12-2015

Toward the Probing of DHQS Activity by Protein Engineering through the Introduction of Unnatural Amino Acids and the Selection of tRNA/tRNA Synthetase Pairs

Shaina E. Ives
University of Nebraska-Lincoln, seives@huskers.unl.edu

TOWARD THE PROBING OF DHQS ACTIVITY BY PROTEIN ENGINEERING THROUGH THE INTRODUCTION OF UNNATURAL AMINO ACIDS AND THE SELECTION OF tRNA/ tRNA SYNTHETASE PAIRS

By

Shaina E. Ives

A Thesis

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Chemistry

Under the Supervision of Professor Jiantao Guo

University of Nebraska

December 2015

TOWARD THE PROBING OF DHQS ACTIVITY BY PROTEIN ENGINEERING THROUGH THE INTRODUCTION OF UNNATURAL AMINO ACIDS AND THE SELECTION OF tRNA/ tRNA SYNTHETASE PAIRS

Shaina E. Ives, M.S.

University of Nebraska, 2015

Advisor: Jiantao Guo

Protein engineering is a valuable tool that allows scientist to explore how an enzyme works by mutation of key residues. This method has been used to improve function or stability of enzymes, thus allowing their use in both the lab and in industry to be expanded. Genetic incorporation of unnatural amino acids (unAA) can be used with protein engineering to exceed the current limitations, due to the limited number of functional groups of the 20 common amino acids.

The majority of this thesis will discuss the progress on incorporating the various unAA into the active site of the enzyme, Dehydroquinate Synthase (DHQS). Previously this enzyme has been understood through the use of inhibitors as well as crystallization of the enzyme with its substrate and cofactors. This work begins to explore the chemical reactions of this enzyme catalysis, as well as enhancing the known tools of genetic incorporation of unAA. This masters thesis focuses on my work the in: (1) the synthesis of the unAAs including hydroxyquinolin-alanine and 2-(5-carboxythienyl)alanine; (2) the identification of aminoacyl tRNA synthetase recognizing 2-(5-bromothienyl)alanine, 2-mercapto-L-histine, diiodo-histine, and 2-

(5-carboxythienyl)alanine; and (3) the incorporation of hydroxyquinolin-alanine and 2-(5-bromothienyl)alanine into DHQS and the evaluation of the resulting mutants.

ACKNOWLEDGEMENTS

Firstly I would like to thank Dr. Guo for giving me this opportunity to work as a graduate student in his lab. You have taught me so much about chemistry and were so helpful in encouraging me to grow as a scientist. I am leaving with so much more knowledge, and I am so grateful that you facilitated an environment where I could learn so much. You are a fantastic supervisor and scientist – so thank you. I would also like to thank Dr. Wei Niu; Wei you are a great person to work with and talk to. You are always willing to give suggestions and help to me if I needed it. You are such an interesting and inspiring individual; I have learned so much from you. I know you will have a great career ahead of you as you start your new lab. I would like to thank my lab mates: Kun, Tong, Xi, Xin, Nanxi, and Erome. You have been so supportive and helpful. I appreciate you guys were always ready to help as well always ready to have fun. I wish you the best in everything you do.

To my dear grad school buddies: thank you for listening, and helping, being supportive, and offering times to just relax (with coffee!!!!). I would have been miserable without you all. Veronika, Katie, Riehannah, and Ben – I will miss you all and I won't forget the time I had with you nor the things I learned from you. You all are great scientists and friends; I know you will do amazing things.

To my family thank you for supporting me and pushing me to explore everything that I am interested in. You have helped to become the person I am today and I think you really developed my interest and love of science. To my fiancée soon to be husband, Lucas McDowell: I am so happy that you were with me through this

journey through graduate school. You have always been so supportive and loving, and you knew how to help when I was very stressed. You have been a constant support for the last five years – and especially the last two. You are my best friend, and I will always love you for your friendship, support, and love you give me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

NOMENCLATURE AND ABBREVIATIONS

aaRS – aminoacyl-tRNA synthetase

Amp – ampicillin

BrTha - 2-(5-bromothienyl)alanine

Cm - chloramphenicol

DAHP - deoxy-*D* –arabino-heptulosonate 7-phostphate

DHQ - 3-dehyroquinate

DHQS - dehyroquinate synthase

GMML – glycerol minimal media with leucine

His tag – polyhistidine –tag

HQAla – hydroxyquinolin-alanine

IPTG – β-D-1-thiogalactopyranoside

Kan – kanamycin

LB – Lysogeny broth

Mb - *Methanosarcina barkeri*

Mj - *Methanococcus jannashii*

NAD^+ - Nicotinamide adenine dinucleotide

OD – optical denisty

RFU – relative fluorescent units

SDM – site directed mutagenesis

sfGFP – superfolded green fluorescent protein

SLIC – sequence and ligation independent cloning

unAA – unnatural amino acid

wt –wild type

CHAPTER 1

BACKGROUND

1.1 Introduction

In nature there are 20 common amino acids that are ordered and then folded in endless combinations to create enzymes that catalyze almost all chemical reactions in biological system. These enzymes sometimes are able to catalyze the synthesis of molecules through amazing and difficult chemistry. As enzymes are found in nature and often can catalyze thermodynamically or entropically unfavorable reactions at room temperature, exploration and manipulation of enzymes has great potential to efficiently produce chemicals needed for our economy today with a greener footprint.

Protein engineering has been used as a tool to generate enzyme with improved properties, such as increased catalytic activity or improved thermo stability. Traditionally protein engineering had only the twenty common amino acids to use with their limited number of functionality of their side chains. An intuitive question is if proteins functions can be improved by introducing novel amino acids with unique physical, chemical, or biological properties. To answer this question, genetic incorporate unnatural amino acids (unAAs) in response to amber nonsense codon has been developed by modifying a cell's translational machinery. With a combination of tradition protein engineering and genetic incorporation of unAAs the method to study a protein's structure and function could be enhanced.

The potential to change an enzyme by the cell's own machinery was first hinted at in 1962 by two papers demonstrating that it is the tRNA not the amino acid itself that dictates the protein code.[1,2] These papers showed that alanine could be incorporated into a peptide chain by the translational machinery after chemically converting a Cys-tRNA^{Cys} into Ala-tRNA^{Cys} that the alanine could still be incorporated into a peptide chain by the translational macheniery.[1,2] This *in vitro* mis-acylation was expanded by showing multiple types of amino acids could be synthetically charged, [3] then incorporated. The Johnson group showed incorporation of an unAA in 1994, showing the incorporation of a fluorescently labeled lysine at a lysine codon.[4]

In 1989 the Schultz group developed a general approach for the site-specific incorporation of unnatural amino acids directly into proteins in response to amber nonsense codon *in vivo*.[5] This method prevented unspecific incorporation of canonical amino acids, as the method involved first changing a specific codon in a target gene into an amber TAG codon through site directed mutagenesis. The following mRNA would be translated entirely by an amber suppressing tRNA that is chemically aminoacylated with an unAA of interest (Figure 1-1); thus allowing for a full-length protein with a site-specific incorporation of the unnatural amino acid.[5] Though many types of unAA could be incorporated there were limitation due to low yields of the acylated tRNA. As such this incorporation method is only particularly useful in situations that microinjections can be used, as shown with *Xenopus* oocytes in the 1995 paper by the Dougherty and Schultz group. [6]

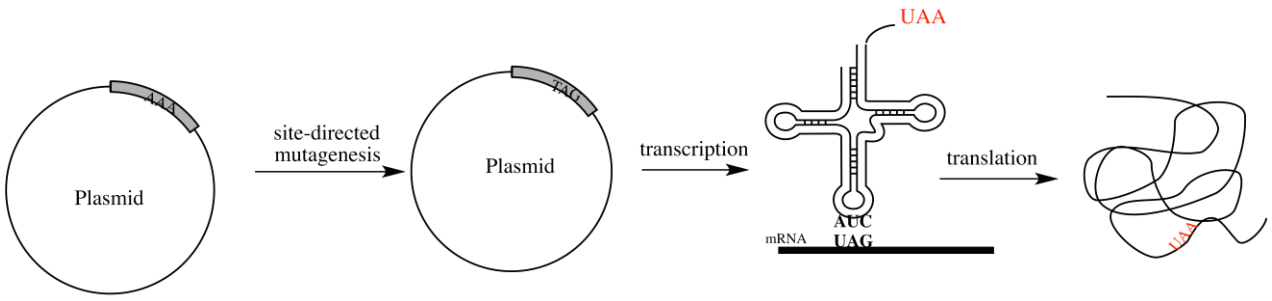


Figure 1-1: Generalized method of site-specific incorporation of unAAs.

As *in vitro* design of incorporating unAA has limitations in how much tRNA can be charged the Schultz group proposed, and demonstrated in 2001, incorporation of a free unAA using a cell's machinery.[7] Incorporation of an *O*-methyl-L-tyrosine was successfully done by use of the *Methanococcus jannashii* (*Mj)* tyrosyltRNA^{Tyr}/TyrRS pair.[7] This archaebacterium tRNA^{Tyr} is different from the *E. coli* tRNA^{Tyr} so after further modifying of *Mj*tRNA and transfer it into an *E. coli* host there was little recognition by the host's endogenous tRNA synthetases.[7] Furthermore the *Mj*TyrRS was mutated for increased specificity of the unAA over naturally occurring tyrosine.[7]

Since this method was first published it has been expanded and improved many times.[8] With over 150 unnatural amino acids (unAAs) reported to be successfully incorporated into proteins, the potential for adding new chemistry to enzymes and creating artificial enzymes has becomes a process of engineering and rationally design.[8,9,10] The ability to incorporate multiple and various unAAs has also expanded since 1989, as now beyond the amber nonsense codon, the low-abundant AGG codon,[11,12] the tRNA^{Pyl} recognizing CUA codon,[13,14] and the quadruplet

codon,[15] all have been shown to successfully incorporate unAAs site specifically into proteins.

1.2 Dehydroquinate synthase

The goal of my research is to apply unAA mutagenesis approach to the engineering of dehydroquinate synthase (DHQS.)

The 7-phospho-3-deoxy-D-*arabino*-heptulosonate phosphate lyase, or dehydroquinate synthase (DHQS) is a sugar phosphate cyclase with a very interesting mechanism. It is found in nature as one of seven enzymes that make up the shikimate pathway, transforming D-erythrose 4-phosphate and phospho*enol*pyruvate to chorismate, the pathway is shown in Figure 1-2.[16] Byproducts and chorismate are used as metabolites in a verity of pathways to form phenylalanine, tyrosine, and tryptophan; essential amino acids to higher order organisms; as well as other aromatic compounds like vitamin E and K, folic acid, ubiquinone, and naturally found coumarins.[16]

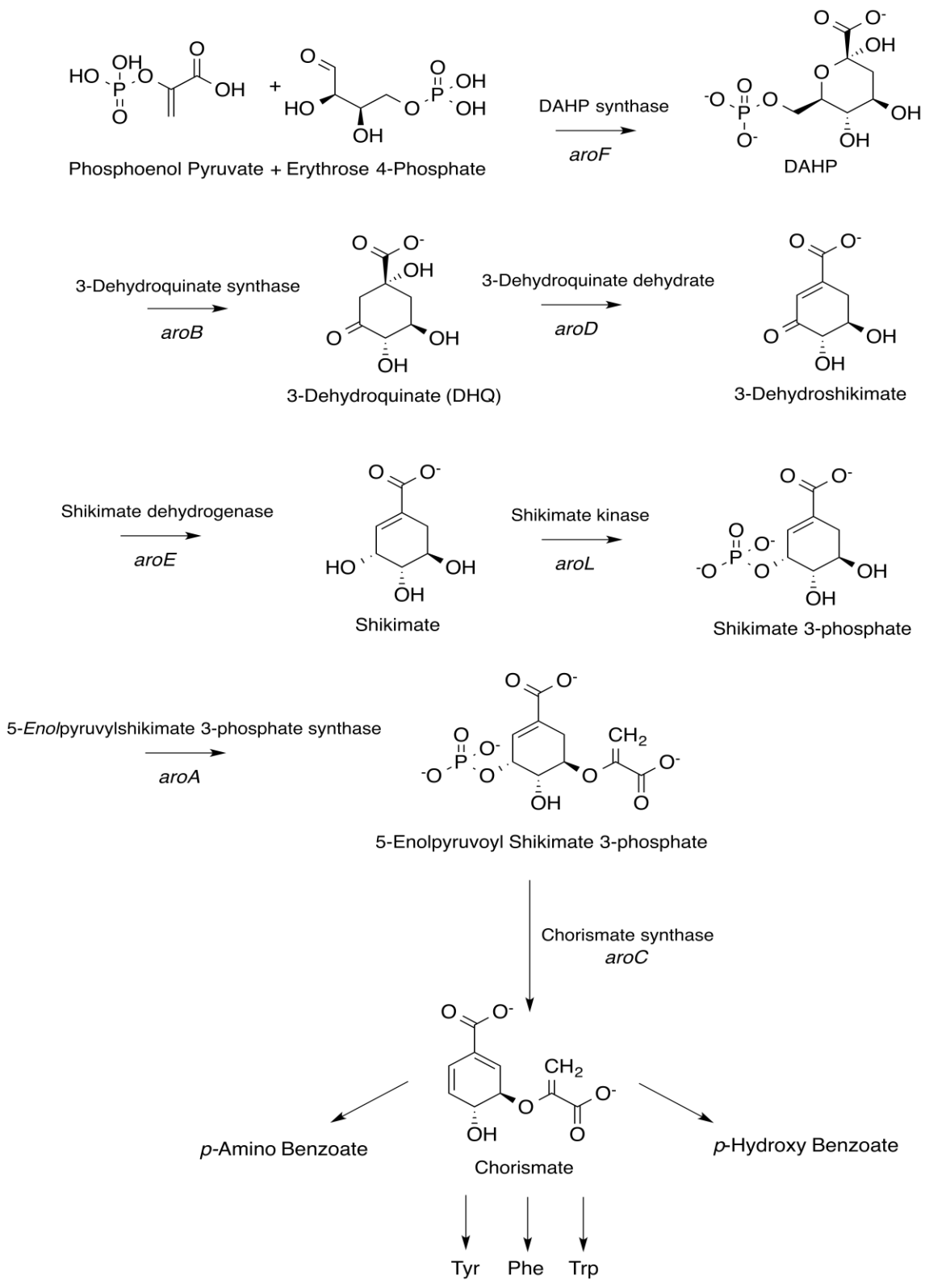


Figure 1-2: Shikimate pathway adapted from Reviews of Herrman[17] and Haslam. [16]

The enzyme DHQS is particularly interesting and the focus of this work as the enzyme transforms deoxy-*D*–arabino-heptulosonate 7-phostphate (DAHP) into 3-dehyroquinate (DHQ) through a series of five reactions. The enzyme catalyzes an alcohol oxidation, phosphate β-elimination, carbonyl reduction, a ring opening, and finally an intramolecular aldol condensation.[17,18] As there are multiple steps catalyzed by DHQS there was a proposed mechanism for the transformation of DAHP to DHQS with a verity of proposals to the steps the substrate went through. The Brown group proposed mechanism of DHQS shown in Figure 1-3 with convincing evidence of the crystallization of DHQS with the cofactors and substrates in 1998, demonstrating how the residues interacted with the substrate in the transformation of DAHP into DHQ.[18]

The DHQS, encoded by the *aroB* gene, is assisted with its multiple reactions with help of the two cofactors, NAD and a divalent metal (either Co^{2+} or Zn^{2+}).[18] The metal ion is dependent of the origin of the DHQS, as from overexpressed *E. coli* DHQS was found to have tightly bound Co^{2+} to have highest activity, while addition of Zn^{2+} has been shown to restore most activity.[19] While the DHQS from the AROM protein of *Aspergillus nidulans*, which was used by the Brown group for crystallization bound Zn^{2+} preferentially.[18] The NAD^+ and the divalent metal act in the first enzymatic step of DHQS oxidizing the C5 alcohol as shown in Figure 1-3.[17,18] The divalent metal ion likely facilitates a hydride transfer from C5 to NAD^+ and a proton loss of the C5 hydroxyl.[18] Though the crystallization and most recent mechanism proposal is from the DHQS from *Aspergillus nidulans*, the *E. coli* DHQS is

comparable in structure and activity and as such all mutants in this work are mutated based on the crystal structure of *Aspergillus nidulans* DHQS (Figure 1-4). [18]

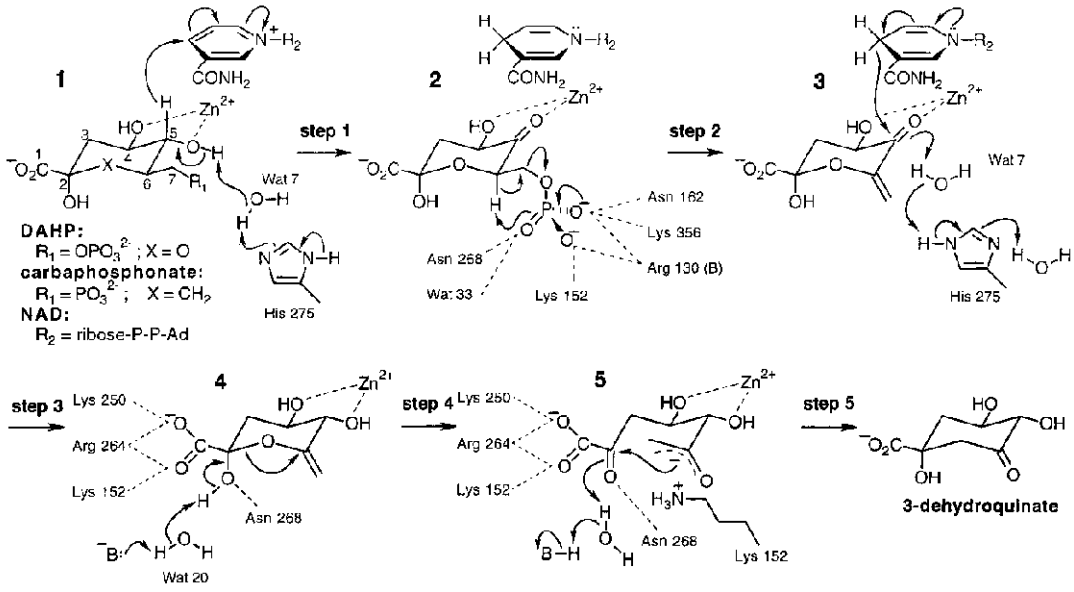


Figure 1-3: The mechanism for DHQS of *Aspergillus nidulans.* [18]

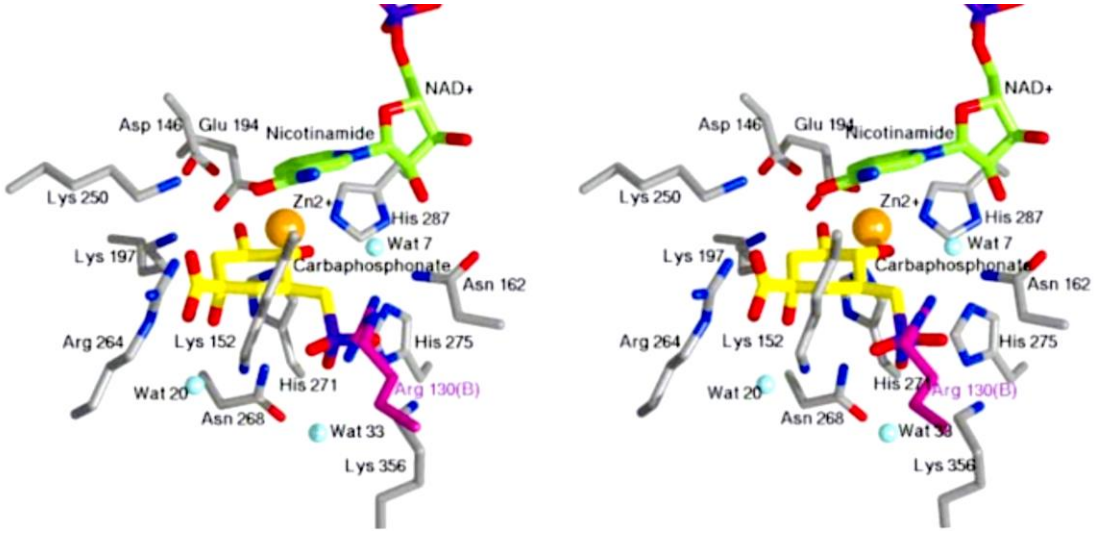


Figure 1-4: Crystal structure of *Aspergillus nidulans* DHQS active site. [18]

The complex and multistep enzyme mechanism of the DHQS makes this enzyme an interesting candidate for investigation, manipulation and introducing new chemical potential by incorporation of unAAs at key active site residues. As there are so many different reactions catalyzed by this enzyme there are a lot of potential to influence the chemistry of the enzyme by what amino acid is incorporated at key residues. Furthermore as this enzyme is a metalloenzyme, so the potential to influence the chemistry through the chelating residues is also an interesting area to explore. This enzyme mechanism was described by Bartlett in a 1988 paper, "The mechanistic details of the transformation reflect both clever functional group manipulation and stereochemical dexterity on the part of the enzyme."[20] The general mechanism is solved now the interesting part is exploring the chemistry behind it.

1.3 **Experimental Design**

The way this work explores the chemistry of the DHQS mechanism is through incorporation of unAAs site specifically in place of key active-site residues. The selection for aminoacyl-tRNA synthetase (aaRS) to charge a tRNA for incorporation of a unAA was done using a similar method to the unAA selection method used by the Schultz group at The Scripps Research Institute, and the experimental design is described below (aminoacyl-tRNA synthetase^{UAA} Selection) and is used in the selection process of this work. To check the efficiency and relative activity of the DHQS mutants, the DHQS gene was inserted downstream of an inducible promoter on a plasmid. These plasmids were then transformed into DHQS knockout hosts to

monitor functionality of mutants (DHQS Mutant Activity Screening). Both methods are described in this section.

Aminoacyl-tRNA Synthetase (aaRS) Selection

The unAA for our interest to incorporate are histidine derivatives, such as 2-(5-bromothienyl)alanine, 2-mercapto-L-histine, diiodo-histine, and 2-(5-carboxythienyl)alanine (Figure 1-5). In order to genetically incorporate these histidine analogs, we chose to evolve a $tRNA^{Pyl}$-pyrrolysyl-tRNA synthetase (PylRS) pair from a *Methanosarcina barkeri* PylRS library. Previous work has shown successful incorporation of a few unAA with structures similar to histidine using a $tRNA^{Pyl}$-PylRS variant derived from *M. barkeri*.[21,22] The $tRNA^{Pyl}$-PylRS pair is responsible for the genetic incorporation of pyrrolysine (the 22^{nd} amino acid) into protein, recognizing the Amber nonsense codon.

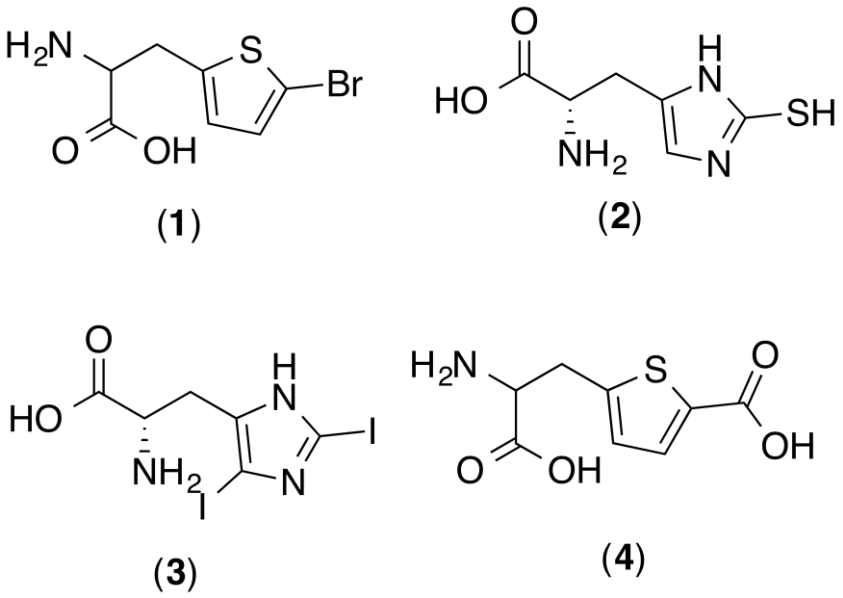


Figure 1-5: 2-(5-bromothienyl)alanine (**1**), 2-mercapto-L-histine (**2**), diiodo-histine (**3**), and 2-(5-carboxythienyl)alanine (**4**)

The basic outline of selection of PylRS/tRNA^{Pyl} pairs is adopted from an established method by the Schultz lab for the *Methanosarcina barkeri* (*Mb*) PylRS 3N library (Figure 1-6A and B). The randomized sites in the 3N *Mb*PylRS library are L270, Y271, L274, C313, and V401. Each amino acid is randomized by the expressed sequence at the codon position, the codons are randomized by NNK degenerative oligonucleotides (N=A,T,G,C and K=G,T). The randomized sequences are then cloned into the pBK plasmid under control of a constitutive *glnS* promoter with a kanamycin resistance marker (Kan^{r}) for plasmid maintenance.

To remove most synthetase mutants that would charge canonical amino acids this plasmid was co-transformed with a negative selection plasmid, pNEG-PylT (Figure 1-6A). This plasmid harbors a ribonuclease barnase gene with two amber nonsense codons at permissive sites. This plasmid also has a copy of tRNA^{Pyl} under

a *Ipp* promoter. When grown without the target unAA any synthetase variants that could charge canonical amino acids would produce full length barnase. As barnase is cytoxic, cells containing these synthetase variants would die, and thus enrich the library with synthetase variants that would be selective against canonical amino acids. These plasmids were harvested and co-transformed with the positive selection plasmid.

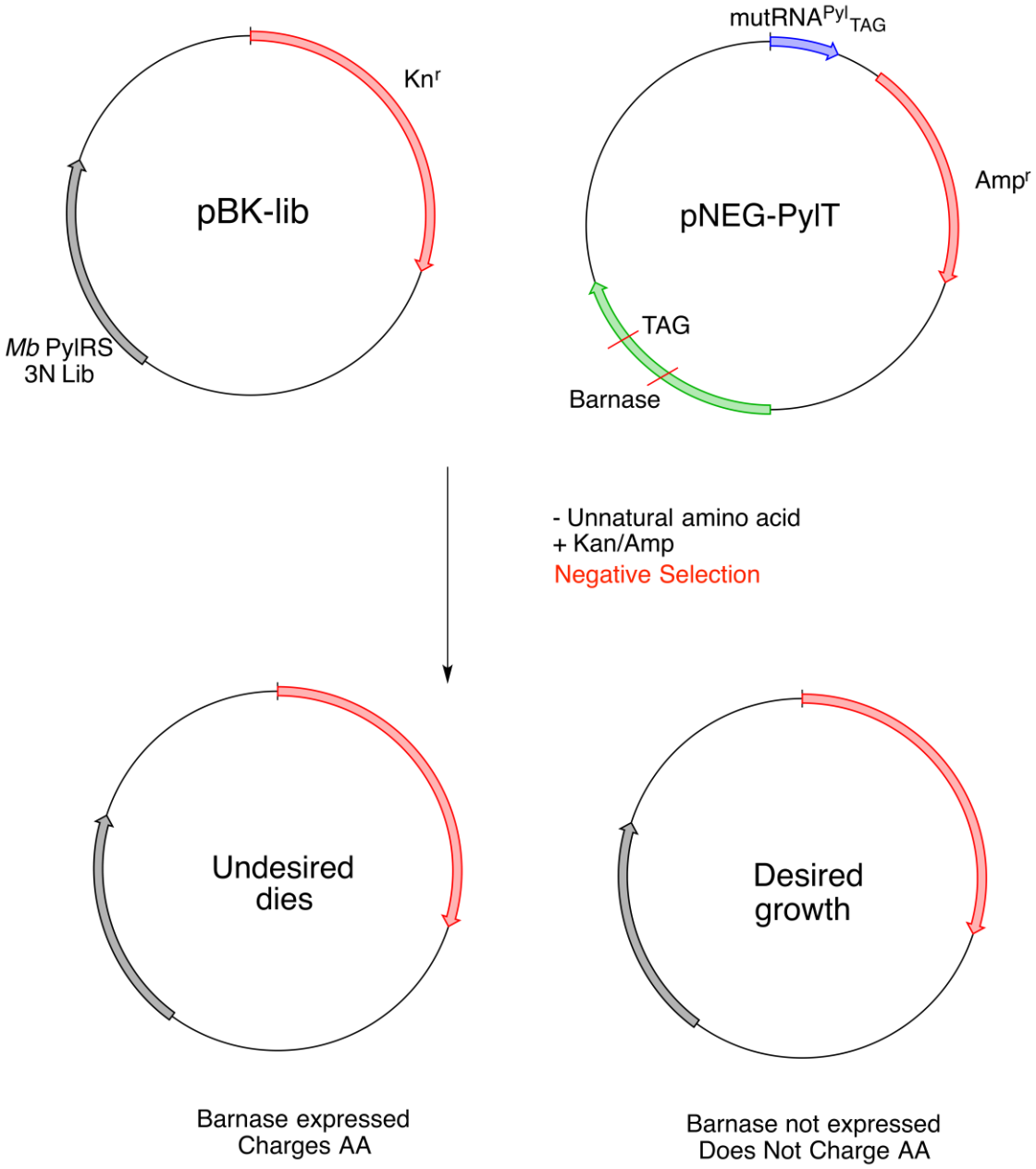


Figure 1-6A: Negative selection outline for unAA against PylRS library

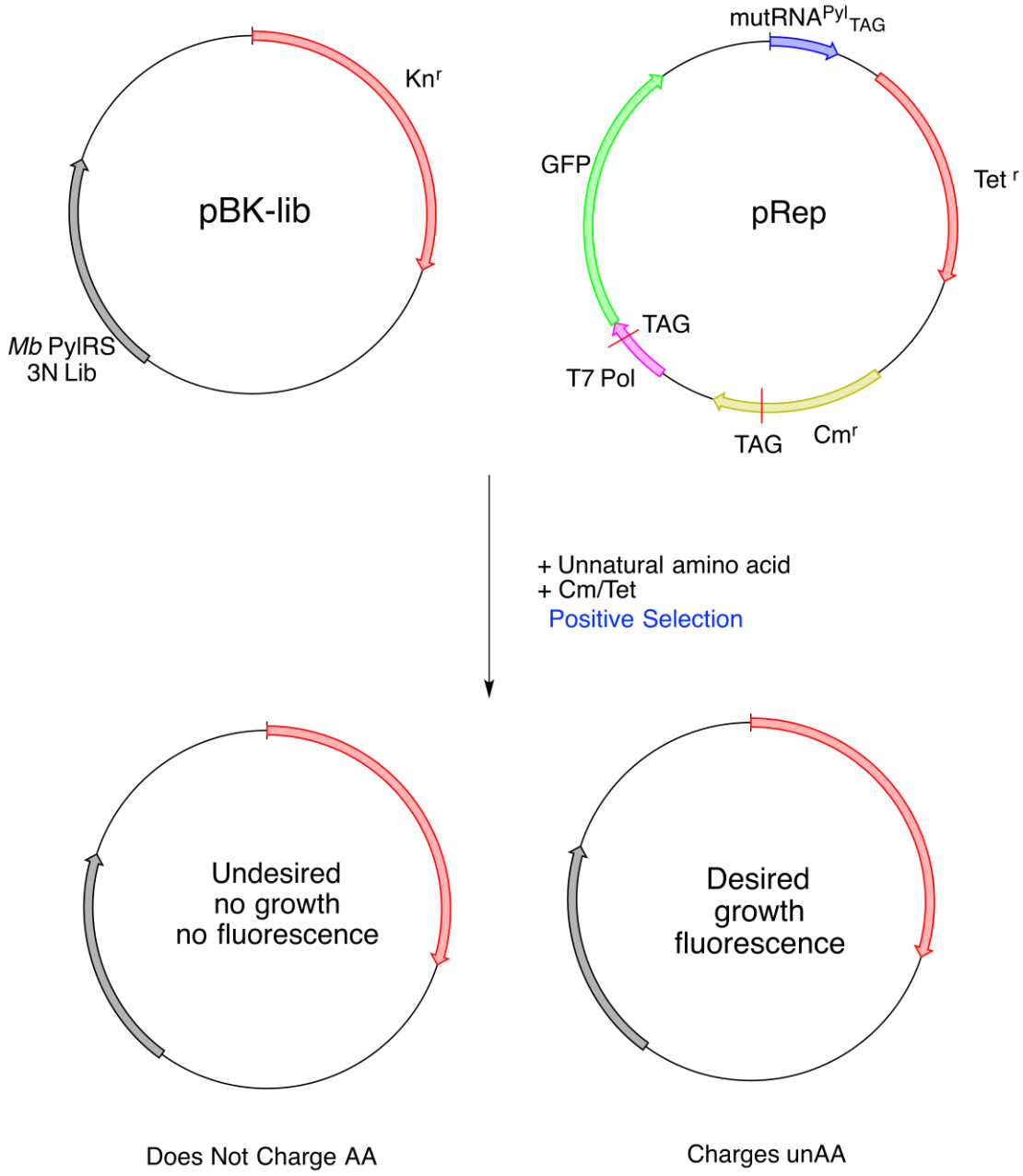


Figure 1-6B: Positive selection outline for unAA against PylRS library

The positive selection plasmid (pRep) has one copy of the *Mb*tRNA^{Pyl} and a chloramphenicol resistance marker (Cm^r) with a TAG codon at a permissive site. This plasmid also contained a T7 RNA polymerase gene with a TAG codon at a

permissive site to act as a second reporter (Figure 1-6B). The cells are grown on various low levels of chloramphenicol concentration with the unAA at 1mM-2mM concentrations. Efficient charging of the amber suppressor $tRNA^{Pyl}$ by PylRS variants with the unAA results in cells that are chloramphenicol resistant and fluorescent. Both growth and fluorescence should indicate a synthetase with high activity to charge an amino acid to the tRNA. Cells harboring PylRS variants capable of utilizing the desired unAA but not endogenous amino acids will survive both negative and positive selections.

Colonies with high relative fluorescence were selected for replication plates with and without the unAA of interest. The colonies that were able to grow and exhibit fluorescence only on the plates with unAA were selected for further characterization. The logic behind this being that only the *Mb*-PylRS that are able to charge the unAA at high efficiency would be able to show strong fluorescence and chloramphenicol resistance. To ensure the hit was not a false positive the *Mb*-$PylRS^{UAA}$ was transformed into a strain with pLei-GFP_{UV}, this GFP gene has an amber codon at a permissive site (Asn149). When induced with the unAA and induced without the UAA, if the *Mb*-$PylRS^{UAA}$ is actually selective then the relative fluorescent signal of the induced culture and unAA will be folds higher. This general method was used consistently in the work following.

DHQS Mutant Activity Screening

To efficiently check the relative activity of the mutant DHQS enzymes, plasmids containing the *aroB* gene with amber mutations of the interest were transformed

into strains with decreased activity of *aroB* or strains with *aroB* being fully knocked out. By growing these strains with the mutant DHQS in cultures without aromatic amino acids and vitamins the cells can only grow if the DHQS is functional. This method is useful as it quickly gives a relation to the activity of the mutant enzyme in comparison to the wild type enzyme.

The strains used in this work were the JWΔaroB and the AB2847 strain. Both were ordered and received from the *E. coli* Genetic Stock Center at Yale University (Section 4.7). The JW strain, originally JW3352-1, is a knockout with the *aroB* gene removed and a kanamycin marker as a placeholder. The marker was removed (Section 4.7 (3)) and was used as a host for the DHQS mutants. The other strain, AB2847, is an older strain that was first reported in 1966 with a full *aroB* gene. There are two mutations compared to the wild-type *aroB* sequence at position 93, His^{CAT} mutated to Gln^{CAG}, and position 254, Glu^{GAA} mutated to Lys^{AAA}.[23] This strain was reported to produce the DHQS with a reduced specific activity of 0.025 compared to the $DHQS_{wt}$ (specific activity, 1.16).[23] Either of these strains were important to in this work not only because it provided an easy method to compare efficiency of the DHQS mutants but also it reduced the background of wild type DHQS protein found in the cell.

1.4 **Conclusion**

As the DHQS is both a metalloenzyme and can catalyze a variety of reactions as described in section 1.2 it creates a very interesting model for unnatural amino acid incorporation. The diversity of reactions creates a platform to test how various

unAA could affect the reactivity of an enzyme. Optimization of DHQS is unlikely to result in any major impact to how aromatic amino acids are synthesized and isolated for commercial use. However the potential to explore what factor effects the activity of the enzyme by use of amino acid analogues would not only be interesting in the interest of science, but also in terms of a bigger picture a way to rationally design an inhibitor of DHQS. Inhibitor of DHQS is important as this pathway is not found in higher order eukaryotes but is found in apicomplexan parasites such as malaria.[24] Information in how this enzyme works could lead to better development of anti-microbial and anti-parasitic compounds.

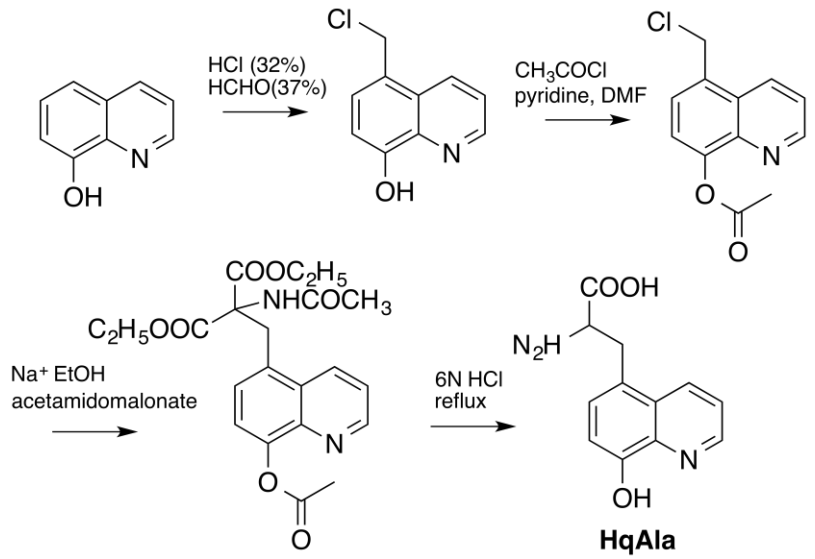
Chapter 2

Metal Chelating Sites in DHQS

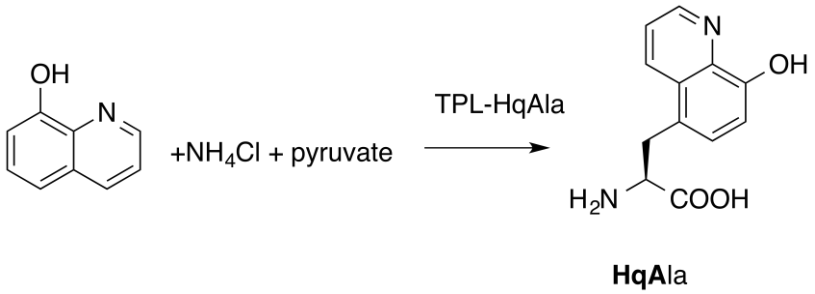
2.1 Introduction

The first sites targeted for the possible incorporation of unAA were the two metal chelating sites in the DHQS. The glutamic acid (E) 184 and histidine (H) 264 sites in *E. coli* DHQS acts to bind cobalt (II) (Co^{2+}) in the active site.[18] The carboxylic acid side chain of the glutamic acid and the imidazole ring of the histidine act as electron donors, binding the Co^{2+}, where it is used by the enzyme to oxidize the C5 alcohol in the first step of the mechanism.[19, 19] As the *Aspergillus nidulans* uses Zn^{2+} as the metal to oxidize the alcohol, there is potential to change the activity of the *E.coli* DHQS by changing what is metal ion is chelated in the active site and how tightly the metal is bound. The unAA used in this project was hydroxyquinolin-

alanine (HQAla) as shown in Figure 2-1. This unAA has been shown to have strong binding to divalent metals, particularly Cu^{2+}.[25] HQAla as a strong metal chelator could have the potential to change the relative activity of the *E. coli* DHQS.



Scheme 2-1: Synthetic scheme of hydroxyquinolin-alanine.[26]



Scheme 2-2: Biosynthesis of hydroxyquinolin-alanine.[25]

The unAA, HQAla, was initially synthesized by the Fridkin group (Figure 2-1) at the Weizmann Institute of Science in hopes of treating Alzheimer Disorder; using its nature as a chelating agent, it assisted a radical scavenger to pass the blood brain barrier successfully.[26] Then a biosynthesis of this unAA was developed by the Wang

group in the Institute of Biophysics, Chinese Academy of Sciences. This biosynthesis started with the precursor of 8-hydroxyquinoline and via a double mutated tyrosine phenol lyase (TPL) transformed it into the unAA HQAla at an overall yield of 40% (Figure 2-2).[25] We followed the organic synthesis route to obtain HQAla at gram scale quantities. Though we referenced the Fridkin paper to synthesize HQAla, the synthesis was optimized as described in section 4.2 to get higher yields in intermediate steps.

2.2 Mutation of Metal Chelating Sites at E184 and H264

The sites targeted for incorporation in the DHQS of the unAA, HQAla, were the E184 and H264 sites. These sites were selected as the crystal structure of *A. nidulans* DHQS showed these residues chelate to the divalent metal ion.[18] The DHQS from *A. nidulans* is analogous to the *E. coli* DHQS, as such the sequence of both DHQS can be compared. The same residues can be targeted for the *E. coli* for mutation, though the residues come at different points in the sequence. To ensure that these two residues are responsible for the binding of the divalent metal, we conducted an alanine scan of the two residues.

The alanine scan was achieved by mutating each of the two residues to alanine through site directed mutagenesis (SDM). As a result, the proposed metal chelating side chains of E184 and H264 would be removed and replaced by a non-reactive methyl group. The relative activity of these alanine DHQS mutants could be monitored by the growth rate of these mutant sequences in the JWΔaroB strain in the presence of these mutants.

After amplifying out the *aroB* gene from Δ^3 *E.coli* plasmid pQZ1.022, the *aroB* gene was cloned into the pBK vector (pBK-aroB$_{wt}$). Three DHQS mutants were constructed including, the DHQS$_{E184A}$ (encoded in plasmid pBK-aroB$_{E184A}$), DHQS$_{H264A}$ (encoded in plasmid pBK-aroB$_{H264A}$), and a double mutant DHQS$_{E184A/H264A}$ (encoded in plasmid pBK-aroB$_{E184A/H264A}$) (Section 4.6). After transforming these mutants into the JWΔaroB strain, they were cultivated under two different conditions: (1) M9 minimal media *with* aromatic amino acids and vitamins (Figure 2-1), and (2) M9 minimal media *without* aromatic amino acids and vitamins (Figure 2-2). The OD of each culture was monitored over 48 hours. The original plasmid pQZ1.022 was used as a negative control as there is no promoter for the expression of the DHQS, so no activity was expected to assist in the growth of the cells. The plasmid, pBK-aroB$_{wt}$, was used as a positive control.

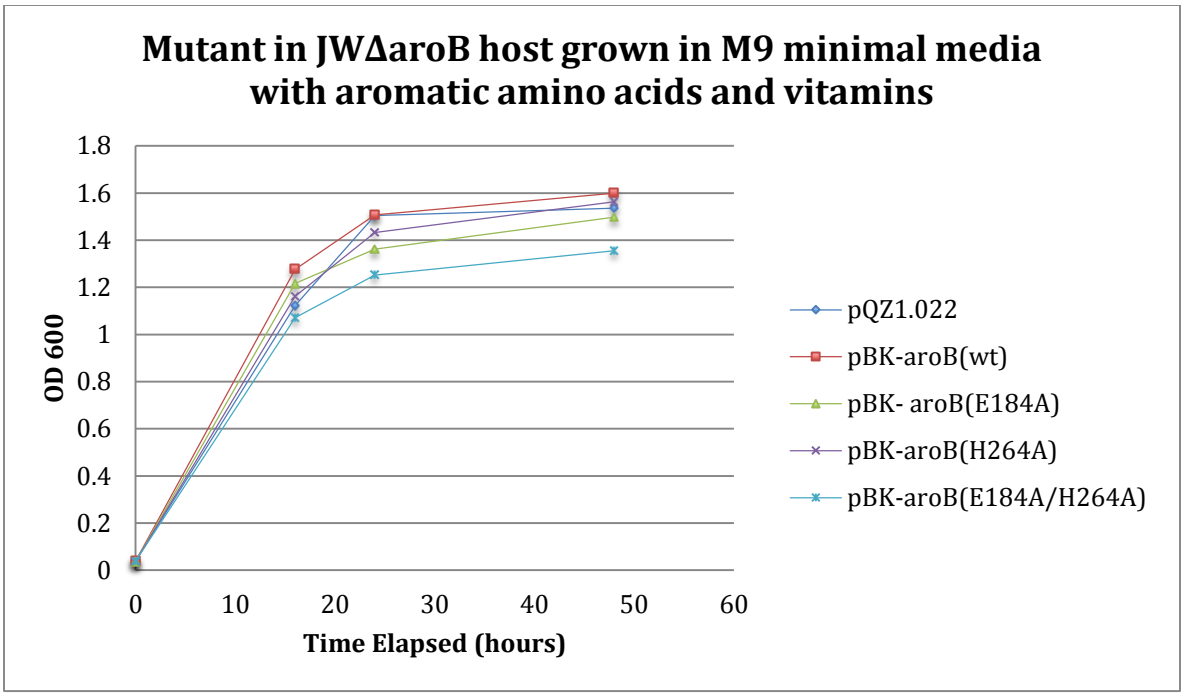


Figure 2-1: Growth JWΔaroB/pBK-aroB variants at 37°C in M9 Minimal Media with aromatic amino acids and vitamins.

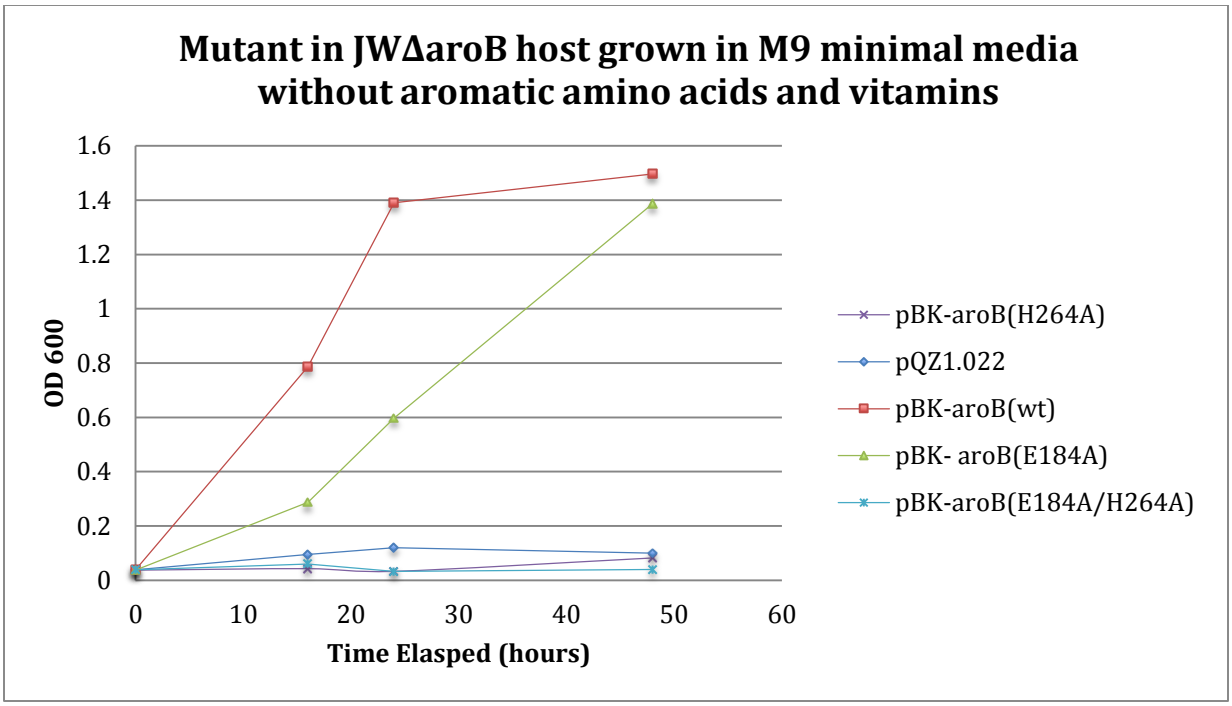


Figure 2-2: Growth of JWΔaroB/pBK-aroB varients at 37°C in M9 Minimal Media without aromatic amino acids and vitamins.

Over the incubation period cultures with aromatic amino acids and vitamins grew at a very similar rate to about the same relative OD, as the functionality of the DHQS mutants expressed does not matter for the growth of the strain. The data shown in Figure 2-1 shows the standard growth potential of the JWΔaroB of each culture. The cultures without the aromatic amino acids and vitamins added however rely on the activity of the DHQS mutants. In the case of the $DHQS_{wt}$ the enzyme was active and the cells were able to reach a similar OD in 24 hours as the cultures with aromatic amino acids and vitamins.

The DHQS variants with alanine mutations did not retain the same activity as the $DHQS_{wt}$. The single Glu184Ala mutant retained some activity as after 48 hours the cells reached a similar OD to the positive control culture. The single His264Ala

mutant did not retain activity. A similar result of non-activity was shown of the double mutant containing Glu184Ala and His264Ala. Comparing each of the mutants, DHQS$_{E184A}$, DHQS$_{H264A}$, and DHQS$_{G84A/H264A}$, shows that the His264 position probably has more effect in the chelation of the Co^{2+}. This is likely due to the ability of not only the nitrogen but also the pi electrons of the imidazole ring that assists in chelation of the metal. Overall the double mutation shows that the DHQS no longer retains any of its activity if these metal chelating sites are removed.

The alanine mutant, pBK-aroB$_{E184A/H264A}$, was then used as the template for another round of SDM to introduce amber nonsense codon. Again three mutants were constructed: DHQS$_{E184TAG/H264A}$, DHQS$_{E184A/H264TAG}$, and DHQS$_{E184TAG/H264TAG}$. These aroB genes were sub-cloned to a pJF vector, forming the plasmids pJF-aroB$_{E184TAG/H264A}$, pJF-aroB$_{E184A/H264TAG}$, and pJF-aroB$_{E184TAG/H264TAG}$ (Section4.6). This construct contains a regulative lac promoter, allowing us to turn on and off the expression of the gene with more control than with the pBK vector. These mutants were co-transformed with the synthetase *Mj*-HqAlaRS encoded on the pEvol plasmid as designed and selected by the Wang group.[25] This aaRS was selected from a *M. jannaschii* TyrRS library and had the following mutations from the wild-type *Mj*TyrRS: Tyr32His, Ile63Val, Leu65His, His70Gly, Phe108Arg, Tyr109Val, Asp158Asn, Leu162Asp, Cal164Gly.[25] This plasmid construct was received from the Wang lab and independently we verified the incorporation efficacy with a co-transformed sfGFP$_{TAG66}$ mutant (Fig 2-3). To ensure a high enough intracellular concentration of HQAla, we used a final concentration of 2mM of HQAla in culture media. Then to induce these cultures, we used 0.5mM of IPTG and 0.2% arabinose.

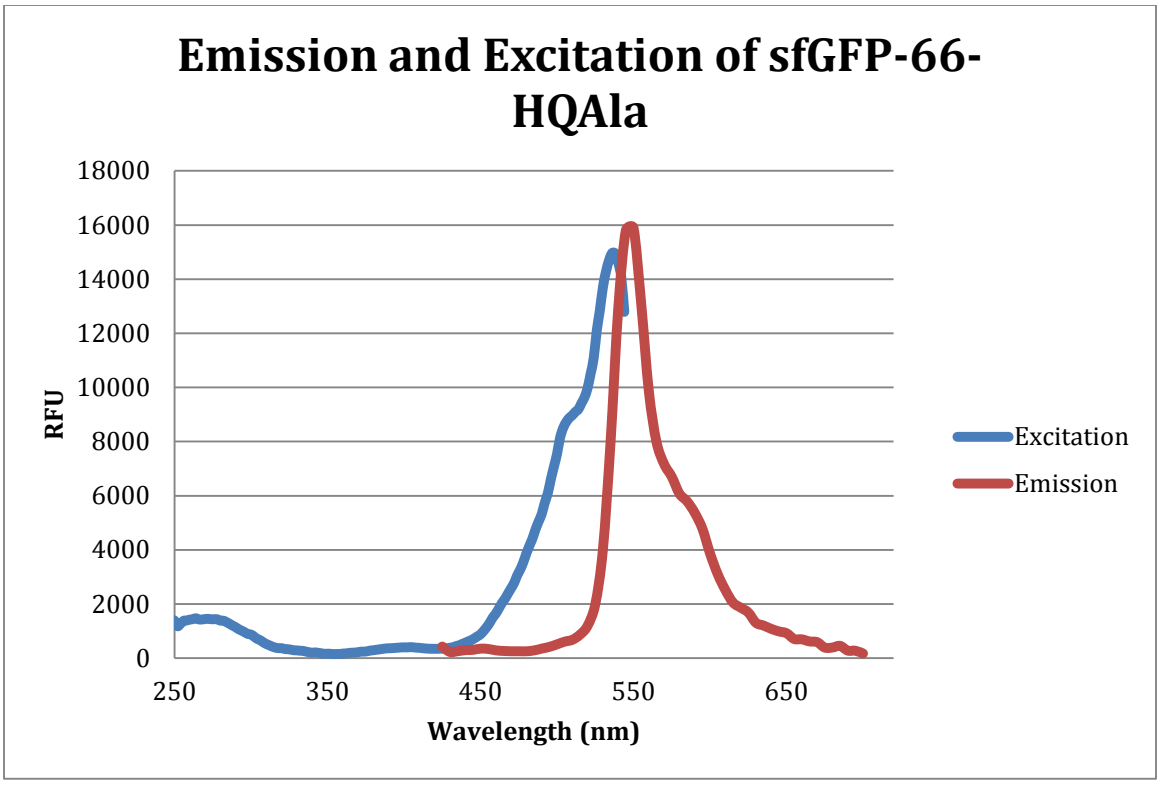


Figure 2-3: Excitation and emission spectra of sfGFP-HQAla66. The replacement of Tyr66 with HQAla66 results in a red-shifted emission maximum of sfGFP-HQAla66 (545 nm) from $sfGFP_{wt}$ (515 nm).[25]

2.3 Growth Test of Mutants $DHQS_{E184TAG/H264A}$ and $DHQS_{E184A/H264TAG}$

The experimental design, by use of the JWΔaroB as a host to express DHQS variants, allows us to effectively test the activity of the mutant enzymes without the need to isolate each enzyme. The mutants $DHQS_{A184TAG/H264A}$ and $DHQS_{E184A/A264TAG}$ transformed in the JWΔaroB strain were inoculated into five different cultures (M9 minimal media) and grown over 48 hours at 37°C. These five cultures included a positive control (an induced culture with aromatic amino acids and vitamins added) and two negative controls (one culture non-induced and the other culture induced, both without aromatic amino acids and vitamins). The two experimental cultures

that were induced with IPTG and 2mM of HQAla without aromatic amino acids and vitamins; though one culture trace elements were added to promote better cells growth. All cultures were inoculated with 100 µL of washed cells (Section 4.6 (3)) and OD was measured at inoculation, 16 hours, 24 hours, and 48 hours (Fig 2-4 A and B).

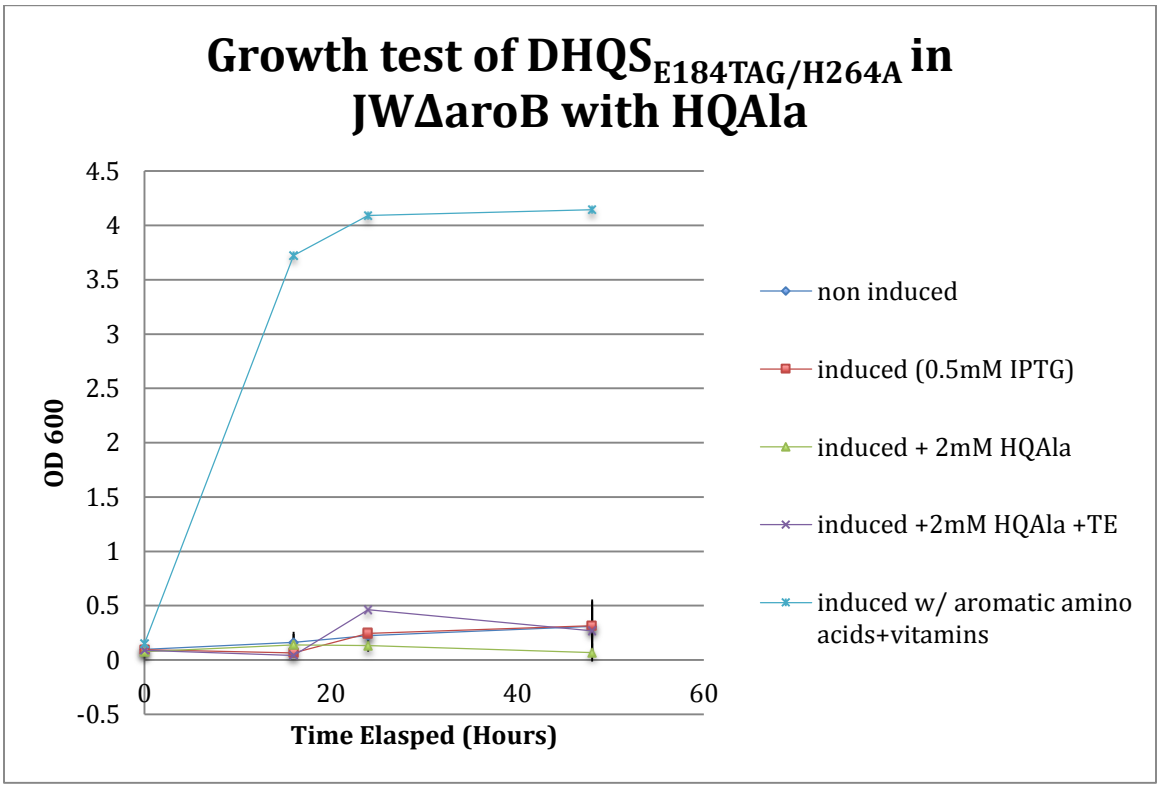


Figure 2-4A: Growth of JWΔaroB/ pJF-aroB$_{E184TAG/H264A}$ at 37°C in M9 Minimal Media without aromatic amino acids and vitamins unless noted.

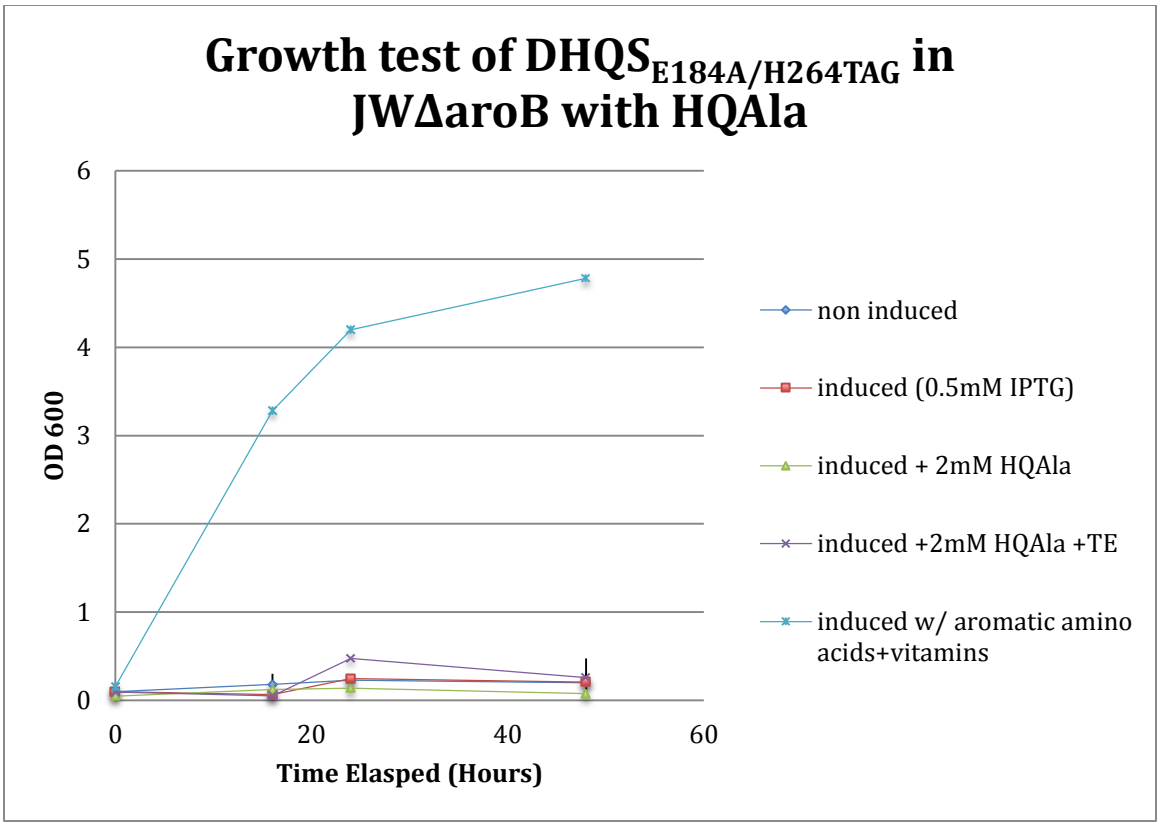


Figure 2-4B: Growth of JWΔaroB/ pJF-aroB$_{E184A/H264TAG}$ at 37°C in M9 Minimal Media without aromatic amino acids and vitamins unless noted.

However, as observed, even after 48 hours, there was little growth of either the DHQS$_{E184TAG/H264A}$ or the DHQS$_{E184A/H264TAG}$ mutant in the absence of aromatic amino acids and vitamins. This meant that there was no activity of the DHQS mutants. To ensure that the issue did not lie in the incorporation efficiency of HQAla into the enzyme, the DHQS$_{E184TAG/H264A}$ and DHQS$_{E184A/H264TAG}$ mutants were expressed in JWΔaroB in rich LB cultures in the presence of the pEvol plasmid and 2 mM HQAla. The cell lysate was partially purified by affinity chromatography. The protein samples were subsequently analyzed with Western Blot using anti-His antibody as the DHQS mutants are modified with a C-terminal histidine tag to assist in purification (Fig 2-5).

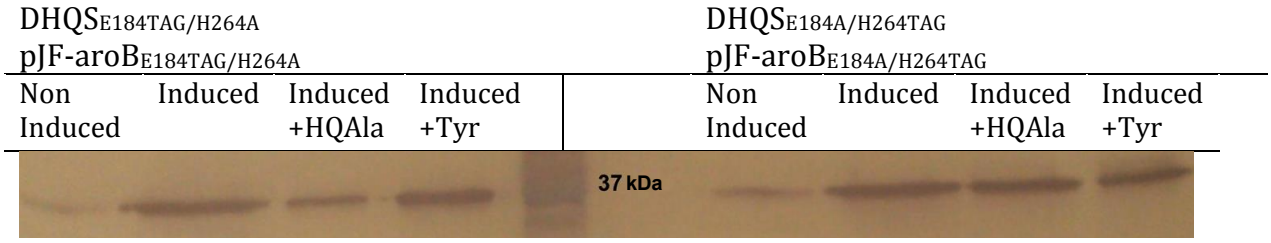


Figure 2-5: Western Blot of $DHQS_{E184TAG/H264A}$ and $DHQS_{E184A/H264TAG}$ mutants, non induced, induced (0.5 mM IPTG and 0.02% arabinose),, induced with HQAla (2mM HQAla, 0.5 mM IPTG and 0.02% arabinose), and induced with Tyrosine (2mM Tyrosine, 0.5 mM IPTG and 0.02% arabinose), in M9 media with aromatic amino acids and vitamins. Full length DHQS is 39kDa.

As Figure 2-5 shows, there was protein produced in the induced cultures with the 2mM of HQAla. However, it is important to note that there is significant background in the induced culture without HQAla, indicating there was incorporation of a canonical amino acid. Checking the incorporation efficiency of the HQAla in the M9 minimal media without aromatic amino acids and vitamins is difficult as there is so little cells in the solution. Attempts at doing a Western Blot of these cultures were too faint to make any key distinctions of DHQS expressed.

The background of unspecific amino acid incorporation in Figure 2-5 was expected, as there is some background in most unAA incorporation and usually other mutations are necessary to increase specificity.[27] The aminoacyl-tRNA synthetase that was mutated to incorporate HQAla is derived from a library developed from tyrosine tRNA synthetase,[25] so it is likely in these rich culture tyrosine is being incorporated due to the nature of the synthetase. However it is not clear why there is high selectivity of the HQAla in the selection experiments, but in

this case there is likely little protein with the incorporated HQAla. One possible reasoning is that there is overcrowding in the active site. As the HQAla is a bicyclical structure, as it is considerably more bulky than either the glutamic acid or histidine. The histidine 264 might have less of an issue in the incorporation of HQAla as it had a bigger footprint, however in both situations incorporation may result in mis-folding and/or truncated proteins

From the Western, shown in Figure 2-5, it is likely that tyrosine is being incorporated at the amber nonsense codon. Significantly more protein is expressed in the $DHQS_{E184A/H264TAG}$, and after incubation in M9 minimal media over 72 hours, a slight increase in growth were observed by these cultures. These factors then lead to asking the question of how the DHQS enzyme will function if Tyrosine is incorporated at position 264 and 184.

2.4 Exploration of Activity of Tyrosine as Chelator by $DHQS_{E184Y}$ and $DHQS_{H264Y}$ mutant

As demonstrated in the growth test shown in figure 2-4 A and B above the DHQS mutants showed little activity if HQAla was incorporated. Analysis of the protein expression with and without the HQAla showed there was non-specific incorporation by the HQAlaRS. It is likely that the amino acid incorporated was Tyrosine, due to the original template of the HQAlaRS. To test whether Tyrosine had any ability in chelating the divalent metal, two additional mutants were made: $DHQS_{E184Y}$ (pJF-$aroB_{E184Y}$) and $DHQS_{H264Y}$ (pJF-$aroB_{H264Y}$) (Section4.6). As Tyrosine

could only potentially donate one electron pair for chelation, unlike HQAla, the other residue remained at the wild type residue to facilitate the chelation of Co^{2+}.

These mutants were then transformed into the JWΔaroB strain and used in a growth test in M9 minimal media following the same experimental procedures as described previously except time points were 0hrs, 24 hrs, and 48 hrs (Figure 2-6 A and B).

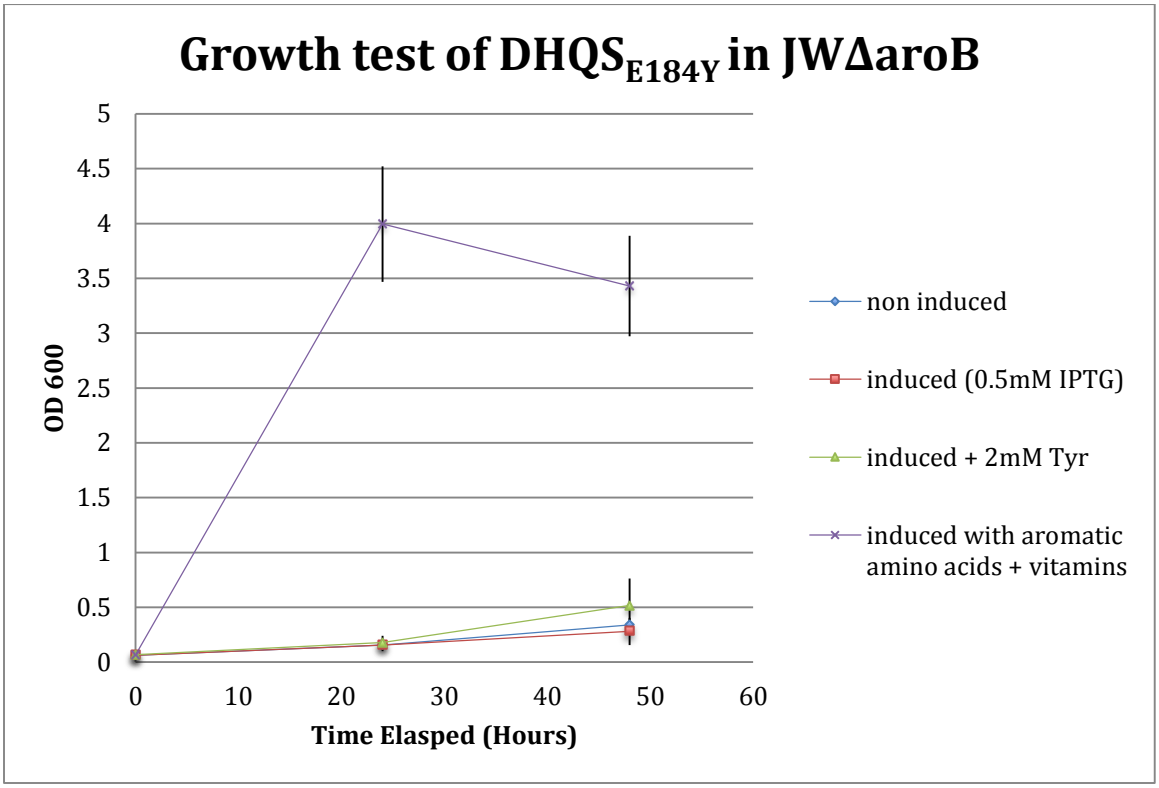


Figure 2-6A: Growth of JWΔaroB/ DHQS$_{E184Y}$ at 37°C in M9 Minimal Media without aromatic amino acids and vitamins unless noted.

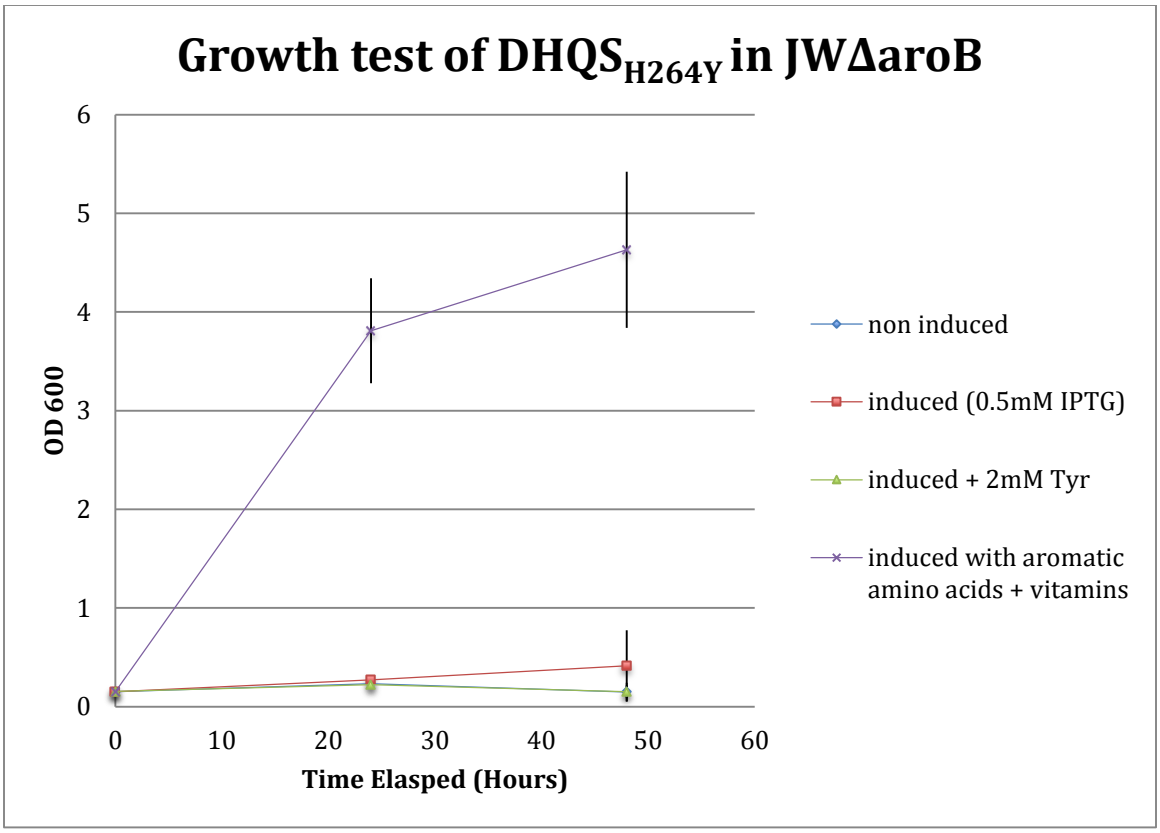


Figure 2-6B: Growth of JWΔaroB/ DHQS$_{H264Y}$ at 37°C in M9 Minimal Media without aromatic amino acids and vitamins unless noted.

Though the H264Y mutants had higher activity there was no significant improvement in activity. The OD over 3 days remained low, with no significant bounce back. This set of experiment showed that though the HQAlaRS was incorporating tyrosine, from solution it did not result in restoring activity. Overall it is likely the unAA HQAla is too large, causing mis-folding, and Tyrosine is unable to act as a chelating residue to replace either H264 of E184.

2.5 Growth of Histidine Analogue, 2-(5-bromothineyl)alanine, at metal Chelating Site H264

As we attempted to mutate the Histidine into a completely different amino acid analogue, the activity drastically decreased. When introducing a histidine analogue at position 264 in the active site, an attempt to reestablish proper folding and metal chelation was made. This unAA could potentially restore some activity to the enzyme. The histidine analogue used in this experiment was 2-(5-bromothienyl)alanine (BrTha) (Figure 2-7). This unnatural amino acid was purchased from Chem-Impex International and the selection process for an aaRS to charge this unAA is described in section 3.2.

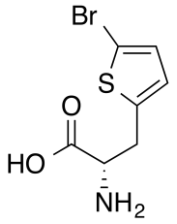


Figure 2-7: L-2-(5-bromothienyl)alanine (BrTha) structure

To set up the plasmids for this mutant was not simple. As the vector with the aaRS encoded does not contain the tRNA to suppress the Amber nonsense codon, the tRNA was moved onto the pJF vector. This gene was inserted by SLIC (Sequence and Ligation Independent Cloning) cloning downstream of the *aroB* gene on the pJF vector, due to problems in amplification extra non coding sequence was required to be added before and after the tRNA. The final vector was pJF118EH-aroB_x-PylT (x= being codon of TAG) (Section 4.6). As the 264 position was once a histidine this

position was chosen for incorporation of the histidine analog. This mutant was made by overlapping PCR to incorporate the aroB$_{E184A/H264TAG}$ into the pJF118EH-aroB$_X$-pylT forming pJF118EH-aroB$_{E184A/H264TAG}$-PylT.

This mutant was then co-transformed with the BrThaRS encoding sequence, pBK-BrThaRS4, into JWΔaroB and used in a growth test in M9 minimal media using the same experimental procedures as described previously (Figure 2-8).

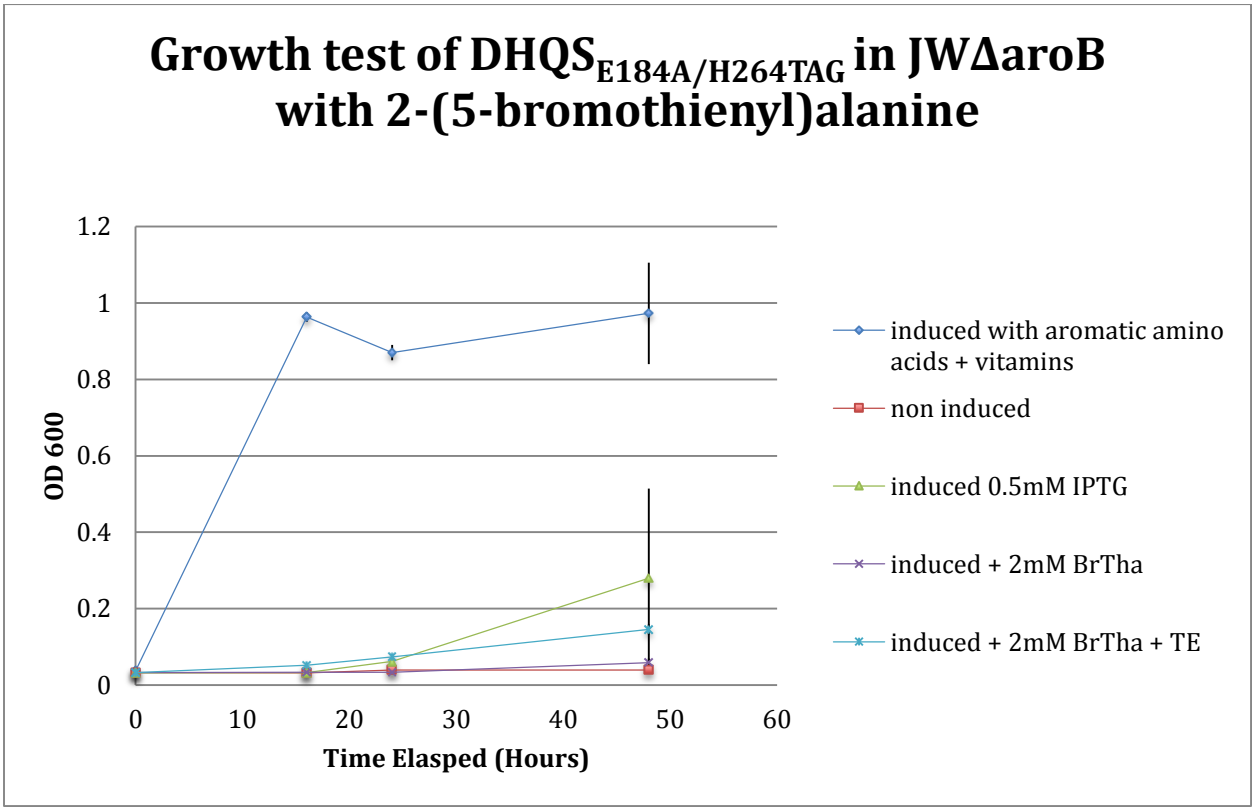


Figure 2-8: Growth of JWΔaroB/DHQS$_{E184A/H264TAG}$ at 37°C in M9 Minimal Media with BrTha without aromatic amino acids and vitamins unless noted.

2.6 Conclusion

Though there were efficient mechanisms to incorporate HQAla into the DHQS, no functional protein were produced with restored enzyme activity. In rich media it appeared that the HQAlaRS incorporated tyrosine or with small amounts of the HQAla. With either HQAla or tyrosine incorporated into either the 184 or the 264 sites, neither showed any major gain toward restoration in the enzyme activity. The biggest apparent problem for the HQAla incorporation into DHQS is the constrained active site. Changing the DHQS active site for the HQAla to be a better fit would be a potential route to increase incorporation of HQAla. However as the DHQS catalyzes multiple reactions, finding the space for HQAla and maintaining the catalytic activity might be a bigger issue.

The incorporation of histidine analogues into the active site seems to be, potentially, a more useful route. As histidine was already at the 264 site by incorporating analogues at this position might create a better understanding in how divalent metal ion facilitates hydride transfer from C5 to NAD^+ as well as proton loss of the C5 hydroxyl. However there are very few histidine analogues that have been shown to be successfully incorporated into proteins.

Chapter 3

Histidine Analogue Incorporation for Acid/Base catalysis in DHQS

3.1 Introduction

Though the oxidation catalyzed by the divalent metal in the DHQS is interesting, it is not the only residue involved in the oxidation of the C5 alcohol. The Histidine 275 of the *Aspergillus nidulans* DHQS acts as a proton acceptor in the acid/base reaction that allows for the oxidation. By changing the basicity of this residue, the reaction rate of this first step would be altered. Histidine analogs are a potential unAA to genetically incorporate as the electronics can be varied, but the relative structure would remain the same.

Currently there are not many histidine analogs that can be used for genetic incorporation by an aaRS,[28] and fewer that can be incorporated into protein using the amber suppression method,[22,22] especially compared to tyrosine analogues. To expand the tools that scientists can use would be extremely beneficial, especially as cells often use histidine in catalysis. The process, however, to find a good tRNA/aaRS pair to recognize these analogues is difficult.

Only a few notable experiments have been done incorporating a histidine analog by amber suppression. In 2014 the Schultz group showed an incorporation of 3-methyl-histidine, 2-furyl-alanine, 3-(2-thienyl)-alanine and 2-(5-bromothineyl)-alanine.[21] These unAA were incorporated using a selected *Mb*PylRS that was initially evolved for the incorporation of 3-methyl-histidine and polyspecific in recognition of unAA, with relatively high efficiency towards the ones mentioned above.[21] In the

same year, the Soll group published a paper showing a polyspecific *Mb*PylRS able to incorporate 2-(3-bromothineyl)-alanine[22] (BrTha) along with several other tyrosine analogues. Though the incorporation efficiency of the BrTha was low, it is a step toward the development in incorporating histidine analogues.

3.2 Library design and selection of tRNA/tRNA synthetase pair for histidine analogs

There are very little known tRNA/aaRS pairs that incorporate histidine analogues, but many that can incorporate tyrosine derivatives. There is a great amount of potential for protein engineering if there were tools to incorporate histidine analogs as often histidine is found in active site with key roles in catalysis. As there has been success in two separate projects incorporating the histidine analogue using an *Mb*PylRS library we used a similar library to get specific incorporation of a histidine analogue rather than selection for a polyspecific aaRS.

Using the 3N *Mb*PylRS library in our lab, we selected for MbPylRS mutants that were able to charge $tRNA^{Pyl}$ with the following histidine derivatives: 2-(5-bromothienyl)alanine (**1**), 2-mercapto-L-histine (**2**), diiodo-histine (**3**), and 2-(5-carboxythienyl)alanine (**4**) (Figure 1-5). The 2-(5-carboxythienyl)alanine was newly synthesized by myself and is based off of the Copper (I)-thiophene-2-carboxylate (CuTC) complex, which is a good metal chelator, particularly to copper. The 3N library has randomized sites at Leu270, Tyr271, Leu274, and Cys313. The Tyr349 was fixed to Phe as it has been previously shown to increase the aminoacylation with tRNA.[29]

Each histidine analogue was used in making GMML selection plates at 2mM concentrations; the GMML as described in section 1.2 was used to reduce the potential background. As minimal media was used, the plates needed to be incubated for 2 days at 37°C as the growth of the cells were reduced. For the 2-mercapto-L-histine in the first round of selection at Cm concentration of 34 μg/mL, there were only four fluorescent colonies, therefore potential hits. After replication, however, all colonies showed to be false positives. As for the diiodo-histine, it showed 15 potential hits after the first selection at Cm concentration of 34 μg/mL, however they were also all false positives. The 2-(5-carboxythienyl)alanine was also used with the 3N library and though there were 12 potential hits after the first round at Cm concentration of 34 μg/mL, none of them were real hits.

The 2-(5-bromothienyl)alanine, however had five successful hits out of the 34 potential colonies. The plasmid containing the aaRS (pBK-*Mb*PylRS) was isolated and sent for sequencing (Figure 3-1). Looking at the sequence of each mutant showed unique sequence compared to the polyspecific aaRS published by the Schultz group.[21] These plasmids were co-transformed with the reporter plasmid pLei-GFP_{149TAG} and inoculated into three cultures, LB non induced, LB induced 0.5 mM IPTG, and LB induced 0.5 mM IPTG with 2mM of 2-(5-bromothienyl)alanine.

Position	270	271	274	313	249
Polyspecific aaRS[21]	Ile (I)	Phe (F)	Gly (G)	Phe (F)	Phe (F)
wt	Leu (L)	Tyr (Y)	Leu (L)	Cys (C)	Tyr (Y)
hit-2	Thr (T)	Phe (F)	Ala (A)	Phe (F)	Phe (F)
	ACG	TTT	GCG	TTT	TTT
hit-3	Leu (L)	Phe (F)	Val (V)	Met (M)	Phe (F)
	CTT	TTT	GTG	ATG	TTT
hit-4	Leu (L)	Phe (F)	Thr (T)	Leu (L)	Phe (F)
	TTG	TTT	ACG	CTT	TTT
hit-6	Val (V)	Phe (F)	Ala (A)	Met (M)	Phe (F)
	GTG	TTT	GCT	ATG	TTT
hit-8	Gly (G)	Phe (F)	Asp (D)	Ser (S)	Phe (F)
	GGG	TTT	GAT	TCG	TTT

Table 3-1: Sequencing of randomized codon for 3N *Mb*PylRS library hits charging 2-(5-bromothienyl)alanine (BrTha).

All five hits were strongly florescent with the addition of the unAA, 2-(5-bromothienyl)alanine. The fluorescent signal of each culture was normalized by the number of cells in that culture. This normalized fluorescence of the induced culture without unAA was compared to the induced culture with unAA. The larger the ratio of the induced culture to the induced culture with unAA correlated to a higher specificity of the aaRS towards the BrTha. The largest ratio of the induced culture

with the unAA to the induced culture without unAA was mutant 4 (Figure 3.2). As the mutant 4 had the highest specificity to BrTha, it was used in the DHQS experiments. These synthetase mutants were then tested for the incorporation of other histidine analogues, mercapto-L-histine, diiodo-histine, and 2-(5-carboxythienyl)alanine. However none of these mutants were able to show indication of incorporation of the other unAA.

Mutant	Fluorescence (RFU)			Normalized Fluorescence (RFU/OD)			Ratio
	non	i	i +BrTha	non	i	i+ BrTha	(i+BrTha)/(i)
2	265	663	9212	128.3	348.0	6533.3	18.7
3	61	121	5452	21.7	84.3	2931.2	34.7
4	26	51	6020	9.7	20.4	2887.3	141.5
6	79	109	5339	41.4	83.5	2591.7	31.0
8	62	124	3614	23.7	69.5	2007.8	28.9

Table 3-2: Incorporation efficiency of MbPylRS hits towards BrTha. [(non) – non induced culture; (i) – induced with 0.5mM IPTG; (i + BrTha) – induced with 0.5mM IPTG and 2mM 2-(5-bromothienyl)alanine]

As mutant 4 (designated as BrThaRS) was able to charge $tRNA^{Pyl}$ with 2-(5-bromothienyl)alanine efficiently, it was used as a template for the construction of a new error-prone PCR library. We hope we can identify *Mb*PylRS mutants from this

library for the genetic incorporation of other histidine analogues that are not recognized by the *Mb*PylRS mutants in the 3N library. Instead of mutating particular codons, the whole aaRS was randomized through a random mutagenesis enzyme, mutazyme. By using this enzyme in amplification of the gene, random errors could occur along the whole sequence. Small changes in the sequence can cause the ability for a synthetase to change its ability to bind a particular amino acid. In the example of selection a TyrRS mutant to bind *para*-acetylphenylalanine, the Schultz group found a small change in the location of an alpha helix at the back of the binding pocket affected specificity greatly.[30] Overall, small changes in the folding of the BrTharRS could increase binding of other histidine analogs.

After amplifying the BrThaRS with a mutazyme, the amplified sequence was cloned into the pBK vector. This mutazyme made small errors randomly as during each round of amplification. The exact size of the library could not be determined as the mutagenesis is random. The library was then transformed into the reporter strain via electroporation. This newly constructed random mutagenesis (RM) library, $BrThaRS_{RM}$, was then used for selection against mercapto-L-histine, diiodo-histine, and 2-(5-carboxythienyl)alanine. No positive hits were obtained for the diiodo-histidine. The selection against mercapto-L-histine resulted in three weak hits, whereas the 2-(5-carboxythienyl)alanine selection resulted in two weak hits (Figure 3-3).

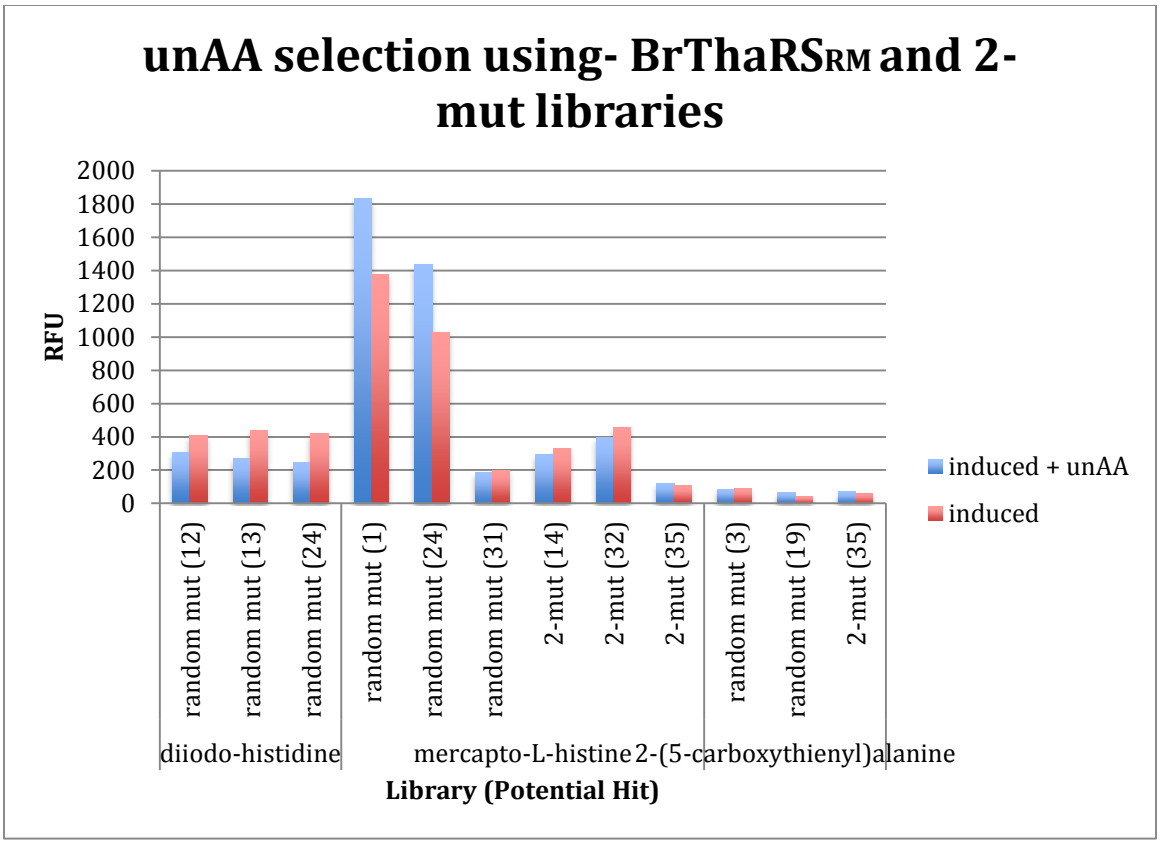


Figure 3-1: Selection of libraries, BrThaRS$_{RM}$ and 2-mut library against mercapto-L-histine, diiodo-histine, and 2-(5-carboxythienyl)alanine

Two other libraries including the WN1, and the 2-mut library, both constructed by Dr. Wei Niu in our lab, were used in selection for the genetic incorporation of other histidine analogues. Unfortunately, no hit was identified from the WN1 library. Though the 2-mut library had three weak hits for the mercapto-L-histine, and one for the 2-(5-carboxythienyl)alanine, there were none for the diiodo-histidine (Figure 3-1). Over the selection process it appeared that the diiodo-histidine is slightly toxic, as there are significantly less colonies that were able to grow, often only up to ten or twenty in comparison to the other selection plates, which grew over a hundred colonies.

As the results from the PylRS selections showed no hits for the incorporation of the other histidine analogues, finding ways to develop a new library to incorporate, specifically, histidine analogues would be beneficial. Previously with the random mutagenesis library, it was hoped that minor changes in folding would result in recognition of the other histidine analogues. But as the work of the Soll group gave some understanding of the mechanism of binding of the 2-(5-bromothienyl)alanine for their polyspecific PylRS;[22] perhaps crystallization of our 2-(5-bromothienyl)alanine could also give some insight in how the synthetase binds histidine analogues.

As crystallization of proteins can give in-depth understanding of how a protein interacts with a substrate, to crystallize BrThaRS would be helpful in determining which codons to randomize to make new libraries for the incorporation of histidine analogs. With a collaboration with the Dr. Wilson lab in the Department of Biochemistry of UNL, we started on the process of crystallizing our BrThaRS that charges 2-(5-bromothienyl)alanine. In previous work, all PylRS structures are crystalized from the *Mm*PylRS, therefore any hits from *Mb*PylRS libraries are analyzed and analogous residues are mutated on the wild-type *Mm*PylRS to match it. As there is no structure for the *Mb*PylRS, we decided to see if we could crystallize our BrThaRS mutant, a derivative of *Mb*PylRS, and develop a method for its crystallization.

The *Mm*PylRS and the *Mb*PylRS are very similar; thereby we designed the expression vector for the BrThaRS much like what was has been reported previously for the *Mm*PylRS. To this end, we inserted the PylRS gene into the

pET28b vector with a C terminus His tag installed.[31] As often, there are issues in solubility of full length proteins, we also amplified a truncated version of the BrThaRS mutant that still maintains the structural components of the active site (Section 4.6). After cloning the amplified sections of the BrThaRS gene into the pET28b vector these plasmids were transformed into the host BL21(DE3) to ensure high protein production.

E. coli BL21(DE3) strain harboring plasmid pET28b-BrThaRS (full-length or truncated version) was cultured in 1 L LB media containing kanamycin. The protein expression was induced by the addition of 1.5 mM IPTG. After overnight growth at room temperature, cells were collected and lysed. Both the full length and the truncated proteins were purified using Ni^{2+} Sepharose 6 Fast Flow resin (GE Healthcare). As shown in the SDS-PAGE analysis (Figure 3-2A and B), a large amount of protein was obtained with high purity. It was found that by keeping the salt concentration at 500mM but removing phosphate and imidazole from solution the protein remains soluble. Therefore there is enough to begin process of crystallization.

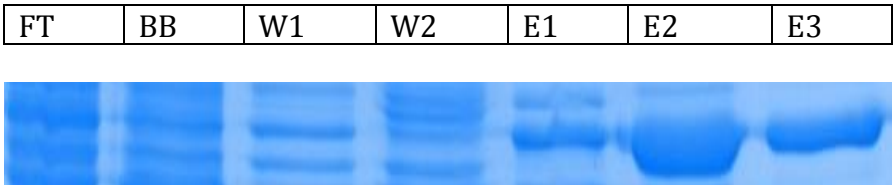


Figure 3-2A: pET28b-BrThaRS$_{\text{full length}}$ expressed in BL21(DE3) cells. Protein size 49.7 kDa

FT	BB	W1	W2	E1	E2	E3

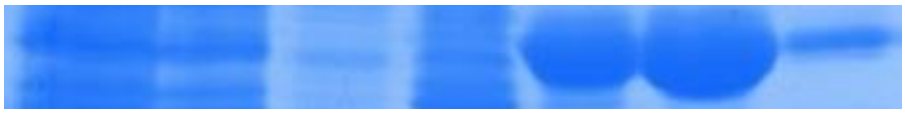


Figure 3-2B: pET28b-BrTha$_{truncated}$ expressed in BL21(DE3) cells. Protein size 33.3 kDa

3.3 Construction of DHQS mutant and expression plasmid

For this work, emphasis on the DHQS and how protein engineering could help understand the mechanism, and affect the activity of the enzyme, was key. Focusing on another active histidine, we mutated the His 251 codon to the Amber nonsense codon. This 251 site is analogous to the 275 site shown on the crystallized DHQS, as published by the Brown group.[18] Using overlapping PCR this site, H251, was mutated to the Amber nonsense codon. This position would be interesting to examine because of its role as a Lewis base in the oxidation process of the C5 oxygen of DAHP.

As previously described, a tRNA suppresses this nonsense codon in the ribosome by the incorporation of an unAA. As our selection process uses the pBK plasmid to encode the aaRS library, there is no tRNA sequence encoded. As the aaRS needs a tRNA that recognizes the nonsense codon, this was installed downstream of the *aroB* sequence on the pJF plasmid (Section 4.6). Initially the tRNA sequence was to be installed via SLIC cloning with minimal excesses non-coding sequence. However amplification and annealing of the minimal tRNA insert proved to be difficult. So the amplified sequence was expanded to include noncoding sequence,

allowing for the successful cloning of the tRNA into the pJF-aroB$_{H251TAG}$ plasmid forming pJF118EH-aroB$_{H251TAG}$-pylT. This expanded plasmid was used for the incorporation experiments with 2-(5-bromothienyl)alanine.

3.4 Growth test of His 251 site mutant of DHQS

While we were only able to genetically incorporate one histidine analogue, 2-(5-bromothienyl)alanine, we decided to use it to probe the catalytic role of His251 of DHQS. After co-transforming both plasmids, pJF118EH-aroB$_{H251TAG}$-pylT and pBK-BrThaRS into the JWΔaroB strain, the cell growth was monitored in the same way as described in the metal chelating mutant experiments. OD was taken after 12 hours, 24 hours, and 48 hours (Figure 3-3). As there was significant jump in growth after 48 hours the cultures were incubated longer till 114 hours.

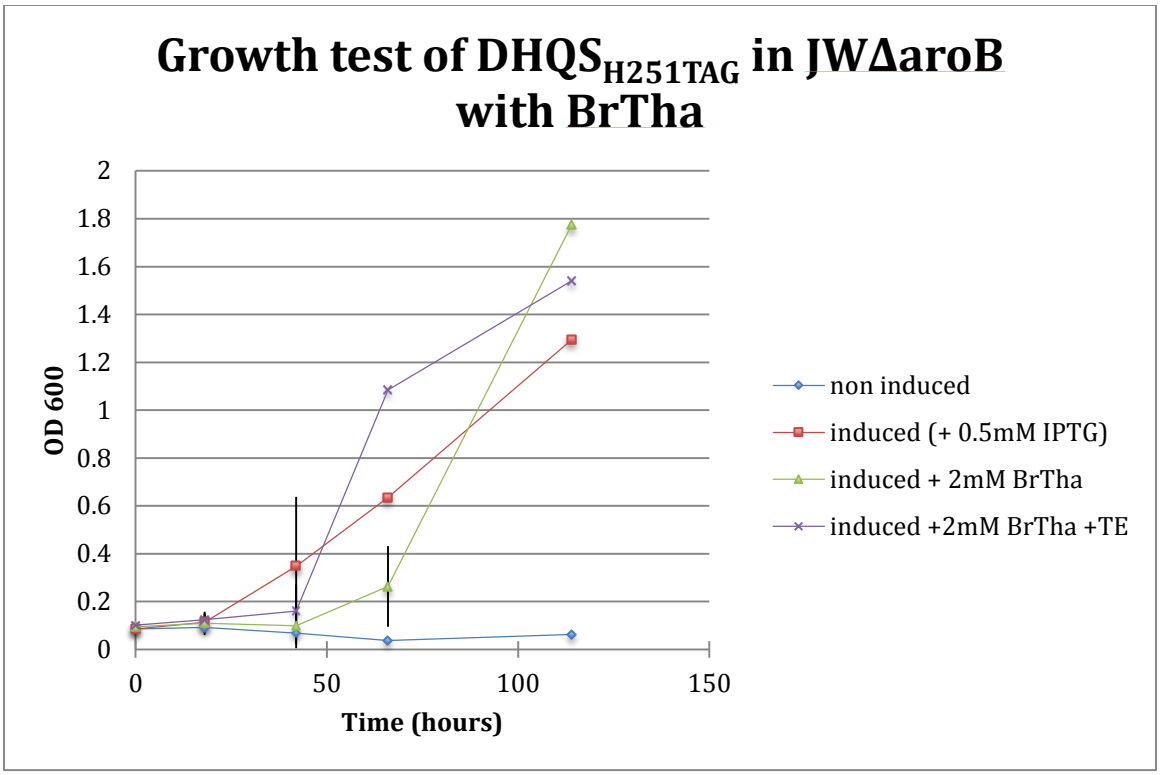


Figure 3-3: Growth of JWΔaroB/ DHQS$_{H251TAG}$ at 37°C in M9 Minimal Media without aromatic amino acids and vitamins.

The DHQS mutant with the BrTha incorporated at the 251 site had an overall lower activity than the wild type DHQS as growth was only seen after about 72 hours. Though this result was promising there was the issue that the induced culture (without unAA) grew more rapidly in the first 48 hours, though plateaued later. As DHQS should be truncated if there is no 2-(5-bromothienyl)alanine added, this means the synthetase is incorporating some other amino acid**.** As there was growth in the induced cultures (with or without unAA), the sequence of each culture was checked after the 114 hour incubation. The initial inoculating culture, the induced culture, and the induced + BrTha + trace elements were compared to see if there were any mutations (Figure 3-4).

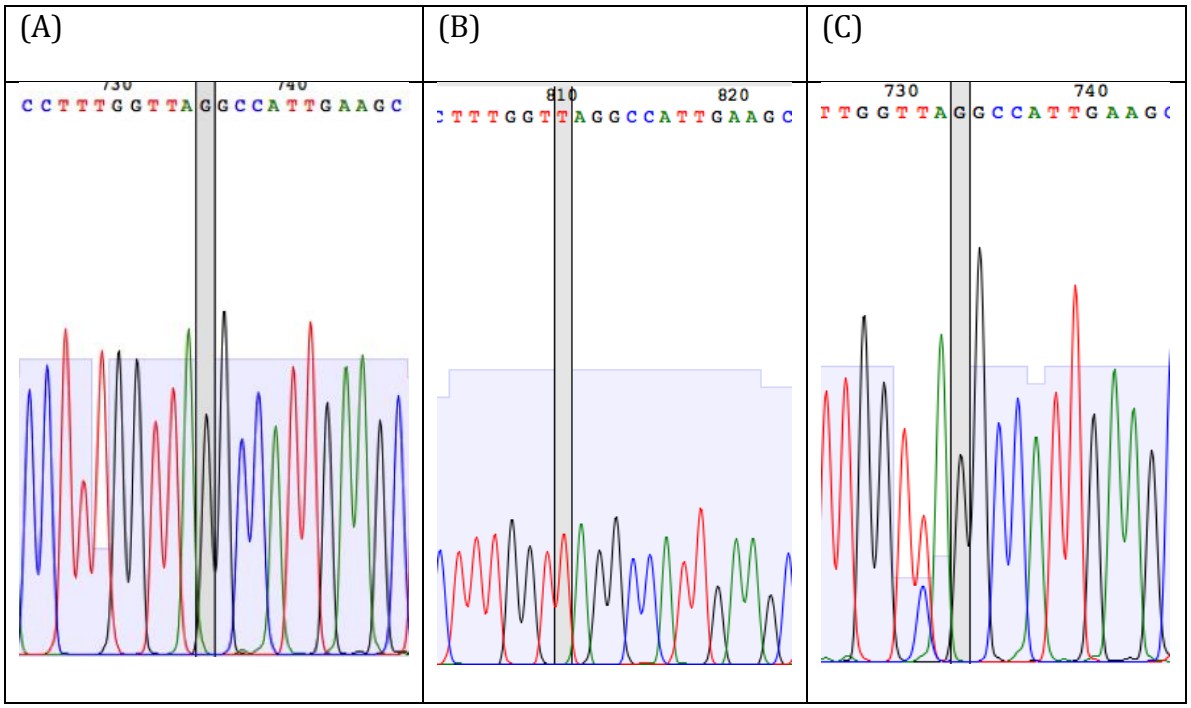


Figure 3-4: Codon sequencing of H251 (A) inoculating culture (B) induced culture, (C) induced culture + 2-(5-bromothienyl)alanine + trace elements

Comparing the sequence of the three cultures showed interesting results, as the induced culture (without unAA) had no mutations and the TAG codon remained unchanged from the initial culture. The induced culture + BrTha+ trace elements, however, showed that there was a mutation from TAG to CAG at about a 60:40 ratio. A mutation to this codon was unexpected, as there is no pressure for mutation with the 2-(5-bromothienyl)alanine in solution. This codon mutation was seen repetitively in multiple grow tests, as such it is hypothesized that it is an error was introduced during the construction of the $aroB_{H251TAG}$ mutant. To purify the culture, the initial culture was streaked out and a single colony was selected. This isolated plasmid was transformed into Genehog. This process was repeated twice more with the final resulting plasmid sent for sequencing. This plasmid was co-transformed into the JWΔaroB host with pBK-aaRSBrH4 and the growth test was repeated

(Figure 3-5). After 66 hours all the cultures began to grow including the induced culture without the unAA. Again the cultures were isolated to check if the codon mutated again, however this time all cultures were the same as the induced culture (Figure 3-6).

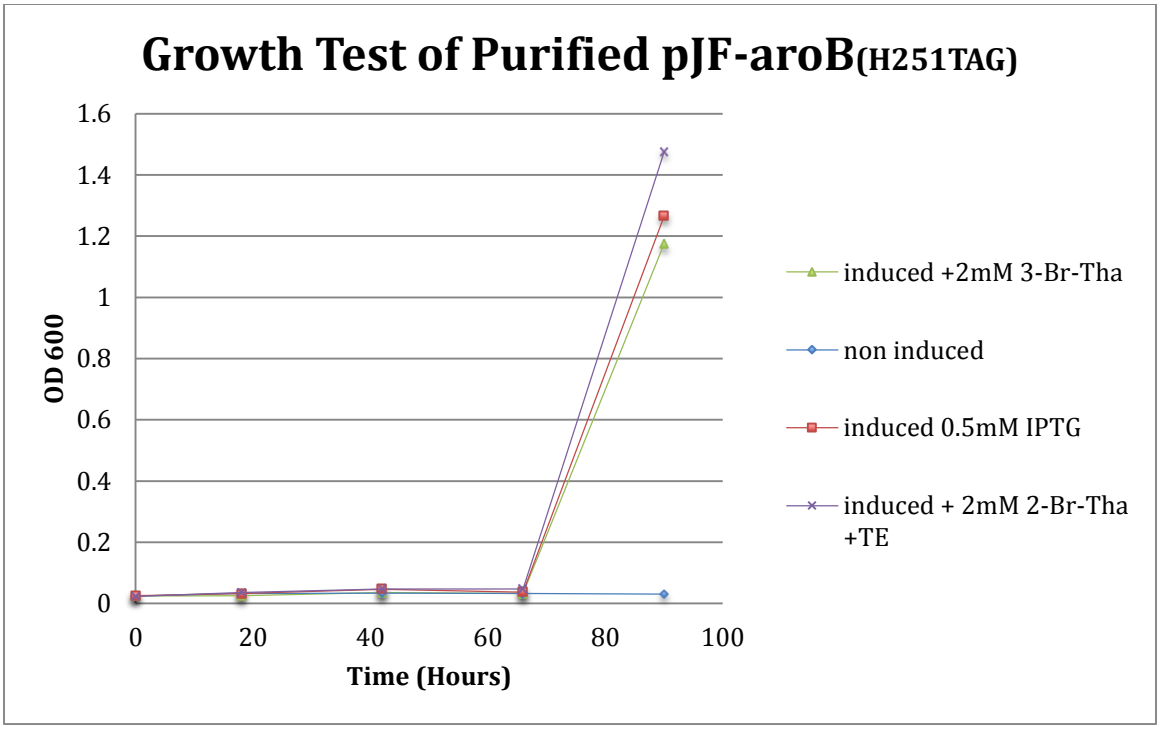


Figure 3-5: Growth test of purified pJF-aroB$_{H251TAG}$ plasmid in JWΔaroB strain grown at 37°C in M9 Minimal Media without aromatic amino acids and vitamins.

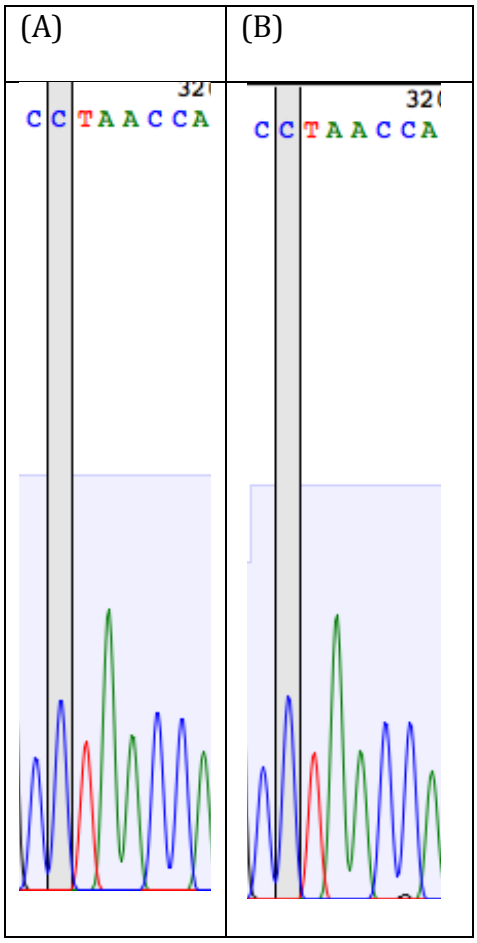


Figure 3-6: Sequence of (A) induced culture and (B) induced culture + 2-(5-bromothienyl)alanine + trace elements

With the culture without the unAA still growing to a similar level to the culture with the unAA, the pJF plasmid was tested by itself without the synthetase. Rationally if there was no synthetase to charge amber suppresser tRNA, then when the culture is induced the DHQS would be truncated and nonfunctional.

Transforming the pJF-aroB$_{H251TAG}$ by itself into the JWΔaroB strain the mutant was monitored during the growth test (Figure 3-7). With no synthetase being co-expressed, the fact that the culture still grew indicated a fault with the plasmid designed or with the strain used. To test this, we used the second knockdown strain of AB2847. This strain still has an *aroB* gene though the specific activity of its encoded protein is much reduced. After co-transforming the pJF-aroB$_{H251TAG}$ and the pBK-BrThaRS together this mutant was examined in a similar growth test as described before (Figure 3-8).

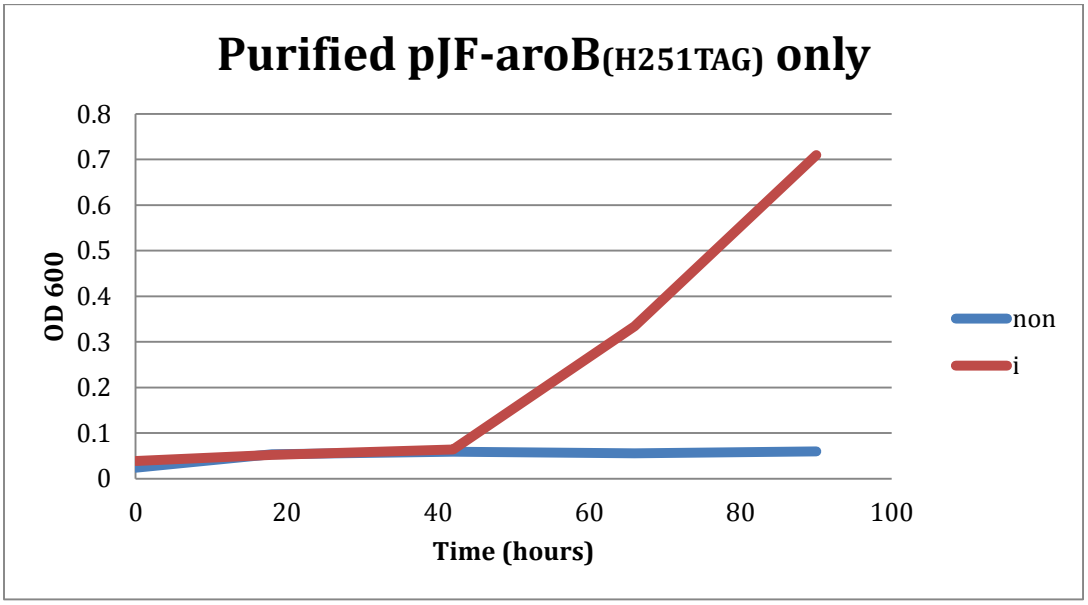


Figure 3-7: Growth test of purified pJF-aroB$_{H251TAG}$ in JWΔaroB strain grown at 37°C in M9 Minimal Media (non: non induced) (i: induced).

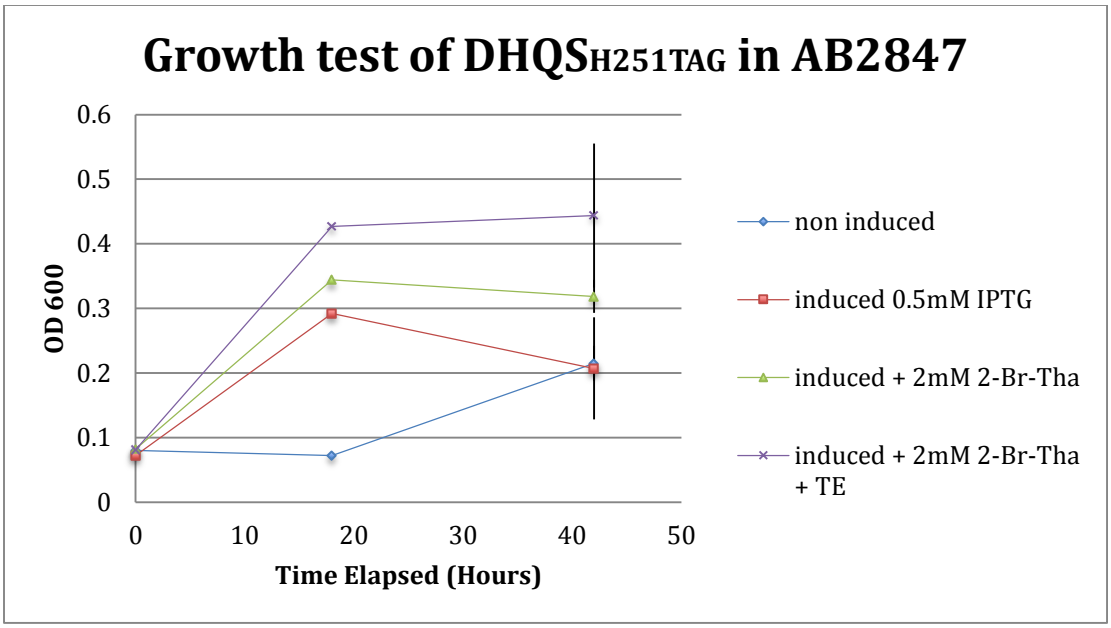


Figure 3-8: Growth test of purified pJF-aroB$_{H251TAG}$-pylT in AB2847 strain grown at 37°C in M9 Minimal Media without aromatic amino acids and vitamins.

From this data, overall it showed that there was an issue of the pJF-aroB$_{H251TAG}$-pylT. This is significantly noticeable as moving hosts and induction by itself results in growth of the culture was observed in the absence of BrThaRS. It is unclear why this phenomenon was occurring, as the same vector and construct design was used with the mutants DHQS$_{E184TAG/H264A}$ and DHQS$_{E184A/H264TAG}$ and this issue was clearly absent there. Comparing the amount of protein by western blot shows that a small amount of DHQS$_{H251TAG}$ is expressed in the presence of unAA, but no noticeable amount of protein is expressed in the induced cultures without unAA. Therefore, since there is an amount expressed, this mutant is still used for the activity assay in section 3.5.

Lastly, in the interest of why the codon mutation from TAG251 to CAG251 happened and if this glutamine has any effect in restoring activity, another mutant

pJF-aroB$_{H251CAG}$-pylT was constructed using overlapping PCR from the pJF-aroB$_{H251TAG}$-pylT vector (Section 4.6). This plasmid was then co-transformed into the JWΔaroB strain for comparison to the DHQS$_{H251TAG}$ enzyme to see the activity difference. After setting up the growth test in the same manner as previously discussed, surprisingly, the DHQS$_{H251Q}$ showed to have high activity, growing significantly after 24 hours (Figure 3-9). As this mutation showed to restore a significant amount of activity, this enzyme was used for the activity assay in section 3.5 as well. Though this mutant was shown to restore activity it is not clear why as the glutamine should not be able to act as a Lewis base needed for the oxidation of the C5 alcohol of DAHP.

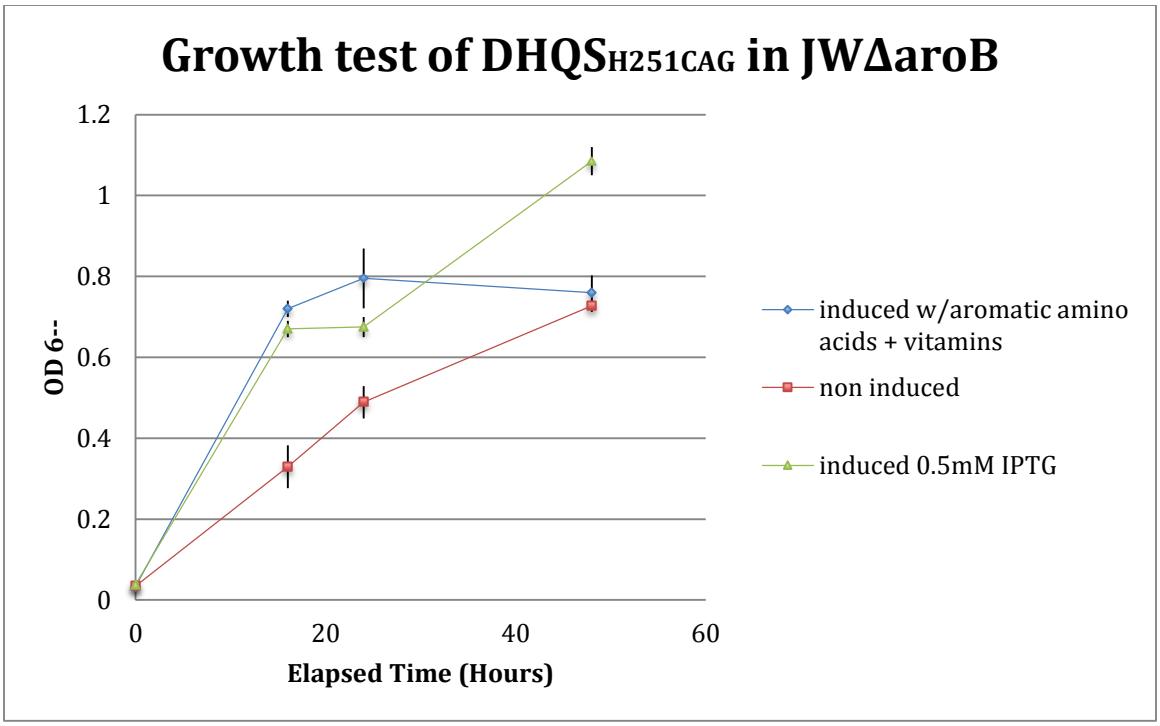


Figure 3-9: Growth test of mutant pJF-aroB$_{H251CAG}$ plasmid in JWΔaroB strain grown at 37°C in M9 Minimal Media without aromatic amino acids and vitamins unless noted.

3.5 Activity Assay of DHQS$_{H251TAG}$

To compare the relative activity of DHQS$_{H251BrTha}$, to the activity of the DHQS$_{wt}$ both enzymes were purified and assayed. The assay is based on the P$_i$per Phosphate Assay that measures the amount of inorganic phosphate in a solution. This is because the DHQS releases a phosphate into solution during the catalysis of DAHP to DHQ. As such, the activity of the DHQS variants can be determined based off of how much phosphate is released into the solution.

This assay was prepared by first incubating the DHQS$_{wt}$, DHQS$_{H251BrTha}$, and the DHQS$_{H251Q}$ with approximately 50 mM of DAHP.[32] The concentration of DAHP

was an estimated value that was based on ^{1}H NMR integration (Section 4.5). The assay was conducted in a buffer containing 50mM MOPS pH 7.75, 50 μM cobalt (II) Zinc, 10 μM NAD^{+}.[32] The assay mixture was incubated at room temperature for 30 minutes in a 96 well plate.[32] After incubation, the reaction was diluted with 50 μL of the $P_i$per Phosphate Assay solution and incubated for another 30 minutes at 37°C. The absorbance of each reaction was taken at 563 nm, as this is the absorbance of Amplex Red formed at the end of the $P_i$per Phosphate Assay. Unfortunately this assay was gave inconsistent results. This was probably due to the extra phosphate in the purified DAHP.

As the $P_i$per Phosphate Assay was not consistent, we decided to use the Couple Enzyme Assay. This was based of the 1997 paper by Frost, the basic buffer solution described above: a buffer containing 50mM MOPS pH 7.75, 50 μM cobalt (II) Zinc, 10 μM NAD^{+}. As 3-dehydroshikimate has an absorbance at 234 nm, the 3-dehydroquinate dehydrate is used in this couple enzyme assay. 0.1 mg of dehydrate is suspended in a solution with DAHP and incubated at room temperature for 5 minutes. Then 0.02 mg of the DHQS mutant is added and the absorbance is monitored over the first 90 seconds (Figure 3-10). Calculating the specific activity of the enzymes, we got the approximate values of 0.625 μmol mg^{-1} min^{-1} for $DHQS_{wt}$ and 0.327 μmol mg^{-1} min^{-1} for $DHQS_{H251TAG}$. The $DHQS_{H251Q}$ was assayed using this same assay however the results were inconsistent. The DHQS of the mutant was half that of the wild-type leading to the reduced activity. The significant lack of growth was probably due to the instability of the BrTha during the catalysis conditions.

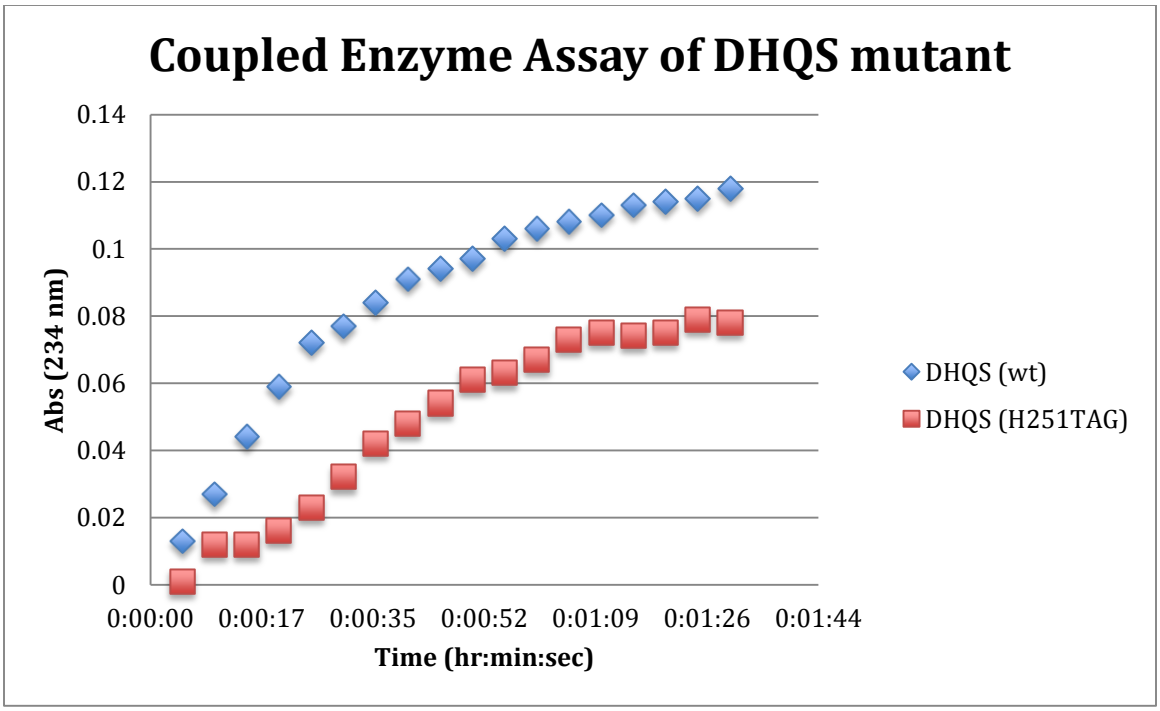


Figure 3-10: Absorbance of 3-dehydroshikimate from the couple enzyme assay of DHQS and 3-dehydroquinate dehydrate.

3.6 Conclusion

In summary all the experimental data shows that incorporation of unAA at the His251 site that there is greater potential in effecting the activity of the DHQS. Though the activity of the DHQS decreased after incorporation of the BrTha, this was expected. As the bromo-thienyl group acts as a weaker base it should be less likely to catalyze the deprotonation of the C5 hydroxyl group. Finding that the replacement of histidine to glutamine maintained the enzymatic activity was surprising. As the amide group acts as a poor Lewis base at neutral and acidic pH, this might say a bit about the relative pK_a of the active site. Or there are more residues that assist in the acid- base catalysis. Overall using histidine analogues

would allow a systematic probing of the active site and the mechanism of the *E. coli* DHQS. This highlights the importance of developing a system to genetically incorporate histidine analogues as a method to probe the catalytic mechanism, both for this enzyme as well as other enzymes that has histidine in the active site.

Chapter 4

Experimental Methods and Procedures

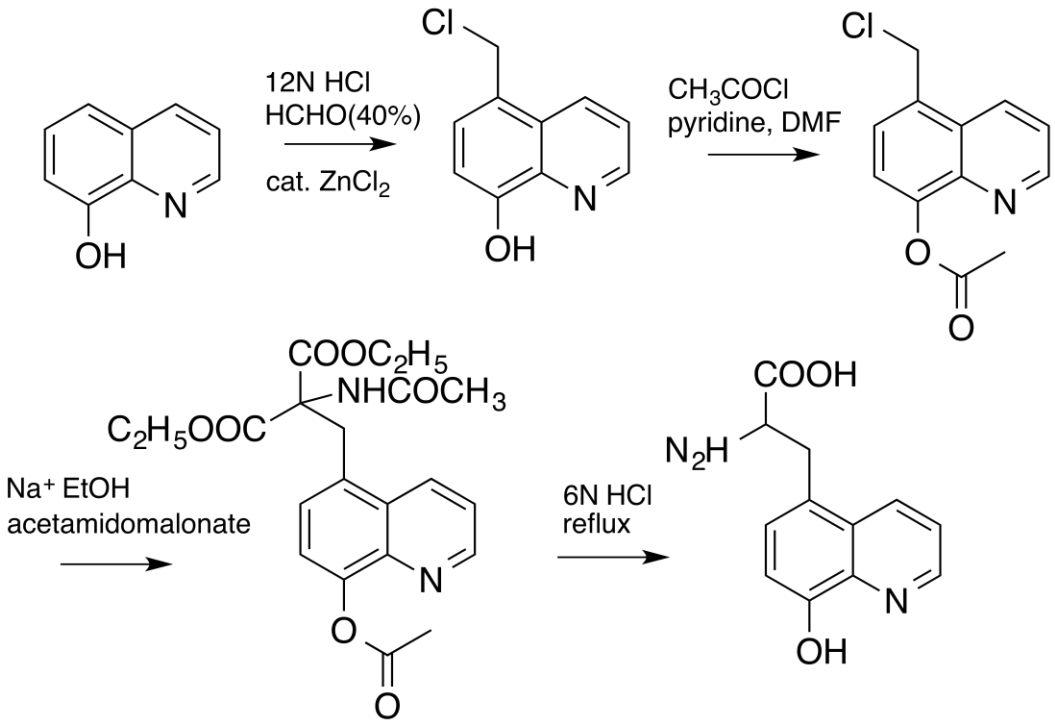
4.1 Introduction

To prepare the amino acid analogues, plasmids, and substrates used for the experiments described above they were synthesized/constructed in as described in this chapter. The two major synthesis projects were the HQAla synthesis mainly based of the Fridkin work. Another synthesis project was the synthesis of 2-(5-carbozylthienyl)alanine. The last synthesis project described was the attempts at synthesize a diazarine-theronine derivative, photo-threonine. The DAHP which was used in qualifying the enzyme activity of the $DHQS_{wt}$, $DHQS_{H251BrTha}$, and $DHQS_{H251Q}$; was isolated, as described, from the knockout strain AB2847. Lastly this section addresses the construction of the various plasmids used in this work.

^1H NMR spectra were recorded on a 300 MHz spectrometer. Chemical shifts for ^1H NMR spectra are reported (in parts per million) relative to internal tetramethylsilane (Me_4Si, δ = 0.0 ppm) with $CDCl_3$ as solvent and to sodium 3-(trimethylsilyl)propionate-2,2,3,3,-d_4 (TSP, δ = 0.0 ppm) when D_2O was the solvent.

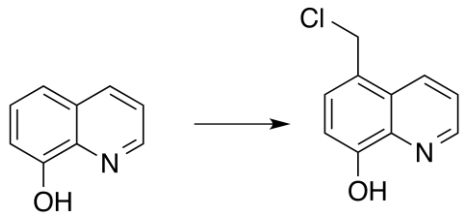
Silica gel was used for flash chromatography. Analytical thin-layer chromatography (TLC) utilized precoated plates on silica gel 60 F-254. TLC plates were visualized by UV or chemical stains. Distilled, deionized water was used to prepare all aqueous solutions. Organic solutions of products were dried over anhydrous agents, such as Na_2SO_4.

4.2 Hydroxyquinolin-alanine synthesis and experimental data



Scheme 4-1: Synthetic route for hydroxyquinolin-alanine (HQAla).

This section addresses the synthesis of the unAA hydroxyquinolin-alanine (HQAla) and synthetic optimizations made in the methodology.



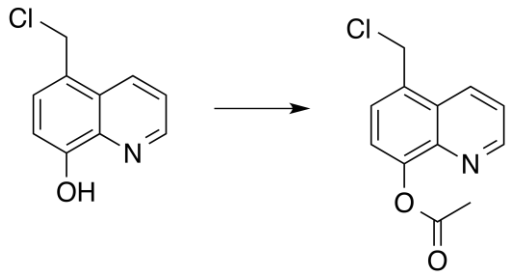
(**1.A**) (**2.A**)

Scheme 4-2: *Synthesis of 5-chloromethyl-8-quinolinol*[33] (**2.A**)

70 mmols 8-quinolinol (**1.A**) was added to 4.4 mmol of Zinc Chloride, and 7 mL 40% Formaldehyde and 50 mL of 12 N HCl. After allowing reaction to stir at room temperature for two days the reaction was filtered and washed with acetone until flow through ran clear. Two days were used instead of one day as yield was significantly higher increased to by 10%. Increasing the concentration of formaldehyde from 20% to 40% improved final yield to 90%

Yield 90%: ^{1}HNMR in D_2O: δ 9.18 (d, 1H), 8.90 (d, 1H), 7.98 (m, 1H), 7.64 (d, 1H), 7.30 (d, 1H), 4.96 (s, 2H).

Though the Fridken group described this transformation they used HCl gas, which is particularly dangerous to handle. This method was selected as the yield is still high but the methodology is much safer.

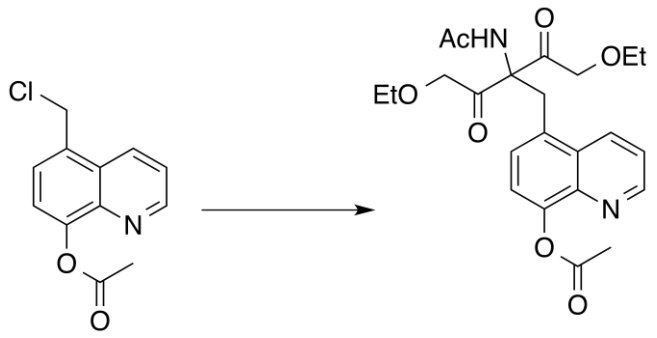


(**2.A**) (**3.A**)

Scheme 4-3: *Synthesis of 8-(5-chloromethyl)quinolyl acetate*[26] (**3.A**)

1 mmol of 5-chloromethyl-8-quinolinol (**2.A**) was placed under N_2 atmosphere and dissolved in 5 mL DMF. The flask was then cooled to 0° C on ice and 2.5 mmol of pyridine and 8mmol acetyl chloride was added simultaneously drop wise under nitrogen. If added too fast total yield of the reaction will decrease significantly. The reaction was stirred for 1 hour at 0° C, then 1 hour at room temperature. The reaction was then cooled to 0° C and 1.1 mmoles of cold water was added and allowed to stir for 20 minutes to quench the reaction. The product was extracted by a chloroform wash (3x20mL) and the organic layers were combined. The organic layer was dried using saturated $NaHCO_3$ (2x20mL) and brine (2x20mL). The solvent was removed by evaporation to yield 8-(5-chloromethyl)quinolyl acetate (**3.A**) about 80% of product though crude.

Yield 80% (crude): ^{1}HNMR in $CDCl_3$: δ 9.05 (d, 2H), 8.62 (d, 2H), 7.65 (d, 4H), 7.64 (d, 2H), 5.02 (s, 4H), 2.57 (s, 5H).

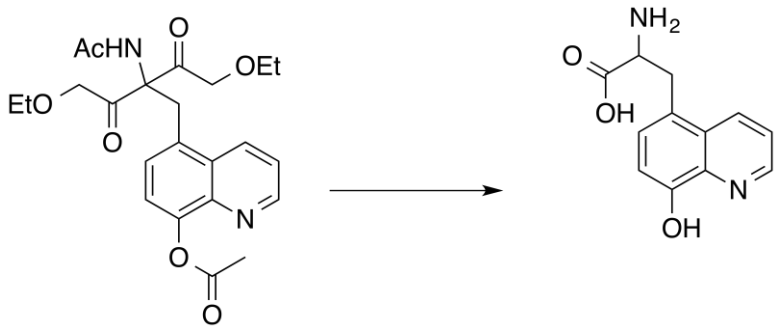


(**3.A**) (**4.A**)

Scheme 4-4: *Synthesis of diethyl (8-hydroxyquinolin-5-yl-methyl)-acetamindomalonate*[26] (**4.A**)

0.762 mmols of 8-(6-chloromethyl)quinolyl acetate (**3.A**) in minimal amounts of chloroform. 0.842 mmols of diethyl amidomalonate was dissolved in dry ethanol with an equal mole amount (.842 mmols) of metallic sodium. After all sodium was dissolved the ethanol solution was added to the 8-(6-chloromethyl)quinolyl acetate solution and allowed to reflux for 8 hours. After completed the ethanol was removed by evaporation 20 mL of water was added to re-solubilize. The product was extracted with 1:1 ethylactate:chloroform (2x20mL) and washed once with brine (10mL) and dried over Na_2SO_4. If more than one wash of brine is used then solution becomes miscible and product cannot be removed. The ethylacetate:chloroform solution was evaporated off to get crude diethyl (8-hydroxyquinolin-5-yl-methyl)-acetamindomalonate (**4.A**). Product cannot be purified.

Yield 59% (crude): ^{1}HNMR in $CDCl_3$: δ 8.77 (d, 1H), 8.33 (s, 1H), 7.43 (m, 1H), 7.13 (m, 1H), 6. 45 (s, 1H), 4.25 (m, 4H), 4.04 (s, 2H), 1.88 (s, 3H), 1.30 (t, 7H).



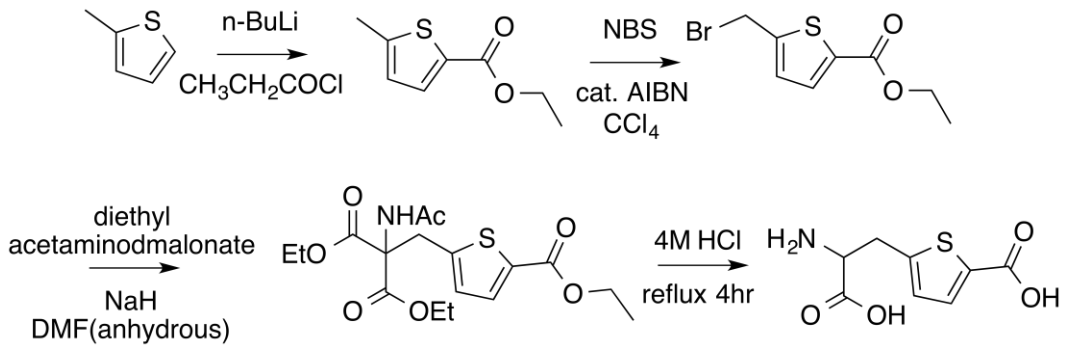
(**4.A**) (**5.A**)

Scheme 4-5: *Synthesis of 3-(8-hydroxyquinolin-5-yl)alanine*[26] (**5.A**)

21.9 mmol of the diethyl (8-hydroxyquinolin-5-yl-methyl) acetamindomalonate (**4.A**) with 150mL of 9 N of HCl and allowing it to reflux for 9 hours. In reference it was suggested 6N of HCl to be used for deprotection, but higher yield of the deprotection was seen when the concentration was raised to 9N. The HCl was removed by evaporation and then the solid was re-dissolved in water. The water then was adjusted to pH 5 - 5.5 with 10% sodium hydroxide. At this time a yellow precipitate is formed and it is then washed with water adjusted to pH 5.5 and then acetone. The precipitate is recrystallized in water pH 5.5 to give final product 3-(8-hydrozyquinolin-5-yl)alanine (**5.A**).

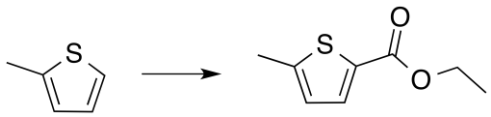
Yield 65%: ^{1}HNMR in D_2O: δ 8.74 (d, 1H), 8.48 (d, 1H), 5.57 (m, 1H), 7.33 (d, 1H), 7.06 (d, 1H), 3.92 (m, 1H), 3.68 (m, 1H), 3.27 (m, 1H).

4.3 2-(5-carboxylthienyl)alanine synthesis and experimental data



Scheme 4-6: Synthetic route for 2-(5-carboxylthienyl)alanine

The unAA 2-(5-carboxylthienyl)alanine is a new compound and has potential for incorporation into proteins. This new molecule was a hopeful for incorporation into the DHQS to the H251 or H264 position. Together with other histidine analogues it would be an interesting addition to examine how histidine and it's unique properties can affect enzyme activity. In addition, as both the thionyl group and carboxyl group can donate electrons, this electron rich histidine analogue has potential as a metal chelator. This compound can be obtained at high yields and high purity. Unfortunately, efforts to evolve an aminoacyl-tRNA synthase to incoperate the unAA failed. In the future, if we can get a crystal structure of BrThaRS, a new library will be designed and examined for the genetic incorporation of this unAA and other histidine analogues.

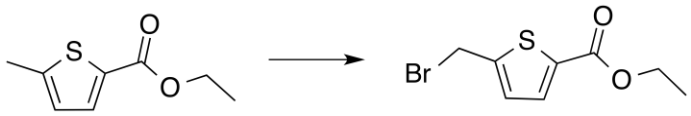


(1.B) **(2.B)**

Scheme 4-7: *Synthesis of 5-methylthiophene-2-carboxylate*[34](**2.B**)

0.925 mL of 2-methylthiophene (**1.B**) was added to dried flask and flushed with Ar. 12 mL of dry THF was added to flask then cooled to -10 C. 8 mL of n-butyllithium (1.6 M in hexanes) was added slowly and stirred for 5 hours at -10 C. After, 1.44 mL ethyl chloroformate was added to 10 mL of dry THF and cooled to -10 C. Ethyl chloroformate mixture was added drop wise to 2-methylthiophene solution. Removed from dry ice bath and allowed to stir at room temperature overnight.

The reaction was poured over ~ 50 mL of ice and stirred at room temperature until ice melted. Water extracted; then 25 mL of 1M HCl used to wash reaction. Washed reaction with 25 mL of saturated sodium carbonate solution. Washed reaction with 25 mL of water then dried with sodium sulfate. Concentrated product under vacuum. Purified by silica flash column (1:9 Ethyl Acetate: Hexanes). Yield 71%: ^{1}H-NMR in $CDCl_3$: δ 1.358 (t, 3H), 2.519 (s, 3H), 4.320 (q, 2H), 6.758 (d, 1H), 7.605 (d, 1H).

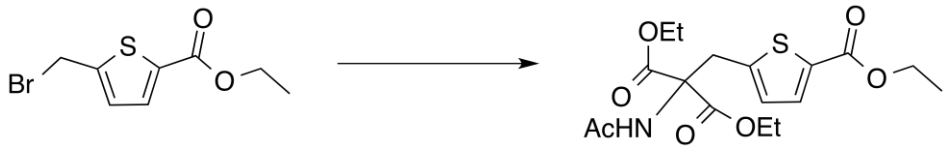


(2.B) **(3.B)**

Scheme 4-8: *Synthesis of Ethyl 5-bromomethylthiophene-2-carboxylate*[35] (**3.B**)

0.638g of 5-methylthiophene-2-carboxylate (**2.B**) and 1.022g of NBS added in dried flask. 10 mL of carbon tetrachloride was added and a catalytic amount of AIBN added. Refluxed 2 hours at 94°C, then another catalytic amount of AIBN added and refluxed for another 2 hours. Removed from heat and filtered. Purified by flash column with 1:9 (Ethyl Acetate: Hexane).

Yield 82%: ^{1}H-NMR in $CDCl_3$: δ 1.368 (t, 3 H), 4.344 (q, 2H), 4.673 (s, 2H), 7.086 (d, 1H), 7.630 (d, 1H).



(3.B) **(4.B)**

Scheme 4-9: *Synthesis of 2-acetamido-2-((5-(ethoxycarbonyl)thiophene-2-yl)methyl)malonate*[36] (**4.B**)

0.861 g (3.96 mmol) of diethyl acetamidomalonate suspended in 8 mL of dry DMF and cooled to 0°C; 0.190 g (2 eq.) of NaH added and flask flushed with Argon. Warmed to room temperature and stirred for 20 minutes. 0.891g (3.57 mmol) of ethyl 5-bromomethylthiophene-2-carboxylate (**3.B**) was added and stirred overnight at room temperature. Reaction quenched with 10 mL of dH_2O and

extracted with ether (16mL x 3) washed with brine and dried over $MgSO_4$ and concentrated. Next step proceeded without purification.

^{1}H-NMR δ 1.368 (3 H, t, *J* = 7.2 Hz, CH_2CH_3), 4.344 (2 H, q, *J* = 7.2 Hz, CH_2), 4.673 (2 H, s, 5 – $BrCH_2$), 7.086 (1H, d, *J* = 4 Hz, 4-H), 7.630 (1 H, d, *J* = 4 Hz, 3-H)

(4.B) **(5.B)**

Scheme 4-10: *Synthesis of 2-(5-carboxylthienyl)alanine* (**5.B**)

1.334g of impure diethyl 2-acetamido-2-((5-(ethoxycarbonyl)thiophen-2-yl)methyl)malonate (**4.B**) was dissolved in 100 mL of 4M HCl and refluxed for 4 hours. Removed from heat and stirred overnight at room temperature.

Concentrated and re-suspended in dH_2O and extracted with ethyl acetate then concentrated.

Yield 70% ^{1}H-NMR in D_2O: δ 3.53 (dd, 2H), 4.186 (t, 1 H), 7.09 (d, 1H), 7.74 (d, 1 H). ^{13}C-NMR in D_2O: δ 30.5, 53.48, 129.54, 133.18 135.10, 143.85, 165.34, 170.5

4.4 3*O*-diazirin-L-threonine synthesis and experimental data

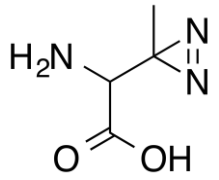
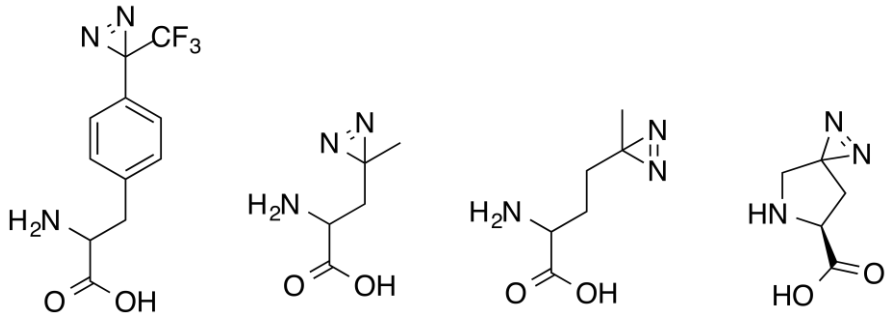


Figure 4-1: Structure of 3*O*-diazirine-L-threonine (photo-threonine)

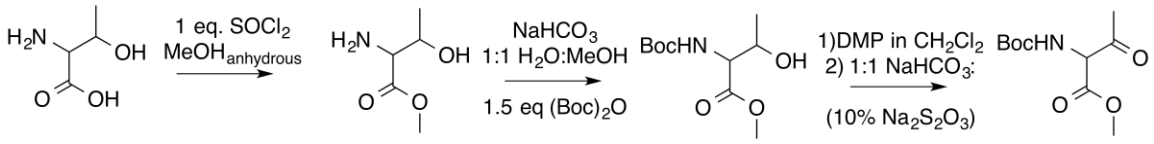
Attempts were made to synthesize the threonine analogue, 3-diazirin-L-threonine (Figure 4-11). The diazirine functionality of this unAA would have the ability to cross-link through a light induced carbene formation. Previous unAA have been designed with the diazarine functionality including 4'-[3-(trifluormethyl)-3*H*-diazirine-3-yl]-L-phenylalanine (A)[37], L-photoleucine[38] (B), L-photomethionine (C)[38, 39] L-photoproline (D)[40] (Figure 4-12). The proposed structure would be the smallest unAA with this diazirine functionality, and would be ideal for incorporation into an active site to replace catalytic cysteine or serine. This was a collaborative project with Dr. Mark Wilson from Biochemistry. Various routes were attempted to synthesize this unAA however none were successful. The inability to synthesize this threonine analogue is likely due to the acidic alpha hydrogen, the reactivity of this hydrogen likely destabilizes the intermediates the diazirine formation.



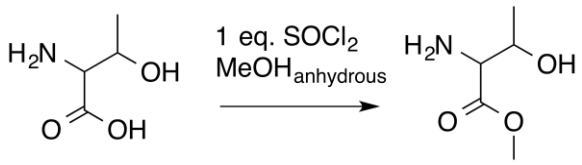
(A) (B) (C) (D)

Figure 4-2: Structures of unAA with diazirine functionality: (A) 4'-[3-(trifluormethyl)-3*H*-diazirine-3-yl]-L-phenylalanine, (B) L-photoleucine, (C) L-photomethione, (D) L-photoproline.

The first synthetic route to this unAA was to create a keto to undergo the diazirine formation in a similar method described in by the Watanabe group shown in the synthesis route shown in Scheme 4-11.[41] This β-keto-threonine derivative was obtained at high yields and used in various attempts to form the diazirine product.



Scheme 4-11: Synthesis route to the β-keto-threonine derivative, Boc-protected β-keto-threonine methyl ester.

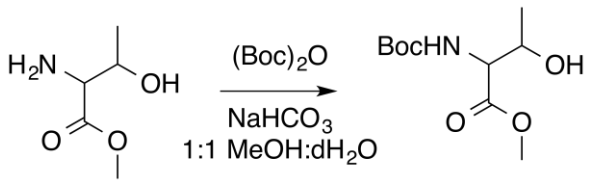


(1.C) **(2.C)**

Scheme 4-12: *Synthesis of L-threonine methyl ester* (**2.C**)

16.8 mmol of L-Threonine (**1.C**) was weighed out and set aside. 25 mL of anhydrous Methanol was cooled to 0°C and 30 mL of thionyl chloride was added to the methanol over 30 minutes then allowed to stir at 0°C for an additional 15 minutes. The weighed out L-threonine was added slowly to the reaction and allowed to stir for 30 minutes at 0°C. The reaction was removed from the ice bath and allowed to stir for 24 hours at room temperature and reaction was checked via TLC (9:1 CH_3Cl:MeOH) for formation of (**2.C**). Solvent removed under reduced pressure.

Yield (Crude) 80% ^{1}H NMR in $CHCl_3$: δ 1.23 (d, 3H), 3.77 (s, 3H), 3.84 (d, 1H), 4.18 (m, 1H)



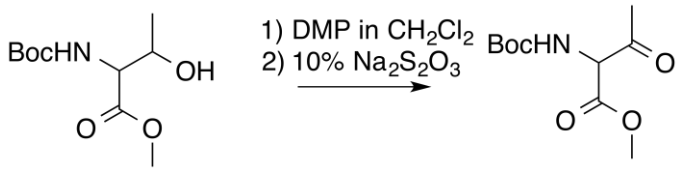
(2.C) **(3.C)**

Scheme 4-13: *Synthesis of Boc-protected-threonine methyl ester* [41](**3.C**)

3.0 g of threonine methyl ester (**2.C**) and 4.607g of $NaHCO_3$ were dissolved in 36 mL of 1:1 MeOH:dH_2O. 5.805g of di-tert-butyl dicarbonate was added slowly over 15 minutes. The reaction was allowed to stir for 24 hours at room temperature. The

solvent was removed under vacuum and solution acidified to pH 4.5 with 1M citric acid. Product (**3.C**) was extracted with 4 mL of ethyl acetate 30 times and solvent was removed in vacuum.

Yield (Crude) 84% ^{1}H NMR in $CDCl_3$: δ 1.22 (d, 3H), 1.43 (s, 9H), 3.74 (s, 3H), 4.26 (d, 1H), 4.33 (d, 1H), 5.26 (bd, 1H)

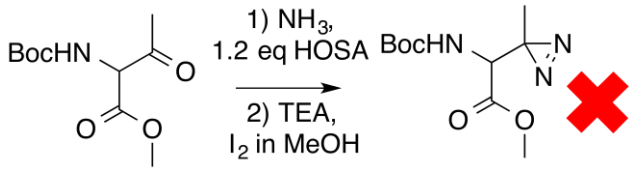


(3.C) **(4.C)**

Scheme 4-14: *Synthesis of Boc-protected β-keto-threonine methyl ester*[41] (**4.C**)

1.7 mmol of Boc-protected-threonine methyl ester (**3.C)** was dissolved in 4 mL of methylene chloride. 0.5 mM of Dess Martin periodinane (DMP) was added and reaction allowed to stir at room temperature for 45 minutes. 16 mL of 1:1 $NaHCO_3$: 10% $Na_2S_2O_3$ added to quench reaction and allowed to stir for 10 minutes at room temperature. Product was extracted with 30 mL of ethyl acetate three times. Solvent removed under vacuum. Boc-protected β-keto-threonine methyl ester (**4.C**) was purified by 9:1 hexane: ethyl acetate column.

Overall yield 33.5% ^{1}HNMR in $CDCl_3$: δ 1.46 (s, 9H), 2.38 (s, 3H), 3.83 (s, 3H), 5.02 (d, 1H), 5.28 (bd 1H).

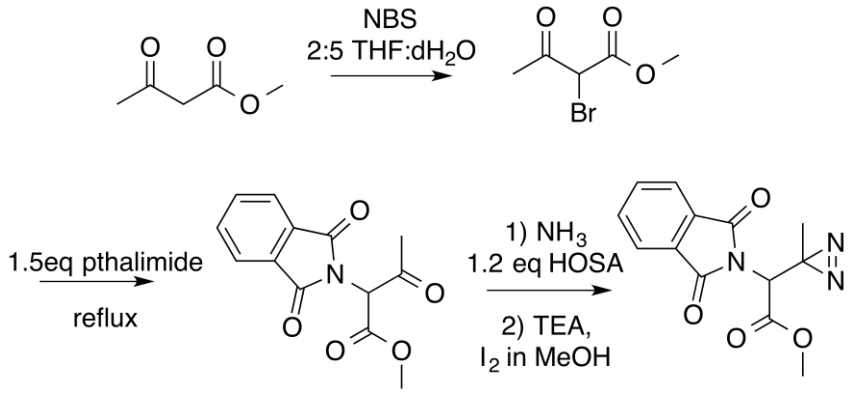


Scheme 4-15: *Synthesis of Boc-protected β-diazirine-threonine methyl ester*.

To transform the keto group to the diazarine was done following the procedure outlined in the synthesis of L-photomethione by the Yao group.[39] 400mg Boc-protected β-keto-threonine methyl ester (**4.C**) was dissolved in 5 mL of liquid ammonia and stirred 7 hours at -78°C. The reaction was cooled to -78°C and 1.2 equivalents of hydroxylamine-O-sulfonic acid in anhydrous methanol was added drop wise over 30 minutes. The reaction mixture was stirred overnight at room temperature allowing the ammonia to evaporate off. The reaction was filtered and washed with methanol. The combined methanol was concentrated to half the volume and 10 mL of fresh methanol was added then reaction cooled to 0°C. 1.2 equivalents of TEA was added slowly. I_2 was added slowly until solution remained light brown. The reaction was stirred at room temperature for 1 hour and concentrated. The product was extracted with ethyl acetate. The organic layer was washed with brine and dried. NMR showed no product formed.

To attempt to get product the procedure was altered slightly in reference to other papers describing the synthesis of diazirines. Varying the concentration of the ammonium was altered from 100% by volume to 50% by volume showed no formation of products. Varying the reaction going for 5 hours, 7 hours, 9 hours, or 12 hours showed no formation of products. Varying the temperature of the reaction from -78°C, 0°C, or 20°C showed no formation of products. The major problem that

we predict prevents the formation of the diazirine product is probably the favored enol formation by the β-keto group. This enol probably is what is preventing the diazirine formation.



Scheme 4-16: Proposed pathway to photo-threoine via methyl acetoacetate.

As it is likely that the β-keto group is forming an enol intermediate, the amine protecting group was changed to a pthalimide to change the electronic properties of this acidic proton. It is key that any protecting groups that are put onto the molecule could be removed in favorable conditions to the diazarine, as it is sensitive to heat and reductive conditions. Though the pthalimide was synthesized, the enol was the major product. However after further reading it was found that the pthalimide protecting group would favorably form an enol at 100%.[42] This pthalimide was synthesized, and though the enol was the major product, it was used to attempt to transform the keto group into a diazarine. As expected no product was formed, this however showed that the enol is not able to undergo the necessary chemistry for this transformation.

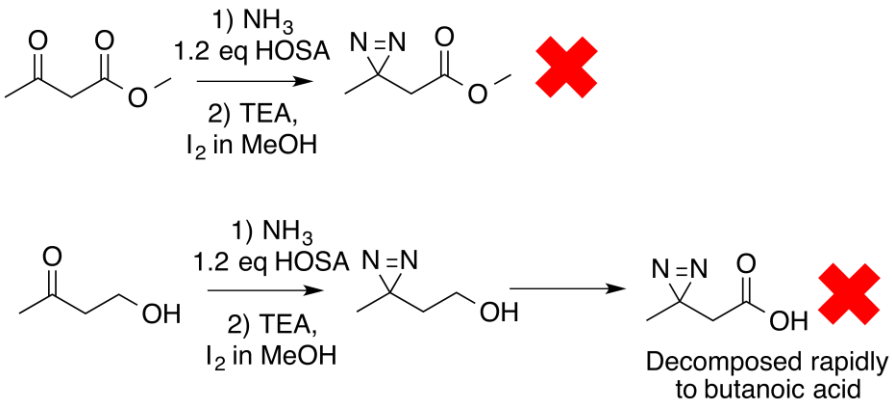


Figure 4-3: *Diazirine-first synthetic approach.*

As the alpha hydrogen is so acidic on we attempted to circumvent this route by adding the diazirine early in the synthesis. Based off of the work by the Behrman group.[43] Initial experiments to add the diazirine starting from the methyl acetoacetate did not work. Again this is probably due to the acidy of the alpha hydrogen. To reduced the acidity of near the keto group we started with 4-hydrozybutan-2-one, and preceded to try and build the molecule from the diazirine molecule. Though the first step was successful, and the diazirine was formed. The following intermediates decomposed rapidly supporting the theory that the alpha hydrogen destabilizes the diazirine functional group. Though it is probably possible to synthesize the photo-threonine if kept in unreactive conditions, it is likely too unstable to be used in biological conditions.

4.5 Biochemical Synthesis of DAHP

To test the relative activity of the enzymes evolved in section 3.5, the assay needed DAHP. The DAHP was biosynthesized by the AB2847 strain in a similar manner outlined by the Frost group. Transforming a plasmid with the aroF gene on a Cm resistant vector, this strain was inoculated into an M9 minimal media culture for overnight incubation at 37°C. Though the Frost group saw excretion of the DAHP from the AB2847 strain without the addition of the *aroF* gene, this addition allows for the suppression of the *aroB* gene of the Shikimate pathway even in high presence of aromatic amino acids and vitamins. The suppression should then produce higher amounts of the DAHP.

This overnight culture was then inoculated into 200 mL of minimal media prepared as following:

K_2HPO_4 (1.4 g), KH_2PO_4 (0.6 g), and $(NH_2)_2SO_4$ (0.2 g) were dissolved in 200 mL of water and autoclaved.[44] 5 mL of 20% glucose was then added with 200 µL of $MgSO_4$ (120mg/mL). Following, L-tyrosine (4mg/mL), L-phenylalanine (4mg/mL), L-tryptophan solutions (4mg/ml) and L-thiamine (1mg/mL) were made and purified by filtration through a membrane. Then 400 µL of L-Tyrosine solution, 400 µL of L-phenylalanine solution, 200 µL of L-tryptophan, and 200 µL of L-thiamine solution were added to the sterile media culture.[44] Finally, this sterile solution was then inoculated with 5 mL of the overnight $AB2947_{aroF}$ culture and incubate at 37°C for 32 hours.[44]

Following the Frost paper[44]: after incubation the culture was spun down and the supernatant was isolated and run through a 100 mL Dowex 50 H^+ column at 4°C.

The pH of the eluent was adjusted to pH 8.0 with 1M LiOH, lithium hydroxide was used as the lithium counter ion which allows for easier solubility in alcohols. This adjusted solution was concentrated and kept at below 30°C. The residue was then dissolved in 200 mL of methanol and cooled to 4°C. This solution was stirred for 1 hour at 4°C until white precipitate formed. The precipitate was removed via filtration and the eluent was concentrated under vacuum. This was then re-suspended in 200 mL of ddH_2O and pH was adjusted to 7.1 with 1M LiOH. The DAHP and DAH in the solution was purified via a DEAE-cellulose column equilibrated with 100mM of triethylammonium bicarbonate (TEAB) pH 7.1. These components were eluted by a gradient of 100mM to 300mM TEAB solution (800 mL – 800 mL).

Each of the column fractions were assayed by a modified Aminoff TBA assay Method. 100 µL of the fraction was incubated with 200 µL of $NaIO_4$ (25mM in 0.125M H_2SO_4) for 10 minutes at room temperature.[44] This step oxidized the monosaccharide. Excess oxidizing reagent was reduced by addition of 400 µL of $NaAsO_4$ (2% wt/vol in 0.5 M HCl), and the reaction went to completion after the solution moved from clear to brown/dark yellow before back to clear.[44] After the reaction was completed 2 mL of thiobarbituric acid (TBA) (0.36% wt/vol, pH 9) was added and the reaction was heated at 100°C for 10 minutes.**Error! Bookmark not defined.**

DAH and DAHP containing fractions were combined and concentrated. The buffer was removed via azeotropic distillation of 2-propanol.[44] The DAHP and DAH fractions were run down a Dowex 50 H^+ column (20 mL) at 4°C. The eluent was

concentrated to dryness and identified by ^{31}P NMR ^{1}H NMR.[44] The DAHP was stored at -20°C and used in the activity assay.

4.6 Conclusion on Synthesis

The synthesis of HqAla, the novel unAA, 2-(5-carboxylthienyl)alanine, and DAHP were straightforward and obtained at good yields. All products were relatively stable and were used in the experiments described in Chapters 2 and 3. The attempt at the synthesis of the diazirine threonine derivative however was not achieved. It is likely that this small molecule is difficult to synthesize or the final product is too unstable for use in a biological system. Though attempts were made to synthesize this unAA, there were no successes.

4.7 Genetic Manipulation

(1) Plasmid Construction

Plasmid pBK-aroB$_{wt}$ was obtained by first amplifying the *aroB* gene from the pQZ.116 plasmid by using the following primers forward: 5'-GTCGAATTCATGGAGAGGATTGTCGTTACTC-3' and reverse primer: 5'-TGTCATAAGCTTTCAGTGGTGGTGGTGGTGGTGCGCTGATTGACAATCGGCAATG-3'.
The PCR product was digested with *EcoRI* and *HindIII* as was installed by the amplification. This was ligated into the pBK vector that was treated with the same restriction enzymes, resulting in the plasmid pBK-aroB$_{wt}$. The resulting plasmid contains a kanamycin resistance marker, and the *aroB* gene was inserted behind the constitutive *glnS* promoter (P$_{glnS}$).

Plasmid pBK-aroB$_{E184A}$ was obtained by modifying the Glu184 codon of the *aroB* gene to CGC of pBK-aroB$_{wt}$ by site-directed mutagenesis using the following primers, forward: 5'-CGTCGGGGCTGGCAGCGGTCATCAAATACGGCA-3' and reverse: 3'-TGCCGTATTTGATGACCGCTGCCAGCCCCGACG-3'.

Plasmid pBK-aroB$_{H264A}$ was obtained by modifying the His 264 codon of the *aroB* gene to CGC of pBK-aroB$_{wt}$ by site-direct mutagenesis using the following primers, forward: 5'-GCTGAAATGGGGTATGGCAATTGGTTAGCGGGTGAAGCGGTCGC-3' and reverse: 5'-GCGACCGCTTCACCCGCTAACCAATTGCCATACCCCATTTCAGC-3'.

Plasmid pBK-aroB$_{E184A/H264A}$ by modifying the His264 codon of the *aroB* gene to CGC of pBK-aroB$_{E184A}$ by site-direct mutagenesis using the following primers, forward: 5'-GCTGAAATGGGGTATGGCAATTGGTTAGCGGGTGAAGCGGTCGC-3' and reverse: 5'-GCGACCGCTTCACCCGCTAACCAATTGCCATACCCCATTTCAGC-3'.

Plasmid pJF-aroB$_{wt}$ was obtained by digesting the pBK-aroB$_{wt}$ with *Hind*III and *EcoR*I. The digested *aroB* gene was ligated with a pJF vector digested with the same restriction enzymes. The resulting plasmid contains the ampicillin resistance marker and the wild type *aroB* gene under the control of a hybrid tac promoter (*Ptac*).

Plasmid pJF-aroB$_{E184TAG/H264A}$ was obtained by amplifying *aroB* gene of pBK-aroB$_{E184A/H264A}$ via overlapping PCR with the following primers: 5'-GTCGAATTCATGGAGAGGATTGTCGTTACTC-3', 5'-TAGCGTCGGGGCTGGCATAGG TCATCAAATACGG-3', 5'-CCGTATTTGATGACCTATGCCAGCCCCGACGCTA-3', and 5'-TGTCATAAGCTTTCAGTGGTGGTGGTGGTGGTGCGCTGATTGACAATCGGCAATG-3'. This changed the Ala 184 codon to TAG and this PCR product was digested with

*Hind*III and *EcoR*I. The digested *aroB* gene was ligated with a pJF vector digested with the same restriction enzymes. The resulting plasmid contains the ampicillin resistance marker and the variant $aroB_{E184TAG/H264A}$ gene under the control of a hybrid tac promoter (*Ptac*).

Plasmid pJF-$aroB_{E184A/H264TAG}$ was obtained by amplifying *aroB* gene of pBK-$aroB_{E184A/H264A}$ via overlapping PCR with the following primers: 5'-GTCGAATTCATGGAGAGGATTGTCGTTACTC-3', 5'-TGAAATGGGGTATGGCAATTGG TTATAGGGTGAAGCGGTC-3', 5'-GACCGCTTCACCCTATAACCAATTGCCATACCC CATTTCA-3', and 5'-TGTCATAAGCTTTCAGTGGTGGTGGTGGTGGTGCGCTGATT GACAATCGGCAATG-3'. This changed the Ala264 codon to TAG and this PCR product was digested with *Hind*III and *EcoR*I. The digested *aroB* gene was ligated with a pJF vector digested with the same restriction enzymes. The resulting plasmid contains the ampicillin resistance marker and the variant $aroB_{E184A/H264TAG}$ gene under the control of a hybrid tac promoter (P_{tac}).

Plasmid pJF-$aroB_{E184Y}$ was obtained by amplifying *aroB* gene of pJF-$aroB_{wt}$ via overlapping PCR with the following primers: 5'-GTCGAATTCATGGAGAGGATTGTCGTTACTC-3', 5'-GCGTCGGGGCTGGCATATGTCAT CAAATACGGCA-3', 5'-TGCCGTATTTGATGACATATGCCAGCCCCGACGC-3', and 5'-TGTCATAAGCTTTCAGTGGTGGTGGTGGTGGTGCGCTGATTGACAATCGGCAATG-3'. This changed the Glu 184 codon to TAG and this PCR product was digested with *Hind*III and *EcoR*I. The digested *aroB* gene was ligated with a pJF vector digested with the same restriction enzymes. The resulting plasmid contains the ampicillin

resistance marker and the variant *aroB*$_{E184Y}$ gene under the control of a hybrid tac promoter (P_{tac}).

Plasmid pJF-aroB$_{H264Y}$ was obtained by amplifying *aroB* gene of pJF-aroB$_{wt}$ via overlapping PCR with the following primers: 5'-GTCGAATTCATGGAGAGGATTGTCGTTACTC-3', 5'-ATGGGGTATGGCAATTGG TTATATGGTGAAGCGG-3', 5'-CCGCTTCACCATATAACCAATTGCCATACCCCAT-3', and 5'-TGTCATAAGCTTTCAGTGGTGGTGGTGGTGGTGCGCTGATTGACAATCGGCAATG-3'. This changed the His 264 codon to TAG and this PCR product was digested with *Hind*III and *EcoR*I. The digested *aroB* gene was ligated with a pJF vector digested with the same restriction enzymes. The resulting plasmid contains the ampicillin resistance marker and the variant *aroB*$_{H264Y}$ gene under the control of a hybrid tac promoter (P_{tac}).

Plasmid pJF-aroB$_{H251TAG}$ was obtained by amplifying *aroB* gene of pJF-aroB$_{wt}$ via overlapping PCR with the following primers: 5'-GTCGAATTCATGGAGAGGATTGTCGTTACTC-3', 5'-GAATCTGGGACACACACCTT TGGTTAGGCCATTGAAGCTGAAAT-3', 5'-ATTTCAGCTTCAATGGCCTAACCAAAGG TGTGTCCCAGATTC-3', and 5'-TGTCATAAGCTTTCAGTGGTGGTGGTGGTGGTGC GCTGATTGACAATCGGCAATG-3'. This changed the His 251 codon to TAG and this PCR product was digested with *Hind*III and *EcoR*I. The digested *aroB* gene was ligated with a pJF vector digested with the same restriction enzymes. The resulting plasmid contains the ampicillin resistance marker and the variant *aroB*$_{H251TAG}$ gene under the control of a hybrid tac promoter (P_{tac}).

Plasmid pJF-aroB$_{H251TAG}$-pylT was obtained by amplifying the pyltRNA sequence from pBK-mmPylT with the following primers: 5'-ACCACCACTGAAAGCTTCCATGGTTCCACAGGGTAGCCAGCAG-3', and 5'-TCCGCCAAAACAGCCAAGCTCTCGAGCAGAACATATCCATC-3'. This PCR product was used with the pJF-aroB$_{H251TAG}$ plasmid digested with *Hind*III for SLIC cloning. The pylT sequence was incorporated behind the *aroB$_{H251TAG}$* gene, and changed the 3' *Hind*III site to a *Nco*I site.

Plasmid pJF-aroB$_{H264TAG}$-pylT was obtained by amplifying *aroB* gene of pJF-aroB$_{E184A/H264TAG}$ via overlapping PCR with the following primers: 5'-GTCGAATTCATGGAGAGGATTGTCGTTACTC-3', and 5'-GGAACCATGGAA GCTTTCAGTGG-3'. This PCR product was digested with *Sac*II and *Nco*I. The digested *aroB* fragment was ligated with a pJF--aroB$_{H251TAG}$-pylT vector digested with the same restriction enzymes. The resulting plasmid contains the ampicillin resistance marker and the variant *aroB$_{E184A/H264TAG}$* gene under the control of a hybrid tac promoter (P_{tac}) with a pylT sequence behind the *aroB* gene.

Plasmid pJF-aroB$_{H251CAG}$-pylT was obtained by amplifying *aroB* gene of pJF-aroB$_{E184A/H264TAG}$ via overlapping PCR with the following primers: 5'-GTCGAATTCATGGAGAGGATTGTCGTTACTC-3', 5'-GAATCTGGGACACACACCTT TGGTCAGGCCATTGAAGCTGAAAT-3', 5'-ATTTCAGCTTCAATGGCCTGACCAAAGG TGTGTCCCAGATTC-3' and 5'-GGAACCATGGAA GCTTTCAGTGG-3'. This changed the His 251 codon to Gln (CAG) and this PCR product was digested with *Sac*II and *Nco*I. The digested *aroB* fragment was ligated with a pJF--aroB$_{H251CAG}$-pylT vector digested with the same restriction enzymes. The resulting plasmid contains the ampicillin

resistance marker and the variant *aroB*$_{E184A/H264TAG}$ gene under the control of a hybrid tac promoter (P_{tac}) with a pylT sequence behind the *aroB* gene.

Plasmid pET-BrThaRS$_{\text{Full Length}}$ was obtained by amplifying out the BrThaRS from the pBK-BrThaRS plasmid using the following primers: 5'-TTCATATGGATAAAAAACCATTAG-3', and 5'-AAGGATCCTTATAGATTGGTT GAAATCCC-3'. This PCR product was digested with *BamH*I and *Nde*I and ligated into the pET28a vector digested with the same enzymes. The resulting plasmid contains the kanamycin resistance marker and the BrThaRS$_{\text{Full Length}}$ sequence under the control of a T_7 promoter.

Plasmid pET-BrThaRS$_{\text{Truncated}}$ was obtained by amplifying out the BrThaRS from the pBK-BrThaRS plasmid using the following primers: 5'-CCCATATGGCTCCTGCTCCTTCACTTAC -3', and 5'-AAGGATCCTTATAGATTGGTT GAAATCCC-3'. This PCR product was digested with *BamH*I and *Nde*I and ligated into the pET28a vector digested with the same enzymes. The resulting plasmid contains the kanamycin resistance marker and the BrThaRS$_{\text{Full Length}}$ sequence under the control of a T_7 promoter.

Plasmid pBK-BrThaRS$_{\text{Random mutagenesis}}$ library was obtained by amplifying out the BrThaRS mutant from pBK-BrThaRS using the following primers and the enzyme mutazyme DNA polymerase that makes small errors during amplification: 5'-GCGGCTTTGTTGAATAAATCGAAC-3', and 5'-CAGACATCATGTAGGCCTG-3'. This PCR product was digested with *Nde*I and *Pst*I and ligated back into a new pBK vector digested with the same enzymes. The resulting plasmid contains a kanamycin

resistance marker, and the randomly mutated BrThaRS sequence was inserted behind the constitutive *glnS* promoter (P_{glnS}).

(2) Strain manipulation

Strains used particular to this work were AB2847 and JWΔaroB. Both strains were received from the *E. coli* Genetic Stock Center at Yale University. The strains specifics are as follows:

AB2847 from *E. coli* Genetic Stock Center

Chromosomal Makers: *tsx-354, glnV42*(AS), *λ-, aroB351, malT354(λ^R)*

JW3352-1 from *E. coli* Genetic Stock Center

Chromosomal Markers: *Δ(araD-araB)567, ΔlacZ4787*(::rrnB-3), *λ-, ΔaroB724::kan, rph-1, Δ(rhaD-rhaB)568, hsdR514*

For use of the JW strain in the work described in this thesis the Kanamycin marker (Kan^r), that was placeholder to aroB724 gene, was removed by the following method:

- Competent cells of JW3351-1 were made by growing inoculated 5mL culture overnight at 30°C until OD 0.4. Centrifuged down for 10 minutes at 4°C and 5000*g.* Re-suspended with cooled CCMB solution and centrifuged again for 10 minutes at 4°C and 5000*g.* Cells re-suspended in 200 µl of CCMB and chilled overnight at -80°C.

- Transformed JW3352 with pCP20 and plated on Amp^{100}, incubated overnight at 30°C. 8 colonies were selected and plated onto pi plate (LB Amp^{100}) and grown at 30°C overnight.
- 1 colony selected from each of the pi and re-streaked on pi plate (LB) and grown overnight at 42°C.
- 2 colonies from each pi taken for replicate on (LB), (LB Amp^{100}), (LB Kan^{50}) and grown overnight at 37°C.
- 4a and 4b on initial 16 colonies did not grow on (LB Amp^{100}) so these colonies taken and grown and prepared for a competent cell line for the experiment

The genotype of the *aroB* gene was confirmed using the following primers for AB2847:

Forward Primer: 5'-GTATGGAGAGATTGCCGACGTGAC-3'

Reverse primer: 5'-GGTTTCAGCTCGTCTTCTGGTTTG-3'

(3) Preparation of cultures for growth tests

5 mL culture of LB (Amp^{100}/Cm^{34}) inoculated with pJF-aroB mutant in JWΔaroB strain and incubated overnight at 37C. 500ul of cells were taken and pelleted down. Re-suspended in 500 µL of M9 minimal media without aromatic amino acids and vitamins. Pelleted down and re-suspended with 500 µL of M9 minimal media without aromatic amino acids and vitamins. 100 µl of washed cells used to inoculate following cultures:

(1) M9 minimal media with aromatic amino acids and vitamins+ 0.25 mM IPTG

(2) M9 minimal media without aromatic amino acids and vitamins

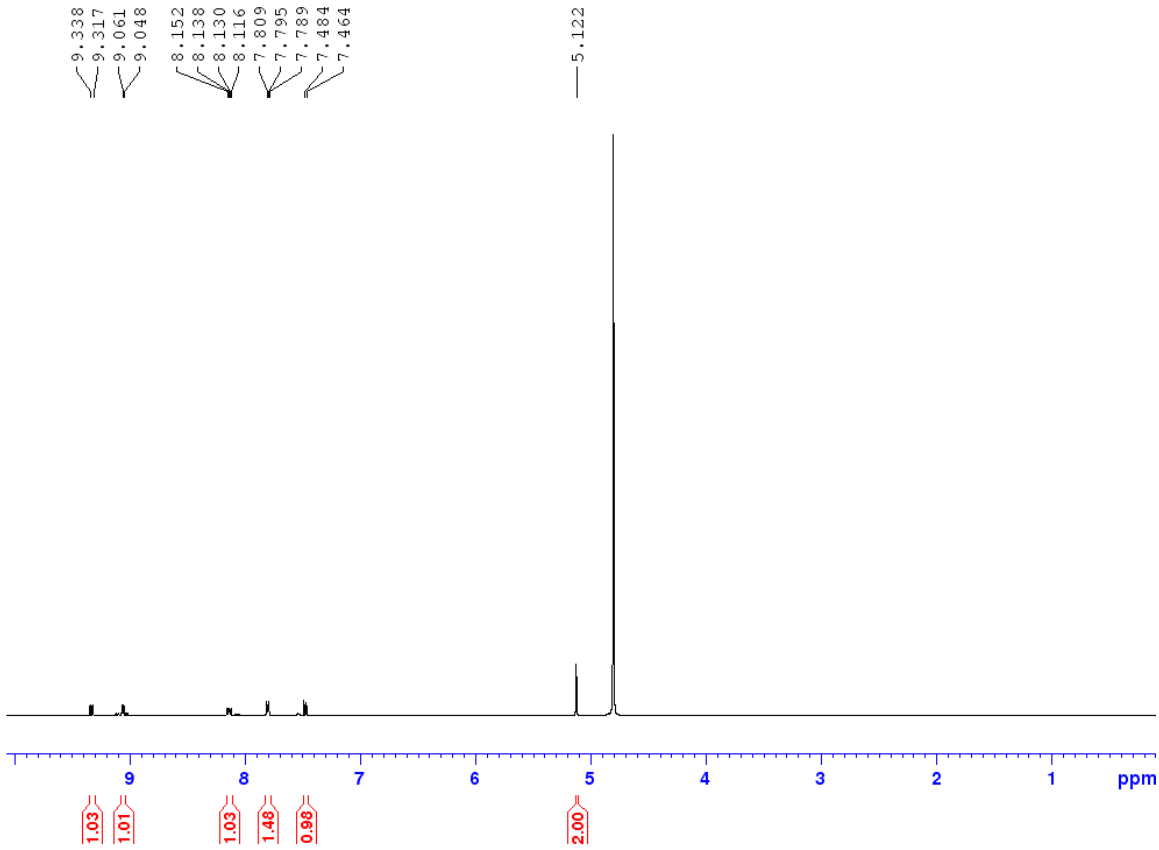
(3) M9 minimal media without aromatic amino acids and vitamins + 0.25 mM IPTG

(4) M9 minimal media without aromatic amino acids and vitamins + 0.25 mM IPTG + 2mM unAA

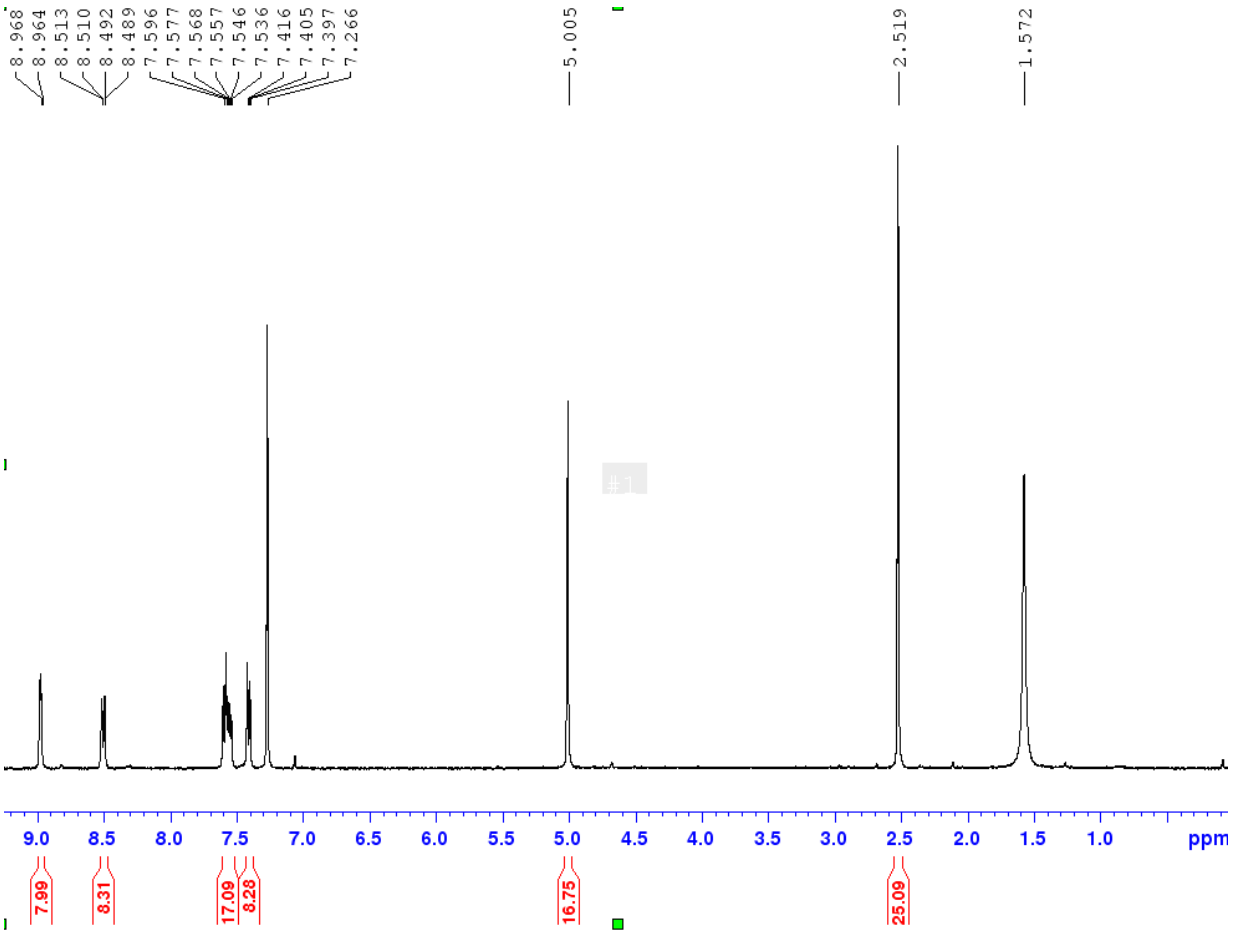
(5) M9 minimal media without aromatic amino acids and vitamins + 0.25 mM IPTG + 2mM unAA + 1x Trace elements

30 µL of each culture was taken for OD reading. Repeated at each time point – in case of cultures with unAA, these were pelleted down and re-suspended in M9 minimal media without aromatic amino acids and vitamins.

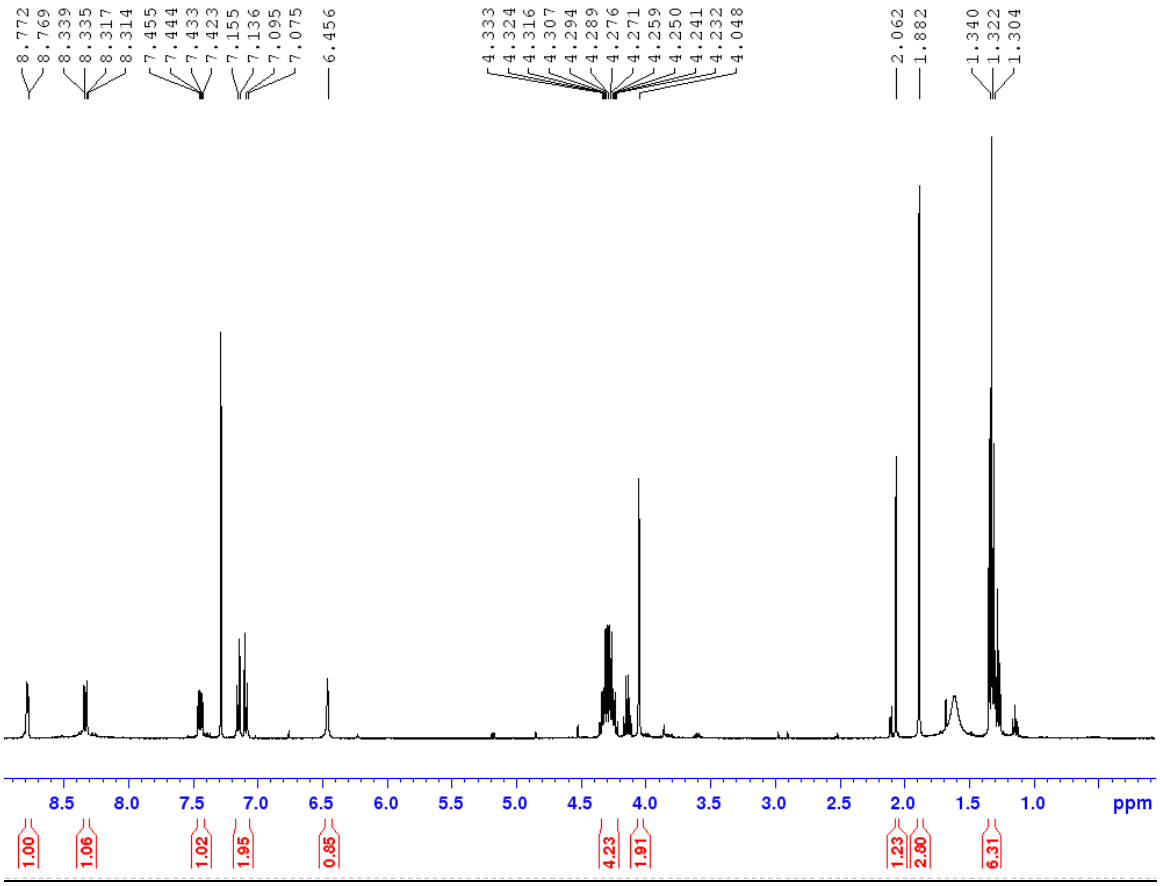
APPENDIX A: List of NMR Spectra



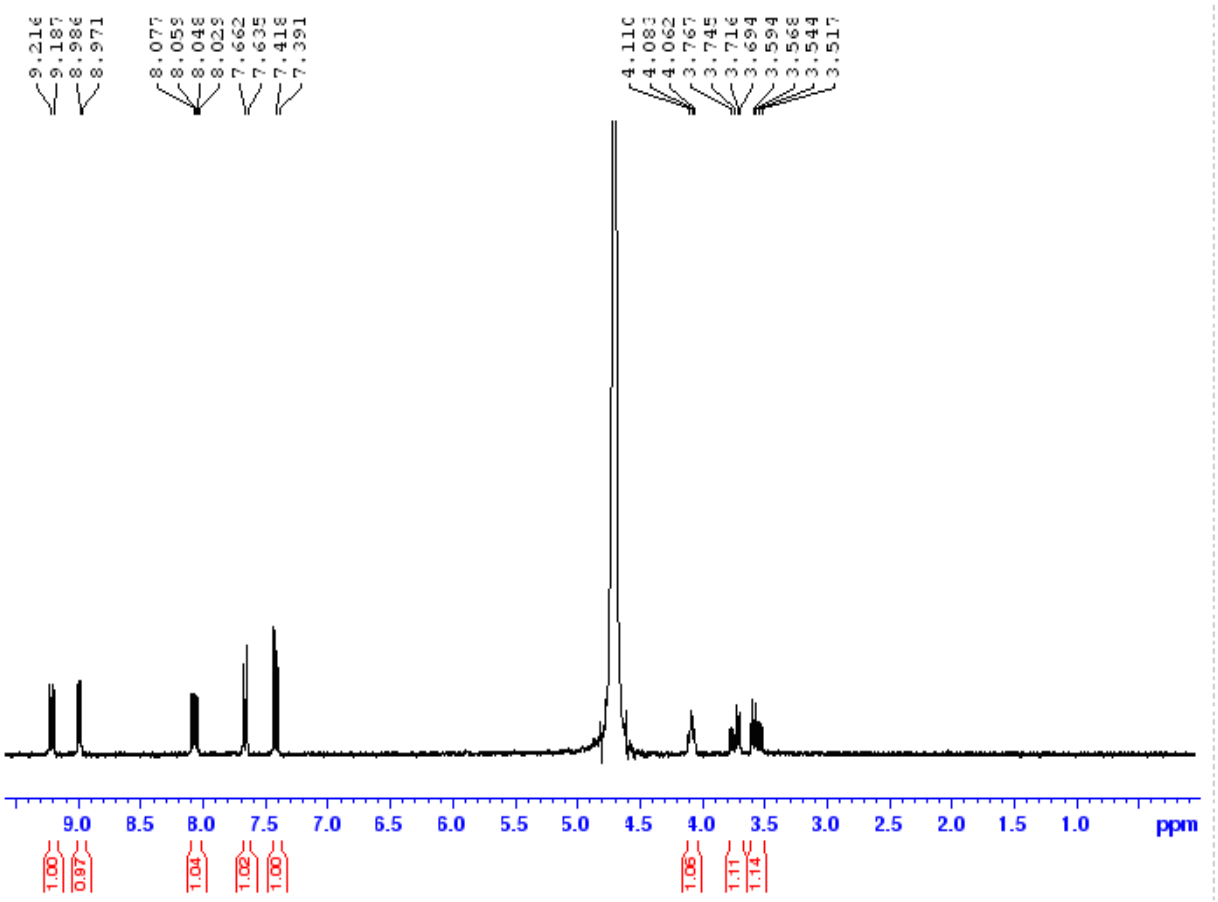
5-chloromethyl-8-quinolinol (2.A) ^{1}H-NMR



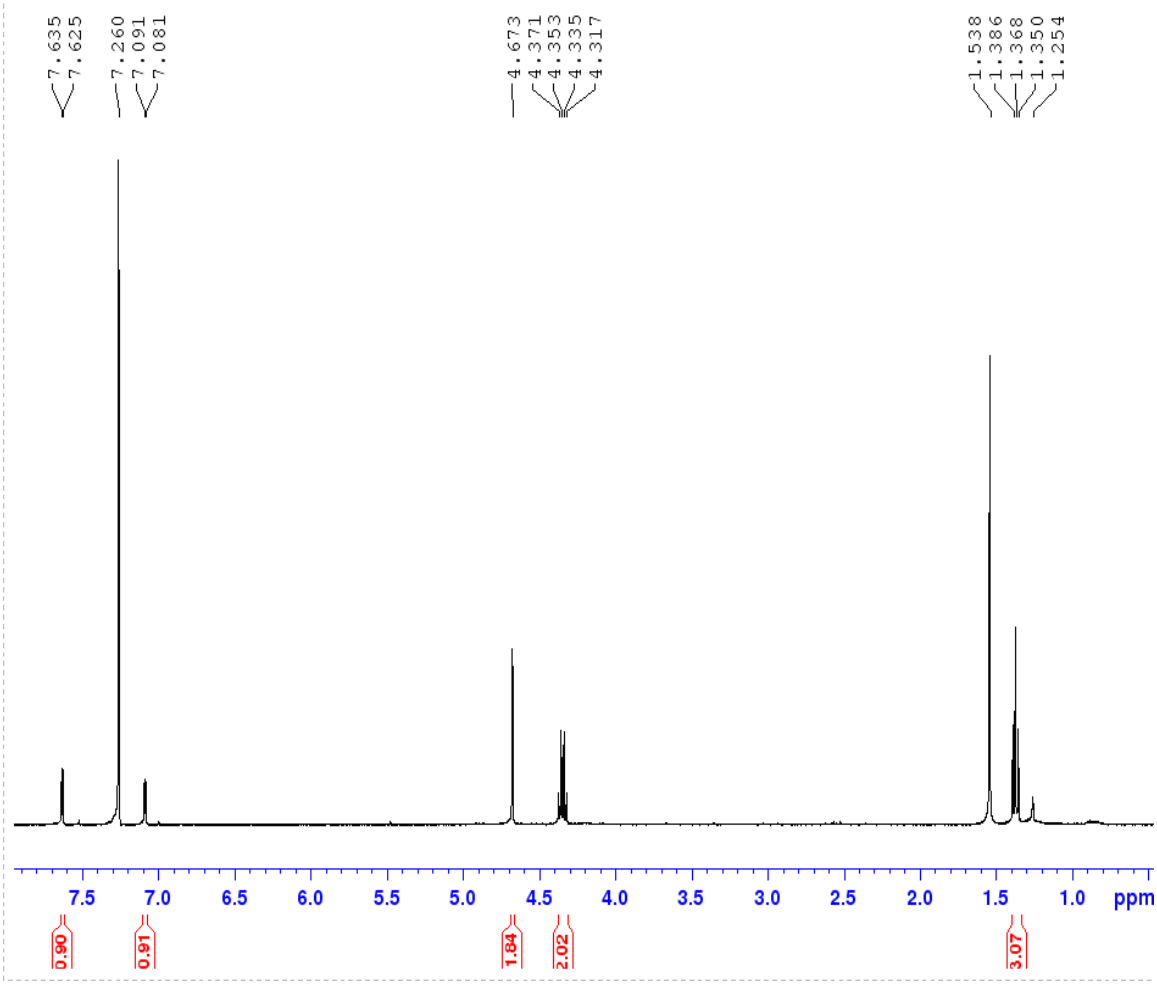
8-(5-chloromethyl)quinolyl acetate (3.A) ^{1}H-NMR



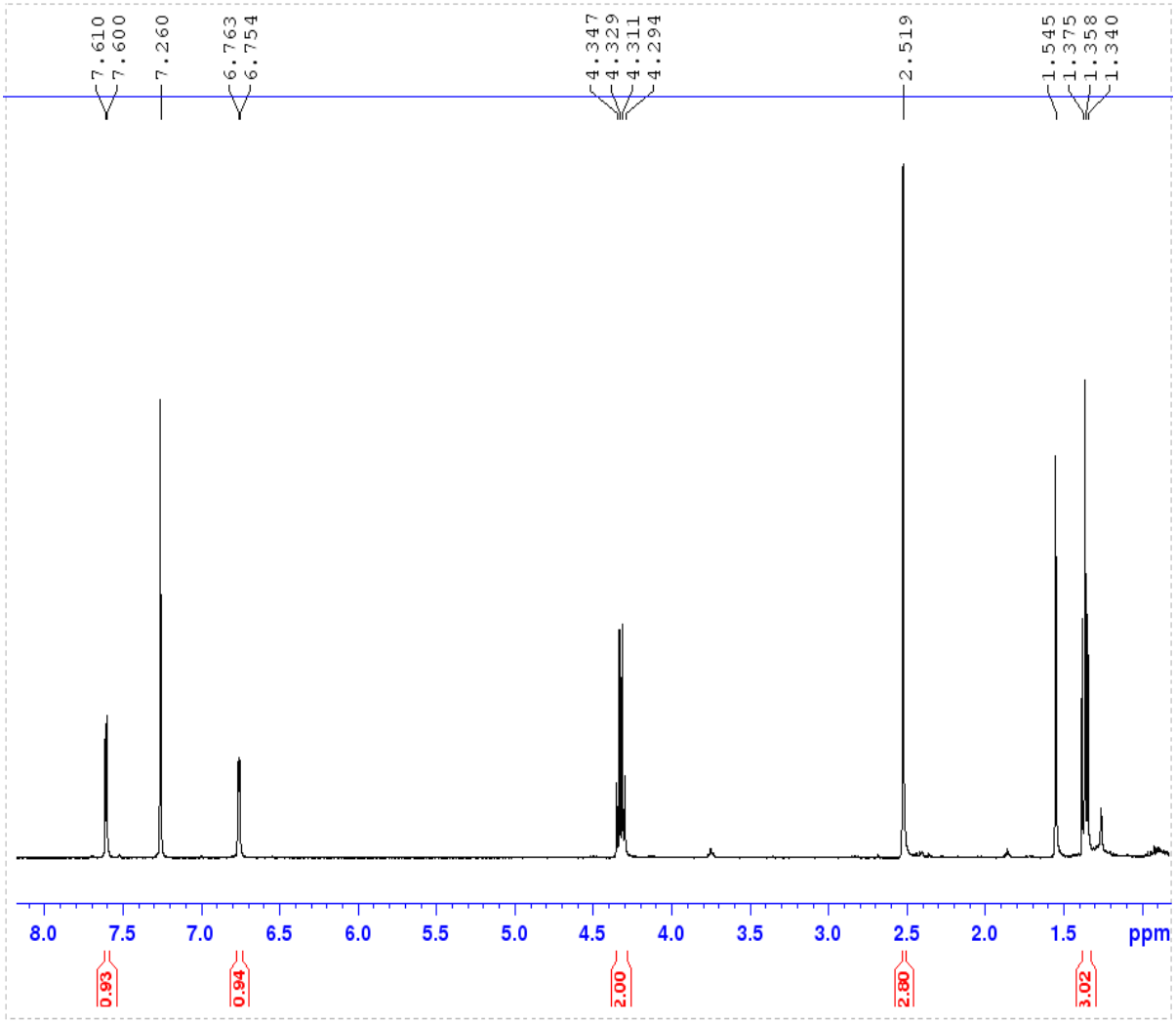
diethyl (8-hydroxyquinolin-5-yl-methyl)-acetamindomalonate (4.A) ^{1}H-NMR



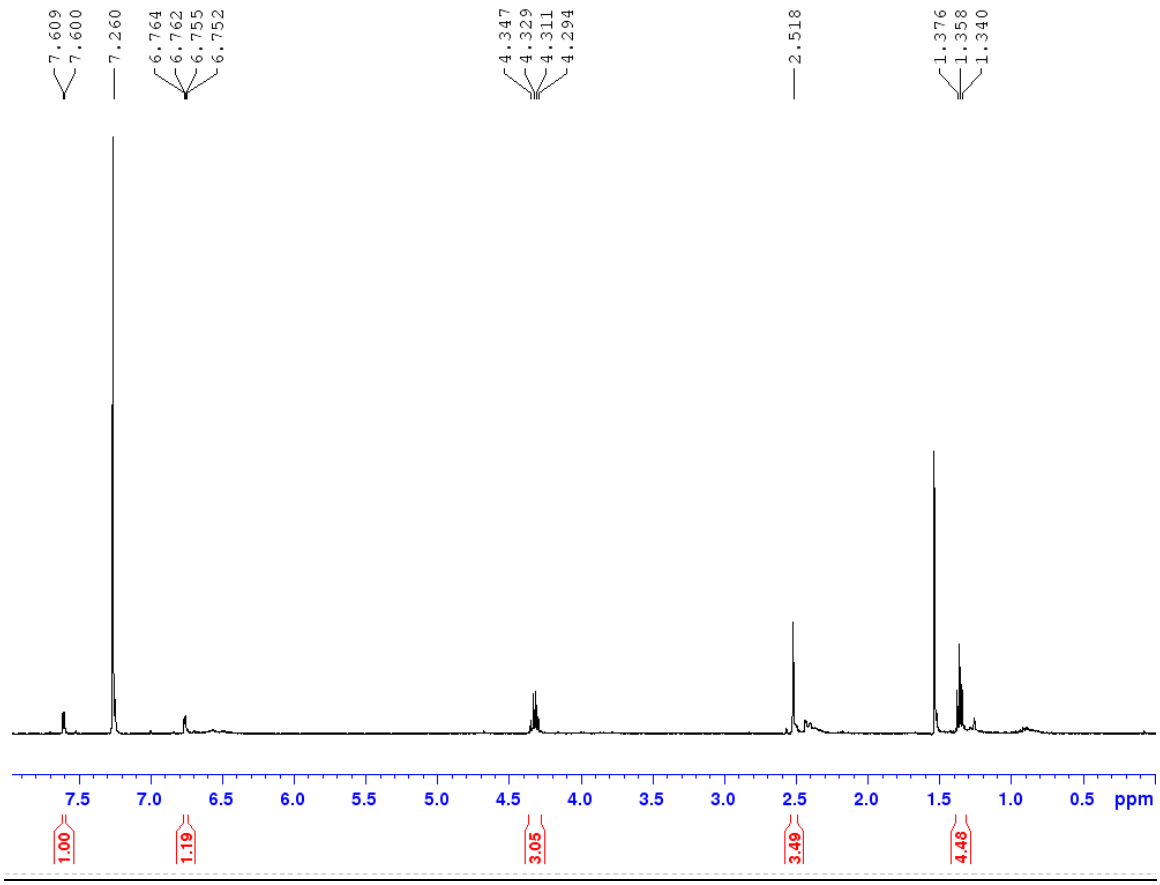
3-(8-hydroxyquinolin-5-yl)alanine (5.A)^{1}H-NMR



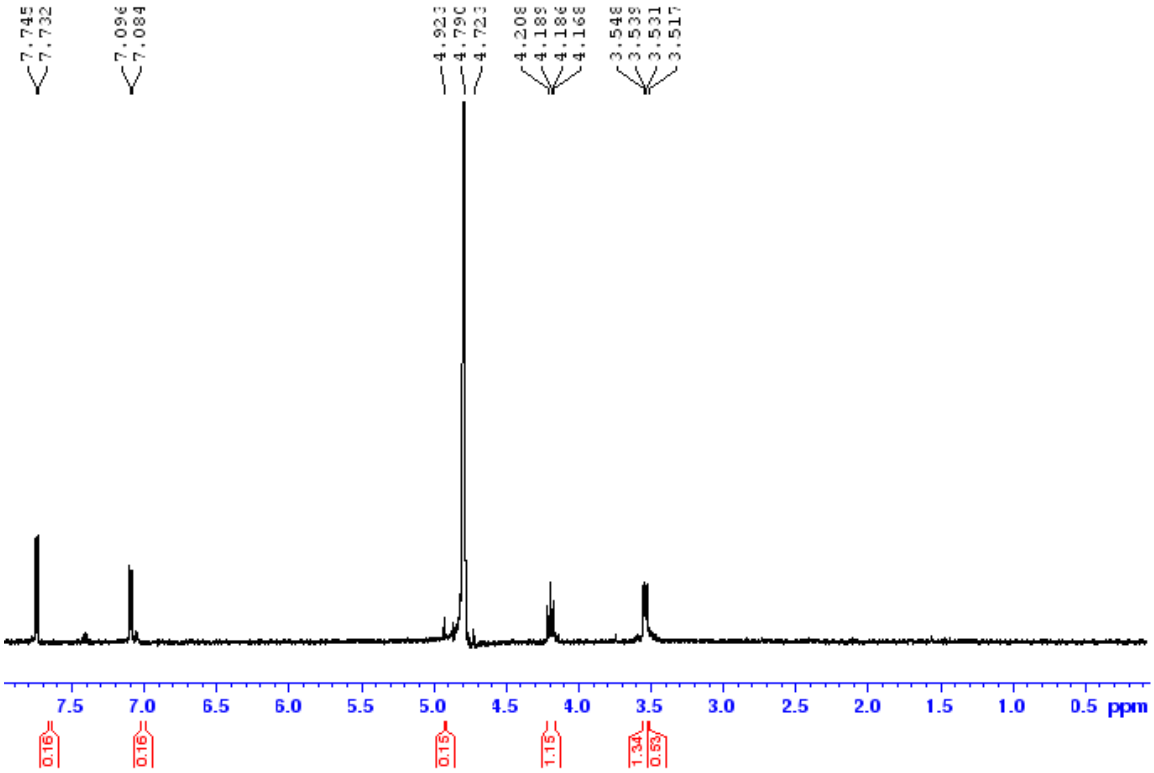
5-methylthiophene-2-carboxylate (2.B) ^{1}H-NMR



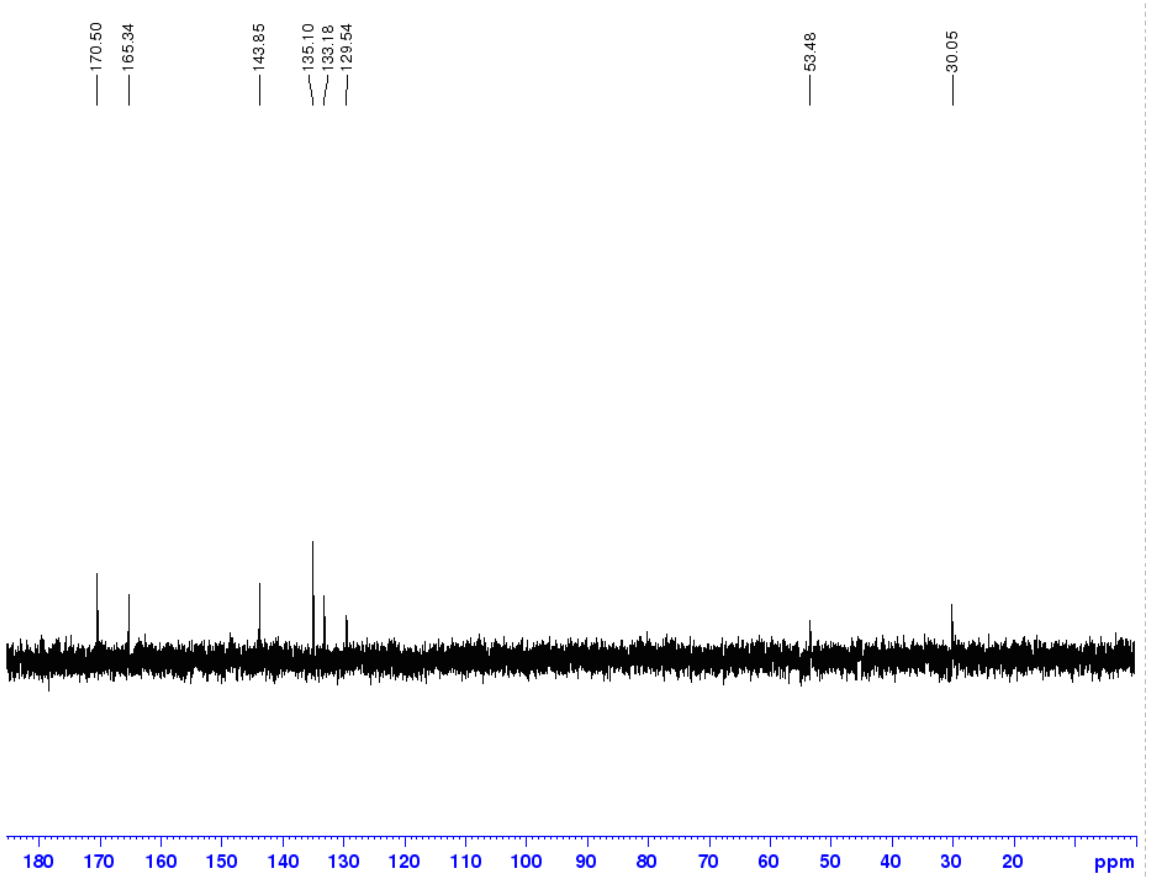
Ethyl 5-bromomethylthiophene-2-carboxylate (3.B) ^{1}H-NMR



2-acetamido-2-((5-(ethoxycarbonyl)thiophene-2-yl)methyl)malonate (4.B)^{1}H-NMR



2-(5-carboxylthienyl)alanine (5.B) ^{1}H-NMR



2-(5-carboxylthienyl)alanine (5.B) ^{13}C-NMR

References:

[1] Chapeville, F.; Lipman, F.; Ehrenstein, G. V.; Weisblum, B.; Ray Jr., W. J.; Benzer, S. "Role of soluble ribonucleic acid in coding for amino acids." *Proc. Natl. Acad. Sci*. **1962**, 48(6), 1086-1092.

[2] Von Ehrenstein, G. Y.; Weisblum, B.; Benzer, S. "The function of soluble ribonucleic acid (sRNA) as amino acid adaptor in the synthesis of hemoglobin." *Proc. Natl. Acad. Sci.* **1963**, 49, 669-675.

[3] Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S. "Chemical aminoacylation of tRNA's." **1978**, *253*, 4517-4520.

[4] Crowley, K. S.; Liao, S.; Worrell, V.; Reinhart, G.; Johnson, A. E. "Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore." **1994**, *78*, 461-471.

[5] Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C., Schultz, P.G. "A General Method for Site-Specific Incorporation of Unnatural Amino Acids into Proteins." *Science*, **1989**, 244:182-188.

[6] Nowak, M.W.; Kearney, P.C.; Sampson, J. R; Saks, M. E. Labarca, C. G.; Silverman, S. K.; Zhong, W.; Thorson, J. Abelson, J. N.; Davidson, N.; Schultz, P. G.; Dougherty, D. A.; Lester, H. A. "Nicotinic Receptor Binding Site Probed with Unnatural Amino-Acid Incorporation in Intact Cells." *Science*. **1995**, *268*, 439-442.

[7] Wang, L.; Brock, A.; Herberich, B.; Schultz, P. "Expanding the Genetic Code of *Escherichia coli.*" *Science,* **2001**, *292*, 498-500.

[8] Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M.C.; Schultz, P.G.; "A General method for site-specific incorporation of unnatural amino acids." *Science*. **1989**, *244*, 182-188.

[9] Liu, C.; Schultz, P. G. "Adding New Chemistries to the Genetic Code." *Annu. Rev. Biochem,* **2010**, *79*, 413-444.

[10] Drienovska, I.; Rioz-Martinez, A.; Draksharapu, A.; Roelfes, G. "Novel artificial metalloenzymes by *in vivo* incorporation of metal-binding unnatural amino acids." *Chem. Sci.* **2015**, *6*, 770-776.

[11] Hohskaka, T.; Sato, K.; Sisido, M.; Takai, K.; and Yokoyama, S. "Site –specific incorporation of photofunctional non-natural amino aids into a polypeptide through in vitro protein biosynthesis. *FEBS Lett*. **1994**, *344*, 171-174.

[12] Zeng, Y.; Wang, W.; Liu, W. R. "Toward reassigning the rare AGG codon in *Escherichia coli.*" *ChemBioChem*. **2014**, *15*, 1750-1754.

[13] Takimoto, J. K.; Dellas, N., Noel; J. P., Wang, L. "Stereochemical basis for engineered pyrrolysyl-tRNA synthetase and the efficient in vivo incorporation of structurally divergent non-native amino acids." *ACS Chem. Biol.* **2011**, *134*, 2950-2953.

[14] Wang, Y. S., Fang, X., Wallace, A. L., Wu, B., Liu, W. R. "A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum." *J. Am. Chem. Soc.,* **2012**, *134*, 2950-2953.

[15] Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. "Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome." *Nature.* **2010**, *464*, 441-444.

[16] Bentley, R.; Haslam, E. "The Shikimate Pathway – A Metabolic Tree with Many Branches." *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 307-384.

[17] Herrmann, K. M. "The Shikimate Pathway as an Entry to Aromatic Secondary Metabolism." *Plant Physiol.* **1995**, *107*, 7-12.

[18] Carpenter, E. P.; Hawkins, A. R.; Frost, J. W.; Brown, K. A. "Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis." *Nature.* **1998**, *394*, 299-302.

[19] Bender, S. L.; Mehdi, S.; Knowles, J. R. "Dehyroquinate synthase: the role of divalent metal cations and of nicotinamide adenine dinucleotide on catalysis." *Biochemistry*. **1989**, *28*, 7555-7560.

[20] Bartlett, P.A.; Satake, K. "Does dehyroquinate synthase synthesize dehydroquinate?" *J. Am. Chem. Soc.,***1988**, *110*, 1628

[21] Xiao, H.; Peters, F. B.; Yang, P.; Reed, S.; Chittuluru, J. R.; Schultz, P. G. "Genetic Incorporation of Histidine Derivatives Using an Engineered Pyrrolysyl-tRNA Synthetase." *ACS. Chem. Biol.* **2014**, *9*, 1092-1096.

[22] Guo. L.; Wang, S.; Nakamura, A.; Eiler, D.; Kavran, J.M.; Wong, M.; Kiessling, L.; Steitz, T. A.; O'Donoghue, P.; Soll, D. "Polyspecific pyrrolysyl-tRNA synthetase from directed evolution." *PNAS*. **2014**, *111*, 16724-16729.

[23] Pittard, J.; Wallace, B. J. "Distribution and function of genes concerned with aromatic biosynthesis in *Escherichia coli*." *J. Bacteriol.* **1966**, *91(4)*, 1494-1508.

[24] Brown, K. A.; Carpenter, E.P.; Watson, K.A.; Coggins, J.R.; Hawkins, A.R.; Koch, M. H. J.; Svergun, D.I. "Twists and turns: a tale of two shikimate-pathway enzymes." *Biochem. Soc. Trans.* **2003**, *31,* 543-547.

[25] Liu, X.; Li, J.; Hu, C.; Zhou, Q.; Zhang, W.; Hu, M.; Zhou, J.; Wang, J. "Significant Expansion of the Fluorescent Protein Chromophore through the Genetic Incoporation of a Metal-Chelating Unnatural Amino Acid." *Angew. Chem. Int. Ed.* **2013**, *52*, 1-6.

[26] Zheng, H.; Youdim, M. B. ,H.; Weiner, L. M.; Fridkin, M. "Novel potential neuroprotective agents with both iron chelating and amino acid-based derivatives targeting central nervous system neurons." *Biochem. Pharma.* **2005**, *70*, 1642-1652.

[27] Wang, N.; Ju, T.; Niu, W.; Guo, J. "Fine-tuning Interaction between Aminoacyl-tRNA synthetase and tRNA for Efficient Synthesis of Proteins Containing Unnatural Amino Acids." *ACS Synth. Biol.* **2015**, *4*, 207-212.

[28] Ikeda, Y.; Kawahara, S.; Taki, M.; Kuno, A.; Hasegawa, T.; Taira, K. "Synthesis of a novel histidine analogue and its efficient incorporation into a protein *in vivo*." *Protein Eng.*, **2003**, *16*, 699-706.

[29] Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Kobayashi, T.; Sakamoto, K.; Yokoyama, S. "Multistep engineering of pyrrolylsyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenyloxycarbonyl) lysine for site-specific protein modification." *Chem. Biol.* **2008**, *15*, 1187-1197.

[30] Young, D.; Young, T.; Jahnz, M.; Ahmad, I.; Spraggon, G.; Schultz, P. "An Evolved Aminoacyl-tRNA Synthetase with Atypical Polysubstrate Specificity." *Biochemistry.* **2011**, *11*, 1894-1899.

[31] Kavran JM, et al. (2007) Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. Proc Natl Acad Sci USA 104(27):11268–11273.

[32] Montchamp, J.; Frost, J. W. "Cyclohexenyl and Cyclohexylidene Inhibitors of 3-Dehydroquinate Synthase: Active Site Interactions Relevant to Enzyme Mechanism and Inhibitor Design." *J. Am. Chem. Soc.* **1997**, *119*, 7645-7653.

[33] Yang, T.; Gao, G.; Zhang, X.; Li, F. "Facile Fabrication of Composition-Tuned Ru-Ni Bimetallics in Ordered Mesoporous Carbon For Levulinic Acid Hydrogenation." *ACS Catal.* **2014**, *4*, 1419-1425.

[34] Suaifan, G.; Goodyer, C.; Threadgill, M. "Synthesis of *N*-(Methoxycarbonylthienylmethyl)thioureas and Evaluation of Their Interaction with Inducible and Neuronal Nitirc Oxide Synthase." *Molecules* **2010,** *15*, 3121-3134.

[35] Choong, I.; Lew, W.; Lee, D.; Pham, P.; Burdett, M.; Lam, J.; Wiesmann, C.; Luong, T.; Fahr, B.; DeLano, W.; McDowell, S.; Allen, D.; Erlanson, D.; Gordon, E.; O'Brien, T. "Identification of Potent and Selective Small-Molecule Inhibitors of Caspase-3 through the Use of Extended Tethering and Structure-Based Drug Design." *J. Med. Chem.* **2002**, *45*, 5005-5022.

[36] Ashkenazi, T.; Pinkert, D.; Nudelman, A.; Widberg, A.; Wexler, B.; Wittenbach, V.; Flint, D.; Nudelman, A. "Aryl chain analogues of the biotin vitamins as potential herbicides. Part 3." *Pes Manag Sci.* **2007**, *63*, 974-1001.

[37] Tippmann, E. M.; Liu, W.; Summerer, D.; Mack, A.V.; Schultz, P.G. "A Genetically Encoded Diazarine Photocrosslinker in *Escherichia coli*." *ChemBioChem.* **2007**, *8*, 2210-2214.

[38] Suchanek, M.; Radzikowska, A.; Thiele, C. "Photo-leucine and Photo-methionine allow identification of protein-protein interactions in living cells." *Nature Methods.* **2005**, *2*, 261-268.

[39] Shi, H.; Zhang, C.; Chen, G.; Yao, S. "Cell-Based Proteom Profiling of Potential Dasatinib Targets by Use of Affinity-Based Probes." *J. Am. Chem. Soc.* **2012**, *134*, 3001-30014.

[40] Van der Meijden, B.; Robinson, J. A. "Synthesis and application of photoproline – a photoactable derivative of proline." *Arkivoc.* **2011**, *Vi,* 130-136.

[41] Kelly, G.; Sharma, V.; Watanabe, C. "An improved method for culturing *Streptomyces sahachiroi*: Biosynthetic origin of the enol fragment of azinomycin B." *Bioorg. Chem.* **2008**, *36*, 4-15.

[42] Thuring, J.; van Gaal, A.; Hornes, S.; de Kok, M.; Nefkens, G.; Zwanenburg, B. "Synthesis and biological ecaluation of strigol analogues modified in the enol ether part." *J. Chem. Soc., Perkin Trans.* **1997**, *1,* 767-774.

[43] Ikeda, Y.; Behrman, E. J. "Improved Synthesis of Photo-leucine." *Synth. Commun.* **2008**, *38*, 2276-2284.

[44] Frost, J.; Knowles, J. "3-Deoxy-D-*arabino*-heptulosonic Acid 7-Phosphate: Chemical Synthesis and Isolation from *Escherichia coli* Auxotrophs." *Biochemistry*. **1984**, *23*, 4465-4469.