5-2021

Investigating the Effects of Catalytic Activity in Mycobacterium Tuberculosis Indole-3-Glycerol Phosphate Synthase in Wildtype, N189L, And E57D Variants

Sarah S. Cho
Montclair State University

Abstract

Indole-3-glycerol phosphate synthesis (IGPS) is an enzyme that catalyzes the ring closure in 1-(*o*-carboxylphenylamino)-1-deoxyribulose 5-phosphate (CdRP). Multiple steps are expected to be involved in formation of the pyrrole ring including dehydration, decarboxylation, cyclization, and condensation. IGPS is an essential protein in the pathogen *Mycobacterium tuberculosis* and a potential target in the treatment of drug-resistant tuberculosis. In order to better understand the function of *Mycobacterium tuberculosis* IGPS (*mt*IGPS), we introduced single-point mutations into active site residues. This was based on the hypothesis that introducing mutations would affect the catalysis in a way where we would be able to obtain more information on our protein. We expressed and purified *Mt*IGPS mutants N189L and E57D. The catalytic activities of the mutants were lower than that of the wild type *mt*IGPS, indicating that these residues play important roles in catalysis. The Michaelis constant, pH-dependence, and temperature dependence of the mutants were also determined. Wildtype *mt*IGPS was found to have a K_M of $5.0 \pm 1.4\mu M$ and a k_{cat} value of $4.8 \pm 0.3 s^{-1}$ whereas E57D kinetics revealed a K_M of $29 \pm 6\ \mu M$ and a k_{cat} value of $0.102 \pm 0.018 s^{-1}$. N189L indicated no measurable activity and could not be further characterized. The data on E57D show an altered pH profile, possibly resulting from the lower pK_a of Asp compared to Glu, and support the hypothesis that E57 serves as the catalytic base in *mt*IGPS mechanism. Together these data contribute to an improved understanding of the *mt*IGPS catalysis and ligand binding.

MONTCLAIR STATE UNIVERSITY

Investigating the Effects of Catalytic Activity in *Mycobacterium tuberculosis* Indole-3-glycerol phosphate synthase in Wildtype, N189L, and E57D Variants

by

Sarah S. Cho

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 2021

College of Science and Mathematics (CSAM)

Department of Chemistry and Biochemistry

Thesis Committee:

[redacted]

Dr. Nina M. Goodey

Thesis Sponsor

[redacted]

Dr. David Konas

Committee Member

[redacted]

Dr. Jaclyn Catalano

Committee Member

INVESTIGATING THE EFFECTS OF CATALYTIC ACTIVITY IN *MYCOBACTERIUM TUBERCULOSIS* INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE IN WILDTYPE, N189L, AND E57D VARIANTS

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

by

Sarah S. Cho

Montclair State University

Montclair, NJ

2021

Table of Contents

Introduction

Tuberculosis is a disease that is still prevalent in today's society. According to the World Health Organization, an estimated total of 1.4 million people succumbed to mortality caused by TB in 2019, and 206,030 people were detected of carrying a strain of multidrug-resistant tuberculosis (MDR-TB)[1]. The number of detectable cases of MDR-TB continue to increase, as current drugs fail to provide any treatment of symptoms. In order to combat this disease, it is crucial to understand how to properly target the bacteria in order to ultimately find a proper inhibitor.

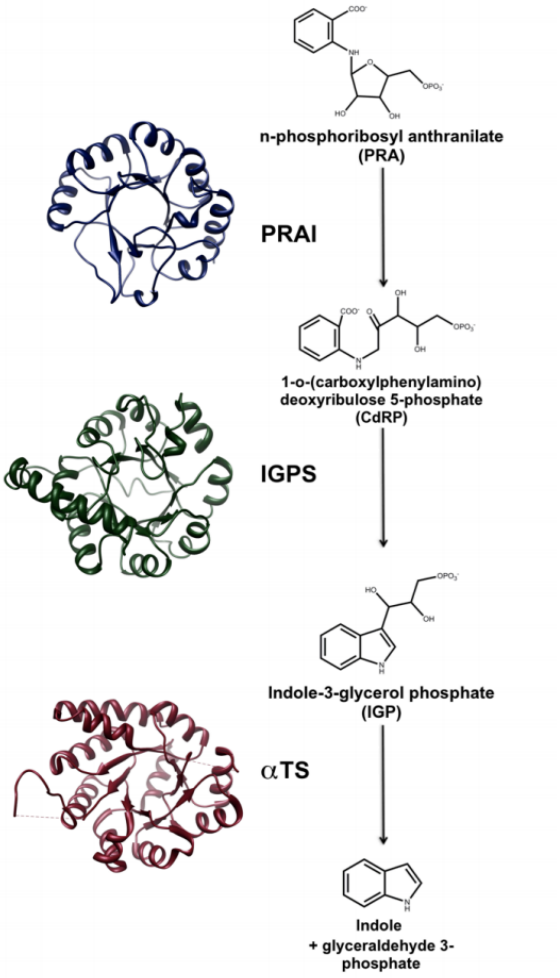


Figure 1: Steps 4-6 of tryptophan synthesis pathway. PRAI, IGPS, and αTS are all (β/α)8-barrel enzymes. IGPS specifically catalyzes conversion of CdRP into IGP

Mycobacterium tuberculosis, along with all other bacteria, is dependent on the essential amino acid, tryptophan. Bacteria utilize various (β/α)8-barrel fold enzymes in the pathway of formation of tryptophan (**Figure 1**)[2]. (β/α)8-barrel fold are popularly studied in the field of biochemistry due to their high success rates in manipulation and engineering than other complexes[3,4]. IGPS is an enzyme that catalyzes the formation of the indole ring seen in the product, indole-3-glycerol phosphate (IGP) by binding to its substrate, 1-o-(carboxylphenylamino) deoxyribulose 5-phosphate (CdRP). This enzyme, essential to all bacteria but not prevalent in humans, is important in furthering studies on

Mycobacterium tuberculosis to provide more insight on how to ultimately find a therapy for MDR-TB. Currently, there is no known inhibitor for *mt*IGPS[5].

Preliminary studies on IGPS include use of *Sulfolobus sulfataricus*, a thermophilic bacterium[5]. By studying the same residues tested to be catalytically active in *ss*IGPS in *Mycobacterium tuberculosis,* it is possible to analyze behavior of the enzyme under various temperature-controlled environments. As shown in **Figure 2**[6],

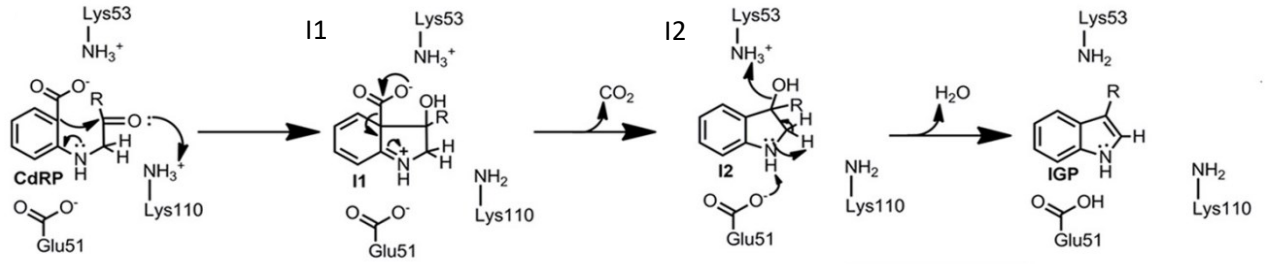


Figure 2: Proposed catalytic pathway for *ss*IGPS, which includes CdRP interacting with IGPS to form I1 and I2 intermediates, and finalize product IGP through decarboxylation, dehydration, and condensation

Zaccardi proposed a mechanism involving residues mapped out from *ss*IGPS[6]. From substrate and enzyme to product formation, two intermediates are projected in which a decarboxylation and dehydration is involved, as well as proton transfer. In the schema, E51 is thought to be involved with the second intermediate in acting as a proton acceptor for the finalization of ring closure and creating the double bond. K53 is proposed to be directly involved in proton donation for dehydration between the second intermediate (I2) and IGP, and K110 is thought to aid as a proton donor between the formation of intermediate 1 (I1) from CdRP. The question remains, if we study this same protein in *M. tuberculosis*, will we be able to obtain results supporting the theory that E51 is conserved to act as the catalytic base for IGPS, and either K53 or K110 acts as the enzyme's catalytic acid?

Only 30% of *mt*IGPS sequence is identical to *ss*IGPS's[5], so it was important to gain perspective on important residues on a map of *mt*IGPS. From *ss*IGPS to *mt*IGPS, E51 was conserved to E57, K53 to K59, and K110 to K119 (**Figures 3**[7] **& 4**). By inducing single-point mutations to these and other valuable residues taken from the homolog *ss*IGPS, it is possible to gather data on IGPS that allows for eventual engineering of inhibitors for the enzyme's active site

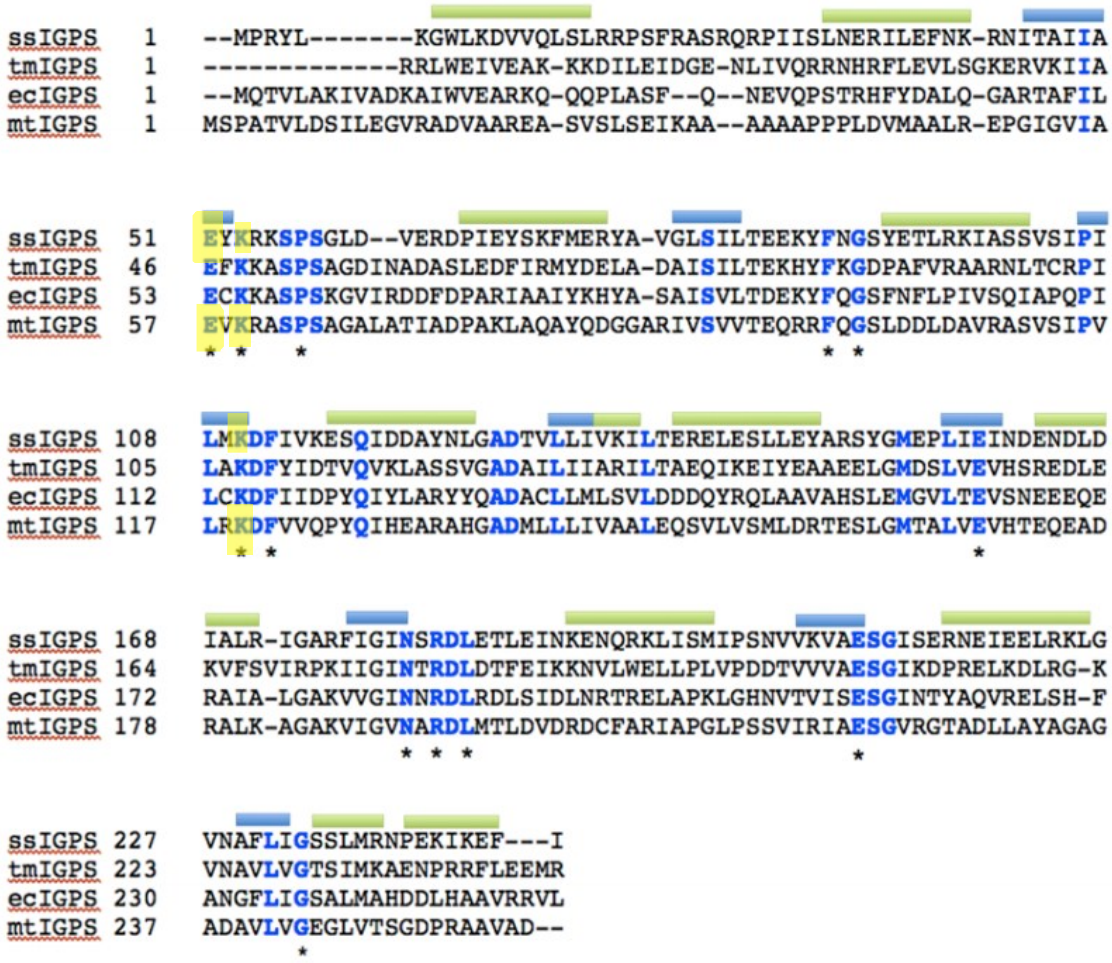


Figure 3: Sequence of IGPS for four different bacteria, including *ss*IGPS and *mt*IGPS (*tm*IGPS is *Thermotoga maritima* IGPS and *ec*IGPS is *Escherichia coli* IGPS . Conserved residues are written in blue. *mt*IGPS shows alignment of E57, K59, and K119 to E51, K53, and K110 in *ss*IGPS respectively (highlighted in yellow)

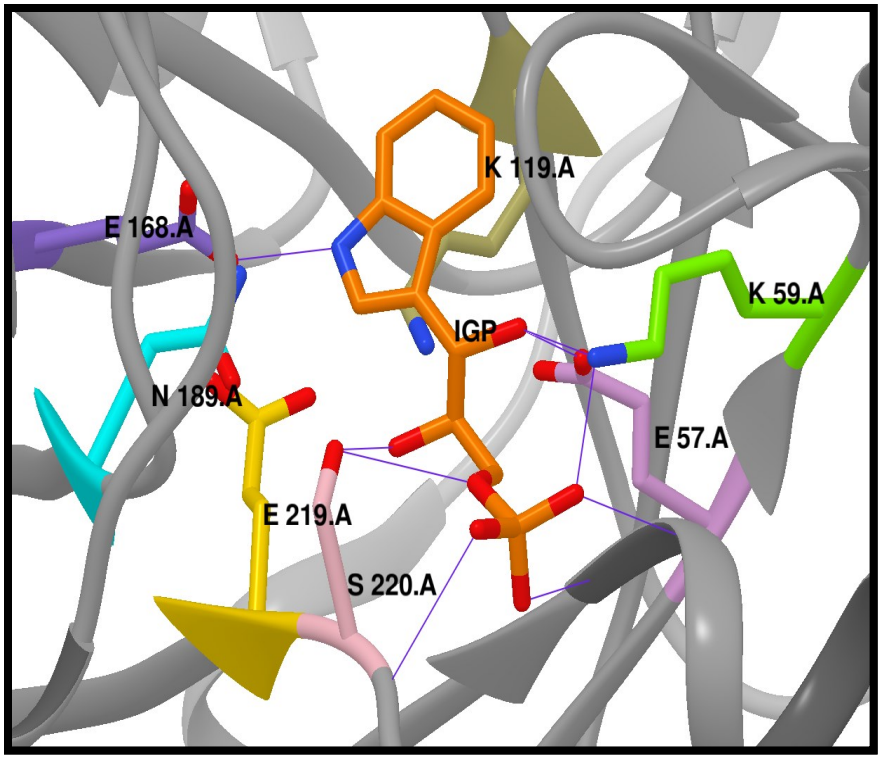


Figure 4: 3-D model of *mt*IGPS mapped with Chimera. IGPS ID was taken from Protein Data Bank (3T44). Colored residues indicate specific targets for Sokol Lab research

By using methods of pH profiling, we can determine where pK_a's lie for *mt*IGPS and whether or not inducing mutations causes a change in those pK_a's. If we see a change in pK_{a1} or pK_{a2}, we can form an idea of which residues act as the catalytic base and acid.

Another large piece of the puzzle missing is where the rate-limiting step lies within IGPS, and whether or not it is maintained with the induction of a mutation. Using experimental methods such as solvent deuterium kinetic isotope effects and solvent viscosity effects allows further understanding of how the enzyme behaves when relations between residues are affected. By manipulating length or presence of H-bonds through selection of certain single-point mutations, it is then likely to gain insight on overall mechanism of IGPS.

Methods

Expression of mtIGPS Wildtype and E57D, N189L Mutants

Past student, Oshane Thomas, transformed *mt*IGPS in a pET30 vector tagged with a His-tag ordered from Genscript into *E.* coli BL21 cells, which were used for expression[8]. Quick Change mutagenesis was used to induce single-point mutations in targeted residues. Cooled 5 mL of autoclaved LB broth was prepared in 15-mL Falcon tubes with 50 µg/mL concentrated kanamycin. *mt*IGPS was inoculated with plastic loops into tubes for culture growth. Wildtype and mutants (N189L, E57D respectively) were cultured separately. A negative culture was used to determine sterility of experiment, which contained only LB broth and kanamycin. Falcon tubes were incubated on a shaker within a time frame of 12-16 hours at 37ºC at a speed of 225 rpm. After no contamination on culture growth was confirmed, all 5 mL of cultures were added to 1.0-L flasks containing 500 mL of 25º C autoclaved LB broth and 50 µg/mL of kanamycin. These cultures were incubated and placed on a shaker at 37º C at 225 rpm. Wildtype and mutant strains were inoculated in separate flasks to an optical density absorbance reading of about 0.600 taken at 600 nm on the spectrophotometer. OD was continuously measured every 20 minutes with a 1-cm cuvette and the machine was blanked with LB broth. In order to ensure that there was no overgrowth, flasks were taken out of heated shaker at an absorbance reading anywhere between 0.500 and 0.550, then placed on ice until the temperature cooled down. They were then induced with 0.1 mM IPTG and placed into the shaker once more at 25º C at 225 rpm for 12-16 hours. After cultures were properly incubated, the 500 mL cultures were centrifuged at 5,000 rpm for 15 minutes at 4º C. The pellet was kept and placed into 50-mL Falcon tubes in preparation for purification. If the pellets were not used for purification in the same day, they were properly labeled and stored in the -80º C freezer.

Gravity flow column-protein purification of mtIGPS Wildtype, N189L, and E57D

15 mL of equilibrium buffer made with 1xPBS at a [10 mM] concentration of imidazole at a pH of 8 was poured into each 50-mL Falcon tube containing the expressed pellets and were resuspended with a vortex. Buffers were kept cold over ice. Resuspended pellets were sonicated on a setting of 3 minutes, 1 second on, 1 second off pulse, with 25% amplification. After cells were lysed through sonication, solution was prepared for centrifugation at 17,000 rpm for 15 minutes at 4^{o} C in order to obtain the supernatant. While cells were centrifuged, columns were set up for the purification procedure. 2 mL of Ni NTA resin was placed into a clean column and washed through with 10 mL of equilibrium buffer. It is important to note that the buffer needs to stay cold, and that resin must not dry out. Once cells were centrifuged fully, 50 µL of supernatant was saved in an Eppendorf tube labeled 'lysate' and the rest was placed into the freshly packed column with a transfer pipette and pellets were discarded properly with bleach. The supernatant was run through the column and 50 µL of this flow-through produced was saved as a fraction in an Eppendorf tube. Wash buffer made with 1xPBS at [40mM] imidazole, [100mM] NaCl, and a pH of 8.0 was then pipetted through the column and tested with the Bradford Assay continuously in intervals of 20 µL until the assay turned up negative. Positive protein turns the reagent blue, whereas a negative fraction does not cause any change in the color of the reagent. After the Coomassie reagent came back negative, 50 µL of washed protein was saved in an Eppendorf tube and elution buffer made with 1xPBS at [250mM] imidazole at a pH of 8.0 was run through the column. Eluted protein was continuously collected in 1.0-mL Eppendorf tubes and tested through Bradford assay. The most concentrated blue wells were indicators of concentrated protein. The most concentrated protein fractions were collected by 1 mL aliquots and mixed together in a dialysis bag for overnight dialysis treatment.

Dialysis of purified Mt. IGPS Wildtype, N189L, and E57D

Dialysis membrane bags needed to be prepped in 2.0 L of pH 7.90 Tris-base dialysis buffer in a cold room to maintain proper temperatures prior to purification. The gathered protein was then placed in dialysis beakers and set on a spinner with a magnetic stirrer for 2 hours. After 2 hours, the dialysis bag was carefully extracted and placed into a second breaker of fresh 1.0 L dialysis buffer for overnight, around 12 to 16 hours. Protein was taken out the next morning and placed in a falcon tube to be measured for determination of concentration with spectrophotometer. If protein concentration was determined to be sufficient, protein was then mixed with 1M DTT along with 10% glycerol for preservation.

Concentration of purified protein

If spectrophotometer readings resulted in an absorbance reading under 0.150, protein was placed into a concentrator and set at 5,000 rpm for 3 minutes. Protein was continuously tested on nanodrop and concentrated in small increments until a sufficient concentration was obtained, in order to ensure the protein did not precipitate out if too concentrated in the filter. Once protein absorbance readings reached above 0.150, protein was then aliquoted according to concentration readings. Absorbances were converted to concentration through Beer's Law; path length was 1 cm, the molar attenuation coefficient was 4595 $M^{-1}cm^{-1}$, and concentration was converted into µM. Aliquots were then stored in -80° C freezer.

SDS-PAGE Gel

All samples to be tested through SDS-PAGE were prepped by adding 14 µL of each sample into 500-µL Eppendorf tubes with 5 µL of 4x LDS sample buffer and 1 µL of [2 mM] DTT. If protein was too concentrated, fractions were diluted by a factor of 10. All tubes were labeled and

placed into a heat bath set at 90° C to heat shock the cells for 5 minutes. The gels were set up into the proper gel apparatus and loaded in the middle and outside chamber with running buffer (1x MES). Gels were then prepped with 5 µL of heat-shocked samples in each well, leaving the first one open for 5 µL of ladder. After all samples were loaded, the gel box was run at 100 V for 45 minutes to an hour, or until the band would fall close to the bottom of each gel. Gels were then stored anywhere from 15 minutes to overnight with instant blue reagent diluted in water.

Kinetic activity readings for Mt. IGPS Wildtype, N189L, and E57D

Collected protein fractions and concentrated proteins were tested with a fluorometer for kinetic activity against its substrate, CdRP, created by Dr. Konas's lab at Montclair State University. CdRP was kept in black Eppendorf tubes to block out light and placed on ice, and protein was prepped in a 1-cm cuvette based on an average of ten separate nanodrop readings. Readings were run with [100 nM] of wildtype IGPS and [1 µM] of mutant respectively. Substrate concentrations per mutant were initially taken at [500uM] prior to finding optimal substrate concentrations through K_m readings. Assay buffer created with [100 mM] PIPES, [2mM] DTT at a pH of 7.5 was also added at [100 mM] in the cuvette per reading. Activity was taken of each reading with a fluorometer at an absorption of 278nm and emission of 340nm over 300 seconds with 5 and 10nm slits as counts per second. Counts per second was a measurement of product formation over a period of 5 minutes. Counts were then calibrated to concentration by using standard activity of CdRP[10]. Temperature was controlled with a water bath at 25° C. Two negative controls were taken testing assay buffer against enzyme only, and then assay buffer against CdRP only. Three trials minimum were taken and averaged for a better indication of k_{cat} and kinetic analysis. Activity was measured as counts per second as a measurement of product formation over a period of 300 seconds.

Measuring K_M for Mt. IGPS Wildtype and E57D

K_M for E57D and WT were measured based off kinetic analysis. WT was kept at [100 μM] and E57D was analyzed at [1.00 mM]. CdRP was analyzed through 8 serial two-fold dilutions ranging from concentrations of 120uM to 0.938uM. Activity was recorded over 300 seconds through 5 and 10 nm slits and plotted through the program Kaleidagraph to be normalized and find a proper K_M along with V_{max}. Three trials were taken of each experiment to average the data. K_M trials were then fitted on Kaleidagraph.

pH profiles for Mt. IGPS Wildtype and E57D

MTEN buffer was made with [50 mM] MES, [25 mM] Tris-base, [25 mM] ethanolamine, [100 mM] NaCl, and [1 mM] DTT at a pH of 7.5 and stored as 4x in fridge, then diluted down to 1x prior to trials being run. 1x tends to degrade so the diluted buffer cannot be used after overnight storage. A series of 1x MTEN buffers were then set at various pH's from pH of 5 to 9 with 0.5 separation between each mixture, and tested in the fluorometer with the same parameters as the kinetic experiments, at a setting of 300 seconds with 5 and 10nm slits. Each experiment tested the WT, mutant, and two negative controls; one without protein, and one without substrate. CdRP was kept at a concentration of [120 μM] while WT was kept at [100 μM], and E57D was kept at [1.00 mM]. pH profiles were then fitted using Kaleidagraph using an equation taken from Zaccardi[5].

Solvent Deuterium Kinetic Isotope Effect

Assay buffer made at [100 mM] PIPES, [2 mM] DTT at a pH of 7.5 was mixed in MilliQ water. A second assay buffer was created with the same concentration of PIPES and DTT but with a pH of 7.1 in deuterium oxide. pH for the deuterium oxide-based buffer was corrected as pD[9].

Wildtype and E57D were tested both in both assay buffers with three trials each. Buffers were maintained at 25° C in a water bath. Wildtype SDKIE was run at [100 nM] WT IGPS, [60 μM] CdRP, and [100 μM] assay buffer. Mutant SDKIE was run at [1 μM] E57D IGPS, [120 μM] CdSRP, and [100 μM] assay buffer. The fluorometer was used for measurements under the settings of 300 second run period with 5 and 10 nm slits (270nm absorbance, 340nm emission).

Solvent Viscosity Effect

Assay buffer was created at [100 mM] PIPES, [2 Mm] DTT, and a pH of 7.5 in water with varying degrees of viscogen added. A set of 0%, 10%, 20%, and 30% glycerol (by weight) were added to buffer in separate falcon tubes. All buffers were kept in a water bath set at 25° C and both WT and E57D variants were tested in all four buffers to measure kinetic activity in the fluorometer. Settings were kept at 300 seconds, with 5 and 10 nm slits, and WTIGPS was run with the parameters [100 nM] IGPS, [60 μM] CdRP, [100 μM] PIPES buffer whereas E57D was run with the parameters [1 μM] IGPS, [120 μM] CdRP, and [100 μM] PIPES buffer.

Results and Discussion

Expression and purification of WT and E57D variants of mtIGPS with dialysis and concentration

Expression was first run with wildtype *mt*IGPS, and then later with E57D. Cells were grown and expressed up to an optical density of 0.600 tested as absorbance on the spectrophotometer. There was trouble maintaining a certain amount of protein necessary to test activity, where loss of protein occurred somewhere between expression, purification, dialysis, and concentrating protein, as indicated below in **Figure 5** labeled E57D batch #4. The minimum absorbance measured necessary for substantial detection on the nanometer is 0.15, with the extinction coefficient 4595 $M^{-1}cm^{-1}$. Many efforts were made in the beginning of the journey to obtain a proper concentration of protein however, protein was lost during expression and purification. SDS-PAGE gels were used as an assay to see where protein was being lost (Fig. 5), and if the problem occurred during dialysis or concentration.

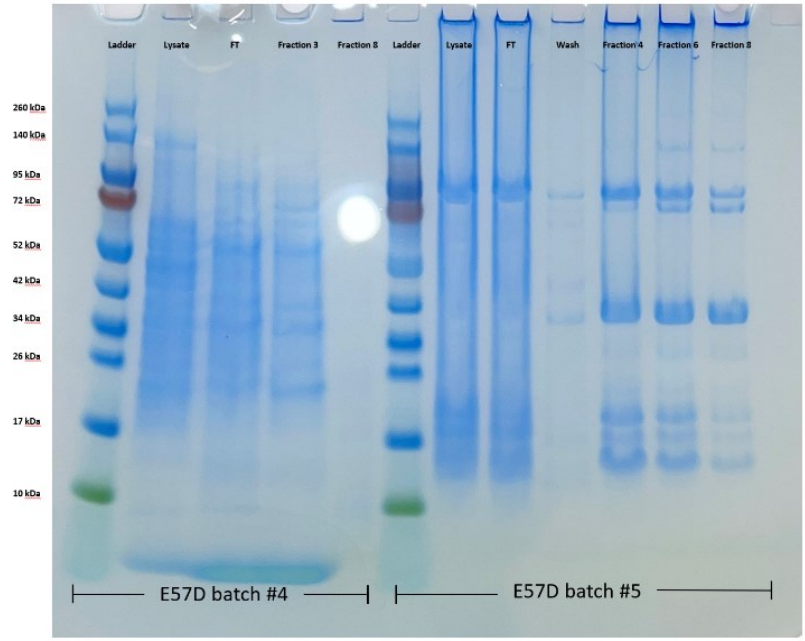


Figure 5: SDS-PAGE gel on *mt*IGPS E57D. IGPS should be indicated by a strong band around 33.5 kDa. Dimer shows a strong band around 68.0 kDa. SDS-PAGE gel of two different batches of E57D IGPS. From left to right: ladder, lysate, flowthrough, and 1.0mL samples of fraction 3 and 8, ladder, lysate, flowthrough, wash, and fractions 4, 6, and 8 of batch 5. Positive protein is indicated by a singular band around 33.5kDa.

A major issue that halted obtaining protein was overgrowth during expression. As seen in Fig. 5, E57D batch 4 shows no protein maintained during purification after three other failed batches. To overcome this issue, IPTG was induced into the culture once density reached anywhere between 0.500-0.550, and to help ensure that the threshold of an OD of 0.600 was not surpassed, temperature was dropped by placing the cultured flask on ice before placing it back on the shaker overnight. More importantly, our gels continuously indicated a strong dimer around 68.0 kDa, so it was important for us to include DTT into all our buffers to break these dimers apart in our protein during studies.

Analysis of K_M, k_{cat}, and efficiency of WT and E57D variants of mtIGPS

The first question we wanted to answer with our experiments was whether or not our residue, Glu57, was catalytically significant to *mt*IGPS. In order to measure this, it was important to compare k_{cat} values, along with efficiency and K_M as indicated in **Table 1**. Once activity was measured in the fluorometer (**Figure 6**), k_{cat} was calculated using data of relative activity with the following equation:

Equation 1

$$k_{cat}\ (s^{-1}) = V_{max}\ (\text{CPS measured in } \mu M/s) \,/\, [E_t]\ (\mu M)$$

Variant	Temp (°C)	k_{cat} (s^{-1})	K_M (µM)	k_{cat}/K_M ($M^{-1}s^{-1}$)	SDKIE	SVE (30:0)%
WT	25	4.8±0.3	5.0±1.4	9.7×10^5	1.26±0.15	0.82±0.03
E57D	25	0.102±0.018	29±6	3.4×10^3	0.63±0.11	0.49±0.04
N189L	25	n.d	n.d	n.d	n.d	n.d

Table 1: Kinetic values determined from experimental data for WT, E57D, and N189L mtIGPS respectively. Not enough results were found for N189L to indicate that there was any activity. For WT, experiments were controlled at a temperature of 25°C. K_{cat}/K_M signifies efficiency of the protein. SDKIE and SVE were tests used to analyze where the rate-limiting step lied for each variant

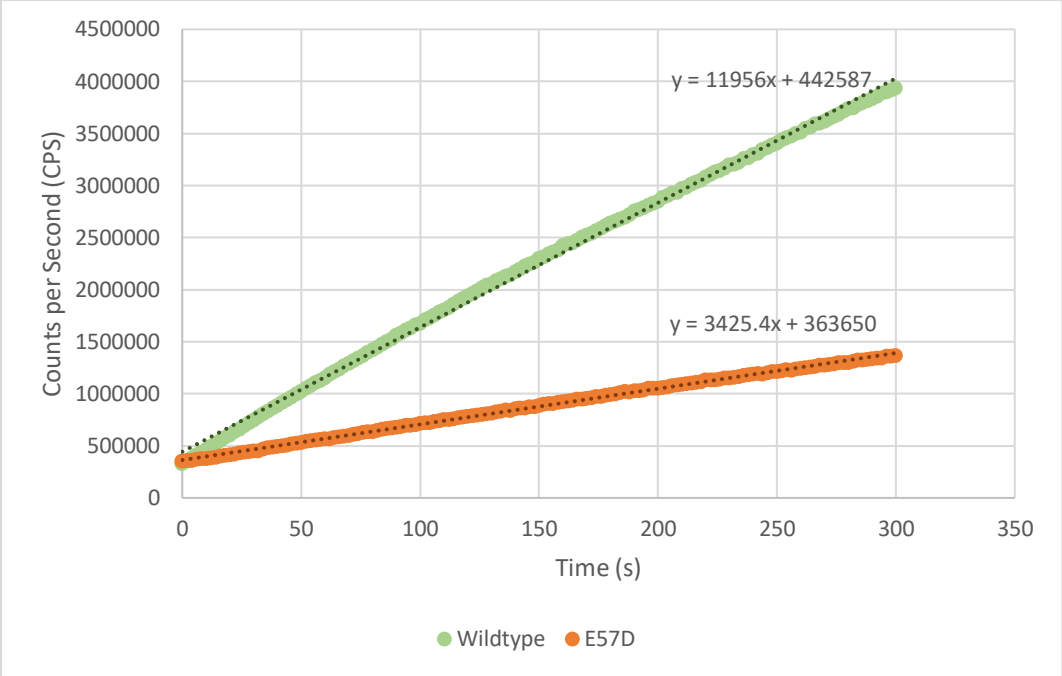


Figure 6: Activity measured for determination of k_{cat}. Green data points represent WT *mt*IGPS, with a velocity of 11,956 CPS/s, and orange data points represent E57D *mt*IGPS activity, with a velocity of 3,425.4 CPS/s. *x*-axis represents time on a scale of 300s and *y*-axis represents counts of products based on fluorescence measured

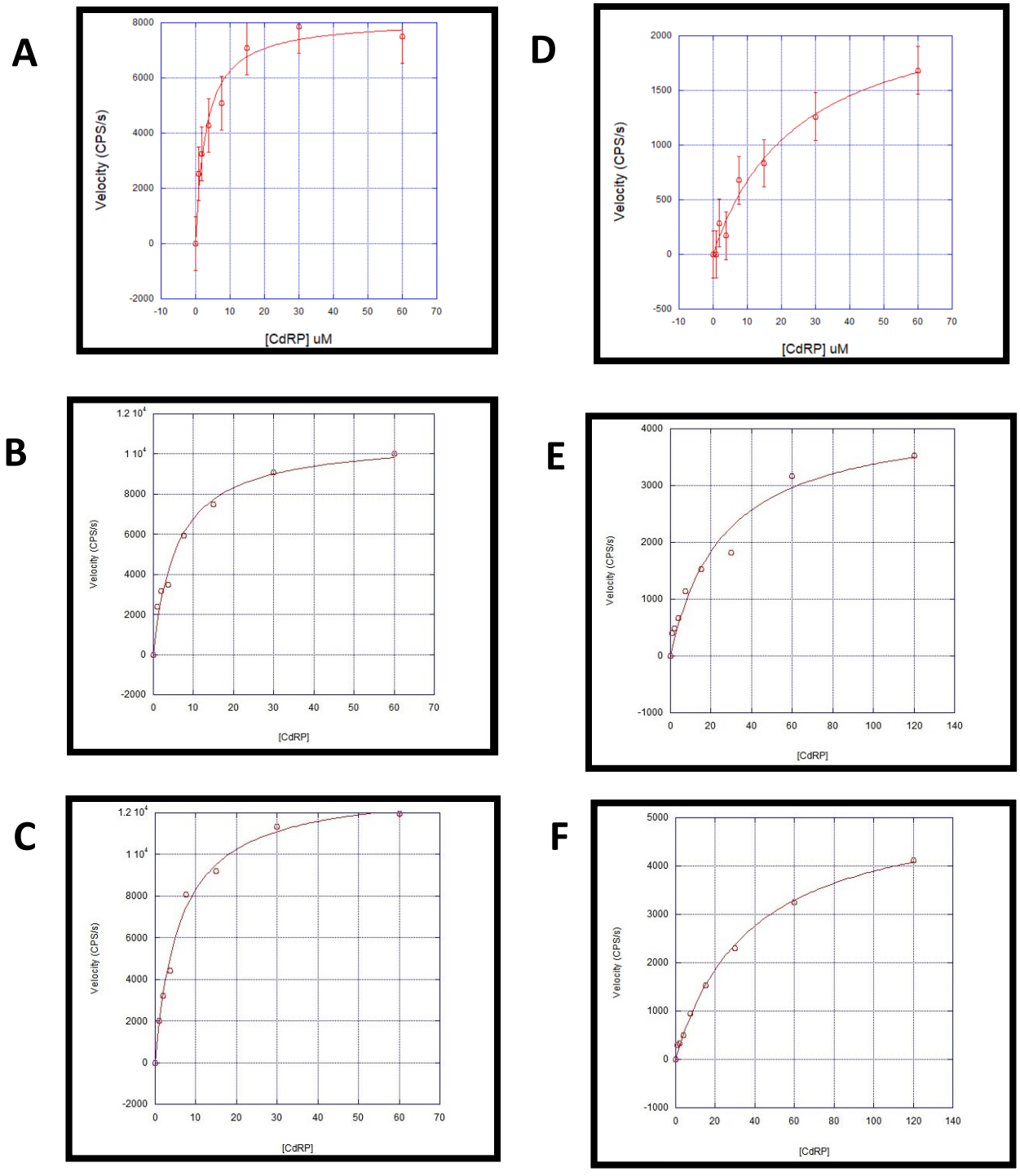


Figure 7 (a-f): K_M trials. Figures a-c indicate K_M trials run for WT IGPS where optimal kinetic activity is indicated at 60µM. K_M is averaged to be 5.0±1.4 µM of CdRP. Assay was taken with [100nM] IGPS, [100µM] PIPES, and various concentrations of substrate diluted twofold. Figures d-f indicate K_M trials run for E57D IGPS where optimal kinetic activity is indicated at 120µM of CdRP. An average K_M of 29±6 is found for the triplicates. Mutant was measured in assay with [1µM] of IGPS, [100µM] of PIPES buffer, and various amounts of substrate diluted twofold.

As represented in Table 1, there is an estimated 50-fold decrease of activity from the original WT to mutant E57D *mt*IGPS. This tells us that Glu57 is significant to IGPS in terms of catalytic activity. N189L mutant did not give a signal similar enough to the positive control (WT) to determine any calculatable data.

K_M of each variant was tested through a series of two-fold diluted CdRP samples (**Figure 7**). K_m data was then normalized on the program Kaleidagraph and as shown in Table 1, there is a 6-fold difference between WT and E57D *mt*IGPS values, indicative that mutant E57D requires a higher concentration of substrate to optimally catalyze the reaction. The first few experiments included testing of up to 240μM of CdRP against the enzymes however, a surprising trend was observed, where a Michaelis-Menten curve was only obtainable below a certain amount of substrate per enzyme variant. Testing enzyme against too large of an amount of substrate compared to enzyme concentration used caused the activity to slow down, which caused a downward curve past the V_{max} as seen in preliminary studies shown in **Figure 8**. Though not shown, the same trend was seen for our WT *mt*IGPS as well. The fact that there is a slowing of activity with higher substrate concentration leads to the hypothesis that product IGP may affect the binding affinity between CdRP and IGPS and may act as a possible inhibitor to enzyme-substrate binding.

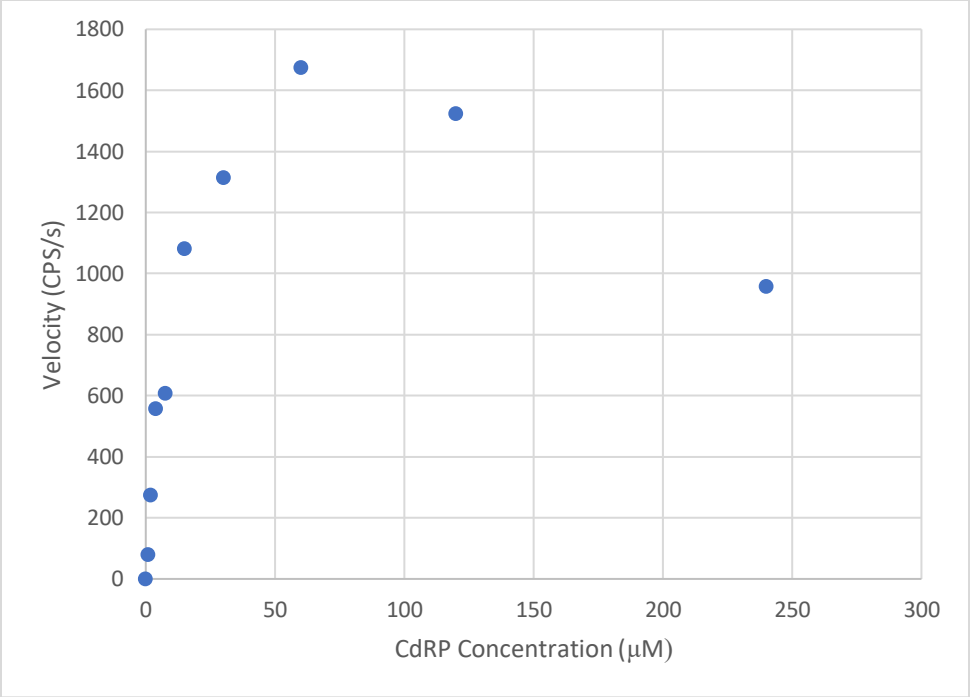


Figure 8: Preliminary K_m run of E57D *mt*IGPS with the set parameters of [1µM] IGPS and [100µM] PIPES buffer, and a range of CdRP from [0.9375-240µM]

By taking a twofold dilution from 240µM to 0.938µM, it was visible within triplicates that WT IGPS met a maximum velocity of activity around 60µM before activity started slowing as seen in **Fig. 7 (a-c)**. E57D IGPS however, met a maximum velocity of activity around 120µM as seen in **Fig. 7 (d-f)**. We assume the increase of substrate necessary to reach optimal activity to be due to the mutation being not as adept to binding to CdRP as the original enzyme. However, similar to the wildtype variant, we failed to see a steady plateau of maxed out activity past a larger concentration of substrate used and instead saw slowing of activity with twofold of CdRP, specifically shown in Fig. 8. The optimal activity taken from each respective variant's K_M study was used that point on as a parameter in further studies; wildtype was tested against 60µM of CdRP and E57D was tested against 120µM of CdRP.

pH profiles of WT and E57D mtIGPS

The k_{cat} and K_M of each variant indicated that Glu57 was catalytically significant, and the next step was to find out how so. This could be tested using pH profiles to find more information as to whether or not Glu57 acts as the catalytic base in *mt*IGPS the way Glu51 is proposed to act in the mechanism proposed for *ss*IGPS. pK_a values were obtained as shown in **Table 2**.

Variant	Optimal pH	pK_{a1}	pK_{a2}
WT	8.00	6.80±0.14	8.68±0.06
E57D	7.00	6.53±0.10	7.68±0.10

Table 2: Table of *mt*IGPS pK_a values from pH profiles conducted. Values are averaged from triplicates and error is taken from standard deviation. T-test conducted reports p-value of 0.115 for pK_{a1} comparison between variants; result is not significant at $p < 0.5$, and a p-value of 0.000051 for pK_{a2}; the result is significant at $p < 0.5$

The pH profiles of both wildtype and E57D variants were taken in triplicates, as seen in **Figure 9**. Profiles were fit using Zaccardi's equation[5]:

Equation 2:

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

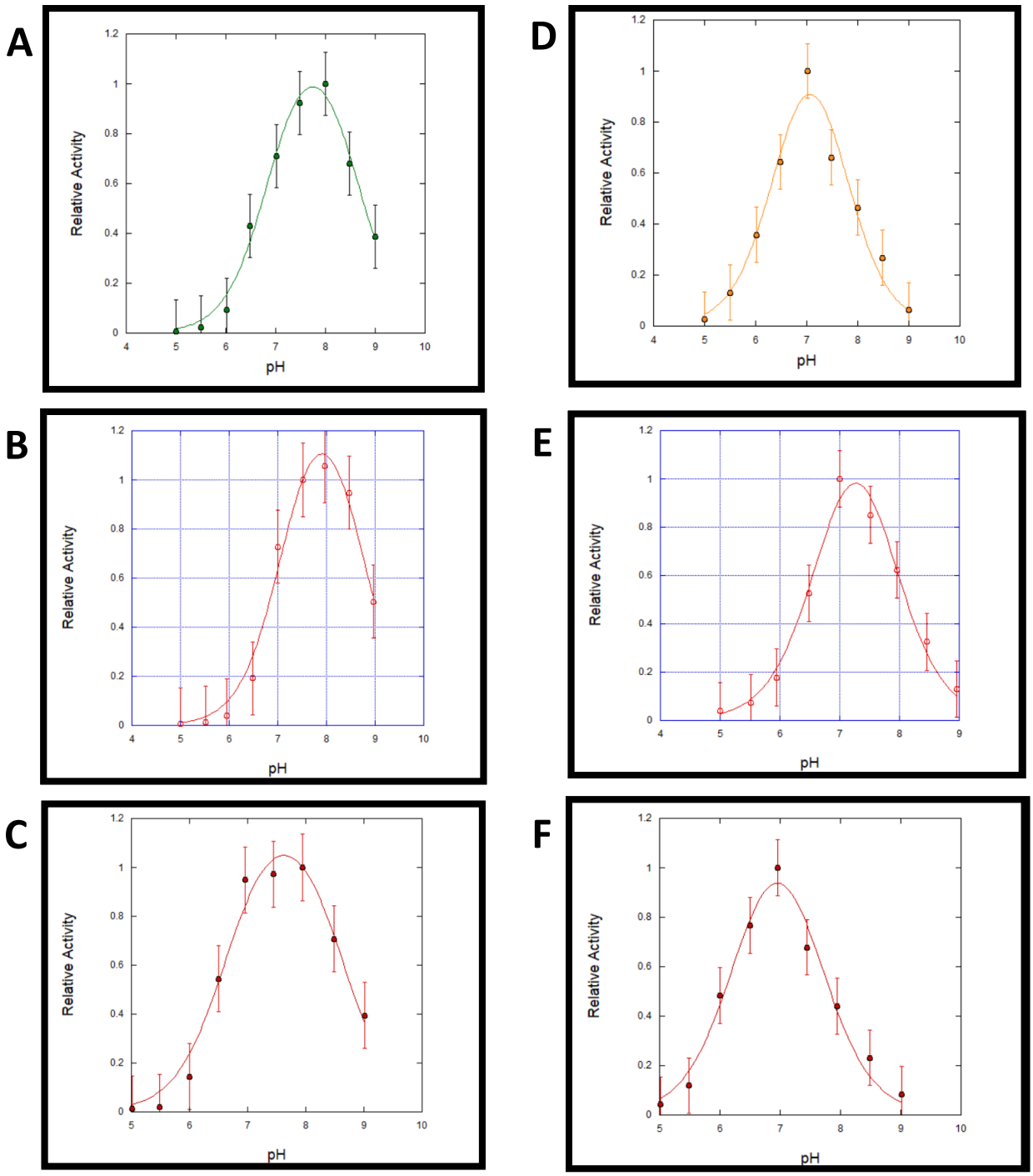


Figure 9: pH profiles of *mt*IGPS variants. (a-c) represent triplicates of WT *mt*IGPS, and trials were run with [100 nM] IGPS, [60 µM] CdRP, and [100 µM] PIPES assay buffer set at various pH's ranging from 5.0-9.0 with 0.5 increment increase. (d-f) represent triplicates of E57D *mt*IGPS and trials were run with [1µM] IGPS, [120 µM] CdRP, and [100 µM] PIPES assay buffer set at various pH's ranging from 5.0-9.0 with 0.5 increment increase. Error bars represent standard error

In previous literature[5], WT *ss*IGPS indicated an optimal activity when measured in an assay buffer with a pH of 8.0 which was replicated in our studies with *mt*IGPS, as shown in **Figure 9 (a-c)**. The trend seen with the ascending limb and descending limb also agreed with past studies done on *ss*IGPS (pK_{a1} = 7.5±0.2 and pK_{a2} = 8.8±0.3 at 37° C)[7], where a steady influx on the ascending limb showed a pK_{a1} of 6.80±0.14 and a pK_{a2} of 8.68±0.06 (**Table 2**). This indicates that the catalytic base for IGPS was half-protonated at the measured pK_{a1} and the catalytic acid was half-protonated at pK_{a2}.

For the mutant E57D, optimal activity had shifted to the left at a pH of 7 in comparison to the optimal activity seen at a pH of 8 for wildtype. pK_{a1} of 6.53±0.10 and a pK_{a2} of 7.68±0.10, both having shifted to the left for the mutant compared to WT *mt*IGPS which supports the hypothesis that Glu57 acts as a catalytic base for *mt*IGPS. The shift in pK_{a1} indicated that the 57th residue acted as a base, more specifically, and although the shift was slightly smaller than seen in the shift for pK_{a2}, this was supported by the specific mutation implemented from glutamate to aspartic acid. Glutamic acid has a pK_{a3} of 4.07 and aspartic acid has a pK_{a3} of 3.90. These two pK_a's have around 0.2 difference between them, similar to the difference we see in our measured pK_{a1}'s, 6.80 and 6.53 for our wildtype and mutant respectively. This can be used to further explain the difference in catalytic activity between the wildtype and mutant because the carboxylic acid of each amino acid's side chain was conserved, however the shift supports the hypothesis that the 57th residue acts as the catalytic base due to observed shifts accordingly with the single-point mutation.

Another significant finding is that the second pKa had shifted within the descending limb to the left as well when glutamate is mutated into aspartic acid. By looking at past literature, Zaccardi

had proposed a schema where either residue K53 or K110 act as possible catalytic acids for *ss*IGPS[6].

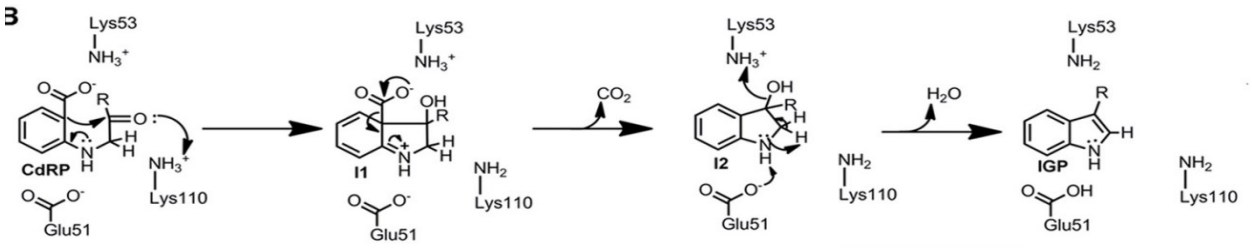


Figure 2: mentioned within introduction

When mapped out, as seen in **Figure 10** represented below, for *mt*IGPS, K110 in *ss*IGPS is conserved as K119, which is depicted in a 3-D model to have interactions between the residue's functional group to E57 residue's functional group.

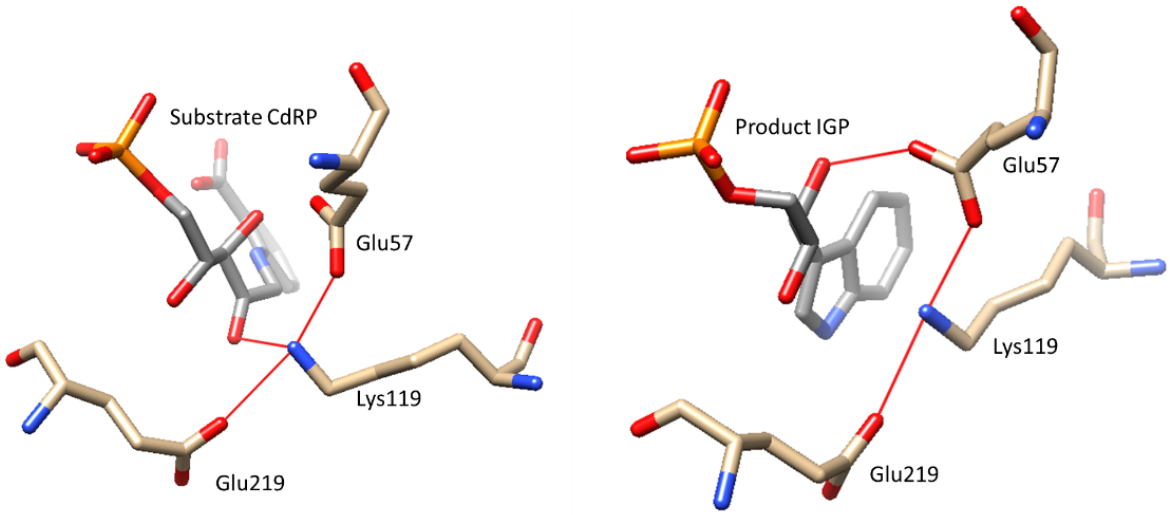


Figure 10: 3-D model of *mt*IGPS interactions between residues. Red indicates oxygens, and blue represents nitrogen. The red lines between residues are hydrogen interactions within IGPS as well as with CdRP, and IGP

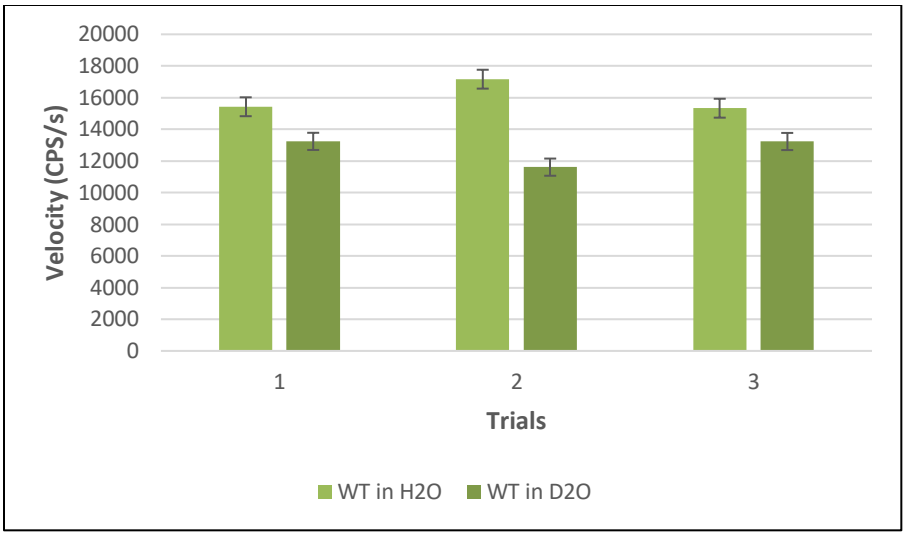
It can be presumed then that if there is a shortening of glutamate to aspartic acid, the interaction between E57 and K119 would weaken. However, we would still expect to see some sort of

interaction due to the conservation of the carboxylic acid. If lysine 119 is in fact the catalytic acid for IGPS, it would explain why there is also a shift in the pKa_2 of our pH profile plotted for the mutant.

Solvent deuterium kinetic isotope effect at room temperature (<25 º C) vs. 25 º C for WT and E57D mtIGPS variants

After finding supporting data that Glu57 may act as the catalytic base for *mt*IGPS, it was then significant to use a series of experiments to characterize the rate-limiting step for *mt*IGPS, and whether or not there is a change in rate-limiting step when the protein is mutated at the catalytic base. By using solvent deuterium kinetic isotope effect, it was possible to see whether or not the protein's rate-determining step was involved in proton transfer, and therefore isotope-sensitive or not. Results are shown in **Figure 11 & 12** for wildtype and E57D mutant respectively.

A



B

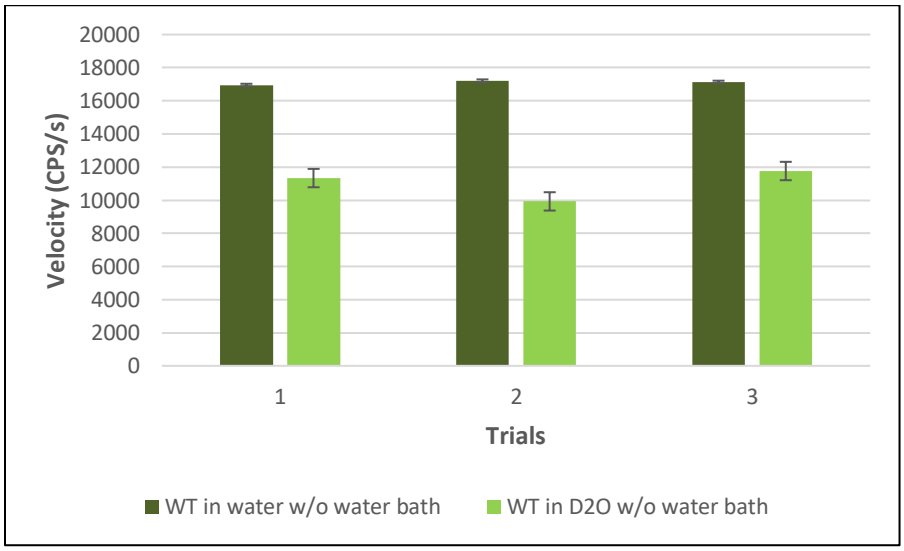


Figure 11 a & b: WT IGPS SDKIE experiments under different temperatures with the parameters of [100nM] IGPS, [60µM] CdRP, and [100µM] PIPES buffer. H_2O-based PIPES buffer's pH was kept at 7.5, and D_2O-based PIPES buffer was kept at a pH of 7.1 A.) Solvent deuterium kinetic isotope effect measured in a controlled 25° C water bath comparing IGPS measured in deuterium oxide to protein measured in water. B.) SDKIE measured in non-controlled environment, where room temperature was measured to be around 4° C less than the controlled water bath

3

Wildtype 25° C				
Runs	1	2	3	average
H_2O	1.5E+04	1.7E+04	1.5E+04	1.6E+04
D_2O	1.3E+04	1.2E+04	1.3E+04	1.3E+04
Scale factor	1.2	1.5	1.2	1.3

4

Wildtype room temp				
Runs	1	2	3	average
H_2O	1.7E+04	1.7E+04	1.7E+04	1.7E+04
D_2O	1.1E+04	9.9E+03	1.2E+04	1.1E+04
Scale factor	1.50	1.73	1.46	1.55

Tables 3 & 4: Table 3 indicates the calculated SDKIE of WT IGPS when conditions are kept at 25° C, with a ratio of activity from water to deuterium oxide of 1.26±0.15. Table 4 indicates the calculated SDKIE value of WT IGPS when conditions were kept under room temperature (< 25° C). The ratio of activity of WT from water to deuterium oxide measured to be 1.55±12.

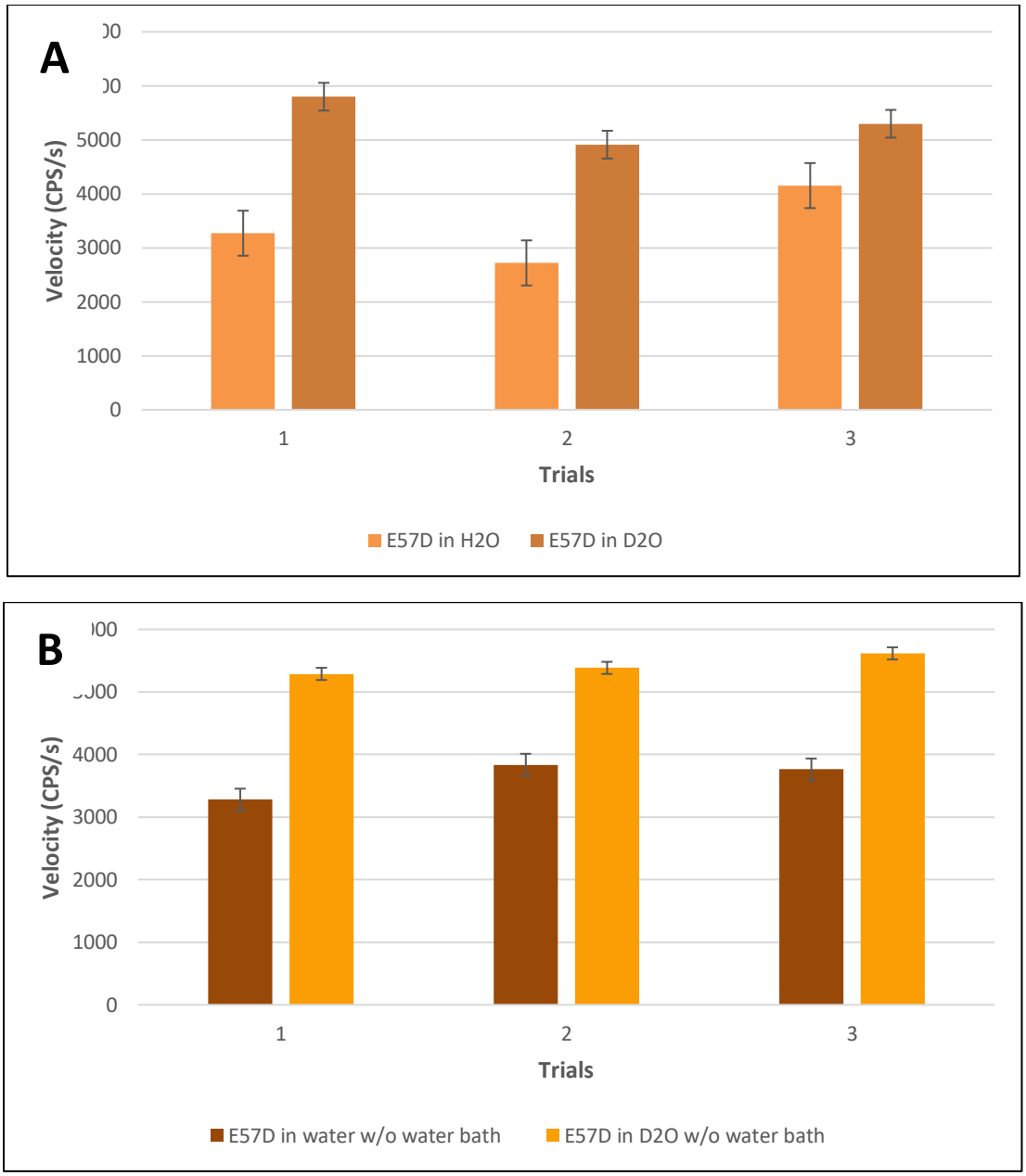


Figure 12 a & b: E57D IGPS SDKIE experiments under different temperatures with the parameters of [1 µM] IGPS, [120µM] CdRP, and [100µM] PIIPES buffer. H_2O-based PIPES buffer's pH was kept at 7.5, and D_2O-based PIPES buffer was kept at a pH of 7.1 A.) Solvent deuterium kinetic isotope effect measured in a controlled 25° C water bath comparing IGPS measured in deuterium oxide to protein measured in water. B.) SDKIE measured in non-controlled environment, where room temperature was measured to be < 25° C

5

E57D 25° C				
Runs	1	2	3	average
H_2O	3.3E+03	2.7E+03	4.2E+03	3.4E+03
D_2O	5.8E+03	4.9E+03	5.3E+03	5.3E+03
Scale factor	0.56	0.55	0.78	0.63

6

E57D room temp				
Runs	1	2	3	average
H2O	3.3E+03	3.8E+03	3.8E+03	3.6E+03
D2O	5.3E+03	5.4E+03	5.6E+03	5.4E+03
Scale factor	0.62	0.71	0.67	0.67

Tables 5 & 6: Table 5 indicates the calculated SDKIE of E57D IGPS when conditions are kept at 25° C, with a ratio of activity from water to deuterium oxide of 0.63±0.11. Table 6 indicates the calculated SDKIE value of E57D IGPS when conditions were kept under room temperature (<25° C). The ratio of activity of WT from water to deuterium oxide measured to be 0.67±0.04.

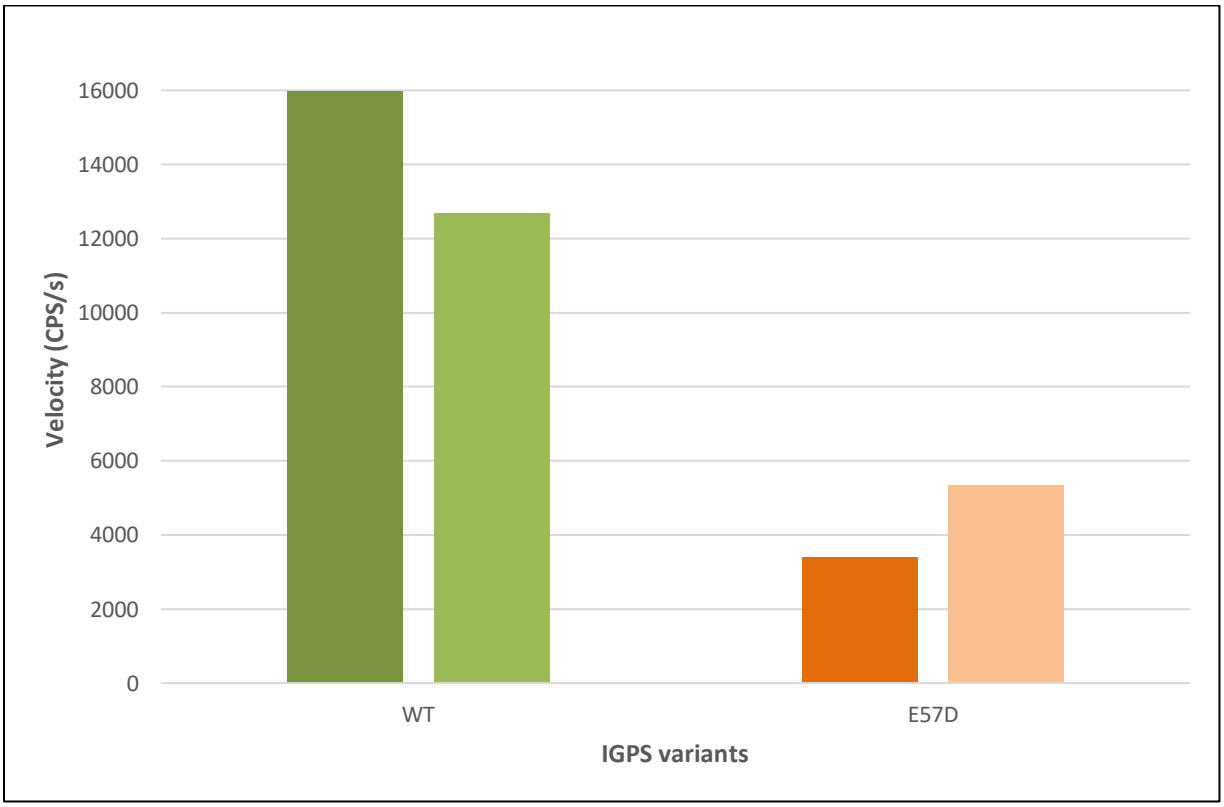


Figure 13: Solvent deuterium kinetic isotope effect comparison of WT IGPS to E57D IGPS. Average values for triplicates of SDKIE runs taken at a controlled environment of 25°C for the two variants of IGPS are plotted. H_2O-based PIPES buffer's pH was kept at 7.5, and D_2O-based PIPES buffer was kept at a pH of 7.1

There were a couple of issues that the research team ran into when first trying to optimize the SDKIE procedure. In trying to replicate original results provided not only by Thomas[10] but also by Zaccardi, several conditions were changed. DTT was made in deuterium oxide, and pH was corrected to measure for pD instead[9]:

Equation 3:

$$pD = pH + 0.4$$

In order to ensure accurate and more reproducible results, we maintained a consistent temperature of 25° C by using a water bath during our rate-limiting step determination experiments.

Several results were observed; when conditions were kept at 25 °C but pH was measured at 7.5 for *both* H_2O- and D_2O-based PIPES buffer, an inverse solvent kinetic isotope effect was observed (**Figure 14**), where activity of IGPS is higher in D_2O than H_2O.

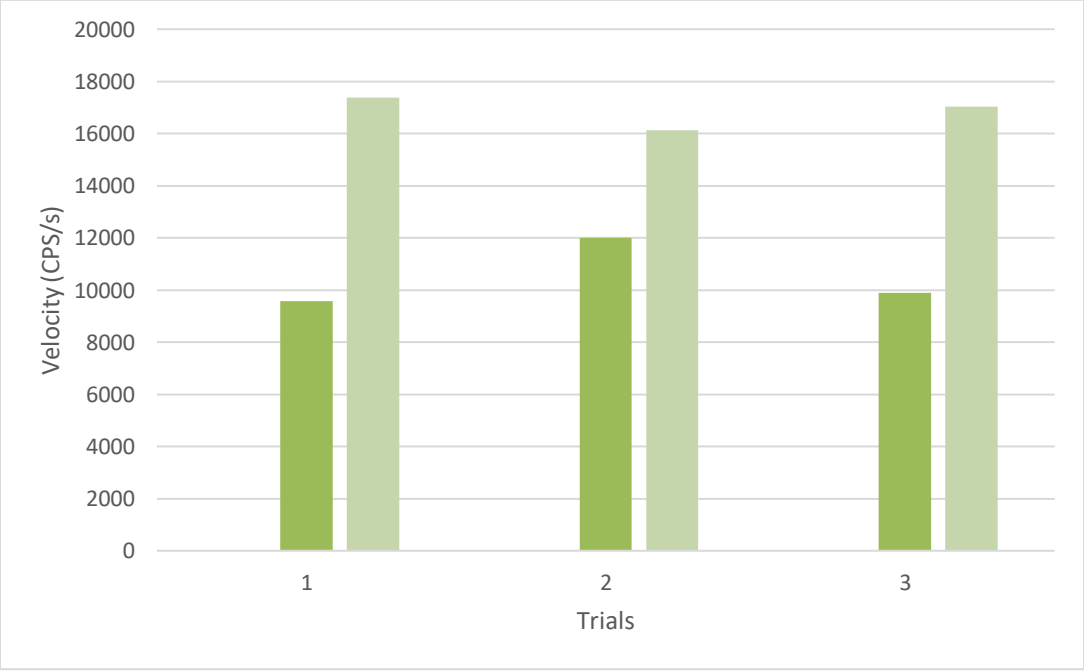


Figure 14: solvent deuterium kinetic isotope effect of WT IGPS with following parameters: [100nM] IGPS, [120μM] CdRP, [100μM] PIPES buffer (H_2O-based PIPES buffer was kept at a pH 7.5, and D_2O-based PIPES buffer was kept at a pH 7.1)

Once temperature and pH were sufficiently kept at necessary conditions, an inverse effect as expected was observed. Interested in repeating Thomas's non-controlled environment[10] to observe effects against a temperature-controlled environment, I chose to take SDKIE triplicates where the water bath was controlled at 25 ºC and another set of SDKIE triplicates where buffer was left out in room temperature. When testing both with the thermometer once buffers were inside the cuvette, the buffers kept in a 25 ºC water bath read to be ~25±0.6 ºC inside the cuvette. The buffers kept out at room temperature however, when directly pipetted inside the cuvette, gave temperature readings of ~21±0.7 ºC, about a 4 ºC drop from the controlled buffers. When buffers were slightly colder inside the cuvette, SDKIE readings were even more significant as seen when comparing scale factors between **Tables 3 and 4**.

When the same experiment was run with E57D IGPS, we observed different results, where an inverse effect was instead visible (**Figure 12**). For both 25 ºC and room temperature (<25 ºC) triplicates, the mutant returned results of an inverse kinetic isotope effect. Although rare, there are many explanations as to why this may be observed in certain experiments, none of which fit the profile of our protein. These included possibilities of cysteine-thiols in protein, the use of metal-bound water, and low-barrier hydrogen bonds[9].

Solvent viscosity effect on WT and E57D mtIGPS

Solvent viscosity effect was used as another indicator to determine where the rate-limiting step may lie for *mt*IGPS variants. If there was a solvent viscosity effect seen, then it supported the hypothesis proposed originally by Zaccardi[7] that the rate-limiting step is found either in enzyme-substrate-binding or product release. Data is shown in **Figure 15 & 16**.

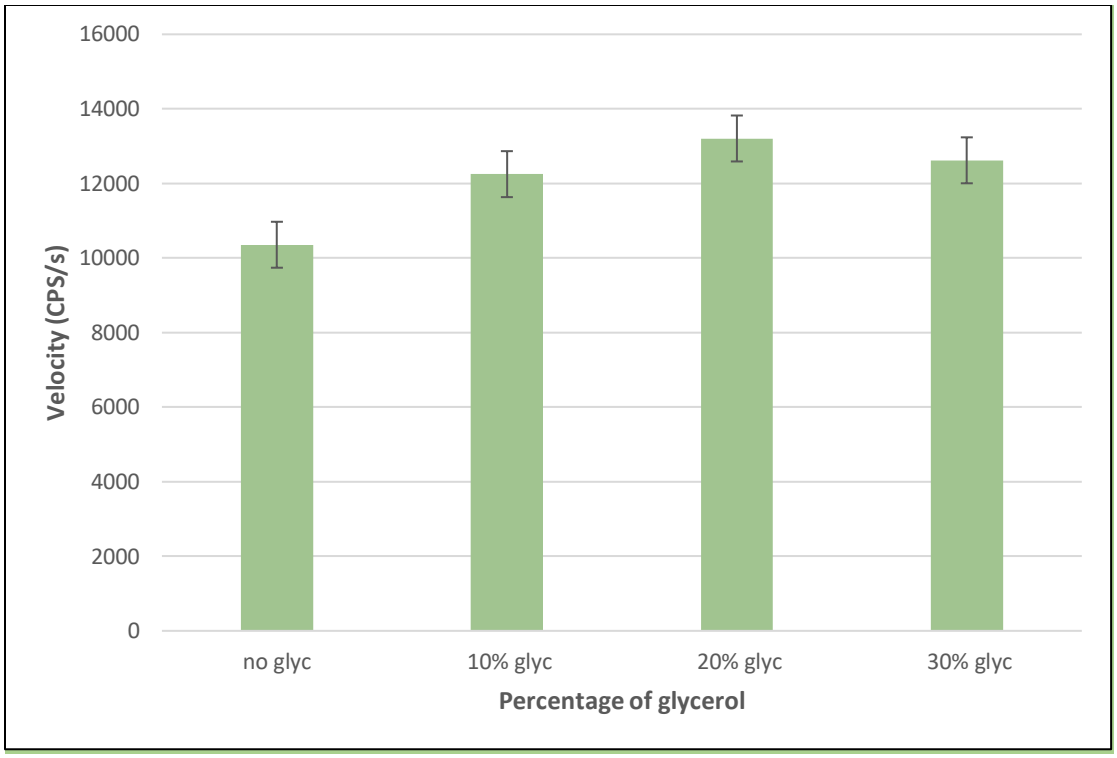


Figure 15: Solvent Viscosity Effect on WT IGPS with the parameters [100nM] IGPS, [60µM] CdRP, and [100µM] PIPES buffer averaged from triplicate run. Enzyme was tested with varying degrees of viscogen mixed in buffer ranging from 0-30% glycerol with a 10% degree of difference between each run.

WT				
Trials	**1**	**2**	**3**	**averages**
0%	1.0	1.0	1.0	1.0
10%	0.92	0.76	0.88	0.85±0.07
20%	0.763	0.796	0.793	0.784±0.015
30%	0.81	0.86	0.80	0.82±0.03

Table 7: Ratio of activity between various amounts of viscogen-based buffer to no (0%) viscogen added-buffer for WT IGPS. Standard deviation was taken of the three trials for calculated averages.

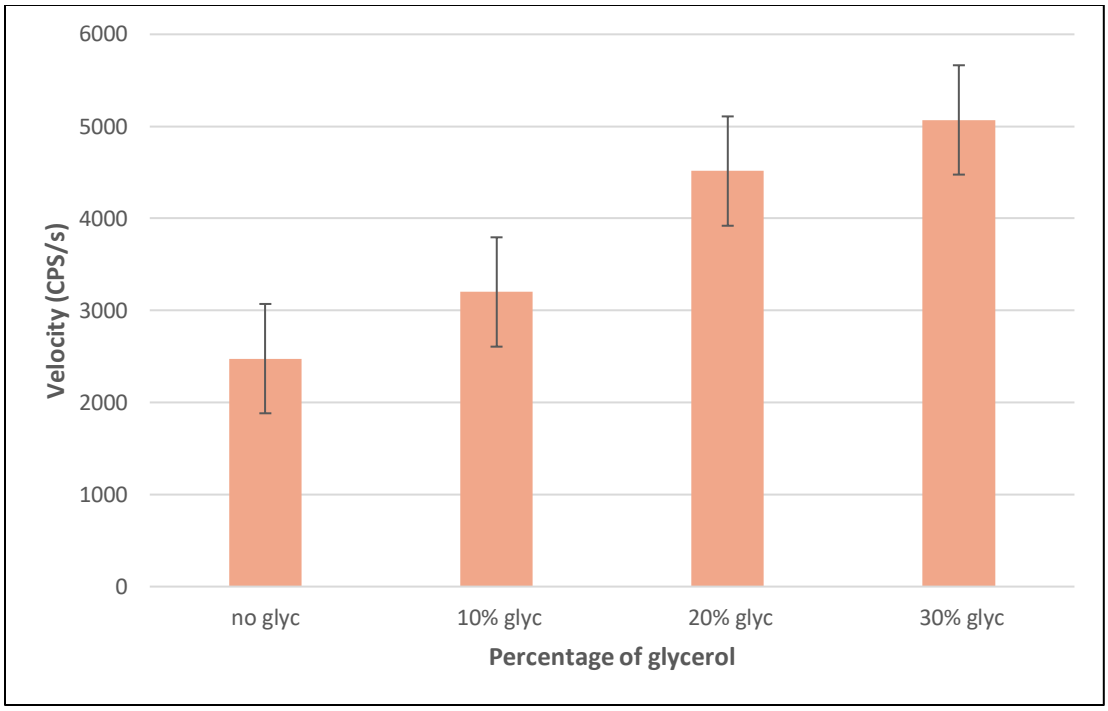


Figure 16: Solvent Viscosity Effect on E57D IGPS with the parameters [1µM] IGPS, [120µM] CdRP, and [100µM] PIPES buffer averaged from triplicate run. Enzyme was tested with varying degrees of viscogen mixed in buffer ranging from 0-30% glycerol with a 10% degree of difference between each run.

E57D				
Trials	**1**	**2**	**3**	**averages**
0%	1.0	1.0	1.0	1.0
10%	0.790	0.753	0.781	0.775±0.016
20%	0.47	0.60	0.60	0.56±0.06
30%	0.52	0.51	0.44	0.49±0.04

Table 8: Ratio of activity between various amounts of viscogen-based buffer to no (0%) viscogen added-buffer for E57D IGPS. Standard deviation was taken of the three trials for calculated averages.

Based on Zaccardi, we chose to measure IGPS against buffers with viscosities ranging from 0-30% glycerol added[6]. The suggestion of using other types of viscogen was discussed however when reading past literature, sucrose was used as another viscogen and results were extremely similar for both glycogen and sucrose[6]. When wildtype was tested against 10%, 20%, and 30% viscous buffers, results were similar across the board, represented by **Figure 15** and **Table 7**.

Based on calculated ratio of 0% viscogen to when viscogen is added per trial, there is no reason to believe that there is any significant SVE for WT IGPS.

The E57D mutant gave unexpected results. As seen in **Figure 16**, there was an influx of activity with more viscogen added. According to classical mechanics, we would expect to see a decrease in activity between enzyme and substrate the more viscogen is added to buffer. However as seen in the bar graph of Fig. 16, there was instead a steady increase the more viscous the solution becomes.

Analysis of rate-determining step from SDKIE and SVE results for WT and E57D mtIGPS

As proposed by Zaccardi in the schema shown below (**Figure 17**[6]), SDKIE and SVE can be used to determine where the rate limiting step lies within each variant. If there is a positive SDKIE but no SVE, then the protein is considered to be isotope-sensitive and therefore its rate-limiting step is suspected to be where proton transfer occurs, during the formation of intermediates. If there is a positive SVE but no SDKIE, then it is assumed that the protein is viscosity-sensitive and therefore its rate-limiting step lies within either enzyme-substrate binding or product release.

Figure 17: Mechanism proposed by Zaccardi to determine where rate-limiting step may lie based on both SDKIE or SVE results. If there is a positive SVE but no SDKIE, then the protein is considered to be viscosity-sensitive and the rate-limiting step is presumed to be in either enzyme-substrate binding or products release. If there is a positive SDKIE but no SVE observed, the protein is considered to be isotope-sensitive, and the rate-limiting step is presumably in step 2, where proton transfer occurs

When analyzing the results of WT IGPS, we see that there is a clear SDKIE and no SVE, which agrees with past literature findings[6]. We can then presume that IGPS when not mutated has a rate-limiting step within the formation of intermediates where the transfer of protons is involved.

When analyzing the data from our mutant E57D IGPS however, it is peculiar to see that there is an inverse effect for both experiments. We see that there is a clear inverse SDKIE, and an inverse SVE. Deuterium oxide is known to be heavier than water, and when looking at the viscosity, deuterium is said to be about 24% heavier, or more viscous than water[9]. This would explain why we see an increase of activity during our SDKIE studies. What we are really seeing, is not so much an inverse SDKIE, or any solvent kinetic isotope effect. In actuality, we are seeing the viscosity effect repeated in our SDKIE study, where we see an increase of activity only due to the fact that the solution E57D is measured in is heavier than regular assay buffer by around the same amount used to test how viscogen-sensitive the enzyme is. We can then presume that there is no SDKIE hence, our mutant is not actually isotope-sensitive, and so the rate-limiting step must lie somewhere within enzyme-substrate binding or product release. This also agrees with our K_M results when comparing wildtype to E57D mutant, where the K_M increased when glutamate was mutated into aspartic acid. This supports further that changing the 57^{th} residue affects the enzyme-substrate binding. As mentioned beforehand, there is an inverse SVE observed instead of a normal SVE, where we would expect the activity to decrease the more viscous the assay buffer gets, as seen in past studies conducted by Zaccardi's team[6]. Perhaps the inverse SVE is a key in giving us more information on determining between whether the rate-limiting step lies in step 1, during enzyme and substrate binding, or in the final step, once product is created and released from enzyme.

Temperature-controlled SDKIE and SVE to further understand behavior or WT and E57D mtIGPS variants

In order to continue characterization of *mt*IGPS variants, it was then important to take findings from activity affected by temperature and create a temperature assay in order to further isolate data in the rate-limiting step determination experiments. Preliminary temperature assays were conducted on wildtype *mt*IGPS and results are shown in **Figure 18 & 19**.

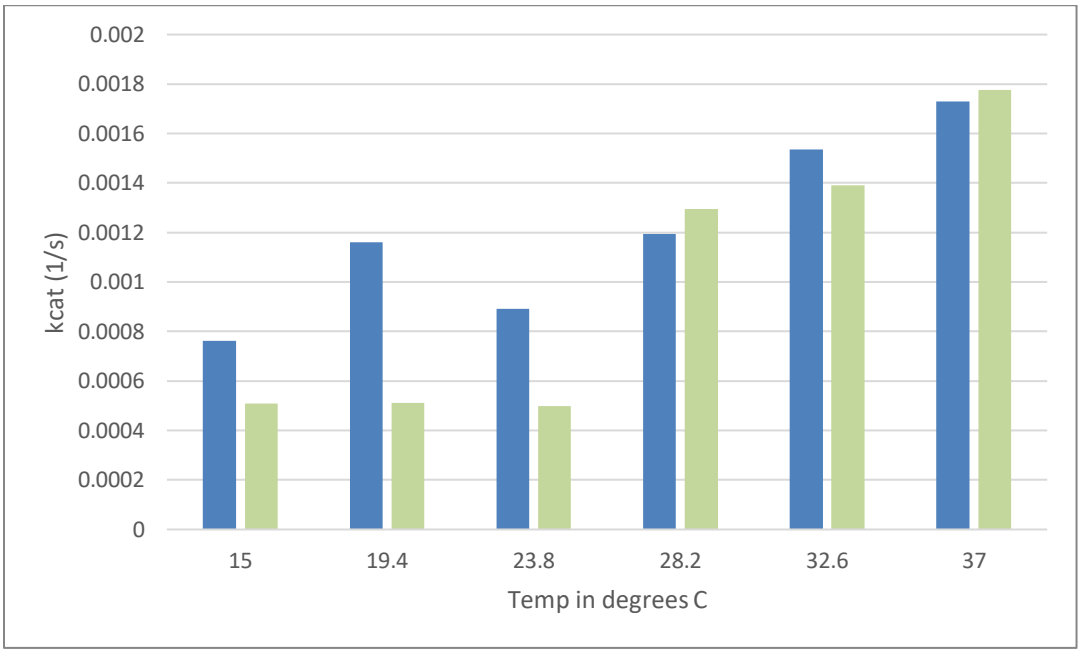


Figure 18: Bar graph of temperature assay for SDKIE study on WT *mt*IGPS. Temperatures were increased in increments of 4.4 ºC and assay settings used were [100 nM] IGPS, [60 µM] CdRP, and [100 µM] PIPES buffer in both H_2O and D_2O respectively

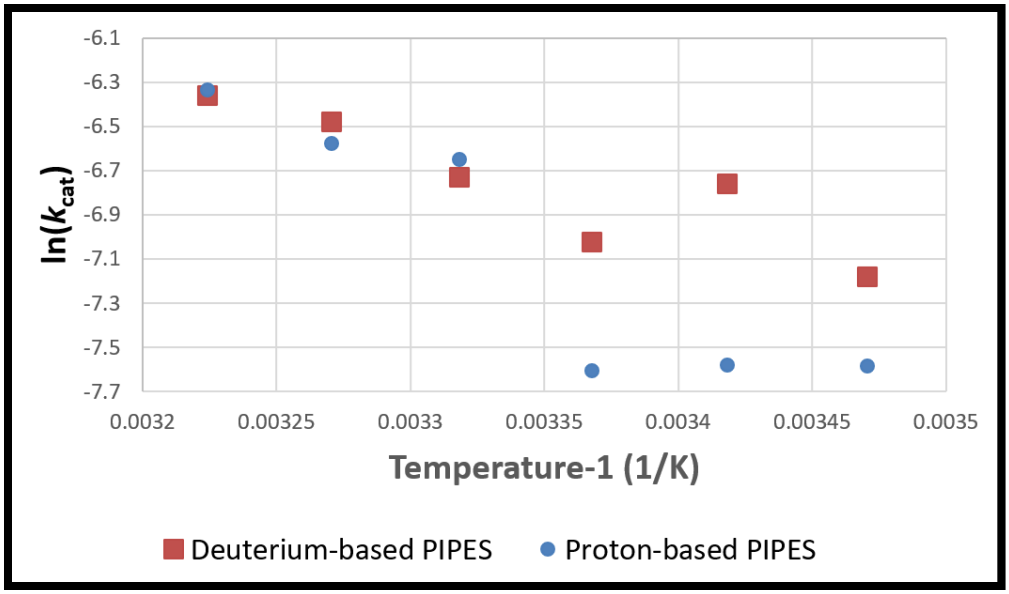


Figure 19: Arrhenius plot of temperature assay for SDKIE study on WT *mt*IGPS. Plot represents one trial run

Because of the myriad of results produced over the course of my time working on this project based solely on temperature, I strongly felt it was important to delve deeper into comprehending how IGPS activity binding to substrate works under various temperatures, especially when taken at a range from 15 ºC-37 ºC. pH was kept consistent to the PIPES assay buffer, of pH = 7.5 to further keep focus on temperature change alone. Currently, only preliminary studies have been conducted in order to gain further insight on the behavior of IGPS during rate-limiting experiments. Specifically it was important to see why IGPS consistently showed an inverse solvent viscosity effect at higher temperatures but a normal effect at regular temperatures.

We chose to study this by using Arrhenius plots, gathered from tightly-controlled temperatures set by a Peltier with the spectrophotometer. By conducting SDKIE studies on WT *mt*IGPS with an increase of 4.4 ºC between every run, we were able to gather enough information to plot a rough Arrhenius plot that leads us to believe the enzyme has a nonlinear correlation overall,

which is highly unusual (Fig. 19) where linearity is only seen from 15º C to about 24º C and then changes drastically when temperature is further increased. In Fig. 18, a trend of positive SDKIE results is observed for WT *mt*IGPS measured in buffer ranging in temperatures from 15 ºC to 22.8 ºC, but after this threshold, the SDKIE pattern shifts and data starts to show a more inverse SDKIE with increasing temperature climbing towards 37 ºC.

Because this is a preliminary run, it is important to fill in the blanks where information is required. More temperature assays will need to be run in triplicates with a smaller increment of temperature between each measurement (2.2 ºC), and pH will be kept consistently at pH = 7.5 for PIPES assay buffer. With more data points, affect of temperature on *mt*IGPS will be observable, especially when studied in comparison to E57D. This temperature assay will also be carried over into measuring for SVE to see whether or not the rate-limiting step is dependent on temperature, and whether or not this dependence carries over into the mutant.

Conclusion

Glu57 is catalytically important to mtIGPS as seen in E57D studies

*mt*IGPS shows similar behavior to past published work on *ss*IGPS however, activity as a function of temperature is crucial in understanding more about this enzyme, as it is not considered thermophilic like *ss*IGPS. When looking at the compilation of data for WT and E57D *mt*IGPS, it is presumable that Glu57 acts as the catalytic base in IGPS. Past studies (unpublished) have been conducted on the same residue with different mutations, such as E57A and E57Q; E57A produced results to lead the team to believe that the mutant is inactive, and E57Q has measurable activity but much lower than seen in E57D when comparing its k_{cat} to WT. The slight difference in amino acid change with conservation of the functional group seems to have helped E57D maintain a certain measurable amount of activity to help characterize *mt*IGPS further. Kinetic studies conducted in determining K_M and k_{cat} for WT and E57D indicate that the 57^{th} residue is in fact catalytically relevant. pH profiles helped to support this even further; the shift in pK_{a1} and pK_{a2} between WT and E57D variants show that Glu57 is most likely the catalytic base in *mt*IGPS. Furthermore, mutating *mt*IGPS at the catalytic base seems to alter the behavior of this enzyme where the rate-limiting step changes from being isotope-sensitive to viscosity-sensitive. The pooled data can only implement our hypothesis in that Glu 57 in *mt*IGPS is in fact catalytically important, and moreover, that it acts as the catalytic base.

Moving forward, there are more studies to be conducted based on the information gathered for E57D. Temperature assays should be run based on the information gathered in our rate-limiting step determination experiments (SDKIE and SVE) and tested for both WT and E57D to see whether or not behavior changes in the mutant. This will give us more insight on the mechanics of *mt*IGPS and the impact temperature has through temperature-controlled experiments.

Asn189 is catalytically important to mtIGPS based on loss of activity

Loss of activity within the single-point mutation N189L indicates that Asn189 plays an important role in catalysis for *mt*IGPS. In order to learn more about how this residue plays a role in our enzyme, we must study mutations at this residue less severe in difference, where the functional group is maintained much like we see in our E57D study. Loss of function leads us to believe that although catalytically relevant, that the change from a polar to nonpolar amino acid was too large, as there was a total alteration in functional groups within the side chain. Using our character analysis through kinetic studies, pH profiles, and rate-limiting experiments, it is possible to obtain more information if we study the mutant N189Q instead.

Is Lys119 the catalytic acid for mtIGPS?

We can only infer through our pH profile combined with 3-D mapping that Lys119 acts as the catalytic acid in *mt*IGPS however, this is only hypothetical. In order to support this hypothesis, it is important to follow up with similar character studies as conducted on E57D. By implementing a mutant on Lys119, we can observe whether or not this residue acts as the acid in the *mt*IGPS mechanism. This specific residue has been studied in Goodey Lab in the past with the single-point mutation K119A and a loss of function was observed. This tells us that K119 is in fact catalytically important, but we are unclear as to exactly *how* it is. A possible mutation to further study this residue is K119H, where histidine will maintain the positive charge from its side chain as found in lysine through the amine and imidazole groups.

By studying residues of *mt*IGPS, we can gain knowledge on the mechanics of this enzyme in order to ultimately use as a basis for future drugs fighting against MDR-TB in hopes that an inhibitor is eventually found to target *M. tuberculosis*.

References

(1) Tuberculosis https://www.who.int/news-room/fact-sheets/detail/tuberculosis (accessed May 1, 2021).

(2) List, F., Sterner, R.; Wilmanns, M., Related (beta alpha) (8) A Barrel Proteins in Histidine and Tryptophan Biosynthesis: A Paradigm to Study Enzyme Evolution. Chembiochem 12 (10), 1487A1494.

(3) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D., De novo computational design of retroaldol enzymes. Science 2008, 319 (5868), 1387A1391.

(4) Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D., Kemp elimination catalysts by computational enzyme design. Nature 2008, 453 (7192), 190AU4.

(5) Zaccardi, M. J.; O'Rourke, K. F.; Yezdimer, E. M.; Loggia, L. J.; Woldt, S.; Boehr, D. D. Loop-Loop Interactions Govern Multiple Steps in Indole-3-Glycerol Phosphate Synthase Catalysis: Loop Interactions in IGPS. *Protein Science* 2014, *23* (3), 302–311. https://doi.org/10.1002/pro.2416.

(6) Zaccardi, Margot & Yezdimer, Eric & Boehr, David. (2013). Functional Identification of the General Acid and Base in the Dehydration Step of Indole-3-glycerol Phosphate Synthase Catalysis. The Journal of biological chemistry. 288. 10.1074/jbc.M113.487447.

(7) Zaccardi, M. J. Mechanistic Analysis of the Tryptophan Biosynthetic Enzyme Indole-3-Glycerol Phosphate Synthase. 2013.

(8) Goodey, N. M., Alapa, M. T., Hagmann, D.F., Korunow, S.G., Mauro, A., Kwon, K.S., Hall, S.M Development of a fluorescently labeled thermostable DHFR for studying conformational changes associated with inhibitor binding, *Biochem. Biophys. Res. Comm.,* 2011, *413*, 442-447.

(9) Fernandez, P. L.; Murkin, A. S. Inverse Solvent Isotope Effects in Enzyme-Catalyzed Reactions. *Molecules* 2020, *25* (8), 1933. https://doi.org/10.3390/molecules25081933.

(10) Thomas, O. Indole-3-Glycerol Phosphate Synthase Ligand Binding Interactions. 43.