TARR, KATHRYN R. M.S. Flavin Dependent Halogenases and SAM-Dependent Methyltransferases in Sporidesmin Biocatalysis. (2025)
Directed by Dr. Jason J. Reddick. 79 pp.

Pseudopithomyces chartarum (also referred to as *Pithomyces chartarum*) is a fungus commonly found in rye grasses after rain spells in the southern hemisphere, known for producing Sporidesmin, a mycotoxin responsible for acute liver toxicity in livestock. Sporidesmin belongs to the epipolythiodioxopiperazines (ETP) class of natural products and is notable for both its toxic effects and the presence of a chlorine atom in its structure. Its biosynthesis is governed by a cluster of 21 genes, referred to as "*spd*." This research project focuses on two key enzymes within the Sporidesmin biosynthetic pathway: Spd1, a homolog of an S-adenosyl-L-methionine (SAM)-dependent methyltransferase, and Spd4, predicted to be a flavin-dependent halogenase. The Reddick Research Group has an interest in the *spd* gene cluster for its potential for biocatalytic activity and for the basic understanding of the biosynthetic pathway of Sporidesmin. Spd1 is of interest for its potential as a methylation biocatalyst, while Spd4 may be responsible for the installation of the chlorine atom in Sporidesmin and holds promise as a general halogenation biocatalyst. The purpose of this research was to better understand the biosynthetic pathway and to potentially use the two enzymes as biocatalysts for other reactions. We are doing this through testing the activity of these enzymes on simple structural analogs of Sporidesmin. These reactions could be applied to other various types of chemistry for the uses in green chemistry and other industrial purposes. To investigate these enzymes, Spd4-MBP, Spd1, and Spd1-MBP were cloned and overexpressed in *Escherichia coli* (*E. coli)* using Gibson assembly and polymerase chain reaction (PCR). The maltose-binding protein, MBP, was fused to two of the proteins to aid in their folding and solubility during purification due to their fungal origins. The resulting proteins were purified, using immobilized metal affinity chromatography, IMAC, resulting in several milliliters, usually around 6 mg of protein per liter of culture grown, that was

then able to be used in subsequent reactions. High-Performance Liquid Chromatography (HPLC) and several enzymatic reactions were used to test Spd4-MBP for the chlorination of a simple indole. Similarly, several enzymatic reactions along with HPLC analysis were conducted on Spd1-MBP for potential methylation products from simple indoles and simple diketopiperazines. These studies have yet to yield any promising results. However, research is ongoing, and modifications are still being made to the investigative process to potentially result in the production of chlorinated or methylated products with good yields and stability.

FLAVIN DEPENDENT HALOGENASES AND SAM-DEPENDENT METHYLTRANSFERASES IN SPORIDESMIN BIOCATALYSIS

by

Kathryn R. Tarr

A Thesis
Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro

2025

Approved by

_____

Dr. Jason J. Reddick
Committee Chair

APPROVAL PAGE

This thesis written by Kathryn R. Tarr has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair

Dr. Jason J. Reddick

Committee Members

Dr. Nicholas H. Oberlies

Dr. Jonathan R. Chekan

November 21, 2025

Date of Acceptance by Committee

October 14, 2025

Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

I. A. Background

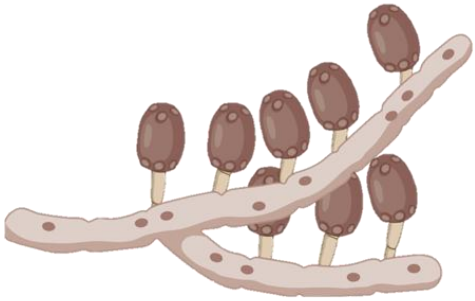


Figure 1: Pseudopithomyces chartarum: A Cartoon Image of the Spores and the Hyphae.

Pseudopithomyces chartarum (**Figure 1**) is a fungus that produces the mycotoxin Sporidesmin **(Figure 2).** The fungus was formerly known as *Pithomyces chartarum* but was reclassified in 1960 by mycologist Martin Ellis. The genus was first described in 1874 by Moses Ashley Curtis and Miles Jospeh Berkeley during their time in New Zealand where they were studying ill livestock.[1] The fungus proliferates in warm, humid climates after rainfall and is seen throughout the southern hemisphere. While it is mostly seen in areas like New Zealand and Australia, it has been documented to be as far north as the Netherlands. The ability of this fungal strain to spread and survive in diverse areas of the globe highlights the potential impact to the health of livestock on a global scale.[2]

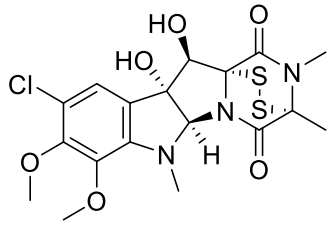


Figure 2: Sporidesmin molecule.

Sporidesmin is a diketopiperazine (DKP) that is made from L-alanine and L-tryptophan. It is fused together in a cyclic fashion, making it one of the smallest dipeptides with known biological activity. DKP's have high biological activity for their size, being found in small drug molecules and other types of potent biochemicals. Structurally, they are comprised of two amide linkages and two carbonyl groups giving them the distinctive core that makes them DKPs. Naturally DKPs are produced by bacteria, fungi, and marine life, but predominately most DKPS are produced by gram-negative bacteria, around 90% of them.[3]

Sporidesmin also belongs to a subclass of DKPs called epipolythiodioxopiperazines (ETPs). ETPs are classified by their distinctive disulfide bridge that spans the diketopiperazine core of the molecule. This disulfide bridge brings more reactivity to the molecule, but also more toxicity as well. This disulfide bridge can inactivate proteins via reactions with thiol groups and can generate reactive oxygen species by redox cycling. ETPs are only made from fungi which makes them interesting to study, since they are not found in any other forms of biological life.[4,5]

This fungal natural product also contains an indoline moiety, which is a heterocyclic compound structurally related to tryptophan. Indolines are like another class of structures called indoles, but they differ in the fact that they have a saturated bond between the carbons 2 and 3 in their heterocyclic ring when indoles have a double bond located in that area. Both compounds are important scaffolds for larger, more complex organic molecules and are incorporated frequently into natural products. They are also biologically significant, showing themselves in many systems as precursor molecules in the synthesis of proteins and other vital compounds.[6,7]

Since Sporidesmin originates from a fungal source and has toxic effects on animals that consume it, it falls into the category of compounds known as mycotoxins.

Mycotoxins are secondary metabolites produced by fungi that are dangerous to animals and humans alike and pose health risks. When *Pseudopithomyces chartarum* fungi are ingested, the bioaccumulation of Sporidesmin leads to a biological cascade of issues including hepatogenous photosensitization, facial necrosis, and liver lesions that can end in death. These issues can be treated using a combination of zinc containing medications for the animals, along with isolating the infected individual, but the illness is difficult to treat.[1,8]

Even though Sporidesmin itself is toxic, its biosynthesis is complex and worthy of investigation. The biosynthesis of Sporidesmin is governed by the gene cluster "*spd*" which is made up of 21 different genes that encode for different functioning proteins. While the full biosynthetic pathway is not fully understood, it is hypothesized that some *spd* genes contribute specialized functions to the final structure of the molecule. For example, *spd1* is thought to encode for a N-methyltransferase while *spd4* is thought to encode for a flavin-dependent halogenase. These educated assumptions have been made based on homology seen with other genes from similar biosynthetic pathways like that of gliotoxin and sirodesmin.[4]

Researchers have proposed a biosynthetic pathway for Sporidesmin that involves the enzymatic steps that make up the final molecule and that account for some of the 21 *spd* genes, that are part of the gene cluster. The initial step is thought to be the cyclization of L-alanine and L-tryptophan by a non-ribosomal peptide synthetase (NRPS), believed to be governed by *spd17*. The rest of the many modifications to the final molecule are done by the other spd genes. For example, *spd1* and *spd21* are predicted to be N-methyltransferases, while *spd7* and *spd11* are expected to be O-methyltransferases. Also, the molecule has two hydroxyl groups made possible by *spd8* and *spd10*, along with the chlorine group brought by the flavin-dependent halogenase *spd4*.

The rest of the *spd's* are thought to be responsible for the implantation of the disulfide bridge in the molecules center. These specialized enzymes convert the original amino acid scaffold into the final product, Sporidesmin. However, the exact sequence and regulation of these steps remains widely unknown.

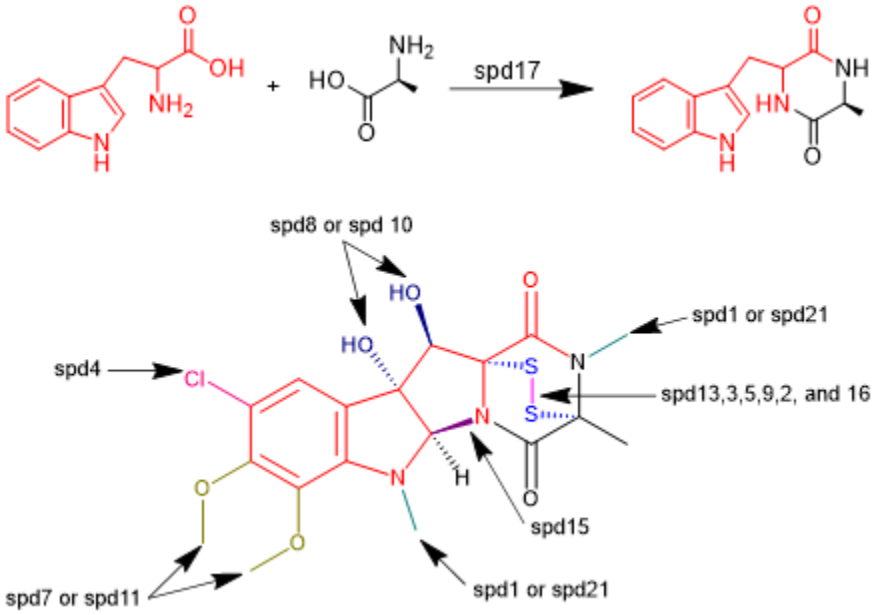


Figure 3: Proposed Roles for the Enzymes from the Biosynthetic Pathway of Sporidesmin.

The long-term goal of the Reddick Research Group is to understand the biosynthetic pathways of fungal species, such as Sporidesmin, with the proposed pathway shown in **Figure 3**. This thesis is on the expression and characterization of proteins involving the halogenation and methylation of substrates, with the aim of understanding how these modifications contribute to the biosynthesis of natural products and how they could potentially be applied to other biocatalytic pathways. The central hypothesis adds that engineered enzymes could potentially be capable of catalyzing halogenation and methylation reactions and how they could be used to selectively functionalize simple indoles and amides.

The rationale behind this work is that a deeper understanding of these enzymatic modifications will facilitate new investigations into natural product biosynthesis, which is essential for the understanding of natural product chemistry and important for the advancements in medicine and the understanding of biological systems.

I. A. i. Mycotoxins

Mycotoxins are compounds produced by fungi and molds that can pose serious health risks to humans and animals.[9] Respiratory problems, gastrointestinal problems, liver damage, and even cancer in some cases can be linked to mycotoxin exposer.[10,11] Most mycotoxins are chemically stable and can be difficult to eradicate once they have been released into an area and are even resistant to many food processing techniques. They thrive in damp, warm, humid conditions and can grow on foodstuffs and grasses. Exposure to these toxins can occur either through direct consumption of the fungus contaminated foodstuffs or by eating animals that have ingested such food or grasses and produced products like milk or cheese, which then passes along the potential risk of the mycotoxin effects to humans.[9]

Hundreds of mycotoxins have been discovered but the ones that are most observed and have been known to cause issues with humans are aflatoxins, ochratoxin A, patulin, fumonisins, zearalenone and nivalenol/deoxynivalenol.[12] Sporidesmin is one of these known mycotoxins, it is mostly known in veterinary spheres since it is more affected in livestock than humans. It does have some similarities with a much more studied mycotoxin, Gliotoxin, which is produced by another fungi. It is in the same family of compounds as Sporidesmin, epipolythiodioxopiperazines, and their biosynthetic gene clusters show homology and some common features.[4,13]

There are organizations that monitor and watch for mycotoxins on the world food web. JECFA, Joint Expert Committee of Food Additives, is a committee jointly created by the WHO (World Health Organization) and the Food and Agricultural Organization of the United Nations, oversees the evaluation of the health risks posed by natural toxins like mycotoxins. These organizations are put in place to help prevent mass outbreaks of mycotoxin illnesses and to help prevent the spread of the fungi that produces these mycotoxins worldwide.[9]

I. A. ii. Flavin Dependent Halogenase

Flavin-dependent halogenases (FDHs) catalyze halogenation reactions by introducing halogen atoms into organic compounds within biological systems. These enzymes can halogenate a variety of substrates including indoles, pyrroles, and phenolic compounds. The reaction typically requires reduced flavin ($FADH_2$), molecular oxygen (O_2), and a halide ion (such as Cl^-, Br^-, or I^-) as the halogen source. [14,15]

The mechanism of the FDH is well understood, yet in the case of Sporidesmin's FDH **(Figure 4)**, research continues. The mechanism begins with the reduced flavin cofactor ($FADH_2$) interacting with molecular oxygen (O_2) through a single-electron transfer, generating a flavin-hydroperoxide intermediate. Next, an ion (such as chloride) enters the reaction and interacts with the flavin-hydroperoxide forming hypochlorous acid and an oxygenated flavin species. The enzyme then leads the hypochlorous acid into the active site preventing it from diffusing or reacting nonspecifically. A lysine residue within the active site interacts with the hypochlorous acid and positions it for the reaction. When a substrate enters the active site, for this case an indole, the ion is transferred from the hypochlorous acid to the appropriate position on the substrate. This then produces the halogenated product while a flavin cofactor is regenerated through the reduction by a partner flavin reductase. [16]

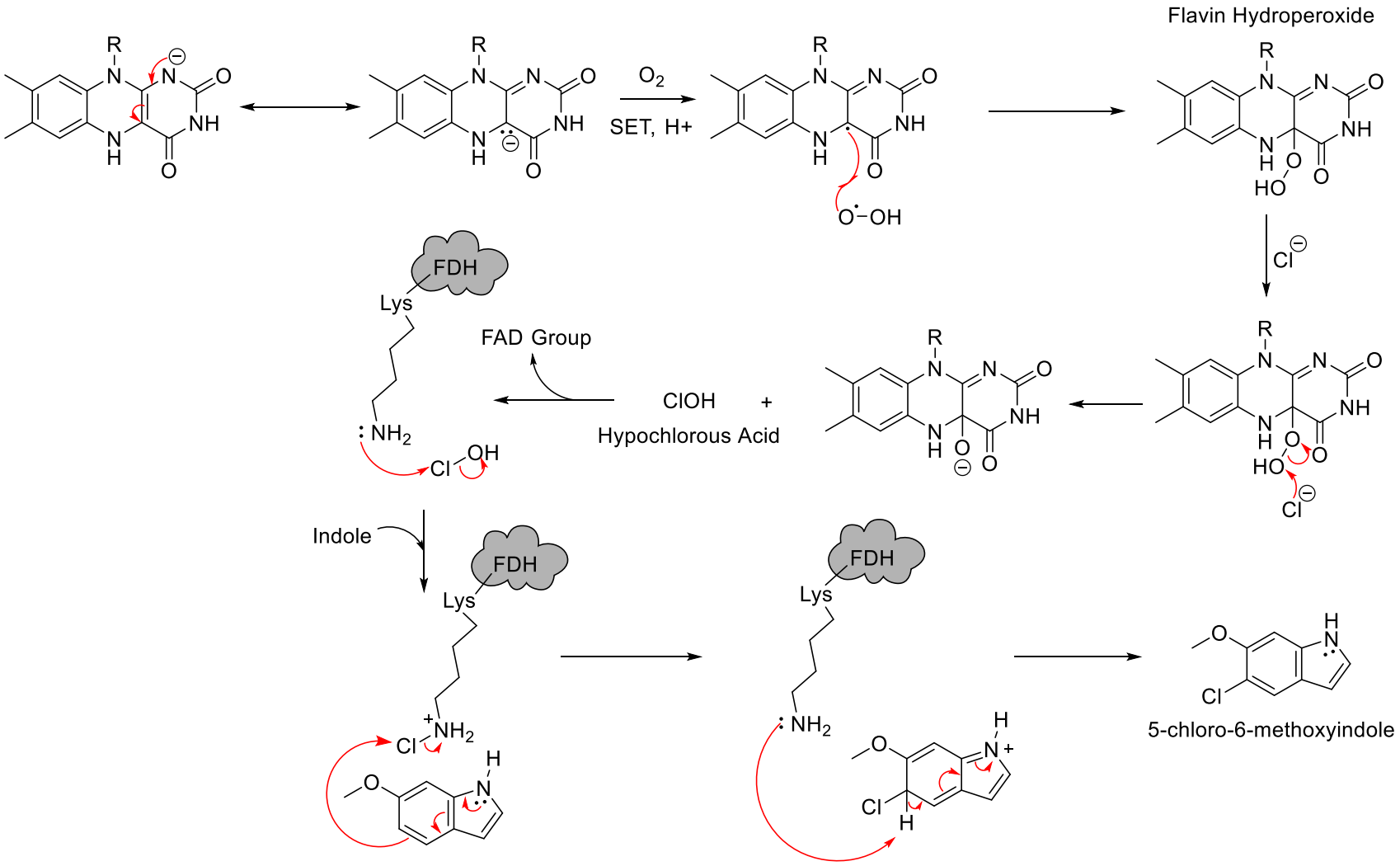


Figure 4: General Mechanism of FDHs in Correspondence with a Sporidesmin Homologue.

FDHs are regioselective and specific for the substrates they act upon. For example, the natural substrate L-tryptophan is the site of enzymatic halogenation by flavin-dependent tryptophan halogenases and is not governed by the electronic forces of the substrate, but the selectivity is determined by the steric involvement of the orientation of the substrate. This all happens in the active site of the enzyme. This allows for the enzyme to be more selective when acting upon certain positions of its substrates surface and accounts for the regioselectivity of this specified FDH.[17]

FDHs operate under biologically safe conditions unlike traditional reactions for halogenation such as electrophilic aromatic halogenation which sometimes involves ferric chloride ($FeCl_3$) as a catalyst and can produce acids like hydrochloric acid (HCl).[18] These types of reactions often require toxic environments and acid-based workups that can lead to toxic byproducts as seen in the previous example. FDHs enable similar reactions without those drawbacks. The products of FDH reactions often can be found in small drug molecules and agrochemicals which make them perfect candidates for biological chemical generation. They can also be made in cellular settings where their more toxic counterparts cannot.[19] Which enables then to potentially be made within a living organism than just in an inanimate object.

A few compounds that have been influenced by FDHs during their biosynthetic pathways are Rebeccamycin, Becatecarin, and Vancomycin. Rebeccamycin is a weak topoisomerase 1 inhibitor from the *Saccharothrix aerocolonigenes* bacteria. It shows significant anti-tumor properties when it was studied against mice melanoma cells and mice leukemia cells. During its biosynthetic pathway, an FDH installs two chlorine ions onto two benzene rings located on either side of the molecule which potentially attributes its usefulness as a potential anti-cancer drug.[20]

Along with this, Becatecarin is the synthetically made analog of Rebeccamycin and has shown the potential to also be used as an anti-cancer agent and both have been used in phase II clinical trials for the treatment of several different types of cancers including lung, liver, and breast.[20,21]

Vancomycin is a glycopeptide antibiotic medication used to treat bacterial infections. It can also be administered intravenously to treat other types of infections if necessary. Vancomycin comes from the soli bacterium *Amycolatopsis orientalis* which was first discovered in 1952 by Eli Lily and company. The vancomycin antibiotic was isolated from the bacteria in 1953. It was then approved for medical use in 1958. Its main use is for treatment of serious life threating infections by Gram-positive bacteria that are unresponsive to other types of medications. There are certain restrictions on using vancomycin today due to the growth of vancomycin resistant bacteria and the World Health Organization (WHO) has set guidelines for its use in modern medicine.[22,23] Vancomycin itself during its biosynthetic pathway has also been affected by a FDH in which chlorination occurred on two benzene rings in its structure that may lead to its increased activity.[24] These compounds, in **Figure 5** below, can serve as helpful agents to human health due to the importance of halogenated compounds in medicine.

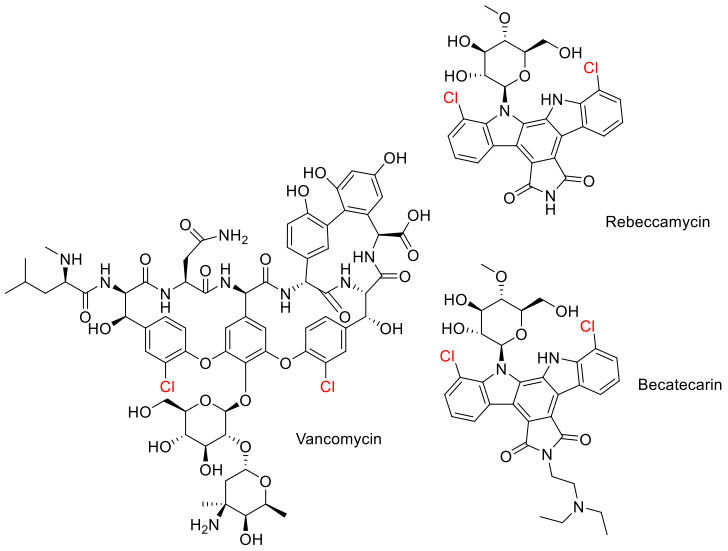


Figure 5: FDHs Halogenated Each Compound Enhancing Their Functionality.

FDHs need cofactors to function properly. These cofactors are known as flavin reductases (FRs). These FRs reduce flavin adenine dinucleotide (FAD) to its active form ($FADH_2$), which is then used by the FDH to halogenate the substrate. Flavin Reductase RebF, generates ($FADH_2$) by oxidizing NADH and then serves as the essential electron donor for the reaction. $FADH_2$ is then able to bind in a separate grove from the active site of the FDH and then becomes oxidized to form the flavin hydroperoxide as described above in the mechanism section.[25] Harnessing FDHs to produce compounds using FRs, in biological environments rather than synthetically, could make the future of biosynthesis applicable in more situations than currently realized.

I. A. iii. Methyltransferases

Methyltransferases are enzymes that methylate substrates for the purpose of prevention of degradation, enhancement of molecular structure, and the addition of methyl groups to better suit a substrate for another enzyme's active site.

These enzymes fall into 3 distinctive classes; class I is described as containing the Rossman fold, which is a tertiary protein structure that binds small molecules like FAD, NAD^+, and $NADP^+$. This fold is composed of alternating beta strands and alpha helical segments where the beta strands are hydrogen bonded to one another forming an extended beta sheet. Class II methyltransferases contain SET domains which are distinctive structures that have exposed amino acids that form the catalytic site of the enzyme. It is theorized that this site acts differently than other methyltransferases by interacting with histones, which is often why they are called "histone methyltransferases". The final group, Class III, are membrane bound methyltransferases and are located on the outside of cells. They regulate cell signaling and gene expression.[26,27]

These enzymes play important roles throughout nature, some of those include the following. DNA methylation is an important part of the DNA lifespan since methylation happens after replication. This is due to the work of DNA methyltransferases adding methyl groups onto freshly replicated DNA for the modifications of gene expression and the maintenance of patterns in this important cycle. Methyl addition to DNA also aids in the preservation of DNA by protecting it from certain nucleases and slowed degradation times. Another example of methyltransferases are histone-methyltransferases. Histone-methyltransferases regulate epigenetics and are key factors in the building blocks of chromatin. They play a critical role in activating or deactivating transcription, and other biological processes like cell regulation, DNA repair, and development. [28,29]

There is yet another group of methyltransferases that are of importance called natural product methyltransferases (NPMTs). They are a diverse enzyme subclass that add methyl groups onto natural product substrates, similarly to Sporidesmin.

These enzymes are specialized for pathways in small groups of species providing secondary metabolites for the organisms in these groups.[30] Luckily, through scientific analysis, scientists can harness these small compounds made by NPMTs, study their biosynthetic pathways, and utilize them in other means.

I. A. iv. S-Adenosyl-l-Methionine (SAM)

S-adenosyl-l-methionine (SAM) is a natural compound used by methyltransferases to facilitate the methyl transfer reaction acting as the prime donor of the methyl group. It is produced by animals, fungi, and bacteria and is also used in transsulfuration and aminopropylation reactions.[31] SAM-dependent methyltransferases are found in DNA regulation, cellular regulation, and the generation of pigments and particles in cells.[32] Over 40 different SAM-dependent methyltransferases have been identified throughout the biological kingdoms.

These specified methyltransferases have been seen to work on a variety of substrates including nucleic acids, lipids, and secondary metabolites (natural products).[31] SAM is generated from a methionine adenosyltransferase when it synthesizes adenosine triphosphate (ATP) and a methionine amino acid. There is also a full cycle that synthesizes and regenerates SAM, called the SAM cycle. In the initial phase of the SAM cycle, SAM dependent methyltransferases utilize SAM as a substrate to generate S-adenosyl homocysteine as a product, the cycle is shown in **Figure 6**. S-adenosyl homocysteine serves as a potent negative regulator of almost all SAM dependent methyltransferases, regardless of their biological variety.[33] It is then hydrolyzed to homocysteine and adenosine by S-adenosine homocysteine hydrolase where the homocysteine subsequently is recycled to methionine through the transfer of a methyl group. This methionine can then be reconverted to SAM, therefore continuing the cycle. It is also

massively important to have two vitamin derivatives, vitamin B_{12} and Folate, in the process as shown below in **Figure 6.**[34]

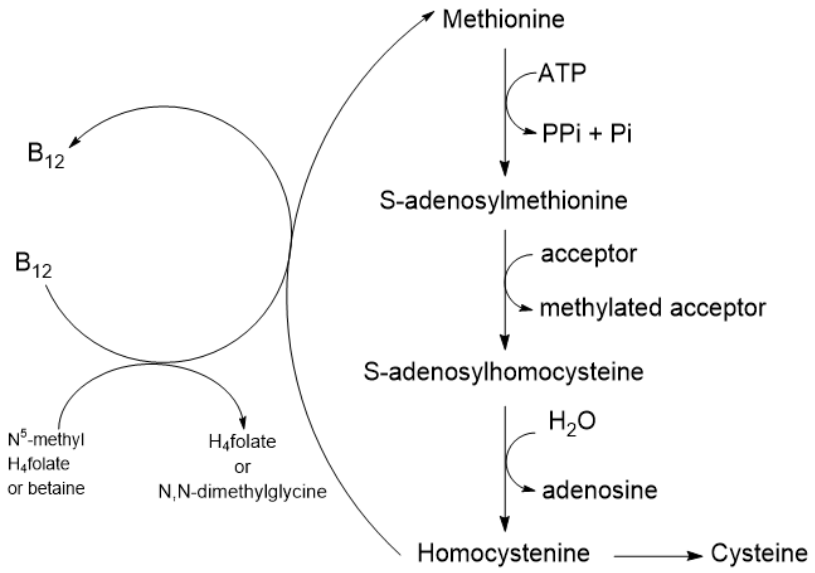


Figure 6: Main Metabolic Pathways of S-adenosylmethionine.

Similarly to the FDH biological mechanism, the SAM mechanism is also more biologically safe than its current industrial counterpart. Traditional methylation reactions, like the Grignard reaction, use harmful reagents like methyl iodide (CH_3I) and dimethyl sulfate ($(CH_3)_2SO_4$) which can produce harmful byproducts.[35,36] By using the biological SAM reactions, hazardous waste can be prevented. One drawback to using SAM is that it is expensive to obtain and purify, which poses a problem for industrial use. If more efficient manners of SAM production can be achieved, then it can be utilized more efficiently on larger scale biological reactions.[37]

SAM's mechanism is rather straightforward; first SAM brings an active methyl group which carries a positively charged sulfonium center which makes the methyl group highly reactive.

The methyltransferase enzyme has two binding pockets, one for SAM and one for the substrate the atom on the substrate for example (oxygen or nitrogen) acts as a nucleophile. The atom of choice then attacks the methyl group attached to the sulfur and SAM, in this single step S_N2 like reaction the methyl group breaks away from SAM's sulfur and attaches to the nucleophilic atom on the substrate. As a result, SAM is converted into S-adenosylhomocysteine which lacks methyl group and leaves the methylated substrate.[38] This is shown in **Figure 7**.

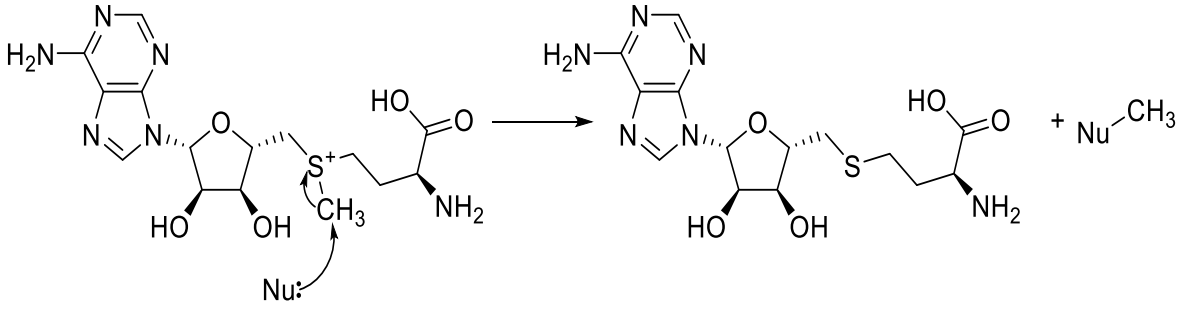


Figure 7: SAM Reaction Mechanism.

I. A. v. Maltose-Binding Protein

Sometimes, when expressing fungal genes in *E. coli* an aid is needed. The maltose binding protein (MBP) is that aid, it enables the proper folding, solubility, and crystallization of fungal and other types of proteins in *E. coli*. MBP is typically used with the gene of interest as a fusion to better express the protein in which the gene encodes for. MBP is a popular choice for many in biotechnology and recombinant protein expression due to its relative ease of use and its ability to be used with metal affinity chromatography to isolate and purify proteins of interest. Along with its usefulness in affinity chromatography, MBP is also useful in preventing aggregation of proteins and inclusion bodies occurring during the protein production process.

Inclusion bodies occur when the protein of interest stays within the cellular debris when the cell is lysed (either by chemical lysing or sonication) rather than being eluted through the process of purification. Aggregation is a process in which the protein forms clusters that stick to one another making it difficult for the protein to be eluted, this is usually observed in SDS-PAGE gel and can be corrected using MBP.[39,40]

The mechanism in which MBP functions is not well understood by science, however it is known that MBP is a monomeric protein that contains two globular domains connected by linker with the maltose binding groove between them. This groove allows for the binding of maltose and for the protein to function as intended. MBP is also encoded by the *malE* gene in *E. coli*, which produces a precursor amino acid sequence that is then cleaved to produce the mature *malE* gene. What is most interesting about the *malE* gene is that in either form, no cysteine residues are present. MBP is a relatively medium sized protein, about 42.5 kDa in weight and relatively easy to spot on SDS-PAGE gels when identification is needed. Its function and abilities to help make difficult proteins express with ease makes it a hallmark of biochemistry.[41,42]

I. A. vi. High Fidelity Gibson Assembly

Gibson Assembly was a DNA assembly technique created by Daniel G. Gibson in 2009 and his colleagues at the J. Craig Venter Institute in Maryland, United States. [43] It is a technique in which several DNA fragments can be linked together in a single isothermal reaction to form one continuous form of DNA. This is widely used to make plasmid DNA since Gibson Assembly can make continuous circular DNA easier than other methods known.[44]

Gibson Assembly relies on the activity of 3 key components to function, an exonuclease, a DNA polymerase, and a DNA ligase.

The endonuclease first attaches itself to the 5' ends of DNA fragments, opening them and exposing the complementary single stranded overhangs. These overlaps guide the fragments to anneal in the correct orientation so DNA can be transcribed from the exposed portions of the DNA strand. Next, DNA polymerase then fills in the gaps and a
DNA ligase seals any nicks that had been left during this process in the sugar phosphate backbone. This process results in fully assembled continuous DNA constructs.[45]

Gibson Assembly has significant advantages over other restriction enzyme-based cloning, which requires more time and multiple steps compared to the streamlined process of Gibson Assembly, shown in **Figure 8**. Instead, fragments can be designed with overlapping edges, making them more flexible and precise and they can be directly adhered to existing DNA by matching them with the correct base code sequence. This allows for the assembly of multiple fragments, sometimes more than 10 in a single reaction, making it specifically useful for constructing large plasmids, synthetic genes, or even entire genomes.[46]

The importance of Gibson assembly lies in both its efficiency and versatility. It has become a model in synthetic biochemistry, metabolic engineering, and genetic research because it streamlines the process of collecting and multiplying DNA for analysis and other testing purposes. Through building customized DNA constructs, researchers can quickly assemble plasmids containing multiple genes and regulatory elements to create fusion proteins without the

constraints of variable restriction sites. This has accelerated workflows in areas from protein engineering to vaccine development.[45]

Gibson assembly is used in both academic and industrial labs due to its reliability and scalability. It can be used for small scale benchtop reactions and even for large scale industrial

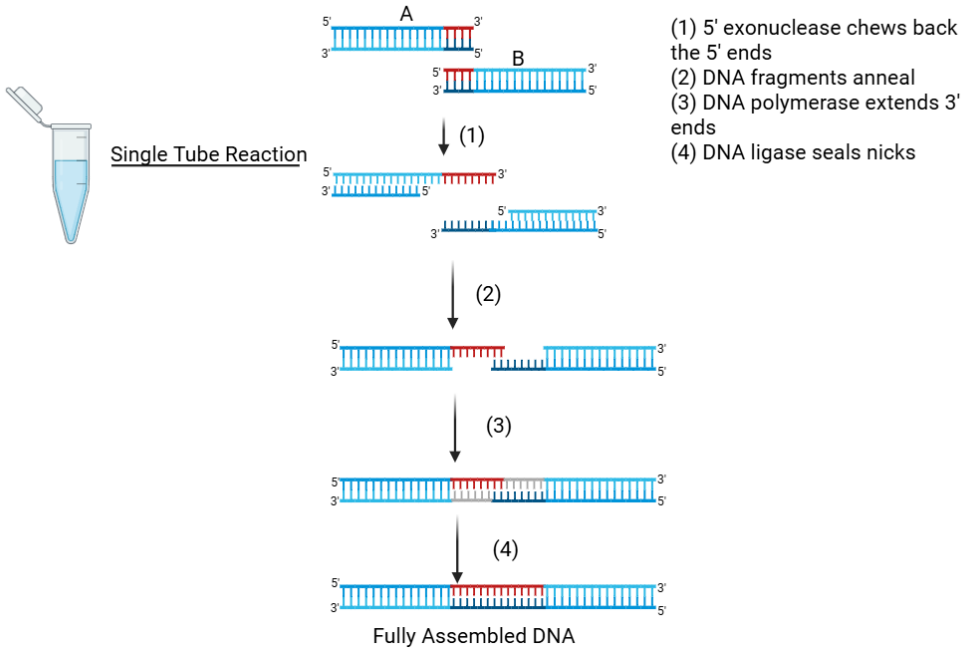


Figure 8: Full Gibson Assembly process showcasing the functionality of the process.

reactions. It's particularly valuable for high throughput projects where many constructs can be generated in parallel, and for projects requiring seamless DNA junctions without extra sequence scars. That's why this method was chosen for the creation of plasmids in this work. Without Gibson assembly, it would have been more difficult to create the MBP fused plasmids that would be needed for protein expression and the generation of purified proteins.[44–46]

I. A. vii. Immobilized Metal Affinity Chromatography

Immobilized Metal Affinity Chromatography (IMAC) is a technique used to purify proteins that can be altered to be selected for during their gene engineering phase. An engineered protein can be separated from a complex mixture using a histidine tag located on either the C-terminus or N-terminus of the protein of interest. This "His-Tag" is specifically engineered onto the protein during the encoding phase and is used in the purification steps to isolate the desired protein from other proteins that are produced during the cell's lifecycle. There are usually 6-10 histidine residues in subsequent order that make up the histidine tag. Histidine, when deprotonated and neutral, can form a coordination complex with a metal ion and allow for the binding of protein to occur between the metal and itself. Which allows the protein of interest to be separated from other debris in the mixture.[47,48]

The method works by exploiting highly selective interactions between the target molecule and a binding molecule, a ligand, that is immobilized on a stationary phase like agarose or polyacrylamide within a plastic column. The specificity of this technique comes from the biological nature of the binding between ligand and protein. Examples of such interactions include antibiotic and antigen binding, enzyme to substrate binding, and receptor to ligand binding. Because of the strong selectivity, affinity chromatography provides much higher selectivity compared to many other chromatography methods making it essentially valuable for protein purification.[49]

The process begins with loading the metal ion, typically nickel or cobalt, onto the stationary phase of the column. Then a series of cleansing washes, typically water and a selected binding buffer that is paired with the desired protein, are performed on the column to prepare it for the introduction of the crude protein extract.

Then the crude mixture which contains biomolecules and other proteins is loaded onto the column and allowed to flow through and is collected. During this, the histidine tag containing protein is adhered to the column and washed with a series of wash buffers to cleanse both the column and the desired protein of other undesired proteins and cellular debris. A final elution step is then done to elude the desired protein from the column which cleaves the bond between the histidine and the column. This elution buffer contains imidazole, a chemical that is specifically used for breaking the coordination complex made between the metal and the protein so that the protein can be collected during the elution phase rather than eluting elsewhere in the IMAC process. The elutes are then taken as fractions to then be used in subsequent reactions and analysis.[47,49] Affinity chromatography enables the efficient purification of proteins and other biomolecules with high efficiency and yield. A pictorial reference to the steps in which are taken to complete a nickel column are shown in **Figure 9.** [50]

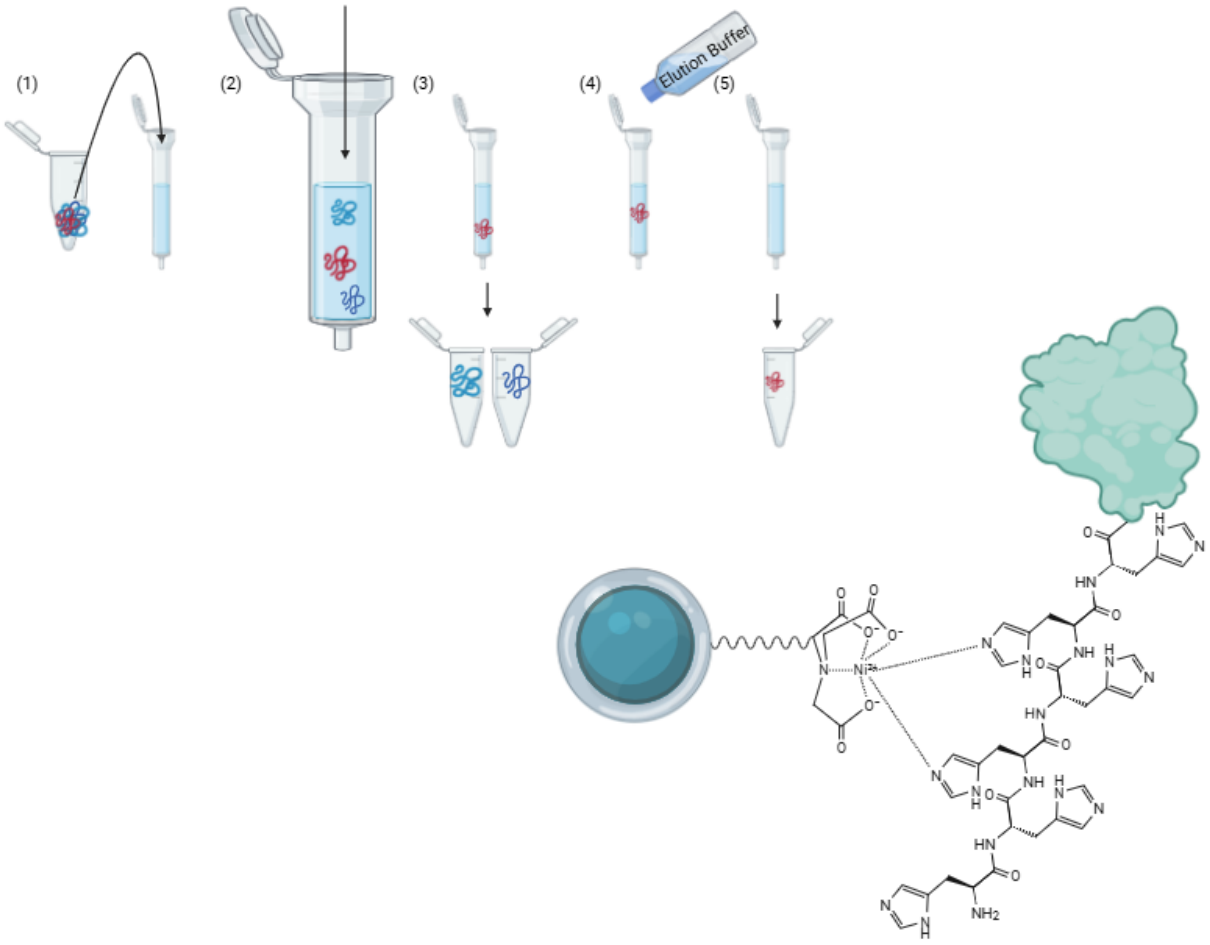


Figure 9: Nickel Chromatography Diagram.

I. A. viii. Escherichia coli

As said in the MBP section, fungal genes sometimes have a difficult time expressing in *E. coli* cells and need optimization to function and express correctly. If placed into bacterial strains without modification, the fungal genes would have difficulties with expression or would not express at all due to the fact they come from differing kingdoms of life and other factors. [51] Producing prokaryotic genes from eukaryotic sources requires genetic optimization in the form of base pair changes and genetic engineering.

Firstly, codons need to be genetically engineered to be suitable for expression in *E. coli*, this is done by replacing base pairs without rewriting the animo acid sequences for which they encode for. The main reason for this is to prevent the codons themselves from being misread, proteins being prematurely or produced later than expected or not even at all. An example of this is promotor sequences in bacteria cannot recognize the promotors of eukaryotes so transcription could not take place.[52]

Another issue is that fungal genes can contain introns in which bacteria, due to its primitive nature, cannot splice them out, so a major issue is that important genetic materials could be lost due to a miss cut along a genetic sequence. To avoid this, the introns are removed before the genes are introduced to the *E. coli*, so the removal of important genes does not occur.[53] Lastly, as stated before, inclusion bodies and aggregation are issues that tend to occur when expressing fungal genes in *E. coli* hosts. This can be negated using MBP which as mentioned, helps with the solubility of the proteins produced, allowing less aggregation to form and inclusion bodied to be less likely.[54]

I. A. ix. Indoles, Diketopiperazines, and Epipolythiodioxopiperazines

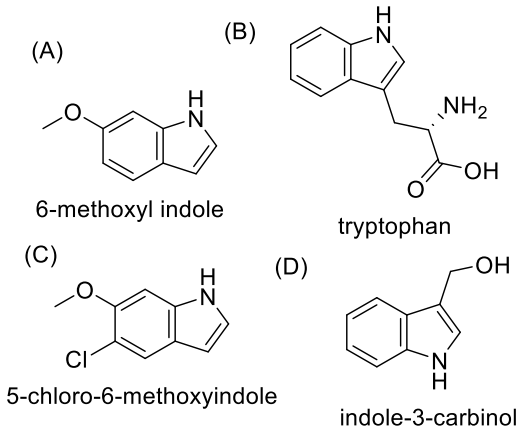


Figure 10:Common indoles B&D, and the testing indoles, A&C used in this project.

Indoles and indolines are bicyclic compounds that contain a benzene ring fused to a pyrrole ring. The key difference between the two structures is that indoles have an unsaturated double bond between the carbons on the pyrrole ring whereas an indoline has a saturated bond at that location. They are widespread in chemistry, being found in natural products, synthetic compounds, and other building blocks of life.[55] They serve as scaffolds for primary and secondary metabolites and are seen in many natural products as the center of bioactivity. In biochemistry, indoles and indole derivatives play critical roles such as signaling molecules, intermediates, and bioactive metabolites. For an example, indoles are involved with bacterial communication and regulation of biofilm formation while at the same time many plants and microorganisms produce indole alkaloids with pharmacological activity.[7,56] In plants, indole-3-carbinol is produced by the breakdown of glucosinolates (sulfur-containing plant compounds found in cruciferous vegetables like broccoli and cabbage) and is one of the most common plant hormones found in the world. It is used by plants as a protective agent, serving to signal root growth and to deter insects.

Indoles and indolines have much versatility, they have been used in antibacterial, anticancer, and psychoactive medications and many other types of drug molecules that aid in modern medicine. Many can be sourced from natural fungi and bacteria and be synthetically created by organic reactions in laboratories. Because of the versatility with both indoles and indolines, several of them have been chosen for study with the Sporidesmin biocatalytic pathway in mind. Due to the Sporidesmin molecule containing an indoline itself, structural indoles like 6-methoxyindole, 5-chloro-6-methoxylindole, and tryptophan, have been decided to be used in testing for potential enzymatic activity, shown in Figure 10.[57]

Diketopiperazines (DKPs) are the smallest cyclic peptides known to science. They are formed from the condensation of two amino acids to form a 6-membered ring and are characterized by their two amide linkages and two carbonyl groups. Even though they are small, they boast high bioactivity and showcase a variety of different functions in the biosphere. They are commonly produced by bacteria and fungi and often have antibiotic, antifungal, and anticancer properties. [3] DKPs are also chemically stable having a ridged structure which makes them great candidates for drug molecules because they are resilient to degradation and can withstand structural adaptations. They are commonly made in biosynthetic pathways by NRPs, or ribosomal pathways, and are useful compact bioactive molecules. [58] A few of these DKPs are shown below, one is a synthetic DKP, Retosiban, which shows that they can be taken from biological sources and brought to the synthetic sphere. Tryprostatin B originates from the fungus *Aspergillus fumigatus* and was isolated in 1995 for potential use as an anticancer drug. Cyclo(L-Leu-L-Hyp) was extracted from *Alternaria alternata* and is said to potentially have antidepressant effects in some studies. [59,60]

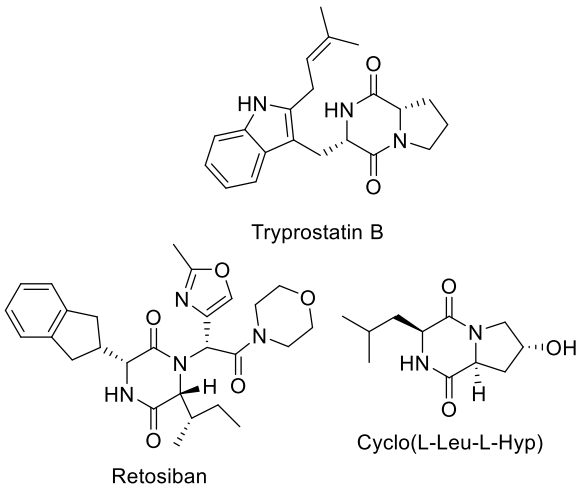


Figure 11: Several Diketopiperazines from Several Different Sources.

Epipolythiodioxopiperazines (ETPs) are a subclass of DKPs that are characterized by a disulfide bridge that spans the piperazine ring of the molecule. This structural component gives many ETPs cytotoxicity but also gives them much in the way of biological activity. [61] They are usually produced by fungi and known for their roles in pathogenic species. Biochemically, ETPs can disrupt cellular systems by generating reactive oxygen species or editing proteins through thiol-disulfide exchanges. [5] One of the most well-known ETPs is gliotoxin, pictured in **Figure 12**, first discovered in 1936 by Weindling and Emerson as a product from *Trichoderma viride* but it is debated if that was the first fungus to produce the compound or not. Gliotoxin has immunosuppressive properties that may suppress and cause apoptosis in certain cells of the immune system, including neutrophils, eosinophils, granulocytes, macrophages, and thymocytes. The immunosuppressant activity is due to the disulfide bridge, and when interactions between the sulfur molecules and amino acids occur, they can deactivate proteins that can be vital for cellular function as mentioned previously. What is most interesting, is that *Trichoderma virenas*, another fungal strain that biosynthesizes gliotoxin, is sold under the name SoilGard as a bio-pesticide.[4] Since ETPs have both bioactivity and cytotoxicity, they both pose an interest to scientific investigation and a concern for human health.

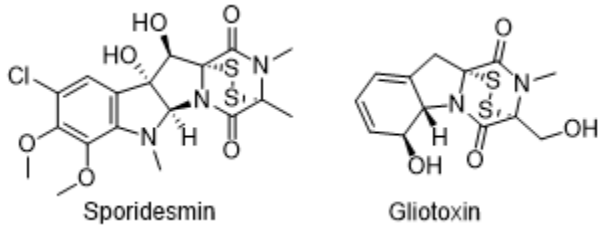


Figure 12: Sporidesmin and Gliotoxin are two molecules from the ETP class of natural products.

Sporidesmin belongs to each of these classes of molecules being comprised of a diketopiperazine, an indoline, and having the distinctive disulfide bridge making it an ETP. Each specific part of the compound gives it bioactivity and plays an important role in how the compound is biosynthesized. Research from the Department of Agriculture in Hamilton New Zealand back in the 1960s and 1970s into the biosynthetic pathway of Sporidesmin has shed light on the final structure of the molecule and has alluded to how it functions biologically but there is still much to be discovered about how it is biosynthesized into the final structure.[62]

I. B. Goals and Hypothesis

There were two goals to be achieved for these projects. One was to chlorinate simple indoles, and the other was to methylate simple indoles and amides to find significant biocatalytic activity from the Spd enzymes. These enzymes that could be genetically engineered for the expression of modified His-tagged proteins could potentially catalyze the reactions necessary for the chlorination and methylation of the simple substrates chosen for study. The main hypothesis of this study was that if these enzymes can be genetically engineered and the proteins that they produce can chlorinate and methylate the substrates of choice, then these enzymes could be applied to other substrates as biocatalysis for other needs. Significantly, these enzymes could assist the growing need for Green Chemistry routes in replacement of industrial organic and inorganic chemistry and highlight the need for more biologically safe chemical practices allocating for their need to be expanded and explored.

CHAPTER II: SPD4-MBP

II. A. Overview and Objectives

This section details the DNA engineering, protein production, and enzymatic reaction makeup for the spd4-MBP project. This section details the procedures and materials used in the DNA assembly of the construct plasmids, of the protein expressions, and of the enzymatic reactions. Information in this section is dedicated to describing the changes that occurred during testing.

II. B. Experimental

II. B. i. Obtaining the pET-28a Plasmid

The first task of the Spd4-MBP project was to create primers for the process of DNA replication by PCR for the following Gibson Assembly process in which the fragments of spd4 and MBP will be conjoined to form the final plasmid. The project was started by obtaining the pET-28a-BURP plasmid, which contains the MBP gene, from the Chekan Research Group from the University of North Carolina at Greensboro. The sequences were then added to Benchling® a DNA sequencing tool, to better understand its sequence and to create the needed primers for the Gibson Assembly, and to find the best location to insert the *spd4* gene. This viewing would also give the opportunity to discover where other important features of the plasmid are located like the LacI, a repressor that regulates the expression of genes, and the KanR gene, the kanamycin resistance factor, are in the plasmids sequence.

II. B. ii. Sequencing the *spd4* Gene

Bioinformatic scientists decoded the gene for *spd4* and uploaded it to the National Library of Medicine (NLM) and the sequence was then ordered in a pET-28a plasmid for use by a PhD student, Oliva Shirley and then further used in this project.

II. B. iii. Creating Primers for the Gibson Assembly of *spd4* and MBP

Both sequences were analyzed on Benchling© to create the most suitable primers for the conjoining of the two fragments into a single plasmid. In Benchling©, primers can be generated, constructs of plasmids can be generated, and other DNA analysis tools can be employed to detail DNA sequences. The construct of *spd4*-MBP is shown in **Figure 13**. The *spd4* gene was to be collected by using the PCR primers made from the information gathered to create the final plasmid.

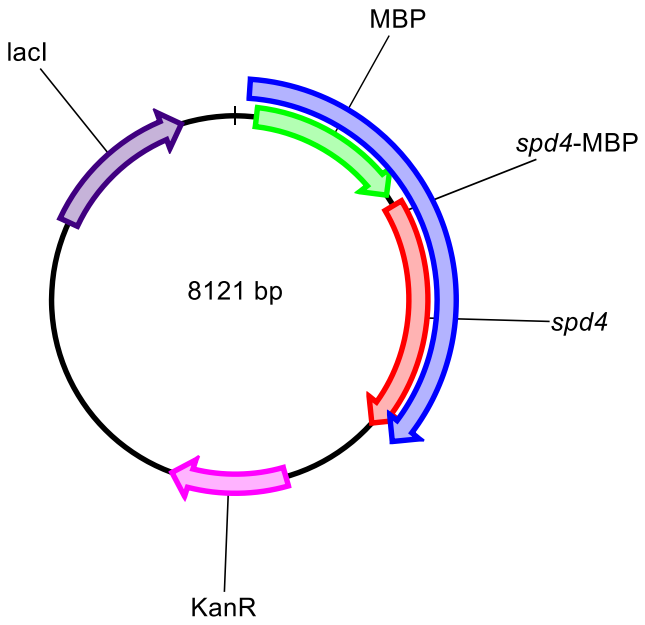


Figure 13: Construct created of the *spd4*-MBP Fusion Plasmid.

The primers generated could not exceed 42 bases in length due to ordering requirements and functionality stipulations. The bases also needed to be longer than 21 base pairs in length or they would not anneal properly or create a long enough overhang. Temperature was also a factor in designing the primers in which the annealing temperature could not exceed 60.0°C since this would be the deciding factor to which the primer would anneal to the template sequence. Once the parameters were met, primers were ordered from Integrated DNA technologies (IDT). The sequences are represented in **Table 1**. The primers were then dissolved in $DDIH_2O$ and were set to a starting concentration of 20μM.

Table 1: Sequences of Primes received from IDT.

Name of Sample	**Sequence of Primer**
MBP_Forward	TAACTCGAGCACCACCACCACCACCACTGAG
MBP_Reverse	GGATCCGGATTGGAAGTACAGGTTCTCAGATCC
*spd4*_Chis_Forward	CTGTACTTCCAATCCGGATCCATGGTGGAGACACCGACATC
*spd4*_Chis_Reverse	GGTGGTGGTGGTGCTCGAGTTAGGTCTTCACACCATGCTTCG

Table showing the sequences of each primer obtained from IDT. Details the sequences in their formats as they were ordered.

II. B. iv. Polymerase Chain Reaction of *spd4* and MBP

Next, was to use the primers in polymerase chain reaction (PCR) to generate copies of the segments of DNA for further use in Gibson Assembly. PCR is a process in which DNA is amplified with the use of thermostable enzymes to create millions of copies of DNA in relatively short time periods. PCR protocols were adapted from New England Biolabs protocols to better suit *spd4* DNA specific needs, i.e. annealing temperature and total runtime. The thermocycler

used for this experiment was a BioRad MyIQ 2 Optics Module. This thermocycler will be used throughout the project.

Preparation of the samples was undertaken by aliquoting 2.5µL of Forward primer at a starting concentration of 20µM, (final concentration 1 µM) 2.5µL of Reverse primer at a starting concentration of 20µM (final concentration 1 µM), 2.0 µL of Template DNA, 18µL Nuclease free H2O (DDI H2O) and 25 µL 2x Phusion Enzyme mix in a mini-microcentrifuge tube. The enzyme mix was added last to maximize time in the thermocycler and allow the reaction to progress effectively. This procedure was carried out for both *spd4* and MBP. For the first experiment, the settings for the thermocycler was that Step 1 had an initial temperature of 98.0°C for 30 seconds with one cycle, Step 2 had an initial temperature of 98.0°C for 10 seconds, an annealing temperature of 60.0°C for 30 seconds and an elongation temperature at 72.0°C for 30 seconds, step 2 was repeated for a total of 30 cycles for 3 hours and 15 minutes. Then, a holding temperature of 72.0°C was set for 10 minutes. If samples were left for an extended period, they would be chilled at a temperature of 4.0°C inside the thermocycler until use. A pictorial reference to the PCR process is shown below in.

Gel electrophoresis was conducted once the PCR reaction had been completed. Gel preparation was conducted with a low-melting agarose powder branded by Thermo Fisher Scientific. First, 0.3 g of the agarose powder was added to 30 mL of 1x TAE buffer, which is a mixture of Tris base, acetic acid, and ethylenediaminetetraacetic acid (EDTA). This created a 1% agarose gel in which 2µL of ethidium bromide was added. The gel solution was then poured and left to solidify for 1 hour. Afterward, the gel was transferred to the BioRad gel electrophoresis device, and 1x TAE buffer was poured over the gel to cover the gel's wells once the comb of the gel had been removed.

To add the samples to the gel, gel loading dye (purple) was first added to the PCR products along with $DDIH_2O$ to allow them to be seen on the gel, the added ethidium bromide allowed for visual confirmation of DNA products under UV. For the preparation for the ladder sample, 2.0 µL of gel loading dye, 4.0 µL of $DDIH_2O$, and 1.0µL of 1 kb DNA Ladder were added together for a total volume of 7.0µL. For testing samples, 2.0µL dye and 7.0µL of the PCR product were used for a total volume of 9.0µL. The volume of sample used in gels was 6.0µL for both ladder and testing samples. Gels were run at 100 volts, 90 mA, for approximately 1 hour. Once the samples were finished running, they were taken to a BioRad UV light device to be examined. Once examined, it was determined that parameters needed to be adjusted for subsequent runs, and thus the PCR procedure was updated:

II. B. v. Gibson Assembly Procedure for *spd4* and MBP

For the Gibson Assembly, concentrations of insert and vector were needed. These would be determined using Nanodrop. The concentrations varied due to PCR production. The calibration of the device was done to ensure correct results and was done using the final buffer solution in the PCR reaction, the elution buffer. A Qubit 4 Fluorometer was also used in finding concentrations of DNA products as well. The final volume of the Gibson Assembly was to be 20µL, so water and enzyme mix were added. A PCR clean up stage was added to further enhance the Gibson Assembly results. These PCR clean up steps were done to the specifications of the manufacturer. The Gibson Assembly was dependent on the concentration of the PCR products of each round. Experiments varied concentrations and techniques in which these concentrations would be expressed.

For example, the first trial was conducted as a concentration ratio. This ratio was the concentration of vector to insert in solution. The ratios were determined using a calculation of base pairs to nanograms. This was that calculation; (pmols) (base pairs) (650)/ (1000) = weight in nanograms. This was then expressed as a concentration in a set amount of $DDIH_2O$. Each added insertion volume increased while the vector value remained constant. The components of these mixtures were the spd4 insert, the MBP vector, DDIH2O, and the NEB DNA Assembly Master Mix for a final total volume of 20 µL. The reaction was then incubated at 50.0°C for 15 minutes, and then 2µL of product was used in transformed *E. coli* cells for plating. The first set of Gibson cloning products had been made from taking samples of inserts that ranged from 2 µL to 10 µL and plating them after the process of transforming them with a set amount of vector. The Gibson Assembly process is shown below in **Figure 14**.

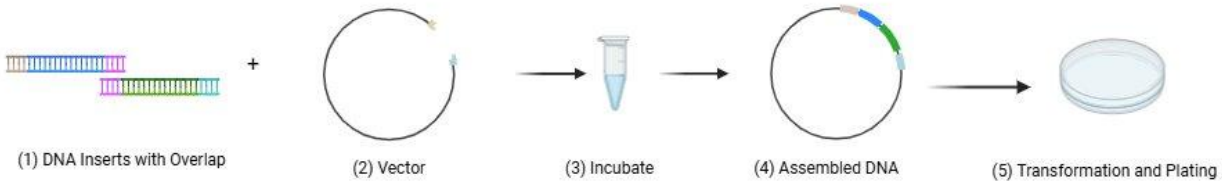


Figure 14: Gibson Process.

II. B. vi. *E. coli* transformation of DH5α and BL21-STAR

Sequencing of the Gibson Assembly products was needed to identify the products created, so a process known as transformation was needed. Transformation takes competent *E. coli* and administers foreign plasmids for uptake and growth in the *E. coli* strain. The *E. coli* culture then can be used to extract the plasmid in larger quantities for sequencing or if a specific strain is used, the *E. coli* can be used for protein production.

The strains used for the sequencing data are DH5α and the strain used for protein production is BL21-STAR. The Reddick Research Group had preexisting stocks of both strains of *E. coli,* and the strains were streaked on Luria Bertani broth (LB) plates for use in the generation of competent *E. coli* cells. The stocks were obtained from the -80.0°C freezer. The plates were then left overnight to grow in an incubator heated at 37. 0°C.The *E. coli* was then to be grown in a liquid culture for the generation of competent cells. Two 15 mL Falcon tubes were selected and 5mL of LB broth was aliquoted into each of them. A colony from each *E. coli* type was then selected with a wooden application stick and swirled into a specified tube broth and then was placed into an incubated New Brunswick Scientific Classic Series Incubation Shaker (C24 Model) for 5 hours to grow until the samples appeared to be a "smokey" consistency. Once the "smokey" consistency was observed, the samples were then spun in a Fisher Scientific Centrific Centrifuge for 5 minutes at speed of 13,000 g.

The supernatant was then discarded into a biowaste container, and the pellets were resuspended in 3mL of autoclaved calcium chloride ($CaCl_2$), which facilitates the uptake of the foreign DNA being administered from the Gibson Assembly in the following steps. The cell suspensions were then sat on ice for 30 minutes for an ice shock treatment and then recentrifuged at the previously discussed parameters to reconstitute the pellets. The supernatant was then discarded again into a biowaste container, and the *E. coli* pellet was resuspended again in 1mL of $CaCl_2$ for final use. This rendered the competent cells completed and ready for the uptake of DNA.

II. B. vii. Gibson transformation into Competent *E. coli*

To ensure uptake of the plasmid, 100µL of fresh competent *E. coli* was added to a microcentrifuge tube along with 2µL of Gibson Assembly product. This was then vortexed briefly and incubated on ice for 30 minutes, incubated at 37.0°C for 2 minutes, and then at room temperature for 10 minutes. To ensure proper growth, 1mL of LB was added to the microcentrifuge tube, and the culture was incubated at 37.0°C for 2 hours until it appeared cloudy. The culture was spun down in an Eppendorf tabletop centrifuge at 8,000 g for 5 minutes, and 1mL of supernatant was discarded. The remainder of the culture was then used in plating. Kanamycin was introduced at this stage due to the plasmid of interest containing the KanR gene encoding for the kanamycin resistance factor, allowing for selective growth of the plasmid of interest. Kanamycin will be added into media in the proceeding steps.

The kanamycin was prepared by using 10mL of $DDIH_2O$ and 300mg of kanamycin, vortexed until dissolved and then filtered through a sterile 0.45-micron syringe filter, it is then aliquoted into 1mL microcentrifuge tubes and froze in a -20.0°C freezer until future use. The competent cell culture was plated by taking 20µL of culture and aliquoting it onto a prepared kanamycin plate and was streaked with a metal scraping tool to spread the cells evenly across the plate. The tool was sterilized in between each use by dipping the tool in 70% ethanol and running the tool under a Bunsen burner for roughly 10-25 seconds then waiting for it to cool to use again. The plates were then left in a 37.0°C incubator overnight.

II. B. viii. Identification of Transformed Products

Sequencing the plasmids was to confirm their identities. This was done by extracting the plasmids from the colonies that had grown on the plates of competent cells.

A Takara low copy plasmid purification kit was used according to the manufacturer's specifications. Once the plasmids were purified, their concentrations were determined by Nanodrop. Samples having concentrations above the 20ng/µL threshold were obtained, those samples were then sent to Plasmidsaurus, a sequencing company, to obtain the final sequencing results of the plasmids. The full competent cell process is shown below in **Figure 15**.

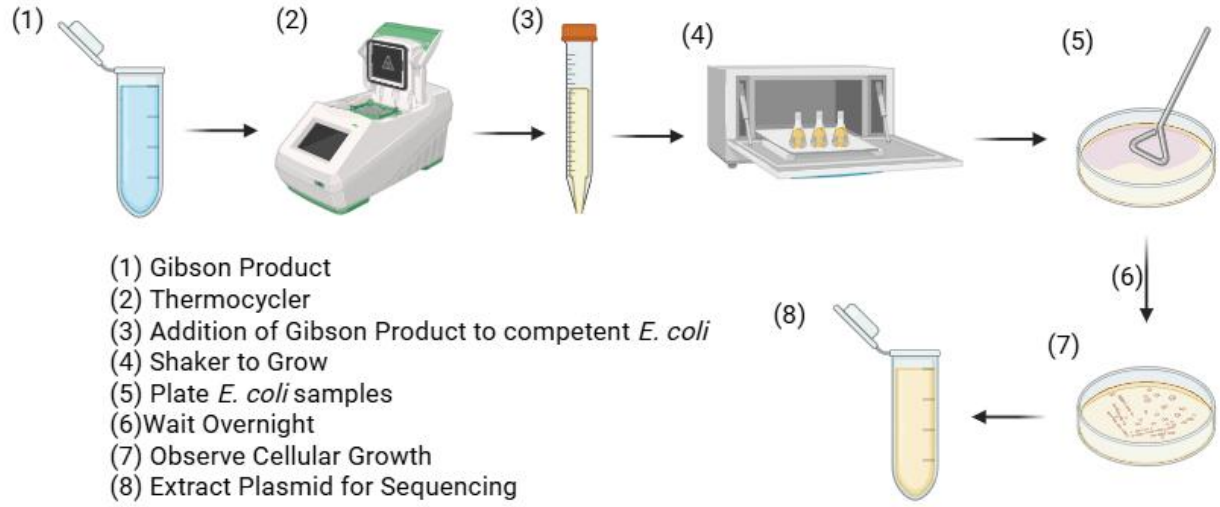


Figure 15: Steps involved in producing the pET-*spd4*-MBP plasmid.

II. C. Experimental

II. C. i. Spd4-MBP – Protein Production

Following the sequencing results, the plasmid B1_3_5 was transformed into the protein strain of *E. coli*. It was done using the same procedure as described before. Freezer -80.0°C stocks of both the DH5α and the BL21-STAR samples were created. The stock procedure is to generate liquid cultures of the plasmid of choice, take 875 µL of stock culture, pipette into a cryogenic tube, and labeled appropriately; next, 125 µL of sterile 80% glycerol was added, and the solution was mixed thoroughly for storage.

These stocks can then be used for longer periods, and the protein production phase can begin.

II. C. ii. Large Scale Protein Growth of Spd4-MBP

The next stage of the protein project was to create a large-scale growth to mass produce protein for testing. This is done by making a 1L of LB and adding 2mL of an overnight starter culture replicating the plasmid. For the LB broth, 10g of tryptone, 10g of NaCl and 5g yeast extract in 1000mL DDIH_2O using a 2800mL Erlenmeyer flask was assembled, stirred for 15 minutes until dissolved, pH adjusted to 7.5, and then autoclaved for sterility. A 1mL sample of the broth is taken for UV-spectrophotometry readings to observe the growth of the culture. This 1mL sample is used as the blank for the UV-vis spectra. The large culture is then cooled and then inoculated with 1mL of 30ng/μL kanamycin and induced with 2mL of the overnight stock of the starter culture of choice. Once this is done, the 1L is then placed into the preheated shaker set at a temperature of 37.0°C. The 1L culture was shaken for 3-5 hours until it reached an absorbance between 0.5-0.6 at 595 nm, then it was induced with 0.2381g of Isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated overnight at the preset temperature.

IPTG is used to induce cells to generate protein, it is added at this stage in the growth cycle due to the number of cells at this point in the growth cycle. If added earlier or later, the IPTG would not be as effective. The following day, the 1L is removed from the shaker and is divided into four centrifuge bottles for the process of separating the cells from the supernatant. The four bottles were balanced, and centrifuged on a Beckman Coulter High Performance centrifuge at 7740g for 1 hour at 4.0°C. Once the centrifuge was completed, the supernatant was discarded and bleached to prevent any potential biological contamination.

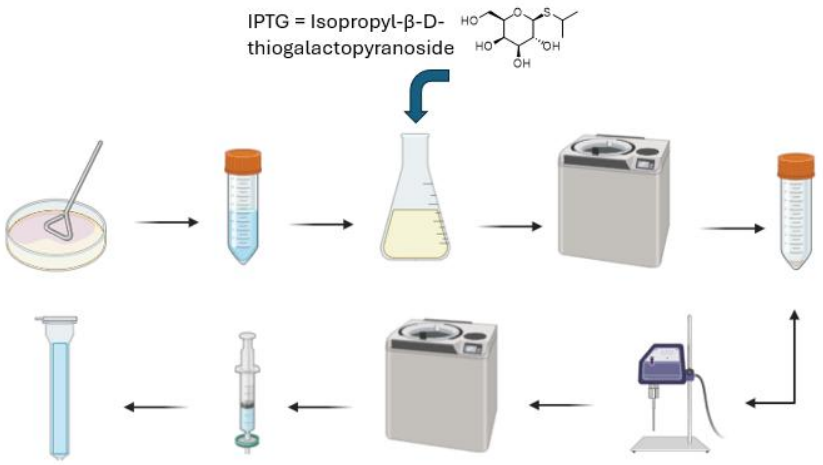


Figure 16: Large-scale protein growth flow diagram.

II. C. iii. Protein Processing of Spd4-MBP

Once done, the pellets were then resuspended in about 30mL of Tris binding buffer made from 0.5M NaCl, 20mM Tris, and 5mM Imidazole at pH 7.9 and then transferred into 50mL Falcon tubes. For storage purposes, the pellets were then recentrifuged for 30 minutes at 10,000 g to reconstitute the pellets for freezing, the supernatant was then discarded into a biohazard container leaving the pellets for storage in the –80.0°C freezer. When the procedure was to be continued, the pellet was resuspended as before and then was sonicated to break open the cells to collect the proteins from inside of the cellular debris. The sonication was done on a Fisher Scientific Sonic Dismembrator on a setting of 16-18 for cycles of 30 seconds on and 30 seconds off for approximately 6 minutes total.

During this process, samples were taken to test the potential of inclusion bodies inside of the cell, this would indicate if there were issues with protein folding during the production process. Once sonication was completed, the sample was then centrifuged at a speed of 16,000 RPM for 20 minutes using a JA-20 rotor in the same centrifuge as used for the large-scale growth samples.

A sample of the "post centrifuge" sample was taken for analysis as well; this will be used later in the process when an SDS-PAGE gel is conducted. The supernatant was then filtered using a 0.45-micron membrane filter to remove debris from the sample. This flow-through was dubbed the "loading sample." This sample was then loaded onto a nickel column to be purified. A full-scale example of the process of large growth is shown above in **Figure 16**.

II. C. iv. Affinity Chromatography of Spd4-MBP

IMAC (Immobilized metal affinity chromatography) will be employed to separate unwanted proteins from the protein of interest. Resin beads loaded with metal, typically nickel, bind to histidine residues on proteins via ligand interactions, allowing target proteins to bind while others flow through the column. A Ni-NTA column (2 mL bed with filter) was equilibrated with binding buffer before loading the crude extract. About 6mL of crude extract was added at a time and was allowed to flow through the column until all the crude extract was added to the column. In some experiments, the collected "crude" sample was rerun through the column for the collection of more potential protein. Next, 20 mL of binding buffer made from the same specifications as previously mentioned, was passed through the column to bind the desired protein to the column for later extraction. The binding flowthrough was also collected for analysis. Once completed, 12 mL of a wash buffer, made with 0.5M NaCl, 20mM Tris, and 60mM Imidazole with a pH of 7.9, was collected and dubbed the "wash" sample, each of these samples was collected in its very own Falcon tube and was set aside for further analysis later.

Once complete, the elution step was to be conducted next, this was the step in which the protein of interest was to be removed from the column and collected. This was done by running 12 mL of elution buffer through the column collected in 1mL fractions for testing.

Later procedures would reduce the amount of total elution buffer to 6mL once the protein was identified. The elution buffer was made from 0.5M NaCl, 20mM Tris-HCl, and 1M Imidazole at a pH of 7.9.

II. C. v. Bradford Assays and Dialysis Procedure

Bradford assays were conducted with 1mL of Bradford dye and 50µL pf protein samples and were left to develop for 10 minutes. They were then blanked and tested for absorbance with a spectrophotometer set at 595nm. The results were then recorded and the elute samples with the highest absorbance values were then to be used in the following dialysis steps. Dialysis was conducted by loading the samples of elutes into Snakeskin dialysis tubing MWCO 7,000 clamped together with two secured dialysis clamps. The device was then laid into a solution of Dialysis buffer (4L of 25mM Tris-HCl at a pH of 8) for overnight dialysis. Once the overnight dialysis process was complete, the sample was removed via pipette and placed into cryogenic tubes for storage. The storage procedure was the same as used with the *E. coli* samples with 80% glycerol.

II. C. vi. SDS-PAGE for Protein Determination

After dialysis, samples were prepared for SDS-PAGE to verify correct protein expression. 50 µL of protein was mixed with 50µL of Laemmli buffer, heated at 65.0°C for 5 minutes, then loaded onto a Bio-Rad gel. Running buffer (0.25M Tris base, 1.9M glycerin, 0.035M SDS in DDI H_2O) was added to submerge the wells. A 10 µL unstained protein ladder standard was loaded into the first and last wells for size reference. Each SDS-PAGE run included 20 µL of crude extract, binding extract, and wash extract, and the final dialyzed elute samples. Occasionally, pre-column samples, like whole cell samples, were also included to assess

expression issues such as inclusion bodies. Gels ran for one hour but sometimes ran for as low as 30 minutes depending on conditions set from the Bio-Rad instrument.

II. C. vii. Other Variables Tested for Protein Optimization

Several variables were tried in optimizing the protein production output. One variable tried was changing the column type used. A cobalt column was also used to determine the affinity of the Spd4-MBP protein. Olivia Shirley, a PhD student in the Reddick Research Group, had success with protein purification using cobalt as the active metal ion in the IMAC system, so it was tried here as well. Her work mainly focused on the non-MBP derivative of the protein. The cobalt affinity column was set up the same way as the nickel column and ran at the same parameters as well. Temperature was also changed during protein production. To determine whether an optimal temperature existed for protein production, an experiment was conducted at a constant 25.0°C for one trial. The eluted protein was tested with nickel and cobalt to determine protein purification methods. In another trial, temperature variation was tried by maintaining the culture at 37.0°C for the first 2-5 hours of growth, then once and absorbance of 0.3-0.4 was achieved the culture was then moved to a pre-cooled shaker set at 18.0°C for the remainder of the growing time. This was ultimately what the final production procedure ended up becoming.

II. C. viii. RebF Protocol and Procedure

Now that the protein was confirmed, to see if it would potentially chlorinate an indole analog, a flavin reductase was going to be needed. RebF was chosen as the flavin reductase for this reaction since it reduces the FAD used in the reactions. We would like to thank the Lewis lab from the University of Indiana for supplying the Reddick Group with the RebF plasmid.

The process in which RebF was obtained is like that of Spd4-MBP but with some slight variations, namely, the use of 100mg/ml ampicillin rather than kanamycin.

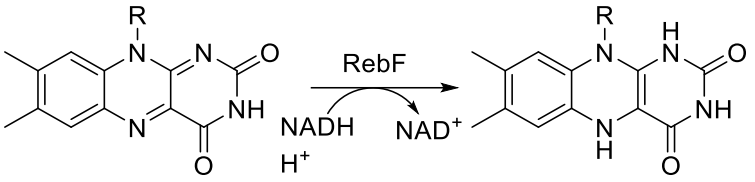


Figure 17: RebF's Functionality.

To grow RebF in excess, the procedure has some slight variations than seen thus far. A liquid culture of 10mL was made with 10 µL ampicillin and one selected colony from the RebF-MBP streak that had been done the night before. This liquid culture was then left to develop overnight. Next day, all 10mL of starter is added to 1L of LB broth along with 1mL of ampicillin which is then grown at a temperature of 37.0°C until an absorbance of 0.6. It was then inoculated with roughly 0.0238 g of IPTG and left to grow for 3 more hours.

Then centrifuged at 4000 g for 30 minutes and pellets were separated and saved in -80.0°C. Pellets used were resuspended in HEPES binding buffer, 25mM at a pH of 7.4, then sonicated for a total of 6 minutes with cycles of 30 second on and 30 seconds off. The sonicated pellet was then centrifuged at 24,000 g for 30 minutes then the supernatant was filtered for use. The loading sample collected was then loaded onto a nickel column for the extraction of the protein. The same column and dialysis process was conducted, along with the same SDS procedure to ensure the correct protein had been created. RebF-MBP has a molecular weight of roughly 60kDa, so the expected band would be located around this marker. The mechanism in which RebF functions is that it uses NADH to reduce FAD to be utilized in the reaction with Spd4-MBP protein to catalyze the chlorination of the substrate in the active site. This reduction is shown above in **Figure 17**.

II. D. Experimental

II. D. i. Enzymatic Reaction Conditions for Spd4-MBP

Once RebF and Spd4-MBP were isolated and purified, enzymatic reactions could be conducted. Enzyme reactions consisted of taking 100 mM NADH, 1 mM FAD, 100 mM NaCl, HEPES Buffer, which is made with 25 mM HEPES and DDIH_2O, Spd4-MBP protein, RebF, and 5 mM of indole. Controls were created by leaving out one of the components and testing them as well.

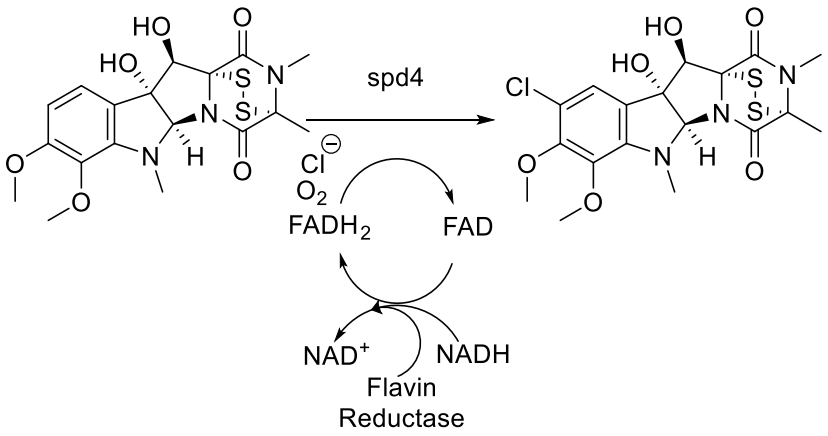


Figure 18:General halogenation of Sporidesmin

Several indoles were proposed for testing, one of which was 6-methoxyindole, shown in **Figure 19**. It was selected due to its similar structure to the indole section of the Sporidesmin molecule and its potential to be chlorinated similarly. It could also have the potential to fit into the active site of the protein in a similar way to that of Sporidesmin, which makes it a good candidate for study.

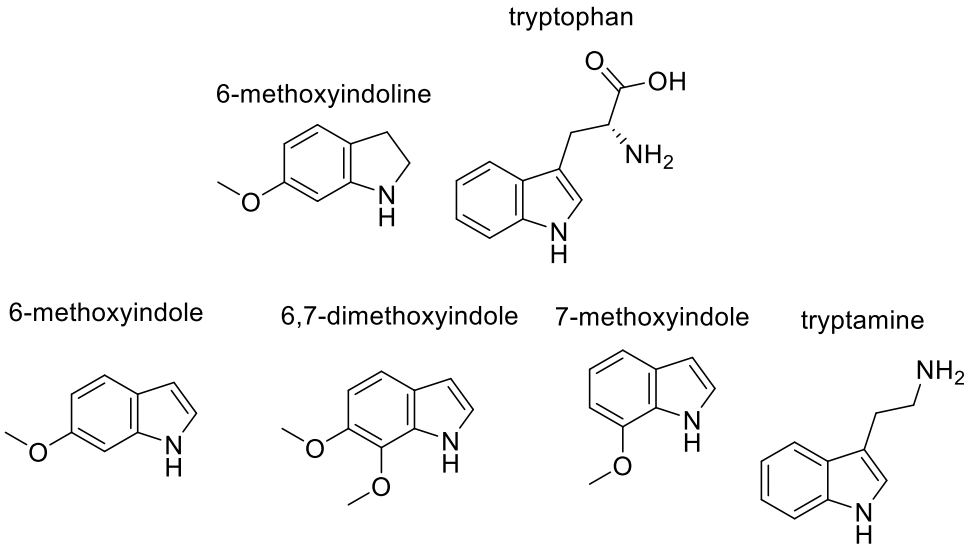


Figure 19: Indoles and Indolines proposed for testing.

Reactions were set up regarding several components to a total reaction mixture. The components are as follows. 100 µL of 100mM NADH, 100 µL of 1mM FAD, 100 µL 100mM NaCl, 400 µL HEPES buffer, 100 µL Spd4-MBP, 100 µL RebF, and finally 100 µL of the testing indole at a concentration of 5mM. These reactions were set up, with one component of each reaction left out to serve as a control for the full reaction and test whether chlorination was occurring on the indole of choice. Once they were set up, they were left in the incubator at 37.0°C for 24, 48, or 72 hours. Once complete, the reactions were placed in a heat block set at 70.0°C for 15 minutes and then centrifuged at max speed (14,000 g) for 5 minutes to process the samples for HPLC analysis. The general expected reaction is pictured above in **Figure 18**, it shows how each of the components act together to form the final halogenated product of Sporidesmin.

II. D. ii. Hight Performance Liquid Chromatography Analysis of Spd4-MBP Reactions

High performance liquid chromatography (HPLC) was used to analyze the components of the experiments. Separation was based on the affinity of each molecule to interact with the column according to its polarity. The system used was an Agilent 1200 equipped with an Eclipse XDB-C18 reverse phase column (5 µm, 4.6 × 150 mm). The system was run isocratically with a 50/50 mobile phase consisting of 1% acetic acid (solvent A) and 1% acetic acid in HPLC grade methanol (solvent B). Samples were injected every 20 minutes and analyzed accordingly period between injections comma a methanol wash was run to cleanse the column comma followed by a 20-minute equilibration period to ensure stable conditions. Additional runs were done with solvent A remaining the same, while solvent B was changed to 100% HPLC grade methanol to compare retention times of products and to enhance chromatogram readings.

II. D. iii. Lewis Lab Spd4-MBP Reaction

A new reaction protocol was derived from the Lewis lab of the University of Indiana after the original protocol showed no definitive results. The reaction contained 10 mM NADH, 100 µM FAD, 10 µM NaCl, 17 µM Spd4-MBP, 17 µM RebF, and 5 mM of the selected indole substrate, all prepared in 25 mM HEPES buffer at pH 7.4. The final reaction volume was 500 µL, composed of 50 µL of each component and 200 µL of HEPES buffer. Six controls were also created similarly to the previous reaction. Reactions were incubated at 30.0°C for 2 hours and were quenched with roughly 100µL of methanol and then were centrifuged for 5 minutes at 14,000g to remove precipitated protein. Several experimental variables were tested, including temperature, concentration, and incubation time. Initial trials extended incubation to 20 to 36 hours, followed by a shorter one-hour reaction to compare yield and conversion efficiency.

With the original batch of 17μM Spd4-MBP ran out, a new preparation was made at a concentration of 8.4μM, it was then used for subsequent tests. Experiments were performed at this lower concentration, which included scaled up reactions for a total volume of 1mL, and tests with LCMS. Samples were also analyzed by HPLC, operating with a flow rate of 1mL/1 min. The mobile phase used consisted of a gradient of 50/50 1% acetic acid in DDIH_2O (solvent A) and HPLC grade methanol (solvent B). The initial conditions were set at 65% A and 35% B, transitioning to 100% B in over 15 minutes. Additional 10-, 20-, and 30-minute runs were performed to identify potential late-eluting compounds and optimize separation quality.

CHAPTER III: *SPD1* AND SPD1-MBP

III. A. Overview and Objectives

The methods and practices for DNA modification, protein preparation, and enzymatic analysis in the *spd1* and *spd1*-MBP projects will be described and outlined. This section covers the methodology behind the practices undergone to further the understanding of the biosynthetic pathway of Sporidesmin and how applicable it could be to other biosynthetic models.

III. B. Experimental

III. B. i. DNA Sequencing and Plasmid Obtainment

The first task was to obtain the genetic sequence of *spd1* and utilize it to produce the Spd1 protein to determine if the assumed homology was correct and if the protein was in fact a SAM-dependent methyltransferase. The sequence for *spd1* was obtained from the NLM, and the gene was ordered as a pET-28a plasmid from Twist Bioscience©.

III. B. ii. Plasmid Preparation for Transformation

The plasmid was redissolved using a preprepared buffer to a concentration of 10 ng/µL using 110.2 µL of 10 mM Tris HCl Elution Buffer. Competent cells DH5α and BL21-STAR were made for the uptake of the *spd1* plasmid. After growth of both *E. coli* types was observed, an overnight starter of 5mL LB was made and after this step, the DH5α was used to make a freezer stock for DNA expression, and the BL21-STAR was used to make a freezer stock along with beginning the protein creation process.

III. B. iii. Spd1 – Protein Production

The small scale and large-scale protein production processes were done the same way as the final protein production phase in the Spd4-MBP project. This was done by preparing the 1L with 2mL of Spd1 starter culture and 1mL of kanamycin. The first 3-5 hours were grown at

37.0°C and then the 1L was then moved to a precooled 18.0°C shaker once an OD_{595} of 0.3-0.4 was obtained. The 1L was then inoculated with 0.2381g of IPTG at an OD_{595} of 0.5-0.6 then left overnight for growth. Cells were processed the same as had been done with Spd4-MBP, with centrifugation and sonication remaining the same. Resuspension of pellets was also kept the same, along with storage requirements. Column chromatography was also to be used in the protein purification of Spd1, so the process was repeated from Spd4-MBP. The procedures remained the same for the dialysis and SDS-PAGE gels. The expected kDa weight of Spd1 is 33 kDa. Contamination was seen in the first gel and inclusion bodies were seen after a follow up purification, so the next step was to employ MBP to try and correct these issues.

III. C. Experimental

III. C. i. *spd1*-MBP – DNA Cloning

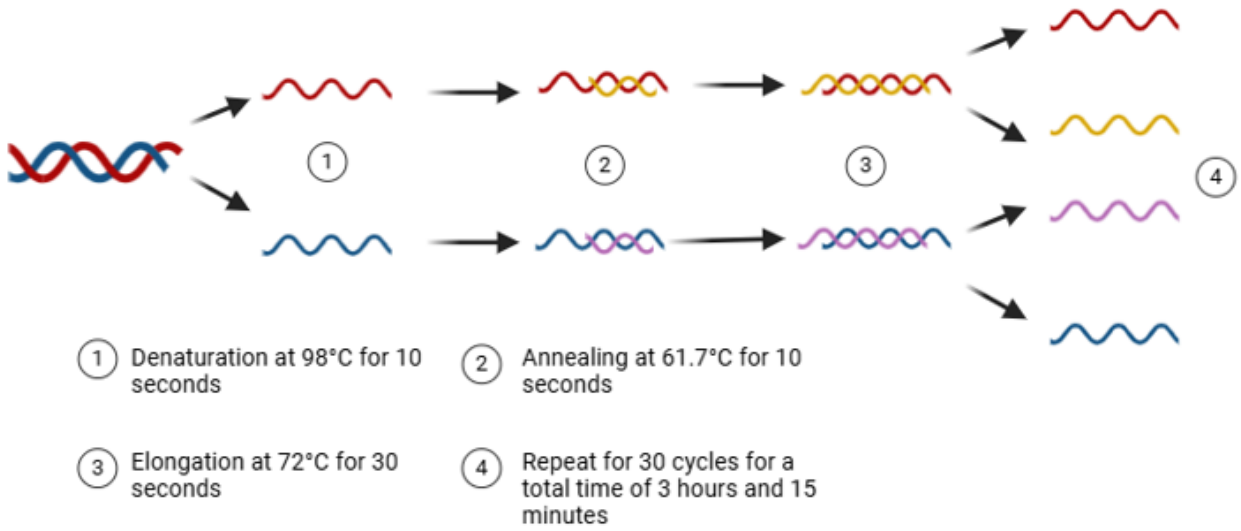


Figure 20: PCR Process of *spd1*-MBP.

First was to obtain both DNA fragments *spd1* and MBP. The goal was to insert the MBP gene into the pET-28-*spd1* plasmid by Gibson Assembly. Primer design was done in Benchling by Stephanine Hernadez for both the *spd1* primers and the MBP primers. The parameters

between *spd1*-MBP and *spd4*-MBP were different due to the enzyme being used in the *spd1*-MBP project. Originally a New England Biolabs DNA polymerase was used, but for *spd1*-MBP project a Primestar DNA polymerase was used. This key change along with the adjustment of PCR parameters, the experiment was able to be conducted to receive the PCR products for the following Gibson Assembly, the PCR process is shown in **Figure 20**. The new parameters for the PCR reaction were a cycle of 98.0°C for 10 seconds, 61.7°C for 40 seconds, and 72.0°C for 1 minute. It was repeated 37 times for a total runtime of approximately 2 hours. A comparison chart is shown in **Table 2**.

Table 2: Comparison Table of the PCR times for *spd4*-MBP and *spd1*-MBP.

	spd4-MBP	*spd1*-MBP
Denaturation	98.0°C for 10 seconds	98.0°C for 10 seconds
Elongation	60.0°C for 30 seconds	61.7.0°C for 40 seconds
Annealing	72.0°C for 30 seconds	72.0°C for 60 seconds
Cycles	30 cycles	37 cycles
Total Time	3 hours and 45 minutes	2 hours

Table describing the differences in PCR conditions for Spd4-MBP and Spd1-MBP. This walks through the different steps of the process from denaturation to annealing.

III. C. ii. Gibson Assembly for *spd1*-MBP

Following the PCR, the fragments need to be assembled. Using Gibson methods, the parameters were calculated using the NEBuilder's online tool and were as follows: 0.5 µL of *spd1* at 0.05 pM, 0.2 µL of MBP at 0.100 pM, 9.3 µL of DDIH2O, and 10.0 µL of the Hifi master mix. The sample was then placed in the thermocycler for 15 minutes at 50.0°C. The Gibson Assembly products were then taken by competent *E. coli*.

During the competent cell uptake process, BL21-STAR had no issues, but DH5α was unsuccessful in up taking the Gibson product. It was determined that using the lab grown stock was inefficient, so a commercially competent strain was ordered and was successful in up taking the Gibson product.

III. C. iii. Transformation of Competent DH5α and BL21-STAR for Spd1-MBP

First, a tube of NEB DH5α competent *E. coli* cells was thawed on ice for 10 minutes. 2µL of the plasmid Gibson product was added to the cell mixture. The tube was flicked 4-5 times to mix the cells and product. Following this, the sample was incubated on ice for 30 minutes and then heat-shocked at 42.0°C for precisely 30 seconds. Then, it was placed back on ice for 5 minutes. Once this cycle of icing and heat shocking was completed, 950µL of room temperature SOC broth (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4, and 20 mM glucose ThermoFisher Scientific) was added to the tube of cells. This was then placed into the 37.0°C shaker for 1 hour at an RPM of 250. Plates were warmed in the incubator for an hour before plating the sample. The cells were mixed thoroughly by flicking the tube; then, a ten-fold dilution was undertaken with the SOC broth. The diluted and undiluted samples were plated, and 75µL of the sample was used in plating onto prewarmed kanamycin plates, and the plates were left overnight for growth to occur. An overnight liquid starter culture was made from the plates to then be used in the extraction kit the following day. A Takara plasmid clean up kit was used to extract the plasmids and then they were sent to Plasmidsaurus© for nanopore sequencing.

III. C. iv. Protein Transformation and Large-Scale Growth

Once sequencing results were obtained, BL21-STAR cells were grown in a 5mL culture for a 1L growth. The protein production and processing phases were carried out how they had

been previously discussed in the final configuration of Spd4-MBP. A nickel column was used to purify the protein, along with dialysis steps and an SDS-PAGE gel to identify the protein generated. The expected kDa weight of Spd1-MBP is 79.5.

III. C. v. Enzymatic Reactions with the Spd1-MBP Protein

Methylation reactions were undertaken with the generated protein. The procedure used was adapted to be used with 6-methoxyindole.[63] The adapted procedure is as follows: A solution of 50 mM of NaH_2PO_4, 300 mM NaCl, and 10% glycerol at a pH of 7.4 was prepared to act as the buffer solution for the ongoing reactions with SAM for both the indole and later the diketopiperazines analogs. 5 mg of procured SAM was diluted in the phosphate buffer to a concentration of 9.5mM. It was then further diluted to a concentration of 1mM for use in further reactions. A stock solution of 153mM 6-methoxyindole was prepared and from that stock, a 5mM stock was prepared. The same was done with the diketopiperazine, however the starting concentration was 33.5mM and the final concentration used for testing was 5mM as well.

III. C. vi. Reaction Conditions with Reagents and HPLC Analysis Conditions

The reaction to see potential methylation was as follows: 50µL of SAM, Indole or Diketopiperazine, 100 µL of Spd1-MBP protein, and 300 µL of sodium phosphate buffer were added together and then incubated for 2 hours at 30.0°C. Once complete, the reaction was then quenched with 200 µL of methanol and centrifuged at 14,000g for 10 minutes. Controls were prepared by leaving one component out of the reaction mixture and were run alongside the full reaction. The reactions were then tested using HPLC and LC-MS. Time, concentration, and temperature were tried to change the outcomes of the reaction conditions. Time was extended to 4,12-,24-, and 48-hour total incubation times, concentrations of substrate and enzyme were varied, and temperature of incubation was altered to 37.0°C to look for potential changes.

III. C. vii. Analysis conditions for the Spd1-MBP reactions

HPLC settings were set to a 50/50 isocratic system of 1%acetic acid in DDH_2O (A) and methanol (B). The HPLC was also run on a gradient which was mentioned previously in Spd4-MBP. This was set at 1 mL/min, and the runs were set at 10 minutes a piece. A methanol "cleansing" run was done in between each sample run to try to eliminate any contamination in between runs. Diketopiperazine trials were done with the original solvents mentioned along with another solvent system; 1% acetic acid in DDH_2O (solvent A) and acetonitrile (solvent B). This was done to potentially better see the DKP since it was unknown at what lambda max would be best for observing the DKP. Methanol absorbs at 205 nm and to negate the potential of losing the DKP to noise, acetonitrile was chosen to see if there was any improvement in visualization.

CHAPTER IV: RESULTS AND CONCLUSIONS

IV. A. *spd4*-MBP Results

This section details the results of the experiments done to obtain the *spd4*-MBP gene fragments and ultimately the *spd4*-MBP plasmid. It explores the avenues taken to obtain the genes of interest and the steps taken to bring them together to form the final plasmid for the expression of protein. By doing this, the research can be replicated if needed by future parties.

IV. A. i. Gel and PCR Results for spd4-MBP

Gel electrophoresis confirmed the PCR products were that of the desired *spd4* insert and the MBP-BURP vector by the location of the bands on the gel show in **Figure 21**. The bands, *spd4* and MBP-BURP, on the gels were estimated to be 1644 base pairs and 6477 base pairs, making the final plasmid 8121 base pairs. The bands shown in lanes 5 and 8 correspond to the correct molecular weight in kilobases for *spd4* and MBP with relatively good intensity for *spd4*, which could indicate that the amount of *spd4* being produced is higher in comparison to MBP since the band for MBP is showing up with less intensity. With this knowledge gained, it was concluded that the products had indeed been made through the PCR process and thus Gibson Assembly could move forward. A positive control test was then done to confirm the results; this is shown in **Figure 22**.

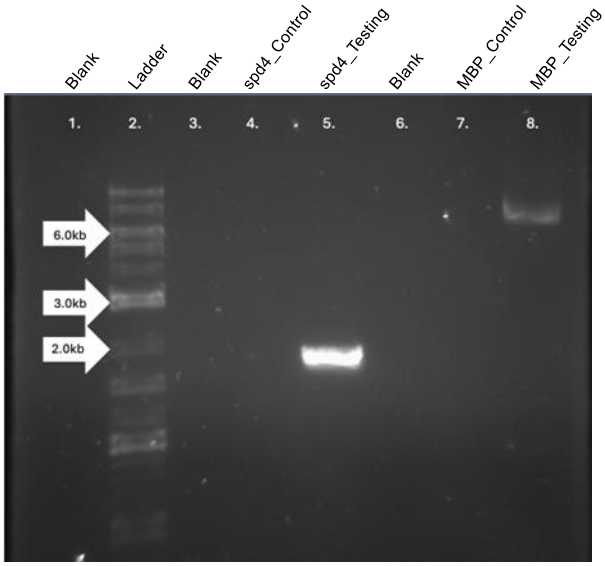


Figure 21: One of the initial gels showing a positive result for both *spd4* and MBP. Gel done on 6/8/23.

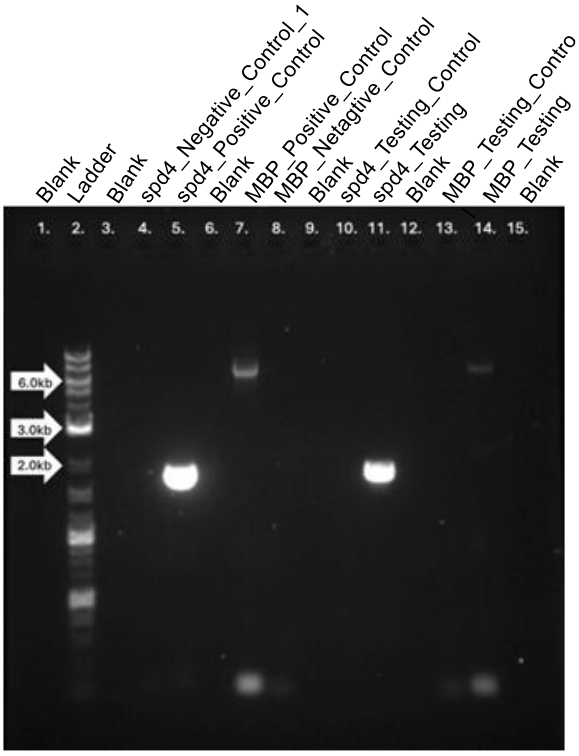


Figure 22: A gel showing positive and negative testing samples for *spd4* and MBP. This gel was taken on 6/15/23.

Figure 22 shows the intensity of *spd4* is retained through the second round of testing, but the intensity of MBP is lessened, in comparison with the positive control, which could be attributed to the size of the fragment being significantly larger than the *spd4* fragment. This could indicate that the fragment has a harder time being produced and is produced less efficiently with lower yields. Nevertheless, the product is still falling at the expected base pair range, therefore the project could move forward with Gibson Assembly of the plasmid of interest. Following the conducted Gibson Assembly, the plasmids were sequenced and the first round of sequences proved unsuccessful. The process was conducted again, and 6 samples were sent for sequencing with 2 returning as potential matches for the plasmid of interest. One was dubbed R1_2 and the other B1_3. R1_2 contained 8118 base pairs, 3 fewer than expected. Sequencing showed that the 3 missing bases would have dropped a stop codon, putting the plasmids functionality at risk. B1_3 did not have any missing base pairs, only 1 silent mutation. **Figure 23** shows the plasmid construct. An alignment revealed that the silent mutation led to another codon but for the same animo acid (His), meaning the mutation had no effect on the final plasmid (shown in **Table 3**).

Table 3: The Segment of the Sequences of Both the MBP-Fusion Construct and the B1_3 Plasmid

Sequence Segment	Sequence
MBP-Spd4-Final	TCGAGCTCCCACCATCACCACCACGCGAAT
B1_3	TCGAGCTCCCACCACCACCACCACGCGAAT

Table showing the mutation found in the sequences of the B1_3 sample and the constructed plasmid. The mutation is notated in red.

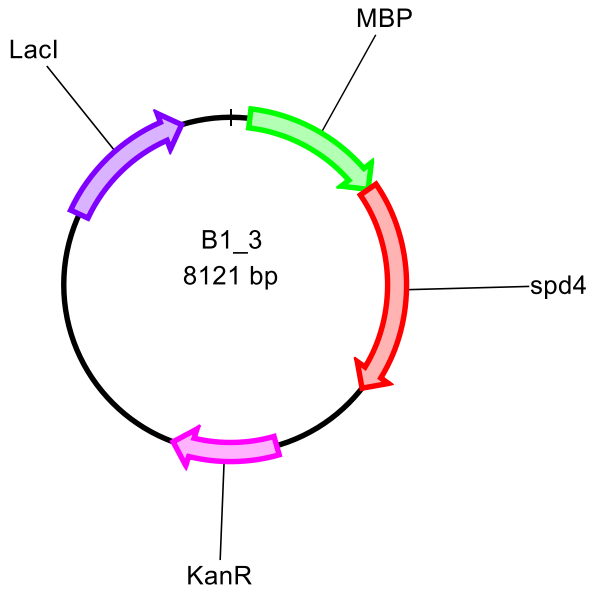


Figure 23: Final sequencing results of B1_3.

IV. A. ii. Protein Production of Spd4-MBP

The large-scale growth for Spd4-MBP was first conducted at 37.0°C, in which the protein purification results that were expected was of a protein falling around the 100 kDa mark since Spd4-MBP is 107kDa in size. But the results in **Figure 24** show protein at roughly the 50 kDa mark in Lane "2/22" along with faint bands located in Lanes "2/29-Elute 1" and "2/29-Elute 2". Indicating that the protein that was being produced was not the correct protein, most likely being just the MBP protein itself which is roughly 43 kDa. Variables like time, temperature, concentration, and purification type would need to be explored to address this issue.

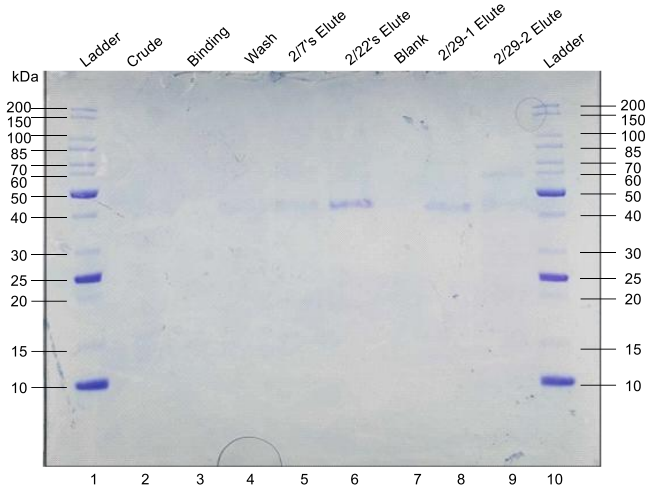


Figure 24: The original protein production methods result; gel was taken on 2/29/24.

Several different conditions were tested for protein production. A protein grown at 37.0°C was conducted but with a change during the growth process that resulted in the expression of viable protein. The large growth was first grown at the original 37.0°C but was moved to an 18.0°C shaker during the process, this ended up being the key to proper protein production. To further optimize the protein production, temperature was investigated again along with induction concentrations, and column chromatography type. Each of these changes resulted in less protein than the newly adapted 37.0°C to 18.0°C method, so the new method was implemented for the remainder of the project. During its purification, the protein showed a brilliant yellow color. This was indicative that the protein was potentially Spd4-MBP due to Spd4-MBP being thought to be flavin dependent halogenase. Flavin proteins produce yellow pigments, so this coloration was indicative of the protein of interest, shown on the column in **Figure 25**. After the dialysis of this protein was conducted, the protein was able to be analyzed for its concentration by using a Bradford assay in conjunction with a previously existing Bovine Albumin calibration curve. With the curve, the calculation of protein concentration could be obtained.

The equation used in this calculation was y=0.001x + 0.005 in which y=absorbance collected from protein sample and x=concentration of the protein sample. Along with this, an SDS-PAGE gel was conducted from the samples collected and this gel shows a band located at 100kDa in Lane “5/22’s Elute” which is where the expected band was to be located, this is shown in **Figure 26.**

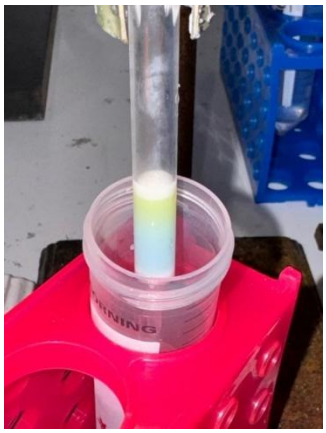


Figure 25: The expression of Spd4-MBP on the nickel column.

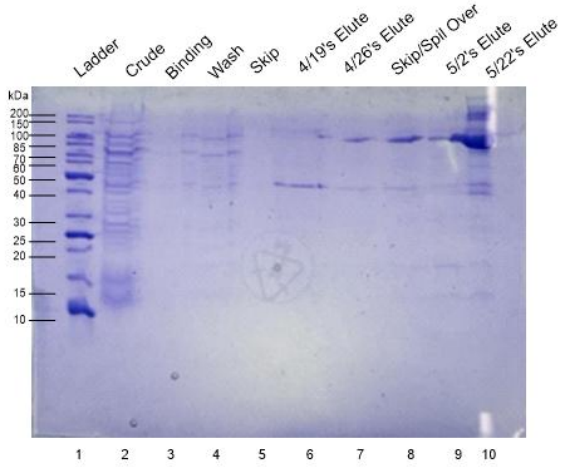


Figure 26: Gel showing the results of the 18C temperature growths and the correct protein molecular weight, gel was taken on 5/22/24.

Following the expression of correct protein, RebF was expressed. **Figure 27** shows an SDS-PAGE gel where RebF and Spd4-MBP have been purified. Spd4-MBP shows a deep band in Lane 2 “Spd4-MBP Elute” indicating a decent protein yield. The gel also shows that a majority of the RebF protein is being lost in the crude extract as shown in Lane “RebF-Crude” which could mean that the purification method could use some needed intervention. Although decent yields of 7.8mg/L for RebF and 6.52mg/L for Spd4-MBP were obtained for the most recent extraction, this is not sufficient for the reactions being conducted. The higher yield of RebF was due to potential purity issues with the Spd4-MBP protein in which the protein being released from the column was not of the highest quality and was being affected by the dialysis step.

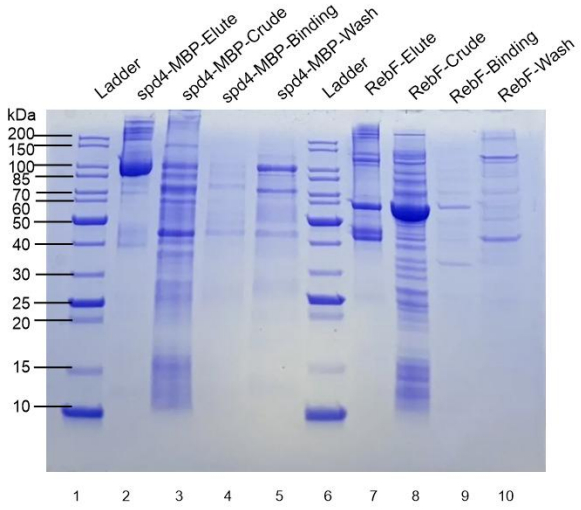


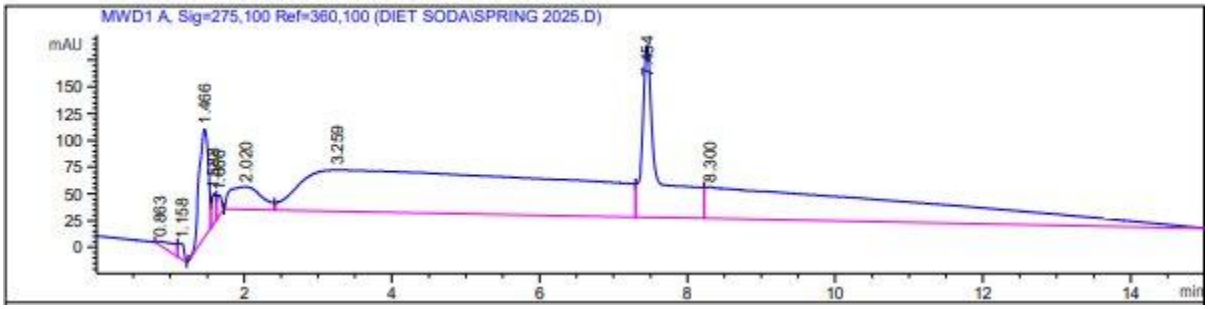
Figure 27: SDS-PAGE gel showing both Spd4-MBP and RebF protein.

IV. A. iii. Reactions with Spd4-MBP

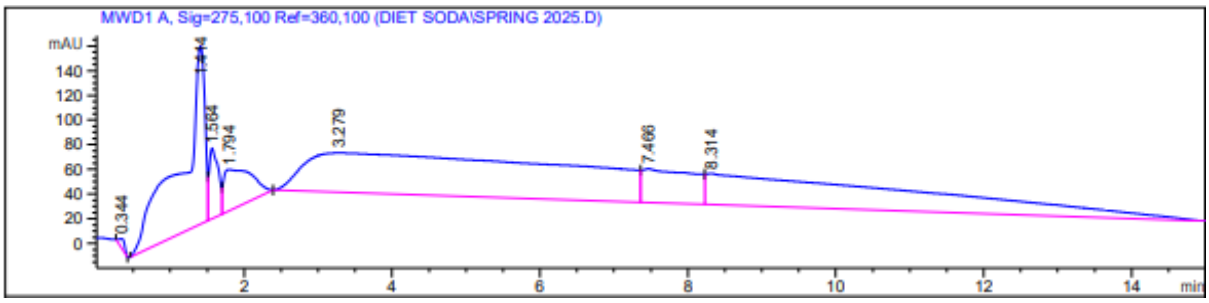
With the conclusion of the DNA work and majority of the protein work accomplished, the remainder of the work needed was to devise reactions to help facilitate the chlorination of the simple indoles chosen for testing.

The ones selected thus far have been 6-methoxyindole, 7-methoxyindole, 6-methoxyindoline, and a few others.

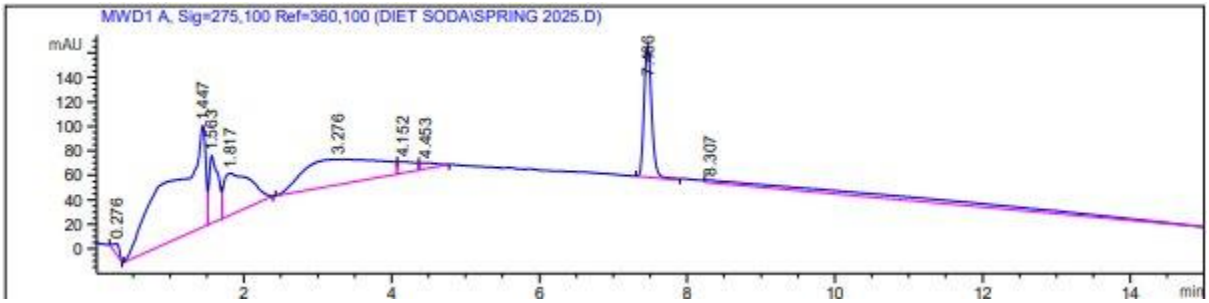
The results of HPLC testing of 6-methoxyindole reveal that no chlorination of note has happened, but some trace amount of some unknown compound has been seen in the new reaction conditions. The retention time of this unknown is 9.7 minutes and can be seen in **Figure 28i**. However, the inability to reproduce this reaction has proven that these results may have been dependent on an unknown factor and may be due to contamination.



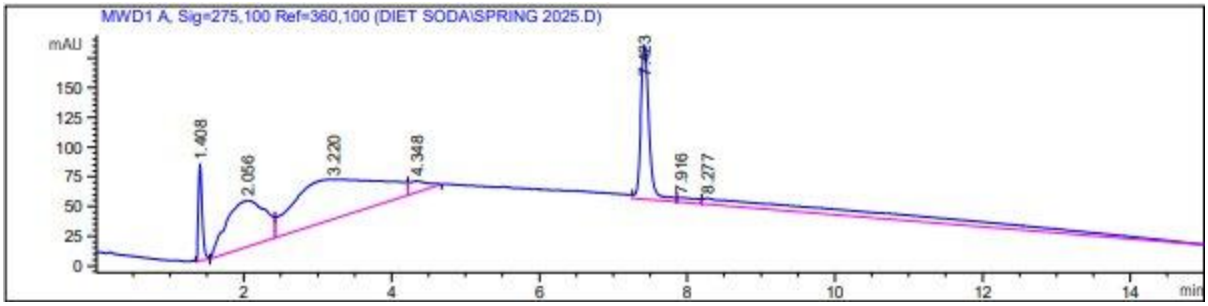
(A) No FAD Control Trial



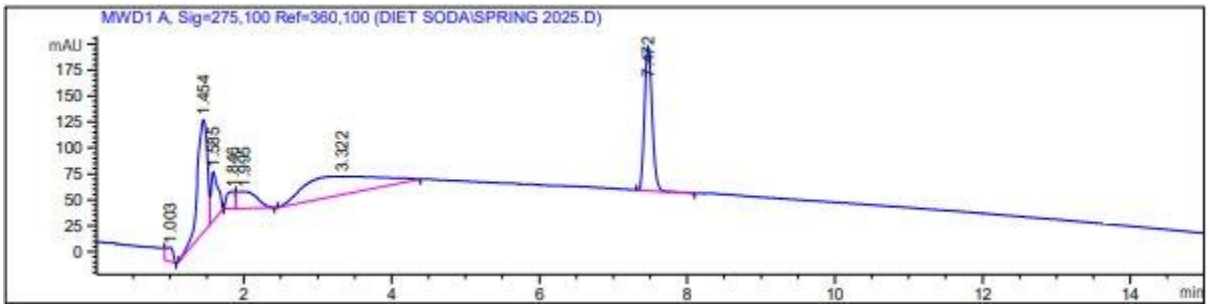
(B) No Indole Control Trial



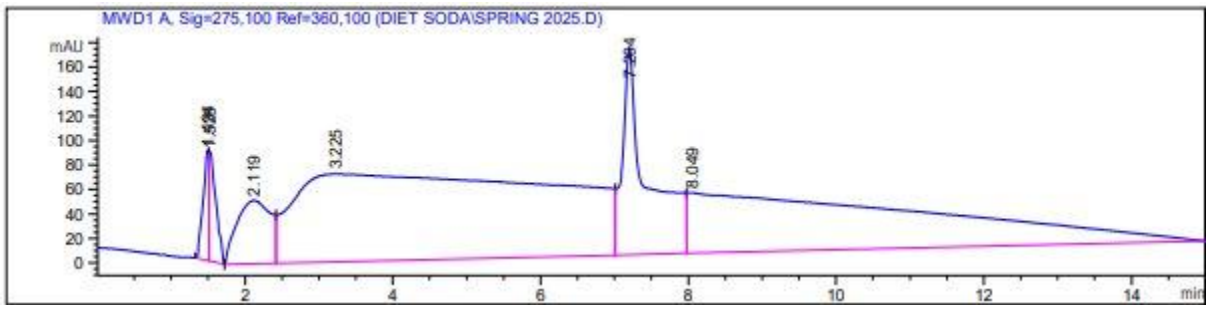
(C) No NaCl Control Trial



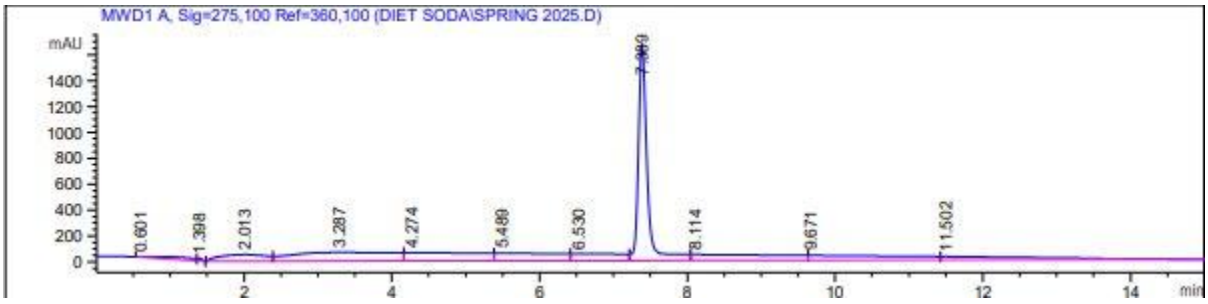
(D) No RebF Control Trial



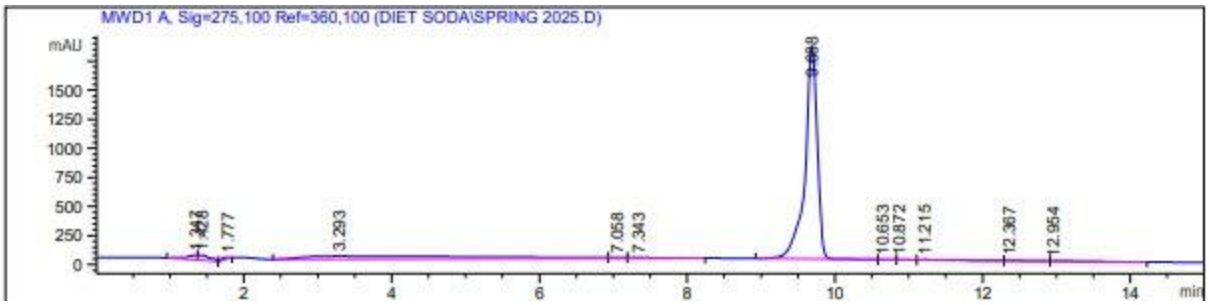
(E) No NADH Control Trial



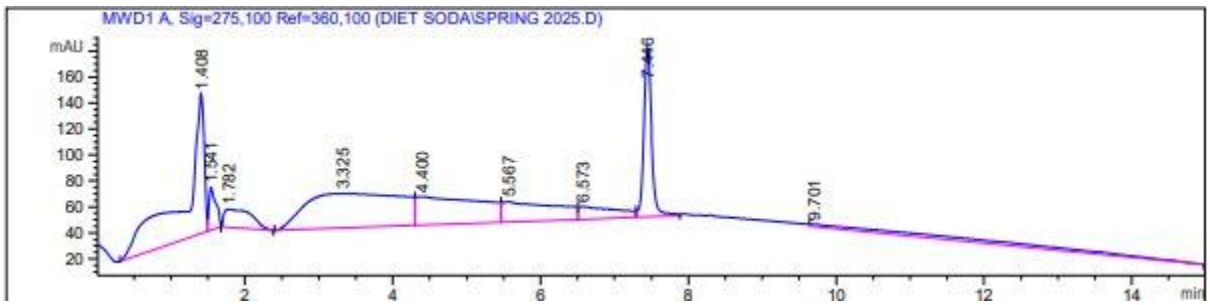
(F) No Spd4-MBP Control Trial



(G) 6-Methoxyindole Control Trial



(H) 5-Chloro-6-Methoxyindole Control Trial



(I) Full Reaction Trial

Figure 28: Full Reaction Panel from the new reaction from the Lewis Lab Paper.

Most of the chromatograms showcase the pure indole coming out at roughly 6.5 or 7.4 minutes depending on when the experiment was run. A chlorinated indole would be predicted to be a more nonpolar compound than its unchlorinated counterpart. Therefore, it would elute after the 6.5/7.5-minute indole peak. What those results point to is that there is potential that a chlorinated indole is forming but at a very little detectable value. As seen from previous tests, the appearance of a reading at the 9.7-minute retention time, in several different reactions, leads to potential of a product being created in the reaction but not enough of said product to be reaching the full limit of detection. One of the issues at hand is the cleanliness of the HPLC itself in cross-contamination and the probability that this trace amount of compound could potentially be nothing of note.

IV. A. iv. LC-MS with Spd4-MBP

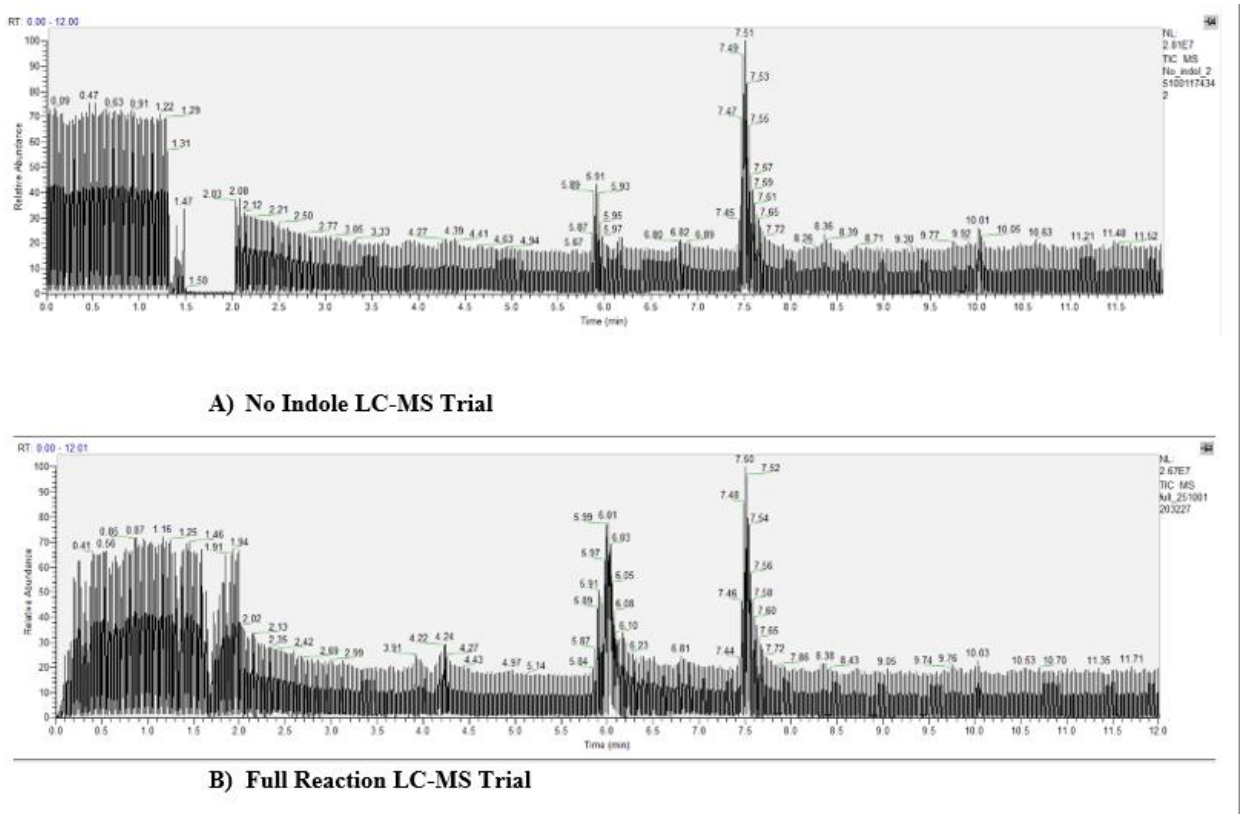


Figure 29: LC-MS Data for the Spd4-MBP trials showing no chlorinated indole in the full reaction trial (B) versus the No indole control (A).

LC-MS Data pertaining to the chlorination reaction of 6-methoxyindole has shown no chlorination. The control reactions show the indole at a m/z of 148.17 appearing at 6.5 minutes and no chlorinated indole, of which would appear at 182.62 is expected farther down the chromatogram from other testing with the chlorinated standard on HPLC testing. Control samples showcase the no indole peak located at the 6.5-minute mark as shown in **Figure 29A** and other components of the reaction. The full reaction run as shown in **Figure 29B** shows no chlorinated product. More LC-MS experiments are planned in the future for this reaction. What is also seen is a large peak around the 7.5-minute mark, with a m/z of 387.12 which has yet to be identified and shows through each trial.

It has been concluded that this fragment may be contamination or a component of the mobile phase of the instrument being used.

IV. B. Spd1 & Spd1-MBP Results

This section details the work done to obtain the *spd1* gene and then the work done to obtain the *spd1*-MBP fusion plasmid. The work described is the culmination of the results obtained from various testing environments done over the period of the project and explained by the facts that were obtained scientifically through the experimental process.

IV. B. i. *spd1* Genetic Obtainment and Protein Growth

The plasmid of choice was able to be ordered then transformed into both strains of *E. coli* to be used in DNA and protein expression work. This was noted by the ability to simply order the plasmid rather than having to order the *spd1* gene than having to Gibson Assembly the plasmid together. DNA work was mostly computerized with codon optimization being done before the ordering of the plasmid took place. The optimization was to better adapt the plasmid to function correctly in the *E. coli* vector. After the process of transforming, large scale growth, and protein purification, a gel was run. Once the gel had been run, **Figure 30** shows 2 bands approximately at the location of 30 kDa in Lane "11/19's Elute" which is roughly the expected mass of Spd1, but since two bands were present there was a high likelihood that these bands were not of the protein of interest.

The process was then redone for a second time to ensure reproducibility. The working hypothesis for the expression of both bands is that potential contamination or the potential presence of another similarly sized protein in the elution phase. A second gel was then performed after a second growth cycle (**Figure 30**). This time however instead of showing two bands in the final elute fractions, the second gel showcased inclusion bodies within the whole cell sample of

the pellet, and no visible protein in the elution sample from this run. This information led to the conclusion being drawn that MBP was going to be needed for the Spd1 protein to solve the issue of the inclusion bodies, and to aid in the general protein folding as seen in the *spd4*-MBP project. This is shown in the sample dubbed "12/22'elute" where there are no significant bands showing in the lane on the gel, and the large, dark band in the "Post Sonicated" sample shown is indicative of inclusion bodies in the sample.

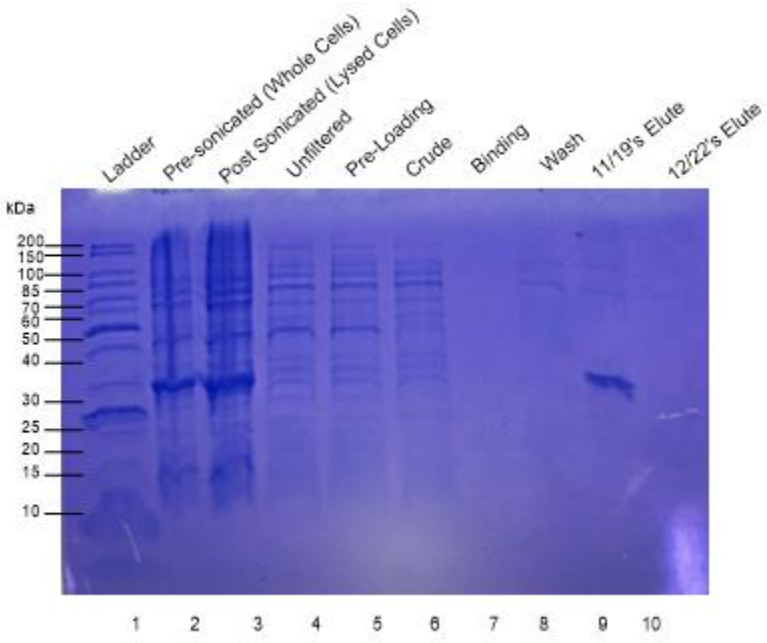


Figure 30: SDS-PAGE gel showing the elutes of Spd1 and the potential contamination seen in the first run, along with the presence of inclusion bodies.

IV. B. ii. MBP Isolation and PCR

Once it was determined that MBP was needed, the *MBP* gene was isolated along with the *spd1* gene, and they were frozen for future use. Isolation techniques used were that of PCR and gel electrophoresis to determine results. A DNA gel showing the presence of both fragments at the expected base pair lengths of 1191 base pairs for *MBP* and 6099 for *spd1* after DNA

extraction, is shown below in **Figure 31**. A second gel was run to create positive and negative controls. This secondary gel is shown below in **Figure 31**.

It shows the Positive and Negative controls and their locations regarding the testing samples, meaning that the testing samples retained the same base pair length as the controls leading to the conclusion that they were the appropriate samples needed for the testing to continue. Along with this, the intensities of the bands remained the same, alluding to the idea that the sample yields were in good standing as well.

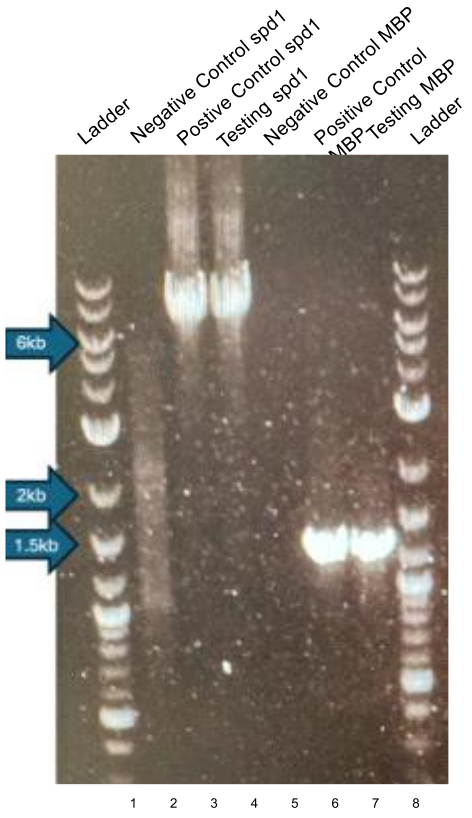


Figure 31: Gel electrophoresis results of PCR products of both *spd1* and MBP.

After creating plasmids and transforming them into both *E. coli* strains, they were plated and placed in the incubator to grow. One plate had been diluted to the specifications of the New England Biolabs procedure, and the other had remained as full concentration as made from the

original sample. The next morning, it was observed that the diluted plate had no colonies, and the non-diluted plate had 4 colonies.

These colonies were then grown in 5mL of LB broth and processed for DNA extraction. The plasmids were then sent for sequencing through Plasmidsaurus. The results of the extraction and sequencing showed that 3 of the 4 plasmids showed an exact match between the base pair length of the construct plasmid produced by Benchling.

The last plasmid had a shorter base pair length than the targeted plasmid sequence. The length of the plasmid of interest was 7284 base pairs. An alignment was performed on one of the new plasmids in comparison to the construct designed in Benchling, and the alignment showed that there were not any missing base pairs or any mutations in the sequence. The plasmid of interest is shown below in **Figure 32**.

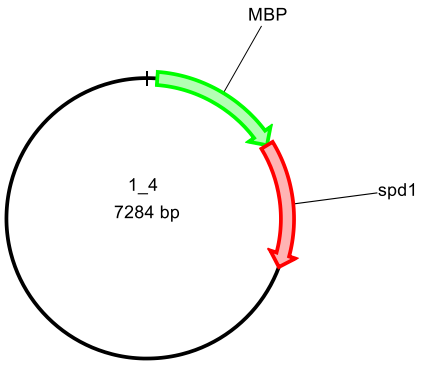


Figure 32: Plasmid that was recovered from the transformations done from PCR products of *spd1* and MBP.

The first sample, dubbed 1_4, was re-streaked onto a kanamycin plate and left overnight for growth. The following day, a 5 mL overnight starter was made for the anticipated plasmid extraction process. The plasmid was extracted using the Takara plasmid miniprep extraction kit according to the manufacturer's specifications. The sample was then used to transform *spd1*-MBP into BL21-STAR for protein expression.

IV. B. iii. Spd1-MBP Protein Production

After 1_4 was transformed into BL21-STAR and underwent the entirety of the protein production process, a gel was run on the eluted extract collected, and the gel (**Figure 33**) shows an intense band at approximately at 79.5kDa.

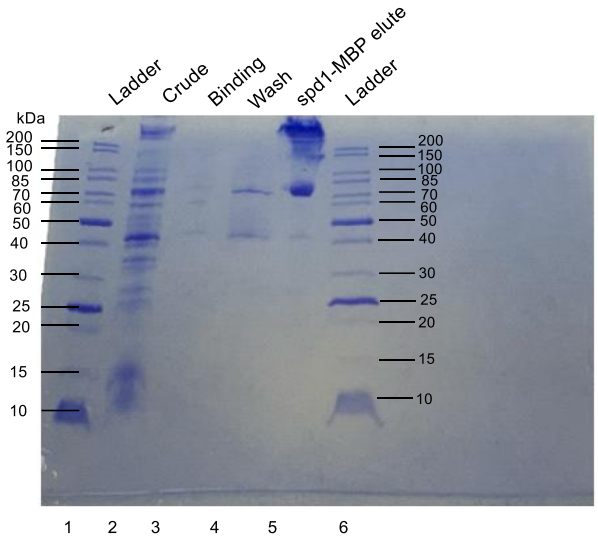


Figure 33: First SDS-PAGE gel showing Spd1-MBP protein samples being detected in the elute sample.

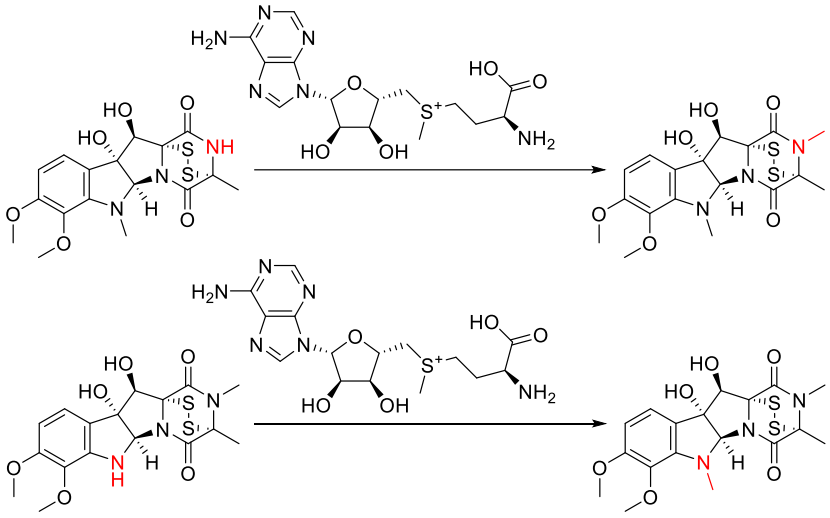


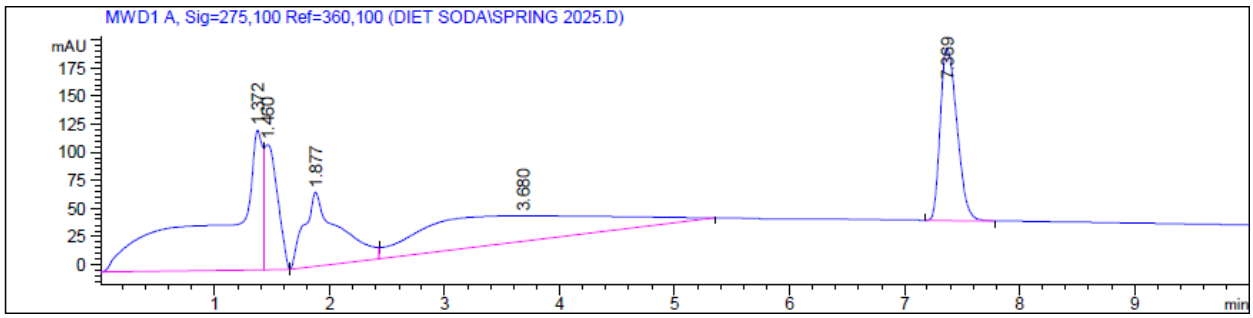
Figure 34: Proposed N-methylation reaction to Sporidesmin

The elutes were then used in chemical reactions that were then analyzed by HPLC. The reactions expected are to follow a similar pathway as derived in **Figure 34**, with an indole of choice instead of the supposed precursor of Sporidesmin. The proposed precursor is not the direct precursor of Sporidesmin, but just one derived from the structure of Sporidesmin. It is not the structure that the fungi use to methylate, since it is currently unknown what part of the biosynthetic pathway the methylation occurs in.

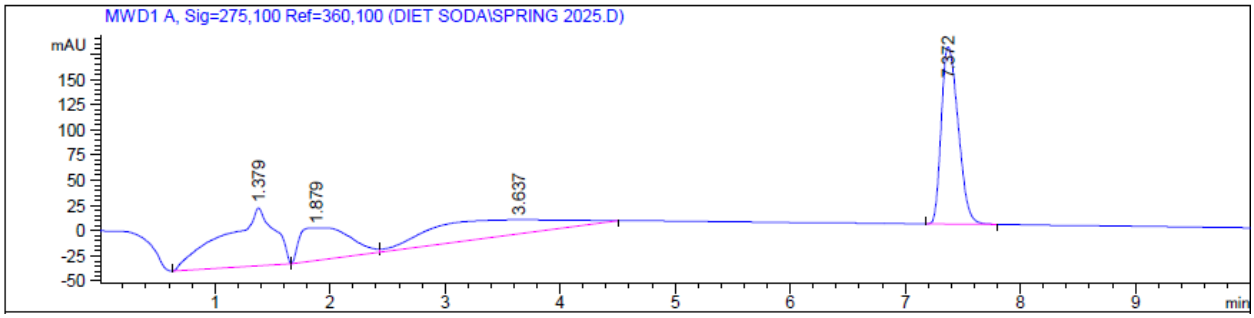
IV. B. iv. Spd1-MBP Enzymatic reactions with Indoles and Diketopiperazines

Reactions with changing concentrations of reagents, temperature, and duration were all tried with various concentrations of enzyme. All reactions showed no results when tested via HPLC. The indole standard, which appeared at 6.4 or 7.4 minutes, did not show any other trace substances beyond this point in the retention times, therefore leading to the conclusion that there were no new products as a methylated compound would be seen as more nonpolar and would have a longer retention time.

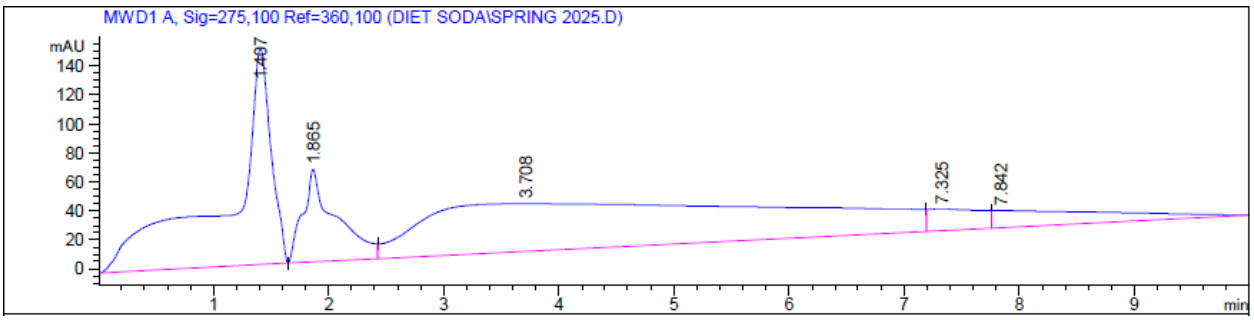
Diketopiperazine (DKP) trials were also tested but showed no results due to solubility issues. The solvent system used mainly for the reactions was 1% acetic acid in DDIH_2O and HPLC grade methanol, but acetic acid and acetonitrile were also tried to see if there were any varying results. Various wavelengths were observed, ranging from 200nm to 275nm to try and observe methylated compounds to no avail. One of the tests run with the DKP once solubilized showed no change in the base peaks leading to the theory that the DKP is becoming insoluble on the column once it is injected into the HPLC, leading to the results shown below in Figure 35. Even using a super saturated solution of 33mM DKP, we see no change in the chromatogram produced. Testing is still ongoing.



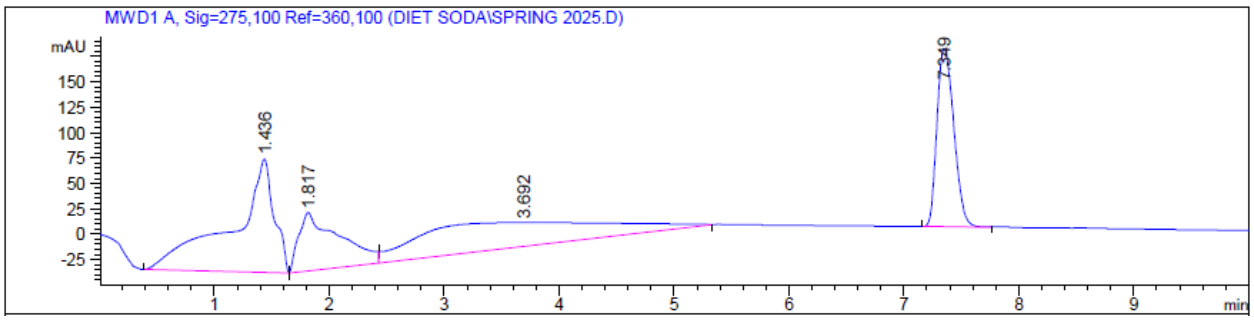
(A) No Spd1-MBP Trial- Spd1-MBP Protein – Indole Test



(B) No SAM Trial-Spd1-MBP Protein – Indole Test



(C) No Indole or Diketopiperazine Trial- Spd1-MBP Protein



(D) Full Indole Reaction with Spd1-MBP Protein

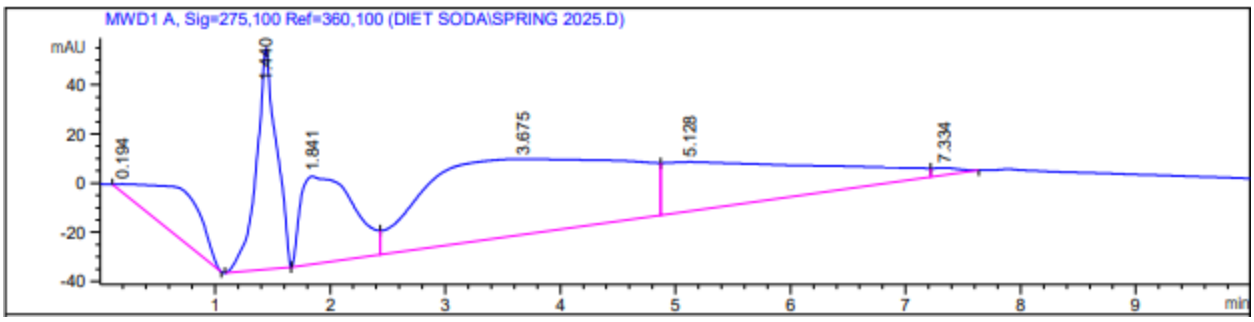


Figure 35E) 33mM Diketopiperazine

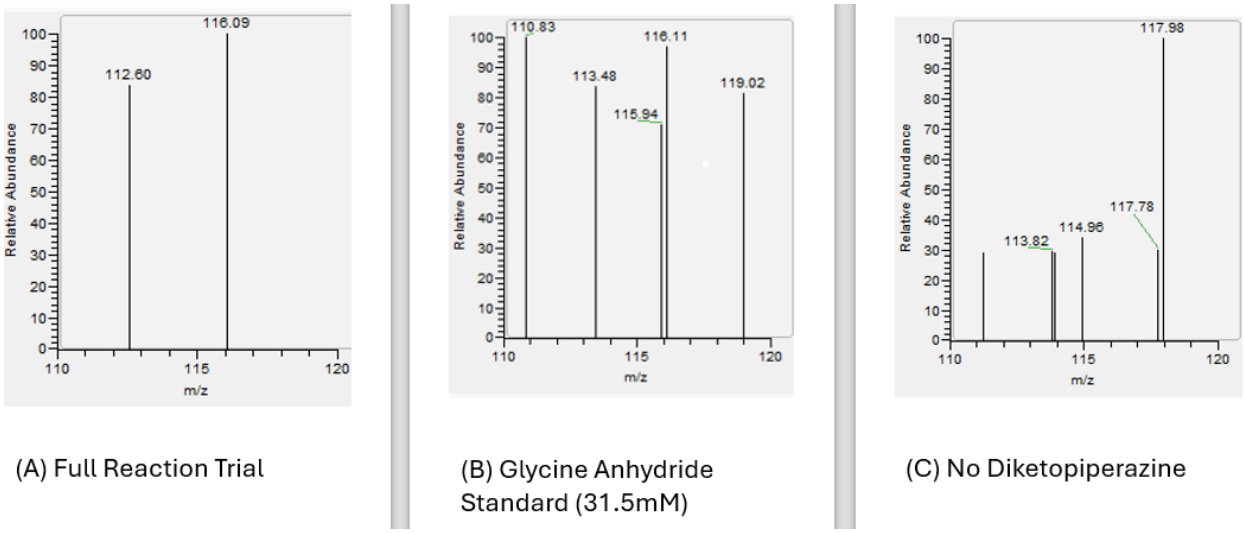


Figure 36: LC-MS Reaction Trials for DKP.

The LC-MS data shows no DKP in the readings and exhibits what was first theorized with the HPLC data; that the DKP is not soluble in the current conditions being used. Since no $[M+1]^+$ peaks of 115.4 (the molecular weight of glycine anhydride is 114.4 g/mol) were detected, as seen in **Figure 36A-C**, it can be concluded that the substrate is not even making it to the detection point. To potentially alleviate this issue, a solution of 0.1%HCl in sodium phosphate buffer was prepared to test for the DKP, only to reveal it was in fact insoluble on the column being used with the HPLC and further testing with a different column type would be needed.

After testing with other solvents like DMSO, acetone, and even 12M HCl, it was concluded that it was in fact possible to get DKP to dissolve in water but just in very small amounts, regarding 1mg per 25mL of water to fully get the DKP to dissolve. Unfortunately, solving this issue still does not solve the issue of the DKP becoming insoluble on the column type of the HPLC or LC-MS, so further testing with other column types is needed in the future to identify this compound for study.

IV. C. Future Work

If possible, Spd4-MBP could potentially work as a biocatalyst for many different types of halogenation reactions if the time can be dedicated to its very intricate biosynthesis and its specific requirements for operations. If the ways in which Spd4-MBP works can be understood and utilized, it would be a valuable enzyme for smaller scale reactions. If the right reagents and the right conditions can be met, the enzyme has the potential to be something great, it could potentially halogenate the substrates we have been looking into, indoles and indolines, and maybe be able to halogenate other small analogs like these like pyrroles and enolates as well. Some work that could be done in the future on this project could entail the expansion into the world of enolates and pyrroles on the substrate scope of enzymatic reactions, working with these structures closely to see if potential halogenation is possible through the work of Spd4-MBP. Alongside this, working with Spd4-MBP in conjunction with another flavin-dependent halogenase would also be a future route that could be taken to assess the enzyme itself and develop the enzyme to better meet functional requirements that may not currently be known now due to the work needed not being yet explored. Yet another avenue to be explored could be the expansion of halogenation to other halogen atoms, like bromine. Using bromine to potentially halogenate a substrate rather than using chlorine to test if there is any difference in the catalytic

value in the reaction would be an asset to this research and would further the understanding of Spd4-MBP as a whole.

Much like Spd4-MBP, Spd1-MBP has great potential as a methylator under the right conditions. If the workings of Spd1-MBP can be understood and the conditions can be met, this enzyme can bring forth new and exciting biocatalytic opportunities for science in which methylation could be done in an enzymatic reaction rather than just done via an organic chemistry reaction. New substrates can be tried, like different types of indolines and others like pyrroles and other more ornate DKPs could potentially be targets of methylation, and the pathway in which the methylation reactions occur could be expanded upon by doing multiple methylations per reaction rather than the expected single methylation reaction as though of with the current reaction conditions. If more time and thought is to be put into this project, furthering the scope of activating the reaction with SAM can be delved into. By doing so, SAM can be better optimized and better utilized as a focal point in the reaction to aid Spd1-MBP in its enzymatic journey.

IV. D. Conclusions

In conclusion the projects of both Spd4-MBP and Spd1-MBP have made insightful impacts on the course of study throughout the duration of the project. As the project's progression was at a steady pace for the first few months, it was hopeful that the project would lead to some gained knowledge for the ever-growing world of biocatalysis and natural products. This project has plenty of avenues left to be explored and many more differing tactics to be employed for testing to reach the final goals of chlorination and methylation. If the right conditions are met and the right reagents are found, these reactions could hold the potential to do great things within the scope of small scale biocatalysis.

REFERENCES

(1) Dingley, J. M. Pithomyces Chartarum, Its Occurrence Morphology, and Taxonomy. *N. Z. J. Agric. Res.* **1962**, *5* (1–2), 49–61. https://doi.org/10.1080/00288233.1962.10419976.

(2) Wearn, J. Pithomyces Chartarum: – A Fungus on the Up? *Field Mycol.* **2009**, *10* (1), 36–37. https://doi.org/10.1016/S1468-1641(10)60497-5.

(3) P. De Carvalho, M.; Abraham, W.-R. Antimicrobial and Biofilm Inhibiting Diketopiperazines. *Curr. Med. Chem.* **2012**, *19* (21), 3564–3577. https://doi.org/10.2174/092986712801323243.

(4) Huber, E. M. Epipolythiodioxopiperazine-Based Natural Products: Building Blocks, Biosynthesis and Biological Activities. *ChemBioChem* **2022**, *23* (23), e202200341. https://doi.org/10.1002/cbic.202200341.

(5) Gardiner, D. M.; Waring, P.; Howlett, B. J. The Epipolythiodioxopiperazine (ETP) Class of Fungal Toxins: Distribution, Mode of Action, Functions and Biosynthesis. *Microbiology* **2005**, *151* (4), 1021–1032. https://doi.org/10.1099/mic.0.27847-0.

(6) Wei, H.; Li, B.; Wang, N.; Ma, Y.; Yu, J.; Wang, X.; Su, J.; Liu, D. Development and Application of Indolines in Pharmaceuticals. *ChemistryOpen* **2023**, *12* (2), e202200235. https://doi.org/10.1002/open.202200235.

(7) Zeng, W.; Han, C.; Mohammed, S.; Li, S.; Song, Y.; Sun, F.; Du, Y. Indole-Containing Pharmaceuticals: Targets, Pharmacological Activities, and SAR Studies. *RSC Med. Chem.* *15* (3), 788–808. https://doi.org/10.1039/d3md00677h.

(8) Di Menna, M. E.; Smith ,B. L.; and Miles, C. O. A History of Facial Eczema (Pithomycotoxicosis) Research. *N. Z. J. Agric. Res.* **2009**, *52* (4), 345–376. https://doi.org/10.1080/00288230909510519.

(9) *Mycotoxins*. https://www.who.int/news-room/fact-sheets/detail/mycotoxins (accessed 2025-09-12).

(10) Omotayo, O. P.; Omotayo, A. O.; Mwanza, M.; Babalola, O. O. Prevalence of Mycotoxins and Their Consequences on Human Health. *Toxicol. Res.* **2019**, *35* (1), 1–7. https://doi.org/10.5487/TR.2019.35.1.001.

(11) Ekwomadu, T.; Mwanza, M.; Musekiwa, A. Mycotoxin-Linked Mutations and Cancer Risk: A Global Health Issue. *Int. J. Environ. Res. Public. Health* **2022**, *19* (13), 7754–7768. https://doi.org/10.3390/ijerph19137754.

(12) Richard, J. L. Some Major Mycotoxins and Their Mycotoxicoses—An Overview. *Int. J. Food Microbiol.* **2007**, *119* (1–2), 3–10. https://doi.org/10.1016/j.ijfoodmicro.2007.07.019.

(13) Munday, R. Studies on the Mechanism of Toxicity of the Mycotoxin, Sporidesmin. V. Generation of Hydroxyl Radical by Sporidesmin. *J. Appl. Toxicol.* **1987**, *7* (1), 17–22. https://doi.org/10.1002/jat.2550070105.

(14) Andorfer, M. C.; Evans, D.; Yang, S.; He, C. Q.; Girlich, A. M.; Vergara-Coll, J.; Sukumar, N.; Houk, K. N.; Lewis, J. C. Analysis of Laboratory-Evolved Flavin-Dependent Halogenases Affords a Computational Model for Predicting Halogenase Site Selectivity. *Chem Catal.* **2022**, *2* (10), 2658–2674. https://doi.org/10.1016/j.checat.2022.07.003.

(15) Phintha, A.; Prakinee, K.; Chaiyen, P. Chapter Eleven - Structures, Mechanisms and Applications of Flavin-Dependent Halogenases. In *The Enzymes*; Chaiyen, P., Tamanoi, F., Eds.; Flavin-Dependent Enzymes: Mechanisms, Structures and Applications; Academic Press, 2020; Vol. 47, pp 327–364. https://doi.org/10.1016/bs.enz.2020.05.009.

(16) Schroeder, L.; Diepold, N.; Gäfe, S.; Niemann, H. H.; Kottke, T. Coupling and Regulation Mechanisms of the Flavin-Dependent Halogenase PyrH Observed by Infrared Difference Spectroscopy. *J. Biol. Chem.* **2024**, *300* (4), 107210–107220. https://doi.org/10.1016/j.jbc.2024.107210.

(17) Moritzer, A.-C.; Minges, H.; Prior, T.; Frese, M.; Sewald, N.; Niemann, H. H. Structure-Based Switch of Regioselectivity in the Flavin-Dependent Tryptophan 6-Halogenase Thal. *J. Biol. Chem.* **2019**, *294* (7), 2529–2542. https://doi.org/10.1074/jbc.RA118.005393.

(18) Lee, C.-H.; Lee, S.-M.; Min, B.-H.; Kim, D.-S.; Jun, C.-H. Ferric(III) Chloride Catalyzed Halogenation Reaction of Alcohols and Carboxylic Acids Using α,α-Dichlorodiphenylmethane. *Org. Lett.* **2018**, *20* (8), 2468–2471. https://doi.org/10.1021/acs.orglett.8b00831.

(19) Büchler, J.; Papadopoulou, A.; Buller, R. Recent Advances in Flavin-Dependent Halogenase Biocatalysis: Sourcing, Engineering, and Application. *Catalysts* **2019**, *9* (12), 1030–1050. https://doi.org/10.3390/catal9121030.

(20) Nock, C. J.; Brell, J. M.; Bokar, J. A.; Cooney, M. M.; Cooper, B.; Gibbons, J.; Krishnamurthi, S.; Manda, S.; Savvides, P.; Remick, S. C.; Ivy, P.; Dowlati, A. A Phase I Study of Rebeccamycin Analog in Combination with Oxaliplatin in Patients with Refractory Solid Tumors. *Invest. New Drugs* **2011**, *29* (1), 126–130. https://doi.org/10.1007/s10637-009-9322-9.

(21) Schwandt, A.; Mekhail, T.; Halmos, B.; O'Brien, T.; Ma, P. C.; Fu, P.; Ivy, P.; Dowlati, A. Phase-II Trial of Rebeccamycin Analog, a Dual Topoisomerase-I and -II Inhibitor, in Relapsed "Sensitive" Small Cell Lung Cancer. *J. Thorac. Oncol.* **2012**, *7* (4), 751–754. https://doi.org/10.1097/JTO.0b013e31824abca2.

(22) Liu, C.; Bayer, A.; Cosgrove, S. E.; Daum, R. S.; Fridkin, S. K.; Gorwitz, R. J.; Kaplan, S. L.; Karchmer, A. W.; Levine, D. P.; Murray, B. E.; J. Rybak, M.; Talan, D. A.; Chambers, H. F. Clinical Practice Guidelines by the Infectious Diseases Society of America for the Treatment of Methicillin-Resistant Staphylococcus Aureus Infections in Adults and Children: Executive Summary. *Clin. Infect. Dis.* **2011**, *52* (3), 285–292. https://doi.org/10.1093/cid/cir034.

(23) Mühlberg, E.; Umstätter, F.; Kleist, C.; Domhan, C.; Mier, W.; Uhl, P. Renaissance of Vancomycin: Approaches for Breaking Antibiotic Resistance in Multidrug-Resistant Bacteria. *Can. J. Microbiol.* **2019**, *66* (1), 11–16. https://doi.org/10.1139/cjm-2019-0309.

(24) Schmartz, P. C.; Zerbe, K.; Abou-Hadeed, K.; Robinson, J. A. Bis-Chlorination of a Hexapeptide–PCP Conjugate by the Halogenase Involved in Vancomycin Biosynthesis. *Org Biomol Chem* **2014**, *12* (30), 5574–5577. https://doi.org/10.1039/C4OB00474D.

(25) Yeh, E.; Garneau, S.; Walsh, C. T. Robust in Vitro Activity of RebF and RebH, a Two-Component Reductase/Halogenase, Generating 7-Chlorotryptophan during Rebeccamycin Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (11), 3960–3965. https://doi.org/10.1073/pnas.0500755102.

(26) Abdelraheem, E.; Thair, B.; Varela, R. F.; Jockmann, E.; Popadić, D.; Hailes, H. C.; Ward, J. M.; Iribarren, A. M.; Lewkowicz, E. S.; Andexer, J. N.; Hagedoorn, P.-L.; Hanefeld, U. Methyltransferases: Functions and Applications. *ChemBioChem* **2022**, *23* (18), e202200212–e202200231. https://doi.org/10.1002/cbic.202200212.

(27) Katz, J. E.; Dlakić, M.; Clarke, S. Automated Identification of Putative Methyltransferases from Genomic Open Reading Frames. *Mol. Cell. Proteomics* **2003**, *2* (8), 525–540. https://doi.org/10.1074/mcp.M300037-MCP200.

(28) Zhang, W.; Xu, J. DNA Methyltransferases and Their Roles in Tumorigenesis. *Biomark. Res.* **2017**, *5* (1), 1–8. https://doi.org/10.1186/s40364-017-0081-z.

(29) Patel, S. Drivers of Bacterial Genomes Plasticity and Roles They Play in Pathogen Virulence, Persistence and Drug Resistance. *Infect. Genet. Evol.* **2016**, *45*, 151–164. https://doi.org/10.1016/j.meegid.2016.08.030.

(30) Liscombe, D. K.; Louie, G. V.; Noel, J. P. Architectures, Mechanisms and Molecular Evolution of Natural Product Methyltransferases. *Nat. Prod. Rep.* **2012**, *29* (10), 1238–1250. https://doi.org/10.1039/c2np20029e.

(31) Cantoni, G. L. The Nature of the Active Methyl Donor Formed Enzymatically from L-Methionine and Adenosinetriphosphate. *J. Am. Chem. Soc.* **1952**, *74* (11), 2942–2943. https://doi.org/10.1021/ja01131a519.

(32) Ding, W.; Smulan, L. J.; Hou, N. S.; Taubert, S.; Watts, J. L.; Walker, A. K. S-Adenosylmethionine Levels Govern Innate Immunity through Distinct Methylation-Dependent Pathways. *Cell Metab.* **2015**, *22* (4), 633–645. https://doi.org/10.1016/j.cmet.2015.07.013.

(33) *SAM and SAH: Key Differences, SAM Cycle, and Methylation Regulation - Creative Proteomics Blog*. https://www.creative-proteomics.com/blog/sam-sah-key-differences-sam-cycle.htm (accessed 2025-10-01).

(34) *S-Adenosylmethionine (sam)*. https://ww2.uthscsa.edu/learning-modules/biochem/08-amino-acid-metabolism/amino-acid-metabolism8.html (accessed 2025-09-19).

(35) Haroon, M.; Ahmad, S.; Fawad Zahoor, A.; Javed, S.; Nadeem Ahmad, M.; Gul Khan, S.; Al-Mutairi, A. A.; Irfan, A.; Al-Hussain, S. A.; Zaki, M. E. A. Grignard Reaction: An 'Old-Yet-Gold' Synthetic Gadget toward the Synthesis of Natural Products: A Review. *Arab. J. Chem.* **2024**, *17* (5), 105715-105742. https://doi.org/10.1016/j.arabjc.2024.105715.

(36) Walter, M.; Ciupak, O.; Biernacki, K.; Rachoń, J.; Witt, D.; Demkowicz, S. Convenient and Efficient N-Methylation of Secondary Amines under Solvent-Free Ball Milling Conditions. *Sci. Rep.* **2024**, *14* (1), 8810-8817. https://doi.org/10.1038/s41598-024-59374-z.

(37) Yan, G.; Li, X.; Yang, J.; Li, Z.; Hou, J.; Rao, B.; Hu, Y.; Ma, L.; Wang, Y. Cost-Effective Production of ATP and S-Adenosylmethionine Using Engineered Multidomain Scaffold Proteins. *Biomolecules* **2021**, *11* (11), 1706-1721. https://doi.org/10.3390/biom11111706.

(38) Lin, H. S-Adenosylmethionine-Dependent Alkylation Reactions: When Are Radical Reactions Used? *Bioorganic Chem.* **2011**, *39* (5–6), 161–170. https://doi.org/10.1016/j.bioorg.2011.06.001.

(39) Kapust, R. B.; Waugh, D. S. Escherichia Coli Maltose-Binding Protein Is Uncommonly Effective at Promoting the Solubility of Polypeptides to Which It Is Fused. *Protein Sci.* **1999**, *8* (8), 1668–1674. https://doi.org/10.1110/ps.8.8.1668.

(40) Sun, P.; Tropea, J. E.; Waugh, D. S. Enhancing the Solubility of Recombinant Proteins in Escherichia Coli by Using Hexahistidine-Tagged Maltose-Binding Protein as a Fusion Partner. In *Heterologous Gene Expression in E.coli: Methods and Protocols*; Evans, Jr., Thomas C., Xu, M.-Q., Eds.; Humana Press: Totowa, NJ, 2011; pp 259–274. https://doi.org/10.1007/978-1-61737-967-3_16.

(41) Duplay, P.; Bedouelle, H.; Fowler, A.; Zabin, I.; Saurin, W.; Hofnung, M. Sequences of the malE Gene and of Its Product, the Maltose-Binding Protein of Escherichia Coli K12. *J. Biol. Chem.* **1984**, *259* (16), 10606–10613. https://doi.org/10.1016/S0021-9258(18)91005-7.

(42) Duan, X.; Hall, J. A.; Nikaido, H.; Quiocho, F. A. Crystal Structures of the Maltodextrin/Maltose-Binding Protein Complexed with Reduced Oligosaccharides: Flexibility of Tertiary Structure and Ligand Binding. *J. Mol. Biol.* **2001**, *306* (5), 1115–1126. https://doi.org/10.1006/jmbi.2001.4456.

(43) Gibson, D. Assembly of Large Nucleic Acids, March 14, 2017. https://patentimages.storage.googleapis.com/b9/75/71/0d2318c09263a2/US9593329.pdf (accessed 2025-11-12)

(44) *How Gibson Assembly® is Changing Synthetic Biology | NEB*. https://www.neb.com/en-us/tools-and-resources/feature-articles/gibson-assembly-building-a-synthetic-biology-toolset (accessed 2025-10-01).

(45) Staff, L. in the L. *Gibson Assembly 101: Expert Cloning Tips You Need to Know*. Life in the Lab. https://www.thermofisher.com/blog/life-in-the-lab/gibson-assembly-101-expert-cloning-tips-you-need-to-know/ (accessed 2025-09-15).

(46) Olszakier, S.; Berlin, S. A Simplified Gibson Assembly Method for Site Directed Mutagenesis by Re-Use of Standard, and Entirely Complementary, Mutagenesis Primers. *BMC Biotechnol.* **2022**, *22* (1), 10-20. https://doi.org/10.1186/s12896-022-00740-y.

(47) *Introduction to Affinity Chromatography | Bio-Rad*. https://www.bio-rad.com/en-us/applications-technologies/introduction-affinity-chromatography?ID=LUSMJIDN (accessed 2025-09-15).

(48) *Basic Theory*. Chemistry LibreTexts. https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_(Analytical_Chemistry)/Instrumentation_and_Analysis/Affinity_Chromatography/02_Theory/01_Basic_Theory (accessed 2025-09-16).

(49) Carter, T. D.; Outten, F. W. Ni-NTA Affinity Chromatography to Characterize Protein–Protein Interactions During Fe-S Cluster Biogenesis. *Methods Mol. Biol. Clifton NJ* **2021**, *2353*, 125–136. https://doi.org/10.1007/978-1-0716-1605-5_7.

(50) *His-tagged Proteins–Production and Purification - US*. https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/his-tagged-proteins-production-purification.html (accessed 2025-09-16).

(51) Khow, O.; Suntrarachun, S. Strategies for Production of Active Eukaryotic Proteins in Bacterial Expression System. *Asian Pac. J. Trop. Biomed.* **2012**, *2* (2), 159–162. https://doi.org/10.1016/S2221-1691(11)60213-X.

(52) Myers, K. S.; Noguera, D. R.; Donohue, T. J. Promoter Architecture Differences among Alphaproteobacteria and Other Bacterial Taxa. *mSystems* **2021**, *6* (4), 10.1128/msystems.00526-21-39. https://doi.org/10.1128/msystems.00526-21.

(53) Edgell, D. R.; Belfort, M.; Shub, D. A. Barriers to Intron Promiscuity in Bacteria. *J. Bacteriol.* **2000**, *182* (19), 5281–5289. https://doi.org/10.1128/jb.182.19.5281-5289.2000.

(54) Reuten, R.; Nikodemus, D.; Oliveira, M. B.; Patel, T. R.; Brachvogel, B.; Breloy, I.; Stetefeld, J.; Koch, M. Maltose-Binding Protein (MBP), a Secretion-Enhancing Tag for Mammalian Protein Expression Systems. *PLoS ONE* **2016**, *11* (3), e0152386. https://doi.org/10.1371/journal.pone.0152386.

(55) Nelson, D. L.; Cox, M. M. Lehninger - Principles of Biochemistry.

(56) Hu, M.; Zhang, C.; Mu, Y.; Shen, Q.; Feng, Y. Indole Affects Biofilm Formation in Bacteria. *Indian J. Microbiol.* **2010**, *50* (4), 362–368. https://doi.org/10.1007/s12088-011-0142-1.

(57) Babalola, B. A.; Malik, M.; Olowokere, O.; Adebesin, A.; Sharma, L. Indoles in Drug Design and Medicinal Chemistry. *Eur. J. Med. Chem. Rep.* **2025**, *13*, 100252-100272. https://doi.org/10.1016/j.ejmcr.2025.100252.

(58) Bojarska, J.; Mieczkowski, A.; Ziora, Z. M.; Skwarczynski, M.; Toth, I.; Shalash, A. O.; Parang, K.; El-Mowafi, S. A.; Mohammed, E. H. M.; Elnagdy, S.; AlKhazindar, M.; Wolf, W. M. Cyclic Dipeptides: The Biological and Structural Landscape with Special Focus on the Anti-Cancer Proline-Based Scaffold. *Biomolecules* **2021**, *11* (10), 1515-1579. https://doi.org/10.3390/biom11101515.

(59) Taga, Y.; Kusubata, M.; Ogawa-Goto, K.; Hattori, S. Identification of Collagen-Derived Hydroxyproline (Hyp)-Containing Cyclic Dipeptides with High Oral Bioavailability: Efficient Formation of Cyclo(X-Hyp) from X-Hyp-Gly-Type Tripeptides by Heating. *J. Agric. Food Chem.* **2017**, *65* (43), 9514–9521. https://doi.org/10.1021/acs.jafc.7b03714.

(60) Wang, H.; Usui, T.; Osada, H.; Ganesan, A. Synthesis and Evaluation of Tryprostatin B and Demethoxyfumitremorgin C Analogues. *J. Med. Chem.* **2000**, *43* (8), 1577–1585. https://doi.org/10.1021/jm9905662.

(61) Chai, C. L. L.; Waring, P. Redox Sensitive Epidithiodioxopiperazines in Biological Mechanisms of Toxicity. *Redox Rep.* **2000**, *5* (5), 257–264. https://doi.org/10.1179/135100000101535799.

(62) Mortimer, P. H.; Collins, B. S. The in Vitro Toxicity of the Sporidesmins and Related Compounds to Tissue-Culture Cells. *Res. Vet. Sci.* **1968**, *9* (2), 136–142. https://doi.org/10.1016/S0034-5288(18)34573-9.

(63) Lee, S.; Kang, J.; Kim, J. Structural and Biochemical Characterization of Rv0187, an O-Methyltransferase from Mycobacterium Tuberculosis. *Sci. Rep.* **2019**, *9* (1), 8059-8071. https://doi.org/10.1038/s41598-019-44592-7.