8-2019

Indole-3-glycerol Phosphate Synthase Ligand Binding Interactions

Oshane Thomas
Montclair State University

Abstract

Indole-3-glycerol phosphate synthase (IGPS) catalyzes the irreversible ring closure of 1-(o-carboxyphenylamino)-1-deoxyribulose 5- phosphate (CdRP) into indole-3-glycerol phosphate (IGP) during the fourth step in tryptophan biosynthesis. Analysis of the crystal structure of IGPS from *Mycobacterium tuberculosis* (*Mt*IGPS) hinted the importance of Lys119 for binding or catalysis. Lys110 from *sulfolobus sulfataricus* IGPS (*ss*IGPS) that aligns with Lys119 from *Mt*IGPS, was proposed to be a general acid for the proton transfer that initiates the ring closure and decarboxylation of CdRP. To study the importance of Lys119 in the chemical mechanism for *Mt*IGPS, an amino acid change was made at this position to study any changes in catalytic function or substrate binding affinity. Kinetic assays were performed and interesting, a mutation to alanine had a dramatic effect on the activity of *Mt*IGPS. Our studies highlight the importance of Lys119 in the *Mt*IGPS-catalyzed chemical mechanism.

MONTCLAIR STATE UNIVERSITY

Indole-3-glycerol Phosphate Synthase Ligand Binding Interactions

by

Oshane Thomas

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

August 2019

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]

John Siekierka, PhD.
Committee Member

INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE LIGAND BINDING INTERACTIONS

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

by

OSHANE THOMAS

Montclair State University

Montclair, NJ

2019

Acknowledgements

I would like to thank all of the lab members that I had the pleasure to work with in Dr. Nina Goodey's lab. It was a thrilling experience and I have learned so much that I could not have possibly dreamt of when I first started. I am grateful for the continual guidance and support of my principal investigator, Dr. Nina Goodey, without her, none of this would have been possible. I would like to acknowledge everyone that has assisted me in the completion of my thesis for my Master of Science degree. This includes my committee members, Dr. John Sierkierka and Dr. David Konas. They have assisted in the troubleshooting of experiments and writing of this thesis. In addition, I would like to thank Hanniah Riley, Katherine Leon Hernandez, Melanie Silva, Thomas Candela, Joseph Lacap, Audrey Tran, Dr. Ueli Gubler, Dr. Nina Goodey, Dr. David Konas and Adam Parker, for their help that had made the completion of this thesis a success. In close, I would like to thank the NIH grant 1R15GM126467-01 titled "Investigation of Substrate Specificity, Mechanism, and Inhibition of IGPS" for funding.

Table of Contents

List of Figures

List of Tables

Introduction

Mycobacterium tuberculosis (*M. tuberculosis*), the causative agent for tuberculosis (TB), is the top infectious killer in the world[13]. Despite advances in the development of antimicrobials, TB drug development remains quite challenging[14]. Most of the drugs used to treat TB are very old and several have never been studied under stringent conditions. As such, there is an urgent need to develop safer, simpler, shorter and more efficacious regimes to treat TB that may be in various forms[14]. TB being among the leading cause of death in several countries, requires continuous innovation of drugs to outcompete with antimicrobial drug resistance (AMR)[11]. According to the World Health Organization, multidrug-resistant tuberculosis (MDR-TB) is a form of TB that is caused by bacteria with resistance to isoniazid and rifampicin, the 2 strongest, primary anti-TB drugs[1]. As such, research on TB will be essential in developing new anti-TB drugs, while combating the future cost of AMR which could cause the global economy to incur losses worth US$ 100 trillion by 2050[12].

The exploration of indole-3-glycerol phosphate synthase (IGPS) as a potential target for TB drug development is of great interest, and has been shown to be essential for the survival of *M. tuberculosis*[16]. In 1998, Cole and co-worker[16] found knockouts in the leucine, proline and tryptophan biosynthetic pathways that are unable to replicate in cultured murine macrophages and had reduced ability to infect mice, suggesting the lack of these amino acids and their importance. They also identified a higher attenuation of virulence especially in the tryptophan auxotrophic strain, highlighting the importance of the tryptophan biosynthetic pathway in the virulence of *M. tuberculosis.* Five years later, Sasetti and co-workers used *M. tuberculosis* transposons to perform *in vitro* mutagenesis to inactivate the gene that encodes *Mt*IGPS, and showed that the IGPS gene is essential for the growth of the pathogen *in vitro*.[21]

Despite the growing interest of using *Mt*IGPS as a potential drug target and for studying the virulence of *M. tuberculosis*, only a few studies on *Mt*IGPS have been reported and one poor inhibitor has been identified[10]. In 2009, Shen and co-workers used homology modeling to deduce the structure of IGPS from a virulent strain, *M. tuberculosis* H37Rv, and screened a library of 60,000 compounds from the Maybridge database it. They later identified, ATB107 (Figure 1), a potential inhibitor that binds tightly to *Mt*IGPS *in vitro* with a K_D of 3 mM and a 50% inhibitory concentration of about 0.41 μM. In addition, ATB107 increased the K_M value of CdRP suggesting it might be a competitive inhibitor for *Mt*IGPS. However, due to the significant differences in the structures of CdRP and ATB107, and the low potency, there's a growing interest in developing novel inhibitors that more closely resembles CdRP. This class of inhibitors would act as transition state analogues, which are more potent inhibitors.

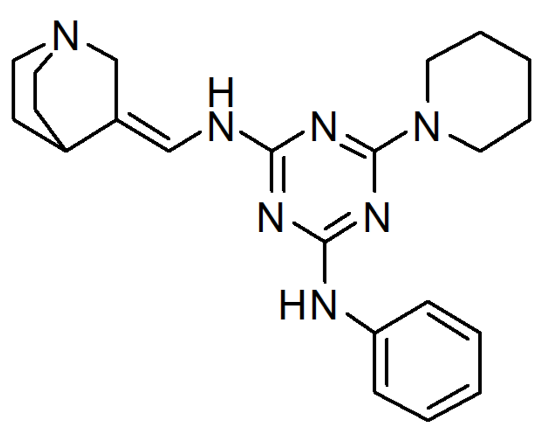


Figure 1. ATB107, a potential inhibitor for *Mt*IGPS that was discovered by Shen and co-workers[10].

IGPS is encoded by the 819 bp *trp*C gene in the tryptophan operon as a polypeptide consisting of 272 amino acids per subunit. IGPS is not found in humans and this fact makes IGPS an ideal candidate for anti-tuberculosis treatment. IGPS catalyzes the irreversible ring closure of 1- (o-carboxyphenylamino)-1-deoxyribulose 5- phosphate (CdRP) to form indole-3-glycerol phosphate (IGP), which is accompanied by the release of CO_2 and H_2O during the fourth step of tryptophan biosynthesis (Figure 2). The reaction is irreversible due to the release of carbon dioxide and formation of the aromatic indole ring in the second step. The similarity between the substrate and product to those in the Bischler-Möhlau indole synthesis, lead Parry[20] to propose that the biotransformation of CdRP to IGP involves the formation of two intermediates, 47 years ago. In the proposed mechanism (Figure 2), the first step involves a reversible enamine cyclization between the aromatic ring and the carbonyl of CdRP to form the first intermediate **3**. The next step includes formation of the second intermediate **4** via decarboxylation of **3**, followed by dehydration of **4** to form the final IGP product. Parry's original suggestion was supported by the crystallography data of IGPS from *S. solfataricus* IGPS (*ss*IGPS)[19] along with electrospray mass spectroscopy studies on IGPS from *M. tuberculosis*[22]. Czekster and co-workers used ESI(+)-MS to monitor the chemically mimicked intramolecular cyclization process of CdRP to IGP. They were able to show the ions attributed to the two reaction intermediates and the other major reaction species.

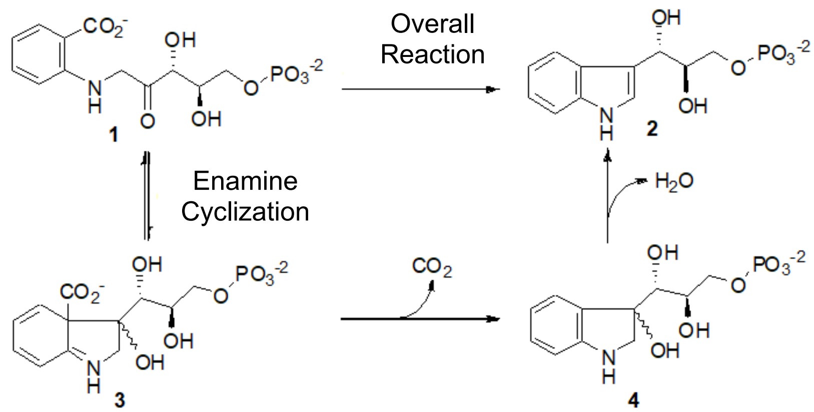


Figure 2. The proposed mechanism for the IGPS-catalyzed reaction. The figure was adopted from Dr. Nina Goodey and includes some slight modifications.

IGPS has the typical $(\beta/\alpha)_8$ barrel, or triosephosphate isomerase (TIM)-barrel fold, which consist of eight parallel β -strands in the shape of a barrel that are surrounded by eight α helices. IGPS shares the similarity with other TIM-barrel enzymes, in which important amino acid residues are located on the β-strands and the loops connecting the β-strands to the α-helices (βα loops)[25]. Hennig and co-workers[19] investigated the enzyme reaction mechanism of *ss*IGPS through X-ray crystallography. Crystal structures of *ss*IGPS complexed with CdRP and a reduced, non-reactive form of CdRP (rCdRP), shows that both CdRP and rCdRP binds to the active site in an extended, unproductive confirmation where the two carbon atoms that will joined are too far apart (4.5 Å) for bond formation. This suggests that conformational changes in the binding pocket are necessary to bring the two carbons closer together[4,19]. Dr. Goodey and co-workers[2] were able use variants of *ss*IGPS that were covalently labeled with fluorescent dyes at the N-terminal extension of the enzyme and the mobile active site loop, to show that the binding and conversion of CdRP to IGP is accompanied by conformational transitions.

Mutagenesis studies of *ss*IGPS have identified the importance of several amino acids, including Lys110 for the biotransformation[4-5]. In an earlier proposed mechanism for the *ss*IGPS-catalyzed reaction, Lys110 was predicted to be a general acid in both the ring closure and dehydration steps. However, there is no clear mechanism outlining why Lys110 could become reprotonated for the second proton transfer (Figure 3B). This inconsistency later lead to an alternative hypothesis that suggests that Lys110 only acts as a general acid in the ring closure step by donating a proton to the carbonyl of CdRP[5] (Figure 3C).

Kinetic experiments performed by Zaccardi *et al.* showed that Lys53 (and not Lys110) donates the second proton to the C2' hydroxyl group to catalyze the dehydration of the second intermediate (**4**) to IGP[4]. Mutations from Lys to Arg and Gln were made at that position to test the importance of the positive charge on Lys53 and its ability to act as a general acid. Both Lys53 substitutions had substantial effects on the k_{cat} and K_M, with K53R having a 10-fold decrease in maximum catalytic turnover. The K53Q variant had an almost 700-fold decrease in activity even at CdRP concentrations 50 x K_M for WT IGPS (12 μM). An alignment of *ss*IGPS and *Mt*IGPS shows that Lys110 corresponds to the Lys119 residue in *Mt*IGPS (Figure 3A).

A

```
S. Solfataricus  ------MPRYLKGWLKDVVQ----LSLRRPSFRASRQRPIISLNERILEFNKRNITAIIA 50
M. Tuberculosis  MSPATVLDSILEGVRADVAAREASVSLSEIKAAAAAAPPPLDV----MAALREPGIGVIA 56
                       :   *:*   **.     :** . .  *:   * :.:    :   :.   .:**

S. Solfataricus  EYKRKSPSGLD--VERDPIEYS-KFMERYAVGLSILTEEKYFNGSYETLRKIASSVSIPI 107
M. Tuberculosis  EVKRASPSAGALATIADPAKLAQAYQDGGARIVSVVTEQRRFQGSLDDLDAVRASVSIPV 116
                 * ** ***.    .  ** : :  : :  *  :*::**:: *:** : *  : :*****:
                  110
                  119
S. Solfataricus  LMKDFIVKESQIDDAYNLGADTVLLIVKILTERELESLLEYARSYGMEPLIEINDENDLD 167
M. Tuberculosis  LRKDFVVQPYQIHEARAHGADMLLLIVAALEQSVLVSMLDRTESLGMTALVEVHTEQEAD 176
                 * ***:*:  **.:*   *** :****  * :  * *:*: :.* **  *:*:: *:: *

S. Solfataricus  IALRIGARFIGINSRDLETLEINKENQRKLISMIPSNVVKVAESGISERNEIEELRKLGV 227
M. Tuberculosis  RALKAGAKVIGVNARDLMTLDVDRDCFARIAPGLPSSVIRIAESGVRGTADLLAYAGAGA 236
                  **: **:.**:*:*** **:::::   ::   :**.*:::****:    ::      *.

S. Solfataricus  NAFLIGSSLMRNPEKIKEFI---L------------248
M. Tuberculosis  DAVLVGEGLVTSGDPRAAVADLVTAGTHPSCPKPAR272
                 :*.*:*..*: . :    .
```

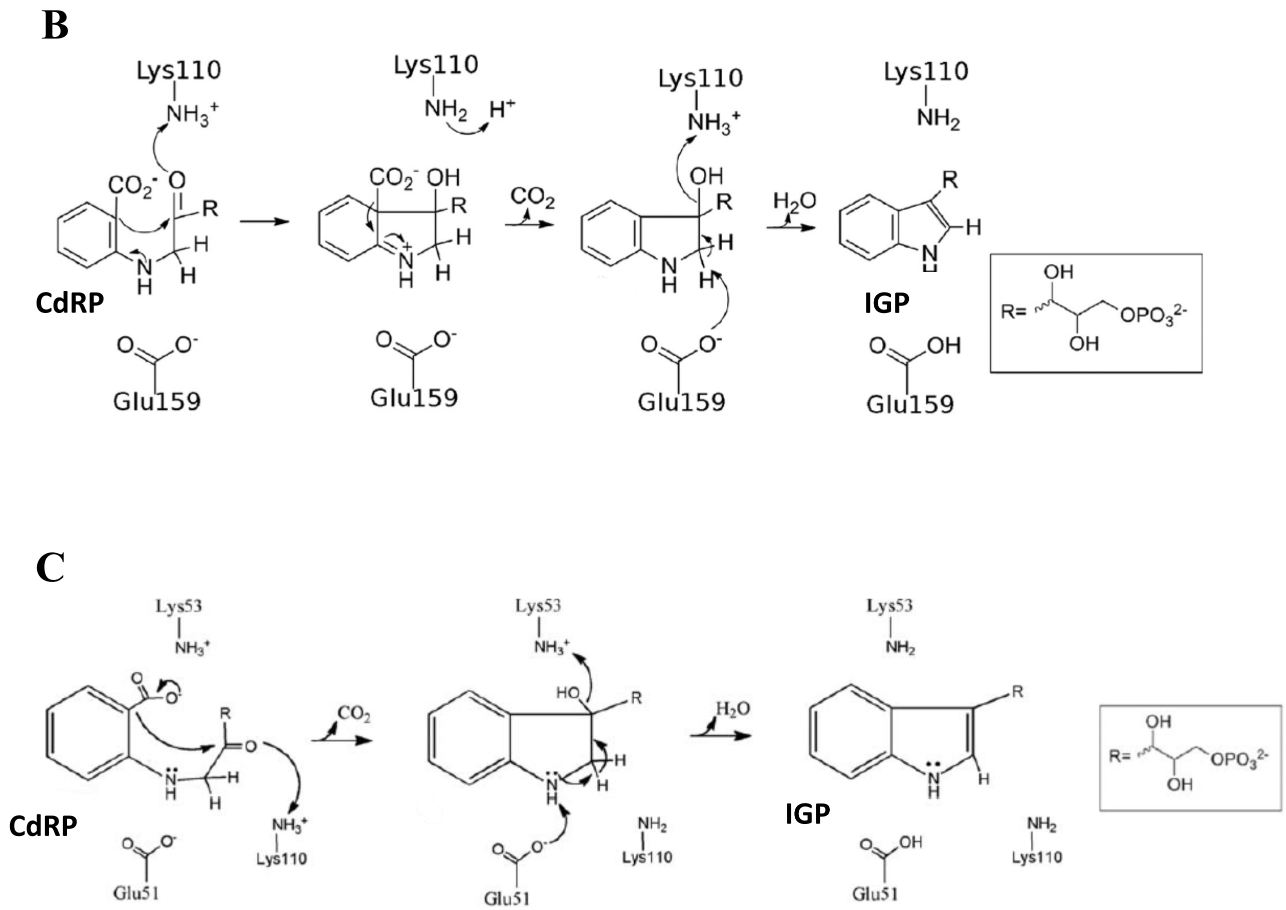


Figure 3: (A) CLUSTAL O (1.2.4) multiple sequence alignment of the amino acid sequences of *ss*IGPS and *Mt*IGPS showing conserved residues that are suspected to play a role in binding or catalysis. The alignment shows that Lys110 from *ss*IGPS corresponds to Lys119 in *Mt*IGPS. (B) The previously propose chemical mechanism for the conversion of CdRP to IGP by IGPS. (C) The newly proposed catalytic mechanism for *ss*IGPS. Figures 3B and 3C were taken from Zaccardi *et al.*[4,5] and include slight modifications.

In this study, we have performed a series of kinetic experiments to elucidate the possible role of Lys119 during binding and catalysis for the *Mt*IGPS-catalyzed reaction. Our experiments revealed that Lys119 is important for catalysis and could be involved in the binding of CdRP. These findings provide a strong insight into the chemical mechanism of *Mt*IGPS.

2. Materials and Methods

Materials – All chemicals were of analytical or reagent grade and were used without further purification unless otherwise specified.

2.1 *Site-directed Mutagenesis*

2.1.1 *Isolation of Plasmid DNA*

*Mt*IGPS exists as a monofunctional enzyme. The *trp*C coding sequence for *Mt*IGPS cloned in pET30a expression vector was purchased from GenScript (New Jersey, USA). The pET30a expression vector was isolated from LOBSTR-BL21 (DE3) host cells using slightly modified version of Promega's Wizard® Plus SV Minipreps DNA Purification System. In this updated protocol, 5 mL of LB broth (25 g/L) and kanamycin (50 µg/mL) was added to a 15 mL Falcon™ conical centrifuge tube, inoculated with host cells previously transformed with the pET30a expression vector. The cells were incubated at 37°C for 16-18 hours with shaking at 225 rpm. The following day, the bacterial culture was centrifuged 1 mL at a time at 13,000 rpm for 60 seconds in 1.5 mL microcentrifuge tube, after which the supernatant was discarded. The pelleted cells were resuspended in 250 µL of a cell resuspension solution followed by 250 µL of a cell lysis solution. The contents were mixed thoroughly by inverting the tube 4–6 times until the solution becomes viscous and slightly clear. Alkaline protease solution (10 µL) was added and the lysis reaction was timed to ensure that it did not proceed for more than five minutes. A neutralization solution (350 µL) was added to the solution and mixed immediately and thoroughly by inverting the tube 4–6 times. The mixture was centrifuged for 10 minutes at 13,000 rpm in a table-top microcentrifuge. The supernatant was removed from the mixture and added to a spin column inserted into a collection tube. The spin column was centrifuge for 60 s and the flow-through was discarded. The spin column was washed by adding 0.75 ml of a wash solution (ethanol added) and centrifuged for 60 s at 13,000 rpm. The flow-through was discarded and the column was centrifuged again at full speed for an additional 60 s to remove residual wash solution. The spin column transferred to a clean 1.5 mL microcentrifuge tube and 30 µL of nuclease-free water was added to the center of the spin column and allowed to stand for 1 minute, after which it was centrifuged for 1 minute, to elute the DNA. The DNA was stored at -80 °C.

2.1.2 *Mutant Strand Synthesis*

Site-directed mutagenesis was performed on the *Mt*IGPS gene to achieve the K119A variant for this study. Oligonucleotides used in the mutant strand synthesis reactions were supplied by Integrated DNA Technologies, Inc (California, USA). The forward primer used to achieve the K119A mutation was: 5'-GATTCCGGTTCTGCGCGCGGATTTCGTCGTGCAG-3' and the reverse primer was: 5'-CTGCACGACGAAATCCGCGCGCAGAACCGGAATC-3'. Each mutagenesis reaction consist of 0.5 µM of both forward and reverse mutagenic primers, 100 ng of DNA template, 0.2 µM dNTPs, and 1X HF buffer (Thermo Fisher Scientific, Cat # F530S) in a total volume of 50 µL. One unit of Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific, Cat # F530S) was added to the reactions, which were performed in a thermal cycler (C1000 Touch^{TM}, Thermal cycler, by BIORAD) with appropriate conditions (Table 1). For the negative control reaction, primers were excluded from the reaction mix and the other reaction components remain unchanged. The final volume of the reaction mix was adjusted to 50 µL using nuclease-free water. Following the PCR reactions, 10 units of *Dpn I* endonuclease (Fisher

Scientific, Cat # FERER1701) were added to each reaction and was gently mixed by pipetting the solution up and down several times after which it was incubated at 37 °C for an hour.

Table 1. The settings used for various steps in the site-directed mutagenesis reactions.

Step	Temperature (°C)	Time (min.)	Cycles
Initial Denaturation	98	0.5	1X
1. Denaturation	98	0.5	22X
2. Anneal	55	1	
3. Extension	72	3.5	
Final	72	10	1X
Storage	4	∞	1X

2.2 *Transformation of XL1 Blue and BL21 strains*

The XL1-Blue competent cells were gently thawed on ice. For each control and sample reaction to be transformed, 100 µL of cells were added to prechilled 1.5 mL microcentrifuge tubes. *Dpn I*-treated DNA (5 µL) from each reaction was added to separate 1.5 mL microcentrifuge tubes containing the XL1-Blue competent cells. The reaction mixtures were gently swirled to mix their contents and incubated on ice for 30 minutes. The transformation reactions were placed in a hot water bath at 42°C for 45 seconds and the immediately placed on ice for 2 minutes. Sterile LB broth (900 µL) was added to each transformation reactions and incubated at 37°C for 1 hour with shaking at 225 rpm. 250 µL of each transformation reaction was placed on preheated agar plates supplemented with kanamycin (50 µg/mL), and spread with inoculating loops. The transformation plates were incubated 37°C for 18 – 20 hours. The plasmids were isolated from two colonies of each transformant as previously described, and sequenced at the Biology Department of Montclair State University.

2.3 *Expression and Purification*

The expression vectors for both the wild type *Mt*IGPS and the K119A variant were transformed in LOBSTR-BL21 (DE3) host cells as previously described for the transformation of XL1-Blue competent cells. Starter cultures for each protein to be expressed were grown to saturation for 16-18 hours at 37 °C with shaking (225 rpm). The following day, previously autoclaved LB media (250 mL) supplemented with kanamycin (50 µg/ml), was inoculated with 1% of the saturated overnight culture (v/v). The expression cultures were grown at 37 °C until the absorbance (A_{600}) reached a value of 0.5-0.6. IPTG was added to the cell culture to a final concentration of 1 mM and the cell culture was incubated for 16-18 hours at 25 °C with shaking (225 rpm). The cells were harvested by centrifugation at 5,000 rpm for 15 minutes at 4 °C. The recombinant IGPS was purified by Ni-NTA His•Binding resin affinity chromatography with buffers at 8.0 pH and 4 °C using a Gravity-flow Column. The harvested cells were resuspended in 15 mL equilibration buffer (1X PBS, 10 mM imidazole) and broken by sonication. The lysate was centrifuged at 15,000 rpm, 4 °C for 10 minutes. The supernatant was carefully removed and used for purification. The Gravity-flow Column was placed on a stand and packed with 2 mL of HisPur™ Ni-NTA resin (Cat # 88222). The storage buffer (buffer in the resin) was allowed to drain from resin by gravity flow. The column was equilibrated with three resin-bed volumes of equilibration buffer and the buffer was allowed to drain from column. The equilibration process was then repeated for a second time. The protein extract (supernatant from the lysate) was added to the resin. The flow-through was collected in a 15 mL Falcon™ conical centrifuge tube. The resin was washed with two resin-bed volumes of wash buffer and the flow-through was collected in separate 15 mL Falcon™ conical centrifuge tubes. This step was repeated using a new collection tube until a Bradford test showed no difference between the colors for the control solution and the flow-through being tested. For the Bradford test, 20 μL of Bradford reagent (1X) and 80 μL of the flow-through was mixed in a 96 well microtiter plate. The color of the flow-through was compared to the control solution (20 μL of 1X Bradford reagent and 80 μL of the wash buffer). The enzymes were eluted from the resin with four resin-bed volumes of elution buffer and collected in 1 mL fractions in separate microcentrifuge tubes, kept on ice. . This step was repeated using a new collection tube until a Bradford test showed no difference between the colors for the control solution (20 μL of 1X Bradford reagent and 80 μL of the elution buffer) and the flow-through. The purified recombinant IGPS was dialyzed against 5 mM Tris-HCl buffer (pH 7.9) and stored at -80 °C. The absorbance of the purified IGPS at 280 nm and its molar extinction coefficient (4595 M^{-1} cm^{-1}) obtained from ExPASy's ProtParam[23] online calculator was used to determine its concentration.

2.4 *CdRP Concentration Determination*

The chemical synthesis of CdRP was optimized by Dr. David Konas' lab (Department of Chemistry and Biochemistry, Montclair State University, NJ), following a previously reported protocol[6-7]. A solution of anthranilic acid (1.0 mmol) in 700 μL of 1:1 (v:v) iPrOH / EtOH was added a solution of ribose-5-phosphate (0.5 mmol) in 2.0 mL of 2:1 (v:v) iPrOH / H2O. The mixture was mixed, stored in the dark at room temperature overnight. The remaining solvents were decanted off and the oil was washed with ethyl acetate and diethyl ether. Barium acetate was added to precipitate the CdRP product. The solid was collected by centrifugation and washed repeatedly with portions of H2O, iPrOH, and EtOAc. With each washing, the solid product was collected by centrifugation and the solvent was decanted. The final product was dried under vacuum at room temperature and obtained as a light yellow solid. Na_2SO_4 was added to the crude CdRP product to change the CdRP counterion from barium to sodium, which increases its solubility in water. The concentration of CdRP was determined by measuring its absorbance at 330 nm[17] and using its molar extinction coefficient, 0.1415 mM^{-1} cm^{-1}, which was determined from preliminary experiments by weighing.

2.5 *Enzymatic Assays*

All assays were performed at 25 °C and measurements were done in triplicates, unless stated otherwise. The activity of *Mt*IGPS was determined by measuring the increase of absorbance intensity at 280 nm. Data points were collected once every two seconds. The assays were performed in 100 mM PIPES, 1 mM DTT, at pH 7.5. The increase in absorbance at 278 nm due to the formation of IGP was monitored for 10 minutes using a spectrophotometer (Evolution 201 UV-visible spectrophotometer, Thermo Scientific) after combining 600 µM CdRP and 3 µM *Mt*IGPS in a 1 mL cuvette. A previously reported molar extinction coefficient of 5,500 M^{-1} cm^{-1} was used for IGP[9]. Control reactions were performed with *Mt*IGPS or CdRP only to determine baseline absorbance values attributed to both species. Fluorescence spectroscopy (280 nm excitation; 340 nm emission) were performed using a spectrofluorometer (FluoroMAax-4, Horiba Jobin Yvon) with 5 nm and 10 nm slits for the excitation and emission wavelengths respectively. Activity assays were performed with 40 nM of *Mt*IGPS and 200 µM of CdRP in 100 mM PIPES, 1 mM DTT, 5mM EDTA at pH 7.5. The K_M values were determined in the fluorometer as described above except that the CdRP concentration ranged from 200 µM to 3.125 µM, and the slits used were 2 times larger for each wavelength. The formation of IGP was monitored by measuring the increase in fluorescence emission in the fluorometer. Initial rate data were fit to the Michaelis-Menten equation using nonlinear regression with the program KaleidaGraph® (version 4.1), using Eqn. 1, where V_0 is the initial velocity, k_{cat} is the enzyme's turn-over number, K_M is the Michaelis-Menten constant, E_T is total enzyme concentration and [S] is the substrate concentration.

$$V_0 = (k_{cat} / E_T) [S] / (K_M + [S]) \qquad \text{(Eqn. 1)}$$

3. Results and Discussion

3.1 *Site-directed Mutagenesis*

Sequencing was performed by Adam Parker at Montclair State University with the T7 promoter and terminator sequencing primers. The sequencing data were viewed in Chromas (version 2.6.6.0), which showed the results as DNA sequences. The DNA sequences were translated to the amino acid sequences using ExPASy's online "translate tool". The amino acid sequence for *Mt*IGPS was aligned with the GenScript quote (Figure 4A). A careful analysis of the results showed that no mutations were introduced in the *Mt*IGPS protein sequence, which was identical to the GenScript quote. The DNA sequence for the sequencing results for K119A was aligned with GenScript quote and a change in the nucleotides that translate to a mutation from lysine to alanine was verified (Figure 4B, green). The amino acid sequence for the sequencing results for K119A was aligned with the Uniprot *Mt*IGPS (Entry: P9WFX7) and the expected amino acid change was identified (Figure 4C, teal). The CLUSTAL O (1.2.4) multiple sequence alignment tool from EMBL-EBI was used for the alignments.

A

```
Sequencing Results MSPATVLDSILEGVRADVAAREASVSLSEIKAAAAAAPPPLDVMAALREPGIGVIAEVKR 60
GenScript Quote    MSPATVLDSILEGVRADVAAREASVSLSEIKAAAAAAPPPLDVMAALREPGIGVIAEVKR 60
                   ************************************************************

Sequencing Results ASPSAGALATIADPAKLAQAYQDGGARIVSVVTEQRRFQGSLDDLDAVRASVSIPVLRKD 120
GenScript Quote    ASPSAGALATIADPAKLAQAYQDGGARIVSVVTEQRRFQGSLDDLDAVRASVSIPVLRKD 120
                   ************************************************************

Sequencing Results FVVQPYQIHEARAHGADMLLLIVAALEQSVLVSMLDRTESLGMTALVEVHTEQEADRALK 180
GenScript Quote    FVVQPYQIHEARAHGADMLLLIVAALEQSVLVSMLDRTESLGMTALVEVHTEQEADRALK 180
                   ************************************************************

Sequencing Results AGAKVIGVNARDLMTLDVDRDCFARIAPGLPSSVIRIAESGVRGTADLLAYAGAGADAVL 240
GenScript Quote    AGAKVIGVNARDLMTLDVDRDCFARIAPGLPSSVIRIAESGVRGTADLLAYAGAGADAVL 240
                   ************************************************************

Sequencing Results VGEGLVTSGDPRAAVADLVTAGTHPSCPKPAR 272
GenScript Quote    VGEGLVTSGDPRAAVADLVTAGTHPSCPKPAR 272
                   ********************************
```

B

```
GenScript Quote     ATGAGTCCGGCGACCGTGCTGGATAGTATTCTGGAAGGTGTTCGTGCTGATGTTGCTGCC 60
Sequencing Results  ATGAGTCCGGCGACCGTGCTGGATAGTATTCTGGAAGGTGTTCGTGCTGATGTTGCTGCC 60
                    ************************************************************

GenScript Quote     CGTGAAGCCTCCGTTTCTCTGTCTGAAATCAAAGCAGCAGCAGCTGCAGCACCGCCGCCG 120
Sequencing Results  CGTGAAGCCTCCGTTTCTCTGTCTGAAATCAAAGCAGCAGCAGCTGCAGCACCGCCGCCG 120
                    ************************************************************

GenScript Quote     CTGGATGTTATGGCAGCTCTGCGTGAACCGGGCATTGGTGTGATCGCAGAAGTTAAACGT 180
Sequencing Results  CTGGATGTTATGGCAGCTCTGCGTGAACCGGGCATTGGTGTGATCGCAGAAGTTAAACGT 180
                    ************************************************************

GenScript Quote     GCAAGTCCGTCCGCAGGTGCTCTGGCAACCATCGCGGACCCGGCAAAGCTGGCCCAGGCA 240
Sequencing Results  GCAAGTCCGTCCGCAGGTGCTCTGGCAACCATCGCGGACCCGGCAAAGCTGGCCCAGGCA 240
                    ************************************************************

GenScript Quote     TATCAAGATGGCGGTGCGCGTATTGTGAGTGTGGTGACGGAACAGCGTCGCTTTCAAGGT 300
Sequencing Results  TATCAAGATGGCGGTGCGCGTATTGTGAGTGTGGTGACGGAACAGCGTCGCTTTCAAGGT 300
```

```
                     ************************************************************

GenScript Quote      TCCCTGGATGACCTGGACGCAGTCCGTGCTTCAGTGTCGATTCCGGTTCTGCGCAAGGAT 360
Sequencing Results   TCCCTGGATGACCTGGACGCAGTCCGTGCTTCAGTGTCGATTCCGGTTCTGCGCGCGGAT 360
                     ******************************************************  ****

GenScript Quote      TTCGTCGTGCAGCCGTACCAAATCCATGAAGCTCGTGCACACGGTGCAGACATGCTGCTG 420
Sequencing Results   TTCGTCGTGCAGCCGTACCAAATCCATGAAGCTCGTGCACACGGTGCAGACATGCTGCTG 420
                     ************************************************************

GenScript Quote      CTGATTGTTGCGGCCCTGGAACAGAGCGTTCT-GGTCTCTATGCTGGATCGCACCGAAAG 479
Sequencing Results   CTGATTGTTGCGGCCCTGGAACAGAGCGTTCTGGGTCTCTATGCTGGATCGCACCGAAAG 480
                     ******************************** ***************************

GenScript Quote      CCTGGGTATGACGGCGCTGGTTGAAGTTCATACCGAACAGGAAGCGGACCGTGCACTGAA 539
Sequencing Results   CCTGGGTATGACGGCGCTGGTTGAAGTTCATACCGAACAGGAAGCGGACCGTGCACTGAA 540
                     ************************************************************

GenScript Quote      AGCAGGTGCAAAGGTTATCGGTGTCAACGCCCGCGATCTGATGACGCTGGATGTTGACCG 599
Sequencing Results   AGCAGGTGCAAAGGTTATCGGTGTCAACGCCCGCGATCTGATGACGCTGGATGTTGACCG 600
                     ************************************************************

GenScript Quote      TGATTGCTTTGCACGCATTGCTCCGGGCCTGCCGAGCTCTGTCATTCGTATCGCAGAAAG 659
Sequencing Results   TGATTGCTTTGCACGCATTGCTCCGGGCCTGCCGAGCTCTGTCATTCGTATCGCAGAAAG 660
                     ************************************************************

GenScript Quote      CGGTGTGCGTGGTACCGCAGACCTGCTGGCTTATGCAGGTGCAGGTGCAGATGCTGTCCT 719
Sequencing Results   CGGTGTGCGTGGTACCGCAGACCTGCTGGCTTATGCAGGTGCAGGTGCAGATGCTGTCCT 720
                     ************************************************************

GenScript Quote      GGTGGGCGAAGGTCTGGTGACGTCTGGTGATCCGCGCGCGGCAGTGGCTGACCTGGTGAC 779
Sequencing Results   GGTGGGCGAAGGTCTGGTGACGTCTGGTGATCCGCGCGCGGCAGTGGCTGACCTGGTGAC 780
                     ************************************************************

GenScript Quote      GGCGGGCACGCATCCGAGTTGTCCGAAACCGGCACGCTAA 819
Sequencing Results   GGCGGGCACGCATCCGAGTTGTCCGAAACCGGCACGCTAA 820
                     ****************************************
```

C

```
Sequencing Results MSPATVLDSILEGVRADVAAREASVSLSEIKAAAAAAPPPLDVMAALREPGIGVIAEVKR 60
Uniprot MtIGPS     MSPATVLDSILEGVRADVAAREASVSLSEIKAAAAAAPPPLDVMAALREPGIGVIAEVKR 60
                   ************************************************************

Sequencing Results ASPSAGALATIADPAKLAQAYQDGGARIVSVVTEQRRFQGSLDDLDAVRASVSIPVLRKD 120
Uniprot MtIGPS     ASPSAGALATIADPAKLAQAYQDGGARIVSVVTEQRRFQGSLDDLDAVRASVSIPVLRAD 120
                   ********************************************************** *

Sequencing Results FVVQPYQIHEARAHGADMLLLIVAALEQSVLVSMLDRTESLGMTALVEVHTEQEADRALK 180
Uniprot MtIGPS     FVVQPYQIHEARAHGADMLLLIVAALEQSVLVSMLDRTESLGMTALVEVHTEQEADRALK 180
                   ************************************************************

Sequencing Results AGAKVIGVNARDLMTLDVDRDCFARIAPGLPSSVIRIAESGVRGTADLLAYAGAGADAVL 240
Uniprot MtIGPS     AGAKVIGVNARDLMTLDVDRDCFARIAPGLPSSVIRIAESGVRGTADLLAYAGAGADAVL 240
                   ************************************************************

Sequencing Results VGEGLVTSGDPRAAVADLVTAGTHPSCPKPAR 272
Uniprot MtIGPS     VGEGLVTSGDPRAAVADLVTAGTHPSCPKPAR 272
                   ********************************
```

Figure 4: (A) The amino acid sequence alignment results for the sequencing results for *Mt*IGPS and the GenScript quote. (B) The alignment of the sequencing results for K119A and the GenScript quote DNA sequences. The nucleotide changes that lead to the K119A variant is highlighted in green. (C) The amino acid sequence alignment results for the sequencing results for K119A and the Uniprot *Mt*IGPS (Entry: P9WFX7). The amino acid change was highlighted in teal. The CLUSTAL O (1.2.4) multiple sequence alignment tool from EMBL-EBI was used for the alignments.

3.2 *Expression and Purification*

*Mt*IGPS and the K119A mutant was expressed in the LOBSTR-BL21 (DE3) strain and purified homogenously to a yield of 18.36 ± 0.03 mg and 48.16 ± 0.08 respectively from a 2 g cell pellet, obtained from 500 mL LB broth (25 g/L). SDS-PAGE was used to approximate the molecular weights of the purified *Mt*IGPS and the K119A mutant (Figure 5), which was similar to the predicted molecular weight for the *Mt*IGPS with the His tag (33.5 kDa), that was determined using ExPASy's ProtParam[23] online tool (Table 2). The result shows that *Mt*IGPS and the K119A mutant were monomers in the solution as previously described[18].

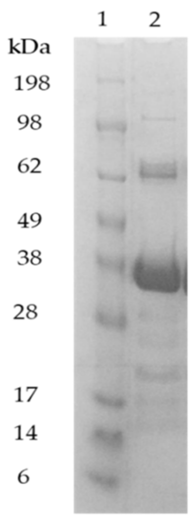


Figure 5. SDS-PAGE results of the purification of wild type IGPS. Lane 1, SeeBlue™ Plus2 Pre-stained Protein Standard (Cat # LC5925); lane 2, elution fraction at the expected molecular weight of 33.5 kDa + His tag.

Table 2. Various physical and chemical parameters determined using ExPASy's ProtParam[23] online tool.

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

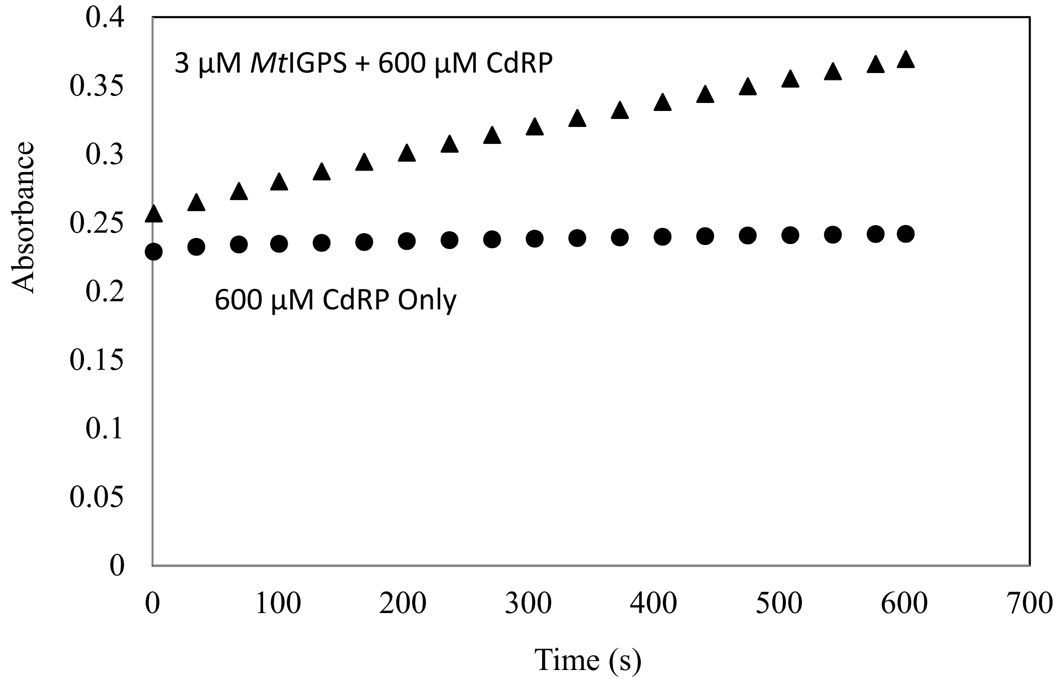
3.3 *Enzymatic Assays*

The biotransformation of CdRP to IGP product was followed by UV-VIS (Figure 6) in a spectrophotometer (Cary 100, version12.00). In a 1 mL cuvette, 3 μM of *Mt*IGPS and 600 μM CdRP were mixed in 100 mM PIPES and 1 mM DTT at pH 7.5. The experiment was repeated for the K119A variant using 3 μM K119A and identical conditions used for the *Mt*IGPS k_{cat} determination. Over the course of the reaction, a steady increase in absorbance owing to the formation of IGP[9] was observed (Figure 6A, triangles). On the other hand, no increase in absorbance was detected for the reaction catalyzed by K119A (Figure 6B, squares). A small increase in absorbance was also observed for the control reactions. This increase in absorbance could be due to the decomposition of CdRP by hydrolysis. The k_{cat} for *Mt*IGPS was determined by converting the slope of the reaction (Figure 6A, triangles) to the maximum velocity (M s^{-1}) using a previously reported molar extinction coefficient for IGP (5,500 M^{-1} cm^{-1})[9], and the total enzyme concentration in the reaction (3 μM). The k_{cat} value for *Mt*IGPS was found to be 0.021 ± 0.003 s^{-1}, which is approximately 10 times lower than the value previously determined (0.16 ± 0.01 s^{-1})[8]. The discrepancy between the k_{cat} values could be attributed to the purity of the CdRP used in the assays; the CdRP could have been contaminated with enzyme inhibitors, which could be the unreacted substrates from the chemical synthesis or the products from the hydrolysis of CdRP. The upper limit for the k_{cat} value for K119A was determined as previously described for *Mt*IGPS and was found to be 0.00121 s^{-1}, which is less than 17 fold lower than k_{cat} value for *Mt*IGPS (Table 3).

Table 3. Steady-state kinetic parameters for *Mt*IGPS and K119A (at pH 7.5). The k_{cat} values were determined using saturating concentration of CdRP (600 μM). 3 μM of *Mt*IGPS and 600 μM CdRP was mixed in 100 mM PIPES and 1 mM DTT at pH 7.5. The k_{cat} was determined by converting the slope of the reaction (Figure 5A, triangles) to M s^{-1} using a previously reported molar extinction coefficient for IGP (5,500 M^{-1} cm^{-1})[9], and the total enzyme concentration in the reaction. The k_{cat} for K119A was calculated as described for *Mt*IGPS. The reported errors are standard deviations of three trials for each reaction.

	k_{cat} (s^{-1})
*Mt*IGPS	0.021 ± 0.003
K119A	< 0.00121

A


3 μM *Mt*IGPS + 600 μM CdRP
600 μM CdRP Only
Absorbance
Time (s)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
100
200
300
400
500
600
700

B

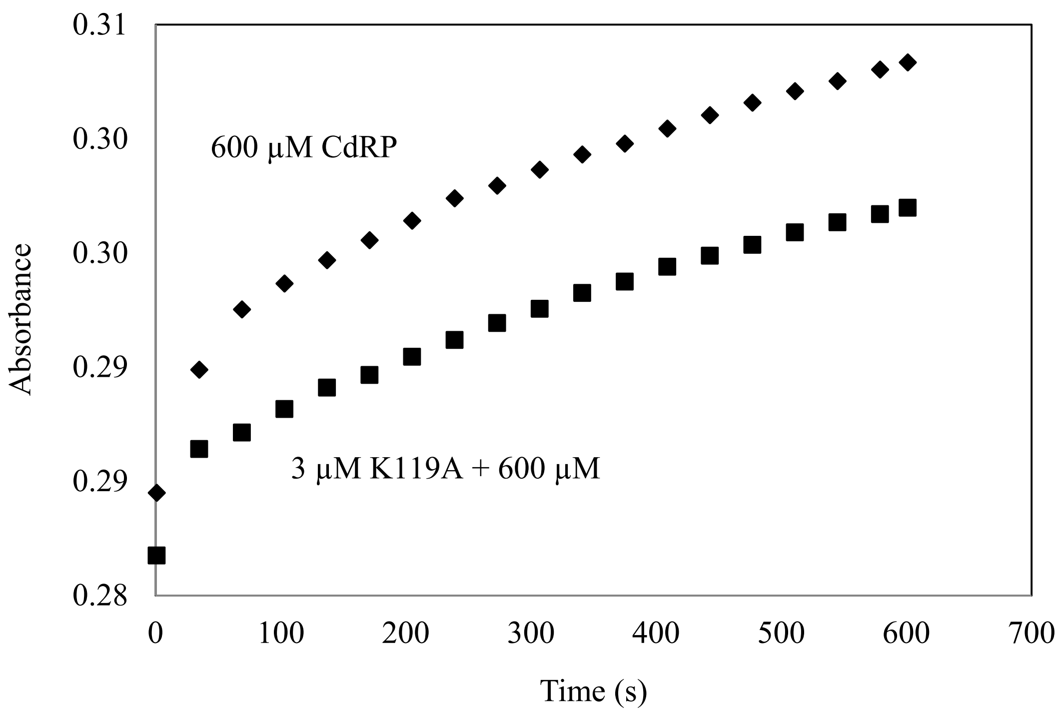

600 μM CdRP
3 μM K119A + 600 μM
Absorbance
Time (s)
0.31
0.30
0.30
0.29
0.29
0.28
0
100
200
300
400
500
600
700

Figure 6: (A) Absorbance versus time in seconds for the conversion of CdRP to IGP catalyzed by *Mt*IGPS over the period of 600 seconds of a solution of 3 μM *Mt*IGPS and 600 μM CdRP (triangles), and 600 μM CdRP only (circles). Both reactions proceeded in 100 mM PIPES and 1 mM DTT at pH 7.5. (B) Reactions for 3 μM K119A and 600 μM CdRP (squares), and 600 μM CdRP only (diamonds). The absorbance data used for the plots was an average of three trials performed under identical conditions. No increase in absorbance was observed for the K119A reactions relative to the CdRP only control.

The K_M value determined from the initial velocity data was found to be 14 ± 3 μM, which is an average of three trials and the error is the standard deviation of the three trials. The K_M value is more than 3 times lower than the value previously obtained by Czekster *et al.* (55 ± 9 μM,)[8] and 60 times lower than the value previously obtained by Shen *et al.* (1.13 mM)[10]. The significant difference in the K_M values could be due to the difference in purity of the CdRP chemically synthesized. The catalytic efficiency, k_{cat}/K_M, for the Mt IGPS is 1.5 x 10^3 M^{-1} s^{-1}, which is more than two times smaller than previously obtained by Czekster and co-workers (2.9 x 10^3 M^{-1} s^{-1})[8].

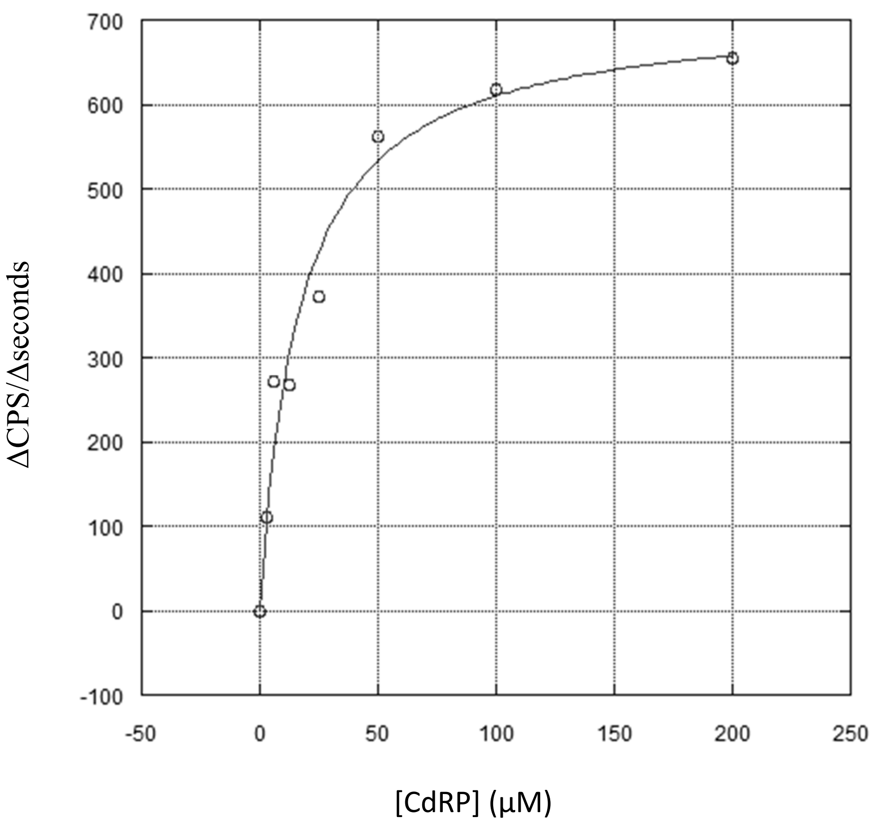
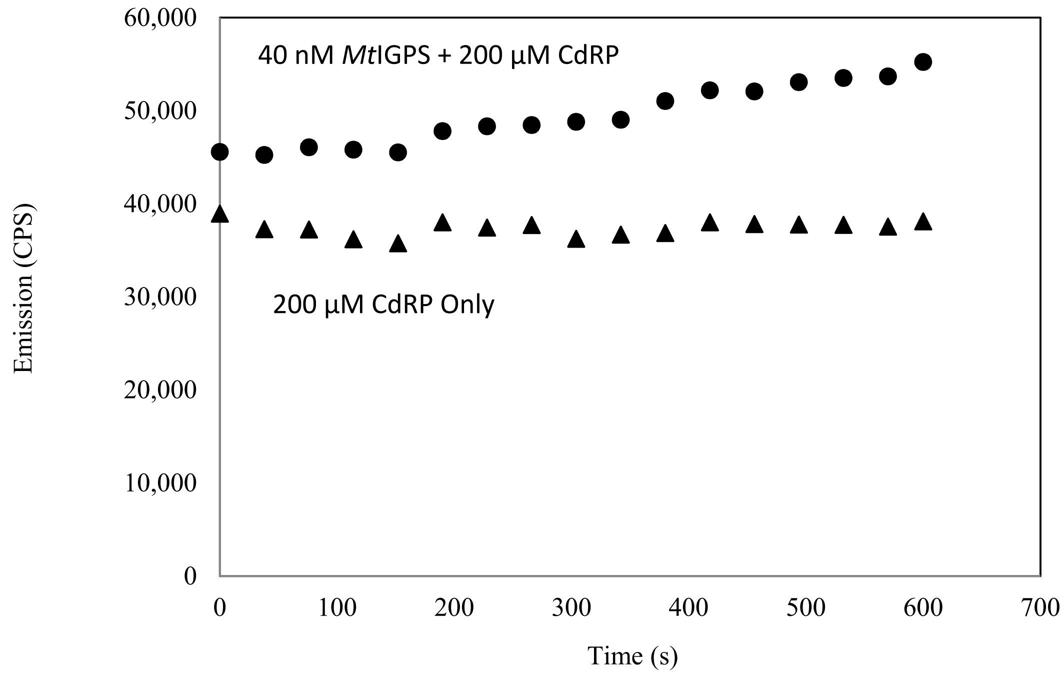


Figure 7. Representative graph of velocity (ΔCPS/Δseconds) versus CdRP concentrations for a trial is shown. Reactions were performed with 40 nM of *Mt*IGPS and various concentration of CdRP ranging from 200 μM to 3.125 μM in 100 mM PIPES, 1 mM DTT, 5mM EDTA at pH 7.5. The slits used were 10 nm and 15 nm for the excitation and emission wavelengths respectively. The slopes (ΔCPS/Δseconds) and respective CdRP concentrations from each reaction was fitted to the Michaelis-Menten equation (Eqn. 1) to determine the K_M value for *Mt*IGPS. The fit for *Mt*IGPS-catalyzed reaction was done using the program KaleidaGraph® version 4.1. The reactions were performed in triplicate and the average of the three K_M values was 14 ± 3 μM.

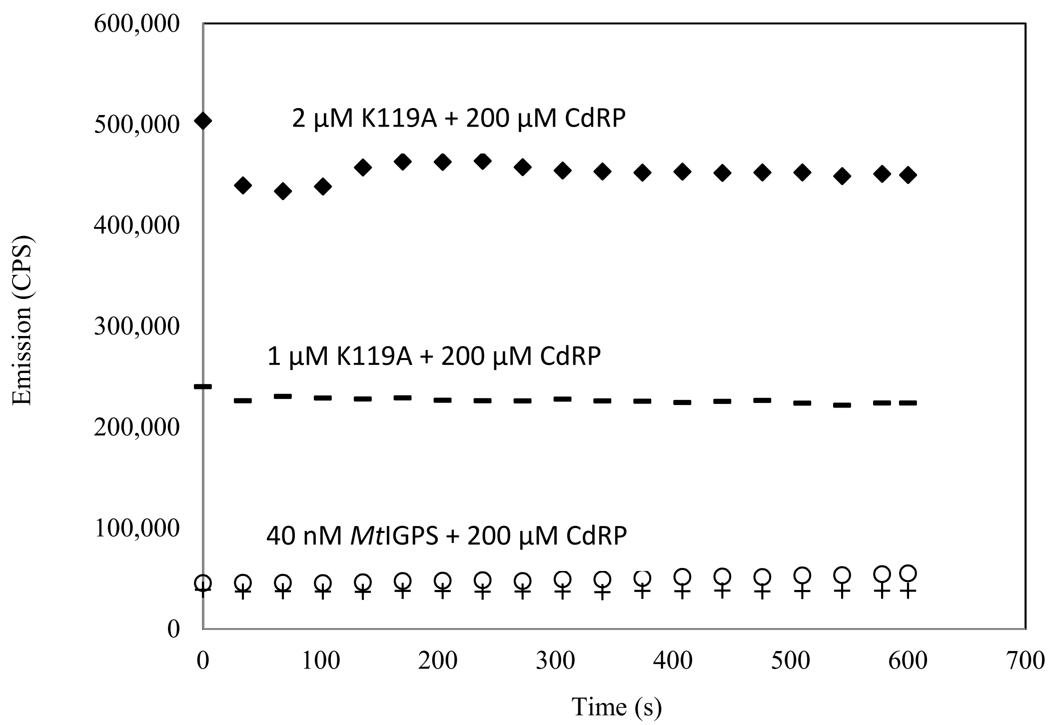
The fluorescence emission (280 nm excitation, 5 nm slit; 340 nm emission, 10 nm slit) attributed to the formation of IGP was measured for 600 seconds for a solution of 40 nM *Mt*IGPS and 200 μM CdRP (Figure 8A, circles), and 200 μM CdRP only (Figure 8B, triangles). Both reactions proceeded in 100 mM PIPES, 1 mM DTT, and 5 mM EDTA at pH 7.5. For the *Mt*IGPS-catalyzed reaction, a steady increase in fluorescence was observed, which indicated the increase in IGP product formation. For the CdRP only control reaction (Figure 8B, triangles), there was little to no change in fluorescence, which was expected because there was no enzyme present in the reaction. The slight increase in fluorescence could be attributed to the decomposition of CdRP. The activity of K119A during the conversion of CdRP to IGP was quite difficult to investigate using UV-VIS and fluorescence spectroscopy; we were unable to measure the K_M and the k_{cat} values for K119A. With increasing amounts of K119A in the assay, there was no increase in the fluorescence emission over the course of 10 minutes. However, the initial fluorescence emission increases with the enzyme concentration and remains steady over the course of the reaction (Figure 8C). Interestingly, UV-VIS spectroscopy data showed an increase in absorbance over a period of 10 minutes for a solution of 3 μM *Mt*IGPS and 600 μM CdRP (Figure 6A, triangle), suggesting that the reaction was less likely to be completed prior to data collection. The increase in initial fluorescence emission with increasing enzyme concentration could be due to the increase in emission from tyrosine and possibly phenylalanine in *Mt*IGPS[26]. *Mt*IGPS lacks tryptophan but has four phenylalanine and three tyrosine residues. Teale and Weber[26] found that tyrosine have two-thirds of the tryptophan fluorescence and tyrosine shows strong excitation at 280 nm and emission at 340 nm. On the other hand, phenylalanine has weak excitation and emission at 280 and 340 nm respectively, and is expected to have a minor effect on the fluorescence emission being recorded by the instrument.

A



60,000
50,000
40,000
30,000
20,000
10,000
0
Emission (CPS)
40 nM *Mt*IGPS + 200 µM CdRP
200 µM CdRP Only
0
100
200
300
400
500
600
700
Time (s)

B



600,000
500,000
400,000
300,000
200,000
100,000
0
Emission (CPS)
2 µM K119A + 200 µM CdRP
1 µM K119A + 200 µM CdRP
40 nM *Mt*IGPS + 200 µM CdRP
0
100
200
300
400
500
600
700
Time (s)

C

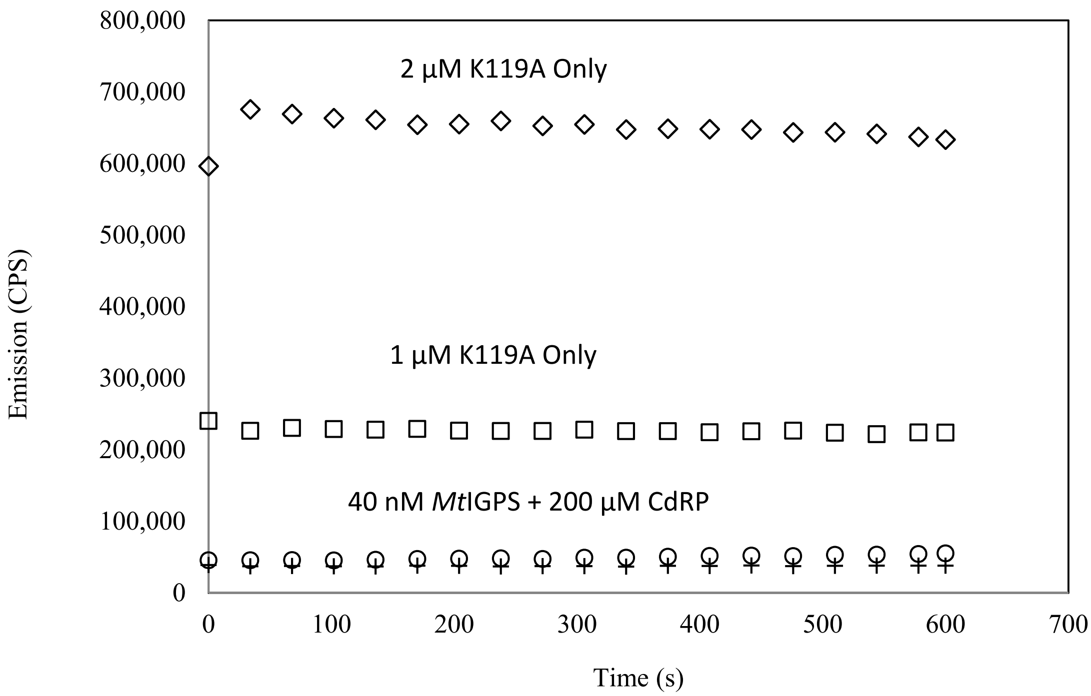
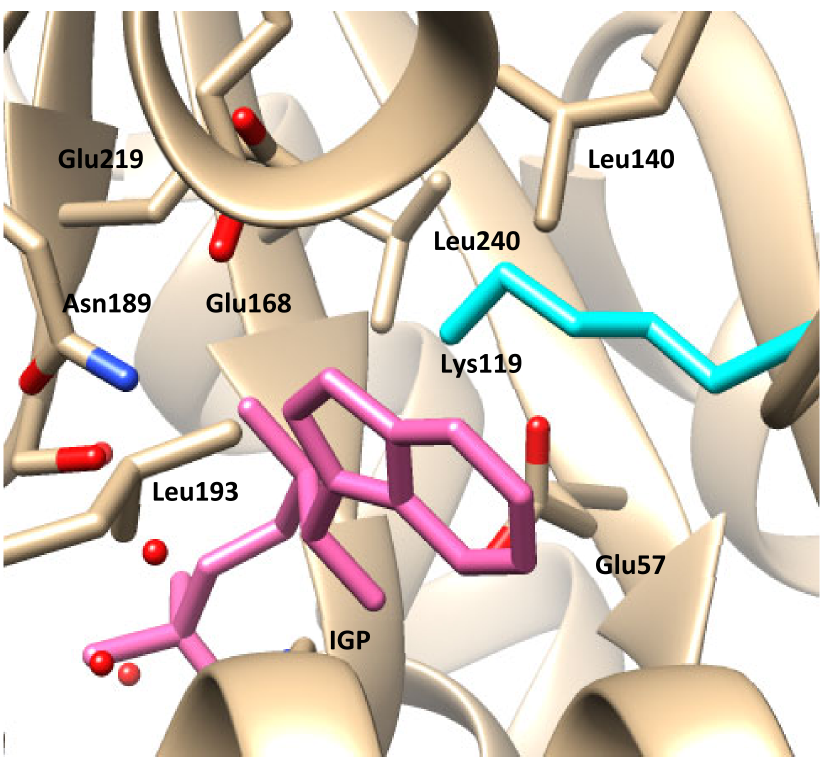


Figure 8: (A) Fluorescence emission (280 nm excitation, 5 nm slit ; 340 nm emission, 10 nm slit) versus time in seconds for the conversion of CdRP to IGP catalyzed by *Mt*IGPS over the period of 600 seconds of a solution of 40 nM *Mt*IGPS and 200 μM CdRP (circles), and 200 μM CdRP only(triangles). Both reactions occurred in 100 mM PIPES, 1 mM DTT, and 5 mM EDTA at pH 7.5. (B) Reactions catalyzed by K119A after the incubation of: 2 μM K119A and 200 μM CdRP (diamonds), 1 μM K119A and 200 μM CdRP (dashed lines), 40 nM *Mt*IGPS + 200 μM CdRP (open circles), and 200 μM only (crosses). Little to no increase in fluorescence emission was observed for K119A reactions relative to the CdRP only control. (C) Control reactions for: 2 μM K119A (open diamonds), 1 μM K119A (open squares), 40 nM *Mt*IGPS + 200 μM CdRP (open circles), and 200 μM only (crosses).

The results of the enzymatic assays show that the Lys119 residue in *Mt*IGPS may be important for catalysis. No increase in fluorescence emission or absorbance due to the formation of IGP was detectable despite significant increases in the enzyme concentration, which made the k_{cat} and K_M determination quite challenging. These findings suggest that Lys119 could be important for catalysis and is consistent with the *ss*IGPS proposed mechanism[5]. Lys119 lies in close proximity to the product IGP which suggests that it could be involved in binding or catalysis (Figure 9A). An overlay of the crystal structures for *ss*IGPS and *Mt*IGPS shows the similarity between the residues in the binding pocket, which could have similar functions in both strains of IGPS (Figure 9B).

A



B

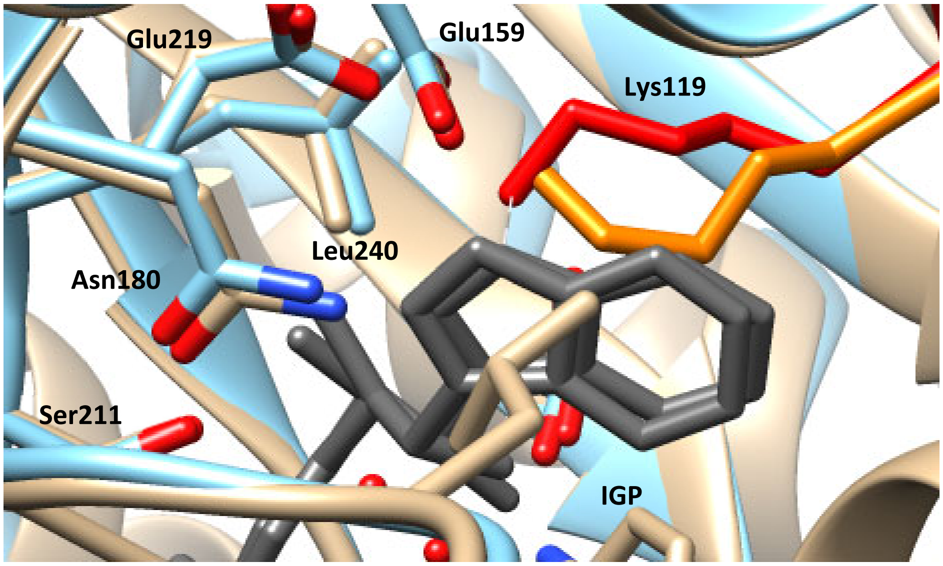


Figure 9: (A) Crystal structure of *Mt*IGPS in complex with IGP and anthranilate (PDB# 3T44) showing interactions between IGP and *Mt*IGPS residues. Lys119 (blue) is 3.54Å from IGP (pink). (B) Structural alignment of *ss*IGPS (blue; PDB# 1A53) and *Mt*IGPS (brown; PDB# 3T44). The amino acid residues are numbered according to *Mt*IGPS; Lys119 (red) and Lys110 (orange; *ss*IGPS) are shown in close proximity to the IGP product (dark grey). Despite only sharing 29% amino acid identity, *ss*IGPS and *Mt*IGPS show strong structural similarity.

The results from the enzymatic assay is consistent with the catalytic mechanism for *ss*IGPS (Figure 3B) proposed by Zaccardi *et al.* in 2013[4]. According to the proposed mechanism, a mutation to alanine is expected to prevent the protonation of the carbonyl of CdRP, preventing IGP formation[4-5]. In addition, the analysis of the crystal structure of *Mt*IGPS complexed with IGP (Figure 9A) shows Lys119 in close proximity to the product IGP (3.54Å), in a position close enough to possibly form a π-cation interaction with the indole ring and may play a role in the mechanism. Recently, Dr. Nina Goodey suggested that there are two binding surfaces in IGPS, one for CdRP and another for IGP. Based on this information, the mutation from lysine to alanine could have affected the substrate or product binding, and the structural rearrangements during catalysis. This would make it quite difficult for *Mt*IGPS to convert CdRP to IGP which is reflected in the upper limit for the k_{cat} for K119A, which was at least 17 fold lower than the k_{cat} value for *Mt*IGPS (Table 3). In future studies, an unreactive substrate analogue, rCdRP will used to investigate the importance of Lys119 during ligand binding as described[2]. The reduced CdRP can be synthesized by selective reduction of CdRP with $NaBH_4$[27]. *Mt*IGPS lacks tryptophan but binding parameters may be determined using intrinsic tyrosine fluorescence owing to the fact that its quantum yield is high enough to give a good fluorescence signal[26]. During binding events, changes in protein folding can be studied because their quantum yield is sensitive to its environment which changes when protein folds/unfolds.

To elucidate the importance of the positive charge on Lys110 and its ability to act as a catalytic acid, mutations to Arg and Gln at that position can be made; Arg retains the positive charge and has the ability to donate a proton, whereas Gln lacks both properties but is able to participate in CdRP binding via hydrogen bonding with the carboxylate moiety. As a result, a mutation to arginine would be expected to have no effect on the enzyme-catalyzed reaction and k_{cat} value is expected to be similar to the value for *Mt*IGPS. On the other hand, if no change in activity is found (represented by the k_{cat} value), this might indicate that the position may be important for binding or orientation of the substrate and/or product. If the mutation to Gln results in a decrease in the k_{cat} value, this might suggest that a positive charge and/or the ability for Lys-119 to act as a general acid are critical for *Mt*IGPS chemical mechanism. However, no effect on the k_{cat} would suggest that neither property is important for catalysis.

In summary, we characterized Lys119 that is suspected of forming important interactions with CdRP to further understand the interactions between IGPS and its ligands. We found that a mutation to alanine at that position had a significant effect on the *Mt*IGPS-catalyzed reaction. The upper limit for the k_{cat} for K119A was found to be 0.00121 s^{-1}, and is less than 17 fold lower than k_{cat} value for *Mt*IGPS (0.021 ± 0.003 s^{-1}). This finding suggests that Lys119 is important for catalysis and is consistent with the previously proposed mechanisms (Figure 3B and 3C). In addition, the results shows the importance of the carboxylate moiety on C2' of CdRP, which provides an increased understanding of binding interactions and structure-function relationships in IGPS and are expected to contribute to the understanding of the mechanism of *Mt*IGPS.

References

1. Yang, Y., Zhang, M., Zhang, H., Lei, J., Jin, R., Xu, S., ... & Wang, H. (2006). Purification and characterization of Mycobacterium tuberculosis indole-3-glycerol phosphate synthase. *Biochemistry (Moscow)*, *71*(1), S38-S43.
2. Schlee, S., Dietrich, S., Kurćon, T., Delaney, P., Goodey, N. M., & Sterner, R. (2012). Kinetic mechanism of indole-3-glycerol phosphate synthase. *Biochemistry*, *52*(1), 132-142.
3. Ramaswamy, S. and Musser, J.M., 1998. Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. Tubercle and Lung disease, 79(1), 3-29.
4. Ramaswamy, S., & Musser, J. M. (1998). Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. *Tubercle and Lung disease*, *79*(1), 3-29.
5. Zaccardi, M. J., O'Rourke, K. F., Yezdimer, E. M., Loggia, L. J., Woldt, S., & Boehr, D. D. (2014). Loop-loop interactions govern multiple steps in indole-3-glycerol phosphate synthase catalysis. *Protein Science*, *23*(3), 302-311.
6. Smith, D. A., Parish, T., Stoker, N. G., & Bancroft, G. J. (2001). Characterization of Auxotrophic Mutants of Mycobacterium tuberculosis and Their Potential as Vaccine Candidates. *Infection and immunity*, *69*(2), 1142-1150.
7. Doy, C. H. (1966). Chemical Synthesis of the Tryptophan Path Intermediate l-(o-Carboxyphenylamino)-l-deoxy-D-ribulose 5-Phosphate. *Nature*, *211*(5050), 736.

8. Czekster, C. M., Neto, B. A., Lapis, A. A., Dupont, J., Santos, D. S., & 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.
9. Kirschner, K., Szadkowski, H., Jardetzky, T. S., & Hager, V. (1987). [48] Phosphoribosylanthranilate isomerase—indoleglycerol-phosphate synthase from Escherichia coli. In *Methods in enzymology* (Vol. 142, pp. 386-397). Academic Press.
10. Shen, H., Wang, F., Zhang, Y., Huang, Q., Xu, S., Hu, H., ... & 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.
11. Gomez, G. B., Dowdy, D. W., Bastos, M. L., Zwerling, A., Sweeney, S., Foster, N., ... & Knight, G. M. (2016). Cost and cost-effectiveness of tuberculosis treatment shortening: a model-based analysis. *BMC infectious diseases*, *16*(1), 726.
12. Henn-Sax, M., Höcker, B., Wilmanns, M., & Sterner, R. (2001). Divergent Evolution of (β/α) 8-Barrel Enzymes. Biological chemistry, 382(9), 1315-1320.
13. "Fight to End Tuberculosis | General Assembly of the United Nations." United Nations, United Nations, 26 Sept. 2018, www.un.org/pga/73/event/fight-to-end-tuberculosis/.
14. Stop, T. B. (2015). Partnership: The Paradigm Shift. 2016-2020. Global Plan to End TB. United Nations Office for Project Services, UNOPS, Geneva Switzerland.
15. "Tuberculosis (TB)." World Health Organization, World Health Organization, www.who.int/news-room/fact-sheets/detail/tuberculosis.
16. Cole, S., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., ... & Tekaia, F. (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature*, *393*(6685), 537.

17. Creighton, T.E. and Yanofsky, C., 1966. Indole-3-glycerol phosphate synthetase of Escherichia coli, an enzyme of the tryptophan operon. Journal of Biological Chemistry, 241(20), 4616-4624.
18. Smith, O.H. and Yanofsky, C., 1962. [107] Enzymes involved in the biosynthesis of tryptophan. In Methods in enzymology (Vol. 5, pp. 794-806). Academic Press.
19. 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.
20. Parry, R.J., 1971. Biosynthesis of compounds containing an indole nucleus. Chemistry of Heterocyclic Compounds: Indoles, Part Two, 25, 1-64.
21. Sassetti, C. M., Boyd, D. H., & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular microbiology*, *48*(1), 77-84.
22. Czekster, C. M., Lapis, A. A., Souza, G. H., Eberlin, M. N., Basso, L. A., Santos, D. S., & 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.
23. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607
24. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, *25*(13), 1605-1612.

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

26. Teale, F. W. J., & Weber, G. (1957). Ultraviolet fluorescence of the aromatic amino acids. *Biochemical Journal*, *65*(3), 476.

27. Bisswanger, H., Kirschner, K., Cohn, W., Hager, V., & Hansson, E. (1979). N-(5-Phosphoribosyl) anthranilate isomerase-indoleglycerol-phosphate synthase. 1. A substrate analog binds to two different binding sites on the bifunctional enzyme from Escherichia coli. *Biochemistry*,*18*(26), 5946-5953.