Exploring BioAB: An Enzyme Involved in the Biosynthesis of Biotin

Presented to the faculty of Lycoming College in partial fulfillment of the requirements for Departmental Honors in Chemistry

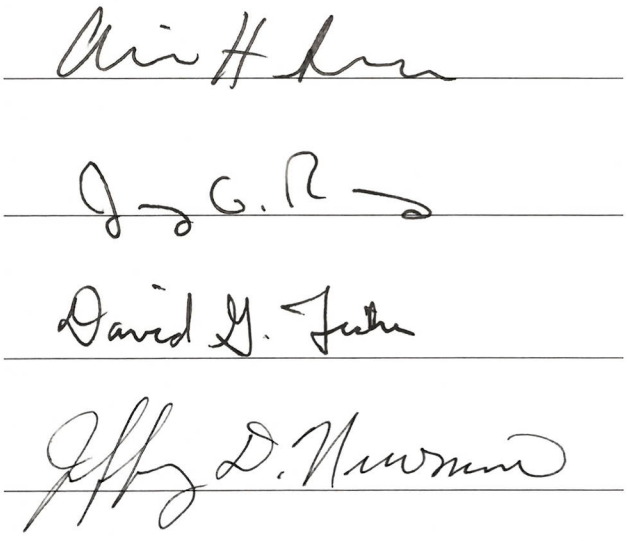
By

Chase Byerly

Lycoming College

May 2022

Approved by:



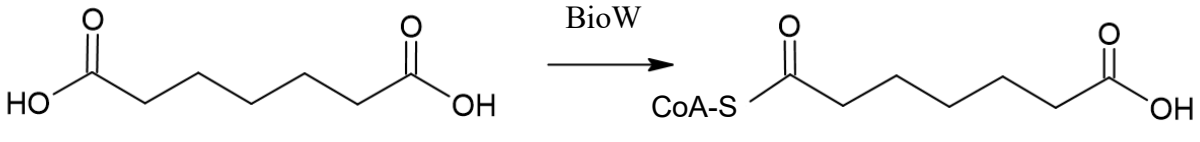
This work was supported by Joanne and Arthur Haberberger Fellowship, awarded to Chase L. Byerly at Lycoming College, Williamsport, PA.

Abstract

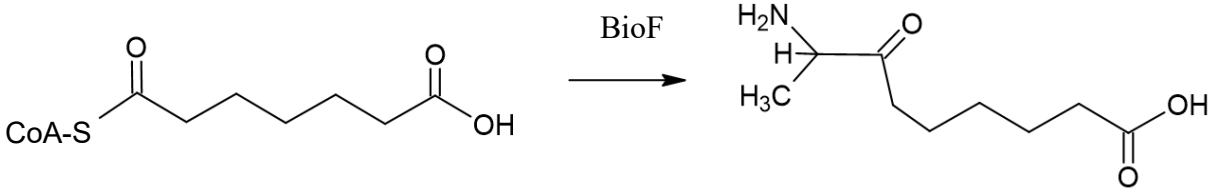
Biotin is an important molecule in the citric acid cycle, synthesis of fatty acids, and various other reactions in humans. However, biotin cannot be synthesized by humans; instead, we rely on our diet and bacteria in our gastrointestinal tract to produce biotin. In *Bacteroides thetaiotaomicron*, an enzyme known as BioAB is used to catalyze aspects of the biosynthesis of biotin. Little is known about the BioAB enzyme and how it functions. To examine its role, the bioAB gene was to be ligated into a pET28a plasmid. However, difficulties arose with purifying the bioAB gene and conducting the transformation. This work focused on troubleshooting the difficulty isolating the bioAB sequence, which was caused by the plasmid and the bioAB gene being similar in length. Multiple methods were used to attempt to isolate the bioAB sequence from the pUC-GW-Amp, such as varying the agarose gel concentration and selecting a third restriction enzyme to reduce the size of the plasmid the bioAB gene shipped in. These methods appeared to fail, so another trial using new plasmid DNA and new media was performed This trial successfully resulted in two correctly ligated colonies that will be used in the future. In future research, the produced, isolated, and solubilized BioAB protein will be tested using Mossbauer and Electron Paramagnetic Resonance spectroscopy. The protein will also be used to check for BioA and BioB protein activity.

Introduction

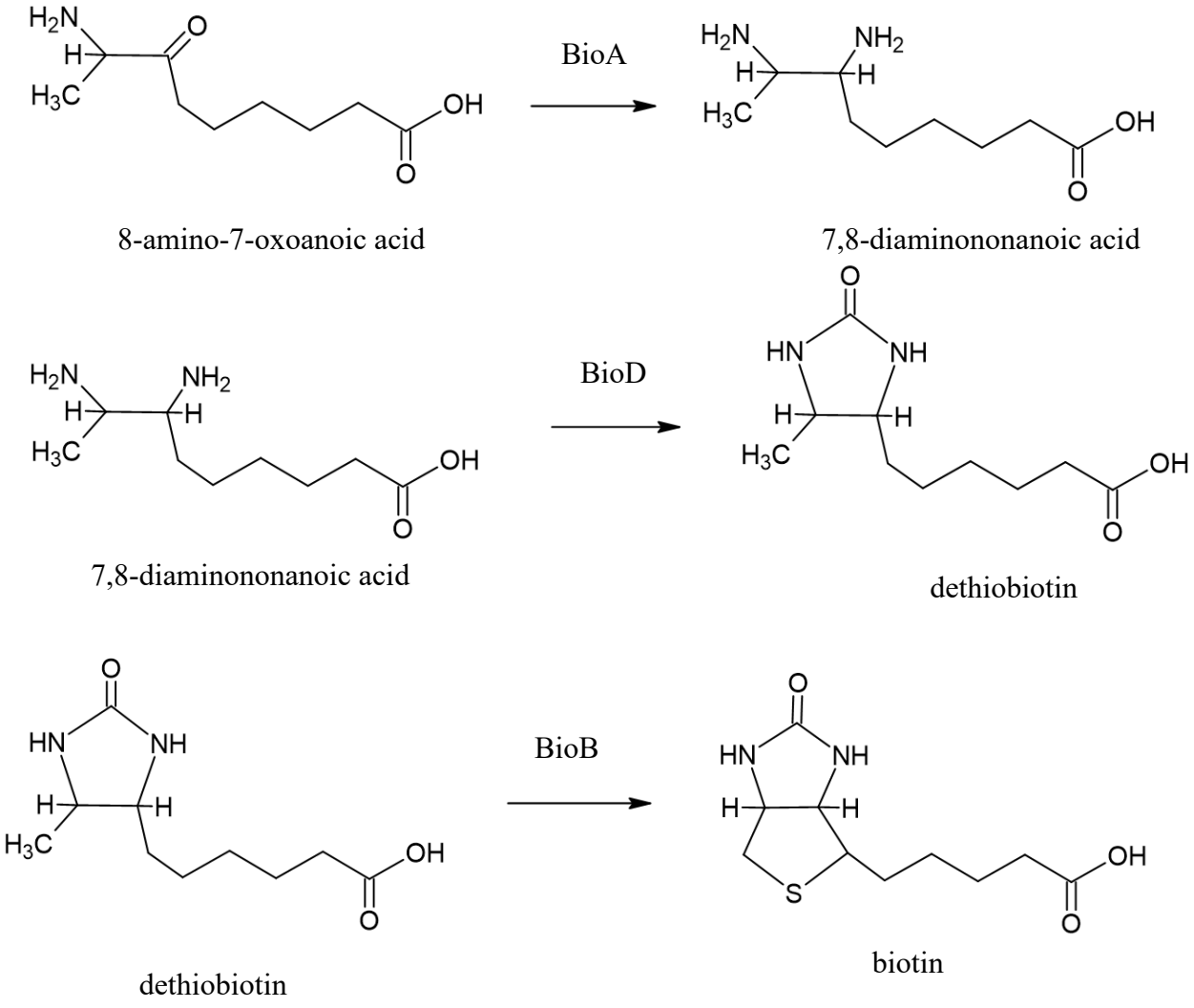
Biotin is an important molecule in enzymes involved in the citric acid cycle as well as for the synthesis of fatty acids. Both of these reactions rely on the transfer of a carbon dioxide to another molecule. In the case of the citric acid cycle, a carbon dioxide molecule is transferred to pyruvate converting it to oxaloacetate; the oxaloacetate can then enter the citric acid cycle.[1] Humans do not have the enzymes to catalyze the reactions to synthesize biotin, so we rely on bacteria in our gastrointestinal tract to produce it. Biotin synthase is a part of the radical S-adenosylmethionine (SAM) family of enzymes. The enzyme is important for the production of biotin, because it catalyzes the last step in the synthesis of biotin. The BioB enzyme contains two iron sulfur clusters which are able to catalyze the reaction. The bioB gene is one of the genes that encodes biotin synthase in some organisms. Examples of the genes that encode enzymes in the pathway include, bioA, bioD, bioF, and bioW. [2] Each of these genes encodes an enzyme that catalyzes a step in the biosynthesis of biotin (Scheme 1).[2] The biotin synthase enzyme functions by making use of its radical SAM cofactors and converting dethiobiotin into biotin. In this process, one of the iron sulfur clusters is sacrificed, as a sulfur atom is donated to the dethiobiotin to produce biotin (Scheme 3).[3]



pimelic acid pimeloyl-CoA



pimeloyl-CoA 8-amino-7-oxoanoic acid

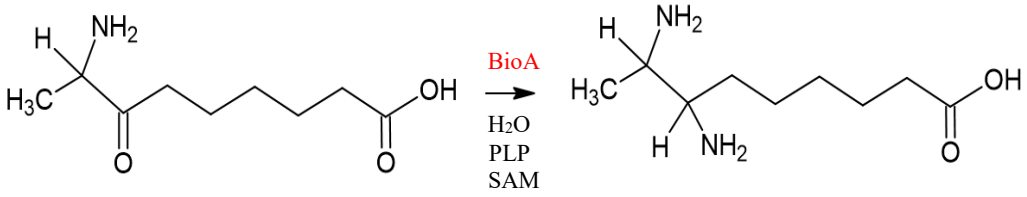


Scheme 1: Entire Biotin Synthesis Pathway[2]

The bioAB gene, the gene of interest in this research, is an understudied gene involved in the production of biotin. The gene of interest is from the bacterial species *Bacteroides thetaiotaomicron. B. thetaiotaomicron* is a gram-negative, anaerobic bacteria commonly found in the human gastrointestinal tract.[4] The confusion around the bioAB gene comes from the lack of knowledge surrounding the correct start codon and annotation of the gene. In the current databases, there are two start codons.[5] Currently, it is not known which is the correct start

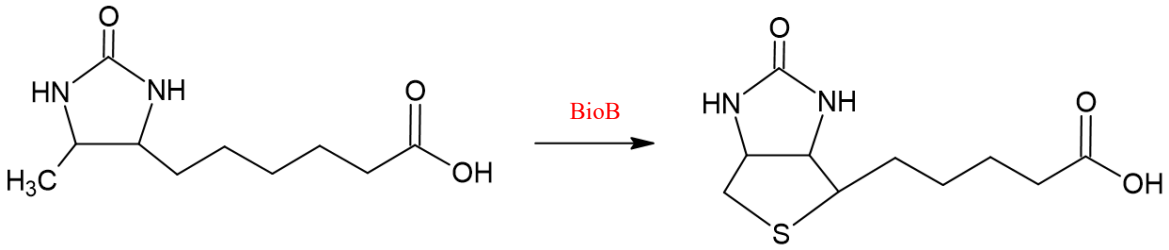
sequence. There is also relatively little known about the active sites of BioAB; it is possible the enzyme has two active sites that function in the production of biotin, just as BioA and BioB would separately. [5]

Although the bioAB gene encodes for what is believed to be one protein, the sequence aligns with two genes from the biosynthetic pathway: bioA and bioB. Both genes are present in other species, however, bioAB is found in *B. thetaiotaomicron*. As stated before, BioB is an enzyme that functions to convert dethiobiotin into biotin by donating a sulfur from the iron sulfur cluster. The mechanism of the reaction functions by first adding the sulfur to the C9 position of the dethiobiotin and later to the C6 position, which completes the ring (Scheme 3). [6] BioA on the other hand functions to convert 8-amino-7-oxonanoic acid to 7,8-diaminononanoic acid (Scheme 2).[2] The reactions are shown below.



8-amino-7-oxoanoic acid 7,8-diaminononanoic acid

Scheme 2: Reaction catalyzed by BioA enzyme [2]



dethiobiotin biotin

Scheme 3: Reaction catalyzed by biotin synthase (BioB protein) [2]

The goal of the research is to create more opportunities into the study of the BioAB protein. However, since the gene is annotated as a combination of two genes, it remains unknown how the enzyme functions in catalyzing the reactions. Most likely the enzyme has two active sites that each catalyze individual reactions. This work aims to express the protein in a manner so it is soluble and can be purified. Then, additional work can be conducted to extend our understanding of the protein. Using Mossbauer and electron paramagnetic resonance spectroscopy the iron sulfur clusters can be characterized and BioA and BioB activity examined.

Materials and Methods

Plasmid Prep and Purification

The bioAB gene, cloned into the EcoRV site of the pUC-GW-Amp vector, was ordered from Genewiz. A transformation of TOP10 *E. coli* cells was conducted with the pUC-GW-Amp and bioAB plasmid, and the cells were plated on an ampicillin containing agar plate and grown overnight at 37ºC. A single colony was selected and used to inoculate 5 mL of lysogeny broth containing 10 µL of ampicillin. The culture was incubated overnight in the shaker at 37ºC. Following the instructions as directed, the plasmid was extracted from the *E. coli* using the Zyppy Plasmid Miniprep kit.[7] One minor change was enacted: after the wash with ethanol, the column was spun a second time at 13,000 rpm for 1 minute to remove excess ethanol from the column.

Restriction Digest

A restriction digest of the plasmid to isolate the bioAB gene was performed using *Xho*I and *Nde*I. It consisted of 1 µL Cutsmart Buffer, 0.3 µL of *Xho*I, 0.3 µL of *Nde*I, 4.4µL of water,

and 4 µL of DNA. The enzymes were purchased from New England Biolabs. Simultaneously, the same restriction enzymes and process were used on a pET28a plasmid for the bioAB gene to later be ligated into. Both of these digests were incubated overnight at 37º C. The digest was conducted in the same manner as the digest in the literature.[8] The digest was separated using a 1% agarose gel conducted via gel electrophoresis for 60 minutes at 120 volts to isolate the desired bioAB sequence, as well as, to purify the pET28a sequence. A restriction digest using the restriction enzyme *AlwN*I was ordered to cleave the pUC-GW-Amp sequence in half so the bioAB sequence could be isolated. The restriction digest with *AlwN*I was completed using the same procedure. The resulting product was used to conduct gel electrophoresis using a 1% agarose gel. The bands were selected and purified out of the agarose gel using a MinElute Gel extraction kit, followed according to the manufacturer's instructions.[9]

Ligation and Transformation

The pET28a and bioAB were then used to conduct a ligation with 1 µL of buffer, 1 µL of ligase, 6.5 µL of bioAB, and 1.5 µL of pET28a. The ligation was kept at 4º C for three days for the ligation to occur. A transformation of TOP10 *E. coli* was then completed according to the procedure. The transformed bacteria were plated on two kanamycin agar plates with 500 microliters of cells added to each plate. The plates were incubated overnight at 37ºC. The resulting colony was used to inoculate 5 mL of luria broth containing 10 µL of kanamycin to select for colonies with the pET28a plasmid. The inoculated media was incubated overnight at 37ºC. The resulting cells were lysed, and the DNA purified from the cells in the same manner as previously explained using the Zyppy plasmid miniprep kit. The purified DNA was then suspended in 30 microliters of 18 MΩ nanopure water. A restriction digest was conducted in a

manner different from that of the previous procedure. Using 1 microliter of cutsmart buffer, 2 microliters of the purified DNA, 1 microliter of *Nde*I, 1 microliter of *Xho*I, and 5 microliters of H_2O, the restriction digest was set up and frozen until use. Prior to being used, the digest was thawed and placed in the 37ºC incubator for two hours. The resulting digest was separated using a 1% agarose gel to determine if the gene was ligated correctly.

Determining DNA Concentration

To determine the DNA concentrations the Qubit 2.0 Fluorometer was used. The instructions were followed exactly to produce the two standards and for the DNA samples. Concentration calculations were also completed to determine the total DNA concentrations.

Colony PCR

The colony PCR was set up in accordance with the protocol. For each reaction it was required that there was 3 µL of 10x reaction buffer, 1.5 µL of forward primer, 1.5 µL of reverse primer, 1 µL of dNTPs, 0.3 µL of Taq polymerase, and 17.7 µL of water, and 5 µL of cell suspension to achieve a desired volume of 30 µL. To make the cell suspensions, a colony was suspended in 10 µL of sterile water. The PCR program consisted of a cycle of 95º C for 1 minute followed by 30 cycles of 1 minute of 95º C, 1 minute of 44º C, and then 2.5 minutes of 72º C. The PCR was finished by 10 minutes of 72º C and then an infinite cool of 4º C until the tubes were removed from the thermocycler for storage. The finished PCR were stored at 4º C until they were verified via gel electrophoresis.

Sequencing

To be sent for sequencing, a dilution of the purified plasmids was made. The plasmids were also sent with primers T7 primers. The prepared samples were sent to Genewiz for sequencing.

Results and Discussion

The bioAB gene in the pUC-GW-Amp plasmid offered many complications that hindered progress of the research. The plasmid was digested using two restriction enzymes, *Nde*I and *Xho*I. The pET28a plasmid was also digested using the same restriction enzymes for later ligation of the bioAB gene. After the digest was completed however, it was determined, via gel electrophoresis and analyzation of the annotated sequences, that both the bioAB and the pUC-GW-Amp sequences were about 2400 base pairs in length. Since the two sequences were about the same size, there was only one DNA band in the gel (Figure 1). The combined sequences made it impossible to purify the bioAB sequence away from the pUC-GW-Amp sequence, so other options were considered.

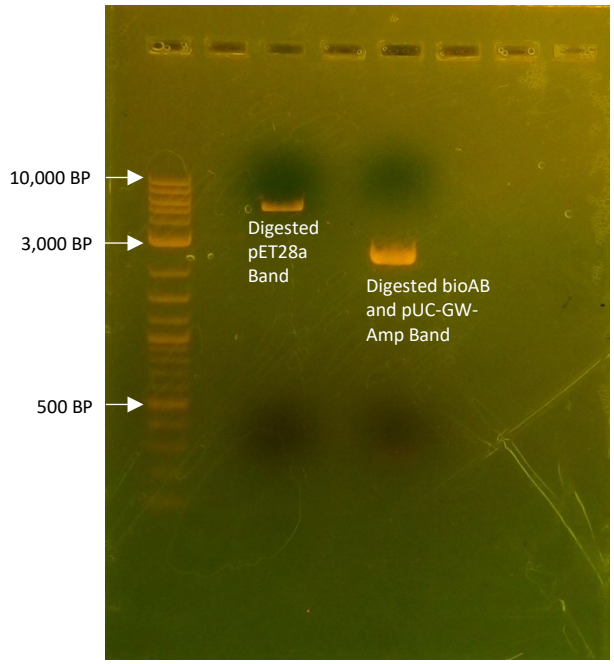


Figure 1. The result of the gel for the digested pET28a plasmid and the digested bioAB and pUC-GW-Amp plasmid.

The first thought was to use a gel with a lower agarose concentration (0.7%). It was believed that a decreased concentration would allow the smaller fragment to move through the gel at an increased rate providing more separation. This method did not succeed and produced a similar result as the original 1% agarose gel. Following this, a different restriction enzyme was used in an attempt to cleave the pUC-GW-Amp sequence in half, thus creating a smaller fragment that would allow for the separation and isolation of the bioAB gene. The first restriction enzyme selected for the previously mentioned restriction digest was *Sma*I. After the restriction digest was performed and the resulting product separated in a gel, it was determined that this method possibly failed because of the lack of a correct restriction site (found after searching the annotated sequence using the Control + F search feature as well as reviewing the sequence manually instead of imputing the sequence for computer analysis). The restriction

enzyme was also from the 2000s so it is possible the enzyme was no longer functional. The results of the restriction digest with *Sma*I also yielded a single, inseparable band (Figure 2). It is still unknown why a restriction site that was not actually present was suggested by NEBCutter 2.0.

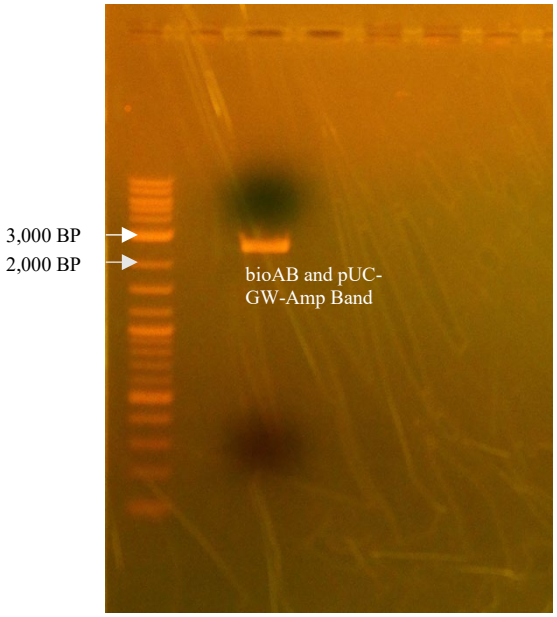


Figure 2. The result of the gel after a second restriction digest was performed on the bioAB gene

Following this attempt, a new restriction enzyme, *AlwN*I, was selected, and a restriction digest performed. Following this digest, multiple bands were present in the agarose gel, suggesting that the bioAB could be isolated from the plasmid this time (Figure 3).

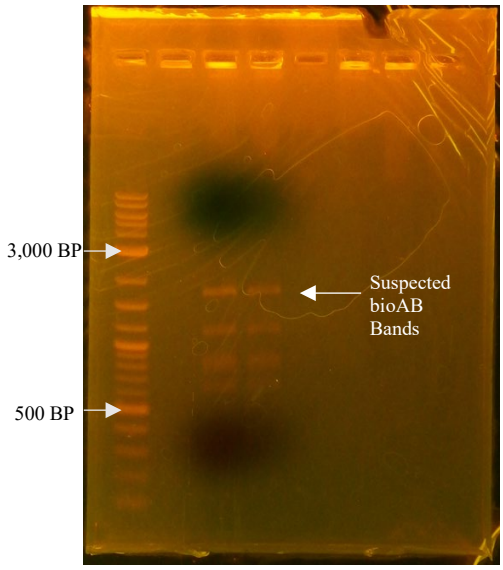


Figure 3. The result of the *AlwN*I restriction digest. Notice the multiple bands present.

The band was removed from the gel, and a ligation performed to ligate the bioAB gene into the pET28a plasmid. Using the results from the ligation a transformation of TOP10 *E. coli* was conducted, and the bacteria plated on kanamycin agar plates. Only one colony grew on the plate, suggesting only one *E. coli* was transformed by the plasmid (Figure 4.)

Figure 4. The single *E. coli* colony grown after the transformation with the ligated pET28a and bioAB plasmid

The colony was used to inoculate luria broth and was grown overnight. The resulting bacteria were lysed, and the DNA purified away from the cell debris. A restriction digest was performed with *Nde*I and *Xho*I to determine if the pET28a and the bioAB ligated correctly (Figure 5).

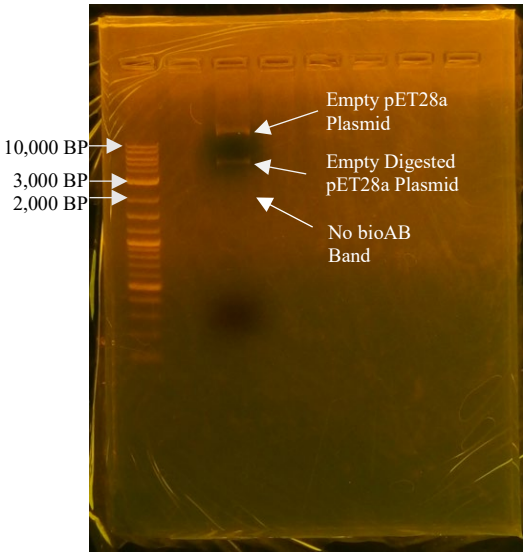


Figure 5. The resulting gel from the digest using the DNA purified from the colony.

The gel shows that the colony did not contain the correctly ligated pET28a and bioAB plasmid and instead most likely contained an empty pET28a plasmid. This was determined by no band being present at the 2400 base pair location, indicating that the bioAB gene was not present.

After this failure, it was determined that we would try other methods for isolating the bioAB gene from the pUC-GW-Amp plasmid. One of these thoughts was to use M13 primers and polymerase chain reaction to create copies of just the bioAB sequence, however this method also proved unsuccessful. It was then determined that we would start over with new plasmid DNA, new media, new miniprep kit, and new competent cells because of the possibility of contamination or the possibility that the enzymes were expired.

For the new trials, one vial of Top10 *E. coli* was transformed using the with the pUC-GW-Amp bioAB plasmid from Genewiz and another with pET28a plasmid. The pUC-GW-Amp *E. coli* were plated on an Ampicillin LB plate and the pET28a *E. coli* on a Kanamycin LB plate. The colonies that grew were then used to inoculate two separate test tubes of LB (containing their respective antibiotic). These steps were taken to produce new samples of both pET28a and pUC-GW-Amp bioAB plasmids that we knew had not been contaminated and were freshly prepared.

As described in the materials and methods, the DNA from these bacterial colonies were purified out of the cells using the ZymoPURE II Midi plasmid prep kit. 200µL of each plasmid were produced and their concentrations determined via the Qubit. The concentrations of the DNA samples were important for the restriction digests, as about 1 µg of DNA should be present in the digest. This step was not taken before, however, it was decided that to promote the likelihood of the bacterial cloning success it was important to take this extra step.

The two plasmids were then digested using *Nde*I and *Xho*I as previously described. The enzymes used this time, however, were new enzymes to prevent another failure. The two resulting digests were then loaded into a 1% agarose gel that was conducted for one hour.

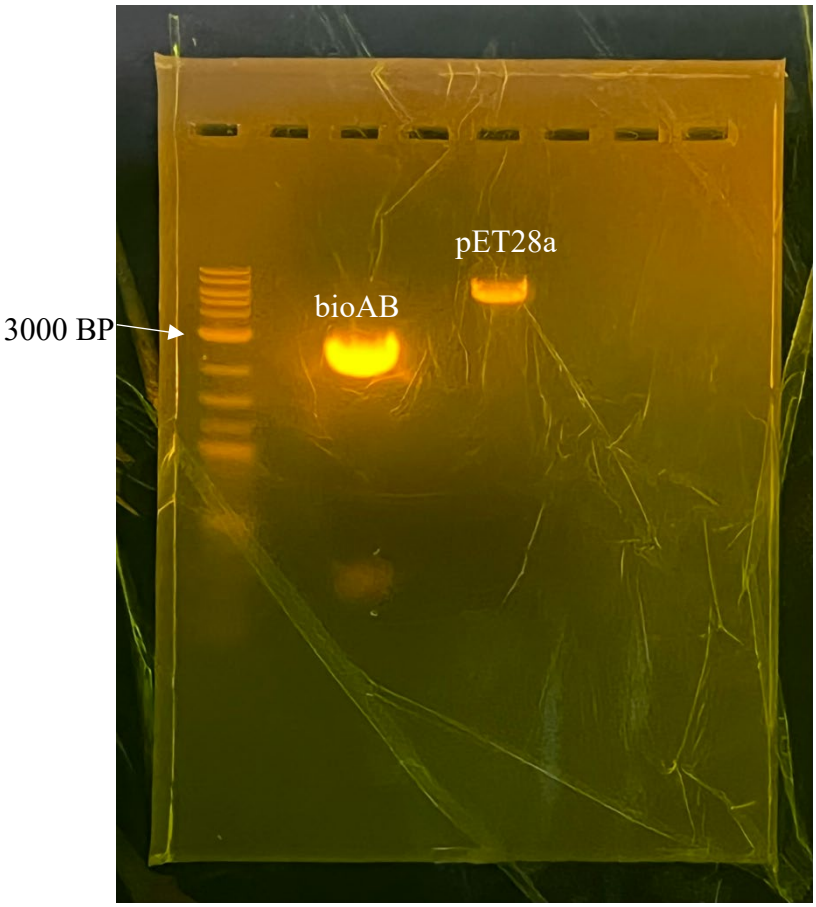


Figure 6. The resulting gel after new plasmid DNA was made and a restriction digest performed.

Note that there was once again no separation between the bioAB band and the pUC-GW-Amp band, just as before, but this time it was decided we would proceed with the two combined (Figure 6). The bioAB band and the pET28a band were removed from the gel and purified using the MinElute Gel Extraction Kit as described by the procedure. The resulting purified DNA was then used in a ligation, also as described. Then, a transformation of TOP10 *E. coli* with the ligated DNA performed. The transformation yielded 14 colonies which was significantly more colonies than before.

These 14 colonies were then used for 14 separate colony PCR reactions. The colony PCR method was chosen to see if the colonies contained the bioAB gene because it allowed for a

quick amplification of the gene as it did not require the lengthy process of culturing the bacteria and then performing a restriction digest to isolate the gene. The colony PCR were then loaded onto a 1% agarose gel and electrophoresed for 20 minutes.

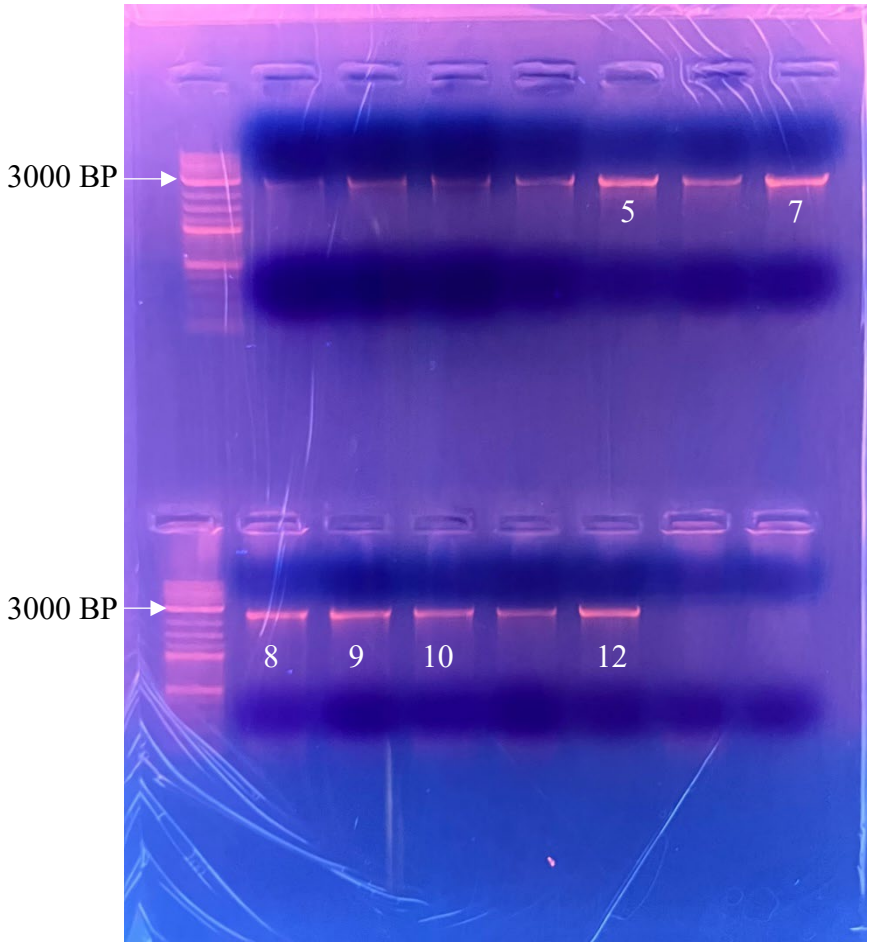


Figure 7. The results of the colony PCR.

Almost all the colonies appeared to contain the bioAB sequence, except those in the last two lanes (Figure 7). The six brightest lanes were selected though because this suggests that they contained the most amount of DNA. Thus, colonies 5, 7, 8, 9, 10, and 12 were chosen to move forward with.

An additional step of plating them on an ampicillin LB plate was taken with these six colonies because there was the possibility that the pUC-GW-Amp sequence remained. The

colonies that grew on the plate could be eliminated, as they were incorrectly ligated, containing both pET28a and pUC-GW-Amp.

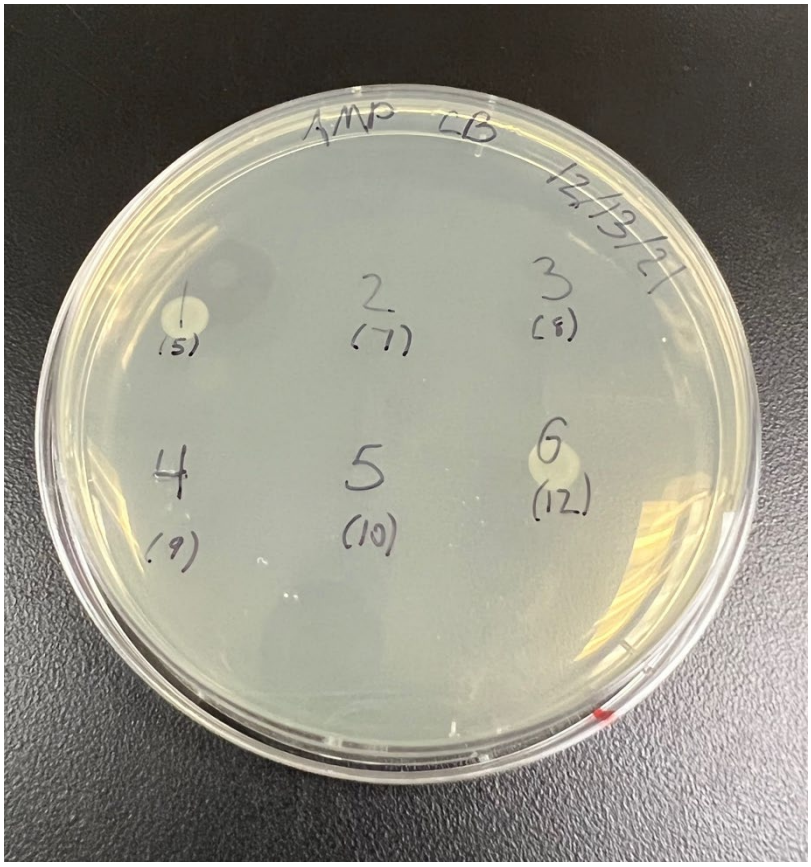


Figure 8. The resulting growth from plating colonies 5, 7, 8, 9, 10, and 12, on an ampicillin LB plate.

Colonies 5 and 12 grew on the plate meaning they contained some form of the pUC-GW-Amp sequence, as the ampicillin resistance gene was present (Figure 8). The colonies 7, 8, 9, and 10 were then used to inoculate 50 mL LB containing 50 µL of Kanamycin so that a MiniPrep of the colonies could be performed and the DNA isolated.

The resulting purified DNA were then used for another restriction digest, once again using *Nde*I and *Xho*I. The digest was performed this time to see if the pET28a and bioAB bands were located at the correct positions. Previously, the gel data showed us that the bioAB sequence was present and plating the colonies on the ampicillin plate showed that the pUC-GW-Amp

sequence was not present, however it was not yet known if the pET28a sequence was present. The restriction digest results were loaded onto a 1% agarose gel, which was then conducted for 20 minutes.

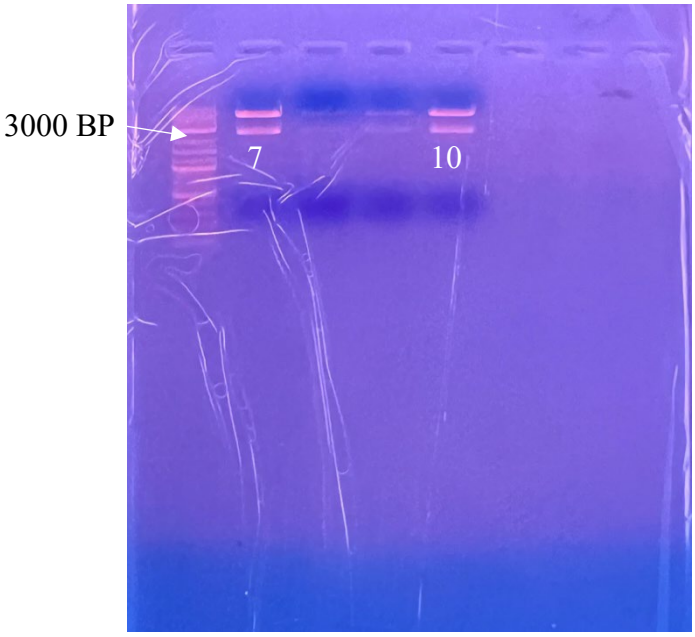


Figure 9. Gel results of the restriction digest with *Nde*I and *Xho*I on colonies 7, 8, 9, and 10.

All the lanes/colonies contained the correct bands because the top bands were roughly at the 5300 base pair location (pET28a), and the lower bands were at the 2400 base pair location (bioAB). However, rows three and four are not as bright, thus they contained less DNA, and thus they were eliminated. Plasmid preps seven and ten were selected to be sequenced because of their brightness.

These two colonies were then used to inoculate 50 mL of LB containing 50 µL of Kanamycin. These colonies were then centrifuged so that a MidiPrep could be performed to produce enough of each colony's plasmids for sequencing. The plasmids were sent to Genewiz for sequencing with T7 and T7 terminator primers. The results came back showing that the start

of the bioAB sequence was present in both colony 7 and 10, suggesting that the colonies contain the correctly ligated plasmids (Figure 10 and 11).

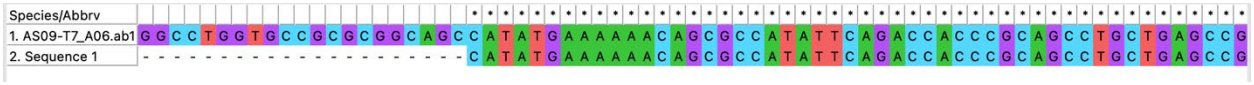


Figure 10. The sequence results for the beginning of the plasmid DNA of colony 7. The top sequence is the sequence received from Genewiz and the bottom sequence is the known bioAB sequence.

Figure 11. The sequence results for the beginning of the plasmid DNA of colony 10. The top sequence is the sequence received from Genewiz and the bottom sequence is the known bioAB sequence.

In a further analysis, new primers will need to be used to verify correct ligation of the entire sequence. The sequence results are verifiable up to about 1000 base pairs, due to limitations caused by primers and sequencing, so it is necessary that the entire sequence is correct for future translation of the protein.

In the future, the correctly transformed *E. coli* will be grown with Isopropyl β-D-1-thiogalactopyranoside (IPTG) to induce transcription of the bioAB gene. The protein produced will then be purified and solubilized. Following this, protein studies will be conducted. Examples of what will be examined are Mossbauer and Electron Paramagnetic spectroscopy to characterize the iron sulfur clusters, as well as to check for BioA and BioB activity. Protein studies will allow for conclusions about the protein to be made, thus allowing for a better understanding of this protein and other similar proteins.

Acknowledgements

The authors thank Dr. Jeffrey Newman and Dr. David Andrew for their suggestions on isolating bioAB from the pUC-GW-Amp plasmid. This work was supported by the Joanne and Arthur Haberberger Fellowship, awarded to the recipient at Lycoming College, Williamsport, Pennsylvania.

References

[1] Nelson, D. L.; Cox, M. M. *Principles of Biochemistry,* 7th ed.; W. H. Freeman and Company, 2017.

[2] Tao, L; Stich, T. A.; Fugate, C. J.; Jarrett, J. T.; Britt, R. D. EPR-Derived Structure of a Paramagnetic Intermediate Generated by Biotin Synthase BioB. *J. Am. Chem. Soc.* 2018, *140*, 12947-12963.

[3] Lotierzo, M.; Raux, E.; Sum Bui, B. T.; Goasdue, N.; Libot, F.; Florentin, D.; Warren, M. J.; Marquet, A. Biotin Synthase mechanism: Mutagenesis of the YNHNLD Conserved Motif, *Biochemistry* 2006, *45*, 12274-12281.

[4] Porter, N. T.; Luis, A. S.; Martens, E. C. Bacteroides thetaiotaomicron. *Trends Microbiol* 2018, *26,* 966-967.

[5] Saunders, A. Lycoming College, Williamsport, PA. Personal Communication, 2021

[6] Fugate, C. J.; Jarrett, J. T.; Biotin synthase: Insights into radical- mediated carbon-sulfur bond formation *Biochim. Biophys. Acta* 2012, *1824*, 1213-1222.

[7] *Zyppy Plasmid Miniprep Kit;* Zymo Research. https://files.zymoresearch.com/protocols/_d4019_ d4020_d4036_d4037_zyppy_plasmid_miniprep_kit.pdf

[8] Biolabs, N. E. Optimizing Restriction Endonuclease Reactions. https://www.neb.com/tools-and-resources/usage-guidelines/optimizing-restriction-endonuclease-reactions (accessed Mar 23, 2021).

[9] *Quick-Start Protocol MinElute Gel Extraction Kit;* QIAGEN, 2016. https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/dna-clean-up/minelute-gel-extraction-kit/