OVEREXPRESSION OF ADENYLATION DOMAINS FOR THE IDENTIFICATION OF THE BIOSYNTHETIC PATHWAY OF K-26, AN ACE INHIBITING METABOLITE

Andrea Monique Sease

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Chemistry

May, 2009

Nashville, Tennessee

Approved:

Professor Brian O. Bachmann

Professor Michael P. Stone

To my father, Kerry Ambrose Sease, who is not here to see. To my mother, Valencia H. Sease, who always encourages and supports. And, to my grandmother, Clara Sease who always cared.

ACKNOWLEDGEMENTS

This work would not have been possible without the financial support of the Vanderbilt Chemistry Department.

I would especially like to thank Dr. Brian Bachmann and the rest of my research group for their patience, wit, and guidance.

I would also like to thank my family and friends for their support and believe in me. Especially my mother and younger sisters, who remind me to pursue what I love.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

CHAPTER I

Introduction

Natural products have had an important place in the history of drug discovery; from the discovery of penicillin to the mass production of Aspirin by the Bayer Company 6 March 1899.[1] Natural products are often hailed as miracle drugs that lead to improvements in the longevity and quality of live. Between January, 1981 and June 2006 ~ 47% of new chemical entities were either natural products or derivatives and analogues of natural products. [2] Some examples of natural products found in use include: antibacterials penicillin G and fosfidomycin, as an herbicide Bialaphos. Aspirin, used as an analgesic to relieve pain, as an antipyretic to reduce fever, and as an anti-inflammatory medication. Fosfomycin is a broad-spectrum antibiotic used for the treatment of urinary tract infections. Cyclosporin is an immunosuppressant drug widely used in organ transplant patients. Surfactin is a very powerful surfactant used as an antibiotic, with other properties including anti-bacterial, anti-viral, anti-fungal, anti-mycoplasma, and hemolytic.

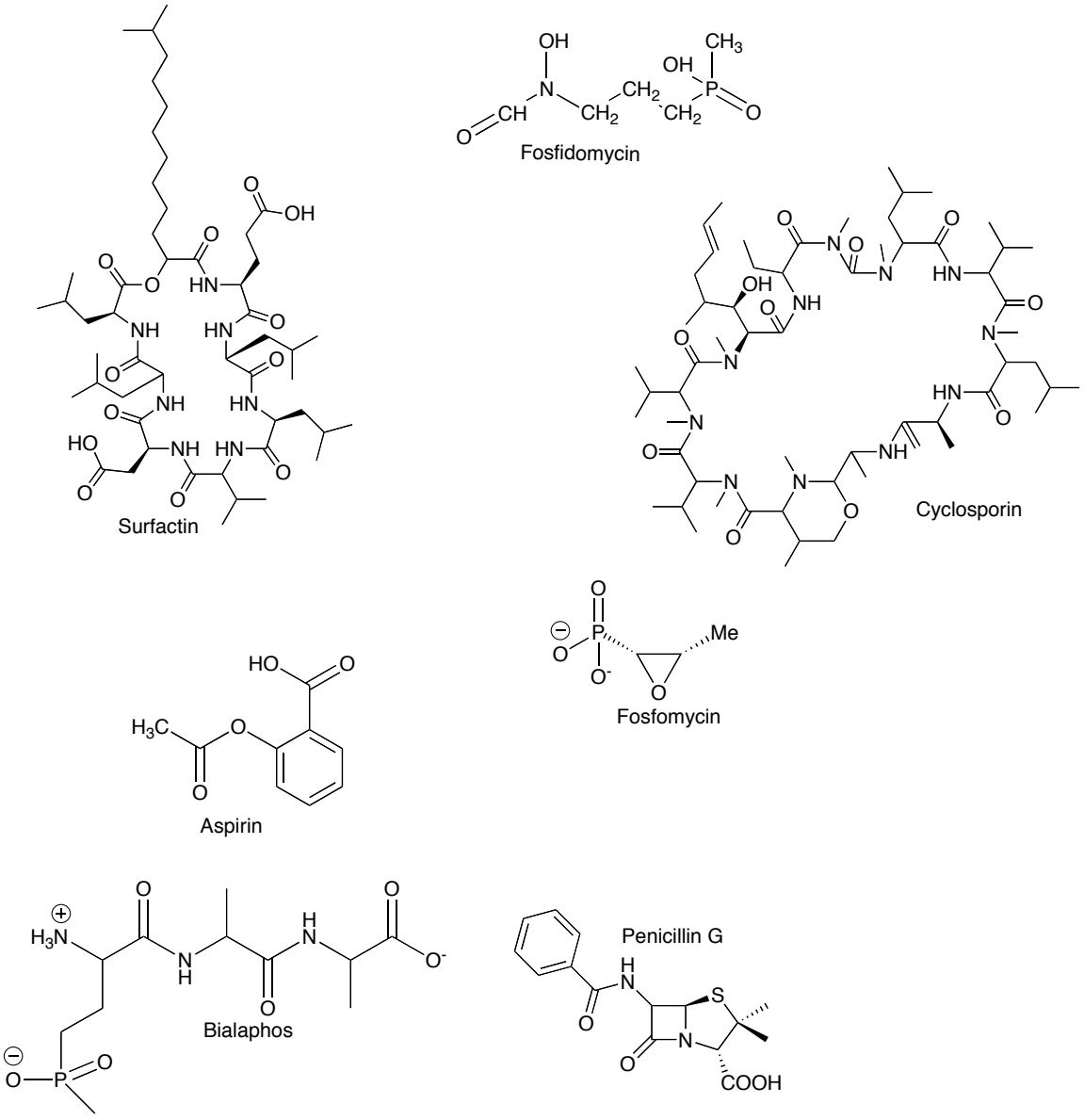


Figure 1 Structures of Natural Products in clinical use

Aspirin

It is important to note the historical precedence for the use of natural products as pharmaceuticals. For example, the use of willow bark for its medicinal properties dates back as far as early Mesopotamia, over 6000 years ago. Willow bark was

often used to cure pain, fever, and inflammation. However, it was not until 1803 until someone actually began to identify and isolate the compound responsible for willow barks activity. In 1829, the crystallized form of salicin, a yellow active compound, was isolated and then prescribed for the treatment of rheumatism.

Salicin was acid hydrolyzed to D-glucose and salicyl alcohol. The salicylic acid can be derived from the oxidation salicylic alcohol, the oxidation of salicylic aldehyde or hydrolysis of methylsalicyl ester. The discovery of Salicylic acid helped in the elucidation of the structure of salicin. Its structure was further elucidated by the French chemist, Cohours, in 1845 and by the Scottish chemist, Couper, in 1858, who both hydrolyzed the methylsalicyl ester in wintergreen (Gaultheria) oil upon treatment with phosphorous perchloride. The hydroxyl radical form readily reacts with the aromatic moiety, which results in hydroxylation. Kolb & Lautemann achieved the chemical synthesis of salicylic acid on a small scale. Later, they devised large-scale chemical reactions, which led to the establishment of a commercial establishment, the Heyden Chemical Company, in Germany for the production of salicylic acid for analgesic and antipyretic purposes. However, salicylic acid upsets most stomachs. Later, Felix Hoffmann preferentially chemically acetylated salicylic acid in 1897 to produce acetylsalicylic acid.

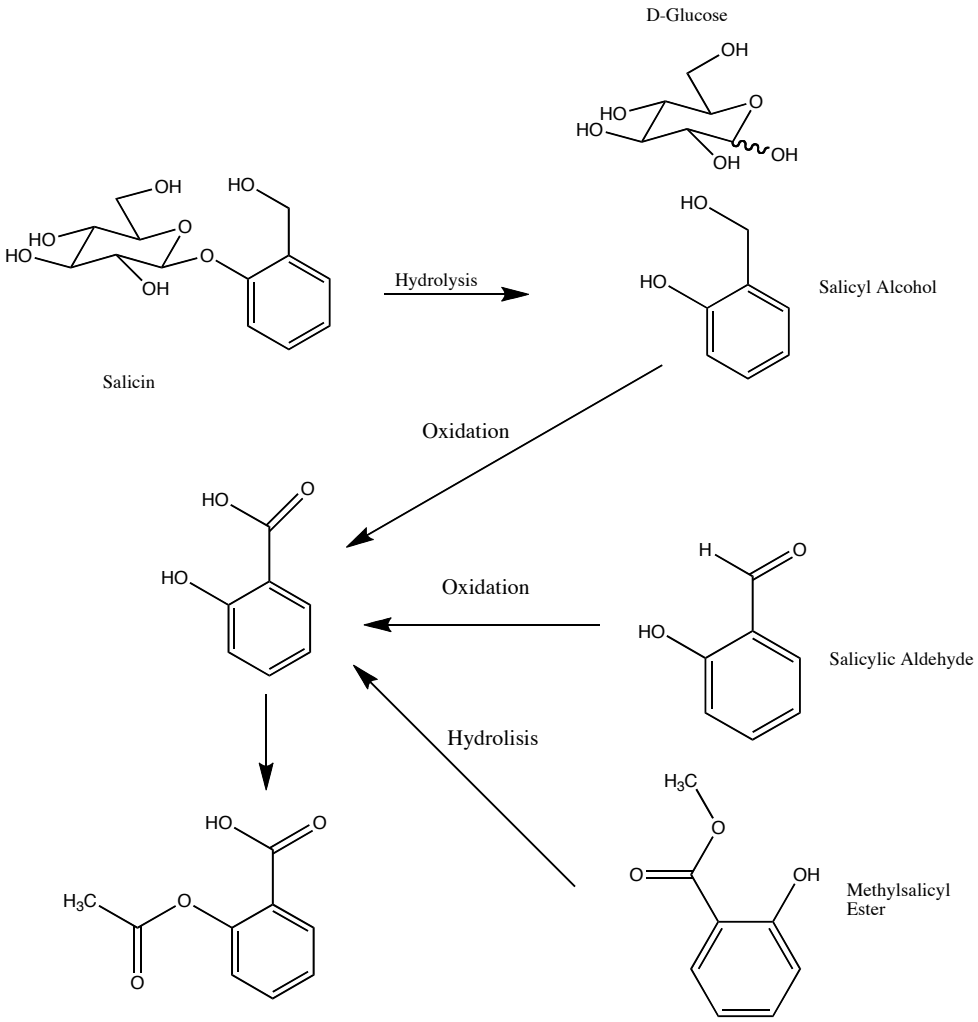


Figure 2: Chemical Synthesis of Aspirin

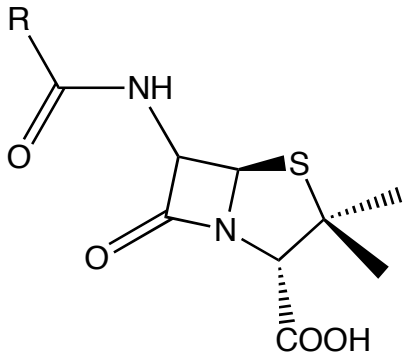
Salicylic acid is not only produced by the willow tree but is common in the plant kingdom. In plants it regulates physiological processes and plant defense the production of has been found to be apart of the shikimic acid cycle, originally discovered in bacteria, it was later confirmed to take place in the chloroplasts of some plants.[3] Once the entire gene cluster and biosynthetic pathway can be elucidated perhaps this will lead to better understanding of aspirin production as well

as other naturally occurring products. Aspirin is an excellent example of the use of a naturally bioactive compound being used in mass production for prescription. This also shows the ability to synthetically improve a natural compound. Another excellent historical discovery, penicillin resulted in the establishment of natural products discovery as a separate science.

Penicillin

In 1928, Alexander Fleming discovered in some Petri dishes, which were supposed to contain bacteria, instead there was a blue-green mold. Even more interesting to note was the ring of nothing surrounding the mold. In the dead zone surrounding the mold, Fleming thought the bacteria were undergoing lysis. The mold was misidentified as *Penicillium rubrum* and broth filtrate named "penicillin."[4] however a number of years passed with Fleming unable to generate enough attention to his findings, until interest grew again under the duress of World War II, WWII. Due to the pressures of WWII, companies and nations bonded together to begin the industrial production, purification, and distribution of penicillin. Despite the industrial production of penicillin, fermentation of the drug was still arduous and time consuming. It was thought that if the structure was known then the steps making a synthetic metabolite could be taken. The process took five years and collaboration between two world powers. While attempting to isolate the active compound from the fermentation broth, it became obvious there was more

than one penicillin. In fact, there was an entire family of penicillins with differing structures and bioactivities. Finally, in 1945 the structure of penicillin was ultimately solved. [5]



R= C_5H_9 or $C_6H_5CH_2$

Figure 3: Penicillin

The biosynthetic pathway of penicillin has also been elucidated. Three amino acids make up the initial structure of penicillin: L-α-amineadipic acid, L-cysteine and, L-valine. The linear peptide ACV is formed by a multi-enzyme synthetase, ACV. The function of this synthetase is similar to modules in nonribosomal peptide synthetases(NRPS) using ATP the amino acid is first activated. The activated amino acid is then bound as a thioester; the L-valine is epimerized, and then the three amino acids are bound to each other. Next, the tri-peptide is oxidized to isopenicillin by isopenicillin N synthase (IPNS). Next, the L-α- aminoadippyl side chain is cleaved giving 6- aminopenicillanic acid, which is then converted to penicillin G via addition of a phenylacetyl group. [6]

Discovery of the penicillin nucleus 6-aminopenicillanic acid [6-APA] in the

late 1950's allowed for synthesis of a multitude of semi synthetic penicillins. After the discovery of the biosynthetic pathway, it was determine that due to ease of production and biological properties, penicillin G was the preferred choice for production.[7] By the addition of phenylacetic acid, the broth would predominantly produce penicillin G. From then on, other precursors were added to fermentation in hopes that other penicillins would result. Only one other showed advantage over penicillin G.

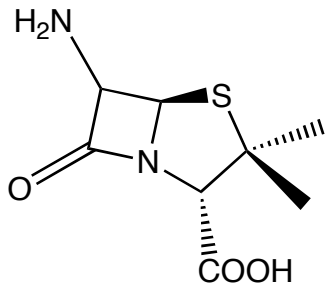


Figure 4: Penicillin nucleus

Another way alternate penicillins could be obtained was to produce one type of penicillin using fermentation, then to chemically modify the product. This approach resulted in what were called semi-synthetic penicillins. The ability to singularly produce and then chemically modify the penicillin nucleus opened the door for the production of even more effective antibiotics, and also solved the problem of resistivity, which had, began soon after mass implementation of penicillin. [8]

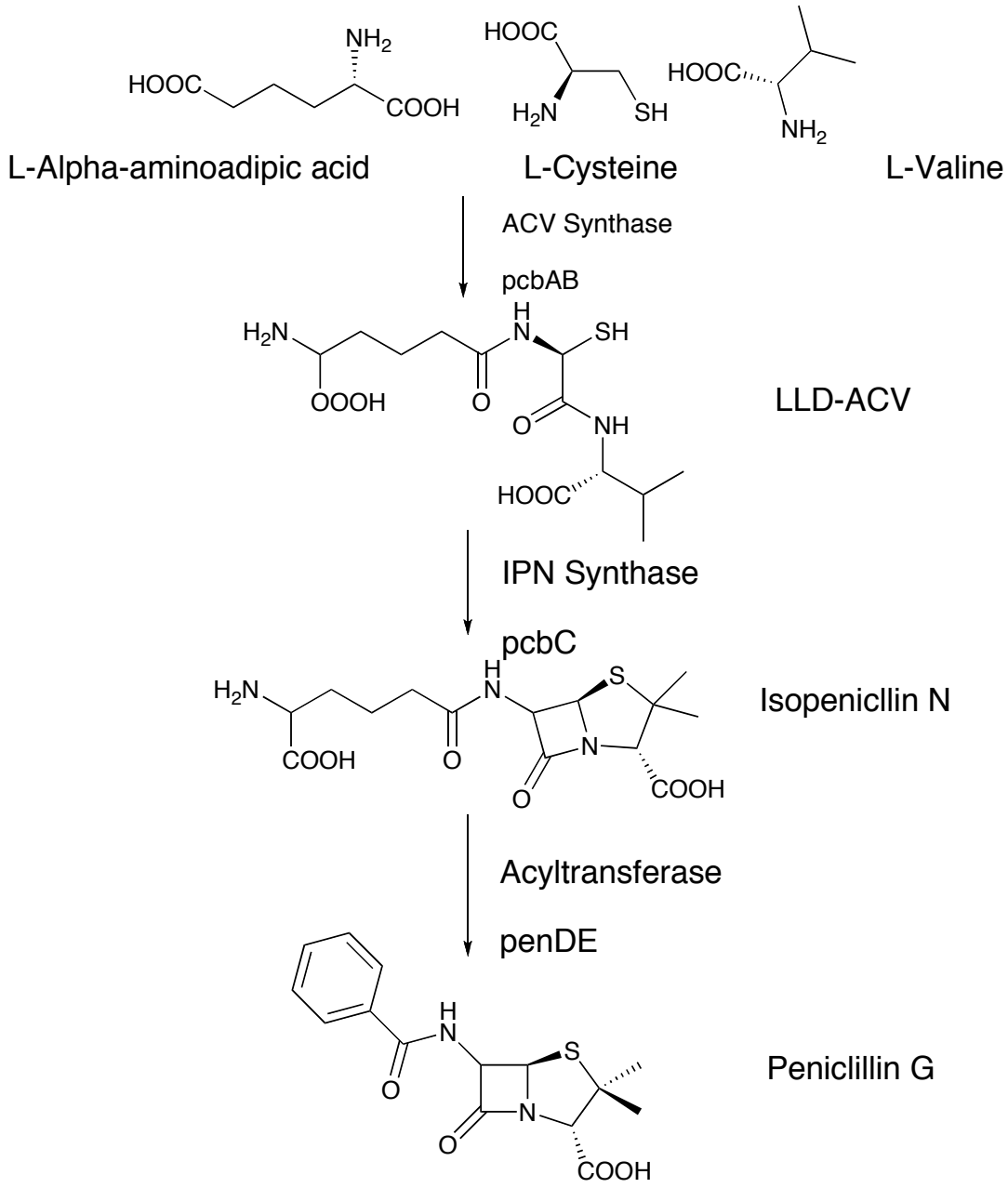


Figure 5: Synthesis Of Penicillin

Drug discovery itself is a newer science in comparison to most. Beginning with the isolation of morphine and other plant based pharmaceuticals, general production and purification was needed. Discovery and isolation of penicillin, led to the first clinical trials and mass fermentation practices for the general produc-

tion of drugs. Following this discovery and subsequent antibiotics, many companies added both a microbiology and fermentation department. Later molecular biology began to influence drug research with the use of cloning and protein expression. Later more advances were made with the use of recombinant proteins. Sequencing of the human genome, high throughput screening, combinatorial chemistry, structural based design and directed evolution has unfortunately led to a decline in the exploration of natural resources for pharmecuticals.[9]

The decline of the exploration of natural products can be attributed to the use of combinatorial chemistry combined with high though put screening. It was thought that sheer numbers could help further the discovery of new chemical entities. Combinatorial chemistry was used to generate huge libraries of chemical structures, which then used an automated system for the determination of activity using biological assays.

There are two general approaches to generating libraries using combinatorial chemistry. There is the focused library and the prospecting library. The focused library uses already identified structures with biological activity and attempts to modify these structures in hopes of improving this activity. The prospecting library is to provide novel structures entirely or later analysis in hopes that something is active.[10] Hits are determined by the use of automated assays using high through put screening. HTS allows the testing of small samples of each product using vitro assays to identify biologically active compounds. Previously only about 3000 selected compounds could be tested using 1mL samples in test

tubes in 1–2 years.[11] Use of high throughout screening has resulted in some hits (discoveries), but many researchers are beginning to advocate, and point out that many of the 'old' drugs have stood the test of time. Many studies also show that there are alternate uses for the older drugs then originally known.

Many biologically produced small molecules contain interesting motifs that not only are singularly produced naturally, but are active for use by humans. Many methods are currently being explored for the use of these. Cloning and the creation of massive gene data banks have played a definite role in the discovery of new entities. As genetic material was sequenced and placed in the databanks, it became obvious that many new natural products could be discovered by simply by isolated the results of areas in the genomes that appeared to be responsible for the production of complex naturalproducts.[12] through the sequencing of gene clusters they are able to identify regions which might encode for novel products as well as areas whose mutations might result in novel products.

With the advent of the science of drug discovery and the introduction of high through out screening, the ways drugs were discovered changed. Before the use of HTPS, compounds were dissolved in aqueous solutions under thermodynamic conditions. Using 96-well plates, compounds are dissolved in DMSO with surfactants to improve solubility. These changes allowed false hits. As a result, compounds that would later be deemed unacceptable were allowed to be tagged as potential hits due to changes in solubility and molecular interaction. Thus, something had to be done to single out these false hits. Looking at the databases of

chemical entities it was found one major reason compounds often failed were due to solubility and permeability issues. Looking at the world drug index a large database of drugs, certain criteria were broken up statistically to determine what factors allowed drugs to get through all the phases of clinical testing. As a result a set of laws that have been assembled concerning what properties a compound must posses before it is considered for use as a prescription called Lipinski's rule of five[13]:

1) The molecular weight can be no greater than 500 g.
2) Have no more than five H-bond donors (expressed as the sum of OH's and NHs).
3) No more than 10 H-bond acceptors (expressed as the sum of Ns and Os)
4) And have a calculated Log P (M LogP) no greater than 5
5) Compound classes that are substrates for biological transporters are exceptions to the rules.

This study preformed by Lipinski et al.[13] finds a correlation between poor adsorption and permeability and compounds that lie outside of this set of rules. It is interesting to note that when this study was conducted it excluded products discovered beforehand. It disregards natural products because often times natural products contain structures that are similar to products already produced by the body, and thus easily incorporated into the body. The natural product often is merely an analogue. This is important because this means that natural products

entering drugs trials already have a greater chance of succeeding over their synthetic counter parts. It sheds light on ways of improving the mechanism of research and discovery for pharmaceuticals.

The field of genetics has also forwarded our understanding of two natural product producing biological systems: polyketide synthetase and non-ribosomal peptide synthetases. Gene cluster mining has lead to many interesting discoveries. for example, most small molecules that are used for pharmaceuticals are produced either non ribosomally, or via polyketide synthesis. These methods of production create molecule with certain structural characteristics that have been shown to be biologically active. Polyketide synthesized and ribosomally produced natural products are identified by the use of a limited group of proteingenic amino acids. When a small molecule contains peptides outside of those 20 proteogenic options, the mechanism of formation can be further elucidated by the presence of other structural motifs.[14]

Non Ribosomal Peptide Synthesis

Nonribosomal peptide synthesis is characterized by the presence non-proteingenic amino acids. Important NRPS compounds with biological activity include vancomycin, actinomycin, and cyclosporine. The structural motifs of NRPS compounds include macrocyclic, branched macrocyclic, and dimers or trimers of identical structural elements. Smaller heterocyclic rings like thiazoline, oxazoline, or thiazole are also common structural features of nonribosomal pep-

tides. Additionally, these peptides may contain N-methylations, N-formylations and glycosylations as well as insertion of acetate or propionate units while fatty acids are also sometimes incorporated.[15]

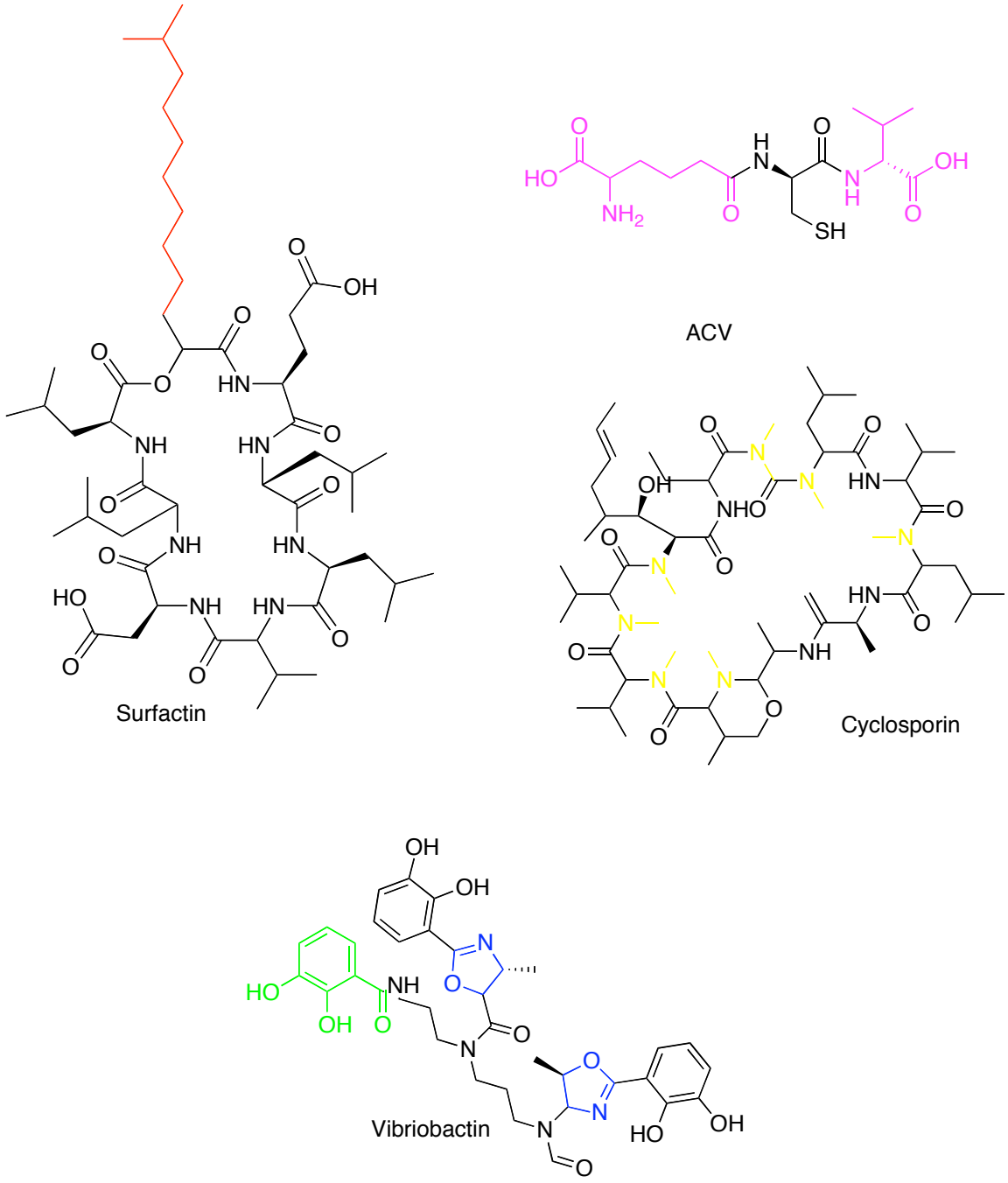


Figure 6: Nonribosomally Produced Compounds And Their Characteristics: Red Are Fatty Acids, Purple Are Non-Proteingenic Amino Acids, Green Are Carboxy Acids, Blue Are Heterocycles And Yellow Are N-Methyls.

NRPS's are multifunction modular enzymes that serve as both the template and the biosynthetic mechanism. Each modular enzyme can then be divided into domains, which are performing one single enzymatic step. There are three main domains that are necessary for each module: the adenylation domain (A), the peptidyle carrier protein domain (PCP) or the thiolation domain (T), and the condensation domain(C).[16] The peptide is first activated by addition of an ATP. Adenylation domains are the highly selective domains that select the peptide, which is activated. Once activated the unstable amino acid is then transferred to the T domain, where is it attached via the pantetheine arm by thiolation. The Apo -PCP domain must also be activated before attaching the activated amino acid. The phosphopantetheine moiety of coenzyme A is covalently attached to a serine residue of T by Sfp, a phosphopantheteine transferase.[17]

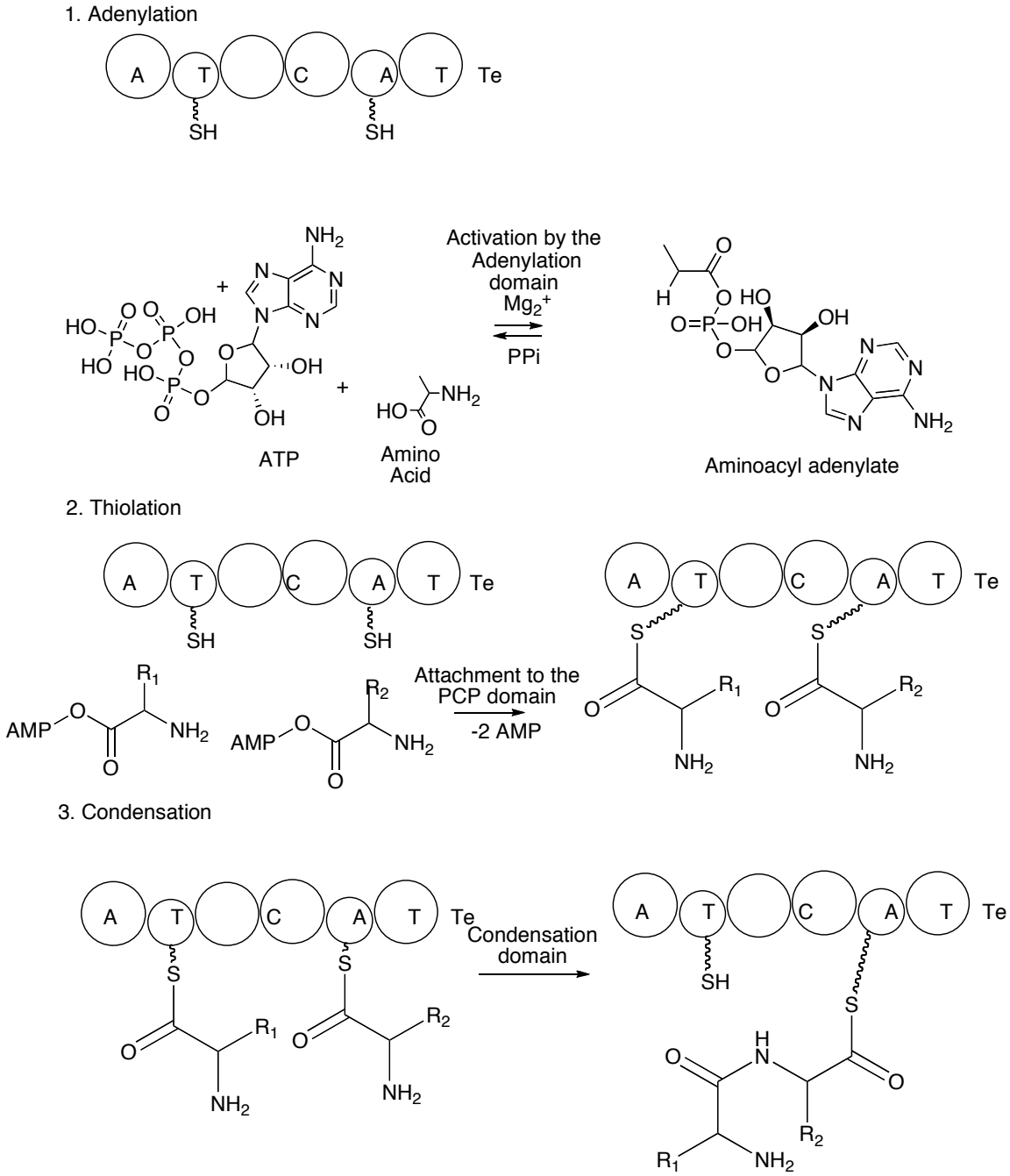


Figure 7: Loading of the NRPS Module

The combination of the A and T domains is known as the initiation module. After activation and covalent binding of the first amino acid substrate by the initiation module, peptide synthesis proceeds by stepwise condensation with each T

bound amino acid. Peptide bond formation is done by the condensation domain (C domain). The C domain catalyzes the nucleophilic attack of the downstream T-bound acceptor amino acid with its free α-amino group on the activated T-bound donor amino acid. The peptide is released by the thioesterase domain (TE). Depending on the identity of the NRPS template, product release can be carried out either by the external nucleophile to give the linear acid product or by an internal nucleophile to yield a cyclic product.

Other modifications are made using other domains that are called editing domains. These modifications include: epimerization, methylation, oxidation, reduction, and modification of the N-terminal peptide end by a N-formyltetrahydrofolate-dependent formyltransferase domain.

Due to the singular enzymatic and mechanistic characteristics of each domain, the structure of a natural product itself can help predict what the genome will look like and aids in the location of the genes responsible for the production of that natural product. For example the small molecule ACV tripeptide mention earlier as a precursor to penicillin, it can be elucidated from the structure, the modules and domains responsible for its production:

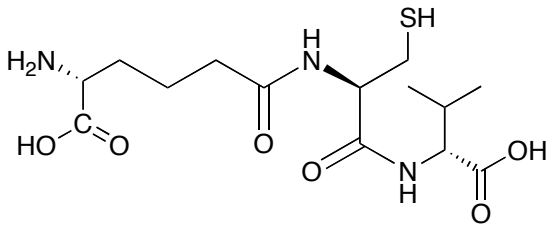


Figure 8:Penicillin Nucleus

There are three amino acids that can be identified by the presence of the three amines: L-α-aminoadipic acid, L-cysteine, D-valine:

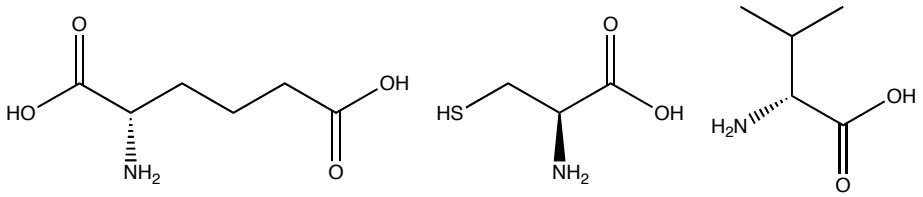


Figure 9: Amino Acids That Make Up Penicillin Nucleus

The presence of the three amino acids means that there must be three modules. One way to verify the incorporation of these amino acids would be by feeding studies, feeding the organism labeled versions of these amino acids and checking for their incorporation into the final product.

All the domains that must be included in these modulate are: three adenylation domains, three T domains, and three condensation domains. There must also be a group to release the tripeptide from the module. The order of the modules can also be elucidated based on the linear organization of NRPS systems. Thus based upon the structure alone we know that the modules are as follows:

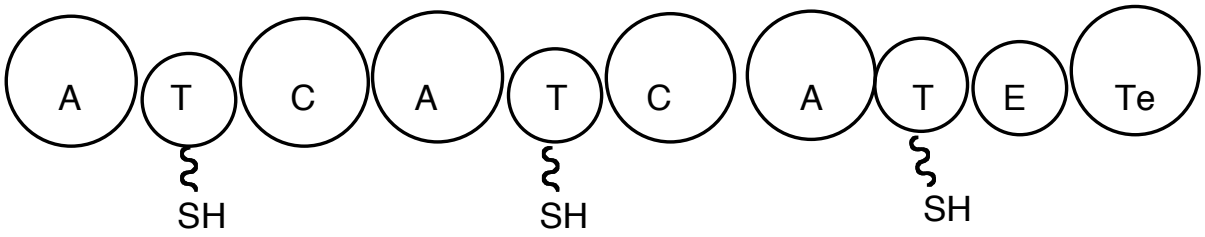


Figure 10:Suggested Modules Of ACV

Each A domain selects and activates each amino acid, which is then attached to the T domain, for there each amino acid is then linked with the previously thiolated amino acid. The third amino acid must be epimerized before it is attached the peptide. Once all three amino acids have then been connected, they are released from the synthetase by use of the Te domain. Having identified the domains and modules to search for as well as the order in which they would appear, one would be able to look at the genome and identify the region responsible for its production.

Following visual inspection, perspective genes are tested with a variety of methods including: use of gene knockouts, or the over expression and characterization of the adenylation domain. One way to generate a gene knock out is transposon mutagenesis. The suspected gene cluster is mutated so that the different mutations have different disruptions in the gene of interest. The mutated gene is then expressed to explore the affect it has on production of the natural product. Another way to determine the correct gene cluster is to clone the adenylation domains and express them. Activation of the adenylation domain is traditionally measured by radioactive ATP- [^{32}P] pyrophosphate (PP) exchange assays. [18]

K-26

An example of a small molecule thought to be a result of a NRPS biosynthetic route is K-26. While searching for ACE inhibiting compounds, K-26 was isolated from the broth filtrate of an actinomycete K-26. K-26 is awater soluble, tri-peptide containing: L-isoleucine, L-tyrosine, and L-(R)-l-amino-2- (4-hydroxyphenyl)-ethylphosphonic acid. K-26 was found to have ACE inhibition properties. Angio-tensin I converting enzyme (ACE) helps in the regulation of blood pressure in mammals. [19]

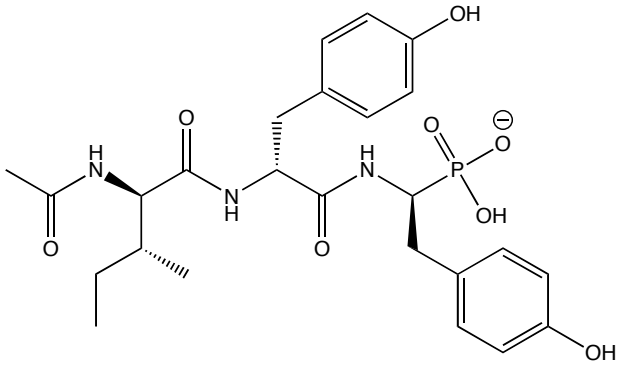


Figure 11: K-26

What is interesting to note in the structure of the ACE inhibitor is the pres-ence of the interesting AHEP moiety. AHEP contains a C –P bond which is un-usual. The small molecule, 2- aminoethylphosphonic acid (AEP) was isolated by Horiguchi and Kandatsu from ciliate protozoa found in sheep rumen, in 1955. [20] Until the discovery of AEP, it was thought that the C-P bond was too unstable to exist naturally in organic compounds. The only way to stabilize a phosphorous was with a heteroatom such as Oxygen. [21] The presence of aminophosphonic

acids has now been detected in many organisms from sea invertebrates to land vertebrates. Interestingly enough it has been found that most of these compounds posses some level of biological activity which includes antiviral, herbicidal, and antibacterial as just a few examples.

This activity has been ascribed to the presence of the phosphonic acid, which acts as an analogue for natural phosphates. The differences include the bond distance between the C-P and the C-O-P, as well as the lower acidity. Enzymes that hydrolyze natural phosphates affect the C-P bond. This implies that if the phosphonic analogue is accepted in place of the phosphate, in a metabolic cycle it cannot be as easily broken down as the phosphate. This means that the analogue might be able inhibit the normal metabolic process, because it does not break down. The analogue might compete for a binding area and thus act as an inhibitor. [22]

In the case of most C-P bond-containing compounds, the incorporation of the C-P bond can be attributed to two enzymes, phosphoenolpyruvate mutase, PEP mutase, and carboxyphosphonoenolpyruvate, CPEP mutase. PEP is responsible for catalyzing the conversion of the phosphate to the phosphonic acid. CPEP functions in the biosynthetic pathway leading to bialaphos; its mechanism remains unknown.

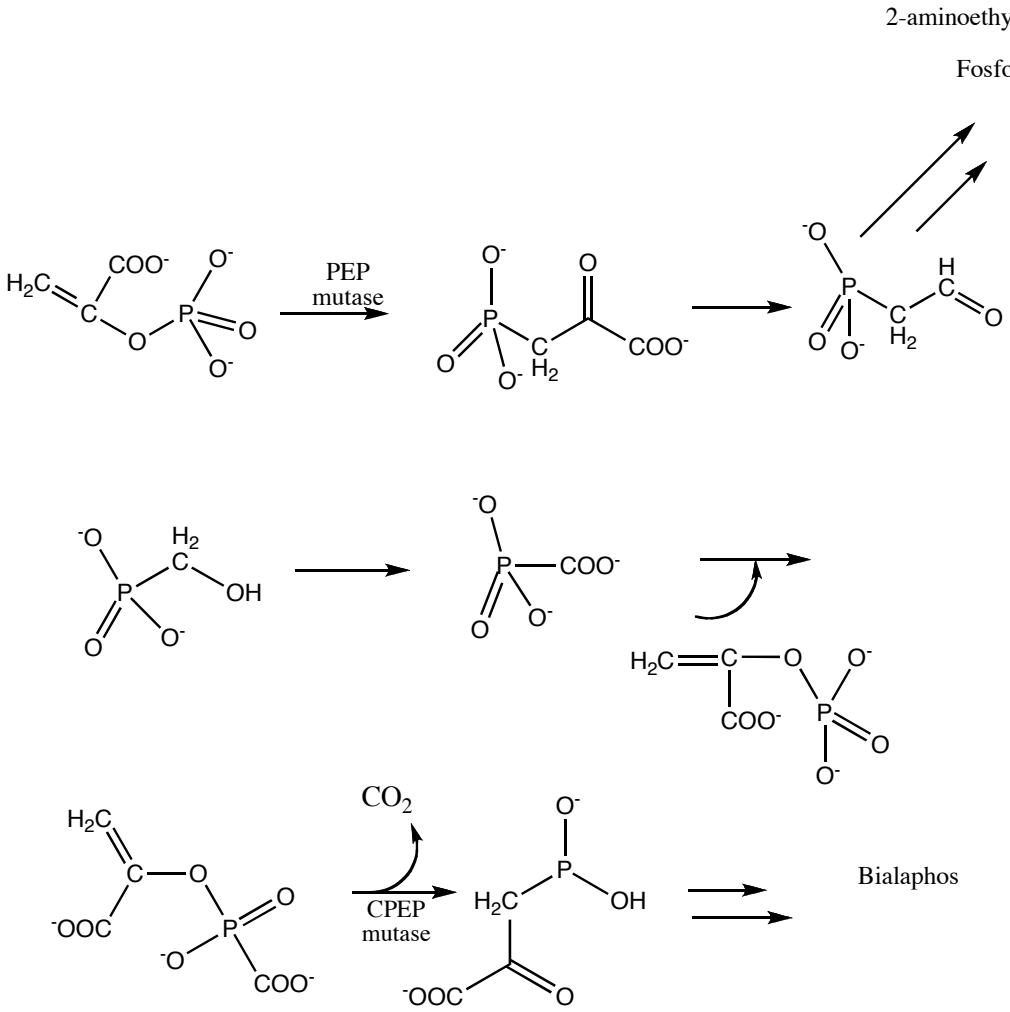
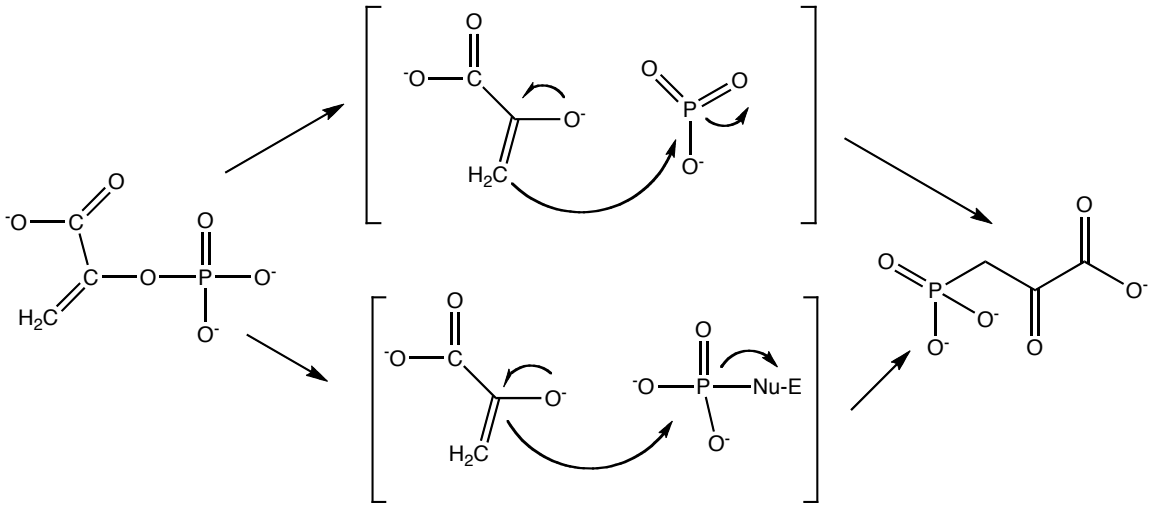


Figure 12: Two C-P Bond Forming Enzymes

Two mechanisms have been proposed for the conversion: In mechanism one the reaction may proceed through an associative, double-displacement pathway in which the enzyme transfers the phosphoryl group from the C (2) oxygen of PEP to an active site residue and, then, to the C (3) of the pyruvate enolate intermediate (mechanism I). The alternative mechanism involves dissociation of metaphosphate from the C (2) oxygen of PEP followed by bond formation to the C (3) of the pyruvate enolate intermediate. [23]

Mechanism 2



Mechanism 1

Figure 13: Mechanism Of C-P Bond Formation

The other enzyme responsible for the formation of C-P bonds CPEP mutase is found exclusively in the biosynthetic pathway of bialaphos. Its mechanism is characterized by an intramolecular rearrangement and a spontaneous decarboxylation. Two key mechanistic questions are whether the decarboxylation event is spontaneous or enzyme catalyzed, and whether it precedes, accompanies or follows the rearrangement. A study later showed that the reaction proceeds *via* a rearrangement of the carboxyphospho group, followed by decarboxylation.[24]

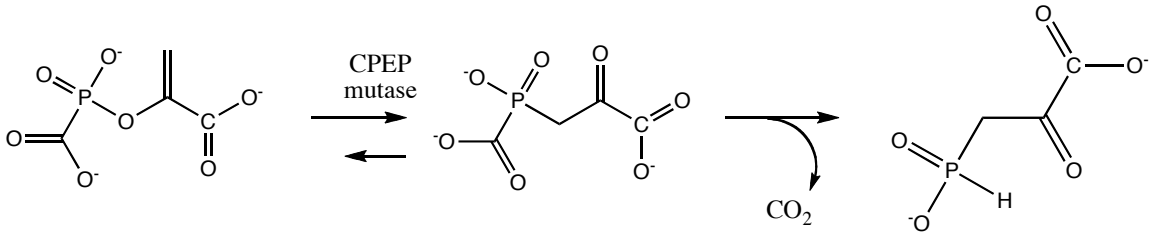


Figure 14: CPEP mutase

The next chapter will delve further in the discovery and biosynthetic pathways of C-P containing compounds. As evidenced by the discovery of penicillin, being able to utilize intermediates that can accumulate during natural production of these compounds can be very useful.

Some interesting phosphonate compounds include: cidofovir for the inhibition of DNA polymerases, bialaphos (an herbicide) and fosfidomycin (an antibacterial). Cidofovir is a synthetic compound that is used for the treatment of HIV. Bialaphos as mentioned is a natural product that can be used as a herbicide. While Fosfidomycin is, an antibacterial used for the treatment of urinary tract infections.

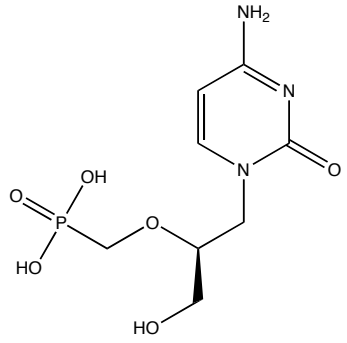


Figure 15: Cidofovir

References

1. J. G. Mahdi. The historical analysis of aspirin discovery, its relation to the willow tree and antiproliferative and anticancer potential. *Cell Proliferation* 2006, *39* (2), 147-155.

2. 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), 461-477.

3. Métraux, J.-P., Recent breakthroughs in the study of salicylic acid biosynthesis. *Trends in Plant Science* 2002, *7* (8), 332-334.

4. Bennett, J. W.; Chung, K.-T., Alexander Fleming And The Discovery Of Penicillin. In *Advances in Applied Microbiology*, Academic Press: 2001; Vol. Volume 49, pp 163-184.

5. Bentley, R., The molecular structure of penicillin. *Journal Of Chemical Education* 2004, 1462.

6. Aharonowitz, Y.; Cohen, G.; Martin, J. F., Penicillin And Cephalosporin Biosynthetic Genes - Structure, Organization, Regulation, And Evolution. *Annu. Rev. Microbiol.* 1992, *46*, 461-495.

7. Rolinson, G. N., 6-APA and the development of the {beta}-lactam antibiotics. *J. Antimicrob. Chemother.* 1979, *5* (1), 7-14.

8. Elander, R. P., Industrial production of ß-lactam antibiotics. *Applied Microbiology and Biotechnology* 2003, *61* (5), 385-392.

9. Drews, J., Drug Discovery: A Historical Perspective. *Science* 2000, *287* (5460), 1960-1964.

10. Spaller, M. R.; Burger, M. T.; Fardis, M.; Bartlett, P. A., Synthetic strategies in combinatorial chemistry. *Current Opinion in Chemical Biology* 1997, *1* (1), 47-53.

11. Pereira, J. A. W., Origin and evolution of high throughput screening. *British Journal of Pharmacology* 2007, *152* (1), 53-61.

12. Challis, G. L., Genome Mining for Novel Natural Product Discovery. *Journal of Medicinal Chemistry* 2008, *51* (9), 2618-2628.

13. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* 1997, *23* (1-3), 3-25.

14. Cane, D. E., Introduction: Polyketide and Nonribosomal Polypeptide Biosynthesis. From Collie to Coli. *Chemical Reviews* 1997, *97* (7), 2463-2464.

15. Finking, R.; Marahiel, M. A., BIOSYNTHESIS OF NONRIBOSOMAL PEPTIDES1. *Annu. Rev. Microbiol.* 2004, *58* (1), 453-488.

16. Sieber, S. A.; Marahiel, M. A., Molecular Mechanisms Underlying Nonribosomal Peptide Synthesis: Approaches to New Antibiotics. *Chemical Reviews* 2005, *105* (2), 715-738.

17. Schwarzer, D.; Finking, R.; Marahiel, M. A., Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* 2003, *20* (3), 275-287.

18. Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D., Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem. Rev.* 1997, *97* (7), 2651-2674.

19. McQuade, T. J.; Shallop, A. D.; Sheoran, A.; DelProposto, J. E.; Tsodikov, O. V.; Garneau-Tsodikova, S., A nonradioactive high-throughput assay for screening and characterization of adenylation domains for nonribosomal peptide combinatorial biosynthesis. *Analytical Biochemistry* 2009, *386* (2), 244-250.

20. Challis, G. L.; Ravel, J.; Townsend, C. A., Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* 2000, *7* (3), 211-224.

21. 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.

22. Isbell, A. F.; Berry, J. P.; Tansey, L. W., Amino Phosphonic Acids. III. The Synthesis and Properties of 2-Aminoethylphosphonic and 3-

Aminopropylphosphonic Acids. *The Journal of Organic Chemistry* 1972, *37* (26), 4399-4401.

23. Fields, S. C., Synthesis of natural products containing a C-P bond. *Tetrahedron* 1999, *55* (42), 12237-12272.

24. Engel, R., Phosphonates as analogues of natural phosphates. *Chemical Reviews* 1977, *77* (3), 349-367.

25. 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), 14165-14173.

CHAPTER II

Introduction

Current natural phosphonates are divided into three categories, each represented as follows: by 2-aminoethylphosphonate, bialaphos, and K-26. Most known phosphonates belong to the first category, in which the C-P bond is formed by phosphoenol-pyruvate mutase catalyzed intramolecular rearrangement of phosphoenol pyruvate to phosphonopyruvate. Phosphinothricin tripeptide (PTT, or bialaphos), the third category is unique among naturally occurring phosphonates in that it has a C-P-C bond. The first C-P bond in PTT is catalyzed by phosphoenol-pyruvate mutase, as in the first category.[25] K-26 is the next category of natural phosphonates, and contains a non-proteogenic amino acid, which excludes the direct involvement of the phosphoenol-pyruvate isomerase, responsible for the beginning synthesis in the previous category. Despite the intriguing biological activity often exemplified by C-P bond containing compounds few biosynthetic pathways has been discovered. Only three phosphonates biosynthetic pathways (as demonstrated by heterologous production) have been uncovered: AEP, bialaphos, and FR900098 [26].

Bialaphos (BA) {Phosphinothricin tripeptide [PTT]}

Bialaphos (BA) representing the third category of C-P bond containing compounds was isolated by Ogawa as a structural component of antibiotic SF-1293 from *Streptomyces hygroscopicus*. [27]It is a tripeptide composed of two alanines and a peptide called phosphinothricin, possessing antibiotic activity because of its interaction with peptide metabolism and nitrogen anabolism. BA is also the result of a non-ribosomal like mechanism. Its three peptides are bonded via condensation domains, after the non-proteogenic peptide PTT is synthesized. The high toxicity of bialaphos in plants also makes it an excellent herbicide. The Phosphinothricin moiety, an analogue of glutamate in plants, acts as an inhibitor resulting in ammonia accumulation, which is the reason for BA's high toxicity.[28]

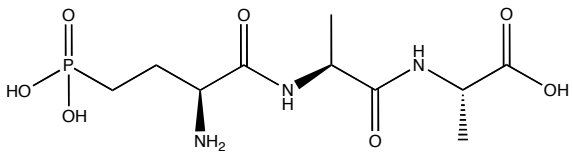


Figure 16: Bialaphos

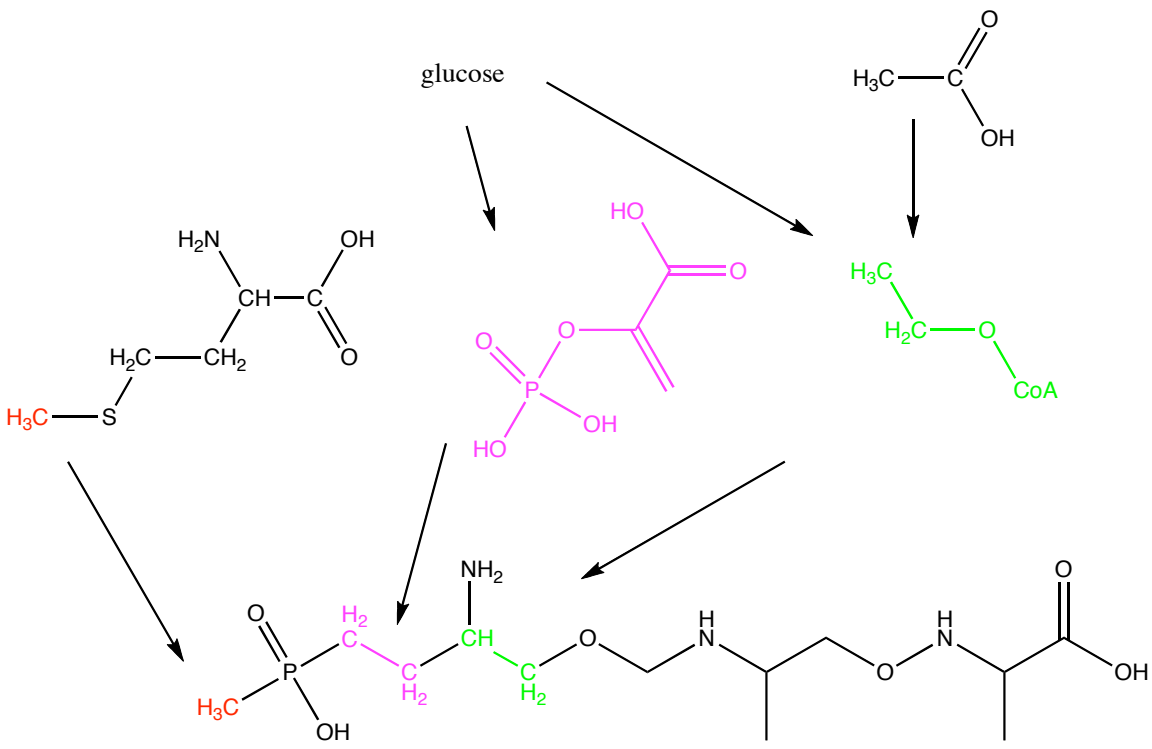


Figure 17: Phosphinothricin Precursors

Biaphos expression has been studied and identified in the strains *Streptomyces viridochromogenes, Streptomyces lividans, or Streptomyces hygroscopicus.* Mutant strains were generated by *Seto et al* [29] to aid in the identification of both the biosynthetic pathway as well as the genes encoding each step. Through use of mutant strains, some of the steps necessary for the production were elucidated.

The first step is the rearrangement reaction discussed in the first chapter of the phosphoenol pyruvate into phosphonopyruvate. The enzyme PEP was proved to responsible for the rearrangement and the first step of the bialaphos pathway with the use of ^{14}C labeled PEP, which was incubated with purified PEP mutase and NADH/malate dehydrogenase. Ironically, the formation of PEP is favored in this reaction; so to prove that this enzyme was indeed reversible the

resulting PPyn was removed to drive the reaction towards its production. To prove that this was indeed the first step in the BA pathway, mutants with the first step blocked were analyzed. The *S. viridochromogenes* genes and enzymes responsible for the initiation of phosphinothricin biosynthesis and production of phosphonoacetaldehyde, involving gene *ppm* and *ppd*, were identified.

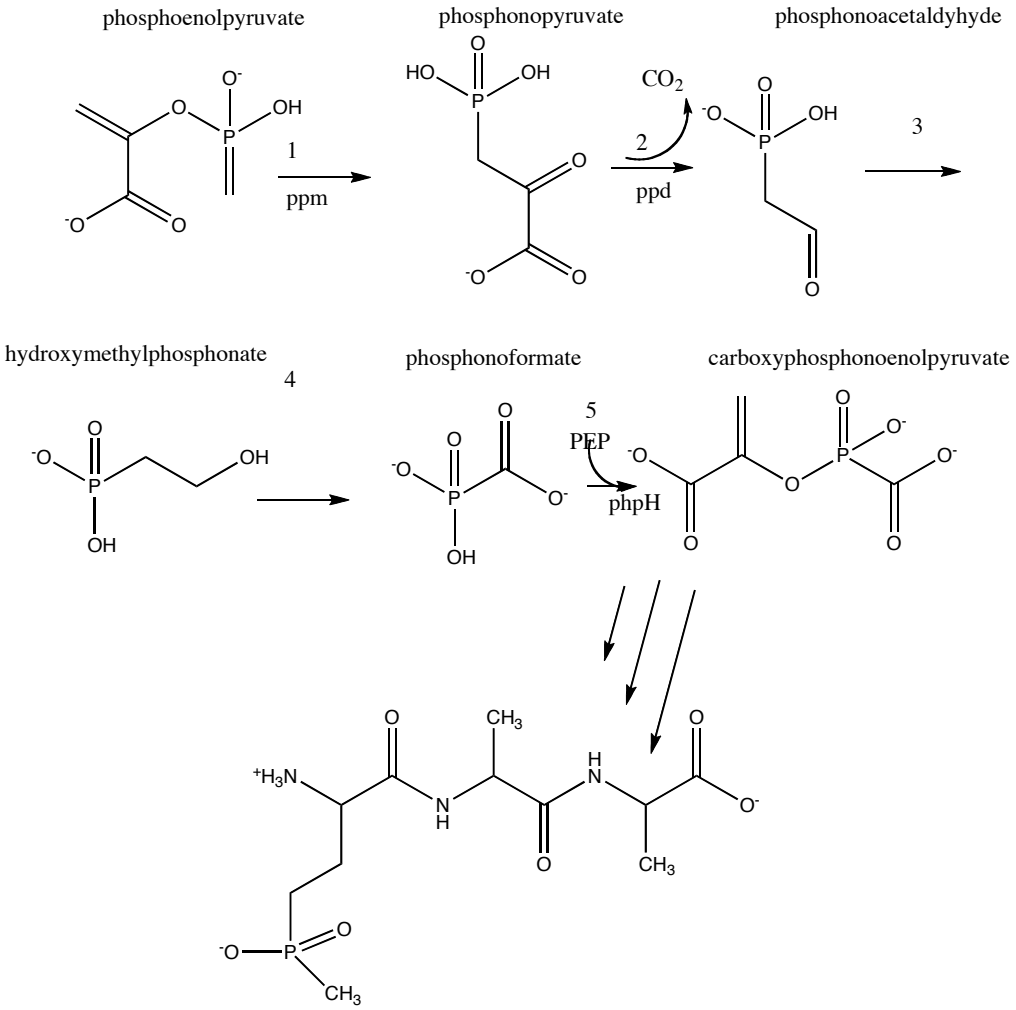


Figure 18 Originally Proposed Pathway

It was originally thought that the catalysis of phosphonoacetaldehyde to carboxy-

phosphonoenolpyruvate took place over three steps starting with cleavage followed by oxidation, and immediately followed by the addition of another molecule of PEP. However a recent study by Blodgett et al[30] suggests that these steps are not so simple nor so few. The CPEP mutase is similar to enolase, which work through dehydration, thus the proposed biosynthetic pathway with expected activity.

Previous studies [31]has shown genes located in a certain region might be involved in these catalytic steps: *phpC*, *phpD*, and *phpE*. Mutants were constructed which blocked each of these individual genes. PTT production took place; however at much lower levels than wild type. The gene we then over expressed with a His tagged end. It was discovered that *phpC* catalyzes the reaction of phosphonoacetaldehyde to hydroxyethlphosphonate using NADH or NADPH as a cofactor. (Step3)

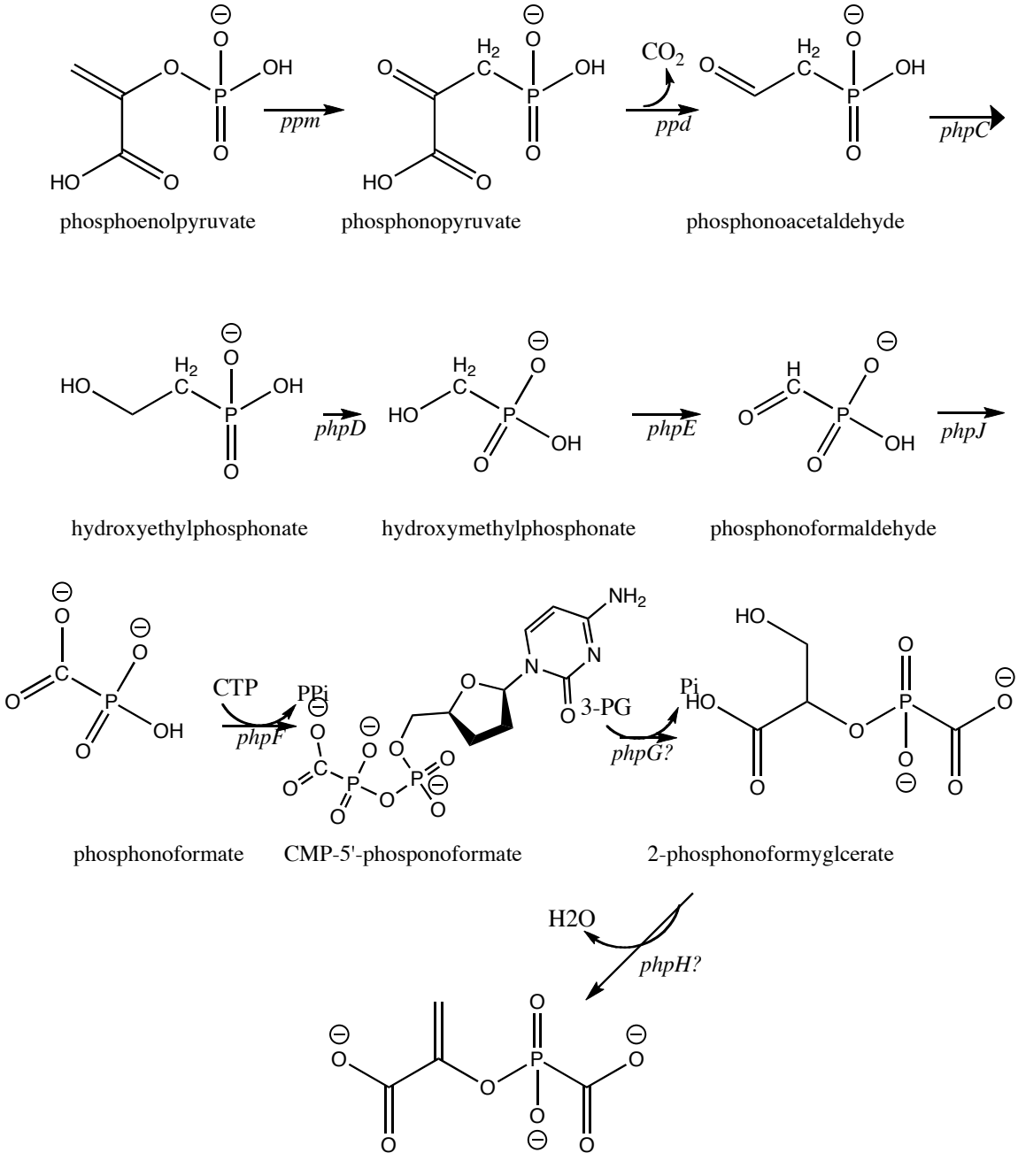
PhpD (step4) was then over expressed as well and it was found through LC/MS analysis that *phpD* is responsible of the conversion of HEP to hydoryxymethylphosphonate. Attempts to characterize *phpE* failed, however bioactivity was detected in the broth cultures were assayed. The *phpE* (step 5) gene encodes a protein similar to the members of the hydoxyacid alcohol dehydrogenase family of enzymes, which suggests that *PhpE* is an alcohol dehydrogenase whose substrate is HMP. If it is an HMP dehydrogenase, the expected product is phosphonoformaldehyde.

To produce the known intermediate phosphonoformate, an aldehyde dehy-

drogenase is required. The sequence of *phpJ* suggests that it is the likely culprit. Mutants of this gene were still able to produce PTT however the accumulated the intermediate aminomethylphosphonate not the expected phosphonoformate. This is reminiscent of the accumulation of other products in previously blocked steps that might be evidence for non-specific aminotransferases. Thus, it was concluded that step 6 is indeed the oxidation of phosphonoformaldehyde to phosphonoformate.

PhpF is responsible for the release of inorganic phosphate when the CMp-5'-phosphonoformat is formed. The identification of this compound proves that the initially proposed pathway cannot take place. It also can be viewed as the activated intermediate for the transfer of the phosphonoformate group. This fits well with the idea that the *PhpG* and *PhpH* proteins could work in concert to produce CPEP in vivo via a reaction series analogous to the phosphoglycerate mutase and enolase reactions of glycolysis. *PhpG* is a close homolog of an autophosphorylating phosphoglycerate mutase from the archaeon *Sulfolobus solfatericus*. It was suggested that *PhpG* catalyzes an analogous reaction using the activated CMP-5'-PF intermediate to donate phosphonoformate to the active site of the enzyme. The enzyme-bound phosphonoformate could then be donated to a compound such as 3-phosphoglycerate, thereby yielding a phosphonoformylated

Intermediate that would then serve as the substrate for the *PhpH* enolase, creating CPEP.



phosphoenolpyruvate phosphonopyruvate phosphonoacetaldehyde

hydroxyethylphosphonate hydroxymethylphosphonate phosphonoformaldehyde

phosphonoformate CMP-5'-phosponoformate 2-phosphonoformyglcerate

Carboxyphosphonoenolpyruvate

Figure 19 Steps 1- 9

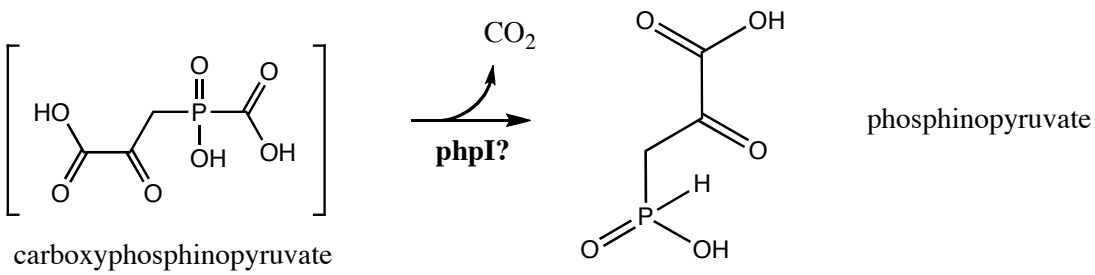


Figure 20: Step 10

Carboxyphosphoenolpyruvate (CPEP) is the substrate for CPEP phosphoenomutase in a reaction that yields phosphonopyruvate as a product. It is unknown whether the decarboxylation of the presumed Carboxyphosphoenolpyruvate intermediate of this reaction is the result of enzyme catalysis or inherent product instability despite *in vitro* study of the enzyme; further enzymatic studies may help clarify the mechanism. CPEP phosphoenomutase is encoded by bcpA in *S. hygroscopicus*, corresponding to phpI in *S. viridochromogenes*.

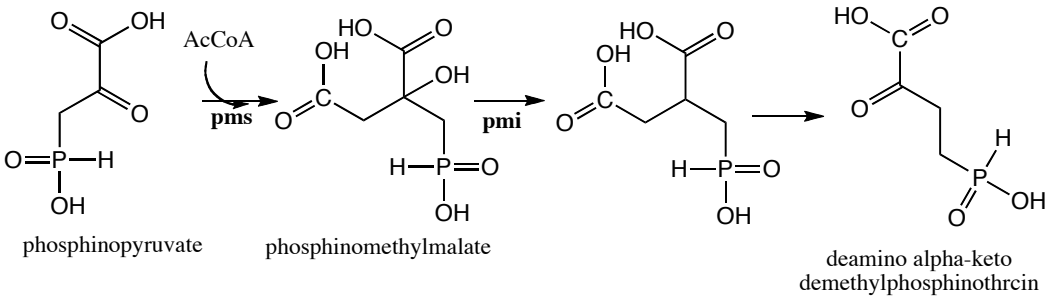


Figure 21: Originally proposed Pathway

Mutants that blocked steps were unable to be made. Another approach was to assume that because the special volume of hydrogen is very small, which allows –PH (=O) OH to be considered an analogue of -C (=O) OH. Thus, it was assumed that PPA and DMPT were analogues of oxalocetic acid, and glutamic

acid thus making this reaction similar to the TCA cycle. Aconitase the catalyst for the first reaction in the TCA cycle is inhibited by monofluroacetic acid. Addition of monofluroacetic acid to the culture resulted in the inhibition of BA synthesis, resulting in an intermediate phosphinomethylmalic (PMM) acid, which bears a structural resemblance to citric acid. This shows that the PPA conversion to PMM is catalyzed by either the citrate synthase or a related enzyme. The related enzyme was later identified as phosphinomethylmalate synthase (PmmS) in *S. hygroscopicus*, a homolog of the *S. viridochromogenes* pms gene product. Phosphinomethylmalate isomerase, the *pmi* gene product, was previously shown to rearrange the structure of phosphinomethylmalate for subsequent oxidation and decarboxylation by an unknown enzyme into deamino-keto-demethylphosphinothricin (DAKDMPT). The enzyme responsible for step X has not been identified. It was previously predicted that this reaction would take place by an enzyme similar to (or perhaps identical to) isocitrate dehydrogenase[32].

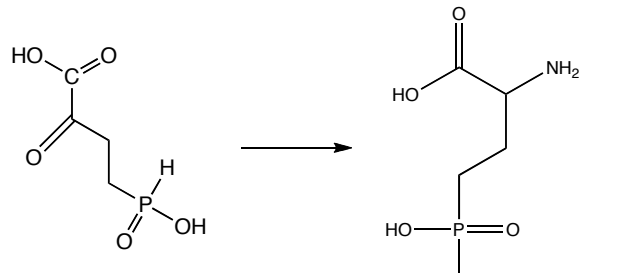
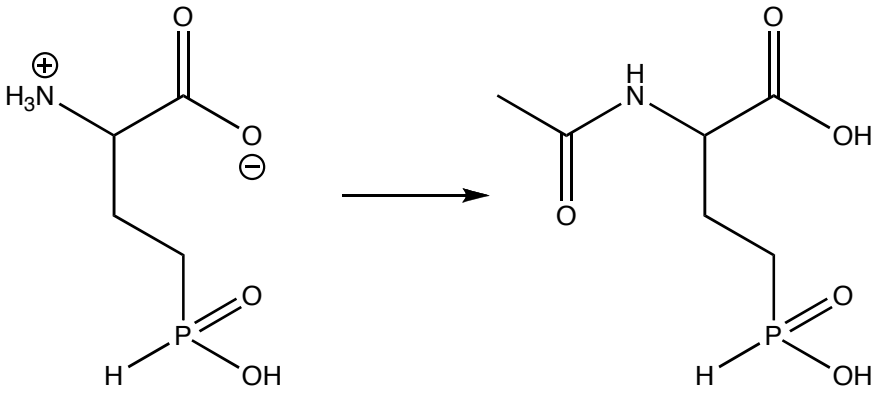


Figure 22:Step 11

Likewise, an aminotransferase homolog was also not found which would likely be required for the conversion of DAKDMPT to demethylphosphinothricin. Unpublished results cited by *Seto and Thompson*[29] indicate that both of these steps could be catalyzed by microorganisms that do not produce PTT; thus, it is probable that these steps are catalyzed by ubiquitous, generic enzymes that can be found in most microorganisms.



Demethylphosphinothricin N-Acetyl demethylphosphinothricin

Figure 23: Step 12

It has previously been shown that the acetylation of demethylphosphinothricin (step XII) is catalyzed by demethylphosphinothricin *N*-acetyltransferase, corresponding to the *pat* [33]or the homologous *S. hygroscopicus bar* gene product.[34]

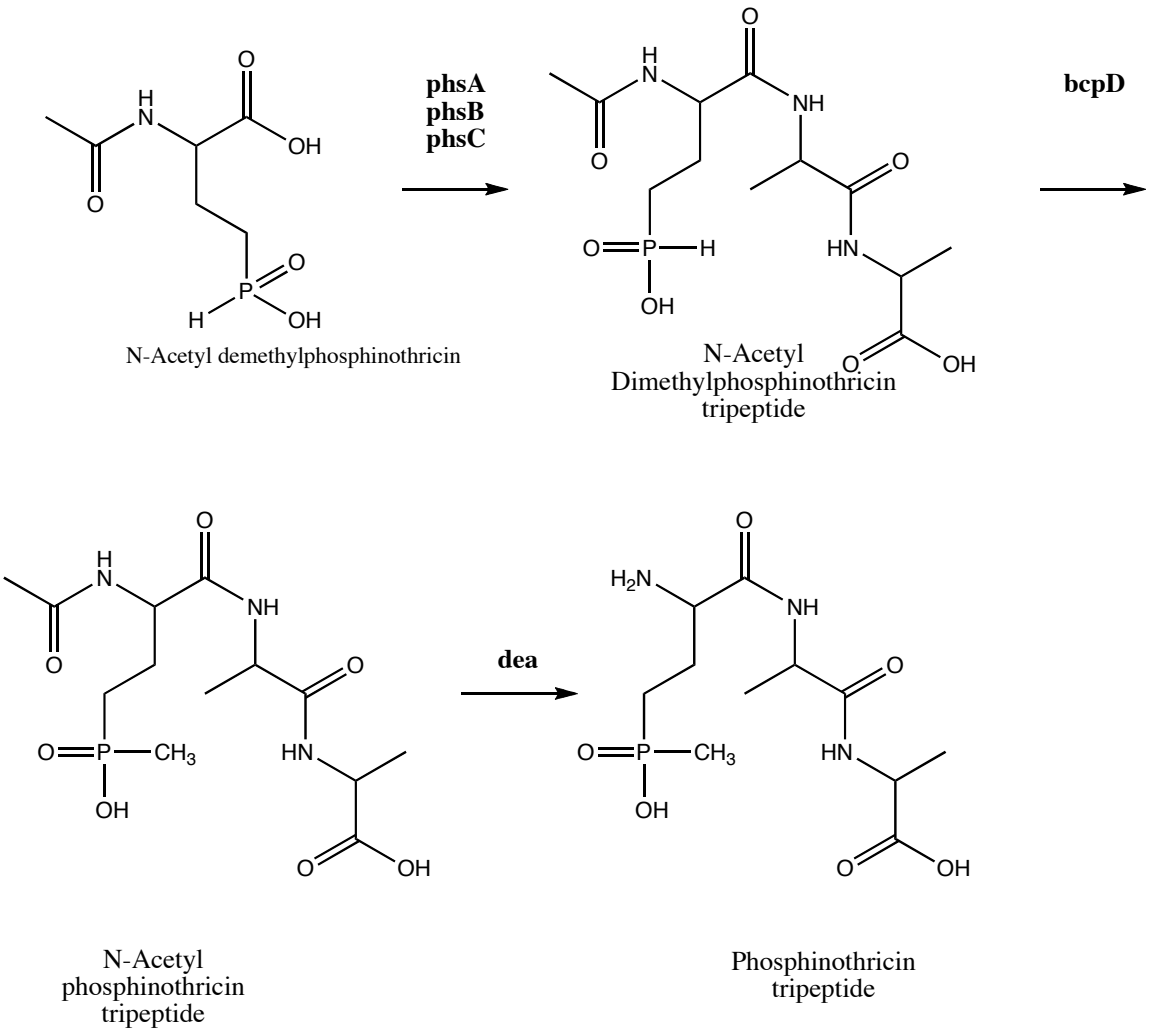


Figure 24: Steps 13- 15

N-Acetyltransferase activity provides the substrate for the alanylation Steps as well as a mechanism of detoxification against free phosphinothricin [35,36]that may be produced within the cell. After *N*-acetylphosphinothricin tripeptide is nonribosomally synthesized, it has been shown that the *phpK* homolog from *S. hygroscopicus*, *bcpD*, encodes the *P*-methyltransferase that creates the second C-P bond [37] [38]yielding *N*-acetylphosphinothricin tripeptide (step XIV). The final step in PTT biosynthesis was found to be the deacetylation of *N*-

acetylphosphinothricin tripeptide (step XV) by the *dea* gene product [39,40] to produce the intact PTT molecule.

The addition of the two alanine residues to Phosphinothricin, producing PTT, has been shown to occur by a nonribosomal peptide synthesis mechanism [41] and a large segment of the minimal gene cluster is dedicated to nonribosomal peptide synthesis activities.

The nonribosomal biosynthesis of PTT differs from that of the known bacterial systems, for one the three NRPS modules are not clustered together but rather scattered in the genome. [42] Sequence analysis has revealed that all three peptide synthetase genes encode only one peptide synthetase module. An *in silico* analysis, as well as biochemical and genetic experiments, proved that *PhsA* is responsible for the activation of the first amino acid, N-Ac-DMPT, the precursor of PT, whereas *PhsB* and *PhsC* each activate one alanine[43]. PhsA consists of A and T domains; PhsB of T, C, and A domains; and PhsC of C, A, and T domains. Currently, it is not possible to determine which of the two alanylylation steps is catalyzed by *PhsB* and *PhsC*. The fact that the two enzymes cannot replace each other suggests a defined positioning for each protein in the PTT assembly line. However, in PTT synthetases, short communication-mediating domains (COM domains) that mediate interactions between peptide synthetases could not be identified.

A typical TE domain at the C terminus of PhsB or PhsC is missing. However, a highly conserved TE GXSXG motif at amino acids (aa) 14 to 18 was lo-

calized at the N-terminus of *PhsA*. Furthermore, the two genes *theA* and *theB* were identified, and their gene products showed high similarity to type II TEs (TEII). TEIIs are autonomous, monomeric proteins, while TEIs typically are integrated in the last module of peptide synthetases. Within polyketide and nonribosomal polypeptide biosyntheses, it was shown that TEII enzymes have editing roles. In most cases, along with the TEI, only one autonomous TE-encoding gene is part of the cluster. Based on these facts, two alternatives for the release of PTT from the peptide synthetase complex can be postulated.

In the first, the TE motif in PhsA is functional. In this case, the three peptide synthetases arrange in a manner such that the C terminus of PhsB or PhsC and the N terminus of PhsA can interact, and the tripeptide can be cleaved off from

The peptide synthetase complex with the participation of the TE motif.

In the second model, the TE motif is not functional. In this case, one of the TEIIs of PTT biosynthesis is responsible for the release of the tripeptide from the peptide synthetase complex. The other TE would have editing roles in PTT biosynthesis specific sequences that exhibit the function of a COM domain

And thus are responsible for the selective interaction of PhsA,

PhsB and PhsC.

Fosfomycin

Fosfomycin representing the first class of C-P bond containing compounds was first isolated from the fermentation broths of *Streptomyces fradiae.* [44] Fosfomycin is a clinically useful antibiotic for the treatment of lower urinary tract infections and limb-threatening diabetic foot infections. The antimicrobial activity of fosfomycin has been attributed to the inactivation of UDP-GlcNAc-3-O-enolpyruvyltransferase (MurA), which catalyzes the biosynthesis of an important cell wall component. [45]

The biosynthetic pathway of fosfomycin is being explored and thus far, minimal gene cluster elucidation has occurred. The biosynthetic pathway has only been fully elucidated in *S. wedmorensis,* however it has never been heterogeously expressed.[46] The genes, *fom1-4,* are thought to be responsible for the first few steps of involving the conversion of PEP into phosphonopyruvate (PnPy) by PEP mutase, followed by a decarboxylation to form phosphonoacetaldehyde (PnAA). These first steps are shared be a number of C-P containing compounds including bialaphos, AEP, and 2-hydroxyethylphosphonate (HEP). [47]

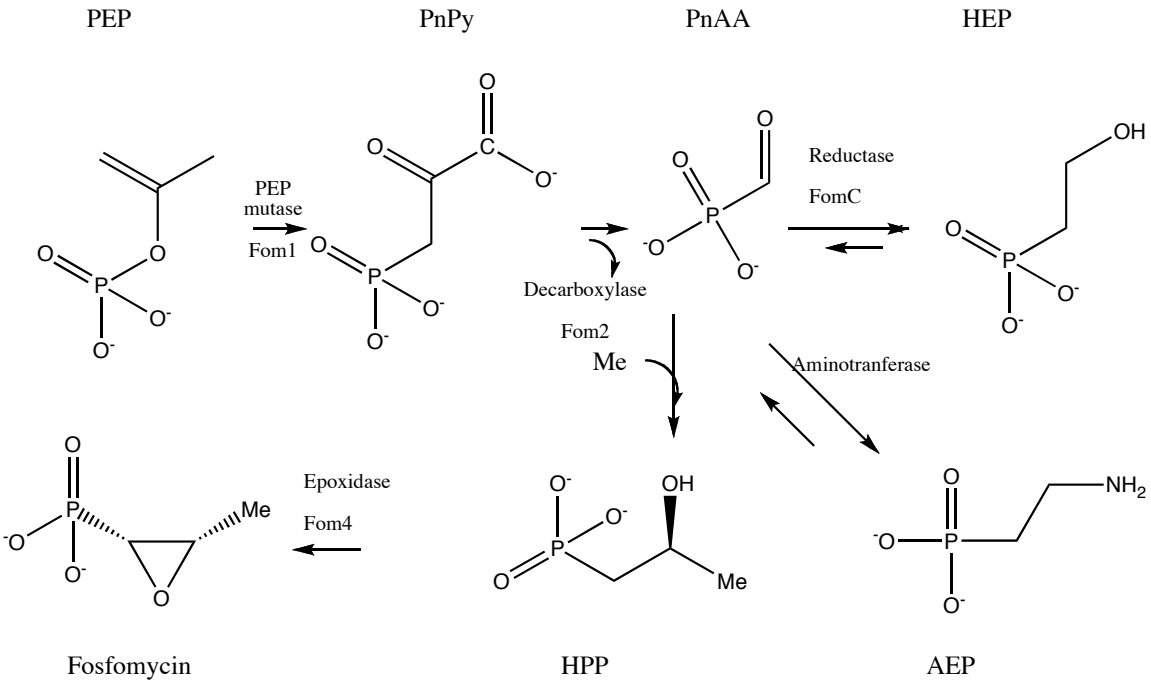


Figure 25: Proposed Biosynthetic pathway of Fosfomycin

From there, further genes were discovered using mutagenesis. The mutated clones were then plated for activity. *S. Fradiae* genomic DNA containing 21 putative Orfs were analyzed. Gene disruption analysis coupled with bioinformatics studies narrowed down the number of Orfs involved to 13 total Orfs including what we believe to be the minimal gene cluster *fom1-4, fomA-D,* and *fomR.* Thus, two genes previously thought to be involved in biosynthesis *fomE* and *fomF*, were found not to be required and a proposed transcriptional activator (*fomR*) was discovered [26]. *FomR* was found to be a regulator for the production of fosfomycin. Additional studies are preformed to determine the functions of *fomC* and D. The biochemical action of both *FomA* and *FomB* from S. wedmorensis as phosphotransferases that inactivate fosfomycin has been demonstrated

The third class of compounds represented by one single compound K-26

has not been heterogeously produced nor has the biosynthetic pathway been elucidated. The next chapter will present the discovery of the K-26 compound as well as discuss the structure moiety of AHEP as well as its ACE inhibitory activity.

References

1. Xiao, Y.; Lee, K.; Liu, P., Syntheses of the P-Methylase Substrates of the Bialaphos Biosynthetic Pathway. *Organic Letters* 2008, *10* (24), 5521-5524.

2. Woodyer, R. D.; Shao, Z.; Thomas, P. M.; Kelleher, N. L.; Blodgett, J. A. V.; Metcalf, W. W.; van der Donk, W. A.; Zhao, H., Heterologous Production of Fosfomycin and Identification of the Minimal Biosynthetic Gene Cluster. 2006, *13* (11), 1171-1182.

3. Seto, H.; Sasaki, T.; Imai, S.; Tsuruoka, T.; Ogawa, H.; Satoh, A.; Inouye, S.; Niida, T.; Otake, N., Studies On The Biosynthesis Of Bialaphos , Isolation Of The First Natural Product With A C-P-H Bond And Their Involvement In The C-P-C Bond. *Journal of Antibiotics* 1983, *36* (1), 96-98.

4. Seto, H., Bioactive natural products with carbon–phosphorus bonds and their biosynthesis. *Nat. Prod. Rep.*, 1999, **16**, 589–596

5. Thompson, C. J., Bialaphos. *Biotechnology* 1995, *28*, 197-222.

6. Blodgett, J. A. V.; 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), 480-485.

7. Blodgett, J. A. V., 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.

8. Heinzelmann, E., G. Kienzlen, S. Kaspar, J. Recktenwald, W. Wohlleben, and D. Schwartz. 2001. The phosphinomethylmalate isomerase gene pmi, encoding an aconitase-like enzyme, is involved in the synthesis of phosphinothricin tripeptide in Streptomyces viridochromogenes. Appl. Environ. Microbiol.67:3603–3609.

9. Strauch, E., W. Wohlleben, and A. Pu¨hler. 1988. Cloning of a phosphinothricin N-acetyltransferase gene from Streptomyces viridochromogenes Tu¨494 and its expression in Streptomyces lividans and Escherichia coli. Gene 63:65–
74.

10. Thompson, C. J., R. N. Movva, R. Tizard, R. Crameri, J. E. Davies, M. Lauwereys, and J. Botterman. 1987. Characterization of of the herbicide gene bar from Streptomyces hygroscopicus. EMBO J. 9:2519–2523.

11. Kumada, Y., H. Anzai, E. Takano, T. Murakami, O. Hara, R. Itoh, S. Imai, A. Satoh, and K. Nagaoka. 1988. The bialaphos resistance gene (bar) plays a role in both self-defense and bialaphos biosynthesis in Streptomyces hygroscopicus.J. Antibiot. 41:1838–1845

12. Wohlleben, W., W. Arnold, I. Broer, D. Hillemann, E. Strauch, and A. Pu¨hler. 1988. Nucleotide sequence of the phosphinothricin N-acetyltransferase gene from Streptomyces viridochromogenes Tu¨494 and its expression in Nicotiana tabacum. Gene 70:25–37

13. Hidaka, T.; Hidaka, M.; Kuzuyama, T.; Seto, H., Sequence of a P-.methyltransferase-encoding gene isolated from a bialaphos-producing Streptomyces hygroscopicus. *Gene* 1995, *158* (1), 149-150.

14. Hidaka, T., M. Hidaka, T. Kuzuyama, and H. Seto. 1995. Sequence of a P-methyltransferase-encoding gene isolated from a bialaphos-producing Streptomyces hygroscopicus. *Gene* 158:149–150.

15. Raibaud, A., M. Zalacain, T. G. Holt, R. Tizard, and C. J. Thompson. 1991. Nucleotide sequence analysis reveals linked N-acetyl hydrolase, thioesterase, transport, and regulatory genes encoded by the bialaphos biosynthetic gene cluster of Streptomyces hygroscopicus. J. Bacteriol. 173:4454–4463

16. Wohlleben, W., R. Alijah, J. Dorendorf, D. Hillemann, B. Nussbaumer, and S. Pelzer. 1992. Identification and characterization of phosphinothricin-tripeptide biosynthetic genes in Streptomyces viridochromogenes. Gene 115:127–132.

17. Grammel, N., D. Schwartz, W. Wohlleben, and U. Keller. 1998. Phosphinothricin-tripeptide synthetases from Streptomyces viridochromogenes. Biochemistry37:1596–1603.

18. Schwartz, D.; Alijah, R.; Nussbaumer, B.; Pelzer, S.; Wohlleben, W., The peptide synthetase gene phsA from Streptomyces viridochromogenes is not juxtaposed with other genes involved in nonribosomal biosynthesis of peptides. *Appl. Environ. Microbiol.* 1996, *62* (2), 570-577.

19. Schwartz, D.; Grammel, N.; Heinzelmann, E.; Keller, U.; Wohlleben, W., Phosphinothricin Tripeptide Synthetases in Streptomyces viridochromogenes Tu494. *Antimicrob. Agents Chemother.* 2005, *49* (11), 4598-4607.

20. Christensen, B. G.; Leanza, W. J.; Beattie, T. R.; Patchett, A. A.; Arison, B. H.; Ormond, R. E.; Kuehl, F. A., Jr.; Albers-Schonberg, G.; Jardetzky, O., Phosphonomycin: Structure and Synthesis. *Science* 1969, *166* (3901), 123-125.

21. Yan, F.; Munos, J. W.; Liu, P.; Liu, H.-w., Biosynthesis of Fosfomycin, Re-Examination and Re-Confirmation of a Unique Fe(II)- and NAD(P)H-Dependent Epoxidation Reaction. *Biochemistry* 2006, *45* (38), 11473-11481.

22. Hidaka, T.; Goda, M.; Kuzuyama, T.; Takei, N.; Hidaka, M.; Seto, H., Cloning and nucleotide sequence of fosfomycin biosynthetic genes of Streptomyces wedmorensis. *Molecular and General Genetics MGG* 1995, *249* (3), 274-280.

23. Hidaka, T., Detection of phosphoenol-pyruvate phosphomutase activity in a fosfomycin high-producing strain of Streptomyces wedmorensis and characterization of its blocked mutant NP-7. *Journal of Antibiotics* 1992, *45* (6), 1008-10.

CHAPTER III

ACE activity

Discovered in 1985, K-26, a novel inhibitor of Angiotensin I converting enzyme (ACE) was isolated from the broth filtrate of an actinomycete K-26. K-26 is a water soluble, acidic peptide composed of an equal mol of L-isoleucine, L-tyrosine, and l(R)-l-amino-2- (4-hydroxyphenyl)-ethylphosphonic acid.

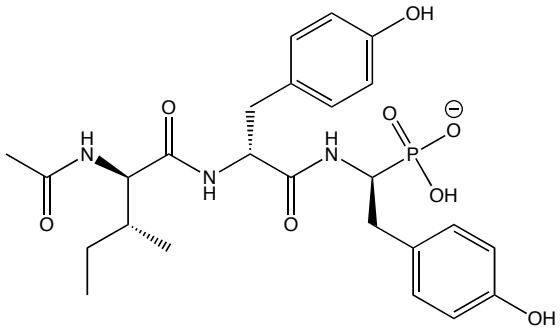


Figure 26: K-26

Angiotensin I converting enzyme (ACE) has been found to play a critical role in the regulation of blood pressure in man and animals. Ever since the discovery of Captopril, a specific inhibitor of ACE, and its successful application to the therapy of hypertension, attempts have been made to develop other ACE inhibitors for use as anti hypertensive agents for clinical use. As a result, a growing number of inhibitors of ACE, synthetic or microbial, have been discovered in recent years. During the course of screening for microbial inhibitors of ACE by

Yamato et al, an active substance, designated K-26, was isolated from the broth filtrate of an actinomycete K-26, and characterized chemically and biologically.

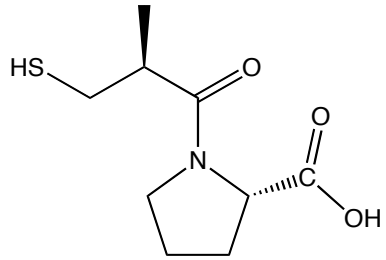


Figure 27: Captopril, Break Through Discovery Ace Inhibitor Drug

ACE inhibitor activity of K-26 was tested in vitro along with Captopril. K-26 proves to be just as potent as Captopril in ACE inhibition in vitro. The IC50 of K-26 was tested in two buffers 0.1 M phosphate buffer, pH 8.3, and 0.1 M borate buffer, pH 8.3, HHL or AG I being used as a substrate. Six determinations were made to obtain an IC50; the average values were represented in table 1.

Table 1: Activity of K-26

()= Captopril, - = not tested

Substrate	IC_{50} (ng/ml)	
	Phosphate	Borate
HHL	6.7(6.0)	12(12)
AG I	3.3(5.0)	-

Each drug was administered intravenously to an anesthetized, normotensive rat at the time indicated by an arrow at the following dose: Angiotensin I (AG I); 300 ng/kg, Angiotensin II (AG II); 100 ng/kg, noradrenalin (NA); 3 ug/kg, K-26; 0.1 mg/kg, Captopril (CAP); 0.1 mg/kg.[19]

It was shown that the introduction o f K-26 into the system had nearly identical affects on the blood pressure as those of Captopril. When ACEI, II, and NA are introduced to the system, a spike in the blood pressure of the rat is recorded. However; proceeding the introduction of the blood pressure raising compounds if either the K-26 of Captopril had been inject the spike caused by the ACE I was reduced. As with Captopril the K-26 compound only inhibits the affects of the type I ACE.

Yamato et al believed that the phosphonate moiety might interact with an active site zinc atom resulting in the inhibitory activity of the compound. However, other small peptides without C-P bonds possess ACE activity. [48] *Wu Et Al* [48] constructed a database of 168 dipeptides and 140 tripeptides. According to the models, the IC50 values of seven new peptides with matchable primary sequences within pea protein, bovine milk protein, and soybean were predicted. The predicted peptides were synthesized, and their IC_{50} values were validated through laboratory determination of inhibition of ACE activity.

To further ascertain the structural component responsible for the biological activity, *Ntai et al* synthezed K-26 and a multiple of analogues (table 2) to determine the affects [49]. ACE activity was extracted from rabbit acetone powder by soaking with 1oomM borate buffer followed by ultra-centrification. Furylacryloyl-Phenylalanyl-Glycyl-Glycine (FAPPGG) was then added. The reaction can be monitored can be monitored based on the hyposcopic shift of the adsorption spectra due to the hydrolysis of FAPGG to FAP and GG. ACE extract was pre-

incubated with a range of inhibitor concentrations in the responsive concentration region for 5 min then 1 mM solution of FAPGG was added. The rate of FAPGG hydrolysis was measured by the rate of hydrolysis by the change in absorbance at 340 nm/time. IC_{50} values were obtained by fitting triplicate measurements to a curve.

Table 2: Synthesized Compounds

Compound	IC_{50}
(K-26) Ac-l-lle-l-Tyr-R-AHEP	14.4
Ac-l-lle-l-Tyr-S-AHEP	139
Ac-l-lle-l-Tyr-R-AHEP $(OEt)_2$	3.83×10^{-5}
L-lle-l-Tyr-R-AHEP	234.3
L-lle-l-Tyr-S-AHEP	5.46×10^{4}
Ac-l-lle-l-Tyr-l-Tyr	2.1×10^{4}
Ac-l-lle-l-Tyr-d-Tyr	2.0×10^{6}
L-lle-l-Tyr-l-Tyr	2.36 × 105
L-lle-l-Tyr-d-Tyr	10^{6}
Captopril	7.7

As demonstrated by the chart the addition of the acetylation of analogues resulted in an increase of activity while the inversion of the AHEP also resulted in a decrease in activity. The inclusion of the AHEP moiety in the analogues improved activity across the board. The only analogue with improved activity was the adduct with the modified AHEP moiety Ac-l-lle-l-Tyr-R-AHEP $(OEt)_{2.}$ This data demonstrates that the affects of ACE inhibition may be due to the strong ionic interaction of the phosphonates anion with the enzyme.

Structure elucidation

The structure of K-26 suggests the three amino acids are condensed using a NRPS like pathway, due to the presence of the amine bonds and the non-proteogenic amino acid. However, it is unclear as to whether the AHEP moiety is modified post or pre adenylation. The aromatic amino acid functionality of AHEP suggests that its origin lies in the shikimic acid pathway. However, it is difficult to rationalize how a PEP mutase generated precursor or analogous mutase reaction can be integrated into classical amino acid metabolism to generate AHEP. To discover the biosynthetic origin of the AHEP, a series of incorporation studies with isotopically labeled tyrosine was undertaken by *Ntai et al*[50]. The labeled tyrosine was introduced by pulse feeding growing cultures of Actinomycete sp. K-26 (1 mM/day for 5 days) with ring-d_4-tyrosine, 3,3-d_2-tyrosine, and ^{15}N-tyrosine(figure 4). The K-26 was then purified and analyzed using MS/MS and subsequently isotopic enrichment measurements using a triple quadrupole mass spectrometer using electrospray ionization (ESI), collision-induced dissociation (CID), and selected reaction monitoring (SRM).

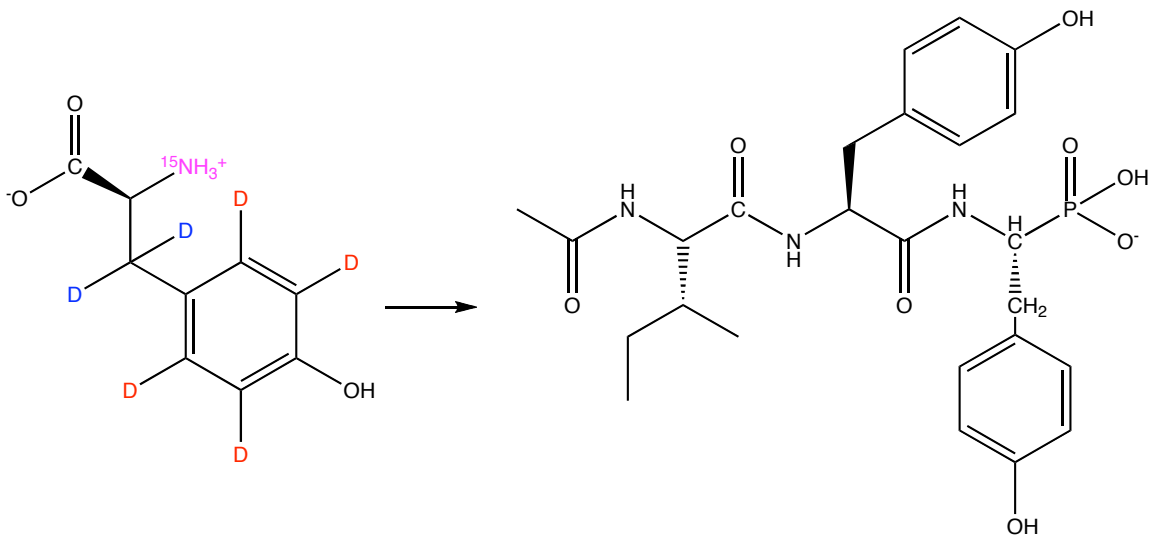


Figure 28: Incorporation Of Labeled Tyrosine Into The AHEP Moiety Of K-26

The data suggests (table 2) that enzymatic phosphonylation occurs after tyrosine biosynthesis and does not proceed via an eliminated intermediate. The AHEP moiety is the result of modifications to Tyr; this study determined that Tyr is an immediate precursor to AHEP.

Table 2: Incorporation Studies On K-26:

a. Determined by subtraction of percentage Ile from percentage Ile-Tyr.
b. Determined by subtraction of percentage AHEP from percentage Tyr-AHEP.

Precursor	Ile (%)	Tyr (%)	AHEP (%)
d_4-tyr	Nd	17.6+	20.2 +/- 0.2
d_2-Tyr	Nd	16.1 +/- 0.6[b]	18.7+/-0.6
15N-Tyr	2.6+/- 0.1	6.2+/- 0.6[a](5.4+/- 1.4)[b]	7.1+/- 0.4

However, these studies were unable to determine if there is a free AHEP intermediate involved, when the C–P bond actually forms and the acetylation of the

Ile. Another study by *Ntai et al* [51] attempts to determine whether AHEP is a free intermediate, the timeline for C-P bond formation and the acetylation of the N terminus. To aid in these determinations $^{13}C_2$ labeled AHEP was synthesized for the purpose of isotopic incorporation studies.

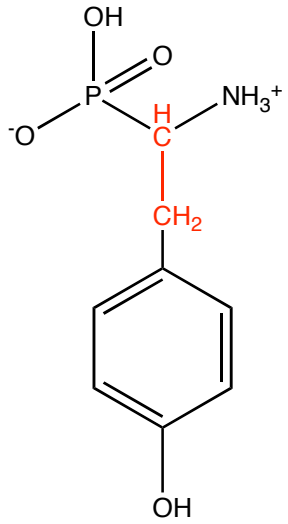


Figure 29: Labeled AHEP

Liquid cultures were fed the isotopically labeled and unlabeled racemic mixtures of AHEP, the unlabeled being used as a control for growth. The incorporation of AHEP was monitored using methods described in the previous paper. A two atomic mass difference between the enriched(536) and unenriched (534)K-26 molecule verifies incorporation of the labeled AHEP. When the data was deconvoluted, it suggested that the labeled K-26 was incorporated at about 85%. This

indicates that AHEP is a discrete precursor and that the substrate of the C-P bonds forming enzyme of Tyr or an analogue.

Decarboxylation is necessary for incorporation of tyrosine into AHEP, they also synthesized labeled tyramine to determine if perhaps that was the direct precursor. Incorporation studies done using the labeled tyramine failed, leading to the belief that tyrosine is a more direct precursor. This suggests that the decarboxylation is combined with the C-P bond formation. To determine the time line for the acetylation of the N- terminus, labeled isoleucine and labeled acetylated isoleucine were synthesized for another feeding study. The cultures were prepared in the same method as before and the data analyzed as before. The data shows that the acetylated Ile was incorporated with one addition mass unit at 43% while the Ile was incorporated at a little more than half of the acetylated compound. The acetylated Ile was only incorporated after being deacetylated, the high level of incorporation of the acetylated Ile vs. the unaceltylated may be the result of the ability of the intracellular transport of the different compounds. This indicates that the acetyltransferase may be a component of the NRPS and potentially the first domain. The data also suggest that the acetylation is independent of the NRPS functionality and occurs after the tripeptide has been released from the Te domain. A third option is the formation of the tripeptide by two free synthetases, in this instance the acetylation occurs after the peptide bond.

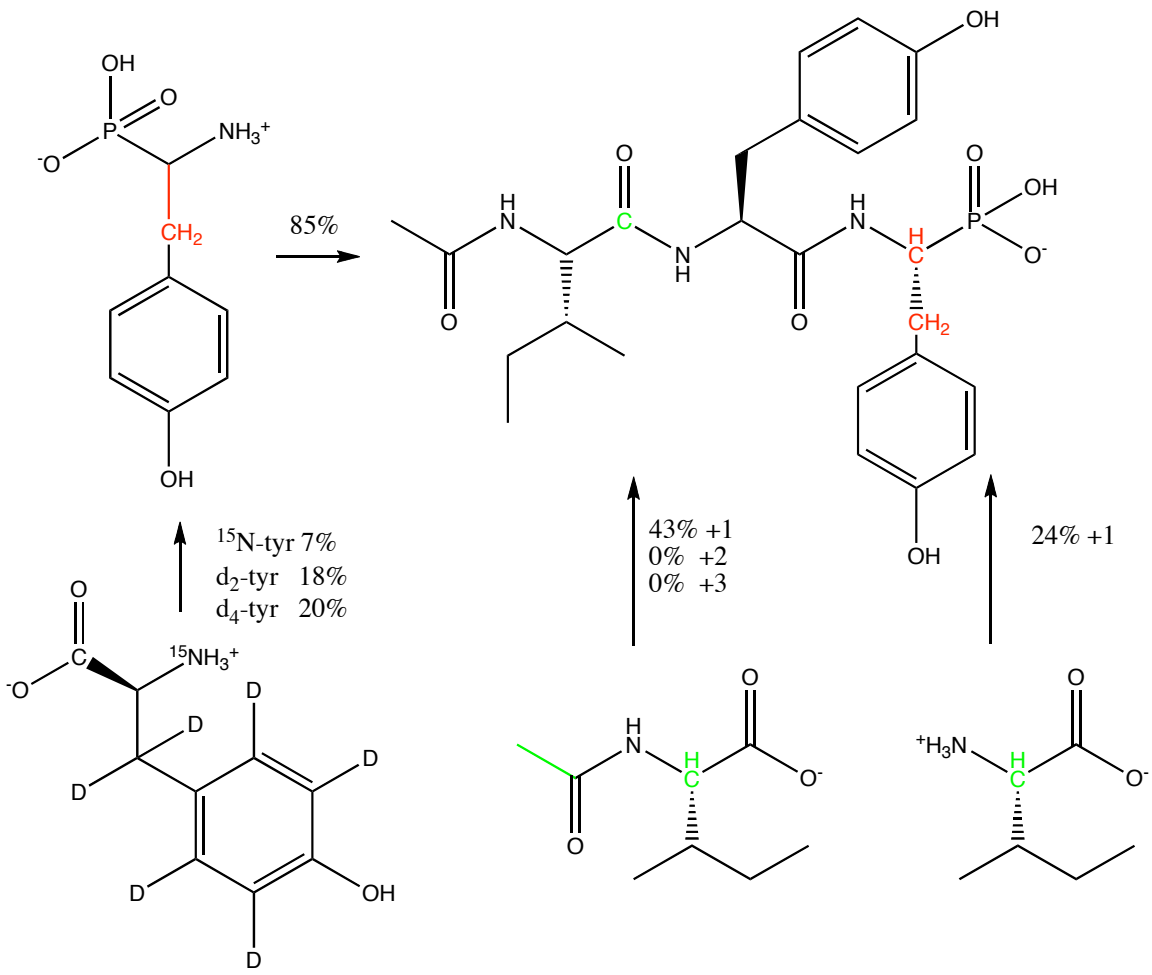


Figure 30: Summary Of Incorporation Studies

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 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. Journal of Agricultural and Food Chemistry 2006, 54 (3), 732-738.

3. Ntai, I.; Bachmann, B. O., Identification of ACE pharmacophore in the phosphonopeptide metabolite K-26. Bioorganic & Medicinal Chemistry Letters 2008, 18 (10), 3068-3071.

4. 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.

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

CHAPTER IV

Introduction

We hypothesize that there is a NRPS system responsible for the biosynthesis of K-26. K-26 an acetylated tripeptide appears to be the result of an NRPS like mechanism, which is evident by the presence of the non-proteogenic amino acid AHEP, as well as the previous incorporation studies. The presence of three peptides leads us to look for three modules. The theoretical organization of domains is shown in figure 33.

To verify the entire genome of the K-26 was shot gun sequenced and BLASTed to aid in locating genes that might be responsible for the biosynthetic pathway of K-26. Areas of high interest were identified in particular NRPS modules, which might encode for acetylation domains and tripeptide producing domains. Blasted sequences did not yield any obvious targets for the gene encoding of K-26 thus the method selected was the over expression of adenylation domains for the purpose of determine which area is responsible for the production of K-26. The activity was verified using a novel MS technique that assays for adenylation activity.

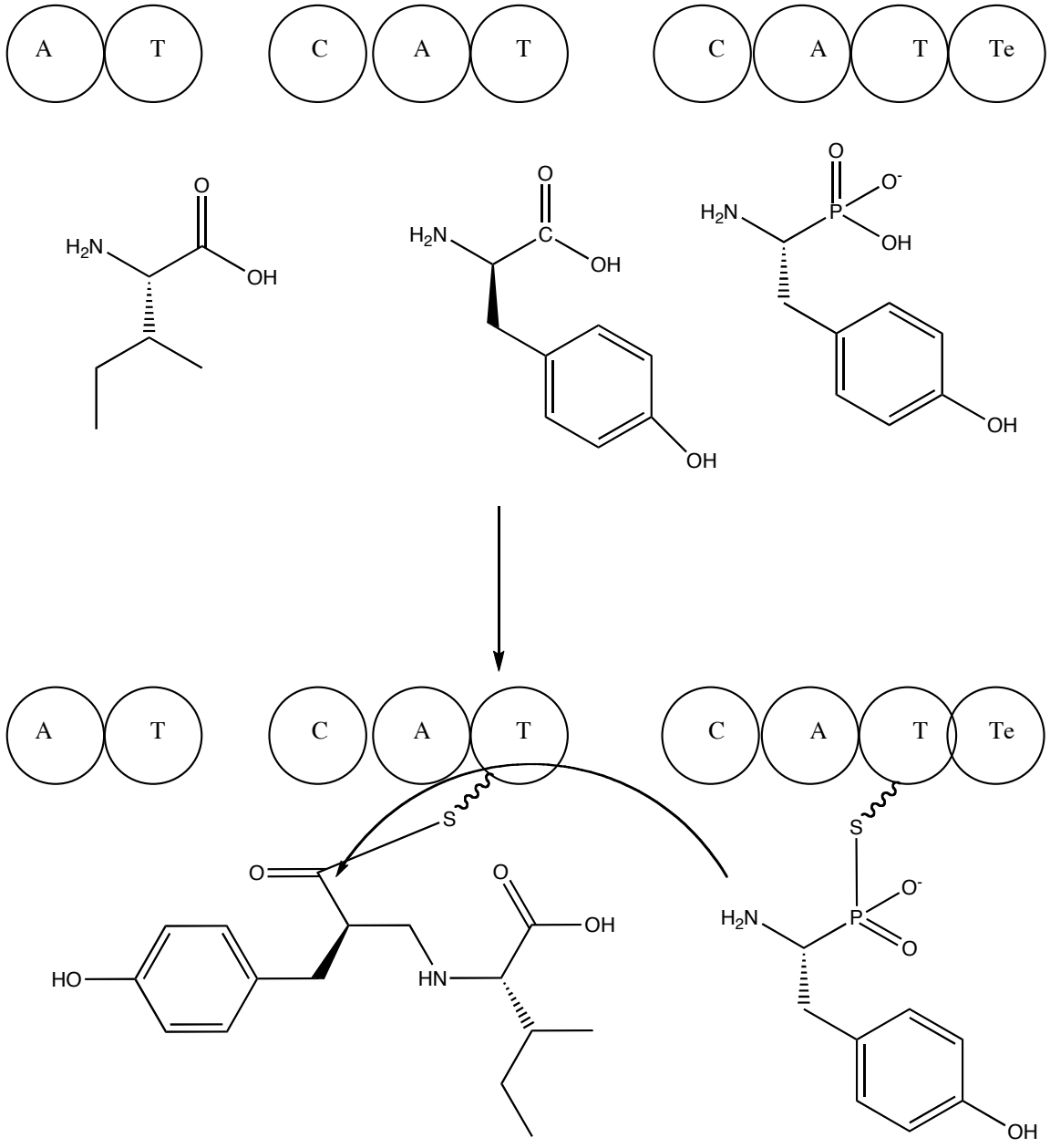


Figure 31: Theoretical Biosynthetic Pathway Of K-26: Three Modules

Looking through the genome of the K-26 species yielded no obvious areas so areas that encoding for NRPS systems were scrupulously examined. Using

Challis 8-amino acid code prediction, areas that were predicted to encode for Ile and Tyr were selected for over expression in *E coli.*

Materials and methods

All primers were purchased from Operon. All enzymes were purchased from the New England Biolab (Ipswich, MA). Chemicals were purchased from Sigma Aldrich (St. Louis, MO). The SUMO vector was purchased from Invitrogen (Carlsbad, CA). The Vanderbilt Chemical Biology Core provided the pBg102 SUMO vector. The Vanderbilt Sequencing Core preformed all sequencing. Media and antibiotics were all purchased from Research Products international or Sigma Aldrich. All SDS gel materials were purchased from Bio-Rad (Hercules, CA). TOPO cloning was preformed using materials from Invitrogen. Both the BL-21 (DE3) strain and the ROSETTA strain were purchased from Novagen (San Diego, CA).

MALDI and ESI MS samples were run by or with the aid of Vanessa Phelan. MALDI-TOF mass spectrometric analysis was performed on a Voyager-DE™ STR Biospectrometry MALDI-TOF MS Workstation. ESI-LC/MS mass spectrometric analysis was performed on a ThermoFinnigan LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific, Waltham, MA) equipped with an ESI interface in negative ion mode

Pps 8 A3

Primers [table 3] were designed with the cleavage sites for the enzymes Xbal and Bsa and for cloning into the SUMO vector. The SUMO vector has a large fusion protein as well as six His tags on the N terminus. They were then amplified for 30 cycles using the concentrations and conditions found in tables 2 and 3. The 1.7Kb PCR product was then ligated into TOPO vector and transformed using electro competent top10 cells and plated for overnight growth on LB plates with kanamycin (50ug/mL).

Table 3 Primers

Primer	Name
TAAGGTCTCAAGGTATGGCGCCGCTGACCCC	2754 FP:1
TAAGGTCTCAAGGTATGGCGCCGCTGACC	2754 FP:2
TAATCTAGATCACTCACGGTCGTCGCG	2754 RP:1
ATAGGATCCATGACCGGCCCCTTCGAG	2753 FP1
ATAGGATCCATGACCGGCCCCTTCGA	2753 FP2
ATAAAGCTTTCATGCGCGTGCGTCGAG	2753 RP1
ATAAAGCTTTCATGCGCGTGCGTC	2753 RP2

Table 4: concentration of PCR reaction

	concentration	Amount used(uL)	Final concentra-tion
Forward primer	1uM	6	
Reverse primer	1uM	6	
Taq plus enzyme	5units/uL	.5	2.5 units
DMSO	-	3	6%
Taq Buffer	10X	5	1X
Water	-	37	-
DNA template	-	.5	-
dNTPs	10mM	1.5	.5uM

Table 5:PCR reaction conditions

cycles	Temperature (Celsius)	Duration (minutes)
1	94	5:00
30	94	0:30
	52	1:30
	72	2:30
1	72	10:00

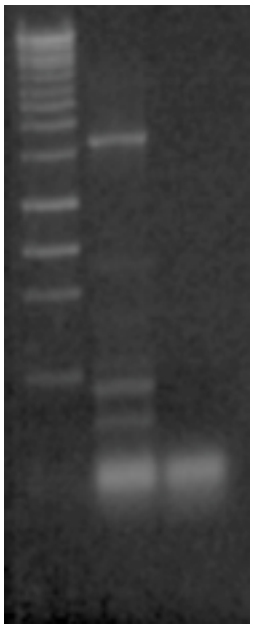


Figure 32: PCR (lane 1- marker , lane 2 - combination of FP1 and RP, lane 3 combination of FP2 and RP)

white overnight colonies were then selected inoculated for over night growth in 2-5mL of LB media and Ampicillin (100ug/mL) . The resulting growths were lysed and digested to confirm the presence of the insert. The insert is then ligated into the expression vector with the SUMO protein.

The new plasmid is then transformed into both BL21 DE3 and Rosetta for expression. Transformation was verified once again via digestion under the following conditions. Expression was first preformed using 250mL cultures grown to OD_{600}= 0.4- 0.6. Then induced using IPTG, under various conditions to determine the best production of protein. Cells were then selected for growth in a 2.8-L baffled flask containing 500 mL LB medium with 100 µg/mL Ampicillin , at 17 degrees with 0.2mM of IPTG for 16 hours in the Rosetta strain. The cells were pelleted (1hr, 3750 rpm, 4°C) and resuspended in binding buffer (500m M NaCl, 20 mM NaH_2PO_4, 20 mM imidazole, pH 7.4). For cell lysis, DNase I was added.

Using a French Press, the cells were lysed and then filtered through a 0.45-µm filter. The protein was affinity purified on a HisTrap FF column on an ÄKTA chromatography system (GE Healthcare) using binding buffer with graduated step increasing imidazole concentration (20–500 mM). The pure protein was then desalted with a HiTrap Desalting column using 20mM Tris, pH7.5 and stored in aliquots. Protein concentrations were measured with Pierce® BCA Protein Assay Kit (Thermo Scientific).

Expression in Rosetta at 17 degrees and induced with .2mM IPTG, and Ampicillin for 16hrs.

75kB ---------->

60kB ---------->

Final
Concentration:
915ug/ml

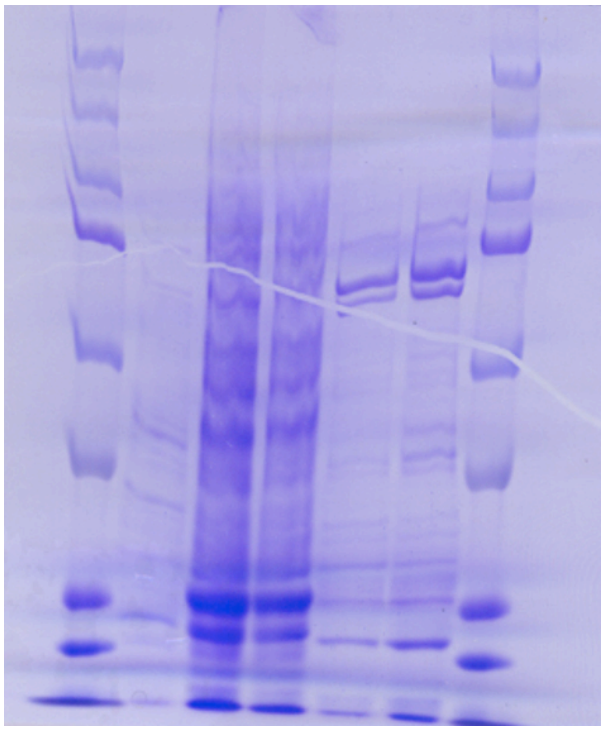


Figure 33: SDS gel lane A: Marker B: uninduced Cell free extract(CFE) C: induced CFE D: Flow through C: Bound concentrated F: Desalted concentrated G: Marker

The protein in the gel above shows two different sizes that cannot be separated and it was bard to determine which protein is indeed ours. This band can be attributed to background production by the strain as the top band is present in the uninduced lane as well.

Pps8 A2

Primers were also purchased and designed with BamHI and HindIII cleavage sites for use in the Vanderbilt Core SUMO vector . The concentrations for the PCR reaction are the same as previously stated. Primers were amplified using the PCR protocol found in table with. The 1.7 Kb PCR product was ligated and

transformed into TOPO vector as per previous protocol. Once again, the plasmid was digested and ligated into the SUMO vector. Then is was transformed into both the Rosetta and the BL-21 DE3 expression strains. Protein was expressed and purified using the same procedure as mentioned previously under the following conditions: 37 degrees for 16 hrs with 5mM IPTG for 16hrs with kanamycin. (50 uL/mL)

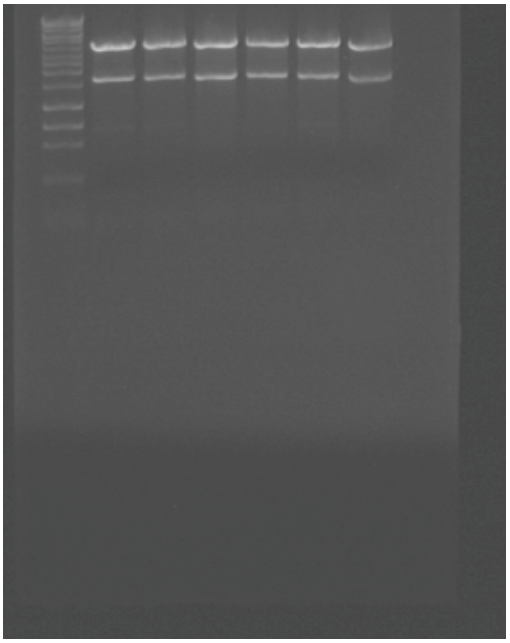


Figure 34: Digestion Of 2753 In Bl-21 (De3) Vector With Insert

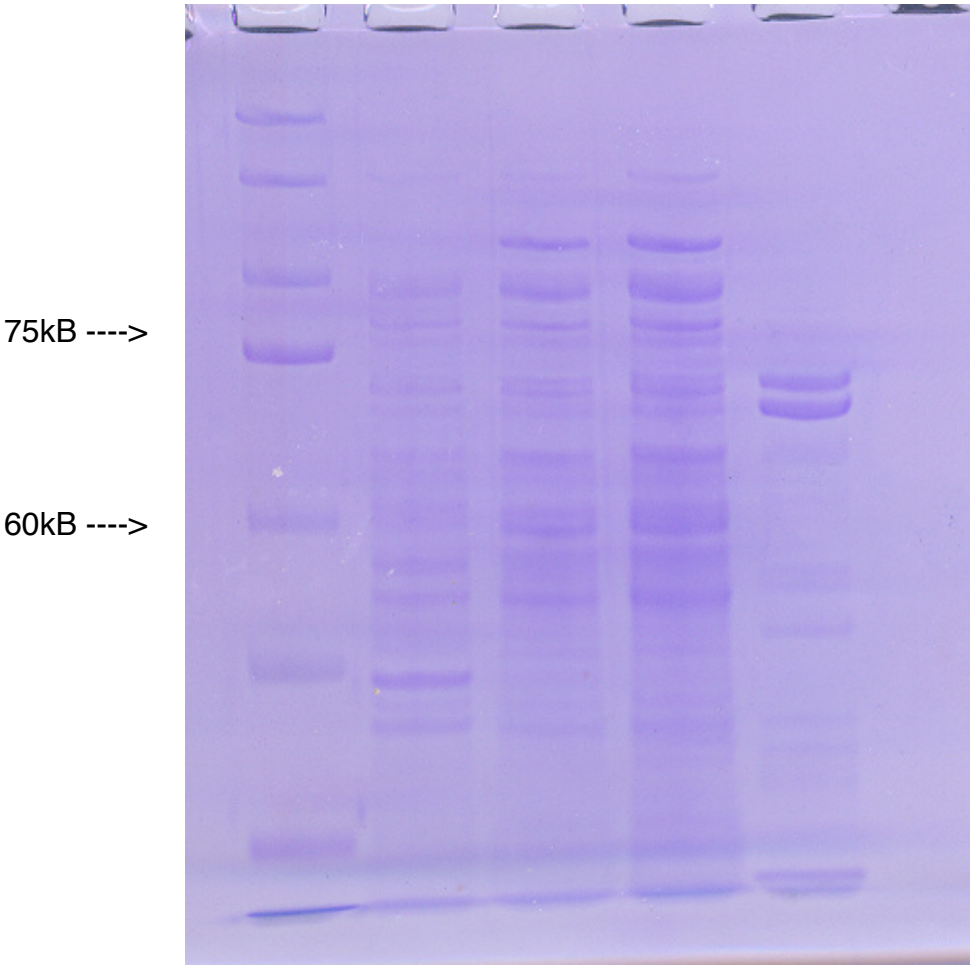


Figure 35: Pps8 A2: Lane A- Marker, Lane B- Uninduced, Lane C- Induced, Lane D- Flow Through, Lane E- Concentrated Desalted Purified Protein

Assay

The activity of the first concentrated protein was then tested for activity using an unpublished method by Phelan et al. involving the use of heavy labeled ATP and ESI MS to record the exchange between the protein and the ATP.

By using y-$^{18}O_4$-ATP isotopically labeled ATP and incubating it with the adenylation protein, and different amino acids we are able to track the exchange through us of MALDI or ESI. the reversible formation of aminoacyl adenylate by the Adenylation domain with the use of the labeled of y-$^{18}O_4$-ATP, can be used to generate unlabeled ATP by the addition of unlabeled Pi.

$$\gamma\text{-}^{18}O_4\text{-ATP} + AA \underset{Mg^{2+}}{\overset{\text{A-domain}}{\rightleftharpoons}} AA\text{-AMP} + {}^{18}O_4\text{-}PP_i$$
$$+ \ \gamma\text{-}^{16}O_4\text{-ATP} \qquad + \ \text{excess } {}^{16}O_4\text{-}PP_i$$

The reactions were carried out in 6uL and with final concentrations of 5 mM $MgCl_2$, 5 mM PP_i, 1 mM g-$^{18}O_4$-ATP, 3mM amino acid, and 20mM Tris-HCl pH 7.5. The reactions were initiated by the addition of 2 mL of 600 nM enzyme in desalting buffer. After a half hour incubation at 25^{o}C, reaction was stopped by the addition of six uL 9-aminoacridine in acetone (10 mg/mL) for the MALDI samples. In the case of the ESI reactions the reactions were quenched with the introduction of pure 6uL of acetone

The mass peak of 514 corresponds to the four labeled ATP, 506 corresponds to the presence of unlabeled ATP. The increase in unlabeled ATP is indicative of an active exchange between the adenylation domain and the labeled ATP, this reaction is reversible. Masses 508, 510 and 512 are present as the result of 1,2, or 3 three respectively labeled O^{18} present on the ATP and if the result of natural abundance and buffer exchange.

Masses 528-536 are the sodium adducts, while the 588 through 596 peaks are the ammonia acetate adducts, a result of the buffer.

For analysis of both methods the ratio of unlabeled ATP was compared to all labeled ATP adducts and then calculated using Microsoft excel.

Using TycA adenylation domain as control we are able to show any activity between adenylation and the amino acids. AS shown in the graphs using the TycA as a control to show the activity of the assay. Compared with TycA domain incubated with no amino acid for background there seems to be some non-specific activity as well as just background. this suggests that none of the amino acids tested were the substrate the protein normally activates.

Discussion

Expression and characterization of these two domains show low activity. The first pps8 A3 show no affinity for any amino acid. The small activity present though non-descript does however verify that the protein is indeed active. This leads us to believe that none of the amino acids tested are the substrate for this domain.

The second protein when tested shows an affinity for aspartate and aspartic acid, which is ideal because of the near identical nature of these AA. However, the third activation is low. However once again because exchange can be

demonstrated it verifies that the protein is indeed active. Once again, the lane shows the presences of a second indistinguishable protein.

DATA

Pps8 A3

Table 6: Activity of Pps A 3: A

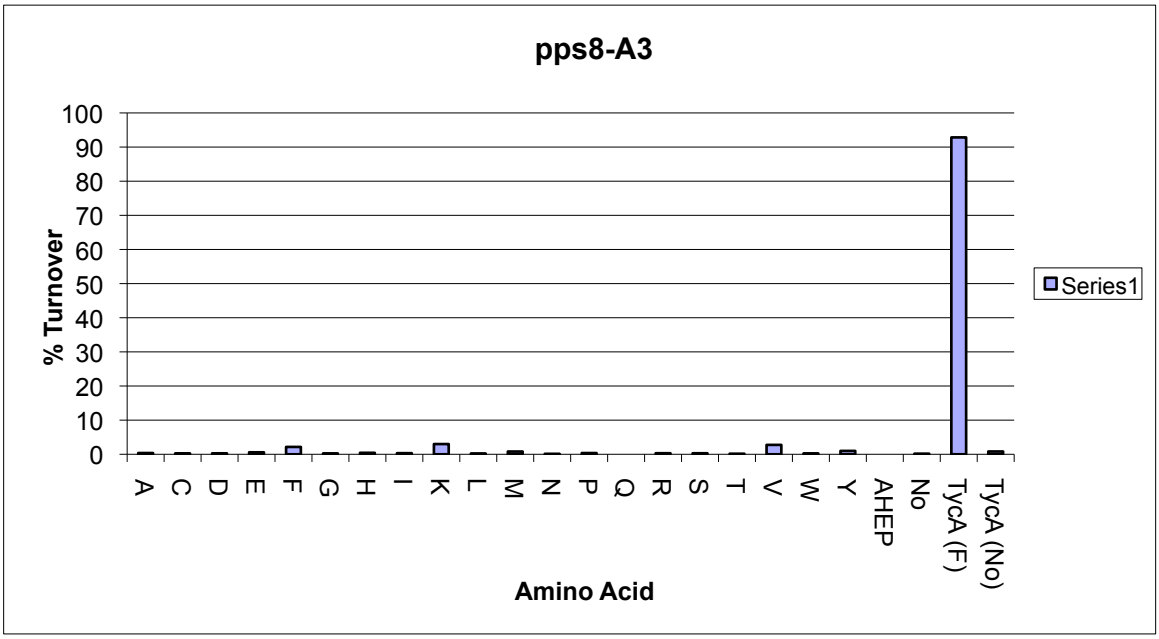


Table 7: Activity of Pps8 A3: B