I. KINETIC ISOTOPE EFFECT IN THE β-SUBUNIT MECHANISM OF TRYPTOPHAN SYNTHASE

II. EXCITED STATE TAUTOMERIZATION OF AZAINDOLE

by

MICHAEL T. CASH

(Under the Direction of Robert S. Phillips)

ABSTRACT

In Chapter I of this thesis, the β-subunit mechanism of tryptophan synthase is investigated. We explored the addition of indole to the preformed aminoacrylate L-serine-enzyme complex to form an indolenine quinonoid intermediate. A kinetic isotope effect (KIE) is observed in the tautomerization of the indolenine quinonoid intermediate, and the KIE is dependent on the presence of allosteric ligands.

Chapter II presents the excited state tautomerization of 4-azaindole and 5-azaindole in certain solvent conditions. The N1-H tautomer is lower in energy in the ground state, while the other tautomer is lower in energy in the excited state.

INDEX WORDS: Tryptophan, Tryptophan synthase, PLP, Azaindole, Tautomerization.

I. KINETIC ISOTOPE EFFECT IN THE β-SUBUNIT MECHANISM OF TRYPTOPHAN SYNTHASE

II. EXCITED STATE TAUTOMERIZATION OF AZAINDOLE

by

MICHAEL T. CASH

B.S. Chemistry, North Carolina State University, 1998

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

MASTER OF SCIENCE

ATHENS, GEORGIA

2003

I. KINETIC ISOTOPE EFFECT IN THE β-SUBUNIT MECHANISM OF TRYPTOPHAN SYNTHASE

II. EXCITED STATE TAUTOMERIZATION OF AZAINDOLE

by

MICHAEL T. CASH

Major Professor: Robert Phillips

Committee: Marly Eidsness
Phillips Bowen

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003

ACKNOWLEDGEMENTS

I would like to give a special thanks to Dr. Phillips for all his guidance throughout my research. He managed to always have a solution or idea for every hurdle blocking the many paths I wandered down. I am continually impressed by his wealth of knowledge in the different areas of chemistry and biochemistry that he shared with me. I would also like to thank Dr. Eidsness for all her help introducing me to the wonderful world of molecular biology and Dr. Bowen for his generous assistance and cooperation.

I owe an extra special thanks to my family for their neverending support in whatever I choose to do. Through the best of times and the worst of times, I could always count on them to listen. They would always go out of their way to assist me in whatever way they could. Thanks to my nephew, Ty Clayton, for keeping me young and reminding me of what is really important in life.

I am fortunate to have worked with some really great people while in the Phillips' lab. Extra thanks to Dr. Roman Khristoforov, Dr. Vijay Gawandi, and Santiago Lima for the many fun and enlightening times we shared both in and out of the lab. I will always remember friday happy hours, hot wings, cold drinks, and the insightful discussions about life and chemistry. Thanks to Dr. Andy Osborne for helping me to get started in the lab.

Many thanks go out to my good friends Michael Trihey and Leitha Bundrick. I could write two books on all the great conversations and adventures we shared, from

camping in the mountains to hanging out in the garage. Thanks to Heather Lynch for convincing me to move to Athens. Thanks to Athens, the 40 Watt, the Engine Room and the Caledonia Lounge too.

Thanks to all the many other friends and family who have shaped and enriched my life in so many ways.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

CHAPTER I

CHAPTER II.

LIST OF FIGURES

Page

CHAPTER II

INTRODUCTION

Enzymes are remarkable miracles of nature. They evolved over millions of years to catalyze chemical transformations at very high rates. As chemists, we continually search for methods and techniques to make molecular transformations as clean and efficient as possible. However, we are far from the synthetic efficiency of enzymes. The recent advances in biochemistry and molecular biology provide us with more insight into the powers of these enzymes. From understanding the mechanism of an enzymatic reaction to using an enzyme as a catalyst for asymmetric synthesis, we are on the verge of breaking new ground in our understanding of life on a molecular level.

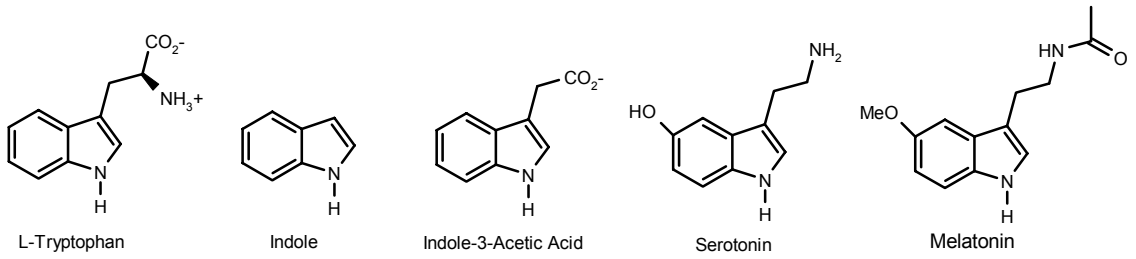
One particular enzyme I am fascinated with is tryptophan synthase. Tryptophan synthase catalyzes the final two steps in the biosynthesis of L-tryptophan.[1] Before discussing the enzyme, I will discuss its product, L-tryptophan. Tryptophan is one of the 21 amino acids found in proteins and is an essential amino acid for humans and other animals. Not only is tryptophan a building block for proteins, but it is also further metabolized into other tryptophan derivatives which are equally essential and impressive. Tryptophan metabolites range from signalling molecules, like hormones, to poisons used as a natural defense mechanism in fungi, plants and animals.

The tryptophan metabolites that act as signaling molecules are what I find most spectacular. Bacteria (*E. coli*), plants and animals all utilize tryptophan metabolites as a signaling molecule. *E. coli* metabolizes tryptophan with the enzyme, tryptophanase, to form indole, which then acts as an extracellular signaling molecule. Indole production signals bacteria to form a biofilm, which allows the bacteria to adhere to various surfaces. The biofilm allows the bacteria to adapt to environmental conditions.[2]

Not only do tryptophan metabolites act as signaling molecules in bacteria, but also in plants and other higher organisms, including humans. In plants, tryptophan is metabolized in a number of steps to indole-3-acetic acid, an auxin involved in growth stimulation.[3-5] Recent observations also suggest that indole-3-acetic acid behaves like a neurotransmitter in plants.[56] As apparent from the name of this molecule, indole is the root part of its structure.

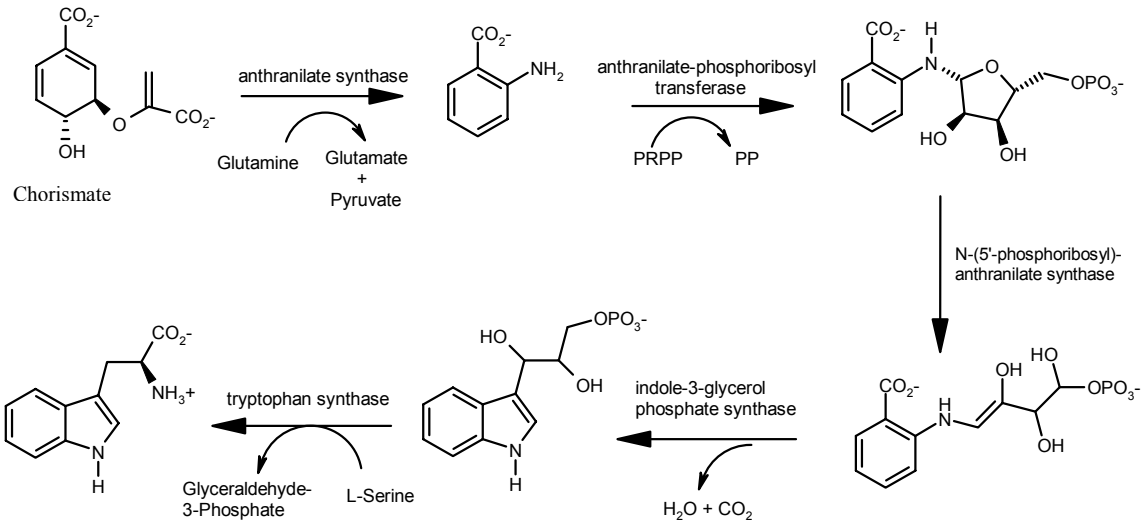
Animals also convert tryptophan into very important signaling molecules, namely serotonin and melatonin. Serotonin has been widely studied for its ablilty to maintain a sense of well-being in humans. It acts as a neurotransmitter which regulates activities from mood to appetite.[6-9] One of the current medications in the treatment of depression is a selective serotonin reuptake inhibitor, which increases the availability of serotonin in the brain.[8,10-12] Also in animals, tryptophan is converted to melatonin, which acts to regulate mammals' biological clock. The production of melatonin begins in the dark, and lets the animal know it is time to sleep.[13-15] All of the tryptophan metabolites discussed are illustrated in **Scheme 1** to illustrate the similiarity in structure, especially the indole moiety.

Scheme 1. Tryptophan Metabolites which act as signalling molecule.



Tryptophan is only synthesized in bacteria, fungi and plants. The synthesis from chorismate to tryptophan occurs from five different enzymes. These five enzymes are made from five different genes which are consecutively aligned on the genome in sequential order and termed the Trp Operon. Charles Yanovsky did much of the pioneering work to understand the sophistication of the Trp Operon.[16-19] Tryptophan is the most expensive amino acid, in terms of energy, for an organism to synthesize and therefore, the regulation of the Trp Operon is highly sophisticated.[20,21] The complete pathway for the synthesis of tryptophan can be seen in **Scheme2**. The five enzymes required for the biosynthesis of tryptophan are anthranilate synthase, anthranilate-phosphoribosyl transferase, N-(5'-phosphoribosyl)-anthranilate synthase, indole-3-glycerol phosphate synthase and tryptophan synthase.

Scheme 2. Pathway for Tryptophan Biosynthesis.

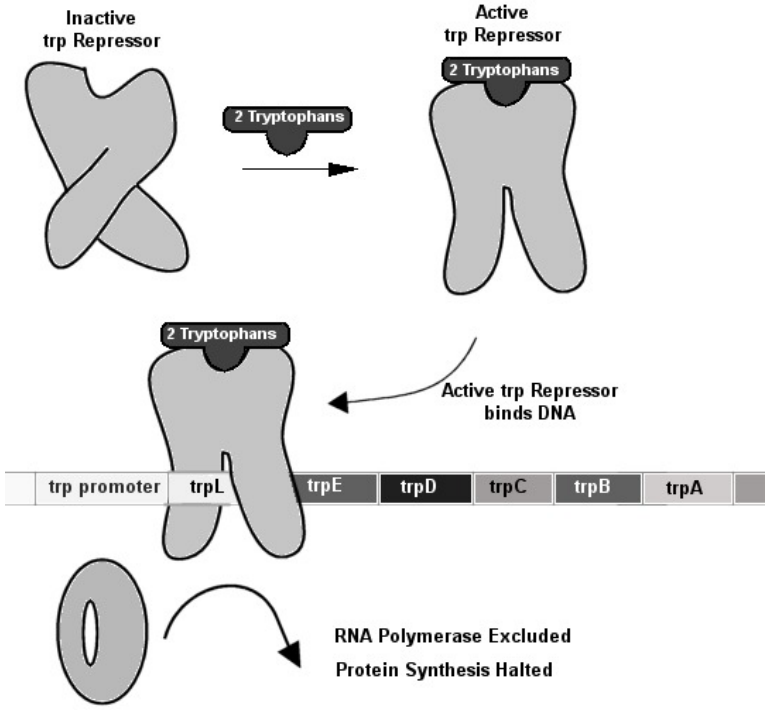


Each of the five genes for tryptophan synthesis (trpE through trpA) and a leader peptide (trpL) are transcribed at once into a single strand of mRNA, and then each enzyme is translated individually.[19] The genes trpE and trpD encode the enzyme anthranilate synthase and anthranilate-phosphoribosyl transferase, respectively. The gene trpC encodes the two enzymes N-(5'-phosphoribosyl)-anthranilate synthase and indole-3-glycerol phosphate synthase. The genes trpB and trpA encode the enzyme tryptophan synthase, where trpB is the β-subunit and trpA is the α-subunit of tryptophan synthase. When tryptophan levels are high in the organism, the synthesis of the proteins in the trp operon are repressed. Two tryptophans bind to the trp repressor protein which then binds to the trp operator located in the trp promoter. The binding of the trp repressor prevents the RNA polymerase from binding to the trp promoter, preventing further protein synthesis. However, if the levels of tryptophan are low in the organism, no tryptophan is available to bind to the trp repressor. Therefore, the tryptophan repressor remains inactive, and protein synthesis can continue.[17,22] **Figure 1** is a simplified cartoon showing the tryptophan repressor in the presence and absence of tryptophan.

Tryptophan synthesis is also regulated by another mechanism termed attenuation. Attenuation is a regulation mechanism that terminates or pauses transcription after it has already begun.[18] The attenuator in the trp operon is located in the leader peptide trpL. The trpL peptide contains two tandem tryptophan codons which test the availability of tryptophan.[23] RNA polymerase begins the transcription of the trpL peptide and pauses at hairpin in the DNA. The ribosome begins the translation of the leader peptide which allows the RNA polymerase to continue.[24] The ribosome then pauses when it comes to the tandem trp codons. If tryptophan is plentiful, the pause is short which allows a

Figure 1. Trp Repressor. Tryptophan Levels Regulate the Trp Repressor

A. Trp Repressor is Active when the levels of Tryptophan are high.



B. When Tryptophan Levels are low, the trp repressor is inactive and transcription can continue.

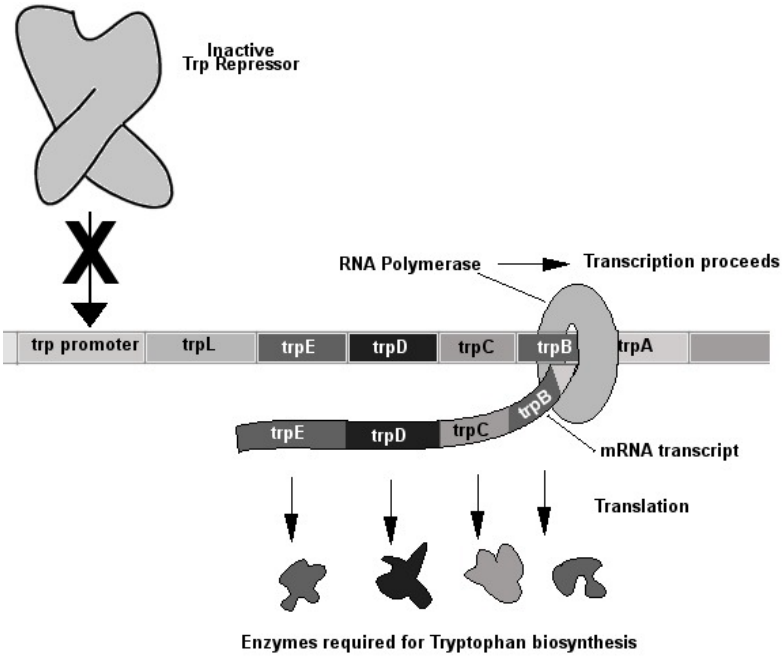
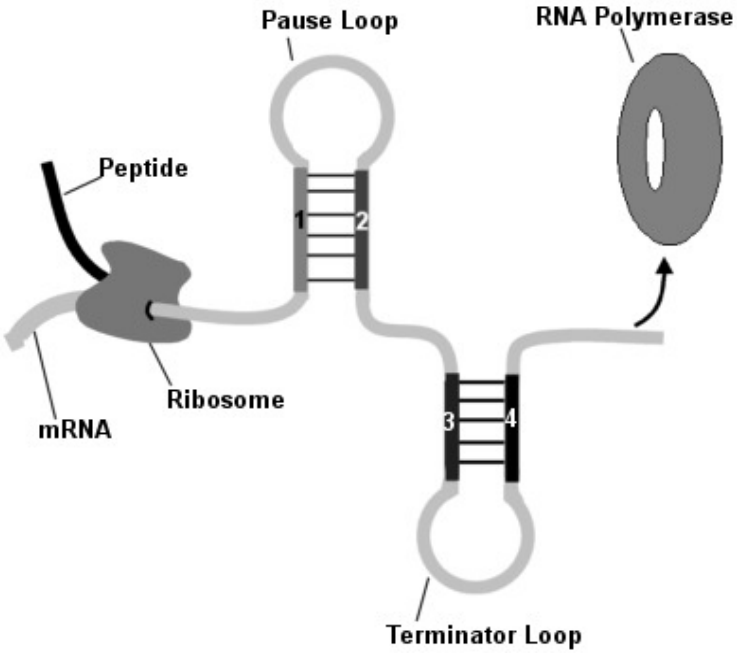
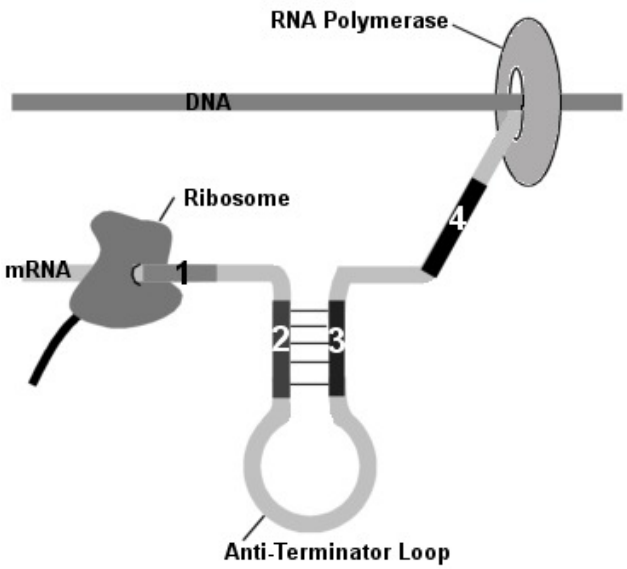


Figure 2. Attenuation Mechanism in the Trp Operon.

A. Loops 1:2 and 3:4 form when Tryptophan Levels are high causing transcription to terminate.



B. Loop 2:3 forms when tryptophan levels are low allowing transcription to continue.



terminator hairpin to form and terminate the transcription by the RNA polymerase. However if tryptophan is low, the ribosome pause lasts longer and allows an antiterminator hairpin to form which allows the RNA poymerase to continue translation of all the trp genes (trpE through trpA). The complete translation of all five enzymes for the biosynthesis of L-Tryptophan can then proceed.[18,25] **Figure 2** illustrates the attenuation of the trp operon.

As noted earlier, many bacteria posses an enzyme to catabolize tryptophan called tryptophan indole lyase (tryptophanase for short). Tryptophanase is located outside of the trp operon, but is also regulated by the availability of tryptophan.[26-28] Tryptophanase converts the tryptophan to indole and ammonium pyruvate. Since the discovery that indole is a signal molecule, there must be other factors regulating the production of tryptophanase. Indole promotes the biofilm formation in *E. coli*, which allows the organism to adhere to different surfaces.[2] Archaea, a recently discovered class of prokaryotes, do not generally contain a gene for tryptophanase. Interestingly, they contain an extra trpB gene (termed trpB2) which lies outside the trp operon. The trpB gene encodes the β-subunit of tryptophan synthase. The β-subunit of tryptophan synthase catalyzes the reaction between indole and L-serine to form L-tryptophan. The question remains, why do archaea contain an extra copy of the trpB gene, and why is it so far away from the Trp Operon? A sequence alignment between the β-subunit of tryptophan synthase from *Salmonella typhymirium* and trpB2 from *Archaeoglobus fulgidus* can be seen in **Figure 3**. We have attempted to clone and express trpB2 from *Archaeoglobus fulgidus* to characterize its activity, but the enzyme is not expressed in *E.*

Figure 3. TrpB Alignment. Alignment of the β-subunit of tryptophan synthase from *S. typhmiyrium* (2wsyb) and trpB2 from *A. fulgidus* (af1240).

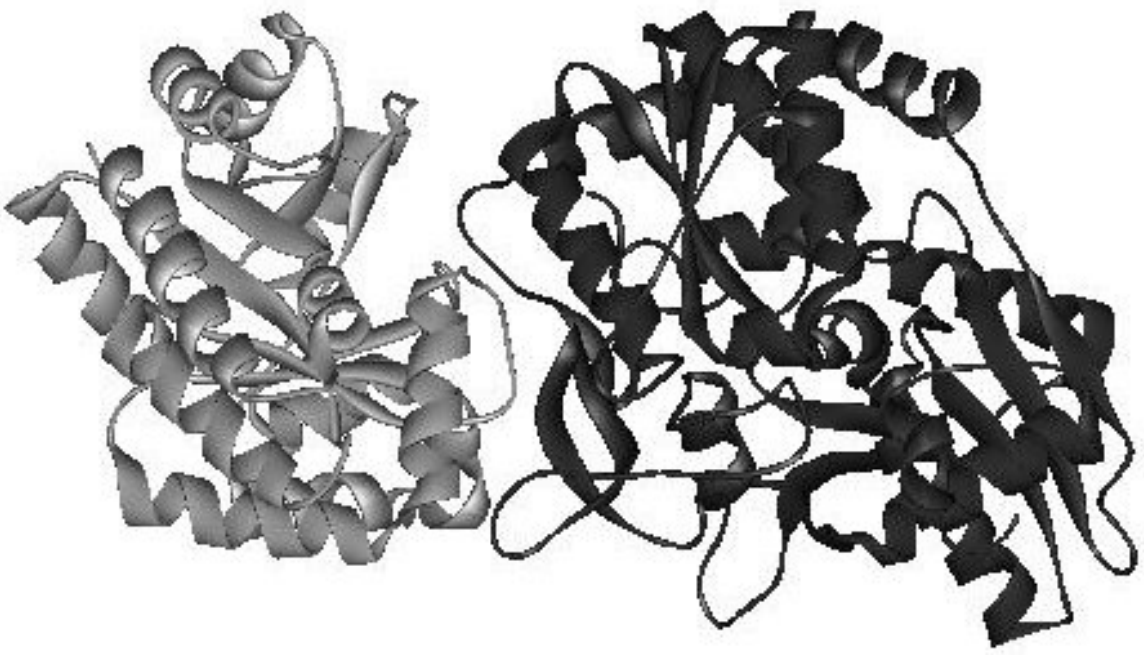
```
2wsyb   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~TTLLNPYFGEFGGMYVPQILMP....ALNQLE
af1240  MLSWCMKKIMLDESEMPKEWYNVLPDLPEPLPPPLHPATQEPVKPEDLEPIFPKGLIQQE

2wsyb   EAFVRAQKDPEFQAQFADLLKNY.AGRPTALTKCQNITAGTRT..TLYLKREDLLHGGAH
af1240  MSGERWIRIPE......DVREIYRIWRPTPLVRAERLEKALKTPARIYFKYEGASPPGSH

2wsyb   KTNQVLGQALLAKRMGKSEIIAETGAGQHGVASALASALLGLKCRIYMGAKDVERQSPNV
af1240  KPNTAVAQAYYNAKEGVERLTTETGAGQWGSALCFATKLFEMACTVYMVKVSFMQKPYRR

2wsyb   FRMRLMGAEVIPVHS.....GSATLKD..........ACNEALRDWSGSYETAHYMLGTA
af1240  VMMETWGGEVIPSPSDRTEVGRKILAENPDTPGSLGIAISEAIED.AAKNENTKYSLGSV

2wsyb   AGPHPYPTIVREFQRMIGEETKAQILDKEGRLPDAVIACVGGGSNAIGMFADFINDTSVG
af1240  LN.H.....VLLHQTVIGLETKAQ.LEKVDEKPDVLIGCVGGGSNFAGLTYPFVNDAKNG

2wsyb   ...LIGVEPGG.HGIETGEHGAPLKHGRVGIYFGMKAPM.MQTADGQIEESYSISAGLDF
af1240  DLEIIAVEPAACPTLTAGEY....KYD.FGDVAGLTPLLKMYTLGHDFIPPPIHAGGLRY

2wsyb   PSVGPQHAYLNSIGRADYVSITDDEALEAFKTLCRHEGIIPALESSHALAHALK..KMMRE
af1240  HGDAPTLCMLVKHGVIKARAVKQLPTFEAGLLFARTEGIIPAPETNHAVRAAIDEAIKAR

2wsyb   QPEKEQLLVVNLSGRGDKDIFTVHDILKARGEI~
af1240  ENNEEKVIVFGFSGHGLLDLQAYDDYLAGRLADT
```

coli. We also attempted to express trpB2 from *Pyrococcus furiosus*, but the enzyme is expressed as an insoluble inclusion body.

The trp operon is quite impressive with its sophisticated regulation mechanism. Tryptophan is equally impressive for the many biologically important metabolites into which it is converted. We would expect then, that the enzymes responsible for the biosynthesis of L-tryptophan to be remarkable. Once again, nature does not dissappoint. The enzyme, tryptophan synthase, that catalyzes the final two reactions in the biosynthesis of tryptophan is one of the most spectacular enzymes known. Tryptophan synthase is most well known for its amazing allosteric interactions and the intramolecular tunnel connecting its two subunits.[1] Tryptophan synthase is a pyridoxal-5'-phosphate (PLP) dependent enzyme that exists as an $\alpha_2\beta_2$ tetramer. The subunits are arranged as $\alpha\beta\beta\alpha$ and span about 150 Å in length.[29-31] The functional unit is the αβ dimer which can

be seen in **Figure 4**. The α-subunit converts indole-3-glycerophosphate to indole and glyceraldehyde-3-phosphate (α-reaction). Indole is then passed to the β-subunit through

Figure 4. Tryptophan Synthase αβ dimer. α-subunit light grey, β-subunit dark grey.



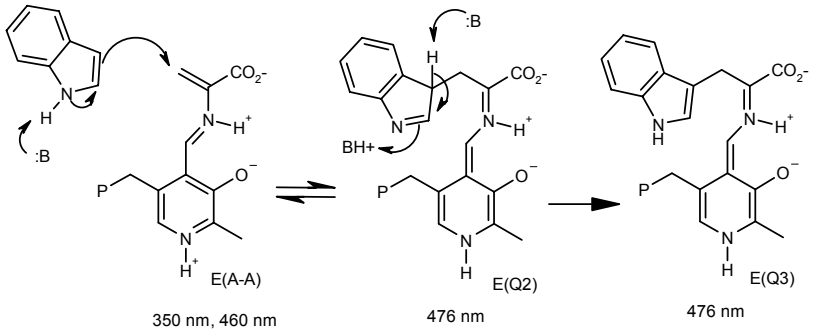
an intramolecular tunnel approximately 25 Å in length. Indole then condenses with an L-Serine-PLP-enzyme complex to form L-Tryptophan (β-reaction). The use of the intramolecular tunnel is a highly efficient and quite remarkable means of coupling two reactions. There has been a great deal of speculation regarding why such a complex system for delivering indole to the β-subunit evolved. Most suggest that the hydrophobic nature of indole make it possible to easily pass through the cell membrane where it wouldbe lost and useless.[1] Considering the fact that tryptophan is the most expensive amino acid to produce, the loss of one of the intermediates would be unfavorable.

However, since indole is a signal which regulates biofilm formation, there is a new reason for the tightly controlled transport of indole.

The allosteric interactions of tryptophan synthase have served as a model of protein-protein interactions. The α-subunit and β-subunit are able to communicate with each other to optimize the coupled reactions.[32,33] In one cycle, L-serine adds to PLP in the β-subunit. Once L-serine is bound, a signal is sent to activate the α-subunit via conformational changes. The activated α-subunit then binds its substrate, IGP, and converts it to indole. IGP enters through a flexible loop in the α-subunit. Once IGP is bound, the α-subunit closes the loop and a signal is sent to close the β-subunit to prepare for the arrival of indole.[34] Indole is then sent to the β-subunit through the hydrophobic intramolecular tunnel. Indole is then converted to L-tryptophan in the β-subunit. The enzyme opens for release of tryptophan and the whole process can then be repeated. The α– and β-subunits are bound together by hydrophobic interactions and a few salt-bridges. The salt bridges appear to play a role in the allosteric regulation. Mutations of the salt bridge residues damage the ability of the subunits to communicate.[33]

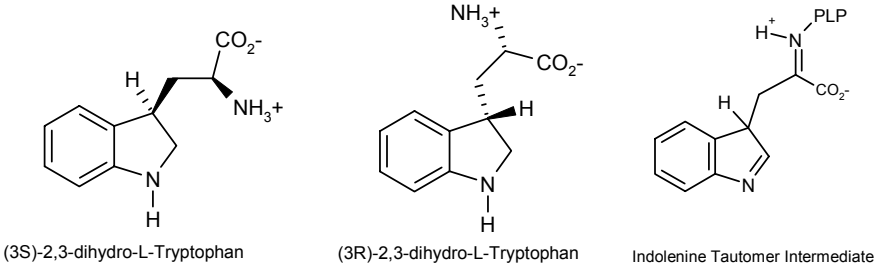
In the following chapter, we will explore part of the β-reaction mechanism. We are specifically interested in the carbon-carbon bond formation followed by tautomerization (Scheme 2). The tautomerization step is virtually irreversible. There is a great interest in why this step is nearly irreversible.[35] Tryptophanase, a PLP-dependent lyase, also

Scheme 2. β-subunit mechanism of tryptophan synthase.



catalyzes the carbon-carbon bond formation followed by tautomerization; however, the tautomerization mechanism is reversible in this case.[36] There are a two main differences that have been observed that distinguish the two different mechanisms. The first difference involves the inhibitor, 2,3-dihydro-L-Tryptophan. One diastereomer, 2,3(S)-dihydro-L-tryptophan is an inhibitor of tryptophan synthase, but not tryptophanase. The other diastereomer 2,3(R)-dihydro-L-Tryptophan is an inhibitor of tryptophanase, but not tryptophan synthase. The inhibitor 2,3-dihydro-L-Tryptophan is a transition state analogue inhibitor, because the key intermediate in each mechanism is the indolenine tautomer (**Scheme 3**). These observations suggest that the indolenine intermediate may possibly be enantiomeric in each enzyme.[37]

Scheme 3. 2,3-dihydro-L-Tryptophan and the indolenine intermediate.



The second major difference in the mechanism involves a stepwise versus concerted mechanism. The tautomerization step in each mechanism involves the deprotonation of the C-3 proton from the indolenine tautomer of tryptophan. The isotope effects of each mechanism have been explored using the substrate, 3-^{2}H-indole. When indole and 3-^{2}H-indole were reacted with tryptophanase, a kinetic isotope effect was observed, $^{H}k/^{D}k$ = 1.88. The presence of a kinetic isotope effect suggests that the addition tautomerization mechanism is concerted.[36] When indole and 3-^{2}H-indole were reacted with tryptophan synthase, no kinetic isotope was originally observed (though this issue was revisited in the next chapter). The lack of a kinetic isotope effect was interpreted to mean the addition-tautomerization mechanism was stepwise.[38]

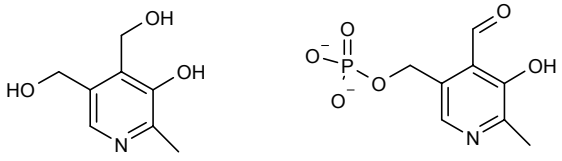
The main difference found between the two mechanisms to date is the enantiomeric indolenine tautomer of tryptophan, and a stepwise versus concerted mechanism. Many other factors contribute to the differences in the mechanism. The amino acid sequences of tryptophan synthase and tryptophanase are not homologous at all. The catalytic residues involved in each mechanism are different, as determined by the crystal structures and site-directed mutagenesis of each enzyme. Also, the other active site residues that participate in stabilizing the transition states are different, all of which can contribute to the differences in the mechanism.

The irreversible synthesis of L-tryptophan by tryptophan synthase makes it a great biocatalysis for the in-vitro synthesis of some L-tryptophan analogues. A number of L-tryptophan analogues have been synthesized using tryptophan synthase.[39] The fact that the reaction proceeds irreversibly allows for nearly quantitative yields. Biocatalysis in organic synthesis provides many advantages over other common catalysts, such as those

involving metal complexes. Enzyme catalyzed reactions are commonly done in water, which limits the amount of hazardous waste generated. The enzyme itself is environmentally friendly, unlike many of the metal catalysts. Also, the production of the enzyme is inexpensive and easy once an expression and purification method has been established.[40,41]

The two enzymes discussed, tryptophanase and tryptophan synthase, are both PLP dependent enzymes. The beauty of PLP dependent enzymes is that you can monitor their mechanisms in real time. The coenzyme, PLP, acts as an electron sink in the mechanism, and the different intermediates of the mechanism absorb light at different wavelengths. Therefore you can directly observe the formation and decay of certain intermediates as they occur.[42,43] PLP is made in the body from Vitamin B-6 (pyridoxine) (**Scheme 4**). PLP is the active form of Vitamin B-6, with two central features, the pyridine ring, which acts as an electron sink, and the aldehyde reaction center. The ε-amine of a lysine residue covalently binds PLP in the active site of the enzyme forming an internal aldimine (E(Ain)). The imine of the internal aldimine is protonated making the complex electrophilic.[44]

Scheme 4.

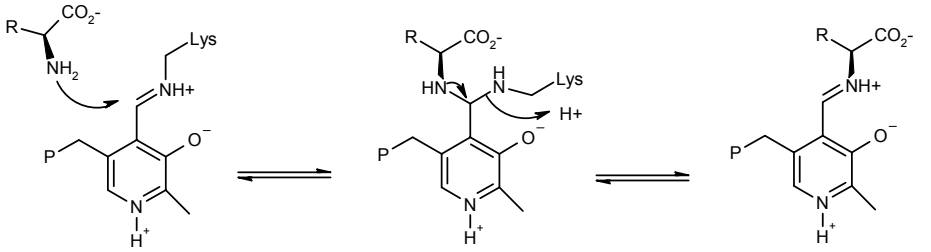


Pyridoxine Pyridoxal-5'-Phosphate

A number of different types of PLP-dependent enzymes exist, including racemases, transaminases, decarboxylases and lyases. The substrates for PLP-dependent

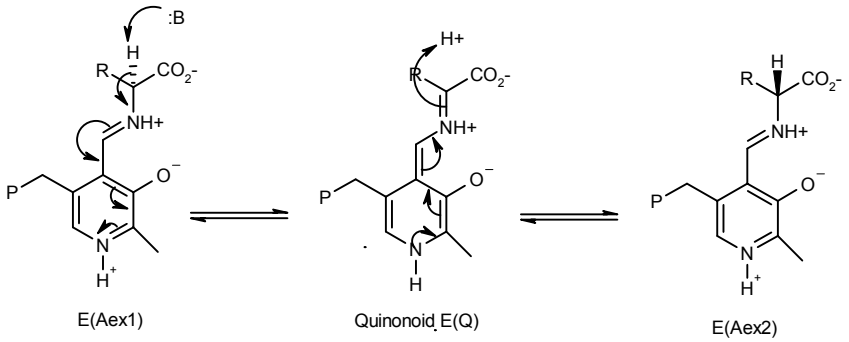
enzymes are generally α-amino acids. The mechanism of the reaction is dependent on which function the enzyme provides.[42] The mechanism always begin with the nucleophilic atttack of the amino group of the α-amino acid on the E(Ain), forming an external aldimine (E(Aex)) (Scheme 5) .

Scheme 5. Mechanism of External Aldimine Formation.



From here the mechanisms diverge, depending on whether the enzyme is a racemase, transaminase, decarboxylase or lyase. The racemases simply deprotonate the α-proton with a basic residue, followed by resonance, forming a quinonoid intermediate.

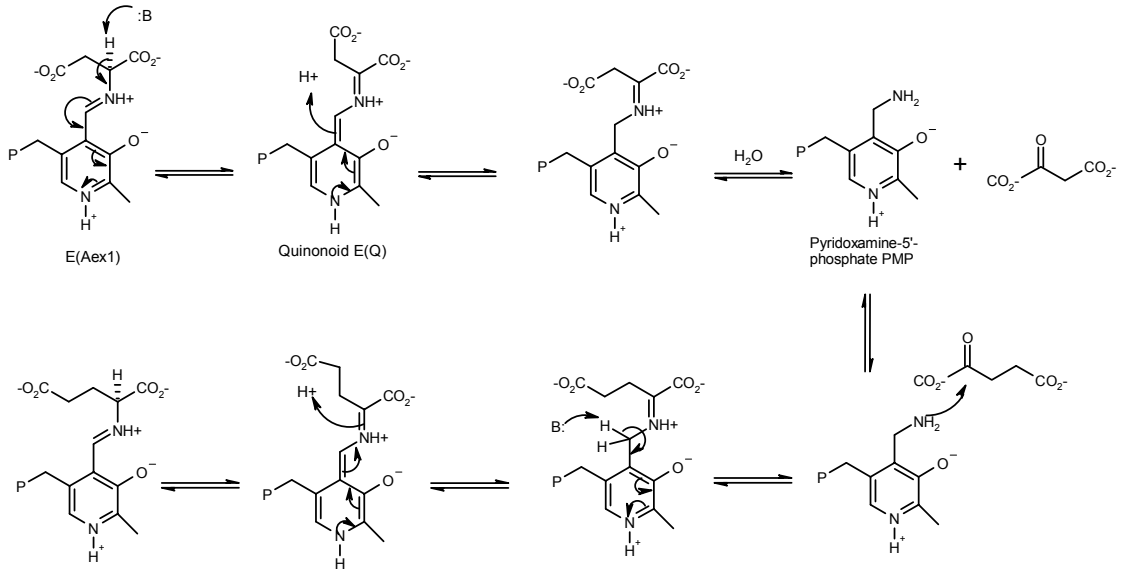
Scheme 6.



The α-position is reprotonated from the opposite face to give the D-amino acid configuration (**Scheme 6**). The enzyme lysine residue then replaces the D-amino acid by reforming E(Ain).[45]

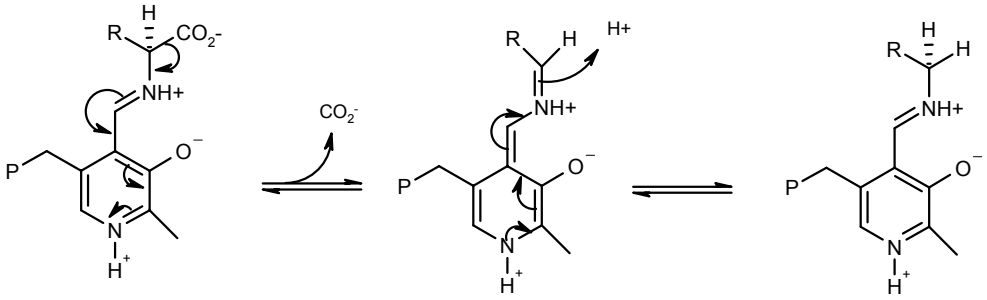
An example of a transaminase is aspartate aminotransferase (Scheme 7). The amino acid aspartate attacks the E(Ain) complex forming the E(Aex) complex. The α-proton is again removed by a basic residue, and the quinonoid internediate is once again formed. The quinonoid intermediate is then protonated on C4' adjacent to the pyridine ring. The product oxaloacetate is released, leaving pyridoxamine-5'-phosphate (PMP). The PMP intermediate then attacks the second substrate, α-ketoglutarate, forming a ketimine, which is deprotonated to give another quinonoid intermediate. The quinonoid intermediate then undergoes protonation on the α-position to give glutamate. The lysine residue of the enzyme then releases the product L-glutamate.[46-48]

Scheme 7.



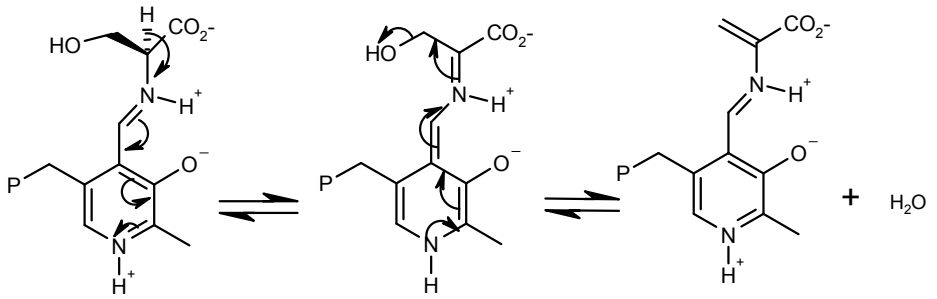
The mechanism for the decarboxylases involves the loss of CO_2 instead of deprotonation of the α-proton (Scheme 8). Beginning with the E(Aex) intermediate, CO_2 is lost, forming a quinonoid intermediate. The α-position is then protonated, followed by the release of the decarboxylated product by the lysine residue of the enzyme.[46,49]

Scheme 8.



The lyase class of PLP-dependent enzymes include tryptophanase and tryptophan synthase. The β-subunit of tryptophan synthase perform a β-elimination reaction on L-serine and a few other substrates. Tryptophanase also catalyzes the β-elimination reaction of L-serine. The β-elimination mechanism is shown in Scheme 9.[50]

Scheme 9.



The coenzyme, PLP, is a useful probe for studying enzyme mechanisms and dynamics. Unfortunately, not all enzymes have a convenient probe like PLP to investigate the enzyme. The amino acid residue, tryptophan, often serves as a probe for investigating protein structure and dynamics. The fluorescent properties of tryptophan make it a useful probe. The fluorescence spectrum of tryptophan is sensitive to the electric field of its local environment and therefore, the fluorescence spectra reveals useful information about the local environment.[51,52] A few tryptophan analogs, including 5-hydroxy-L-tryptophan, 7-aza-tryptophan and 6-azatryptophan, have been incorporated into protein as alternative fluorescent probes. Each of these offer some advantages over natural tryptophan. They each have a greater fluorescence intensity, greater sensitivity to the electric field in the environment and a red shifted fluorescence spectrum compared to natural tryptophan.[53-55] In chapter II, we explore the tautomers of the chromophoric moiety of the azatryptophans, azaindole, and to characterize the fluorescent spectra of 4-azaindole and 5-azaindole.

References:

(1) Miles, E. W. *Chem Rec* **2001**, *1*, 140-151.

(2) Di Martino, P.; Merieau, A.; Phillips, R.; Orange, N.; Hulen, C. *Can J Microbiol* **2002**, *48*, 132-137.

(3) Becker, D.; Hedrich, R. *Plant Mol Biol* **2002**, *49*, 349-356.

(4) Evans, M. L.; Rayle, D. L. *Plant Physiol* **1970**, *45*, 240-243.

(5) Tagliani, L.; Nissen, S.; Blake, T. K. *Biochem Genet* **1986**, *24*, 839-848.

(6) Price, L. H.; Charney, D. S.; Delgado, P. L.; Heninger, G. R. *Am J Psychiatry* **1991**, *148*, 1518-1525.

(7) Young, S. N.; Pihl, R. O.; Benkelfat, C.; Palmour, R.; Ellenbogen, M.; Lemarquand, D. *Adv Exp Med Biol* **1996**, *398*, 45-50.

(8) Allan, T. R.; Goldstein, L. *Conn Med* **1996**, *60*, 215-219.

(9) Young, S. N.; Leyton, M. *Pharmacol Biochem Behav* **2002**, *71*, 857-865.

(10) Amsterdam, J. D. *J Psychopharmacol* **1998**, *12*, S99-111.

(11) Gregor, K. J.; Overhage, J. M.; Coons, S. J.; McDonald, R. C. *Clin Ther* **1994**, *16*, 306-315; discussion 271-302.

(12) Sherman, T. L.; McDougle, C. J.; Price, L. H. *Conn Med* **1993**, *57*, 587-592.

(13) Nystrand, A. *Lakartidningen* **1996**, *93*, 3048.

(14) Nir, I. *Biomed Environ Sci* **1995**, *8*, 90-105.

(15) Reppert, S. M.; Weaver, D. R.; Rivkees, S. A.; Stopa, E. G. *Science* **1988**, *242*, 78-81.

(16) Denney, R. M.; Yanofsky, C. *J Bacteriol* **1974**, *118*, 505-513.

(17) Squires, C. L.; Lee, F. D.; Yanofsky, C. *J Mol Biol* **1975**, *92*, 93-111.

(18) Kuroda, M. I.; Yanofsky, C. *J Biol Chem* **1984**, *259*, 12838-12843.

(19) Morse, D. E.; Mosteller, R. D.; Yanofsky, C. *Cold Spring Harb Symp Quant Biol* **1969**, *34*, 725-740.

(20) Lee, F.; Yanofsky, C. *Proc Natl Acad Sci U S A* **1977**, *74*, 4365-4369.

(21) Oppenheim, D. S.; Bennett, G. N.; Yanofsky, C. *J Mol Biol* **1980**, *144*, 133-142.

(22) Shimizu, Y.; Shimizu, N.; Hayashi, M. *Proc Natl Acad Sci U S A* **1973**, *70*, 1990-1994.

(23) Landick, R.; Yanofsky, C.; Choo, K.; Phung, L. *J Mol Biol* **1990**, *216*, 25-37.

(24) Landick, R.; Carey, J.; Yanofsky, C. *Proc Natl Acad Sci U S A* **1985**, *82*, 4663-4667.

(25) Yanofsky, C.; Kelley, R. L.; Horn, V. *J Bacteriol* **1984**, *158*, 1018-1024.

(26) Yanofsky, C.; Horn, V.; Gollnick, P. *J Bacteriol* **1991**, *173*, 6009-6017.

(27) Nurmikko, V.; Pyykko, K. *Acta Chem Scand* **1971**, *25*, 2311-2319.

(28) Stewart, V.; Yanofsky, C. *J Bacteriol* **1985**, *164*, 731-740.

(29) Hyde, C. C.; Ahmed, S. A.; Padlan, E. A.; Miles, E. W.; Davies, D. R. *J Biol Chem* **1988**, *263*, 17857-17871.

(30) Lane, A. N.; Kirschner, K. *Eur J Biochem* **1983**, *129*, 675-684.

(31) Miles, E. W. *Adv Enzymol Relat Areas Mol Biol* **1979**, *49*, 127-186.

(32) Dunn, M. F.; Agular, V.; Drewe, W. F., Jr.; Houben, K.; Robustell, B.; Roy, M. *Indian J Biochem Biophys* **1987**, *24*, suppl 44-51.

(33) Weber-Ban, E.; Hur, O.; Bagwell, C.; Banik, U.; Yang, L. H.; Miles, E. W.; Dunn, M. F. *Biochemistry* **2001**, *40*, 3497-3511.

(34) Miles, E. W. *J Biol Chem* **1991**, *266*, 10715-10718.

(35) Ahmed, S. A.; Martin, B.; Miles, E. W. *Biochemistry* **1986**, *25*, 4233-4240.

(36) Phillips, R. S.; Sundararaju, B.; Faleev, N. G. *Journal of the American Chemical Society* **2000**, *122*, 1008-1014.

(37) Phillips, R. S.; Miles, E. W.; Cohen, L. A. *J Biol Chem* **1985**, *260*, 14665-14670.

(38) Lane, A. N.; Kirschner, K. *Eur J Biochem* **1983**, *129*, 571-582.

(39) Sloan, M. J.; Phillips, R. S. *Bioorganic & Medicinal Chemistry Letters* **1992**, *2*, 1053-1056.

(40) Schoemaker, H. E.; Mink, D.; Wubbolts, M. G. *Science (Washington, DC, United States)* **2003**, *299*, 1694-1698.

(41) Saha, B. C.; Demirjian, D. C. *ACS Symposium Series* **2001**, *776*, 2-12.

(42) Hayashi, H. *Journal of Biochemistry (Tokyo)* **1995**, *118*, 463-473.

(43) Hayashi, H.; Wada, H.; Yoshimura, T.; Esaki, N.; Soda, K. *Annual Review of Biochemistry* **1990**, *59*, 87-110.

(44) Bugg, Tim; An Introduction to Enzyme and Coenzyme Chemistry; Oxford Mead, Oxford; Blackwell Sciences Ltd.; 1997

(45) Faraci, W. S.; Walsh, C. T. *Biochemistry* **1988**, *27*, 3267-3276.

(46) Smith, D. M.; Thomas, N. R.; Gani, D. *Experientia* **1991**, *47*, 1104-1118.

(47) Julin, D. A.; Kirsch, J. F. *Biochemistry* **1989**, *28*, 3825-3833.

(48) Churchich, J. E.; Farrelly, J. G. *J Biol Chem* **1969**, *244*, 3685-3690.

(49) Abell, L. M.; O'Leary, M. H. *Biochemistry* **1988**, *27*, 5927-5933.

(50) Mahdi, J. G.; Kelly, D. R. *Biotechnology (2nd Edition)* **2000**, *8b*, 41-171.

(51) Clayton, A. H. A.; Sawyer, W. H. *European Biophysics Journal* **2002**, *31*, 9-13.

(52) Weljie, A. M.; Vogel, H. J. *Methods in Molecular Biology (Totowa, NJ, United States)* **2002**, *173*, 75-87.

(53) Li, Q.; Du, H.-N.; Hu, H.-Y. *Biopo lymers* **2003**, *72*, 116-122.

(54) Twine, S. M.; Murphy, L.; Phillips, R. S.; Callis, P.; Cash, M. T.; Szabo, A. G. *Journal of Physical Chemistry B* **2003**, *107*, 637-645.

(55) Mohammadi, F.; Prentice, G. A.; Merrill, A. R. *Biochemistry* **2001**, *40*, 10273-10283.

(56) Baluska, F.; Samaj, J.; Menzel, D. *Trends Cell Biol.* **2003,** 6, 282-285.

CHAPTER I

KINETIC ISOTOPE EFFECT IN THE β-SUBUNIT MECHANISM OF TRYPTOPHAN SYNTHASE

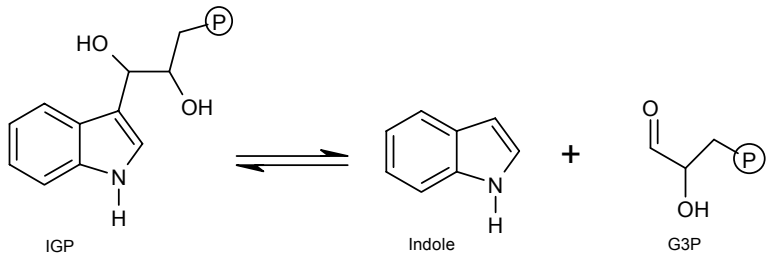
Introduction.

The enzyme, tryptophan synthase, has captured the imagination of many scientists. Tryptophan synthase is an ideal model for illustrating the beauty of protein-protein interactions, allosteric interactions and substrate channeling.[1,2] The tetrameric structure of tryptophan synthase from *Salmonella typhimurium* exists as an $\alpha_2\beta_2$ dimer, with the substrates arranged as αββα, spanning approximately 150 Å in diameter.[3-5] The overall reaction that tryptophan synthase catalyzes is the conversion of indole-3-glycerol phosphate and L-serine to L-tryptophan and glyceraldehyde phosphate.[6-8] Each subunit catalyzes a different reaction. The α-subunit substrate, indole-3-glycerol phosphate, is cleaved to form indole and glyceraldehyde-3-phosphate. Indole is channeled from the α-subunit to the β-subunit through an intramolecular tunnel approximately 25 Å in length.[1,9,10] Indole then condenses with a preformed aminoacrylate intermediate composed of L-serine and pyridoxal-5-phosphate (PLP) to form the product L-tryptophan.[11] The overall reaction is orchestrated by a series of allosteric interactions between the α- and β-subunits, which switches each subunit to an open (low-activity) conformation or a closed (high-activity) conformation.[2,12]

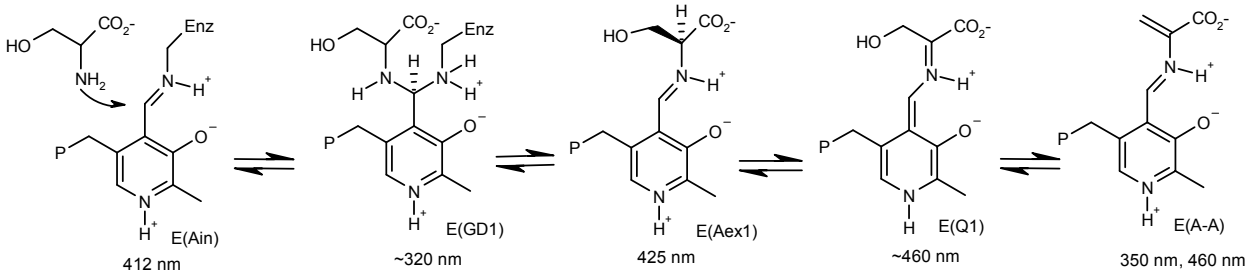
Within the last ten years, tryptophan synthase was found to be regulated by the presence or absence of different monovalent cations (MVC).[13,14] With the advances in high resolution crystallography, there are more and more proteins that appear to be regulated by MVCs. These studies show that the metal ions usually act as an allosteric effector and the metal binding site is greater than 6 Å away from the substrate active site. The activity of tryptophan synthase is greatly effected by the presence of different metal ions. The presence of Na^+ and K^+ stabilizes the open (low activity) conformation, while Cs^+ and NH_4^+ stabilize the closed (high activity) conformation of the β-subunit.[15-17] The presence or absence of each of these MVCs alters the conformation of the β-subunit as observed in crystallographic structures.[14]

Scheme 1

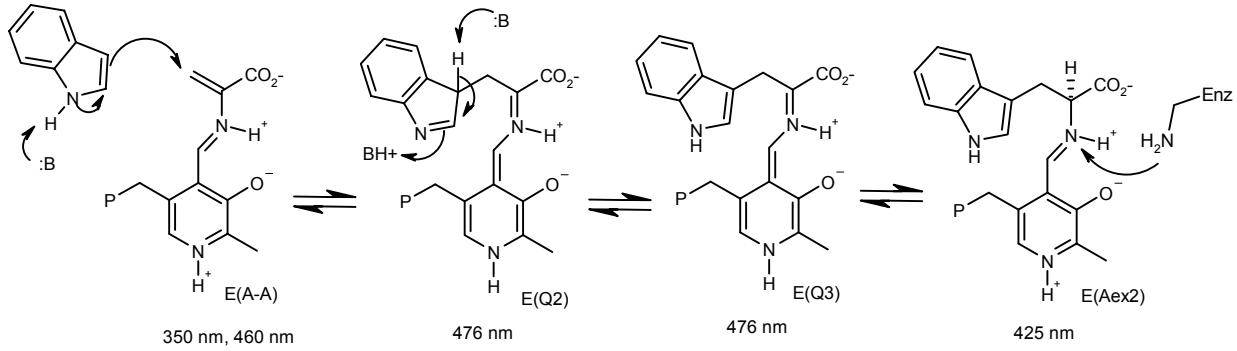
α-Reaction

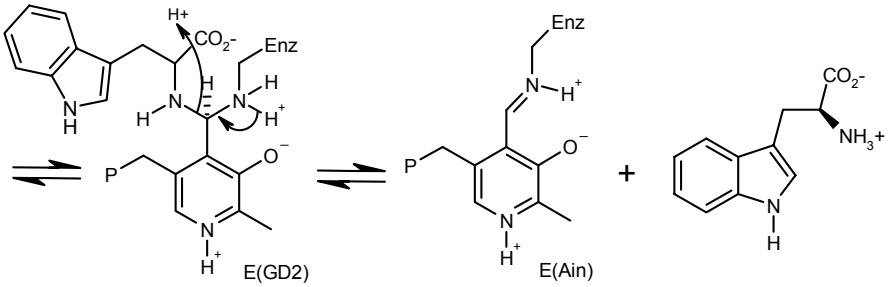


Stage I β-Reaction Mechanism



Stage II β-Reaction Mechanism



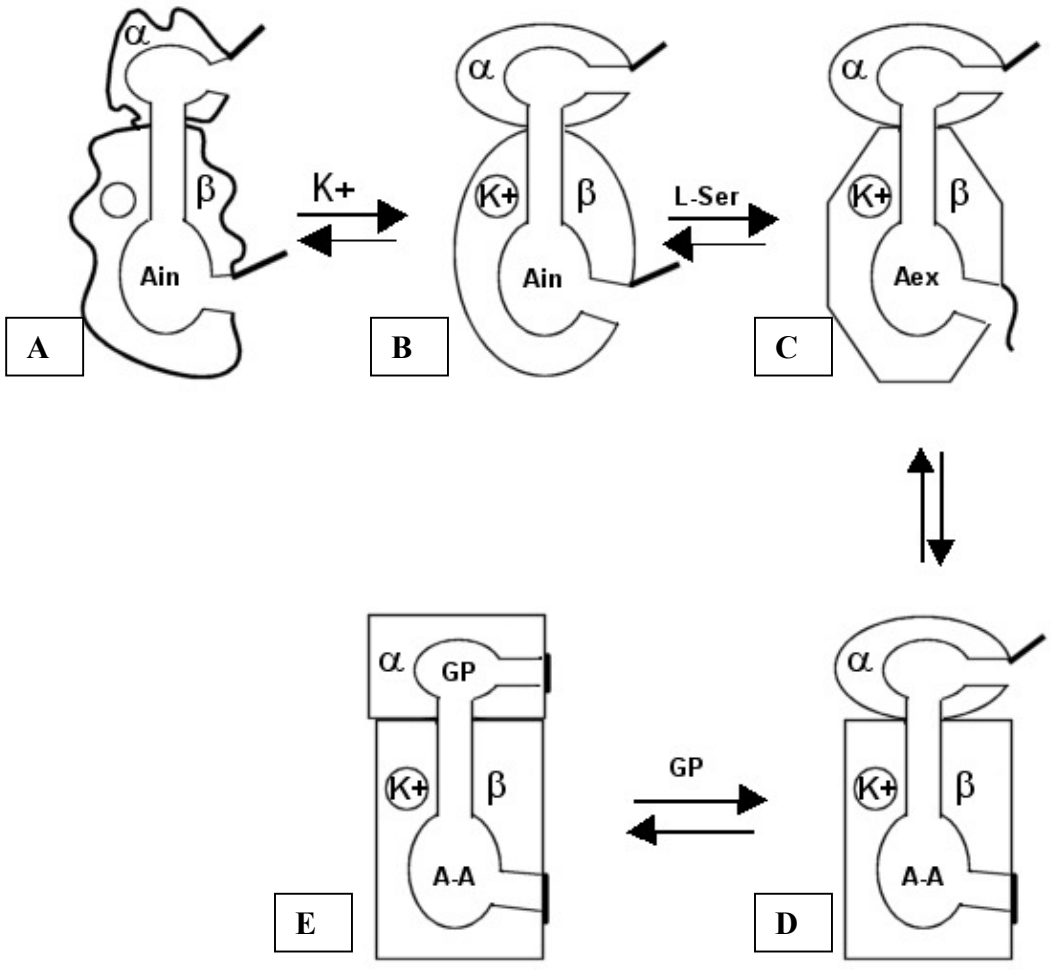


Substrates or ligands of the α-subunit activate the β-subunit to bind the substrate L-serine. Once the β-subunit binds L-serine, the α-subunit is activated to cleave its substrate, and channel the product, indole, to the β-subunit. Indole then condenses with the aminoacrylate intermediate to form L-tryptophan, which again activates the α-subunit to open and bind another substrate.[18] The coenzyme, pyridoxal-5'-phosphate (PLP), provides a probe for monitoring the reaction and activity. The reaction intermediates bound to PLP have been well characterized, and the rates of formation and decay of the different intermediates provide a wealth of information.[11] By taking advantage of the allosteric interactions, we can monitor the activity of different steps in the mechanism. For example, the α-subunit ligand α-glycerophosphate (GP) binds in the α-subunit, though no reaction is observed. Instead, α-glycerophosphate remains trapped in the α-subunit, activating the β-subunit to bind L-serine and form the aminoacrylate intermediate which closes the β-subunit. Also, since α-glycerophophate remains in the α-subunit, it also remain closed which makes it difficult for the β-subunit substrate, indole, to enter.[1,19,20] These allosteric interaction provide us with tools for exploring the β-subunit mechanism.

Stage II of the β-subunit mechanism has been well established, although a question remains regarding one step of the mechanism. Indole reacts with the aminoacrylate intermediate to form quinonoid 2 ($E(Q_2)$), followed by deprotonation at C-3 of indole to form quinonoid 3 ($E(Q_3)$). The reaction intermediates E(Q2) and $E(Q_3)$) are indistinguishable spectroscopically. Therefore, a direct observation of either intermediate has not been possible.[11] However, this stage in the mechanism is critical, because the reaction is virtually irreversible. In fact, another enzyme exists, tryptophanase, simply to catalyze the reverse reaction, the cleavage of L-tryptophan to indole and ammonium pyruvate.[21] The question remains, what is the difference in the mechanisms that allow one enzyme to catalyze the reaction reversibly and the other irreversibly. Lane and Kirschner investigated stage II of the β-subunit mechanism by

exploring the possibility of a Kinetic Isotope Effect (KIE) on the deprotonation step. [11] They did not observe a KIE, and concluded that a step wise mechanism of addition-deprotonation occurred in the mechanism. However, Tryptophanase has a large KIE in the deprotonation step, and the mechanism is considered to be concerted.[21] These differences in the two mechanisms may suggest a possible reason for the reversibility issue.

Scheme 2. Simplified cartoon illustrating the conformational changes caused by different allosteric ligands and MVCs. Enzyme conformations D and E are explored in this work by pre-equilibrating tryptophan synthase with L-Serine and a MVC (D) and pre-equilibration of the tryptophan synthase with L-Serine and Na2GP (E).



In this work, we reinvestigated the β-subunit mechanism, by exploring the KIE in tryptophan synthase utilizing recent observations concerning allosteric regulation. Scheme 2 illustrates the active states of tryptophan synthase in the presence of MVCs, L-Serine and α-glycerophosphate (GP). The substrate indole enters the β-subunit active site via the opening in the α-subunit. In the presence of only MVCs and L-serine, the α-subunit is open to receive and tunnel indole to the beta subunit. In the presence of MVCs, L-serine and GP, the α-subunit is closed which makes it more difficult for indole to enter the β-subunit. .[1,20] We found that the KIE is observed in the reaction of indole with the α-aminoacrylate intermediate, and the KIE is affected by the MVCs.

Material and Methods.

Synthesis of 3-^{2}H-Indole.

Indole (1g) was added to 20 mL of 0.01 M DCl in D_2O. The mixture was heated to 60 °C and stirred for 4 hours. The reaction mixture was allowed to cool, and was then extracted twice with 10 mL of diethyl ether. The organic layers were combined and evaporated to give 0.8 g of 3-^{2}H-indole. NMR analysis showed that the product was 92% deuterated by the disapperance of the C-3 proton peak at 6.5 ppm. The product was then treated with the 0.01 M DCl in D_2O twice more as stated above, yielding 0.6 g of >95% dideutero N,3-^{2}H-indole. The product was then dissolved in methanol and allowed to sit stirring for one hour to exchange the N1 deuterium for hydrogen. The methanol was evaporated leaving 0.6 g of 3-^{2}H-indole.

Materials.

L-Serine, sodium α-glycerophosphate, indole, NaCl, KCl, CsCl, NH_4Cl, triethanolamine, were purchased from Sigma. All experiments were carried out in 50 mM triethanolamine buffer at pH 8. The temperature of all experiments were carried out between 24-25 °C,

with less than 0.1 °C difference between the individual indole and 3-^{2}H-indole experiments. The enzyme Tryptophan Synthase from *Salmonella typhimurium* was a generous gift from Dr. Edith W. Miles from The National Institutes of Health. Tryptophan Synthase was dialyzed on a column to remove the potassium phosphate buffer. The enzyme was eluted with 50 mM triethanolamine hydrochloride at pH 8.0.

Spectroscopic methods.

Rapid Scanning Stopped Flow Spectrophotometry (RSSF) was used to monitor the reaction of the addition of indole to the pre-equilibrated L-serine-enzyme complex. All reactions were monitored for 1 second at the rate of 1000 scans per second. The RSSF instrument (RSM-1000) and data analysis software (Globalworks) were made by OLIS, located in Bogart, GA. The concentrations of indole were verified by a Cary 1E UV/Vis spectrophotometer using the extinction coefficient 5500. The concentration of tryptophan synthase was adjusted to 1.5 mg/mL for the experiments involving Na_2GP, Na^+ and K^+. For the experiments involving Cs^+ and NH_4^+, the concentration was adjusted to 3 mg/mL. The concentration of the MVC was constant at 100 mM. The concentration of L-serine was saturating at 40 mM. The concentration of the indoles was varied from 50 μM to 800 μM in the KIE experiments.

Results.

Equilibrium of E(Aex) and E(A-A) at Time Zero. Stage II of the β-subunit mechanism was investigated by RSSF in the presence of different MVCs and α-subunit ligands. One syringe contained tryptophan synthase (1.5 mg/mL), L-serine (40 mM), triethanolamine hydrochloride buffer (50 mM, pH=8.0), and one of the four MVCs (200 mM). The other syringe contained indole (50 μM to 800 μM), L-serine (40 mM), triethanolamine buffer (50 mM, pH=8.0) and the same MVCs (200 mM) as the enzyme syringe. The four monovalent cations used were Na^+, K^+, NH_4^+ and Cs^+. The α-subunit

ligand was α-glycerophosphate which was used as the disodium salt. Figure 1 shows the spectra at zero time in the presence of each of the allosteric effectors. Since L-serine and the enzyme were premixed prior to injection, the primary species observed at zero time is the aminoacrylate intermediate (350 nm). Table 1 shows the absorbance values of aminoacrylate (350 nm) and external aldimine (425 nm) at zero time. The external aldimine present at zero time could be a mixture of the L-serine external aldimine (E(Aex1)) and the indolenine external aldimine ($E(Aex_2)$). The presence of Na^+ and GP strongly favors the E(A-A) intermediate at zero time of the reaction.

Table 1. Absorbance Values of Intermediates at Time Zero

MVC	GP	Abs. 350 nm	Abs. 425 nm
Na^+	-	0.10	0.044
Na^+	+	0.076	0.025
K^+	-	0.073	0.05
Cs^+	-	0.10	0.053
$NH4^+$	-	0.058	0.03

Conentration Dependence of Indole and 3-^{2}H-Indole in the Presence of Na^+. The concentration dependence of the reaction of indole and 3-^{2}H-indole in the presence of Na^+ can be seen in Figure 2. The observed rates of the reaction were determined by global analysis between 320 nm and 550 nm with a two species fit. The rate for the formation of the first species involved the decay of E(A-A) at 350 nm and the formation of $E(Q_3)$ at 476 nm. The rate for the second species involved the decay of $E(Q_3)$ and the formation of the $E(Aex_2)$ at 425 nm. The rate of the reaction for the formation of $E(Aex_2)$ was constant at all concentrations of indole with an average rate of 20 sec^{-1} for 3-^{2}H-indole and 22 sec^{-1} for indole. The observed rate for the formation of $E(Q_3)$ increased hyperbolically with increasing indole concentration. Figure 2 shows the concentration dependence of the reaction of indole and 3-^{2}H-Indole in the presence of

Na^+ . There is clearly a significant difference in rates of reaction of indole and 3-^{2}H-indole. The concentration dependence was fit with the best fit theoretical curve according to Equation 1. The V_{max} for indole was determined to be 308 s^{-1} and the V_{max} for 3-^{2}H-indole was 140 s^{-1}. The K_I for Indole was 204 μM and 103 μM for 3-^{2}H-indole. The scans for the first 100 ms of the reaction can be seen in figure 3. The formation of E(Q3) at 476 nm, and the subsequent decay can clearly be seen. An isobestic point at 450 nm can also be seen between two of the scans. The time course for the formation and decay at 476 nm can be seen in Figure 4. The faster formation for indole can be seen along with the converging decay rate.

$$k_{obs} = \frac{V_{\max}[I]}{(K_I + [I])} \qquad (1)$$

Concentration Dependence of Indole and 3-^{2}H-Indole in the Presence of K^+. The reaction between the aminoacrylate intermediate of tryptophan synthase and indole in the presence of K^+ behaves quite similarly to the reaction in the presence of Na^+. The observed rates for the reaction were determined by global analysis between 320 and 550 nm. Again, a two species fit was used, where the first species corresponded to E(Q3), at 476 nm, and the second species formed was $E(Aex_2)$, at 425 nm. The observed rate constant for the formation of $E(Aex_2)$ is again constant across all concentrations of indole with an average of 33 ± 3 s^{-1} for 3-^{2}H-indole and 36 ± 3 s^{-1} for indole. The observed rates constants for the formation of E(Q2) increase hyperbolically with increasing indole concentration, as seen in Figure 5. The data were fit to Equation 1. The V_{max} for indole is 388 s^{-1} and the V_{max} for 3-^{2}H-indole is 193 s^{-1}. Here, in the presence of K^+, each rate constant is faster than those in the presence of Na^+. Once again there is a significant difference between the V_{max} of indole versus 3-^{2}H-indole. The KIE is somewhat smaller than the reaction in the presence of Na^+ at $k^H/k^D = 2.0$. The K_I for the reaction in the presence of K^+ is 214 μM for 3-^{2}H-indole and 251 μM for indole. The difference between the K_I values is not as great as it is in the presence of Na^+, but once again indole

has the higher K_I. Figure 6 shows the scans for the first 100 ms of the reaction. As was the case with Na^+, the formation and decay of $E(Q_3)$ can clearly be seen at 476 nm with an isobestic point again at 450 nm. The time course for the formation and decay at 476 nm can be seen in Figure 7. The faster rate for the reaction with indole can be seen followed by the convergence during the decay.

Concentration Dependence of Indole and 3-^{2}H-Indole in the Presence of NH_4^+. The reaction of the aminoacrylate complex of tryptophan synthase and indole in the presence of NH_4^+ behaves differently from Na^+ and K^+. The reaction in the presence of NH_4^+ is much faster than in the presence of Na^+ and K^+. Attempts at a global analysis using two species between the 320 nm and 550 nm gave rate constants for the first phase and the second phase that were always the same. However, by monitoring the reaction at 476 nm, we were able to determine the rate constant for the formation and decay of the 476 nm species. Again, the rate of the decay remained constant at all concentrations of indole. The average rate for indole was $9 \pm 7\ s^{-1}$ and the average rate for 3-^{2}H-indole was also $9 \pm 8\ s^{-1}$. The observed rate constant of the formation of $E(Q_3)$ at 476 nm increases hyperbolically with increasing indole concentration as seen in Figure 8. Equation 1 was again used for the best fit. The V_{max} for indole was $556\ s^{-1}$ and the V_{max} for 3-^{2}H-indole was $264\ s^{-1}$. The individual V_{max} for indole and 3-^{2}H-indole are nearly double the rates in the presence of Na^+. The KIE is comparable to the Na^+ and K^+ reactions at $^Hk/^Dk = 2.1$. The K_I for the reaction in the presence of NH_4^+ was 517 μM for 3-^{2}H-indole and 665 μM for indole. The K_I for indole is again greater than that of 3-^{2}H-indole. Also, the K_I's in the presence of NH_4^+ are three times or more the K_I's in the presence of Na^+ and K^+. Figure 10 shows the scans for the first 100 ms of the reaction. In the presence of NH_4^+, a distinct formation and decay of the quinonoid species at 476 nm is no longer observed. The reaction has become so fast that the tryptophan external aldimine E(Aex2) is being

formed as fast as $E(Q_3)$. The time course at 476 nm can be seen in Figure 11. The time course shows the formation of the quinonoid species, but as seen in the scans, there is no apparent decay.

Concentration Dependence of Indole and 3-^{2}H-Indole in the Presence of Cs^+. The reaction of indole and the aminoacrylate complex of tryptophan synthase in the presence of Cs^+ appears to be behave intermediate between K^+ and NH_4^+, based on the time course at 476 nm seen in Figure 12. Unlike the reaction in the presence of NH_4^+, the formation and decay of $E(Q_3)$ is observed. However, the decay amplitude is much smaller than in the presence of Na^+ and K^+. The reaction appears to be fast, and all attempts at global analysis and single wavelength fits gave indentical rates for both phases, indicating a coupled reaction. The observed rates did show a hyperbolic increase with increasing indole concentration. The rate of the reaction with indole is significantly faster than the rate of the reaction with 3-^{2}H-indole. Actual V_{max} and K_I values could not be calculated due to the presence of the coupled reaction. Figure 13 shows the scans for the first 50 ms of the reaction. $E(Q_3)$ and E(Aex2) appear to be forming simultaneously due to the increased rate. An isosbestic point at 380 nm is seen between the decaying E(A-A) band at 350 nm and the formation of $E(Aex_2)$ at 425 nm.

Concentration Dependence of Indole and 3-^{2}H-Indole in the Presence of Na^+ and a-Glycerophosphate (GP). The reaction of indole and the aminoacrylate intermediate of tryptophan synthase in the presence of both Na^+ and GP is completely different from the reactions with only an MVC present. GP is an α-subunit ligand that stabilizes a closed conformation of the α-subunit, which is the entrance point for indole to reach the β-subunit. The binding of indole is much slower, due to the blocked passage way in the α-subunit and other conformational changes that are induced in the β-subunit. The observed rates for the reaction of tryptophan synthase with indole and 3-^{2}H-indole were

determined by global analysis between 320 nm 550 nm. In this case, a one species fit was used because of the quasi-stable quinonoid peak formed at 476 nm. The observed rates for the formation of $E(Q_3)$ increase linearly with increasing indole concentration, as seen in Figure 14. There is virtually no difference between the observed rates of indole and 3-^{2}H-indole. Figure 15 shows the scans from the first 200 ms of the reaction. The two primary intermediates observed are E(A-A) at 350nm and $E(Q_3)$ at 476 nm. A clear isobestic point is seen again at 380 nm. The $E(Aex_2)$ intermediate is also observed as a shoulder on the $E(Q_3)$ band. The time course for the reaction at 476 nm is seen in figure 16. The figure shows the slow formation of $E(Q_3)$ with no difference between the reaction of indole and 3-^{2}H-indole.

Discussion.

The reaction of indole with the aminoacrylate intermediate is a biphasic reaction. The first phase is the rapid formation of the two quinonoid species followed by a slow phase which is the decay of the quinonoid $E(Q_3)$ to form the external aldimine of L-Tryptophan. A simple mechanism for the reaction of indole with the aminoacrylate intermediate of tryptophan synthase in the relevant portion of Stage II of the β-subunit reaction is shown in Equation 2.

$$\text{E(A-A) + Indole} \underset{}{\overset{K_1}{\rightleftharpoons}} \text{E(A-A)-Indole} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \text{E}(Q_2) \xrightarrow{k_3} \text{E}(Q_3) \qquad (2)$$

After the formation of the non-covalent tryptophan synthase aminoacrylate-indole complex (E(A-A)-Indole), indole undergoes a covalent reaction with E(A-A), via a Michael type addition, to form the indolenine quinonoid species $E(Q_2)$. The indolenine quinonoid intermediate, $E(Q_2)$, then tautomerizes to form the indole quinonoid intermediate $E(Q_3)$. $E(Q_2)$ and $E(Q_3)$ both absorb light at the same wavelength (476 nm) preventing a direct observation of each intermediate. The conversion of $E(Q_2)$ to $E(Q_3)$ is

considered to be virtually irreversible, with the assumption that tautomerization is the barrier to reversal of the reaction. In exploring the KIE of this part of the β-subunit mechanism, we were interested in the first rapid phase. This phase includes the formation of both $E(Q_2)$ and $E(Q_3)$ which are the isotopically sensitive steps.

Based on the mechanism in Equation 2, a rate equation can be derived as seen in Equation 3 assuming a quasi-steady state on $E(Q_2)$.

$$k_{obs} = \frac{k_2 k_3 [Indole]}{(k_{-2} + k_3)(K_1 + [Indole])} \qquad (3)$$

The rate constant k_3 is the isotopically sensitive step in the mechanism, so the expression of the isotope effect depends on the relative magnitude of k_{-2} and k_3. If the rate of k_3 is much greater than the rate k_{-2}, then Equation 3 reduces to Equation 4.

$$k_{obs} = \frac{k_2 [Indole]}{K_1 + [Indole]} \qquad (4)$$

Since k_3 has dropped out of the equation, no kinetic isotope effect is expected in this situation. On the other hand, if the rate of k_{-2} is much greater than the rate of k_3, then Equation 3 is reduced to Equation 5.

$$k_{obs} = \frac{k_2 k_3 [Indole]}{k_{-2}(K_1 + [Indole])} \qquad (5)$$

Now that the isotopically sensitive step is part of the equation an isotope effect may be observed.

Table 2 shows the values for the combined rate constants derived from Equations 4 and 5. The values were obtained from the concentration dependence studies involving

indole and 3-^{2}H-indole between the concentrations of 50 μM and 800 μM. The reaction between indole (and 3-^{2}H-indole) with the E(A-A) intermediate in the presence of MVCs does produce an observable KIE. The magnitude of the KIE is approximately 2 in the presence of Na^{+}, K^{+} and NH_4^{+}. The presence of a small isotope effect suggests that the tautomerization is indeed partially a rate limiting step. When we change the enzyme conformation by the addition of α-glycerophosphate, a KIE is no longer observed. The presence of the α-subunit ligand apparently slows the addition of indole to the preformed aminoacrylate intermediate. The rate for the formation of E(A-A) may have also slowed, which adds to the absence of an observable KIE.

Lane and Kirschner originally proposed a stepwise mechanism between $E(Q_2)$ and $E(Q_3)$ due to the absence of a KIE on the reaction of 3-^{2}H-indole.[11] At the time, little was known about the role of MVCs in enzymes. More than 100 enzymes have been shown to be regulated by the presence of metal ions. [17] We were able to use these effects as a tool to observe a KIE. The fact that a KIE is observable in the case where no α-subunit is present, but a KIE is not observable in the presence of α-subunit ligands suggests that conversion of $E(Q_2)$ to $E(Q_3)$ is a stepwise process which supports the originally proposed mechanism. The presence and absence of the KIE under different reaction conditions provides an indirect observation of the two intermediates.

Table 2. KIE in the Presence of MVCs

Substrate	k_2k_3/k_{-2} (sec-1)	k_2 (sec-1)	K_1 (mM)	KIE
Indole + Na^{+}	308	-	0.20	2.2
3-^{2}H-Indole + Na^{+}	140	-	0.10	
Indole + K^{+}	388	-	0.25	2.0
3-^{2}H-Indole + K^{+}	193	-	0.21	
Indole + NH_4^{+}	556	-	0.67	2.1
3-^{2}H-Indole + NH_4^{+}	264	-	0.52	
Indole + Na^{+} + GP	-	7.2	-	none
3-^{2}H-Indole + Na^{+} + GP	-	7.0	-	

The apparent rate constant, k_2k_3/k_{-2}, varies depending on the MVC present, but the isotope effect remains constant. This could possibly suggest that the rate constant, k_3, is independent of the MVC present, but only k_2 and k_{-2} are effected. The rate K_2 therefore increases according to the MVC in the following order $NH_4^+ >> K^+ > Na^+$. The formation of $E(Q_3)$ is fastest in the presence of NH_4^+ with an almost two-fold increase in both the reaction with indole and 3-^{2}H-indole compared to the rate constant in the presence of Na^+. The rate constant of the reaction of indole and 3-^{2}H-indole in the presence of K^+ shows a slight increase in rate of formation $E(Q_3)$ compared to Na^+. The binding constant K_I of indole is also effected by the presence of different MVCs. While Na^+ and K^+ have nearly the same binding constant for indole, NH_4^+ has an almost three-fold increase in the binding constant.

The decay of $E(Q_3)$ is independent of the indole concentration. Table 3 shows the observed rates for the decay of $E(Q_3)$ in the presence of Na^+, K^+ and NH_4^+. There is a slight increase in the observed decay rate constant between Na^+ and K^+. In the presence of NH_4^+, the observed rate constant is more the steady state formation of the $E(Q_3)$. As shown in the time course of Figure 11, the decay is no longer observed at 476 nm. The amplitude of the decay of $E(Q_3)$ decreased in the following order: $Na^+ > K^+ > Cs^+ > NH_4^+$.

Table 3. Observed Rates for the Decay of E(Q3).

Substrates	Observed rate (sec^{-1})
Indole + Na^+	22
3-^{2}H-Indole + Na^+	20
Indole + K^+	36
3-^{2}H-Indole + K^+	33
Indole + $NH4^+$	9
3-^{2}H-Indole + $NH4^+$	9

Our results agree with the current theory about the conformations of tryptophan synthase in the presence of different MVCs. Previous studies suggested that the presence of Na^+ and K^+ stabilize the open (low activity) conformation, while Cs^+ and $NH4^+$ stabilize the closed (high activity) conformation of the beta subunit. The fact that the observed rate for the formation of $E(Q_3)$ was much higher in the presence of NH_4^+ than Na^+ or K^+ agrees with the closed conformation of NH_4^+. Also, the binding of indole was weaker in the presence of NH_4^+ versus Na^+ and K^+. The closed conformation in the presence of NH_4^+ may weaken the binding.

Conclusion.

In stage II of the β-subunit mechanism, a KIE ~ 2 is observed when indole condenses with the L-serine-enzyme α-aminoacrylate intermediate in the presence of Na^+, K^+, Cs^+ and NH_4^+. No KIE is observed in the presence of the α-subunit ligand α-glycerophosphate and Na^+. The allosteric ligands change the relative rates of the two steps of addition and tautomerization, allowing us to indirectly observe the presence of both $E(Q_2)$ and $E(Q_3)$. We conclude that this supports the previously proposed stepwise mechanism.

References:

(1) Dunn, M. F.; Aguilar, V.; Brzovic, P.; Drewe, W. F., Jr.; Houben, K. F.; Leja, C. A.; Roy, M. *Biochemistry* **1990**, *29*, 8598-8607.

(2) Pan, P.; Woehl, E.; Dunn, M. F. *Trends Biochem Sci* **1997**, *22*, 22-27.

(3) Lane, A. N.; Kirschner, K. *Eur J Biochem* **1983**, *129*, 675-684.

(4) Wilhelm, P.; Pilz, I.; Lane, A. N.; Kirschner, K. *Eur J Biochem* **1982**, *129*, 51-56.

(5) Hyde, C. C.; Ahmed, S. A.; Padlan, E. A.; Miles, E. W.; Davies, D. R. *J Biol Chem* **1988**, *263*, 17857-17871.

(6) Miles, E. W. *Adv Enzymol Relat Areas Mol Biol* **1979**, *49*, 127-186.

(7) Miles, E. W. *Adv Enzymol Relat Areas Mol Biol* **1991**, *64*, 93-172.

(8) Miles, E. W. *J Biol Chem* **1991**, *266*, 10715-10718.

(9) Lane, A. N.; Kirschner, K. *Biochemistry* **1991**, *30*, 479-484.

(10) Anderson, K. S.; Miles, E. W.; Johnson, K. A. *J Biol Chem* **1991**, *266*, 8020-8033.

(11) Lane, A. N.; Kirschner, K. *Eur J Biochem* **1983**, *129*, 571-582.

(12) Fan, Y. X.; McPhie, P.; Miles, E. W. *Biochemistry* **2000**, *39*, 4692-4703.

(13) Rhee, S.; Parris, K. D.; Hyde, C. C.; Ahmed, S. A.; Miles, E. W.; Davies, D. R. *Biochemistry* **1997**, *36*, 7664-7680.

(14) Rhee, S.; Parris, K. D.; Ahmed, S. A.; Miles, E. W.; Davies, D. R. *Biochemistry* **1996**, *35*, 4211-4221.

(15) Weber-Ban, E.; Hur, O.; Bagwell, C.; Banik, U.; Yang, L. H.; Miles, E. W.; Dunn, M. F. *Biochemistry* **2001**, *40*, 3497-3511.

(16) Woehl, E.; Dunn, M. F. *Biochemistry* **1999**, *38*, 7131-7141.

(17) Woehl, E.; Dunn, M. F. *Biochemistry* **1999**, *38*, 7118-7130.

(18) Miles, E. W. *Chem Rec* **2001**, *1*, 140-151.

(19) Houben, K. F.; Dunn, M. F. *Biochemistry* **1990**, *29*, 2421-2429.

(20) Milne, J. J.; Malthouse, J. P. *Biochem J* **1995**, *311*, 1015-1019.

(21) Kiick, D. M.; Phillips, R. S. *Biochemistry* **1988**, *27*, 7339-7344.

Figure 1. Zero Time Spectra in the Presence of Different MVCs.

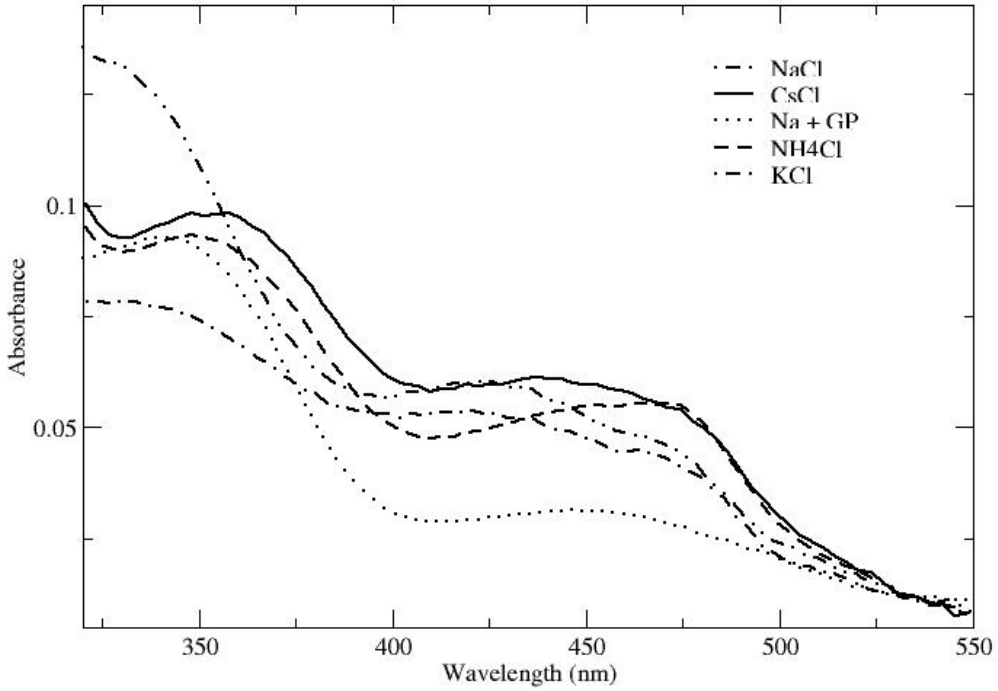


Figure 2. Concentration Dependence in the Presence of Na^{+}.

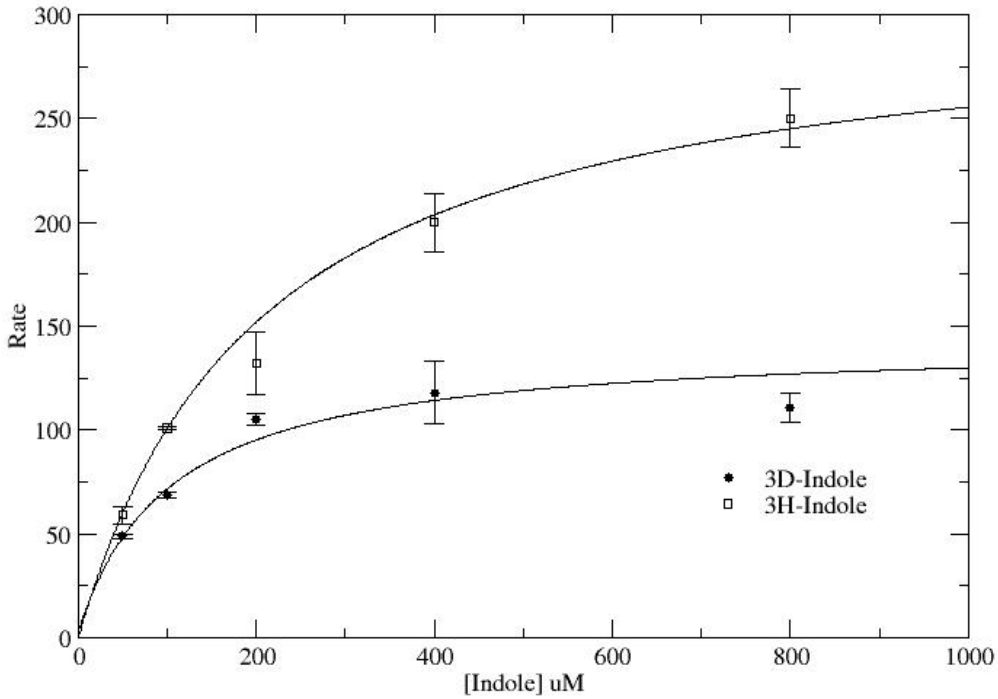


Figure 3. Spectra in the Presence of Na^+. Scans shown are from the first 100 ms in 10 ms intervals.

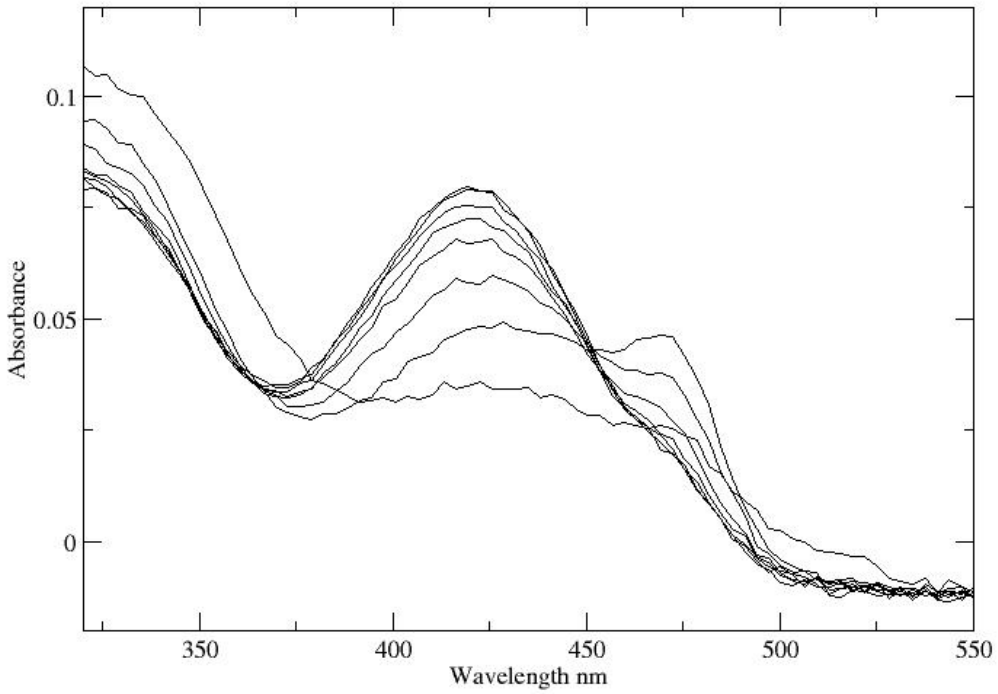


Figure 4. Time Course in the Presence of Na^+. Monitored at $\lambda = 476$ nm.

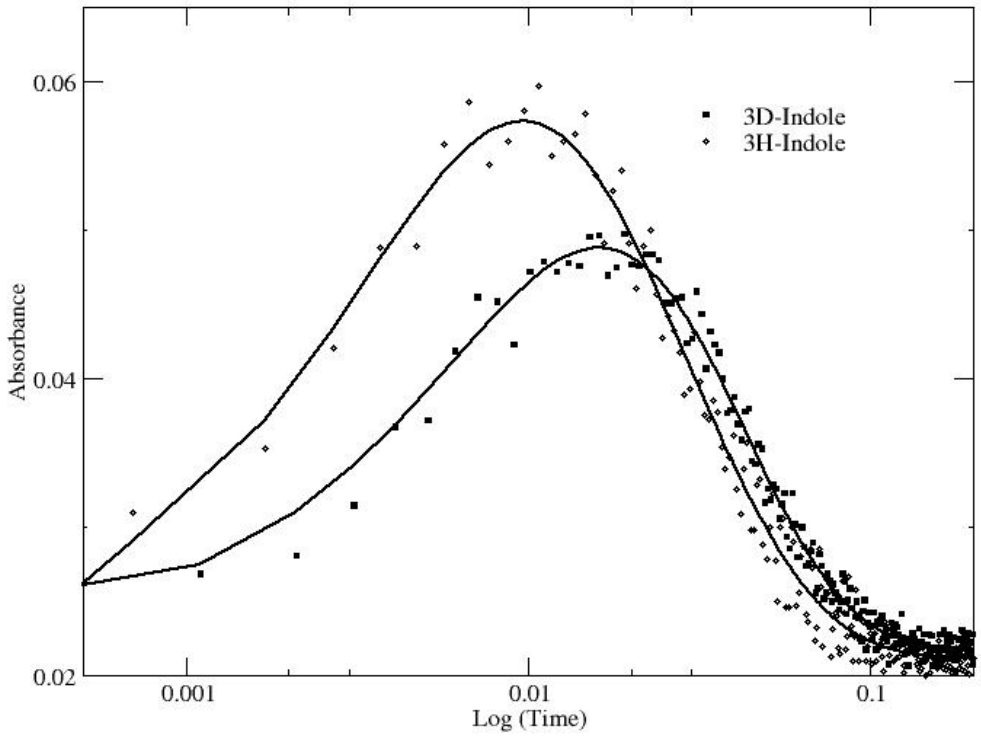


Figure 5. Concentration Dependence in the Presence of K^+

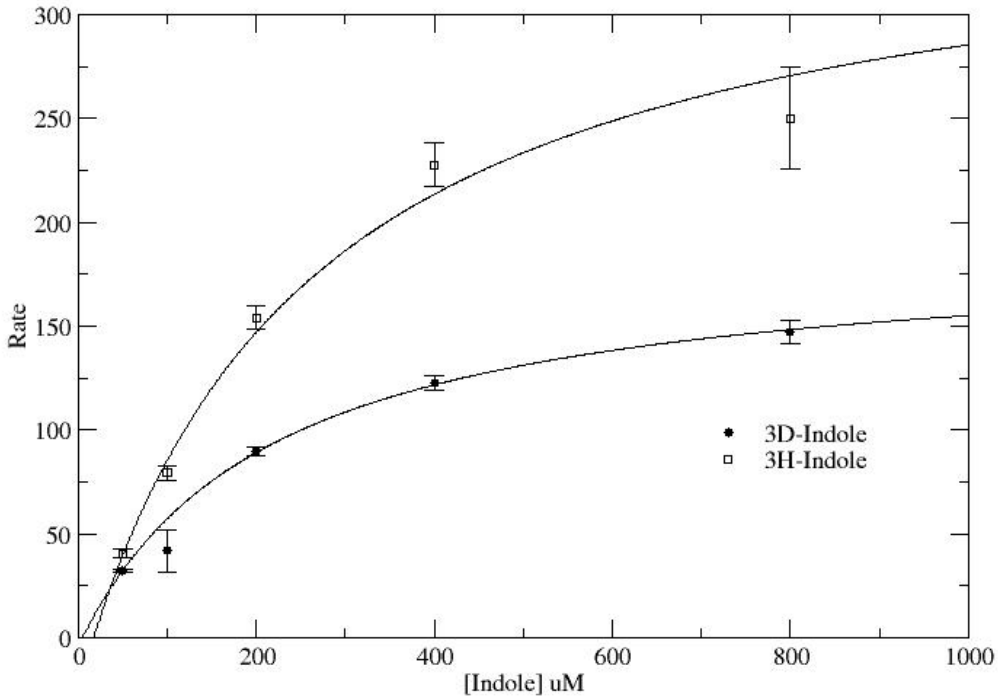


Figure 6. Spectra in the Presence of K^+. Scans shown are from the first 100 ms in 10 ms intervals.

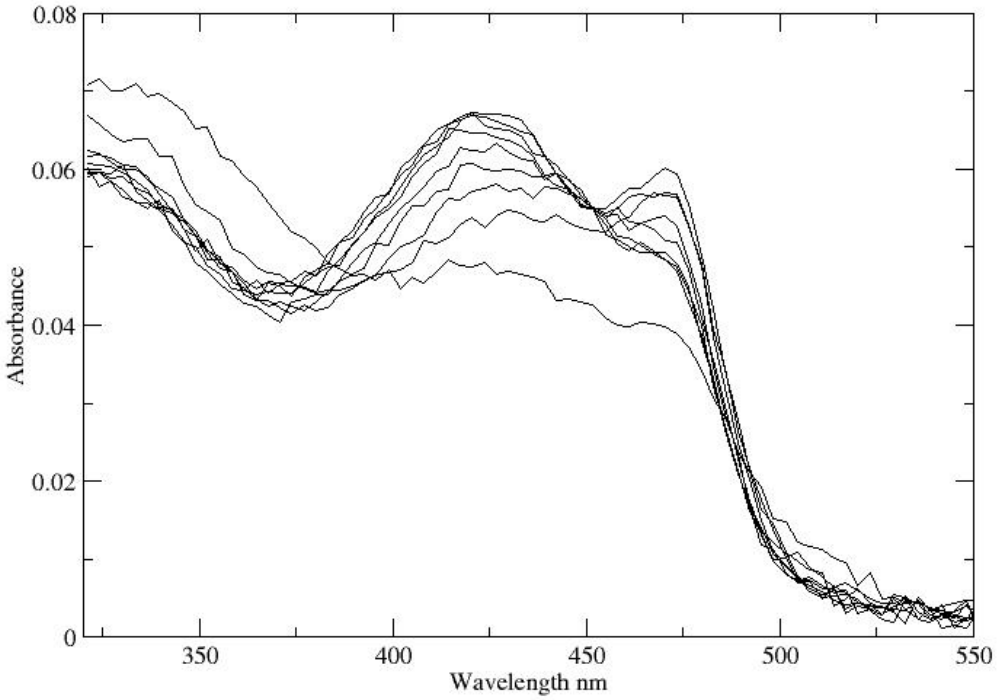


Figure 7. Time Course in the Presence of K^+. Monitored at $\lambda = 476$ nm.

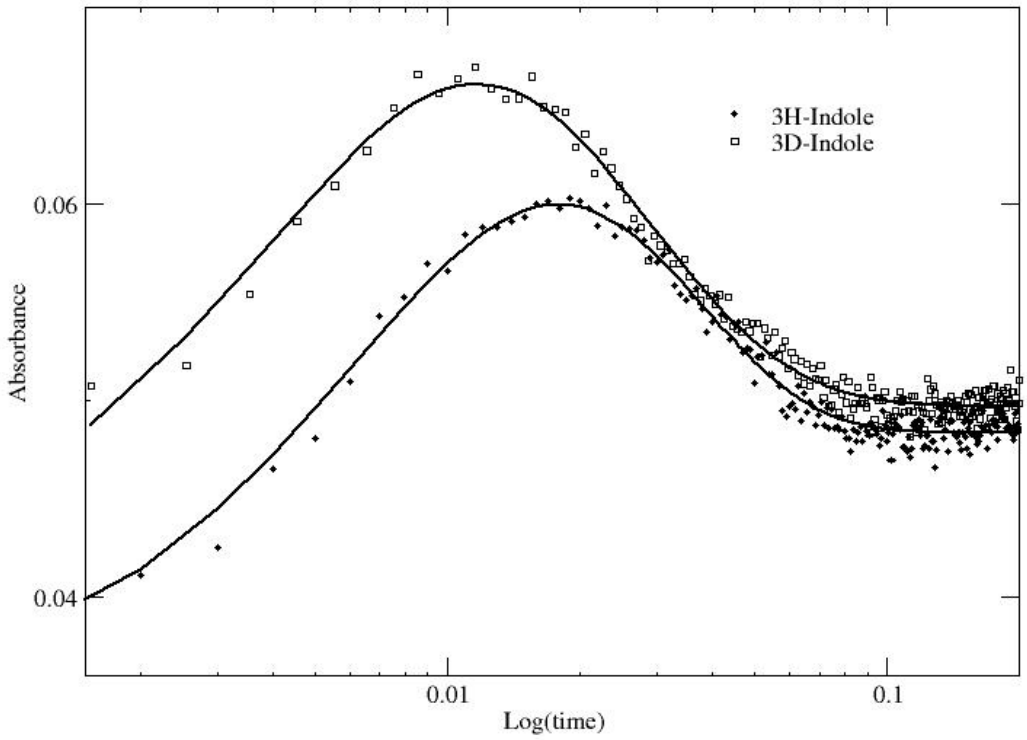


Figure 8. Concentration Dependence in the Presence of NH_4^+.

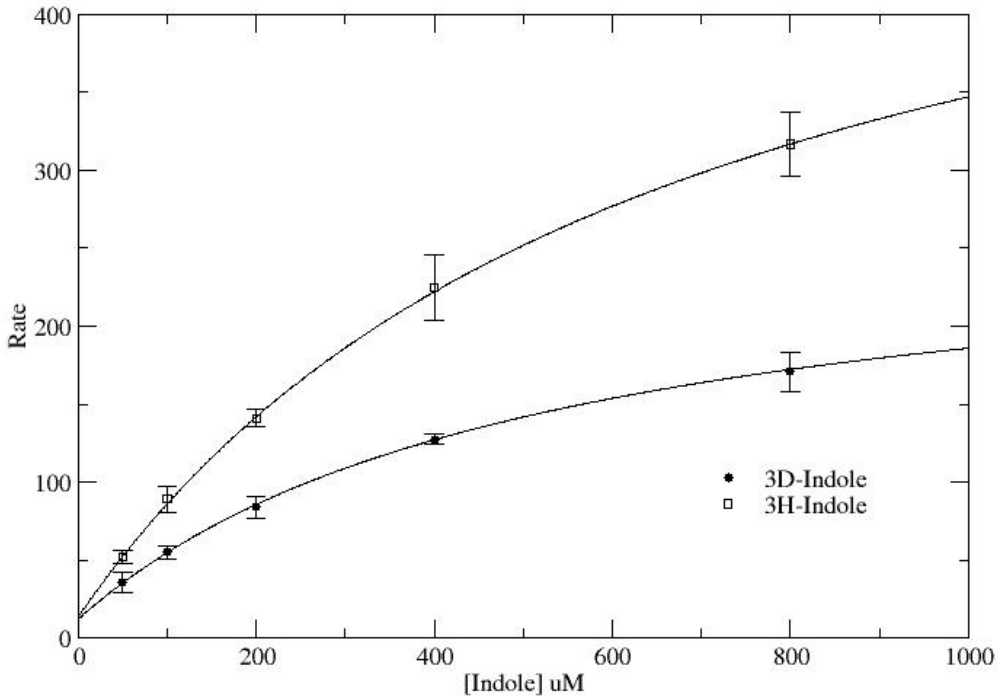


Figure 10. Spectra in the Presence of NH_4^+. Scans shown are from the first 100 ms in 10 ms intervals.

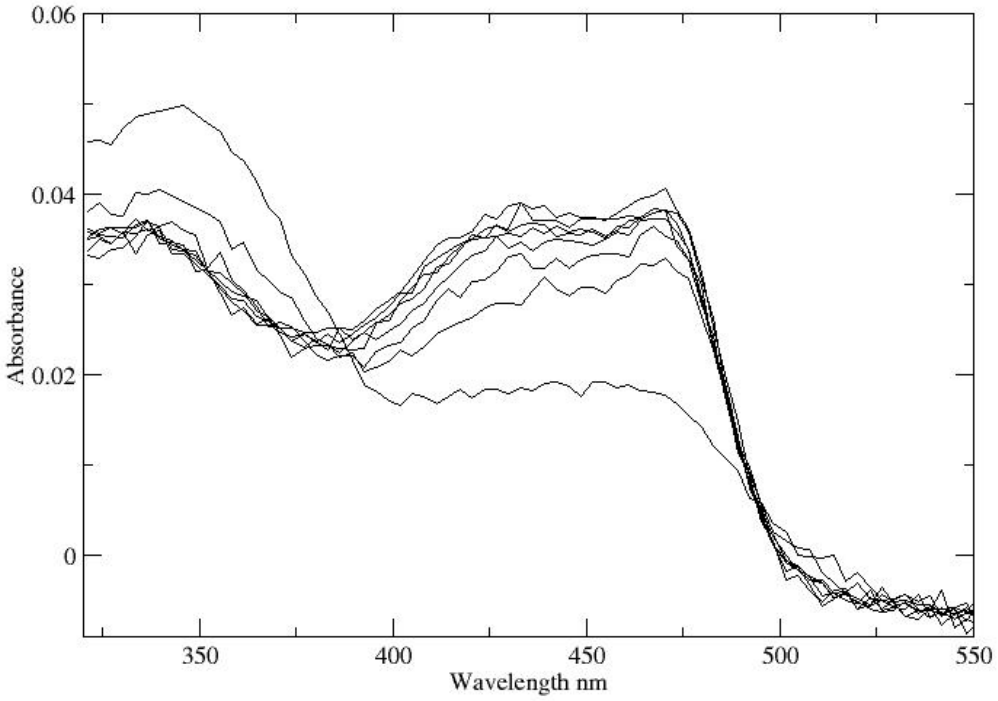


Figure 11. Time Course in the Presence of NH_4^+. Monitored at $\lambda = 476$ nm.

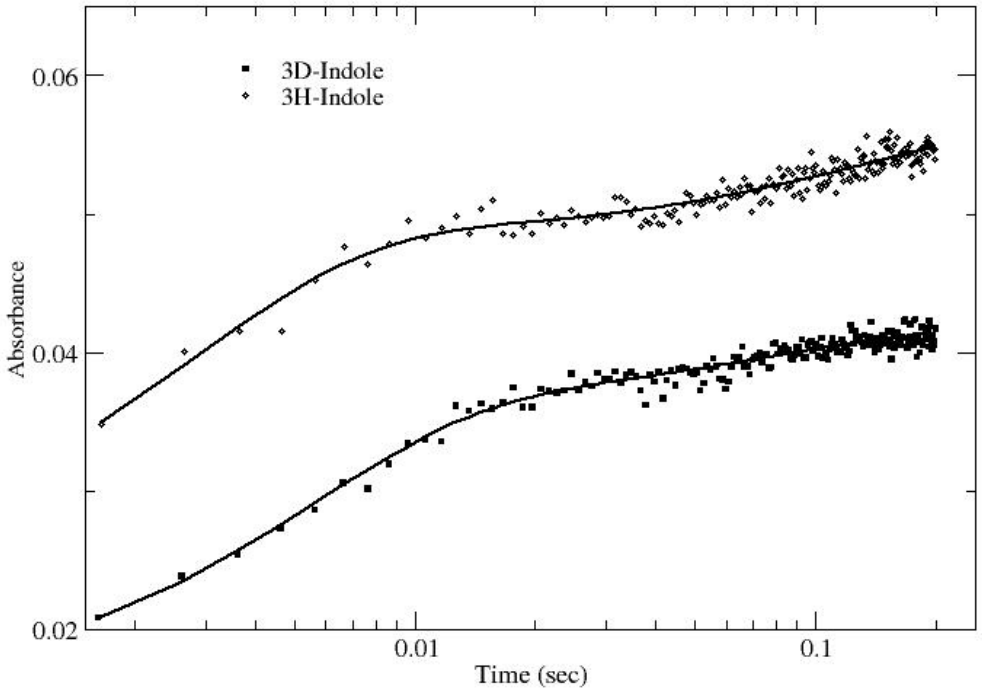


Figure 12. Time Course in the Presence of Cs^+. Monitored at $\lambda = 476$ nm.

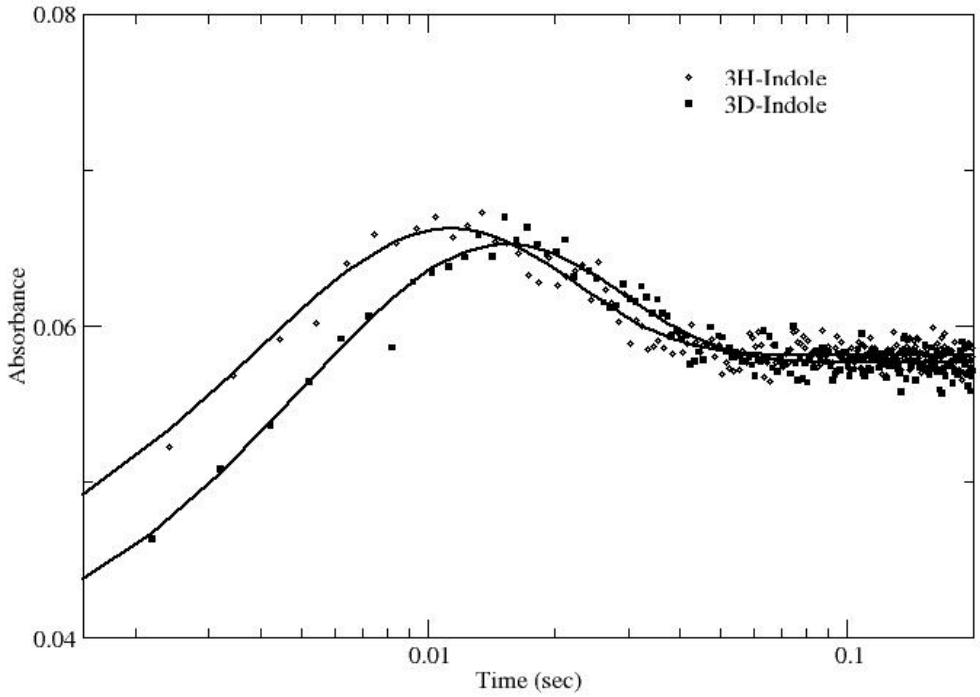


Figure 13. Spectra in the Presence of Cs^+. Scans shown are from the first 50 ms in 6 ms intervals.

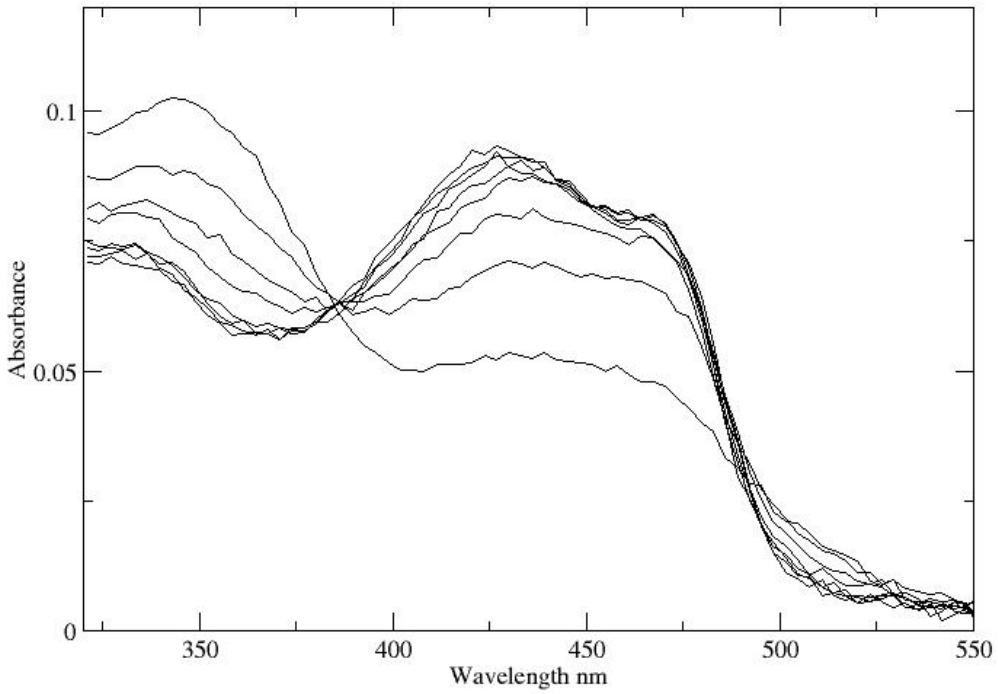


Figure 14. Concentration Dependence in the Presence of Na_2GP.

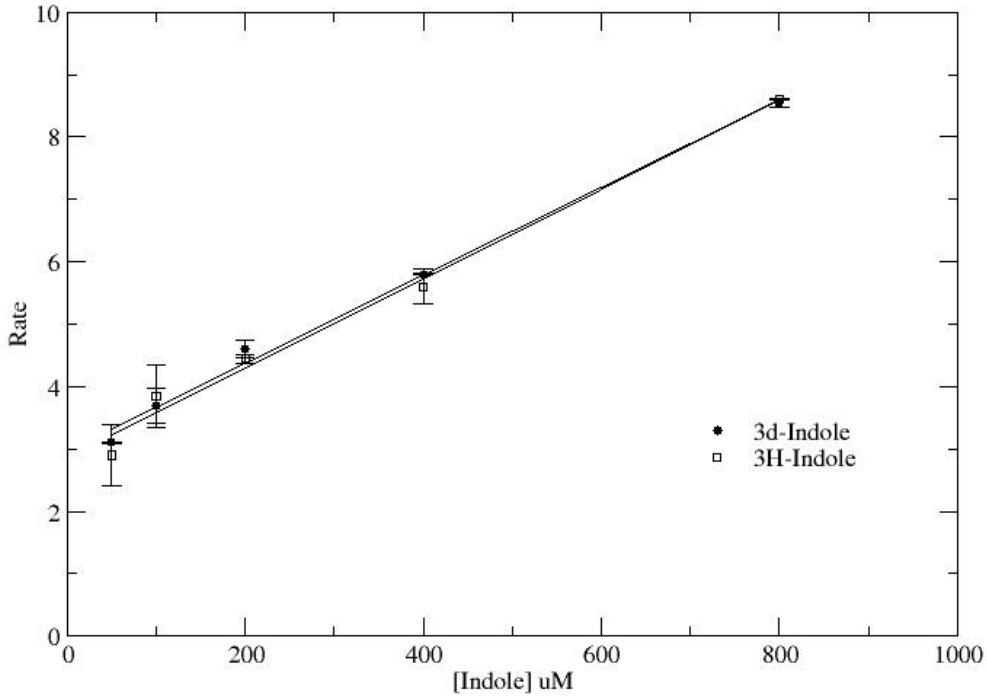


Figure 15. Spectra in the Presence of Na_2GP. Scans shown are from the first 200 ms in 25 ms intervals..

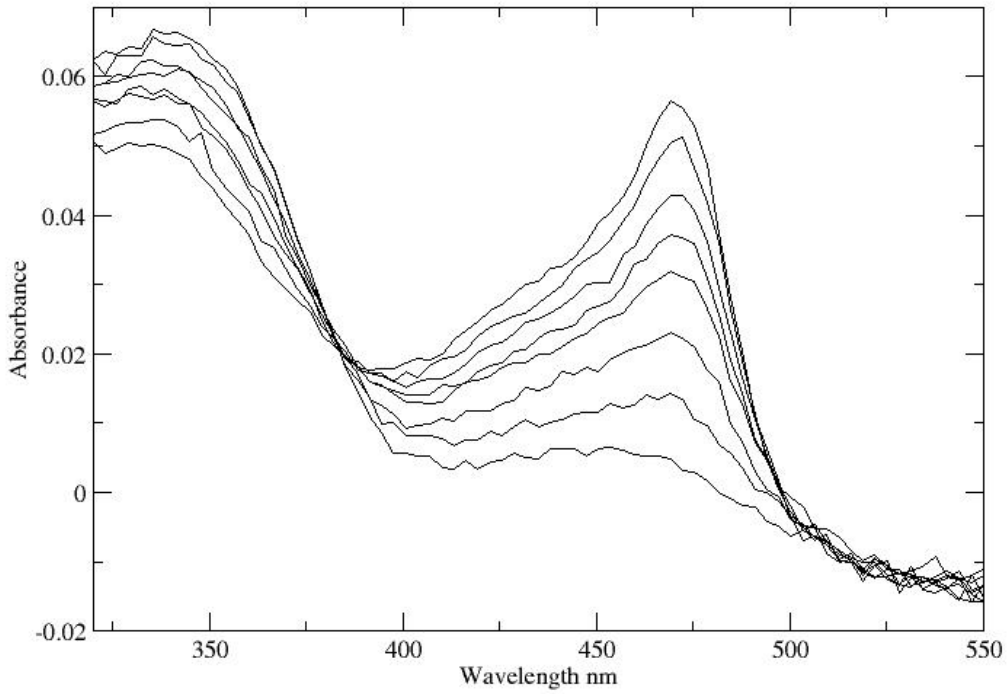
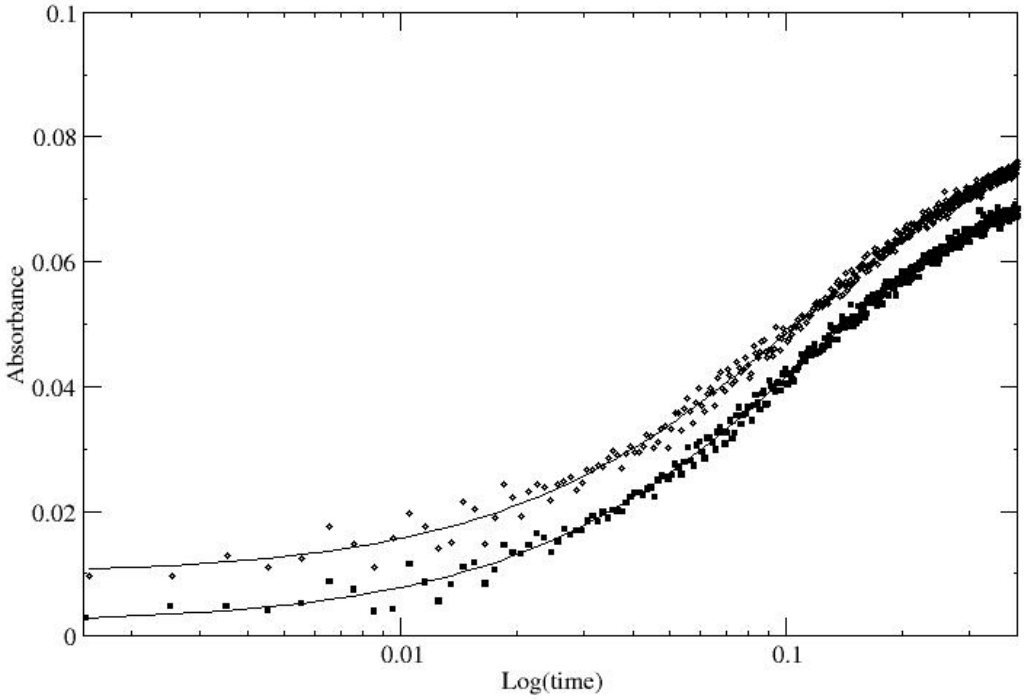


Figure 16. Time Course in the Presence of Na_2GP. Monitored at $\lambda = 476$ nm.



CHAPTER II

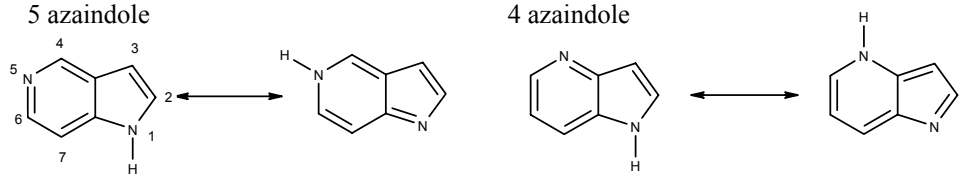
EXCITED STATE TAUTOMERIZATION OF AZAINDOLE

Introduction

Tryptophan has been used for many years as an intrinsic fluorescent probe to investigate protein dynamics and structutre. Tryptophan fluorescence is sensitive to the electric field of its local environment. A red shift in the fluorescence emission wavelength is observed when tryptophan is on the inside of a protein in a hydrophobic pocket relative to a tryptophan residue exposed to solvent. Charged amino acid side chains also quench the fluorescence decay of tryptophan, providing additional information about the local environment of the tryptophan residue.[1-5] These characteristics make tryptophan a useful probe for understanding protein dynamics and protein-protein interactions. Recently, several tryptophan analogs have been investigated for incorporation into proteins. These include, 5-hydroxy-tryptophan, 7-azatryptophan, and 6-azatryptophan.[6-8] Each of these analogs offer advantages over natural tryptophan for exploring protein dynamics. The advantages include greater fluorescence intensity, red-shifted fluorescence, and greater sensitivity to the electric field of the local environment.[7]

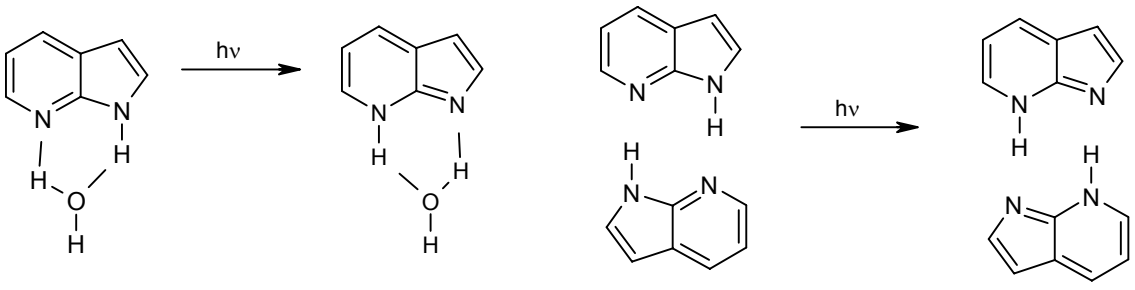
The azatryptophan isomers are of great interest because of their close structural similiarity to natural tryptophan and their ability to undergo excited state tautomerization. The incorporation of 7-azatryptophan and 6-azatryptophan into proteins has proven to be relatively simple compared to other synthetic amino acid probes.[6,7] The azatryptophan analogs can be incorporated by expressing the protein in a tryptophan auxotroph host while supplying the azatryptophan in the media. Both 7-azatryptophan and 6-azatryptophan have been successfully incorporated into proteins and peptides with promising results. Two other azatryptophan analogs, 4-azatryptophan and 5-azatryptophan, may also be useful as probes for protein dynamics and structure.

Scheme 1



Indole is the chromophoric moiety of tryptophan and the photophysics of indole are explored to avoid complications of the zwitterionic form of tryptophan in experimental studies.[3] In this work, we expand the toolbox of fluorescent probes by investigating the tautomers of two azaindole isomers (Scheme 1): 4-azaindole (4AI) and 5-azaindole (5AI). Through comparison of 4AI and 5AI with the previously studied 6-azaindole (6AI) and 7azaindole (7AI), we can gain further insight into the photophysics and tautomerism of all the azaindole isomers. 7AI has been thoroughly studied regarding its ability to undergo excited state tautomerization. 7AI has been shown to tautomerize by both a dimer excited state double proton exchange and also a cyclic intermediate involving one water molecule (Scheme 2).[9,10] Recently, it was shown that 6AI also undergoes excited state tautomerization evident by the presence of fluorescent band at 440 nm in certain solvent conditions, including strong base and methanol-water mixtures. The tautomerism of 6AI occurs by protonation at N6 in the excited state, followed by deprotonation at N1 also in the excited state.[7]

Scheme 2



Indole is the base structure of a number of very important biological molecules. The azaindole isomers could also be useful for other applications besides peptide and protein dynamics. 7AI is currently being explored as a molecular probe for use in auxin physiology. In this case, 7AI is the chromophoric moeity of 7-azaindole-3-acetic acid. Considering the great potential for use as a fluorescent probe, we are interested in characterizing the fluorescence properties of 4AI and 5AI. The ability of the azaindole

isomers to tautomerize in solution is an important aspect regarding the fluorescence, because each tautomer has a unique fluorescent spectra. We calculated and compared each isomer and its tautomer to determine which isomer and tautomer were the lowest in energy in both ground state and excited state. We also looked at the effects of solvation to see what role the solvent has in preference for a given tautomer. The three main questions we look to answer are: 1. Do 4AI and 5AI also undergo excited state tautomerization? 2. Do 4AI and 5AI also look promising as biological probes? 3. Why is one tautomer more stable than the other in the ground state?

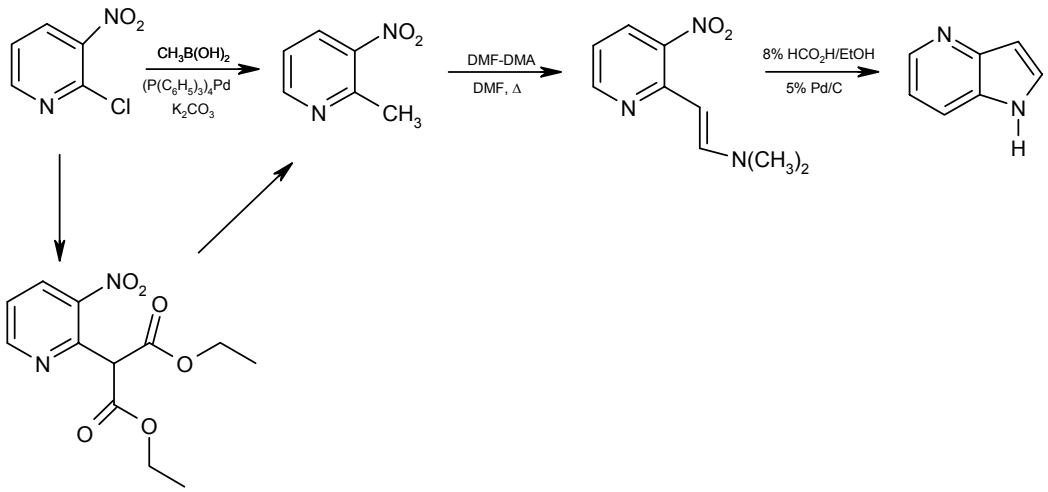
Theoretical and Experimental:

A. Computations. The ground state studies of the azaindole isomers in both gas phase and solvent were done using density functional theory. Nucleus Independent Chemical Shifts (NICS) were also done using density functional theory. The level of the computations used is B3LYP/6-31+G**, which utilizes Becke's three parameter exchange functional[11], and the Lee-Yang-Parr nonlocal correlation functional.[12]

Electronic solvent effects were performed with the Self Consistent Reaction Field (SCRF) calculations using the Isodensity Polarized Continuum Model (IPCM).[13] The IPCM model was chosen because the effects of solvation are incorporated into the iterative SCF calculation, which results in a proper variational condition. The dielectrics of the three solvents mimic chloroform (4.8), methanol (32.7) and water (78.4).

Both time dependent DFT and ab intio were used for the excited state transition energies. The ab initio method chosen is Configuration Interaction with Single excitations (CIS) procedure.[14] CIS was used to obtain the optimized excited state energies and geometries. The same basis set was used (CIS/6-31+G**) as for the DFT calculations. All calculations were done using GAUSSIAN 94.[15]

B. Synthesis of 4-azaindole. Two routes were explored for obtaining 2-methyl-3-nitro-pyridine from 2-chloro-3-nitropyridine.



Method 1.[16]

Preparation of diethyl-2-(3-nitropyridyl)malonate.

2-chloro-3-nitropyridine (2.0 g) was added to 0.3 g NaH in 10 mL of dry DMF with stirring under nitrogen. Diethyl malonate (2.0 g) was added dropwise, resulting in a brownish-red mixture. The reaction proceeded for 1 hr, and the solvent was then evaporated giving a brown oil. The resulting oil was diluted with 50 mL of water, followed by neutralization with a few drops of acetic acid. The precipitate was extracted with dichloromethane three times, and the combined organic layers were dried with $MgSO_4$. The solvent was removed and the remaining oil was then passed over a silica column using ethyl acetate/hexane mixture. The product diethyl-2-(3-nitropyridyl)malonate (1.8g) was isolated as yellow crystals (50% yield).

Preparation of 2-methyl-3-nitropyridine.

Diethyl-2-(3-nitropyridyl)malonate (1.0 g) was dissolved in 50 mL of 6N HCl and refluxed for 8 hrs. The solvent was evaporated leaving a brownish oil that solidified

upon cooling. 50 mL of saturated sodium carbonate was added, and the solution was extracted three times with 25 mL of dichloromethane. The organic layers were combined and dried with $MgSO_4$ followed by removal of solvent yielding 0.42 g of 2-methyl-3-nitropyridine as a brown solid (86% yield).

Method 2.[17]

Preparation of 2-methyl-3-nitropyridine.

A mixture of (0.16 g) 2-chloro-3-nitropyridine, tetrakis(triphenylphosphine)Pd (0.12 g), methylboronic acid (0.07 g) and potassium carbonate (0.4 g) was added to dioxane and refluxed for two days. The solvent was evaporated and the mixture was separated on a silica column using an ethyl acetate/hexanes solvent. 2-methyl-3-nitropyridine (0.06 g) was obtained as a light brown solid (43% yield).

Preparation of the enamine.[16]

2-Methyl-3-nitropyridine (0.5 g) was dissolved in 10 mL of dry DMF and stirred under nitrogen. DMF-DMA (1 g) was added dropwise. The reaction was heated to 90 °C. After 15 minutes of heating, a deep reddish color began to appear, and it was allowed to react for 4 hrs. After 4 hrs, the solvent was removed by evaporation. The red oil obtained (0.65 g) was used without further purification .

Preparation of 4-azaindole.

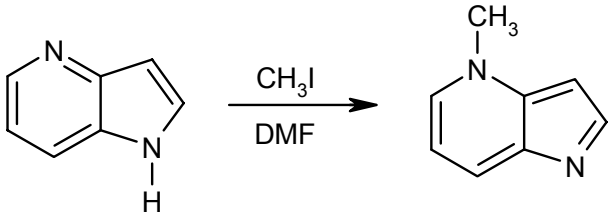
The crude enamine (0.65) was dissolved in 10 mL of 8.8% formic acid in methanol. The mixture was added dropwise to a flask containing 10 mL of 8.8% Formic acid in methanol with 0.2 g of 5% Pd/C, stirred under nitrogen. The reaction proceeded for 4 hrs, until the red color had completely disappeared. The Pd catalyst was removed by filtration through celite, and the filtrate was then concentrated. After sitting overnight, the product, 4-azaindole, crystallized (0.12 g). The crystallized product was filtered and

washed with hexanes. The filtrate was then evaporated to dryness giving 0.26 g of 4-azaindole for an overall yield of 0.31 g (78%).

C. Synthesis of 5-azaindole.

5-azaindole was synthesized as previously described.[16]

D. Synthesis of the N-Methyl-Azaindole isomers.[18]



The azaindole isomer (100 mg) was dissolved in 2 mL dry DMF. Iodomethane (120 mg) was added to the mixture, and the reaction was allowed to stir for 3 hrs. The solvent was evaporated, leaving a brown oil. The remaining brown oil was dissolved in 25 mL dichloromethane and extracted three times with 15 mL of saturated sodium carbonate. The organic layer was dried with $MgSO_4$ follwed by evaporation and the mixture was then loaded onto an alumina column. The parent azaindole was eluted first with chloroform. The N-methyl azaindole was eluted with 20% ethanol/chloroform.

Absorbance and Fluorescence Spectroscopy.

The absorbance spectra were obtained on a Cary 1E UV/Vis spectrophotometer. The path length was 1 cm and the temperature was 20 °C. The fluorescence experiments were obtained on a Fluorolog Spectrofluorimeter by Instruments S.A., Inc. The path length was 1 cm and all experiments were performed at 20 °C. Each of the azaindole isomers was analyzed in water and 1 M NaOH solution. The excitation wavelength for 4AI was 293 nm. The excitation wavelength for 5AI was 270 nm.

Results and Discussion:

Many theoretical and experimental studies have been reported on 7-AI. It has been well established that the N1-H tautomer of both 7-AI and 6-AI is the most stable tautomer in the ground state.[6,7,10,19,20] Rasmussen and Mahadevan reported that all four azaindole isomers are more stable in the N1-H tautomer form based on ab-initio calculations using the basis set STO-3G.[18] In Table 1, the optimized ground state energy differences of the tautomers of the azaindole isomers were calculated using density functional theory with the 6-31+G** basis set. The N1-H tautomer is more stable by more than 10 kcal/mol for each of the azaindole isomers as seen in Table 1. The N1-H tautomer has often been referred to as the normal tautomer for this very reason.[6] Among the normal tautomers, 6-azaindole is the lowest in energy while 7-azaindole is the highest, with 2.4 kcal/mol separating the two. On the other hand, in the excited state the other tautomer (referred to as Nn-H, where n=4,5,6, or 7 depending on the azaindole isomer) is more stable, by as least 15 kcal/mol. Previous experimental work agrees with the calculations, as both 7-azaindole and 6-azaindole do tautomerize in the excited state, provided that the solvent conditions are suitable (i.e., alkaline solution or combination of water with a less polar co-solvent).[6,7]

Table 1. Energy Differences Between Tautomers. (Energy of Nn-H - Energy of N1-H).

Isomer	Ground State (kcal/mol)	Excited State (kcal/mol)
4-azaindole	11.6	-15.5
5-azaindole	11.4	-17.6
6-azaindole	10.1	-14.0
7-azaindole	12.5	-18.1

The following investigation was to see if 5-azaindole and 4-azaindole also tautomerize in the excited state. Calculations indicate that the Nn-H tautomers are more stable by more than 15 kcal/mol in the excited state and are therefore expected to

tautomerize. However, there is considerable distance between the pyrrole nitrogen and pyridine nitrogen in the case of 4AI and 5AI. Therefore, the possibility of any proton transfer as thoroughly explored with 7-azaindole and seen in Scheme 2, is absent here. In the case of 6-azaindole, a proposed mechanism of the normal tautomer being excited, following by solvent proton extraction at N6 in the excited state, and the deprotonation of N1 also in the excited state.[7]

The relative energies of the N1-H azaindole isomers can be seen Table 2. 7AI is lower in energy by about 4 kcal/mol in the ground state compared to other N1-H azaindole isomers. 4AI, 5AI and 6AI all have nearly the same ground state energy as the N1-H tautomer. Likewise, in the excited state, 7AI is again the lower in energy by as much as 9.4 kcal/mol in the N1-H tautomer.

Table 2. Relative Energies of N1-H Azaindole Isomers

Isomer	Ground State (kcal/mol)	Excited State (kcal/mol)
4-azaindole	-0.4	-2.8
5-azaindole	-0.7	-1.8
6-azaindole	0.0	0.0
7-azaindole	-4.8	-9.4

The relative energies of the Nn-H tautomers can be seen in Table 3. N5-H 5AI is more stable by about 5 kcal/mol in the ground state compared to other Nn-H azaindole isomers. N4-H 4AI and N6-H 6AI have nearly the same ground state energy as N5-H 5AI, while N7-H 7AI is highest in energy However, in the excited state, N7-H 7AI is the lowest in energy by as much as 13.6 kcal/mol.

Table 3. Relative Energies of Nn-H Azaindole Isomers

Isomer	Ground State (kcal/mol)	Excited State (kcal/mol)
N4-H 4AI	-3.6	9.2
N5-H 5AI	-5.0	8.1
N6-H 6AI	-4.6	13.6
N7-H 7AI	0.0	0.0

Spectroscopic Studies

The fluorescence of 5-azaindole and 4-azaindole were measured to determine if these isomers can tautomerize in the excited state and to characterize the fluorescence wavelength of the tautomer and the protonated species. The N4-methyl and N5-methyl derivatives were synthesized to help characterize the fluorescence spectra, since the presence of a methyl group locks the azaindole into the Nn-H tautomeric form. As seen from the spectra in Figures 1 and 2, at high pH (i.e., pH=14), 4AI and 5AI both tautomerize. In the case of 4AI, we see the presence of two bands, one at 390 nm and another at 480 nm. The N-4-methyl-4AI fluorescence spectrum shows a single band at 480 nm in the presence of 1 M NaOH. The band at 480 nm is assigned to the N4-H 4AI tautomer based on the fluorescence spectrum of N-4-methyl-4AI. The fluorescence spectrum of 5AI in 1 M NaOH shows only one broad band with a maximum at 455 nm. The fluorescence spectrum of 5-methyl-5AI, has a virtually identical spectra with a maximum at 460 nm. The fluorescent band at 455 nm is assigned to the N5-H 5AI tautomer. For comparison, the spectra of 5AI and 4AI in water are shown. These fluorescent bands are most likely the protonated form of the azaindole isomer. 4AI has a maximum at 415 nm, and 5AI has a maximum at 410nm. Table 4 summarizes the fluorescence data for each of the azaindole isomers.

Table 4. Fluorescence Maxima. Fluorescence maxima of the tautomers and protonated azaindole isomers. 6AI and 7AI fluorescenece maxima taken from the specified reference.

Isomer	Wavelength nm
4AI	na
protonated 4AI	415
N4-H 4AI	480
5AI	na
protonated 5AI	410
N5-H 5AI	450
6AI	350[7]
protonated 6AI	380[7]
N6-H 6AI	440[7]
7AI	385[6]
protonated 7AI	440[6]
N7-H 7AI	500[6]

The excitation energies can be qualitatively calculated. We explored the excitation energies using two different methods. First, we used time-dependent DFT, which calculates a vertical excitation from the optimized ground state.[21] The optimized excited state geometries and energies were calculated using the CI Singles method, which also calculates the vertical excitation energies.[14] Both TDDFT and CIS calculations were done using the 6-31+G** basis set. The calculated and experimental excitation wavelength and excited state energies can be seen in Table 5. The experimental wavelength were taken from reference [18]. There is a good qualitative agreement between the calculated and experimental. The main note seen here is the other tautomer requires less energy to be excited than the normal tautomer, suggesting a smaller HOMO-LUMO gap.

Table 5. Calculated and Experimental Absorbance Wavelengths.

Isomer	Experimental (nm)[18]	CIS/6-31+G** (nm)	TDDFT/6-31+G** (nm)
4AI	293	242	270
N4-H 4AI	370	336	383
5AI	270	248	260
N5-H 5AI	310	358	355
6AI	300	240	264
N6-H 6AI	355	318	345
7AI	296	251	271
N7-H 7AI	350	373	387

Solvent Effects.

The effects of solvent on the tautomer energy difference were explored to see if the dielectric alone from the solvent would stabilize the energy difference allowing for the other tautomer to be more stable. All the azaindole isomers were stabilized by the dielectric constant of the solvent, but the most stabilization occurred for 5AI and 6AI as seen in Figure 3 and Table 6. The energy difference of both isomers were stabilized 5.8 kcal/mol and 4.8 kcal/mol respectively. 4AI and 7AI were stabilized by 3.4 kcal/mol and 3.5 kcal/mol respectively. Table 6 also shows the effect on the dipole moment associated with each tautomer. Each isomer is stabilized by about 1 kcal/mol between the dielectric of chloroform and that of water. Though the stabilization is minimal, the increased dielectric does provide stabilization between the tautomers.

Table 6. Solvent Effects. Energy differences between tautomers in the presence of different dielectrics. Dipole differences in parenthesis.

Isomer	Gas	Chloroform	Methanol	Water
4AI	11.6 (1.4)	9.1 (2.0)	8.3 (2.3)	8.2 (2.3)
5AI	11.4 (1.8)	7.4 (2.7)	5.8 (3.1)	5.6 (3.2)
6AI	10.1 (1.5)	6.7 (2.3)	5.5 (2.7)	5.3 (2.7)
7AI	12.5 (2.2)	10.0 (3.2)	9.1 (3.7)	9.0 (3.7)

The effect of solvation was greater on the Nn-H tautomer than on the N1-H tautomer. The presence of a dielectric constant on the azaindole isomer causes an increase in the dipole of the molecule. As the solvent dielectric increases, the dipole difference between the tautomers also increases. The increasing dipole differences between the tautomers are linear in relation to the decreasing energy differences between the tautomers. Each of the Nn-H isomers have a higher dipole moment than the N1-H tautomer. The stabilization of that tautomer is greater, which lowers the energy difference between each tautomer. Table 7 shows the solvation stabilization energies of each isomer and their tautomer by the water dielectric. Each of the N1-H tautomers are stabilized by about 8 kcal/mol, with the exception of 7AI. The solvation energy of the Nn-H tautomer varies. The N5-H 5AI tautomer solvation energy is the greatest at 14.2 kcal/mol, while N7-H-7AI is the lowest with a solvation energy of 9.9 kcal/mol.

Table 7. Solvation Energies. Energy differences from gas phase to solvent dielectric equal to water (78.4).

Isomer	Solvation energy (kcal/mol)
4AI	8.1
N4-H 4AI	11.5
5AI	8.3
N5-H 5AI	14.2
6AI	8.0
N6-H 6AI	12.8
7AI	6.5
N7-H 7AI	9.9

NMR Spectroscopy

Other calculations include comparison of experimental and calculated NMR chemical shifts. Table 8 shows the comparison in experimental vs. calculated NMR. The

experimental NMRs were performed in $CDCl_3$ and taken from reference [18]. Evaluation of the pyrrole protons (C2-H and C3-H) are shown in the table, because they are most independent of the symmetry in the pyridine ring. The NMR calculations are in good agreement with the experimental chemical shifts regarding trends. The C-2 proton shows a change in chemical shift of approximately 1 ppm between the normal tautomers and the other tautomer. The Nn-H tautomer is developing more pi bond character between N1 and C2, which sheilds the C2 proton more causing the downfield shift. The C3 proton chemical shifts are moved downfield by about 0.2 to 0.3 ppm for 5AI, 6AI and 7AI between the N1-H and Nn-H tautomer. The C3 proton on 4AI is shifted upfield going from the normal to other tautomer. The pyridine nitrogen in the N4-H tautomer of 4AI is positioned to donate electrons to C3 which deshields the C3 proton causing the upfield shift.

Table 8. **Azaindole Chemical Shifts.** Chemical shifts of the C2 and C3 protons of the azaindole isomers.

Proton	4-AI	N4H-4AI	5-AI	N5H-5AI	6AI	N6H-6AI	7AI	N7H-7AI
C2(Exp)	7.37	8.29	7.34	8.05	7.45	8.37	7.21	8.2
C2(Calc)	7.43	8.44	6.98	8.27	7.07	8.71	7.12	8.31
C3(Exp)	6.76	6.48	6.5	6.8	6.56	6.74	6.57	6.75
C3(Calc)	7.09	6.35	6.63	6.79	6.56	6.99	6.48	6.88

Chemical shifts can also be calculated at any position around a molecule, not only at a nucleus. These type of calculations have been given the term NICS, nucleus independent chemical shifts. Schleyer and co-workers have shown that using NICS to

calculate the isotropy at the center of an aromatic ring can provide a quantitative measurement of the aromaticity of the ring.[22] Gordon and co-workers suggested that part of the reason for the greater stability of the N1-H tautomer over the N7-H tautomer of 7-Azaindole was due to aromatic stabilization of the N1-H tautomer.[20] In the N7-H tautomer, the pyrrole portion has only 4 pi electrons which suggests it would be anti-aromatic. However, NICS calculations taken from 1 Å above the center of each ring shows that both N1-H tautomer and the Nn-H tautomer are almost equally aromatic. The isotropy values are shown in Table 9. The isotropy value of 10.2 for benzene was included as a reference point.

Table 9. NICS. Isotropy from 1 Å above the center of the ring.

Molecule	Isotropy (Pyridine Ring)	Isotropy (Pyrrole Ring)
4AI	10.64	10.19
N4-H 4AI	9.11	12.92
5AI	10.81	10.31
N5-H 5AI	9.16	13.29
6AI	10.91	10.50
N6-H 6AI	8.69	13.44
7AI	10.68	10.25
N7-H 7AI	9.44	12.55
Benzene	10.20	

Conclusion

4AI and 5AI both experience excited state tautomerization in the presence of strong base. The protonated form of 4AI has an emission maximum at 415 nm and the N4-H tautomer has a fluorescence maximum at 480 nm. The protonated form of 5AI has an emission maximum at 410 nm and the N4-H tautomer has a fluorescence maximum at 450 nm. All the azaindole isomers are more stable in the N1-H tautomer by at least 10 kcal/mol in the ground state. In the excited state, all the Nn-H tautomers are lower in energy. The dielectric force of solvent helps to reduce the energy difference between the tautomer in

the ground state by as much as 5.8 kcal/mol, but not enough to reverse the ground state tautomer preference. The preference for the N1-H tautomer appears to not be due to aromatic stabilization considering both tautomers are equally aromatic.

References:

(1) Clayton, A. H. A.; Sawyer, W. H. *European Biophysics Journal* **2002**, *31*, 9-13.

(2) Royer, C. A. *Biophysical Journal* **1993**, *65*, 9-10.

(3) Lakowicz, J. R. *Photochemistry and Photobiology* **2000**, *72*, 421-437.

(4) Engelborghs, Y. *Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy* **2001**, *57A*, 2255-2270.

(5) Creed, D. *Photochemistry and Photobiology* **1984**, *39*, 537-562.

(6) Smirnov, A. V.; English, D. S.; Rich, R. L.; Lane, J.; Teyton, L.; Schwabacher, A. W.; Luo, S.; Thornburg, R. W.; Petrich, J. W. *Journal of Physical Chemistry B* **1997**, *101*, 2758-2769.

(7) Twine, S. M.; Murphy, L.; Phillips, R. S.; Callis, P.; Cash, M. T.; Szabo, A. G. *Journal of Physical Chemistry B* **2003**, *107*, 637-645.

(8) Li, Q.; Du, H. N.; Hu, H. Y. *Biopolymers* **2003**, *72*, 116-122.

(9) Chang, C.-P.; Lin, T.-C. *Huaxue* **1997**, *55*, 83-94.

(10) Catalan, J.; Kasha, M. *Journal of Physical Chemistry A* **2000**, *104*, 10812-10820.

(11) Becke, A. D. *Journal of Chemical Physics* **1993**, *98*, 5648.

(12) Lee, C.; Yang, W.; Parr, R. G. *Physical Review B* **1988**, *37*, 785.

(13) Curutchet, C.; Cramer, C. J.; Truhlar, D. G.; Ruiz-Lopez, M. F.; Rinaldi, D.; Orozco, M.; Luque, F. J. *J Comput Chem* **2003**, *24*, 284-297.

(14) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Journal of Physical Chemistry* **1992**, *96*, 135.

(15) M. J. Frisch, G. W. T., H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C.

Gonzalez, and J. A. Pople In;; Gaussian, Inc.: Pittsburgh, PA, 1995.

(16) Sloan, M. J.; Phillips, R. S. *Bioorganic and Medicinal Chemistry Letters* **1992**, *2*, 1053-1056.

(17) Niu, C.; Li, J.; Doyle, T. W.; Chen, S.-H. *Tetrahedron* **1998**, *54*, 6311-6318.

(18) Mahadevan, I.; Rasmussen, M. *Tetrahedron* **1993**, *49*, 7337-7352.

(19) Gordon, M. S. *Journal of Physical Chemistry* **1996**, *100*, 3974-3979.

(20) Chaban, G. M.; Gordon, M. S. *Journal of Physical Chemistry A* **1999**, *103*, 185-189.

(21) Jamorski Jodicke, C.; Luthi, H. P. *J Am Chem Soc* **2003**, *125*, 252-264.

(22) von Rague Schleyer, P.; Manoharan, M.; Wang, Z. X.; Kiran, B.; Jiao, H.; Puchta, R.; van Eikema Hommes, N. J. *Org Lett* **2001**, *3*, 2465-2468.

Figure 1. Fluorescence spectra of 4AI and N4-Methyl-4AI.

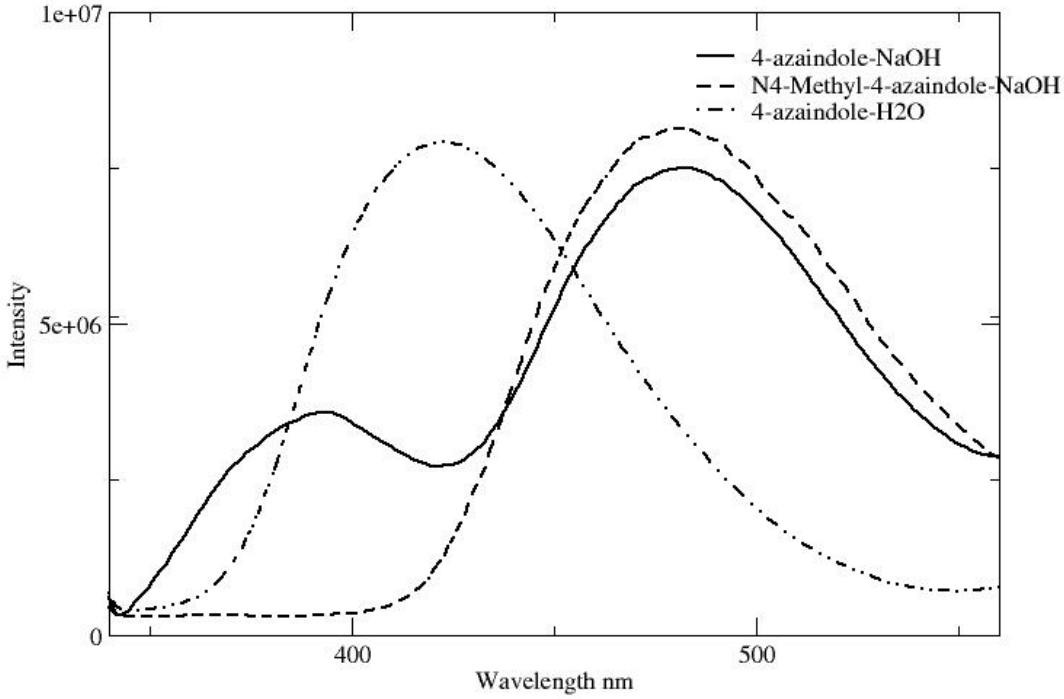


Figure 2. Fluorescence spectra of 5AI and N5-Methyl-5AI.

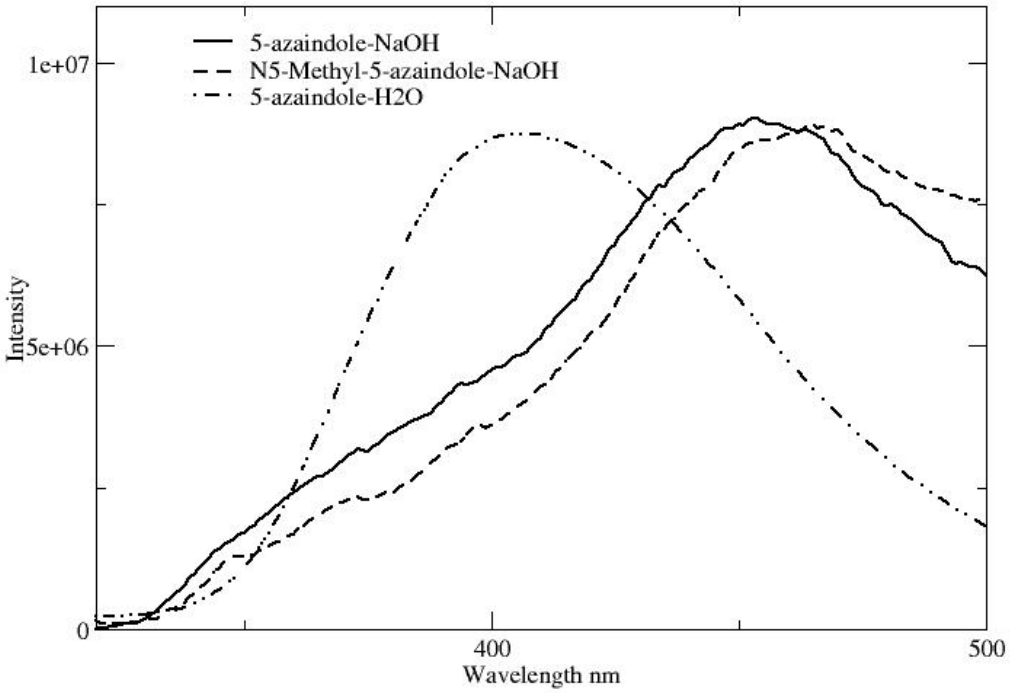


Figure 3. Solvent effects on the Azaindole Isomers.