Montclair State University

Montclair State University Digital Commons

Theses, Dissertations and Culminating Projects

5-2022

Role of E168D in M. tuberculosis IGPS Catalysis

Hedda Booter

Follow this and additional works at: https://digitalcommons.montclair.edu/etd

 Part of the Biochemistry Commons, and the Chemistry Commons

Abstract

Mycobacterium tuberculosis is a bacterium that affects the lungs and causes tuberculosis, infecting over 10 million people worldwide each year. Owing to the organism's evolving drug resistance, existing antibiotics no longer combat the disease. The need to discover new drug targets has never been greater. Previous findings suggest that the protein IGPS (indole-3-glycerol phosphate synthase) in *M. tuberculosis* (*Mt*IGPS) may be a target in the treatment of the disease. *Mt*IGPS catalyzes the fourth step in the tryptophan biosynthetic process. The ability to create their own tryptophan is a major advantage that many bacteria have. The role of E168, in the active site of *Mt*IGPS was investigated. Wildtype and E168D *Mt*IGPS proteins were then purified and K_M values were determined to be 6.9 ± 1.4 μM and 26 ± 3 μM, respectively. After this, Wildtype and E168D k_{cat} values were determined to be 0.022 ± 0.002 s^{-1} and 0.011 ± 0.001 s^{-1}, respectively. An enhanced understanding of ligand binding and catalysis in the *Mt*IGPS active site will aid in the design of *Mt*IGPS inhibitors.

MONTCLAIR STATE UNIVERSITY

Role of E168D in *M. tuberculosis* IGPS Catalysis

by

Hedda Booter

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

May 2022

College of Science and Mathematics
Department of Chemistry and Biochemistry

Thesis Committee:

[redacted]
Nina Goodey, PhD.
Thesis Sponsor

[redacted]
David Konas, PhD.
Committee Member

[redacted]
Jaclyn Catalano, PhD.
Committee Member

ROLE OF E168D IN *M. TUBERCULOSIS* IGPS CATALYSIS

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

by

HEDDA BOOTER

Montclair State University

Montclair, NJ

2022

Copyright © 2022 by *Hedda Booter*. All rights reserved.

Acknowledgements

It has been a thrilling experience to be a member of Dr. Nina Goodey's IGPS lab. She has an undying passion and determination to help her students, which I appreciate wholeheartedly. Additionally, I would like to acknowledge my distinguished committee members, Dr. David Konas and Dr. Jaclyn Catalano, for assisting me in the competition of this thesis. They played a huge part in this project, and I would love to recognize them for all their contributions. I would also like to show my appreciation for multiple members of the lab including Dr. Kevin Olsen, Katherine Leon-Hernandez, Sarah Cho, Oshane Thomas, Dr. Cheryl Janson, Cinthya Moran, Tyler Eck, Bhagyashree Vaidya, Maryum Bhatti and Patrycja Marin for their assistance in helping me complete this research. Additionally, I would like to voice my appreciation for my single mother who has sacrificed so much for me and supported me in every way for my education. Furthermore, I would like to acknowledge my grandfather who passed away near the end of my master's degree for sending me heartfelt prayers that I feel every day. Also, a huge thank you to God for providing me with countless opportunities and blessing me with the most supportive family. To conclude, I would like to sincerely express my gratitude to the National Institute of Health for providing funding through grant 1R15GM126467-01: "Investigation of Substrate Specificity, Mechanism, and Inhibition of IGPS."

Table of Contents

Page

List of Figures

List of Tables

List of Equations

1. Introduction

1.1 *An Overview of Mycobacterium tuberculosis*

The disease tuberculosis is caused by a bacterium called *Mycobacterium tuberculosis* (*Mt*). *Mt* targets the lungs of humans but can also target other organs, such as the brain, kidneys, or spine (Zaman, 2010). This disease is said to cause a "global emergency" and kills 5,000 people per day, worldwide. Being a top cause of death, this illness affects children and adults alike (Anochie et al., 2018).

One of the main reasons tuberculosis is so deadly is that multiple bacterial strains have become drug or multiple drug resistant (MDR). These MDR strains can result from patients being given incorrect antibiotics. MDR strains can also be caused by improper administration of medication by patients. These inconsistencies may be caused by supply shortages or because patient symptoms have subsided (Ormerod, 2005). There is always a need to find effective drugs in the "arms race" against the deadly disease tuberculosis (Barberis et al., 2017). This prevalent need to treat tuberculosis has been studied by identifying the genes related to its drug resistance as part of a sequencing study conducted by H. Zhang. The study aimed to signify the degree of drug resistance of *Mycobacterium tuberculosis* and the potential of its virulence over time (Zhang, H. et al., 2013).

Unfortunately, not much progress has been made to develop anti-tuberculosis drugs, which is a huge problem shared by the international community (Esposito et al., 2021). The discovery of the illness by Koch back in 1882 and Schonlein coining the name "tuberculosis" in 1834 paved the way for scientists to explore this emerging disease (Barberis et al., 2017). Studying tuberculosis and its drug targets are necessary for the welfare of the global community.

1.2 *The Importance of MtIGPS and How It Can Inhibit the Disease*

IGPS (indole-3-glycerol phosphate synthase) is a protein discovered in *Mt* that catalyzes the fourth step of tryptophan biosynthesis (Shen et al., 2009). IGPS reacts with CdRP (1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate) to create the product IGP (1-C-(3-indolyl)-glycerol 3-phosphate) while creating the byproducts, H_2O (water) and CO_2 (carbon dioxide) as represented in the figure below:

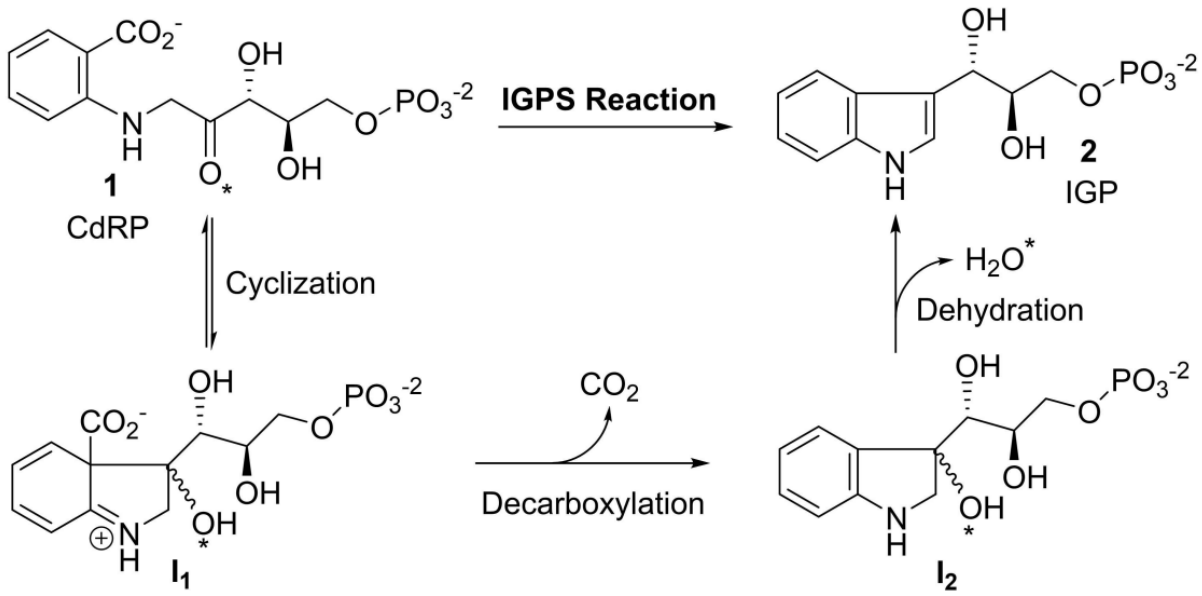

Figure 1. Mechanism of IGPS catalyzing the reaction after binding to CdRP and producing IGP. Figure adopted from Esposito et al., 2021.

The reaction initiated by IGPS is irreversible because it releases CO_2 as well as closes the indole ring in step 2 of the mechanism, as seen in Figure 1 (Parry, 1971). Due to similar comparisons among the substrates and products with those of indole synthesis (Bischler-Möhlau synthesis), these findings assisted Parry to the conclusion that the substrate CdRP and enzyme IGPS produce two intermediates (Figure 1). Step 1 in the mechanism introduces a reversible indole cyclization reaction where the CdRP carbonyl and aromatic ring form intermediate 1 (I_1).

Afterwards, intermediate 2 (I_2) is formed by decarboxylation of I_1. This step is ensued by a dehydration reaction of I_2 creating IGP. IGP is the final product in the fourth phase of the tryptophan biosynthetic process. The idea of the mechanism containing two intermediates was then further confirmed by the crystallization of *S. solfataricus* IGPS (*Ss*IGPS) in complex with CdRP and r-CdRP (non-reactive). The crystallization showed that carbon bonds between the ligands and the active site were too far away at 4.5 Å. This finding suggested that conformational changes (in the form of intermediates) were required in the active site to mobilize and bind the carbons (Hennig et al., 2002). Additionally, another confirmation for these intermediates was done via electrospray mass spectroscopy on *M. tuberculosis.* The purpose of using electrospray mass spectroscopy was to observe how intramolecular cyclization helps CdRP to create IGP. This analysis was done by showing ions and other reaction species of the intermediates themselves (Czekster et al., 2008).

Much excitement fills the scientific community about the protein indole-3-glycerol-phosphate synthase (IGPS) and the potential role it may play in stopping the disease tuberculosis. Sansetti and coworkers suggest that IGPS in *Mycobacterium tuberculosis* (*Mt*IGPS) may be a target in the treatment of *Mt*. They hypothesize this because *Mt*IGPS catalyzes the fourth step of tryptophan biosynthesis. The amino acid tryptophan is an essential amino acid to aid bacteria in survival (Sansetti et al., 2003). A huge advantage for targeting IGPS is that humans cannot produce it. This fact makes IGPS the perfect enzyme to target for anti-tuberculosis drugs (Shen et al., 2009). By targeting the inhibition of *Mt*IGPS, tryptophan can become limited in the bacteria creating an environment where the organism can barely survive and grow. Tryptophan and other amino acids are essential for *Mt* to build proteins and help the organism carry out other vital bacterial functions. (Collins et al., 2020).

Zhang et al. found that *Mt* is a self-sufficient bacterium that can survive successfully through its tryptophan biosynthesis regardless of host T-cell disruption (Zhang, Y. et al., 2013). This tryptophan biosynthesis comes in handy for the bacteria when T-cells in the human body try to fight off an infection caused by *Mt*. T-cells try to deplete amino acids such as tryptophan in bacteria. Since *Mt* can create its own tryptophan, the T-cells of the human body alone cannot prevent the bacteria from growing. The natural ability of *Mt* to create its own tryptophan during biosynthesis is a huge edge in survival that many bacteria share. This natural immunity that *Mt* has to human T-cells affects human beings in a deadly way (Collins et al., 2020). Not only does *M. tuberculosis* have to survive human T-cells but also needs to survive phagocytosis by human neutrophils, macrophages, and dendritic cells. It is additionally vital for *Mt* to have a tryptophan biosynthetic pathway to survive in the conditions of this phagosome (Lott, 2020).

As previously stated, IGPS is known to catalyze the ring closure reaction of CdRP to create IGP irreversibly, as shown in Figure 1. It is important to grasp the concept that the ring closure step affects overall tryptophan biosynthesis. It is from the IGP that is formed after this ring closure that contributes to the creation of the final product, tryptophan. Inhibiting this ring closure step of the tryptophan biosynthetic process can therefore inhibit IGP formation. *Mt* would therefore be unable to survive or grow because of this amino acid deficiency. Scientists such as Yang and coworkers have found that recombinant IGPS showed decreased activity by kanamycin, geomycine, and streptomycin. Since these antibiotics inhibited activity in recombinant IGPS, this led the researchers to believe that IGPS is in fact a novel target for antibiotics

To add, scientists such as Cole and coworkers have targeted amino acids and their biosynthetic processes to find if they could inhibit the disease, tuberculosis. The scientists

previously mentioned touched upon the importance of these biosynthetic processes by testing the leucine, proline and tryptophan biosynthetic pathways. It was shown that when cultured in murine macrophages, these three amino acids did not replicate as rapidly or infect mice as effectively as before. These results suggested that these amino acid biosynthetic processes were important in bacterial growth. To add, a high reduction in virulence was observed in the tryptophan strain, leading Cole et al. to believe that the tryptophan biosynthetic process was most important to *M. tuberculosis* growth (Cole et al., 1998). Therefore, this study is essential in shedding light on the importance of the tryptophan biosynthetic process. Through knowledge of tryptophan biosynthesis, researchers pinpointed inhibition centers to prevent the virulence of the disease overall.

Even though IGPS looks like a promising drug target, very few novel inhibitors have been found to be successful in inhibiting *Mt*IGPS. The only inhibitor scientists have come close to discovering is the molecule ATB107, found by Shen and co-authors in 2009 (Shen et al., 2009). This molecule acts as a poor inhibitor for *Mt*IGPS and was obtained through homology modeling. This modeling was then used to infer the structure of IGPS from a *Mt* strain called H37Rv. After this, Shen et al. used the Maybridge database to screen 60,000 compounds to eventually discover the compound ATB107. Inhibitor ATB107 binds to *Mt*IGPS at 3 mM K_D and has an IC_{50} (50% inhibitory concentration) of 0.41 µM. This difference between K_D and IC_{50} values indicated that more than just *Mt*IGPS inhibition was occurring. Additionally, the K_M value of CdRP was increased by ATB107 while V_{max} remained the same, which lead Shen and coworkers to infer that ATB107 acts as a competitive inhibitor of *Mt*IGPS. Since there are many differences between the structures and potency of ATB107 and CdRP, researchers are still far in their search for a similar and effective inhibitor of *Mt*IGPS. Furthermore, a molecule called r-

CdRP, a reduced substrate analog of CdRP, was also found to be a competitive inhibitor of *Ec*IGPS (*E. coli*). This r-CdRP had a K_i of 0.19 µM and was also assumed to inhibit *Mt*IGPS as well (Eberhard et al., 1995). Scientists would benefit from finding novel inhibitors such as transition state analogs that closely resemble CdRP and that may inhibit *Mt*IGPS better (Thomas, 2019). As more data on the mechanism and structure of *Mt*IGPS unfolds, drug screening will provide the means to finding novel drug leads and targets through IGP and CdRP framework (Esposito et al., 2021).

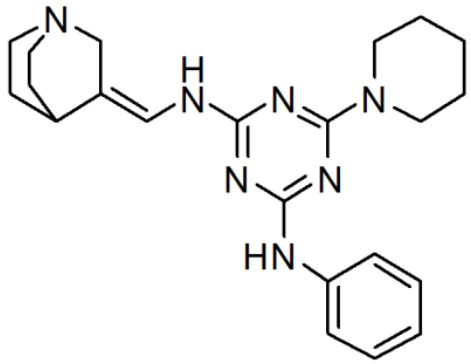

Figure 2. Inhibitor ATB107. Obtained from Shen and coworkers, 2009. A compound containing a substituted triazine and a bicyclic [2.2.2] heterocycle.

1.3 *Structure Determination of MtIGPS*

IGPS is a part of a group of proteins called $(\beta/\alpha)_8$-barrel proteins. These types of proteins are also commonly called triosephosphate isomerase (TIM)-barrel fold protein. In this type of protein, eight α helices are surrounding eight β-strands forming a barrel shape (Henig et al., 1995). TIM-barrel proteins contain vital residues of amino acids situated on its β-strands. These proteins also consist of loops that connect its α-helices and its β-strands together (O'Rourke et al., 2016). This type of protein can act as a scaffold for other types of TIM-barrel proteins. For

instance, *E. Coli* (*Ec*), *Mt*, *B. subtilis* (*Bs*), *P. aeruginosa* (*Pa*), *S. solfataricus* (*Ss*), and *T. maritima* (*Tm*) IGPS are all TIM-barrel proteins that may act as scaffolds for each other as seen in figure 3 (Esposito et al., 2021).

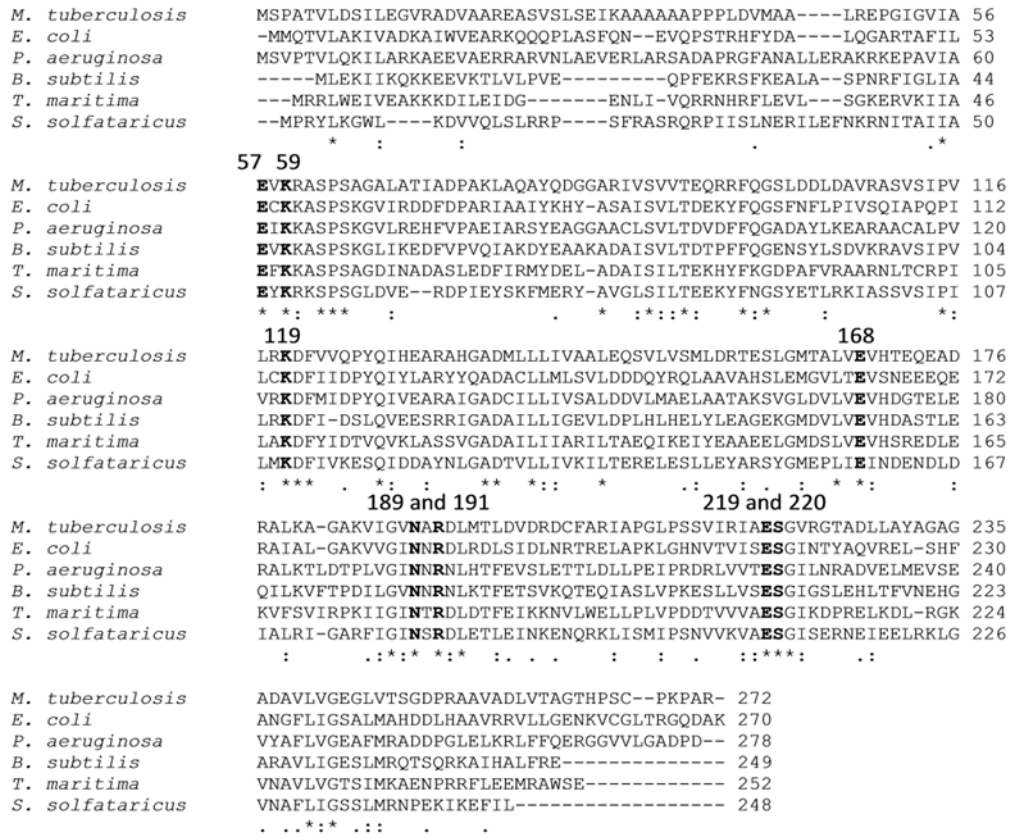

Figure 3. CLUSTAL O of *E. Coli, M. tuberculosis, B. subtilis, P. aeruginosa, S. solfataricus,* and *T. maritima* amino acid sequences. The bolded residues are proposed to be important in CdRP conversion to IGP labeled with *Mt*IGPS labeling. Figure obtained from Esposito et al., 2021.

For more context, the mechanism of reaction in *Ss*IGPS was investigated using X-ray crystallography (Hennig et al., 2002). When compounded with the substrate CdRP and r-CdRP

(the reduced and non-reactive version of CdRP), Hennig and coworkers noticed that two carbons on the ligand and active site of *Mt*IGPS were binding far away from each other (4.5 Å). This length was deemed to be too far for bond formation to occur but also implied that conformational changes were taking place in the binding pocket that would be bring the two carbons together (Hennig et al., 2002). To confirm this idea, modified *Ss*IGPS variants were used. The variants were covalently labeled using fluorescent dyes. This technique was conducted on the N-terminal side of these variants and their active site loops. In conclusion, this showed that the ability of converting and binding CdRP into IGP was afforded by conformational changes (Schlee et al., 2012). Being able to use the *Ss*IGPS scaffold, assisted in the understanding of the *Mt*IGPS structure and its binding (Leisola and Turunen, 2007).

There have been many structures found on the RCSB Protein Data Bank that allowed IGPS to be explored in how it interacts with different ligands. Obtaining clues about the binding of IGPS may lead to newly formulated hypotheses on ligands binding to it (Parry, 1971). In the case of IGPS, the structures of one *B. abortus* IGPS, one *P. aeruginosa* IGPS, two *Tm*IGPS, 20 *Ss*IGPS, and two artificial protein constructs have been determined. One of the artificial constructs was designed by utilizing *Ss*IGPS as the protein scaffold (Giger et al., 2013). In addition, six structures of *Mt*IGPS not bound to any substrate were found and included two apo structures that included PDB 3QJA and PDB 4FB7 (Esposito et al., 2021). To add, structures such as PDB 3T44, contained *Mt*IGPS bound to anthranilate and IGP substrates. Other substrates in complex with *Mt*IGPS found include phenoxymethyl benzoic acid (PDB 3T55), 5-fluoroanthranilate (PDB 3T78) and N-2-carboxyphenyl glycine (PDB 3T40) (Esposito et al., 2021). All the structures of *Mt*IGPS and various ligands stated above have only been deposited in PDB, but the structures have not yet been published in peer review papers. The structures stated

previously, clarified how researchers could target IGPS through drug design and how IGPS bonded to its potential ligands (Esposito et al., 2021). In figure 3, UCSF Chimera's tool align function was applied to evaluate and further investigate the active sites of *Mt*IGPS (PDB 3T44) and *Ss*IGPS (PDB 1A53) and (Esposito et al., 2021).

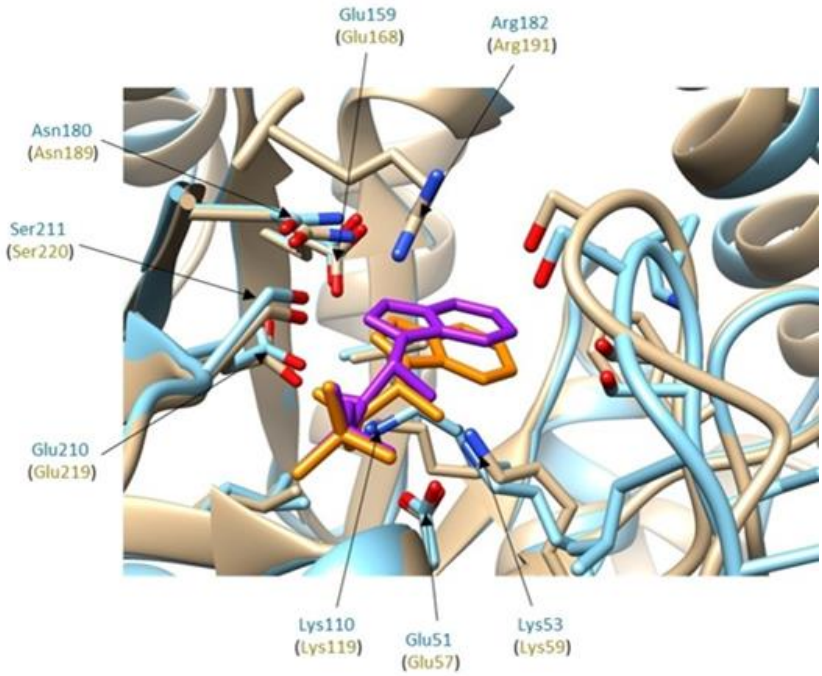

Figure 4. The structure of PDB 3T44, *Mt*IGPS (gold) in complex with IGP (orange) compared to the structure of PDB 1A53 *Ss*IGPS (light blue) and IGP (purple). Correlating active site residues are labeled with blue for PDB 1A53 and gold for PDB 3T44. Figure obtained from Esposito et al., 2021.

1.4 *Hypothesis on the Role of the E168 Residue of MtIGPS*

E168 in *Mt*IGPS was studied by Shen and coworkers when they came across an increased K_M after mutating the glutamic acid to alanine (Shen et al., 2009). The increase in K_M was 19-fold and this led the researchers to hypothesize that the E168 residue may be playing an essential

part in the catalytic mechanism of *Mt*IGPS (Shen et al., 2009). After further analysis, Czekster and coworkers branched off this idea by Shen and researchers by hypothesizing that E168 was the necessary active site base in the mechanism (Czekster et al., 2009). Esposito noted from structure PDB 3T44 that the negative oxygen on the side chain of E168 was located in proximity to the nitrogen of IGP and could form a hydrogen bond that was 2.85 Å long as observed in figure 5.

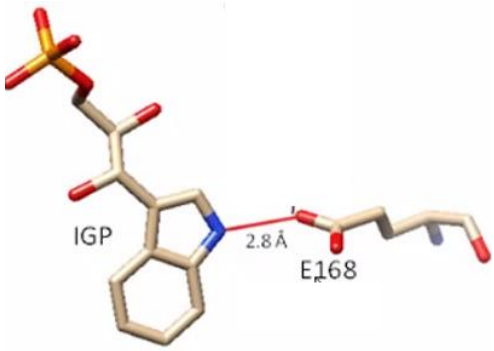

Figure 5. Structure PDB 3T44's negative oxygen of E168 of *Mt*IGPS in proximity (2.8 Å) to the nitrogen on the product, IGP. Figure obtained from Esposito et al., 2021.

To add, in intermediate 1 (I_1) of the mechanism shown in Figure 1, the intermediate contains a positive charge on the nitrogen. Esposito et al. proposed even further that the E168 residue might be the catalytic base in the mechanism just as Czekster et al. hypothesized. The negatively charged oxygen on the E168 side chain might be working to stabilize the transient charge of the positively charged nitrogen on intermediate 1 of the mechanism (refer to figure 1). It also seemed as if E168 was positioned to deprotonate the adjacent carbon in the dehydration step (I_2) of the mechanism in figure 1 and in the end, facilitated the key catalytic ring closure in step 2 to ultimately form the end product, IGP (Esposito et al., 2021).

Just as Esposito et al., I also hypothesized that the E168 side chain anion could provide stabilization for the transient charge on I_1 and assist in the main catalytic ring closure step to finalize the creation of IGP, the product of the reaction (see figure 1). The idea behind replacing Glu with Asp was to see how a change to another acidic side chain would affect binding as well as the rate of conversion of the substrate into IGP. The only difference between Glu and Asp is that Asp is shorter by one carbon. This Asp change may not allow *Mt*IGPS to bind as well to CdRP anymore, hence affecting IGP production. I decided to mutate the residue E168 to E168D because I wanted to see if this mutation would affect the activity, binding, pK_{a1}, pK_{a2}, and rate determining steps in the mechanism of *Mt*IGPS.

I proposed that E168D would end up lowering in activity (k_{cat}), but not dramatically lowering it due to the side chain still being acidic and the only difference being the one shorter carbon chain. I also proposed that the K_M of the mutant would increase and that forming the hydrogen bond between CdRP and E168D would become more difficult due to its shorter carbon chain. In terms of pH profiles and the determination of pK_{a1} and pK_{a2}, I proposed that if E168 is in fact a catalytic base, there would be little to no shift or change in the pH profile. My hypothesis for SVE (solvent viscosity effects) is that at higher glycerol content, there should be a decrease in both wildtype and E168D activities. In a viscous solution, the rate at which the substrate and enzyme would bind becomes slower. Additionally, decreased activity in E168D compared to the wildtype should be observed in a viscous solution due to its shorter carbon chain. My hypothesis for SDKIE (solvent deuterium kinetic isotope effects) would be that since D_2O is a heavier isotope with a lower vibrational frequency than H_2O, it would require more energy to reach the transition state and decrease the overall rate of the reaction (Northrop, 1975).

2. Materials and Methods

2.1 *Expression*

The expression vector, pET30a was purchased from Genscript (NJ, USA) and contained cloned *trp*C coding sequences. To obtain E168D mutated proteins, site directed mutagenesis was done by Oshane Thomas who obtained oligonucleotides (purchased from Integrated DNA Technologies, Inc.) necessary for mutant strand synthesis. A strain of *E. coli* called LOBSTR-BL21(DE3) cells was used as they are resistant to kanamycin. LOBSTR cells were grown in 8 mL of autoclaved LB Broth containing 50 mg/mL kanamycin. These cultures were then incubated in a Thermo-Scientific shaker at 37°C for 16 hours at 225 rpm. After incubation, the 8 mL cell cultures were placed into 1 L flasks containing 500 mL of autoclaved LB nutrient broth 50 µg/mL kanamycin. Cells were grown at 37°C and 225 rpm until an optical density at 600 nm of 0.600 was reached. At this time 1.0 mM IPTG was added to begin protein expression and the temperature was lowered to 25°C at 225 rpm for 17 hours. The LB media was then centrifuged for 15 minutes at 4°C and 5,000 rpm in a JA-10 rotor. Once centrifuged, the supernatant and pellet were quickly separated, and the pellet (~2 grams) was stored at -80°C.

2.2 *Purification*

Approximately 2 grams of the *Mt*IGPS pellet was suspended in equilibrium buffer (1x PBS buffer + 10 mM imidazole) using a vortex machine. The bacterial cell suspension was then sonicated for three minutes with an amplitude of 25% for 1 second on and 2 seconds off at 4°C. The mixture was then centrifuged at 4°C for 15 minutes at 17,000 rpm in a JA-17 rotor. Nickel affinity chromatography was used to bind to 6 histidine tag present in the IGPS protein. After the resin was loaded to the column, 10 mL of equilibrium buffer (1X PBS and 10 mM imidazole, pH

8) was used to prime it. The supernatant was then added. Afterwards wash buffer (1X PBS, pH 8, 40 mM imidazole, 100 mM NaCl) was loaded. A Bradford Coomassie assay test was done to ensure the impurities were eliminated. To elute *Mt*IGPS out of the resin, 20 mL of elution buffer (1X PBS, 250 mM imidazole, pH 8) was added to the column. Fractions of *Mt*IGPS were then collected. Another Bradford assay test was conducted to select the fractions with the highest concentration of *Mt*IGPS. The concentrated fractions of *Mt*IGPS were then dialyzed (molecular weight cut off (MWCU) of 10 kDa (kilodaltons)) in two liters of dialysis buffer (pH 7.9, 5 mM Tris Base, 20 mM NaCl, 1 mM DTT) at 4°C for 2 hours. The bags were then placed into new containers of dialysis buffers at 4°C for another 2 hours to remove any imidazole and exchange the buffer. DTT (2 mM) was then added to the protein to reduce dimerization. The protein was then placed in a centrifugal filter (Amicon® MWCO 10,000) and inside a ST-16R centrifuge and set to 5,000 RPM for 15 minutes to concentrate the protein to approximately 100 µM. Glycerol (10%) was then used to stabilize and preserve the structure of *Mt*IGPS for freezing at -80°C. A nanodrop was then used to measure the absorbance of IGPS at a wavelength of 280 nm (detects aromatic structures in *Mt*IGPS). The extinction coefficient (Ɛ) for the nanodrop was 4595 $M^{-1}cm^{-1}$ and was utilized to calculate the concentration of the protein (Thomas, 2019). Gel electrophoresis (SDS-PAGE) in a Mini Cell Novex gel box was then utilized to observe the molecular weight, purity, and monomer formation of *Mt*IGPS. The electrophoresis gel used was NuPAGE™ 4 to 12% Bis-Tris. Thermo Scientific™ Spectra Multicolor Broad Range Protein Ladder (Catalog # 26623) was used in lane one as shown in figure 6. Lanes 2-6 show treated with 5% β-mercaptoethanol (Bio-Rad Laboratories). The SDS-PAGE was started after loading 90°C heat shocked proteins into the gel for 1 ½ hours at a setting of 107 volts at 25°C. The concentration of wildtype and E168D proteins in each lane was approximately 30 µM.

2.3 *Steady State Kinetics*

To conduct K_M experiments, a Synergy H1 Biotek plate reader set to 25°C was used to measure the absorbance of *Mt*IGPS at 278 nm. Readings were taken in clear Costar 200 µL plates where the pathlength was determined to be 0.46 cm. In each well, either 250 nM of wildtype or 1 µM of E168D was mixed with PIPES assay buffer (2 mM DTT, 100 mM PIPES, pH 7.5) with varying concentrations of substrate (CdRP serial dilutions). Serial dilution concentrations consisted of 463, 278, 167, 100, 60, 36, 22, 13, 8, 5, 3, and 0 µM of CdRP substrate. PIPES assay buffer was used in this and other experiments because of its effective buffering range being 6.1 to 7.5 at 25°C (Good et al., 1966). Triplicates at each CdRP concentration were run on the plate reader for one hour and shown in figure 8. From the data, initial velocity (υ_0) was calculated by dividing the slope of the absorbance versus time graph by the extinction coefficient (Ɛ) 5500 $M^{-1}cm^{-1}$ that was multiplied by the pathlength (0.46 cm) Kirschner et al., 1987). The υ_0 values were then averaged from three trials and the K_M values were generated from the data (substrate concentration versus average υ_0) via input to the program Kaleidagraph® (version 4.1). Standard error was also calculated by taking the standard deviation of all three trials and dividing those values by the square root of 3 (total number of trials). The standard error was used to create the error bars is Kaleidagraph. The data was then fit to the Michaelis-Menten equation to automatically generate the K_M, V_{max} and their respective error values:

$$v = \frac{V_{\max}\,[S]}{K_M + [S]}$$

Equation 1. The Michaelis-Menten equation used in the software Kaleidagraph to obtain K_M and V_{max} values from the data. The value of k_{cat} was then obtained by dividing V_{max} by the total

enzyme concentration as seen in figure 8. The overall rate of substrate molecules converted to product is the k_{cat}.

2.4 *Rate-Limiting Step Determination*

2.4.1 *Solvent Deuterium Kinetic Isotope Effects (SDKIE)*

A Biotek plate reader set to 25°C and 37°C was used to observe solvent deuterium kinetic isotope effects (SDKIE). Readings were measured in absorbance at 278 nm using Greiner clear half-area plates (1 cm pathlength). For wildtype and E168D proteins, 1 µM of protein was mixed with 400 µM of CdRP and PIPES assay buffer (2 mM DTT, 100 mM PIPES, pH 7.5). A separate PIPES assay buffer, made with D_2O, was set to a pH of 7.1 and pD of 7.5 (pD = pH + 0.4) (Zaccardi et al., 2013). From absorbance versus time graphs, slopes were divided by the extinction coefficient (Ɛ) 5500 $M^{-1}cm^{-1}$ multiplied by enzyme concentration (M) and further divided by the pathlength (1 cm) to calculate k_{cat} (Kirschner et al., 1987). The Microsoft Excel generated bar graphs seen in figure 9, show the average of five replicates done on each addition of D_2O (using percentage) over the course of 10 minutes. Error bars were also added to figure 9 by taking the standard deviation of all trials and dividing by the square root of 5 (total number of trials) to obtain the standard error which was input to the graph.

2.4.2 *Solvent Viscosity Effects (SVE)*

A Biotek plate reader set to 25° and 37°C was used to observe solvent viscosity effects (SVE). Readings were taken with absorbance at 278 nm from Greiner half-area, clear plates (1 cm pathlength) containing 1 µM of wildtype or mutant proteins mixed with 400 µM CdRP and PIPES assay buffer (2 mM DTT, 100 mM PIPES, pH 7.5). The PIPES buffer was made by

varying the amount of 100% v/v glycerol. The glycerol was diluted in H_2O to create 60%, 40% and 30%. Respectively, 30%, 20% and 15% glycerol buffers were made afterwards by mixing the dilutions in a 1:1 ratio with PIPES buffer. From absorbance versus time graphs, slopes were divided by the extinction coefficient (Ɛ) 5500 $M^{-1}cm^{-1}$ multiplied by the enzyme concentration (M) then further divided by the pathlength (1 cm) to calculate k_{cat} (Kirschner et al., 1987). A bar graph of k_{cat} versus percentage of glycerol was plotted in Microsoft Excel to show viscosity effects (see figure 10). Five replicates of each percentage of glycerol were run for 15 minutes to obtain the data and standard error to the create the graph with error bars.

SVE was also conducted in sucrose using the same machine settings, methods and calculations seen in glycerol SVE methods discussed in the previous paragraph. To prepare different concentrations of the sucrose PIPES buffers, solid sucrose (342.3 g/mol) was measured (in grams) and added to PIPES buffer (2 mM DTT, 100 mM PIPES, pH 7.5). Five replicates of each sucrose concentration were conducted to obtain the data and standard error to create the graph with error bars (see figure 11).

2.5 *pK_{a1} and pK_{a2} Determination: pH Profiles*

A Biotek plate reader was set to 25°C to measure the fluorescence of pH profile experiments using a 278 nm excitation wavelength and 340 nm emission wavelength. Black Costar fluorescence plates were used to monitor pH profiles (0.46 cm pathlength). Activity becomes minuscule at very low or very high pH values, and for this reason, fluorescence (method used for sensitive experiments) was used for monitoring pH profile experiments (Szmacinski and Lakowicz, 1993). In each well, 250 nM of wildtype or 1 µM of E168D was mixed with 60 µM

CdRP. The pH profile buffers were made with 4X MTEN (100 mM NaCl, 25 mM Tris Base, 1 mM DTT, 25 mM ethanolamine, 50 mM MES) which was diluted in H_2O to create a 1X solution. The 1X MTEN buffer was then separated into different falcon tubes and set to varying pH values as seen in figure 12. MTEN buffer was used in this experiment for its versatile buffering range at the different pH values seen in the figure previously mentioned. Four replicates per pH value, run for 30 minutes each, were used to obtain the data. From fluorescence versus time graphs, slopes (counts per second or CPS per second) were divided by each individual trials' highest value slope to normalize the data. After this, each of the four trials normalized values were averaged and the standard deviation was calculated for each pH value ranging from 5-10 pH. The standard error was finally taken by dividing the standard deviation by the square root of 4 (number of trials taken). Standard error was used to create the error bars as seen in figure 13. The data of average normalized velocity versus pH was then inserted to Kaleidagraph and fit to the Zaccardi pH equation (equation 2) to predict the pK_{a1} and pK_{a2} values (Zaccardi et al., 2013):

$$v = C/\left(1 + 10^{pKa1-pH} + 10^{pH-pKa2}\right)$$

Equation 2. Zaccardi pH equation utilized in Kaleidagraph to fit the data obtained and predict pK_{a1} and pK_{a2} values.

3. Results and Discussion

3.1 *Expression and Purification*

A typical expression of wildtype and E168D resulting from LOBSTR-BL21 (DE3) strains yielded approximately 2 grams of pellet from 500 mL LB nutrient broth. From purification, a yield of approximately 11.7 mg (4 mg/mL or 113 µM) was obtained for wildtype while

approximately 15.5 mg (5.2 mg/ mL or 155 µM) was obtained for E168D from 500 mL LB broth. To determine protein purity, monomeric presence, and molecular weight, SDS-PAGE was conducted on the wildtype and E168D proteins. The wavelength of 280 nm on a nanodrop is used to detect the aromatic structures within the *Mt*IGPS protein, but also other proteins. Hence, it is necessary to conduct SDS-PAGE on all proteins to check for impurities that affect overall k_{cat} of the mechanism. The molecular weight of wildtype and E168D with His-tag was approximately 40 kDa as seen in figure 6, which is higher than the predicted molecular weight determined to be 33.5 kDa (Yang et al., 2006). For *E. coli* β-barrel membrane proteins, folded forms migrate faster down the gel, but a completely denatured ones will migrate slower (Rath et al., 2009). It seems as if my protein has fully denatured and stopped migrating at 40 kDa (compared to 33.5 kDa) and that light blue bands (may be degradation products) seen at 25 and 13 kDa (add up to 38 kDa) may be aggregates possibly caused by heat shocking my proteins for longer than five minutes at 90°C (refer to figure 6). This gel shows predominantly pure *Mt*IGPS even if other impurities (at 25 and 13 kDa) exist. The possible impurities observed at 25 and 13 kDa are light blue (not as concentrated) compared to the deep blue bands see at 40 kDa (more concentrated). The findings of the SDS-PAGE also revealed that wildtype and E168D to be monomers. In figure 6, lanes 2-6 were treated with β-mercaptoethanol versus lanes 7-11 that were not. β-mercaptoethanol breaks disulfide bonds and if two or more bands form instead of one, dimers or trimers might be present (Smith and Yanofsky, 1962). There was no difference between the β-mercaptoethanol treatments and therefore it can be seen that the *Mt*IGPS proteins exist as monomers. It is observed from this data that upon addition of standard 2 mM DTT added to proteins, dimers do not form. Monomers are needed to obtain the most accurate kinetic results. If impurities or dimers are present, then the effective concentration obtained from the nanodrop

280 nm will be skewed to have a higher concentration, resulting in an increased and inaccurate k_{cat} value.

Table 1. Parameters of *Mt*IGPS obtained through online program, ExPASy's ProtParam (Gasteiger et al., 2005). Table obtained by Oshane Thomas (Thomas, 2019). The extinction coefficient (Ɛ) was generated by Oshane Thomas. The wavelength of 280 nm on the nanodrop machine was used to detect absorbance of the protein *Mt*IGPS itself. The predicted molecular weight of *Mt*IGPS matches the established one by Yang and coworkers in 2006.

Number of Amino Acids	322
Molecular Weight	33455.03 Da
Theoretical pI	5.35
Extinction Coefficient	4595 M^{-1} cm^{-1}

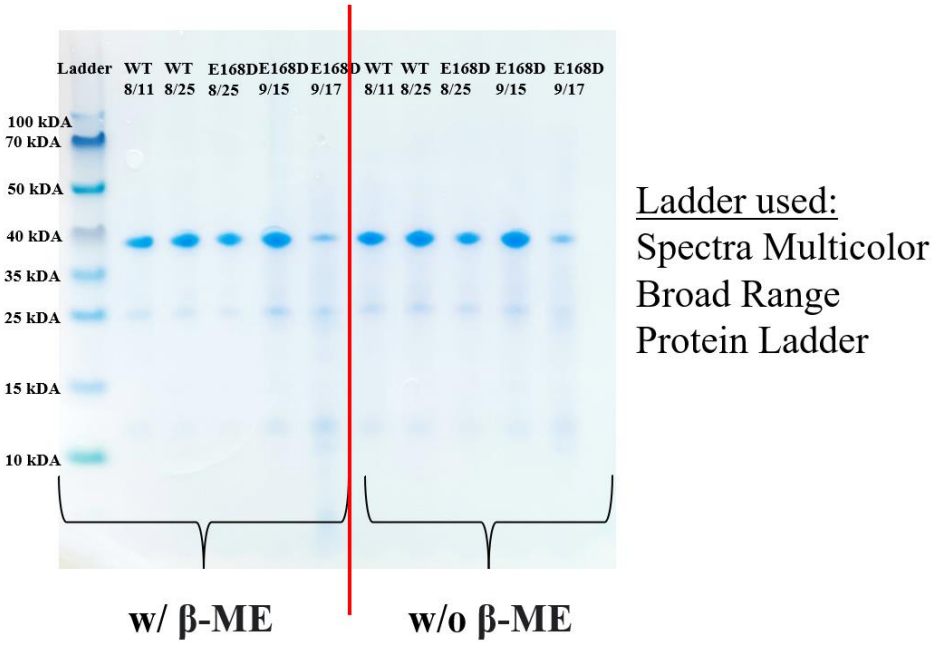

Figure 6. SDS-PAGE gel results of wildtype and E168D purifications. The machine used was a Mini Cell Novex gel box. The electrophoresis gel used was NuPAGE™ 4 to 12% Bis-Tris. Thermo Scientific™ Spectra Multicolor Broad Range Protein Ladder was used in lane 1. Lanes 2-6 show samples treated with 5% β-mercaptoethanol (Bio-Rad Laboratories). Lanes 7-11 show no treatment with 5% β-mercaptoethanol. Approximately 30 µM of each protein was loaded into each lane after being heat shocked at 90°C for 5 minutes. The SDS-PAGE was run for 1 ½ hours at a setting of 107 volts at 25°C. Lanes 2-11 showed 40 kDa to be the molecular weight in both types of protein. No dimers seem to be present in the proteins. Light bands are seen at 25 and 13 kDa and might be degradation products of the *Mt*IGPS due to their sum being 38 kDa (approximate to 40 kDa seen for the most concentrated band). The light bands can also be representative of impurities and different proteins entirely.

3.2 *Steady State Kinetics*

Steady state kinetics of wildtype and E168D proteins were monitored via K_M experiments, which were necessary to determine the affinity of enzymes to bind the substrate. K_M may also test for stability of enzymes and is defined as the amount of substrate needed to reach ½ V_{max}. The K_M for wildtype was found to be 6.9 ± 1.4 µM through experiments done by Sarah Cho. The K_M for E168D was found by me to be 25 ± 3 µM (see table 2). In table 2, there is a 3.62-fold increase in the amount of substrate needed for the E168D mutant to reach ½ V_{max}, indicating a weak binding affinity to the substrate after the mutation. This may mean that the carbon length of the E168 residue is important to the mechanism and may be helping the E168 anionic oxygen reach the cationic nitrogen in I_1 of the mechanism (Reddy et al., 2012). The k_{cat}, which is defined as the conversion of substrate molecules into product per second, is a necessary value to show the

overall rate of the catalytic mechanism. As explained in the methods, K_M, V_{max} and their relative errors were generated by Kaleidagraph through the Michaelis-Menten equation, while k_{cat} was produced after dividing V_{max} by enzyme concentration (M). The k_{cat} for wildtype was found by Sarah Cho to be 0.022 ± 0.002 s^{-1} and for E168D was found to be 0.011 ± 0.001 s^{-1} through my experiments. Since the Glu to Asp substitution still has an acidic side chain, a dramatic decrease in the k_{cat} was not seen. These results also indicate that E168D is 2-fold less active than wildtype, which may hint that there is a rate limiting step (slow step) in the mechanism, which decreases the overall k_{cat} of the reaction.

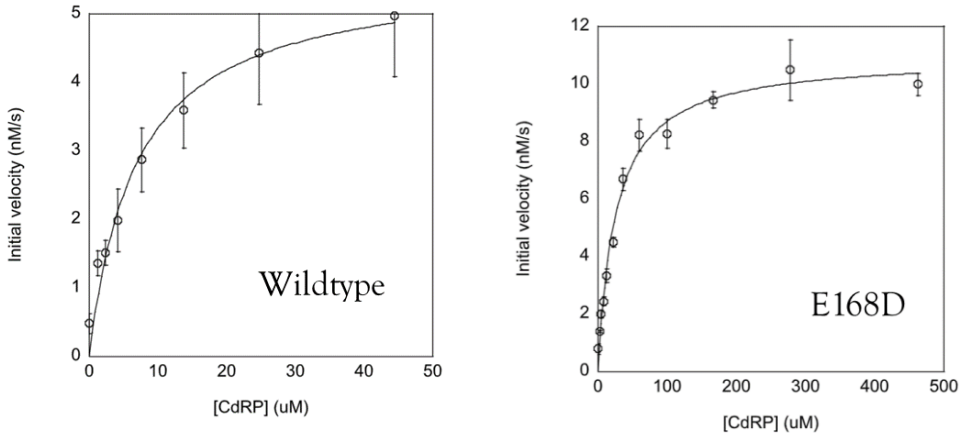

Figure 7. K_M graphs for wildtype (obtained by Sarah Cho) and E168D *Mt*IGPS proteins (obtained by me). A Biotek plate reader was set to absorbance and readings were taken in clear Costar 120 µL well plates. The wavelength used was 278 nm, 0.46 cm was the pathlength and the extinction coefficient (Ɛ) was 5500 $M^{-1}cm^{-1}$. The experiments were conducted at 25°C with PIPES assay buffer (2 mM DTT, 100 mM PIPES, pH 7.5) In each well, 250 nM of wildtype or 1 µM of E168D was used in the experiments in conjunction with serial dilutions of CdRP. Triplicates at each CdRP concentration were taken to obtain the data and standard errors were calculated to obtain error bars. Lastly, the K_M, V_{max} and their respective error values were

generated via the software Kaleidagraph by fitting the data from the plate reader to the Michaelis-Menten equation. The K_M value for wildtype was 6.9 ± 1.4 (obtained by Sarah Cho) and for E168D (obtained by me) was 26 ± 3 μM, a 3.62-fold increase after the mutation. The k_{cat} value was obtained by dividing V_{max} by the enzyme concentration. for wildtype was 0.022 ± 0.002 s^{-1} (Sarah Cho) and for E168D (by me) was found to be 0.011 ± 0.001 s^{-1} via the Michaelis-Menten equation.

Table 2. Comparison of the k_{cat} values for wildtype (obtained by Sarah Cho) and E168D (obtained by me) *Mt*IGPS proteins. The values of k_{cat} were generated by Kaleidagraph via the Michaelis-Menten equation using V_{max}. The k_{cat} value for wildtype was 0.022 ± 0.002 s^{-1} (Sarah Cho) and for E168D (by me) was found to be 0.011 ± 0.001 s^{-1} via the Michaelis-Menten equation.

*Mt*IGPS	k_{cat} (s^{-1})	K_M (μM)
Wildtype	0.022 ± 0.002	6.9 ± 1.4
E168D	0.011 ± 0.001	26 ± 3

In the PDB structure, 1LBF, the ligand (r-CdRP) is located in close proximity to E158 of *Ss*IGPS (correlates to E168 in *Mt*IGPS) to form hydrogen bonds. Since the K_M value increased with the E168D mutation, the E168 residue may be involved in substrate capture. To capitalize on this finding, when E158 *Ss*IGPS was investigated on PDB structure 1LBF, imagining the carbon chain of E158to be shorter by one methyl group, helps to picture two important hydrogen bonds (shown in figure 8). These hydrogen bonds have the capability of sitting further back on the structure after the mutation and not being able to reach the H_2O that mediates hydrogen bonding

to the r-CdRP (reference figure 8). The hydrogen bonds cannot form after the mutation due to the increased distance, and an inference can be made that this could lead to weak binding affinity to the substrate. Just as figure 5 shows, the E168D mutation in *Mt*IGPS could also lead to weak binding affinity for the substrate CdRP because of the shorter methyl, giving insight into the larger K_M value obtained from the data.

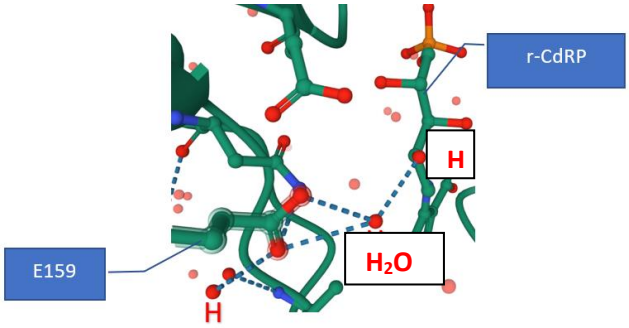

Figure 8. PDB structure 1LBF. *Ss*IGPS in complex with r-CdRP ligand. Two hydrogen bonds on the E158 residue are shown to have a possible hydrogen bonding network through an H_2O molecule with r-CdRP. The figure relates to the E168 residue of *Mt*IGPS and can infer why there is a weaker binding affinity after the E168D mutation.

3.3 *Rate Limiting Determination Steps*

Understanding where the rate limiting step (slow step) lies in a mechanism, may hint at the overall role of the residue in the mechanism. If the rate limiting step is different after the mutation, this may give insight into which step the residue may be involved in. To add, if the rate limiting step is known and there is a change in k_{cat}, the role a particular residue can be determined. The rate limiting step may also help in drug design and inhibit certain steps in the IGPS mechanism. If the rate limiting step is known, this helps to predict transition state analogs

as well. Additionally, this prediction may give a clear image as to what the transition state looks like and helps interpret the changes of the k_{cat} of the mutation.

Solvent deuterium kinetic isotope effect (SDKIE) experiments were conducted to determine if the rate limiting step lied in step 2 or 4 of the mechanism (reference figure 9).

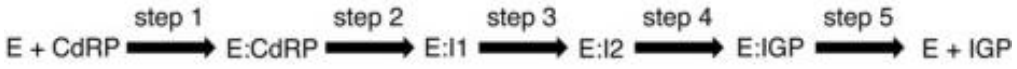

Figure 9. Steps of the IGPS catalytic pathway mapping out locations of where rate limiting steps may lie using SDKIE and SVE (solvent viscosity effect) experiments. SDKIE experiments may hint that the rate limiting step occurs at step 2 or 4 (proton transfer steps). SVE experiments may hint that the rate limiting step occurs at step 1 and/or step 5 (substrate binding and/or product release, respectively). Figure taken from Zaccardi et al., 2013.

In SDKIE experiments, if the rate of the reaction (k_{cat}) decreases upon addition of the heavy isotope, D_2O, then the rate limiting step may lie in either of the proton transfer steps (see step 2 and 4 in figure 9). In, D_2O, the isotope has a lower vibrational frequency than H_2O and therefore requires a larger input of energy to reach the transition state. Due to this large input of energy, it can be inferred that upon addition of D_2O, a slower overall rate of the reaction should be observed. If a decreased rate is in fact seen, it would again mean that the proton transfer steps could be the rate limiting steps in the mechanism (Northrop, 1975).

Figure 10 shows wildtype and E168D SDKIE experiment results in their preliminary stages. A consistent pattern is not observed upon the addition of D_2O while error bars overlap in the

wildtype runs. The temperatures 25 and 37°C were also used to see if rate limiting steps varied with temperature (seen in Sarah Cho's previous experiments). Comparing wildtype 0 and 100% D_2O amounts in both temperatures, it is seen that there is a potential inverse kinetic isotope effect (IKIE) at 25°C. An IKIE may appear in multistep reactions where rapid equilibrium occurs before the rate limiting step (Fernandez and Murkin, 2020). Fernandez and Murkin also explain that D_2O has a viscosity that is 24% more than H_2O. This variability in viscosity may even be the reason for the detected inverse kinetic isotope effect in the figure 10 data. It is important to create equal viscosity conditions in further experiments with D_2O and H_2O to rule out any viscosity effects (Fernandez and Murkin, 2020). To add, at the wildtype 37°C graph in figure 10, no change (error bars overlap) is seen between 0 and 100% D_2O, meaning that the rate limiting step may not be determined or the proton transfer step may not be the rate limiting step in the mechanism. Further experimentation should also be conducted on the wildtype at both temperatures since the pattern does not follow a steady increase or decrease upon addition of D_2O buffer.

For preliminary E168D SDKIE experiments, k_{cat} values do not match the established 0.011 s^{-1}, so further trials need to be conducted. A reason for the discrepancy in turnover number could be constant thawing of proteins from -80°C freezing or other experimental errors. Moreover, at 25°C E168D, results do not seem like the rate limiting step can be determined as the proton transfer steps (due to error bars overlapping). Lastly, for E168D at 37°C, another IKIE is seen, but again k_{cat} values are not in proximity to the established values and therefore further investigation needs to be done as discussed previously. Overall, this preliminary data for

wildtype and E168D needs to be replicated to hint where the rate limiting step may lie in the mechanism and how the E168 residue affects the mechanism.

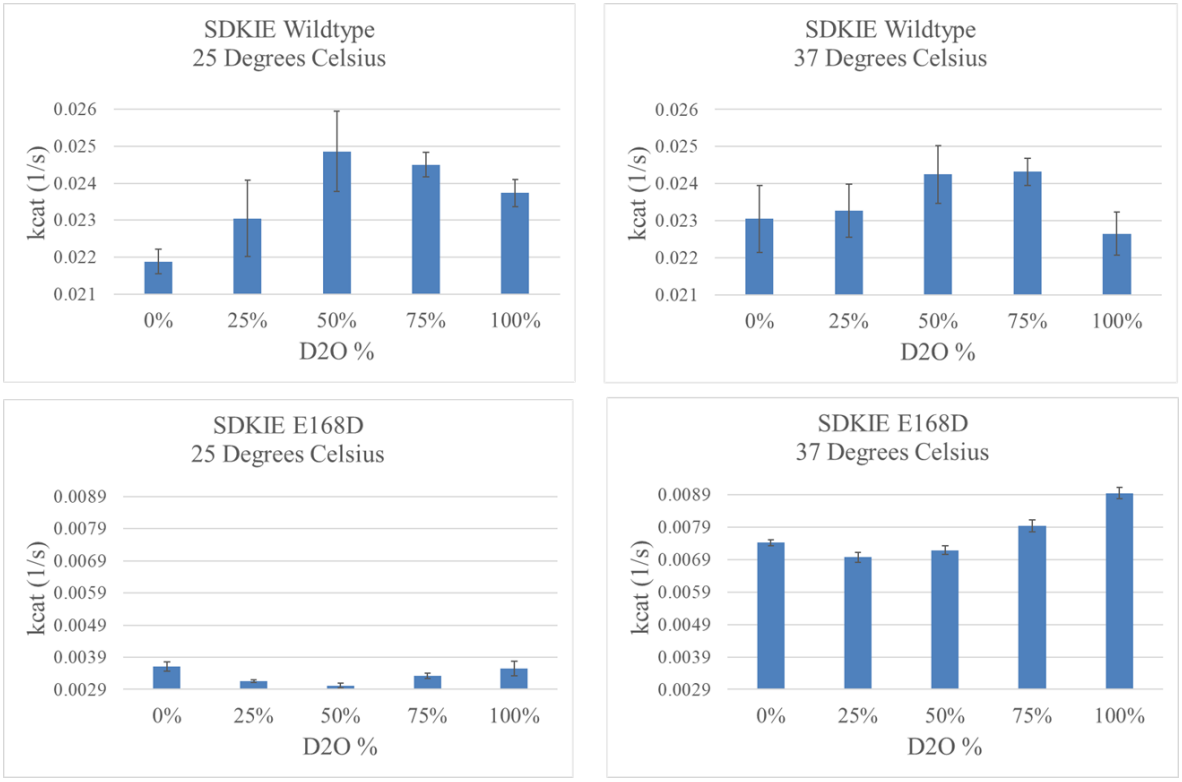

Figure 10. SDKIE of wildtype and E168D *Mt*IGPS proteins. A Biotek plate reader was used to measure the absorbance of proteins at 278 nm using Greiner clear half-area plates. The extinction coefficient was 5500 M^{-1} cm^{-1} and the pathlength was 1 cm. The temperature for the experiments was set to 25°C and 37°C. For wildtype and E168D proteins, 1 µM of protein was used in the wells as well as 400 µM of CdRP. In these experiments, H_2O or D_2O based PIPES assay buffer (2 mM DTT, 100 mM PIPES, pH 7.5) was used. The results show five replicates done on each addition of D_2O with error bars over the course of 10 minutes. Note that E168D k_{cat} values at both temperatures are not in proximity to the established 0.011 s^{-1}. Further experimentation needs to be conducted on wildtype and E168D on where the rate limiting step could lie in the mechanism.

Solvent viscosity effect (SVE) experiments were conducted in glycerol and sucrose to determine if the rate limiting steps in the IGPS mechanism were diffusion-controlled via variable viscosity. Glycerol and sucrose molecules were used in these experiments because they produced preferential hydration of proteins. Preferential hydration is a method used to discover the mechanism of stabilization or denaturation in proteins. For glycerol buffers, this hydration was caused by the glycerol molecules repelling the proteins. On the other hand, sucrose buffers cause preferential hydration of proteins via an increase of surface tension and steric impediment (Uribe and Sampedro, 2003). Moreover, diffusion-controlled steps of the mechanism may consist of the substrate capture or product release steps of the mechanism. Respectively, steps 1 and 5 in figure 9. Commonly, substrate capture is a fast step and product release is most likely the rate limiting step.

For the results seen in figure 11, the wildtype protein in glycerol shows a decrease in catalytic activity upon addition of glycerol. From the data, it can be inferred that there might be an actual SVE at both 25 and 37°C. This result may indicate that the rate limiting step does not change based upon the temperature. The opposite was observed in preliminarily SDKIE experiments. There is also a dramatic increase in k_{cat} after a 10% increase in glycerol. This increase might be seen due to glycerol stabilizing the protein structure conformation. This was an unexpected result, but more experiments are needed for the preliminary results seen. It may also be inferred by this wildtype in glycerol result that step 1 and/or step 5 of the mechanism might be the rate limiting steps in the mechanism. To compare, in the E168D mutation, error bars are seen overlapping in the data, leading to an inference that no SVE's may be seen (reference figure 11).

In this preliminary result, it could be also inferred that since no SVE's are seen after the mutation, that the E168 residue may play an important role in substrate capture and/ or product release of steps 1 and 5 of the mechanism (figure 9). Furthermore, there is an approximate 2-fold decrease in the activity of the wildtype and approximate 4-fold decrease in activity for E168D at 25°C in 0% glycerol, possibly caused by multiple thawing of old protein fractions where denaturation of the protein may have occurred. For this reason, further experimentation needs to be conducted on enzymes that contain activity values that are similar to established.

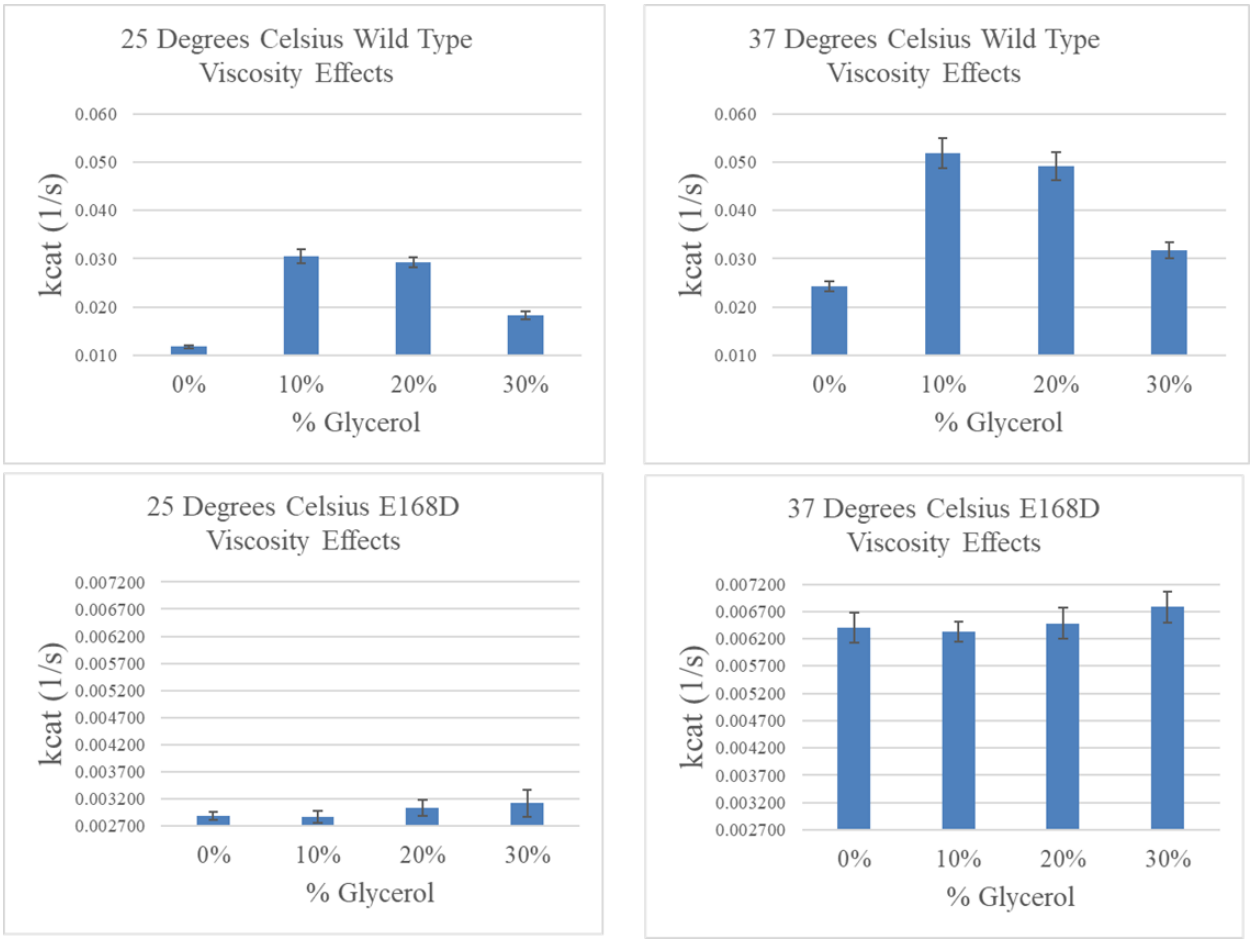

Figure 11. Wildtype and E168D tested in differing amounts of glycerol PIPES buffers at 25 and 37°C. A Biotek plate reader was used to take absorbance readings at 278 nm from Greiner half-area clear plates. The pathlength was 1 cm and the extinction coefficient (Ɛ) used to calculate k_{cat}

from the absorbance readings was 5500 M^{-1} cm^{-1}. For wildtype and E168D proteins, 1 µM of protein was used in the wells. The concentration of CdRP used in the experiment was 400 µM. Five replicates of each percentage of glycerol were run for 15 minutes to obtain the data and error bars. An increase in of rate in wildtype 0 – 10% was seen, meaning that the glycerol may be stabilizing the structure of IGPS to create a higher rate. A decrease in rate was seen upon addition of glycerol from 10 to 30% meaning the slow step in the mechanism could lie in the substrate binding or product release step. Since the error bars are overlapping in E168D experiments, the results are inconclusive.

Furthermore, an experiment for wildtype tested in various concentrations of sucrose was also conducted. From the data seen in figure 12, there is a potential inverse SVE seen in the preliminary experiment when comparing 0 M and 1.5 M sucrose at both temperatures. In this experiment, sucrose may possibly be serving to stabilize the wildtype conformation, lending to its increased activity, but more experimentation needs to be conducted for the preliminary data especially because a gradual increase is not seen in the data upon the addition of sucrose. To add, for E168D experiments conducted in varying sucrose concentrations, SVE's are seen at both temperatures. It may also be inferred by this result that step 1 and/or step 5 are the rate limiting steps in the mechanism. Moreover, there is an approximate 4-fold decrease in the activity of the wildtype and approximate 5-fold decrease in activity for E168D at 25°C in 0% glycerol, possibly caused by multiple thawing of old protein fractions where denaturation of the protein may have occurred. For this reason, further experimentation needs to be conducted on enzymes that contain activity values that are similar to established values.

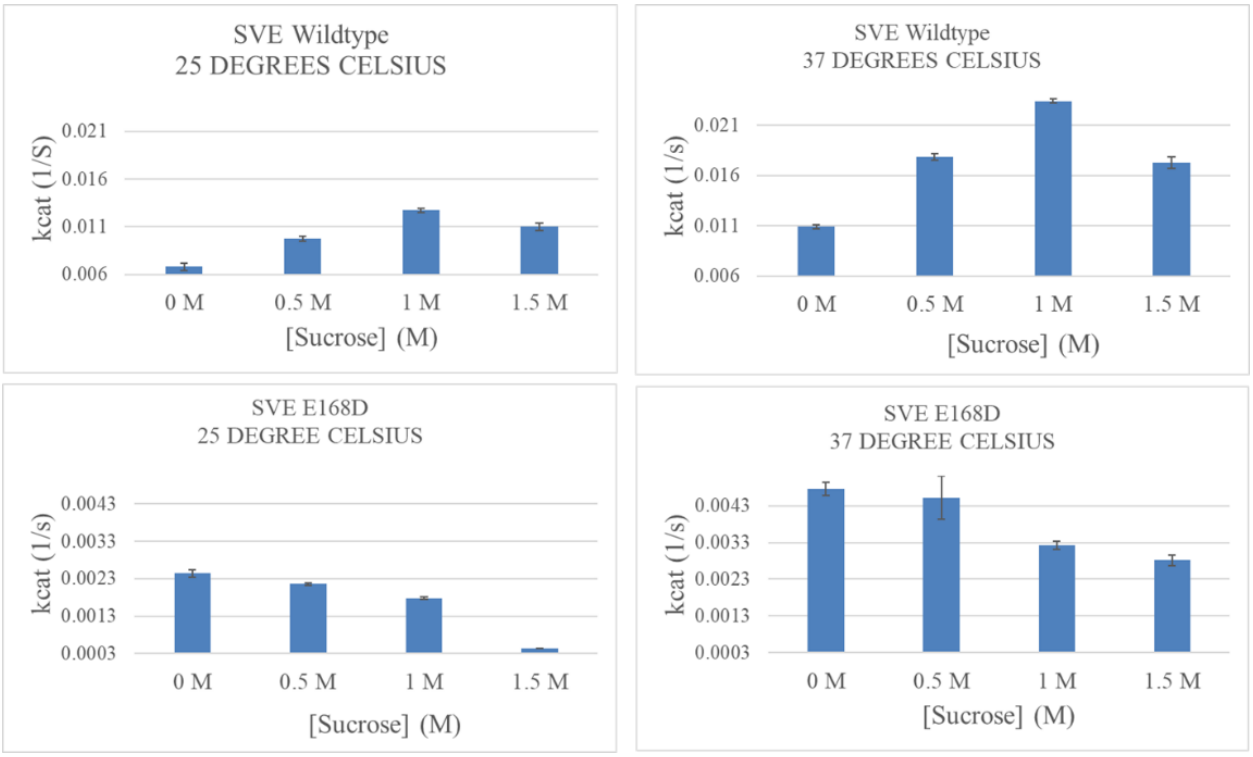

Figure 12. Wildtype and E168D tested in differing concentrations of sucrose in PIPES buffers (2 mM DTT, 100 mM PIPES, pH 7.5) at 25 and 37°C. A Biotek plate reader was used to take absorbance readings at 278 nm from Greiner half-area clear plates. The pathlength was 1 cm and the extinction coefficient (Ɛ) used to calculate k_{cat} from the absorbance readings was 5500 M^{-1} cm^{-1}. For wildtype and E168D proteins, 1 µM of protein and 400 µM CdRP was used in the experiments. Five replicates of each concentration of sucrose were run for 15 minutes to obtain the data and error bars. SVE was also conducted in sucrose using the same machine settings and methods seen in glycerol SVE discussed in the previous paragraph. For wildtype, the k_{cat} seems reduced overall showing the value of 0.007 s^{-1}. This unexpected change could be caused by structure denaturation from re-thawing the same protein fraction multiple times. This is a 3-fold reduction in the established 0.02 s^{-1} k_{cat} seen in wildtype.

For future SVE experiments, it would be helpful to use a large molecule such as PEG (polyethylene glycol) as a macroviscogen or crowding agent. A macromolecular crowding agent is a viscogen that takes up space in the buffer, causing inaccurate concentrations of components in the assay. Glycerol and sucrose are small molecules where only viscosity and not crowding is not introduced. Using PEG could help to uncover ideas as to why increased activity might be seen in SVE experiments. To add, future K_M experiments need to be conducted to specifically study the substrate binding in the mechanism. Variable viscosity and variable substrate concentration will be able to help interpret if the rate limiting step lies in the substrate binding step or product release and if E168 is important in binding to the substrate.

3.4 *pK_{a1} and pK_{a2} determinations: pH Profiles*

A determination of pK_{a1} and pK_{a2} were made via pH profile experiments. In these experiments, the pH of PIPES buffer was varied while k_{cat} values were calculated (reference figure 13). For the pH profile experiments, I was expecting both the wildtype and E168D mutation to be within similar pK_a values and not shift. This is because I mutated an acidic side chain to another acidic side chain so low difference in pK_a should be seen. From assumptions made by Czekster and Shen, I also hypothesized that E168 could be the catalytic base in the mechanism that could loan its negative oxygen to the cationic nitrogen in intermediate 1 of the mechanism shown in figure 1. For a catalytic base graph, I expected to see protonation at lower pH values (high activity) and deprotonation at higher pH values (low activity).

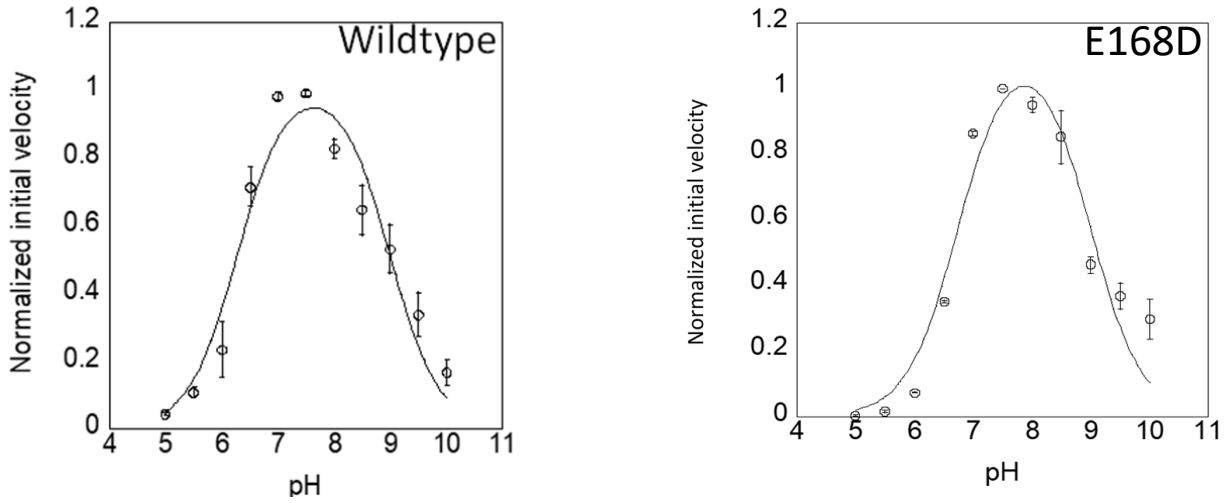

Figure 13. Wildtype (obtained from Sarah Cho) and E168D *Mt*IGPS protein pH profiles conducted at 25°C. The contents of each run consisted of 250 nM wildtype or 1 µM E168D and 60 µM CdRP in MTEN buffer (100 mM NaCl, 25 mM Tris Base, 1 mM DTT, 25 mM ethanolamine, 50 mM MES). The experiments were conducted on a Biotek plate reader set to fluorescence where black Costar plates were used. The excitation wavelength was set to 278 nm while the emission wavelength was set to 340 nm. Four replicates per pH value, run for 30 minutes each, were used to obtain the data. From fluorescence versus time graphs, slopes (CPS per second) were divided by each individual trial's highest value slope to normalize the data. After this, each of the 4 trials normalized values were averaged and the standard deviation was calculated for each pH value ranging from 5-10 pH. The standard error was finally taken by dividing the standard deviation by the square root of 4 (number of trials taken). This data of average normalized velocity versus pH was then inserted to Kaleidagraph and fit to the Zaccardi pH equation (equation 2) to predict the pK_{a1} and pK_{a2} values. Standard error was also shown as error bars in the data graphs.

In my experiments for wildtype and E168D (seen in figure 13), I obtained a bell curve in both data sets that showed a peak for both proteins at pH 7.5. To add, for wildtype, I obtained the values for pK_{a1} of 6.3 ± 0.1 and pK_{a2} of 9.0 ± 0.1 through the help of the Zaccardi pH equation (see equation 2). For E168D, the proteins exhibited a pK_{a1} of 6.8 ± 0.2 and pK_{a2} of 9.0 ± 0.2. The non-shifting of the pK_a values was expected due to the acidic side chains of both proteins, as explained previously, while hinting that the E168 residue might not be important in catalysis of the mechanism.

Due to the bell curve in both my data sets (figure 13), it can be inferred that no catalytic base behavior is being seen. One explanation of the observed behavior could be that the activity decreased at low pH values because the possible catalytic base becomes protonated and therefore decreased in k_{cat}. Another explanation could be that the proteins unfold at low pH values. To back this finding, a preliminary investigation by Dr. Catalano with circular dichroism (CD) was conducted and showed that there might be misfolding at low pH values for the proteins.

Additionally, a previous mutant studied by Maryum Bhatti called E168Q was found to have no activity. This finding may suggest that an ionic stabilization may be taking place in the E168 residue between the protein and intermediate 1 of the mechanism. This stabilization of the positive nitrogen on the intermediate via the oxygen anion on the residue may be causing the energy in the intermediate to decrease. Even though the E and Q amino acids are of the same length, the amide functional group cannot behave as an acid or proton acceptor. The Gln mutation may also be able to form the same hydrogen bonds in the active site but can never be

the catalytic base. Through studying the E168Q mutant, the ability of the residue to accept protons was removed and ultimately showed no activity. It can be inferred from this study that the charge of the E168 residue is vital in the mechanism to serve as the catalytic base stabilizing intermediate 1.

Even though the E168 residue does not seem to be the catalytic base from the data shown in figure 12, experiments conducted by Dr. Catalano and Maryum Bhatti indicate that there is still a possibility of E168 being in fact the catalytic base as Czekster and Shen hypothesized. Further investigation on these preliminary experiments needs to be conducted.

4. Conclusion

In conclusion, the E168 residue of *Mt*IGPS may be playing an important role in the IGPS mechanism. Mutating the side chain of E168 to D affected the overall rate (k_{cat}) of the reaction two-fold from 0.022 ± 0.002 s^{-1} to 0.011 ± 0.001 s^{-1} after the mutation. The E168D mutation also showed more substrate was needed to reach ½ V_{max} after the mutation. Shortening the acidic side chain by one methyl group could be affecting the binding affinity of the protein to the substrate. The K_M value before the mutation was found to be 6.9 ± 1.4 µM and after became 26 ± 3 µM, which is approximately a 4-fold increase that supports the hypothesis that if the side chain is shortened, it may not be able to reach the substrate as readily to bind. Further SDKIE and SVE experiments need to be conducted to predict the rate limiting steps in the mechanism to hint at the overall role of the residue in the mechanism. Lastly, pH profile experiments do not indicate any catalytic base behavior as previously hypothesized, but further investigation of other residues as

well as circular dichroism experiments may help to uncover the role of the residue in the mechanism. Much can be learned about the IGPS catalytic pathway through experimentation with the E168 residue and its D mutation. The comparisons made between the two have provided a heightened understanding of not only the structure-function relationships but also the binding interactions in the IGPS catalytic pathway.

Works Cited

Anochie, P. I., Ajogwu, A., Kalu, G. O., Akpan, M. I.- O., Onyeneke, E. C., and Onyeozirila, A. C. (2018) How to control the Tuberculosis and HIV/AIDS dual epidemic, *Archives of Community Medicine and Public Health* 4(2), 26-37.

Barberis, I., Bragazzi, N. L., Galluzzo, L., and Martini, M. (2017) The history of tuberculosis: from the first historical records to the isolation of Koch's bacillus, *Journal of Preventive Medicine and Hygiene* 58(1), E9-E12.

Cole, S., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., ... and Tekaia, F. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence, *Nature* 393(6685), 537.

Collins, J. M., Siddiqa, A., Jones, D. P., Liu, K., Kempker, R. R., Nizam, A., Shah, N. S., Ismail, N., Ouma, S. G., Tukvadze, N., Li, S., Day, C. L., Rengarajan, J., Brust, J. C., Gandhi, N. R., Ernst, J. D., Blumberg, H. M., and Ziegler, T. R. (2020) Tryptophan catabolism reflects disease activity in human tuberculosis, *JCI insight* 5(10), 1-16.

Czekster, C. M., Lapis, A. A., Souza, G. H., Eberlin, M. N., Basso, L. A., Santos, D. S., and Neto, B. A. (2008) The catalytic mechanism of indole-3-glycerol phosphate synthase (IGPS) investigated by electrospray ionization (tandem) mass spectrometry, *Tetrahedron Letters* 49(41), 5914-5917.

Czekster, C. M., Neto, B. A., Lapis, A. A., Dupont, J., Santos, D. S., and Basso, L. A. (2009) Steady-state kinetics of indole-3-glycerol phosphate synthase from Mycobacterium tuberculosis, *Archives of biochemistry and biophysics* 486(1), 19-26.

Eberhard, M., Tsai-Pflugfelder, M., Bolewska, K., Hommel, U., and Kirschner, K. (1995) Indole glycerol phosphate synthase-phosphoribosyl anthranilate isomerase: comparison of the bifunctional enzyme from *Escherichia coli* with engineered monofunctional domains, *Biochemistry* 34(16), 5419-5428.

Esposito, N., Konas, D. W., and Goodey, N. M. (2021) Indole-3-Glycerol Phosphate Synthase From *Mycobacterium tuberculosis*: A Potential New Drug Target, *ChemBioChem* 22, 1-11.

Fernandez, P. L., & Murkin, A. S. (2020) Inverse Solvent Isotope Effects in Enzyme-Catalyzed Reactions, *Molecules (Basel, Switzerland)* 25(8), 1933.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, *The Proteomics Protocols Handbook*, 571-607.

Giger L., Caner S., Obexer R., Kast P., Baker D., Ban N., Hilvert D. (2013) Evolution of a designed retro-aldolase leads to complete active site remodeling, *Nature Chemical Biology* 9, 494–498.

Good, Norman E.; Winget, G. Douglas; Winter, Wilhelmina; Connolly, Thomas N.; Izawa, Seikichi; Singh, Raizada M. M. (1966) Hydrogen Ion Buffers for Biological Research, *Biochemistry* 5(2), 467-77.

Hennig M., Darimont B., Sterner R., Kirschner K., Jansonius J.N. (1995) 2.0 å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability, *Structure* 3(12), 1295–1306.

Hennig, M., Darimont, B. D., Jansonius , J. N., and Kirschner, K. (2002) The catalytic mechanism of indole-3-glycerol phosphate synthase: Crystal structures of complexes of the enzyme from Sulfolobus solfataricus with substrate analogue, substrate, and product, *Journal of Molecular Biology* 319(3), 757-766.

Kirschner, K., Szadkowski, H., Jardetzky, T. S., and Hager, V. (1987) [48] Phosphoribosylanthranilate isomerase—indoleglycerol-phosphate synthase from Escherichia coli, *Methods in Enzymology* 142, 386-397.

Leisola M., and Turunen O. (2007) Protein engineering: opportunities and challenges, *Applied Microbiology and Biotechnology* 75, 1225-1232.

Lott, J. S. (2020) The tryptophan biosynthetic pathway is essential for mycobacterium tuberculosis to cause disease, *Biochemical Society Transactions* 48(5), 2029-2037.

Northrop, D. B. (1975) Steady-state analysis of kinetic isotope effects in enzymic reactions, *Biochemistry* 14(12), 2644-2651.

Ormerod, L. P. (2005) Multidrug-resistant tuberculosis (MDR-TB): epidemiology, prevention and treatment, *British Medical Bulletin* 73-74(1), 17-24.

O'Rourke, K., Jelowicki, A., and Boehr, D. (2016) Controlling active site loop dynamics in the (β/α) 8 barrel enzyme indole-3-glycerol phosphate synthase, *Catalysts, 6(9),* 129.

Parry, R.J. (1971) Biosynthesis of compounds containing an indole nucleus, *Chemistry of Heterocyclic Compounds: Indoles, Part Two* 25, 1-64.

Rath, A., Glibowicka M., Nadeau, V., Chen, G. and Deber, C. (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins, *The Proceedings of the National Academy of Sciences* 106(6), 1760-1765.

Reddy, M. C. M., Bruning, J. B., Thurman, C., and Sacchettini, J. C. (2012) 3T44: Crystal structure of mycobacterium tuberculosis indole glycerol phosphate synthase (IGPS) in complex with indole glycerol phosphate (IGP) AMD anthranilate, *RCSB Protein Data Bank*.

Sansetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis, *Molecular microbiology*, *48(1)*, 77-84.

Schlee, S., Dietrich, S., Kurćon, T., Delaney, P., Goodey, N. M., and Sterner, R. (2012) Kinetic mechanism of indole-3-glycerol phosphate synthase, *Biochemistry* 52(1), 132-142.

Shen, H., Wang, F., Zhang, Y., Huang, Q., Xu, S., Hu, H., ... and Wang, H. (2009) A novel inhibitor of indole-3-glycerol phosphate synthase with activity against multidrug-resistant Mycobacterium tuberculosis, *The FEBS journal* 276(1), 144-154.

Smith, O.H. and Yanofsky, C., 1962 [107] Enzymes involved in the biosynthesis of tryptophan, *Methods in enzymology* 5, 794-806.

Szmacinski, H. and Lakowicz, J. R. (1993) Optical measurements of pH using fluorescence lifetimes and phase-modulation fluorometry, *Analytical chemistry* 65(13), 1668–1674.

Thomas, O. (2019) Indole-3-glycerol Phosphate Synthase Ligand Binding Interactions, *Montclair State University Digital Commons,* 322.

Uribe, S. and Sampedro, J. G. (2003) Measuring Solution Viscosity and its Effects on Enzyme Activity, *Biological Procedures Online* 5(1), 108-115.

Yang, Y., Zhang, M., Zhang, H., Lei, J., Jin, R., Xu, S., Bao, J., Zhang, L. and Wang, H. (2006) Purification and Characterization of Mycobacterium tuberculosis Indole-3-glycerol Phosphate Synthase, *Biochemistry (Moscow)* 71(1), S38-S43.

Zaccardi, M. J., Yezdimer, E. M., and Boehr, D. D. (2013) Functional Identification of the General Acid and Base in the Dehydration Step of Indole-3-glycerol Phosphate Synthase Catalysis*, *Enzymology* 288(37), 26350-26356.

Zaman, K. (2010) Tuberculosis: A global health problem, *Journal of Health, Population, and Nutrition* 28(2), 111-113.

Zhang, H., Li, D., Zhao, L., Fleming, J., Lin, N., Wang, T., Liu, Z., Li, C., Galwey, N., Deng, J., Zhou, Y., Zhu, Y., Gao, Y., Wang, T., Wang, S., Huang, Y., Wang, M., Zhong, Q., Zhou, L., Chen, T., Zhou, J., Yang, R., Zhu, G., Hang, H., Zhang, J., Li, F., Wan, K., Wang, J., Zhang, X.-E., and Bi, L. (2013) Genome sequencing of 161 mycobacterium tuberculosis

isolates from China identifies genes and intergenic regions associated with drug resistance, *Nature Genetics* 45(10), 1255-1260.

Zhang, Y. J., Reddy, M. C., Ioerger, T. R., Rothchild, A. C., Dartois, V., Schuster, B. M., Trauner, A., Wallis, D., Galaviz, S., Huttenhower, C., Sacchettini, J. C., Behar, S. M., and Rubin, E. J. (2013) Tryptophan biosynthesis protects mycobacteria from CD4 T-cell-mediated killing, *Cell* 155(6), 1296-1308.