THE SYNTHESIS OF ANTHRANILIC ACID, TRYPTOPHAN, AND SULFENYL CHLORIDE ANALOGUES, AND ENZYMATIC STUDIES

By

Phanneth Som

(Under the Direction of ROBERT S. PHILLIPS)

Abstract

This dissertation includes four chapters. Chapter 1 includes the introduction and literature review. Chapter 2 covers the enzymatic synthesis of tryptophan via anthranilic acid analogues. Chapter 3 covers the nitration and resolution of tryptophan and synthesis of tryptophan derivatives. Chapter 4 covers the synthesis of sulfenyl chloride derivatives.

Tryptophan is important in many aspects in the studies of proteins and also serves as precursors for important compounds. The enzymatic synthesis of tryptophan analogues from anthranilic acid analogues is important and could provide a more efficient route for incorporation of non-canonical amino acids onto proteins. Nitration of tryptophan is also important because it provides a synthetic route for synthesizing other derivatives of tryptophan. Sulfenyl chloride derivatives can be used for labeling proteins and the tryptophan analogues can be determined by mass spectrometry analysis with proteins with a limited number of tryptophan residues makes this application useful.

INDEX WORDS: Tryptophan, Anthranilic acid, Enzymatic synthesis, Non-canonical amino acid, Nitration, Sulfenyl chloride, Incorporation

THE SYNTHESIS OF ANTHRANILIC ACID, TRYPTOPHAN, AND SULFENYL CHLORIDE ANALOGUES, AND ENZYMATIC STUDIES

by

Phanneth Som

B.S., Georgia State University, 2003

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2009

THE SYNTHESIS OF ANTHRANILIC ACID, TRYPTOPHAN, AND SULFENYL CHLORIDE ANALOGUES, AND ENZYMATIC STUDIES

By

Phanneth Som

Major Professor: Robert S. Phillips

Committee: George F. Majetich

Jason J. Locklin

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2009

DEDICATION

To my parents, brothers and sister for their love and patience over the years that has helped in molding me into the person I am today. I would also like to dedicate this to Professor Robert S. Phillips for his guidance and friendship over the years.

ACKNOWLEDGEMENTS

I would like to thank my research advisor, Prof. Robert S. Phillips, for allowing me to join his group and offering me his much appreciated guidance over the years. His advice proved to be invaluable for me when it came to anything pertaining towards my research or goals in life if it be biochemistry, chemistry or world politics. His knowledge about everything in general astounded me on many occasions. I would also like to thank him for years of patience and understanding. I am very fortunate to be able to work for a man of this caliber and will have good memories as I move towards other goals in my life. I would also like to thank my advisory committee members, Prof. George F. Majetich for his years of synthetic advice and banter about everyday things, and Dr. Jason J. Locklin, for the shared humor and advice on how to improve as a chemist. Overall, I would like to thank my committee for their guidance, patience and encouragement.

I would like to thank my family for their years of support on many different levels and aspects of my life throughout the years. They have done for me more than anyone can ask for from family members. Most notably I would like to thank Kethanda Som for his moral support over the years when I would need advice about work or life. Also, I would like to thank my uncle and aunt, John and Kesa Simon for helping to support me when I was an undergraduate.

My lab partners throughout the years of my stay at UGA: Austin Harris, Musa Musa, Santiago Lima, Sunil Kumar, Chandan Maintrani, Nathan Lott, Johnny Do, Vijay Gawandi, and Jallendar Borra have contributed on many different things for me throughout the years. We have

shared many ups and downs pertaining to research and life. I would also like to say thanks to Brandon Quillian for years of friendship and advice. I would also like to say thanks to the members of the Majetich lab for their help they have given me over the years. I would like to say thanks to the new group members of the Phillips group, Kyle Beard, Chris Nealon, and Bryan Linn for the advice on my presentation and insights.

I wish to thank the graduate students and staff at the Chemistry Department, University of Georgia, for their help in getting my work done.

I would like to express my gratitude to my undergraduate research advisors, Prof. David W. Boykin and Dr. Chad E. Stephens for teaching and giving me the opportunity to conduct research in their laboratory at Georgia State University. Their support and encouragement helped me progress to where I am now and will continue to help me in the future. I would also like to thank my friends, Jeremy Rodgers and Subrata Mishra, for being such good friends throughout the years. I would like to thank all of my friends in Athens and Atlanta for their support.

I would like to give a special thanks to my wife for her support and love for the past year. She has encouraged me to pursue all of my goals. For this I am truly grateful to her.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
AA	Anthranilic Acid
ASase	Anthranilate Synthase
GC	Gas Chromatography
g	gram
HPLC	High Performance Liquid Chromatography
IGPSase	Indole Glycerol Phosphate Synthase
mmol	millimole
mL	milliliter
MOPS	3-(n-Morpholino)Propanesulfonic Acid
NAD^+	Nicotinamide Adenine Dinucleotide
NADH	Nicotinamide Adenine Dinucleotide, Reduced
$NADP^+$	Nicotinamide Adenine Dinucleotide Phosphate

NADPH	Nicotinamide Adenine Dinucleotide Phosphate, Reduced
NMR	Nuclear Magnetic Resonance
P_i	Phosphate
PEP	Phosphoenolpyruvate
PRAIase	Phosphoribosyl Anthranilate Isomerase
PRPP	Phosphoribosyl Pyrophosphate
PRTase	Phosphoribosyl Transferase
PLP	Pyridoxal-5-Phosphate
PP	Pyrophosphate
SPI	Selective pressure incorporation
TLC	Thin Layer Chromatography
TEA	Triethylamine
TFA	Trifluoroacetic Acid
TFAA	Trifluoroacetic Anhydride
TSase	Tryptophan Synthase
TPL	Tyrosine Phenol Lyase
VB	Vogel-Bonner

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Anthranilic Acid

Carl Julius Fritzsche (1808-1871) published a paper in 1840 where he identified a series of compounds. Further degradation of one of these compounds resulted in the discovery of a compound he named "Chrysanilsaure." He also obtained a compound he named "Anthranilsaure" [1]. He determined the composition of this compound accurately and it was found to be acidic. He noticed that the compound decomposed to aniline and carbon dioxide when heating above the melting point (145°C). He was unable to determine the actual structure, which was established later through the work of Adolf von Baeyer [2].

Anthranilic acid (2-aminobenzoic acid, AA, **1**) is a white to pale yellow crystalline powder with a sweetish taste [3]. AA is an inexpensive, important and versatile starting material for the synthesis of benzofused heterocycles and other molecules. It also plays an important role in the biosynthesis of tryptophan, tyrosine, phenylalanine, and several types of alkaloids. AA exhibits an amethyst fluorescence when dissolved in ether, alcohol or glycerol [3]. AA has properties that can be used for bioanalytical purposes such as monitoring the glycosylation of proteins [4]. AA is an amphoteric compound like other amino acids, which behaves as an acid and base [5]. An intramolecular proton exchange has been documented, which occurs in AA between the amine and carboxylic acid [6]. AA contains a hydrogen acceptor and a hydrogen donor that are

in a conjugated system. The acid and base strengths of AA differ in different media. Tables 1.1

and 1.2 illustrate the differing pK_a values in different media as both the free base and protonated

forms [5].

Table 1.1: pK_a of Anthranilic Acid

pK_a	Medium
6.53	50% dioxane in H_2O
6.51	84.2 % wt. MeOH in H_2O
5.40	EtOH/H_2O
5.00	H_2O
4.95	H_2O
3.71	pyridine

Table 1.2: pK_a of Protonated Anthranilic Acid

pK_a	Medium
2.36	Nitromethane
2.15	H_2O
2.11	H_2O
1.57	84.2 % wt. MeOH in H_2O

Some AA derivatives have photophysical properties and have been used as sunscreens [7].

Other derivatives of AA shown in Figure 1.1 below also have important commercial implications

such as ester methyl anthranilate (**2**), which is an important ingredient in perfumes and serves as

a flavor additive in soft drinks, such as grape flavoring [8]. Ester methyl anthranilate is prepared

by performing a Fisher esterification with methanol or from isatoic anhydride and methanol.

Another derivative, anthranilamide (2-aminobenzamide, **3**), has been used as a starting material

in synthetic procedures for pharmaceuticals and natural products [3]. Anthranilonitrile (2-

aminobenzonitrile, **4**) is used industrially and is prepared from 2-nitrotoluene from ammonia

catalyzed by silica in a gas phase reaction [9]. Isatoic anhydride (2*H*-3,1-benzoxazine-2,4(1*H*)-

dione (**5**) is used as a starting material for various syntheses and is synthesized from performing

a condensation of AA and phosgene [10]. AA has been used to form coordination complexes

with many metals such as gallium, aluminum, lithium, and other metals, and they were

determined using X-ray crystallography [11].
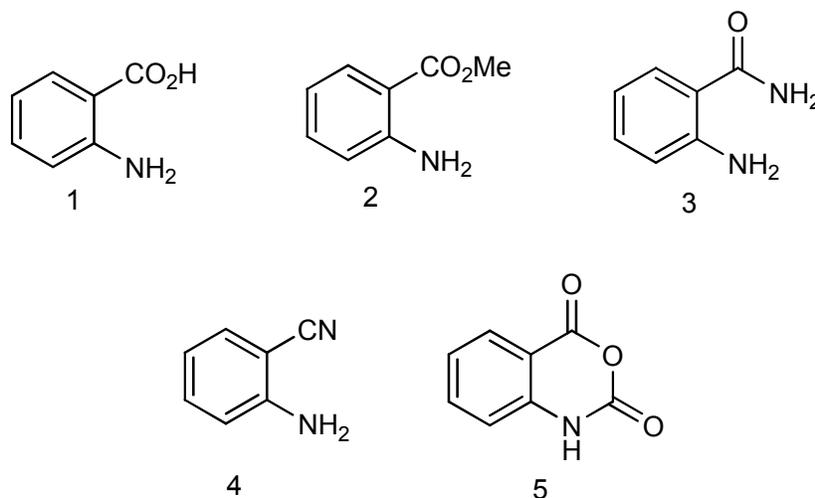
Figure 1.1: Important anthranilic acid analogues

Derivatives of AA have been used for medicinal purposes such as flufenamic acid, glafenine, and

etofenamate (**6**, **7**, and **8** respectively), which have non-steroidal antiflamatory and analgesic

properties [12]. The 3-substituted anthranilate hydroxamic acids (**9**) are known to inhibit matrix

metalloproteinases (MMPs), which are enzymes that play a role in remodeling during the

degradation of extracellular matrix proteins. They have also been used in the studies of several

diseases such as cancer and rheumatoid arthritis [13]. The anthranilate amides (**10**) have been

shown to be good cholecystokinin receptor antagonists [14]. The amides (**11**) have been shown

to behave as competitive inhibitors of protein-tyrosine phosphatases, which are a class of

enzymes that are therapeutic targets for diseases such as cancer, diabetes, and inflammation [15].
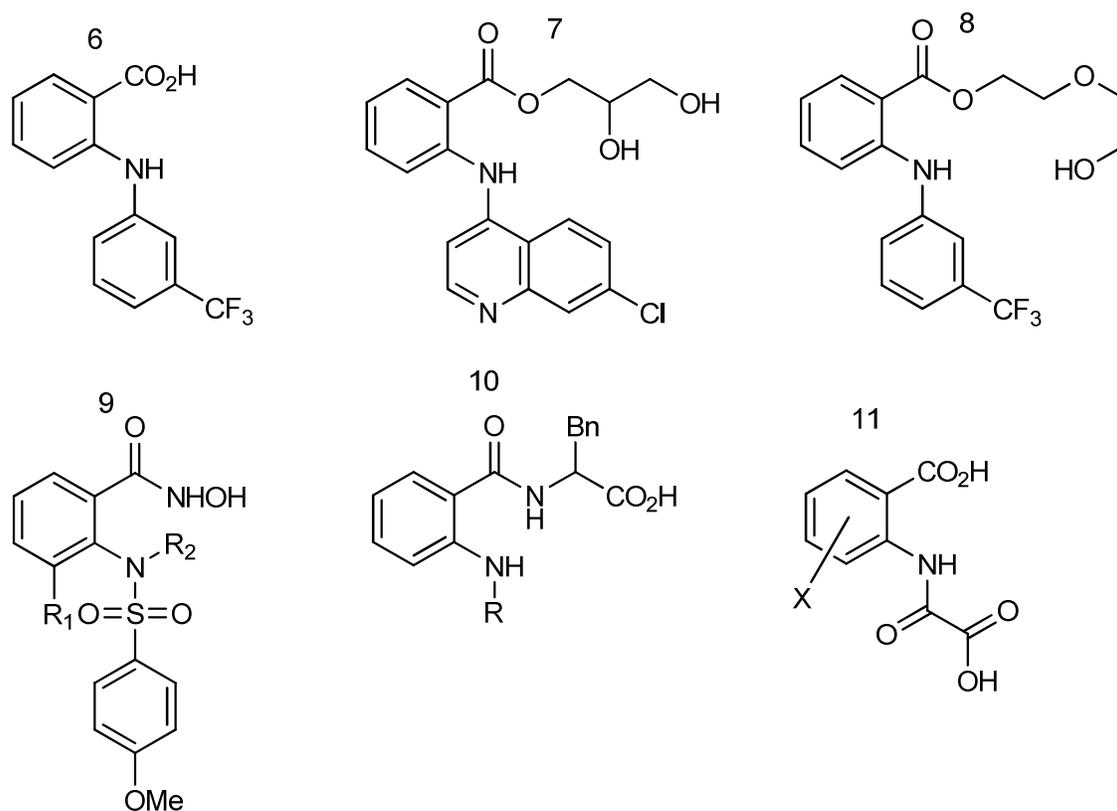
Figure 1.2: Medicinal derivatives of anthranilic acid

The synthesis of AA and analogues is shown below in Scheme 1.1. Typically, AA is prepared

by the reduction of *o*-nitrobenzoic acid [1]. The industrial synthesis of AA begins with the

oxidation of *o*-xylene to give phthalic anhydride using V_2O_5 [16]. The anhydride is opened

using ammonia to give both carboxylic acid and amide functionalities. Performing the Hofman

rearrangement under basic conditions using sodium hypochlorite gives AA on an industrial scale.

The synthesis of substituted AA analogues begins by treating aniline with chloral hydrate and

hydroxylamine to give isonitrosoacetanilide, which upon treatment with strong acid is converted

to isatin [17]. Then, oxidation of isatin using alkaline hydrogen peroxide gives the desired

substituted AA analogues.

Scheme 1.1: Industrial synthesis of anthranilic acid and anthranilic acid analogues
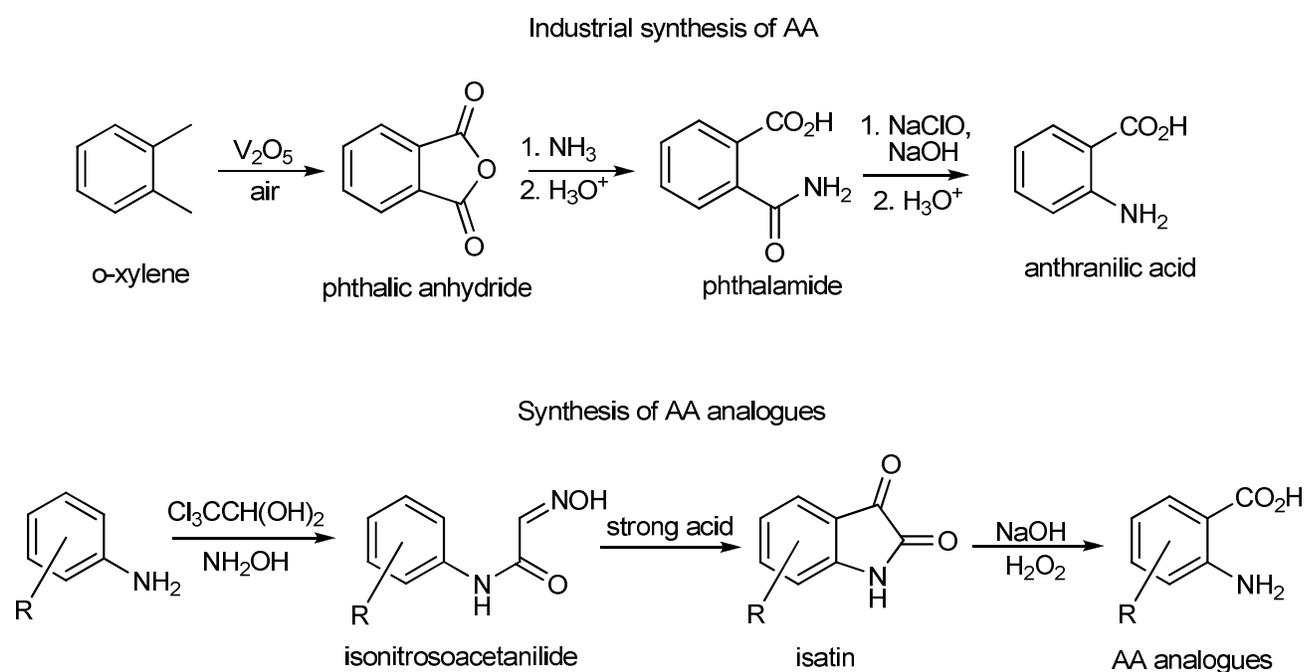
AA is important for the biosynthesis of aromatic amino acids [18]. The aromatic amino acids,

phenylalanine, tyrosine, and tryptophan, are derived from a shared pathway that has chorismic

acid as a key intermediate for their biosynthesis [19]. Chorismate is a common molecule for the

synthesis of the aromatic amino acids, the fat-soluble vitamins E and K, folic acid and other

aromatic compounds like enterochelin and plastoquinone [20].

Scheme 1.2: Biosynthetic pathway of chorismic acid.
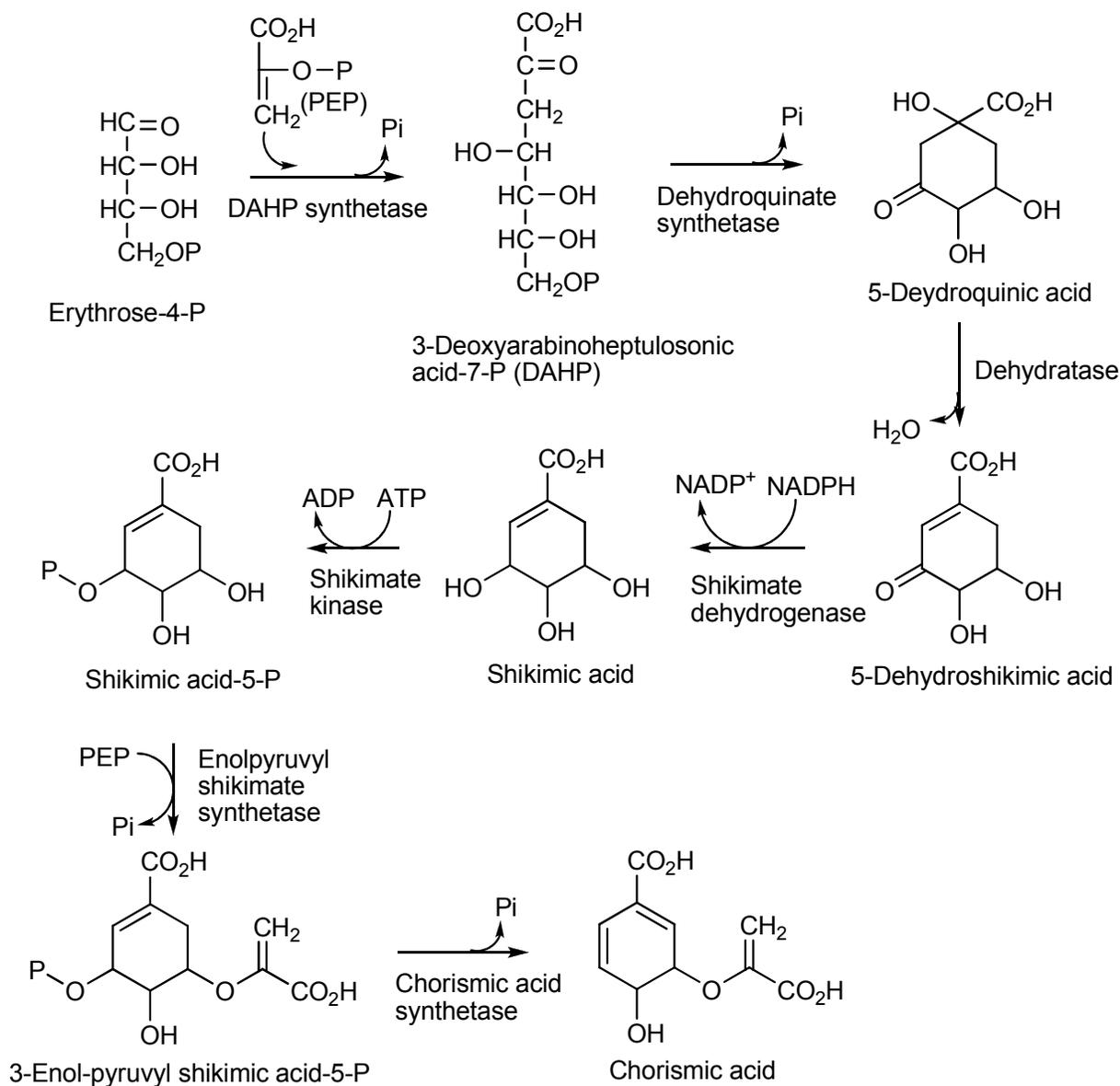
Chorismate biosynthesis occurs via the shikimate pathway shown in Scheme 1.2 above [19].

The precursors for this pathway are the metabolic intermediates phosphoenolpyruvate (PEP) and

erythrose-4-phosphate. These intermediates are linked to form 3-deoxyarabinoheptulosonate-7-

phosphate (DAHP) by DAHP synthase. This is considered the most important point for

regulation of aromatic amino acid biosynthesis despite being one of the early steps. In the next

step, DAHP is cyclized to form a six-membered saturated ring structure, 5-dehydroquinate by

dehydroquinate synthetase. In the following step, 5-dehydroquinate is converted into α,β-

unsaturated 5-dehydroshikimic acid by a dehydratase, liberating water. Then 5-dehydroshikimic

acid is reduced to shikimic acid, which is then catalyzed by shikimate dehydrogenase with

NADPH, releasing $NADP^+$. Shikimic acid is converted into shikimic acid-5-P by a

phosphorylation reaction by shikimate kinase and ATP, releasing ADP. Shikimic acid-5-P is

then condensed with PEP to give phosphate and 3-enol-pyruvyl-shikimic acid-5-P. After these

steps, shikimic acid is converted *in vivo* into chorismic acid by chorismic acid synthetase [19].

Scheme 1.3: Conversion of chorismic acid to aromatic amino acids.
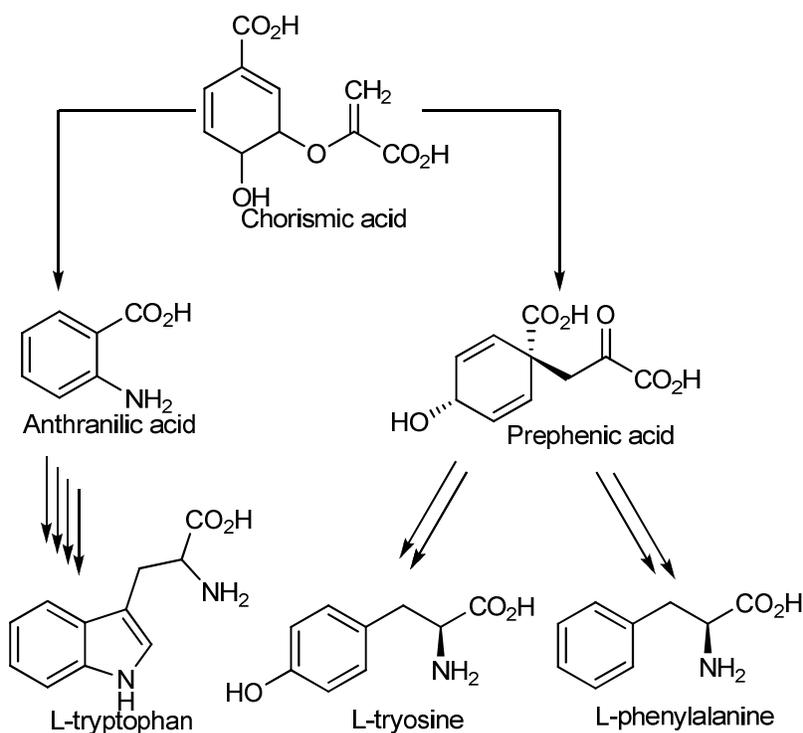
Chorismic acid can then be converted into AA or prephenic acid, which can give L-tryptophan

from AA, or L-tyrosine or L-phenylalanine from prephenic acid through a series of enzymatic

reactions shown in Scheme 1.3 above [19]. AA is formed from chorismic acid and glutamine by anthranilate synthetase, releasing glutamate and pyruvate. Prephenic acid is formed by a Claisen rearrangement of chorismic acid by chorismate mutase.

Tryptophan

Tryptophan is one of the eight essential amino acids, with the others being isoleucine, leucine, lysine, methionine, phenylalanine, threonine, and valine. These amino acids cannot be synthesized in the body and must be obtained from the diet. Tryptophan is necessary for normal growth and development, and it serves as the precursor for several metabolites such as serotonin, kynurenine, melatonin, and niacin [21, 22]. Serotonin (5-hydroxytryptamine, 5-HT) is a neurotransmitter in the central nervous system (CNS) that is responsible for the regulation of mood, anger, aggression, body temperature, and appetite [23]. Kynurenine is a key component for the synthesis of several metabolites such as kynurenic and quinolic acids which could have an effect on neurotransmitters [23]. Melatonin is an important metabolite that is involved in sleep regulation [23]. Niacin (vitamin B_3 or nicotinic acid) is essential for energy metabolism, DNA repair, and production of steroid hormones in the adrenal gland [23]. Deficiencies in niacin could lead to the development of pellagra. Niacin is also a precursor for NADH, NAD^+, $NADP^+$ and NADPH, which all play important metabolic roles in living cells and involved in all biosynthetic and energy pathways in the body [24].

Tryptophan is a heterocyclic compound that is found in small amounts in most proteins. It is highly hydrophobic in nature and often occupies critical positions in proteins [21]. It plays roles in the structure and function of proteins, peptides, and in the biosynthesis of hormones, such as serotonin in animals, indole alkaloids in plants, and other biologically important products [25].

Tryptophan has been extensively reviewed over the years [26]. Figure 1.3 illustrates the metabolites formed from the metabolism of tryptophan.
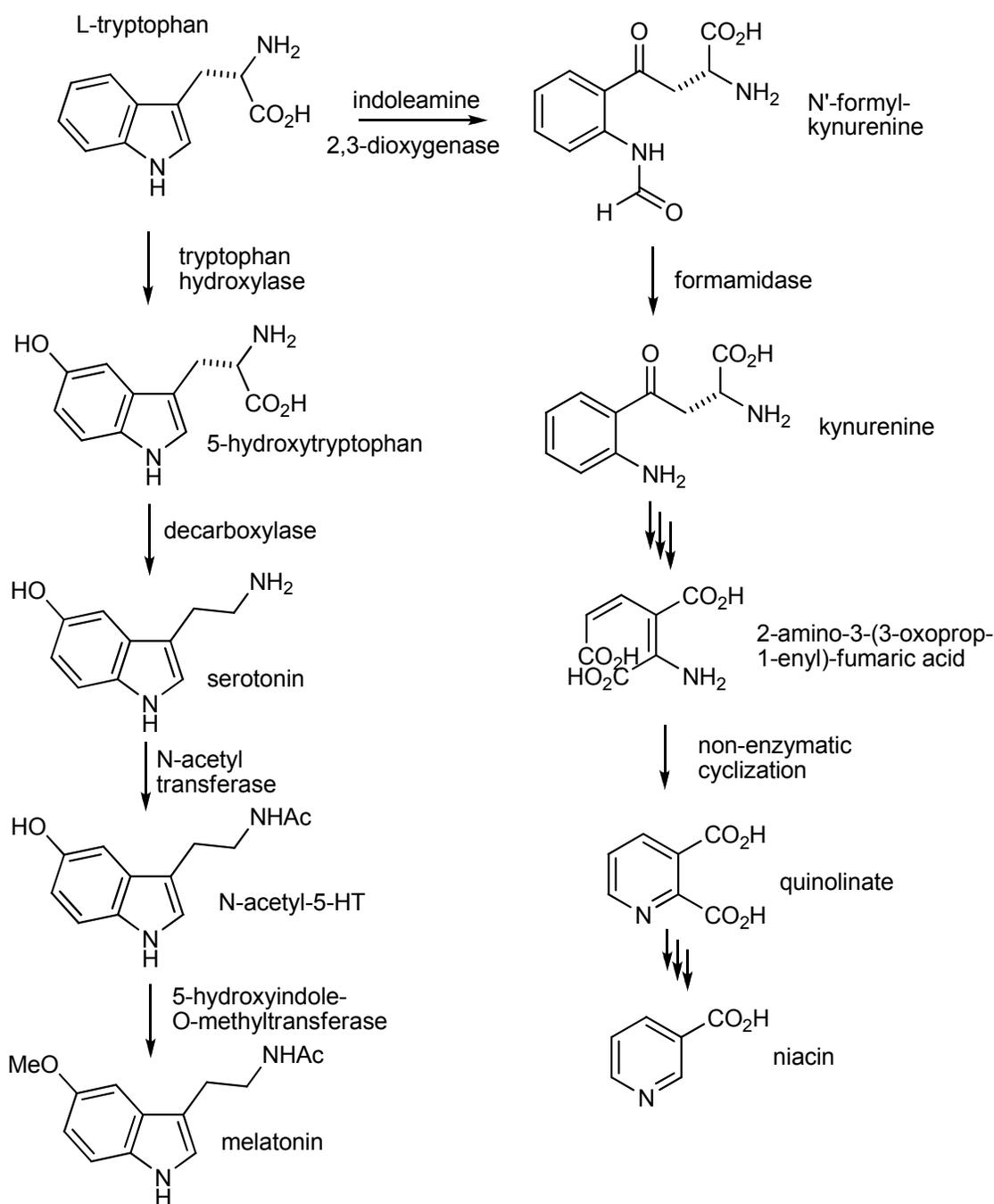
Figure 1.3: Biosynthetic derivatives of tryptophan

The metabolism of L-tryptophan has been found recently to protect the fetus from rejection by the mother and regulate the human immune response [27]. Tryptophan is the largest and the least abundant of the amino acid of the 20 common amino acids, usually constituting 1% or less of the average protein mass [20]. In some proteins, there are no tryptophan residues present. In humans, tryptophan has low tissue storage and has the lowest overall concentration of the common amino acids in the body. Because of the low concentration of tryptophan in the human body, it is thought to play a rate-limiting role in protein synthesis [28]. The L isomer of tryptophan is used for protein synthesis, as is the same for all amino acids, and tryptophan can pass through the blood brain barrier. Tryptophan is the most costly of the common amino acids requiring 78 moles of ATP to synthesize one mole of tryptophan. The energy cost is 20% higher than the biosynthesis of phenylalanine, which is the second most costly [20]. The biosynthesis of tryptophan requires approximately double the biosynthetic energy of the average of the common amino acids. Research in the tryptophan field has contributed to the resolution of problems in the areas of protein structure and synthesis, coding, suppression, transcriptional and translational control, protein-protein and protein-nucleic interaction, catalysis, mutagenesis, genetic recombination, viral genetics, molecular cloning, and transmembrane phenomena [20]. Tryptophan also has therapeutic applications, such as being used in the treatment of depression, mood disorders, and sleep disorders [23].

Tryptophan was first isolated in 1901 by Hopkins and Cole from solutions of tryptic casein hydrolysate [29]. The structure was first reported by Ellinger and Flamand shortly afterwards. They reported the first synthesis of tryptophan in 1908, shown in Scheme 2.1 [30]. They synthesized indole-3-aldehyde from indole using the Reimer-Tiemann reaction, then combined indole-3-aldehyde, sodium acetate, acetic anhydride, and hippuric acid to do an Erlenmeyer

azlactone synthesis to give the desired intermediate. Mild alkaline hydrolysis and treatment with

sodium metal in ethanol gave tryptophan (10-17% overall yield).

Boyd and Robson reported a better synthesis using hydantoin instead of hippuric acid in 1935

with an overall yield of 50-60% shown in Scheme 2.2 [31]. They reacted indole-3-aldehyde with

hydantoin to give the indole hydantoin intermediate. Treatment of indole hydantoin with

ammonium sulfide and ammonium hydroxide gave tryptophan in a better yield and fewer steps

to finalize this synthesis. The chemical synthesis of tryptophan was replaced in the 1980s by a

fermentation processes that produced it on a larger scale, thus making tryptophan supplements

more widely available [32].

Scheme 1.4: Synthesis of tryptophan via hippuric acid
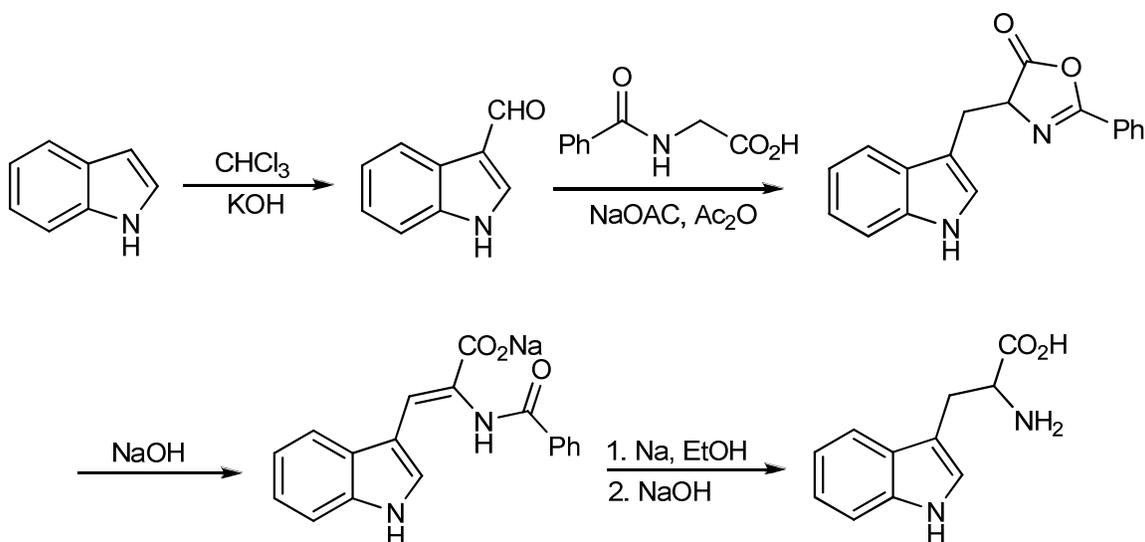
Scheme 1.5: Synthesis of tryptophan via hydantoin
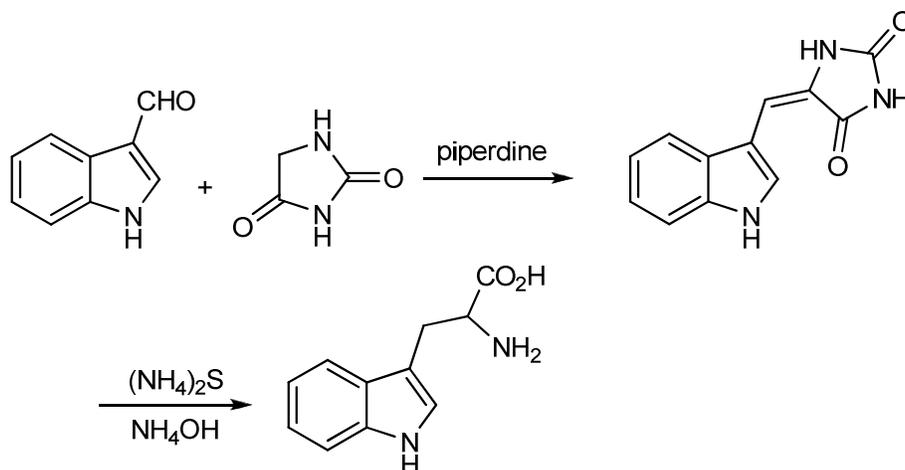
Enzymatic Studies

Enzymatic reactions are nature's way of making life possible and organic chemists constantly strive to mimic and surpass them. Nature has been perfecting her reactions over millions of years and the evolution of enzymatic reactions continues today. Organic chemists find it difficult to mimic nature in areas such as the total synthesis of metabolites and compounds. Chemical pathways usually give low yields after expensive, multi-step syntheses. Nature accomplishes these tasks with an efficiency that we have yet to match, and the desired compounds are usually isolated in much higher yields after a few extraction techniques are applied. Enzymes are generally natural proteins that catalyze chemical reactions [33]. Enzymes can have molecular weights that range between several thousand to several million Daltons and have the ability to catalyze transformations on small molecules such as carbon dioxide and nitrogen [33]. Enzymes have the unique ability to accelerate reaction speeds up to 10^8-10^{10} [34]. An example of this is hydrolyzing an unactivated amide, which usually requires a ten hour reflux

in 8N HCl, but using chymotrypsin this is done in minutes at neutral pH and ambient temperature [34].

Tryptophan exhibits unique spectroscopic properties, which enables it to serve as an intrinsic probe for protein structures, protein dynamics, and intermolecular interactions between proteins and other molecules [20]. Rationally designed protein mutants, that contain electron-rich atoms and retain a well-ordered three-dimensional structure in the crystalline state, have been shown in the past several years to be one of the most promising approaches to solving the phase problem in protein X-ray crystallography [35]. 7-Azatryptophan (**6b**), 5-hydroxytryptophan (**7a**) and fluorotryptophan analogues (**7b**-**7d**) have been used in various studies [36]. In addition, 7-Azatryptophan and 5-hydroxytryptophan have been used as spectroscopic probes studying protein-protein and protein-DNA complexes and fluorotryptophan analogues have been used in NMR studies [36]. Figure 1.4 shows the tryptophan derivatives that have been used in enzymatic studies.

6a R_1 = N, R_2 = C
6b R_1 = C, R_2 = N

7a R_1 = H, R_2 = OH, R_3 = H
7b R_1 = F, R_2 = H, R_3 = H
7c R_1 = H, R_2 = F, R_3 = H
7d R_1 = H, R_2 = H, R_3 = F
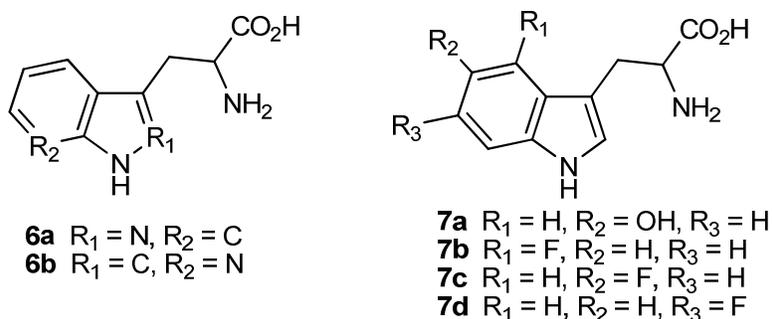
Figure 1.4: Tryptophan derivatives used in enzymatic studies

Enzymatic studies on tryptophan and its analogues have gained great importance over the years, with tryptophan having unique and desirable spectroscopic properties in the study of proteins. Tryptophan analogues, such as 5-hydroxytryptophan, 7-azatryptophan, 2-azatryptophan, 6-

fluorotryptophan, 5-fluorotryptophan, 4-fluorotryptophan, and other compounds, have been incorporated into proteins and have been studied [35, 36]. The enzymatic synthesis of both azatryptophan and chlorotryptophan analogues have been reported using L-serine and tryptophan synthase [37, 38]. In addition, the enzymatic syntheses of selenolopyrroles, furopyrroles, and thienopyrroles have been reported using similar conditions [25, 39]. 5-Hydroxytryptophan and the azatryptophan analogues (**6a**, **6b**) alter the absorption and fluorescence properties of the protein they have been incorporated into with a red shift of between 10-15 nm. This shift provides a unique spectral window for monitoring the absorption and exciting the fluorescence of the protein that contains the incorporated unnatural amino acid [36]. The fluorotryptophan analogues that are incorporated allow for ^{19}F NMR studies on proteins with essentially no alteration of biological activity [36].

The incorporation of non-canonical amino acids poses many drawbacks. Non-canonical amino acids are usually toxic and there is a preferential incorporation of the canonical amino acids. The most common non-canonical amino acid being used for incorporation is selenomethionine [35]. Telluromethionine incorporation has been reported and exhibits the same problems as selenomethionine [35]. The problems associated with these non-canonical amino acids are the high level of toxicity and the low levels of incorporation [35, 40].

Selective pressure incorporation (SPI) is a method for *in vivo* or *in vitro* incorporation of non-canonical amino acids and it allows for an expansion beyond the 20 canonical amino acids [40, 41]. The *in vivo* incorporation is based on the use of intact host expression cells that are subjected to enough pressure to incorporate the non-canonical amino acid and the *in vitro*

incorporation is based on the natural phenomena of suppression, using various cellular extracts [40].

The requirements for the SPI method are: (i) a strong host auxotrophism, (ii) a controlled amino acid supply, (iii) optimal fermentation conditions, (iv) an efficient expression system, and (v) a robust protein expression [35]. The main drawback associated with the SPI method is that only a limited number of amino acids can be incorporated into proteins [40].

The SPI method is closely related to the problem associated with expressing toxic gene products, and this is done by keeping silent the target gene activity while the host cells are allowed to grow to appropriate levels [35]. Then, after enough healthy cells have been grown, the protein can be synthesized by inducing the cells. The host cells will then serve as factories for the production of the recombinant protein [35]. The suppression of the metabolic pathway for the canonical amino acid for the cells is important to ensure that the non-canonical amino acid is incorporated. This is done in a minimal medium with growth-limiting concentrations of the native canonical amino acid that is the natural substrate of the cells. Then the non-canonical amino acid is introduced and protein synthesis with the incorporated non-canonical amino acid can be induced [35]. The other methods of incorporation are partial incorporation by exposure of a non-canonical amino acid and site-specific incorporation [42]. All three incorporation methods have the drawback of producing low yields.

phosphoribosyl anthranilate

indoleglyercolphosphate

PRTase

PP

PRPP

IGPSase

CO_2

anthranilic acid

CO_2H

NH_2

glutamic acid

+

pyruvic acid

glutamine

ASase

chorismate

CO_2H

CO_2H

HO

1-(o-carboxyphenylamino)-1-deoxyribulose-5-phosphate

PRAlase

TSase

L-serine

gly-3-P
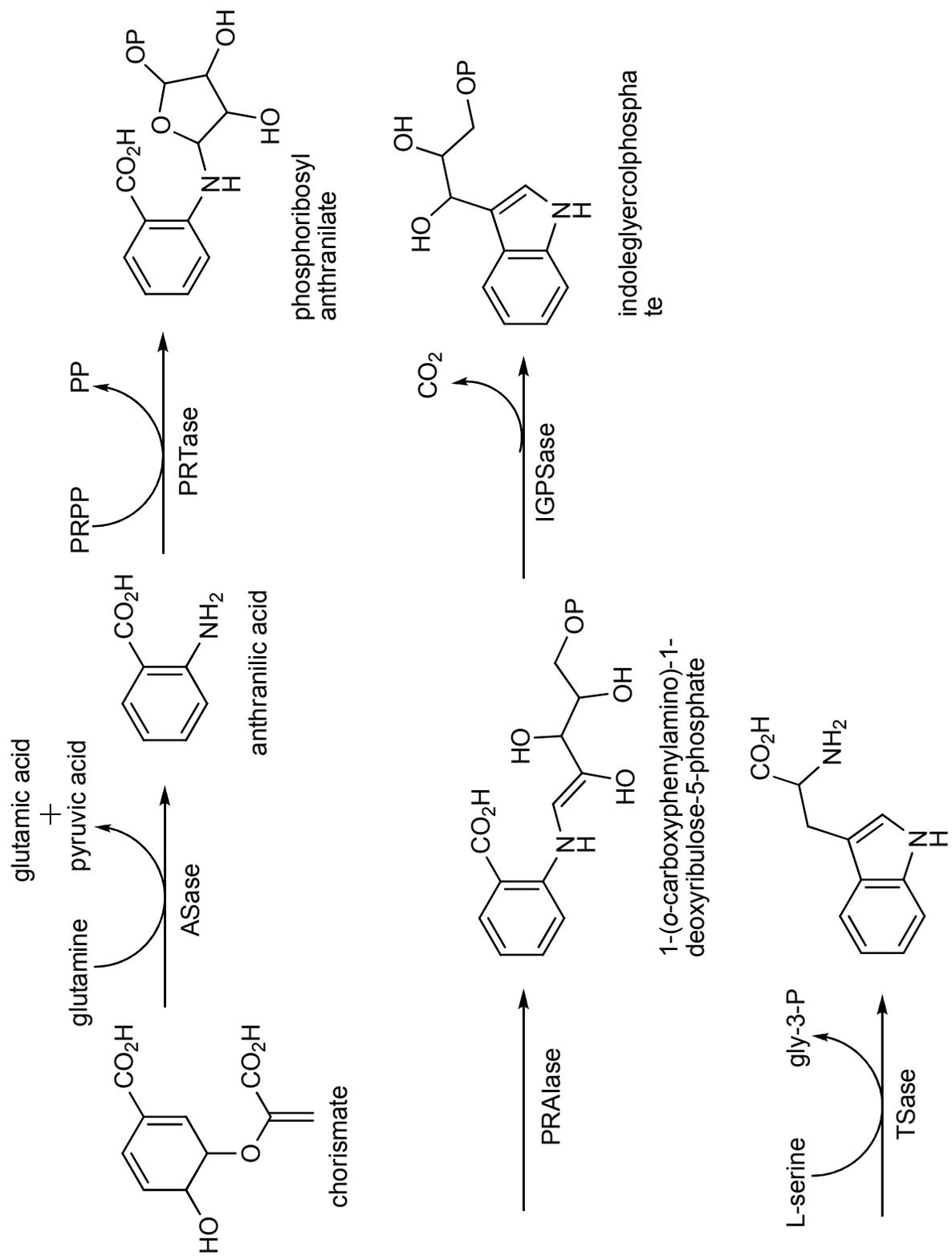
The tryptophan biosynthetic pathway shown in Figure 1.5 has been used extensively for studying biological systems. Many people have contributed to the field of tryptophan research, such as the research groups of C. Yanofsky, B Davis, D. Sprinson, F. Gibson, A. J. Pittard, and K. M. Herrmann [18]. The study of the tryptophan biosynthetic pathway has been extensive over years with many reviews [26]. The tryptophan biosynthetic pathway involves the five key enzymes ASase (Anthranilate Synthase), PRTase (Phosphoriboysl Transferase), PRAIase (Phosphoribosyl Anthranilate Isomerase), IGPSase (Indoleglycerol Phosphate Synthase), and TSase (Tryptophan Synthase), which each play a crucial role in the pathway [20, 21]. There is still much to understand and explore using tryptophan and its biosynthetic pathway. The field continues to expand in many directions to help explain biological and chemical questions.

References

1. Sheibley, F. E. *J. Chem. Educ*. **1943**, *20*, 115-117.

2. (a) Baeyer, A. *Ber*. **1878**, *11*, 1296-1297. (b) Baeyer, A. *Ber*. **1882**, *15*, 50-56. (c) Baeyer, A.; Drewsen, V. *Ber*. **1882**, *15*, 2856-2864.

3. *The Merck Index*, 13th ed.; Merck Research Laboratories: Whitehouse Station, NJ, **2001**, p. 74.

4. Dhume, S. T.; Ebert, M. B.; Saddic, G. N.; Anumula, K. R. *Met. Mol. Biol.* **2002**, *194*, 127.

5. Wiklund, P.; Bergman, J. *Curr. Org. Syn*. **2006**, *3*, 379-402.

6. Reiter, J.; Beyer, A.; Potschka, M.; Schuster, P. *J. Phys. Chem*. **1989**, *93*, 442-451.

7. (a) Beeby, A.; Jones, A. E. *J. Photochem. Photobiol., B* **2001**, *64*, 109. (b) Urbach, F. *J. Photochem. Photobiol., B* **2001**, *64*, 99.

8. Yadav, G. D.; Krishnan, M. S. *Org. Process Res. Dev.* **1998**, *2*, 86.

9. Bakke, J.; Heikman, H.; Nystrom, G. *Acta Chem. Scand.* **1972**, *26*, 355.

10. Coppola, G. M. *Synthesis*. **1980**, 505-536.

11. (a) Branch, C. S.; Lewinski, J.; Justyniak, I.; Bott, S. G.; Lipkowski, J.; Barron, A. R. *J. Chem. Soc., Dalton Trans.* **2001**, 1253. (b) Wiesbrock, F.; Schmidbaur, H. *J. Chem. Soc., Dalton Trans*. **2002**, 4703-4708. (c) Wiesbrock, F.; Schier, A.; Schmidbaur, H. *Z. Naturforsch., B: Chem. Sci*. **2002**, *57*, 251-254. (d) Wiesbrock, F.; Schmidbaur, H. *J. Am. Chem. Soc*. **2003**, *125*, 3622-3630. (e) Wiesbrock, F.; Schmidbaur, H. *Inorg. Chem*. **2003**, *42*, 7283-7289.

12. (a) Brogden, R. N. *Drugs*. **1986**, *32 Suppl 4*, 27-45. (b) Badwan, A. A.; Zughul, M. B.; Al Omari, M. *Analytical Profiles of Drug Substances and Excipients*. **1992**, *21*, 197-232.

13. Levin, J. I.; Du, M. T.; DiJoseph, J. F.; Killar, L. M.; Sung, A.; Walter, T.; Sharr, M. A.; Roth, C. E.; Moy, F. J.; Powers, R.; Jin, G.; Cowling, R.; Skotnicki, J. *Bioorg. Med. Chem. Lett*. **2001**, *11*, 235-248.

14. Varnavas, A.; Lassiani, L.; Valenta, V.; Berti, F.; Mennuni, L.; Makovec, F. *Bioorg. Med. Chem*. **2003**, *11*, 741-751.

15. Andersen, H. S.; Iversen, L. F.; Jeppesen, C. B.; Branner, S.; Norris, K.; Rasmussen, H. B.; Moller, K. B.; Moller, N. P. H. *J. Biol. Chem.* **2000**, *275*, 7101-7108.

16. (a) Kutyanin, L. I.; Kuznetsov, A. A.; Poddubnyi, I. S.; Kulikova, O. A.; Gura, L. N.; Mil'otin, I. M.; Sergeev, S. A. *In RU 2187496, CA 139:22498*; **2002**. (b) Chapman, S. *J. Chem. Soc.* **1925**, *127*, 1791-1797.

17. Popp, F. D. *Adv. Heterocycl. Chem.* **1975**, *18*, 1-58.

18. *Amino Acid Biosynthesis – Pathways, Regulation and Metabolic Engineering*; Wendisch, V. F., Ed.; Microbiology Monographs; Springer: Berlin, **2007**.

19. Bender, D. A. *Amino Acid Metabolism*; John Wiley & Sons Ltd.: London, **1985**.

20. *Amino Acids Biosynthesis and Genetic Regulation*, Herrmann, K. M.; Somerville, R. L., Ed.; Biotechnology; Addison Wesley Publishing Company: Massachusetts, **1983**.

21. *The Microbial Production of Amino Acids*, Yamada, K., Kinoshita, S., Tsunoda, T., and Aida, K., Ed.; Kodansha Ltd: Tokyo, **1972**.

22. *L-Tryptophan Current Prospects in Medicine and Drug Safety*, Kochen, W. and Steinhart, H. L., Ed.; Walter de Gruyter: Berlin, **1994**.

23. Dawn, R. M.; Dawes, M. A.; Mathias, C. W.; Acheson, A.; Hill-Kapturczak, N.; Dougherty, D. M. *International Journal of Tryptophan Research*, **2009**, *2*, 45-60.

24. Weidel, H. *Justus Liebig's Annalen der Chemie und Pharmacie.* **2006**, *165*, 330-349.

25. Phillips, R. S.; Cohen, L. A.; Annby, U.; Wensbo, D.; Gronowitz, S. *Bioorg. Med. Chem. Letts.* **1995**, *5*, 1133-1134.

26. (a) Somerville, R. L.; Yanofsky, C. *J. Mol. Biol*. **1965**, *11*, 747-759. (b) Crawford, I. P.; Stauffer, G. V. *Ann. Rev. Biochem*. **1980**, *49*, 163-195.

27. (a) Munn, D. H.; Zhou, M.; Attwood, J. T.; Bondarex, I.; Conway, S. J.; Marshall, B.; Brown, C.; Mellor, A. L. *Science*, **1998**, *281*, 1191-1193. (b)Mellor, A. L.; Munn, D. H. *Immunology Today*, **1999**, *20*, 469-473.

28. Reilly, J. G.; McTavish, S. F. B.; Young, A. H. *J. Psychopharmcol*. **1997**, *11*, 381-392.

29. Hopkins, F. G.; Cole, S. W. *J. Physiol*. **1901**, *27*, 418-428.

30. Ellinger, A.; Flamand, C. *Z. physiol. Chem.* **1908**, *55*, 8.

31. (a) Boyd, W. J.; Robson, W. *Biochem. J.* **1935**, *29*, 555. (b) Boyd, W. J.; Robson, W. *Biochem. J.* **1935**, *29*, 2256.

32. Sidransky, H. *Tryptophan: Biochemical and Health Implications*. CRC Press, Boca Raton, FL, **2002**.

33. Silverman, R. B. *The Organic Chemistry of Enzyme-Catalyzed Reactions*; Academic Press: London, **2002**.

34. Menger, F. M. *Acc. Chem. Res*. **1993**, *26*, 206-212.

35. Bae, J. H.; Alefelder, S.; Kaiser, J. T.; Friedrich, R.; Moroder, L.; Huber, R.; Budisa, N. *J. Mol. Biol*. **2001**, *309*, 925-936.

36. (a) Senear, D. F.; Mendelson, R. A.; Stone, D. B.; Luck, L. A.; Rusinova, E.; Ross, J. B. A. *Anal. Biochem*. **2002**, *300*, 77-86. (b) Wong, C-Y.; Eftink, M. R. *Protein Science*. **1997**, *6*, 689-697. (c) Broos, J.; Gabellieri, E.; Biemans-Oldehinkel, E.; Strambini, G. B. *Protein Science*, **2003**, *12*, 1991-2000.

37. Sloan, M. J.; Phillips, R. S. *Bioorg. Med. Chem. Letts*. **1992**, *2*, 1053-1056.

38. Lee, M.; Phillips, R. S. *Bioorg. Med. Chem. Letts*. **1992**, *2*, 1563-1564.

39. Welch, M.; Phillips, R. S. *Bioorg. Med. Chem. Letts*. **1999**, *9*, 637-640.

40. Minks, C.; Alefelder, S.; Moroder, L.; Huber, R.; Budisa, N. *Tetrahedron*. **2000**, *56*, 9431-9442.

41. Budisa, N.; Alefelder, S.; Bae, J. H.; Golbik, R.; Minks, C.; Huber, R.; Moroder, L. *Protein Science*, **2001**, *10*, 1281-1292.

42. (a) Dougherty, D. A. *Curr. Opin. Chem. Biol*. **2000**, *4*, 645-652. (b) Bacher, J. M.; Ellington, A. D. *J. Bacteriol*. **2001**, *183*, 5414-5425.

CHAPTER 2

SYNTHESIS OF TRYPTOPHAN ANALOGUES VIA ENZYMATIC REACTIONS IN MINIMAL MEDIA

Introduction

Tryptophan is unique among amino acids because it has spectroscopic properties which can be used in the studying proteins [1]. There are very few tryptophan residues present in proteins. Some proteins are devoid of tryptophan, and this can be exploited for studying protein structures and interactions [1]. The unique properties of tryptophan offer a method of solving the phase problem in X-ray crystallography [2]. Using tryptophan analogues containing electron rich atoms, such as selenium, could be more useful for solving the phase problem and studying proteins [3, 12]. The incorporation of tryptophan analogues into proteins being studied has gained importance over the years. Incorporation of tryptophan analogues such as azatryptophan and fluorotryptophan has been reported as an application for ^{19}F-NMR and fluorescence studies [4]. Early attempts at incorporating unnatural amino acids were performed in order to gain insight into several basic biochemistry related problems, such as directed evolution and metabolic alternatives for bioengineered organisms [5]. Compstatin is a 13-residue cyclic peptide inhibitor that binds to the third component of complement and cancels the activation of complement cascade. Incorporation of tryptophan analogues into compstatin in *Escherichia coli* increases the activity three-fold relative to wild-type compstatin and in some cases lowered the

activity [6]. Exposure of the cells to an unnatural amino acid will result in partial incorporation despite the unnatural amino acid exhibiting toxicity to the cells [6]. The methods for incorporation of non-canonical amino acids are site-specific incorporation, exposure of the unnatural amino acid for partial incorporation, and the SPI method [7, 8]. The main drawback for these methods of incorporation is that the quantity produced from each method is usually low [3].

Enzymatic reactions using thienopyrroles, selenolopyrroles, and furopyrroles have been reported with L-serine and tryptophan synthase (PLP dependent enzyme), from *Salmonella typhimurium,* to give tryptophan analogues [9, 10]. Also, these conditions have been used for enzymatic reactions for azatryptophan, chlorotryptophan, and other analogues [11].

Using the tryptophan biosynthetic pathway to incorporate non-canonical amino acids is of interest because it could circumvent the main drawbacks with the other methods of incorporation. It has been shown that the final step of reacting thienopyrroles with L-serine and tryptophan synthase is successful as shown in Figure 2.1 below [9]. The main drawbacks are the tedious synthetic route, low yields and the instability of the thienopyrroles, which decompose easily. Replacing the benzene ring of anthranilate with AA analogues containing heterocyclic ring systems, such as thiophene and selenophene derivatives, would be of interest due to the electron rich atoms present. Replacing tryptophan residues with analogues could produce potential antagonists, drugs, and antiobiotics that could serve as therapeutic agents [8, 12]. Budisa and coworkers reported the incorporation of the thienopyrroles into proteins with a comparison to tryptophan containing proteins shown in Figure 2.2 below [13]. They also reported the incorporation of selenolopyrroles and this would be of more interest due to the more

electron dense selenium atom [3]. It would be of interest to synthesize AA analogues to use for incorporation of non-canonical amino acids that could potentially increase the yields and circumvent the other problems associated with the known methods of incorporation.
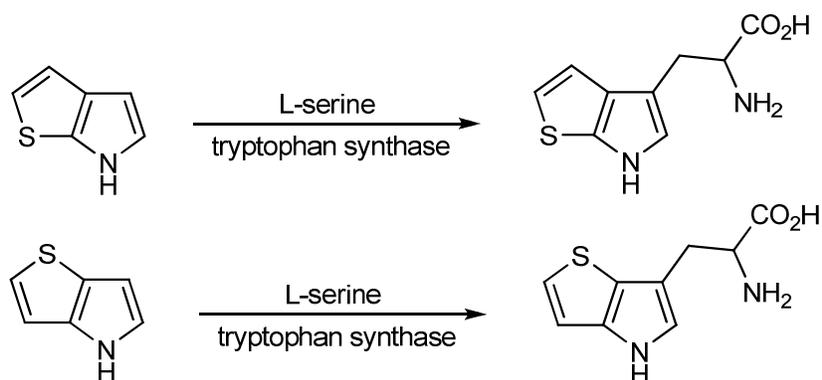
Figure 2.1: Enzymatic reaction of thienopyrroles.
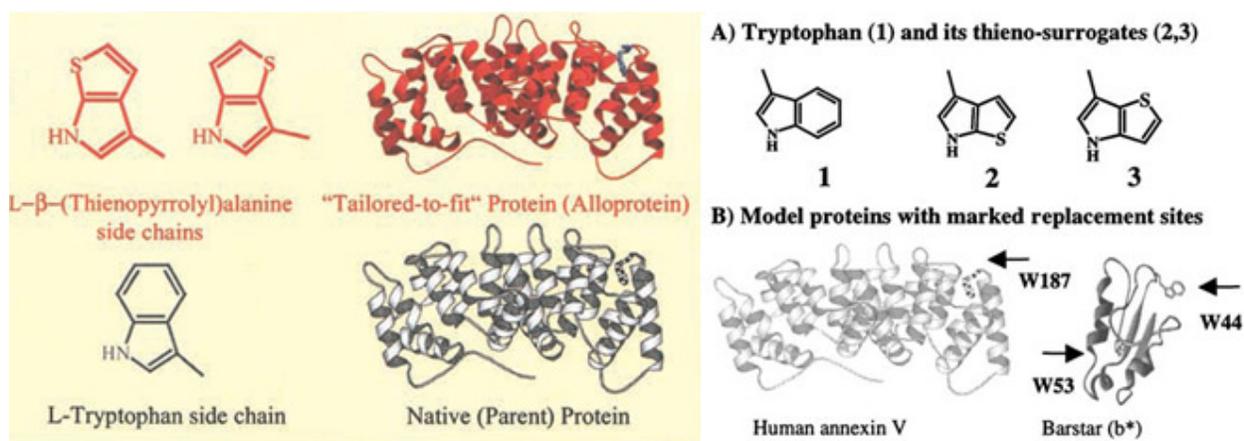
Figure 2.2: Incorporation of thienopyrroles comparison to tryptophan

Proposed Plan of Research

The synthesis of the AA analogues, 2-amino-3-thiophenecarboxylic acid and 3-amino-2-thiophenecarboxylic acid were envisioned to be prepared using Scheme 2.1 and 2.2, respectively. The 2-amino-3-thiophenecarboxylic acid synthesis was to start from 3-methylthiophene with a

selective nitration at the C2 position [14]. Then oxidation of the methyl group to a carboxylic

acid,followed by reduction of the nitro moiety to an amine by catalytic hydrogenation should

give the desired product [15].

Scheme 2.1: Proposed synthesis of 2-amino-3-thiophenecarboxylic acid
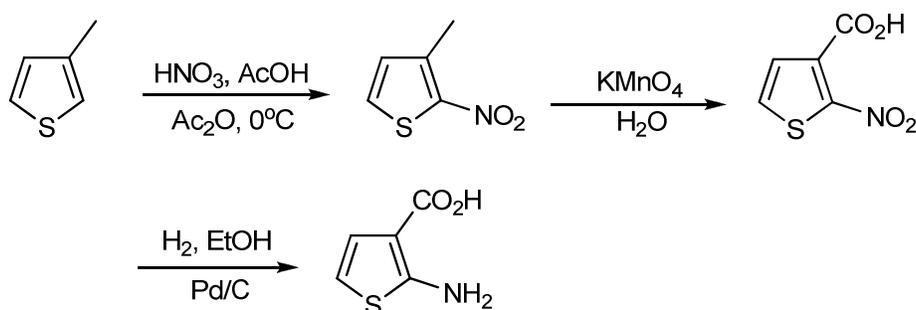
The 3-amino-2-thiophenecarboxylic acid synthesis starts with the selective nitration of 3-methyl-

2-thiophenecarboxylic acid [16]. The product undergoes a decarboxylation in a copper and

quinoline system [17]. Then oxidation of the methyl group to a carboxylic acid and reduction of

the nitro moiety to an amine by catalytic hydrogenation should give the desired product [15].

Scheme 2.2: Proposed synthesis of 3-amino-2-thiophenecarboxylic acid
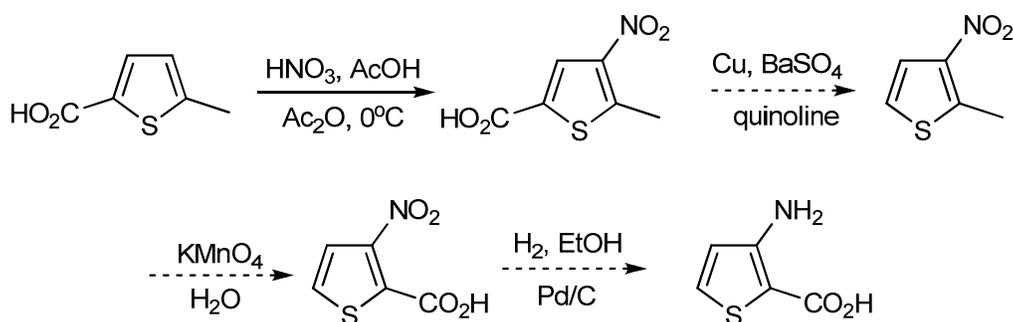
The AA analogues will then be placed in a minimal medium, MOPS, with the cell line W3110

tna A_2 Δ trp E_5 to wean the cells and induce growth with the non-canonical amino acid. This

will be done by a series of cell culture growths of varying concentrations of the natural substrate, AA, which will be lowered from one growth to the next growth. After growth of the cells, the proteins in the cells will be isolated to determination if activity is present and if incorporation was successful. Then further studies of the proteins could be performed, such as X-ray crystallography.

Results and Discussion

The synthesis of 2-amino-3-thiophenecarboxylic acid was successful with good yields from each step. The selective nitration of 3-methylthiophene directed the nitro group to the C2 position [14]. Both the oxidation of the methyl group to the carboxylic acid and reduction of the nitro group to an amine were efficient. The synthesis of 2-amino-3-thiophenecarboxylic acid was more tedious than expected. The nitration step was very efficient with good yields. The decarboxylation step proved to be very difficult. Decarboxylation in different copper and quinoline systems was attempted, which resulted in no reactions, but using the $BaSO_4$ promoted copper/quinoline gave partial decarboxylation with the desired product and starting material still present [18]. The decarboxylation in glycerol did not give the desired results [17]. The decarboxylation using the sublimination glassware gave white needles as a pure product, but the problem encountered with this procedure was the very low yield, usually less than 1%. The low yield is a result of the starting material or product decomposing at high temperatures. We envisioned another route to synthesize 2-amino-3-thiophenecarboxylic acid starting from 2-methylthiophene, shown in Scheme 2.3 below. Blocking the C5 position with another group that could be removed later could potentially give the desired nitro group at the C3 position.

Bromine was a selected choice because it could be removed in the final step under

hydrogentation conditions [19].

Scheme 2.3: Proposed alternate synthesis of 2-amino-3-thiophenecarboxylic acid
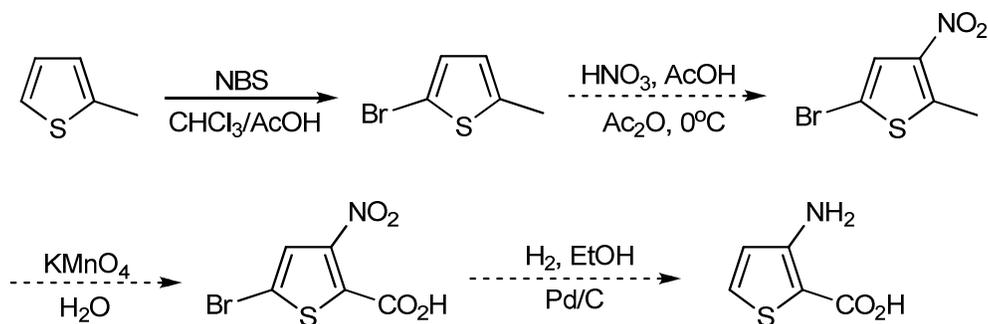
The selective bromination of the C5 position of thiophene was efficient with good yields. At the

nitration step, we encountered the major stumbling block. We determined, from NMR and TLC

analysis, that multiple products were formed. The nitration reaction was not selective. It

occurred not only at the C3 and C4 position of the thiophene ring system, but it also resulted in

substitution of the bromine.

Another route for the synthesis of 2-amino-3-thiophenecarboxylic acid proved to be more

successful. Starting with methyl thioglycolate and 2-chloroacrylonitrile, a variation of the

Gewald thiophene synthesis was used, and sodium methoxide in methanol gave a yellow oil in

good yields [20].

Scheme 2.4: Synthesis of methyl 3-aminothiophene-2-carboxylate
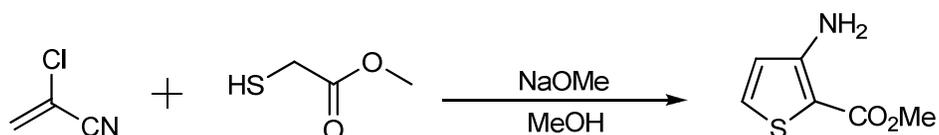
To finalize the synthesis of 2-amino-3-thiophenecarboxylic acid, the methyl ester needed to be hydrolyzed to the carboxylic acid. This was accomplished using both acidic and basic conditions. We first used 6N HCl and 1,4 dioxane to remove the methyl ester, which gave a yield of 40% [21]. Using basic conditions, KOH and H_2O/MeOH, the hydrolysis was more efficient. After adjusting the pH to 5.6, the desired product was obtained with a better yield of 63%, shown in Scheme 2.5 [22].

Scheme 2.5: Acid and base hydrolysis of methyl 3-aminothiophene-2-carboxylate
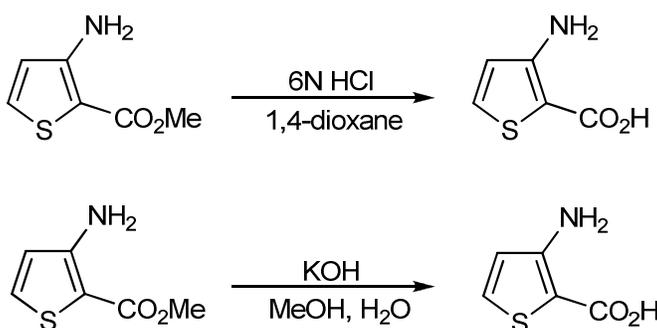
The cell line used for the incorporation experiments is W3310 tna A_2 Δ trp E_5, express tyrosine phenol lyase (TPL), which contains only one tryptophan residue. First, cell cultures are grown overnight on agar plates, and from these a single colony is transferred to a 15 mL BIOLGIX tube to grow 5 mL cultures. A suitable minimal medium was to be determined for growth. Minimal MOPS medium was the first to be used and proved to be too rich in nutrients for the cells because both the positive and negative controls grew very well [23]. Another minimal medium explored was the Vogel-Bonner minimal media which proved to be stringent and gave good results for the positive and negative controls [24]. The toxicity of the AA analogues was tested by introducing the cells with a carbon source and nutrients to determine if growth is occurring and the AA analogues showed no growth. This required the cells to be weaned to be able to

incorporate the AA analogues. Each growth contained 100 µL AA analogue (0.01 mg/mL), 100

µL of the 10% carbon source solution, 100 µL of a cell stock solution (5 mL), and the amount of

AA (0.1 mg/mL) was lowered from each growth from 100 µL to 1 µL. These growths were

diluted to 5 mL with the VB minimal media. Attempts that were made to lower the amount of

AA to go below 1 µL (0.1 mg/mL) showed no growth for the AA analogues. We concluded that

1 µL of AA (0.1 mg/mL) was necessary for the incorporation of the AA analogues. After

weaning the cells with the AA analogues, the cells were transformed with a plasmid and grown

on ampicillin agar plates [25]. These cells were then grown in cultures consisting of 100 µL AA

analogue, 100 µL of the carbon source, 100 µL of a cell stock solution, 100 µL ampicillin

solution (5 mg/mL), 1 µL AA, and diluted to 5 mL with VB minimal media. The cell cultures

were then grown on a larger scale and the protein from each growth was purified and tested for

activity. Using glucose as a carbon source, cultures showed TPL activity only for the 2-amino-3-

thiophenecarboxylic acid compound. For 2-amino-3-thiophenecarboxylic acid, the carbon

sources were expanded to growing with casein, glycerol, lactic acid, and succinic acid due to

having little to no activity using glucose. TPL activity was noticed using glycerol but the

activity was not robust. Further testing revealed that the absorbance and fluorescence of the

purified protein did not match what was expected of the thienopyrroles but did match tryptophan.

Experimental

General procedures: ^{1}H-NMR spectra were recorded on a Bruker 400 MHz spectrometer at

room temperature in the required deuterated solvent using tetramethylsilane as the internal

standard. Column chromatographies were carried out on standard grade silica gel (60Å, 32-63

µm) in the required solvent system. Thin layer chromatographies were carried out on Whatman

Al/Sil G/UV plates (250 µm) in the required solvent system. Solvent removal was performed using a Buchi Rotavapor R-114 and Waterbath B-480. A Parr shaker hydrogenation apparatus was used for hydrogenation reactions. Pure water for cell growth and protein purification experiments were obtained from Milli-Q Plus. Solutions were sterilized using a Steris Amsco Century SG-120 Scientific Gravity Sterilizer. Cell growth experiments were carried out in a LABCONCO Tissue culture enclosure UV/Vis to minimize contamination. Cell cultures were grown in a Type 37900 Culture Incubator. Cell growth experiments were grown using a Lab-Line shaker bath. Centrifuging was carried out using IEC B-20A Centrifuge Damon/IEC Division and Marathon micro A Fisher Scientific centrifuge. Protein activity was determined using the Cary 1E UV-Visible spectrophotometer. For protein purification a Fisher Scientific Isotemp Constant Temperature Circulator Model 800, Branson Sonifier 450, and BIO-RAD Model 2110 Fraction Collector was used for experiments. The Olis RSM-1000 Rapid-Scanning Monochromator was used for determining absorbance.

3-Methyl-2-nitrothiophene: In a 100 mL round-bottom flask, 30 mL of Ac_2O and 20 mL of AcOH is added and chilled in an ice-salt bath while stirring. Then 2.5 mL of nitric acid is added dropwise to the solution and stirring is continued for 5 minutes. Then, 0.5 g (5 mmol) of 3-methylthiophene is added dropwise over a period of 5 minutes. The reaction is allowed to stir for 45 minutes in the ice salt bath. Ice water is added to quench the reaction and an orange-yellow solid precipitates. The solid is filtered and collected to give 0.64 g (86%). Mp 61-62 ^{o}C. Lit mp 62-63 ^{o}C [26]. ^{1}H NMR ($CDCl_3$, ppm, 400 MHz), δ: 2.61 (s, 3H), 7.14 (d, 1H, J = 4.4), 7.81 (d, 1H, J = 4.4)

2-Nitro-3-thiophenecarboxylic acid: In a 125 mL Erlenmeyer flask 0.5 g (3.5 mmol) of 3-methyl-2-nitrothiophene, 2.0 g of $MgSO_4 \cdot 7H_2O$, and water and heated. In another 125 mL, Erlenmeyer flask 1.7 g of $KMnO_4$ is dissolved in water while heating. The $KMnO_4$ solution is added dropwise to the 3-methyl-2-nitrothiophene solution while heating, generating a brown precipitate. The mixture was gravity filtered to give a clear solution and the solvent was removed *in vacuo* to give a white solid 0.41 g (68%). Mp 155-158 ^{o}C. Lit mp 155-156 ^{o}C [27]. ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 6.81 (d, 1H, J = 4.3), 7.33 (d, 1h, J = 4.3).

2-Amino-3-thiophenecarboxylic acid: In a hydrogenation glass vessel, 0.3 g (1.7 mmol) of 2-nitro-3-thiophenecarboxylic acid, 30 mg of 10% Pd/C, and 15 mL of EtOH was mixed. The mixture was hydrogenated overnight at room temperature. The next day the reaction mixture was filtered over Celite and the solvent was removed *in vacuo* to give a tan solid, 0.2 g (81%). Mp 59-61 ^{o}C. ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 7.54 (d, 1H, J = 4.5), 7.76 (d, 1H, J = 4.5).

5-Methyl-4-nitrothiophene-2-carboxylic acid: In a100 mL round-bottom flask, 30 mL of Ac$_2$O and 20 mL of AcOH is added, and chilled in an ice-salt bath while stirring. Then, 3.0 mL of nitric acid is added dropwise to the solution and stirring is continued for 5 minutes. Then 0.5 g (3.5 mmol) of 3-methyl-2-thiophenecarboxylic acid is added in increments. The reaction is allowed to stir for 45 minutes in the ice salt bath. Ice water is added to quench the reaction and an orange-yellow solid precipitates. The solid is filtered and collected to give 0.58 g (88%). Mp 178-182 ^{o}C. Lit mp 180-181 ^{o}C [16]. ^{1}H NMR (acetone-d$_6$, ppm, 400 MHz), δ: 2.11 (s, 3H), 8.16 (s, 1H).

2-Methyl-3-nitrothiophene: A sublimation apparatus is setup and 1.0 g (5.3 mmol) 5-methyl-4-nitrothiophene-2-carboxylic acid is added and heated. White needles formed and are collected to

give 0.05 g (7%). Mp 43-47 $^{\circ}$C. Lit mp 44-46 $^{\circ}$C [16]. ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ:

2.34 (s, 3H), 7.64 (d, 1H, J = 5.3), 8.21 (d, 1H, J = 5.3).

2-Bromo-5-methylthiophene: In a 50 mL round-bottom flask 30 mL of chloroform/acetic acid

(1:1), 2.0 g NBS, and 1.0 g (10 mmol) 2-methylthiophene were mixed and the reaction was

refluxed in the absence of light. The mixture was diluted with water and extracted. The solvent

was removed in vacuo to give an oil 1.03 g (57%). ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 2.41 (s,

3H), 6.53 (d, 1H, J = 3.9), 6.85 (d, 1H, J = 3.9).

Methyl 3-aminothiophene-2-carboxylate: In a 50 mL round-bottom flask, 1.0 g (43 mmol) of

sodium metal is dissolved in 40 mL of MeOH and then chilled in an ice-salt bath. Then, 0.75 g

(7.1 mmol) of methyl thioglycolate was added dropwise and allowed to stir for 10 minutes.

Next, 0.85 g (9.7 mmol) of 2-chloroacrylonitrile in a solution of MeOH was added dropwise

while the reaction is being chilled. The reaction was allowed to stir for 5 minutes in an ice-salt

bath. The reaction was then allowed to stir at room temperature for 1 hour. The reaction was

extracted with diethyl ether, the solvent was dried over Na$_2$SO$_4$, and the solvent was removed *in*

vacuo to give yellow oil. The oil was crystallized from petroleum ether to give a yellow solid,

0.9 g (81%). Mp = 65-68 $^{\circ}$C. Lit mp = 64-66 $^{\circ}$C [28]. ^{1}H NMR (acetone-d$_6$, ppm, 400 MHz), δ:

3.41 (s, 3H), 5.47 (s, 2H), 6.54 (d, 1H, J = 5.2 Hz), 7.45 (d, 1H, J = 5.2 Hz).

3-Amino-2-thiophenecarboxylic acid (A): In a 50 mL round-bottom flask, 0.3 g (1.9 mmol) of

methyl 3-aminothiophene-2-carboxylate, 15 mL 6N HCl, and 20 mL 1,4-dioxane was added and

refluxed for 3 hours. The solvent was removed in vacuo to give a dark tan solid as the HCl salt,

0.11 g (40%). Mp = 79-83 $^{\circ}$C. Lit mp = 88-89 $^{\circ}$C [27]. ^{1}H NMR (D$_2$O, ppm, 400 MHz), δ: 6.59

(d, 1H, J = 4.2), 7.39 (d, 1H, J = 4.2).

3-Amino-2-thiophenecarboxylic acid (B): In a 50 mL round-bottom flask, 0.4 g (2.5 mmol) of methyl 3-aminothiophene-2-carboxylate was added to a solution of KOH solution in MeOH/H_2O. The mixture was allowed to reflux overnight. The mixture was extracted with diethyl ether. The pH of the aqueous layer was adjusted to 5.6 to precipitate a brown solid, 0.23 g (63%). Mp = 80-85 ^{o}C. Lit mp = 88-89 ^{o}C [29]. ^{1}H NMR (D_2O, ppm, 400 MHz), δ: 6.58 (d, 1H, J = 4.2), 7.21 (d, 1H, J = 4.2)

References

1. *Amino Acids Biosynthesis and Genetic Regulation*, Herrmann, K. M.; Somerville, R. L., Ed.; Biotechnology; Addison Wesley Publishing Company: Massachusetts, **1983**.

2. (a) Taylor, G. *Acta Cryst*. **2003**, *D59*, 1881-1890. (b) Uson, I.; Schmidt, B.; von Bulow, R.; Grimme, S.; von Figura, K.; Dauter, M.; Rajashankar, K. R.; Dauter, Z.; Sheldrick. G. M. *Acta Cryst*. **2003**, *D59*, 57-66. (c) Weiss, M. S.; Mander, G.; Hedderich, R.; Diederichs, K.; Ermler, U.; Warkentin, E. *Acta Cryst*. **2004**, *D60*, 686-695.

3. Bae, J. H.; Alefelder, S.; Kaiser, J. T.; Friedrich, R.; Moroder, L.; Huber, R.; Budisa, N. *J. Mol. Biol*. **2001**, *309*, 925-936.

4. Senear, D. F.; Mendelson, R. A.; Stone, D. B.; Luck, L. A.; Rusinova, E.; Ross, J. B. A. *Anal. Biochem*. **2002**, *300*, 77-86.

5. Bacher, J. M.; Ellington, A. D. *J. Bacteriol*. **2001**, *183*, 5414-5425.

6. Katragadda, M.; Lambris, J. D. *Protein Expression and Purification*. **2006**, *47*, 289-295.

7. Dougherty, D. A. *Curr. Opin. Chem. Biol*. **2000**, *4*, 645-652

8. Minks, C.; Alefelder, S.; Moroder, L.; Huber, R.; Budisa, N. *Tetrahedron*. **2000**, *56*, 9431-9442.

9. Welch, M.; Phillips, R. S. *Bio. Med. Chem. Letts*. **1999**, *9*, 637-640.

10. Phillips, R. S.; Cohen, L. A.; Annby, U.; Wensbo, D.; Gronowitz, S. *Bio. Med. Chem. Letts*. **1995**, *5*, 1134-1135.

11. (a) Sloan, M. J.; Phillips, R. S. *Bio. Med. Chem. Letts*. **1992**, *2*, 1053-1056. (b) Lee, M.; Phillips, R. S. *Bio. Med. Chem. Letts*. **1992**, *2*, 1563-1564. (c) Wilcox, M. *Anal. Biochem*. **1974**, *59*, 436-440.

12. Budisa, N.; Minks, C.; Medrano, F. J.; Lutz, J.; Huber, R.; and Moroder, L. *Proc. Natl. Acad. Sci.* **1998**, *95*, 455–459.

13. Budisa, N.; Alefelder, S.; Bae, J. H.; Golbik, R.; Minks, C.; Huber, R.; Moroder, L. *Protein Science*, **2001**, *10*, 1281-1292.

14. Gronowitz, S.; Ander, I. *Chemica Scripta*. **1980**, *15*, 20-22.

15. Wilcox, C. F.; Wilcox, M. F. *Experimental Organic Chemistry A Small-Scale Approach*. Prentice-Hall: New Jersey, **1995**.

16. Snyder, H. R.; Carpino, L. A.; Zack, J. F.; Mills, J. F. *J. Am. Chem. Soc*. **1957**, *79*, 2556-2559.

17. Welch, M. W. M.S. Thesis, University of Georgia of Athens, **1998.**

18. Jones, G. B.; Chapman, B. J. *J. Org. Chem*. **1993**, *58*, 5558-5559.

19. Luo, Q.; Sheng, S.; Cheng, S.; Tian, H. *Aust. J. Chem*. **2005**, *58*, 321-326.

20. Huddleston, P. R; Barker, J. M. *Synth. Commun*. **1979**, *9*, 731-734.

21. Phillips, R. S.; Cohen, L. A. *J. Heterocyc. Chem*. **1988**, *25*, 191-192.

22. Elliott, R. L.; O'Hanlon, P. J.; Rogers, N. H. *Tetrahedron*. **1987**, *43*, 3295-3302.

23. Neidhardt, F. C.; Bloch, P. L.; Smith, D. F. *J. Bacteriol*. **1974**, *119*, 736-747.

24. Vogel, H. J.; Bonner, D. M. *J. Biol. Chem*. **1956**, *218*, 97-98.

25. Chen, H.; Gollnick, P.; Phillips, R. S. *Eur. J. Biochem*. **1995**, *229*, 540-549.

26. Ballini, R.; Bartoli, G.; Bosco, M.; Dalpozzo, R.; Marcantoni, E. *Tetrahedron*. **1988**, *44*, 6435-6440.

27. Raich, W. J.; Hamilton, C. S. *J. Am. Chem. Soc*. **1957**, *79*, 3800-3804.

28. Gronowitz, S.; Konar, A.; Hornfeldt, A. –B. *Chemica Scripta*. **1976**, *10*, 159-162.

29. Ogawva, K.; Yamawaki, I.; Matsusita, YI.; Nomura, N.; Kador, PF.; Kinoshita, JH. *Eur. J. Med. Chem*. **1993**, *28*, 769-781.

SUPPORTING DATA
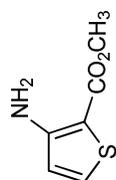
^{1}H-NMR Methyl 3-aminothiophene-2-carboxylate
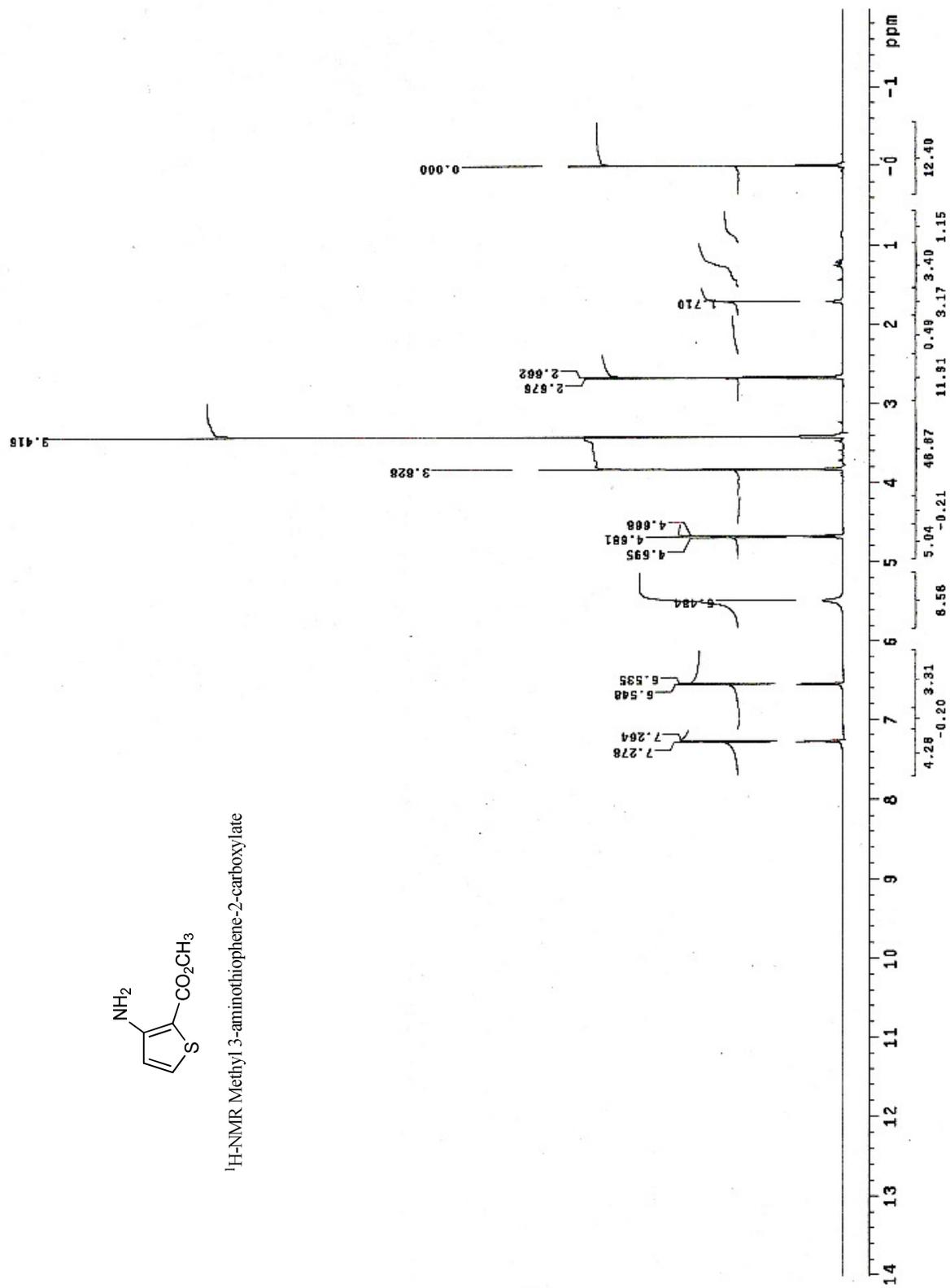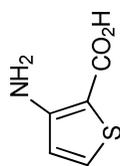
¹H-NMR 3-Amino-2-thiophenecarboxylic acid
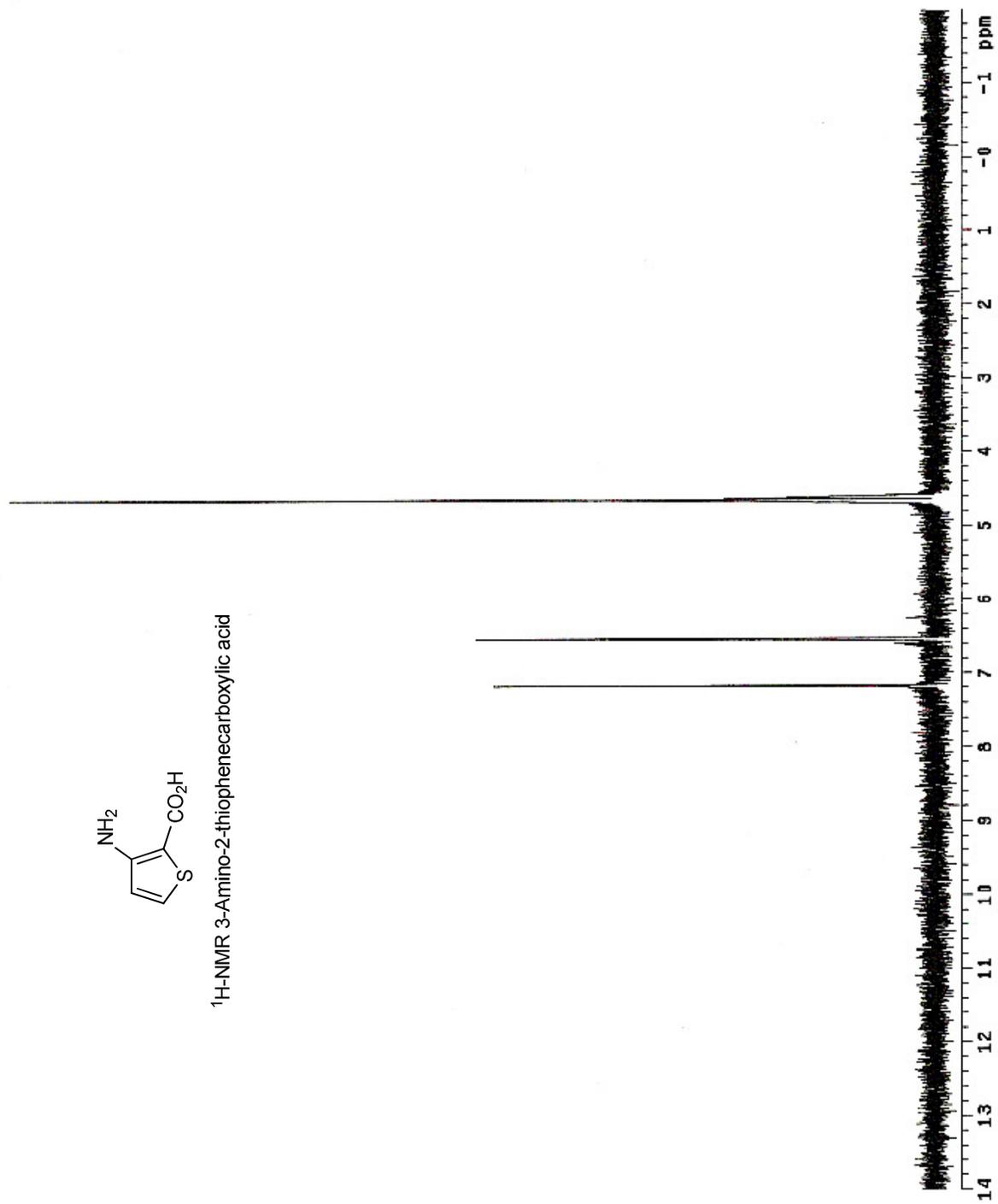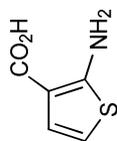
^{1}H-NMR 2-Amino-3-thiophenecarboxylic acid
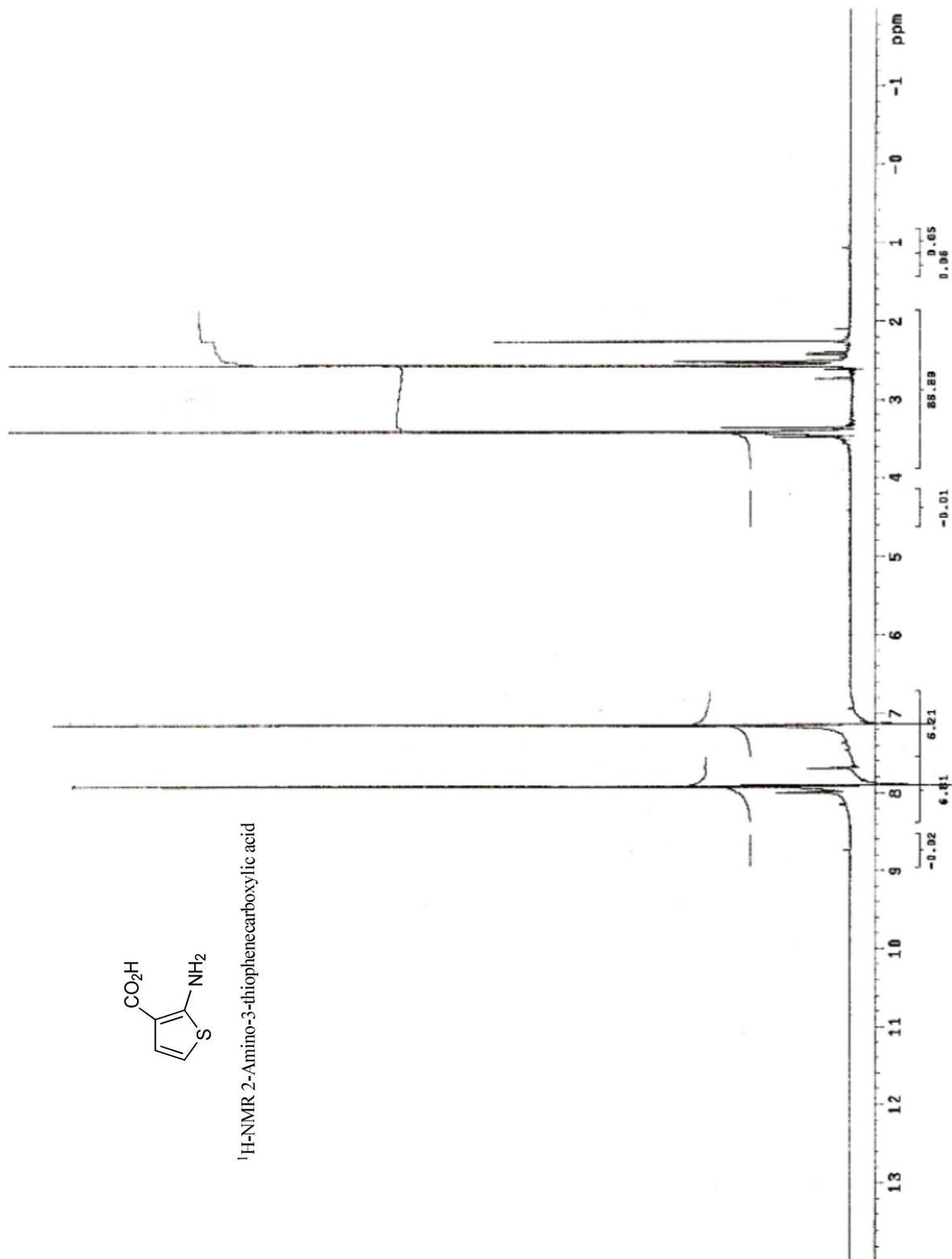
CHAPTER 3

SYNTHESIS AND SELECTIVE NITRATION OF TRYPTOPHAN ANALOGUES AND RESOLUTION

Introduction

Amino acids are indispensable in synthesis because they are used as inexpensive members of a chiral pool, and many chemical transformations of amino acids have been reported [1]. Tryptophan is an important molecule in a variety of ways such, as being a precursor for important metabolites including serotonin and melatonin. Tryptophan is a building block for proteins and nonribosomal peptides and it plays roles in protein function. The indole group, specifically, is also important due to its amphiphilic and electrostatic properties [2]. It has been reported that, in toxic peptides and peptide derived natural products from marine organisms, bromotryptophans are common components [3]. Bromotryptophan derivatives are shown in Figure 3.1. Other halogenated derivatives of tryptophan have a significantly modified bioactivity and bioavailability and are used as precursors for *in vitro* synthesis. They also serve as substrates for selective functionalization through coupling reactions [4]. Tryptophan derivatives also have luminescence properties. They are used as probes in systems for the determination of protein environment and conformational dynamics by measuring inter- and intramolecular distances in fluorescence resonance energy transfer experiments and dielectric-dependent changes in emission spectra [5]. It has been shown that the metabolism of L-tryptophan by indoleamine-2,3-

dioxygenasae protects the fetus from being rejected by the mother and regulates the human

immune responses [6]. The synthesis of tryptophan analogues that have a wide range of

functionality and electronic properties is highly desirable.
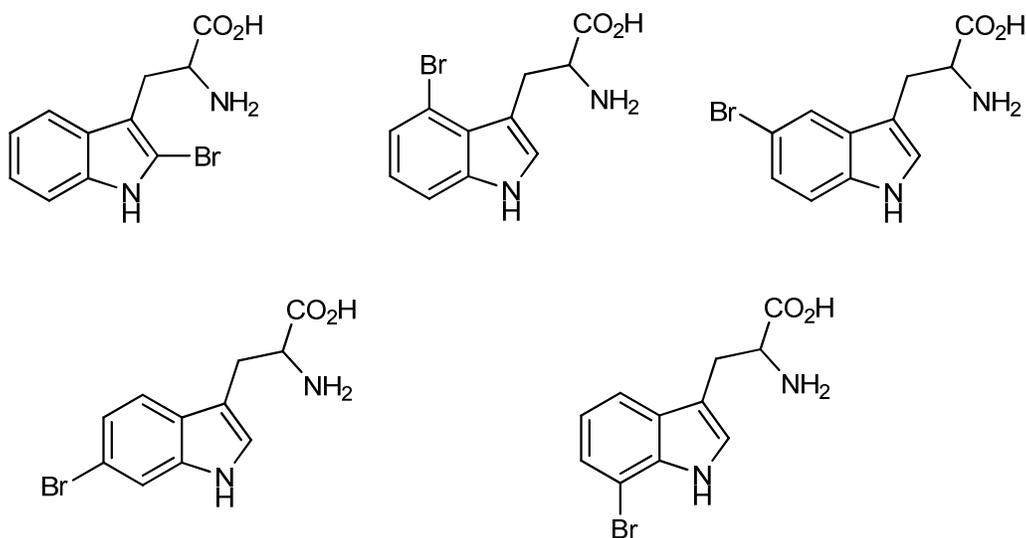
Figure 3.1: Bromotryptophan analogues

Tryptophan has been synthesized using chemical and enzymatic reactions such as using

tryptophan synthase [7]. In that enzymatic reaction, the enzyme uses indole-3-glycerol

phosphate and L-serine to produce L-tryptophan in two reactions. The restriction with this

method is the accessibility of the enzyme's active site because the indole must pass through the

α-subunit then passes through a 2.5 nm tunnel before reaching the active site in the β-subunit

[8]. Chemical synthesis involves electrophilic substitution reactions at the C3 position of indole,

where it is favored for resonance and electrostatic reasons [9]. Using serine as the electrophile

and favorable reaction conditions will result in the formation of tryptophan analogues. The

electrophilic addition of indole is shown below in Figure 3.1, where the cation formed by the

attack on the 3 position can be delocalized without involving the benzenoid [9]. By blocking the

3 position, reactions can be carried out to synthesize tryptophan derivatives selectively. A

variety of tryptophan derivatives have been synthesized utilizing enzymatic and chemical

strategies [7]. Functionalizing tryptophan with a nitro group is of interest because nitro groups

serve as intermediates for other functional groups through conversion to a diazonium and further

chemical transformation to tryptophan analogues [10]. Determining if racemization is occurring

is important to ensure that a pure product is obtained from the multistep reactions in acidic

conditions.
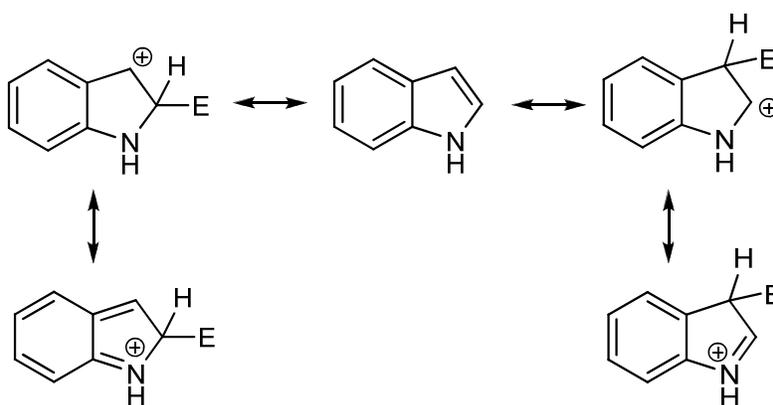
Figure 3.2: Electrophilic addition at the C2 and C3 positions

Proposed Plan of Research

The selective nitration of tryptophan requires protection of the free amine and carboxylic acid

groups. The nitration occurs directly at the C6 position and the desire to nitrate the C2 position

is of interest. The carboxylic acid is converted into an ester [11] and the free amine as well as

the indole nitrogen is protected with TFA protecting groups [10]. The protections are done in a

two-step synthesis shown below in Scheme 3.1.

Scheme 3.1: Synthesis of TFA-protected tryptophan
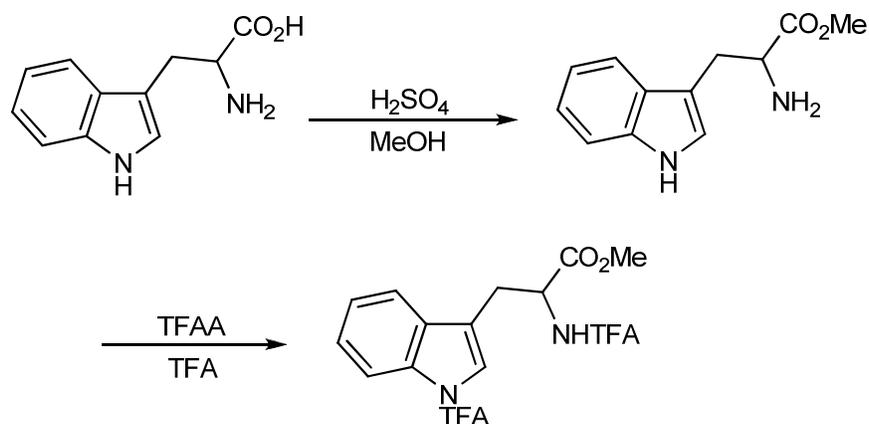
The selective nitration of the TFA-protected tryptophan is dependent on the conditions to obtain the desired nitrated tryptophan analogue. This is illustrated below in Scheme 3.2. With nitric acid, acetic anhydride, and acetic acid, you observe the nitration at the C2 position and when using nitric acid and trifluoroacetic acid you obtain the nitro group at the C6 position predominantly.

Scheme 3.2: Regioselective nitration of TFA-protected tryptophan
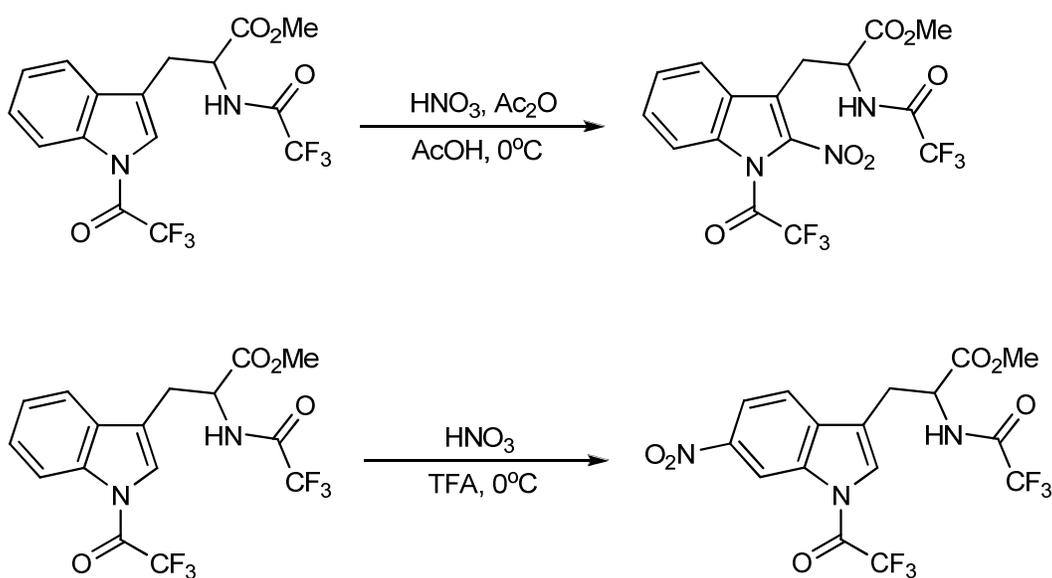
The Phillips group has reported the nitration of the 2-nitro and 6-nitrotryptophan [12]. To

determine if racemization occurred for the nitrated TFA-protected tryptophan, an option would

be to synthesize diastereomers of the C2- and C6-nitrated isomers. This is done by converting

the methyl ester into a better leaving group and then converting it into the menthol ester with

newly generated stereogenic centers [13]. NMR or GC analysis will then determine the isomers.

This is shown below in Scheme 3.3 for each isomer.

Scheme 3.3: Synthesis of diasteromeric tryptophan isomers using menthol

Cmpd 1 = R_1 = NO_2, R_2 = H
Cmpd 2 = R_1 = H, R_2 = NO_2
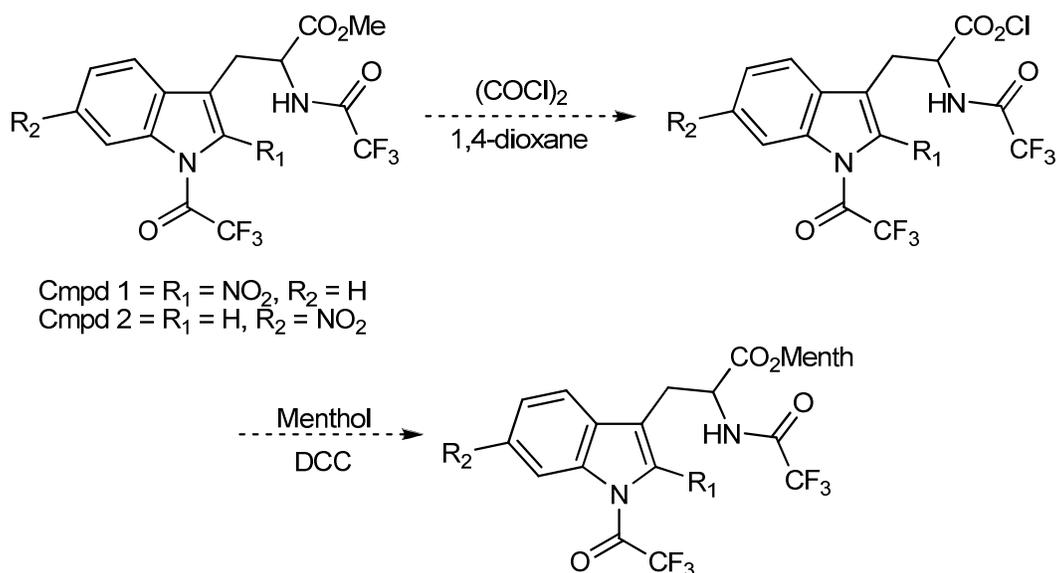
The synthesis of 2-fluorotryptophan starts by converting the nitro group of *bis*-TFA protected 2-

nitrotryptophan into an aryl diazonium fluoroborate [14]. Due to the instability of the amino

group at the C2 position, we decided to try this reaction as a one-pot reaction [23]. We first

reduce the nitro group into an amine, but this is not isolated due to the instability of the 2-

aminoindoles [23]. We then use the Schiemann reaction by turning the amine into a diazonium

fluoroborate, followed by heating in fluoroboric acid. This reaction should give the 2-

fluorotryptophan, which is shown in Scheme 3.4 below [15].

Scheme 3.4: Synthesis of 2-fluorotryptophan
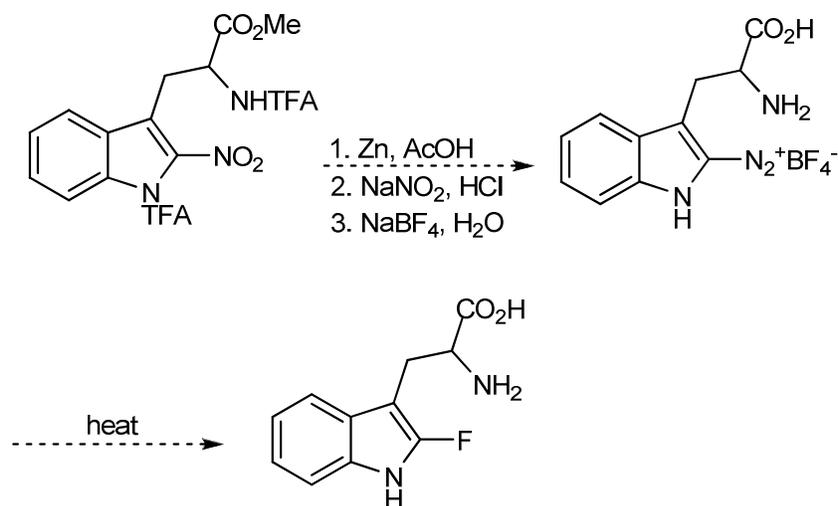
To obtain the D isomers of tryptophan derivatives, we begin by reacting the indole derivatives

with DL-serine, acetic anhydride, and acetic acid to form the DL isomers of the tryptophan

derivatives. The D isomer can then be isolated by resolution using acylase. This is shown in

Scheme 3.5 below [16, 17].

Scheme 3.5: Synthesis of D-tryptophan from indole derivatives
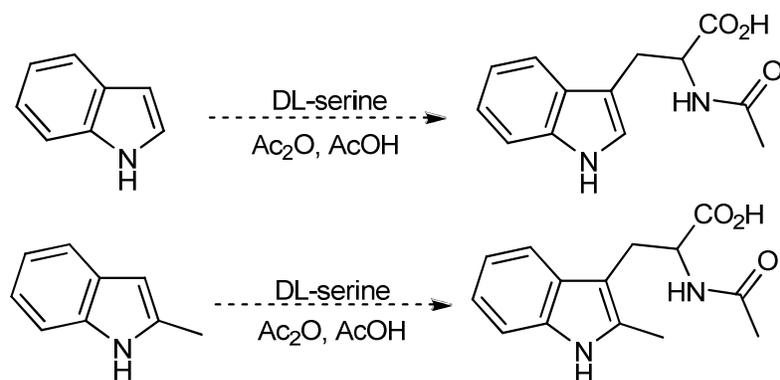
Results and Discussion:

The Fischer esterification of tryptophan in sulfuric acid and methanol gave low yields. It has

been reported that tryptophan decomposes in strongly acidic conditions, which may account for

the low yields that were observed, < 50% [18]. An alternative route for the synthesis of the

methyl ester tryptophan is shown below in Scheme 3.6 using thionyl chloride and methanol at

0°C [19]. The esterification using sulfuric acid gave the free base, which has to be carried over

into the next reaction to avoid decomposition. The thionyl chloride esterification gave the

chloride salt, then adjusting the pH gave a white solid, which is stable and can be stored for

future use. The reaction also resulted in a better yield of 85%.

Scheme 3.6: Alternate synthesis of tryptophan methyl ester
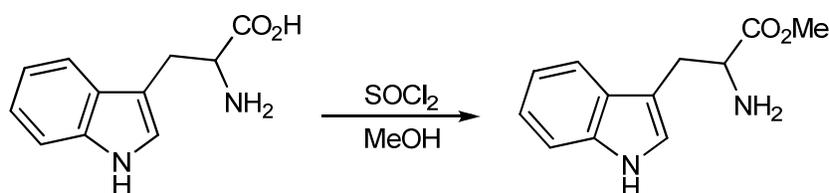
The selective nitration required conditions that played an important role in determining which

isomer formed predominantly. The 2-nitrotryptophan formed predominantly when using acetic

anhydride and acetic acid. The 6-nitrotryptophan formed predominantly when using

trifluoroacetic acid, but the order of the addition of reagents played a role in determining which

isomer formed. When the nitric acid and trifluoroacetic acid were mixed and the protected

tryptophan was added last, the 2-nitrotryptophan formed as the major product. When the

protected tryptophan was dissolved in trifluoroacetic acid and the nitric acid was added last, the

6-nitrotryptophan formed as the major product. This suggests there is a different nitrating

species for each reaction, NO_2^+ for the acetic anhydride and acetic acid and $CF_3CO_2NO_2$ for the

trifluoroacetic acid. The order of addition is important to form each isomer.

Table 3.1: Nitration experiments with N_α,N_1-bis-(trifluoroacetyl)-L-tryptophan methyl ester
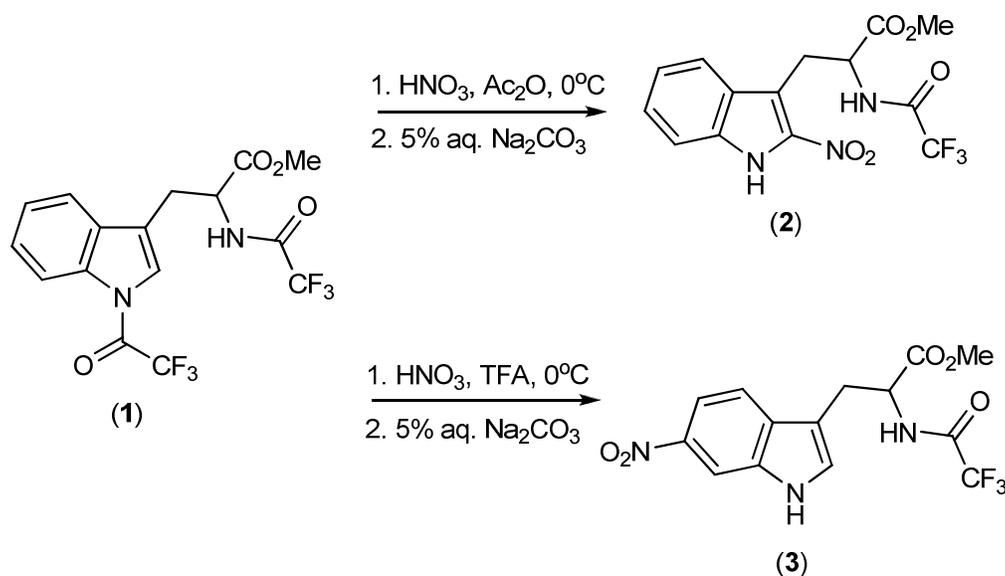
Experiment	Temp.	Solvent	Time	(1), %	(2) %	(3), %	other, %
1	25°C	AcOH	24 hours	58	0	0	42
2	25°C	Ac₂O	15 min	40	33	15.4	11.6
3	25°C	TFA	15 min	0	11.8	66.8	21.4
4	25°C	TFAA	15 min	0	4	8	88
5	0°C	AcOH	24 hours	45	2	3	50
6	0°C	Ac₂O	1 hour	0	66.8	14.4	18.8
7	0°C	TFA	1 hour	0	4	68.8	27.2
8	0°C	TFAA	4 hours	9	7	10	74
9	60°C	Ac₂O	4 hours	0	6	12	82
10	60°C	TFA	4 hours	0	0	10	90

Table 3.1 are the nitration experiments done by Andrew Osborne. The experiments are

compared based on the solvent, length of time, and temperature to determine which set of

conditions are optimal for nitration at the C2 and C6 position.

The synthesis of diasteromeric isomers of the protected nitrotryptophan was done to determine if the compound was undergoing racemization in acidic conditions. The synthesis of the menthol ester was inconclusive using NMR spectra to determine if racemization occurred. Then an alternative route was to synthesize isomers using *S*-methylbenzylamine [20]. Using NMR spectra and GC analysis was also inconclusive in determining if racemization occurred. This is shown in Scheme 3.7 below.

Scheme 3.7: Alternate synthesis of diasteromeric isomers of protected nitrotryptophan

Cmpd 1: $R_1 = NO_2$, $R_2 = H$
Cmpd 2: $R_1 = H$, $R_2 = NO_2$

We decided to use a chiral column for the HPLC, using $CuSO_4$ solution, to determine if racemization did occur [21]. Then we deprotected the protected nitrotryptophan isomers shown below in Scheme 3.8 using 6N HCl and 1,4-dioxane [12].

Scheme 3.8: Deprotection of protected nitrotryptophan isomers

Cmpd 1: $R_1 = NO_2$, $R_2 = H$
Cmpd 2: $R_1 = H$, $R_2 = NO_2$

The deprotection was efficient and using the chiral column on the HPLC showed that synthesis of the L isomers showed no racemization. This was confirmed by doing the same reactions on the DL and L isomers. The HPLC analysis of the DL isomers showed two well defined peaks that separated in comparison to the L isomers that contained only one peak. With no detectable amount of the nitrotryptophans present it can be concluded that the products are greater than 98% homochiral.

The problems encountered during the synthesis of 2-fluorotryptophan were due to the instability of the 2-aminotryptophan when performing the reduction using zinc and acetic acid. To circumvent this, the reaction was carried out using stannous chloride and HCl, as shown in Scheme 3.9 below [22]. The reaction was carried out on the *bis*-TFA protected 6-nitro-L-tryptophan with good results. Using the *bis*-TFA protected tryptophan was to help with the stability of the 2-aminotryptophan formation. The reaction can be carried forward to do the Schiemann reaction using fluoroboric acid. Then these reaction conditions can be used to synthesize the 2-fluorotryptophan.

Scheme 3.9: Alternate synthesis of tryptophan diazonium fluoroborate

The synthesis of the DL-2-methyltryptophan is a one-step synthesis but the previously reported procedures used to make analogues of the desired compound gave us mixed results based on NMR data. The procedure was altered to the following in order to obtain the desired product. The reaction was refluxed for three hours under nitrogen and then extracted with ethyl acetate, the solvent was washed with water, and the solvent was removed to give an oil, which was the desired product. To isolate the D isomer from the racemic mixture the next reaction with acylase has to be carried out.

Experimental

General procedures: ^{1}H-NMR and ^{19}F spectra were recorded on a Bruker 400 MHz spectrometer at room temperature in the required deuterated solvent using tetramethylsilane as the internal standard. Column chromatographies were carried out on standard grade silica gel (60Å, 32-63 µm) in the required solvent system. Thin layer chromatographies were carried out on Whatman Al/Sil G/UV plates (250 µm) in the required solvent system. Solvent removal was performed using a Buchi Rotavapor R-114 and Waterbath B-480. Resolution was determined using a SpectraSYSTEM P2000 HPLC with a 4x250 mm Chiral ProCu=Si100 Polyol 3-hydroxy-D-proline colum (Serva). Pure water for HPLC runs were obtained from a Milli-Q Plus.

L-Tryptophan methyl ester (A): In a 250 mL round-bottom flask 5.0 g (25 mmol) of L-tryptophan is added to a solution of 125 mL of MeOH. Then, 2.5 mL of H_2SO_4 is added dropwise to the solution and the tryptophan dissolves in the solution. The reaction is allowed to stir overnight at room temperature. The solvent is removed in vacuo to give an oil. The oil is dissolved in ethyl acetate and diluted with H_2O. $NaHCO_3$ is added in portions until no more

liberation of gas is noticed. The mixture is extracted and the organic layer is removed in vacuo to give the free base as an oil, 2.1 g (43%). ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 3.19 (m, 2H, CH$_2$), 3.70 (s, 3 H, OCH$_3$), 3.86 (m, 1 H, CH), 7.07 (s, 1H, CH), 7.35 (m, 4H, C$_{4-7}$-H), 8.19 ppm (s, 1H, indole NH).

L-Tryptophan methyl ester (B): In a 250 mL round-bottom flask 6.0 g (50 mmol) of thionyl chloride is added dropwise to a solution of 100 mL of MeOH and stirred in an ice-salt bath for 30 minutes. Then 5.0 g (25 mmol) of L-tryptophan was added in portions to the solution and stirred while cooling for 10 minutes. The ice-salt bath was removed and the reaction was allowed to stir at room temperature for 5 hours. Then the mixture was refluxed for 15 minutes and the solvent was removed *in vacuo* to give a white solid. The solid was recrystallized from ethanol to give a white powder, 4.5 g (85%). Mp 91-94. Lit mp 90-91 [19]. ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 3.10 (m, 2H, CH$_2$), 3.65 (s, 3 H, OCH$_3$), 3.79 (m, 1 H, CH), 7.11 (s, 1H, CH), 7.42 (m, 4H, C$_{4-7}$-H), 8.36 ppm (s, 1H, indole NH).

N$_α$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester: In a 50 mL round-bottom flask 2.5 g (12 mmol) of tryptophan methyl ester is dissolved in 15 mL of trifluoroacetic acid. Then 5.0 g (24 mmol) of trifluoroacetic anhydride is added dropwise and the solution is allowed to stir overnight at room temperature. The mixture is quenched with H$_2$O and a grey solid precipitates. The solid is washed with diethyl ether to give a pure white solid 3.1 g (66%). Mp 165-166 ^{o}C. ^{1}H NMR (acetone-d$_6$, ppm, 400 MHz), δ: 3.38 (1H, dd, J$_{ββ'}$ = 14.9 Hz, J$_{αβ'}$ = 9.4 Hz, β'-H), 3.50 (1H, dd, J$_{ββ'}$ = 14.9 Hz, J$_{αβ}$ = 4.70 Hz, β-H), 3.75 (3H, s, -OCH$_3$), 4.95 (1H, dd, J$_α$ = 9.4 Hz, J$_{αβ}$ = 4.7 Hz, α-H), 7.44 (1H, t, J = 7.1 Hz, C$_5$-H), 7.49 (1H, t, J = 7.49 Hz, C$_6$-H), 7.67 (1H, d, J = 7.0 Hz, C$_2$-H), 7.77 (1H, d, J = 7.0 Hz, C$_4$-H), 8.37 (1H, d, J = 7.1 Hz, C$_7$-H), 8.95 (1H, br s, indole NH). ^{19}F NMR (acetone-d$_6$, ppm, 400 MHz), δ: -70.76 (indole COCF$_3$), -76.74 (α-COCF$_3$).

2-Nitro-N$_\alpha$-trifluoroacetyl-L-tryptophan methyl ester: In a 50 mL round-bottom flask, 20 mL of acetic anhydride and 15 mL of acetic acid is added and chilled in an ice-salt bath for 10 minutes. 4.0 mL of nitric acid is added dropwise to the solution and allowed to stir for 5 minutes. Then 1.0 g (2.4 mmol) of N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester is added and allowed to stir chilled for 45 minutes. The reaction is diluted with ice-water and a yellow-orange precipitate forms. The precipitate is filtered and collected to give 0.52 g (47%). Mp 199-202 $^\circ$C. Lit mp 201-203 $^\circ$C [12]. ^{1}H NMR (acetone-d$_6$, ppm, 400 MHz), δ: 3.70 (1H,dd, J$_{\beta\beta'}$ = 21.9, J$_{\alpha\beta}$ = 12.2 Hz, β-H), 3.72 (3H, s, OCH$_3$), 3.92 (1H, dd, J$_{\beta\beta'}$ = 21.9, J$_{\alpha\beta'}$ = 9.4 Hz, β'-H), 5.00 (1H, dd, J = 9.4, 12.2 Hz, α-H), 7.22 (1H, t, J = 8.1 Hz, C$_6$-H), 7.46 (1H, t, J = 7.8 Hz, C$_5$-H), 7.52 (1H, dd, J$_{45}$ = 8.4, J$_{46}$ = 0.9 Hz, C$_4$-H), 7.83 (1H, d, J = 8.6 Hz, C$_7$-H), 8.83 (1H, br d, α-NH), 11.74 (1H, br s, indole NH). ^{19}F NMR (acetone-d$_6$, ppm, 400 MHz), δ: -76.73 (α-COCF$_3$).

6-Nitro-N$_\alpha$-trifluoroacetyl-L-tryptophan methyl ester: In a 50 mL round-bottom flask 1.0 g (2.4 mmol) of N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester, 20 mL of acetic anhydride and 15 mL of acetic acid is added and chilled in an ice-salt bath for 10 minutes. Then 4.0 mL of nitric acid is added dropwise to the solution and allowed to stir for 45 minutes. The reaction is diluted with ice-water and a orange precipitate forms. The precipitate is filtered and collected to give 0.45 g (41%). Mp 201-204 $^\circ$C. Lit mp 200-202 $^\circ$C [12]. ^{1}H NMR (acetone-d$_6$, ppm, 400 MHz), δ: 3.38 (1H, dd, J$_{\beta\beta'}$ = 14.9, J$_{\alpha\beta}$ = 8.6 Hz, β-H), 3.52 (1H, dd, J$_{\beta\beta'}$ = 14.9, J$_{\alpha\beta'}$ = 5.5 Hz, β-H), 3.73 (3H, S, -OCH$_3$), 4.95 (1H, dd, J = 8.6, 5.5 Hz, α-H), 7.78 (1H, d, J=8.6 Hz, C$_4$-H), 7.94 (1H, dd, J$_{45}$ = 8.6, J$_{56}$ = 1.6 Hz , C$_5$-H), 7.68 (1H, d, J = 2.35 Hz, C$_2$-H), 8.38 (1H, d, J = 2.3 Hz, C$_7$-H), 8.75 (1H, br d, α-NH), 10.87 (1H, br s, indole N-H). ^{19}F NMR (acetone-d$_6$, ppm, 400 MHz), δ: -76.78 (α-COCF$_3$).

2-Nitro-L-tryptophan: In a 50 mL round-bottom flask 0.5 g (1.1 mmol) of 2-Nitro-N-trifluoroacetyl-L-tryptophan is added to a solution of 10 mL of 3N HCl and 20 mL of 1,4-dioxane and stirred. The mixture is allowed to reflux for 2 hours. The solvent is removed in vacuo to give a reddish-brown solid, 0.16 (58%). Mp 242-245 ^{o}C. Lit. mp 240 ^{o}C [12]. ^{1}H NMR (D$_2$O, ppm, 400 MHz), δ: 3.23 (dd, 1H, $J_{\alpha\beta}$ = 6.1 Hz, $J_{\beta\beta'}$ = 11.2 Hz, β-H), 3.31 (dd, 1H, $J_{\alpha\beta}$ = 6.1 Hz, $J_{\beta\beta'}$ = 11.2 Hz, β-H), 4.05 (dd, 1H, J = 6.8 Hz, α-H), 7.21 (ddd, 1H, J_1 =2.3 Hz, J_2 = 5.2, J_3 = 7.9 Hz, Ar-H), 7.57 (m, 2H, Ar-H), 7.80 (d, 1H, J = 7.9 Hz, Ar-H).

6-Nitro-L-tryptophan: In a 50 mL round-bottom flask 0.4 g (0.88 mmol) of 6-Nitro-N-trifluoroacetyl-L-tryptophan is added to a solution of 10 mL of 3N HCl and 20 mL of 1,4-dioxane and stirred. The mixture is allowed to reflux for 2 hours. The solvent is removed in vacuo to give a reddish-brown solid 0.11 g (50%). Mp 260-266 ^{o}C. Lit. mp 265 ^{o}C dec [24]. ^{1}H NMR (D$_2$O, ppm, 400 MHz), δ: 3.20 (dd, $J_{\alpha\beta}$ = 5.4 Hz, $J_{\beta\beta'}$ = 12.2 Hz, 1H, β-H), 3.32 (dd, $J_{\alpha\beta}$ = 4.9 Hz, $J_{\beta\beta'}$ = 12.2 Hz, 1H, β-H), 3.97 (dd, J = 5.4 Hz, 1H, α-H), 7.45 (s, 1H, Ar-H), 7.60 (d, J = 7.8 Hz, 1H, Ar-H), 7.85 (dd, J = 2.0, 7.8 Hz, 1H, Ar-H), 8.28 (s, 1H, Ar-H).

N-Acetyl-DL-2-methyltryptophan: In a 100 mL round-bottom flask 0.5 g (3.8 mmol) of 2-methylindole, 1.0 g (9.5 mmol) of DL-serine, 30 mL of acetic anhydride, and 15 mL of acetic acid was combined and refluxed for three hours under nitrogen. The reaction was extracted with ethyl acetate and washed with H$_2$O. The solvent was removed *in vacuo* to give a dark reddish brown oil, 0.82 g (41%). Mp 210-215 ^{o}C. Lit. mp 208-209 ^{o}C [25]. ^{1}H NMR (acetone-d$_6$, ppm, 400 MHz), δ: 1.89 (s, 3H, CH$_3$), 2.38 (s, 3H, CH$_3$), 3.17 (dd, 1H, $J_{\beta\beta'}$ = 6.0, $J_{\alpha\beta}$ = 6.4 Hz, β-H), 3.25 (dd, 1H, $J_{\beta\beta'}$ = 6.0, $J_{\alpha\beta'}$ = 6.0 Hz, β'-H), 4.75 (dd, 1H, J = 6.4, 6.4 Hz, α-H), 6.97 (m, 2H, J = 14.8, 15.2 Hz, Ar-H), 7.19 (1H, t, J = 8.0 Hz, Ar-H), 7.24 (1H, dd, J = 7.6 Hz, Ar-H), 7.52 (1H, s, NH), 9.91 (br s, 1H, indole NH).

2-Nitro-N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester: In a 100 mL round-bottom flask, 30 mL of acetic anhydride and 20 mL of acetic acid is added and chilled in an ice-salt bath for 10 minutes. Then 4.5 mL of nitric acid is added dropwise to the solution and stirred for 5 minutes. Then 1.5 g (3.6 mmol) of N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester was added in portions. 4.5 mL of nitric acid is added dropwise to the solution and allowed to stir for 30 minutes. The reaction is diluted with ice-water and a yellow-orange precipitate forms. The precipitate is filtered and collected to give 0.62 g (38%). Mp 157-160 ^{o}C. ^{1}H NMR (acetone-d$_6$, ppm, 400 MHz), δ: 3.23 (1H,dd, J$_{\beta\beta'}$ = 10.0, J$_{\alpha\beta}$ = 9.6 Hz, β-H), 3.39 (1H, dd, J$_{\beta\beta'}$ = 13.6, J$_{\alpha\beta'}$ = 10.0 Hz, β'-H), 3.60 (3H, s, OCH$_3$), 4.83 (1H, dd, J = 8.8, 8.8 Hz, α-H), 7.36 (2H, m, J$_1$ = 6.8 Hz, J$_2$ = 6.8 Hz, Ar-H), 7.64 (1H, d, J = 6.8 Hz, Ar-H), 8.25 (1H, d, J = 8.0 Hz, Ar-H), 8.75 (1H, br d, α-NH). ^{19}F NMR (acetone-d$_6$, ppm, 400 MHz), δ: -70.77 (indole COCF$_3$), -76.75 (α-COCF$_3$).

6-Nitro-N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester: In a 100 mL round-bottom flask 1.0 g (2.4 mmol) of N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester, 30 mL of acetic anhydride and 20 mL of acetic acid is added and chilled in an ice-salt bath for 10 minutes. 4.5 mL of nitric acid is added dropwise to the solution and allowed to stir for 30 minutes. The reaction is diluted with ice-water and a orange precipitate forms. The precipitate is filtered and collected to give 0.45 g (41%). Mp 134-137 ^{o}C. ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 3.35 (1H, dd, J$_{\beta\beta'}$ = 12.2, J$_{\alpha\beta}$ = 7.8 Hz, β-H), 3.54 (1H, dd, J$_{\beta\beta'}$ = 12.2, J$_{\alpha\beta'}$ = 5.2 Hz, β-H), 3.85 (3H, s, -OCH$_3$), 4.96 (1H, dd, J = 7.8, 4.9 Hz, α-H), 7.04 (1H, d, J = 6.5 Hz, Ar-H), 7.52 (1H, s, Ar-H), 7.69 (1H, d, J = 6.3 Hz, Ar-H), 8.37 (1H, d, J = 2.3 Hz, Ar-H), 9.39 (1H, s, α-NH). ^{19}F NMR (acetone-d$_6$, ppm, 400 MHz), δ: -70.75 (indole COCF$_3$), -76.73 (α-COCF$_3$).

N$_\alpha$-trifluoroacetyl-L-6-Tryptophan diazonium fluoroborate: In a 3-neck 50 mL round-bottom flask 0.3 g (0.66 mmol) of 6-Nitro-N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester is dissolved in 10 mL 6N HCl and 0.5 g (2.6 mmol) of stannous chloride and purged with nitrogen. The reaction was refluxed under nitrogen for 3 hours. The reaction was chilled in an ice-salt bath and 0.3 g (3.3 mmol) of NaNO$_2$ dissolved in 3 mL of H$_2$O was added by a syringe and stirred for 30 minutes. The mixture was filtered over Celite to remove any precipitates that formed. The solution was chilled in an ice-salt bath and 0.37 g (3.3 mmol) of NaBF$_4$ was added and an orange precipitate formed. The precipitate was filtered and collected to give an orange solid, 0.064 g (31%). ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 3.36 (dd, J$_{\alpha\beta}$ = 5.2 Hz, J$_{\beta\beta'}$ = 5.2 Hz, 1H, β-H), 3.50 (dd, J$_{\alpha\beta}$ = 6.4 Hz, J$_{\beta\beta'}$ = 8.8 Hz, 1H, β-H), 4.95 (dd, J = 6.0, 12.4 Hz, 1H, α-H), 7.05 (s, 1H, Ar-H), 7.52 (d, J = 7.8 Hz, 1H, Ar-H), 7.68 (dd, J = 2.0, 7.8 Hz, 1H, Ar-H), 8.33 (s, 1H, Ar-H), 9.35 (s, 1H, indole NH). ^{19}F NMR (acetone-d$_6$, ppm, 400 MHz), δ: -76,04 (N$_2$BF$_4$, -76.27 (α-COCF$_3$).

References

1. Coppole, G. M.; Schuster, H. F. *Asymmetric Synthesis. Construction of Chiral Molecules Using Amino Acids*; John Wiley & Sons: New York, **1987**.

2. Sanderson, J. M. *Org. Biomol. Chem.* **2005**, *3*, 201-212.

3. Bittner, S.; Scherzer, R.; Harlev, E. *Amino Acids*. **2006**, *33*, 19-42.

4. Neidleman, N. L.; Geigert, J. *Biohalogenation: Principles, Basic roles and Applications*; Ellis Horwood Ltd: Chichester, **1986**.

5. Royer, C. A. *Chem. Rev.* **2006**, *106*, 1769-1784.

6. (a) Munn, D. H.; Zhou, M.; Attwood, J. T.; Bondarev, I.; Conway, S. J.; Marshall, B.; Brown, C.; Mellor, A. L. *Science*. **1998**, *281*, 1191-1193. (b) Mellor, A. L.; Munn, D. H. *Immunology Today*. **1999**, *20*, 469-473.

7. (a) Lee, M. S.; Phillips, R. S. *Bio. Med. Chem. Letts*. **1991**, *1*, 477. (b) Sloan, M. J.; Phillips, R. S. *Bio. Med. Chem. Letts*. **1992**, *2*, 1053. (c) Lee, M.; Phillips, R. S. *Bio. Med. Chem. Letts*. **1992**, *2*, 1563. (d) Dua, R. K.; Phillips, R. S. *Tetrahedron Letts*. **1992**, *33*, 29. (e) Welch, M.; Phillips, R. S. *Bio. Med. Chem. Letts*. **1999**, *9*, 637. (f) Ma, C.; Liu, X.; Li, X.; Flippen-Anderson, J.; Yu. S.; Cook, J. M. *J. Org. Chem*. **2001**, *66*, 4525. (g) Phillips, R. S. *Tetrahedron: Asymmetry*. **2004**, *15*, 2787. (h) Goss, R. J. M.; Newill, P. L. A. *Chem. Commun*. **2006**, *47*, 4924-4925.

8. Hyde, C. C.; Ahmed, S. A.; Padiun, E. A.; Miles, E. W.; Davies, D. R. *J. Biol. Chem*. **1988**, *263*, 17587-17871.

9. Gilchrist, T. L. *Heterocyclic Chemistry*; Addison Wesley Longman Ltd: London, **1997**.

10. Osborne, A. S.; Som, P.; Metcalf, J. L.; Phillips, R. S. *Bioorganic & Medicinal Chemistry Letters*. **2008**, *18*, 5750-5752.

11. Qian, Z-S.; Zhou, C-J.; Li, G.; Cao, L-H. *Youj: Huaxue*, **2008**, *28*, 467-471.

12. Phillips, R. S.; Cohen, L. A. *J. Heterocyc. Chem*. **1988**, *25*, 191-192.

13. Katayama, M.; Kato, Y.; Marumo, S. *Biosci. Biotechnol. Biochem*. **2004**, *68*, 1287-1292.

14. Bandgar, B. P.; Sadavarte, V. S.; Shinde, S. D.; Kamble, V. T. *Sulfur Letters*. **2000**, *24*, 123-126.

15. Balz, G.; Schiemann, G. *Ber*. **1927**, *60*, 1186-1190.

16. Yokoyama, U.; Kikawa, H.; Mitsuhashi, M.; Uyama, A.; Murakami, Y. *Tetrahedron Letters*. **1999**, *40*, 7803-7806.

17. Blaser, G.; Sanderson, J. M.; Batsanov, A. S.; Howard. J. A. *Tetrahedron Letters*. **2008**, *49*, 2795-2798.

18. Onslow, H. *Biochem. J*. **1921**, *15*, 383-391.

19. Song, Q.; Tang, W.; Hei, X.; Wang, H.; Guo, Q.; Yu, S. *Eur. J. Org. Chem*. **2005**, 1097-1106.

20. Lopez-Serrano, P.; Jongejan, J. A.; van Rantwijk, F.; Sheldon, R. A. *Tetrahedron: Asymmetry*. **2001**, *12*, 219-228.

21. Lee, M. Ph.D. Dissertation, University of Georgia of Athens, **1992.**

22. *Org. Syntheses*. **1943**, *2*, 130.

23. (a) Roy, S.; Gribble, G. W. *Heterocycles*. **2006**, *70*, 51-56. (b) Madelung, W. Justus *Liebig Ann. Chem*. **1914**, *405*, 58-95.

24. De Fazi, R.; Berti, G.; Da Settimo, A. *Gazz. Chim. Ital*. **1959**, *89*, 2238-2242.

25. Snyder, H. R.; MacDonald, J. A. *J. Am. Chem. Soc*. **1955**, *55*, 1257-1259.

SUPPORTING DATA

Figure 3.3: Chiral analysis of 2-nitrotryptophan. The free amino acids were obtained by complete hydrolysis under acid catalysis. The DL and L 2-nitrotryptophans were equilibrated with 5 mM CuSO$_4$, pH 3.15, and run at 1 mL/min.

DL and L-2NO2Trp
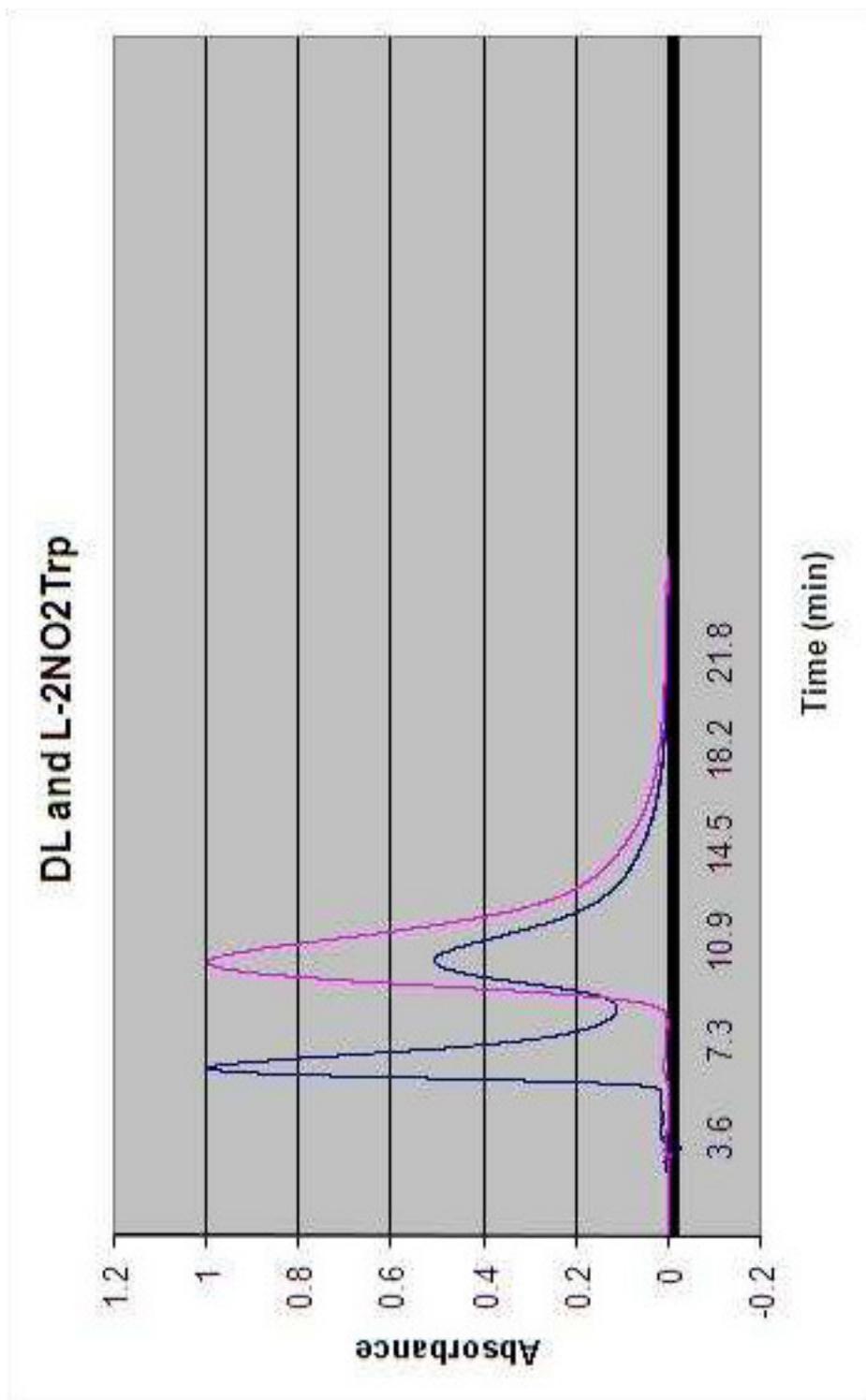
Figure 3.4: Chiral analysis of 6-nitrotryptophan. The free amino acids were obtained by complete hydrolysis under acid catalysis. The DL and L 6-nitrotryptophans were equilibrated with 5 mM $CuSO_4$, pH 3.15, and run at 1 mL/min.

DL and L-6NO2Trp

^{1}H-NMR N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester

3.754
3.897
3.911
3.692

3.931
3.946

4.978
4.993
5.000
5.014
5.022
5.036

7.205
7.221
7.241
7.438
7.459
7.476
7.517
7.536

7.824
7.846

8.821
8.841

11.743

^{1}H-NMR 2-Nitro-N$_\alpha$-trifluoroacetyl-L-tryptophan methyl ester

3.352
3.374
3.389
3.411
3.489
3.503
3.526
3.540

3.731

4.858
4.879
4.891
4.912

CO$_2$CH$_3$

NHTFA

O$_2$N

^{1}H-NMR 6-Nitro-N$_\alpha$-trifluoroacetyl-L-tryptophan methyl ester

7.762
7.783

7.943
7.970

8.393

8.787
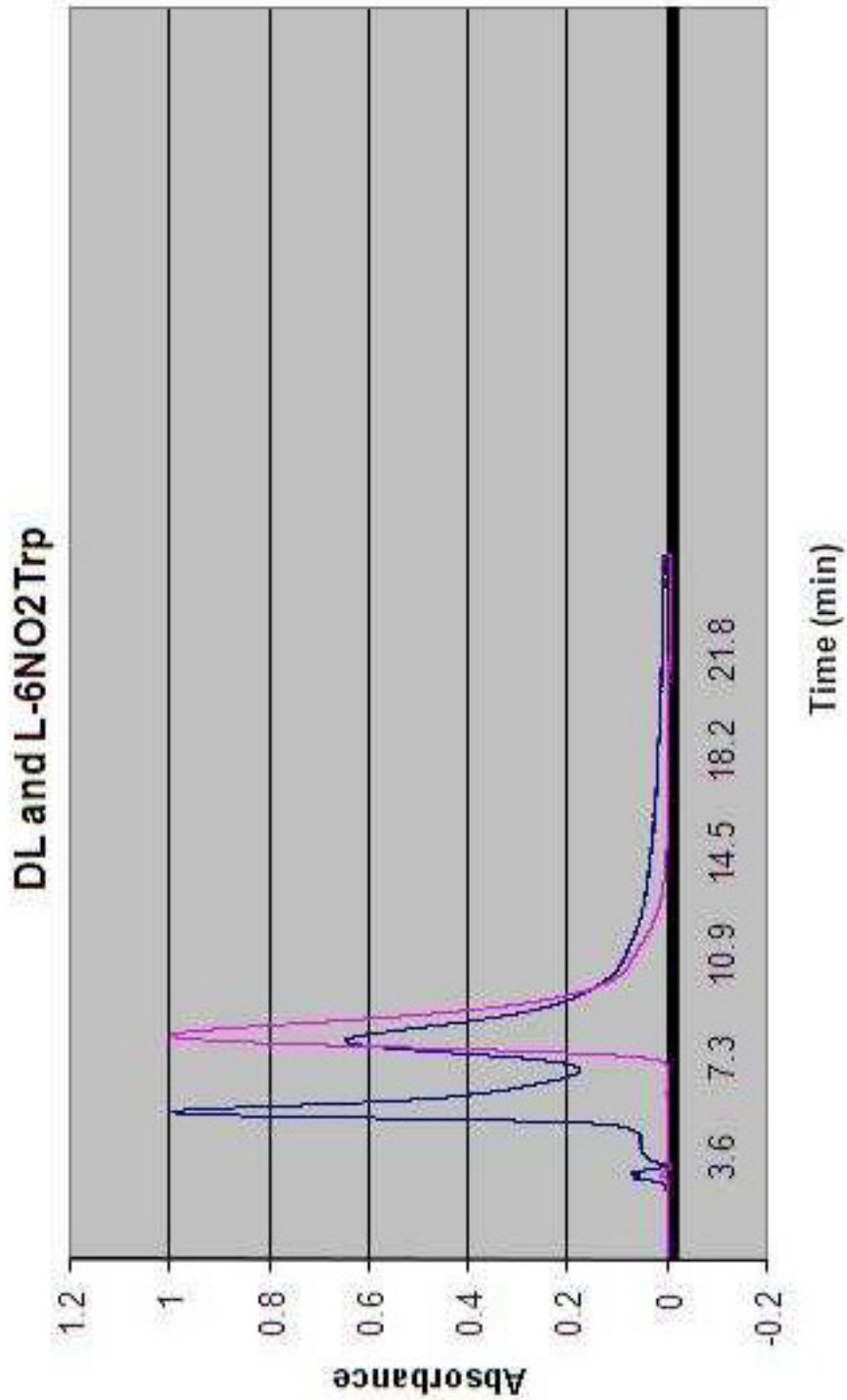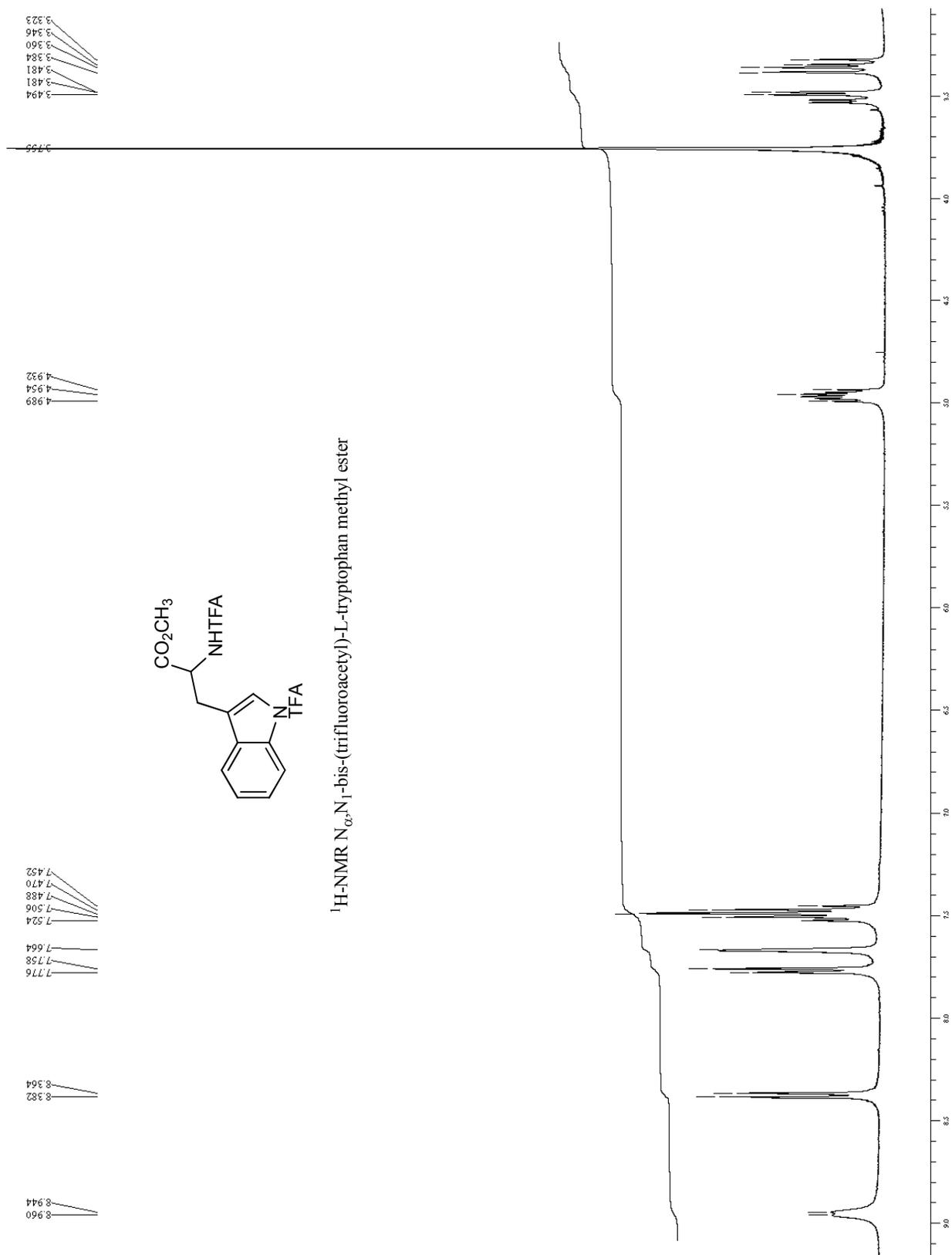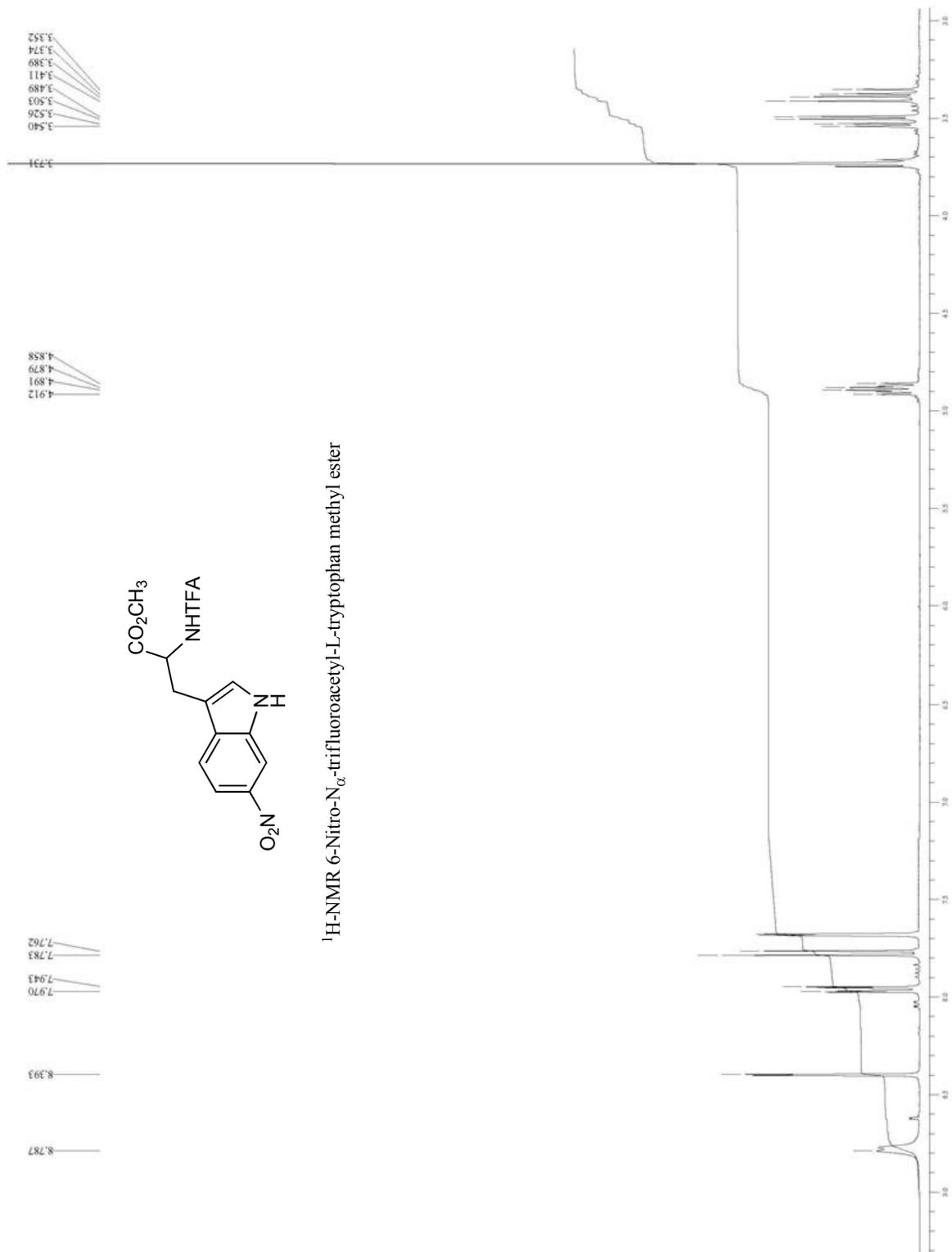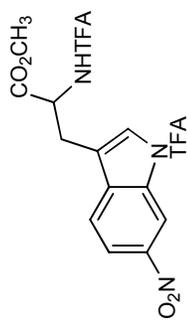
^{19}F-NMR 6-Nitro-N_α,N_1-bis-(trifluoroacetyl)-L-tryptophan methyl ester

^{19}F-NMR 6-Nitro-N$_\alpha$-trifluoroacetyl-L-tryptophan methyl ester

ppm

^{1}H-NMR 6-Nitro-N$_\alpha$,N$_1$-bis-(trifluoroacetyl)-L-tryptophan methyl ester
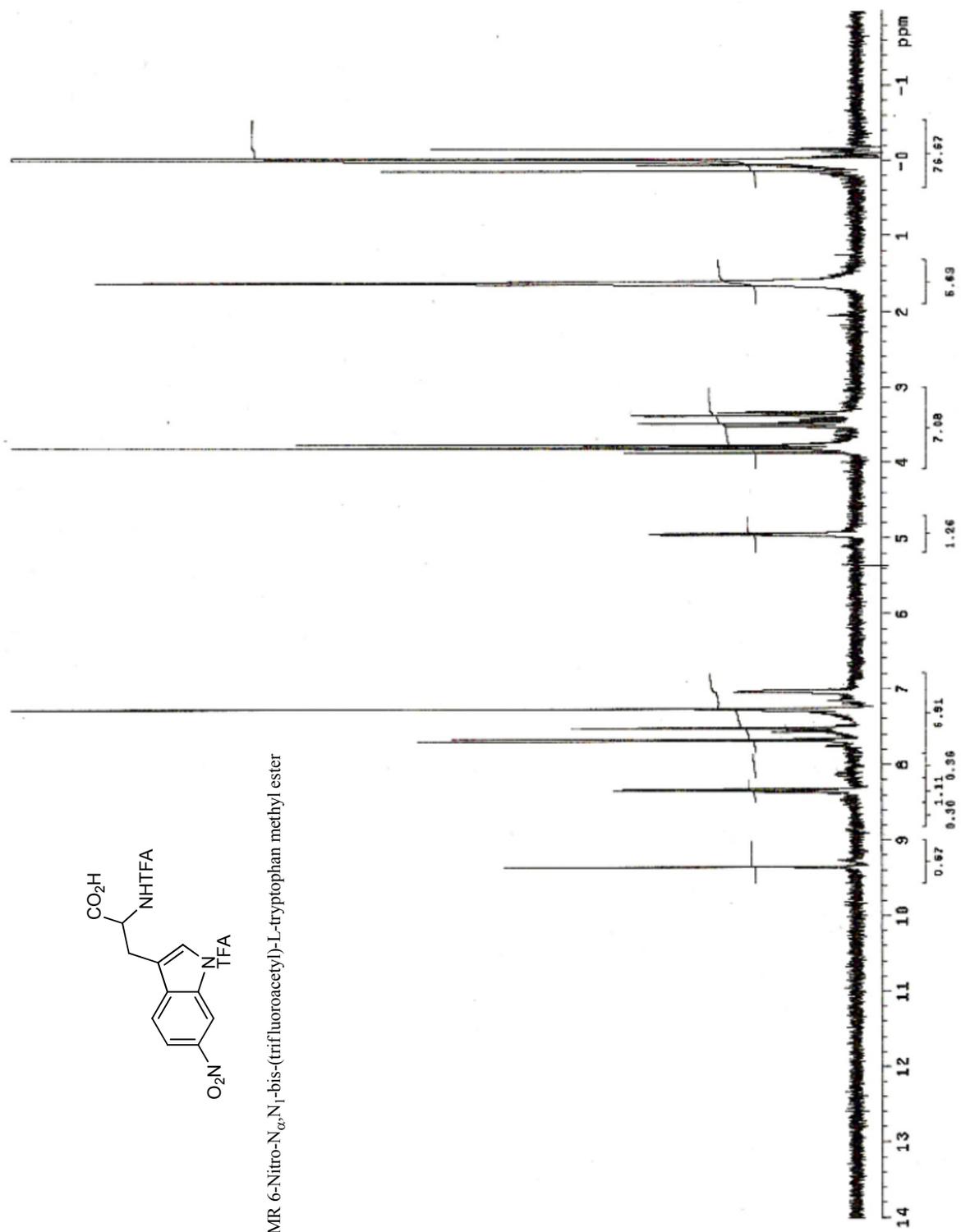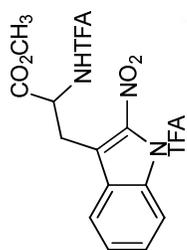
^{1}H-NMR 2-Nitro-N$_\alpha$,N$_i$-bis-(trifluoroacetyl)-L-tryptophan methyl ester
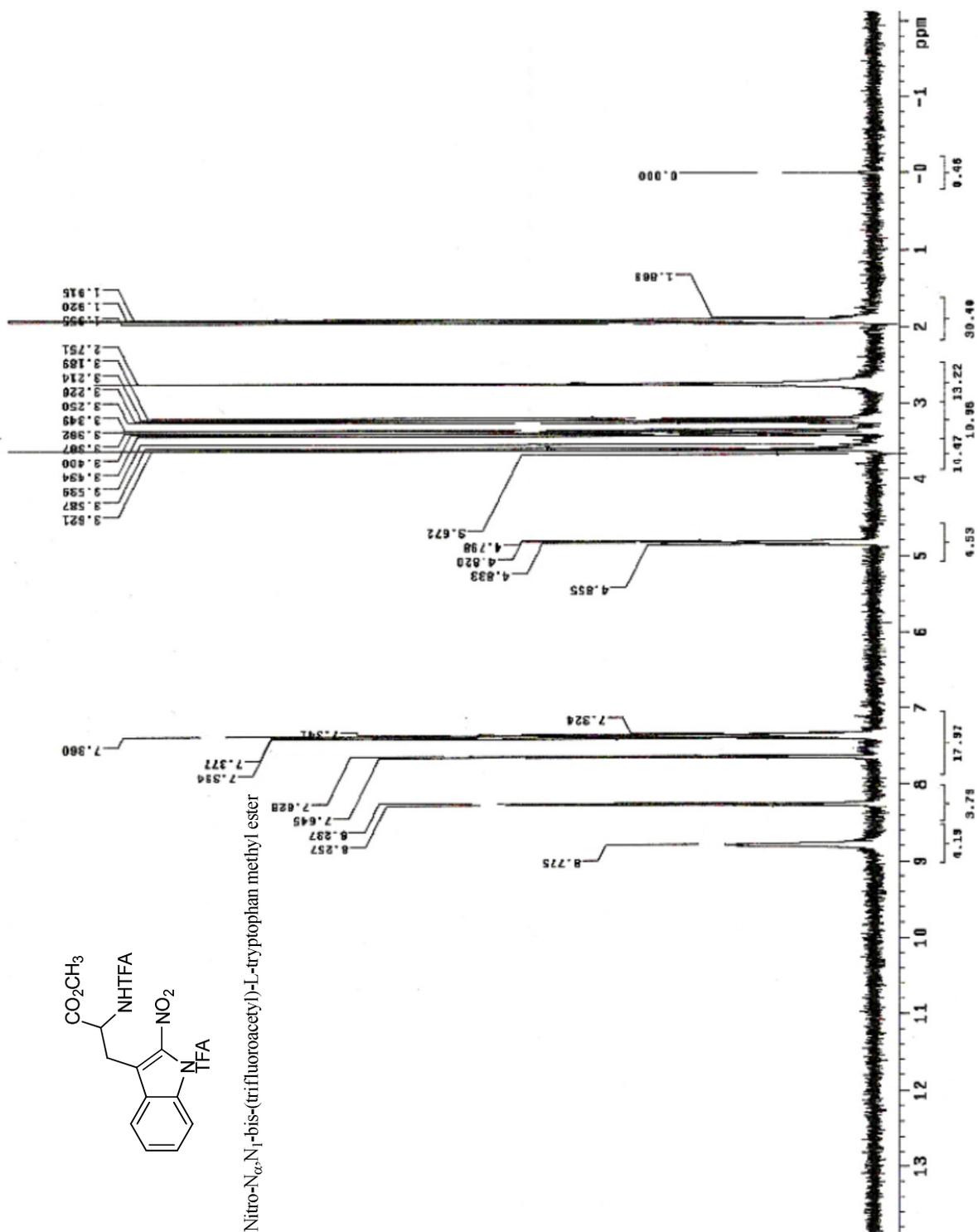
Figure 3.5: GC-MS analysis of N-Acetyl-DL-2-methyltryptophan as a crude product.

C:\GCMSsolution\Data\R. Philips\ATME-04-19-2009.qgd

Chromatogram (Zoom)

3,756,009

TIC*1.00

9.0   10.0   11.0   12.0
min

Spectrum

Line#:1 R.Time:10.985(Scan#:1588)
MassPeaks:446
RawMode:Single 10.985(1588) BasePeak:144.05(2199170)
BG Mode:None

2000000

1000000

144

242

45 63 77 91 102 113 128 155 171 182 197 213 233 255 275 295 316 335 355 368 388 401 415 429 443 473 489 503513 535 548

20   50   80   110   140   170   200   230   260   290   320   350   380   410   440   470   500   530   560
m/z

N_α-Acetyl-DL-2-methyltryptophan

^{1}H-NMR N-Acetyl-DL-2-methyltryptophan

^{1}H-NMR N$_\alpha$-trifluoroacetyl-L-L-6-Tryptophan diazonium fluoroborate
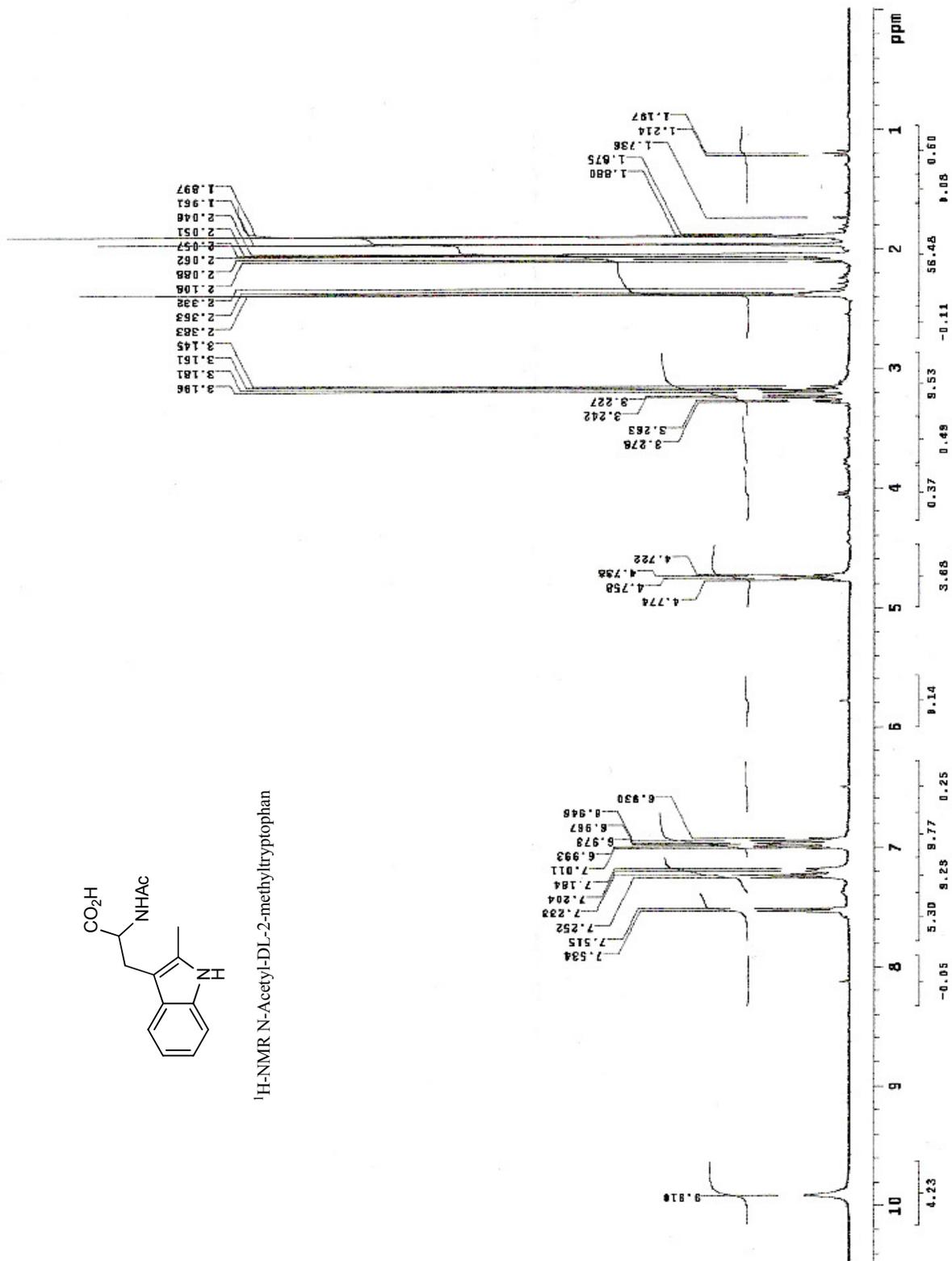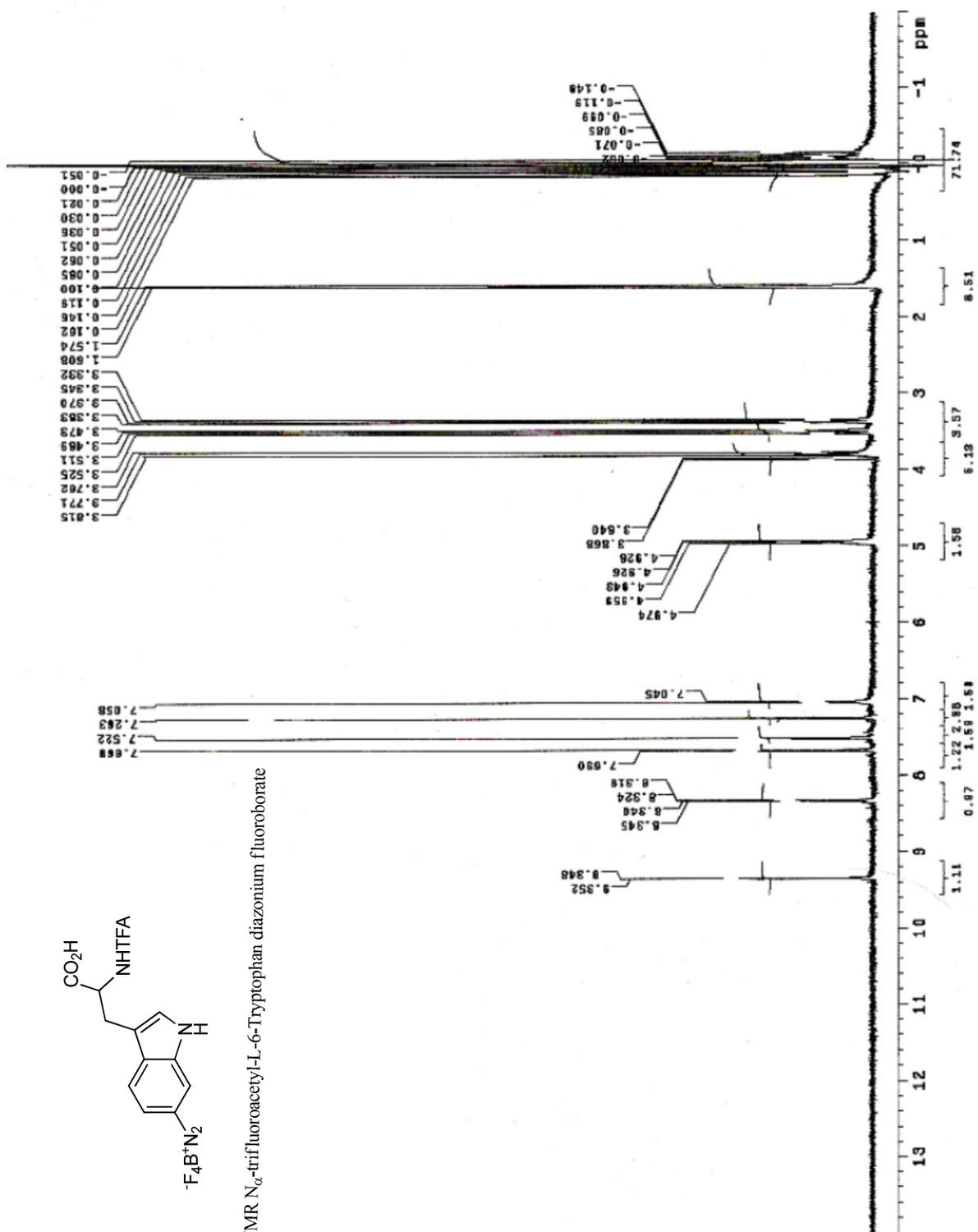
CHAPTER 4

SYNTHESIS OF SULFENYL CHLORIDE ANALOGUES

Introduction

Tryptophan usually has few or no residues on proteins and this makes it an attractive target for labeling. It is the least abundant of the common amino acids, usually occurring at a rate of 1.7% in the composition of a protein and even less than 10% of tryptic peptides [1]. Chemical tagging of tryptophan and other amino acids is a useful tool in proteomics analysis and it is used to introduce isotope labels and mass defect labels into proteolytic peptides by derivatization of amino acid residues [2]. By targeting tryptophan-containing peptides, the proteomic approach will exhibit a considerable reduction in sample complexity [2]. There have been a number of tagging strategies that have been developed that target specific amino acid residues [3]. It has been reported that derivatives of tryptophan in proteins have been achieved using photo-oxidation, iodination, and ozonization [4]. A convenient spectrophotometric method used for determining tryptophan content in proteins is using N-bromosuccinimide, but it is relegated to quantitative analysis of tryptophan in proteins [5]. It has been reported that 2-hydroxy-5-nitrobenzyl bromide has been used as a selective reagent for tryptophan at acidic pH, but it is very sensitive to hydrolysis and has solubility issues under aqueous conditions [6]. It has been reported that sulfenyl halides have a specific reactivity with tryptophan residues of proteins in acidic media [7]. Also, it has been shown that sulfenyl chlorides will react with tryptophan

residues to form thioether bonds at the 2 position of the indole ring [2]. A method for relative quantification of proteins using light and heavy 2-nitrobenzenesulfenyl chloride for mass spectrometric analysis of tryptophan residues has been reported [8]. Therefore, synthesizing sulfenyl chloride analogues that have a chemical selectivity for tryptophan residues and are stable for MALDI analysis is of interest [2]. Figure 4.1 shows the reaction of sulfenyl chlorides with tryptophan residues.
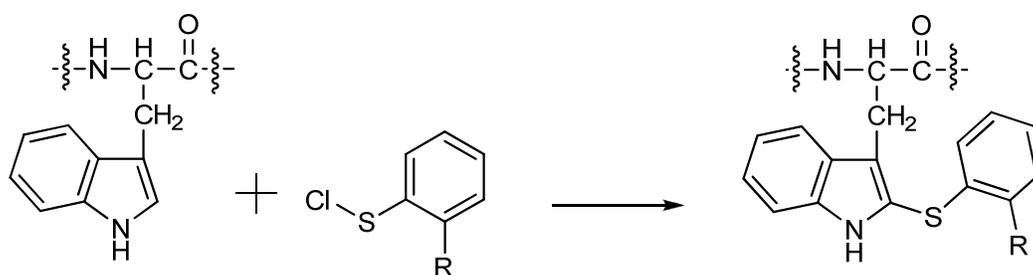
Figure 4.1: The reaction of sulfenyl chlorides with tryptophan residues

Proposed Plan of Research

The proposed synthesis of 2-(trifluoromethyl)benzenesulfenyl chloride is a one step synthesis in sulfuryl chloride, pentane and triethylamine. Schemes 4.1 and 4.2 below illustrate the proposed synthesis of 2-(trifluoromethyl)benzenesulfenyl chloride and 2,4-dibromo-6-(trifluoromethyl)benzenesulfenyl chloride respectively. The synthesis of 2,4-dibromo-6-(trifluoromethyl)benzenesulfenyl chloride requires a four step synthesis. First, bromination of the benzene ring to give 2,4-dibromo-6-(trifluoromethyl)aniline, then converting the amine into the aryl diazonium fluoroborate and coupling it with the resin-SH bound to give 2,4-dibromo-6-(trifluoromethyl)benzenethiol. Conversion of the thiol to the desired sulfenyl chloride for the

final step completes the synthesis. Sulfenyl chlorides containing bromines could provide a tool

as mass-defect labels.

Scheme 4.1: Synthesis of 2-(trifluoromethyl)benzenesulfenyl chloride
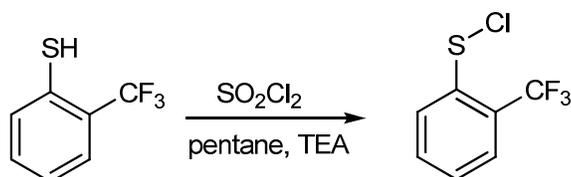
Scheme 4.2: Synthesis of 2,4-dibromo-6-(trifluoromethyl)benzenesulfenyl chloride
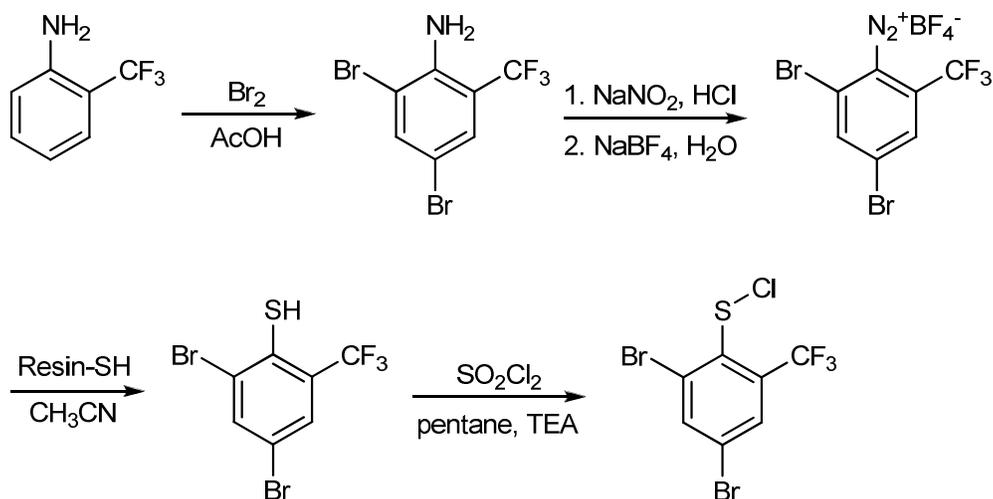
After the synthesis of the sulfenyl chlorides are completed then Chunyan Li will conduct the

mass spectrometry labeling experiments to determine if the compounds are viable derivatization

reagents.

Results and Discussion

The synthesis of 2-(trifluoromethyl)benzenesulfenyl chloride (Scheme 4.1) was straightforward.

The problems with the sulfenyl chlorides are that they are hard to purify. There are side products

that formed from this reaction but predominantly the desired product formed. The synthesis of

4-dibromo-6-(trifluoromethyl)benzenesulfenyl chloride (Scheme 4.2) was more tedious. The

bromination of 2-(trifluoromethyl)aniline was efficient but in some cases did not go to

completion if the reaction was not allowed to react long enough, and this was shown by data

from a mass spectrum. The diazonium fluoroborate was synthesized from the aniline in two

steps to give a white salt, which when treated with the Resin-SH gave the benzenethiol [9]. The

synthesis of the sulfenyl chloride proved to be more difficult but showed partial labeling.

Adding a strong electron-withdrawing group onto the benzene ring would help stabilize the

benzenesulfenyl chloride, making it more applicable as a derivatization reagent [2]. Also, it is

desirable that the new substituent does not produce an absorption band above 300 nm which is

exhibited by a nitro group that is used conventionally for tryptophan derivatization [2].

The mass spectra data was obtained by Chunyan Li of the Amster research group. In Figure 4.2

below shows the unlabelled and labelled reaction of bombesin with 2-

(trifluoromethyl)benzenesulfenyl chloride (2-TFBSCl) using DHB as matrix. The shift is

observed in the (b) spectra below at 1845.7808 m/z where the unlabeled peak is observed at

1669.7817 m/z. The data suggests the stability of the 2-TFBSCl as a derivatization reagent and a

completion of the reaction with the observed peak at 1845.7807 m/z, which matches the

predicted mass of 1845.7725 m/z [2].
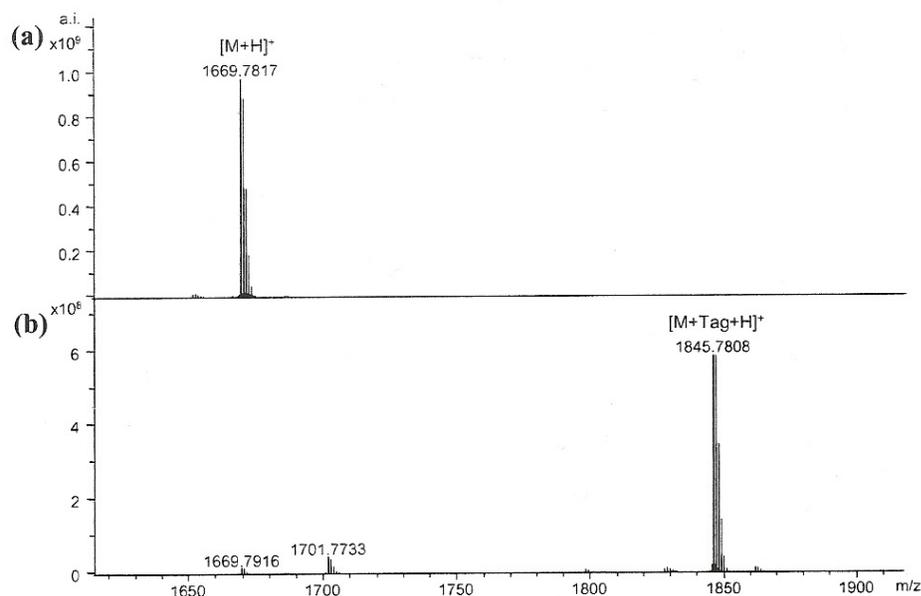
Figure 4.2: MALDI-FTMS spectra of reaction of bombesin (a) before and (b) after with 2-TFBSCl using DHB as matrix

Figure 4.3 below shows the tryptic fragments of labeled and unlabeled myoglobin in a reaction with 2-TFBSCl using DHB as a matrix. The spectrum (b) shows the shift of the single and double labeled peptides at 2023.8898 and 2167.9005 m/z where the unlabeled peptide is observed at 1815.904 m/z in spectrum (a). These experiments were carried out to test for the specificity of the 2-TFBSCl. Myoglobin has two tryptophan residues and both tryptophan residues occur in the same tryptic fragment. The tryptic fragments were analyzed and it was observed that both residues were labeled, but there were some singly labeled residues [2]. This indicates that the reaction nearly went to completion with the observed peak at 2167.9005 m/z which shifted from 1815.9041 m/z.
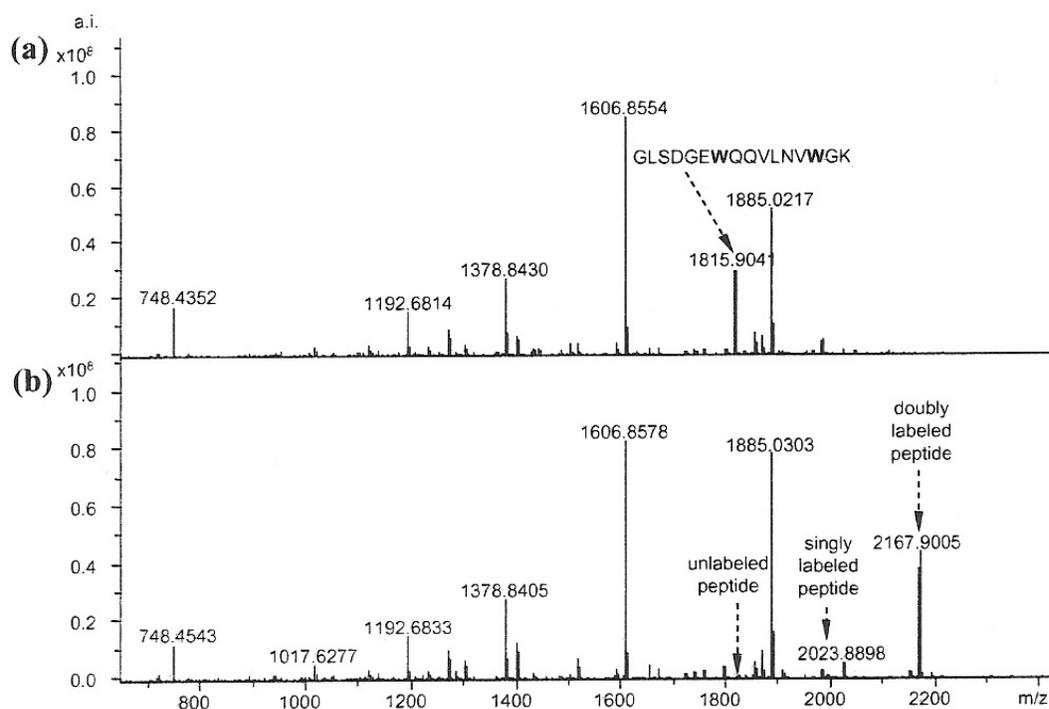
Figure 4.3: MALDI-FTMS spectra of the tryptic fragments of (a) underivatized myoglobin and (b) 2-TFBS-derivatized myoglobin using DHB as matrix

Figure 4.4 below shows the CAD and IRMPD spectra of the TFBS labeled peptide in myoglobin fragmentation. These experiments were carried out to further test for the specificity of 2-TFBS-Cl as a derivatization reagent. The data observed confirmed the high specificity that 2-TFBS-Cl has for tryptophan residues [2]. In Table 4.1 are the peak assignments in the CAD spectrum and in Table 4.2 are the peak assignments in the IRMPD spectrum.
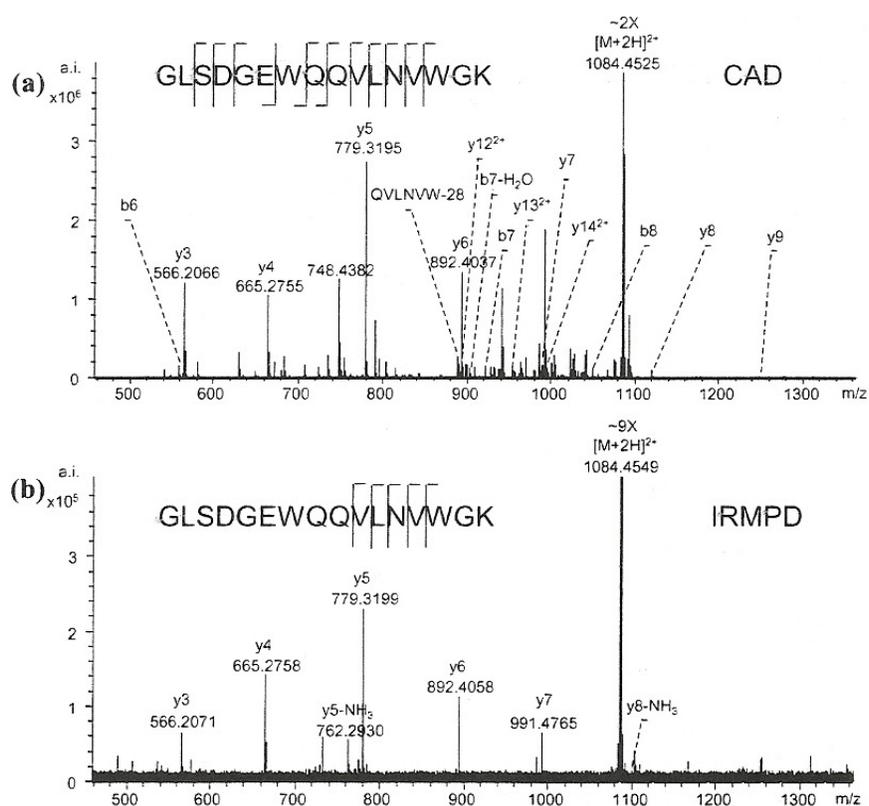
Figure 4.4: TFBS-labeled peptide in myoglobin (a) CAD spectrum and (b) IRMPD spectrum

Table 4.1: Assignment of peaks in CAD spectrum labeled with TFBSCl

Fragment ion	Theoretical mass	Experimental mass	R.I. (%)	Error in ppm
b6	599.2364	599.2375	2.4	2.0
y3	566.2049	566.2066	18.3	3.0
y4	665.2733	665.2755	15.7	3.3
y5	779.3163	779.3195	41.7	4.1
QVLNVW-28	888.4054	888.4032	4.3	-2.5
y6	892.4003	892.4037	20.3	3.8
y12-2	898.3639	898.3673	2.7	3.8
b7-H2O	903.2959	903.3008	2.0	5.4
b7	921.3065	921.3099	2.6	3.7
y13-2	955.8774	955.8824	1.3	5.2
y7	991.4687	991.4736	4.8	4.9
y14-2	999.3934	999.3971	2.9	3.7
b8	1049.3651	1049.3723	2.1	6.9
y8	1119.5273	1119.5341	1.7	6.1
y9	1247.5859	1247.5936	0.4	6.2

Table 4.2: Assignment of peaks in IRMPD spectrum labeled with TFBSCl

Fragment ion	Theoretical mass	Experimental mass	R.I. (%)	Error in ppm
y3	566.2049	566.2071	1.8	3.9
y4	665.2733	665.2758	3.8	3.8
y5-NH$_3$	762.2897	762.293	1.5	4.3
y5	779.3163	779.3199	6.2	4.6
y6	892.4003	892.4058	3.1	6.2
y7	991.4687	991.4765	1.8	7.9
y8-NH$_3$	1102.5008	1102.5075	1.2	6.1

Experimental

General procedures: ^{1}H-NMR and spectra were recorded on a Bruker 400 MHz spectrometer at room temperature in the required deuterated solvent using tetramethylsilane as the internal standard. Thin layer chromatographies were carried out on Whatman Al/Sil G/UV plates (250 µm) in the required solvent system. Solvent removal was performed using a Buchi Rotavapor R-114 and Waterbath B-480.

2-(Trifluoromethyl)benzenesulfenyl chloride: In a 10 mL round-bottom flask 0.4 g (2.2 mmol) of 2-(trifluoromethyl)benzenethiol was added to a solution of pentane while being chilled in an ice-salt bath. Dry triethylamine was added dropwise to the solution, then sulfuryl chloride was added dropwise to the solution and the reaction was allowed to stir for 2 hours. The solvent was removed *in vacuo* to give a light brown oil, 0.24 g, (60%).

2,4-Dibromo-6-(trifluoromethyl)aniline: In a 100 mL round-bottom flask, 1.0 g (6.2 mmol) of 2-(trifluoromethyl)aniline was dissolved in 40 mL of AcOH. Then 1.98 g (12.4 mmol) bromine in AcOH was added dropwise to the solution and the reaction was allowed to stir for 30 minutes. The reaction was quenched with H$_2$O, which precipitated an oil, and the layers were separated to

give an orange oil. The oil was crystallized to give an orange solid, 1.1 g (56%). Mp 39-41°C. Lit mp 40-43°C. ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 4.76 (s, 2H), 7.53 (s, 1H), 7.72 (s, 1H).

2,4-dibromo-6-(trifluoromethyl)benzenediazonium fluoroborate: In a 50 mL round-bottom flask, 1.0 g (3.1 mmol) of 2,4-dibromo-6-(trifluoromethyl)aniline was added to a solution of 6N HCl. A chilled solution of NaNO$_2$ was added dropwise while stirring. Then a solution of NaBF$_4$ was added. A white solid precipitated, which was filtered and collected 0.95 g (73%). ^{1}H NMR (D$_2$O, ppm, 400 MHz), δ: 8.35 (s, 1H), 8.64 (s, 1H).

Polymer supported hydrogen sulfide: In a 100 mL round-bottom flask 1.4 g (25 mmol) NaSH was added to 50 mL MeOH. Then 5.0 g Amberlite IRA-400 was added and the mixture was allowed to stir vigorously for 1 hour. The resin-SH was filtered and washed with distilled H$_2$O to remove any excess NaSH and chloride ions.

2,4-dibromo-6-(trifluoromethyl)benzenethiol: In a 50 mL round-bottom flask a mixture of 1.0 g (2.4 mmol) of aryl diazonium fluoroborate, 2.5 g of hydrosulfide exchange resin, and 15 mL CH$_3$CN was stirred for 10 minutes at room temperature. The resin changed colors to indicate completion of the reaction. The solution was filtered and the solvent was removed *in vacuo* to give an oil, 0.54 g (67%). ^{1}H NMR (CDCl$_3$, ppm, 400 MHz), δ: 4.18 (s, 1H), 7.39 (s, 1H), 7.58 (s, 1H).

2,4-dibromo-6-(trifluoromethyl)benzenesulfenyl chloride: In a 10 mL round-bottom flask, 0.3 g (0.89 mmol) of 2,4-dibromo-6-(trifluoromethyl)benzenethiol was added to a solution of pentane while being chilled in an ice-salt bath. Dry triethylamine was added dropwise to the

solution, then sulfuryl chloride was added dropwise to the solution and the reaction was allowed to stir for 2 hours. The solvent was removed *in vacuo* to give a golden brown oil, 0.22 g (68%).

References

1. Creighton, T. E. *Proteins: Structures and Molecular Properties*; W. H. Freeman and Company: New York, **1993**.

2. Li, C.; Gawandi, V.; Protos, A.; Phillips, R. S.; Amster, I. J. *Eur. J. Mass Spectrum*. **2006**, *12*, 213-221.

3. (a) Hamdan, M.; Righetti, P. G. *Mass Spectrom. Rev*. **2002**, *21*, 287-302. (b) Leitner, A.; Lindner, W. *J. Chromatogr. B*. **2004**, *813*, 1-26.

4. (a) Weil, L.; James, S.; Buchert, A. R. *Biochem. Biophys*. **1953**, *46*, 266-278. (b) Ramachandran, L. K. *Chem. Rev*. **1956**, *56*, 109-218. (c) Previero, A.; Coletti, M. A.; Galzigna, L. *Biochem. Biophys. Res. Commun*. **1964**, *16*, 195-198.

5. Patchornik, A.; Lawson, W. B.; Witkop, B. *J. Am. Chem. Soc*. **1964**, *80*, 4747-4748.

6. Koshland, D. E.; Karkhanis, Y. D.; Latham, H. G. *J. Am. Chem. Soc*. **1964**, *86*, 1448-1450.

7. (a) Scoffone, E.; Fontana, A.; Rocchi, R. *Biochem. Biophys. Res. Commun*. **1966**, *25*, 170-174. (b) Scoffone, E.; Fontana, A.; Rocchi, R. *Biochemistry*. **1968**, *7*, 971-979. (c) Fontana, A.; Scoffone, E. *Methods Enzymol*. **1972**, *25*, 482. (d) Fontana, A.; Veronese, F. M.; Boccu, E. *Febs. Lett*. **1973**, *32*, 135-138. (e) Tanaka, S.; Mohri, N.; Kihara, H.; Ohno, M. *J. Biochem*. **1985**, *97*, 1377-1384.

8. Kuyama, H.; Watanabe, M.; Toda, C.; Ando, E.; Tanaka, K.; Nishimura, O. *Rapid Commun. Mass. Spectrom.* **2003**, *17*, 1642-1650.

9. Bandgar, B. P.; Sadavarte, V. S.; Shinde, S. D.; Kamble, V. T. *Sulfur Letters.* **2000**, *24*, 123-126.

10. (a) Sologub, L. S.; Moshchitskii, S. D.; Markovskii, L. N.; Ivashchenko, Y. N. *Chem. Heterocycl. Compd.* **1970**, *6*, 1150-1153. (b) Borisov, A. V.; Bel'skii, V. K.; Borisova, V. N.; Osmanov, V. K.; Matsulevich, Z. V.; Goncharova, T. V. *Chem. Heterocycl. Compd.* **2001**, *37*, 702-706.

SUPPORTING DATA

Figure 4.5: Labeling of tryptophan with 2,4-dibromo-6-(trifluoromethyl)benzenesulfenyl chloride with partial labeling.
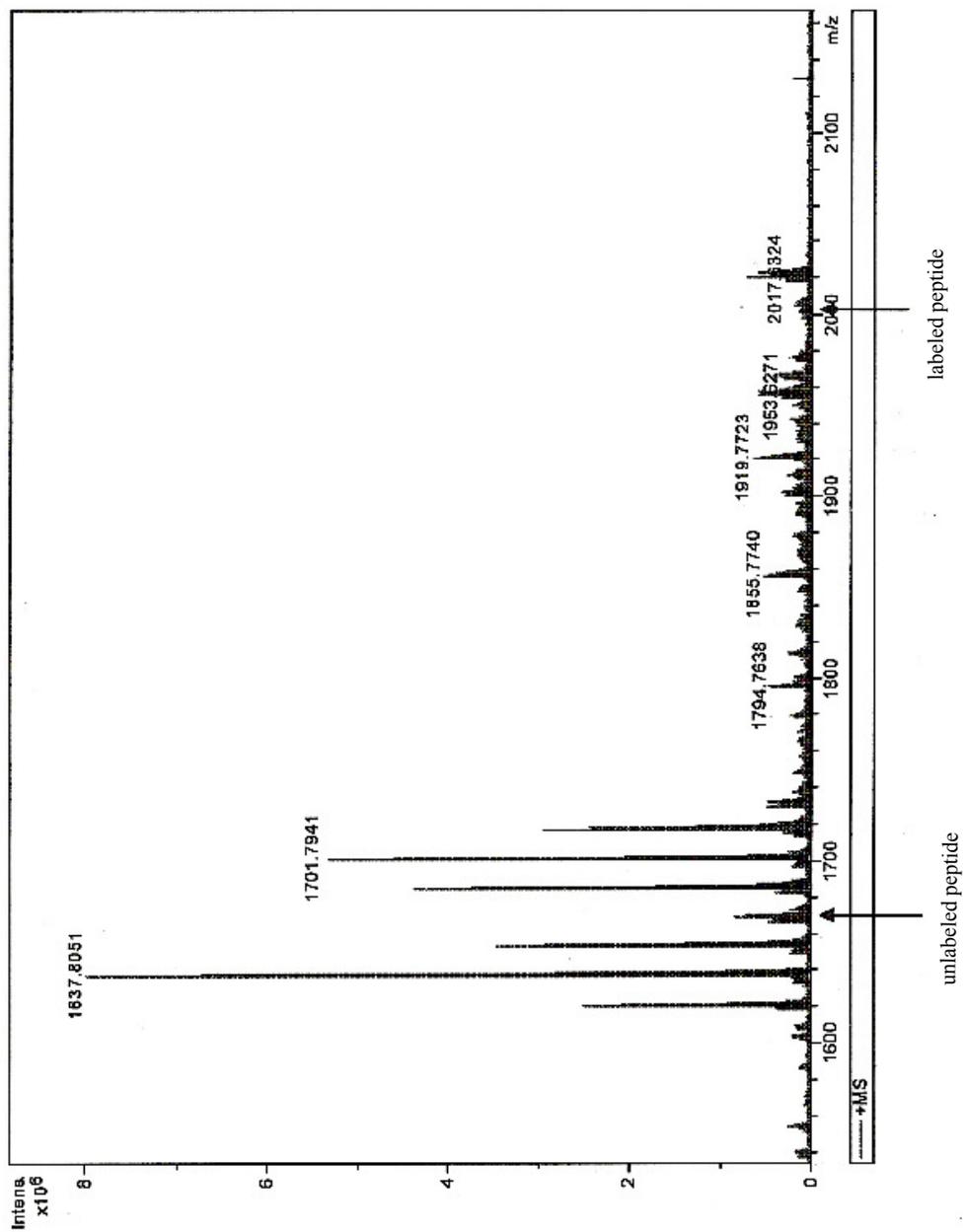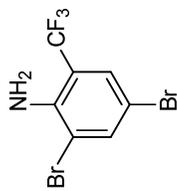
^{1}H-NMR 2,4-Dibromo-6-(trifluoromethyl)aniline
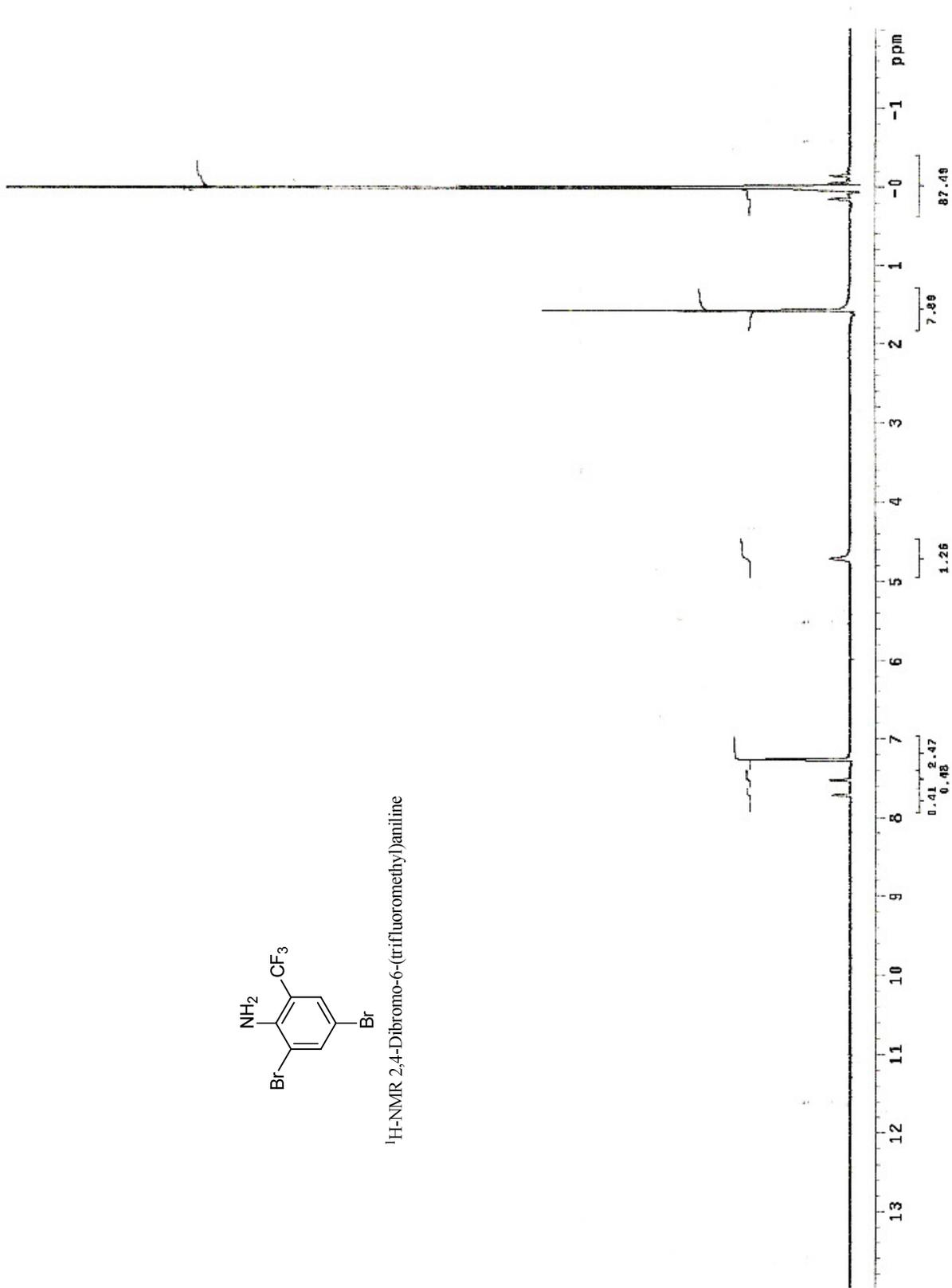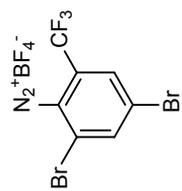
^{1}H-NMR 2,4-dibromo-6-(trifluoromethyl)benzenediazonium fluoroborate
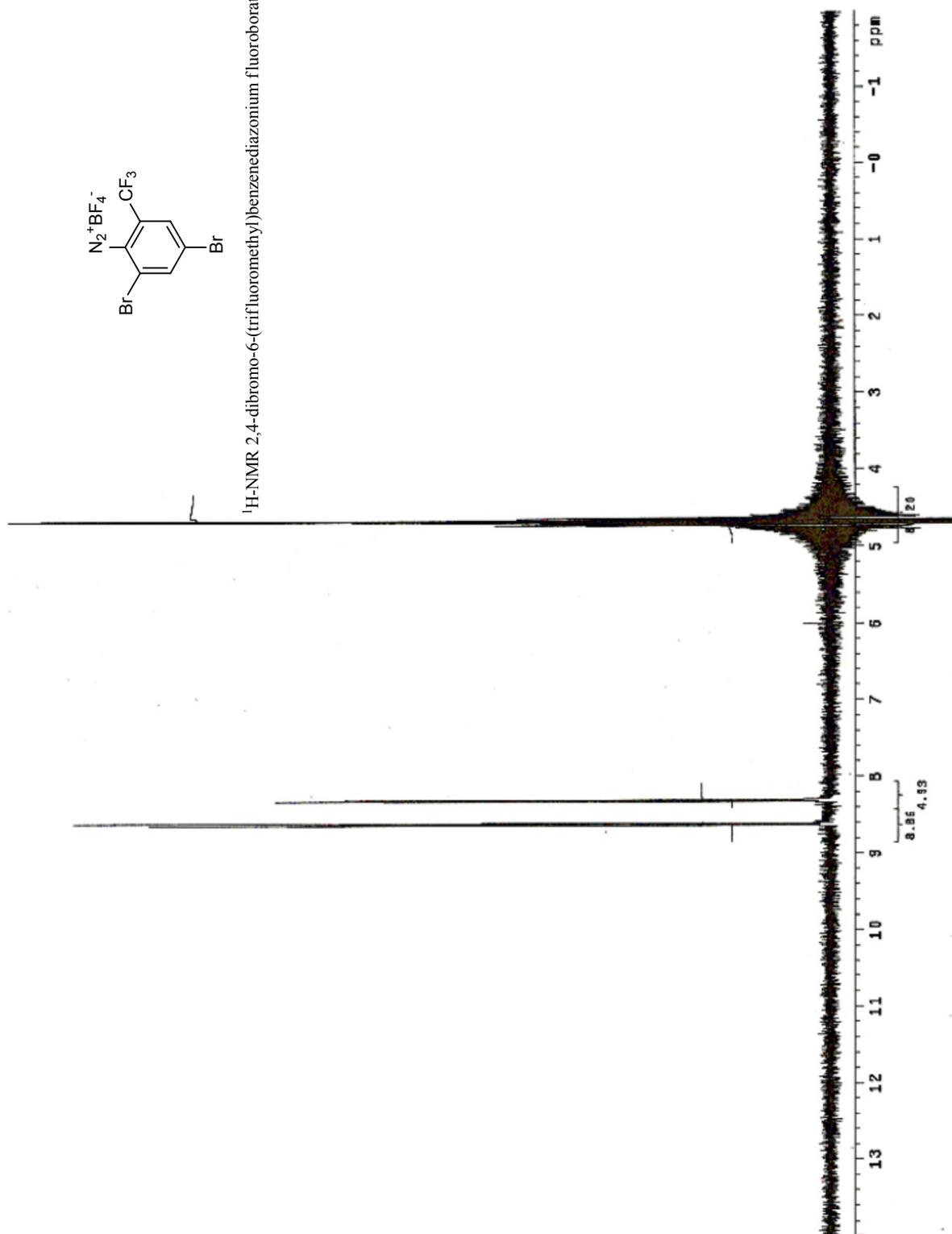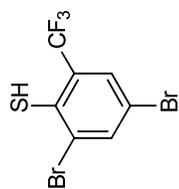
^{1}H-NMR 2,4-dibromo-6-(trifluoromethyl)benzenethiol
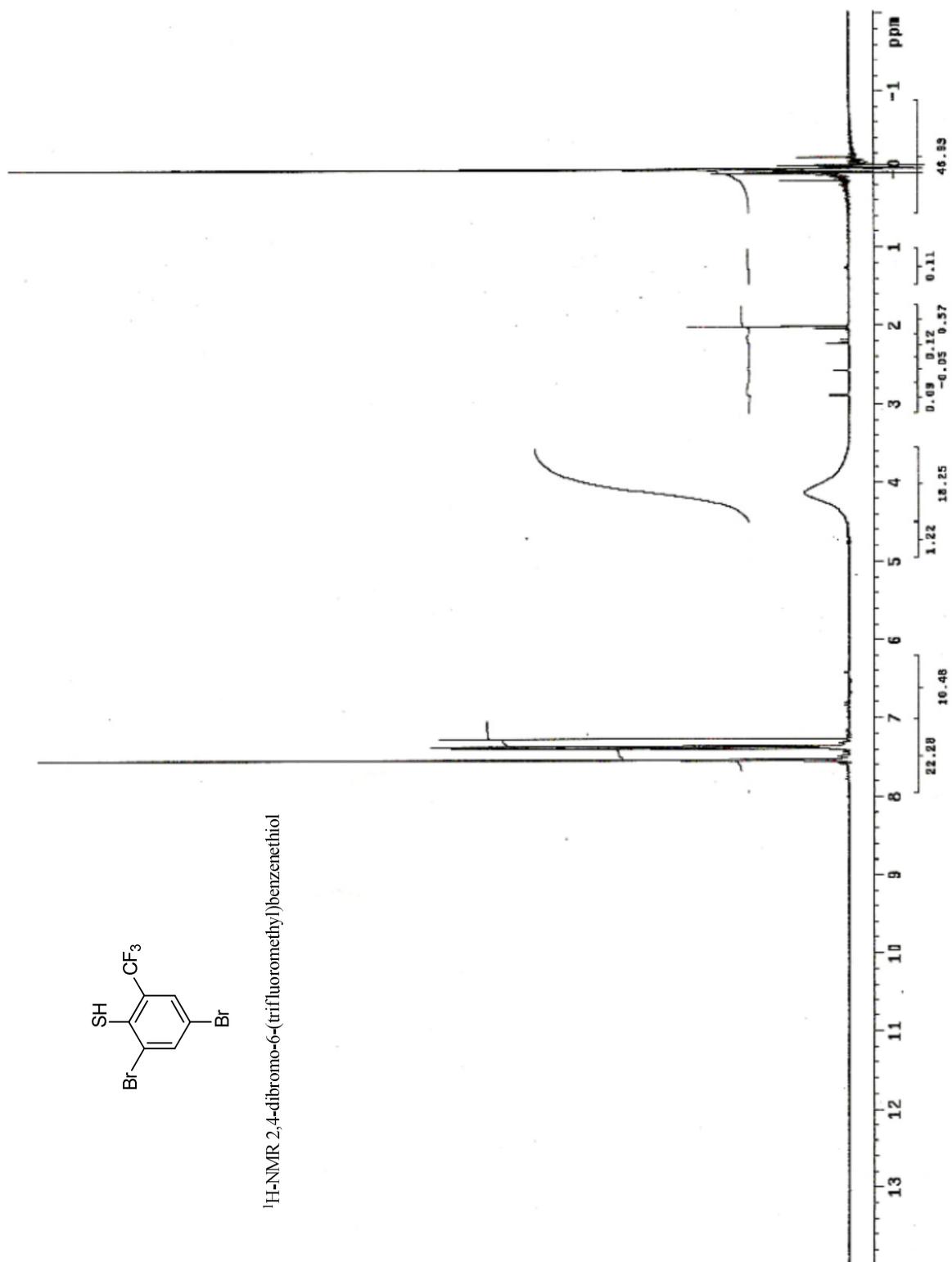
CONCLUSIONS

The synthesis of the thiophene analogues was successful and required procedure modifications to efficiently achieve the first part of the project. The thienopyrroles were synthesized in the past and are known to work, but had problems such as tedious synthesis and decomposing. The synthesis designed for these analogues required less steps and a mild workup. The second part of the project for the incorporation of the thiophene analogues in place of AA required manipulating the cells to uptake the compounds. The incorporation experiments did not give the desired results as expected, but does give us insight into other possibilities pertaining to the non-canonical incorporation project. Knowing the growth conditions and limitations could help further other non-canonical amino acid projects. The possibility of exploring electron rich atoms such as selenium and tellurium could be of interest. The synthetic routes for the thiophene analogues could be applied towards selenophene and tellurophene analogues. The selenophene and tellurophene analogues would be of interest for incorporation experiments.

We were able to successfully determine if racemization occurred in the tryptophan analogues when doing a nitration or trifluoroacetylation reaction. The DL and L isomers were synthesized successfully. At first, synthesizing diasteromeric isomers for the GC analysis showed no separation and it could not be resolved using NMR analysis. It was concluded that racemization does not occur for the nitrotryptophan products by comparing the DL and L isomers by HPLC analysis.

The nitration of tryptophan provides a synthetic route for other derivatives by diazotization and substitution via Sandmeyer or Schiemann reactions. The problem of the instability of the 2-aminoindoles can be circumvented by synthesizing the diazonium fluoroborate, but this has to be done in a one pot reaction without isolating the 2-aminotryptophan. NMR data shows that the

diazonium fluoroborate can be synthesized at the C6 position of tryptophan, which we can apply the same conditions to the C2 position. This would solve the instability problem of the intermediate and provide a synthetic route for other tryptophan derivatives with functionalities at the C2 and C6 position.

To synthesize D-2-methyltryptophan directly would be useful because there is not a published chemical synthesis of the D-2-methyltryptophan. The published procedures for tryptophan analogues related to D-2-methyltryptophan gave undesirable products. We changed the procedure and were able to isolate the DL-2-methyltryptophan. The final step to resolve the D isomer is a reaction with acylase and has not been done. Other D-tryptophan analogues could be synthesized and isolated using this same method and would be of interest.

The synthesis of sulfenyl chlorides analogues was successful. The main problem associated with the sulfenyl chloride synthesis was the final step to introduce the sulfenyl chloride functionality. The reaction with sulfuryl chloride generates by-products and purification for the sulfenyl chlorides was difficult. Due to this, the prior step was ensured to be pure before generating the sulfenyl chloride. More research is needed to develop a method of purifying the sulfenyl chlorides. The labeling of tryptophan experiments conducted by Chunyan Li went to completion for the TFBS-Cl. The brominated TFBS-Cl only exhibited partial labeling, but was shown to have the ability to label. The sulfenyl chlorides provide a tool for proteomics analysis as mass defect labels.