THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF CHEMISTRY

LOOP DYNAMICS IN THE FUNCTION OF E. COLI INDOLE GLYCEROL PHOSPHATE SYNTHASE

AJAY PADAKI

Spring 2011

A thesis
submitted in partial fulfillment
of the requirements for
a baccalaureate degree
in Chemical Engineering
with honors in Chemistry

Reviewed and approved* by the following:

David Boehr
Assistant Professor of Chemistry
Thesis Supervisor

Kenneth S. Feldman
Professor of Chemistry
Honors Adviser

Howard M. Salis
Assistant Professor of Chemical Engineering
Faculty Reader

*Signatures are on file in the Schreyer Honors College.

ABSTRACT

A co-evolving interaction occurs between the β1α1 and β2α2 loops of *Escherischia coli* indole-3-glycerol phosphate synthase (ecIGPS) that may help direct enzyme catalysis. The interaction specifically occurs between Lys56 (β1α1) and Gln94 (β2α2) and is hypothesized to function in rearranging the extended formation of the 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5–phosphate (CdRP) substrate. To gain more insight into this interaction, Lys56 and Gln94 were separately mutated to Ala. The mutant enzymes were characterized in terms of their Michaelis-Menten steady-state kinetics, pH dependence, and solvent viscosity effects. While the k_{cat} Lys56Ala IGPS decreased only 13.6% compared to wild-type (WT) IGPS, the K_M, a measure of substrate binding affinity, was almost four times weaker than the WT value. In contrast, Gln94Ala IGPS displayed a slightly higher K_M (41.7% greater) than the WT IGPS value, but displayed a much lower k_{cat} (70.6% lower than the WT). The resulting catalytic efficiencies (k_{cat}/K_M) of the mutant IGPS enzymes were similarly low with a 4.5 fold decrease compared to WT IGPS. The increase in the Gln94Ala IGPS k_{cat} of 555% as the temperature increased from 25 ^oC to 37 ^oC resulted in a more similar value to the WT IGPS. The tremendous impact of the temperature on the mutant's catalysis speed suggests that conformational flexibility may help to mitigate the hindrances imposed by the mutated residue. While distinctions between the WT IGPS and mutant IGPS were observed in Michaelis-Menten dynamics, no such trends were observed in the pH or solvent viscosity effect (SVE) enzyme assays. Both Gln94Ala and WT IGPS displayed a bell-shaped pH dependence with pKa values close to 6.8 and 8.7 for the ascending and descending limbs of the pH rate profile, respectively. The distinct limbs are consistent with general base and acid catalysis, as previously proposed for the IGPS enzymatic mechanism. The lack of any observable SVE indicates the chemical step is likely the rate-determining step, which simplified

the analysis of k_{cat} and K_M values. The low K_M of the Lys56Ala mutant and the slow catalysis of the Gln94Ala mutant suggest that the interaction between the β1α1 and β2α2 loops is intimately involved in the binding and ring enclosure of the CdRP substrate.

Table of Contents

Acknowledgements

I would like to thank Dr. David Boehr for his patience in teaching me the diverse range of laboratory techniques I have utilized over the past year of research. I would also like to thank the entire Boehr laboratory group for assisting me with procedures - especially Jennifer Axe, who patiently walked me through the entire protein purification protocol.

Additionally, I would like to thank my parents for encouraging me to pursue the sciences and my brother for helping me balance the medical school application process with the honors thesis workload.

Chapter I – Introduction

The tryptophan biosynthesis pathway generates the aromatic amino acid through a series of systematic reactions. Proper expression of this pathway is essential to the survival of bacteria[1] in the absence of environmentally accessible tryptophan. Interestingly, three consecutive steps of the biosynthesis pathway utilize enzymes with similar overall structures (Figure 1). Specifically, N-(5-phospho-β-D-ribosyl)-anthranilate isomerase (PRAI), indole-3-glycerol phosphate synthase (IGPS), and the α-subunit of tryptophan synthase (αTS) catalyze the synthesis of indole from N-(5-phospho-β-D-ribosyl)-anthranilate. Indole is subsequently converted to tryptophan by the β-subunit of tryptophan synthase (βTS) in the final step of catalysis[2]. PRAI, IGPS, and αTS all belong to the $(\beta/\alpha)_8$ barrel family of proteins, which are also known as triosephosphate isomerase (TIM) barrel proteins. The TIM-barrel structure is present in approximately 10% of proteins whose structures have been identified, making it the most common fold currently known[3]. Monomeric TIM-barrel proteins typically consist of between 200 and 250 amino acid residues with the central eight β-strands surrounded by eight α-helices[3].

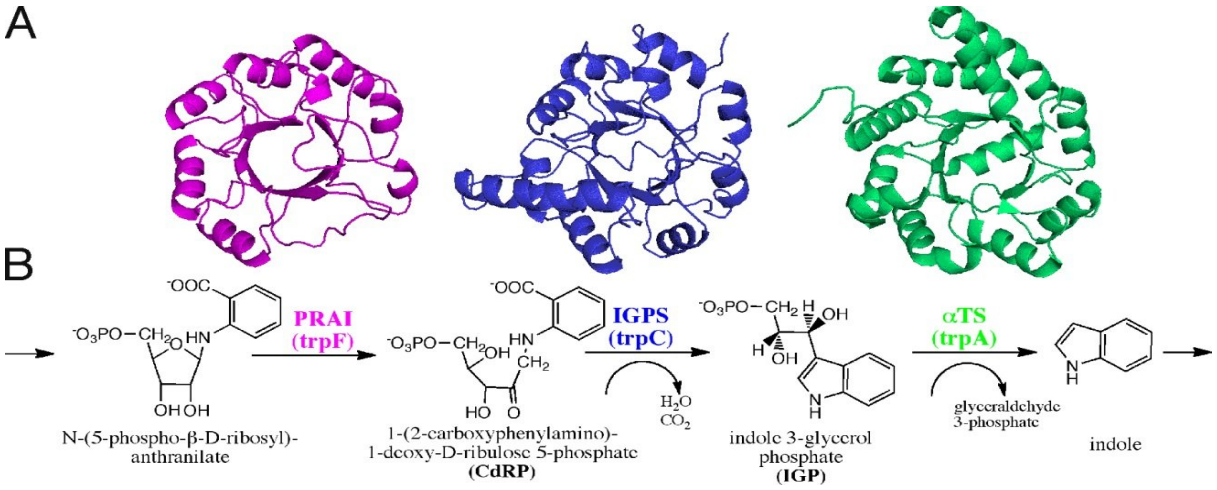

Figure 1: The structure (A) and catalyzed reactions (B) of the TIM-barrel proteins involved in the tryptophan biosynthesis pathway.

The evolutionary relationship between these three TIM-barrel proteins has been acutely studied in *Escherichia coli*. The superimposition of the carbon backbones of PRAI, IGPS and αTS show a very strong structural conservation despite the minimal amino acid sequence conservation[4]. In addition to the overall similarities, the phosphate-binding site in particular has shown incredible structural conservation. With approximately a 15% conserved amino acid sequence in the three enzymes' phosphate binding site region (comprised of the β7α7 and β8α8 N-termini), Wilmanns et al found a structural conservation of between 50 and 75%.

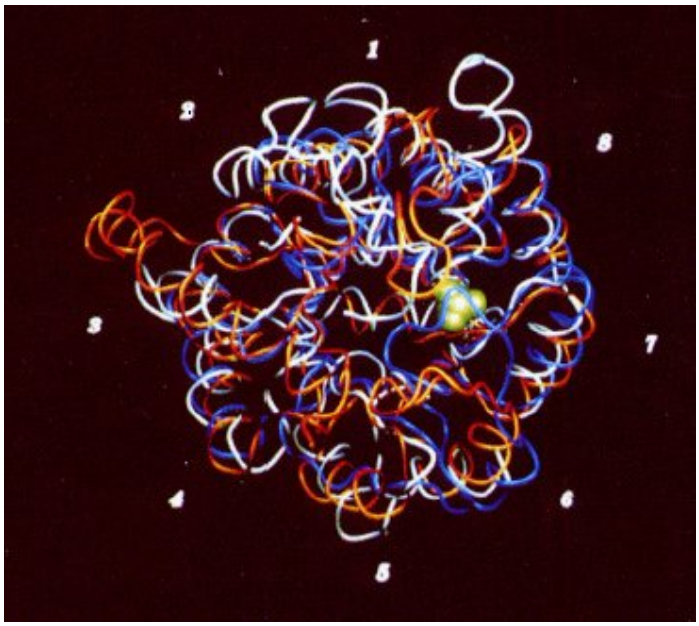

Figure 2: Superimposed carbon backbones of PRAI (blue), IGPS(red), and αTS (white) with alpha helices numbered 1-8[4].

The IGPS enzyme is particularly interesting because it has been found to be absolutely required for *Mycobacterium tuberculosis*[6], suggesting that it is potentially a novel anti-tuberculosis drug target. The specific role of IGPS in the tryptophan biosynthesis pathway is ring closure of 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5 –phosphate (CdRP) to indole 3-glycerol phosphate

(IGP). IGPS has been extensively studied to locate amino acid residues crucial to catalysis. The C-termini of the β7 and β8 strands and the C-termini of the β1α1, β3α3 and β6α6 compose the catalytic pocket within IGPS. The binding pocket consists of salt bridges between residues K55, E53 and K114 on the phosphate ion side and hydrogen-bonding between the anthranilate and the residues E163 and N184 in ecIGPS[7]. The corresponding salt bridge in ssIGPS consists of the K53, E51 and K110 residues while the corresponding hydrogen bonds occur between E159, N180 and S210. The selectivity of the IGPS active sites in only utilizing the C-termini of β-strands is consistent with the observation that all known TIM-barrel proteins solely use this region to formulate the catalytic site[8].

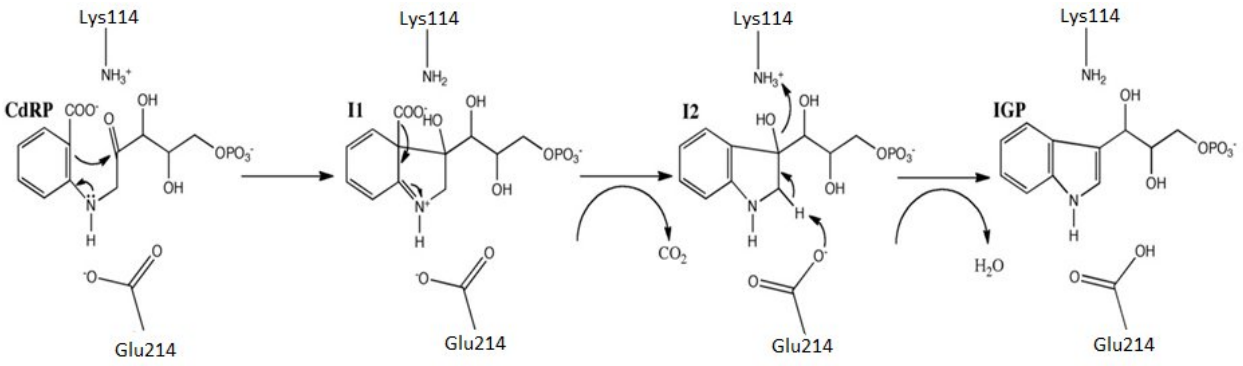

Figure 3: *E. coli* condensation, decarboxylation and dehydration of CdRP to produce indole-3-glycerol phosphate

The proposed mechanism for the IGPS catalysis involves three steps and consequently, two intermediates[9]. Initially, the newly formed carbonyl is attacked to form a five-member ring, with Lys114 providing the hydrogen to the reduced carbonyl. The carboxylate is then released as carbon dioxide, with the re-establishment of the carbon-carbon double bond also returning the lone pair of electrons back to the ring-bound nitrogen. Finally, Lys114 deprotonates a second time to

facilitate the dehydration step. In this last step, Glu214 acts as the general base by accepting a hydrogen ion from an α-carbon to the amine.

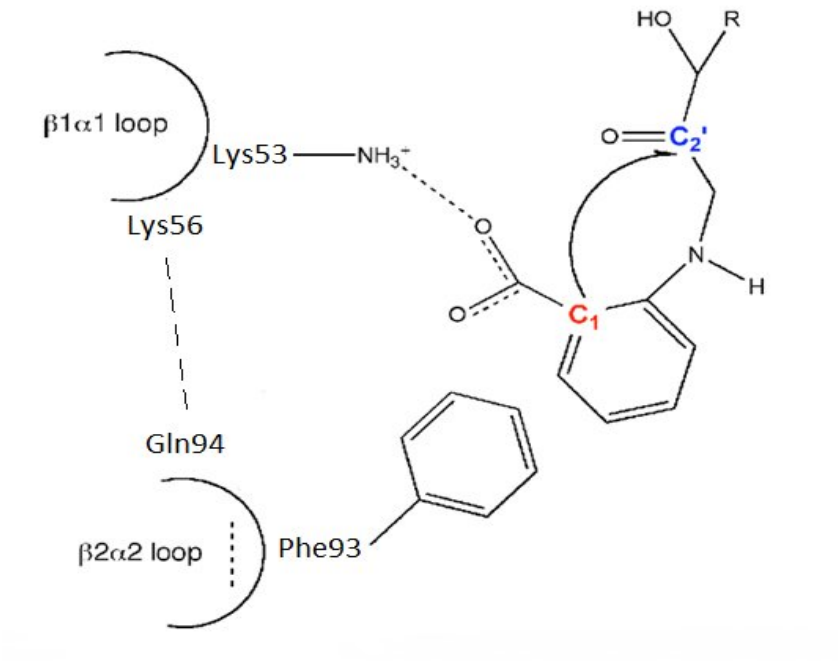

Figure 4: Proposed mechanism of *E. coli* IGPS (ecIGPS) binding to CdRP

X-ray crystallography was previously used to determine the structure of IGPS in the presence of both the substrate and a substrate analog[10,11]. The substrate analog was generated by reducing the CdRP C_2' from a carbonyl to an alcohol, thus allowing the reduced CdRP (rCdRP) to bind to the IGPS enzyme but preventing the condensation required to complete the ring enclosure. One intriguing aspect of the substrate-bound IGPS structure is that the distance between C_1 and C_2' (4.5 Angstroms) is clearly beyond the typical bonding radius between carbons (1.5 Angstrom[12]). This suggests that the substrate needs to undergo a conformational change to facilitate carbon-carbon bond formation. The β1α1 loop may be involved in closing the distance between the C_1 and C_2'. Lys53, a residue belonging to the β1α1 loop, forms a hydrogen bond with the carboxyl group of

the anthranilate ring. Amino acid residues between 55 and 59 in β1α1 loop are also involved in the binding pockets of both anthranilate and phosphate[8]. Molecular dynamic simulations have also shown that the β1α1 loop displays the greatest mobility and therefore this loop is very likely involved in the stabilization of anthranilate as it shifts conformation to bring C_1 and C_2' closer together[7]. MD simulations have also suggested that motions in the β1α1 and β2α2 loops are correlated. One interaction that may facilitate loop coordination in ssIGPS is the interaction between Arg54 (Lys56 in E. coli) in β1α1 loop and Asn90 (Gln94 in E. coli) in β2α2 loop.

To examine the relative importance of this interaction, the degree of its conservation in the IGPS of varying species is of great importance. Co-evolution measures how well conserved amino acids at a given position of an enzyme are conserved across different organisms utilizing the enzyme. While the overall structure of many proteins is often conserved, the co-evolution of residues in enzymes involved in amino acid synthesis has been shown to be especially robust[13]. Studying the co-evolution of amino acids allows for further insight into the folding mechanisms of proteins and the catalytic residues in enzymes[14]. Co-evolution studies also allow for the examination of conserved motions within conformational changes that occur in enzymes during catalysis[15]. While differing methods exist to study the co-evolution of residues, statistical coupling analysis is often performed to detect the specific amino acids involved in catalysis[16]. Statistical coupling analysis specifically analyzes whether changing a given amino acid will cause a corresponding change in a neighboring amino acid within a family of proteins.

The preserved binding region involving the β1α1 and β2α2 loops in ecGIPS and ssIGPS also applies to a vast number of IGPS encoded in other organisms. Examination of over 400 different IGPS's displayed that approximately 15% of the protein co-evolved, with almost the entirety of the

co-evolution occurring in the active site formed by the C-termini of β-strands[17]. Molecular dynamic simulations done by Shen et al also have shown that the movements that take place in the binding and release of CdRP is also conserved within the binding site of IGPS. Antibiotics developed to target this antigenic site would therefore allow for diverse applications and target a broader range of infectious species which are dependent on tryptophan biosynthesis.

To determine the importance of the β1α1 - β2α2 interaction, site-directed mutagenesis was performed on *E. coli* IGPS. The Lys56 and Gln94 residues, which constitute the major interaction between the loops, were separately mutated to Ala. Following purification of the mutated enzymes, assays were performed to analyze the kinetics of the mutants with respect to the wild-type (WT) IGPS. Strong variations in Michaelis-Menten parameters would indicate that these residues are crucial in the binding and catalysis of CdRP. If the kinetic dynamics were altered sufficiently, the β1α1 - β2α2 interaction would thus represent a region of the enzyme essential to catalytic function.

Chapter II – Materials and Methods

Site directed mutagenesis

The QuickChange protocol was utilized to create the mutant plasmids expressing Lys56Ala and Gln94Ala IGPS, according to the procedure supplied by the manufacturer. Following PCR mutagenesis, overnight mutant plasmid transformation and *E.coli* growth on LB-kanamycin (50 mg/mL) agar at 37^{o}C, single colonies were used to inoculate 5 mL LB-kanamycin media. DNA was then isolated and sequenced to confirm the mutagenesis.

Protein purification

For protein overexpression, *E.coli* BL21(DE3) carrying plasmids expressing WT, Lys56Ala or Gln94Ala IGPS were grown in 5 mL of LB-kanamycin media overnight and then used to inoculate 1 L LB-kanamycin media. Cultures were allowed to grow for approximately 3 hours at 37ºC, until the optical density (A_{600}) reading reached 0.60. Protein overexpression was induced by the addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG). After induction, the culture was incubated at 25ºC for 16 hours in a shaker. The culture was then centrifuged (10 000xg for 10 minutes), and the pellet was dissolved in 35 mL of vacuum-filtrated buffer A [25 mM HEPES buffer, pH = 7.5, 1 mM ethylenediamenetetraacetic acid (EDTA)]. Following cell lysis by sonication, the sample was centrifuged again (30 000xg for 30 minutes) to separate cellular debris from the supernatant containing IGPS enzyme.

The dialyzed supernatant was then run through a 20 mL Q-sepharose anion exchange chromatography column. The buffer strength was varied over time to increase the salt concentration. Specifically, a gradient that initially began as 100% buffer A was altered to a 100%

buffer B solution [1M NaCl, 25 mM HEPES buffer, pH = 7.5, 1 mM ethylenediamenetetraacetic acid (EDTA)]. Fractions were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to locate the IGPS. Selected fractions were pooled, dialyzed against buffer A overnight to remove NaCl, and then run through a 1 mL Mono-Q anion exchange chromatography column which also utilized buffer A and B to provide the necessary salt gradient. Fractions containing IGPS were pooled, concentrated and then loaded onto a Sephacryl S100 gel filtration (GF) chromatography column. Purified IGPS was diluted to 25 μM with 50 mM HEPPS pH 7.5, 4 mM EDTA.

Enzyme assays

Michaelis-Menten assays, pH effect assays, and solvent viscosity effect (SVE) assays were performed on all IGPS enzymes (WT, Gln94Ala, Lys56Ala) using a 96-well microtitre plate assay. The final volume for each assay was 250 μL (usually 10 μL 25 μM IGPS, 10 μL CdRP, 230 μL HEPPS buffer). A final 1 μM IGPS concentration was used for all assays, and all assays were run in triplicates. To measure the rate of IGP production, assays were measured at 278 nm for 3-5 minutes. Most assays were conducted at both 25°C and 37°C. Enzyme and buffer were allowed to equilibrate at test temperature for four minutes before CdRP was added to initiate the assay.

For the Michaelis-Menten steady-state kinetic assays, substrate concentrations were varied between 0.125 and 12 μM and assays were performed at both 25ºC and 37ºC. Assays testing pH dependence were conducted using a saturating concentration of CdRP (12 μM) with different buffers used to vary the pH between 5.5 and 10 (100 mM MES buffers at pH 5.5, 6 and 6.5; 100 mM HEPES buffers at pH 6.5, 7.0, 7.5, and 8.0; 100 mM CHES buffers at pH 8.5, 9.0, 9.5, and 10.0). Solvent viscosity effect assays were also conducted using 12 μM CdRP with glycerol as

the viscogen (varied between 0 and 30%). Viscosities were measured using an Ostwald viscometer, according to manufacturer instructions.

Chapter III – Results and Discussion

The IGPS domain of the PRAI-IGPS complex was over-expressed and purified for both the WT enzyme and the Gln94Ala and Lys56Ala mutants. Various enzyme assays were performed to analyze the importance of the β1α1 and β2α2 interaction with respect to substrate binding and catalysis. Michaelis-Menten assays tested these parameters while the pH effect assays tested for acid-base catalysis and the solvent viscosity effect assays analyzed the importance of diffusion-limiting steps to the catalytic rate. A saturating substrate concentration of 12 μM CdRP was utilized for all assays other than those testing Michaelis-Menten parameters.

Overexpression and Purification of ecIGPS:

Despite previous protocols utilizing phosphate-based buffers for chromatography[18], HEPES buffer was used for all chromatography steps. While Tris-HCl was initially employed, the IGPS did not bind to the Q-sepharose column and eluted rapidly. SDS-PAGE confirmed that the majority of the protein eluted immediately into the 'flowthrough' beaker. By using the identical protocol to overexpress IGPS but instead using HEPES, a much stronger separation was achieved. SDS-PAGE confirmed that the IGPS was not lost to the 'flowthrough' and that the Q-sepharose adequately bound the IGPS.

While significant IGPS was lost throughout the protein purification protocol, the concentrations produced were ample for kinetic studies using enzyme assays. Specifically, IGPS was lost after sonication, as gel electrophoresis testing showed that the pellet presented a strong band at the IGPS molecular weight. SDS-PAGE also displayed similar band intensity for the supernatant and pellet, which reveals that a much higher concentration of IGPS could have been achieved. Further spin-

concentration could increase the 25 µM concentrations of IGPS used in enzyme assays to the 200 µM, suggesting the potential for future NMR studies based on these constructs.

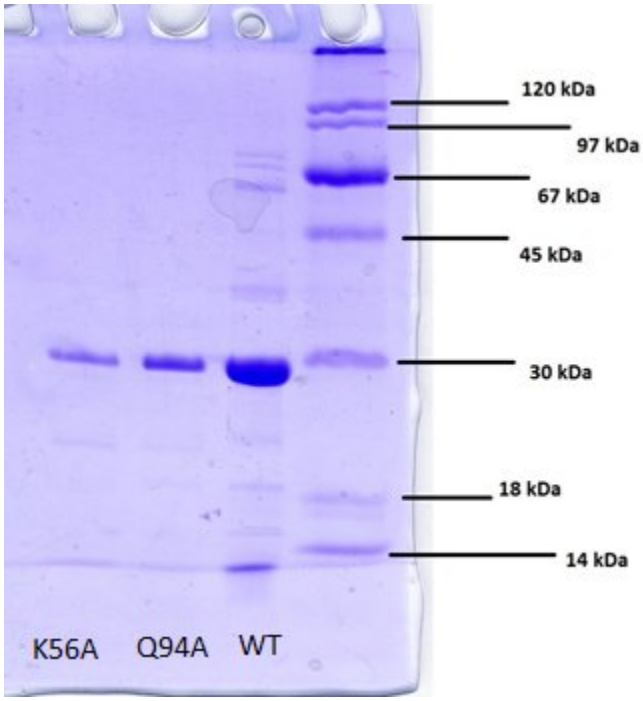

Figure 5: SDS-PAGE of WT, Gln94Ala, and Lys56Ala at a concentration of 29.7 µM

Following purification, all three samples were spin-concentrated to a concentration of 29.7 µM and Figure 5 shows the side-by-side comparison. Despite the appearance of more impurities in the WT lane, the most IGPS is also present in this lane.

Steady-state kinetics of WT, Gln94Ala, and Lys56Ala ecIGPS:

The kinetic parameters obtained for the WT IGPS closely match previous studies. Past research has estimated an *E. coli* IGPS k_{cat} between 2.2 and 7.2 s^{-1} and has estimated the K_M value between 0.34 and 1.2 µM[19]. While the 4.08 s^{-1} k_{cat} observed at 25°C fits within the given range, the

temperature dependence of the Michaelis constant is unknown for IGPS. Therefore, the accuracy of the 1.63 μM value obtained at 37°C cannot be determined.

The catalytic rate of IGPS and the two mutants were also tested with varying substrate concentrations (CdRP) to determine Michaelis-Menten kinetic parameters.

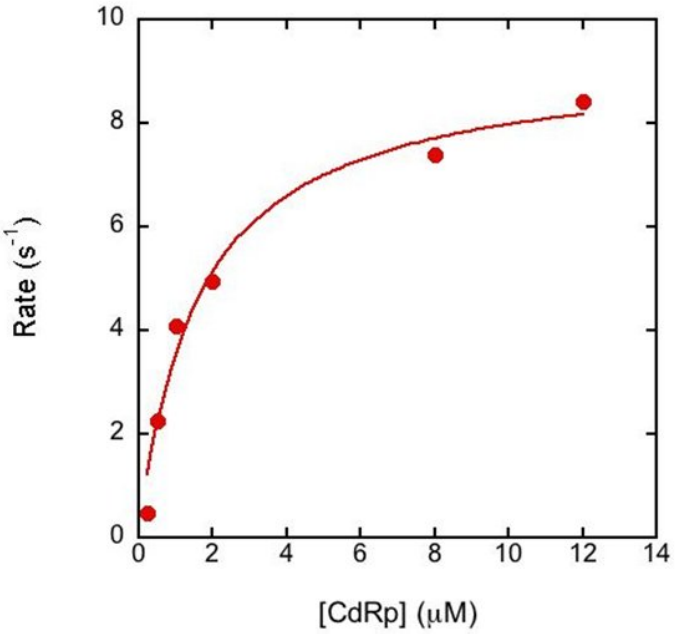

**Figure 6: Michaelis-Menten curve for WT IGPS (37 °C) **

The k_{cat} of the WT IGPS was found to be 9.78 times greater than the Gln94Ala mutant when both rates were measured at 25°C and 3.40 times greater when the rates were measured at 37 °C. The catalytic efficiency of the WT IGPS was 4.82 times greater than the Gln94Ala mutant at 37 °C. In contrast, the kinetic parameters of the Lys56Ala mutant were only tested at 37 °C due to a lower recovery of this mutant. The WT IGPS k_{cat} was found to be 1.16 times greater than the mutant's value at this temperature. The WT IGPS catalytic efficiency was greater than the Lys56Ala mutant by a factor of 4.34.

Table 1: Steady-state kinetic parameters for WT and Lys56Ala + Gln94Ala IGPS mutants

Protein	Temperature (°C)	$k_{cat}(s^{-1})$	K_M (μM)	k_{cat}/K_M ($x10^6 s^{-1}M^{-1}$)	$k_{cat,WT}/k_{cat,mut}$	$(k_{cat}/K_M)_{WT}/(k_{cat}/K_M)_{mut}$
WT	25	4.08± 0.20[1]				
	37	9.28± 0.60	1.63±0.34	5.69		
Gln94Ala	25	0.417± 0.07[1]			9.78	
	37	2.73± 0.34	2.31±0.72	1.18	3.40	4.82
Lys56Ala	37	8.02± 2.89	6.08±4.50	1.31	1.16	4.34

[1] estimated values using [CdRP]=12 μM

Michaelis-Menten curves showed significant differences in the kinetic constants of the WT enzyme and the Gln94Ala and Lys56Ala mutants. While increasing the temperature from 25°C to 37°C only increased the WT k_{cat} by 127%, the corresponding increase in the Gln94Ala mutant was 555%. The difference between the WT and mutant IGPS catalytic rates was significantly reduced at 37°C as the WT k_{cat} was 3.4 times greater than Gln94Ala and 1.16 times greater than Lys56Ala.

The large differential in catalytic speed between the IGPS k_{cat} values at 25°C indicates that disrupting the interactions between the β1α1 and β2α2 loops significantly reduces the catalytic activity of the enzyme. As the k_{cat} of the WT IGPS is almost 10 times greater than the k_{cat} of the Gln94Ala mutant as 25°C but only 3.4 times greater at 37 °C, higher temperatures must decrease the hindrance to catalysis that the mutated residues provide. Additionally, the similar k_{cat} of the Lys56Ala mutant to the WT k_{cat} displays that this mutation does not substantially decrease the

catalytic ability of IGPS. Therefore, the β1α1 loop is likely less crucial in the rate-limiting step of catalysis than the β2α2 loop.

Analysis of the K_M values also shows distinctive differences between the WT and mutant IGPS kinetics. The WT IGPS maintains the smallest K_M, at 1.63 μM, indicating that this IGPS has the greatest affinity with CdRP. The Gln94Ala K_M value is approximately 50% greater and the Lys56Ala value is almost 300% greater. Therefore, mutating the residues that link the β1α1 and β2α2 loops clearly decreases the affinity between IGPS and CdRP. The impact on the K_M from this mutation displays that the β1α1-β2α2 loop interaction is strongly involved in the biding of CdRP. Because the Lys56Ala mutation decreases affinity much more drastically than the Gln94Ala mutation, however, the β1α1 loop may be more important than the β2α2 concerning binding affinity. These parameters are consistent with the hypothesis that the β1α1 loop facilitates the rearrangement of the extended-conformation substrate to facilitate ring closure[20].

The catalytic efficiency of the WT and mutant IGPS enzymes were also calculated utilizing the Michaelis-Menten constants. The WT IGPS catalytic efficiency was calculated to be 5.69×10^6 $s^{-1}M^{-1}$, a value 382% greater than the Gln94Ala mutant and 334% greater than the Lys56Ala mutant. The mutants' catalytic efficiency values were found to be very similar at 1.18 (Gln94Ala) and 1.31 (Lys56Ala) $s^{-1}M^{-1}$.

Running a pH assay on the WT IGPS showed a very clear optimum pH range between 7.5 and 8, as seen in Figure 7. This range was also consistent with the Gln94Ala IGPS mutant. The distinct bell-shaped pH curve was also present for both the WT and mutant IGPS.

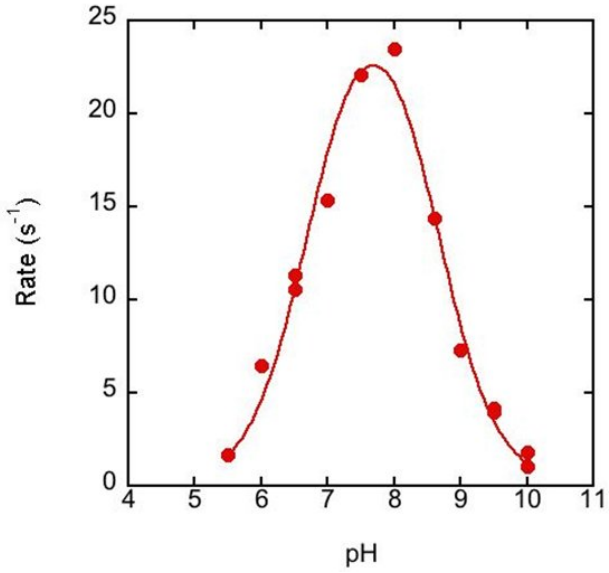

Figure 7: The pH dependence of ecIGPS catalysis at 37°C

Further analysis of the pH curve provides the two pKa points of IGPS. The changes in concavity correspond to the pKa values displayed in Table 2. The values were ascertained by employing non-linear regression to the curve fit. The pKa_1 and pKa_2 values of the Gln94Ala mutant closely mirrored the WT values.

Table 2: pKa values for ecIGPS enzymes

Protein	pK_{a1}	pK_{a2}
WT	6.70 ± 0.08	8.67 ± 0.08
Gln94Ala	7.00 ± 0.59	8.83 ± 0.59

The distinct ascending and descending limbs of the pH curve support the presence of a catalytic base and acid in the IGPS active site. The ascending portion likely corresponds to the catalytic base while the descending portion corresponds to the catalytic acid. The highest catalytic activity

occurs between these two pKa's, indicating IGPS is most efficient when the basic residue is deprotonated and the acidic residue is protonated.

In contrast to the IGPS finding, previous studies have shown a logarithmic growth of the catalysis rate as the pH is increased in *Mycobacterium tuberculosis* IGPS (mtIGPS)[18]. This distinctly differs from the bell-shaped curve of ecIGPS and suggests the importance of only the deprotonation of a residue in mtIGPS. Because the rate does not decrease with increased pH in the *Mycobacterium*, it is very likely that the dehydration step (final chemical step) is the rate-limiting step. Protonating the basic residue drastically decreases catalysis speed while pH's even above 12 do not damage the IGPS's catalytic ability. As this step is the only step in catalysis that requires proton abstraction, it is also likely the rate-limiting step in *E. coli*. The bell-shaped curve indicates that that protonation of the acidic residue is equally important in ecIGPS, however.

Kinetic studies were also performed on the WT and mutant IGPS proteins to test the solvent viscosity effect (SVE). By varying the viscosity of the solution that the IGPS is performing catalysis in, the effects of diffusion can be studied. The two diffusion-controlled steps in CdRP catalysis are the binding of CdRP and the release of IGP.

The results are displayed for the three IGPS variants in Table 3. Values are intended to vary between 0 (no SVE) and 1 (very strong SVE). The WT IGPS at 37 ^{o}C displays the strongest solvent viscosity effect with a value of 0.20. The WT enzyme also displays the lowest, non-negative SVE of 0.01 at 25^{o}C. The SVE's of the Gln94Ala at both temperatures and the Lys56Ala mutant as 25 ^{o}C displayed negative values.

Table 3: Solvent viscosity effects for WT and Gln94Ala IGPS using glycerol as a viscogen

Protein	Temperature (°C)	SVE[1]
WT	25	0.01
	37	0.20
Gln94Ala	25	-0.43
	37	-0.04
Lys56Ala	25	-0.09

[1] SVE defined as the slope of the plot relative viscosity ($\eta_o / \eta_{viscogen}$) vs. relative rate ($k_0/k_{viscogen}$)

The 0.20 value of the WT IGPS is still too low to support a significant SVE finding within the IGPS mechanism. The lack of a significant SVE within IGPS indicates that the rate-limiting step of catalysis is not diffusion-controlled. Therefore, the ring-enclosure, decarboxylation, or dehydration of CdRP represents the rate-limiting step in catalysis.

Chapter IV – Conclusion

While a portion of the overexpressed IGPS was lost in the purification protocol previously described, a sufficient amount was synthesized to perform an array of enzyme assays. Additionally, spin-concentration of the IGPS recovered could produce viable concentrations (>200 μM) for NMR studies, which can provide more insight into the conformations assumed by IGPS throughout catalysis.

The assays performed revealed various aspects about the IGPS catalytic mechanism. The lower K_M value of the Gln94Ala and Lys56Ala mutants indicates the β1α1 - β2α2 loop interaction is essential to the binding affinity between IGPS and CdRP. Additionally, a disproportionate increase in the k_{cat} values of the mutants due to the temperature increase was observed. The greater conformational flexibility of the IGPS mutant at the higher temperature appeared to mitigate the structural defect from the mutated residue, consistent with previous studies[21]. The varying results of the two mutants support molecular simulations displaying the β1α1 loop as the most flexible and most likely loop to facilitate ring closure of the CdRP.

The pH effect assays and SVE assays also revealed critical information about the IGPS mechanism. Specifically, the lack of a significant solvent viscosity effect in either the WT or mutant IGPS enzymes indicates that chemical catalysis and not a diffusion-mediated phase is the rate-limiting step. Additionally, the distinct ascending and descending limbs of the pH curve support the presence of a catalytic base and catalytic acid within the active site of IGPS.

The combination of the lack of a solvent viscosity effect and the pH curve obtained by *Mycobacterium* studies indicates that the dehydration of CdRP is the rate-limiting step. Because mutating the β1α1-β2α2 loop interaction decreases both the catalytic rate and substrate affinity, this

interaction is very likely intimately involved in CdRP catalysis. Therefore, a novel antibiotic should strive to target this antigen on IGPS to inactivate the enzyme and inhibit bacterial growth.

Chapter V – Appendix

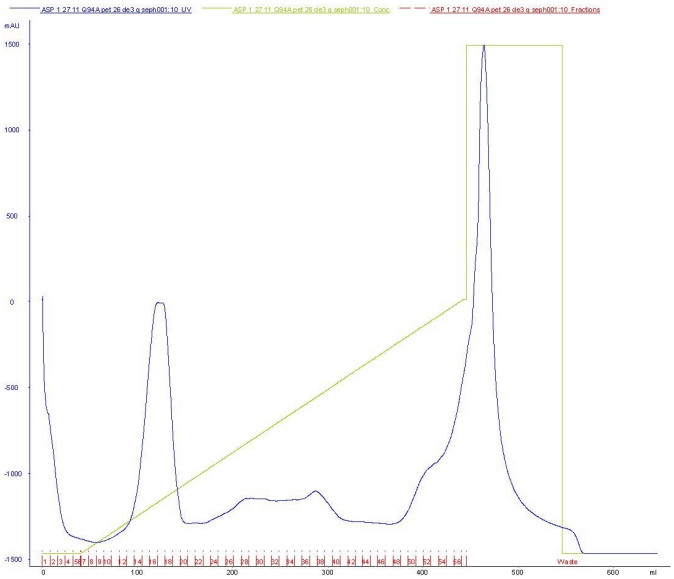

Figure 8: Q-sepharose anion exchange chromatogram of Gln94Ala

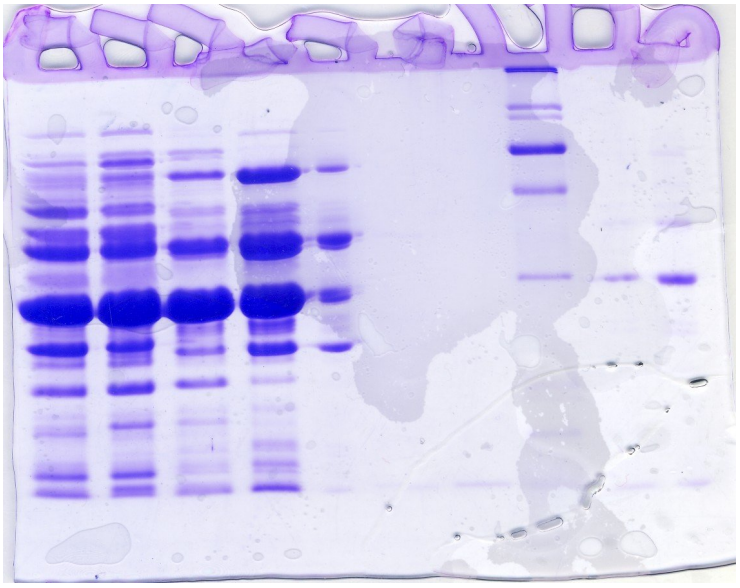

Figure 9: Polyacrylamide gel electrophoresis (PAGE) of Gln94Ala Q-sepharose

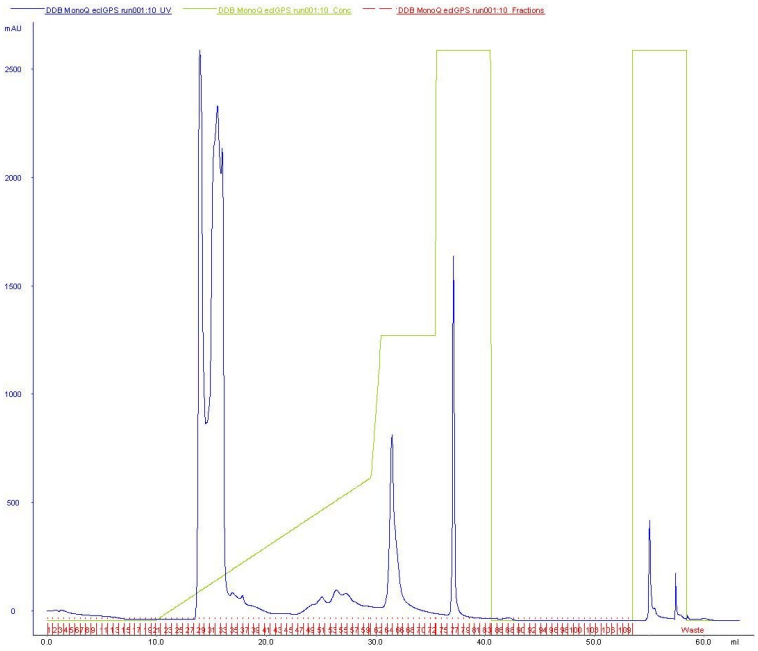

Figure 10: Mono-Q Chromatogram of Gln94Ala Q-sepharose fractions 14-19

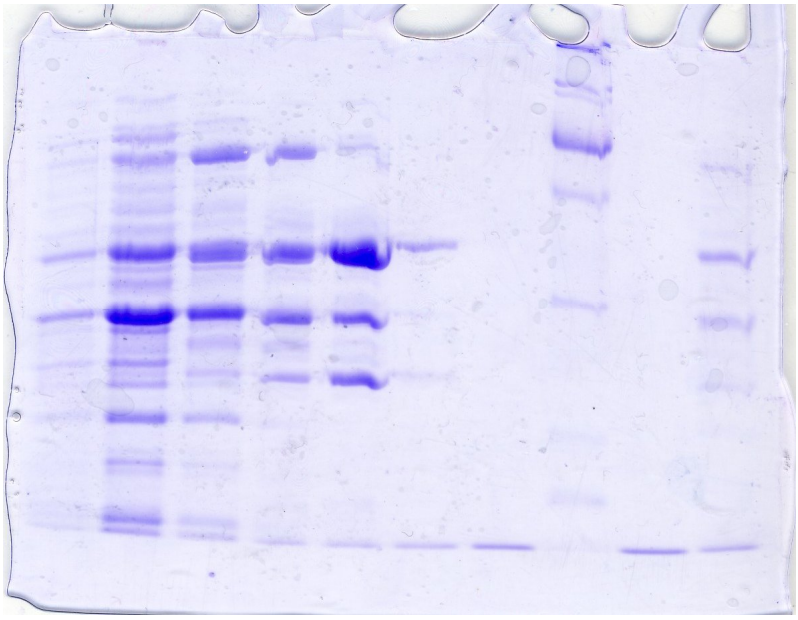

Figure 11: PAGE of Gln94Ala Mono-Q column fractions 29-33 in lanes 6-10 respectively

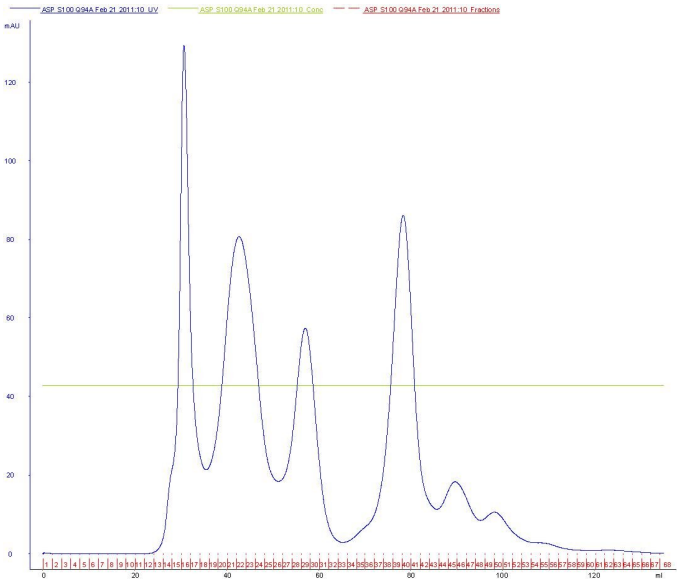

Figure 12: Gel filtration (GF) chromatogram of Gln94Ala mono-Q column fractions 29-33

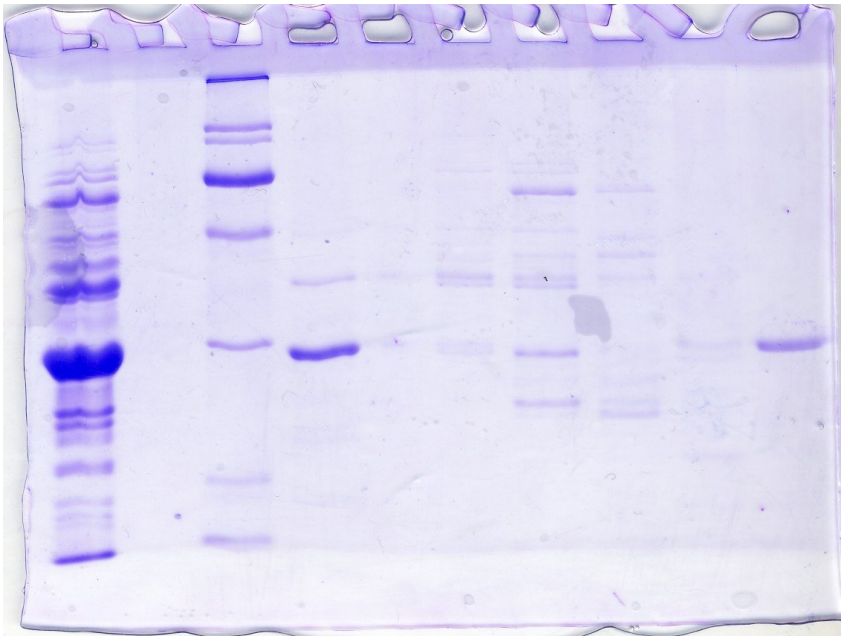

Figure 13: PAGE of Gln94Ala GF column fractions 16, 18, 20, 22, 24, 26, 28 in lanes 1-7, respectively

Bibliography

[1]Somerville, R. L and C Yanofsky. 1965. *Studies on the Regulation of Tryptophan Biosynthesis in Escherichia Coli.* Journal of Molecular Biology. Vol. 11, p. 747 – 749.

[2]Yanofsky, C, I.P. Crawford, B.P. Nichols, G.E. Christie, H. Horowitz, M. VanCleemput and A.M. Wu. 1981. *The Complete Nucleotide Sequence of the Tryptophan Operon of Escherichia Coli.* Nucleic Acids Research. Vol. 9, p. 6647-6668.

[3]Wierenga, R. K. 2001. *The TIM-barrel Fold: a Versatile Framework for E¤cient Enzymes.* FEBS Journal. Vol. 492, p. 193-198.

[4]Wilmanns, Matthias, C. Hyde, D. Ravies, K. Kirschner and J. Jansonius. 1990. *Structural Conservation in Parallel Plar-Barrel Enzymes That Catalyze Three.* Biochemistry. Vol. 30, p. 9161-9169.

[5]Forsyth, William and C. Matthews. 2002. *Folding Mechanism of Indole-3-glycerol Phosphate Synthase from Sulfolobus Solfataricus: A Test of the Conservation of Folding Mechanisms Hypothesis in (β/a)8.* Journal of Molecular Biology. Vol. 320, p. 1119-1133.

[6]Shen, Hongbo, F. Wang, Y. Zhang, Q. Huang, S. Xu, H. Hu and J. Yue. 2009. *A Novel Inhibitor of Indole-3-glycerol Phosphate Synthase with Activity against Multidrug-resistant Mycobacterium Tuberculosis.* FEBS Journal. Vol. 276, p. 144-154.

[7]Darimont, Beatrice, C. Stehlin, H. Szadkowski and K. Kirschner. 1998. *Mutational Analysis of the Active Site of Indole Glycerol Phosphate Synthase from Escerichia Coli.* Protein Science. Vol. 7, p. 1221-1232.

[8]Sterner, R and B. Hocker. 2005. "*Catalytic Versatility, Stability, and Evolution of the (β /a)8-barrel Fold.*" Chemistry. Vol. 105, p. 4038-4055.

[9]Parry, R. J. 1972. *Biosynthesis of Compounds Containing an Indole Nucleus, in Chemistry of Heterocyclic Compounds: Indoles, Part Two.* Vol. 25, John Wiley & Sons, Inc., Hoboken, NJ, USA.

[10]Hennig, M, B. Darimont, J. Jasonius and K. Kirschner. 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. Vol. 319, p. 757-766.

[11] Hennig, M, B. Darimont, R. Sterner, J. Jasonius and K. Kirschner. 1995. "*2.0 å Structure of Indole-3-glycerol Phosphate Synthase from the Hyperthermophile Sulfolobus Solfataricus: Possible Determinants of Protein Stability.*" Structure. Vol. 3, p. 1295-1306.

[12]Carravetta, Marina, M. Eden, O Johnannessen and H. Luthman. 2001. *Estimation of Carbon−Carbon Bond Lengths and Medium-Range Internuclear Distances by Solid-State Nuclear Magnetic Resonance.* Journal of the American Chemical Society. Vol. 123, p. 10628-10638.

[13]Wong, J Tze-Fei. 2003. *Coevolution of Genetic Code and Amino Acid Biosynthesis.* Trends in Biochemical Sciences. Vol. 6, p.33-36.

[14]Williams, S. G and S. Lovell. 2009. "*The Effect of Sequence Evolution on Protein Structural Divergence.*" Molecular Biology and Evolution. Vol. 26, p.1055-1065.

[15]Little, D. Y and L. Chen. 2009. "*Identification of Coevolving Residues and Coevolution Potentials Emphasizing Structure, Bond Formation and Catalytic Coordination in Protein Evolution.*" PLoS One. Vol. 4.

[16]Estabrook, R. A, J. Luo, M. Purdy, V. Sharma, P. Weitzliem, T. Bruice and N. Reich. 2004. *Statistical Coevolution Analysis and Molecular Dynamics: Identification of Amino Acid Pairs Essential for Catalysis.* National Academy of Sciences. Vol. 102, p. 994-999.

[17]Shen, Hongbo, F. Wang, Q. Huang, F. Xu and H. Hu. 2008. *Coevolving Residues of (b/a)8-barrel Proteins Play Roles in Stabalizing Active Site Architecture and Coordinating Protein Dynamics.* Journal of Structural Biology. Vol. 164, p.281-292.

[18]Czekster, CM, B. Neto, A. Lapis, J. Dupont and D. Santos. 2009. *Steady-state Kinetics of Indole-3-glycerol Phosphate Synthase from Mycobacterium Tuberculosis*. Archives of Biochemistry and Biophysics. Vol. 486, p. 19-26.

[19]"EC 4.1.1.48 Indole-3-Glycerol Phosphate Synthase." 2010. BRENDA. <http://www.brenda-enzymes.info/php/result_flat.php4?ecno=4.1.1.48&Suchword=&organism%5B%5D=Escherichia+coli&show_tm=0>

[20]Mazumder-Shivakumar, Devleena and T. Bruice. 2004. "*Molecular Dynamics Studies of Ground State and Intermediate of the Hyperthermophilic Indole-3-glycerol Phosphate Synthase.*" Proceedings of the National Academy of Sciences. Vol. 101, p.14379-14384.

[21]Merz, Astrid, M. Yee and H. Szadkowski. 1999. *Improving the Catalytic Activity of a Thermophilic Enzyme at Low Temperatures.* Biochemistry. Vol. 39, p. 880-889.

Academic Vita

EDUCATION

Columbia University (Class of 2015)
Doctor of Medicine (M.D.)

The Pennsylvania State University - Schreyer Honors College (August 2007 – May 2011)
College of Engineering – Chemical Engineering with honors in Chemistry

EXPERIENCE

Merck and Co., Inc. – Engineering and Technology Fellowship (May – August 2010)

- Selected as most outstanding Penn State undergraduate engineering student
- Implemented vaccine vial glass-testing protocol, revised cleaning SOP

Air Products and Chemicals, Inc. – Chemical Process Internship (May – August 2009)

- Worked in downstream purification of hydrogen plant (HYCO)
- Synchronized logic of HYCO pressure-swing adsorption unit
- Wrote SOP for R&D experimental burner lab using P&ID

RESEARCH

Department of Chemistry (August 2010 – May 2011)

- Advisor: Dr. David Boehr – Pennsylvania State University
- Conducted site-directed mutagenesis on indole glycerol phosphate synthase
- Performed anion-exchange chromatography, PCR, gel electrophoresis, and dialysis

Department of Biobehavioral Health (January 2009 – May 2009)

- Advisor: Dr. David Vandenbergh – Pennsylvania State University
- Amplified bisulfite-treated Agouti mice DNA to analyze polymorphisms
- Performed PCR, gel electrophoresis and micropipetting

Department of Biology (December 2009 – January 2010)

- Designed and executed research projects; presented findings
- Worked in rain forest, cloud forest, beach and mountain climates across Costa Rica

AWARDS

- Penn State Intramural Flag Football Campus Champions (2007-2008)
- Academic Excellence Award (August 2007 – May 2011)
- Dean's List – 7/7 semesters (December 2007 – December 2010)
- President's Freshmen Award (February 2008)
- Merck Engineering and Technology Fellowship (November 2009)