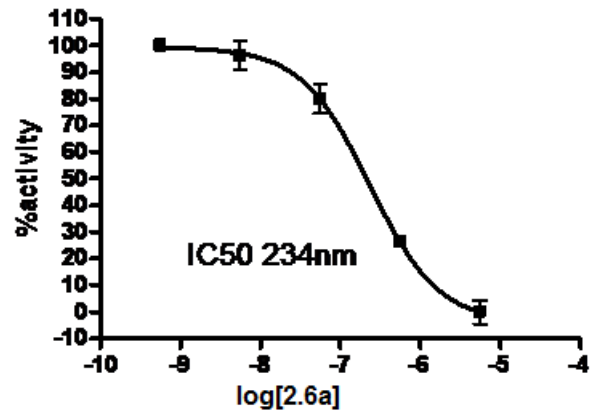
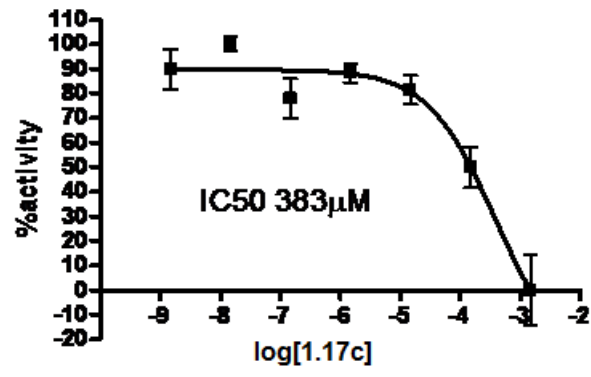
F2 - Processing parameters
 SI: 327.5
 SF: 300.1300800 MHz
 WDW: EM
 SSB: 0
 GB: 0.00 Hz
 PC: 1.00

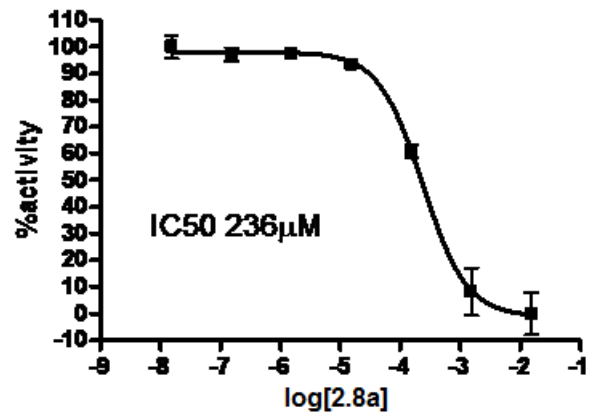
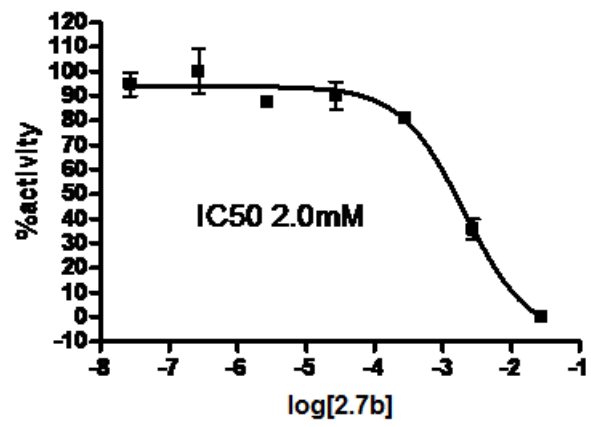
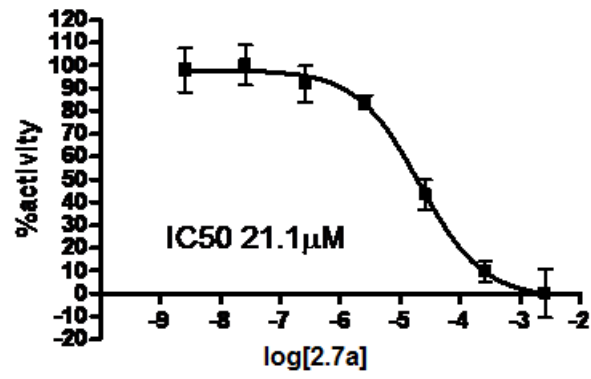


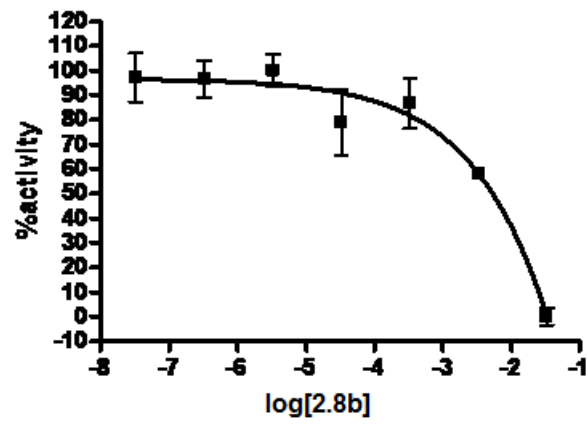
APPENDIX B

GRAPHS USED IN IC₅₀ CALCULATIONS









APPENDIX C

CALCULATIONS OF LABELED PRECURSOR'S INCORPORATION

Calculation of ^{15}N enrichment example. The tabulated numbers below (Calculation 1-4) resulted from the following calculations. Since all the calculations (1-4) are similar, the deconvolution of tandem MS data in order to determine ^{15}N enrichment in Tyr-AHEP (Calculation 4) exemplifies the method. In this example, the **theoretical isotope matrix** was constructed using the relative mass spectral ion intensities of transitions 534→379, 535→380, and 536→381 measured for an unenriched sample. In the absence of a pure standard for $^{15}\text{N}_1\text{-K-26}$ and $^{15}\text{N}_2\text{-K-26}$, we assumed that the theoretical distribution of unenriched K-26 would be as follows.

m/z	Theoretical Isotope Ratio (x100)		
	$^{15}\text{N}_0\text{-PA}$	$^{15}\text{N}_1\text{-PA}$	$^{15}\text{N}_2\text{-PA}$
534 -> 379	100.00	0.00	0.00
535 -> 380	16.89	100.00	0.00
536 -> 381	1.86	16.53	100.00
		1.86	16.17
			1.86

The **mole fraction matrix** was obtained by normalizing the relative ion intensities to the sum of the ion intensities for each species ($^{15}\text{N}_0$, $^{15}\text{N}_1$, and $^{15}\text{N}_2$, the tracer substances). The mole fraction matrix was transposed resulting in the **isotope design matrix**, in order that the final result would be along the same row instead of down the same column. This set of normalized mass spectral ion intensities can be used as a set of fixed coefficients in a system of linear equations that describe the relative ion abundances of the different tracers. The general form of the equation is

$$I_j = A_{i,j} \times B_i \quad \text{equation 1}$$

Where I_j is the intensity of transition j , $A_{i,j}$ is the normalized ion abundance of tracer substance i during transition j (isotope design matrix), and B_i is the unknown mole fraction of tracer substance i . The system has to include one equation for each tracer and each equation must contain one term for each tracer. This results in the following set of linear equations:

$$I_{534 \rightarrow 379} = 84.213B_{15N0} + 14.219B_{15N1} + 1.568B_{15N2}$$

$$I_{535 \rightarrow 380} = 0.000B_{15N0} + 84.467B_{15N1} + 13.962B_{15N2}$$

$$I_{536 \rightarrow 381} = 0.000B_{15N0} + 0.000B_{15N1} + 84.724B_{15N2}$$

In order to solve this set of linear equations, the **pseudoinverse matrix** was calculated according to the following equation.

$$P = (A^T \times A)^{-1} \times A^T \quad \text{equation 2}$$

Where P is the pseudoinverse matrix, A is the isotope design matrix and A^T is the transpose of A . Multiplying P with the **measured isotope ratio matrix** (normalized measured ion intensities) gives the **corrected isotopic abundances and ratio**. This calculation was performed for both unenriched and enriched samples to check for errors in the method. The corrected isotopic abundances were normalized to 100. The unenriched samples had corrected tracer: tracee ratio of 100:0:0 for $^{15}N_0$: $^{15}N_1$: $^{15}N_2$ while the enriched sample had a ratio of 100:12.49:1.43, clearly indicating ^{15}N enrichment.

Calculation 1: Presence of ¹⁵N in the N-terminal Isoleucine or the central Tyrosine

Note: Neutral loss isotopic enrichment model

m/z	Theoretical Isotope Ratio (x100)			m/z	Mole Fraction (x100)		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
534 -> 216	100.00	0.00	0.00	534 -> 216	83.13	0.00	0.00
535 -> 216	17.88	100.00	0.00	535 -> 216	14.87	83.39	0.00
536 -> 216	2.40	17.52	100.00	536 -> 216	2.00	14.61	83.64
		2.40	17.16				
			2.40				

	Isotope Design Matrix				Pseudoinverse Matrix		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
534 -> 216	83.135	14.866	1.999	534 -> 216	0.0120	-0.0021	0.0001
535 -> 216	0.000	83.389	14.610	535 -> 216	0.0000	0.0120	-0.0021
536 -> 216	0.000	0.000	83.640	536 -> 216	0.0000	0.0000	0.0120

Sample	Measured Isotope Ratio (x100)				Corrected, Tracer:Tracee Ratio (x100)		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
1	100.000	17.831	2.370	1	100.00	-0.05	-0.03
2	100.000	17.823	2.427	2	100.00	-0.06	0.03
3	100.000	17.992	2.416	3	100.00	0.11	-0.01
Mean	100.000	17.882	2.405	Mean	100.00	0.00	0.00
S.D.	0.000	0.095	0.030	S.D.	0.00	0.09	0.03
1	100.000	26.970	4.595	1	100.00	9.06	0.60
2	100.000	26.181	4.526	2	100.00	8.27	0.66
3	100.000	26.848	4.421	3	100.00	8.94	0.44
Mean	100.000	26.666	4.514	Mean	100.00	8.76	0.57
S.D.	0.000	0.424	0.088	S.D.	0.00	0.42	0.11

Corrected Isotopic Abundances & Ratio

	[M]	[M + 1]	[M + 2]
	1.203	-0.001	0.000
	1.203	-0.001	0.000
	1.203	0.001	0.000
	1.203	0.109	0.007
	1.203	0.100	0.008
	1.203	0.108	0.005

Calculation 2: Presence of ¹⁵N in the AHEP moiety

Note: Charged residue isotopic enrichment model

m/z	Theoretical Isotope Ratio (x100)			m/z	Mole Fraction (x100)		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
534 -> 216	100.00	0.00	0.00	534 -> 216	92.12	0.00	0.00
535 -> 217	8.56	100.00	0.00	535 -> 217	7.88	92.42	0.00
536 -> 218	0.00	8.20	100.00	536 -> 218	0.00	7.58	92.42
		0.00	8.20				
			0.00				

	Isotope Design Matrix				Pseudoinverse Matrix		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
534 -> 216	92.115	7.885	0.000	534 -> 216	0.0109	-0.0009	0.0001
535 -> 217	0.000	92.421	7.579	535 -> 217	0.0000	0.0108	-0.0009
536 -> 218	0.000	0.000	92.421	536 -> 218	0.0000	0.0000	0.0108

Sample	Measured Isotope Ratio (x100)				Corrected, Tracer:Tracee Ratio (x100)		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
1	100.000	8.509	0.000	1	100.00	-0.05	0.00
2	100.000	8.656	0.000	2	100.00	0.10	-0.01
3	100.000	8.515	0.000	3	100.00	-0.04	0.00
Mean	100.000	8.560	0.000	Mean	100.00	0.00	0.00
S.D.	0.000	0.083	0.000	S.D.	0.00	0.08	0.01
1	100.000	16.107	0.000	1	100.00	7.52	-0.62
2	100.000	15.395	0.000	2	100.00	6.81	-0.56
3	100.000	15.439	0.000	3	100.00	6.86	-0.56
Mean	100.000	15.647	0.000	Mean	100.00	7.06	-0.58
S.D.	0.000	0.399	0.000	S.D.	0.00	0.40	0.03

Corrected Isotopic Abundances & Ratio

	[M]	[M + 1]	[M + 2]
	1.086	-0.001	0.000
	1.086	0.001	0.000
	1.086	0.000	0.000
	1.086	0.082	-0.007
	1.086	0.074	-0.006
	1.086	0.074	-0.006

Calculation 3: Presence of ^{15}N in the N-terminal Isoleucine

Note: Neutral loss isotopic enrichment model

m/z	Theoretical Isotope Ratio (x100)			m/z	Mole Fraction (x100)		
	$^{15}\text{N}_0\text{-PA}$	$^{15}\text{N}_1\text{-PA}$	$^{15}\text{N}_2\text{-PA}$		$^{15}\text{N}_0\text{-PA}$	$^{15}\text{N}_1\text{-PA}$	$^{15}\text{N}_2\text{-PA}$
534 -> 379	100.00	0.00	0.00	534 -> 379	92.36	0.00	0.00
535 -> 379	8.01	100.00	0.00	535 -> 379	7.40	92.67	0.00
536 -> 379	0.26	7.65	100.00	536 -> 379	0.24	7.09	92.98
		0.26	7.29				
			0.26				

	Isotope Design Matrix				Pseudoinverse Matrix		
	$^{15}\text{N}_0\text{-PA}$	$^{15}\text{N}_1\text{-PA}$	$^{15}\text{N}_2\text{-PA}$		$^{15}\text{N}_0\text{-PA}$	$^{15}\text{N}_1\text{-PA}$	$^{15}\text{N}_2\text{-PA}$
534 -> 379	92.361	7.395	0.244	534 -> 379	0.0108	-0.0009	0.0000
535 -> 379	0.000	92.670	7.089	535 -> 379	0.0000	0.0108	-0.0008
536 -> 379	0.000	0.000	92.980	536 -> 379	0.0000	0.0000	0.0108

Sample	Measured Isotope Ratio (x100)				Corrected, Tracer:Tracee Ratio (x100)		
	$^{15}\text{N}_0\text{-PA}$	$^{15}\text{N}_1\text{-PA}$	$^{15}\text{N}_2\text{-PA}$		$^{15}\text{N}_0\text{-PA}$	$^{15}\text{N}_1\text{-PA}$	$^{15}\text{N}_2\text{-PA}$
1	100.000	7.880	0.180	1	100.00	-0.13	-0.07
2	100.000	7.621	0.283	2	100.00	-0.38	0.05
3	100.000	8.520	0.331	3	100.00	0.51	0.03
Mean	100.000	8.007	0.265	Mean	100.000	0.000	0.000
S.D.	0.000	0.463	0.077	S.D.	0.000	0.461	0.065
1	100.000	10.484	0.355	1	100.00	2.47	-0.10
2	100.000	11.149	0.373	2	100.00	3.13	-0.13
3	100.000	10.290	0.457	3	100.00	2.28	0.02
Mean	100.000	10.641	0.395	Mean	100.000	2.625	-0.070
S.D.	0.000	0.451	0.054	S.D.	0.000	0.449	0.078

Corrected Isotopic Abundances & Ratio

	[M]	[M + 1]	[M + 2]
	1.083	-0.001	-0.001
	1.083	-0.004	0.001
	1.083	0.006	0.000
	1.083	0.027	-0.001
	1.083	0.034	-0.001
	1.083	0.025	0.000

Calculation 4: Presence of ¹⁵N in the central Tyrosine or the AHEP moiety

Note: Charged residue isotopic enrichment model

m/z	Theoretical Isotope Ratio (x100)			m/z	Mole Fraction (x100)		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
534 -> 379	100.00	0.00	0.00	534 -> 379	84.21	0.00	0.00
535 -> 380	16.89	100.00	0.00	535 -> 380	14.22	84.47	0.00
536 -> 381	1.86	16.53	100.00	536 -> 381	1.57	13.96	84.72
		1.86	16.17				
			1.86				

	Isotope Design Matrix				Pseudoinverse Matrix		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
534 -> 379	84.213	14.219	1.568	534 -> 379	0.0119	-0.0020	0.0001
535 -> 380	0.000	84.467	13.962	535 -> 380	0.0000	0.0118	-0.0020
536 -> 381	0.000	0.000	84.724	536 -> 381	0.0000	0.0000	0.0118

Sample	Measured Isotope Ratio (x100)				Corrected, Tracer:Tracee Ratio (x100)		
	¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA		¹⁵ N ₀ -PA	¹⁵ N ₁ -PA	¹⁵ N ₂ -PA
1	100.000	16.567	1.841	1	100.00	-0.32	0.03
2	100.000	16.904	1.821	2	100.00	0.02	-0.04
3	100.000	17.185	1.923	3	100.00	0.30	0.01
Mean	100.000	16.885	1.862	Mean	100.000	0.000	0.000
S.D.	0.000	0.310	0.054	S.D.	0.000	0.309	0.039
1	100.000	31.980	5.447	1	100.00	15.05	1.08
2	100.000	27.237	5.079	2	100.00	10.32	1.50
3	100.000	29.020	5.584	3	100.00	12.10	1.71
Mean	100.000	29.412	5.370	Mean	100.000	12.489	1.429
S.D.	0.000	2.396	0.261	S.D.	0.000	2.388	0.317

Corrected Isotopic Abundances & Ratio

	[M]	[M + 1]	[M + 2]
	1.187	-0.004	0.000
	1.187	0.000	-0.001
	1.187	0.004	0.000
	1.187	0.179	0.013
	1.187	0.123	0.018
	1.187	0.144	0.020