Identifying Novel Inhibitors of RpFabG in Typhus-inducing *Rickettsia prowazekii*

Presented by: Panatpur, Aparna[1]

In partial fulfillment of the requirements for graduation with the ***Dean's Scholars Honors Degree*** in *Neuroscience*

Acknowledgements

Thank you to Dr. Beckham (RE) and Dr. Pollak (Advisor) for their support and guidance in this project and Dr. Fast (PI) for providing the facilities to conduct this research. Thank you to the CNS FRI TIDES department for giving undergraduates the opportunities to perform research.

ABSTRACT

Epidemic typhus is a rickettsial disease that is contracted via ticks and lice found on the flying squirrel. The disease is caused by *Rickettsia prowazekii*, an intracellular, gram-negative coccobacillus. The biosynthetic pathways of *Rickettsia prowazekii* and its host are intertwined. Therefore, the ideal antipathogenic drugs would not target a protein that is found within the biochemical pathway of the human host. There is a type II fatty acid synthase pathway that is unique to Rickettsia prowazekii, which can be distinguished from the multienzyme type I fatty-acid synthase pathway used in humans. Fab G 3-ketoacyl-(acyl-carrier-protein) reductase (RpFabG) is a protein that is specific to this type II pathway. Hence, this research is focused on finding a small molecule drug that inhibits RpFabG. The coding DNA sequence for the RpFabG protein was previously cloned into a pNIC-Bsa4 plasmid, which was transformed into BL21(DE3) Escherichia coli and Dh5α competent cells. The transformed bacteria were cultured in LB media, and the cells were harvested. The expressed protein was purified via Ni-NTA affinity chromatography, made possible by a His6 tag on the vector. Gel electrophoresis was performed to determine the purity of the obtained protein sample; due to indication of slight contamination, gel filtration fast protein liquid chromatography was run on a concentrated protein sample. Genetic Optimization of Ligand Docking (GOLD) is a molecular docking software package that was used to rank potential inhibitors of RpFabG according to binding strength; ethyl acetoacetate (EAA) and acetoacyl coenzyme A (AAC) were determined to have high binding strengths and thus determined to be strong potential inhibitors. An enzyme assay was run to determine the functionality of the enzyme, with EAA and AAC as substrates. AAC successfully decreased the enzymatic activity of RpFabG, suggesting its potential as a novel drug. DSF assay results were equivocal, indicating that inhibition assays should be run on AAC to further assess its potential as an inhibitor.

SIGNIFICANCE

Proteins play a central role in disease processes. By targeting and inhibiting protein function within a cell, it is possible to mitigate the deleterious outcomes of disease states in humans. The most effective indicator of an inhibitor's potency is its performance in *in vitro* or *in vivo* wet lab assays. However, such assays are expensive and time consuming to perform. Progressing a drug though initial binding assays, animal models and eventual human clinical trials can take 10-15 years and upwards of 800 million dollars. The Virtual Cures stream streamlines the aforementioned process by virtually analyzing the structural and chemical bonding relationships between hundreds of potential inhibitors and the protein target. Inhibitors with low binding scores can be screened out, and inhibitors with high scores can be purchased and then tested in wet lab assays.

1. INTRODUCTION

Epidemic typhus is a rickettsial disease that results in the death of 10-60 percent of all patients who contract it. Typhus is most prevalent in cold, unhygienic geographical regions [3]. Common symptoms of typhus include but are not limited to delirium, dry cough, high fever, aching of the joints and muscles, decreased blood pressure, and rashes. Humans contract typhus via ticks and

lice that are found on the flying squirrel. The disease is caused by *Rickettsia prowazekii*, an intracellular, gram-negative coccobacillus. *Rickettsia prowazekii* is the only member of the genus *Rickettsia* to cause a latent infection, which can manifest years to decades later. This milder recurrence of the disease is known as Brill-Zinsser disease [7]. The relatively high fatality rate of patients who contract typhus makes finding a potential inhibitor of *Rickettsia prowazekii* advantageous.

Rickettsia prowazekii relies heavily on its host to carry out its biosynthetic activities, because it has a limited genome [6]. Therefore, the biosynthetic pathways of *Rickettsia prowazekii* and its host are intertwined. The ideal antipathogenic drugs would not target a protein that is found within the biochemical pathway of the human host. *Rickettsia prowazekii* utilizes a unique biochemical pathway for the biosynthesis of fatty acids; the type II fatty acid synthase pathway. The end products of this pathway are the elongated fatty acids that are used to construct the lipid bilayer. Without this enzyme the cell wall would be compromised, leading to cell death. This biochemical pathway can be distinguished from the human biosynthesis of the lipid bilayer, which relies on a multienzyme type I fatty-acid synthase – a large peptide in humans that serves multiple biochemical purposes. The fatty acids are synthesized within the different domains of the peptide.

3-Ketoacyl-(acyl-carrier-protein) [3-ketoacyl-(ACP)] reductase, more commonly known as RpFabG, is specific to the type II pathway found in *Rickettsia prowazekii*. More specifically, it is the third enzyme that catalyzes the necessary chemical reaction [4]. Because of it's potential for use as a bioterrorist, Fab G 3-ketoacyl-(acyl-carrier-protein) reductase is placed in the second highest pathogen biodefense category as delineated by the National Institute of Allergy and Infectious Diseases. There is no known homolog of RpFabG in humans. A BLAST search revealed the highest homology to have a 93% query cover, with a 38% identity match – a relatively low percentage [1].

Enzyme Name	3-oxoacyl-[acyl-carrier-protein] reductase
Enzyme Number	1.1.1.100
Isoelectric Point (pI)	9.07
Molecular Weight	25.9 kD
Molar Extinction Coefficient	10,680 $M^{-1}cm^{-}$
Substrate	(3R)-3-hydroxyacyl-[acyl-carrier protein]
Cofactor	NADPH

Table 1: Table displaying characteristics of 3-ketoacyl-(acyl-carrier-protein) reductase, more commonly known as FabG, as well as the substrates and cofactors involved in the reaction mechanism associated with FabG [4].

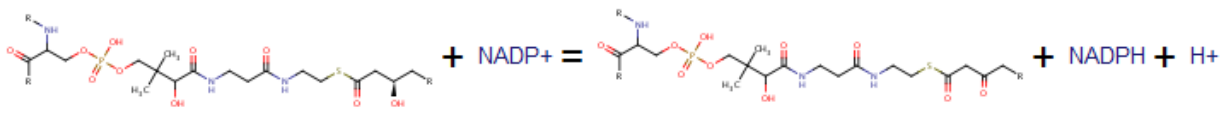

Figure 1: Reaction mechanism involving 3-ketoacyl-(acyl-carrier-protein) reductase. The substrate in this reaction is (3R)-3-hydroxyacyl-[acyl-carrier protein], while the cofactor in this reaction is NADPH [4].

Current available antibiotics for typhus include Doxycycline, Tetracycline, and Chloramphenicol [3]. However, there aren't any known, or commercially available inhibitors specifically for 3-ketoacyl-(acyl-carrier-protein) reductase. There are however known inhibitors for homologs of the protein that differ slightly in structure from RpFabG. Flavonoids have been discovered as an efficient inhibitor in *Plasmodium falciparum* and *Escherichia coli*; there is experimental evidence that the flavonoids inhibit the synthesis of fatty acids in the organism. T butein, isoliquirtigenin, 2,29,49-trihydroxychalcone and fisetin, are four flavonoids that have been determined as inhibitors of β-hydroxyacyl-ACP dehydratase in *Mycobacterium bovis.* However elevated Rv0636 concentrations dampen the effect of these flavonoids as inhibitors [2].

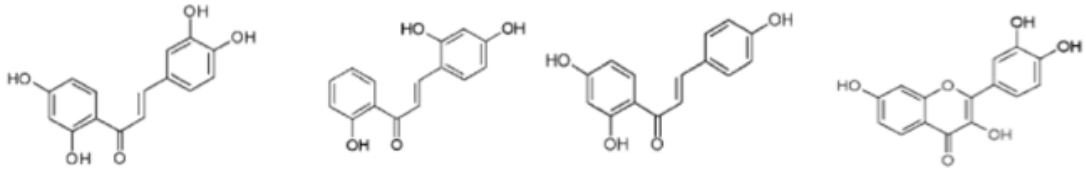

Figure 2: Chemical structure of T butein, isoliquirtigenin, 2,29,49-trihydroxychalcone, and fisetin

2. MATERIALS & METHODS

2.1 Chemical Reagents

All reagents were purchased from Sigma-Aldrich (St. Louis, MO), unless otherwise specified. Black 384-well polypropylene plates (catalog no. 264576) were from Nalge Nunc (Rochester, NY). Microcentrifuge tubes of 1.5 mL and 50mL were purchased from Fisher Scientific (Pittsburgh, PA). A custom HEPES buffer was created to minimize protein precipitation. This custom HEPES buffer solution was prepared with 20 mM HEPES (pH 7.5), 500 mM NaCl, and 5% glycerol. The protein was maintained in this buffer throughout the experimental protocol.

2.2 *Virtual Screening*

Virtual Screening is a computational method that scans libraries of compounds and determines which structures are most likely to bind to the active site of a target; these structures can then be used in *in vitro* assays. There are two main types of virtual screening – structure based, and ligand based. Ligand based screening integrates the pharmacophoric features of multiple ligands to extract common chemical features essential for substrate binding. This information is then used to screen potential compounds. Structure based screening employs docking algorithms and scoring functions to estimate the binding affinity of a substrate to the active site of a target of interest.

Structure based screening was used to screen potential novel inhibitors of 3-ketoacyl-(acyl-carrier-protein) reductase (RpFabG) in *Rickettsia prowazekii.* Specifically, high throughput

molecular docking software (GOLD) was used to screen a commercial library of ligands (Chembridge Diversity Set) against the acetoacetyl-ACP region of the active site (PDB ID 3OP4). Water was removed and hydrogen was added to the virtual structure before docking. A Ti3D Drug Discovery high performance computer (DDFE) consisting of 64 parallel cores was used to run screening jobs. Each compounds was assigned a fitness score indicative of substrate binding strength. The ligands with the best GOLD scores were selected and tested in vitro.

Although a 3-dimensional crystal structure is present for 3-ketoacyl-(acyl-carrier-protein) reductase [PDB 3F9I], the available structure has not been crystallized with a natural substrate. Crystallization of a protein with a natural substrate aids in the identification of the active site. The software used to develop a 3-dimensional model of the protein, PyMOL, uses a naturally bound substrate to identify the initial active site. For this reason, a homologous protein [PDB 3OP4] was used for the creation of the following PyMOL image [4].

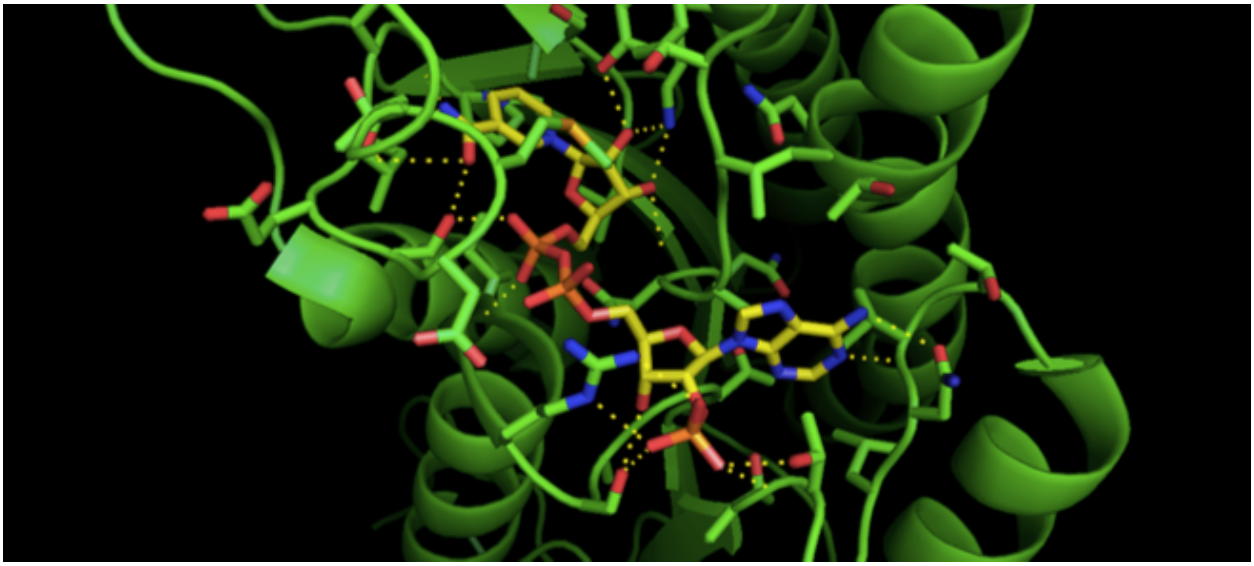

Figure 3: 3-ketoacyl-(acyl-carrier-protein) reductase from *Vibrio cholerae O1* biovar eltor str. N16961 in complex with NADP+; PDB ID (3OP4).

2.3 Gene Construction via PCR

The gene of interest was inserted via ligation independent cloning into the pNIC-Bsa4 vector. Complimentary cohesive ends were generated for the gene of interest and pNIC-Bsa4 vector via T4 DNA polymerase. Following annealing of the vector and gene of interest, the plasmid was transformed into BL21(DE3) *Escherichia coli* as well as Dh5α. BL21(DE3) cells are expression hosts, and therefore possess the cellular machinery required for expression. Dh5α is a cloning host, and allows isolate on of the plasmid without degradation. The plasmid contained a T7 promoter, which is inducible by Isopropyl β-D-1-thiogalactopyranoside (IPTG) and activates gene expression. The vector also contained a His6 tag, which enabled purification via Ni-NTA affinity chromatography. Additionally, the vector contains a kanamyacin resistance gene, which allowed bacterium containing the gene of interest to be isolated.

2.4 RpFabG Protein Expression and Purification

The transformed cells were cultured overnight in Luria broth (LB) media at 37^{o}C. Kanamyacin was added to ensure that only bacteria containing the gene of interest were isolated. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added at an OD600 absorbance of 0.5, which was measured using a Vernier Visible Spectrophotometer. This ensured that the IPTG is added when the bacteria are in log phase, optimizing protein yield. The bacteria were harvested via centrifugation using a JA10 rotor and Beckman Centrifuge. The obtained pellet was resuspended in the custom HEPES buffer and then lysed via sonication to release the protein. The protein was then purified using Ni-NTA affinity chromatography. The His6 tag allows binding of the protein to the Ni-NTA resin. A Wash buffer with a low concentration of imidazole was run through the chromatography column to displace extraneous proteins that didn't contain the His6 tag, and were therefore bound less tightly to the resin. An elution buffer with a high concentration of imidazole was then run through the column. The high concentration of imidazole eluted the protein of interest by competing with the protein for binding sites to the resin. The eluted samples were run through a 0.45μM PES syringe to remove large particulate matter. Protein concentration was determined by UV-VIS spectrophotometry at a wavelength of 250nm. Once the protein was obtained, it was concentrated; half the sample was stored in glycerol stored at 4°C, while the other half was snap frozen using liquid nitrogen and stored at -80°C.

2.5 Spectrophotometric Enzyme Assay

Inhibitor efficacy was assessed *in vitro* via a spectrophotometric enzyme assay. NADPH and the two ligands with the highest GOLD scores – acetoacyl Coenzyme A, and ethylacetoacetate – were used as substrates for the enzymatic assay of RpFabG. The rate of NADPH oxidation was obtained by measuring the decrease in absorbance of NADPH at 30°C and a wavelength of 420nm using a Vernier Visible Spectrophotometer. Assays were performed with 0.2 to 0.6 μg RpFabG in a custom HEPES buffer in a final volume of 80μl. Samples were sealed in a polypropylene tube shielded from light to prevent premature oxidation of NADPH.

3. RESULTS

The coding DNA sequence for the gene of interest, which had already been cloned into pNIC-Bsa4, was transformed into BL21(DE3) *Escherichia coli* as well as Dh5α. The transformed bacteria were then plated on kanamyacin agar plates to isolate the bacteria containing the plasmid. The bacteria were plated at two different concentrations – 10μL and 50 μL – with the latter yielding more bacterial colonies. Protein expression yielded a concentration of 3.465mg/mL, indicating a sufficient concentration to proceed with characterization.

Gel electrophoresis was performed and produced the gel in Figure 4. The first five samples should not produce a visible band at the molecular weight of RpFabG, since there should be little to no protein in these samples. A distinct band should only be visible for samples six through nine, since these are the samples for elution 1 and 2. Elution 1, in lanes 6 and 8, produced a distinct band corresponding to the molecular weight of RpFabG (25.9 kD). There is slight contamination surrounding the band, indicating that additional purification protocols such as

FPLC are necessary. Elution 2, in lanes 7 and 9, also produced a distinct band with significantly less contamination. Purification of the protein produced a protein concentration of 3.285mg/mL in elution 1, and 0.365mg/mL in elution 2. A protein concentration above 1mg/mL is generally sufficient to proceed with enzyme assays. However, evidence of slight protein contamination indicated further purification protocols. Fast protein liquid chromatography (FPLC) was performed on a combined sample of elution 1 and 2. The protein exists as both a monomer and a dimer, as shown in Figure 5. Both samples were collected, and the active dimerized form was used in spectrophotometric enzyme assays.

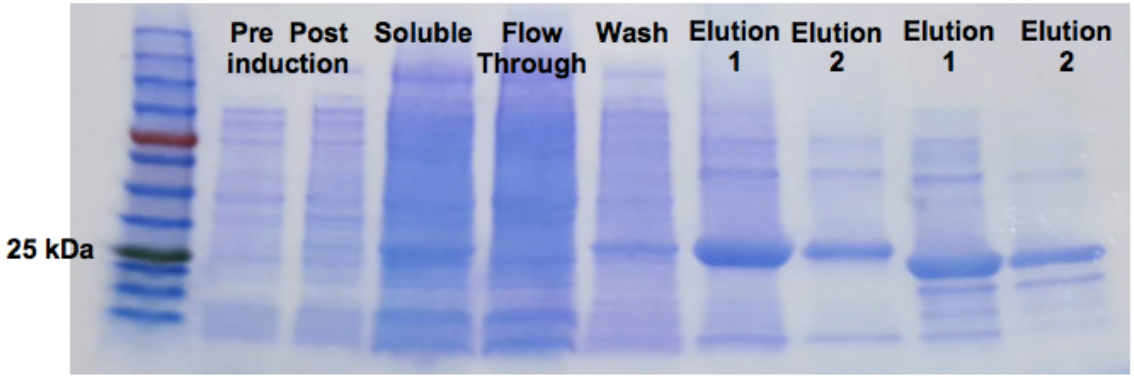

Figure 4: SDS-PAGE gel characterizing RpFabG purified by Ni-NTA. Lane 1 contains the ladder (11-245 kDa). Lane 6 & 8 contains Elution 1 (250 mM Imidazole) and Lane 7 & 9 contains Elution 2 (50mM Imidazole).

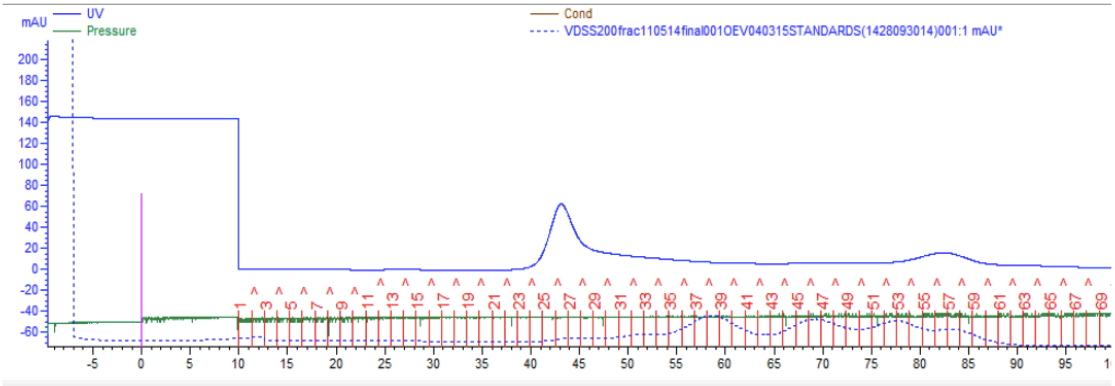

Figure 5: SEC-FPLC graph for RpFabG run on an S200 column. The first peak at 40mL corresponds to the dimerized form of the protein, while the latter corresponds to the monomeric form at 82mL.

The enzymatic activity of RpFabG was measured with acetoacyl coenzyme A (AAC) and ethylacetoacetate (EAA) as the substrate, in the presence of NADPH. A spectrophotometric assay was used to measure the decrease in absorbance of NADPH, and thus the efficacy of the inhibitor. Figure 6, which corresponds to the enzyme assay performed with acetoacyl coenzyme A, shows a decrease in absorbance at 340nm. This decrease in absorbance indicates that NADPH is being oxidized by RpFabG to $NADP^+$ and that the enzyme of interest, RpFabG, is active. Additionally, this decrease in absorbance suggests that the affinity of RpFabG for NADPH is therefore comparable both with and without the presence of AAC. Figure 7, which corresponds to the enzyme assay performed with ethylacetoacetate (EAA), does not show a decrease in absorbance at 340nm. The lack of a decrease in absorbance indicates that NADPH is not being oxidized by RpFabG, and that RpFabG has a reduced affinity for NADPH in the presence of ethylacetoacetate.

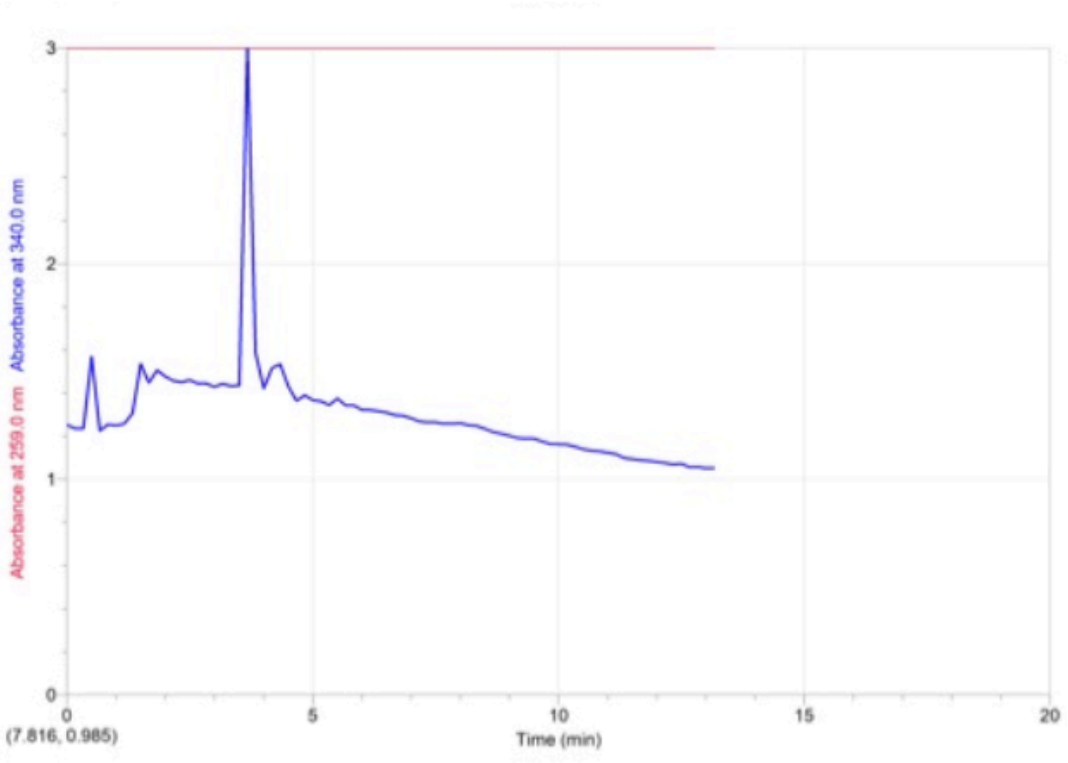

Figure 6: Enzyme assay run with RpFabG, Water, Custom HEPES Buffer, NADPH (6µL), AAC (4.5µL*)*.

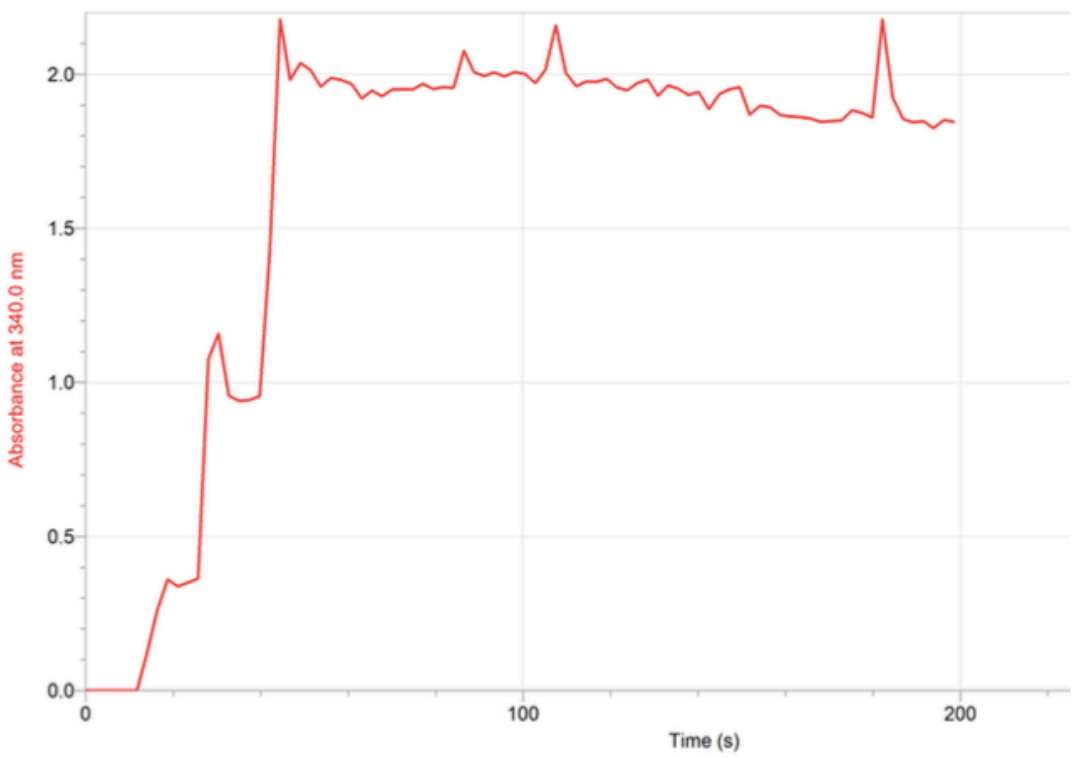

Figure 7: Enzyme assay run with RpFabG, Water, Custom HEPES Buffer, NADPH (6µL), EAA (5µL).

4. DISCUSSION

4.1 Optimizing Protein Yield

The experimental protocol was adjusted several times in order to optimize protein yield. In order to optimize protein expression, the culture had to reach an OD600 absorbance of 0.5 prior to induction. Across multiple rounds of expression, the culture was induced just under an OD500 of 0.5. If induced too quickly, the bacteria may not be in log phase, and may not be reproducing rapidly enough. Therefore in future trials, induction at an OD500 of 0.5 may produce higher protein yield. However, the obtained concentration of expressed protein wasn't a significant impediment to further protocols.

Furthermore, the harvested protein tended to precipitate out of solution throughout the expression and purification process. This precipitated protein was not functional, and could not be used in enzyme assays. Given that the protein of interest was stored in solution for several days between expression and subsequent purification protocols, measures were taken to minimize protein precipitation. It has been reported that a 0.5 mM solution in 0.02 M NaOH(pH 12.3) showed no loss of RpFabG purity in a week at 4 °C or - 85 °C, but a 13% loss at –20 °C. Storage at –20 °C resulted in decomposition of the compound to produce b-NADP and monophosphoadenosine 5-diphosphoribose, even when the protein was protected from light [4].

Therefore, once the protein was concentrated, half the sample was stored in glycerol stored at 4°C, while the other half was snap frozen using liquid nitrogen and stored at -80°C. Additionally, the protein was stored in a custom HEPES buffer prepared with 20 mM HEPES (pH 7.5), 500 mM NaCl, and 5% glycerol to minimize precipitation. Overall, the experimental protocol was successfully modified to yield a high enough protein concentration to proceed with enzyme assays. However, a significant amount of time was devoted to repeated trials in order to develop this modified protocol.

4.2 Future Directions

Preliminary in vitro assays indicate EAA as a potential inhibitor of RpFabG. The next step would be to run a Michaelis Menten assay and calculate kinetic coefficients to determine if EAA is a competitive inhibitor of RpFabG. This kinetic analysis of RpFabG may be performed with various concentrations of NADPH in the presence of constant concentrations of EAA. A similar kinetic analysis can be performed with various concentrations of EAA in the presence of constant concentrations of NADPH.

A Differential Scanning Fluorimetry (DSF) binding assays may be performed on RpFabG to assess protein functionality prior to running enzyme assays. DSF assays may also be used to test inhibitor efficacy, as successful ligand binding would be indicated by the increased thermal stability of a protein. If DSF assays provide clarification as to EAA's potential as an inhibitor of RpFabG, the next step would be to run inhibition assays with RpFabG and EAA. In an inhibition assay, RpFabG will be incubated at room temperature with the determined inhibitor to yield IC50 values, which will provide a measure of inhibitor efficacy.

Overall, this initial work is a good starting point to continue testing potential novel inhibitors of RpFabG. Site directed mutagenesis could be performed in order to pinpoint chemical properties required for a compound to successfully inhibit RpFabG. Furthermore, additional rounds of virtual screening could identify additional compounds for wet lab testing.

REFERENCES

1. BLAST ®. BLAST: Basic Local Alignment Search Tool, http://blast.ncbi.nlm.nih.gov/blast.cgi (accessed Apr 12, 2016).
2. Brown, A. K.; Papaemmanouil, A.; Bhowruth, V.; Bhatt, A.; Dover, L. G.; Besra, G. S. Flavonoid Inhibitors as Novel Antimycobacterial Agents Targeting Rv0636, a Putative Dehydratase Enzyme Involved in Mycobacterium Tuberculosis Fatty Acid Synthase II. *Microbiology*. 2007, *153*, 3314–3322.
3. Cowan, G. Rickettsial Diseases: the Typhus Group of Fevers---a Review. *Postgraduate Medical Journal*. 2000, *76*, 269–272.
4. Hou, J.; Zheng, H.; Chruszcz, M.; Zimmerman, M. D.; Shumilin, I. A.; Osinski, T.; Demas, M.; Grimshaw, S.; Minor, W. Dissecting The Structural Elements for the Activation of β-Ketoacyl-(Acyl Carrier Protein) Reductase from Vibrio Cholerae. Journal of Bacteriology J. Bacteriol. 2015, 198, 463–476.
5. ProtParam. ExPASy, http://ca.expasy.org/tools/protparam.html (accessed Apr 17, 2016).

6. Subramanian, S.; Abendroth, J.; Phan, I. Q. H.; Olsen, C.; Staker, B. L.; Napuli, A.; Voorhis, W. C. V.; Stacy, R.; Myler, P. J. Structure Of 3-Ketoacyl-(Acyl-Carrier-Protein) Reductase from Rickettsia Prowazekii at 2.25 Å Resolution. *Acta Cryst Sect F Acta Crystallogr F Struct Biol Cryst Commun* **2011**, *67*, 1118–1122.
7. Typhus: MedlinePlus Medical Encyclopedia. U.S National Library of Medicine, https://www.nlm.nih.gov/medlineplus/ency/article/001363.htm (accessed Apr 11, 2016).
8. BRENDA. www.brenda-enzymes.org (accessed Apr 17, 2016).