Mechanistic Studies on (*S*)-Norcoclaurine Synthase and Dimethylallyltryptophan Synthase

by

Louis Yu Pan Luk

B.Sc., The University of British Columbia, 2004

A THESIS SUMBITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Chemistry)

The University of British Columbia
(Vancouver)

September, 2010

Abstract

In alkaloid biosynthesis, there are a limited number of enzymes that can catalyze an aromatic electrophilic substitution. One example is norcoclaurine synthase, which catalyzes an asymmetric Pictet-Spengler condensation of dopamine and 4-hydroxyphenylacetaldehyde to give (*S*)-norcoclaurine. This is the first committed step in the biosynthesis of the benzylisoquinoline alkaloids that include morphine and codeine. In this work, the gene encoding for the *Thalictrum flavum* norcoclaurine synthase is highly overexpressed in *Escherichia coli* and the His-tagged recombinant enzyme is purified for the first time. A continuous assay based on circular dichroism spectroscopy is developed and used to monitor the kinetics of the enzymatic reaction. Dopamine analogues bearing a methoxy or hydrogen substituent in place of the C-1 phenolic group were readily accepted by the enzyme whereas those bearing the same substituents at C-2 were not. This supports a mechanism involving a two-step cyclization of the putative iminium ion intermediate that does not proceed via a spirocyclic intermediate. The reaction of [3,5,6-$^{2}H_3$]-dopamine was found to be slowed by a kinetic isotope effect of 1.7 ± 0.2 on the value of k_{cat}/K_M. This is interpreted as showing that the deprotonation step causing re-aromatization is partially rate determining in the overall reaction.

Dimethylallyltryptophan synthase is an aromatic prenyltransferase that catalyzes an electrophilic aromatic substitution between dimethylallyl diphosphate (DMAPP) and L-tryptophan. The synthase catalyzes the first committed step in the ergot alkaloid biosynthesis. The enzymatic reaction could follow either an S_N1 reaction involving a discrete dimethylallyl cation intermediate or an S_N2 mechanism in which the indole ring directly displaces diphosphate in a single step. In this work, positional isotope exchange experiments are presented in support of an S_N1 pathway. When [1-^{18}O]-DMAPP is subjected to the synthase reaction, 15% of the ^{18}O-label is found to have scrambled from a bridging to a non-bridging position on the α-phosphorus.

Kinetic isotope effect studies show that steps involved in the formation of the arenium ion intermediate are rate-determining, and therefore the scrambling occurs during the lifetime of the dimethylallyl cation/diphosphate ion pair. Similarly, when the unreactive substrate analogue, 6-fluorotryptophan, was employed, complete scrambling of the ^{18}O-label in DMAPP was observed.

Preface

A version of Chapter 2 has been published and some of the experimental results are reproduced with permission from: Luk, L. Y. P.; Bunn, S.; Liscombe, D. K.; Facchini, P. J.; Tanner, M. E. *Biochemistry* **2007**, *46*, 10153-10161 (© 2007 American Chemical Society). The plasmid carrying the norcoclaurine synthase gene was donated by Dr. Liscombe and Professor Facchini at the University of Calgary. However, the gene expression and mechanistic investigations reported in this publication and Chapter 2 were performed by the author of this thesis under the supervision of Professor Tanner. In addition, some of the studies were performed with the assistance of Ms. Bunn, who was an undergraduate research assistant during May to August 2006 in the Tanner group.

A version of Chapter 4 has been published, and some of the experimental results are reproduced with permission from: Luk, L. Y. P.; Tanner, M. E. *J. Am. Chem. Soc.* **2009**, *131*, 13932-13933 (© 2009 American Chemical Society). The studies reported in this publication and Chapter 4 were performed by the author of this thesis under the supervision of Professor Tanner.

Table of Contents

List of Tables

List of Figures

Abbreviations and Symbols

4-HPAA	4-hydroxyphenylacetaldehyde
40-50mer	oligomers of 40 to 50 nucleotides
$[\alpha]^{25}_{D}$	specific rotation of a chemical compound using sodium D line (589 nm) at 25 ^{o}C
ΔE_H, ΔE_D	activation energy involving protiated and deuterated substrates, respectively
δ	chemical shift (ppm)
$[\theta]$	molar ellipticity (mdeg cm^{-1} M^{-1}) measured in circular dichroism spectroscopy
γ-Glu-P	γ-glutamyl phosphate
ε	molar absorptivity
ε_{360}	extinction coefficient at 360 nm
A_{360}	absorbance of light at 360 nm
ABBA	a unique protein fold found in certain aromatic prenyltransferases, such as NphB and DMATS
Ac	acetyl
AcOH	acetic acid
Ado	adenosine
ADP	adenosine diphosphate
ATP	adenosine triphosphate
BINOL	1,1'-bi-2-naphthol
Bn	benzyl
BSA	bovine serum albumin
CD	circular dichroism
CH_3CO_2AMP	acetyl adenylate
D	deuterium (^2H)

Da	Dalton
DMAPP	dimethylallyl diphosphate
DMAT	4-dimethylallyltryptophan
DMATS	4-dimethylallyltryptophan synthase
dmaW	gene of dimethylallyltryptophan synthase from *Claviceps* sp.
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
k_H/k_D	deuterium kinetic isotope effect (on k_{cat}/K_M)
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetate, disodium salt
ESI-MS	electrospray ionization mass spectrometry
Et_3N	triethylamine
$EtNH_2$	ethylamine
Et_2O	diethyl ether
EtOAc	ethyl acetate
fgaPT2	gene of dimethylallyltryptophan synthase from *Aspergillus fumigatus*
F_H	fractional conversion of a protiated compound to products
FPP	farnesyl diphosphate
GPP	geranyl diphosphate
HPLC	high performance liquid chromatography
*i*PrOH	isopropyl alcohol
IPP	isopentenyl diphosphate
IPTG	isopropyl 1-thio-β-D-galactopyranoside
J	coupling constant (NMR); subscripts indicate coupling partners

kb	kilobases (1000 base pair)
k_{cat}	catalytic rate constant
k_{cat}/K_M	specificity constant; second-order rate constant
k_H, k_D	rate of reaction involving protiated and deuterated substrates
kDa	kilodalton
KIE	kinetic isotope effect
K_M	Michaelis constant
LAD	lithium aluminum deuteride
LB	Luria-Bertani medium
MeOH	methanol
MESG	2-amino-6-mercapto-7-methylpurine ribonucleoside
mRNA	messenger ribonucleic acid
MWCO	molecular weight cut-off
n	number of substrate binding sites (allostery)
NCS	norcoclaurine synthase
NMR	nuclear magnetic resonance
NphB	bacterial aromatic prenyltransferase isolated from *Streptomyces* sp. strain CL190
OD_{600}	optical dispersion at 600 nm
P	natural abundance of ^{13}C isotope
pRARE	a vector carrying unconventional tRNA genes for heterologous expression in *E. coli* cells
p-TosOH	*p*-toluenesulfonic acid
PCR	polymerase chain reaction
pet. ether	petroleum ether

P_i	inorganic phosphate
PP_i	pyrophosphate
PIX	positional isotope exchange
PT	prenyltransferase
R	final ratio of isotopically labeled to unlabeled compound
R_o	initial ratio of isotopically labeled to unlabeled compound
rpm	revolutions per minute
RNA	ribonucleic acid
RT	room temperature
SDM	site-directed mutagenesis
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
ssp	subspecies
TEAP	*bis*-triethylammonium phosphate
TFA	trifluoroacetic acid
*tfNCS*Δ19	gene of norcoclaurine synthase isolated from *Thalictrum flavum*
TB	Terrific Broth
THF	tetrahydrofuran
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
tRNA	transfer ribonucleic acid
UV	ultraviolet
v	initial reaction velocity (rate)
v_{chem}	percentage of starting material (DMAPP) converted to product
v_{pix}	percentage of PIX extent
v_{pix}/v_{chem}	partitioning ratio measured in PIX experiment

WT	wild type

Common Amino Acid Abbreviations

A	Ala	alanine
C	Cys	cysteine
D	Asp	aspartate
E	Glu	glutamate
F	Phe	phenylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	methionine
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	serine
T	Thr	threonine
V	Val	valine
W	Trp	tryptophan
Y	Tyr	tyrosine

Nucleotide Base Abbreviations

A	adenine
C	cytosine
G	guanine
T	thymine
U	uracil

Acknowledgements

It has almost been eight years since I've become a proud member of the Tanner Group, and as the longest serving member so far on the team, I have been very lucky to have received the unwavering support and guidance from some of the most talented individuals. First and foremost, I would like to extend my sincere gratitude to my supervisor, Professor Martin Tanner, for his inspiration, wisdom and enthusiasm. Since my undergraduate directed study, Professor Tanner has guided me through the long and often challenging programs with patience and encouragement. Without his guidance my path towards a Ph.D. would certainly be much tougher, so 'Thank You'.

I would also need to thank those who have made this past six years truly enjoyable and inspiring. I would like to express my appreciation to all the past and present members of the Tanner group, with whom I have enjoyed many memorable and comical moments, and for the assistance and friendship they have bestowed. I would like to give my special thanks to Dr. Wayne Chou, Dr. Andrew Murkin, Dr. James Morrison, and Dr. Jay Read, for their useful guidance in laboratory techniques. I would also like to thank Qi Qian for his efforts to the works on the site-directed mutagenesis. I owe my gratitude to the Sherman group, including Jon Freeman and Grant Bare, for their assistance with HPLC, and to the Facchini group for providing the plasmid carrying the gene of norcoclaurine synthase. Additionally, the invaluable aid I have received from the staff members in the Biological Services Laboratory, Center for Blood Research, NMR and mass spectrometry facilities, including Dr. Elena Polishchuk, Candice Martins, Jesse Chen, Dr. Fred Rosell, Dr. Maria Ezhova, Dr Yun Ling, Minaz Lakha, David Wong, was much appreciated. Lastly, I am thankful for the endless support of my family and friends.

Dedicated to

the readers

Chapter 1

Alkaloids and Enzyme-catalyzed Pictet-Spengler Reactions

1.1 Alkaloids: Brief History and Definition

Alkaloids have been intricately intertwined with the history of humans, as they possess a vast variety of physiological properties. This has led to their exploitation as pharmaceuticals, stimulants, narcotics and poisons.[1–3] One major historical event related to the use of an alkaloid was the death of the philosopher Socrates, who was accused of corrupting the minds of the youth of Athenians. He was sentenced to death in 399 BC by drinking a solution of hemlock extract which contained the potent neurotoxin alkaloid coniine (Figure 1.1).[3] The epidemics of ergotism between the years of 1400 and 1668 were another historical event that was related to alkaloids. The fungus *Claviceps purpurea* that periodically infests rye and produces ergot alkaloids as secondary metabolites caused hallucination and even death to the people consuming it. Historians have speculated that the phenomenon was the main cause for the witch hunts that unfortunately led to the death of about 30,000 people.[2, 3] On the other hand, alkaloids have also been exploited for good purposes. For instance, in the 1600s the native Americans showed the Spanish missionaries that cinchona bark can be used as an anti-malaria agent. This herb was later shown to contain the alkaloid quinine, a medicine that is still currently used for treating malaria.[1–3] Even though a variety of alkaloids have been utilized for over thousands of years, the term 'alkaloid' was not actually introduced until 1819 when Carle Meissner discovered the alkaline property of the famous analgesic morphine isolated from the opium poppy *Papaver somniferum*.[1–3] To date, the alkaloids are known to be an enormous group of natural products composed of about 13,000 compounds, many of which have been extensively examined. A wide variety of these compounds have been used for clinical purposes. Some examples include the analgesics morphine and codeine, the anticancer drugs vinblastine and paclitaxel (taxol), the muscle relaxant tubocurarine, the antiarrhythmic ajmaline and the antibiotic

sanguinarine (Figure 1.1). Other well-known alkaloids related to humans are caffeine, nicotine, cocaine and the semi-synthetic morphine derivative heroin (Figure 1.1&1.2).[1–3]

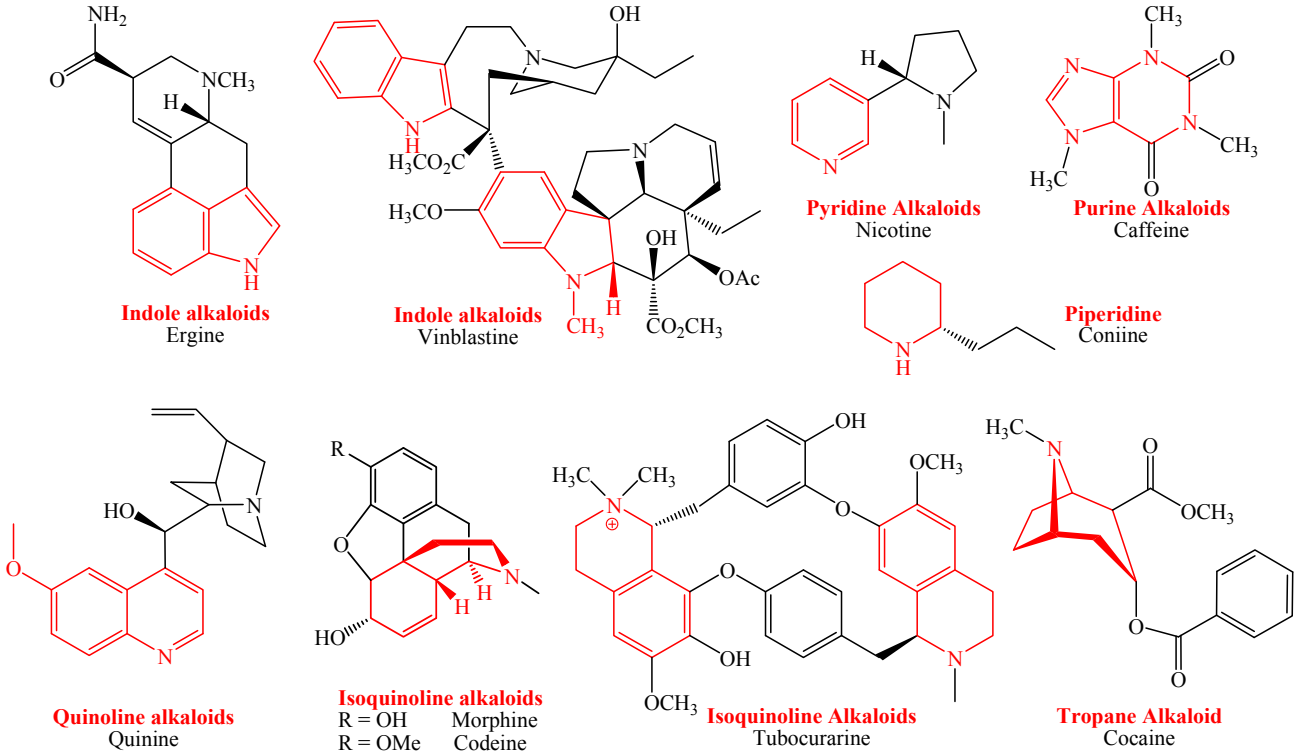


Figure 1.1 Examples of alkaloids (Red chemical bonds denote the chemical moiety to which the alkaloid is sub-classified)

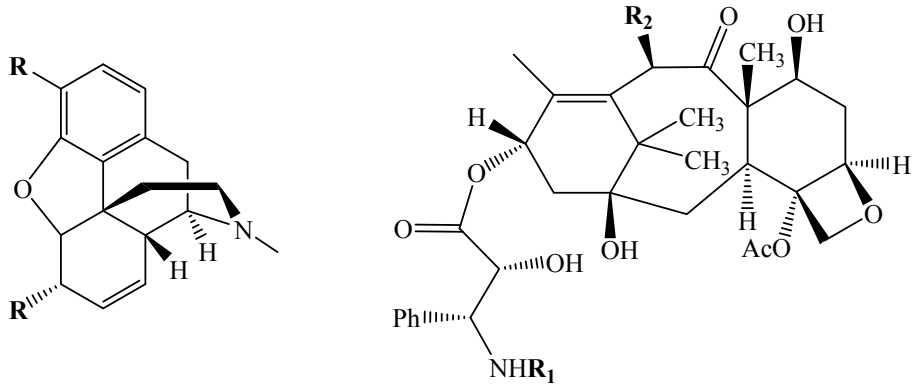
Because the alkaloids are one large family of natural products with highly diversified chemical structures, there is actually no single definition for the term 'alkaloid' that could be established without contradictions and/or exceptions.[1, 3] Most researchers now generally think of alkaloids as amino-acid derived, basic and nitrogen-containing secondary metabolites that often have a high physiological activity as most of them play a defensive role in the host organisms. However, amino acids, peptides, proteins, nucleotides, nucleic acids and amino sugars are typically not considered as alkaloids.[3] Furthermore, alkaloids can be sorted into different classes according to the location of the nitrogen(s) within the structure. Interestingly, the classification system itself is quite complicated as it further branches into many subcategories based on their featured chemical

functionalities and the organisms from which they are isolated.[4] One example is the class of heterocyclic alkaloids, which refer to natural products that possesses nitrogen atom(s) inside a cyclic carbon ring. According to their featured ring moieties, these alkaloids are further subcategorized as follows: the indole alkaloids such as ergine and vinblastine, the pyridine alkaloids such as nicotine, the purine alkaloids such as caffeine, the piperidine alkaloids such as coniine, the quinoline/isoquinoline alkaloids such as quinine, tubocurarine, morphine and codeine, and finally the tropane alkaloids such as cocaine (Figure 1.1). Other alkaloid classes include exocyclic nitrogen, terpene and steroidal, and peptide alkaloids.[4] In spite of the obscure definition and the complicated classification system, extensive research has focused on the properties and biosynthesis of alkaloids as they possess potent biological activity and interesting complex structures.

1.2 Biosynthesis of Alkaloids

Despite the wealth of information that is available on their pharmacological effects, obtaining natural products with good yield is often perceived as a challenging task. While the extraction of an alkaloid from its natural source can be very tedious, time-consuming and expensive, its chemical synthesis can also be too complicated and/or too expensive to be optimized on an industrial scale. For example, only 3 kg of morphine can be extracted from a one acreage field of the opium poppy *P. somniferum*. On the other hand, developing an efficient synthetic pathway for morphine has also been a long-standing goal in organic chemistry,[5, 6] and the most recent asymmetric synthesis that has been developed by Parker consists of 13 steps with an overall yield of 1.7% (Figure 1.2).[7] In another illustration, one single dose of the anti-tumor drug paclitaxel (taxol) consumes about one

yew tree of *Taxus brevifolia* with an isolation yield of ~0.02%,[8] whereas the synthesis of this compound that was first reported by Holton in 1994 gave an overall yield of 2.7% (Figure 1.2).[9, 10]



R = OH Morphine
R = OAc Heroin

R1 = $CO_2C_6H_5$ **R2** = CO_2CH_3 Paclitaxel
R1 = $CO_2C(CH_3)_3$ **R2** = OH Taxotere

Figure 1.2 Morphine, taxol and their analogs

A vast amount of research has been focused on gaining insights into the biosynthesis of complicated natural products, because this knowledge could be used to develop a new method for generating these compounds with a high yield. For instance, the enzymes cloned from the alkaloid-producing host could be employed to establish an efficient chemo-enzymatic synthetic pathway. In addition, current molecular biology research has focused on developing cell-lines derived from the alkaloid-producing organism in order to generate these natural products at a high yield *in vitro*. Lastly, different types of alkaloid analogs have been prepared in an attempt to moderate the toxicity of the original compound while still retaining its curative effect as a medication. Indeed, heroin was originally made with the hope that it could replace morphine by maintaining its analgesic property but be less addictive.[11] Similarly, a paclitaxel analog, taxotere, is another anti-cancer agent that might have a greater efficacy and better solubility than the original natural product itself (Figure 1.2).[12] In recent years, rational design and/or directed evolution have focused on enzyme-engineering in an attempt to widen an enzyme's substrate specificity and thereby allow for the generation of a variety of natural product analogs.

The study of alkaloid biosynthesis has a history of only about 60 years. One major obstacle was probably the elucidation of the chemical structure of alkaloids, particularly of those containing multiple stereogenic centers. For example, although nicotine contains only one stereogenic center and was discovered in 1828, its structure was not determined until it was synthesized in 1904. The chemical structure of morphine contains five stereogenic centers and was not elucidated until 1952, almost 150 years after its isolation.[3] Studies on alkaloid biosynthesis were finally underway in the late 1950s. Initially, radiolabeled precursors were introduced to plants and the resultant radioactive alkaloids were chemically degraded to identify the positions of the label. As the development of analytical instrumentation advanced, the isotopically labeled natural products were analyzed by mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy instead of chemical degradation.[3] Nevertheless, the actual enzymes responsible for alkaloid biosynthesis had not yet been identified until the 1970s when the plant and fungal cell cultures were exploited. Since then, over a hundred novel enzymes that participate in alkaloid biosynthesis have been identified. With the recent improvements in the techniques of molecular biology, many enzymes can be overexpressed heterologously in organisms, such as *Escherichia coli*, that have better fermentation characteristics than their original hosts (e.g. plant and fungal cell cultures). These technological advances greatly enhance the opportunities to study each of the enzymes in alkaloid biosynthesis.[1–3]

1.3 Aromatic Electrophilic Substitutions in Nature

As is evident in the previous examples, many alkaloids possess polycyclic structures containing both aliphatic and aromatic moieties (Figure 1.1). Consequently, one can anticipate that there are a variety of enzymes capable of catalyzing aromatic electrophilic substitutions in alkaloid

biosynthetic pathways and thereby forming C-C bond between an aromatic ring and an alkyl group. In organic synthesis, the prototypical aromatic electrophilic substitution reaction is the Friedel-Crafts alkylation, which involves an alkyl halide, an aromatic molecule and a Lewis acid catalyst such as aluminum chloride (Figure 1.3). The reaction is initiated when the Lewis acid abstracts a halide to generate an electrophilic carbocation intermediate. Electrophilic addition by the aromatic ring then yields an arenium ion intermediate. Lastly, deprotonation causes re-aromatization and yields the substituted aromatic as product. Of course, in most biological systems alkyl halides and Lewis acid catalysts are not commonly available and do not serve as substrate and catalyst, respectively. Instead, an enzyme serves as a surrogate for the chemical Lewis acid catalyst, and substrates other than alkyl halides are employed. This leads to some interesting questions in the field of natural product chemistry and biochemistry: how do enzymes replace the role of Lewis acid catalyst to initiate a wide variety of reactions? Moreover, what type of molecules serve as substrates in these reactions?

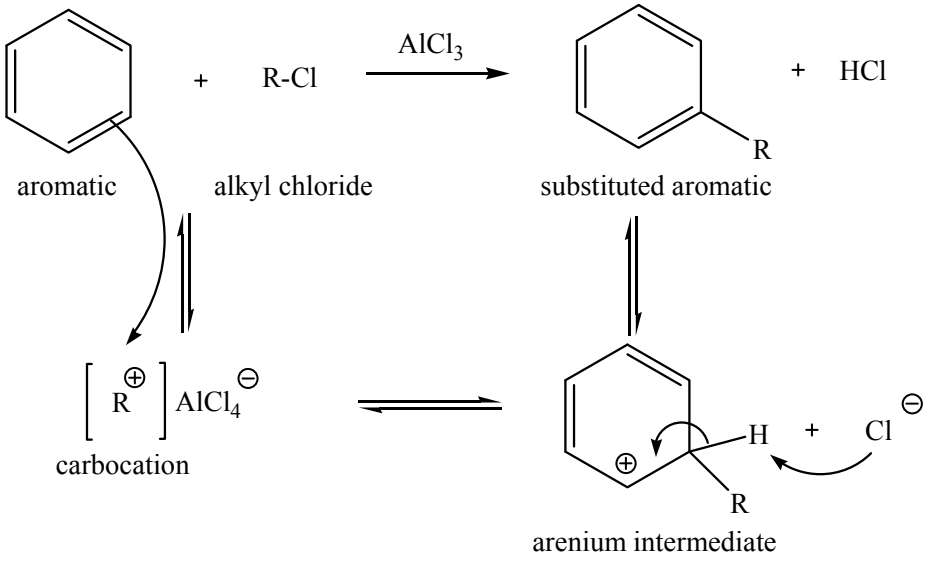


Figure 1.3 General Friedel-Crafts reaction mechanism

The discoveries of enzymes that catalyze aromatic electrophilic substitutions are somewhat rare, and few have been studied in detail. The enzymes histidine ammonia-lyase and phenylalanine

ammonia-lyase remove the ammonia moiety from the corresponding amino acids to give urocanic acid and cinnamic acid, respectively (Figure 1.4). Even though these enzymes do not catalyze an overall substitution reaction, they have been referred as the first reported set of enzymes that catalyze a Friedel-Crafts-like reaction. Interestingly, these enzymes promote an electrophilic addition of the methylene group from the 5-methylene-3,5-dihydroimidazol-4-one (MIO) cofactor to the aromatic ring of the substrate (histidine or phenylalanine) as a key step in the catalysis.[13] Porphobilinogen deaminase is another example that catalyzes three sequential electrophilic aromatic substitution reactions in the biosynthesis of tetrapyrroles.[13, 14] Several Friedel-Crafts type enzymes have recently been identified in the alkaloid biosynthetic pathways. The norcoclaurine synthase in benzylisoquinoline alkaloid production and the dimethylallyltryptophan synthase in ergot alkaloid production are two of these examples, and the investigation of their mechanisms will be the focus of this thesis.[15–18] Background information related to these two enzymes will be elaborated below.

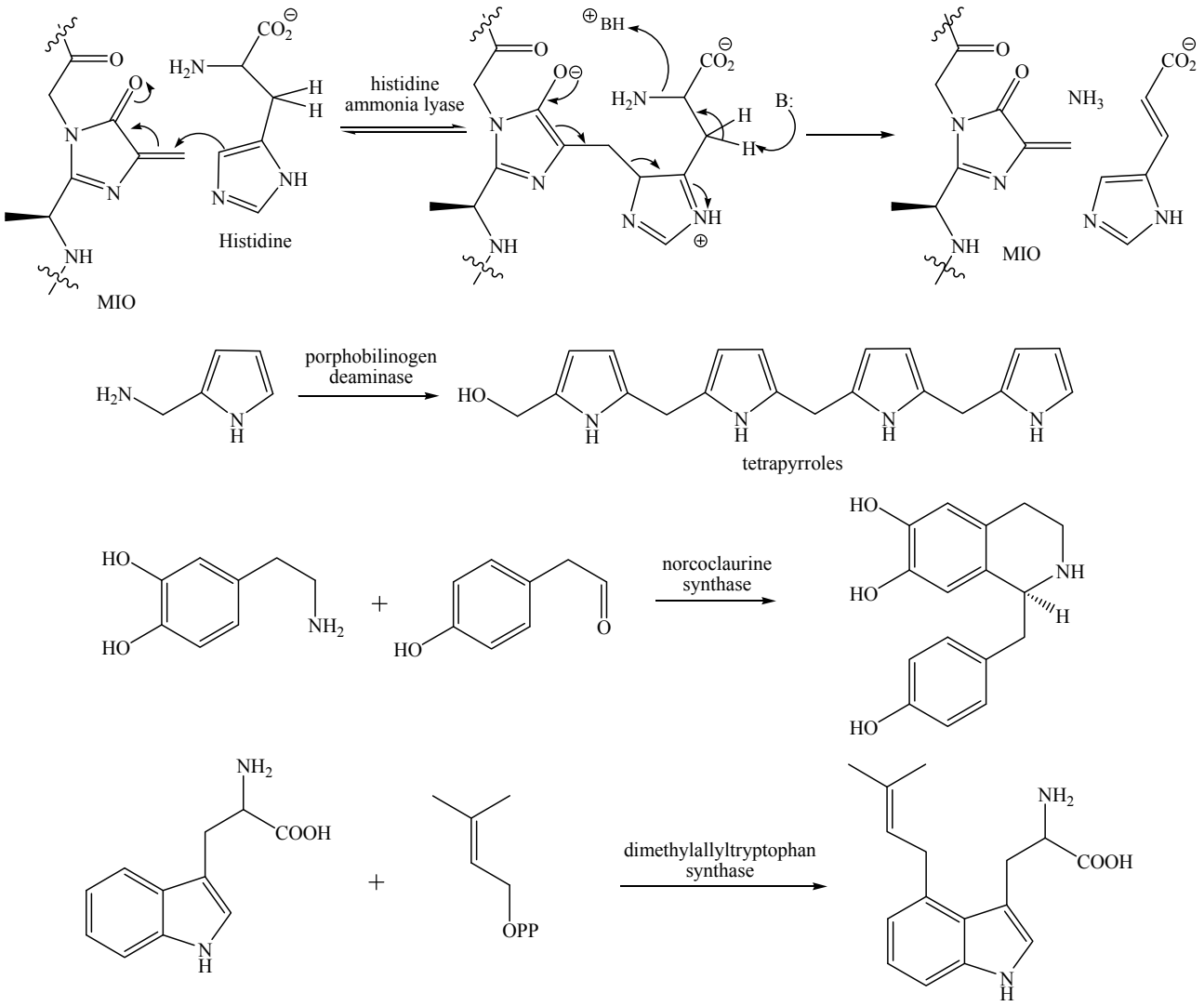


Figure 1.4 Examples of enzymes catalyzing Friedel-Crafts-type reactions

1.4 The Pictet-Spengler Reaction: Discovery and Development

The Pictet-Spengler reaction was first reported by Amé Pictet and Theodor Spengler in 1911, and this reaction is essentially a hybrid between an aromatic electrophilic substitution and the Mannich reaction.[19, 20] In this acid-catalyzed reaction an electron-rich β-arylethylamine molecule, such as tryptamine or phenethylamine, undergoes a condensation with an aldehyde to form an

iminium intermediate (Figure 1.5). Subsequently, an intramolecular ring closure step creates the products, β-carboline and isoquinoline, respectively.[19, 20]

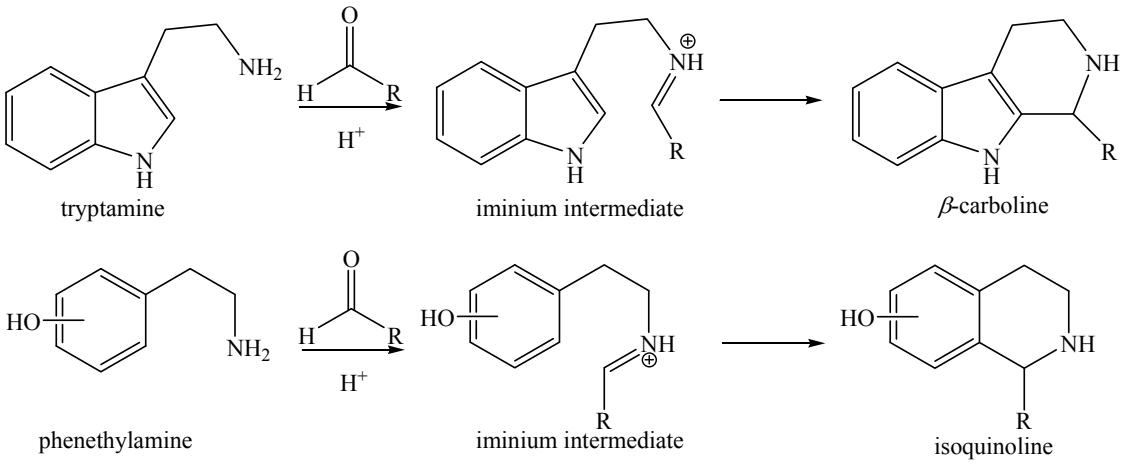


Figure 1.5 Pictet-Spengler reactions that produce β-carboline and isoquinoline

The Pictet-Spengler reaction has become a useful tool to create a large variety of β-carboline and isoquinoline derivatives, which can serve as a key intermediate in the synthesis of different useful alkaloids.[20] Some examples of β-carboline alkaloids are yohimbine and ajmaline that have been shown to possess potent stimulatory and cardiovascular effects, respectively (Figure 1.6).[21, 22] Macroline has been proposed to have an anti-malarial and an anti-tumor activity.[23] Some isoquinoline alkaloids are also valuable natural products, such as the analgesics morphine and codeine and the heart stimulating reagent (*S*)-norcoclaurine.[1, 24] Macarpine was shown to possess anti-tumor activity, whereas sanguinarine and berberine were shown to possess antimicrobial activities.[1] Many of the β-carboline and isoquinoline alkaloids possess multiple stereogenic centers and synthesizing them is not trivial. Hence, a lot of effort has been invested in developing an efficient asymmetric Pictet-Spengler reaction.

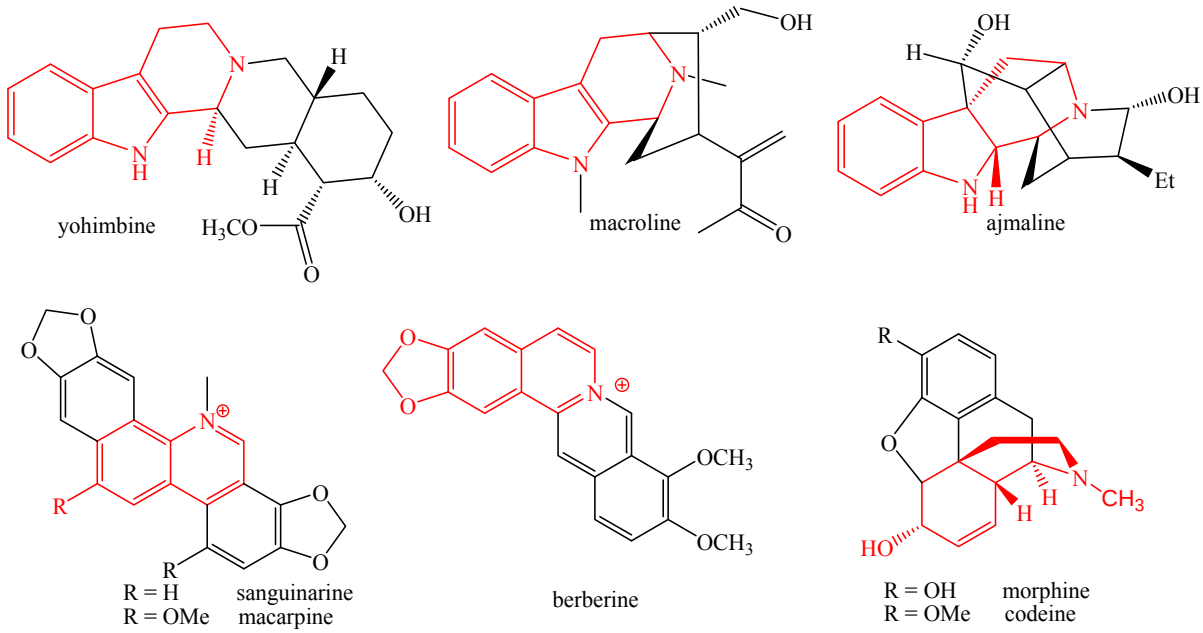


Figure 1.6 Different isoquinoline and β-carboline analogs. Red denotes the β-carboline and isoquinoline backbones.

1.4.1 Non-enzymatic Pictet-Spengler Reactions and their Mechanistic Studies

In order to develop an asymmetric Pictet-Spengler reaction, it is crucial to study the mechanism of the reaction. The Pictet-Spengler reaction with an indole-containing starting material has been extensively studied, and the mechanism turns out to be not as simple as a typical aromatic electrophilic substitution. Two plausible mechanistic pathways have been proposed for this reaction. In both cases the reaction begins with the formation of an iminium ion (Figure 1.7).[20] Both C-2 and C-3 of tryptamine are nucleophilic, although substitutions on free indole preferentially occur at C-3. Hence, the iminium functionality can alkylate the C-3 of the indole ring to yield a spiroindolenine intermediate and a subsequent 1,2-alkyl shift then generates a benzylic carbocation intermediate (pathway A, Figure 1.7). Alternatively, alkylation can occur directly at C-2 to give the same carbocation intermediate (pathway B). In either case, deprotonation of the carbocation intermediate

results in re-aromatization to form the β-carboline product.[20] Evidence that supports both mechanisms has been obtained and the predominant mechanism has not yet been identified.

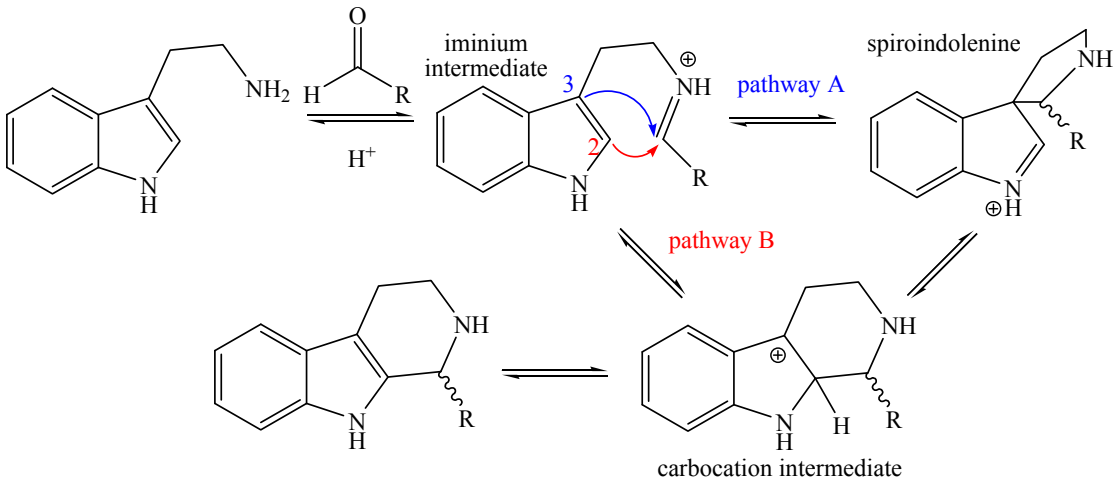


Figure 1.7 Two proposed Pictet-Spengler reaction pathways

In general, the indole-based Pictet-Spengler reaction is believed to proceed via the spiroindolenine intermediate (pathway A, Figure 1.7).[20] This mechanism was first suggested by Woodward and supported by his famous study in the synthesis of the toxin strychnine.[25–27] A condensed product 2-veratrylspiroindolenine (**2**) would form upon reflux with 2-veratryltryptamine (**1**) and the glyoxylic ester (Figure 1.8). This product was suggested to be a 'trapped'-spiroindolenine intermediate, because the C-2 position carries an aromatic moiety.[25–27] A similar observation was made when the tryptophan methyl ester (**3**) was allowed to undergo a Pictet-Spengler reaction in the presence of palladium and excess hydrogen, which trapped the spiroindolenine intermediate **4** by reducing the iminium bond to an amine functionality.[28] Perhaps the most convincing evidence is the observation of positional isotope exchange, when a hydrazine tryptamine analog **5** reacted with deuterium-labeled formaldehyde and gave the Pictet-Spengler products that have the deuterated methylene group at different positions. This suggested that the reaction involved a symmetrical spiroindolenine intermediate **6**.[29, 30]

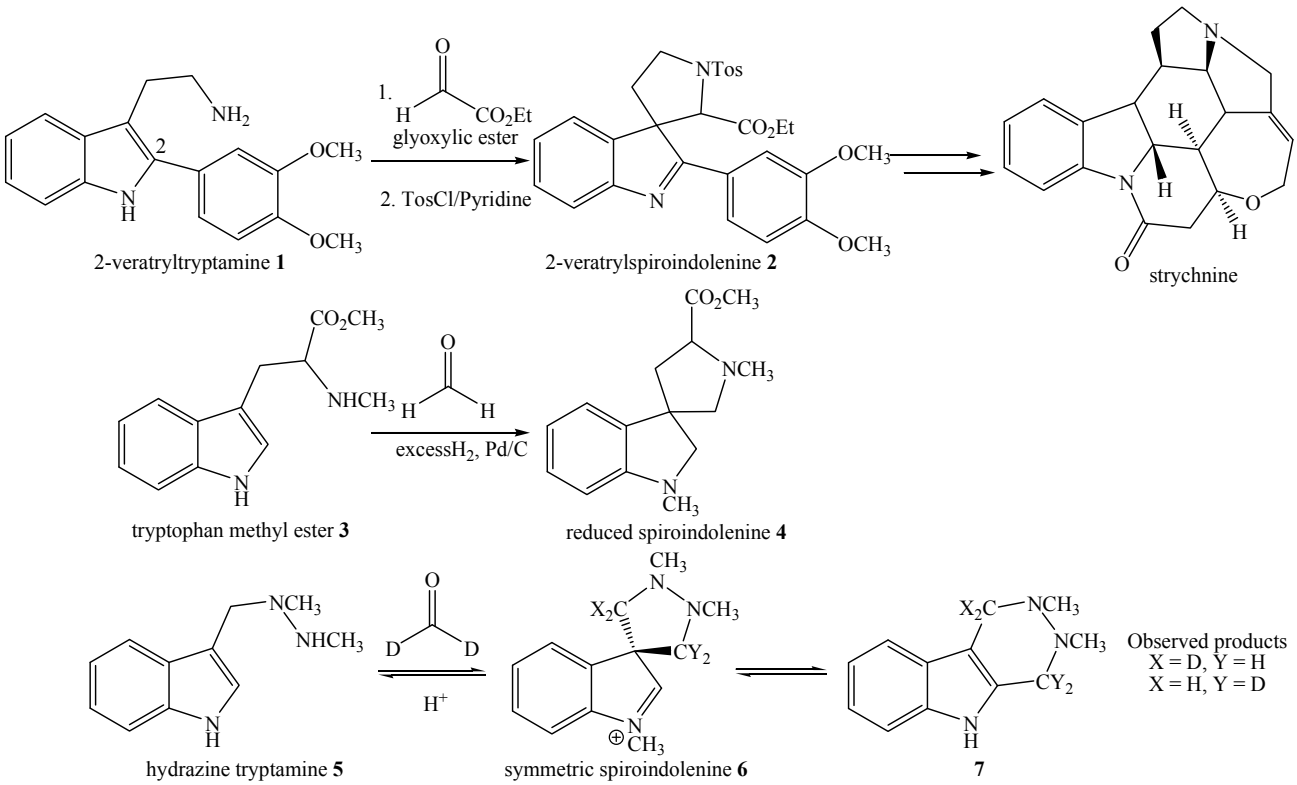


Figure 1.8 Experiments supporting formation of the spiroindolenine intermediate

In contrast to the observations mentioned above, experimental evidence that supports the direct nucleophilic attack in the Pictet-Spengler reaction (pathway B, Figure 1.7) has also been obtained. When *N*-(4,4-diethoxybutyl)tryptamine (**8**) was hydrolyzed in acidic media, it reacted intramolecularly to give the corresponding Pictet-Spengler product **9** (Figure 1.9). This strongly supported the direct nucleophilic attack, because formation of a strained spiroindolenine intermediate **10** with a *trans*-double bond in a 7-membered ring would be very unlikely.[31] Moreover, a computer modeling study suggested that the rearrangement of the spiroindolenine intermediate required a higher transition state energy than the direct nucleophilic attack, thereby favoring the mechanism of direct nucleophilic attack.[32] In fact, the generation of a spiroindolenine intermediate (pathway A, Figure 1.7) does not follow Baldwin's rules as the 5-*endo*-trig ring closure is disfavored, whereas a favorable 6-*endo*-trig ring closure is responsible for the direct nucleophilic attack

(pathway B).[31] In conclusion, although many studies have been conducted, the predominant mechanism for the Pictet-Spengler reaction has not yet been clearly verified.

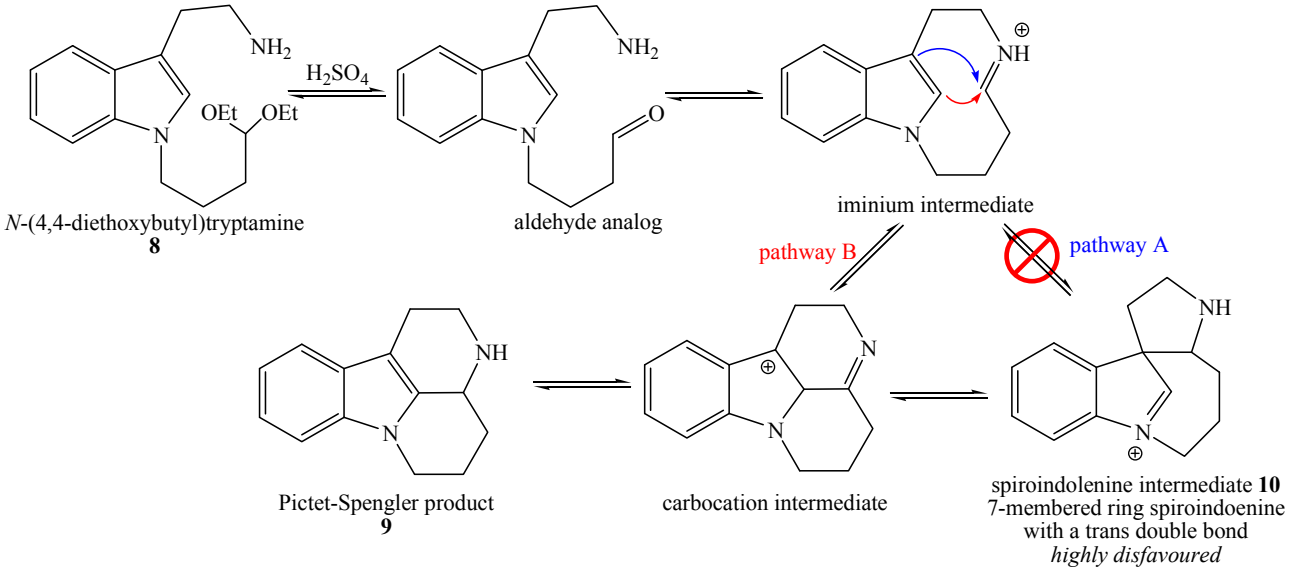


Figure 1.9 Intramolecular Pictet-Spengler reaction supporting the direct nucleophilic pathway

Mechanistic studies on the isoquinoline-generating Pictet-Spengler reaction have not been as extensive as those on the β-carboline-generating reaction. However, similar to the indole-ring system, there are also two mechanistic possibilities for the phenethylamine reaction that diverges after the formation of the iminium intermediate. The first possibility involves the generation of a spirocyclic intermediate followed by a 1,2-alkyl shift (pathway A, Figure 1.10) to produce a σ-intermediate. The second pathway forms the same σ-intermediate via a direct nucleophilic attack (pathway B). As mentioned above, both pathways are plausible for the indole-ring-based Pictet-Spengler reaction, because both the C-2 and the C-3 of the indole ring are relatively reactive. On the other hand, the mechanistic studies on isoquinoline production are complicated by the fact that the position of different substituents on the phenyl ring will affect its electronic distribution and possibly alter the mechanism of the Pictet-Spengler reaction.

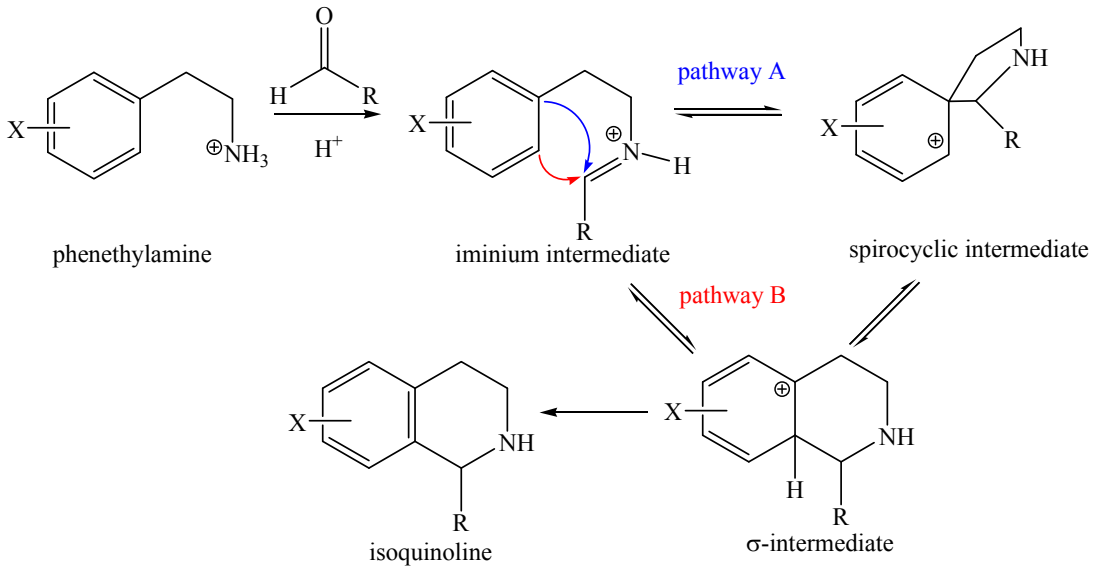


Figure 1.10 Pictet-Spengler reaction for the formation of isoquinoline

The mechanism of the 'formal' phenethylamine Pictet-Spengler reaction has not been inspected. Thus far, only a series of 'Pictet-Spengler-like' reactions that employed different benzylaminonitrile derivatives have been investigated by the research groups of Taylor and Waigh.[33-35] The authors developed an acid-catalyzed cyclization reaction of 1-(3,4-dimethoxybenzylamino)benzyl-carbonitrile (**11**, Figure 1.11).[34] Upon acidification, the acetonitrile functionality was protonated and underwent a direct nucleophilic attack from the C-6 of the phenyl ring to give the product dimethoxyisoquinolone (**12**). The reaction also created an alternative product 2-benzazepine (**13**) with a yield of 33%, which was postulated to come from an indirect nucleophilic attack pathway. The C-1 position of the 3,4-dimethoxybenzyl moiety attacks intramolecularly to give a spirocyclic intermediate, which further rearranges to an iminium intermediate **14** and give the corresponding product **13**. This observation is particularly fascinating, because the proposed indirect alkylation reaction essentially contains two back-to-back Pictet-Spengler reactions that begin with the spirocyclic intermediate.[34] This observation revealed that both proposed pathways are possible for the isoquinoline-producing Pictet-Spengler reaction.

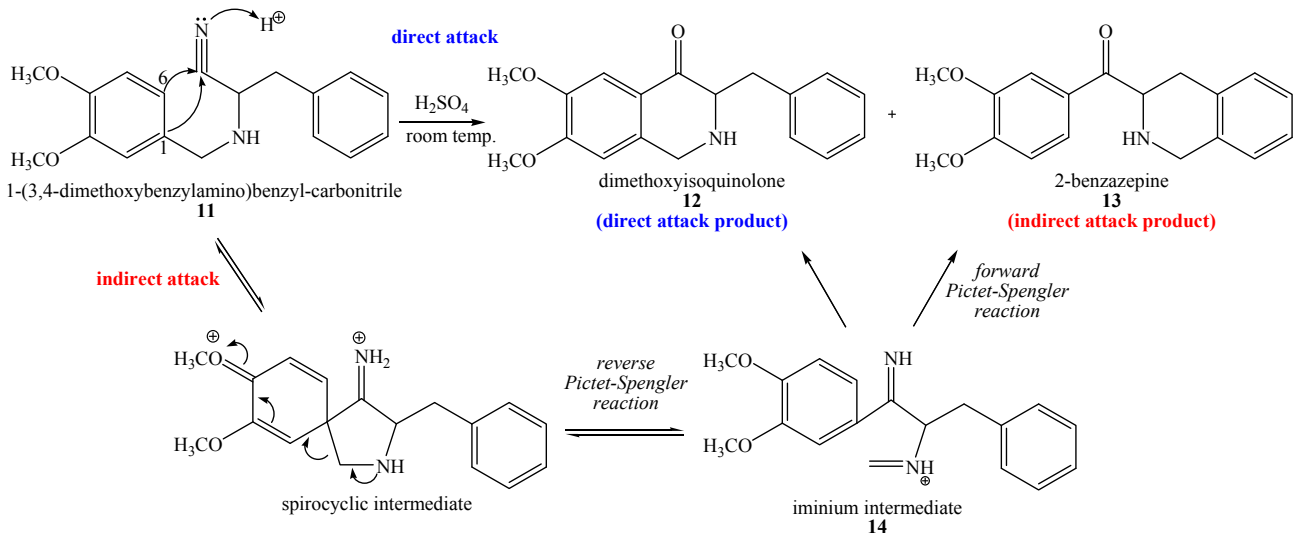


Figure 1.11 Pictet-Spengler-like reaction for isoquinoline derivatives. Blue denotes the direct nucleophilic attack pathway, and red denotes the indirect nucleophilic attack.

1.4.2 Development of Asymmetric Pictet-Spengler Reactions

The development of an asymmetric Pictet-Spengler reaction is often complicated by the high tendency of stereochemically pure β-carbolines to racemize. In a study carried out by Gözler and his co-workers, (*R*)-roeharmine (**15**) was an alkaloid that was isolated from its natural source, and it was claimed to be stereochemically pure with an optical activity of $[\alpha]^{25}_{D} = -4^{o}$ (c = 0.12 g/mL, methanol).[36] Interestingly, the optical activity of the synthetically prepared (*R*)-roeharmine was $[\alpha]^{25}_{D} = -18^{o}$ (c = 0.12 g/mL, methanol).[37] Hence, (*R*)-roeharmine was proposed to be partially racemized during the acid/base mediated isolation procedure. To verify this hypothesis, the chemically made (*R*)-roeharmine was exposed to a catalytic amount of trifluoroacetic acid in dichloromethane at room temperature. While the proton NMR spectrum and the R_f value of the alkaloid remained unchanged, the optical activity dropped to $[\alpha]^{25}_{D} = -0.8^{o}$ (c = 0.12 g/mL, methanol). In order to explain this observation, two potential racemization mechanisms were proposed. In the first mechanism, the sigma C1-N bond is broken and regenerated leading to a

mixture of stereoisomers (pathway A, Figure 1.12). In an alternative pathway, the indole ring can be protonated, leading to the deprotonation at the C-1 position and the consequent racemization of roeharmine (pathway B, Figure 1.12). The validity of these two pathways was tested by running the racemization reaction in the presence of deuterated trifluoroacetic acid. Because deuterium was not incorporated at the C-1 position, the first racemization pathway (A) is favored over the second.[36, 37] Another plausible explanation that might account for the racemization is a reverse Pictet-Spengler reaction, and upon re-forming the product a racemic mixture would result.[20] However, it was believed that the reversal of a Pictet-Spengler reaction should require a high temperature environment to proceed, thus this hypothesis is considered less likely to occur. Furthermore, other indole alkaloids such as resperine (**16**) and alloyohimbine (**17**) also lose their stereochemical purities in acidic conditions. To sum up, the presence of acid should be avoided when carrying out an asymmetric Pictet-Spengler reaction and during the purification of the product.[20, 38] Lastly, the possibility for a similar racemization reaction in isoquinoline derivatives has never been explored.

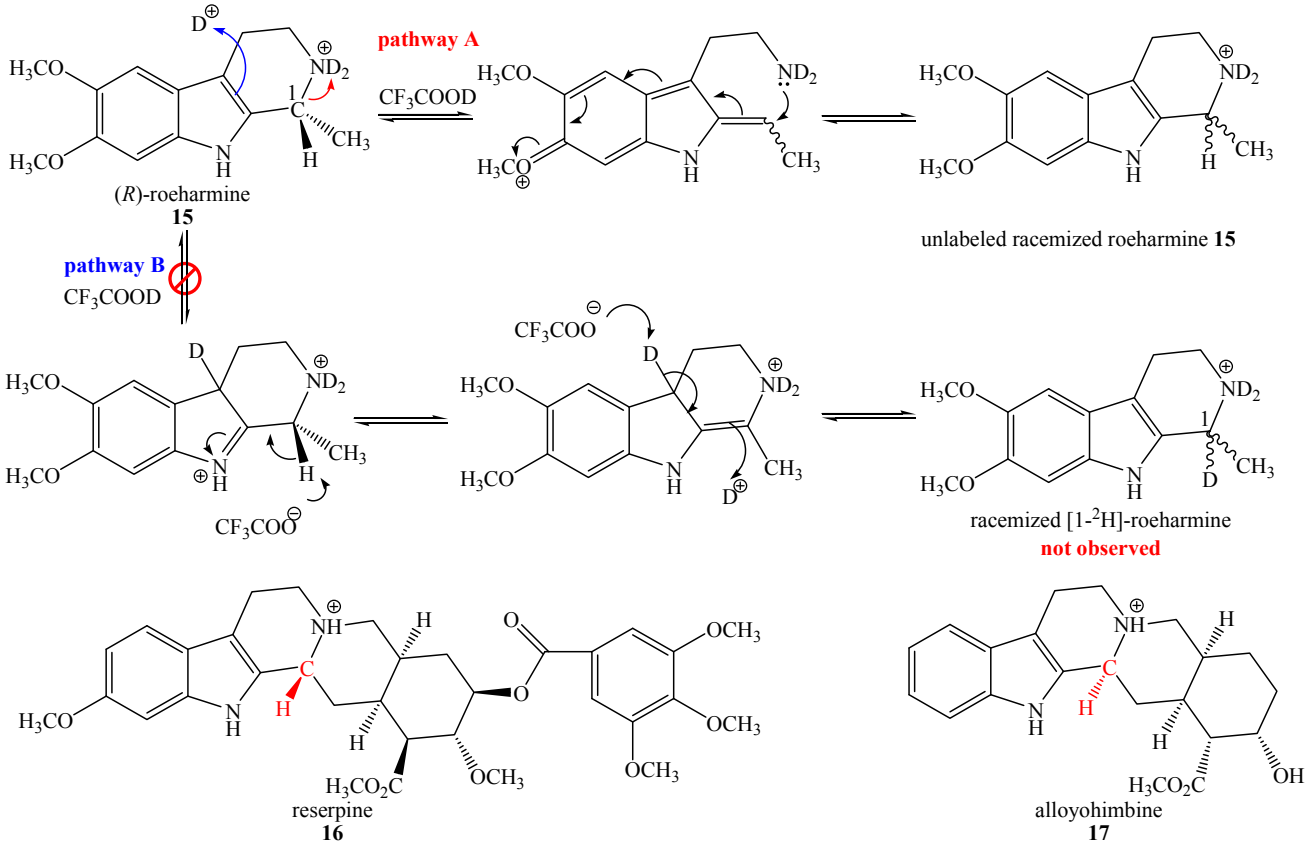


Figure 1.12 Epimerization in asymmetric β-carboline natural products by cleavage across the carbon-nitrogen bond. Red denotes the carbon center being racemized upon acidification

Despite the mechanistic questions and the high tendency to racemize, research has been done on developing an asymmetric Pictet-Spengler catalyst because of the many useful β-carboline and isoquinoline alkaloids that contain stereogenic centers. The chiral thiourea derivatives developed by Jacobsen, and the BINOL-phosphoric acid by List are the only two asymmetric Pictet-Spengler reaction catalysts available, and they are designed to produce β-carboline molecules in a stereoselective fashion (Figure 1.13). Although the catalytic mechanisms have not yet been determined, it is likely that both catalysts control the stereochemistry of reactions via steric interactions with the iminium intermediate.[39–44] Moreover, a catalyst for the asymmetric Pictet-Spengler reaction producing the benzylisoquinoline group has not yet been developed.

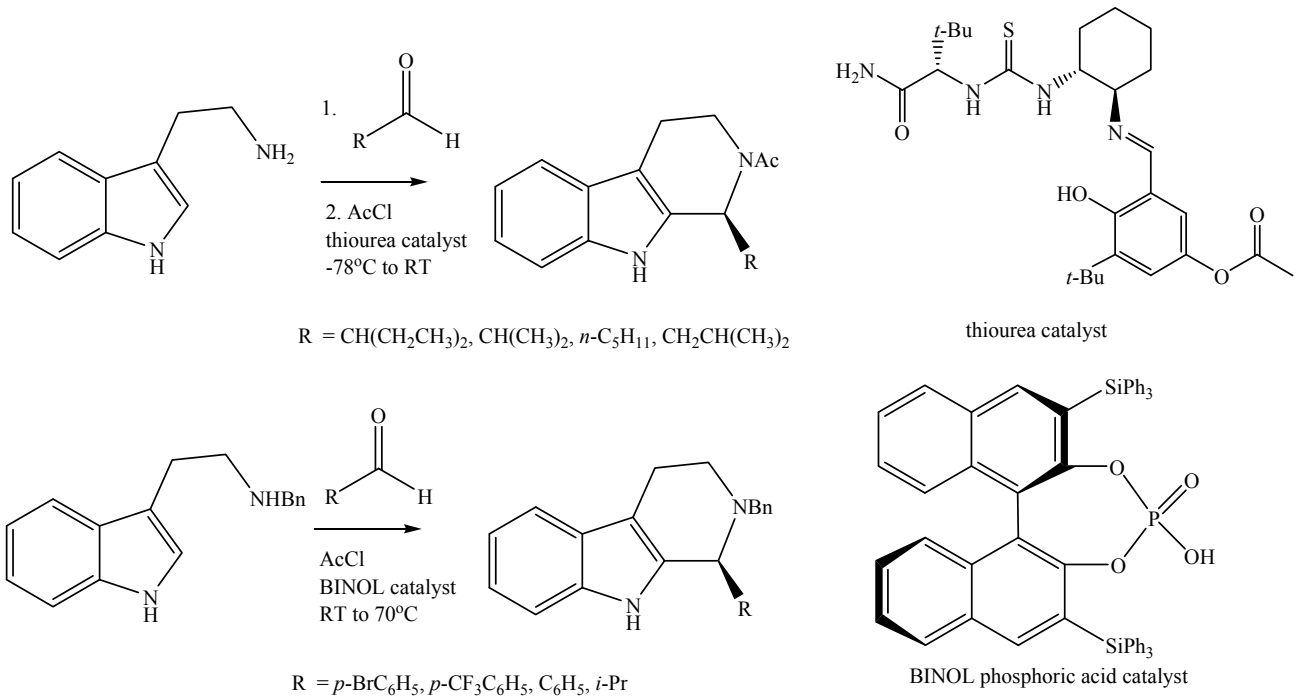


Figure 1.13 Asymmetric Pictet-Spengler reactions

Another approach to search for an efficient asymmetric Pictet-Spengler catalyst is to isolate enzymes in nature that catalyze such a reaction. These enzymes can potentially accept different starting material analogs, thereupon creating a large pool of different β-carboline and tetrahydroisoquinoline building blocks with high stereochemical purities. Thus far, only three of these 'Pictet-Spenglerases', have been identified in alkaloid biosynthetic pathways. These include: strictosidine synthase, norcoclaurine synthase, and deacetylipecoside synthase. Strictosidine synthase catalyzes the condensation of secologanin and tryptamine to form strictosidine,[11] while norcoclaurine synthase brings dopamine and 4-hydroxyphenylacetaldehyde together to form (*S*)-norcoclaurine (Figure 1.14).[45] Interestingly, deacetylipecoside synthase acts on dopamine and secologanin to give a hybrid Pictet-Spengler product, (*R*)-deacetylipecoside.[46] In recent years, the genes encoding for strictosidine synthase and norcoclaurine synthase have been cloned and the enzymes have been biochemically studied.[47–51]

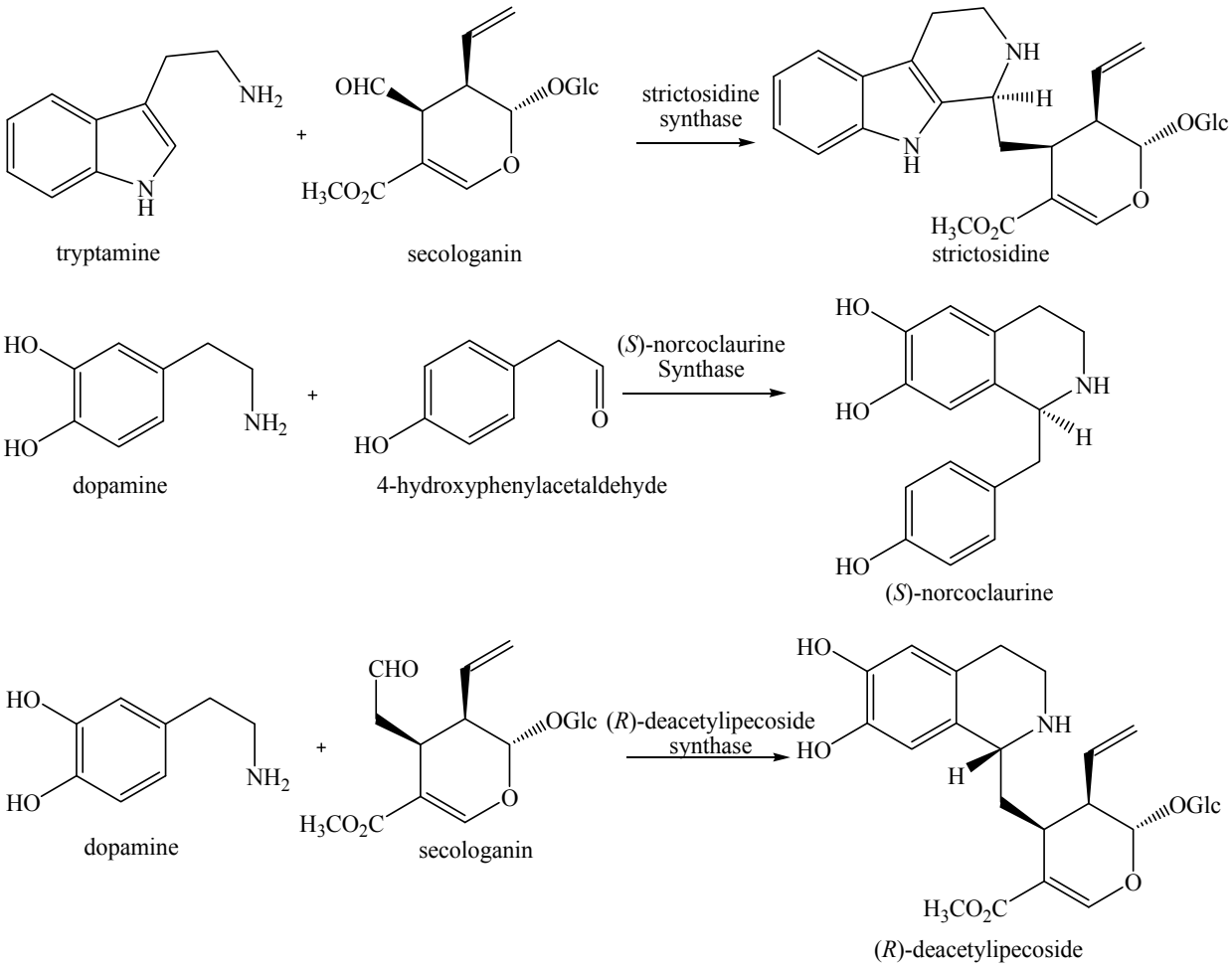


Figure 1.14 Enzymes catalyzing the Pictet-Spengler reaction

1.4.3 Enzyme-catalyzed Pictet-Spengler Reaction: Strictosidine Synthase

Strictosidine synthase is a 'Pictet-Spenglerase' that catalyzes the diastereoselective condensation of tryptamine and secologanin to give strictosidine with a (*S*)-configuration at C-1 (Figure 1.15). Along with its capability to catalyze this unique reaction with high stereochemical purity, strictosidine synthase also serves to produce the β-carboline backbone in the terpene indole alkaloid biosynthetic pathway. These alkaloids belong to a large and structurally diverse group of natural products found in a wide variety of plant species.[52, 53] Perhaps, the most well-known terpene

indole alkaloids are vinblastine and vincristine that have been isolated from *Catharanthus roseus* and have long been established for their anti-tumor activities.[54] Other known examples include yohimbine, macroline and ajmaline, many of them have potent pharmacological effects (see Section 1.4, Figure 1.6). Hence, the study of the strictosidine synthase would contribute greatly to our understanding of terpene indole alkaloid syntheses in Nature.[21, 22, 52, 53]

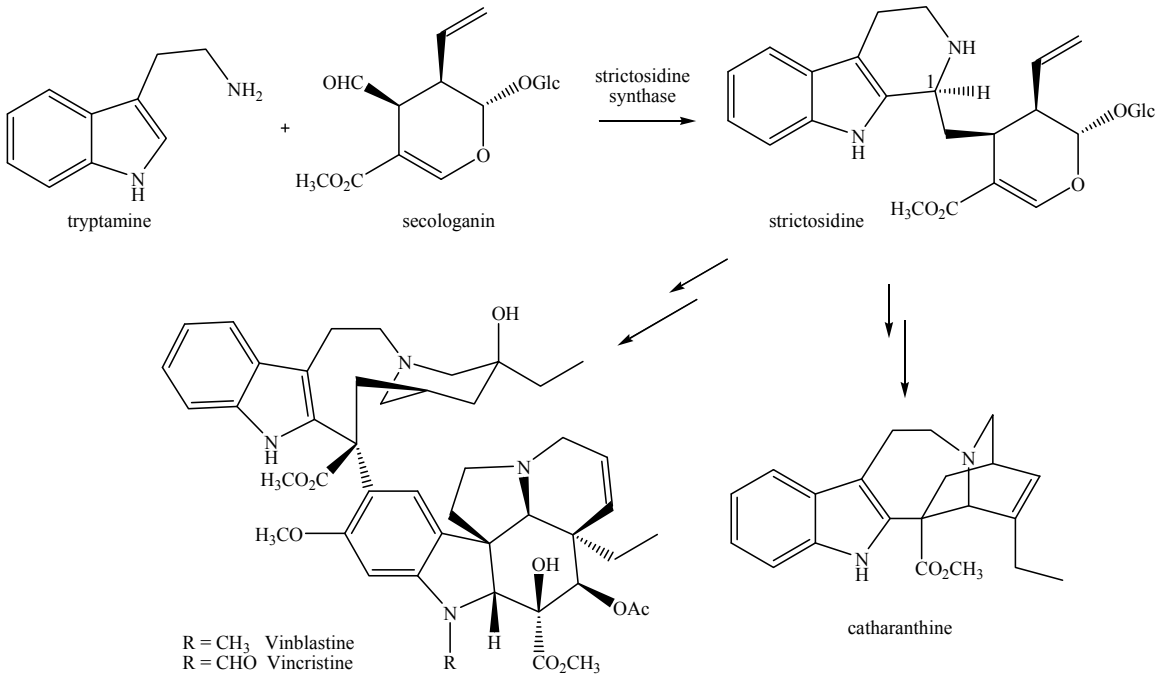


Figure 1.15 Strictosidine synthase and its role in nature for indole alkaloid biosynthesis

Strictosidine synthase, originally isolated from *Catharanthus roseus* and *Rauvolfia serpentina* almost 30 years ago, has been the subject of numerous steady-state kinetic analyses by both the research groups of Scott and Zenk, but mechanistic studies are lacking.[55, 56] The first insight into the nature of substrate binding and the enzymatic mechanism was acquired in 2006, when Stöckigt and his co-workers obtained the crystal structure of *R. serpentina* strictosidine synthase in complex with both secologanin and tryptamine.[48] The enzyme consists of a novel six-bladed β-propeller fold, in which each propeller fold is formed by four strands of antiparallel β-sheet. Only three ionizable residues, His307, Glu309 and Tyr151, are located in the active site (Figure 1.16).

His307 appears to be involved in binding to the glucose moiety of secologanin, and its mutation to alanine resulted in a 130-fold increase in the value of the Michaelis constant K_M (39 to 5070 μM) for secologanin. Glu309 was proposed to play a critical role in catalysis, as it is located closest to the amino group of tryptamine (2.4 Å) and the Glu309Ala mutant showed a 900-fold drop in the value of k_{cat} (1.3 to 0.015 s^{-1}). The kinetic parameters of the Tyr151Phe mutant are similar to the ones of the wild type enzymes. This suggested that the Tyr151 probably does not participate in catalysis or in substrate binding.[48]

Figure 1.16 Topology representation of the *Rauvolfia serpentina* strictosidine synthase. Diagrams obtained from the crystallographic studies on strictosidine synthase.[48, 49]

Two years after the studies performed by Stöckigt, the strictosidine synthase from *C. roseus* was investigated. Two reaction pathways for the catalysis of strictosidine synthase were proposed by O'Connor in 2008 (subsequent to the work described in this thesis), and were based on the preceding mechanistic studies of the non-enzymatic Pictet-Spengler reaction. One pathway involved the spiroindolenine intermediate (pathway A, Figure 1.17), while the other employed a direct alkylation of the iminium functionality (pathway B). As mentioned in the previous section, evidence exists for both mechanisms in solution, and thus far the predominant mechanism remains unidentified. Thus, a real question remains as to the mechanism of the strictosidine synthase reaction.[49]

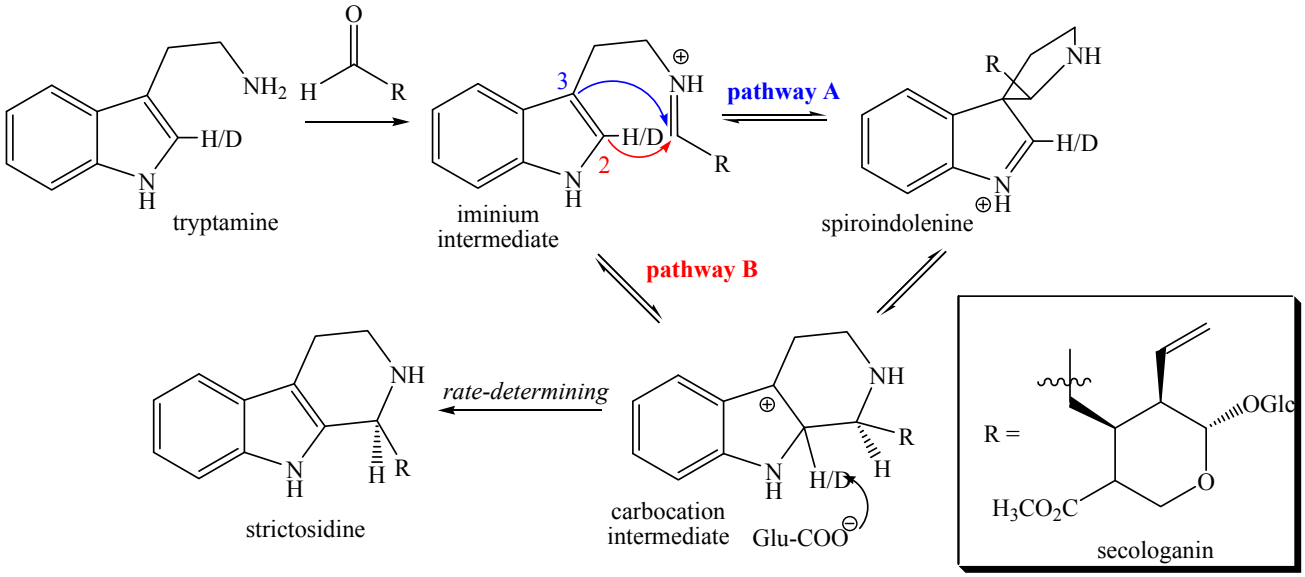


Figure 1.17 Proposed mechanisms for the strictosidine synthase reaction

Several experiments were carried out to analyze the reaction catalyzed by the strictosidine synthase. Initially, the *C. roseus* strictosidine synthase was crystallized with a bound inhibitor, iminium analog **18**. This compound was found to bind in proximity to the same three ionizable residues previously discussed.[48, 49] Unfortunately, the crystal structure showed neither C-2 nor C-3 of the indole ring positioned in close proximity to the carbon atom corresponding to the iminium functionality (Figure 1.18). Thus, it gave no clue as to the mechanism of the enzymatic reaction. In the next experiment, [2-^{2}H]-tryptamine was employed in a kinetic isotope effect (KIE) measurement and showed a primary kinetic isotope effect of 2.70, indicating that deprotonation of the positively charged carbocation intermediate is the rate-determining step of the enzymatic reaction. This observation is somewhat counter-intuitive, as the cyclization step might be expected to be rate-determining when the aromaticity of the indole ring is broken. Additional studies suggested that Glu309 acts as a base during the reaction. It is positioned in close proximity to the C-2 proton of the carbocation intermediate and a significant reduction in the reaction rate at low pH also agreed with the role of Glu309 as a catalytic base. Finally, the reaction rates also decreased at a higher pH, implying that tryptamine binds to the enzyme with the amine in its protonated form. Nonetheless,

neither the X-ray crystal structure nor the kinetic isotope effect experiment could discern between the two potential pathways.[49]

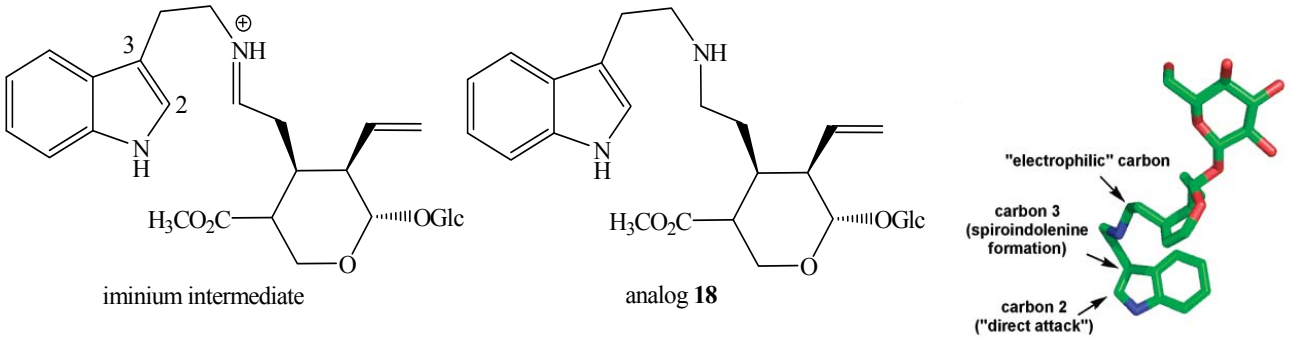


Figure 1.18 Iminium inhibitor analog 18 and a graphic representation of it in the active site of strictosidine synthase. Pictorial diagram obtained from Maresh, *et. al.*[49]

In an attempt to elucidate the reaction mechanism, O'Connor and her co-workers employed *Ab Initio* calculations. In order to simplify the calculations, a smaller aldehyde molecule, ethanal, and a *trans*-iminium intermediate (which is relatively more stable than the *cis*-isomer) were used in the model (Figure 1.19). According to this model, the activation energy required to go from the iminium ion to the carbocation intermediate (~4 kcal/mol) is relatively lower than that for spiroindolenine intermediate formation (~8 kcal/mol). Hence, the 1,2-alkyl shift is likely not a productive step and the authors favor the direct nucleophilic attack mechanism (pathway B, Figure 1.19).[49] Because the model study employed a greatly simplified aldehyde, the indole moiety can react with more than one orientation, and the reaction environment is significantly different from the active site of strictosidine synthase, the ability of the calculation to accurately model the enzymatic reaction is debatable.[49]

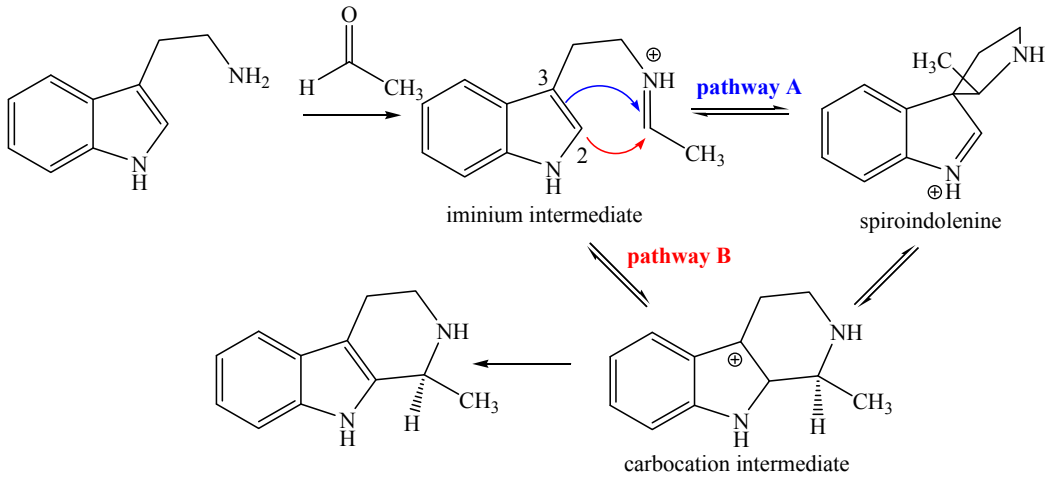


Figure 1.19 The Pictet-Spengler reaction modeled by *Ab Initio* calculations

Another interesting aspect of this study is that the strictosidine synthase Pictet-Spengler reaction can readily proceed non-enzymatically to give a mixture of diastereomers with either an (*R*) or an (*S*)-stereoconfiguration at the C-1 position (Figure 1.20).[49] To investigate the rate-determining step of the non-enzymatic reaction, acetic acid was used as a buffer, since its pK_a (4.6) is similar to that of glutamic acid. Also, because the secologanin is hydrolytically labile in an acidic environment, propanal was used. With [2-^{2}H]-tryptamine, the Pictet-Spengler reaction in solution at pH 4.6 gave a KIE of 2.65, a similar value to that of the enzymatic reaction. Furthermore, the authors observed that at a high pH (>5) where acetate concentration is high, the KIE value dropped to 1, showing that the deprotonation step is no longer rate-determining. From this observation, the authors argued that strictosidine synthase employs the same mechanism as the non-enzymatic Pictet-Spengler reaction.[49]

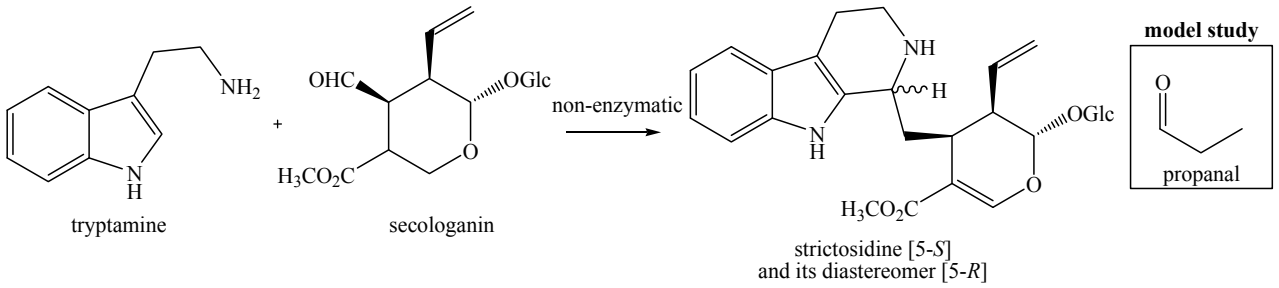


Figure 1.20 Non-enzymatic Pictet-Spengler reaction

As mentioned previously, an enzyme that catalyzes a Pictet-Spengler reaction can be an excellent tool for producing different stereochemically pure β-carbolines.[53, 57] It turns out that strictosidine synthase has a high fidelity for its substrates, and only a few analogs can serve as alternative substrates. Substrates including methylated tryptamine, benzofuran and benzothiophene heterocycles and 4-, 5-, 6-, 7-fluoro-tryptamine were accepted as substrates, whereas tryptophan, phenethylamine and pyrrole derivates were not (Figure 1.21). Also, the only secologanin analogs that are accepted by the strictosidine synthase were those with the methyl group of the ester functionality replaced with an ethyl or an allyl group.[53, 57]

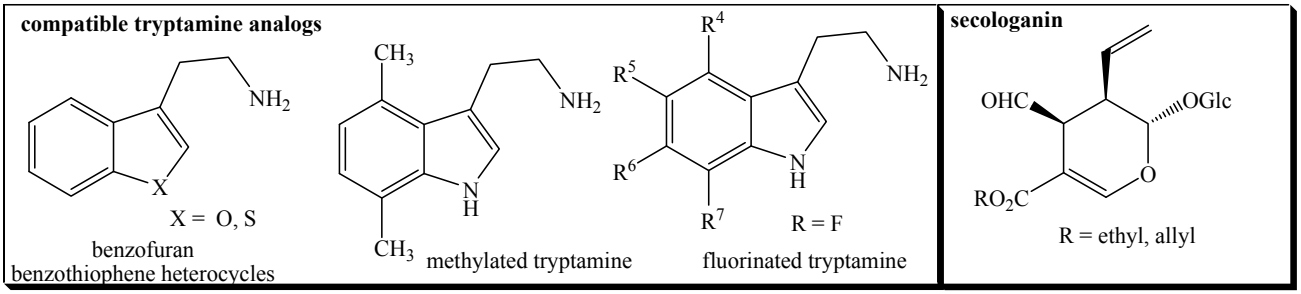


Figure 1.21 Substrate analogs compatible with strictosidine synthase

1.4.4 Enzyme-catalyzed Pictet-Spengler Reaction: Norcoclaurine Synthase

Norcoclaurine synthase is another asymmetric 'Pictet Spenglerase' in Nature that catalyzes the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) to give (*S*)-

norcoclaurine (Figure 1.22). The product sets up the characteristic backbone for the biosynthesis of benzylisoquinoline alkaloids, which comprise a group of more than 2500 plant secondary metabolites found in several related plant families.[47] This group of compounds has been widely investigated, as they possess divergent structural scaffolds and exhibit a range of pharmacological properties.[1, 17] Both dopamine and 4-hydroxyphenylacetaldehyde are derived from the chemical modifications of tyrosine via the reactions of several enzymes,[58] but norcoclaurine synthase is considered to catalyze the first committed step in the biosynthesis of benzylisoquinoline alkaloids.[18, 45, 59] The (*S*) configuration is required for subsequent biosynthetic steps and is retained in several alkaloids of this pathway, including the key branch-point intermediate (*S*)-reticuline. However, this configuration is often inverted later in biosynthetic steps to compounds like the analgesic morphine and the muscle relaxant tubocurarine, which bear an (*R*) configuration at this position. In addition, the configuration of this carbon is lost in some benzylisoquinoline alkaloids, such as the fungicide berberine, as it is oxidized to an sp^2 hybridized center.[1, 17]

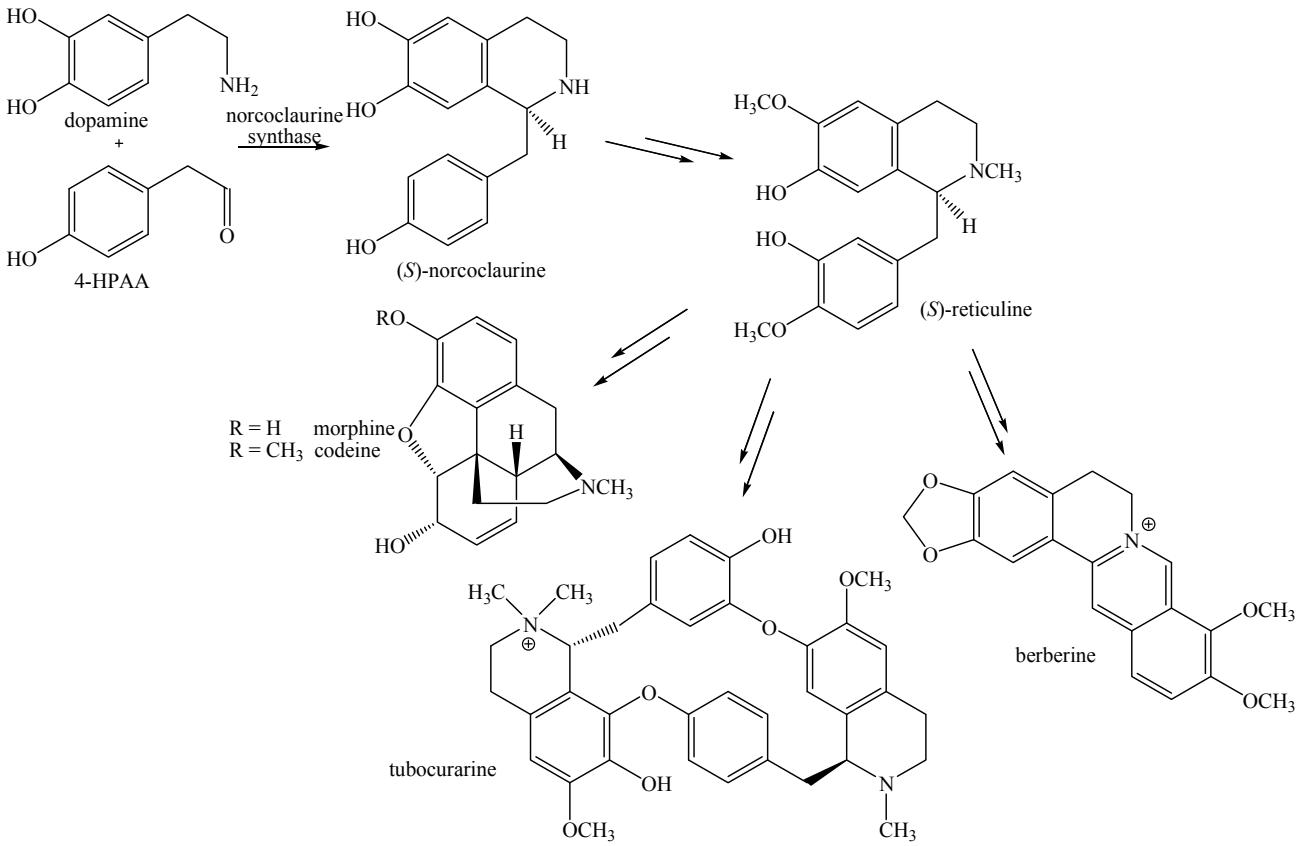


Figure 1.22 Norcoclaurine synthase and its participation in benzylisoquinoline alkaloid biosynthesis

The activity of norcoclaurine synthase was first detected in a variety of plant extracts such as *Eschscholtzia* sp. in 1981 by Zenk.[49, 50,57] It was originally characterized as a norlaudanosoline synthase based on its ability to use 3,4-dihydroxyphenylacetaldehyde as an alternate substrate (Figure 1.23).[50] However, Zenk and his co-workers later discovered that the mono-hydroxylated 4-hydroxyphenylacetaldehyde was the precursor of morphine in the plant *Papaver somniferum*. Accordingly, they re-named the enzyme as norcoclaurine synthase (NCS).[18, 45, 59, 60] Despite this mistake, some initial biochemical characterizations were achieved using 3,4-dihydroxyphenylacetaldehyde (DHPAA) as a substrate. First, the pH rate profile showed a maximum at around pH 7.8. Secondly, radioactively labeled [3,5-^{3}H]-dopamine was employed to measure enzyme kinetics and showed hyperbolic kinetics with K_M values of 1.5 and 0.7 mM for dopamine and DHPAA, respectively (Figure 1.23).[61] In another experiment, the enzymatic reaction

product was peracetylated and its circular dichroism spectrum was measured. Interestingly, the peracetylated product was only 20-25% optically pure as compared to the authentic samples of the sterochemically pure (*S*)- or (*R*)-peracetylated norlaudanosoline. The authors concluded that the (*R*)-norlaudanosoline impurity was a by-product from the non-enzymatic Pictet-Spengler reaction.[50]

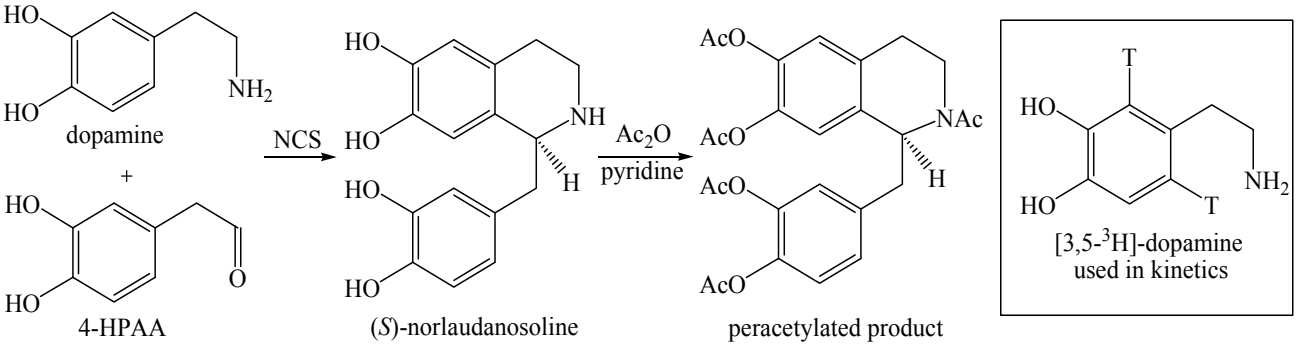


Figure 1.23 Norlaudanosoline production by norcoclaurine synthase

In 2001, Facchini and his research group also purified norcoclaurine synthase from the opium poppy *Papaver somniferum*. They discovered that this enzyme exhibits hyperbolic saturation kinetics with 4-HPAA (K_M = 1.0 mM), but that unlike the results from the Zenk group, it shows sigmoidal kinetics for dopamine with a Hill coefficient of 1.84.[51, 62] It was not until 2004 that the gene encoding for the truncated form of the enzyme from *Thalictrum flavum* ssp. glaucum (that lacks a putative signal peptide) was cloned and expressed in *E. coli* (the NCS gene that was previously cloned from *P. somniferum* did not yield an active enzyme). Notably, there is no significant sequence homology between strictosidine synthase and norcoclaurine synthase, suggesting that they probably originated from different ancestors as a consequence of convergent evolution.[48] While the recombinant enzyme was found to be active in the cell lysate, the purification of the norcoclaurine synthase was not achieved and only limited studies were performed on this enzyme.[47]

Two general mechanistic possibilities may be considered for the norcoclaurine synthase reaction (Figure 1.24).[63] In each case, the first step involves the formation of an iminium ion

intermediate that may either be enzyme-catalyzed or spontaneously occur in solution. The most direct path to product involves an ionization of the phenolic group at the C-2 position of the catechol moiety to generate the zwitterionic intermediate **19** that cyclizes to give the σ-intermediate (pathway A). A subsequent deprotonation causes re-aromatization and product formation. A second possibility involves ionization of the phenolic group at C-1 of the catechol moiety to generate the zwitterionic intermediate **20** that cyclizes to form a spirocyclic intermediate (pathway B). A subsequent ionization of the C-2 hydroxyl group and protonation of the ketone induces a semi-pinacol rearrangement to give the σ-intermediate that proceeds to product as described. While the second mechanism involves an additional step and a potentially unfavorable 5-endo-trig cyclization, it should not be discounted as several lines of evidence support an analogous mechanism for some non-enzymatic Pictet-Spengler reactions.[33–35] While this mechanism has only been postulated for the indole-based reactions in the literature, it is possible that catechol-based cyclizations could also proceed in this manner.[63]

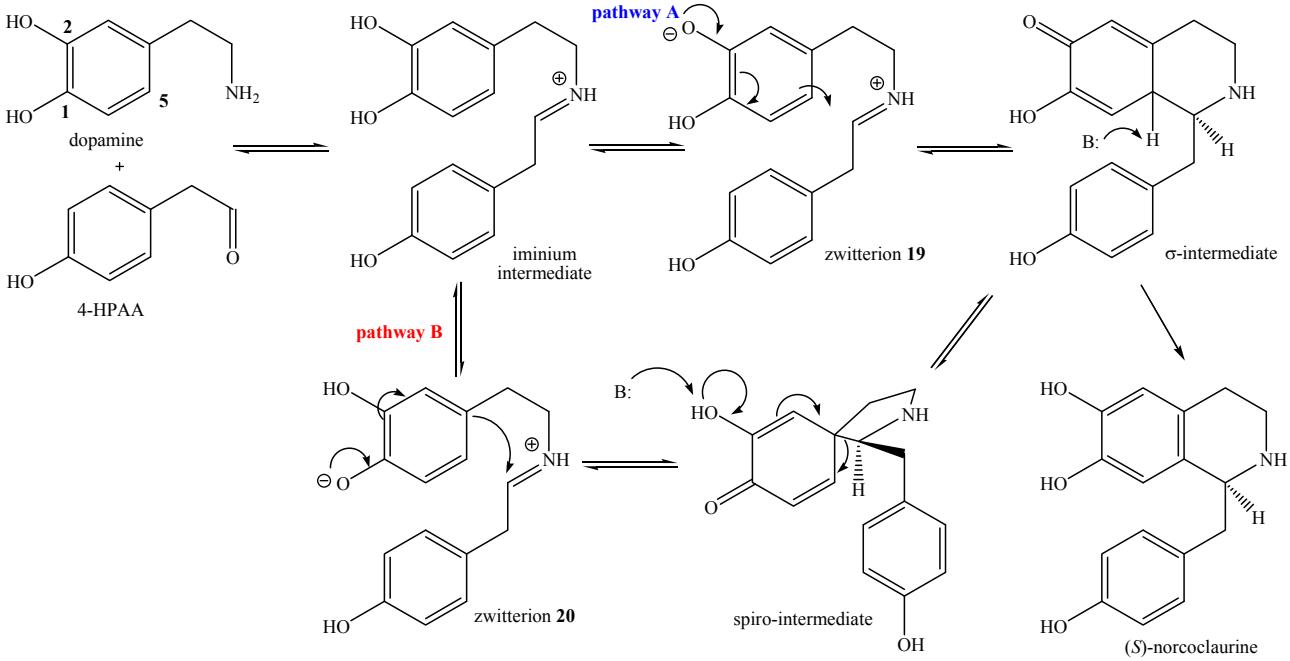


Figure 1.24 Two mechanistic possibilities for the norcoclaurine synthase reaction

1.5 Thesis Aims

The overall aim of this thesis is to probe the Friedel-Crafts type reaction mechanisms by which norcoclaurine synthase and 4-dimethylallyltryptophan synthase might operate (Figure 1.4). Although both enzymes are involved in alkaloid biosyntheses, the enzymatic reactions are notably distinct from each other and warranted separate analyses. Accordingly, a detailed background on dimethylallyltryptophan synthase is presented in Chapter 3, followed by the mechanistic characterization of this enzyme in Chapter 4. In order to provide the appropriate context, goals related to this specific project have been deferred until the end of Chapter 3.

The unique reaction catalyzed by the norcoclaurine synthase makes it an attractive target for investigation. Furthermore, the role of this enzyme in the biosynthesis of benzylisoquinoline alkaloid adds additional intrigue. Therefore, Chapter 2 is dedicated to addressing the mechanism and the substrate specificity of this enzyme. First, the biochemical characterization of the protein including its overexpression, and purification are discussed. The reaction products are analyzed by NMR spectroscopy and mass spectrometry, and a kinetic analysis of the enzymatic reaction is presented. The two potential mechanistic pathways are probed by the use of several substrate analogs. In an attempt to gain insights into the mechanism employed by the norcoclaurine synthase, an isotopically labeled substrate was made and subjected to kinetic isotope effect measurements. Finally, the mechanistic similarities and differences between norcoclaurine synthase and strictosidine synthase are discussed.

Chapter 2

Mechanistic Studies on Norcoclaurine Synthase (NCS)

A version of this Chapter has been published and some figures are reproduced with permission from: Luk, L. Y. P.; Bunn, S.; Liscombe, D. K.; Facchini, P. J.; Tanner, M. E. *Biochemistry* **2007**, *46*, 10153-10161.

As discussed in Chapter 1, norcoclaurine synthase (NCS) catalyzes an asymmetric Pictet-Spengler condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) to give (*S*)-norcoclaurine. This is the first committed step in the biosynthesis of the benzylisoquinoline alkaloids that include morphine and codeine. Furthermore, norcoclaurine synthase is one of the few 'Pictet-Spenglerases' found in Nature. Hence, mechanistic investigations on this enzyme would be of interest.

The initial work by Facchini and co-workers in cloning the *Thalictrum flavum* norcoclaurine synthase showed that the crude enzyme is active in the bacterial lysate, but attempts to optimize the expression and to purify the enzyme were not reported. In this Chapter, a new approach for NCS expression, product characterization via the use of NMR spectroscopy and mass spectrometry, and the development of a continuous assay based on circular dichroism spectroscopy are presented. Also, a set of dopamine analogs and isotopically labeled dopamine are employed in probing the mechanism of the NCS reaction.

These mechanistic studies of NCS were the first reported for any 'Pictet-Spenglerase'. Soon after this work was published, the crystallographic studies of NCS and the biochemical characterizations of strictosidine synthase were reported as well. Hence, a detailed discussion and comparison among all these studies will be presented at the end of this Chapter.

2.1 Enzymatic Activity of Norcoclaurine Synthase

2.1.1 Cloning and Expression of Norcoclaurine Synthase

In the past work on norcoclaurine synthase (NCS), the full length cDNA encoding for the enzyme was cloned and sequenced from a cell culture of *Thalictrum flavum* by Professor Facchini

and his research group.[47] Expression of this gene in ER2566 pLysS *E. coli* did not result in any detectable level of recombinant protein. However, when the NCS gene was truncated to remove the first 10 or 19 amino acids at the N-terminal end of the corresponding protein (pNCSΔ10 and pNCSΔ19, respectively), a small amount of active enzyme could be generated. Kinetic studies of the crude truncated enzymes in bacterial lysate revealed that the Michaelis constants (K_M) and the Hill coefficients were similar to those of NCS isolated from natural sources, indicating that the truncations did not greatly affect its catalytic properties. Unfortunately, the level of NCS expression was extremely low and prohibited its isolation. With an eye for solving this problem, we established a collaboration with the research group of Professor Facchini.[47]

In order to optimize the expression level and permit isolation of the recombinant protein, the gene sequence of NCS was first inspected. A fundamental element of basic biology states that a codon is composed of three nucleotides and usually codes for an amino acid. Since four nucleotides (A, T, G, C) are available in the genetic code, there are 64 (4^3) possible codons, 61 of which specify the 20 amino acids used during protein synthesis.[65] Accordingly, most of the amino acids (except methionine and tryptophan) are coded by a degenerate set of two to six synonymous codons. During the process of translation, the codon in mRNA is recognized by the corresponding anticodon in a tRNA molecule, which carries the correct amino acid at the 3' end. As a consequence of evolution, genes within a taxonomic group exhibit similarities in codon choice regardless of the gene function.[65] The codons for arginine (AGA, AGG), isoleucine (AUA), leucine (CUA), glycine (GGA, GGU) and proline (CCC) are considered 'rare' in *E. coli* genes, because the corresponding aminoacyl-tRNA genes that carry the matching anti-codon base pairs are expressed at a minimal level. In contrast, these rare *E. coli* codons are abundant in most plant systems, as the suitable aminoacyl-tRNA genes are expressed at an optimal level.[66, 67] These taxonomy-specific codon biases have been referred as a 'codon dialect'.[67] In particular, the NCS gene contains 15 of these

rare *E. coli* codons out of a total of 211 amino acids (Figure 2.1). Thus, a regular *E. coli* 'codon dialect' would likely experience difficulty in translating the gene that is originated from the plant species *T. flavum*.[64] Consequently, if there are extra copies of tRNA genes available to cope with these rare codon translations, NCS protein expression should theoretically be greatly enhanced.

5' – ATG ATG AAG ATG GAA GTT GTA TTT GTT TTC TTA ATG TTG TTA GGA ACA ATA AAT
TGC CAG AAA CTG ATT CTG ACA GGT AGG CCG TTT CTG CAC CAC CAG GGC ATA ATA AAC CAG
GTG TCT ACA GTC ACA AAA GTG ATT CAT CAT GAG TTG GAA GTT GCT GCT TCA GCT GAT GAT
ATA TGG ACT GTT TAT AGC TGG CCT GGC TTG GCC AAG CAT CTT CCT GAC TTG CTC CCT GGC
GCT TTT GAA AAG CTA GAA ATC ATT GGT GAT GGA GGT GTT GGT ACC ATC CTA GAC ATG ACA
TTT GTA CCA GGT GAA TTT CCT CAT GAA TAC AAG GAG AAG TTT ATA TTA GTC GAT AAT GAG
CAT CGT TTA AAG AAG GTG CAA ATG ATT GAG GGA GGT TAT CTG GAC TTG GGA GTA ACA TAC
TAC ATG GAC ACA ATC CAT GTT GTT CCA ACT GGT AAA GAT TCA TGT GTT ATT AAA TCC TCA
ACT GAG TAC CAT GTG AAA CCT GAG TTT GTC AAA ATC GTT GAA CCA CTT ATC ACC ACC GGT
CCA TTA GCT GCC ATG GCA GAC GCC ATC TCA AAA CTT GTT CTA GAA CAC AAA TCC AAA AGC
AAC TCA GAT GAA ATT GAG GCC GCA ATA ATA ACA GTC TGA–3'

Figure 2.1 Gene sequence of *Thalictrum flavum* norcoclaurine synthase. 15 rare *E. coli* codons, arginine (AGG), isoleucine (ATA), isoleucine (CTA) and glycine (GGA), were identified.[64]

To verify this hypothesis and analyze the reaction catalyzed by NCS, Dr. Peter Facchini sent us the plasmid pNCSΔ19,[47] which contains the N-terminally truncated *ncs* gene that had been cloned into a pET29b system from Novagen.[68] Several features in this vector are helpful to the overexpression protocol. These include a T7 promoter region that allowed for the induction of overexpression upon addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG), a region coding for a C-terminal hexahistidine tag (His-tag) that was used for purification of the overexpressed protein by Ni^{2+} affinity chromatography, and a gene coding for kanamycin resistance which selected for the bacteria transformed with the plasmid in antibiotic-containing media.[68] Extensive attempts were made to optimize the expression levels with pNCSΔ19. A notable increase in protein production

was observed when the plasmid was transformed into a specialized *E. coli* strain BL21 (DE3) RIL. This strain bears an additional plasmid encoded with extra copies of the tRNA genes, *argU*, *ileY*, and *leuW*, that can compensate for most of the codon biases during heterologous expression.[69] Further increases were observed when an enriched media, Terrific Broth, was used in the growth of the starter cultures and when the bacteria were grown at 25 ˚C following IPTG induction at an optical density (OD_{600}) of 0.5 to 0.6.

The purification of the recombinant synthase was carried out using Ni^{2+} affinity chromatography. The harvested *E. coli* cells were lysed by the use of a French pressure cell. The crude lysate was then loaded onto an immobilized metal affinity column that had been charged with a solution of $NiSO_4$. Protein that was bound non-specifically to the column through interactions other than a His-tag was eluted with 125 mM imidazole buffer. Recombinant NCS was eluted with 500 mM imidazole buffer and concentrated by ultrafiltration. Ultimately it was possible to obtain the recombinant synthase at levels of 15 mg/L under optimal conditions. In this fashion, protein could be isolated with >90% purity as estimated by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis (Figure 2.2). This protein has a molecular weight of about 28 kDa, was quite stable, and could be stored at -78 ˚C for ~6 months without significant loss of activity.

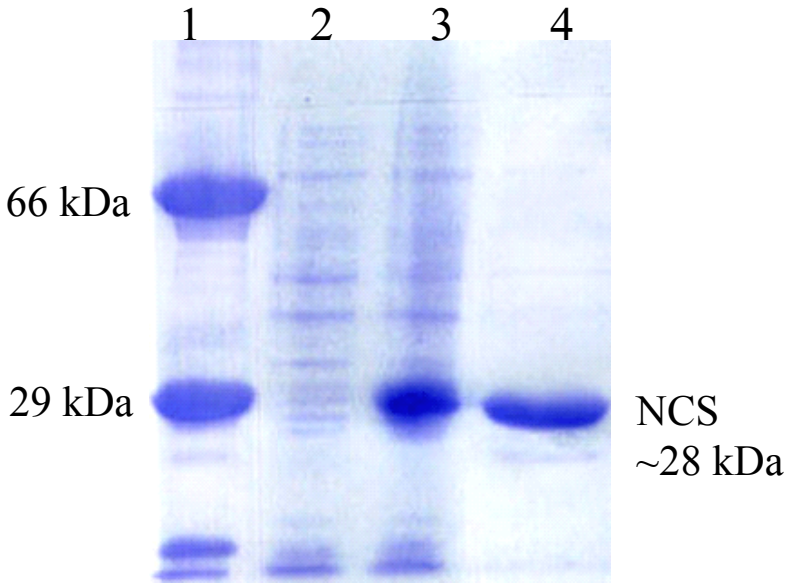


Figure 2.2 SDS-PAGE gel showing the purification of norcoclaurine synthase (TfNCSΔ19). Lane 1 contains molecular mass standards of 66 kDa (bovin serum album) and 29 kDa (carbonic anhydrase), lane 2 shows crude cell extract before induction, lane 3 shows crude cell extract after isopropyl-β-D-1-thiogalactopyranoside induction, and lane 4 shows the purified norcoclaurine synthase.

2.1.2 Norcoclaurine Synthase Activity

Because the recombinant His-tagged enzyme had never been purified previously, the activity of NCS was first tested using a ^{1}H NMR-based assay with dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) in a deuterated buffer. After a 30 min incubation with NCS, it was observed that the starting materials were converted into (*S*)-norcoclaurine. As depicted in Figure 2.3, the signal corresponding to 1-H proton of (*S*)-norcoclaurine appeared as a multiplet at δ 4.54 ppm. Also, the aromatic proton signals for both substrates disappeared as the reaction progressed. These included the δ 7.14 ppm doublet belonging to the 2-/6-H of 4-HPAA, the δ 6.70 ppm doublet of doublets belonging to the 5-H of dopamine, and other aromatic protons for both substrates overlapping at around δ 6.84 ppm. In turn, these signals were replaced by two doublets at δ 7.10 and 6.83 ppm corresponding to the 2'-/6'-H and 3'-/5'-H of the phenolic moiety, and two singlets at δ 6.69 and 6.61 ppm corresponding to the 5-H and 8-H of the catecholic moiety in (*S*)-

norcoclaurine, respectively. Furthermore, the positive ESI-MS spectrum of the crude reaction showed a peak at 272 m/z ($M + H^+$), which corresponded to the molecular weight (MW) of protonated (*S*)-norcoclaurine. Lastly, TLC of the crude reaction using a 4:1:5 1-butanol:acetic acid:methanol solvent system showed a UV-active spot at $R_f = 0.6$, which corresponded to (*S*)-norcoclaurine as previously reported.[47] Thus, a method for high level overexpression of the NCS gene and purification of active recombinant enzyme was developed for the first time.

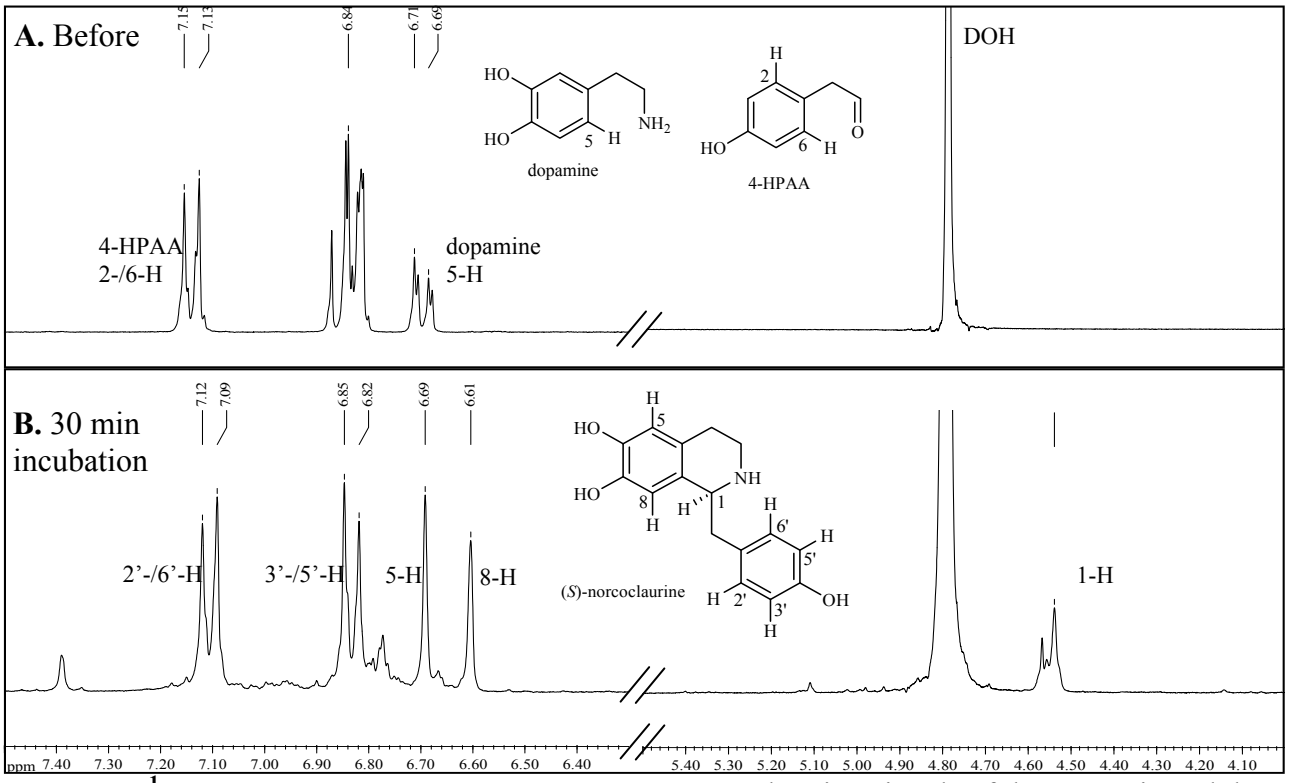


Figure 2.3 ^{1}H NMR assay of the NCS reaction. Spectra showing signals of the aromatic and the 1-H signal of (*S*)-norcoclaurine: A) spectrum refers to the solution before NCS was added, and B) spectrum refers to the mixture after 30 min incubation with NCS.

2.1.3 Isolation and Characterization of (*S*)-Norcoclaurine

Several attempts were made to purify the (*S*)-norcoclaurine generated in the NCS reaction. A purification protocol was thought to be important, because the enzyme could be useful as an

asymmetric Pictet-Spengler reaction catalyst and provide valuable chiral compounds. The starting materials, dopamine and 4-HPAA (50 mM each), were incubated with NCS (0.22 mg) in potassium phosphate buffer for 30 min at 37 ^{o}C. The crude product was subjected to silica gel chromatography in acidic solvent 50:50:1 acetone/methanol/acetic acid, and the resulting norcoclaurine appeared pure as analyzed by ^{1}H NMR and ESI-MS. However, the material did not show any signal in the CD spectrometer, implying that the compound had racemized during the chromatography. To assess the optical activity of (*S*)-norcoclaurine, a CD spectrum of the enzymatic reaction was taken before chromatography and showed maximal signal at 285 nm. The loss of stereochemical purity probably occurred as a result of exposing (*S*)-norcoclaurine to acidic solvent and/or silica gel, which led to C-N sigma bond dissociation through a path similar to the one in β-carboline roeharmine racemization (Figure 2.4&1.12, see Chapter 1).[20, 36, 37] Indeed, the cationic intermediate **21** of the proposed racemization pathway is stabilized by the electron-donating nature of the hydroxyl groups. Modifications of the purification protocol were made in an attempt to avoid the racemization. In one attempt, acetic acid was not included in the silica chromatography. However, this was not successful as norcoclaurine streaked in the column and could not be eluted. Neutral alumina was used instead of the weakly acidic silica gel in one purification protocol, yet it still gave racemized product.

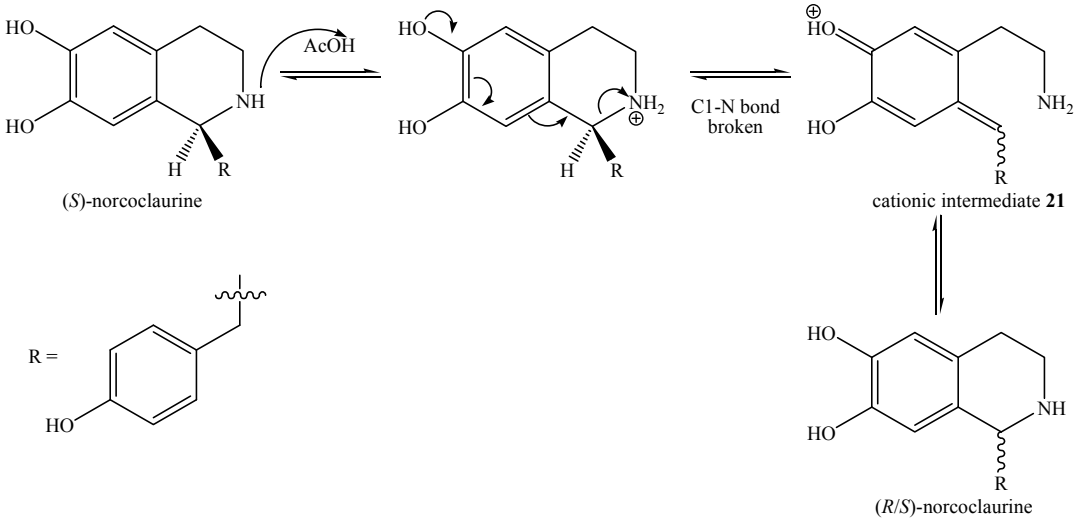


Figure 2.4 Possible racemization reaction pathway for (*S*)-norcoclaurine during purification

In order to obtain (*S*)-norcoclaurine without racemization, a crude purification method was developed. The enzyme and phosphate buffer were removed by ultrafiltration and precipitation by ice-cold ethanol, respectively, and crude (*S*)-norcoclaurine was obtained by removal of solvent. The CD spectrum of the semi-purified crude illustrated the same molar ellipticity as before, thereby suggesting that (*S*)-norcoclaurine prepared in this fashion had not yet lost its optical activity.

2.2 Characterization of Norcoclaurine Synthase

2.2.1 Kinetic Characterization of NCS by a Continuous Assay

Previously, two stopped radioactive assays using [2,6-^{3}H]-dopamine and [8-^{14}C]-dopamine were developed separately to study the kinetics of the NCS-catalyzed reaction. The former method monitored the amount of tritium released into the solution, while the latter method required the use of a TLC separation.[47, 50] A continuous assay that does not require radiolabeled substrate or a separation step could greatly facilitate studies on this enzyme. Moreover, it would allow one to measure the kinetics of the reaction with dopamine analogs that are not available in an isotopically labeled form. Since the substrates dopamine and 4-HPAA are achiral whereas the product (*S*)-norcoclaurine is chiral, circular dichroism (CD) spectroscopy presents an attractive method for following product formation.

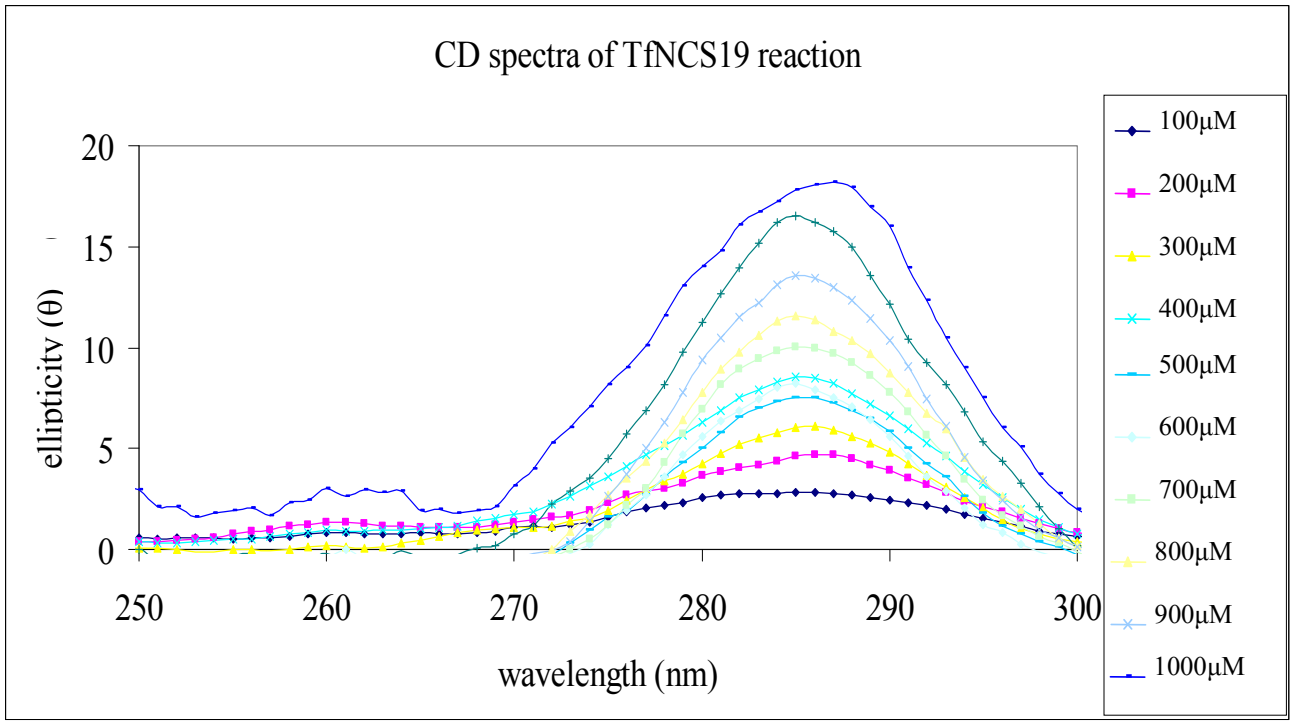


Figure 2.5 Circular dichroism spectra of (*S*)-norcoclaurine formed in the NCS reaction using various concentrations of substrates

The observed CD signal at a particular wavelength, or more precisely the ellipticity θ_λ reported in degrees, is dictated by the molar ellipticity $[\theta]_\lambda$ with units of mdeg cm^{-1} M^{-1}. The relationship between the measured ellipticity and the molar ellipticity is given by:

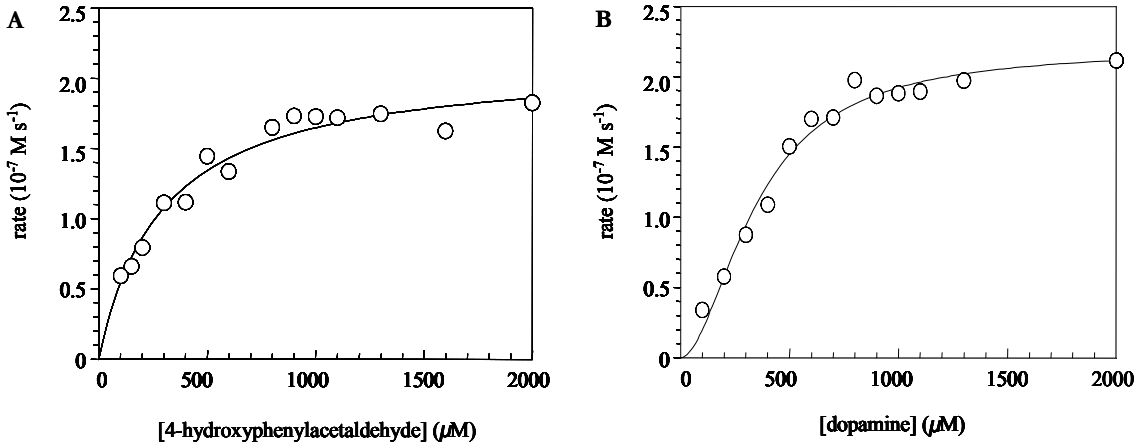
$$[\theta]_\lambda = \frac{\theta_\lambda}{l \cdot C} \quad \text{equation (2.1)}$$

where l is the path length in cm and C is the concentration in mol L^{-1}. The CD spectra of the (*S*)-norcoclaurine formed in the enzymatic reaction with various equimolar concentrations of substrates (100 to 1000 μM, Figure 2.5) were obtained and found to exhibit a maximum at 285 nm with a molar ellipticity of 12 500 mdeg cm^{-1} M^{-1}. This analysis assumes that the NCS reaction proceeds to completion, as is corroborated by ^{1}H NMR spectroscopic studies.

In past studies using NCS that was obtained from natural sources, all kinetic data had to be corrected for a low background reaction that can occur non-enzymatically.[51, 61] In this study, the

background reaction is no longer a problem as the relative abundance of recombinant enzyme permitted a much more rapid kinetic analysis (5 min) during which time the extent of non-enzymatic reaction is negligible. Furthermore, CD spectroscopy measures the difference between the absorptions of left-handed circularly polarized light and right-handed circularly polarized light that arise due to structural asymmetry. Hence, only the enzymatic product will give a signal in the spectrum, as any norcoclaurine produced in a non-enzymatic fashion will be racemic.

The kinetic parameters were evaluated in a pH 7.5 potassium phosphate buffer at 37 ^{o}C. At saturating levels of both dopamine and 4-HPAA (2000 μM), the kinetic constant k_{cat} was determined to be 6.4 ± 0.3 s^{-1} (Figure 2.6). When the concentration of dopamine was held at a saturating level (2000 μM) and the concentration of 4-HPAA was varied, a hyperbolic kinetic profile was observed that led to a value of K_M = 288 ± 38 μM. On the other hand, when the concentration 4-HPAA was held constant (2000 μM) and the concentration of dopamine was varied, a sigmoidal kinetic profile was observed indicative of positive cooperativity. Fitting this data to the Hill equation gave values of K' = 350 ± 48 μM and n = 1.8 ± 0.2. Before interpreting the data, the theory underlying allostery and cooperativity are elaborated in the following section.



Varying [substrate]	K_M (μM)	k_{cat} (s^{-1})	Hill coefficient
Dopamine	350 ± 58*	6.6 ± 0.5*	1.8 ± 0.2*
4-HPAA	288 ± 40		1

Figure 2.6 Enzyme kinetics for norcoclaurine synthase
* The kinetic constants were extrapolated from the corresponding Lineweaver-Burk or Hill plot, and the standard errors were obtained by a linear least squares analysis.
* In each experiment, each concentration point was carried out in quadruplicate and the measured initial velocities never differed by more than 3%.
* Every experiment was repeated 3 times and the resulting kinetic constants never differed by more than 20%.

Regulatory enzymes and the Hill equation

Many enzymes are multimeric and possess multiple substrate-binding sites, and they can be categorized as either non-regulatory or regulatory enzymes. In a non-regulatory enzyme, the catalysis follows Michaelis-Menten kinetics and the plot of reaction rate (v) versus substrate concentration ([S]) gives a rectangular hyperbolic line (or a Lineweaver-Burk plot of $1/v$ versus $1/[S]$ gives a linear plot). If the enzyme is multimeric, this type of plot suggests that substrate binding and catalysis at one site has no effect on other subunits. In contrast, many regulatory enzymes do not follow Michaelis-Menten kinetics, but rather give a sigmoidal curve in the plot of v versus [S]. This observation can be explained as a result of cooperativity, in which the substrate (or product) behaves as an allosteric modulator and regulates binding to the subsequent substrate molecule. This type of

enzyme is often classified as ‘controlled’ or ‘regulatory’ and is located at the beginning or at the branch point of a metabolic pathway. When the binding of one substrate molecule increases the affinities of other vacant binding sites towards additional substrate molecules, it is defined as positive cooperativity. In contrast, negativity cooperativity refers to the situation where the substrate affinity decreases as the substrate concentration increases.

The homo-tetrameric hemogloblin was the first recorded example of a protein that shows (a positive) cooperativity towards ligand (oxygen) binding.[70] It was postulated that hemoglobin and other enzymes with cooperativity have at least two conformations and each conformation has a different degree of substrate affinity. To explain the observed sigmoidal trend in the oxygen dissociation curve of hemoglobin, the symmetric Monod-Wyman-Changeux (MWC) and the asymmetric (or induced fit) Koshland-Némethy-Filmer (KNF) models have been proposed.[70]

In the MWC model, the enzyme can exist in two conformational states: the predominant T (tense) and the minor R (relaxed) states.[70] While the T state has less affinity for the substrate than the R state, these two states are in equilibrium and differ in the energies and numbers of bonds between subunits. Moreover, the transition between one conformation (T) and the other (R) is an all-or-none event (i.e. all subunits in one enzyme must be in the same state), thus this model has only two dissociation constants for substrate binding. In other words, the enzyme is always ‘symmetric’ as the subunits are always in the same state. At low substrate concentration, enzymes are mostly in the inactive T state, resulting in slow initial reaction rates. However, as the substrate concentration increases, more enzyme molecules will transform into the R state thereby enhancing the overall affinity to the substrate. Hence, this explains the observation of a sharp increase at the midpoint of a sigmoidal rate curve.

In the alternative KNF model, the all-or-none event was not assumed, and subunits within an enzyme can be in different conformational states (T or R). Indeed, substrate binding induces a

conformational change from the T to R state in the subunit to which the substrate is bound. Subsequently, this conformational change affects other vacant subunits within the enzyme, and thus changes the substrate affinity of the subsequent enzyme subunit. Since the dissociation constants for each consecutive substrate binding event are different due to the absence of symmetry in the KNF model, the mathematical system is much more complicated than the one developed in the MWC model. Nonetheless, because only the KNF model can account for enzymes with negative cooperativity, this model is generally considered to provide a better illustration for most regulatory enzyme systems.[70]

In order to investigate the sigmoidicity in enzyme kinetics, the empirical Hill equation was employed. The Hill equation can be used to correlate catalysis to substrate binding in a regulatory enzyme as follows:

$$v = \frac{V_{max}\,[S]^{\mathbf{n}}}{(K' + [S]^{\mathbf{n}})} \qquad \text{equation (2.2)}$$

where v is the reaction rate, V_{max} is the maximal reaction rate, **n** is the Hill coefficient and K' is a constant comprising interaction factors and intrinsic dissociation constants (i.e. an apparent overall binding constant). K' is often called 'apparent K_M', but it should be noted that K' does not equal to the concentration of substrate at half of the V_{max} (except when **n** = 1 in non-regulatory enzymes). In order to determine **n**, the Hill equation can be rearranged into a logarithm form:

$$\log \frac{v}{(V_{max} - v)} = \mathbf{n} \log [S] - \log K' \qquad \text{equation (2.3)}$$

A plot of the function $\log [v/(V_{max} - v)]$ versus $\log [S]$ gives a straight line with a slope corresponding to **n**. The observed **n** value refers to the Hill coefficient and is almost always less than the theoretical **n** value, which should equal to the minimum available binding sites in a multimeric regulatory enzyme. The observed Hill coefficient is often lower than this value because complete cooperativity is rare.

When the Hill coefficient **n** = 1, the substrate binding to each site is independent of one another, the enzyme is non-cooperative and Michaelis-Menten kinetics are followed. When **n** is larger than 1, the subunits within the enzyme are positively cooperative. The classic example hemoglobin shows a value of **n** ~ 2.8. In contrast, when **n** is less than 1, the enzyme is considered to be negatively cooperative, suggesting that the substrate affinity decreases after the substrate binds to the first subunit.

A continuous kinetic assay for the reaction catalyzed by NCS was developed with the use of CD spectroscopy. The observation of a hyperbolic kinetic profile with varying levels of 4-HPAA and a sigmoidal kinetic profile with varying concentrations of dopamine is in agreement with previous studies using crude plant extracts or purified exogenous enzyme from several sources.[51] The value of the Hill coefficient (**n**) in the kinetics of varying dopamine concentrations was found to be 1.8, indicating that there is positive cooperativity between two dopamine binding sites within a NCS enzyme molecule. The report that this enzyme exists as a homodimer in solution helps to explain how this cooperativity can occur as binding to one subunit could affect binding to the other subunit.[51]

This assay was only useful at substrate concentrations higher than 100 μM or lower than 1000 μM, because of low signal-to-noise ratios. This led to errors in the values of K_M that were typically ±15%. Nevertheless, the value of K_M determined for 4-HPAA agreed reasonably well with those previously reported for the wild type enzyme (335 μM) and the crude recombinant enzyme (700 μM).[47, 51] The sigmoidal dependence of rate versus dopamine concentration was also observed in previous studies and the Hill coefficients (**n**) of 1.8 and 1.98 were reported for the wild type and recombinant enzymes, respectively. However, the *K'* value for dopamine in the wild type and the recombinant enzymes (40 μM and 2.5 μM, respectively) are significantly lower than that obtained in

this study. This may be due to the narrow range of dopamine concentrations used in the previous kinetic assays (55-280 μM for the wild type enzyme and 20-70 μM for the recombinant enzyme) that would almost certainly lead to errors.[47, 51] In addition, differences in pH values and buffers among these assays could account for some of the discrepancy.

2.2.2 Alternate Substrate Testing and Mechanistic Analysis

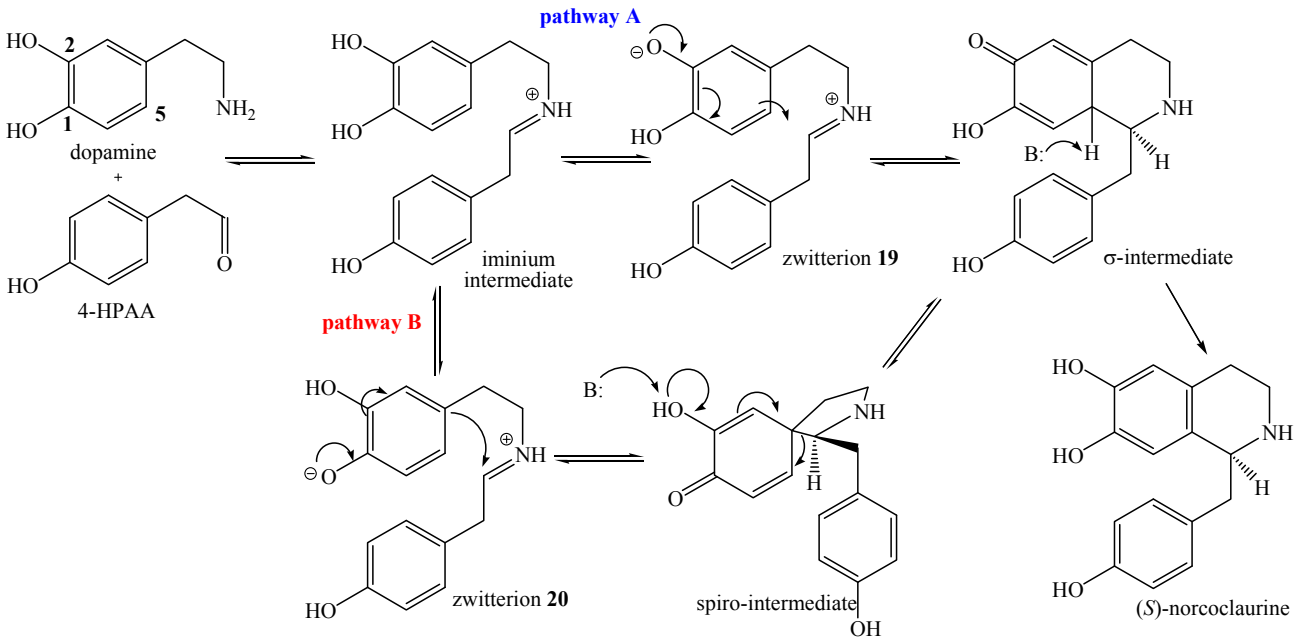


Figure 2.7 Two mechanistic possibilities for the NCS Reaction

Several substrate analogs were employed to probe the mechanism and the substrate specificity of the norcoclaurine synthase reaction. As shown in Figure 2.7, deprotonation of the C-2 phenol plays a key step in catalysis in both proposed mechanisms. The function of the C-2 hydroxyl group was examined by testing dopamine analogs either lacking such functionality (**22**, tyramine) or bearing a methoxy group at the C-2 position (**23**) as alternative substrates (Figure 2.8). The C-1 hydroxyl group of dopamine is crucial in the first step of the spirocyclic intermediate mechanism, as its deprotonation promotes C-5 alkylation (pathway B, Figure 2.7). Therefore, dopamine analogs

lacking the C-1 hydroxyl or bearing a methoxy group at the C-1 position (**24** and **25**, respectively) were also analyzed. The reactions were first monitored by ^{1}H NMR spectroscopy and mass spectrometry to determine whether these compounds would be accepted by the enzyme. It turned out that compounds **24** and **25** could serve as alternate substrates, and therefore they were subjected to kinetic analysis using CD spectrometry. The kinetic constants obtained for these compounds were essentially identical to those obtained with dopamine indicating that they were excellent substrate analogs (Table 2.1). On the other hand, compounds **22** and **23** did not serve as substitutes for dopamine, and no condensation products could be observed upon prolonged incubations. The fact that compound **22** is unreactive was not surprising, since it does not contain an electron-donating substituent at the C-2 position of dopamine that would be required for either reaction mechanism. Compound **23**, in contrast, carries a C-2 oxygen in the form of an ether, and this type of compound is known to be a good substrate for acid-catalyzed non-enzymatic Pictet-Spengler reactions.[71] A possible explanation is that the formation of the phenolate ion at C-2 is required during catalysis. This would greatly increase the nucleophilicity of the catechol ring. Alternatively, steric interactions with the methyl group of the ether may have prevented the analog from binding to the enzyme. Nevertheless, the observation that compounds **24** and **25** exhibit wild type-like kinetic constants in the NCS reaction rules out the mechanism involving a spirocylic intermediate (pathway B, Figure 2.7, Table 2.1), because the phenolic functionality at C-1 would be required for the initial spirocyclization.

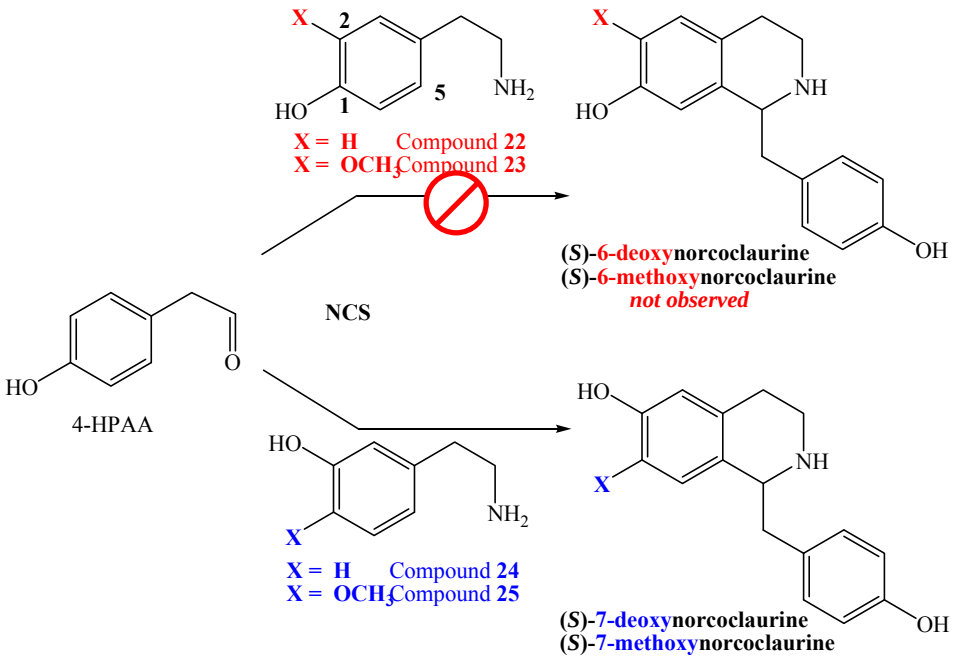
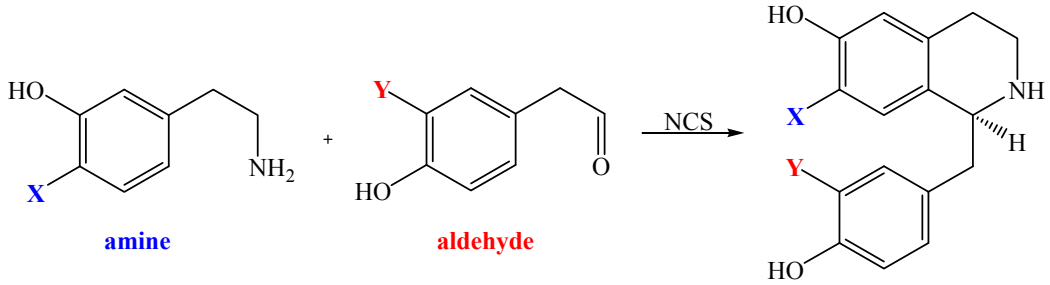


Figure 2.8 Alternative substrate testing in the NCS reaction

Table 2.1 Kinetic parameters for substrate analogs in NCS-catalyzed reactions



Amine	Aldehyde (Y)	k_{cat} (s^{-1})	*K*' (μM)	Hill coefficient (n)	K_M (μM)
X = OH (**dopamine**)	**Y** = H (**4-HPAA**)	6.6 ± 0.5*	350 ± 48*	1.8 ± 0.2*	288 ± 38*
X = H (**24**)	**Y** = H (**4-HPAA**)	6.0 ± 0.5	407 ± 54	1.9 ± 0.5	N/A
X = OMe (**25**)	**Y** = H (**4-HPAA**)	6.0 ± 0.6	303 ± 54	1.9 ± 0.5	N/A
X = OH (**dopamine**)	**Y** = OH (**26**)	7 ± 2	N/A	N/A	335 ± 221

* The kinetic constants were extrapolated from the corresponding Lineweaver-Burk or Hill plot, and the standard errors were obtained by a linear least squares analysis.
* In each experiment, each concentration point was carried out in quadruplicate and the measured initial velocities never differed by more than 3%.
* Every experiment was repeated 3 times and the resulting kinetic constants never differed by more than 20% with the exception of DHPAA (**26**), which had large error due to weak molar ellipticity $[\theta]_\lambda$.

In addition, an aldehyde analog, 3,4-dihydroxyphenylacetaldehyde (DHPAA, **26**), was investigated as an alternate substrate. This compound, interestingly, was first thought to be the natural substrate of the enzyme, but later it was shown that the additional hydroxyl group of the catechol was added in subsequent biosynthetic steps.[18, 45, 50, 59] It is also of synthetic interest as many benzylisoquinoline alkaloids are hydroxylated at that position. As expected from previous studies on crude plant extracts, this aldehyde analog is an excellent substrate for the synthase and displays kinetic constants quite similar to those of 4-HPAA.

2.3 Solvent Isotope Incorporation

In an attempt to investigate how NCS may facilitate the electrophilic addition to the aromatic ring of dopamine, non-reactive substrate analogs were employed to analyze for enzyme-catalyzed solvent-derived isotope incorporation. Since the enzyme activates the C-5 position to act as a nucleophile (presumably by deprotonating the C-2 hydroxyl group), it may also catalyze deuterium incorporation into the C-5 position of dopamine or unreactive analogs. The first deuterium incorporation experiment was accomplished by incubating dopamine (~0.1 M) with NCS (6.4 mg) in deuterated potassium phosphate buffer at pD 7.5 and monitoring with ^{1}H NMR spectroscopy and ESI-MS methodology. At the beginning of the incubation, the ^{1}H NMR spectrum showed a doublet of doublets at δ 6.69 ppm which corresponded to the 5-H of dopamine, and integrated to one proton (Figure 2.9, left panel). After 24 h incubation with NCS the proton signal remained, and its integration had not changed. Furthermore, the peak of the unlabeled dopamine at 154 m/z ($M + H^+$ in methanol) in the ESI-MS spectrum appeared as the major peak before and after the incubation. The absence of deuterium incorporation could be due to the inability of dopamine to bind the enzyme active site alone. To address this hypothesis, dopamine and the 4-HPAA analog tyrosol (**27**)

were co-incubated in another deuterium incorporation experiment under the same conditions. However, as indicated by the ^{1}H NMR and ESI-MS spectra, deuterium was not incorporated into the C-5 position of dopamine (Figure 2.9, right panel).

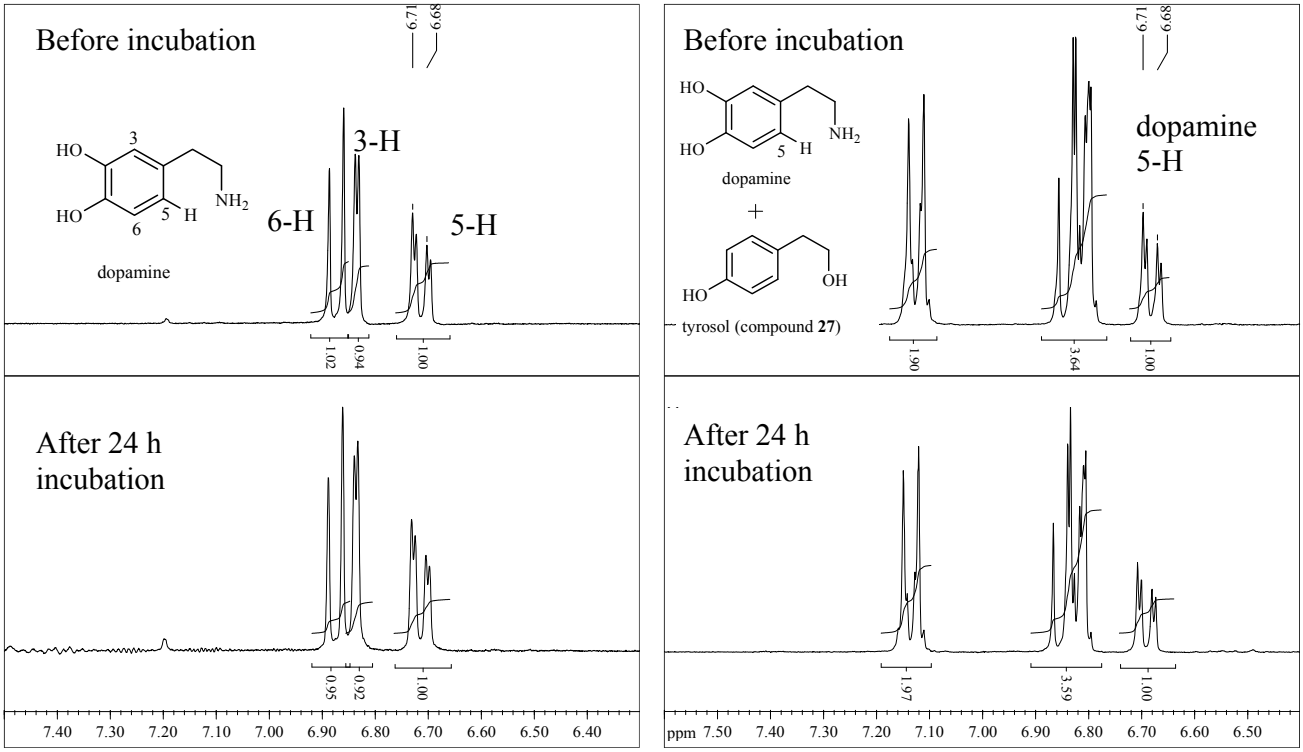


Figure 2.9 ^{1}H NMR spectra of deuterium incorporation experiment with dopamine (left panel**) and dopamine + tyrosol (27) (**right panel**)**

.

The reduced iminium intermediate analog **28**, was also synthesized and subjected to the deuterium incorporation experiment. Again, deuterium was not incorporated into the C-5 of the iminium intermediate analog **28**. In the ^{1}H NMR spectrum, the H-5 signal at δ 6.67 ppm and its integration did not change after 24 h of incubation (Figure 2.10). Also, the peak of the protonated analog **28** at 274 m/z ($M + H^{+}$) remained as the major peak in the ESI-MS spectra before and after the incubation.

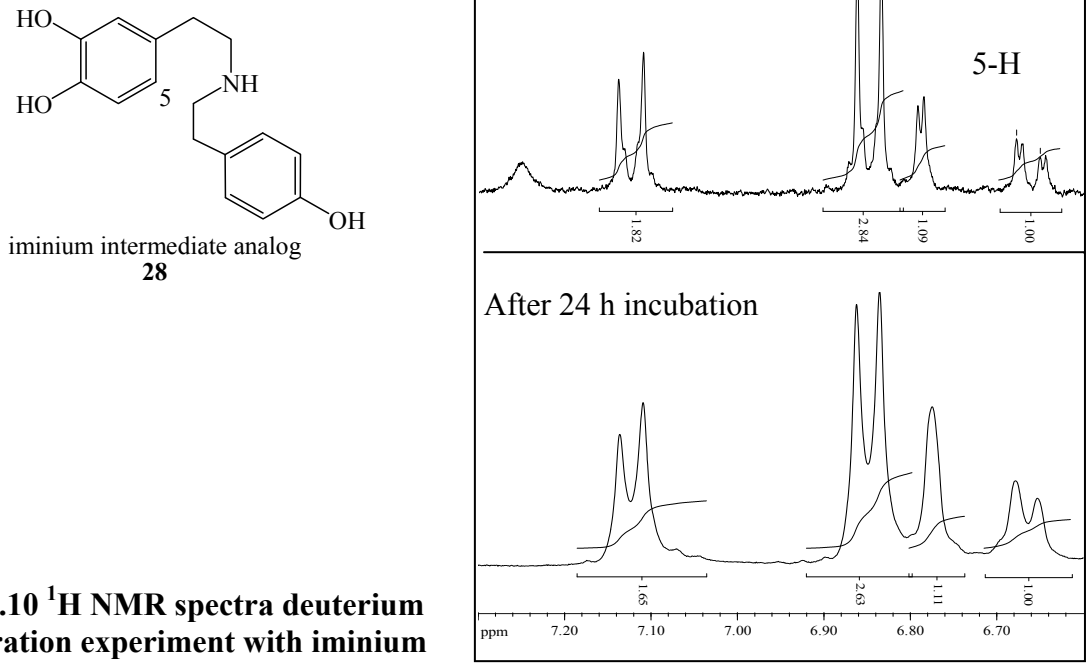


Figure 2.10 ^{1}H NMR spectra deuterium incorporation experiment with iminium intermediate analog (28)

Several possible scenarios may account for the absence of deuterium incorporation. First, it may simply reflect the absence of a suitably positioned acidic residue in the active site that could promote deuterium exchange. However, there should be an acid/base residue in close proximity to C-5 that normally deprotonates the σ-intermediate to form (*S*)-norcoclaurine. Another scenario is that neither dopamine alone nor the substrate analogs bind to the active site of NCS in a productive manner, thereby prohibiting deuterium incorporation into these molecules. Perhaps the most likely explanation is that in the NCS active site, protonation of dopamine (and/or the analog **28**) is enantiospecific and the subsequent deprotonation occurs on the same face. For instance, once the 2-OH group of dopamine (and/or the analog **28**) is deprotonated, protonation may occur on the *pro-S* face of C-5 to form a transient [5-^{2}H]-(*S*) intermediate (Figure 2.11). Subsequently, the same active site base may remove the deuterium and yield the unlabeled starting material. In fact, this explanation was supported by the data obtained in the crystallographic study of NCS. The acidic

residue is thought to be Glu110, and it positioned near the *pro-S* face of the C-5 of dopamine (see Discussion of this Chapter).

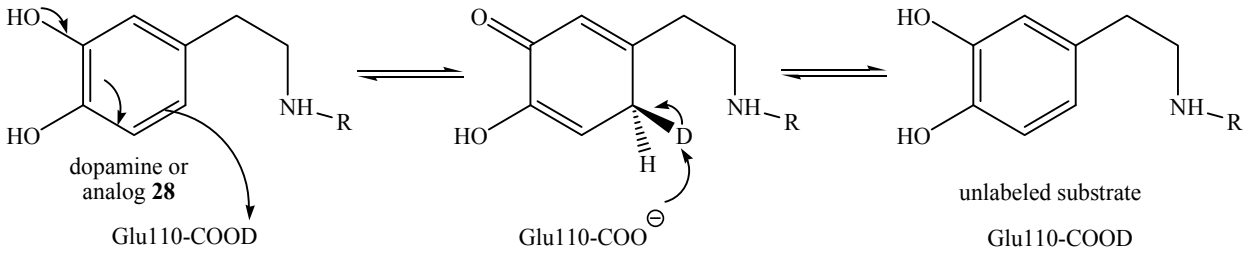


Figure 2.11 A possible explanation for the lack of deuterium incorporation

2.4 Kinetic Isotope Effect Measurement

2.4.1 Introduction and Hypothesis

A possible method to gain mechanistic insights into a reaction is to compare the kinetic parameters between an unlabeled substrate and an isotopically labeled substrate when bonding changes at the site of isotopic substitution are involved in the rate-determining step.[72] This phenomenon is collectively known as kinetic isotope effects (KIEs) and occurs as a result of the difference in bond strengths between unlabeled and isotopically labeled bonds. A KIE involving hydrogen and deuterium is generally represented as k_H/k_D, where k_H and k_D are the reaction rate constants for the corresponding isotopologues.[72] In general, KIEs can be classified as either primary or secondary, and the theory will be briefly described in the following sections.

A primary deuterium KIE is observed when the C-H bond that is broken in the rate-determining step is replaced by a C-D bond.[72, 73] The increase in the reduced mass of the two atoms due to deuterium substitution gives the C-D bond a lower vibrational zero-point energy (Figure 2.12).

Thus, a higher activation energy is needed to remove the deuterium and reach the transition state which is presumed to be of similar energy for both isotopologues. A typical C-H stretching vibration absorbs at 3000 cm^{-1} in IR spectroscopy and gives a zero-point energy of 4.3 kcal mol^{-1}.[72, 74] Deuterium substitution reduces the infrared spectral absorption to 2000 cm^{-1} and the zero-point energy to 3.2 kcal mol^{-1}. The difference between the two stretching modes is ΔE_o = 1.1 kcal mol^{-1}, and the application of the Arrhenius equation provides the theoretical limit for an 'intrinsic' primary KIE as $k_H/k_D = \exp(\Delta E_o/RT) = 6.4$. The intrinsic isotope effect refers to the full effect originated from a single isotopically sensitive step, and it is exclusive of all interference from isotopically insensitive steps. As the C-H bond vibration ranges from 2800 cm^{-1} to 3300 cm^{-1}, the model predicts that the primary KIE values will range from 5 to 8.[72, 74] Nonetheless, the model does not account for the wide variation in the magnitude of primary isotope effects in different types of reactions. Specifically, because enzymatic reactions are often composed of multiple reversible steps, the rate-determining step of the reaction could be masked by other isotopically sensitive step(s), resulting in a lower KIE value. Primary KIEs for cleavage of a C-H bond in enzymatic reactions are typically on the order of 2-4. Large isotope effects can also be observed as a result of quantum mechanical tunneling. In such cases, the wavelength of the proton is comparable to the width of the activation energy barrier and molecules with insufficient energy to overcome the energy barrier may 'tunnel' through it leading to isotope effects of 20 and above.[72, 73]

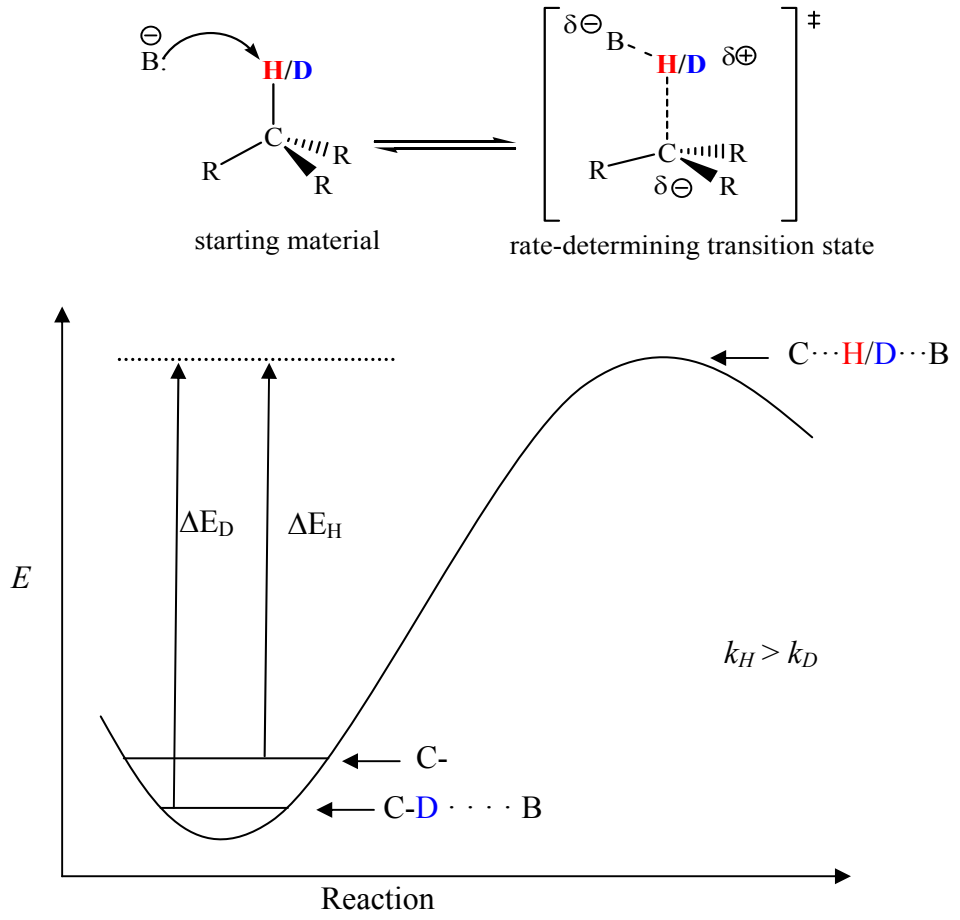


Figure 2.12 Reaction coordinate diagram illustrating the origin of primary kinetic isotope effects. ΔE_H, ΔE_D = activation energy required to break a C-H or C-D bond bearing hydrogen or deuterium, respectively. Zero-point and transition state energy levels are indicated by horizontal lines. B = enzymatic base, R = substituents and k_H and k_D = reaction rates of the corresponding isotopologues.

Secondary KIEs can be observed when a deuterium is substituted at a position where cleavage of the deuterium bond takes place.[74–76] This type of KIE can be further classified as α, β, or remote (γ and δ), depending on the position of the isotopic label with respect to the reactive center.[76] An α-KIE refers to the situation where deuterium is attached directly to a carbon that is undergoing a bonding change. Although not directly involved in the reaction, the isotope can still influence the internal vibrational system, which in turn affects the zero point energy and the observed reaction rate. The changes in internal vibrational energy can result from a change in orbital hybridization at the

reactive carbon. C-H bonds involving either sp^3 or sp^2 orbitals possess a number of vibrational modes, including stretches, in-plane bending and out-of-plane bending. The common belief is that changes in these vibrational modes during a reaction are primarily responsible for the secondary KIEs. In an sp^3 hybridized carbon, the in-plane and out-of-plane bending are similar in frequency (~1380 cm^{-1}) due to the symmetry of the tetrahedral arrangement. However, this is not the case in a trigonal planar sp^2 hybridized carbon; the out-of-plane bending in an IR spectrum appears at only ~800 cm^{-1}, whereas the in-plane bending at ~1340 cm^{-1}.[74, 76] Hence, when the hybridization of the reactive carbon changes from sp^3 to sp^2 in the rate-determining step, the force constant of the C-H bond will be weakened (lower bending frequency) in the transition state. Upon deuterium substitution, a normal secondary α-deuterium effect of around 1.02 to 1.40 will be observed (case 1, Figure 2.13). By the same token, when the hybridization of the reactive carbon changes from sp^2 to sp^3 orbital, the bending of the C-H bond becomes less energetically favorable (higher bending frequency) in the transition state. Hence, deuterium substitution will cause an inverse secondary KIE ranging from 0.8 to 0.9 (case 2, Figure 2.13).[74, 76]

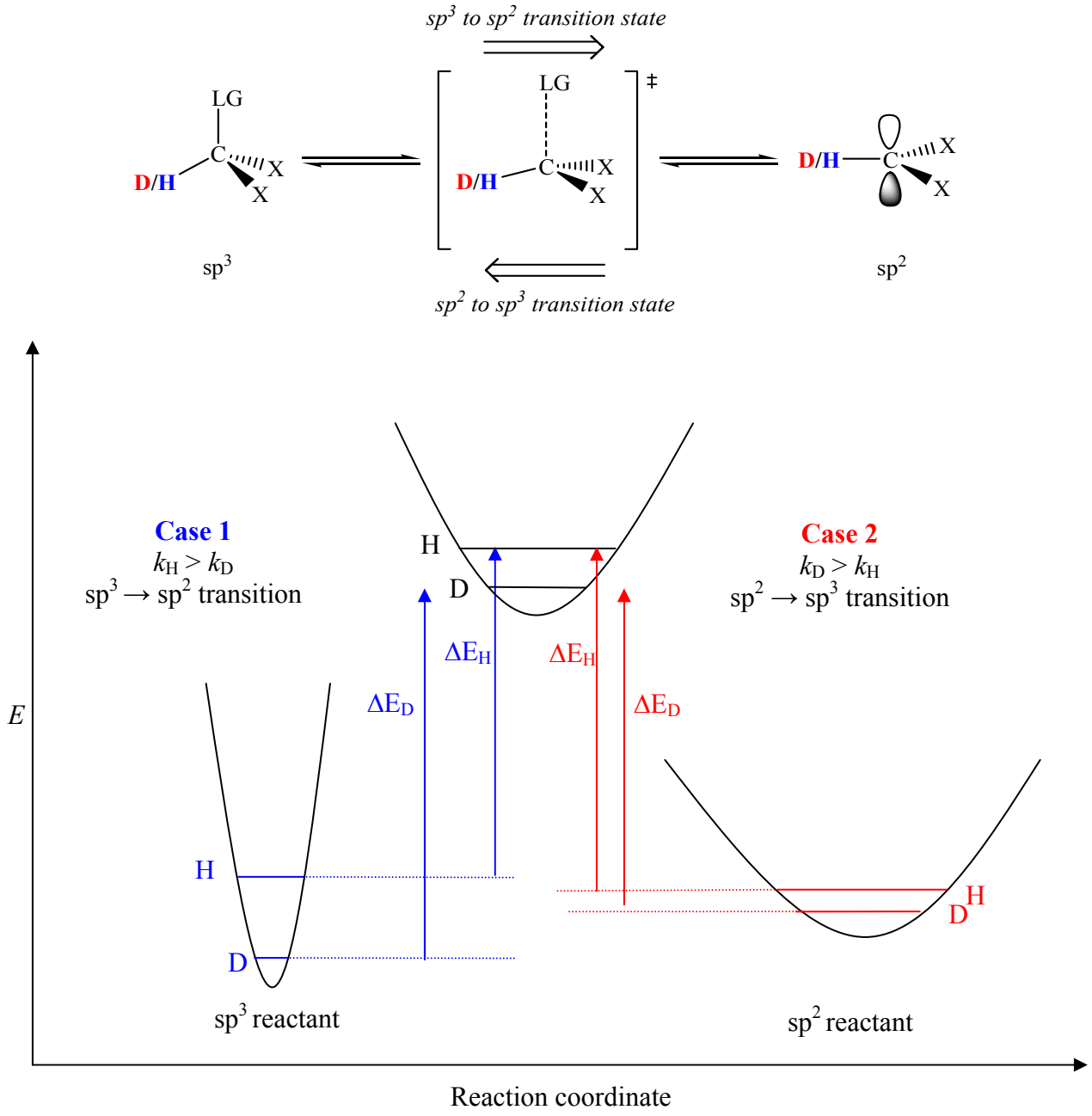


Figure 2.13 Reaction coordinate diagram illustrating the origin of secondary α-deuterium kinetic isotope effects. ΔE_H, ΔE_D = activation energy required to change the hybridization of the reactive carbon, respectively. Zero-point energies of reaction and transition state levels are indicated. LG = leaving group and k_H and k_D = reaction rates of the corresponding isotopologues.

2.4.2 Kinetic Isotope Effect on k_{cat}/K_M

In order to probe the rate-determining step of the NCS reaction, the KIE measurement was carried out by comparing the rate of reaction using [3,5,6-$^2H_3$]-dopamine (**29**) and unlabeled dopamine. The tri-labeled dopamine was chosen because it is readily synthesized compared to the singly labeled [5-^{2}H]-dopamine, and the assumption was made that only the deuterium at the site of alkylation (C-5) will affect the reaction rate. Several possible outcomes can be anticipated from this study. First, if deprotonation of the σ-intermediate is rate-determining, a primary KIE (2 to 8) will be observed. In turn, an inverse secondary KIE (< 1) would imply that the formation of σ-intermediate is the rate-determining step of the catalysis, because the carbon being alkylated experiences a hybridization change from sp^2 to sp^3 orbital. Lastly, if the observed KIE is close to unity, it would imply other steps such as iminium formation, or the binding of substrates and/or the release of product are rate-determining. Because formation of the σ-intermediate disrupts the aromaticity of the iminium intermediate, this step was originally expected to be rate-determining and an inverse secondary KIE was anticipated. Hence, the KIE measurement on k_{cat} was not performed, as it requires a direct comparison method that employs circular dichroism spectroscopy and the inherent errors could possibly mask small secondary isotope effects. Instead, a competitive method was employed that gives the KIE on the value of k_{cat}/K_M.

An intermolecular competitive method was employed to measure the KIE, symbolized as k_H/k_D, on the second-order rate constant k_{cat}/K_M (For details regarding the KIE on k_{cat} and k_{cat}/K_M, please see appendix). A mixture of the two dopamine isotopologues of known isotopic composition is treated with the enzyme. After a certain fractional conversion to products has occurred, the isotopic composition of the remaining starting material is determined. If there is a KIE, the faster

reacting isotopologue will be depleted to a greater extent than the slower reacting isotopologue. By accurate measurement of the reaction progress and the initial and final isotopic ratios of the unlabeled and isotopically labeled substrates, the k_H/k_D can be extrapolated using the equation (2.4) derived by Melander and Saunders:[74]

$$k_H/k_D = \frac{\ln(1-F_H)}{\ln[(1-F_H)R/R_o]} \qquad \text{equation (2.4)}$$

where F_H refers to the fractional conversion of the protiated species to products, and R_0 and R are the initial and final ratios of the protiated to deuterated substrate, respectively. Ratio (R_o and R) of deuterated substrate molecules, a_D, to protiated substrate molecules, a_H, are given in equation (2.5):

$$R = \frac{a_D}{a_H} \qquad \text{equation (2.5)}$$

This equation can be rearranged into an exponential format (equation 2.6), which can give an instructive graphical representation as depicted in Figure 2.13.[74]

$$\frac{R}{R_o} = (1-F_H)^{\left[\frac{1}{k_H/k_D} - 1\right]} \qquad \text{equation (2.6)}$$

In the case of primary and normal secondary KIEs where $k_H/k_D > 1$, the isotopic enrichment of the heavier isotope (R/R_o) in the recovered starting material increases exponentially as the reaction progresses to completion (Figure 2.14).[74] That is, as the value of F_H approaches unity, R/R_o accelerates exponentially to infinity. Alternatively, in the case of an inverse secondary KIE where k_H/k_D is less than 1, the isotopic enrichment of the lighter isotope (R/R_o) in the recovered starting material decreases exponentially towards a value of zero as the reaction reaches completion. Hence, in order to maximize the vertical spread of the curve thereby allowing an accurate measurement of KIEs, it is helpful to allow the reaction (i.e. F_H) to proceed to a significant extent. This would be

particularly helpful for determining smaller secondary KIEs, as the R/R_o spreads less dramatically in these cases.[74] Nevertheless, if the reaction proceeds too close to completion, the accuracy of R and F_H measurements might also be compromised by the low concentrations of the recovered starting materials. Hence, a balance between reaction progress and accurate data measurement needs to be struck in order to access the value of k_H/k_D in an unambiguous manner.

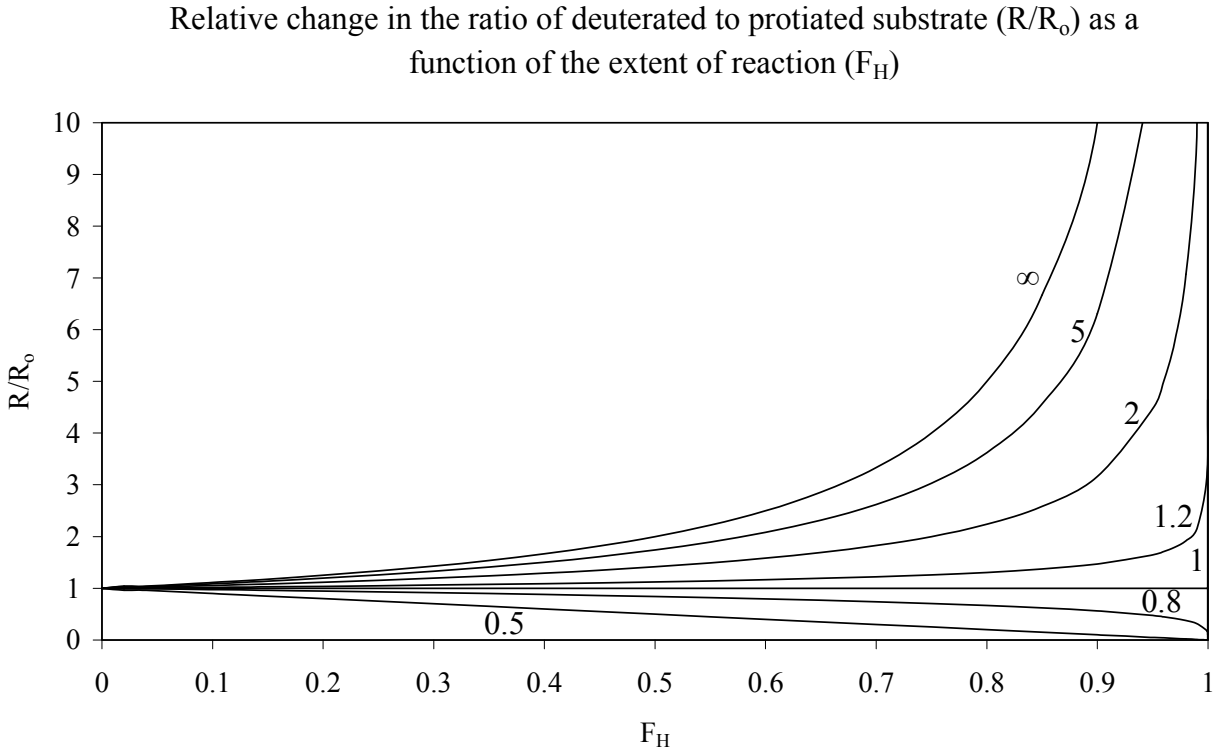


Figure 2.14 Relative change in the ratio of deuterated to protiated substrate (R/R_o) as a function of the extent of the reaction (F_H). Isotope effect values are indicated for each curve.

In the case of a reversible reaction, significant error may be introduced if the reaction is allowed to proceed too far, because equation (2.6) does not take into account any reversal of the reaction and its accompanied isotope effect.[74] Fortunately, the condensation reaction catalyzed by norcoclaurine synthase is considered as irreversible, since no starting materials can be detected by NMR and mass spectral methods upon an extended period of incubation with NCS.

An approximately equimolar solution (as determined by ESI-MS) of the deuterated [3,5,6-$^2H_3$]-dopamine (**29**) and unlabeled dopamine was enzymatically converted into norcoclaurine, and the ratio of the two isotopically labeled starting materials was monitored as a function of the reaction progress. The value of F_H was determined by using ^{1}H NMR spectroscopy and comparing the integrals obtained for the dopamine benzylic signals with those of the internal standard, dioxane, as a function of time. The values of R_0 and R were determined using positive ESI-MS by comparing the relative peak intensities of the labeled and unlabeled substrates as a function of time.

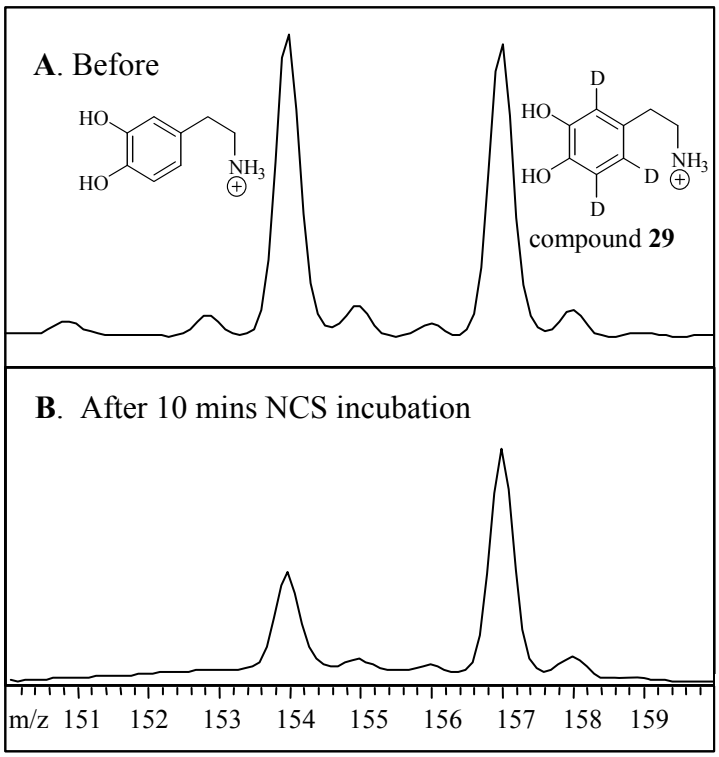


Figure 2.15 Mass spectra following the consumption of a mixture of unlabeled dopamine and [3,5,6-$^2H_3$]dopamine (29) in the reaction catalyzed by NCS. (Top) spectrum taken prior to the addition of enzyme. (Bottom) Spectrum taken after 86% of the unlabeled dopamine has been converted to product.

The competitive method for determining KIEs is much more accurate than the direct method, as reaction rates involving different isotopologues do not have to be measured individually and most of the random errors are cancelled. Figure 2.15 shows representative mass spectral traces that were used in determining the KIE. Prior to the addition of enzyme, the ratio of labeled to unlabeled

dopamine was about 51:49 (Figure 2.15, top spectrum). After 86% of the unlabeled material had been converted to product, the ratio changed to 32:68 (Figure 2.15, bottom spectrum). Since the two isotopologues are 3 m/z apart on the mass spectra, the heavier isotopologue peak does not need to be corrected for natural abundances. It is clear that the unlabeled material is reacting faster than the deuterated material. When applying the data to equation (2.6), a value of k_H/k_D was determined to be 1.7 ± 0.2 (Table 2.2).

Table 2.2 Summary of k_H/k_D calculations in the NCS reaction

Trials	a_D/a_H	R/R_o	F_H	k_H/k_D
1	0.706 ± 0.035*	1.000 ± 0.025	0	-
	2.196 ± 0.110	3.110 ± 0.008	0.936 ± 0.094**	1.68 ± 0.094***
2	0.695 ± 0.035	1.000 ± 0.025	0	
	1.715 ± 0.186	2.467 ± 0.006	0.860 ± 0.086	1.86 ± 0.102
3	0.966 ± 0.048	1.000 ± 0.025	0	
	2.149 ± 0.107	2.224 ± 0.006	0.861 ± 0.086	1.69 ± 0.092
			Average	1.7 ± 0.2◊

* An approximately equimolar solution of dopamine and [3,5,6-$^2H_3$]-dopamine (**29**) was subjected to positive ESI-MS for 6 separate trials, and relative error for a_D/a_H was estimated to be ~5%.

** Three separate ^{1}H NMR spectra (50 scans each) of a mixture of dopamine and dioxane were obtained (5 mM each in 1 mL), and the ratio of the integrals for the dopamine benzylic signals to the dioxane signals showed 5% deviation among the spectra. The relative error value for F_H was then estimated to be ~10%.

*** The propagated errors due to measurement limitations were indicated in the k_H/k_D calculated in each trial.[77]

◊ The average value was reported within the 95% confidence interval.

2.5 Discussion and Related Studies

The Pictet-Spengler reaction is a specialized form of electrophilic aromatic substitution reaction that involves a cyclization of an iminium ion with an electron rich phenolic or indole ring. In Nature only three known 'Pictet-Spenglerases', norcoclaurine synthase, strictosidine synthase, and deacetylipecoside synthase, have been identified. This thesis presents the first mechanistic investigations to be carried out on the NCS enzyme.

The reaction catalyzed by NCS likely proceeds through a direct electrophilic aromatic substitution without a spirocyclic intermediate (pathway A, Figure 2.7). This is strongly supported by the fact that dopamine analogs lacking a C-1 hydroxyl (**24** and **25**) served as efficient alternate substrates with catalytic constants comparable to those of dopamine itself. Indeed, in a non-enzymatic direct electrophilic substitution mechanism the replacement of the C-1 hydroxyl group for a hydrogen (analog **24**) would be expected to increase the reaction rate, since a meta hydroxyl group is slightly electron-withdrawing ($\sigma_m = 0.12$).[78] This rate enhancement was not observed with NCS, probably because the alkylation step is not rate-limiting in the enzymatic reaction. The inability of analogs **22** and **23** to serve as alternative substrates for the enzymatic reaction indicated the importance of the C-2 hydroxyl group in dopamine. The existence of a C-2 phenolate ion in the enzymatic reaction probably enhances the rate of cyclization onto the iminium ion in the first step of the substitution process.

The deuterium KIE of ~1.7 is interpreted as a partially masked primary KIE on the step involving deprotonation of the σ-intermediate that re-aromatizes to norcoclaurine. The observed KIE for non-enzymatic electrophilic aromatic substitution reactions (alkylations, chlorinations, and nitrations) often range from ~0.7–1.2.[79, 80] In these reactions, the difference in rates is generally attributed to a secondary kinetic isotope effect where the electrophilic addition step is rate-

determining. As mentioned in Section 2.4.1, inverse isotope effects would be expected because the carbon bearing the isotopic substituent experiences a hybridization state change from sp^2 to sp^3 and leads to an increase in zero point energy.[80] When the KIEs for some reactions were found to be close to or slightly greater than unity, the argument of hyperconjugation has been suggested.[80, 81] The out-of-plane carbon-hydrogen bond and the p-orbitals of the other five ring carbons are interacting with one another via hyperconjugation, and deuterium substitution may serve to decrease the zero point energy. This would counteract the effects resulting from re-hybridization at the transition state. An example of such a case is the alkylation of anisole with the diphenylmethyl cation ($k_H/k_D = 1.0$) where these opposing effects simply cancel one another.[82] In the case of norcoclaurine synthase, the isotope effect of $k_H/k_D = 1.7$ is notably larger than ones normally observed in non-enzymatic reactions. Thus, it is inconsistent with a secondary KIE on the electrophilic addition step. Instead, it reflects a primary isotope effect on the step involving deprotonation of the σ-intermediate. Nonetheless, this value is somewhat lower than what is expected for an intrinsic rate-determining primary kinetic isotope effect (~2−8). A likely explanation is that previous non-isotopically sensitive steps, such as iminium intermediate formation, are partially rate-determining and mask the intrinsic primary kinetic isotope effect. Alternatively, the observed value may reflect an intrinsic primary kinetic isotope that is smaller than expected because of an asymmetric transition state in an exothermic step of the reaction.[83, 84]

2.5.1 Crystallographic Studies on NCS

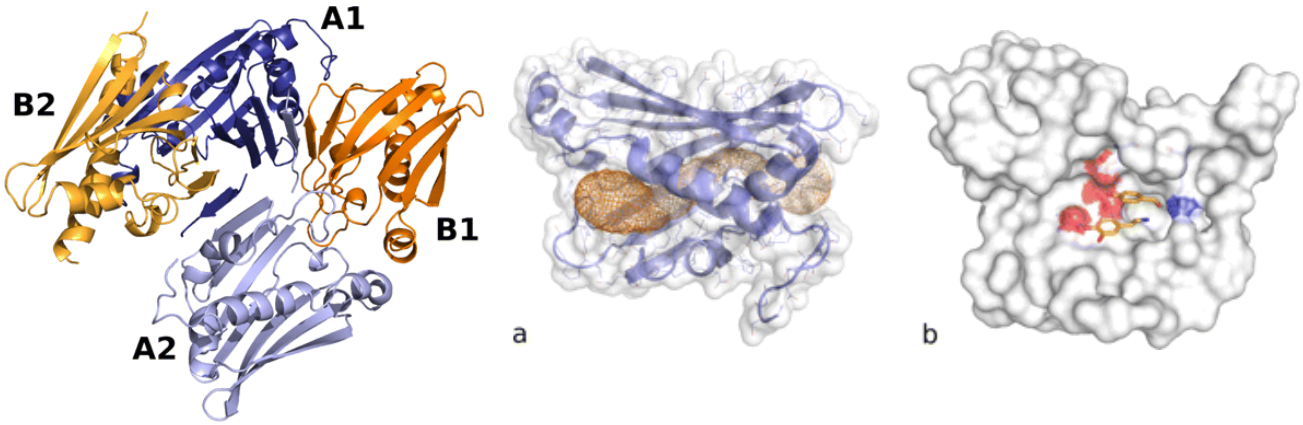


Figure 2.16 Model of NCS obtained from crystallographic studies. Diagrams obtained from Ilari *et. al.*[85]

The structure of NCS has also been extensively studied, and in 2009 the crystal structure of the enzyme was solved by the research group of Boffi.[85, 86] Although NCS appears a dimer in solution, analysis of the X-ray structures of NCS revealed a tetrameric assembly (Figure 2.16).[85] Each subunit possesses a seven-stranded antiparallel β-sheet that wraps around a long C-terminal helix and two smaller α-helical segments. Additional data were also obtained when the crystal of NCS was soaked with the natural substrate, dopamine, and a non-reactive aldehyde analog, *p*-hydroxybenzaldehyde (PHB, **30**). These molecules bound to the active site of NCS and provided a snapshot of the initial step of catalysis. The active site of the enzyme is located in an accessible cavity that is formed at the interfaces between the β-sheets and the three α-helices in each monomer. Inside the cavity, the dopamine and PHB molecules adopt a stacked configuration with respect to their aromatic moieties. The binding of substrate and analog is stabilized via a series of polar and hydrogen-bonding networks. Specifically, the carbonyl oxygen of PHB is hydrogen-bonded to Lys122, while the phenolic hydroxyl is hydrogen-bonded to the carboxylate of Asp141 (Figure 2.17). With respect to the dopamine molecule, it is held in place by a stacking interaction with PHB and by a hydrogen bond between the C-1 hydroxyl functionality and Tyr108.[85, 86] Furthermore, the C-5

position of dopamine lies 2.7 Å away from the carboxylate group of Glu110. A comparison between the structures of the unliganded and the substrate-bound NCS indicated that the position of the residues which were believed to participate in substrate binding and catalysis experienced very small adjustments (Figure 2.17).[85, 86]

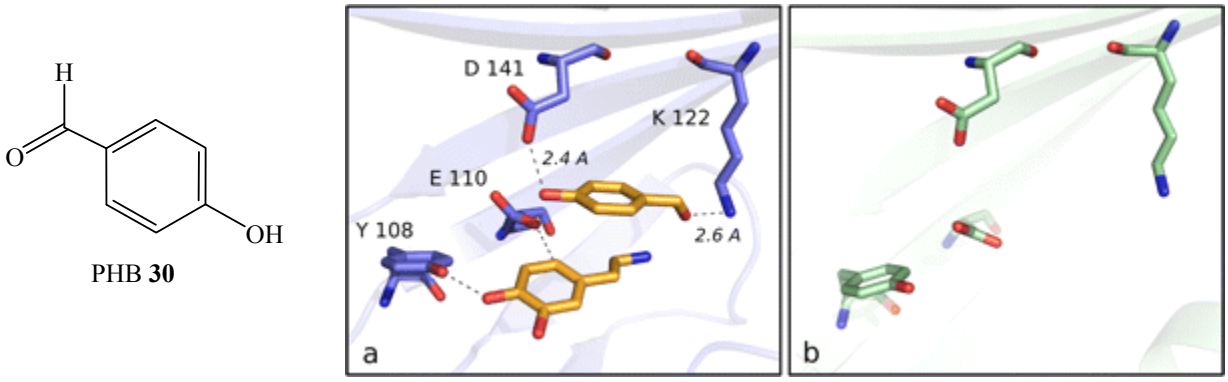


Figure 2.17 The active site of norcoclaurine synthase. Left diagram *p*-hydroxybenzaldehyde (**30**), middle diagram is the NCS active site containing 4-HPAA and PHB, and right diagram is the NCS active site without any ligands. Color code: nitrogen, blue; oxygen, red. Diagram obtained from the crystallographic studies of NCS. [85]

The Boffi group performed kinetic measurements on the wild type NCS and the mutants K122A, E110A and Y108F using the CD spectroscopy assay described earlier in this thesis.[85] No activity could be observed with the K122A mutant, whereas low but significant activity was detected with both Y108F and E110A mutants (Table 2.3). Based on this data, the authors proposed that Lys122 is responsible for binding 4-HPAA and catalyzes the formation of the iminium intermediate, while Glu110 deprotonates the σ-intermediate and leads to the re-aromatization of (*S*)-norcoclaurine.

Table 2.3 Kinetic parameters for wild type NCS and mutants

NCS	k_{cat}	K' of dopamine (μM)	K_M of 4-HPAA (μM)
Wild type	4.5 ± 0.4 (6.4 ± 0.3 in this thesis)*	380 ± 55 (350 ± 48)	330 ± 35 (288 ± 38)
K122A	-	-	-
E110A	0.8 ± 0.3	ND	505 ± 50
Y108F	1.7 ± 0.5	773 ± 62	360 ± 44

Kinetic data taken from the studies conducted by Ilari, *et. al.*,[85] unless otherwise indicated by parentheses.

Even though the crystallographic studies on NCS did not indicate whether the enzymatic reaction proceeds via a direct or indirect nucleophilic pathway, some subtle and interesting insights into the substrate binding of the active site were provided. In particular, the C-1 hydroxyl group of dopamine is hydrogen-bonded to the Tyr108 residue of the enzyme. This observation is somewhat counter-intuitive to the proposed mechanism (pathway A in Figure 2.7), since the C-2 hydroxyl group of dopamine should be deprotonated to facilitate the attack of C-5 alkylation onto the iminium functionality. Consequently, the C-2 hydroxyl group, instead of the one at the C-1 position, should interact with the enzyme. Several possible scenarios are offered to explain this observation. First, the data extrapolated from the crystallographic studies only provided insights into the initial step of the reaction mechanism. A significant change in the enzyme conformation and/or the orientation of the iminium intermediate might occur and allow the C-2 hydroxyl group to be deprotonated by NCS. Another possibility is that the Tyr108 residue first removes the proton at the C-1 hydroxyl group to yield a catecholate, which can then abstract the proton from the C-2 hydroxyl group and promotes formation of the σ-intermediate (Figure 2.18). However, this scenario is contradicted by the high catalytic competency of dopamine analogs, compounds **24** and **25**, which lack the C-1 hydroxyl group (Table 2.2). Finally, it is possible that a base is not required for the deprotonation of the C-2 hydroxyl group and that the enzyme simply binds the substrate in the phenolate form. For dopamine

in aqueous solution the first pK_a of the catechol moiety is expected to be near 9.5,[87] and this value could easily be lowered by the electrostatic interactions in the active site.

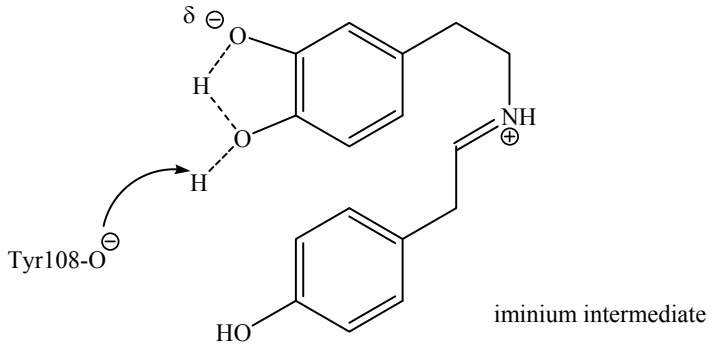


Figure 2.18 A potential proton-relay mechanism for the formation of the C-2 phenolate

Significant insights into the stereochemistry of the NCS reaction are also revealed in the crystallographic study. Given that the *Re*-face of the C-5 of dopamine faces toward 4-HPAA (or PHB) in the active site of the enzyme,[85] there are only two possible outcomes for alkylation, depending on the stereoconfiguration of the iminium functional group. If a *Z*-iminium functionality participates in the catalysis, electrophilic addition onto the *Si*-face of the iminium would yield a σ-intermediate that carries (*R*)- and (*S*)-stereoconfigurations at the C-1 and C-8a positions, respectively (Figure 2.19). Deprotonation at the C-8a by the basic residue Glu110 would yield (*R*)-norcoclaurine as product. In contrast, if an *E*-iminium group is involved, electrophilic addition has to occur onto the *Re*-face and would yield the diastereomeric (1*S*, 8a*S*)-σ-intermediate. Subsequent deprotonation would give (*S*)-norcoclaurine as product. Since NCS catalysis generates (*S*)-norcoclaurine, an *E*-iminium intermediate is likely involved in the reaction and the C-8a of the σ-intermediate should contain an (*S*)-stereoconfiguration.

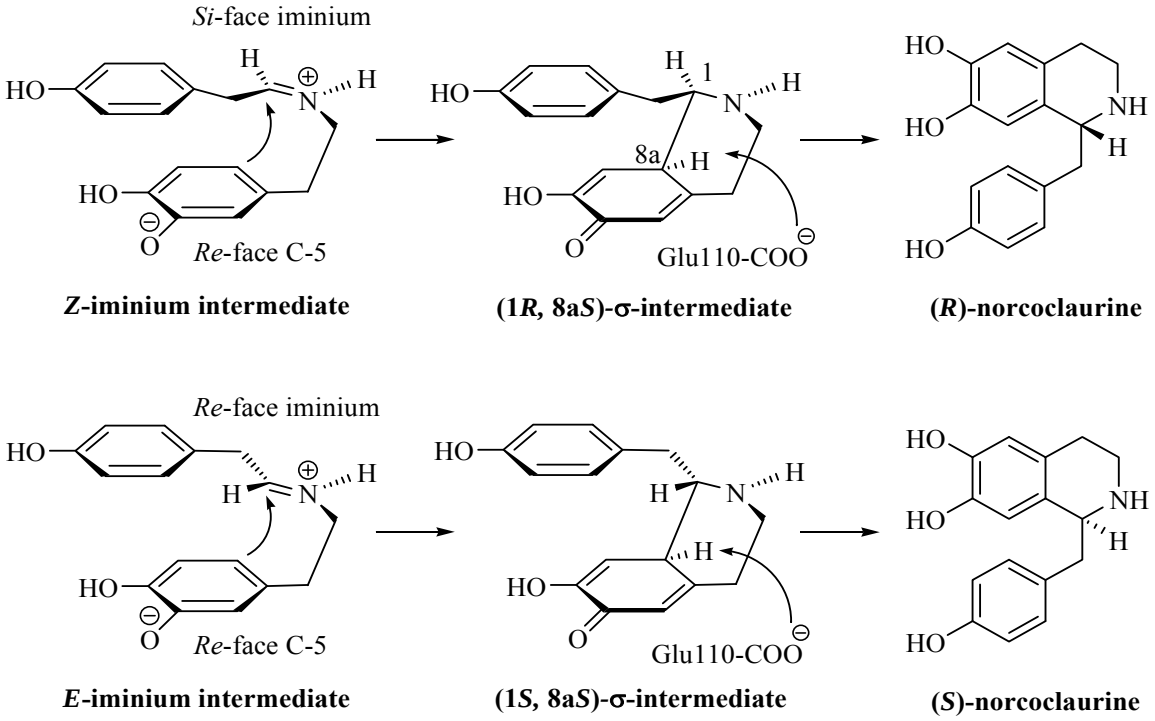


Figure 2.19 Two possible orientations in the electrophilic addition of NCS catalysis

2.5.2 Comparisons between the Mechanistic Studies and Crystal Structures of Norcoclaurine Synthase and Strictosidine Synthase

Approximately one year after our mechanistic studies on NCS were published,[63] studies on another 'Pictet-Spenglerase', strictosidine synthase, were published by the research group of O'Connor.[49] Interestingly, given the low sequence similarity and the poor structural homology between NCS and strictosidine synthase, many mechanistic similarities that were uncovered between these enzymes probably arise as a consequence of convergent evolution.[49, 53, 85, 86]

In the strictosidine synthase reaction, tryptamine and secologanin are condensed to yield strictosidine (see Section 1.4.3). A direct and an indirect nucleophilic pathway were also proposed by O'Connor (pathway A and B in Figure 1.17).[49] To probe the catalytic mechanism, O'Connor and

her coworkers employed a computer model study of a simplified Pictet-Spengler reaction between tryptamine and propanal. Although no direct experimental evidence could be obtained in this manner, the model study suggested that formation of the spiroindolenine intermediate is energetically unfavorable, thereby supporting the direct pathway (A, Figure 1.17).[49]

Competitive KIE studies were also carried out on the strictosidine synthase reaction. The substrate [2-^{2}H]-tryptamine gave a primary deuterium KIE value of $k_H/k_D \sim 2.6$ (Figure 1.17). The authors also concluded that re-aromatization of strictosidine is the rate-determining step in enzyme catalysis.[49] Interestingly, this observed KIE value was significantly higher than the corresponding k_H/k_D measured in NCS catalysis (~1.7). The difference between the magnitudes of these KIE values implies that there are differences between the reaction coordinates of the NCS and strictosidine synthase reactions. In the strictosidine synthase reaction deprotonation of the carbocation intermediate is a clearly (or fully) rate-determining step, whereas in the NCS reaction the corresponding step is only partially rate-determining. One possible explanation would be that deprotonating the C-2 hydroxyl group of dopamine is an extra step in the reaction catalyzed by NCS and may be partially rate-determining. Because there is no basic residue in proximity to this hydroxyl group as indicated in the crystallographic study (Figure 2.16), it is possible that removal of this proton is somewhat rate-limiting. In contrast, the corresponding step is absent in the reaction catalyzed by strictosidine synthase, thus making the step of re-aromatization to be cleanly rate-determining (Figure 1.17). Alternatively, the difference in the KIE values might be due to the differences in the rates of forming the iminium intermediates. Analysis of the crystal structure of strictosidine synthase led to the proposal that the negatively charged residue, Glu309, first binds tryptamine and then promotes the formation of the iminium functionality (Figure 1.16).[49] In the case of NCS, in turn, the positively charged residue, Lys122, was suggested to first bind to 4-HPAA and then catalyze the formation of the iminium group (Figure 2.16).[85, 86] Hence, the KIE differences

might suggest that formation of the iminium intermediate might proceed more readily in the active site of strictosidine synthase than the active site of NCS.

The non-enzymatic Pictet-Spengler reaction between [2-^{2}H]-tryptamine and propanal has also been inspected by O'Connor. The reaction was carried out in an acetate buffer (pH 4.6) that was meant to mimick the glutamate in strictosidine synthase catalysis, and it showed a KIE of ~2.6. The authors interpreted this observation to mean that strictosidine synthase uses the same mechanism as the non-enzymatic reaction.[49] This conclusion is somewhat ambiguous, as the mechanism of the non-enzymatic Pictet-Spengler reaction might involve a spiroindolenine intermediate, and the pKa of the active site glutamate is likely perturbed away from a value of 4.6. In the NCS reaction, it should also be possible to monitor the kinetics of the non-enzymatic reaction and determine if C-5 deprotonation is rate-limiting. However, deprotonation of the C-2 hydroxyl group of dopamine was believed to be important in the NCS catalysis, whereas the corresponding mechanism in the non-enzymatic reaction remained undetermined. Hence, future experiments will focus on elucidating the mechanism of non-enzymatic norcoclaurine formation.

2.6 Summary

In summary, the facile intramolecular ring closure of a phenolate onto an iminium ion explains why the rate-determining steps differ between the Pictet-Spenglerases and most non-enzymatic aromatic electrophilic substitution reactions. Formation of the iminium intermediate and the subsequent alkylation step are relatively fast chemical steps in the reaction as indicated in the KIE studies. Most electrophilic aromatic substitution reactions, in contrast, are carried out under acidic conditions where a phenolate will not exist, and the alkylation steps are often the slow step. The electrophilic addition of dopamine in the NCS reaction is relatively fast, because formation of

the 6-membered ring involves a favorable 6-*endo*-trig ring closure.[31] Additionally, the reactivity is further enhanced by the possibility of simultaneously having a phenolate and an iminium ion in the zwitterion intermediate **19**. The first pK_a of the catechol moiety is expected to be near 9.5 and this may account for the ability of a background non-enzymatic reaction to occur at measurable rates.[87] In the case of the enzyme-catalyzed reaction, this pK_a value is likely perturbed to a lower value since maximal activity is observed at a pH of 7.0.[47]

2.7 Future Directions

In order to complete our mechanistic studies, the precise roles of the amino acid residues involved in the NCS reaction will be investigated. In the crystallographic studies, the authors proposed that Glu110 and Tyr108 play crucial roles in catalysis, and site-directed mutation of these residues reduces the activity to different extents.[85] The functional roles of these residues can be further examined using a combination of techniques, including site-directed mutagenesis and KIE measurement. Results obtained from these experiments should provide more insights into the Pictet-Spengler reaction catalyzed by NCS.

Due to the close proximity between the carboxylate group of Glu110 and the C-5 of dopamine, this residue was proposed to deprotonate the σ-intermediate (Figure 2.17).[85] If Glu110 is mutated to a weak hydrogen-bond acceptor, E110Q, or a non-polar residue, E110A, the energy barrier to re-aromatization of (*S*)-norcoclaurine should increase considerably. If sufficient activity permits, a KIE will be measured with these mutants using [3,5,6-$^2H_3$]-dopamine (**29**), and an intrinsic primary KIE ($k_H/k_D > 1.7$) will be anticipated because deprotonation of the σ-intermediate should become fully rate-determining in the reaction.

Our mechanistic studies have indicated that deprotonation of the C-2 hydroxyl of dopamine is crucial for catalysis, but the crystal structure shows no residue in close proximity to this functional group.[85] Instead, the residue Tyr108 was found to be hydrogen-bonded to the C-1 hydroxyl of dopamine, and site-directed mutation to a phenylalanine residue (Y108F) decreased the catalytic activity (k_{cat}) significantly.[85] It was proposed that this residue is either responsible for substrate binding and/or for abstracting the proton from the C-2 hydroxyl group via a proton-relay mechanism and promoting the addition of C-5 to the iminium functionality (Figure 2.18). To differentiate these possibilities, the Y108F mutant will be subjected to KIE measurement using [3,5,6-$^2H_3$]-dopamine (**29**). If Tyr108 is only responsible for substrate-binding, the magnitude of k_H/k_D measured with the Y108F mutant should remain at ~1.7. In contrast, if the tyrosine residue is responsible for deprotonating the C-2 hydroxyl group, the magnitude of k_H/k_D should drop considerably (< 1.7) because the electrophilic addition step should become either partially or fully rate-determining.

2.8 Experimental Procedures

2.8.1 Materials

All chemicals were purchased from Sigma-Aldrich and used without further refinement unless otherwise noted. 4-Hydroxyphenylacetaldehyde (4-HPAA) was synthesized by the Doering−Parikh oxidation of 2-(4-hydroxyphenyl)ethanol as described previously.[88] It was stable indefinitely if stored at -78 °C as a neat oil.

Thin layer chromatography (TLC) was performed on aluminum-backed sheets of silica gel $60F_{254}$ (Merck) of thickness 0.2 mm. Compounds were visualized by UV or by spraying with a solution containing H_2SO_4 (31 mL), ammonium molybdate (21 g), and $Ce(SO_4)$ (1 g) in water (500

mL) and then heating at 110 ^{o}C for 1 min. Silica gel (230 – 400 mesh, BDH) was used for column chromatography.

^{1}H NMR spectra were obtained on a Bruker AV300 and AV400 spectrometer at field strengths of 300 or 400 MHz, respectively. All chemical shifts were reported using the δ scale in ppm. Mass spectrometry was performed by the Mass Spectrometry Center at UBC by electrospray ionization (ESI-MS) using a Waters Micromass LCT mass spectrometer.

The plasmid pNCSΔ19, encoding for the *Thalictrum flavum* NCS with the first 19 N-terminal amino acids truncated was provided by Dr. Peter J. Facchini at the University of Calgary, Alberta. [47] JM109 (DE3) and BL21 (DE3) RIL *E. coli* cells were purchased from Promega and Stratgene, respectively.

2.8.2 General Methods

2.8.2.1 General Enzyme Methods

Centrifugal filters (4 mL 10 000 MWCO) were purchased from Millipore. Acryl-cuvettes for use in enzyme kinetic assays were from Sarstedt. Chelating Sepharose® Fast resin was purchased from Pharmacia Biotech. Protein concentrations were determined by the method of Bradford on a Cary 3E UV-Vis spectrophotometer using bovine serum albumin as standard.[89] All measurements were performed at room temperature. Protein purity was assessed using SDS-PAGE, stained with Coomassie blue according to Laemmli.[90] Molecular weight standards for SDS-PAGE were BSA (66 kDa) and carbonic anhydrase (29 kDa), both purchased from Sigma.

Deuterated buffers were made by dissolving buffer components first in H_2O, adjusting to a desired pH with dilute HCl, followed by concentration to dryness by vacuum centrifugation and reconstituted with an equal volume of D_2O. The pD values of the resulting buffer solutions were

determined by adding 0.4 units to the observed pH meter readings, as has been empirically established by Glasoe and Long.[91] General microbiology techniques used are those described by Ausubel *et. al.*[92]

2.8.2.2 Overexpression and Purification of Norcoclaurine Synthase (TfNCSΔ19)

The recombinant pNCSΔ19 plasmid was transformed into BL21 (DE3) RIL $CaCl_2$-competent *E. coli* (Stratagene).[47] The cells were incubated overnight at 37 °C in 10 mL Terrific Broth (TB) medium containing 35 μg/mL chloramphenicol and 30 μg mL^{-1} kanamycin with shaking at 225 rpm. The overnight cultures were poured into 1 L of Luria-Bertani (LB) medium containing 37 μg/mL chloramphenicol and 30 μg/mL kanamycin and grown at 37 °C with shaking at 225 rpm until an OD_{600} of 0.6 was reached. Cells were induced for overexpression by addition of 238 mg (1 mM) of isopropyl β-D-galactopyranoside (IPTG), and the cultures were allowed to continue growth at 24 °C until an OD_{600} of 3.0−4.0 was reached (~24 h). Cells were harvested by centrifugation at 6000 rpm in Sorvall® SLC-1500 rotor for 30 min and resuspended in 20 mL of 200 mM potassium phosphate buffer (pH 7.5) containing 1 μg/mL pepstatin A and 1 μg/mL aprotinin. The cells were lysed at 20 000 psi in an ice-cooled French pressure cell. The cell lysate was clarified by centrifugation at 6000 rpm for 60 min and filtered through a 0.22 μm membrane prior to affinity chromatography.

A 10 mL column containing Chelating Sepharose® Fast Flow resin (Pharmacia Biotech) was charged with 2 column volumes (CV) of 100 mM $NiSO_4$ followed by washing with 2 CV of distilled H_2O and 3 CV of phosphate buffer (200 mM potassium phosphate buffer, pH 7.5) containing 5 mM imidazole. The filtered cell lysate was loaded at 2 mL/min, and phosphate buffer containing 5 mM imidazole (~10 CV) was passed through the column at 3 mL/min until no more flow-through protein eluted, as determined by monitoring A_{280}. A wash with phosphate buffer containing 125 mM

imidazole (~5 CV) was used to remove nonspecifically bound proteins. Histidine-tagged protein was finally eluted with 3−4 CV of phosphate buffer containing 500 mM imidazole. The enzyme was concentrated by ultrafiltration (Amicon Ultra-4, 10 000 MWCO) to 5−30 mg/mL, and 25−50 μL aliquots were flash-frozen with liquid N_2 in the presence of 10% glycerol. The enzyme could be stored at -70 °C for at least 6 months without significant loss of activity.

2.8.2.3 ^{1}H NMR Assay of Norcoclaurine Synthase Activity.

A solution of deuterated potassium phosphate buffer (200 mM, pD 7.5, 1.00 mL) containing dopamine (10 mM) and 4-HPAA (10 mM, prepared from the addition of 10 μL of a 1 M stock in d^4-methanol) was placed in an NMR tube, and an initial ^{1}H NMR spectrum was taken. To the solution was then added 100 μg of norcoclaurine synthase (TfNCSΔ19), and after incubating the reaction mixture for 30 min at 25 °C a ^{1}H NMR spectrum was taken.

2.8.2.4 Isolation and Characterization of Norcoclaurine

To a solution of potassium phosphate buffer (200 mM, pH 7.5, 10 mL) containing dopamine (50 mM) and 4-hydroxyphenylacetaldehyde (50 mM, prepared from the addition of 500 μL of a 1 M stock in methanol) was added 220 μg of norcoclaurine synthase (TfNCSΔ19), and the mixture was incubated for 30 min at 37 °C. Enzyme was removed by centrifugal ultrafiltration, and the resulting filtrate was evaporated to dryness under reduced pressure. In order to isolate (*S*)-norcoclaurine, the resulting solid was suspended in 100 mL of ice-cold EtOH and filtered to remove the phosphate salts. Removal of the solvent under reduced pressure gave (*S*)-norcoclaurine that was free of phosphate salts as analyzed by ^{31}P NMR spectroscopy. When analyzed by circular dichroism spectroscopy, material isolated in this fashion retained the same molar ellipticity as was originally observed during the enzymatic reaction.

In attempt to purify the compound from minor impurities, the crude was subjected a silica gel column. However, the product obtained in this fashion was found to be racemic. The solid obtained following ultracentrifugation and evaporation was dissolved in 10 mL of MeOH, and 600 mg of silica gel was added. After MeOH was removed under reduced pressure, the silica gel-adsorbed sample was dry-loaded onto a silica gel column (50 mL). The column was washed with 1000 mL of 100:1 acetone/acetic acid, and the product was eluted with 200 mL of 50:50:1 acetone/methanol/acetic acid. The resulting (*R*,*S*)-norcoclaurine acetate salt was characterized using ^{1}H NMR in D_2O and positive ESI-MS mass spectrometry: ^{1}H NMR (D_2O) δ 1.96 (s, 3H), 2.96 (m, 3H), 3.28 (m, 1H), 3.47 (m, 2H), 4.60 (dd, 1H, *J*= 9.2 Hz, *J* = 5.5 Hz), 6.68 (s, 1H), 6.74 (s, 1H), 6.90 (d, 2H, *J* = 8.5 Hz), 7.18 (d, 2H, *J* = 8.5 Hz); positive ESI-MS (MeOH) *m*/*z* 272 ($M + H^+$).

2.8.3 Reactions with Substrate Analogs.

To test the four dopamine analogs **22-25** as alternate substrates, solutions containing 30 mM of the analog and 30 mM 4-hydroxyphenylacetaldehyde in 200 mM potassium phosphate buffer (pH 7.5, 5 mL total volume) were incubated with 150 μg of norcoclaurine synthase (TfNCSΔ19) at 37 °C for 24 h. The reaction progress was monitored by positive ESI-MS. To test 3,4-hydroxyphenylacetaldehyde **26** as an alternate substrate, a similar incubation containing 30 mM **26** and 30 mM dopamine was monitored. Products from the successful reactions (obtained with analogs **22**−**25**) were isolated in an identical manner as described for (*S*)-norcoclaurine. 7-Deoxynorcoclaurine acetate (product from analog **24**): ^{1}H NMR (D_2O) δ 1.96 (s, 3H), 3.04 (m, 3H), 3.50 (m, 3H), 4.70 (dd, 1 H, *J* = 8.5 Hz *J* = 6.2 Hz), 6.77 (s, 1H), 6.79 (m, 1 H, *J* = 8.5 Hz), 6.91 (d, 2 H, *J* = 8.3 Hz), 7.12 (d, 1 H, *J* = 8.5 Hz), 7.19 (d, 2 H, *J* = 8.3 Hz); +ve ESI-MS (MeOH) *m*/*z* 256 ($M + H^+$). 7-Methoxynorcoclaurine acetate (product from analog **25**): ^{1}H NMR (D_2O) δ 1.96 (s, 3H), 3.05 (m, 1H), 3.30 (m, 3H), 3.50 (m, 2H), 3.65 (s, 3H), 4.60 (t, 1H, *J* = 8.5 Hz), 6.68 (s, 1H),

6.74 (s, 1H), 6.90 (d, 2H, J = 8.5 Hz), 7.18 (d, 2H, J = 8.5 Hz); +ve ESI-MS (MeOH) m/z 286 ($M + H^+$).

2.8.4 Deuterium Incorporation Experiment

The intermediate analog, *N*-(4-hydroxyphenethyl)-*N*-(3,4-dihydroxyphenethyl)amine hydrochloride **28**, was synthesized as previously described.[93] To three NMR tubes containing 900 μL of deuterated potassium phosphate buffer (200 mM, pD 7.5) was added either 50 μL of a solution containing 0.1 M dopamine, or 50 μL of a solution containing 0.1 M dopamine and 0.1 M 2-(4-hydroxyphenyl)-ethanol (tyrosol, **27**), or 50 μL of a solution containing 0.1 M of compound **28** (final concentrations of 5 mM). The incubations were initiated by the addition of 50 μL of norcoclaurine synthase (TfNCSΔ19, 150 μg) and incubated at 37 °C. ^{1}H NMR spectra and ESI-MS (following dilution of an aliquot into MeOH) were acquired at 0 and 24 h.

2.8.5 Circular Dichroism Spectroscopic Assay

Enzyme kinetics were measured by monitoring the formation of (*S*)-norcoclaurine using circular dichroism spectroscopy. An initial determination of the molar ellipticity of (*S*)-norcoclaurine was made by incubating dopamine and 4-HPAA (100 to 1000 μM each) in 200 mM phosphate buffer, pH 7.5, with 30 μg of norcoclaurine synthase (TfNCSΔ19) for 30 min at 37 °C. The enzyme was removed by ultracentrifugation and a circular dichroism spectrum was taken that showed a maximal signal at 285 nm with [θ] = 12541 mdeg cm^{-1} M^{-1}. Control reactions run in D_2O and monitored by ^{1}H NMR spectroscopy confirmed that the enzymatic reaction had proceeded to completion during this incubation time.

To measure the enzyme kinetics, a 0.5 cm cuvette containing 200 mM potassium phosphate buffer (pH 7.5), 4-HPAA (variable) and dopamine (variable) was thermally equilibrated for 3 min at

37 °C. The enzymatic reaction was initiated by the addition of norcoclaurine synthase (TfNCSΔ19, 2.1 μg) for a total assay volume of 1000 μL, and the enzymatic rate was calculated from the observed increase of signal at 285 nm (using [θ] = 12541 mdeg cm^{-1} M^{-1}). The K_m value for 4-HPAA was measured in the presence of 2 mM dopamine (saturating) with the concentration of the aldehyde varying between 100 μM and 2000 μM. The K_M value for dopamine was measured in the presence of 2 mM 4-HPAA (saturating) with the concentration of dopamine varying between 100 μM and 2000 μM. Kinetic parameters were determined from initial velocities fit to the Michaelis−Menten (4-HPAA) and sigmoidal (dopamine) kinetics using the programs GraFit and Sigma Plot, respectively.

Kinetic constants for the dopamine analogs **24** and **25** were measured under identical conditions except that they were monitored at 280 and 285 nm and the molar ellipticities of the products were found to be [θ] = 6590 and 6175 mdeg cm^{-1} M^{-1}, respectively. Kinetic constants for the 4-hydroxyphenylacetaldehyde analog **26** were measured under identical conditions except that the reaction was monitored at 280 nm and the molar ellipticity of the product was found to be [θ] = 5963 mdeg cm^{-1} M^{-1}.

2.8.6 Kinetic Isotope Effect Studies

[3,5,6-$^2H_3$]-Dopamine **29** was prepared according to the procedure of Vining *et. al.*[94] The extent of ^2H incorporation was determined to be ≥97% by ^1H NMR spectroscopy and ESI-MS. To a solution of an approximately 1:1 molar ratio of unlabeled dopamine and [3,5,6-$^2H_3$]-dopamine (12.5 mM each) and 30 mM 4-HPAA in 200 mM potassium phosphate buffer (pH 7.5, 5 mL total) was added norcoclaurine synthase (TfNCSΔ19, 340 μg in 100 μL of 200 mM potassium phosphate buffer, pH 7.5). 100 μL aliquots were collected at 0, 2.5, 5, 10, and 15 min time intervals. To each sample was added 200 μL of 1 M HCl to inactivate the enzyme, followed by neutralization with 200 μL of 1

M NaOH. The samples were immediately diluted with 20 mL of H_2O, flash frozen with liquid N_2, and lyophilized. Final ^1H NMR (D_2O containing 1 mM dioxane) and ESI-MS spectra were acquired. The KIE on k_{cat}/K_M was determined from the initial and final spectra using the equation KIE = ln(1 − F_H)/ln[(1 − F_H)R/R_0],[74] where F_H is the fractional conversion of the protiated species to products and R_0 and R are the initial and final ratios of protiated to deuterated substrate, respectively. F_H was calculated from the NMR spectra before and after addition of the enzyme by integration of the benzylic dopamine signals at δ 3.11 ppm relative to the internal standard of 1 mM dioxane. R_0 and R were calculated from the mass spectra before and after addition of the enzyme using the relative peak intensities of the protiated substrate [*m/z* 154 ($M + H^+$)] to the deuterated substrate [*m/z* 157 ($M + H^+$)]. The reported value represents the average of three independent determinations. A control reaction lacking enzyme showed that <5% product was formed from non-enzymatic background reaction during these incubation times.

Chapter 3

Alkaloids and Enzyme-catalyzed Aromatic Prenylation in Nature

Dimethylallyltryptophan synthase (DMATS) is another example of an enzyme that catalyzes an electrophilic aromatic substitution in Nature. This enzyme participates in the biosynthesis of ergot alkaloids, and the reaction catalyzed by DMATS is significantly different from the one catalyzed by norcoclaurine synthase. Hence, DMATS warranted further mechanistic investigation, and the rest of this thesis will elaborate on the background (Chapter 3) and our studies of this enzyme (Chapter 4). Since this enzyme belongs to the family of prenyltransferases (PTs), mechanistic investigations carried out on different PTs will also be summarized in this chapter.

3.1 Prenyltransferases

In Nature, prenylation refers to the addition of an isoprenoid group to an acceptor molecule, and these reactions are catalyzed by enzymes called prenyltransferases (PTs).[95, 96] PTs are ubiquitously distributed among prokaryotes and eukaryotes, and serve many crucial physiological functions, including cell signaling and protein trafficking. Prenylation is also a key step in the biosynthetic pathways of many natural products found in a wide variety of microorganisms. These prenylated natural products often have a wide range of biological activities. Whereas some of them are considered to be toxins and/or hallucinogenic reagents, such as certain ergot alkaloids, many others have been exploited as medicinal agents, such as the antibiotic clorobiocin.[95, 96] Hence, the studies of PTs have attracted widespread attention from the scientific community.

In a PT-catalyzed reaction, an acceptor serves as a nucleophile, and is alkylated by an isoprenyl diphosphate to form a prenylated molecule and pyrophosphate as products (Figure 3.1). Depending on the type of acceptors involved in catalysis, PTs can be classified into three main groups as follows: isoprenyl diphosphate synthases, protein prenyltransferases and aromatic

prenyltransferases.[95, 96] Many PTs have been investigated mechanistically and these studies will be presented in the following sections.

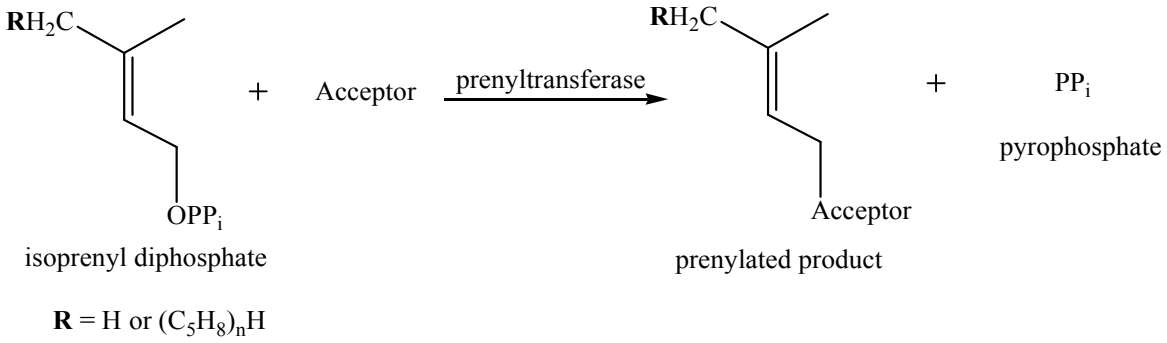


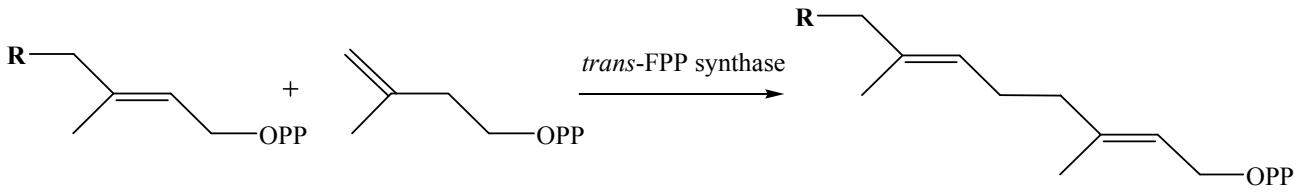
Figure 3.1 The reaction catalyzed by prenyltransferases

3.2 Isoprenyl diphosphate Synthases: *trans*-Farnesyl Diphosphate Synthase

Isoprenyl diphosphate synthases comprise a group of PTs that employ different isoprenyl diphosphates as electrophiles to produce allyl diphosphates of increasing carbon lengths. All of these enzymes show amino acid sequence homology and possess two aspartate-rich conserved DDxxD motifs, suggesting that they have evolved from a common ancestor. Based on the stereochemical outcome of the reaction, these enzymes can be further subdivided into *trans*- and *cis*-isoprenyl diphosphate synthases.[95] In particular, *trans*-farnesyl diphosphate (FPP) synthase is an enzyme that participates in the biosynthesis of sterol compounds. This enzyme was the first PT to be discovered and is arguably the most thoroughly studied PT as well.[97]

Studies on the isolation and initial characterization of the *trans*-FPP synthase were performed by the research groups of Porter and Popják in the 1960s.[98, 99] The *trans*-FPP synthase catalyzes two sequential 1'-4 condensation reactions (Figure 3.2). In the first reaction, the hydrocarbon moiety of dimethylallyl diphosphate (DMAPP) is added to isopentenyl diphosphate (IPP) to give a C-10

geranyl diphosphate (GPP), which in turn serves as a substrate in the second reaction and reacts with another IPP molecule to give the C-15 FPP. The enzymatic activity was found to require the presence of a divalent cation, Mn^{2+} or Mg^{2+},[99] that is chelated by the aspartate residues in the DDxxD motif and facilitates the binding of the diphosphate moieties.[95]



isoprenyl diphosphate IPP product

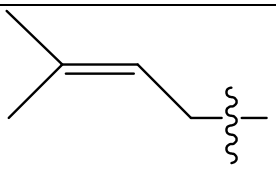
trans-FPP synthase	**R**	**Isoprenyl diphosphate**	Overall Reaction
First reaction	H	DMAPP	DMAPP + IPP → GPP
Second reaction		GPP	GPP + IPP → FPP

Figure 3.2 Two sequential reactions catalyzed *trans*-FPP synthase reaction

The stereochemistry of the *trans*-FPP synthase catalysis has been investigated by Popják and Cornforth.[100–102] When the tritiated (*S*)-[1-$^3H_1$]-DMAPP was used in catalysis, the C-1 position of the starting material experienced an inversion in stereoconfiguration (Figure 3.3).[100, 102] Additionally, studies with *trans*-[4-$^3H_1$]-IPP showed that the *Si*-face of the acceptor alkene adds to DMAPP, because the enzyme-catalyzed reaction only generated the *S*-stereoisomer as product.[101] Finally, the tritium of (*R*)-[2-$^3H_1$]-IPP is always lost to solvent during catalysis to give a product containing a *trans* double bond.[100, 101]

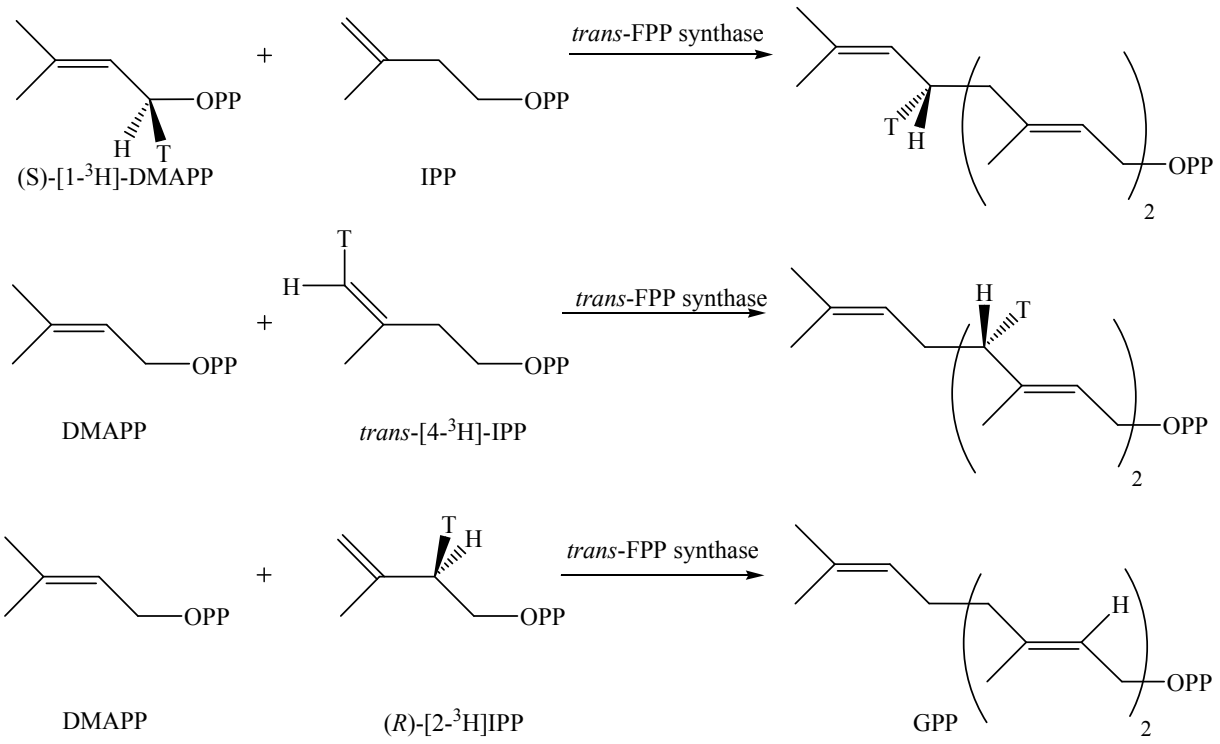
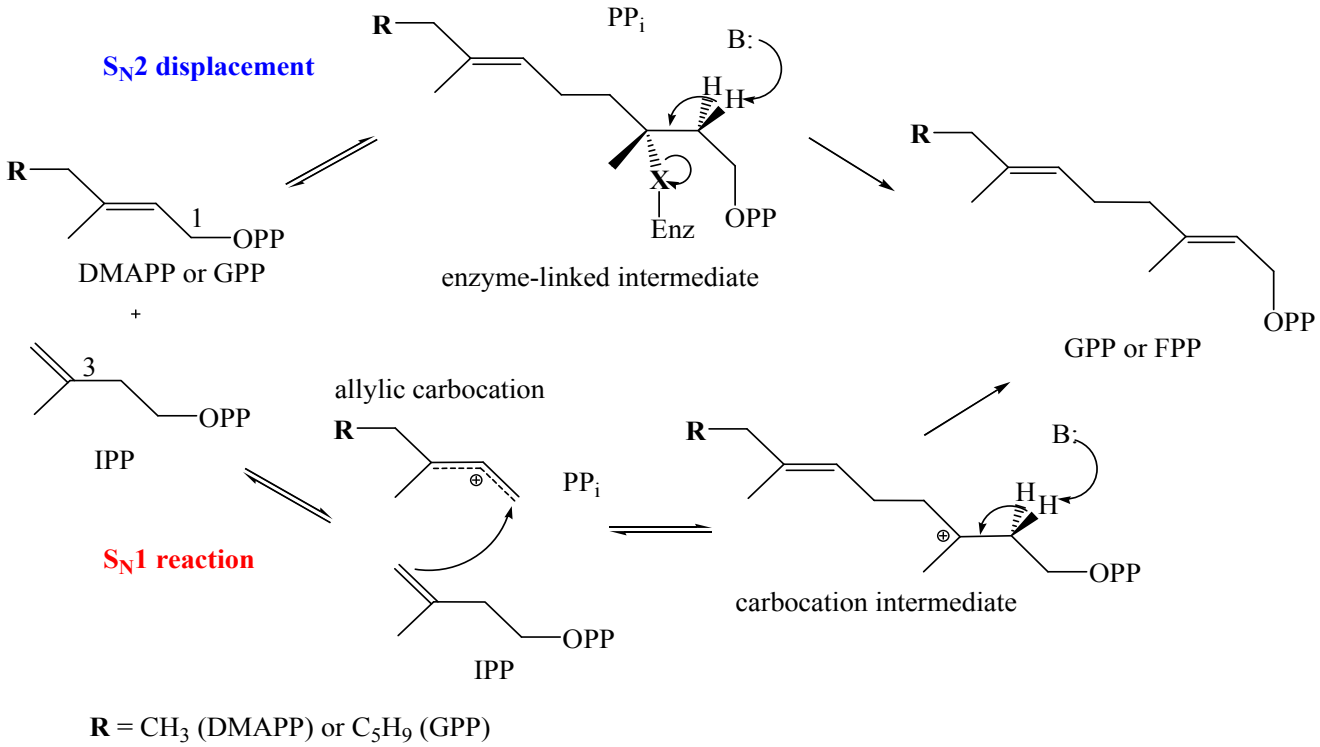


Figure 3.3 Stereochemical analysis of the *trans*-FPP synthase reaction

Based on these observations, Popják and Cornforth proposed an S_N2 displacement mechanism for the *trans*-FPP synthase reaction (Figure 3.4). An unknown nucleophile (the **X** group) attacks the C-3 of IPP, thereby promoting the subsequent nucleophilic attack of the C-4 of IPP onto the C-1 of DMAPP (or GPP) and the displacement of pyrophosphate. In a subsequent step, elimination of the **X** group of the enzyme-linked intermediate yields the product.[100]

In addition to this S_N2 displacement pathway, Poulter has also proposed an alternative S_N1 mechanistic pathway, which was also referred as an 'ionization-condensation-elimination' mechanism.[97] This pathway begins with dissociation of the diphosphate moiety from DMAPP (or GPP) forming a tertiary allyl carbocation, which then adds to the double bond of IPP to generate a tertiary carbocation intermediate. Removal of the C-2 proton of this intermediate generates a new *trans*-double bond in the product GPP (or FPP). Poulter argued that the stereochemical studies were insufficient to distinguish the S_N2 and the S_N1 pathways, because enzyme-catalyzed reactions tend to

be stereoselective regardless of the mechanism. Hence, further studies were needed in order to elucidate the mechanism of the *trans*-FPP synthase reaction.[97]



R = CH_3 (DMAPP) or C_5H_9 (GPP)

Figure 3.4 Proposed mechanisms for the *trans*- FPP synthase reaction

In order to verify the S_N1 pathway, Poulter and his co-workers synthesized a series of fluorinated DMAPP analogs and tested them as alternative substrates.[97] Trifluoromethyl (CF_3) analogs of DMAPP, (*E*)- and (*Z*)-3-trifluoromethyl-2-butenyl diphosphate (CF_3-DMAPP, *E*-**31** and *Z*-**31**), were employed (Figure 3.5).[97, 103] Formation of the corresponding allyl cations from these analogs was believed to be highly unfavorable, because they carry a strongly electron-withdrawing CF_3 group which has a substituent constant of $\sigma^+ = 0.612$.[103, 104] In contrast, the CF_3 group has been shown to slightly enhance the rate of nucleophilic displacement of chloride by iodide in a (*E*)-2-butenyl system.[105] Hence, the authors predicted that the reactivity of analog **31** would be slightly accelerated in a S_N2 displacement but greatly depressed in the case of the S_N1 reaction.[97, 103, 105] The prediction made by the Poulter group regarding the increased rate at the S_N2 reaction is somewhat

contentious, because fluorine substitution could also cause rate depression in the case of an 'exploded'-S_N2 mechanism as described in the later sections. Also, the enzyme might change its mechanism in order to accommodate the fluorinated analogs as substrate. Nonetheless, the authors observed some interesting experimental results when employing the fluorinated substrate analogs and were one of the pioneers in the use of this technique in mechanistic enzymology.

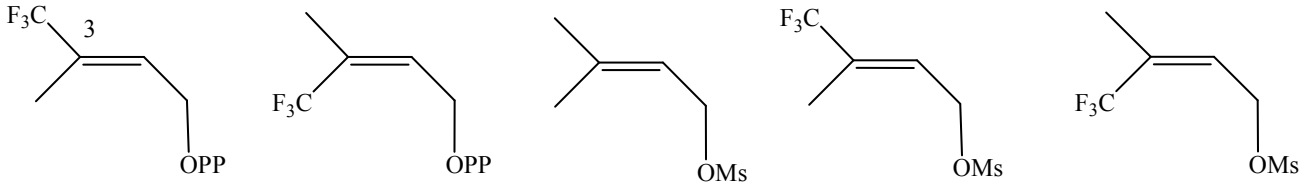
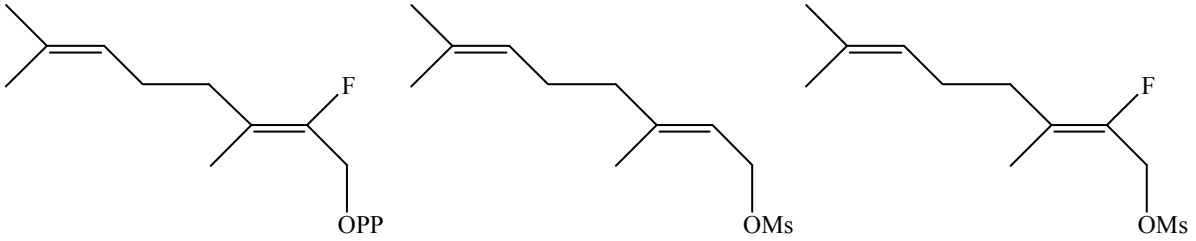


Figure 3.5 Trifluoromethyl (CF3) analogs for dimethylallyl diphosphate (CF_3-DMAPP, *E*-31 and *Z*-31), dimethylallyl methanesulfonate (DMA-Ms, 32) and its fluorinated analogs (*E*-33 and *Z*-33)

To analyze the reactivity of these analogs, the relative reaction rate (k^{rel}) was calculated by dividing the rate of catalysis obtained with the fluorinated analog by the one obtained with the natural substrate DMAPP.[97, 103, 105] It turned out that the k^{rel} values were only ~3 x 10^{-7}, showing a great rate depression as a result of the replacement of the C-3 methyl group by the CF_3 group.[97] Additionally, the non-enzymatic cationic solvolysis of methanesulfonate analog (DMA-Ms, **32**) and its fluorinated analogs (CF_3-DMA-Ms, compound ***E***- and ***Z*-33**) were carried out, and the corresponding k^{rel} values were also evaluated.[103] The k^{rel} values of **33** in the solvolytic reaction were found to be ~1 x 10^{-7}, similar to those of the enzymatic reaction. Accordingly, these experimental data supported the S_N1 reaction pathway, in which an allyl carbocation intermediate was formed during the *trans*-FPP synthase catalysis.[103] Nevertheless, some ambiguities remained. The Michaelis constants (K_M) for the ***E***- and ***Z*-31** were 51 and 61 μM, respectively, whereas the K_M value of DMAPP was determined to be 1.8 μM. The difference implied that the binding affinities for these analogs were substantially reduced. Further kinetic analysis showed that the fluorinated

inhibitors gave a mixed linear pattern in a Lineweaver-Burk plot. That is, the high K_M values arise from the inhibitory binding of the fluorinated substrates to the IPP binding site instead of the DMAPP/GPP binding site. This unexpected inhibition effect could possibly contribute to the observed rate depression.[103]

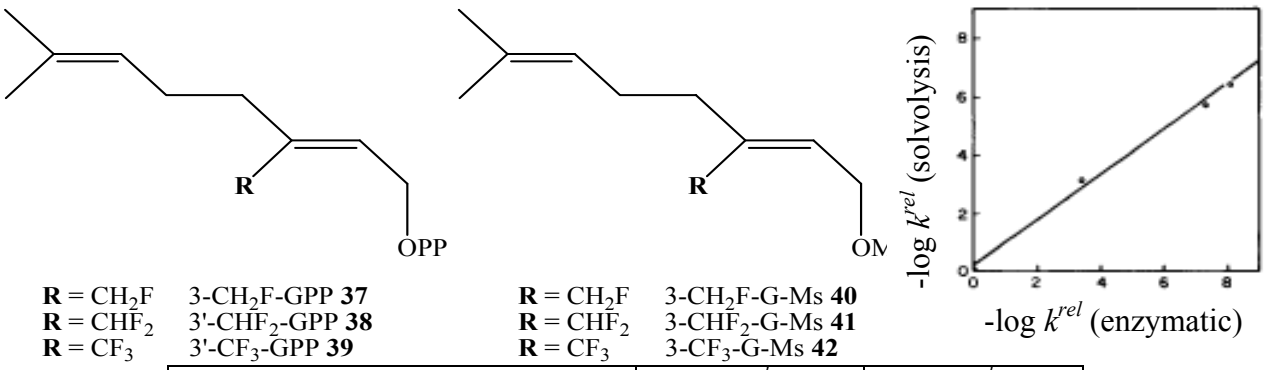
To resolve the mixed inhibitor binding in the previous experiment, 2-fluorogeranyl diphosphate (**34**) was tested as an alternative substrate for *trans*-FPP synthase (Figure 3.6).[106, 107] This analog contains a longer hydrocarbon (C-10) chain, thus it is unlikely to occupy the IPP binding site (C-5). In addition, a single fluorine substitution at the C2 vinylic position should depress the rate of cation formation but not as severely as the CF_3 analogs, ***E*-** and ***Z*-31**, thus allowing a more accurate measurement of the kinetic parameter.[107] These predictions were indeed confirmed by the experimental results. The K_M value for the analog was 1.1 μM, which is similar to that for GPP (0.7 μM).[107] Similar to the experiment mentioned above, the reactivity of the fluorinated GPP **34** was evaluated by calculating the k^{rel} value and this value was compared with the analogous k^{rel} value measured in the non-enzymatic solvolysis of the methanesulfonate derivatives (**35**) and its fluorinated analog (**36**). The k^{rel} value of compound **34** in the enzymatic reaction was 0.84×10^{-3}, while the k^{rel} value of compound **36** in the solvolysis was 4.4×10^{-3}.[106, 107] The similarity between these values implied that *trans*-FPP synthase employs the S_N1 pathway during catalysis.



34 **35** **36**

Figure 3.6 2-Fluorogeranyl diphosphate (34), geranyl methanesulfonate (35) and 2-fluorogeranyl methanesulfonate (36)

To complete the studies of the substrate analog experiments, an additional set of GPP analogs that carried either a CH_2F (**37**), a CHF_2 (**38**) or a CF_3 group (**39**) at the C-3 position were tested as alternative substrates.[108] Similar to the previous experiments, the k^{rel} values of analogs **37**–**39** in the enzymatic reactions were comparable to the k^{rel} values of the corresponding methanesulfonate derivatives **40**–**42** in the non-enzymatic solvolysis (Figure 3.7). To gain more insight into the catalytic mechanism, a Hammet plot was constructed by plotting the -log k^{rel} values for the *trans*-FPP synthase-catalyzed reactions against those of the non-enzymatic solvolysis. The slope extrapolated from the Hammett plot was ~0.77, which is lower than the threshold value for a reaction involving 'anchimeric assistance' (Figure 3.4). With respect to the reaction catalyzed by *trans*-FPP synthase, the 'anchimeric assistance' refers to the interaction between the double bond of IPP and the C-1 of DMAPP during the release of pyrophosphate.[108, 109] Hence, the authors concluded that pyrophosphate dissociation and the subsequent condensation step by IPP are distinct in the enzymatic reaction, and an S_N1 pathway is operative in *trans*-FPP catalysis.[108]



R = CH_2F 3-CH_2F-GPP **37**
R = CHF_2 3'-CHF_2-GPP **38**
R = CF_3 3'-CF_3-GPP **39**

R = CH_2F 3-CH_2F-G-Ms **40**
R = CHF_2 3-CHF_2-G-Ms **41**
R = CF_3 3-CF_3-G-Ms **42**

Analogs	k^{rel} (FPP synthase)	k^{rel} (solvolysis)
GPP and **G-Ms** (**35**)	1	1
Compounds **37** and **40**	3.7×10^{-4}	7.7×10^{-4}
Compounds **38** and **41**	4.0×10^{-8}	2.2×10^{-6}
Compounds **39** and **42**	7.7×10^{-9}	4.0×10^{-7}

Figure 3.7 Fluorinated geranyl diphosphates (37–39), the corresponding methanesulfonate derivatives (40-42) and the corresponding Hammet plot of the *trans*-FPP synthase reaction. The graph was obtained from Mash *et. al.*[108]

The Poulter group also probed for the potential involvement of an X group in the S_N2 displacement pathway, which had been proposed by Popják and Cornforth (Figure 3.4). A fluorinated analog of the nucleophile IPP (**43**), was proposed to be a potential irreversible inhibitor.[110] According to the S_N2 displacement reaction, the IPP analog **43** should be attacked by the **X** group of the enzyme to yield a covalently-linked fluorinated intermediate (Figure 3.4 and 3.8). The basic residue that was originally proposed to deprotonate the 2-H of this intermediate might act as a nucleophile and add to the C-2 position. Upon release of the fluoride ion, a stable molecule would be attached to the enzyme via two covalent bonds (inactivated enzyme **44**) and render the enzyme inactive.[100, 110] Alternatively, if the S_N1 mechanism is operative, compound **43** might serve as a slow reacting substrate. It is noteworthy to mention that compound **43** was used as a racemic mixture, thus the authors also proposed that only 50% of the material could serve as substrate because the enzyme could only deprotonate the (*S*)-enantiomer. As it turned out, IPP analog **43** was an alternative substrate that was converted to the corresponding FPP analog, and the enzyme

retained most of its activity even after prolonged incubation. Due to the lack of inactivation by formation of **44**, the possibility of the S_N2 displacement pathway was disfavored. Nonetheless, the fluorinated IPP analog **43** also served as a competitive inhibitor, because the fluorine atom greatly weakened the nucleophilicity.[110]

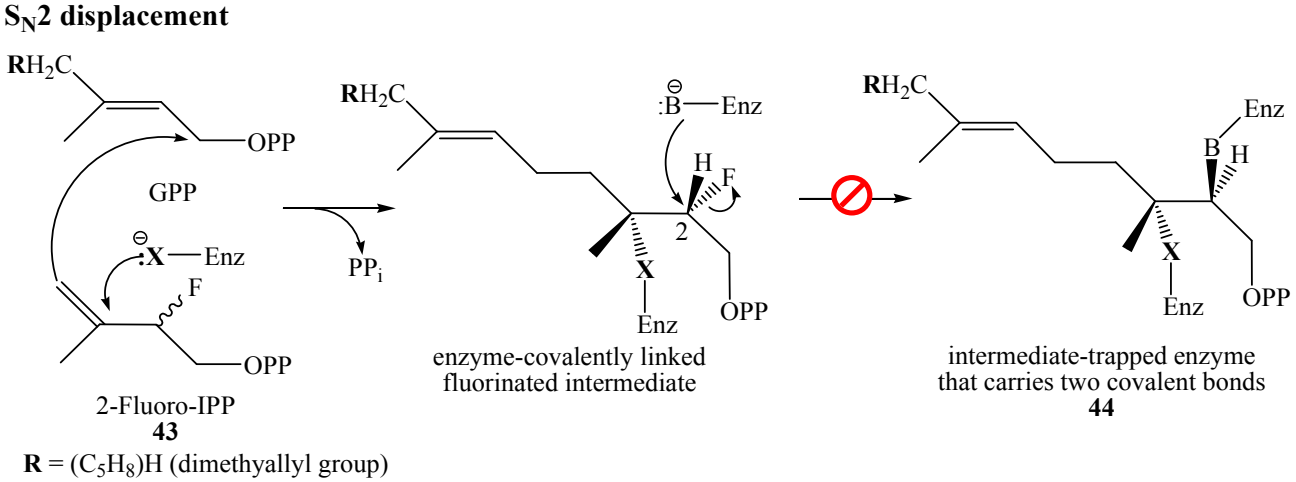


Figure 3.8 Potential inactivation of *trans*-FPP synthase by fluorinated analog of IPP, 2-fluoro-IPP (43)

An alternative approach to probe for the formation of the allyl carbocation was to test for positional isotope exchange (PIX) that might occur when using the isotopically labeled bridging-[1-^{18}O]-GPP (**45**).[108] This compound carries an ^{18}O-label at the bridging position between the hydrocarbon and the diphosphate moieties (Figure 3.9). It was proposed that C1-^{18}O bond of **45** would dissociate in the enzyme active site and give the corresponding geranyl carbocation/^{18}O-pyrophosphate ion pair. At this point, rotation of the P-O bond between the phosphate moieties and re-association with the carbocation could yield the starting material GPP that carries the ^{18}O isotope at the non-bridging position. Formation of this non-bridging ^{18}O-labeled GPP in the recovered starting material would provide direct evidence for the formation of the allyl carbocation, thereby supporting the S_N1 pathway. To observe isotopic scrambling, formation of the allyl carbocation needs to be reversible and its lifetime must be long enough to allow torsional equilibration of the

^{18}O-labeled pyrophosphate. In order to maximize the possibility that these conditions would exist, 2-fluoro-IPP (**43**) was employed in this experiment. This should raise the activation energy of the nucleophilic alkylation, thereby increasing the lifetime of the allyl cation/pyrophosphate ion pair and the reversibility of its formation.[108, 110] Unfortunately, isotopic scrambling in compound **45** was not observed. A likely explanation was that pyrophosphate was prevented from rotating by its chelation to the Mg^{2+} ion cofactor in the enzyme active site, thus prohibiting the possibility of PIX.[108]

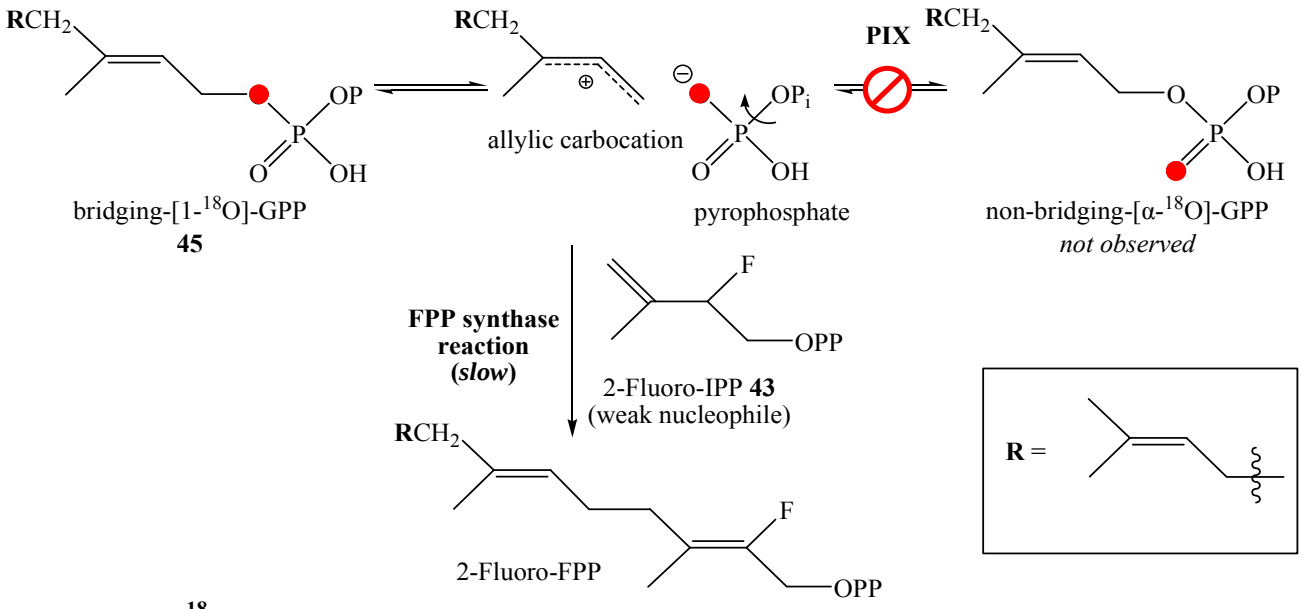


Figure 3.9 [1-^{18}O]-Geranyl diphosphate (45) and the possibility of positional isotope exchange in the *trans*-FPP synthase reaction. Solid red circle represents the ^{18}O-isotope.

In addition to the mechanistic studies, the structures of *trans*-FPP synthases isolated from different species have been widely investigated in X-ray crystallographic studies.[95] The first reported crystal structure among these studies was that of the avian enzyme solved to a resolution at 2.6 Å (Figure 3.10).[95, 111] The enzyme possesses 13 α-helices, 10 of which form a large central cavity containing the active site. There are two conserved DDxxD motifs located within the cavity that are 12 Å apart from each other and create the substrate-binding pockets for IPP and DMAPP/GPP. Additionally, the aspartate motifs are chelated to a divalent cation cofactor such as Mg^{2+}, which in turn binds to the diphosphate moieties of the substrates. Mutation of these aspartate

residues to alanine significantly reduced the k_{cat} values by 4-10 orders of magnitude.[111] Interestingly, the aromatic residue Phe112 that is located about 12 Å from the GPP-binding DDxxD motif was proposed to dictate the chain length of the product. Mutation of the Phe112 residue to a serine residue relieved the steric hindrance in binding a larger isoprenyl diphosphate; this mutant was consequently transformed into a C-25 geranylfarnesyl diphosphate synthase.[112]

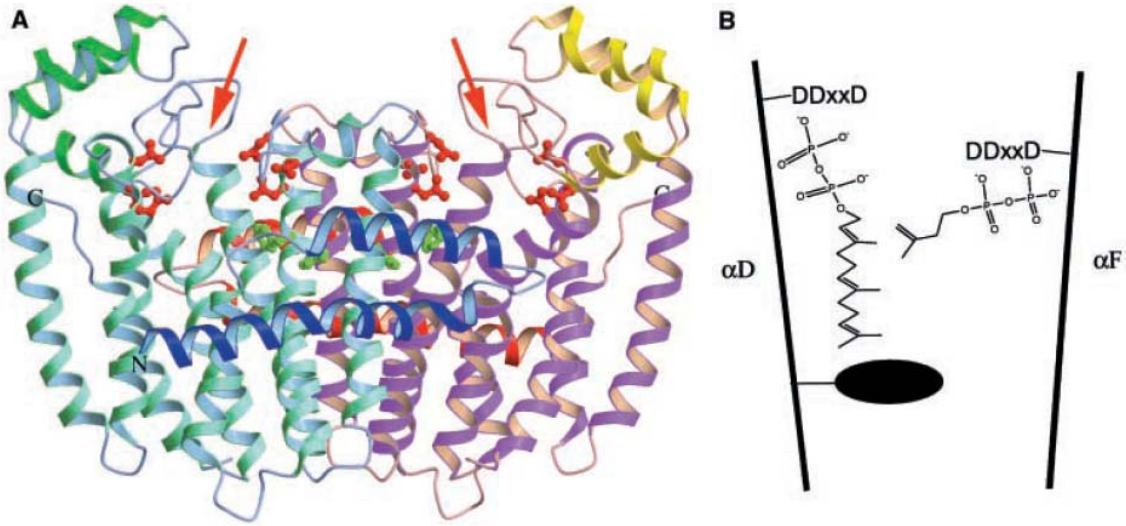


Figure 3.10 Crystal structure of *trans*-FPP synthase and graphic representation of substrate binding. A) refers to the model of the enzyme and orange arrows refer to the active site. B) is the graphic representation of the enzyme, and the black circle represents Phe112. Diagram obtained from the crystallographic report on the enzyme.[95]

3.3 Protein Prenyltransferases: Protein Farnesyltransferase

Protein prenyltransferases constitutes the second family of PTs, and they catalyze the addition of an isoprenyl diphosphate to a cysteine residue that is located within the conserved CaaX motif of a protein (C is the conserved cysteine reside, a's are usually aliphatic residues and X is frequently M, Q or S). Protein prenylation is an important post-translational modification for a number of proteins, including Ras and G protein families that are vital in a series of cell signaling systems (Figure 3.11).[95] Specifically, the prenylated protein can be anchored to the cell membrane

via the farnesyl hydrocarbon chain, thereby localizing them and controlling their interactions with other proteins at a subcellular level.[113] The activities of all protein PTs require a Zn^{2+} ion cofactor that coordinates with the cysteine thiolate. However, these enzymes do not contain the conserved DDxxD motif for Mg^{2+}-diphosphate binding that is found in *trans*-FPP synthase. In fact, there is no sequence homology between the genes encoding for any of the isoprenyl diphosphate synthases and the protein PTs, suggesting an alternative mechanism for catalysis. Most studies have focused on the mammalian protein farnesyltransferase, because an oncogenic form of the farnesylated Ras protein is found in nearly 30% of human cancers. Since the oncogenic Ras protein requires post-translational farnesylation for full cellular function, inhibition of protein farnesyltransferase has become a potential strategy for anticancer therapy.[114]

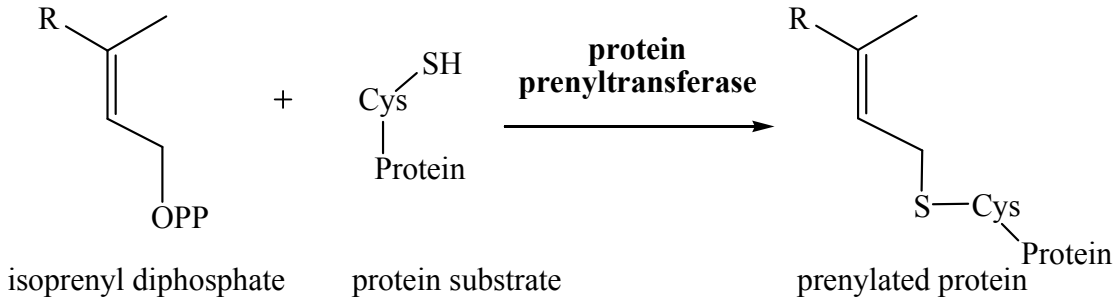
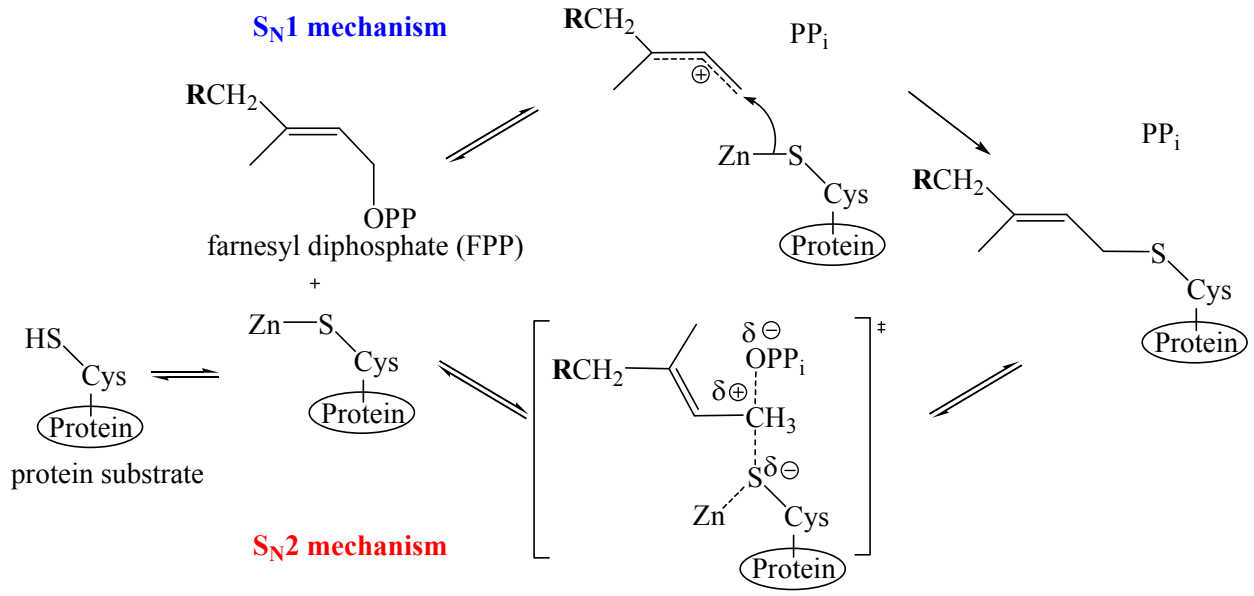


Figure 3.11 The reaction catalyzed by protein prenyltransferases

Protein farnesyltransferase catalyzes the nucleophilic substitution reaction between FPP and the cysteine residue of the substrate protein. Two distinct mechanistic pathways, S_N1 and S_N2, have been suggested for catalysis (Figure 3.12).[113] Both pathways begin with the deprotonation of the cysteine residue, followed by coordination of the thiolate to the Zn^{2+} cofactor in the enzyme active site. In the S_N1 pathway, pyrophosphate then dissociates from FPP to generate the corresponding allyl carbocation intermediate. To yield the farnesylated protein as product, the cysteine thiolate anion is added to carbocationic intermediate in a distinct chemical step. In contrast, the S_N2 mechanistic pathway involves a single transition state, in which the steps of pyrophosphate

dissociation and thiolate alkylation occur concurrently. Different approaches have been employed to elucidate the mechanism employed by protein farnesyltransferase, and it is now generally accepted to involve an S_N2-like transition state, where the hydrocarbon moiety of FPP contains a significant degree of carbocationic character.[113]



R = $(C_5H_8)_2H$ (geranyl group)

Figure 3.12 Proposed S_N1 and S_N2 mechanisms for the protein farnesyltransferase reaction

Mechanistic investigations on protein farnesyltransferase isolated from yeast were first performed in the laboratory of Poulter.[115] Later on, the mammalian (rat) homolog was also isolated and its mechanism was studied by the research group of Fierke.[113] This section will focus on the studies with the mammalian protein farnesyltransferase because the Fierke group has more fully characterized the enzyme. The CH_2F- and CF_3-FPP analogs (**46** and **47**, respectively) were tested as alternative substrates in the protein farnesyltransferase reaction (Figure 3.13). It was hypothesized that the fluorine substitution would destabilize the corresponding allyl carbocations and consequently reduce the reaction rates.[108, 113, 115] The relative catalytic rate (k^{rel}) of analogs **46** and **47** were calculated to be 1.8×10^{-2} and 2.6×10^{-4}, respectively, suggesting that significant

carbocationic character is generated in the allylic group of FPP during catalysis (Table 3.1).[113] However, the magnitudes of the rate depressions were less severe than those of the corresponding GPP analogs, **37** and **39**, in the *trans*-FPP synthase reaction and of the methanesulfonate derivatives, **40** and **42**, in the non-enzymatic hydrolysis (Table 3.1). Both of these latter reactions are believed to proceed through an S_N1 pathway. When the non-enzymatic solvolysis of **40** and **42** was carried out in the presence of a strong nucleophile azide, the measured k^{rel} values matched more closely to the k^{rel} of the fluorinated analogs **46** and **47** in the protein farnesylation reaction.[115] Hence, it was believed that protein farnesyltransferase employs an S_N2 pathway for catalysis, but the associative transition state was considered to be 'exploded' and carries a considerable carbocationic character at the allylic position.[116, 117]

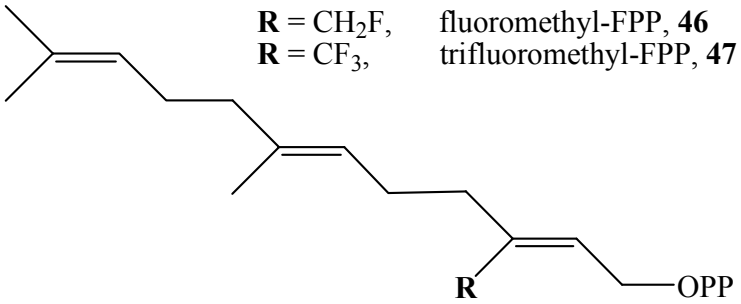


Figure 3.13 Fluorinated analogs of farnesyl diphosphate, fluoromethyl-FPP (46) and trifluoromethyl-FPP (47)

Table 3.1 Effects of fluorine substitution on the relative reaction rates (k^{rel})

Reactants	k^{rel} of FPP, GPP and their Fluorinated Analogs			
	protein farnesyltransferase	***trans*-FPP synthase**	**solvolysis in water**	**solvolysis in aqueous azide**
CH_3	1 (FPP)	1 (GPP)	1 (G-Ms)	1 (G-Ms)
CH_2F	1.8×10^{-2} (**46**)	3.7×10^{-4} (**37**)	7.7×10^{-4} (**40**)	6.1×10^{-2} (**40**)
CF_3	2.6×10^{-4} (**47**)	7.7×10^{-9} (**39**)	4.0×10^{-7} (**42**)	6.5×10^{-3} (**42**)

To further analyze the protein farnesyltransferase reaction, a deuterium-labeled substrate, [1,1-$^2H_2$]-FPP (**48**), was employed in the measurement of a kinetic isotope effect (KIE) by the Distefano research group (Figure 3.14). In the S_N1 pathway, the release of pyrophosphate from FPP was predicted to be rate-determining.[118] Hence, a large secondary isotope effect (~1.1–1.4) would be anticipated, because the C-1 position of FPP experiences a change in hybridization from sp^3 to sp^2 (Figure 3.14).[118] However, an isotope effect near unity (~1) was observed in the experiment, implying that deuterium substitution caused no significant change in the barrier to the rate-limiting step of the reaction.[118] Similar observations have been made previously for a variety of concerted nucleophilic substitutions, thus the authors concluded that the KIE result agreed more closely with the S_N2 mechanistic pathway.[118, 119]

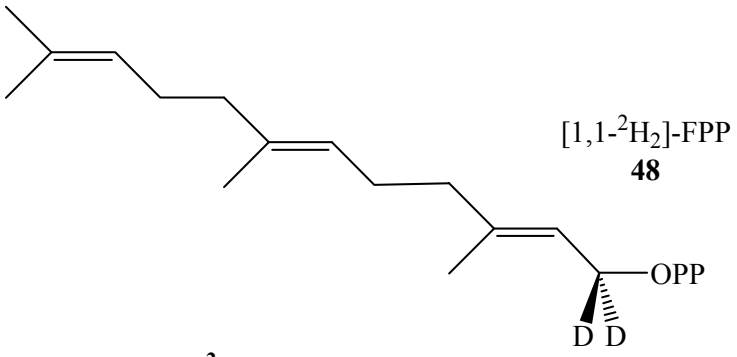


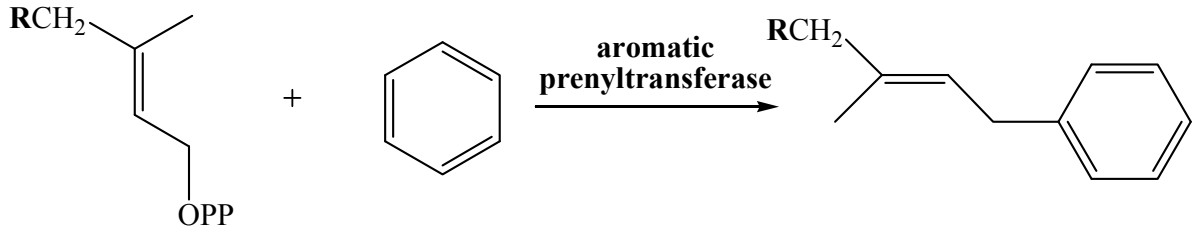
Figure 3.14 Deuterium labeled [1,1-$^2H_2$]-farnesyl diphosphate (48)

The protein farnesyltransferase reaction was further characterized by incorporating different metals into the active site.[120] The replacement of the Zn^{2+} cofactor by a stronger thiolate-chelating metal cation, cadmium Cd^{2+}, was expected to decrease both the pK_a and the nucleophilicity of the thiolate. This hypothesis was confirmed by the fact that the K_M value for peptide binding was increased by ~5 fold, while the k_{cat} value decreased by ~6 fold.[113] Lastly, the pH-rate profile data showed that the peptide substrate is bound as a thiolate anion in the active site, and the reaction

proceeded with an inversion of stereoconfiguration at the C-1 of FPP. Although these studies do not provide direct evidence in support of the mechanism, all of them agree well with the proposed S_N2 pathway (Figure 3.12).[121, 122]

3.4 Small-molecule Aromatic Prenyltransferases

The last category of PTs, namely the small-molecule aromatic prenyltransferases, use a variety of aromatics as substrates for prenylation (Figure 3.15). These enzymes can be further subcategorized into the membrane-associated aromatic PTs and the soluble aromatic PTs.[96] The former has been identified as playing a role in the biosynthesis of membrane lipids in prokaryotes and archaea and contain the same DDxxD motif for Mg^{2+}-isopentenyl diphosphate binding that is observed in *trans*-FPP synthase.[123, 124] The latter group has been identified from different bacteria and fungi species but no sequence homology with other known family of PTs was discovered. The lack of the DDxxD motif in the soluble aromatic PTs suggested that the mechanism of substrate binding is likely to be different from that of *trans*-FPP synthase. Interestingly, almost all of the soluble aromatic PTs (except NphB, see below) can catalyze their reactions in the absence of a divalent metal cation. The crystal structure of these enzymes has a characteristic antiparallel β/α-barrel peptide (α-β-β-α) arrangement that is distinctively different from the tertiary arrangement of the previously studied PTs.[95, 96, 111, 112, 125, 126] Based on this information, soluble aromatic PTs represent a new class of enzyme on which mechanistic studies are limited.



allylic diphosphate aromatic substrate allylic aromatic substrate

R = $(C_5H_8)_nH$

Figure 3.15 Reaction catalyzed by aromatic prenyltransferases

3.4.1 Bacterial Aromatic Prenyltransferases

Based on their amino acid sequences and origins, the currently discovered soluble aromatic PTs can be further sub-classified into two remotely related subgroups, the bacterial aromatic PTs and the fungal indole PTs. The bacterial soluble aromatic PTs have recently been discovered in several species including the Gram positive soil bacteria *Streptomyces* sp. and the marine bacteria *Salinispora arenicola*.[96, 127] They contain a unique protein architecture that has an α-β-β-α rearrangement, and are thus entitled the ABBA family and/or the PT barrel enzyme. Several ABBA enzymes have been identified and characterized,[96] including NphB, CloQ and NovQ that are found in the biosynthesis of the anti-oxidant naphterpin and the antibiotics clorobiocin and novobiocin, respectively (Figure 3.16).[125, 128, 129]

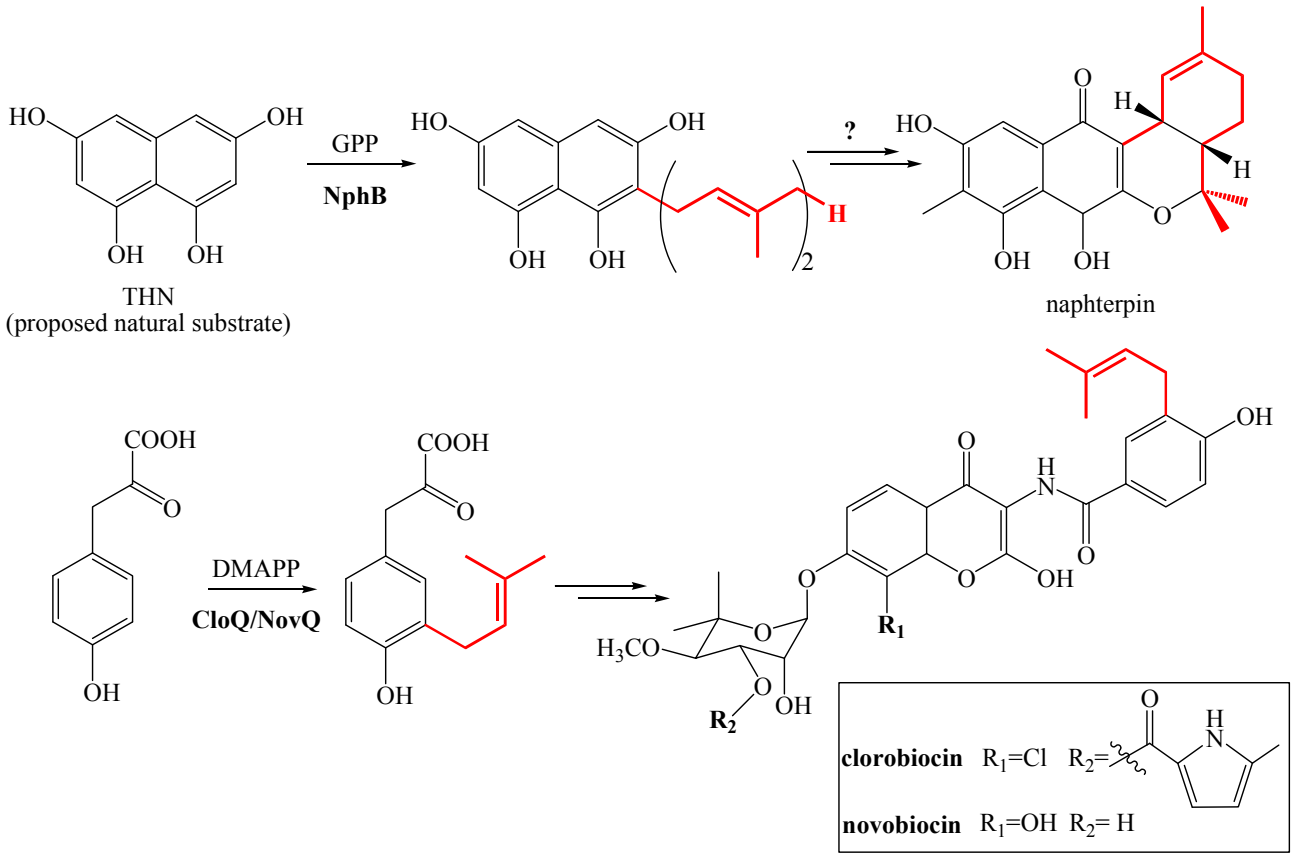


Figure 3.16 Participation of aromatic prenyltransferases, NphB, CloQ and NovQ, in natural product biosynthesis. Red denotes the chemical bonds derived from the prenyl groups.

The first bacterial aromatic PT discovered was NphB and its biochemical characterization has been conducted. The true physiological substrates for this enzyme remain somewhat obscure, since it is able to catalyze GPP or FPP prenylations on a large variety of aromatics *in vitro* (Figure 3.16).[125] However, NphB has been proposed to play to a role in the biosynthesis of the anti-oxidant naphterpin, and it catalyzes the addition of the carbon backbone of GPP onto the aromatic substrate, 1,3,6,8-tetrahydroxynaphthalene (THN). Interestingly, NphB is the only known soluble aromatic PT that requires a Mg^{2+} cofactor for catalysis, even though all of these enzymes share a certain extent of sequence homology. The crystal structure of NphB complexed with 1,6-dihydroxynaphthalene (1,6-DHN) and a non-reactive thiol GPP analog **49** was obtained and revealed the characteristic ABBA arrangement (or PT barrel) of the enzyme (Figure 3.17). The structure of NphB is reminiscent of the

classical structure of triosephosphate isomerase (TIM barrel), because both consist of a β-barrel that is wrapped by a series of α-helices. However, the cavity in the PT barrel is formed by antiparallel β-strands, instead of the parallel β-strands in the TIM barrel. Hence, the connectivity of the secondary structure elements in the PT barrel are indeed significantly different from that of the TIM barrels.[96, 125]

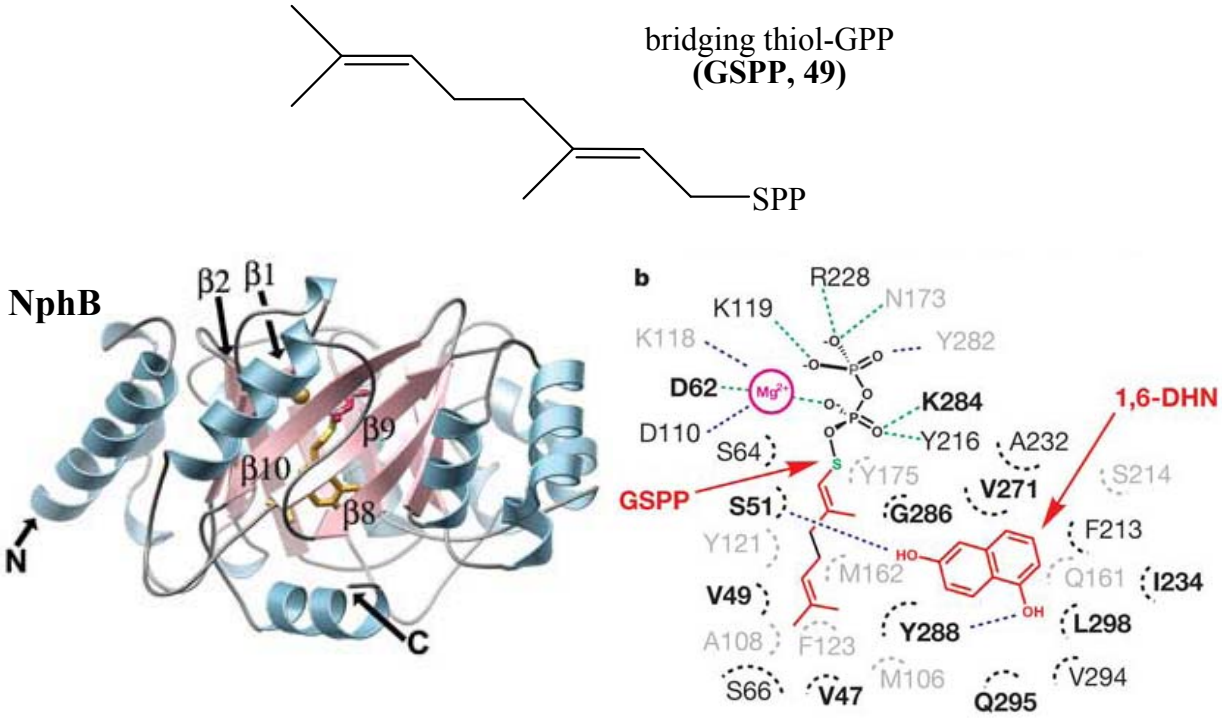


Figure 3.17 Thiol analog of GPP (49), crystal structure of NphB and model of active site. Diagram was obtained from the crystallographic study of NphB.[125]

Two mechanistic possibilities have been proposed for the NphB reaction based on the crystallographic studies.[125] In the S_N1 pathway, a discrete geranyl carbocationic intermediate was formed as a result of dissociation of the pyrophosphate from GPP (Figure 3.18). Subsequently, THN attacks the carbocation and yields a keto-intermediate. Finally, deprotonation leads to re-aromatization of the prenylated aromatic. In the S_N2 pathway, the release of pyrophosphate and the nucleophilic attack by THN are combined into one transition state. No direct study has been conducted to investigate the catalytic mechanism, however the crystallographic study showed that

the distance between the nucleophilic carbon and the C-1 of the thiol-analog **49** is about 4 Å. Due to the long distance between these two carbons, the authors suggested a dissociative reaction where the pyrophosphate is first released from GPP and the resultant allyl carbocation is then attacked by THN.[96, 125, 128]

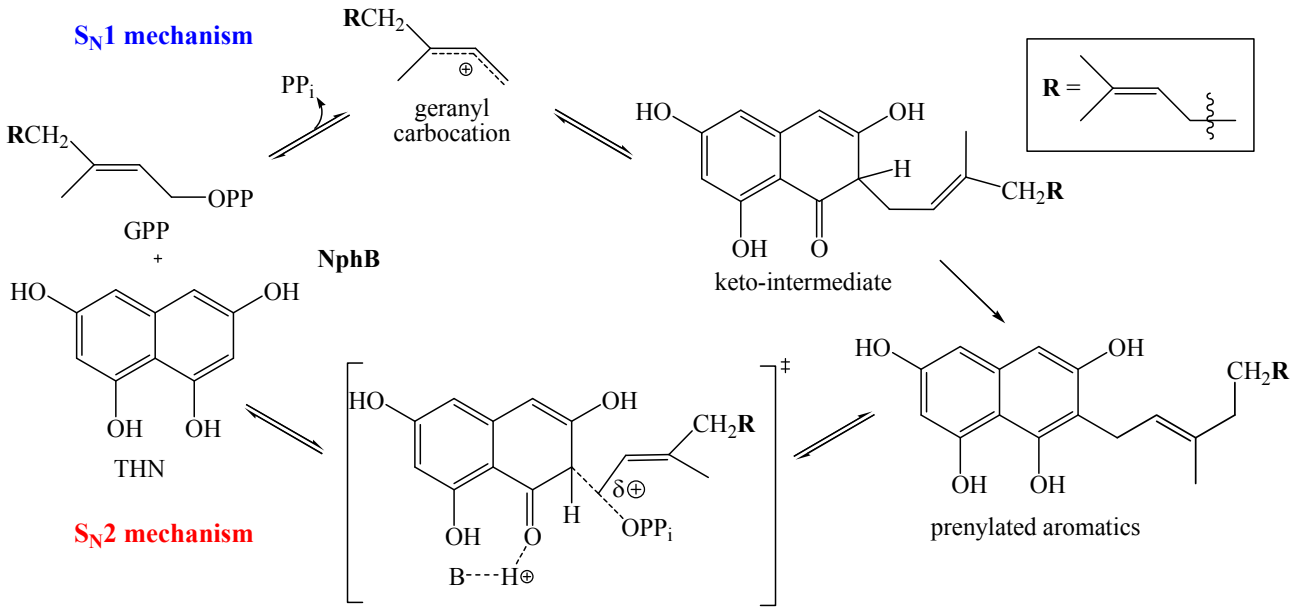


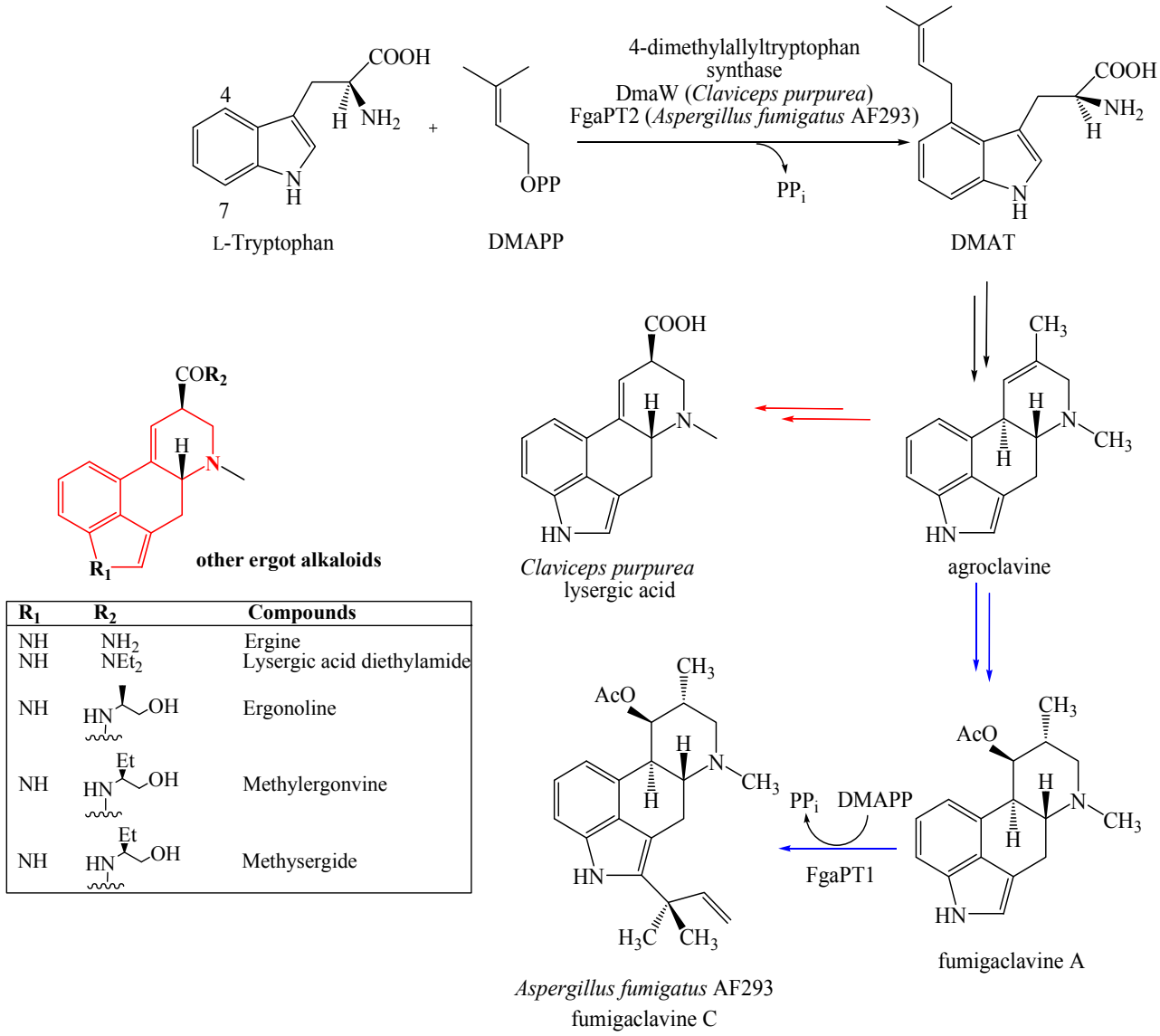
Figure 3.18 The proposed S_N1 and S_N2 mechanism for the NphB reaction

3.4.2 Fungal Indole Prenyltransferases: Dimethylallyltryptophan Synthase

The fungal indole PTs are another subgroup of soluble aromatic PTs and are remotely related to the bacterial soluble aromatic PTs (~15-20% identity). All of these fungal aromatic PTs add a prenyl group to substrates carrying an indole ring-system, and they have been discovered in different fungal species.[96, 130, 131] For instance, 4-dimethylallyltryptophan synthase (DmaW) was isolated from *Claviceps purpurea* and catalyzes the first step in the biosynthesis of lysergic acid, a precursor to the synthesis of the recreational drug, lysergic acid diethylamide (LSD) (Figure 3.19).[16, 132] Other examples include fumigaclavin C synthase (FgaPT1) and another 4-dimethylallyltryptophan

synthase (FgaPT2) that were discovered in *Aspergillus fumigatus*. Both of them are involved in the biosynthesis of fumigaclavin C, which was first identified as a toxin but was later found to possess useful properties for cardiovascular protection and weight loss control.[130, 131, 133, 134] Another paralog, 7-dimethylallyltryptophan synthase, has also been identified from *A. fumigatus*, but its role in nature remains undetermined.[135] In addition to DmaW and FgaPT2, a third 4-dimethylallyltryptophan synthase, MaPT, has been cloned from the fungal species *Malbranchea aurentiaca*.[136] All of these enzymes have high sequence identities with one another (≥35%) and do not require divalent metal ions for catalysis, although the reaction rates are mildly enhanced by them.[128–131, 135, 136]

Many fungal indole PTs have been proposed to play crucial roles in the biosynthesis of ergot alkaloids, all of which carry an ergoline backbone in their structures (Figure 3.19). Because many of the ergot alkaloids possess potent biological activities, these compounds are actually a popular area of research in natural products chemistry. Indeed, many ergot alkaloids have been exploited for different purposes due to their potent biological activities. For instance, ergine and LSD are psychedelic drugs used for recreational purposes.[132] Ergonovine, methergine and methylergonovine are uterocontractants employed during child labor and/or elective abortion.[137] Methysergide is a prescriptional medicine for migraine treatment.[138] Research focused on the biosynthesis of ergot alkaloids is ongoing, but to-date only a few enzymes have been identified and cloned.[139]



R_1	R_2	Compounds
NH	NH_2	Ergine
NH	NEt_2	Lysergic acid diethylamide
NH		Ergonoline
NH		Methylergonvine
NH		Methysergide

Figure 3.19 Reactions catalyzed by different fungal indole prenyltransferases. Red chemical bonds denote the ergoline backbone. Red arrows indicate the biosynthesis of lysergic acid in *Claviceps purpurea*, and blue arrows indicate the biosynthesis of fumigaclavine C in *Aspergillus fumigatus* AF293.

4-Dimethylallyltryptophan synthase (including DmaW and FgaPT2, and generalized as DMATS hereafter) is the most well-known fungal indole PT among all that have been discovered. It catalyzes a Friedel-Crafts-type reaction by transferring the allyl moiety of DMAPP to the C-4 of L-tryptophan. DMATS has been found in several fungal species in the genera of *Claviceps, Penicillium,* and *Aspergillus*. Furthermore, this enzyme was proposed to catalyze the first committed step in the biosynthetic pathway of ergot alkaloids because the reaction product,

dimethylallyltryptophan (DMAT), provides the basic carbon skeleton for building the ergoline backbone (Figure 3.19). The biochemical properties of DMATS have been investigated by several research groups using different producer fungal strains.[16, 100, 101, 103]

Initially, the Floss group investigated the stereochemistry of the reaction by feeding isotopically labeled substrate to *Claviceps sp.* SD58 (Figure 3.20).[16] When deuterium-labeled DMAPP, (*Z*)-[3-$^2H_3$]-methyl-2-butenyl diphosphate (**50**), was used in the DMATS reaction, the deuterated methyl group did not experience any positional exchange in the product. Additionally, the C-1 of the stereospecifically labeled (*S*)-[1-^{2}H]-DMAPP (**51**) experienced an inversion of stereoconfiguration during catalysis.[16] Although these studies did not reveal much about the catalytic mechanism, they demonstrated that DMATS catalyzes a highly stereospecific reaction.

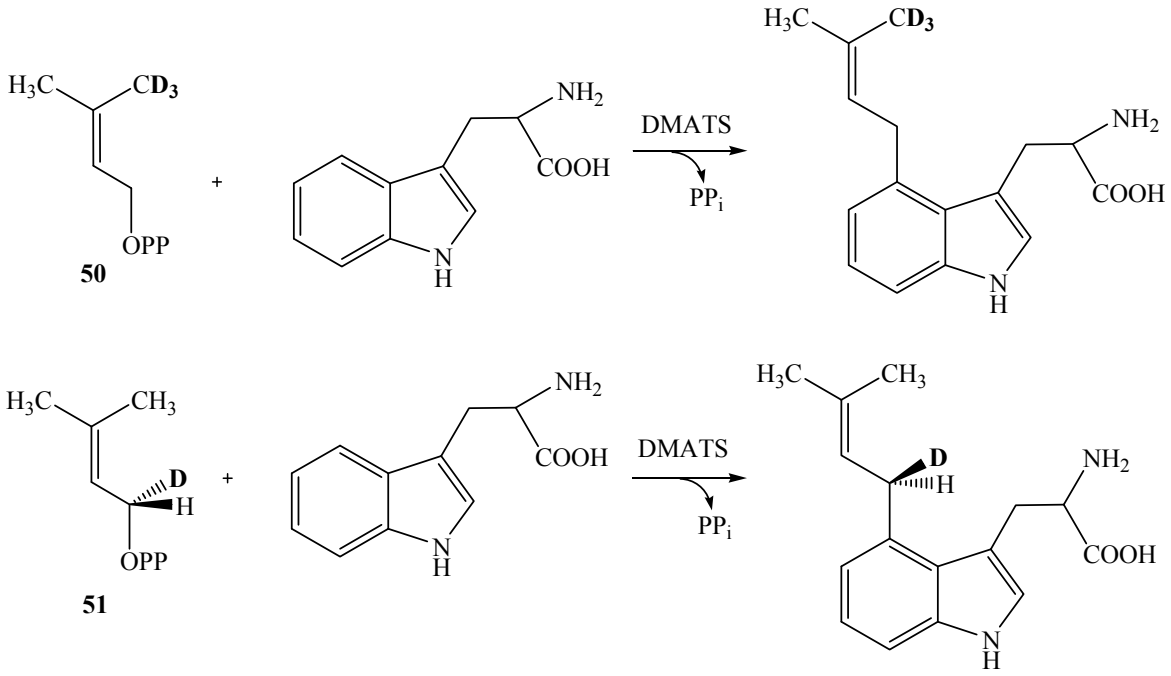


Figure 3.20 Investigations on the stereochemistry of the DMATS reaction

In addition to the labeling studies, the Floss group proposed two possible reaction pathways, involving either S_N1 or S_N2 reaction pathway, for the electrophilic aromatic substitution reaction catalyzed by DMATS (Figure 3.21). In the S_N1 pathway, the reaction begins with the formation of

the pyrophosphate/dimethylallyl carbocation ion pair. Subsequently, the C-4 position of L-tryptophan adds to the carbocation intermediate to yield an arenium intermediate, and a final deprotonation at H-4 causes re-aromatization and forms DMAT.[15, 16] The same arenium ion intermediate is involved in the S_N2 reaction as well, however, it is generated via an associative transition state, where the nucleophilic attack of the aromatic ring onto DMAPP is synchronized with the dissociation of pyrophosphate.[16]

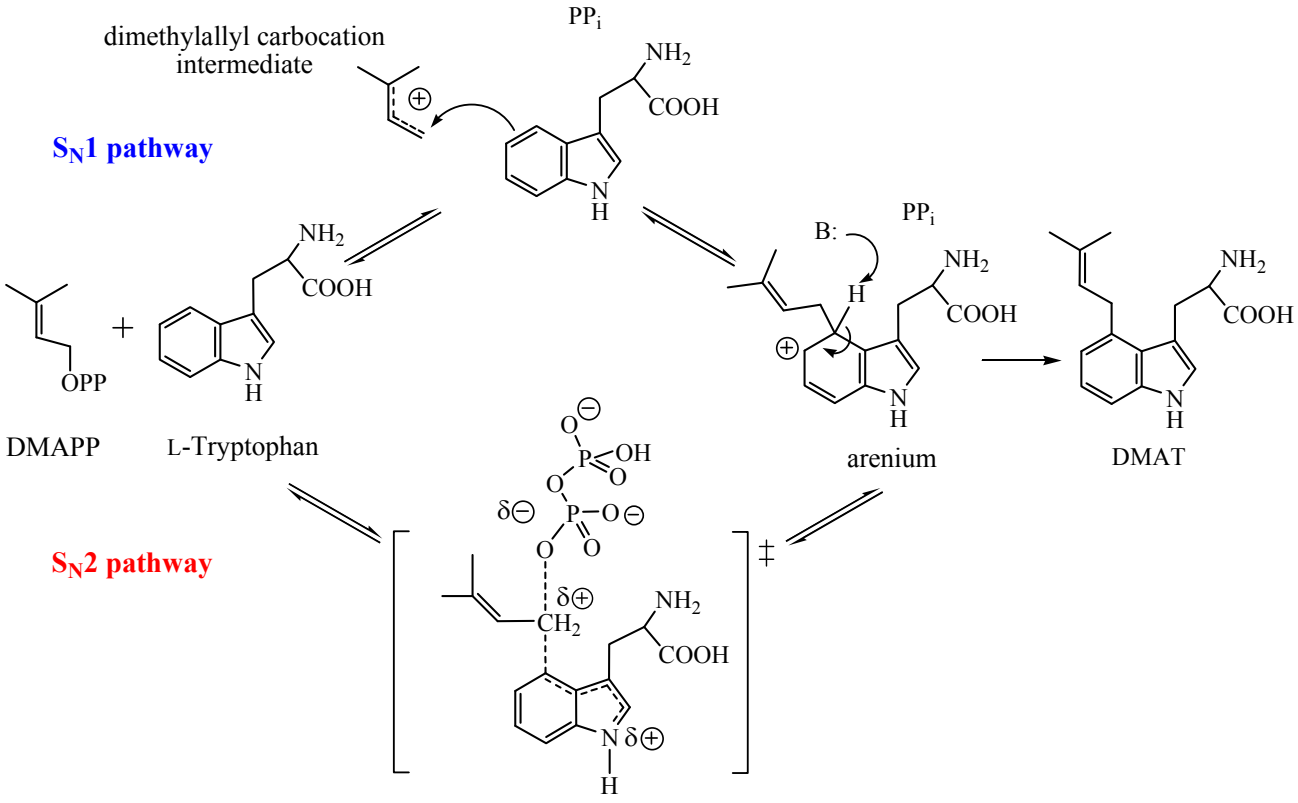


Figure 3.21 Proposed S_N1 and S_N2 mechanistic pathways for the DMATS reaction

Elucidating the mechanistic pathway employed by DMATS is particularly interesting from the perspective of a mechanistic enzymologist. In the reaction catalyzed by *trans*-FPP synthase, the nucleophilicity of the double bond is relatively weak, thus it is reasonable to expect an S_N1 reaction pathway is employed since it generates a relatively reactive allyl carbocation (Figure 3.4). On the other hand, in the reaction catalyzed by protein farnesyltransferase, the nucleophilicity of thiolate is

considerably greater, thus one would predict that an S_N2 reaction is employed (Figure 3.12). Since the nucleophilicity of an indole is greater than that of an alkene but less than that of a thiolate,[140] the possibility of either reaction mechanism exists in the DMATS reaction. Although the Floss group has shown that the catalysis proceeds with an inversion of stereochemistry at the C-1 position of DMAPP, this observation does not eliminate either mechanism because the enzyme active site can control the orientation of the attack. Hence, further investigation is needed to reveal more insight into the mechanism of the enzyme-catalyzed reaction.

In order to further study the catalytic mechanism, an efficient method to generate a significant amount of pure DMATS would be needed, thus a vast amount of effort has been devoted to overexpression and purification of the enzyme.[132, 141, 142] Initially, both the laboratories of Rilling and Poulter have developed a system of isolating DMATS from different *Claviceps* sp. Unfortunately, both methods required multiple chromatographic procedures and provided only a low yield of protein.[132, 141] Later on, Schardl isolated the cDNA of the enzyme from *Claviceps fusiformis* and expressed it in the yeast *Saccharomyces cerevisiae*. However, purification of the DMATS in the heterologous system remained unsuccessful and only minimal enzymatic activity could be detected.[131, 142]

In spite of the problems in obtaining the purified enzyme, different types of mechanistic studies have been conducted on the reaction catalyzed by DMATS. Poulter and his co-workers have investigated the catalysis with alternative substrate analogs bearing various electron-donating/withdrawing groups.[15] The reaction rates were measured using the substrate L-tryptophan and its 7-substituted derivatives (**X** = OCH_3, CH_3, F, CF_3, and NO_2) (Figure 3.22). When a Hammett plot was constructed by plotting the logarithms of the k^{rel} values against the corresponding substituent constants σ^+, a linear correlation of ρ = -2.0 was observed. The negative value implied that attack by L-tryptophan is likely rate-determining.[15] However, the authors could not explain why

the data point for the natural substrate L-tryptophan does not lie on the linear correlation plot. Also, the K_M values of these analogs varied considerably (0.6 to 368 μM), suggesting that the binding affinities were affected by the substitutions.[15]

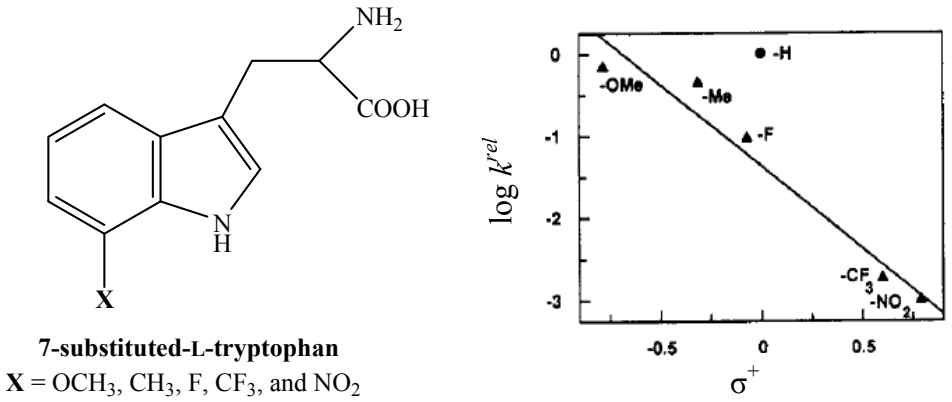
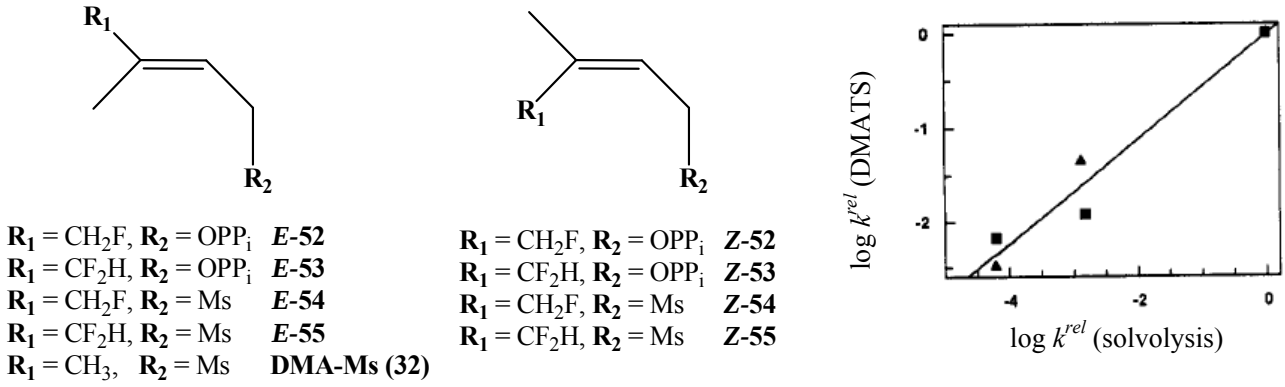


Figure 3.22 7-Substituted derivatives of L-tryptophan and the Hammett Plot for the DMATS reaction. Graph was obtained from the studies conducted by Gebler, *et. al.*[15]

The CH_2F- and CHF_2-analogs of DMAPP were also tested as alternative substrates and evaluated in the kinetic studies. Specifically, both the *E*- and *Z*-isomers of compounds **52** and **53** were employed.[15] The k^{rel} values of the CH_2F-analogs ***E*-** and ***Z*-52** were on the order of 10^{-2}, whereas the k^{rel} values of the CHF_2-analogs ***E*-** and ***Z*-53** were on the order of 10^{-3} (Figure 3.23 and Table 3.2). These observations suggested that a significant carbocationic character is created at the allylic carbon of DMAPP during catalysis. In addition, non-enzymatic solvolysis of the corresponding fluorinated methanesulfonate derivatives, **54** and **55** (both *E- and Z*-isomers for each compound) was carried out, and the values of k^{rel} were on the order of 10^{-3} and 10^{-5}, respectively (Table 3.2). A Hammett plot between the logarithms of the k^{rel} values of the enzymatic catalysis and the non-enzymatic solvolysis were plotted and gave a positive correlation (Figure 3.23). Based on these results, the authors concluded that the DMATS reaction proceeds through a 'highly electrophilic transition state,' where the dimethylallyl moiety possesses a significant carbocationic

character.[15] The authors never explored why the k^{rel} values observed in the enzymatic reaction were significantly higher than those in the solvolysis of the methanesulfonate derivatives. Indeed, this study has not clearly elucidated the catalytic mechanism.



R_1 = CH_2F, **R_2** = OPP_i ***E*-52**
R_1 = CF_2H, **R_2** = OPP_i ***E*-53**
R_1 = CH_2F, **R_2** = Ms ***E*-54**
R_1 = CF_2H, **R_2** = Ms ***E*-55**
R_1 = CH_3, **R_2** = Ms **DMA-Ms (32)**

R_1 = CH_2F, **R_2** = OPP_i ***Z*-52**
R_1 = CF_2H, **R_2** = OPP_i ***Z*-53**
R_1 = CH_2F, **R_2** = Ms ***Z*-54**
R_1 = CF_2H, **R_2** = Ms ***Z*-55**

Figure 3.23 Fluorinated derivatives of DMAPP (52 and 53), their methanesulfonate derivatives (54 and 55) and the corresponding Hammett plot for the DMATS reaction. Square and triangle represent the *Z*- and *E*-fluorinated DMAPP analogs and methanesulfonate derivates, respectively. Graph was obtained from the studies conducted by Gebler *et. al.*[15]

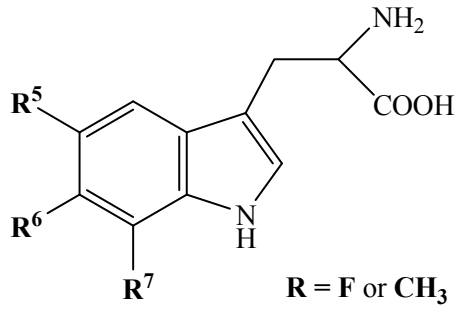
Table 3.2 The k^{rel} values of the DMATS reactions and the corresponding non-enzymatic solvolysis

Reactants	k^{rel}	
	DMATS	**solvolysis**
DMAPP and DMA-Ms	1	1
***Z*-52** and ***Z*-54**	4.4×10^{-2}	1.3×10^{-3}
***E*-52** and ***E*-54**	1.1×10^{-2}	1.5×10^{-3}
***Z*-53** and ***Z*-55**	3.8×10^{-3}	6.0×10^{-5}
***E*-53** and ***E*-55**	6.9×10^{-3}	6.1×10^{-5}

As mentioned in the previous section, difficulty in producing a sufficient amount of DMATS had become an obstacle to carrying out further mechanistic studies. Recently, the whole genome of an ergot alkaloid producing strain *Aspergillus fumigatus* AF293 was sequenced.[130, 131] From this data, Li and his co-workers identified a putative DMATS gene, namely *fgaPT2*, which encodes a protein with ~50% amino acid sequence identity to the DMATS (DmaW) isolated from *C. fusiformis*.[130, 131] Moreover, a third DMATS homolog, *maPT*, was isolated from the genome of

Malbranchea aurentiaca.[136] Fortunately, the cloning and heterologous expressions of these DMATS genes were successful in an *E. coli* system, thus creating an opportunity for further biochemical characterization of this enzyme.

The amino acid sequences of the newly discovered DMATS were compared with other remotely related bacterial aromatic PTs, such as NphB, CloQ and NovQ.[136] Several conserved residues, Asp179, Lys189 and Lys261 (based on the amino acid sequence of MaPT), were identified, and shown to be essential for activity by site-directed mutagenesis. It was proposed that Lys189 and Lys261 might be responsible for binding the diphosphate group of DMAPP.[136] The reactivity of different commercially available tryptophan analogs were tested in the DMATS catalysis.[130, 131] In particular, 5-, 6- or 7-methyl-D,L-tryptophan derivatives were converted to the corresponding DMAT analogs with relative yields ranging from 8 to 16% of that of the regular reaction using L-tryptophan (Figure 3.24). On the other hand, a fluorine substitution at any of these positions did not generate the corresponding product at any detectable level.[130] However, no mechanistic experiments have been performed on any of these newly identified DMATS.



R-substituted tryptophan

Figure 3.24 The analogs 5-, 6- or 7-methyl or fluorinated tryptophan tested in the DMATS reaction

3.5 Thesis Aims

Since the mechanism of 4-dimethylallyltryptophan synthase (DMATS) has not yet been determined, the second goal of this thesis is to provide experimental evidence that elucidates the mechanism by which the DMATS from *Aspergillus fumigatus* (*fgaPT2*) operates. In order to attain this goal, a series of experiments will be carried out that are designated to distinguish between an S_N1 and an S_N2 pathway. A more detailed discussion of these approaches as well as a proposal for further studies will be addressed in the final chapter.

Chapter 4

Mechanistic Studies on Dimethylallyltryptophan Synthase (DMATS)

A version of this Chapter has been published, and some figures are reproduced with permission from: Luk, L. Y. P.; Tanner, M. E. *J. Am. Chem. Soc.* **2009**, *131*, 13932-13933.

As described in Chapter 3, 4-dimethylallyltryptophan synthase (DMATS) plays a fundamental role in the biosynthesis of ergot alkaloids. There are two mechanistic possibilities that can account for the formation of DMAT. The S_N1 mechanism involves the formation of a discrete dimethylallyl carbocation intermediate, while the S_N2 mechanism involves a single associative transition state (Figure 3.21). Previously, studies that employed substrate analogs bearing electron-donating/withdrawing groups have been forwarded as evidence in support of an electrophilic reaction.[15] The results showed that electron-withdrawing groups on either substrate slowed the reaction considerably, as expected for either an S_N1 mechanism involving a carbocation intermediate, or an S_N2 mechanism involving an "exploded" transition state with considerable carbocation character. Unfortunately, the difficulty in obtaining DMAT from *Claviceps* sp. had prohibited additional studies on this enzyme.

Recently, the gene encoding for DMATS (*fgaPT2*) was cloned from another fungal species, *Aspergillus fumigatus* AF293, and was heterologously expressed in *E. coli* to give active enzyme by Li and his co-workers.[130, 131] Different substrate analogs were tested as alternative substrates for this enzyme, but the catalytic mechanism remained unelucidated. We envisioned that this was an opportunity to conduct mechanistic investigations on this newly identified DMATS. This Chapter will describe the gene synthesis and the expression of *fgaPT2* and unambiguously demonstrate the activity of recombinant DMATS. In addition, the characterization of the HPLC-purified reaction product by NMR spectroscopy and the development of a continuous steady-state kinetic assay by use of UV-Vis spectroscopy will be described. Lastly, the experimental results obtained from positional isotope exchange and kinetic isotope effect studies will be presented.

At approximately the same time as our mechanistic works were published, the crystal structure of DMATS was solved and reported.[126] A detailed discussion and comparison between

these two studies will be presented. Finally, our preliminary site-directed mutagenesis studies that are designed to probe the roles of active site residues will be described.

4.1 Purification of Dimethylallyltryptophan Synthase

4.1.1 Synthesis of the *fgaPT2* Gene

The gene encoding for DMATS, *fgaPT2*, was prepared by the method of gene synthesis at the Genscript® Corporation. Gene synthesis employs a combination of organic chemistry and molecular biology methods to conduct *de novo* production of a gene without the use of biological templates. In 1968, Khorana and his co-workers established the first gene synthesis by chemically synthesizing short oligonucleotides and then covalently linking them together by the enzymatic activity of a ligase.[143] This method often leads to defects in making a gene with a long nucleotide sequence (≥ ~2 kb), because errors are prone to be incorporated during oligonucleotide synthesis and ligases often do not work properly. Fortunately, several technological advances have led to a new gene synthesis protocol that can efficiently create polynucleotides with high fidelity. First, improved automated solid phase synthesis has become a reliable way to generate oligonucleotide of 40 to 50 nucleotide sequences (40-50mers).[144] Secondly, the invention of the polymerase chain reaction (PCR) by Kary Mullis has created an extremely easy protocol to produce polynucleotides that are a kilobase in length.[65] Lastly, it has been shown that a PCR reaction can be used to reconstitute a gene that has been digested into small fragments (10-50mers) by restriction enzymes.[145] By bringing all these advances together, an improved gene synthetic method has been established. A set of overlapping 40-50mers can be synthesized based on the target gene sequence (i.e. *fgaPT2* in this thesis), and then subjected to a PCR reaction (Figure 4.1). Appropriately spaced 40-50mers anneal

to one another during the reaction and form a tertiary complex, that is elongated by the action of a DNA polymerase. Eventually, the targeted gene is made and is further amplified/purified by running a final set of PCR reactions using primers that bind to the ends of the gene. This modified method can efficiently synthesize a gene with a longer nucleotide sequence (2 kb to 10 kb).[146–148]

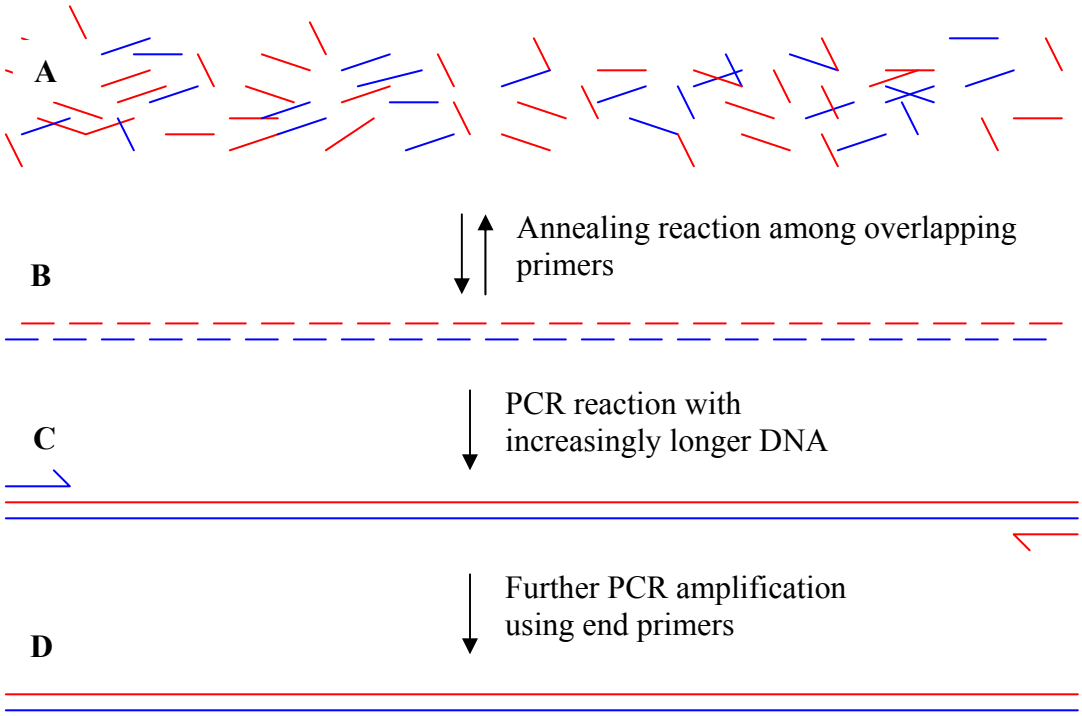


Figure 4.1 Pictorial representation of gene synthesis. A) shows oligomers (primers) of overlapping genes, B) primers annealing to one another, C) PCR reaction creates the gene after DNA synthesis, D) last set of PCR amplification with the use of end primers. The blue and red colors denote two opposite senses of DNA sequence within a gene. Diagram modified from Young *et. al.*[146]

Further modifications have also been made to increase the yield and the accuracy of the gene synthesis protocol. For example, a specialized approach called thermodynamically balanced inside-out (TBIO) PCR reaction begins with the DNA synthesis that corresponds to the center of a gene.[149] The oligomers are then added in a consecutive fashion and the double-stranded DNA elongates bidirectionally. This modified gene synthesis reaction has been shown to proceed with increased yield.[149] The sequence fidelity of the product can be improved by adding a DNA proof-reading

enzyme, MutS, which binds to mismatching nucleotide pairs that may form during a gene synthesis reaction thereby preventing the error-containing polynucleotide from further amplification. This modification was reported to reduce errors by 15-fold.[150] In this thesis, the gene *fgaPT2* was synthesized based on the cDNA sequence that was reported by Unsöld *et. al.*[131]

4.1.2 DMATS Protein Expression

The plasmid *fgaPT2*/pET28a containing the synthesized *fgaPT2* gene was prepared by Genscript®. The pET28a system contains several useful cloning features, including a sequence encoding for a C-terminal histidine tag for protein purification, a kanamycin resistance gene for selection of transformed bacteria and a T7 promoter/lac operon system for IPTG induction and protein expression.[68] Although Li and his co-workers demonstrated that expression of this gene in a regular *E. coli* strain gave a decent yield of DMATS (~23 mg per litre of bacterial culture),[131] modifications of the overexpression protocol were still made in an attempt to further increase the protein yield. As depicted in Figure 4.2, the *fgaPT2* gene carries 41 'rare' *E. coli* codons out of a total of 460 amino acids. Therefore, a regular *E. coli* 'codon dialect' might not be optimal for the expression of this gene (see Section 2.1.1 for discussion on codon bias).[131] Hence, the plasmid *fgaPT2*/pET28a was transformed into a specialized *E. coli* strain, Rosetta (DE3) pLysS. This expression strain was chosen, because it carries an additional plasmid called pRARE, which expresses extra copies of tRNA genes, *proL*, *leuW*, *metT*, *argW*, *thrT*, *glyT*, *tyrU*, *thrU*, *argU* and *ilex* to cope with nearly all of the rare codon translations.[151]

5’ – ATG AAG GCA GCC AAT GCC TCC AGT GCG GAG GCC TAT CGA GTT CTT AGT CGC GCC
TTT AGA TTC GAT AAT GAA GAT CAG AAG CTG TGG TGG CAC AGC ACT GCC CCG ATG TTT GCA
AAA ATG CTG GAA ACT GCC AAC TAC ACC ACA CCT TGT CAG TAT CAA TAC CTC ATC ACC TAT
AAG GAG TGC GTA ATT CCC AGT CTC GGA TGC TAT CCG ACC AAC AGC GCC CCC CGC TGG TTG
AGC ATC CTC ACT CGA TAC GGC ACT CCG TTC GAA TTG AGC CUA AAT TGC TCT AAT TCA AUA
GTG AGA TAC ACA TTC GAG CCG ATC AAT CAA CAT ACC GGA ACA GAT AAA GAC CCA TTC AAT
ACG CAC GCC ATC TGG GAG AGC CTG CAG CAC CTG CTT CCA CTG GAG AAG AGC ATT GAT CTG
GAG TGG TTC CGC CAC TTC AAG CAC GAT CTC ACC CTC AAC AGT GAA GAA TCT GCT TTT CTG
GCT CAT AAT GAT CGC CTC GTG GGC GGC ACT ATC AGG ACG CAG AAC AAG CTC GCG CTC GAT
CTG AAG GAT GGC CGC TTT GCA CTT AAG ACG TAC AUA TAC CCG GCT CTC AAA GCT GTC GTC
ACC GGC AAG ACA ATT CAT GAG TTG GTC TTT GGC TCA GTC CGC CGG CTG GCA GTG AGG GAG
CCC CGA ATC TTG CCC CCA CTC AAC ATG CTG GAG GAA TAC ATC CGA TCA CGC GGT TCC AAG
AGC ACT GCC AGT CCC CGC CUA GTG TCC TGT GAT CTG ACC AGT CCT GCC AAG TCG AGA ATC
AAG ATC TAC CTG CTG GAG CAG ATG GTT TCA CUA GAA GCC ATG GAG GAC CTG TGG ACT CTG
GGC GGA CGG CGC CGA GAC GCT TCC ACT TTA GAG GGG CTC TCT CTG GTG CGT GAG CTT TGG
GAT CTG ATC CAA CTG TCG CCG GGA TTG AAG TCC TAT CCG GCG CCG TAT CTG CCT CTC GGG
GTT ATC CCA GAC GAG AGG CTG CCG CTT ATG GCC AAT TTC ACC CTG CAC CAG AAT GAC CCG
GTC CCA GAG CCG CAA GTA TAT TTC ACA ACC TTC GGC ATG AAC GAC ATG GCG GTG GCG GAT
GCC CTG ACG ACG TTC TTC GAG CGC CGG GGT TGG AGT GAA ATG GCC CGC ACC TAC GAA ACT
ACT TTG AAG TCG TAC TAC CCC CAT GCG GAT CAT GAC AAA CTT AAC TAC CTC CAC GCC TAC
AUA TCC TTC TCC TAC AGG GAC CGT ACC CCT TAT CTG AGT GTC TAT CTT CAA TCC TTC GAG
ACA GGG GAC TGG GCA GTT GCA AAC TTA TCC GAA TCA AAG GTC AAG TGT CAG GAT GCG GCC
TGT CAA CCC ACA GCT TTA CCT CCA GAT CTG TCA AAG ACA GGG GTA TAT TAT TCC GGT CTC
CAC TGA –3’

Figure 4.2 Gene sequence of *fgaPT2* isolated from *Aspergillus fumigatus* AF293. 15 rare *E. coli* codons, arginine (CGA, CGG, AGG, AGA), isoleucine (AUA), leucine (CUA), threoline (ACG), proline (CCC) and glycine (GGA, GGG), were identified using online rare codon calculator.[64]

Further increase of protein yield was observed when Terrific Broth was used in the growth of the starter culture and when the bacteria were grown at 25 °C following IPTG induction at an optical density (OD_{600}) of 0.5 to 0.6. The *E. coli* cells were lysed in a French pressure cell, and the crude lysate was applied to a column of immobilized metal ion affinity resin which had been charged with a solution of $NiSO_4$. Eventually, the His-tagged DMATS was eluted using a step gradient of imidazole buffer. As shown in Figure 4.3, the isolated recombinant protein was >90% pure as judged by SDS-PAGE.

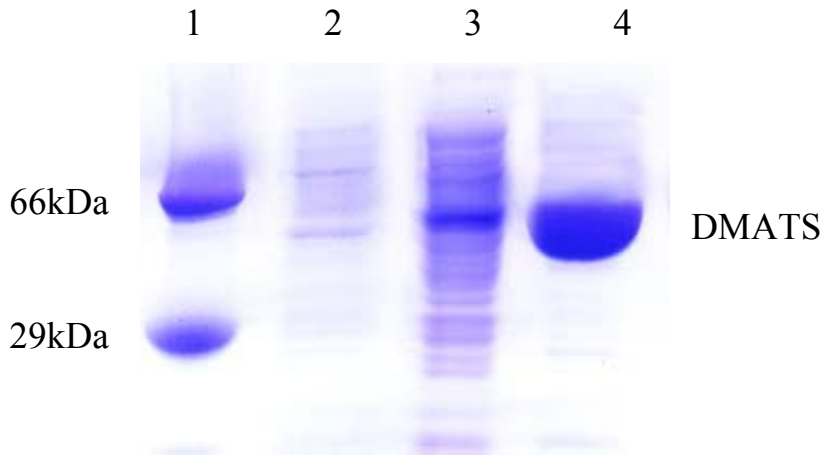


Figure 4.3 SDS-PAGE of the purified His-tagged DMATS. Lane 1 contains molecular mass standards of 66 kDa (bovin serum album) and 29 kDa (carbonic peptidase), lane 2 shows crude cell extract before isopropyl-β-D-1-thiogalactopyranoside induction, lane 3 shows crude cell extract after induction, and lane 4 shows the purified DMATS.

Initially, the pooled column fractions were centrifuged through a membrane (10 kDa MWCO) in an attempt to remove imidazole and to concentrate the protein. Unfortunately, this protocol led to a significant loss in total protein, which very likely adsorbed to the membrane. To get around this problem, the imidazole was instead removed by passing the enzyme through a Sephadex™ G-25 gel filtration column. In this fashion ~35 mg of the synthase was typically obtained from 1 litre of cell culture, and was found to be stable to storage at -80 °C for 6 months without significant loss of activity.

4.2 Characterization of Dimethylallyltryptophan Synthase Activity

4.2.1 Characterization of the Product 4-Dimethylallyltryptophan

The activity of DMATS was monitored by HPLC and NMR spectroscopic analysis. The substrates L-tryptophan and dimethylallyl diphosphate (DMAPP) (10 mM each) were incubated with the synthase, and a ^{31}P NMR spectrum was acquired. After 5 h incubation at 37 ^{o}C, the singlet corresponding to pyrophosphate appeared at δ -6.2 ppm, suggesting that DMAPP had been consumed in the reaction. The enzyme was removed by centrifugation through a membrane (10 kDa MWCO), and the crude reaction was subjected to a reverse phase HPLC separation. The last peak (~20 min) eluting from the column was collected, and a ^{1}H NMR spectrum (D_2O) was obtained after solvent evaporation. The proton signals agreed well with those previously reported for the ^{1}H NMR spectrum of 4-dimethylallyltryptophan (DMAT).[131] In particular, the H-4 signal of L-tryptophan had disappeared and was replaced by the signals for a dimethylallyl moiety. These include the allylic H-2' at δ 5.30 ppm, the methylene H-1' at δ 3.70 ppm and the methyl groups H-4' and H-5' at δ 1.67-1.70 ppm (Figure 4.4). Finally, a peak corresponding to DMAT was observed in the positive ESI-MS spectrum at 273 m/z [$M + H^+$]. As a control without enzyme, incubations of L-tryptophan and DMAPP do not show any product formation.

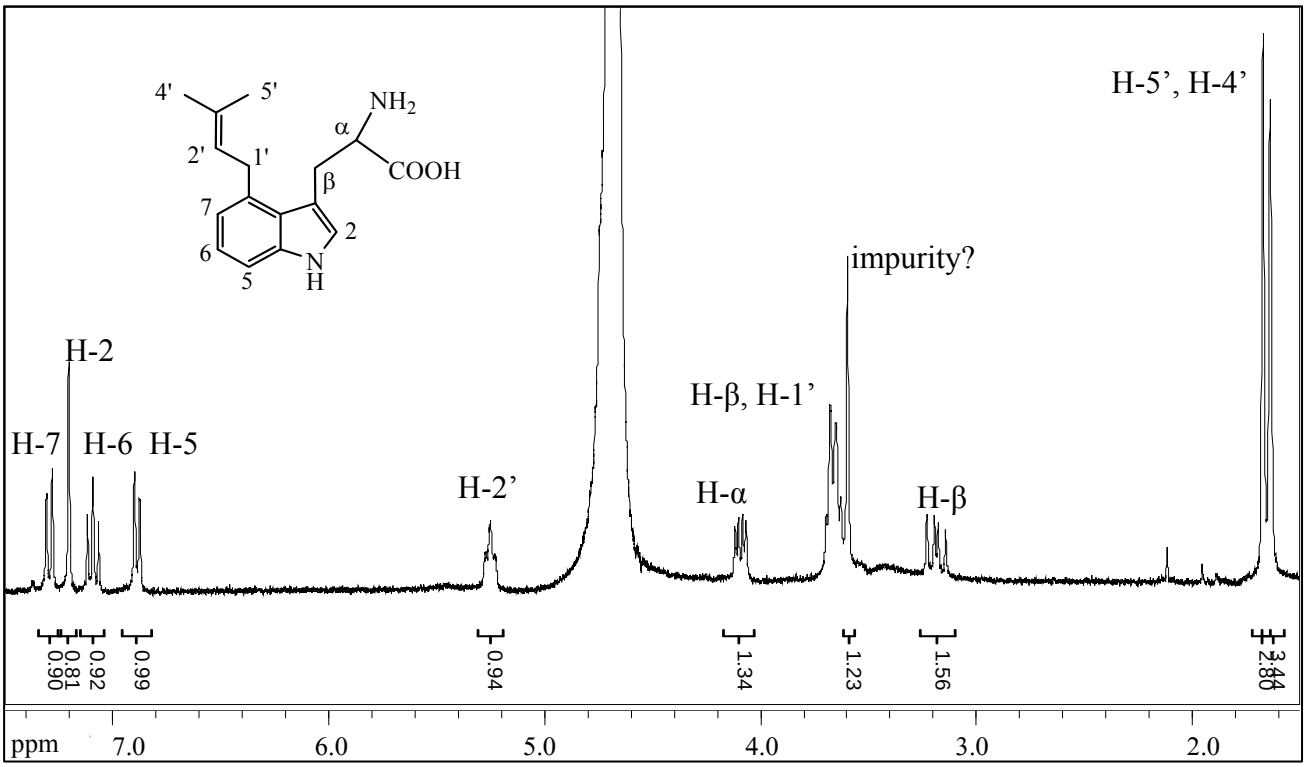


Figure 4.4 ^{1}H NMR spectrum (D_2O) of 4-dimethylallyltryptophan isolated from reverse phase HPLC

4.3 Steady-state Kinetic Characterization of DMATS by a Continuous Coupled Assay

An enzyme coupled assay for pyrophosphate formation was employed to quantitate the kinetic activity of DMATS. In past studies, the kinetics of the reaction were analyzed using a stopped radioactive or HPLC assay.[15, 131, 132] In the former case [1-^{3}H]-DMAPP was employed and the tritium incorporated into the product was measured,[15, 132] whereas in the latter case the retention peak of the product in HPLC was quantitated by the use of an authentic sample of DMAT (4-dimethylallyltryptophan) as standard.[131] A continuous assay that does not require radiolabeled substrate or a separation step could greatly facilitate studies on the synthase. Furthermore, it would allow one to measure the kinetics of the reaction with starting material analogs that are not available

as isotopically labeled compounds or that would create a product analog with no authentic sample available. Since the enzymatic reaction generates pyrophosphate as one of the products, UV-Vis spectroscopy is a potential method for monitoring its kinetics. In this continuous assay, pyrophosphate formation was indirectly measured by first hydrolyzing it into two equivalents of phosphate via the action of the enzyme pyrophosphatase (Figure 4.5). The resulting phosphate is a substrate for another coupling enzyme purine ribonucleoside phosphorylase (PNP) that converts 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) into ribose-1-phosphate and 2-amino-6-mercapto-7-methylpurine. The formation of 2-amino-6-mercapto-7-methylpurine is then monitored spectrophotometrically by recording an increase in absorbance at 360 nm.

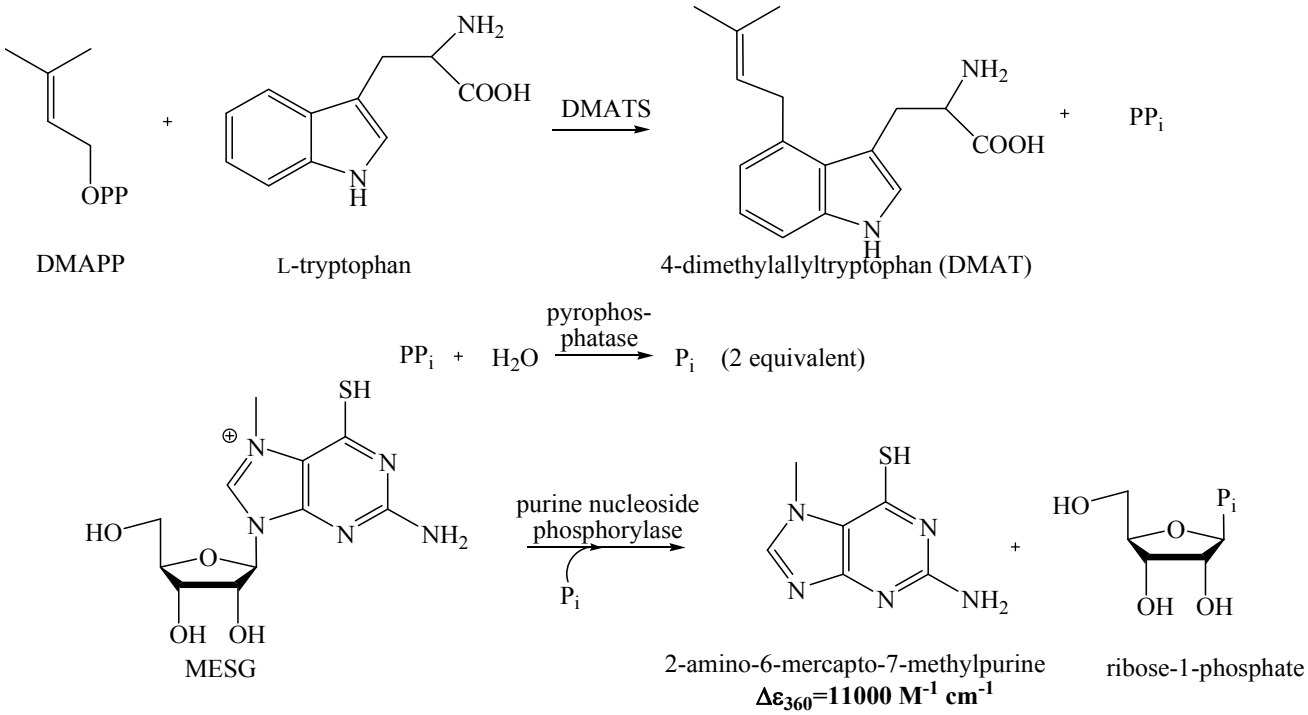
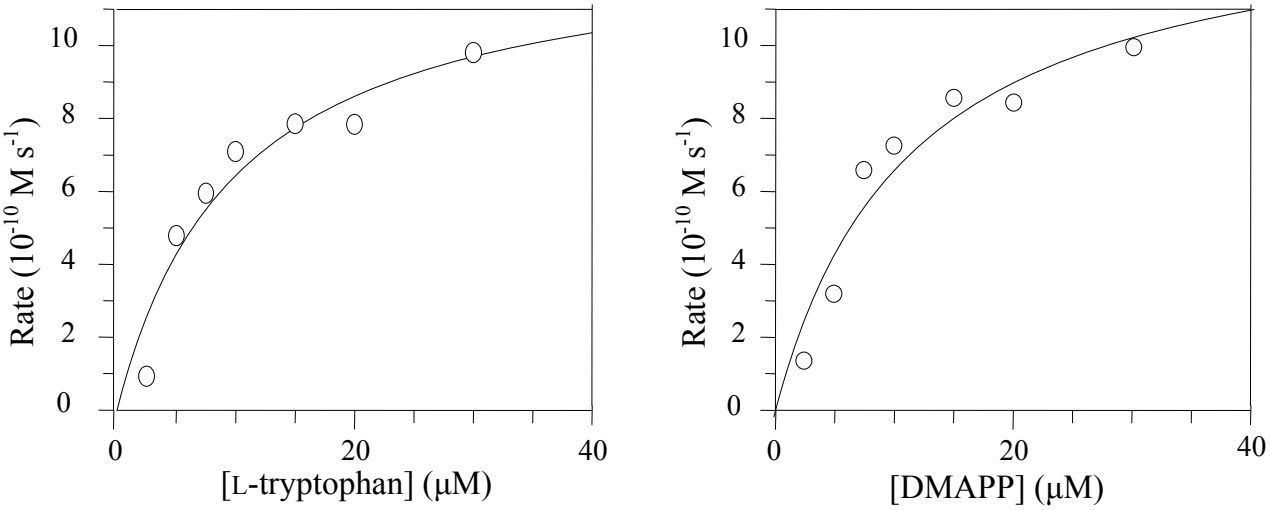


Figure 4.5 A continuous coupled assay for the DMATS reaction

At saturating concentrations of both substrates (30 μM), the kinetic constant k_{cat} was determined to be 0.27 ± 0.05 s^{-1} (Figure 4.6). When the concentration of L-tryptophan was held constant and the concentration of DMAPP was varied, a hyperbolic kinetic profile was observed leading to a value of K_M = 11 ± 4 μM. When the concentration of DMAPP was held constant and

the concentration of L-tryptophan was varied, a hyperbolic kinetic profile was observed leading to a value of $K_M = 10 \pm 4$ μM. These K_M values agree reasonably well with those previously reported in kinetic studies on the same DMATS (*fgaPT2*, 4 μM for DMAPP and 8 μM DMAPP L-tryptophan).[131] Furthermore, this enzyme assay was efficiently coupled, as doubling the synthase concentration caused a doubling in the rate of increase in absorbance indicating that pyrophosphate production was rate-determining. Therefore, the continuous coupled assay developed in this study appears to be a reliable method to measure the activity of DMATS.



Varying [substrate]	K_M (μM)	k_{cat} (s^{-1})
DMAPP	11 ± 4	0.27 ± 0.05
L-tryptophan	10 ± 4	

Figure 4.6 Enzyme kinetics for DMATS with natural substrates.

* The kinetic constants were extrapolated from the corresponding Lineweaver-Burk plot, and the standard errors were obtained by a linear least squares analysis.
* In each experiment, each concentration point was carried out in triplicate and the measured initial velocities never differed by more than 5%.
* Every experiment was repeated 3 times and the resulting kinetic constants never differed by more than 20%.

Kinetic studies have also been reported for the DMATS isolated from *Claviceps* strain SD58.[132, 152] Although the K' constants (7.2 and 8.8 μM for DMAPP and L-tryptophan, respectively)

are similar to the K_M values measured in this study, the *Claviceps* enzyme showed a mixed cooperativity in the enzyme kinetics.[132, 152] Moreover, in the kinetic studies on another DMATS isolated from *Claviceps fusiformis*, the K_M values (67 μM for L-tryptophan and 200 μM for DMAPP) are notably higher than the ones obtained in this study.[152] These differences may be a consequence of divergent evolution, because the DMATS used in this study (*fgaPT2* isolated from *Aspergillus fumigatus*) has approximately 50% sequence homology to other DMTAS from different *Claviceps* species. Differences in K_M values could reflect the different concentrations of substrates present in the parent organism. In addition, differences in pH and buffer among these assays could account for some of the discrepancies.

4.3.1 Alternate Substrate Testing and Mechanistic Analysis

In many past studies on different PTs, fluorinated analogs of isoprenyl diphosphate have been used to probe the mechanism of the prenylation reaction. The first example of these studies was performed by Poulter and his co-workers in their investigations on *trans*-farnesyl diphosphate synthase.[103] Originally, the authors argued that a fluorine substitution would greatly reduce the catalytic rate of an S_N1 pathway but slightly enhance the rate of an S_N2 pathway.[103] However, it was later shown that a fluorinated diphosphate analog greatly depressed the rate of an S_N2 reaction that involves an 'exploded'-associative transition state with a considerable amount of carbocationic character.[113] For instance, the cysteine prenylation reaction catalyzed by protein farnesyltransferase is believed to involve an S_N2-like pathway, yet the reaction rates were significantly depressed when the fluorinated FPP analogs **46** and **47** were employed as substrates (Table 3.1 see page 94).[113] Therefore, measuring the reaction rate with a fluorinated analog might not unambiguously elucidate the reaction mechanism. However, fluorinated analogs have been used to gain information on the

electrophilic nature of the reaction. Therefore, ***E*-52** was prepared as an alternative substrate in these studies.

The synthesis of the fluorinated DMAPP analog, ***E*-52**, was developed by combining two synthetic procedures that were previously reported by Poulter and his co-workers.[15, 106] To begin with, ethyl 3-(fluoromethyl)-2-butenoate (**56**) was made via a Horner-Wadsworth Emmons reaction between fluoroacetone and triethyl phosphonoacetate (Figure 4.7). A DIBAL reduction then gave the corresponding alcohol **57** as a mixture of geometric isomers (~6:4 of *E*- and *Z*-isomer).[15] The *E*-isomer was isolated by silica gel chromatography, and was subsequently coupled with *bis*-triethylammoniumphosphate in the presence of trichloroacetonitrile to give the resulting diphosphate ***E*-52**.[106] Silica gel chromatography yielded the product with reasonable purity (~95%) as determined by ^{1}H, ^{19}F and ^{31}P NMR and ESI-MS methods.

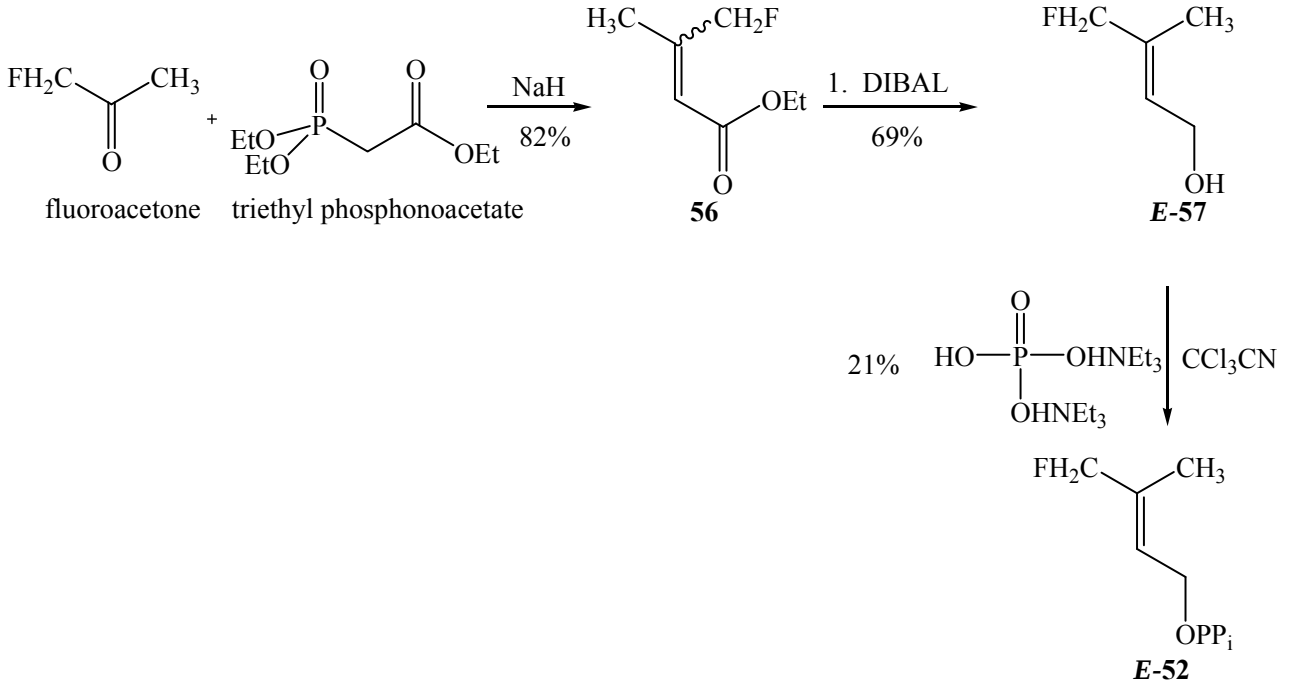
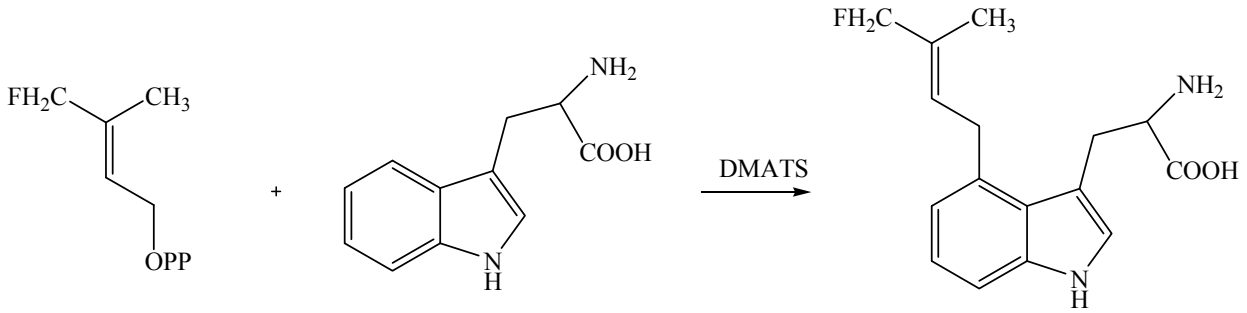


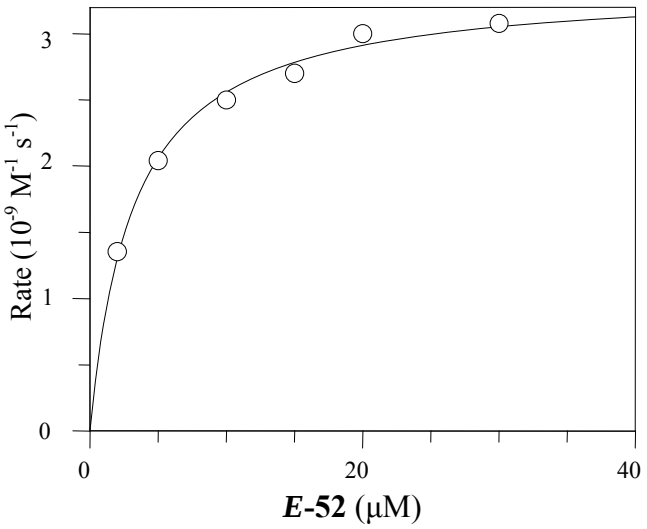
Figure 4.7 Synthesis of (*E*)-3-fluoromethyl-2-butenyl diphosphate (*E*-52)

The reactivity of compound ***E*-52** in the DMATS reaction was analyzed by reverse phase HPLC. It was found that the fluorinated substrate was converted to the corresponding product

analog, as a new peak was observed in the HPLC trace. The ^{1}H and ^{19}F NMR and the positive ESI-MS spectra (291 m/z, $M + H^{+}$) of the purified retention peak confirmed the identity of the fluorinated dimethylallyltryptophan (**58**) (Figure 4.8).



***E*-52** L-tryptophan fluorinated dimethylallyltryptophan **58**



Varying [substrate]	K_M (μM)	k_{cat} (s^{-1})	k^{rel}
DMAPP	11 ± 4	0.27 ± 0.05	1
***E*-52**	3.2 ± 0.3*	0.014 ± 0.003*	0.05**

Figure 4.8 Kinetics of DMATS using fluorinated DMAPP analog *E*-52

* The kinetic constants were extrapolated from the corresponding Lineweaver-Burk plot, and the standard errors were obtained by a linear least squares analysis.
* In each experiment, each concentration point was carried out in triplicate and the measured initial velocities never differed by more than 5%.
* Every experiment was repeated 3 times and the resulting kinetic constants never differed by more than 20%.
** Relative error value for k^{rel} was not reported as it has little relevance to this discussion.

The kinetic constants with the substrate analog ***E*-52** were determined using the enzyme coupled assay as elaborated in the previous section. At saturating concentrations of the fluorinated analog and L-tryptophan (30 μM each), the k_{cat} parameter was found to be 0.014 s^{-1}. The k^{rel} value used to quantitate the reactivity of ***E*-52** was calculated by dividing the k_{cat} value measured with the fluorinated analog by the k_{cat} value measured with DMAPP (Figure 4.8 and Table 4.1). The k^{rel} value was determined to be 0.05, which agrees reasonably well with the same studies on the DMATS isolated from *Claviceps* strain SD58 (k^{rel} = 0.044 for ***Z*-52** and 0.014 for ***E*-52**, Table 4.1).[15]

Similar monofluorinated analogs have been kinetically analyzed in other related reactions. Specifically, the hydrolysis of geranyl methanesulfonate and the *trans*-farnesyl diphosphate synthase reaction are considered to proceed through an S_N1-like pathway. In these reactions, the k^{rel} values of the corresponding CH_2F-substrate analogs are on the order of 10^{-4} (Table 4.1 and Table 3.1 on page 94).[15] These values are significantly lower than the k^{rel} values measured in the DMATS reaction. In turn, the k^{rel} value of ***E*-52** matches more closely to the k^{rel} values of the CH_2F-substrate analogs that were measured in the protein farnesyltransferase reaction and the nucleophilic reaction between geranyl methanesulfonate and azide (k^{rel} = 0.018 and 0.062, respectively) (Table 4.1 and Table 3.1 on page 94).[15, 113] These reactions are known to proceed through an S_N2-like pathway, which involves an associative transition state with significant carbocationic character. Hence, the similarities imply that DMATS might employ an associative transition state when ***E*-52** serves as substrate. However, it is conceivable that the enzymatic mechanism could have switched to an S_N2 pathway for catalysis when this unnatural substrate is employed, because the fluorinated dimethylallyl carbocation is much less stable than the corresponding non-fluorinated carbocation. Hence, more mechanistic studies are needed to further characterize the DMATS reaction in the absence of such a dramatic perturbation.

Table 4.1 The k^{rel} of monofluorinated analog in different S_N1 or S_N2 reactions

Enzymatic Reactions or Solvolysis	Mechanism Involved	k^{rel}
DMATS (*fgaPT2*) reaction (this study)	S_N2 (**?**)	0.05
DMATS (*dmaW*) reaction (Poulter study)*	S_N2 (**?**)	0.014
Solvolysis of geranyl methanesulfonate in water*	S_N1	7.7×10^{-4}
trans-FPP synthase reaction*	S_N1	3.7×10^{-4}
Nucleophilic reaction of geranyl methanesulfonate in azide*	S_N2	0.061
Protein farnesyltransferase reaction*	S_N2	0.018

Asterik refers to values obtained from the previous studies conducted by other research groups.[15, 113]

4.4 Application of Positional Isotope Exchange (PIX) in the Study of DMATS

4.4.1 Introduction

In order to gain additional insight into the mechanism, positional isotopic exchange (PIX) experiments were applied to the studies on DMATS. PIX is an approach that follows the movement of an isotope from a particular position in the starting material to another position within the recovered starting material after a partial reaction. This characteristic isotopic scrambling can provide evidence for formation of a transient enzyme-bound intermediate, thus making PIX a particularly useful technique in mechanistic studies.

A PIX experiment is generally applied to substrate containing a carboxylate or a phosphate ester group. When an ^{18}O label is placed at the bridging position of the starting material and the ^{18}O-R_1 bond is broken to give a carboxylate or phosphate intermediate, the two C-O bonds or the three P-O bonds may become torsionally equivalent via rotation about the R_2-C or R_2O-P bond, respectively (Figure 4.9). Provided that the reverse reaction and the torsional equilibration occur at a rate higher than the subsequent step of the enzymatic reaction, the labeled oxygen should be 'scrambled'

randomly among the bridging and the non-bridging positions (two in carboxylate and three in phosphate) in recovered starting material.

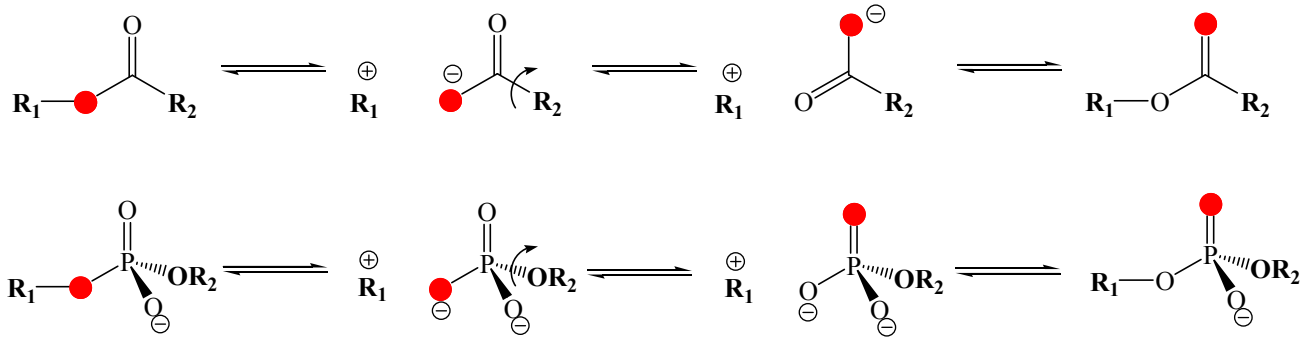


Figure 4.9 Positional isotope exchange in carboxylate and phosphate esters. Solid red circle represents ^{18}O label, and R_1 was cleaved as a carbocationic intermediate.

In 1976, Midelfort and Rose were the first to demonstrate the feasibility of applying a PIX experiment in mechanistic enzymology studies. Specifically, they investigated the reaction catalyzed by glutamine synthetase and showed that γ-glutamyl phosphate (γ-Glu-P) was an intermediate formed during the reaction (Figure 4.10).[153] When two of the substrates, [γ-$^{18}O_4$]-ATP (**59**) and glutamate, were incubated with the enzyme in the absence of the third substrate, ammonia, the isotope label scrambled from the β-γ bridging (**59**) to the β non-bridging position (**60**) in the recovered ATP mixture. Accordingly, the PIX result was indicative of the reversible formation of enzyme-bound γ-Glu-P and ADP.

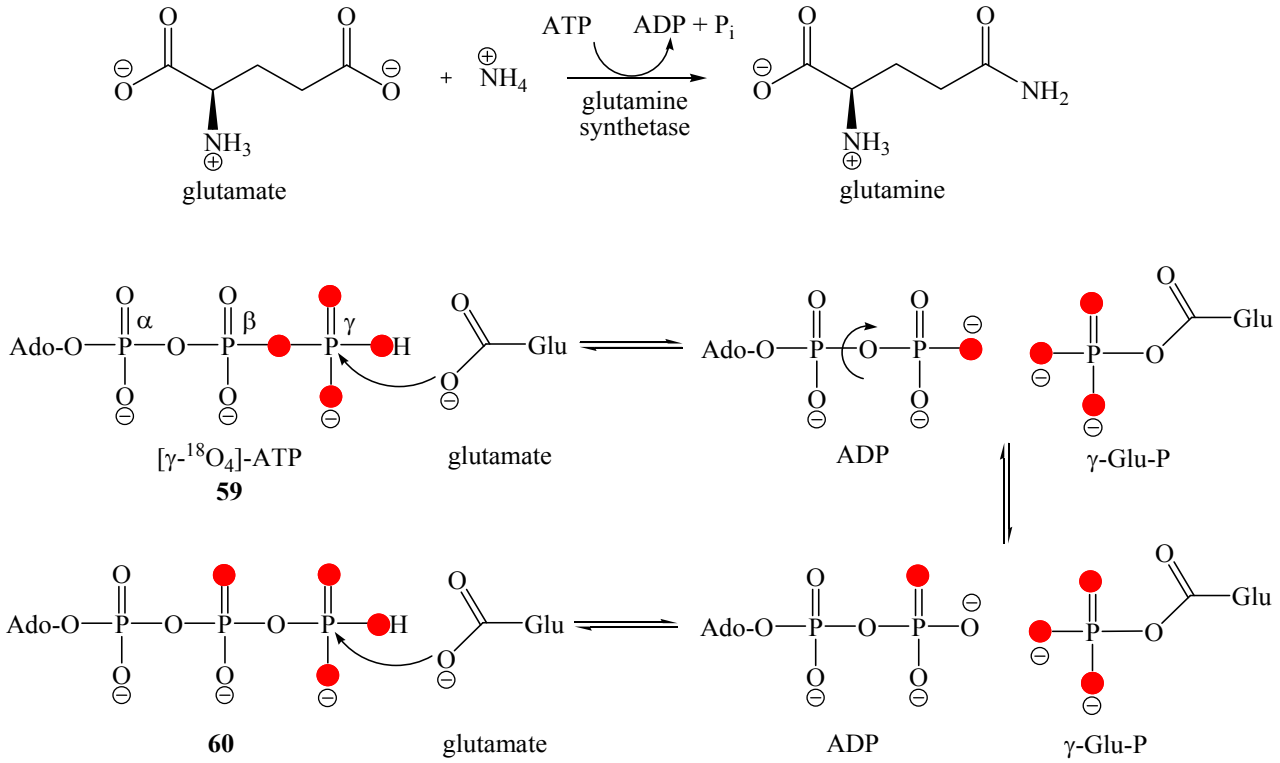


Figure 4.10 The glutamine synthetase reaction and the formation of γ-glutamyl phosphate as determined by PIX. Solid red circle represents ^{18}O-label.

At the time when the first PIX experiment was carried out, there was no analytical method that could directly differentiate between compounds **59** and **60**. Thus, the positions of the ^{18}O-labels in the recovered ATP mixture were determined by an indirect method that involved a series of enzymatic reactions and mass spectrometric analysis.[153] To begin with, the recovered ATP mixture was subjected to another enzyme-catalyzed PIX reaction. Specifically, the ATP mixture (**59** and **60**) was treated with acetyl-CoA synthetase, which catalyzes the reversible adenylation of acetate. The terminal diphosphate of **59** and **60** would be displaced to give ^{18}O-labeled pyrophosphate, **61** and **62**, respectively (Figure 4.11). The second PIX reaction occurred, when these pyrophosphate isotopologues dissociated from the active site and then re-associated with the enzyme with the phosphate moieties inverted. Subsequently, two new forms of ^{18}O-labeled ATP, **63** and **64**, were generated. At this point, the extent of PIX caused by glutamine synthetase could be analyzed, because compound **64** (derived from **62**) is the only ATP isotopologue that carried one ^{18}O-isotope

at the γ-phosphate position. Upon treatment with dihydroxyacetone and glycerokinase, the γ-phosphate in **64** was then transferred to give dihydroxyacetone phosphate (**65**). Finally, inorganic phosphate was released from compound **65** and then permethylated to give trimethylphosphate ester (**66**) for gas chromatography MS analysis.

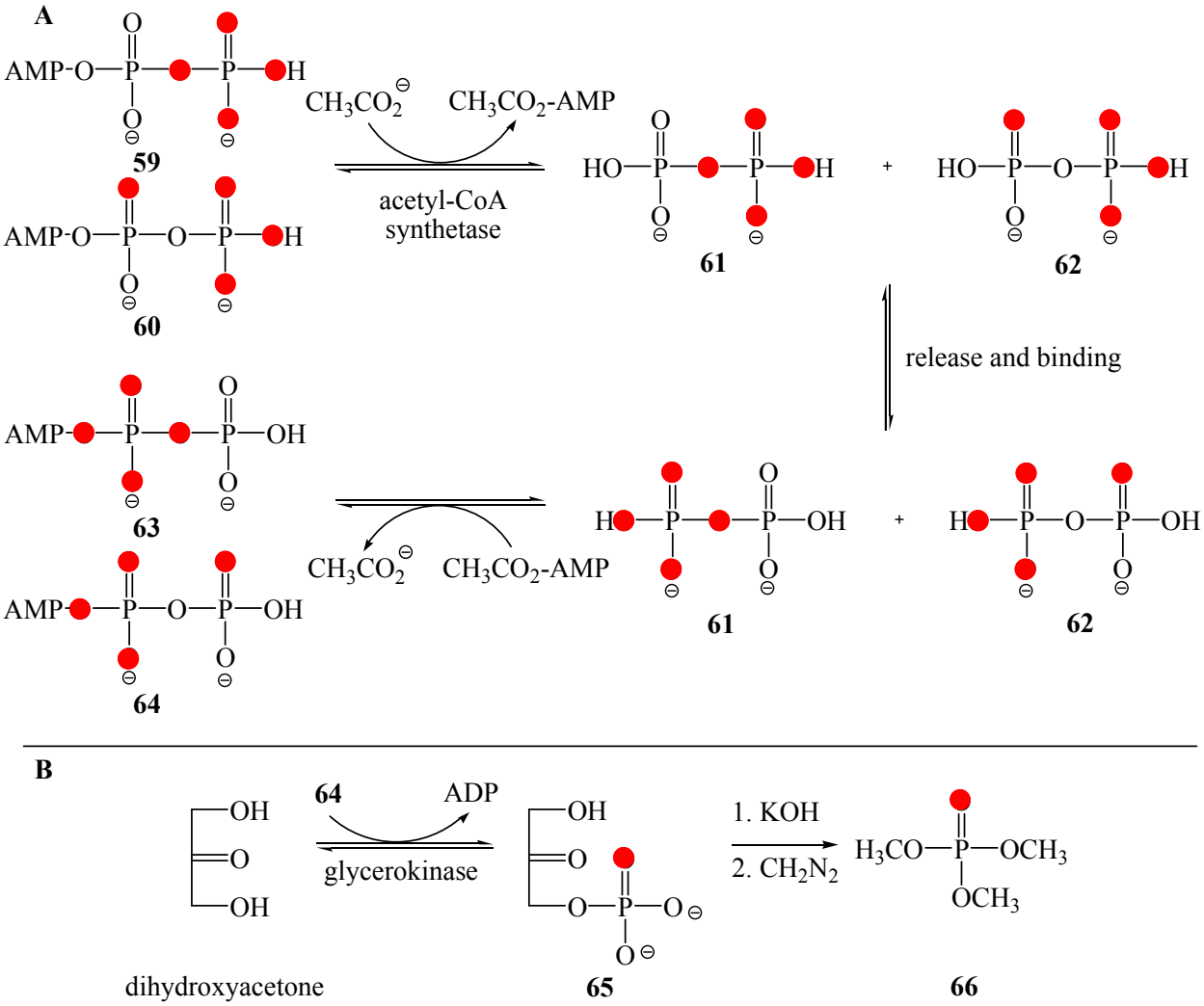


Figure 4.11 Analysis of the glutamate synthetase-catalyzed PIX reaction via an indirect method. Solid red circle represents ^{18}O label.

This rather time-consuming and tedious method was eventually circumvented by the use of NMR spectroscopy in the analysis of PIX. The recovered starting material that carries the scrambled isotope will produce a new signal in high resolution NMR spectra. This is based on the observation that different isotopes, such as ^{16}O and ^{18}O, perturb the electronic environment of the attached

phosphorus (or carbon) nucleus to different extents and thereby induce a measurable difference in chemical shifts.[154] The replacement of an ^{18}O-isotope for each of the ^{16}O atoms in a phosphate or phosphate ester causes a small upfield shift (~0.02 ppm). As demonstrated by Cohn and Hu in 1978, the ^{31}P NMR spectrum of inorganic phosphate randomly labeled with 44% ^{18}O contained five peaks, which represent the statistical ratio of molecules carrying from 0 to 4 isotopic labels (Figure 4.12).[154]

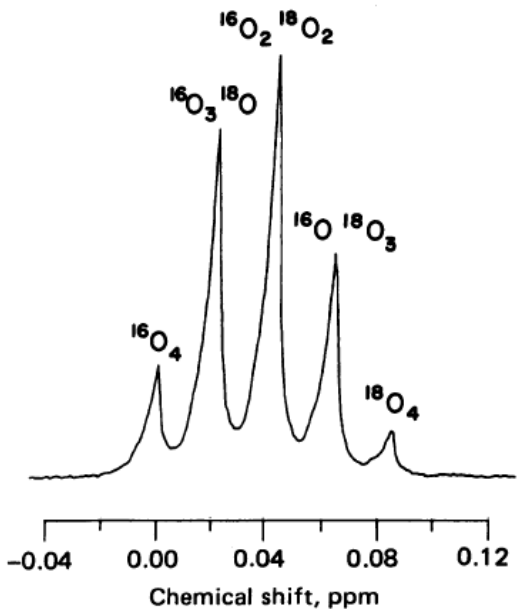


Figure 4.12 ^{31}P NMR spectrum of inorganic phosphate randomly labeled with 44% ^{18}O-isotopes. Spectrum was obtained from the NMR studies conducted by Cohn and Hu.[154]

In addition to the number of isotopically labeled atoms, the bond order between the isotopic label and the NMR active nucleus also dictates the extent of the ^{18}O-induced chemical shift.[155] In further investigation by Cohn and Hu, ATP and ADP carrying ^{18}O labels at various positions in the phosphate chain were synthesized and their ^{31}P NMR spectra were obtained to determine the changes in chemical shift. While ^{18}O in a bridging position (P-O bond order of one) causes a smaller upfield shift of 0.017 ppm, the isotope in a non-bridging position (P-O bond order of 1.33 or 1.5) induces a higher upfield shift ranging from 0.022 to 0.028 ppm. Consequently, when an ^{18}O-isotope in a labeled substrate scrambles from a bridging position to a non-bridging position during a PIX reaction, a new upfield signal will appear in the ^{31}P NMR spectrum.[155]

4.4.2 Application of PIX in the Study of DMATS

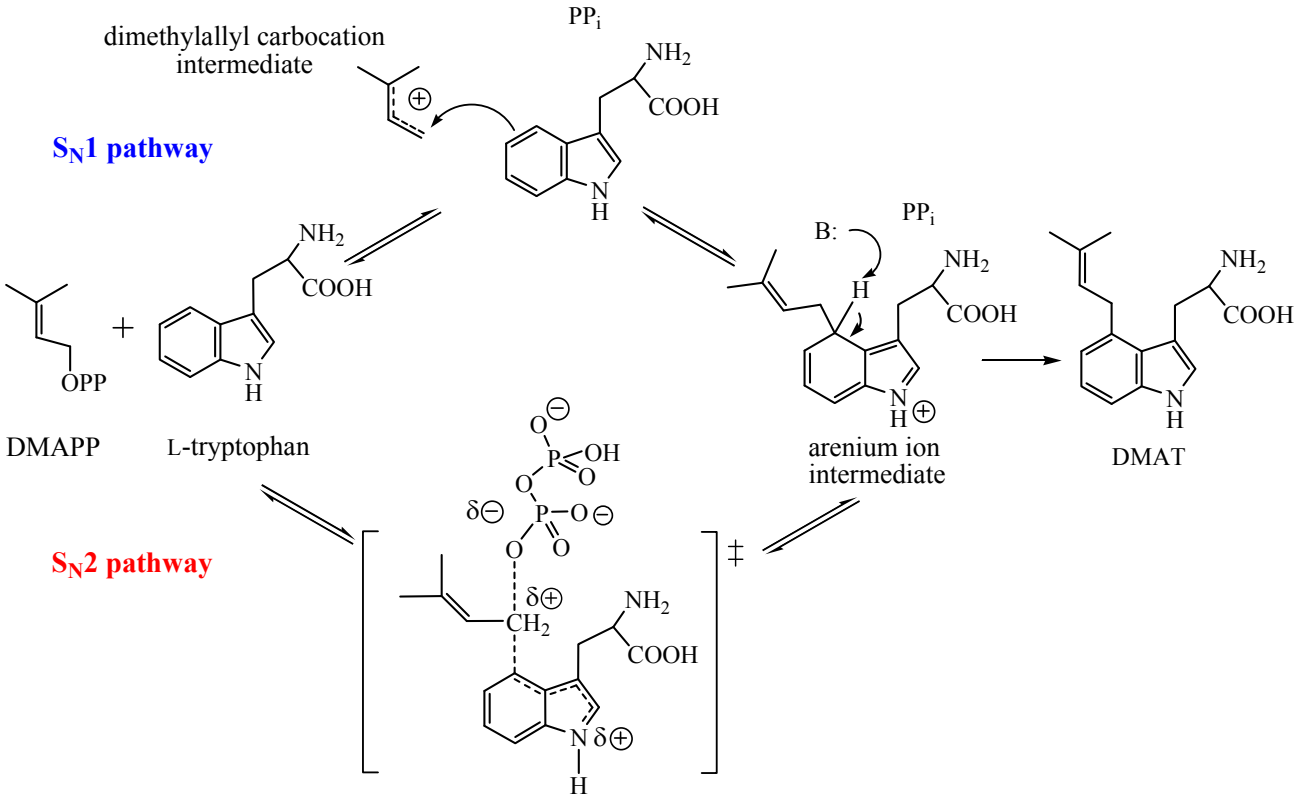


Figure 4.13 Proposed S_N1 and S_N2 mechanistic pathways of the DMATS reaction

As described in Chapter 3, there are two possible mechanistic pathways that DMATS can use to catalyze the condensation between L-tryptophan and DMAPP (Figure 4.13). The dissociation of pyrophosphate from DMAPP can be a distinct chemical step and form an allylic carbocation intermediate (S_N1), or it can be combined with the nucleophilic attack into one transition state (S_N2). In order to probe the transient formation of the allylic carbocation, it was considered relevant to perform positional isotope exchange (PIX) studies. The PIX experiments employ synthetic [1-^{18}O]-DMAPP (**67**), which contains an ^{18}O-isotope at the position bridging the hydrocarbon and diphosphate moieties (Figure 4.14). When the enzyme is incubated with the labeled substrate and substoichiometric L-tryptophan, partial conversion into the product DMAT will be expected. The remaining isotopically labeled DMAPP will be recovered and examined for evidence of isotope

scrambling. Assuming the formation of pyrophosphate is reversible and it has a life time comparable or greater than that required for the torsional equilibration of the labeled atom, the recovered DMAPP should have the ^{18}O isotope 'scrambled' between the bridging and non-bridging positions of the α-phosphate group ([1-^{18}O]- and [α-^{18}O]-DMAPP, respectively).

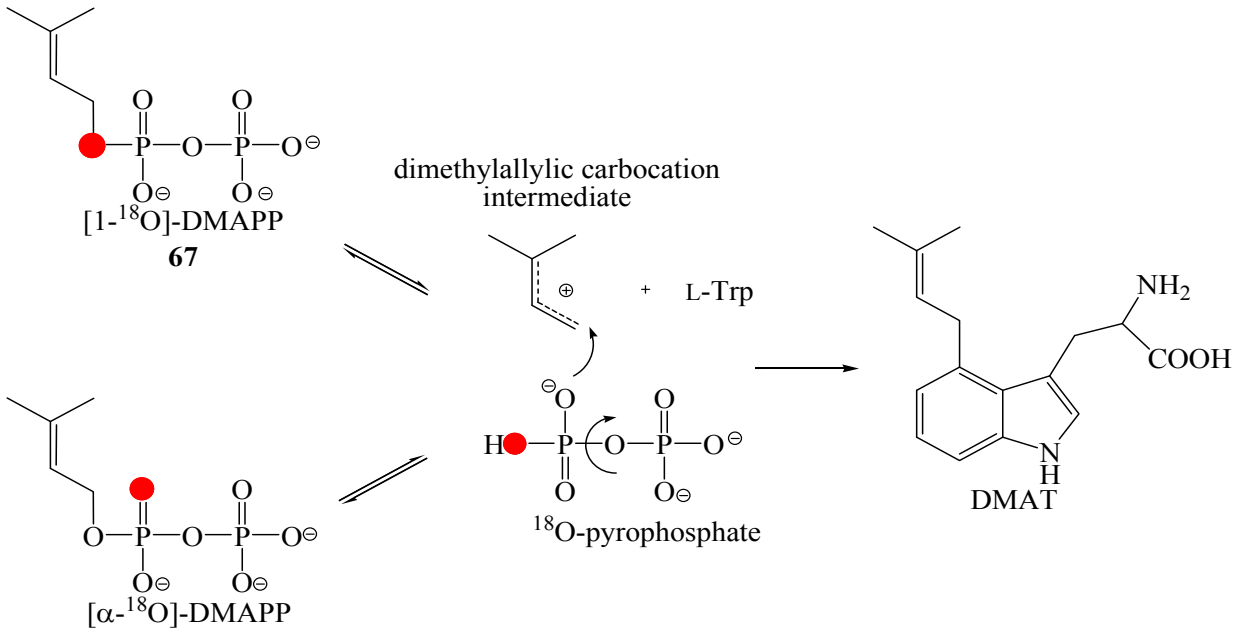


Figure 4.14 Test for C-O bond cleavage by PIX during the DMATS reaction. Solid red circle represents ^{18}O label.

4.4.3 Synthesis of Labeled Substrate [1-^{18}O]-DMAPP for the PIX Experiment

The strategy for the synthesis of [1-^{18}O]-DMAPP (**67**) began with the incorporation of the ^{18}O-isotope into the allylic position to give the labeled [1-^{18}O]-dimethylallyl alcohol (**70**). An initial attempt to prepare **70** was based on the reported nucleophilic reaction between dimethylallyl chloride (**68**) and ^{18}O-water in the presence of silver carbonate.[156] However, this reaction gave a low percentage of ^{18}O incorporation (≤25%). An alternative approach involved modifying a procedure that was reported for the synthesis of a similar compound.[157, 158] To begin with, triethylorthoacetate was refluxed with 95% enriched ^{18}O water and *p*-toluenesulfonic acid to yield the sodium [1,1-$^{18}O_2$]-

acetate (Figure 4.15).[158] The ^{18}O-labeled acetate was heated with dimethylallyl chloride (**68**) and a small amount of triethylbenzylammonium chloride to produce the ^{18}O-labeled ester **69**. Deacetylation in anhydrous methoxide gave the [1-^{18}O]-dimethylallyl alcohol **70**. Lastly, the coupling reaction between **70** and triethylammonium phosphate (TEAP) in the presence of trichloroacetonitrile (CCl_3CN) generated [1-^{18}O]-DMAPP (**67**). This method gave a reasonable yield (overall ~8%) and a substantially higher level of ^{18}O-incorporation as determined by negative ESI-MS spectrometry (m/z = 247 and ~91%).

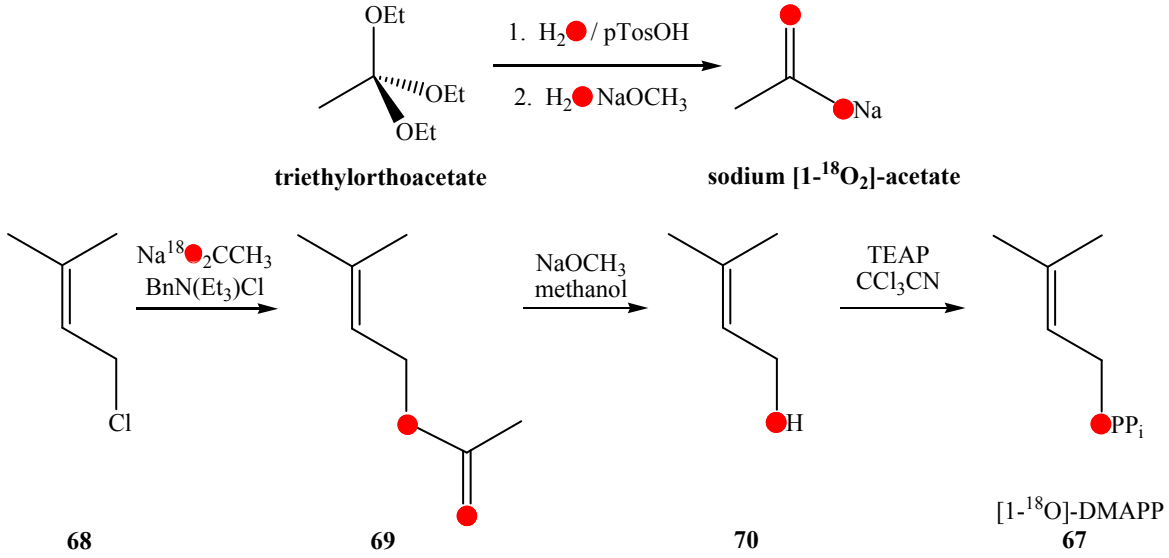


Figure 4.15 Synthesis of [1-^{18}O]-DMAPP (67). TEAP is *bis*-triethylammonium phosphate, and solid red circle represents ^{18}O label.

4.4.4 PIX Experiment with Natural Substrate L-Tryptophan

The PIX experiment described in Section 4.4.2 was used to probe for the formation of a dimethylallyl carbocation during catalysis. A sample of labeled [1-^{18}O]-DMAPP (**67**) bearing 63% enrichment of ^{18}O-isotope was incubated with L-tryptophan (0.8 equivalent) and DMATS (1 mg) for an extended period of time in Tris-HCl buffer containing either EDTA (3 mM) or $CaCl_2$ (5 mM). The reaction progress was monitored by ^{31}P NMR spectroscopy, which showed the accumulation of

pyrophosphate as [1-^{18}O]-DMAPP was consumed during the DMATS reaction. After approximately 3 h incubation with the enzyme (~1 mg), 57% of the substrates had been converted to product. The enzyme was then removed by centrifugation through a membrane filter (10 kDa MWCO). The ^{31}P NMR spectra of the crude reaction mixture were acquired to evaluate the extent of PIX.

Since a high resolution is required to observe the small difference in chemical shifts (~0.01 ppm) in the ^{31}P NMR spectra, it is crucial to properly set up the acquisition parameters of the spectrometer. First, the sweep width was reduced to 20 ppm (centered at -5 ppm), accompanied by a prolonged extension of the acquisition time (aq = 27 s, see experimental section for details). Second, additional mathematical processing of the acquired FID (i.e. line broadening) was employed to further resolve the small variations in the chemical shifts. In the reaction containing $CaCl_2$, the presence of divalent cations caused an undesired broadening of the ^{31}P NMR signal. In order to solve this problem, a chelating agent such as EDTA can be employed. In this particular experiment, however, a small amount of specialized chelating resin (Chelex-100) was added to the crude reaction mixture prior to the acquisition, and proved to efficiently enhance the resolution of the ^{31}P NMR spectra. This resin is particularly convenient since it does not need to be removed from the NMR tube and causes no notable interference during data acquisition.

The high resolution proton-decoupled ^{31}P NMR spectra of the reaction mixture before and after the addition of DMATS are shown in Figure 4.16. Before the addition of DMATS, the spectrum showed a doublet at δ -5.51 ppm, which corresponds to the β-phosphorus nuclei of both unlabeled and [1-^{18}O]-DMAPP (Figure 4.16A). The α-phosphorus nuclei of the DMAPP mixture appeared as two set of doublets at δ -9.28 and 9.30 ppm. Since a ^{31}P signal will be shifted slightly upfield by an attached ^{18}O-isotope,[154, 159] the major doublet (63%) corresponds to the ^{18}O-labeled substrate **67**, and the minor doublet (37%) corresponds to the unlabeled substrate which served as a

reference peak for the analysis. All of these signals appear as doublets due to the coupling to the neighboring phosphorus nuclei ($J_{P\text{-}P}$ = 21.4 Hz).

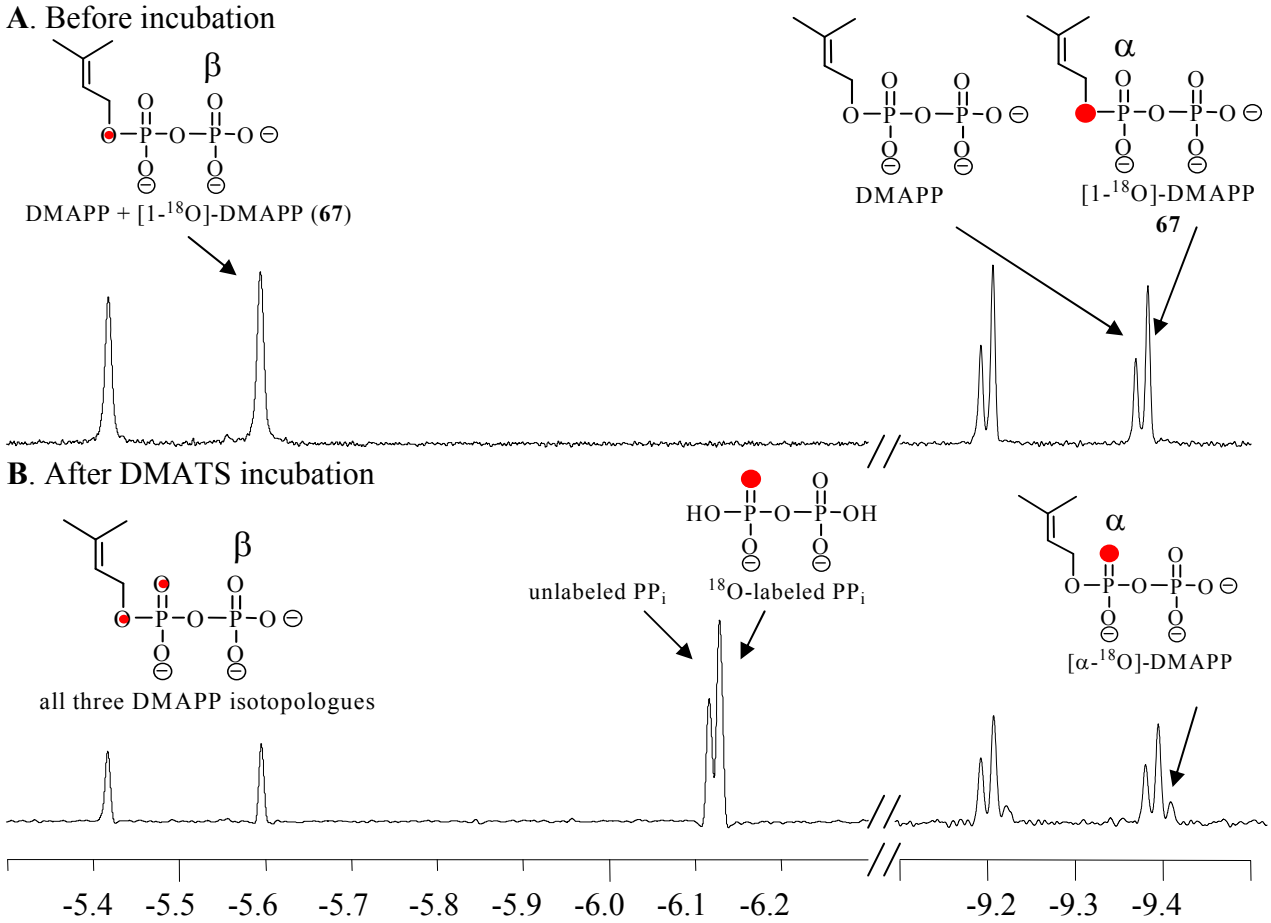


Figure 4.16 ^{31}P spectra (D_2O) the DMATS-catalyzed PIX reaction. A) Before addition of enzyme, and B) after ~57% reaction completion. The pyrophosphate signal and the α- and β-phosphorus signal of DMAPP and its ^{18}O-labeled isotopologues are indicated. Red denotes ^{18}O-label. A small red circle refers to mixture of ^{16}O- and ^{18}O-labeled DMAPP isotopologues.

After the mixture was incubated with DMATS, two new signals appeared in the ^{31}P NMR spectrum. Two overlapping singlets emerged at around δ -6.12 and -6.13 ppm, both of which represent the pyrophosphate molecules that were produced during the enzymatic reaction (Figure 4.16B). The minor peak was assigned as the phosphorus nuclei of unlabeled pyrophosphate derived from DMAPP. In turn, the major peak was assigned as the phosphorus nuclei of ^{18}O-labeled pyrophosphate that was derived from bridging-[1-^{18}O]-DMAPP (**67**). This was somewhat surprising, as the ^{18}O-labeled pyrophosphate was initially expected to give rise to two coupled signals that

display a non-equivalent AB splitting pattern. Previous studies have indeed reported on this phenomenon, and it is believed that the ^{18}O-isotope causes an unexpected β-isotope shift of the unlabeled phosphorus signal, which in turn overlaps with the signal of the phosphorus nucleus that is attached to the ^{18}O-isotope. The phosphorus nuclei actually couple with each other, but the coupling constant is only 0.5 Hz.[160–163] Thus, the ^{18}O-labeled pyrophosphate signal appears as an apparent singlet peak in the spectrum.

The key region of the spectrum that is important for this PIX study contains the signals of the α-phosphorus nuclei of [1-^{18}O]-DMAPP (-9.20 to -9.38) and is magnified in Figure 4.17. As described earlier, the two doublets at δ -9.28 ppm and -9.30 ppm correspond to the unlabeled and labeled starting material (**67**), respectively (Figure 4.17A). A spectrum of the DMAPP remaining after the enzymatic reaction showed the presence of a new signal shifted further upfield at δ -9.31 ppm that accounts for 15% of the ^{18}O-labeled material (Figure 4.17B). Because the magnitude of the chemical shift is directly proportional to the bond order of the ^{18}O-P bond,[154, 159] this new doublet indicates that the ^{18}O isotope has moved into the non-bridging position of the phosphate moiety creating [α-^{18}O]-DMAPP. The amount of ^{18}O-scrambled product was evaluated by dividing the integral of the [α-^{18}O]-DMAPP (non-bridging) signal over the combined integrals of the [α-^{18}O]-DMAPP (non-bridging) and the [1-^{18}O]-DMAPP (bridging) signals. According to these measurements, the enzymatic reaction run in buffer containing EDTA resulted in a slightly larger extent of ^{18}O scrambling at the α-phosphate (~15% [α-^{18}O]-DMAPP) than the reaction run in buffer containing $CaCl_2$ (~12% [α-^{18}O]-DMAPP). However, neither reaction showed the statistical 1:2 distribution between the bridging and non-bridging positions, implying that the positional exchange of the isotope was incomplete. As a control, a PIX experiment was carried out in the absence of L-tryptophan and formation of this new doublet was not observed. It should be noted that no isotopic splitting was observed in the signals of the β-phosphorus nuclei of the recovered starting material.

This indicates that the pyrophosphate could not flip in the active site or be released and rebound during catalysis so that the position of the phosphate moieties interchange.

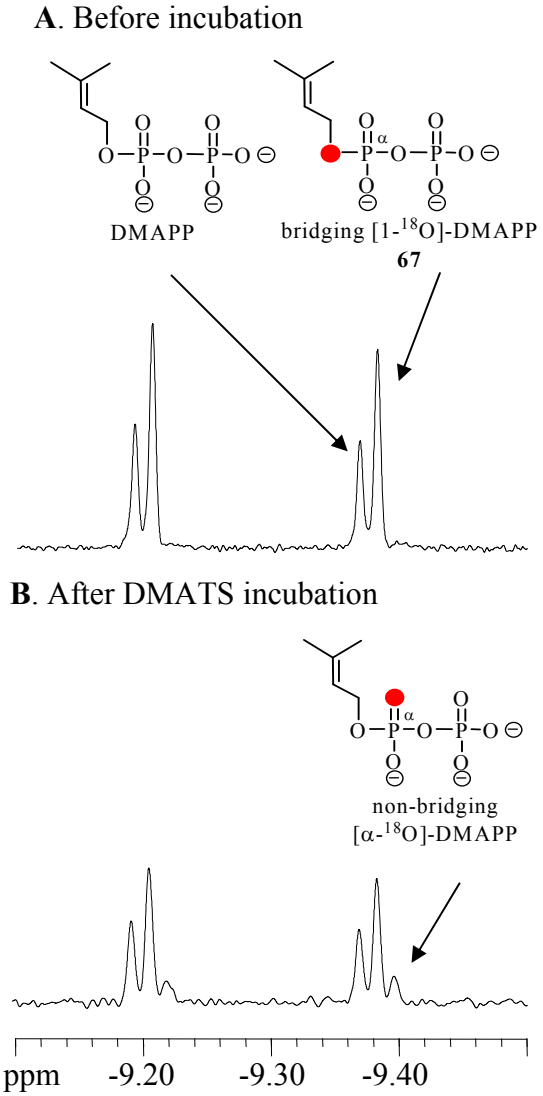


Figure 4.17 The α-phosphorus signals in the ^{31}P spectra (D_2O) of the DMATS-catalyzed PIX reaction. Red denotes the ^{18}O-isotope.

In the DMATS-catalyzed PIX reaction, the isotopic scrambling reflects a partitioning of the carbocation/pyrophosphate ion pair (dimethylallyl carbocation in the S_N1 pathway or arenium ion in the S_N2 pathway) between an irreversible conversion to products and a reversible return to the DMAPP starting material. Since no scrambling is observed in the absence of tryptophan, the initial rate of PIX will correlate linearly with the rate of product formation at all substrate concentrations.

Therefore, the partitioning ratio can be calculated from the ratio of the percentage of scrambled DMAPP to the percentage of DMAPP that has been converted to product. This parameter is crucial in quantifying the amount of PIX that occurs upon each turnover event. In the PIX experiment with [1-^{18}O]-DMAPP (**67**) and tryptophan, the amount of ^{18}O-scrambled product does not exceed 15% over the course of the first 57% of the reaction (effectively initial velocity conditions), and an approximate value of the partitioning ratio (v_{pix}/v_{chem}) can be obtained from equation (4.1):

$$v_{pix}/v_{chem} = \frac{\text{percentage of PIX extent}}{\text{percentage of DMAPP converted to product}} \qquad \text{equation (4.1)}$$

Theoretically, when the extent of PIX reaches completion, the integrals of the non-bridging signal to the bridging signal will be in the ratio of 2:1 and the maximal amount of ^{18}O-scrambled product is 66%. Hence, the percentage of PIX extent could be described as v_{pix} = 15%/66% = 23%. The percentage of DMAPP converted to product was approximated by dividing the integration of the pyrophosphate product signals by the combined integrations of the pyrophosphate signals and the total signals due to the DMAPP α-phosphorus (both labeled and unlabeled) X 2. This value was found to be 57%. Therefore, the observed v_{pix}/v_{chem} = 23/57 = 0.40. If the full extent of PIX had been observed, the theoretical maximal v_{pix}/v_{chem} (full PIX) = 100/57 = 1.75, which is ~5x higher than the observed v_{pix}/v_{chem} values.

There are two scenarios that could account for the observed PIX process. As mentioned earlier, the reaction could follow the S_N1 pathway (Figure 4.18), in which generation of the dimethylallyl cation/pyrophosphate ion pair is reversible and its lifetime is long enough for rotation of the pyrophosphate to occur. If collapse of the ion pair is competitive with product formation, isotopic scrambling would be observed. However, the reaction could also follow the S_N2 pathway

with a reversible first step (Figure 4.18). This could occur if a subsequent step, such as deprotonation of the arenium ion intermediate or product release, was rate-determining.

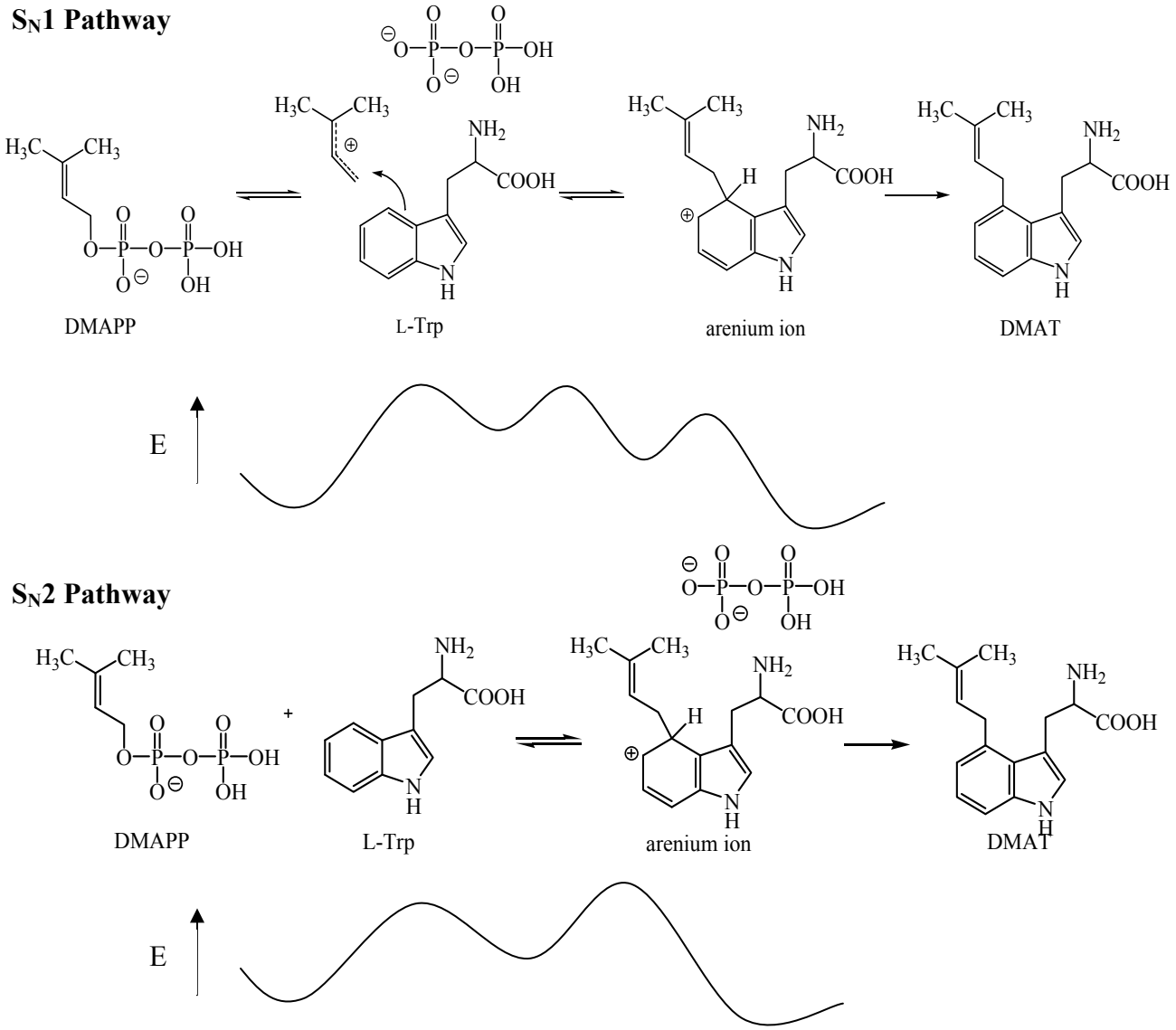


Figure 4.18 Possible scenarios to explain the observation of PIX. Conceptual reaction coordinate diagrams are shown below mechanistic steps.

4.5 Kinetic Isotope Effect on k_{cat}/K_M by Intermolecular Competition

In order to analyze the observation of ^{18}O-scrambling in the PIX experiment, the nature of the rate-determining step(s) were investigated using kinetic isotope effect (KIE) measurements. The KIE was measured on the second-order rate constant, k_{cat}/K_M, symbolized as k_H/k_D. As described in Section 2.4.2, a mixture of two isotopologues of known isotopic composition is treated with the enzyme until a certain fraction is converted to product. Then, the isotope composition of the substrate is recovered and investigated. In the cases of a primary and a secondary normal KIE the lighter isotopologue will be consumed to a greater extent than the heavier isotopologue, while in the case of a secondary inverse KIE the heavier isotopologue will be consumed to a greater extent. The KIE can be quantitated by equation (4.2):

$$k_H/k_D = \frac{\ln(1-F_H)}{\ln[(1-F_H)R/R_o]}$$

equation (4.2, same as equation 2.4)

where F_H is the fractional conversion of the protiated species to products, and R_0 and R are the initial and final ratios of the protiated to deuterated substrate, respectively. Specifically, R values can be determined by the ratio of deuterated to protiated substrate ($R = a_D/a_H$). To probe the rate-determining step(s) in the DMATS reaction, two isotopically labeled substrates [1,1-$^2H_2$]-DMAPP (**71**) and [4-^{2}H]-L-tryptophan (**72**) were used. The results and discussions for these investigations will be presented in the following sections.

4.5.1 Kinetic Isotope Effect Measurement with [1,1-$^2H_2$]-DMAPP

The initial KIE measurement employed the deuterated substrate, [1,1-$^2H_2$]-DMAPP (**71**), in probing the rate-determining step in the DMATS reaction (Figure 4.19). In the S_N1 pathway, if dissociation of the pyrophosphate is the slow step, a normal secondary isotope effect is expected because the C-1 atom experiences a change in hybridization from sp^3 to sp^2. A similar KIE would be expected in the S_N2 pathway, since the allylic position of DMAPP possesses a significant carbocationic character in the 'exploded'-associative transition state. If other chemical steps or non-chemical steps are rate-determining, a KIE of unity will be observed. For example, deprotonation of the arenium ion intermediate could be the slow step of the reaction, as was observed in the reactions catalyzed by norcoclaurine synthase and strictosidine synthase (see Chapter 2).[49, 63] Also, the binding of substrate(s) and/or the release of product(s) can be a rate-limiting process. This has actually been observed in the kinetic studies on another prenyltransferase, *trans*-farnesyl diphosphate synthase. The release of the condensed product, FPP, was suggested to be the slowest step of catalysis.[164]

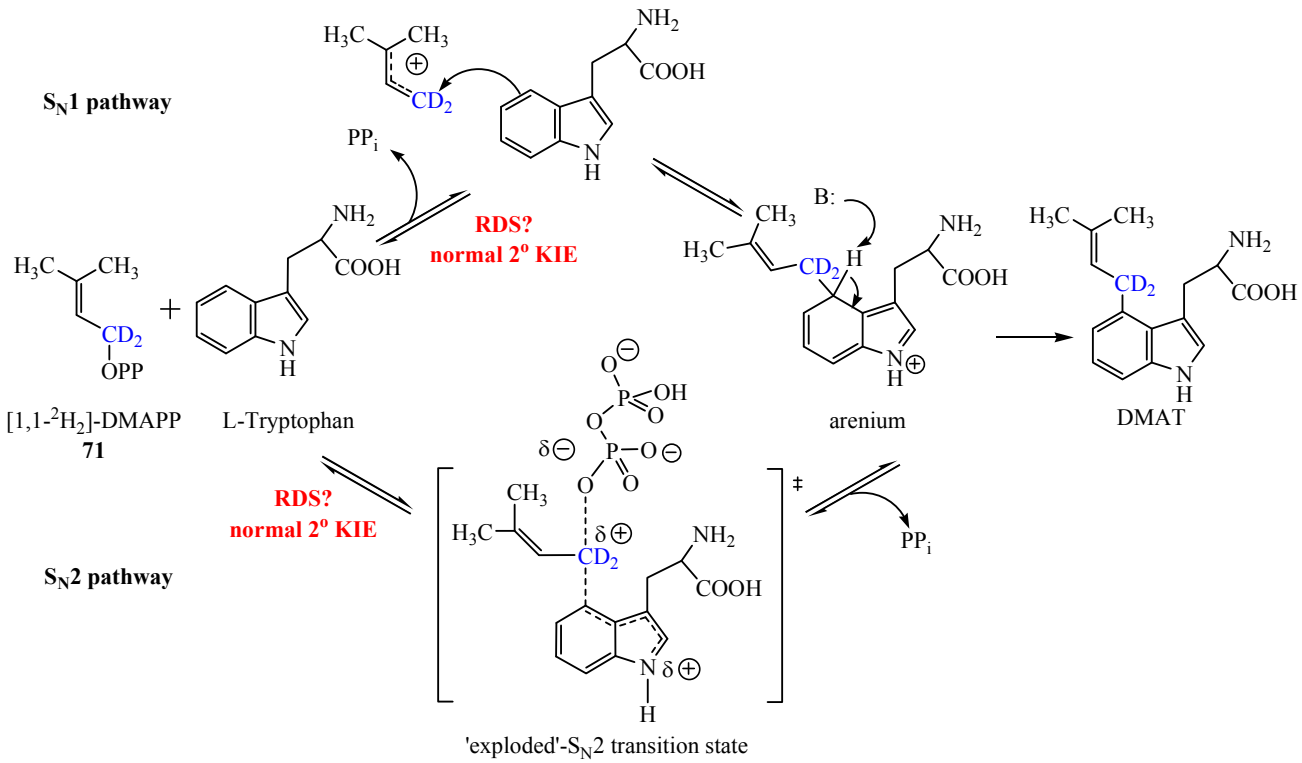


Figure 4.19 Possible isotope effects that can occur upon the use of [1,1-$^2H_2$]-DMAPP (71)

The synthesis of [1,1-$^2H_2$]-DMAPP (**71**) was modified from the synthesis of a similar deuterated compound (Figure 4.20).[165] Dimethylallylic acid (**73**) was reduced by lithium aluminum deuteride (LAD) to give the corresponding alcohol (**74**). After purification by Kugelrohr distillation, the deuterated alcohol was coupled with TEAP in the presence of trichloroacetonitrile to yield the product **71**.[165] The overall yield was ~15% and deuterium incorporation was measured to be 92% as determined by negative ESI-MS.

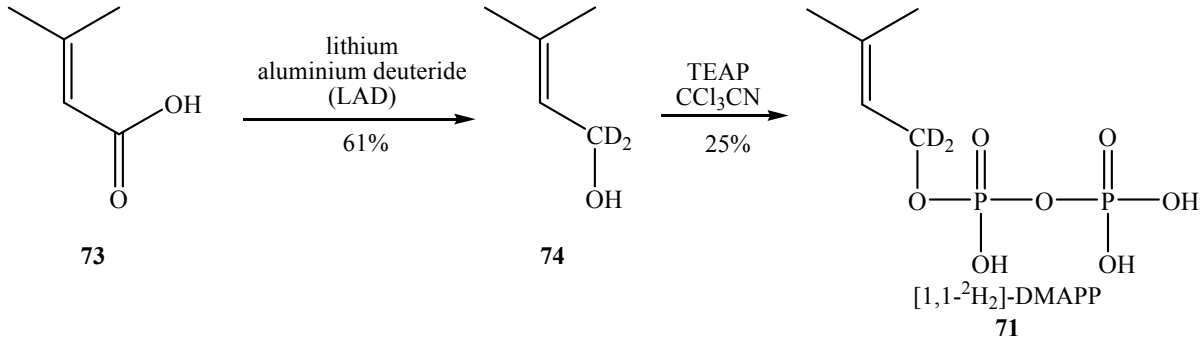


Figure 4.20 Synthesis of [1,1-$^2H_2$]-DMAPP (71)

An equimolar solution (based on negative ESI-MS) of DMAPP and [1,1-$^2H_2$]-DMAPP (**71**) (~12 mM each) was prepared in Tris-HCl buffer containing L-tryptophan (26 mM) and 5mM $CaCl_2$, DMATS (1.2 mg) was then added. ^{31}P NMR spectroscopy was used to determine the value of F_H by comparing the integrals between the α-phosphate signal and an internal standard of 2-deoxy-glucose-6-phosphate as a function of time. To ensure that the integrations of the phosphorus signals were truly proportional to the concentrations of the phosphate compounds, the relaxation time (t_2) between each scan was prolonged to 4.5 s. The values of R_o and R were measured using negative ESI-MS and comparing the relative peak area for the labeled (m/z 247 M-H^+) and unlabeled (m/z 245 M-H^+) substrates as a function of time (Figure 4.21).

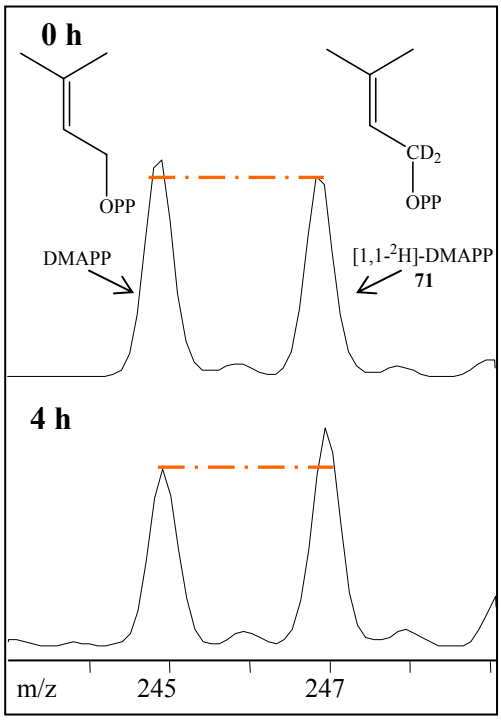


Figure 4.21 Mass spectra following the consumption of a mixture of unlabeled DMAPP and [1,1-$^2H_2$]-DMAPP (71) in the DMATS reaction. Top spectrum was taken prior to the addition of the enzyme, and bottom spectrum after 75% of the unlabeled DMAPP has been converted to product.

Prior to the addition of enzyme, the ratio of labeled to unlabeled DMAPP was approximately 50.2:49.8 as shown in the mass spectral traces (Figure 4.21). After 75% of the unlabeled material

had been converted to product, the ratio became 45.2:54.8. Although the change in ratio was relatively subtle, this observation was reproducible. Because the two isotopologues are 2 m/z apart on the mass spectra, the peak intensity representing the deuterated species did not need to be corrected for the natural abundance of heavy isotopes. The fact that the protiated substrate was consumed faster than the deuterated substrate during catalysis indicates that a normal secondary isotope effect is observed. The measurements were performed in triplicate (Table 4.2), and an average value was calculated to be 1.16 ± 0.06.

Table 4.2 Summary of k_H/k_D calculations obtained with [1,1-$^2H_2$]-DMAPP (71)

Trials	a_D/a_H	R/R_o	F_H	k_H/k_D
1	0.781 ± 0.044*	1.000 ± 0.071	0	-
	0.953 ± 0.048	1.082 ± 0.076	0.751 ± 0.05**	1.167 ± 0.075***
2	0.995 ± 0.050	1.000 ± 0.071	0	-
	1.226 ± 0.061	1.232 ± 0.087	0.733 ± 0.05	1.188 ± 0.083
3	0.958 ± 0.048	1.000 ± 0.071	0	-
	1.116 ± 0.056	1.165 ± 0.082	0.738 ± 0.05	1.128 ± 0.077
			Average	1.16 ± 0.06◊

* An approximately equimolar solution of DMAPP and [1-$^2H_2$]-DMAPP (**71**) was subjected to negative ESI-MS for 5-6 separate trials, and relative error for a_D/a_H was estimated to be ~5%.

** Three separate ^{31}P NMR spectra (100 scan each) of a mixture of DMAPP and 2-deoxy-glucose-6-phosphate (10 mM each in 1 mL) were obtained, and the ratio of the integrals for the DMAPP α-phosphorus signal to the anomeric phosphate signal showed 3% deviation among the spectra. The relative error value for F_H.was then estimated to be ~7%.

*** The propagated errors due to measurement limitations were indicated in the k_H/k_D value calculated in each trial.[77]

◊ The average value was reported within the 95% confidence interval.

4.5.2 Kinetic Isotope Effect Measurement with [4-^{2}H]-L-Tryptophan

To continue our investigation, [4-^{2}H]-L-tryptophan (**72**) was employed in another KIE measurement (Figure 4.22). In order to explain the isotope scrambling that was observed in the

previously described PIX experiment, formation of the carbocation/pyrophosphate ion pair (dimethylallyl carbocation in S_N1 or arenium ion in S_N2) needs to be a reversible step of the overall catalysis (Figure 4.18). When deuterated L-tryptophan (**72**) is employed as a substrate, the magnitude of KIE will reflect the reaction pathway that is involved. In the case of an S_N1 pathway, an inverse secondary KIE is anticipated as formation of the C-C bond is also a second partially rate-limiting step of catalysis (Figure 4.22). In the case of a reversible S_N2 pathway, however, a primary deuterium KIE is expected because re-aromatization should be the rate-limiting step. Hence, taken together, the PIX and KIE experimental results should elucidate the reaction pathway and probe the rate-determining step(s) of the DMATS reaction.

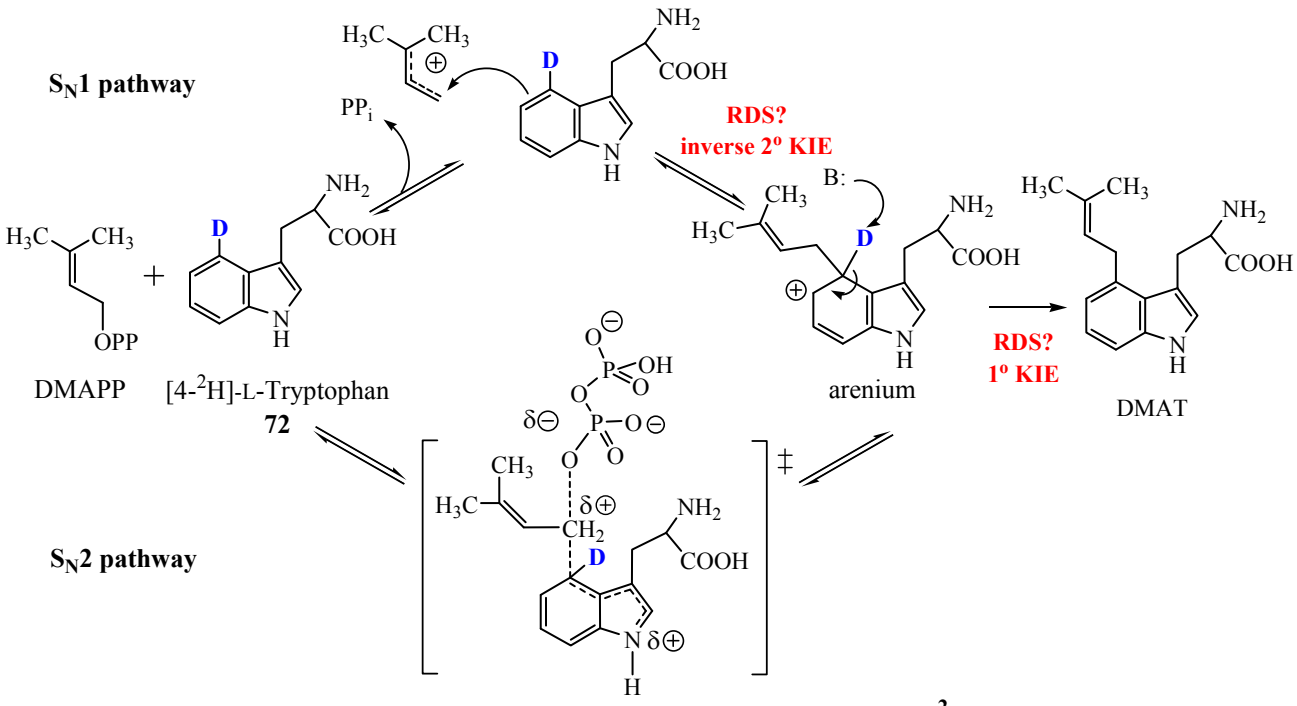
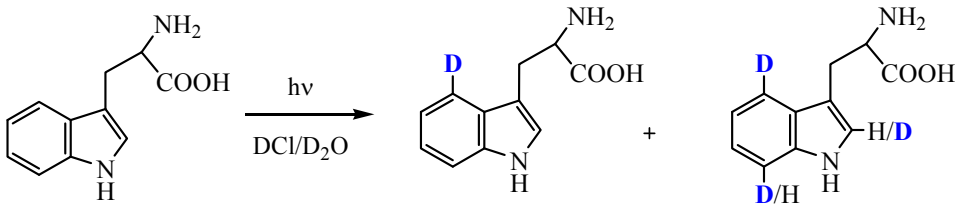


Figure 4.22 Possible isotope effects that can occur upon the use of [4-^{2}H]-L-tryptophan (72) and are consistent with the observation of PIX

[4-^{2}H]-L-tryptophan (**72**) was made by subjecting L-tryptophan to a photoreaction in a mixture of degassed DCl/D_2O (Figure 4.23).[166] The extent of deuterium incorporation at C-4 was ~90% as determined by ^{1}H NMR spectroscopy. Minor deuteration was also observed at the C-2 and C-7 positions, and caused 16% of the labeled product to be di-deuterated. The calculations used to

determine the KIE were adjusted to account for the presence of this di-deuterated isotopologue (see below).



L-tryptophan | [4-^{2}H]-L-tryptophan (**72**) ~74% | di-deuterated-L-tryptophan ~16%

Figure 4.23 Photoreaction of [4-^{2}H]-L-tryptophan (72)

A solution containing equimolar amounts (based on positive ESI-MS) of L-tryptophan and [4-^{2}H]-L-tryptophan (**72**) (~12 mM each) was allowed to react with DMAPP (26 mM) and DMATS (1.2 mg). In this set of experiments, the F_H value was accessed by subjecting the quenched reaction to HPLC separation and comparing the area of the peaks corresponding to tryptophan and an internal standard, 1-methyl-D,L-tryptophan.

Positive ESI-MS was used to evaluate the initial R_o and final R values of the deuterated (m/z 206 M+H$^+$) and protiated (m/z 205 M+H$^+$) substrates (Figure 4.24). Because the mass of the two isotopologues only differed by 1 m/z, it is necessary to correct for the natural abundance of the ^{13}C isotopes. That is, the observed intensities of the mass spectral peaks (i_{205} and i_{206}) needs to reflect the relative number of molecules of the two isotopologues (a_H and a_D). The percentage of undeuterated tryptophan that will appear at m/z 206 due to the presence of ^{13}C isotope at natural abundance was designated as P and determined to be 11.6%. Furthermore, di-deuterated tryptophan (i_{207}) was made as an undesired by-product during the synthesis of [4-^{2}H]-L-tryptophan (**72**). It was assumed that all (or almost all) of this di-deuterated L-tryptophan contains a deuterium at the C-4 position. The a_H and a_D can subsequently be calculated as follows:

$$a_H = \frac{i_{205}}{(1 - P)}$$

$$a_D = i_{206} - a_H (P) + i_{207}$$

equations (4.3) and (4.4)

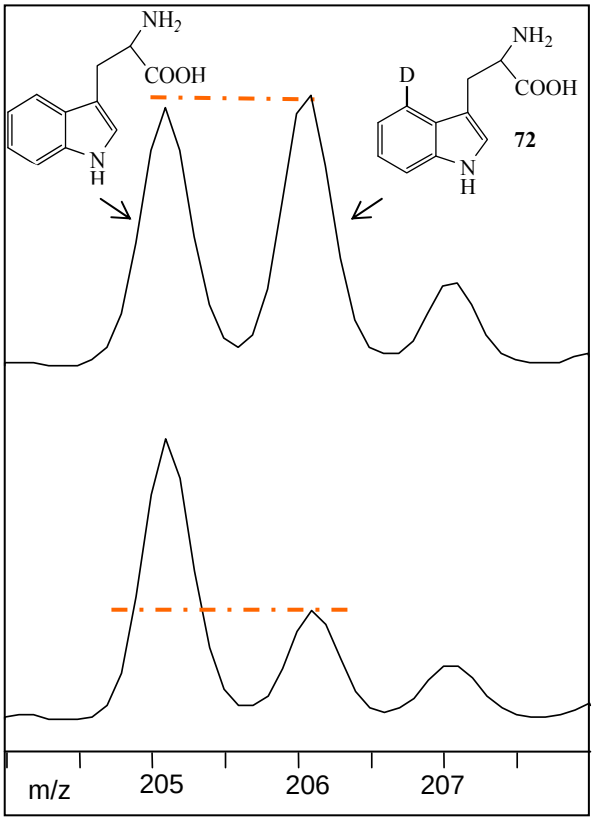


Figure 4.24 Mass spectra following the consumption of a mixture of unlabeled L-tryptophan and [4-^{2}H]-L-tryptophan (72) in the DMATS reaction. Top spectrum was taken prior to the addition of the enzyme, and bottom spectrum after 70% of the unlabeled L-tryptophan has been converted to product.

As indicated in Figure 4.24, the deuterated isotopologue was consumed more rapidly than the protiated one in the DMATS reaction, indicating that an inverse isotope effect has occurred. The KIE measurements were performed in triplicate and an average k_H/k_D value was calculated to be 0.81 ± 0.04. A sample of data collection is shown in Table 4.3.

Table 4.3 Summary of k_H/k_D calculations obtained with [4-^{2}H]-L-tryptophan (72)

Trials	a_D/a_H	R/R_o	F_H	k_H/k_D
1	0.942 ± 0.047*	1.000 ± 0.071	0	-
	0.594 ± 0.030	0.631 ± 0.045	0.821 ± 0.057**	0.789 ± 0.048***
2	0.947 ± 0.047	1.000 ± 0.071	0	-
	0.622 ± 0.031	0.656 ± 0.046	0.827 ± 0.057	0.807 ± 0.049
3	0.851 ± 0.043	1.000 ± 0.071	0	-
	0.648 ± 0.032	0.761 ± 0.054	0.738 ± 0.057	0.831 ± 0.067
			Average	0.81 ± 0.04$^{\lozenge}$

* An approximately equimolar solution of L-tryptophan and [4-^{2}H]-L-tryptophan (**72**) was subjected to positive ESI-MS for 5 separate trials, and relative error for a_D/a_H was estimated to be ~5%.

** Three separate HPLC runs of a mixture of L-tryptophan and 1-methyl-L-tryptophan (8 mM each in 2 mL water) were carried out, and the ratio of the areas of the tryptophan peak to the 1-methyl-trytophan peak showed ~3% deviation. The relative error value for F_H.was then estimated to be ~8%.

*** The propagated errors due to measurement limitations were indicated in the k_H/k_D value calculated in each trial.[77]

$^{\lozenge}$ The average value was reported within the 95% confidence interval.

According to the experimental results, the use of the deuterated substrates [1,1-$^2H_2$]-DMAPP (**71**) and [4-^{2}H]-L-tryptophan (**72**) resulted in normal secondary and inverse secondary KIEs, respectively. This implies that both the formation of the dimethylallyl carbocation and the arenium intermediates are each partially rate-determining steps in the reaction catalyzed by DMATS. The occurrence of PIX, therefore, must occur due to the reversible formation of the carbocation/pyrophosphate pair in an S_N1 process (Figure 4.18). The absence of a primary KIE with [4-^{2}H]-L-tryptophan (**72**) indicated that the re-aromatization of the arenium intermediate is not the rate-determining step in the reaction. However, this step needs to be rate-determining in order to explain the PIX reaction, thus the reversible S_N2 reaction pathway can be ruled out by the KIE results.

4.6 PIX Experiment with Fluorinated L-Tryptophan Analogs

In an attempt to further characterize the S_N1 pathway of the DMATS reaction (Figure 4.18), another set of PIX experiments was conducted using different fluorinated tryptophan inhibitors (**75** and **76**, Figure 4.25). If the ^{18}O-scrambling in the PIX reaction could proceed to completion, the ^{18}O-label on the non-bridging position to the bridging position should reach the statistical distribution ratio of 2:1. However, this was not observed in the previously described PIX reaction (see Section 4.4.4). An incomplete isotope scrambling could be due to either a restriction on the torsional equilibration of the pyrophosphate and/or a short life-time of the allylic cation intermediate before the nucleophilic attack by L-tryptophan and the irreversible deprotonation of the arenium intermediate. To test this theory, in a new set of PIX experiments, L-tryptophan was replaced by its fluorinated analogs, which are likely to bind but would have greatly reduced nucleophilicity (Figure 4.25). Since the nucleophilic attach by tryptophan is prohibited, this may allow for reversible ion pair formation that would increase the effective life time of the dimethylallyl carbocation leading to a full extent of PIX. 4-Fluoro-D,L-tryptophan (**75**) should not serve as a substrate, because the fluorine atom replaces the proton that needs to be deprotonated in the enzyme-catalyzed reaction. In the substrate analog studies conducted by Li and his co-workers, 6-fluoro-D,L-tryptophan (**76**) was found to be non-reactive in the DMATS reaction, whereas the analog carrying a methyl group at the C-6 position reacted with ~15% of the activity observed with L-tryptophan.[130] This implied that low reactivity of the C-6 fluorinated analog is mainly due to the electron-withdrawing effect of the fluorine atom.

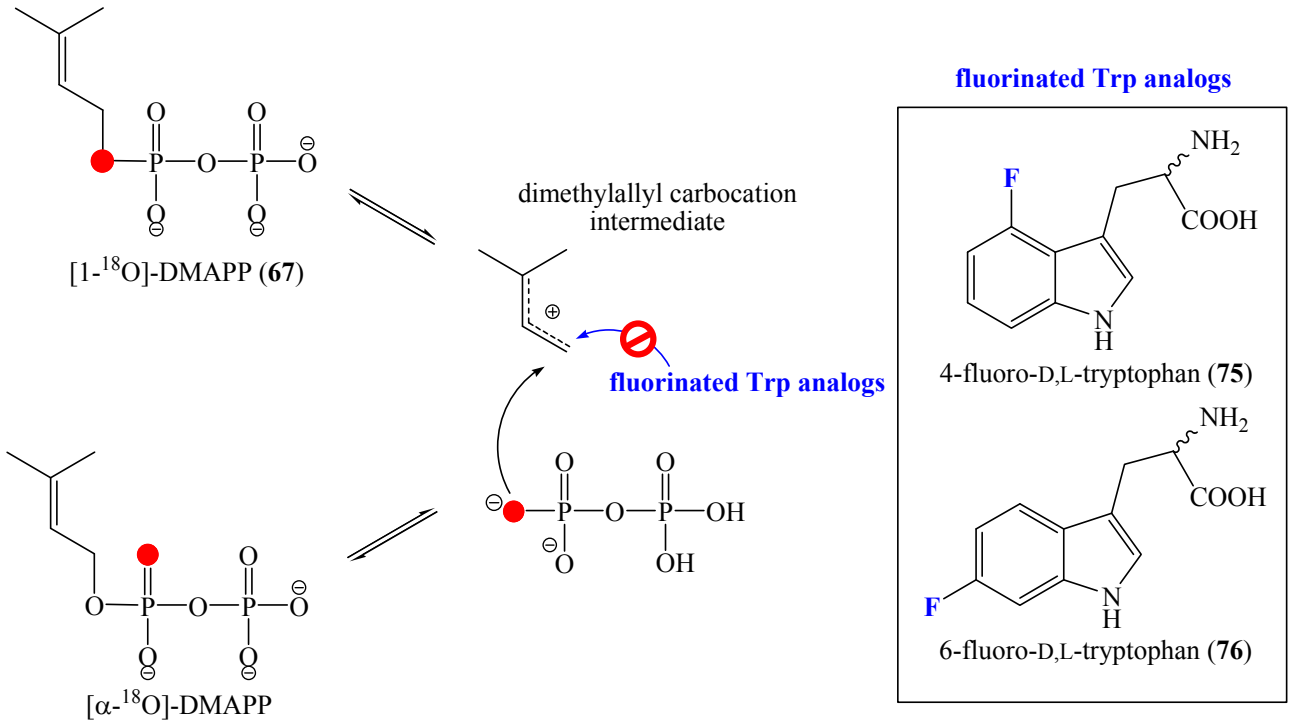


Figure 4.25 PIX experiment with 4-fluoro- and 6-fluoro-D,L-tryptophan analogs (75 and 76). Solid red circle represents ^{18}O label.

The analogs **75** and **76** were both commercially available in racemic forms and the PIX reactions were carried out in Tris-HCl buffer containing EDTA (3 mM). The reaction conditions were similar to those described in the previous PIX reaction (Section 4.4.4), but some modifications were made to facilitate the use of these analogs. In previous experiments only 0.8 equivalents of L-tryptophan was employed, because it readily reacts to give DMAT and the concentration of the recovered DMAPP might become too low to provide a well-resolved ^{31}P NMR spectrum. Since the fluorinated analogs are not reactive, this was no longer a concern and the concentrations of the analogs were increased to 2.3 equivalents. Lastly, the amount of DMATS was increased two-fold (~2 mg) in the hope of optimizing the isotopic scrambling reaction.

In the PIX reaction employing analog **75**, analysis of the ^{31}P spectra indicated that no isotopic scrambling had occurred. After 24 h incubation with the synthase, the two α-phosphorus doublets representing the unlabeled and [1-^{18}O]-DMAPP (**67**) (at -9.28 and -9.30 ppm, respectively) remained

unchanged and formation of the non-bridging [α-^{18}O]-DMAPP (doublet at -9.31 ppm) was not observed (Figure 4.26). A likely explanation is that the binding of **75** may prevent the binding of DMAPP, or vice versa, due to the steric effect of the fluorine substitution, thus prohibiting the isotope exchange.

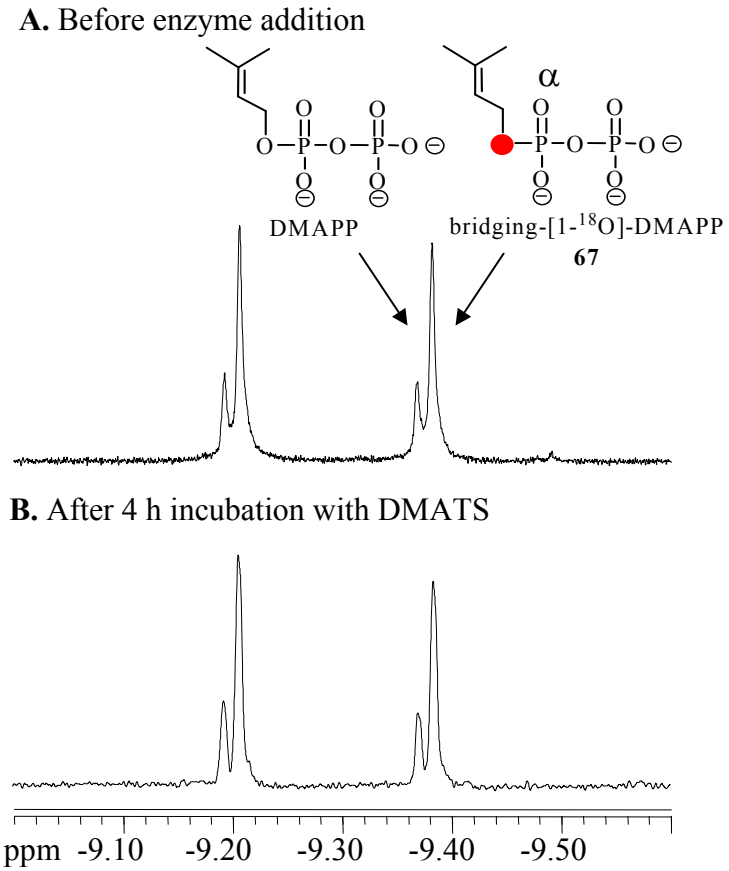


Figure 4.26 ^{31}P spectra (D_2O) showing the α-phosphorus signals of the DMAPP mixture in the PIX reaction employing 4-fluoro-D,L-tryptophan (75). A) before enzyme addition, and B) after 24 h incubation with DMATS. Solid red circle represents ^{18}O label.

The use of 6-fluoro-D,L-tryptophan (**76**) in another PIX reaction fortunately showed the full extent of isotope scrambling. Prior to enzyme addition, the α phosphorus signals of the unlabeled DMAPP and [1-^{18}O]-DMAPP (**67**) appeared as two separate doublets at δ -9.28 and -9.30 ppm in the ^{31}P NMR spectrum, respectively (Figure 4.27A). After 4 h incubation with DMATS at 37 ^{o}C, the new upfield doublet at δ -9.31 ppm that represents [α-^{18}O]-DMAPP was observed (Figure 4.27B).

The distribution of the ^{18}O isotope obeyed the statistical 2:1 ratio between the non-bridging position and the bridging position, implying that isotopic scrambling had reached completion.

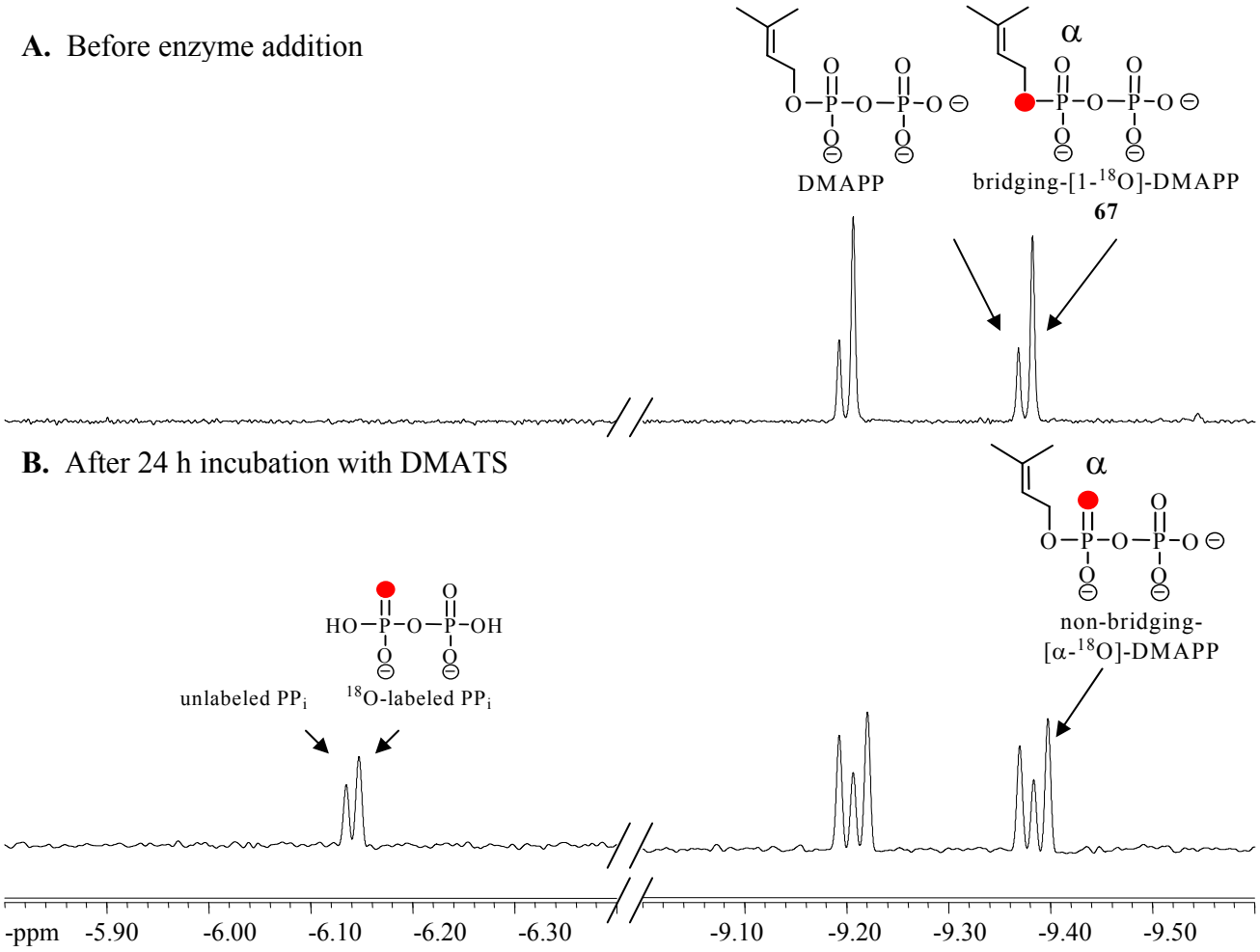


Figure 4.27 ^{31}P spectra (D_2O) of the DMAPP and the [1-^{18}O]-DMAPP (67) in the PIX reaction employing 6-fluoro-D,L-tryptophan (76). A) Prior to enzyme addition, and B) after 24 h incubation with the enzyme.

Surprisingly, approximately 9% of the phosphorus signals appeared as a set of singlets at δ -6.12 and -6.13 ppm, which correspond to the unlabeled and the non-bridging ^{18}O-labeled pyrophosphate, respectively (Figure 4.27B). It was first suspected that the fluorinated tryptophan analog **76** might have served as a substrate so that DMAPP has been consumed to give pyrophosphate. However, HPLC analysis did not show any peak corresponding to fluorinated DMAT. Other possibilities are the dimethylallyl carbocation was occasionally intercepted by a

water molecule inside the active site of the enzyme or non-enzymatic hydrolysis was occurring during incubation/work up; further experiments are needed to test these theories.

4.7 Measurements of Kinetic Isotope Effect using [4-^{2}H]-L-Tryptophan and Fluorinated DMAPP Analogs

Taken together, the PIX and KIE experimental results supported the formation of a dimethylallyl carbocation in an S_N1 pathway. However, in the kinetic studies employing the fluorinated DMAPP analog ***E*-52**, the k^{rel} parameter was indicative of an associative transition state, because its value on the same order of magnitude as the k^{rel} parameters that were measured in different S_N2 reactions (see Section 4.3.1).[113] A likely explanation is that DMATS adopted an S_N2 pathway in order to accommodate analog **52**, because the corresponding fluorinated dimethylallyl carbocation is highly unstable. To examine this theory, the KIE obtained with [4-^{2}H]-L-tryptophan (**72**) was determined using the fluorinated analog **52** instead of DMAPP. If an S_N2 mechanism were operative, an inverse isotope effect will be anticipated because the associative transition state should be rate-determining and have considerable carbocationic character. If an S_N1 pathway is operative, however, a KIE near unity will be expected because formation of the unstable fluorinated dimethylallyl carbocation should be fully rate-determining in catalysis.

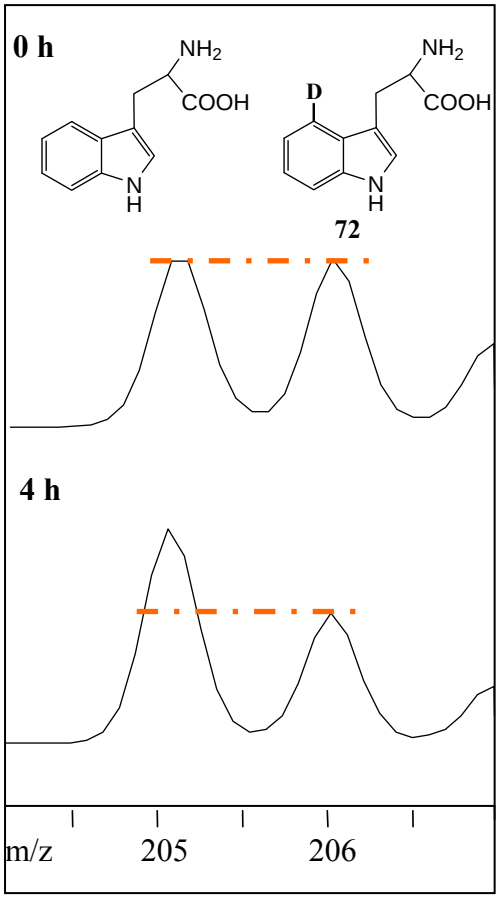


Figure 4.28 Mass spectra following the consumption of a mixture of unlabeled L-tryptophan and [4-^{2}H]-L-tryptophan (72) in the DMATS reaction. Top spectrum was taken prior to the addition of DMATS, and bottom spectrum after 70% of the unlabeled DMAPP has been converted to product.

The fluorinated analog **52** reacts ~20 fold slower than the natural substrate DMAPP in catalysis (see Section 4.2), thus the concentrations of the diphosphate substrates used in the KIE measurement were reduced to 2 mM. The F_H and *R*'s values were estimated by HPLC separations and by positive ESI-MS spectral methodologies as elaborated in Section 4.5.2. As depicted in the mass spectra in Figure 4.28, a secondary inverse isotopic effect was observed in the reaction between analog **52** and [4-^{2}H]-L-tryptophan because the deuterated species was consumed to a greater extent. The measurement was carried out twice and the k_H/k_D value was determined to be 0.70 ± 0.05 (Table 4.4). Hence, an associative transition state was likely involved when fluorinated DMAPP (**52**) served as a substrate in the DMATS reaction.

Table 4.4 Summary of k_H/k_D calculations for L-tryptophan in DMATS reaction using *E*-3-fluoromethyl-2-butenyl diphosphate (*E*-52)

Trial	a_D/a_H	R/R_o	F_H	k_H/k_D
1	0.926 ± 0.046*	1.000 ± 0.071	0	-
	0.712 ± 0.036	0.556 ± 0.054	0.472 ± 0.033**	0.708 ± 0.110***
2	0.744 ± 0.037	1.000 ± 0.071	0	-
	0.393 ± 0.020	0.472 ± 0.037	0.762 ± 0.053	0.692 ± 0.047
			Average	0.70 ± 0.02◊

* An approximately equimolar solution of L-tryptophan and [4-^{2}H]-L-tryptophan (**72**) was subjected to positive ESI-MS for 5 separate trials, and relative error for a_D/a_H was estimated to be ~5%.

** Three separate HPLC runs of a mixture of L-tryptophan and 1-methyl-L-tryptophan (8 mM each in 2 mL water) were carried out, and the ratio of the areas of the tryptophan peak to the 1-methyl-trytophan peak showed ~3% deviation. The relative error value for F_H.was then estimated to be ~8%.

*** The propagated errors due to measurement limitations were indicated in the k_H/k_D value calculated in each trial.[77]

◊ The average value was reported within the 95% confidence interval.

4.8 Discussion and Related Studies

In nature, prenyltransferases (PTs) represent a diverse group of enzymes that catalyze the transfer of an isoprenyl group from its diphosphate form onto a variety of acceptors. These enzymes can be categorized into three groups: isoprenyl diphosphate synthases, protein PTs and aromatic PTs. Dimethylallyltryptophan synthase (DMATS) was the first aromatic PT to be discovered about 30 years ago. The enzyme can either employ an S_N1 pathway that involves formation of a dimethylallyl carbocation or an S_N2 pathway that involves an 'exploded' transition state (Figure 4.13). Mechanistic investigations on this enzyme have been conducted,[15, 16] but the reaction pathway employed by this enzyme for catalysis remained unidentified. In this thesis, we present experimental results that strongly support the S_N1 pathway for the DMATS reaction.

The results obtained from the PIX and the KIE experiments clearly illustrate that a discrete dimethylallyl cation is formed in the DMATS-catalyzed reaction. The most convincing evidence was provided by the fact that DMATS catalyzed an isotopic scrambling reaction of [1-^{18}O]-DMAPP (**67**) to give [α-^{18}O]-DMAPP. When L-tryptophan was used as the complement substrate in the PIX reaction, a small amount of [α-^{18}O]-DMAPP was generated, implying that the dimethylallyl carbocation/pyrophosphate pair was reversibly formed during catalysis. Furthermore, the carbocationic character generated in the allylic position of DMAPP was supported by the normal deuterium secondary KIE observed upon the use of [1,1-$^{2}H_2$]-DMAPP (**71**). The reversible dissociation of pyrophosphate from DMAPP indicates that one of the chemical steps subsequent to it should be partially rate-determining as well. The inverse secondary KIE observed with [4-^{2}H]-L-tryptophan (**72**) supported this idea, as it showed that formation of the arenium ion intermediate is the second partially rate-determining step of the reaction. The reversible S_N2 mechanism can also be ruled out by the KIE results, because re-aromatization of DMAT would need to be rate-determining in order to explain the PIX reaction. Lastly, the calculated partitioning ratio, v_{pix}/v_{chem}, in the PIX reaction (0.40) is notably lower than the theoretical v_{pix}/v_{chem} (1.75) calculated for the full extent of PIX. This implies that the lifetime of the dimethylallyl carbocation/pyrophosphate intermediate was too transient for the oxygens within pyrophosphate to experience complete torsional equilibration.

This transient formation of a dimethylallyl carbocation is further supported by the observation of a full extent of PIX in the reaction using an unreactive analog, 6-fluorinated-D,L-tryptophan (**76**). It is inconsistent with the S_N2 pathway, since the fluorine substitution should increase the barrier to arenium ion formation rendering this irreversible and preventing PIX. Instead, it strongly supported the S_N1 mechanism in which a dimethylallyl cation/pyrophosphate ion pair is formed and collapses at a rate that is much faster than attack by the electron deficient fluorotryptophan. No scrambling of the isotopic label from the α- to the β-phosphorus of DMAPP

could be detected in any of these PIX reactions, indicating that pyrophosphate can not tumble within the active site and is not released/rebound during the lifetime of the carbocation intermediate.[167]

To our knowledge this is the first observation of PIX catalyzed by a prenyltransferase. Previously, similar PIX experiments have been performed on *trans*-farnesyl diphosphate synthase (see Figure 3.9 and Section 3.2) and its paralog, bornyl diphosphate synthase, but ^{18}O-scrambling was not observed (Figure 4.29).[108, 167] While these enzymes are thought to generate a cationic intermediate, they are also metal-dependent and it is likely that the metal binds to the phosphate groups of the pyrophosphate intermediate and restricts bond rotations that could lead to PIX. Since DMATS does not require a divalent cation for catalysis, it is possible to observe the scrambling phenomenon. Perhaps the closest precedence to this study can be found in the enzyme farnesyl diphosphate isomerase where PIX was used to establish the existence of a carbocation intermediate (Figure 4.29).[167]

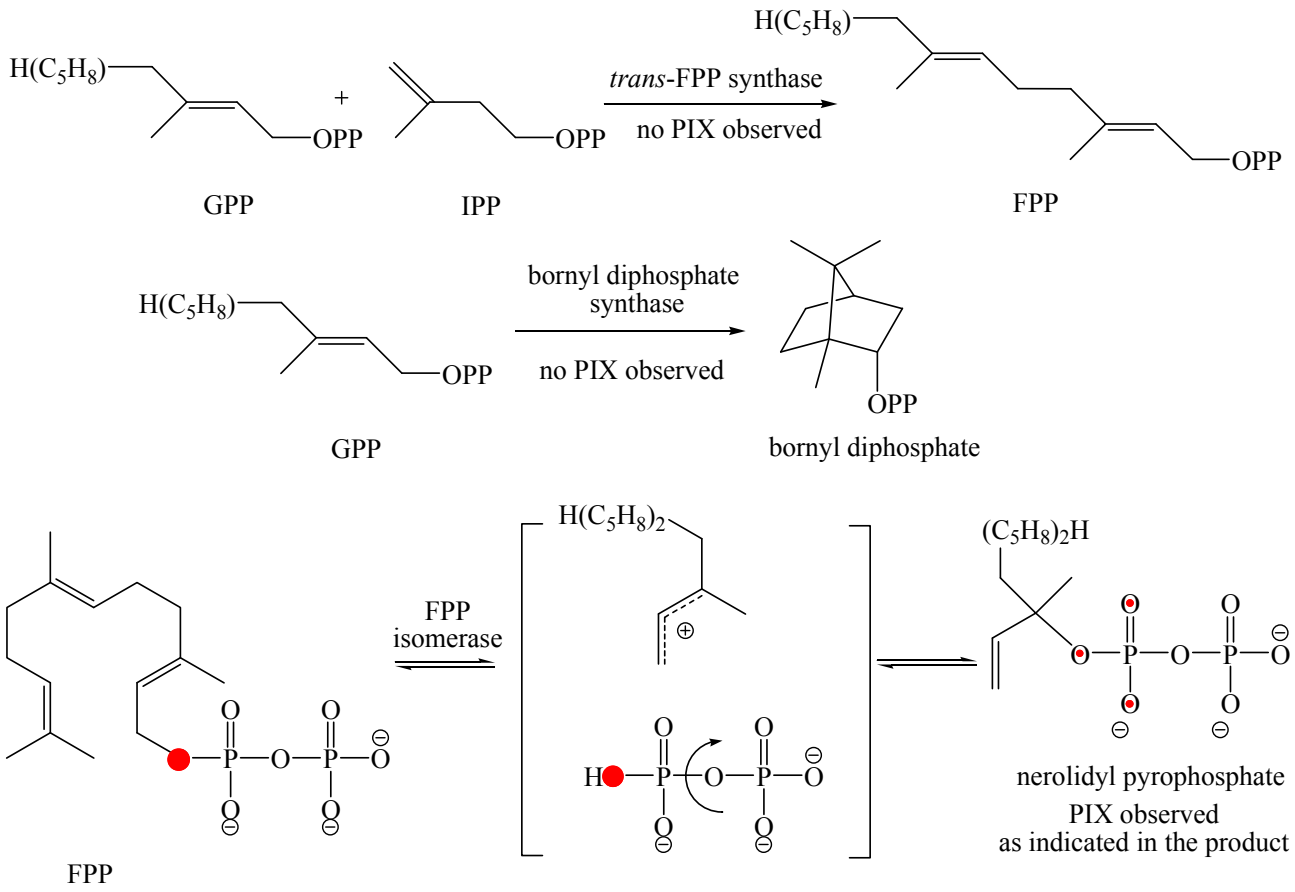


Figure 4.29 PIX experiments that were carried out on other prenyltransferases and related enzymes.[108, 167]

4.8.1 Crystallographic Studies on DMATS

The crystal structure of DMATS was also reported by the research group of Stehle,[126] at approximately the same time as our mechanistic studies on this enzyme were published[168]. While different homologs of DMATS have been discovered and the family of known indole prenyltransferases is rapidly expanding, information regarding the protein structure was lacking. Therefore, solving the structure of DMATS represents an important milestone in our understanding of indole prenyltransferases.

The structure of DMATS is comprised of a rare ABBA rearrangement, which is also known as the prenyltransferase (PT) barrel fold (Figure 4.30).[126] The active site of the enzyme is shaped by a central barrel of 10 antiparallel β-strands, which is surrounded by a ring of solvent-exposed α-helices. The protein fold is reminiscent of the TIM barrel fold of the enzyme triosephosphate isomerase.[169] However, the cavity of TIM barrel is constructed from parallel β-strands, thus it is not structurally homologus to DMATS. The ABBA rearrangement was first discovered in another aromatic PT, NphB, which catalyzes the prenylation reaction of naphterpin and was isolated from a Gram-positive bacterial species *Streptomyces* sp. strain CL190 (see Section 3.4.1).[126] Interestingly, a protein sequence alignment of DMATS and NphB reveals only 9% identity and 29% similarity. Given the highly diverged primary sequence, the conservation of the protein architectures probably arises as a consequence of convergent evolution. However, DMATS (459 residues) is significantly larger than NphB (307 residues), as the fungal enzyme contains extended loop regions and extra α-helices.[126]

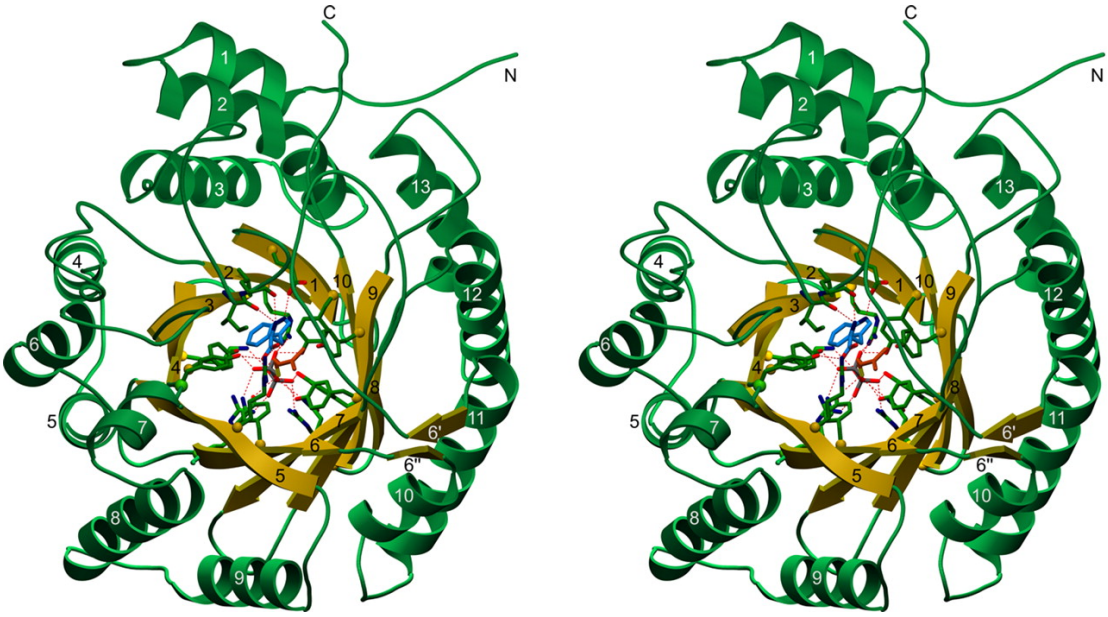
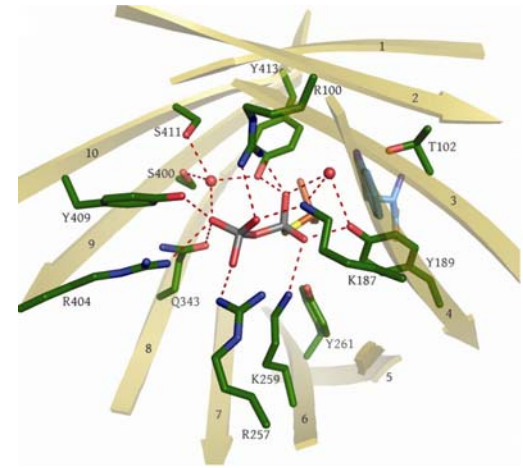
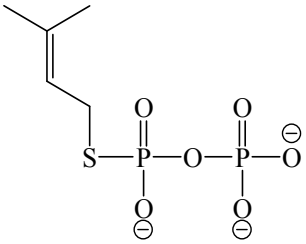


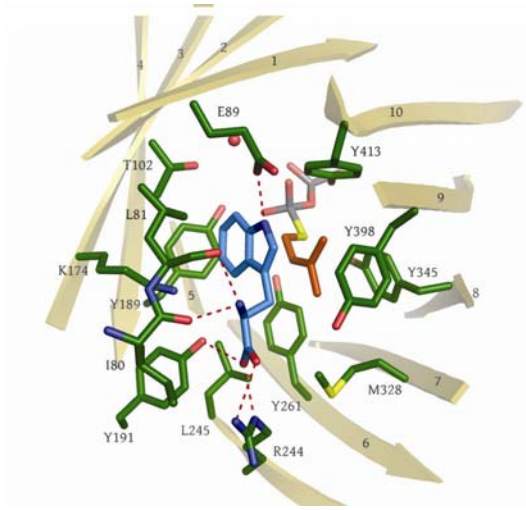
Figure 4.30 The ABBA-fold of DMATS. L-Tryptophan and a substrate analog of DMAPP, **77**, as well as residues that contact them via polar interactions are shown as ball-and-stick models. Diagram obtained from the crystallographic report of the enzyme.[126]

The active site of DMATS was uncovered when the crystal was soaked with the natural substrate, L-tryptophan, and a thiol analog of DMAPP (**77**) (Figure 4.31).[126] A series of positive residues, R100, R257, R404, K187 and K259 were found to interact with the diphosphate group of the DMAPP analog **77**. The positively charged residues are probably responsible for substrate binding, thereby rescuing the enzyme from being metal-dependent. Furthermore, the authors proposed that dissociation of pyrophosphate is promoted by these ionic interactions. This idea is supported by the fact that site-directed mutation of K187 or K259 to a glutamine residue abolished 99.6% of the activity. There is an additional network of hydrogen-bonds between the diphosphate group and a group of aromatic tyrosine residues (Y189, Y261, Y345, Y409 and Y413), which are considered to be important for the binding of substrate as well as stabilizing the carbocationic intermediates via cation-π interactions (Figure 4.31).[126]





thiol analog of DMAPP
77



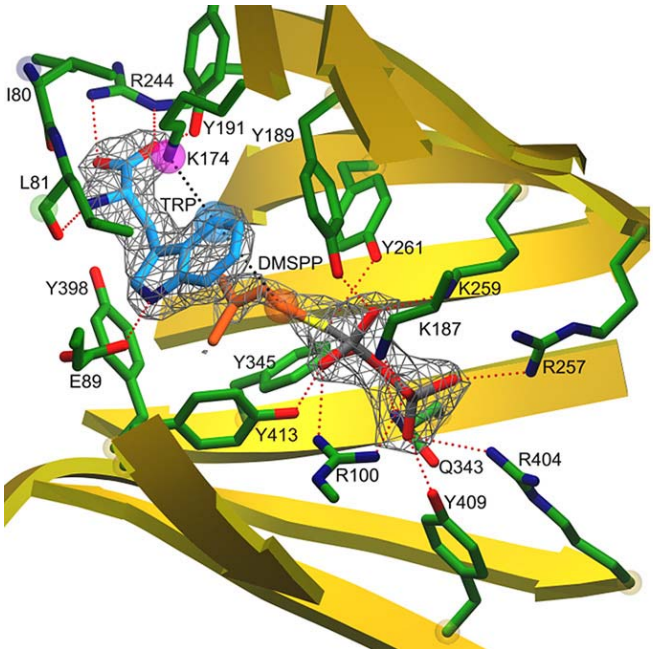


Figure 4.31 The thiol analog of DMAPP (77) and the active site of DMATS from different perspectives. Color code: nitrogen, blue; oxygen, red; phosphate, gray, and sulfur, yellow. Diagram obtained from the crystallographic report on the enzyme.[126]

Some subtle insights into the mechanism of the DMATS reaction were also revealed in the crystallographic studies, and two residues, E89 and K174, were proposed to be crucial for catalysis (Figure 4.32).[126] The carboxylate group of E89 was hydrogen-bonded to the indole N-H, thus enhancing the nucleophilicity of C-4 and promoting the C-C bond formation. When E89 was mutated to an alanine residue, catalytic activity was not detected. In addition, the amine group of K174 was proposed to deprotonate the arenium intermediate leading to the re-aromatization of DMAT (Figure 4.32). Surprisingly, mutation of K174 to a glutamine residue resulted in only a 60% loss of activity. The authors suggested that the carboxylate functionality of tryptophan might replace

the role of K174 and deprotonate the arenium ion (Figure 4.32).[126] However, a more likely explanation is that the pK_a of the arenium proton is probably comparable to the pK_a (~0.5) of a carbonyl amide,[170] thus the K174Q mutant retained a substantial amount of activity.

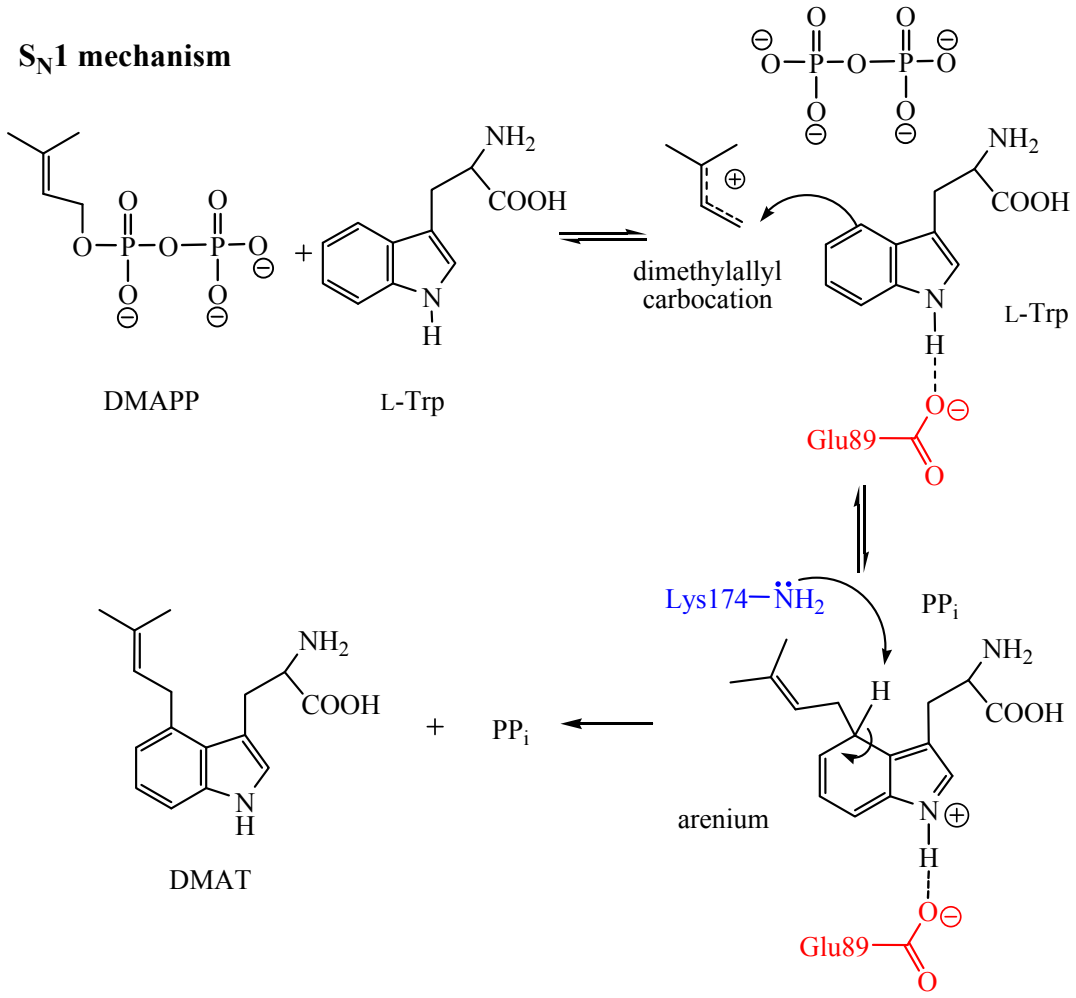


Figure 4.32 The proposed roles of Glu89 and Lys174 in the DMATS reaction

In general, the observations discovered in the crystallographic studies on DMATS agree reasonably well with our experimental results. Specifically, the observation of the interactions between the positively charged residues and the diphosphate moiety of DMAPP might explain the low partitioning ratio, v_{pix}/v_{chem}, calculated in the PIX reaction.[126] The interactions between the positive residues and pyrophosphate might be relatively tight, thereby preventing the oxygen atoms from freely rotating before the dimethylallyl carbocation/pyrophosphate pair collapses. On the other

hand, since a PIX reaction is observed with DMATS but not with the Mg^{2+}-dependent *trans*-FPP synthase, the pyrophosphate is probably less rigidly held than when it is coordinated to a Mg^{2+} cation cofactor.[108, 167]

4.9 Ongoing Studies and Future Directions

In order to complete our mechanistic studies, the precise roles of the amino acid residues involved in the DMATS reaction will be investigated. In the crystallographic report, Glu89 and Lys174 were proposed to play crucial roles in catalysis and site-directed mutation of these residues reduced the catalytic activities to different extents.[126] We envision that the functional roles of Glu89 and Lys174 can be further examined using a combination of techniques, including site-directed mutagenesis, PIX reaction and KIE measurement. Also, results obtained from these experiments should further characterize the carbocationic intermediates formed during the reaction.

It is believed that the nucleophilicity of C-4 is enhanced by the hydrogen-bond between the carboxylate group of Glu89 and the indole N-H (Figure 4.33).[126] If Glu89 is mutated to a weak hydrogen-bond acceptor, E89Q, or a non-polar residue, E89A, the energy barrier to the formation of arenium intermediate should increase and lifetime of the dimethylallyl carbocation intermediate should be extended. Because an alanine residue is non-polar, the enzymatic activity of E89A should be completely (or almost) abolished as has been observed in the crystallographic studies.[126] In turn, a glutamine residue is a weaker hydrogen-bond acceptor, thus the enzymatic activity of E89Q should be retained to a certain extent. Furthermore, PIX reactions will be carried out with these mutants using L-tryptophan as the complement substrate. The extent of PIX, or more precisely the partitioning ratio (v_{pix}/v_{chem} (E89Q and E89A)), should increase substantially as compared to the

v_{pix}/v_{chem} (WT) of the wild type reaction (see Section 4.4.4). Finally, if sufficient activity permits, KIEs will be measured with the mutant enzymes. The inverse secondary KIE observed with [4-^{2}H]-L-tryptophan (**72**) should be greatly enhanced ($k_H/k_D \leq 0.80$) in the E89Q reaction. The normal secondary KIE caused by [1,1-$^2H_2$]-DMAPP (**71**) should be eliminated, as the C-C bond formation will theoretically be cleanly rate-determining (Figure 4.33).

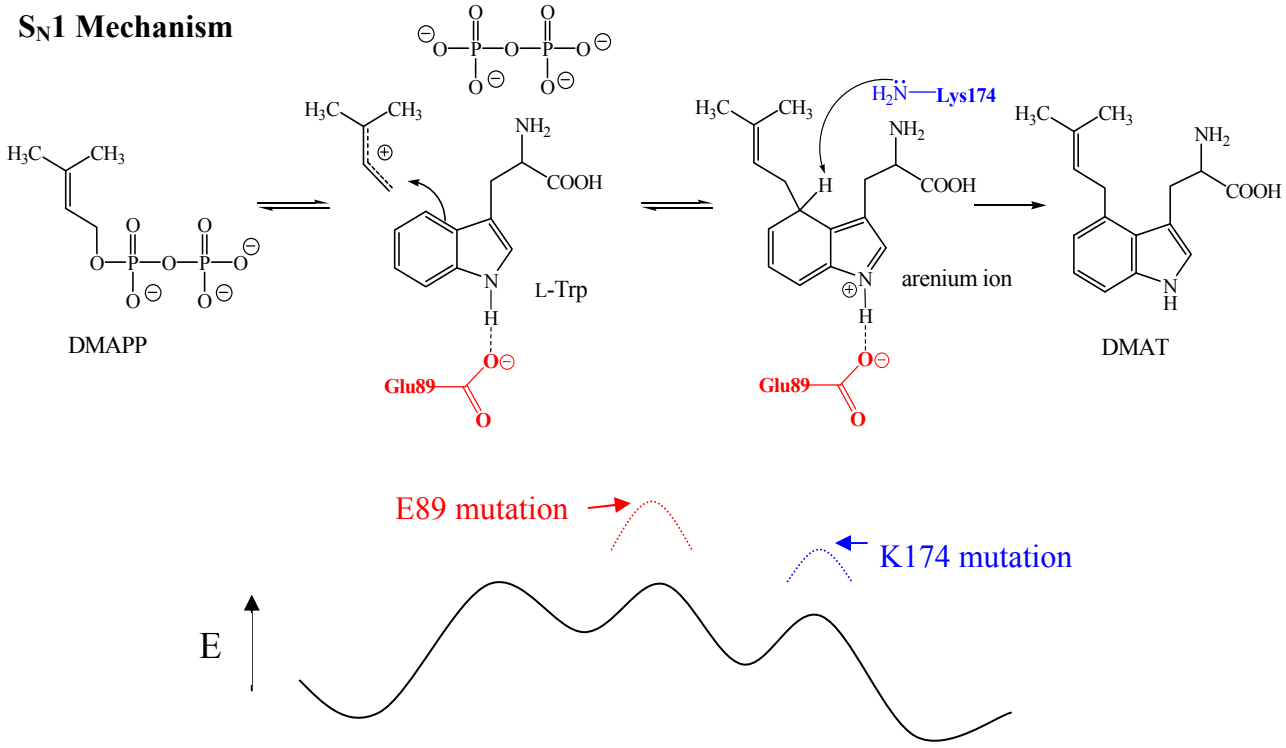


Figure 4.33 The proposed changes in the reaction coordinate of the DMATS reaction upon site-directed mutation of Glu89 and Lys174. Red dotted line refers to the hypothesized change caused by an E89 mutation, and blue dotted line by a K174A mutation.

The amino group of Lys174 was proposed to deprotonate the arenium ion leading to the re-aromatization of DMAT. Mutation of Lys174 to a non-basic residue, therefore, should increase the activation energy of this chemical step and should extend the lifetimes of the carbocationic intermediates (Figure 4.33). The mutants, K174Q and K174A, will be made to test this hypothesis. In the previous studies, K174Q retained ~40% activity, probably because the arenium ion is relatively acidic and was able to protonate the amide functionality of the glutamine residue.

However, the K174A mutant has not been made and its activity remains to be determined. When PIX reactions are carried out with these mutants, a notable increase in v_{pix}/v_{chem} (K174Q and K174A) is anticipated. The deprotonation step should become (partially) rate-determining in the mutant reaction, and ^{18}O-scrambling could occur during the lifetime of either the dimethylallyl carbocation/pyrophosphate ion pair or the reversibly formed arenium/pyrophosphate ion pair. In order to test this theory, KIE measurements will also be carried out. The secondary deuterium KIEs caused by isotopologues **71** and **72** should be eliminated in the mutant reactions. Instead, a primary KIE would be seen upon the use of deuterated tryptophan **72 (Figure 4.33)**.

4.9.1 Site-Directed Mutagenesis and Overexpression of the DMATS Mutants

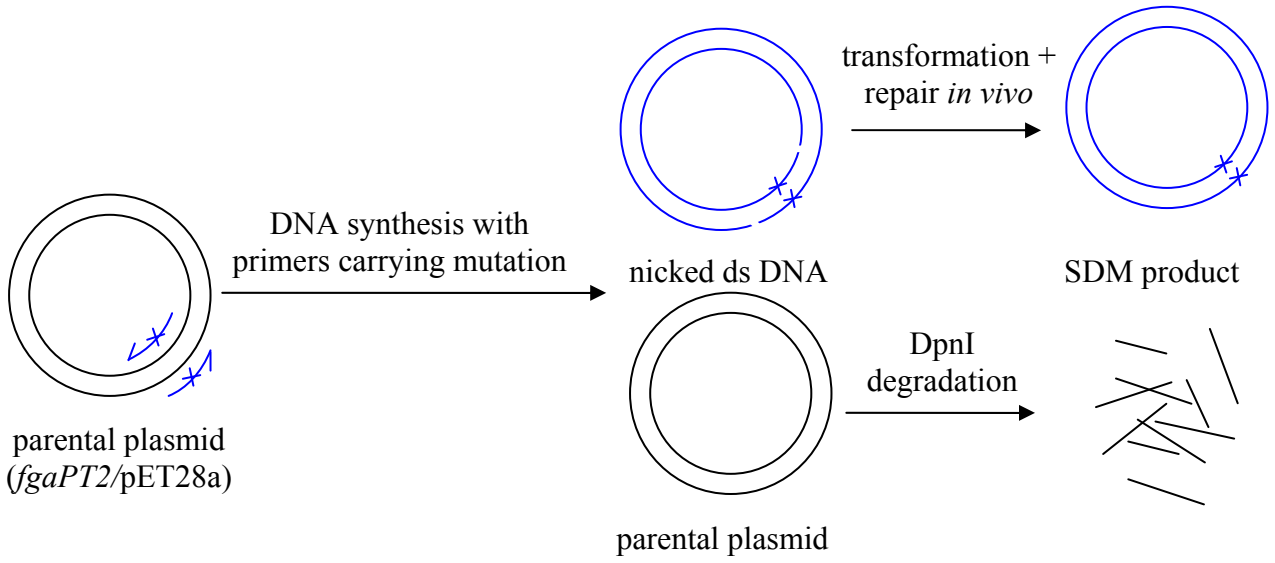


Figure 4.34 Steps involved in the QuikChange® site-directed mutagenesis protocol

The DMATS mutants, E89A, E89Q, K174A and K174Q, were prepared according to the protocol of the QuikChange® Site-Directed Mutagenesis kit from Stratagene (Figure 4.34). Complementary oligonucleotide pairs carrying the mutations are listed in Table 4.5 with altered nucleotides highlighted. For each mutant, the WT plasmid was subjected to polymerase

amplification using the appropriate primers. The crude reaction was then treated with DpnI restriction enzyme, which digested the methylated and non-mutated parental plasmid, and was subsequently transformed into DH5α *E. coli* cells. The transformed bacterial colonies were selected by growing on LB agar plates containing 30 μg/mL kanamycin, and plasmids were isolated from overnight culture. Lastly, the DNA sequence of the DMATS mutants was confirmed by sequencing.

Table 4.5 Primers used in site-directed mutagenesis

Mutants	primer sequences
E89A	(forward) 5'-CGGCGCTCCGTTCGCATTGAGCCTAAATTG-3' (reverse) 5'-CAATTTAGGCTCAATGCGAACGGAGTGCCG-3'
E89Q	(forward) 5'-ATACGGCACTCCGTTCCAATTGAGCCTAAATTGC-3' (reverse) 5'-GCAATTTAGGCTCAATTGGAACGGAGTGCGGTAT-3'
K174A	(forward) 5'-TATCAGGACGCAGAACGCGCTCGCGCTCGATCTG-3' (reverse) 5'-CAGATCGAGCGCGAGCGCGTTCTGCGTCCTGATA-3'
K174Q	(forward) 5'-CTATCAGGACGCAGAACCAGCTCGCGCTCGATCTG-3' (reverse) 5'-CAGATCGAGCGCGAGCTGGTTCTGCGTCCTGATAG-3'

The DMATS mutants were overexpressed and purified by the protocol elaborated in Section 4.1.2. The yields of the mutants were notably lower than that of the wild type, ranging from approximately 10 to 15 mg/L of cell culture (versus ~35 mg of wild type DMATS/L cell culture). Nevertheless, sufficient amounts of enzymes were generated to conduct the PIX experiments described below.

4.9.2 Application of PIX in the Study of DMATS Mutants

The catalytic roles of E89 and K174 were first inspected by running PIX experiments with the corresponding site-directed mutants. The reaction conditions for the mutant-catalyzed PIX

reactions were similar to the ones mentioned in previous studies (see Section 4.4.4 and Section 4.6). In brief, [1-^{18}O]-DMAPP (**67**) and L-tryptophan (0.8 equivalent) were incubated with the DMATS mutants for 24 h in Tris-HCl buffer containing EDTA. The mutants were predicted to be much less active than the wild type DMATS. Thus, the amount of each mutant was increased to ~5 mg in each of the PIX reactions.

4.9.2.1 The PIX Reactions Catalyzed by DMATS Mutants E89A and E89Q

Some fascinating observations were discovered in the isotope scrambling reaction catalyzed by the Glu89 mutants. The α-phosphorus nuclei of the unlabeled DMAPP and [1-^{18}O]-DMAPP (**67**) appear as two separate doublets at δ -9.28 and -9.30 ppm, respectively (Figure 4.35A). After 24 h incubation with the mutants at 37 ^{o}C, the signals that correspond to the unlabeled and ^{18}O-labeled pyrophosphate appear (Figure 4.35B&C). The percentage of DMAPP that had been consumed in the reactions was estimated using ^{31}P NMR integration as described in Section 4.4.4. Only about 8% of the DMAPP mixture was converted to the pyrophosphate in the E89A reaction, while 19% was converted to pyrophosphate in the E89Q reaction. The observation of pyrophosphate formation in the E89A reaction contradicts the site-directed mutagenesis studies reported by Metzger *et. al.*,[126] because the mutant E89A was previously claimed to be inactive. It is possible that the active site of the mutants is less protective than that of the wild type enzyme, and that water could intercept the dimethylallyl carbocation to give dimethylallyl alcohol. Alternatively, some pyrophosphate could also be formed by non-enzymatic hydrolysis during the incubation and lyophilization process. The amount of DMAT produced in each mutant reaction will be accessed by HPLC analysis as proposed in the 'future directions' section.

Isotope scrambling was observed in both of the E89A- and E89Q-catalyzed PIX reactions (Figure 4.35). After the reaction was carried out by the Glu89 mutants, the region of the α-phosphorus signals showed a new doublet at δ 9.31 ppm, indicating that [α-^{18}O]-DMAPP has formed. The amount of ^{18}O-scrambled product was calculated to be 21% and 18%, and the percentage of PIX (v_{PIX}) was estimated to be 32% and 27% in the E89A and E89Q reactions, respectively. Hence, the partitioning ratios (v_{pix}/v_{chem}) of the E89A and E89Q reactions were determined to be 4.0 and 1.5, showing a 10x and 3.8x increase over the v_{pix}/v_{chem}(WT), respectively (Table 4.6).

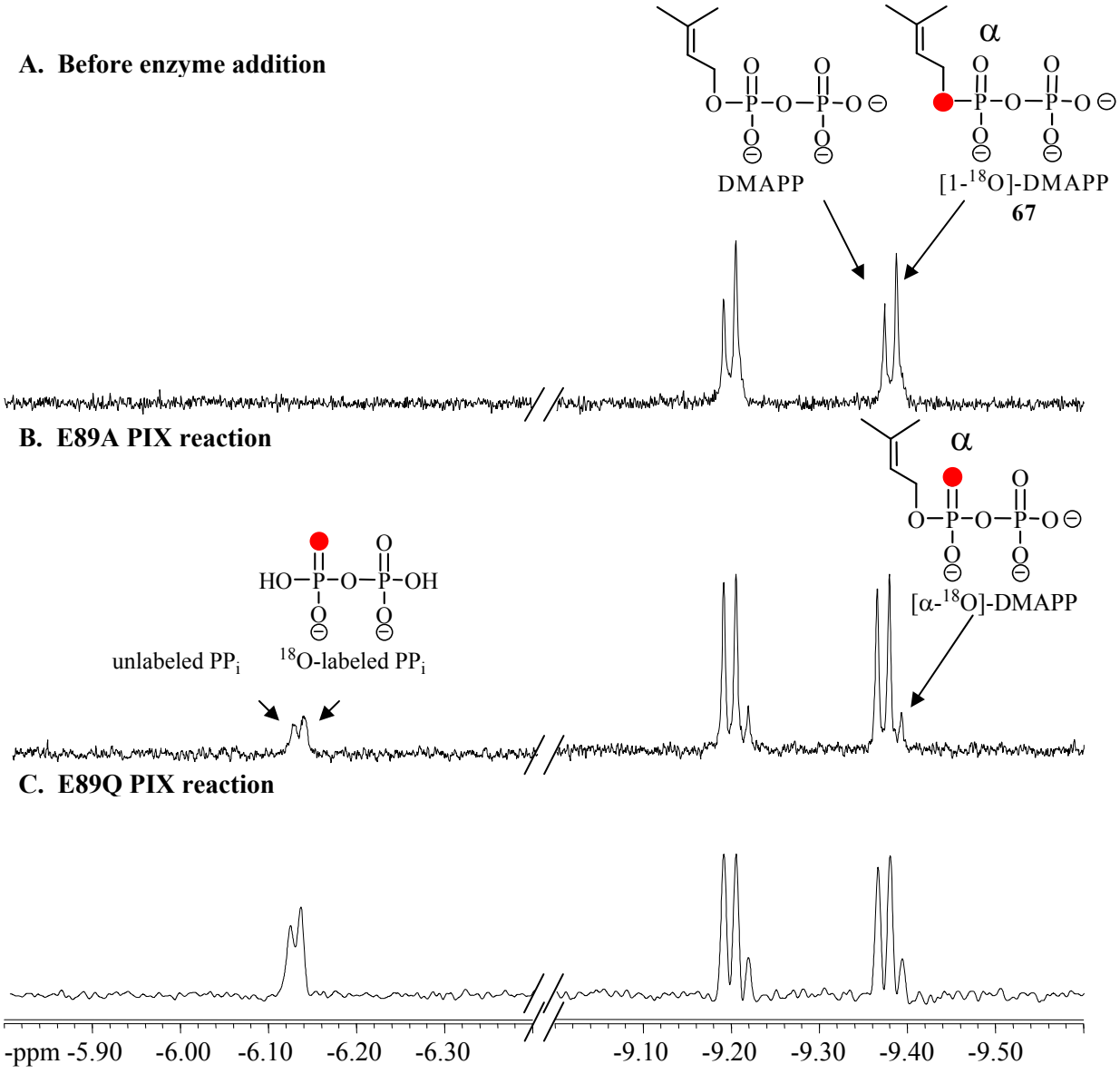


Figure 4.35 ^{31}P NMR (D_2O) spectra of the PIX reactions catalyzed by mutants of DMATS, E89A and E89Q. A) Before reaction, B) reaction catalyzed by E89A and C) by E89Q.

Surprisingly, another ^{18}O-scrambed product was observed in the PIX reactions catalyzed by the Glu89 mutants. Before enzyme addition, there was only a doublet at δ -5.51 ppm that corresponds to the β-phosphorus nuclei of both unlabeled and ^{18}O-labeled DMAPP (**67**). A new doublet at -5.53 ppm was observed after incubation with the mutants (Figure 4.36). Since the magnitude of the chemical shift is directly proportional to the bond order of ^{18}O-P, this new doublet indicates that [β-^{18}O]-DMAPP has formed in the PIX reaction. It is likely that the ^{18}O-labeled pyrophosphate has 'flipped' inside and/or outside of active site and then re-attached to the dimethylallyl carbocation. The extent of '^{18}O-flipping' was evaluated by dividing the integral of the [β-^{18}O]-DMAPP (flip) signal over the combined integrals of the [β-^{18}O]-DMAPP (flip), [1-^{18}O]-DMAPP (non-bridging) and [α-^{18}O]-DMAPP (bridging) signals. It turned out that the extent of '^{18}O-flipping' was notably larger in the E89A-catalyzed PIX reaction (14%) than the one in the E89Q-catalyzed reaction (8%). Also, it should be noted that the calculations of PIX extent (v_{pix} (E89A and E89Q)) did not include the observation of '^{18}O-flipping'. Subsequently, the partitioning ratio v_{pix}/v_{chem} (E89A and E89Q) are probably underestimated by ~10%.

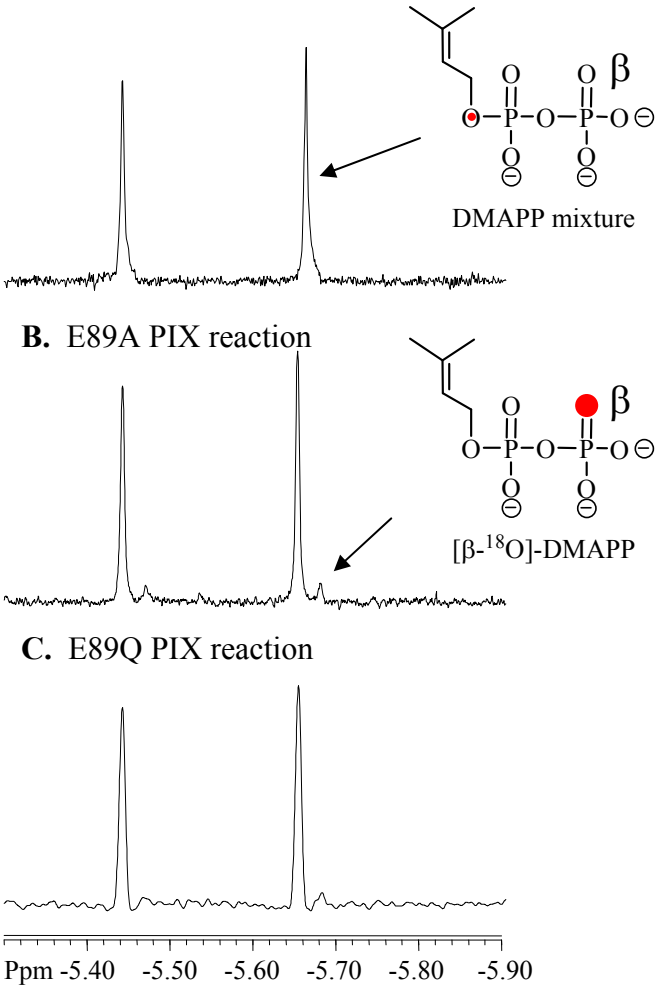


Figure 4.36 ^{31}P NMR (D_2O) spectra showing the β-phosphorus signals of the DMAPP mixture in the PIX reactions catalyzed by mutants of DMATS, E89A and E89Q. A) Before reaction, B) reaction catalyzed by E89A and C) by E89Q.

4.9.2.2 The PIX Reactions Catalyzed by DMATS Mutants K174A and K174Q

The PIX reactions catalyzed by the Lys174 mutants were also inspected. Prior to enzyme addition, the α-phosphorus signals of the unlabeled DMAPP and [1-^{18}O]-DMAPP (**67**) appear as two

separate doublets at δ -9.28 and -9.30 ppm (Figure 4.37). After extensive incubation with the mutants at 37 ^{o}C, the formation of unlabeled and ^{18}O-labeled pyrophosphate appeared at -6.12 and -6.13 ppm, respectively. It turned out that only 6% of the DMAPP had been consumed in the K174A reaction, whereas 27% of the DMAPP had been consumed in the K174Q reaction. As mentioned previously, DMAPP could either be used to form the product DMAT or be hydrolyzed in the mutant reactions, thus HPLC analysis will be used for further analysis.

The scrambled starting material, [α-^{18}O]-DMAPP, was formed in both of the K174A- and K174Q-catalyzed PIX reactions (Figure 4.37). The α-phosphorus signal of [α-^{18}O]-DMAPP emerged as a doublet at δ -9.31 ppm in the ^{31}P spectrum that was collected after the enzyme incubation. The amount of ^{18}O-scrambled product was determined to be 17% and 24%, and the percentage of PIX reaction (v_{pix}) was calculated to be 26% and 37% in the K174A and K174Q reactions, respectively. Based on these values, the partitioning ratios (v_{pix}/v_{chem}) of the K174A and K174Q reaction are determined to be 4.4 and 1.4, showing a 11x and 3.5x increase over the v_{pix}/v_{chem}(WT), respectively (Table 4.6). Lastly, the ^{31}P NMR spectra of the Lys174 mutant reactions did not show any sign of '^{18}O-flipping', as was observed in the Glu89 mutant reactions.

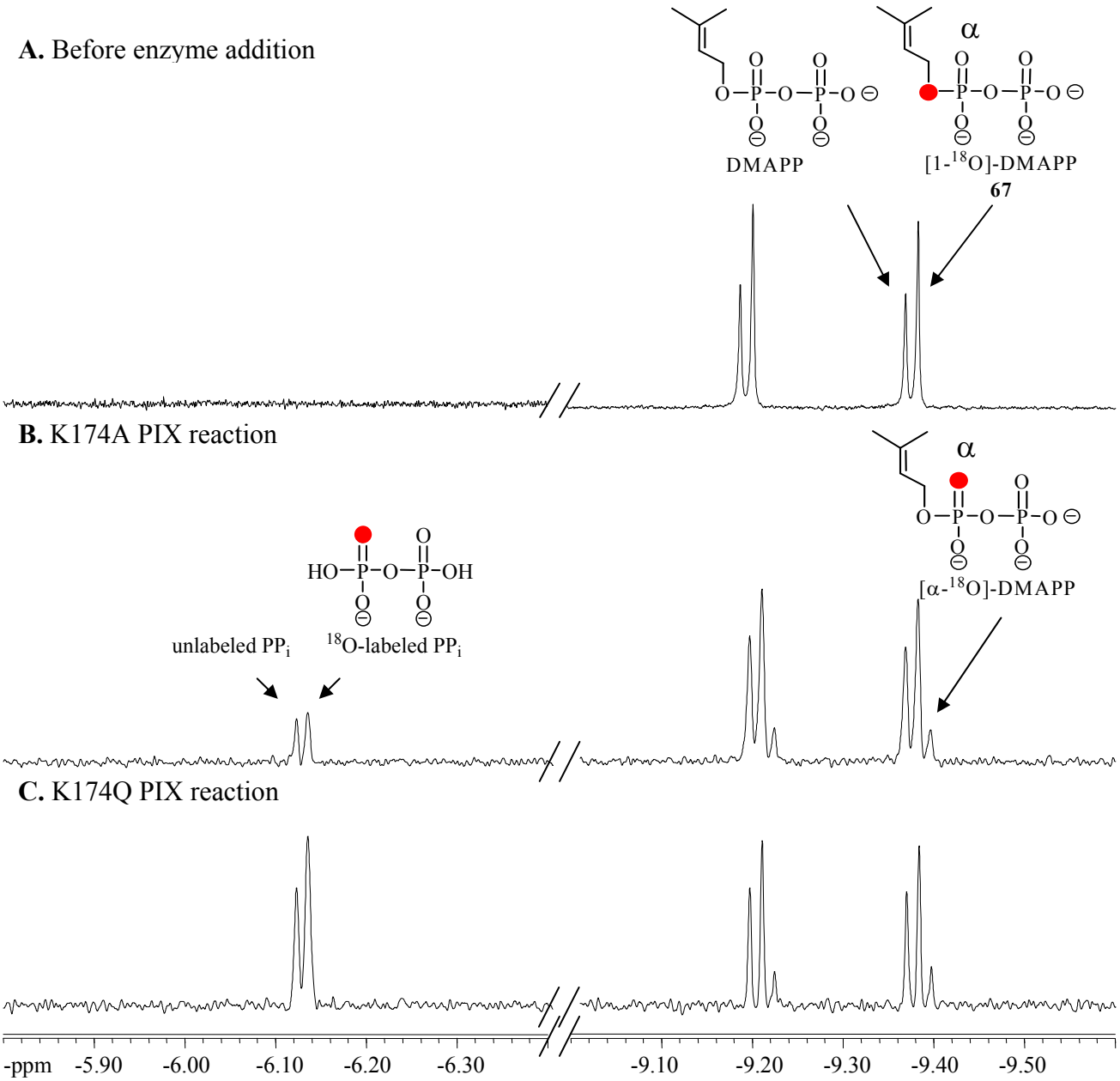


Figure 4.37 ^{31}P NMR (D_2O) spectra of the PIX reactions catalyzed by mutants of DMATS, K174A and K174Q. A) Before reaction, B) reaction catalyzed by K174A and C) by K174Q.

Table 4.6 The measured partitioning ratios for the WT and mutant-catalyzed PIX reaction

DMATS	Partitioning Ratio (v_{pix}/v_{chem})	Relative increase in v_{pix}/v_{chem}
Wild type	0.40	1x
E89A	4.0	10x
E89Q	1.5	3.8x
K174A	4.4	11x
K174Q	1.4	3.5x

4.9.3 Discussion of the Mutant-catalyzed PIX Reactions and Future Directions.

The residue Glu89 in the active site of DMATS is likely responsible for promoting formation of the arenium intermediate by hydrogen-bonding to the indole N-H, as previously suggested in the crystallographic studies.[126] This is supported by the observation of a 10-fold increase in the partitioning ratio (v_{pix}/v_{chem}) of the E89A-catalyzed PIX reaction, which implies that the barrier to the formation of the arenium/pyrophosphate ion pair increased substantially. The nucleophilicity of the indole C-4 is greatly depressed in the active site of the E89A mutant, due to the lack of polar interaction with the indole N-H. Hence, the dimethylallyl carbocation/pyrophosphate ion pair would tend to collapse to the starting material rather than proceed to the arenium intermediate. In contrast, the v_{pix}/v_{chem}(E89Q) only increased 3.8-fold, suggesting that the barrier to the arenium/pyrophosphate ion pair formation increased to a smaller extent. Given that glutamine is a weak hydrogen-bond acceptor, it is reasonable to expect the E89Q mutant could slowly catalyze the C-C bond formation. Furthermore, the increase in the barrier to the arenium/pyrophosphate ion pair formation is further supported by the formation of [β-^{18}O]-DMAPP. The observation suggested that the pyrophosphate could tumble inside the mutant's active site, thereby leading to the re-attachment of the unlabeled phosphate moiety to the dimethylallyl carbocation. The E89A mutant could catalyze a greater extent of '^{18}O-flipping', probably because the change in the reaction coordinate is more dramatic as speculated above. Because the C-C bond formation should become fully rate-determining in the E89Q reaction (Figure 4.33), the extent of KIE will be measured using [1,1-$^{2}H_2$]-DMAPP (**71**) and [4-^{2}H]-L-tryptophan (**72**). This should further characterize the catalytic role of Glu89 in the reaction.

The basic residue, Lys174, likely deprotonates the arenium ion intermediate in the DMATS reaction. An 11-fold increase in the v_{pix}/v_{chem} of the K174A-catalyzed PIX reaction can be

rationalized as a consequence of a considerable increase in the energy barrier to DMAT re-aromatization. Because the mutant carries a non-basic alanine residue, the arenium/pyrophosphate ion pair would tend to reverse to the formation of L-tryptophan/DMAPP, instead of forming products. On the other hand, significant activity and only a 3.5x increase in the v_{pix}/v_{chem}(K174Q) was observed with the K174Q mutant. This implies that the pK_a of the arenium ion is close to the pK_a of a carbonyl amide functionality (~0.5) and the glutamine can serve as a base.[170] Previously, in the crystallographic studies on DMATS, it was proposed that the carboxylate group of tryptophan might replace the role of Lys174, thus the K174Q mutant could retain a significant amount of activity.[126] Our results can rule out this hypothesis, because the activity of K174A was much lower and the v_{pix}/v_{chem} of the K174A and the K174Q reactions were quite different. Additional studies are needed to analyze the functional role of Lys174 in catalysis. As mentioned above, [1,1-$^2H_2$]-DMAPP (**71**) and [4-^{2}H]-L-tryptophan (**72**) will be used to measure the extent of KIE in the K174Q reaction, because the deprotonation step should become either partially or fully rate-limiting in catalysis (Figure 4.33). Furthermore, the K174A mutant will be subjected to a 'chemical rescue' experiment (Figure 4.38). The activity of K174A (k_{cat}) should be greatly enhanced when ethylamine is added in the reaction buffer. Specifically, it was proposed that ethylamine can enter the active site and 'rescue' the catalytic activity by deprotonating the arenium ion (Figure 4.38). Lastly, the kinetic parameters of all of the mutants will be measured using the enzyme-coupled assay as described in Section 4.3.

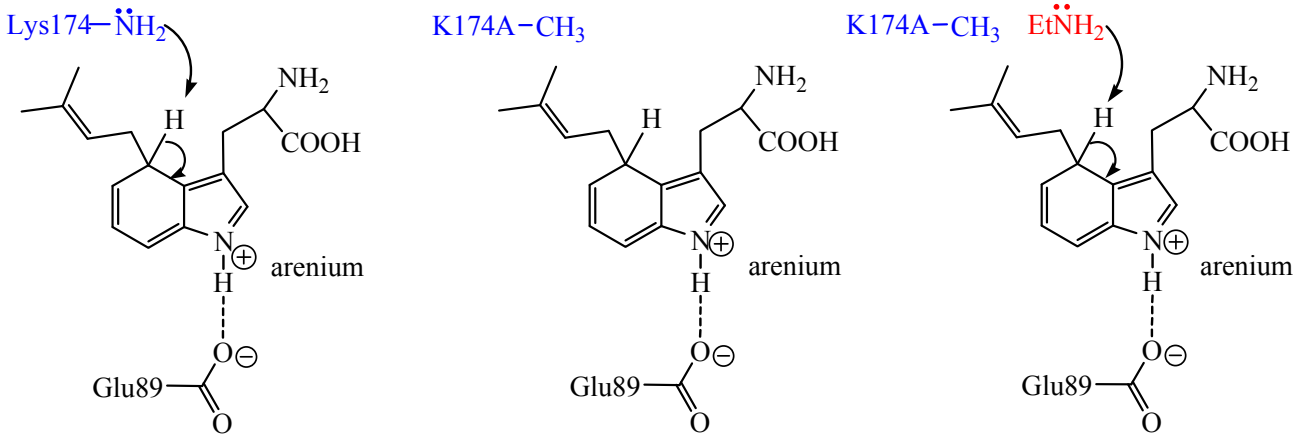


Figure 4.38 The proposed 'chemical rescue' experiment for K174A. $EtNH_2$ refers to ethylamine.

In each of the mutant-catalyzed PIX reactions, formation of pyrophosphate was observed indicating that DMAPP has been consumed. However, the active sites of the mutants might be less protective from solvent than that of the wild type enzyme. DMAPP consumption in the mutant reactions could result either in the production of DMAT or in hydrolysis and/or elimination to give the corresponding alcohol and/or diene. To distinguish these two possibilities, it is considered relevant to analyze the product of the PIX reaction by reverse phase HPLC. Comparison between the extent of DMAT production measured in HPLC analysis and the extent of DMAPP consumption measured by ^{31}P NMR spectroscopy will provide a clue as to whether dimethylallyl alcohol was formed in the reactions.

In conclusion, the proposed catalytic roles of E89 and K174 are supported by the results obtained from the mutant-catalyzed PIX reactions.[126] When these residues were mutated to a non-ionic residue (alanine or glutamine), the catalytic activity decreased but the lifetimes of the carbocation intermediate/pyrophosphate pair(s) were extended. However, additional experiments are still needed to fully characterize the functions of these residues.

4.10 Summary

In summary, DMATS employs a unique mechanism to catalyze the indole prenylation reaction. The steps in forming the dimethylallyl carbocation and the arenium ion intermediates are both partially rate-determining as indicated in the PIX and KIE studies. This may be compared to the reaction catalyzed by *trans*-farnesyl diphosphate synthase which employs a similar catalytic mechanism and generates a reactive allyl carbocation intermediate.[97] However, the rate-limiting process in this reaction is the release of the product GPP (or FPP), because the Mg^{2+} cofactor is tightly coordinated to the diphosphate moiety.[164] In the protein farnesyltransferase reaction, change of protein conformation and/or product release is a rate-limiting process as well.[113, 118] However, this prenylation reaction is promoted by the formation of a Zn^{2+}-thiolate complex, and the enzyme employs an associative transition state for catalysis.[113] Although all of these enzymes are capable of catalyzing a prenylation reaction, the catalytic mechanisms appear to be significantly different from one another probably as a consequence of convergent evolution.

With respect to the non-enzymatic reaction, the closest example found in the literature is the addition of a benzhydrylium ion onto the C-3 position of an unsubstituted indole ring. It was shown that the non-enzymatic reaction follows first-order kinetics and is dictated by the rate of benzhydrylium ion formation.[140, 171] With the DMATS reaction, formation of both the dimethylallyl carbocation and the arenium ion are partially rate-determining. This difference suggests that DMATS promotes the reaction by decreasing the energy barrier to dimethylallyl carbocation formation. Furthermore, the crystallographic studies on DMATS have indicated that there are catalytic residues involved in each of the chemical steps in the reaction as well as residues that stabilize the intermediates via π-carbocation interactions.[126] Hence, this leads to the conclusion that DMATS adapts the same mechanism as the non-enzymatic reaction, but the energy barriers in each

of the chemical steps are lowered in catalysis. In particular, the rate of formation and stability of the carbocation intermediates are greatly enhanced in the DMATS reaction.

4.11 Experimental Procedures

4.11.1 Materials

All reagents were purchased from Sigma-Aldrich and used without further purification unless otherwise noted. D_2O (99.9%) and $H_2{}^{18}O$ (97%) were purchased from Cambridge Isotope Laboratories, Inc.

Thin layer chromatography was performed on aluminum-backed sheets of silica gel $60F_{254}$ (Merck) of thickness 0.2 mm. Compounds were visualized by UV or by spraying with a solution containing H_2SO_4 (31 mL), ammonium molybdate (21 g), and $Ce(SO_4)$ (1 g) in water (500 mL) and then heating at 110 ^oC for 1 min. Silica gel (230 – 400 mesh, BDH) was used for column chromatography.

^1H NMR spectra were obtained on the Bruker AV300 and AV400 spectrometer at field strengths of 300 and 400 MHz, respectively. Proton-decoupled ^{31}P NMR spectra were recorded on the Bruker AV300 spectrometers at 121.5 MHz. ^{19}F NMR spectra were obtained on the Bruker AV300 spectrometer at 282.3 MHz. HPLC analysis was carried out on a Waters Delta 600 equipped with a model Waters 2996 photodiode array detector, reverse phase Delta Pak 15µ C18 100Å Waters HPLC column at room temperature (~24 ^oC). Rosetta (DE3) pLysS *E. coli* cells was purchased from Novagen.

4.11.2 General Methods

4.11.2.1 General Enzyme Methods

Centrifugal filters (4 mL 10 000 MWCO) were purchased from Millipore. Acryl-cuvettes for use in enzyme kinetic assays were from Sarstedt. Chelating Sepharose® Fast resin and Sephadex™ G-25 resin were purchased from Pharmacia Biotech. Protein concentrations were determined by the method of Bradford on a Cary 3E UV-Vis spectrophotometer using bovine serum albumin as standard.[89] All measurements were performed at room temperature. Protein purity was assessed using SDS-PAGE, stained with Coomassie blue according to Laemmli.[90] Molecular weight standards for SDS-PAGE were BSA (66 kDa) and carbonic anhydrase (29 kDa), both purchased from Sigma.

4.11.2.2 Overexpression and Purification of Dimethylallyltryptophan Synthase (FgaPT2)

The *fgaPT2* gene of *Aspergillus fumigatus* (strain AF 293) (GenBank accession no. AY775787.1) flanked with a BamH1 and a HindIII restriction sites was synthesized and cloned into a pET28a vector (Novagen) by GenScript. Overexpression of dimethylallyltryptophan synthase (DMATS) was modified from the protocol as described by Steffan *et. al.*[172] Rosetta(DE3) pLysS (Novagen) harboring the recombinant *fgaPT2*/pET28a construct were grown at 37 ^{o}C in 1 L of Terrific Broth (TB) medium containing 35 μg/mL chloramphenicol and 30 μg/mL kanamycin until an OD_{600} of 0.6 was reached. Cells were induced for overexpression by the addition of 119 mg (0.5 mM) of isopropyl-1-thio-*β*-D-galactopyranoside (IPTG). After growing at 24 ^{o}C for an additional 24 h, cells were harvested by and DMATS was purified following a previously reported protocol.[172] Typically ~35 mg of enzyme was purified from 1 L culture, and it could be stored at -80 ^{o}C for at least 6 months without significant loss in activity.

4.11.2.3 Assay of DMATS Activity and HPLC Purification of Dimethylallyltryptophan

To a solution of Tris-HCl buffer (50 mM, 5 mM $CaCl_2$ pH 7.5, 1.00 mL) containing DMAPP and L-tryptophan (10 mM each) was added 1 mg of DMATS. After incubating the reaction mixture for 4 h at 35 ^{o}C, enzyme was removed by centrifugal ultrafiltration and the filtrate was subjected to reverse phase HPLC purification using a linear gradient of 20-70% acetonitrile in 0.1% aqueous trifluoroacetic acid as previously described by Unsöld *et. al.*[131] The peak corresponding to dimethylallyltryptophan (retention time ~20 min) was collected and dried. The ^{1}H NMR and ESI-MS spectral data were the same as previously reported.[131]

4.11.3 Unlabeled Dimethylallyl Diphosphate, Isotopically Labeled [1-^{18}O]- and [1,1-$^{2}H_2$]-Dimethylallyl Diphosphate (67 and 71) and *E*-3-(Fluoromethyl)-2-butenyl Diphosphate (*E*-52)

DMAPP, its isotopologues **67** and **71**, and fluorinated analog ***E*-52** were synthesized by coupling the corresponding alcohols with *bis*-triethylammonium phosphate in the presence of trichloroacetonitrile as previously described by Poulter and his co-worker.[173, 174] Synthesis of the isotopically labeled alcohols will be elaborated below.

4.11.3.1 [1,1-^{2}H]-3-Methyl-2-buten-1-ol (74)

[1,1-^{2}H]-3-methyl-2-buten-1-ol was prepared as previously described.[165] The extent of ^{2}H incorporation was determined to be ≥97% by ^{1}H NMR spectroscopy and ESI-MS.

4.11.3.2 [1-^{18}O]-3-Methyl-2-buten-1-ol (**70**)

The synthesis of [1-^{18}O]-3-methyl-2-buten-1-ol (**70**) was modified from the procedure described by Vani *et. al.*[157] 1-Chloro-3-methyl-2-butene (4.27 g, 40.8 mmol) was heated at 100 ^{o}C for 4 h in the presence of triethylbenzylammonium chloride (TEBA) (0.104g, 0.45 mmol) and sodium [1-$^{18}O_2$]-acetate (3.87 g, 47.1 mmol), which was prepared as previously described.[158] The resultant acetylated product was then extracted with dichloromethane (3 x 10 mL) and concentrated to dryness under vacuum. The residue was dissolved in 30 mL of anhydrous methanol solution containing sodium methoxide (1.3 M) and refluxed for 4 h. The reaction was quenched by addition 20 mL saturated NH_4Cl solution and extracted three times with 20 mL diethyl ether. The organic layer was separated, dried with $MgSO_4$ and evaporated to dryness under reduced pressure. The product [1-^{18}O]-3-methyl-2-buten-1-ol (**70**) was purified by Kugelrohr distillation. The extent of ^{18}O incorporation was determined to be ~92% by ESI-MS.

4.11.3.3 *E*-3-(Fluoromethyl)-2-buten-1-ol (*E*-57)

E-3-(Fluoromethyl)-2-buten-1-ol (***E*-57**) was prepared as previously described by Gebler *et. al.*[15]

4.11.4 Continuous Two-Enzymes Coupled Enzymatic Assay

Enzyme kinetics were measured by a modification of a previously reported continuous coupled phosphate assay.[175] A cuvette containing 50 mM Tris-HCl buffer (pH 7.5, final volume 1000 μL), L-tryptophan (variable), dimethylallyl diphosphate (variable), 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) (200 μM), purine nucleoside phosphorylase (PNPase) (1 units exchanged twice into 50 mM Tris-HCl buffer pH 7.5), inorganic pyrophosphatase (PPase) (0.5 units), $MgCl_2$ (5 mM), $CaCl_2$ (5 mM) was thermally equilibrated for 5 min at 35 ^{o}C. The enzymatic

reaction was initiated by the addition of DMAT synthase (1 μg) and the rate was calculated from the observed increase of absorption at 360 nm (using ε = 11,000 $M^{-1} \cdot cm^{-1}$). The K_M value for L-tryptophan was measured in the presence of 200 μM DMAPP (saturating), and that for DMAPP was measured in the presence of 200 μM L-tryptophan (saturating). Kinetic parameters were determined from the initial velocities fit to Michaelis-Menten kinetics using the program *Grafit*.

To determine the kinetic parameters for *E*-3-(Fluoromethyl)-2-butenyl Diphosphate (***E*-52**), the same conditions were used but with 10 μg of DMATS. The K_M value for ***E*-52** was measured in the presence of 200 μM L-tryptophan (saturating).

4.11.5 Kinetic Isotope Effect Studies

An intermolecular competitive method was employed to measure a KIE on the second-order rate constant k_{cat}/K_M, symbolized as k_H/k_D. A mixture of two isotopologues of known isotopic composition was treated with the enzyme. After a certain fractional conversion to products has occurred, the isotopic composition of the starting material was recovered and determined. The KIE on k_{cat}/K_M was determined from the initial and final spectra using the equation KIE = $\ln(1 - F_H)/\ln[(1 - F_H)R/R_0]$,[74] where F_H is the fractional conversion of the protiated species to products and R_0 and R are the initial and final ratios of protiated to deuterated substrate, respectively.

4.11.5.1 Kinetic Isotope Effect Measurement for [1,1-$^2H_2$]-Dimethylallyl Diphosphate (71)

To determine the k_H/k_D of dimethylallyl diphosphate, a solution of an approximately 1:1 molar ratio of unlabeled and [1,1-$^2H_2$]-DMAPP **71** (12 mM each), 26 mM L-tryptophan and 5 mM $CaCl_2$ in Tris-HCl (pH 7.5, 50 mM, final volume 12 mL) was prepared and DMATS (1.2 mg) was added. Two aliquots (1 mL each) were collected at 0, 90, 120 and 240 min time intervals for ESI-MS and NMR analyses. The ESI-MS samples were quenched by adding 1 mL of dry-ice-chilled

methanol and the enzyme was removed by ultrafiltration (Amicon Ultra-4, 10 000 MWCO) prior to negative ESI-MS data collection. The NMR samples were immediately diluted with 10 mL of distilled water, flash frozen with liquid N_2 and lyophilized. ^{31}P NMR spectra (D_2O containing 10 mM 2-deoxy-glucose-6-phosphate) were collected with an AV300 NMR spectrometer at 4.5 s delay time. F_H was calculated by integration the α phosphate signal at -10.0 ppm to the internal standard 2-deoxy-glucose-6-phosphate (10 mM) at 4.0 ppm in the ^{31}P NMR spectra before and after enzyme additions. R_o and R were calculated from the mass spectra before and after addition of the enzyme using the relative peak intensities of the protiated substrate [DMAPP *m/z* 245 ($M - H^-$)] to the deuterated substrates [[1,1-^{2}H]-DMAPP *m/z* 247 ($M - H^-$)]. The measurements were performed in triplet and an average value was reported. The values of individual trials gave the following results: k_H/k_D = 1.13, 1.16, and 1.18.

4.11.5.2 Kinetic Isotope Effect Measurement for [4-^{2}H]-L-Tryptophan (72)

[4-^{2}H]-L-tryptophan (**72**) was synthesized by photosubstitution as previously described.[176] The extent of ^{2}H incorporation at C-4 was determined to be 82% by ^{1}H NMR spectroscopy and ESI-MS. Minor deuteration was also observed at the C-2 and C-7 positions during the reaction, and caused 16% of the labeled product to be di-deuterated.

To determine the k_H/k_D of L-tryptophan, a solution of an approximately 1:1 molar ratio of [4-^{1}H] and [4-^{2}H]-L-tryptophan **72** (12 mM each), 26 mM dimethylallyl diphosphate and 5 mM $CaCl_2$ in Tris-HCl (pH 7.50, 50 mM, final volume 12 mL) was prepared and 1.2 mg of DMAT synthase was added. 1 mL aliquots were collected at 0, 90, 120 and 240 min time intervals. Each sample was quenched by the addition of 1 mL of dry-ice-chilled methanol containing 10 mM L-1-methyl-tryptophan and the protein was removed by ultrafiltration (Amicon Ultra-4, 10 000 MWCO). The samples were subjected to HPLC analysis under the previously described conditions,[131] and the L-

tryptophan peaks were collected for positive ESI-MS analysis. F_H values were calculated from the HPLC data before and after addition of the enzyme by integration of the L-tryptophan peak relative to the internal standard peak of L-1-methyl-tryptophan. R_o and R were calculated from the intensity of the mass spectral signals before and after addition of the enzyme (L-tryptophan *m/z* 205 ($M + H^+$) and [4-^{2}H]-L-tryptophan *m/z* 206 ($M + H^+$)). The intensities of the signals were corrected for the natural abundance of ^{13}C and for the presence of the di-deuterated tryptophan (m/z 207 $M + H^+$) before R_o and R were calculated. The measurements were performed in triplicate and an average value was reported. The values of individual trials were gave the following results $k_H/k_D = 0.78$, 0.82, and 0.83.

4.11.6 PIX Experiment

4.11.6.1 PIX Experiment with Natural Substrate L-Tryptophan

A 1 mL solution containing [1-^{18}O]-DMAPP (**67**) and unlabeled DMAPP (30 mM total concentration, 63% ^{18}O incorporation) and L-tryptophan (24 mM, 0.8 equivalent) in Tris-HCl prepared using D_2O (50 mM, pD 7.9, containing 3 mM EDTA) was placed in a 5 mm NMR tube and a ^{31}P NMR spectrum was collected. A solution of DMATS (1 mg in 1 mL Tris-HCl 50 mM pH 8) was added and incubated for 4 h at 35 ^oC. The enzyme was removed by ultrafiltration (Amicon Ultra-4, 10 000 MWCO) and to the filtrate was added Chelex-100 resin (20 mg of 200-400 mesh, Na^+ form, previously rinsed with D_2O). After extensive mixing, a second ^{31}P NMR spectrum was obtained. The proton-decoupled ^{31}P NMR spectra were obtained on a Bruker AV300 multinuclear spectrometer operating at a frequency 121.5 MHz. All spectra were externally referenced to H_3PO_4 (0 ppm). Typical acquisition parameters were 2437 Hz (20 ppm) sweep width centered at -5 ppm,

27 s acquisition time, 2 s delay time and 13.5 μs pulse width. Well resolved spectra were achieved after 200 to 1000 scans. The FIDs were apodized using an Exponential/Gaussian resolution enhancement function.

4.11.6.2 PIX Experiment with Fluorinated L-Tryptophan Analogs (75 and 76)

PIX experiments using 4-fluoro- and 6-fluoro-D,L-tryptophan (**75** and **76**) were performed as described above, but with an excess of tryptophan analogs (70 mM, 2.3 equivalents relative to DMAPP) and 2 mg DMATS. Proton-decoupled ^{31}P NMR spectroscopy was conducted as described above except that the delay time between each scan was increased to 3 s.

4.11.6.3 PIX Experiment with DMATS Mutants

PIX experiments with DMATS mutants, E89A, E89Q, K174A and K174Q, were performed as described above but with 5 mg of mutant enzyme used in each trial.

4.11.7 Site-directed Mutagenesis

Plasmids coding for the mutant enzymes were constructed using the QuikChange™ site-directed mutagenesis kit (Stratagene). The *fgaPT2*/pET28a plasmid bearing the wild-type DMATS gene was employed as template. DNA synthesis primers were designed as follows (altered bases are underlined): **E89A**, 5′-CGG CGC TCC GTT CGC ATT GAG CCT AAA TTG-3′ and 5′-CAA TTT AGG CTC AAT GCG AAC GGA GTG CCG -3′; **E89Q**, 5′-ATA CGG CAC TCC GTT CCA ATT GAG CCT AAA TTG C -3′ and 5’-GCA ATT TAG GCT CAA TTG GAA CGG AGT GCG GTA T-3’; **K174A**, 5′-TAT CAG GAC GCA GAA CGC GCT CGC GCT CGA TCT G-3′ and 5′-CAG ATC GAG CGC GAG CGC GTT CTG CGT CCT GAT A-3′; **K174Q**, 5′-CTA TCA GGA CGC AGA ACC AGC TCG CGC TCG ATC TG-3′ and 5′-CAG ATC GAG CGC GAG CTG GTT CTG

CGT CCT GAT AG-3′. The site-directed mutagenesis reaction mixtures contained 5 μL of 10× polymerase buffer, 5 μL of 2.5 mM dNTP mix, 125 ng of each primer, 25 or 50 ng of *fgaPT2*/pET28a, and 2.5 units of Pfu polymerase in a total volume of 50 μL. The reaction mixture was topped with 30 μL of mineral oil and passed through the following cycles: one cycle of 2.5 min at 95 ^{o}C followed by 16 cycles of 1 min at 95 ^{o}C, 1 min at 55 ^{o}C, and 12 min at 68 ^{o}C. After synthesis, the mixture was incubated at 37 °C for 1.5 h with 10 units of *Dpn*I restriction enzyme. The PCR product was semi-purified by DNA precipitation and then transformed (4 μL) into 100 μL XL1-Blue (Stratagene) supercompetent cells and plated on an LB agar plate containing 30 μg/mL kanamycin. Individual colonies were used to inoculate a 5-mL overnight culture of LB containing 30 μg/mL kanamycin. Plasmid DNA was isolated using the Wizard™ Miniprep DNA purification system (Promega). Four plasmids containing the correct mutation were obtained and in each case the entire gene was sequenced to ensure no other errors were present.

4.11.7.1 Overexpression, Purification and Activity Assay of the DMATS Mutants (E89A, E89Q, K174A and K174Q)

Each of the DMATS mutants (E89A, E89Q, K174A and K174Q) was overexpressed and purified as described for the wild type enzyme.[172] Approximately 10 mg of enzyme was purified from 1 L culture, and it could be stored at -80 ^{o}C for at least 1 month without significant loss in activity.

References

1. Facchini, P. J. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2001**, *52*, 29-66.

2. Ziegler, J.; Facchini, P. J. *Annu. Rev. Plant Biol.* **2008**, *59*, 735-769.

3. Kutchan, T. M. *Plant Cell* **1995**, *7*, 1059-1070.

4. Aripov, K. N. *Chem. Nat. Compd.* **1977**, *13*, 743-753.

5. Blakemore, P. R.; White, J. D. *Chem. Commun.* **2002**, 1159-1168.

6. Novak, B. H.; Hudlicky, T.; Reed, J. W.; Mulzer, J.; Trauner, D. *Curr. Org. Chem.* **2000**, *4*, 343-363.

7. Parker, K. A.; Fokas, D. *J. Org. Chem.* **2006**, *71*, 449-455.

8. Guenard, D.; Gueritte-Voegelein, F.; Potier, P. *Acc. Chem. Res.* **1993**, *26*, 160-167.

9. Holton, R. A.; Somoza, C.; Kim, H. B.; Liang, F.; Biediger, R. J.; Boatman, P. D.; Shindo, M.; Smith, C. C.; Kim, S. *J. Am. Chem. Soc.* **1994**, *116*, 1597-1598.

10. Holton, R. A.; Kim, H. B.; Somoza, C.; Liang, F.; Biediger, R. J.; Boatman, P. D.; Shindo, M.; Smith, C. C.; Kim, S. *J. Am. Chem. Soc.* **1994**, *116*, 1599-1600.

11. Kutchan, T. M.; Hampp, N.; Lottspeich, F.; Beyreuther, K.; Zenk, M. H. *FEBS Lett.* **1988**, *237*, 40-44.

12. Boven, E.; Venema-Gaberscek, E.; Erkelens, C. A. M.; Bissery, M. C.; Pinedo, H. M. *Ann. Oncol.* **1993**, *4*, 321-324.

13. Poppe, L.; Rétey, J. *Angew. Chem. Int. Ed.* **2005**, *44*, 3668-3688.

14. Galpin, J. D.; Ellis, B. E.; Tanner, M. E. *J. Am. Chem. Soc.* **1999**, *121*, 10840-10841.

15. Gebler, J. C.; Woodside, A. B.; Poulter, C. D. *J. Am. Chem. Soc.* **1992**, *114*, 7354-7360.

16. Shibuya, M.; Chou, H. M.; Fountoulakis, M.; Hassam, S.; Kim, S. U.; Kobayashi, K.; Otsuka, H.; Rogalska, E.; Cassady, J. M.; Floss, H. G. *J. Am. Chem. Soc.* **1990**, *112*, 297-304.

17. Zenk, M. H. *Pure Appl. Chem.* **1994**, *66*, 2023-2028.

18. Stadler, R.; Kutchan, T. M.; Loeffler, S.; Nagakura, N.; Cassels, B.; Zenk, M. H. *Tetrahedron Lett.* **1987**, *28*, 1251-1254.

19. Pictet, A.; Spengler, T. *Ber. Dtsch. Chem. Ges.* **1911**, *44*, 2030-2036.

20. Cox, E. D.; Cook, J. M. *Chem. Rev.* **1995**, *95*, 1797-1842.

21. Li, J.; Wang, T.; Yu, P.; Peterson, A.; Weber, R.; Soerens, D.; Grubisha, D.; Bennett, D.; Cook, J. M. *J. Am. Chem. Soc.* **1999**, *121*, 6998-7010.

22. Dorn, G. W.; Oswald, K. J.; McCluskey, T. S.; Kuhel, D. G.; Liggett, S. B. *Biochemistry* **1997**, *36*, 6415-6423.

23. Liao, X.; Zhou, H.; Yu, J.; Cook, J. M. *J. Org. Chem.* **2006**, *71*, 8884-8890.

24. Kosuge, T.; Yokota, M.; Nagasawa, M. *Yakugaku Zasshi* **1978**, *98*, 1370-1375.

25. Woodward, R. B.; Cava, M. P.; Ollis, W. D.; Hunger, A.; Daeniker, H. U.; Schenker, K. *J. Am. Chem. Soc.* **1954**, *76*, 4749-4751.

26. Jackson, A. H.; Smith, A. E. *Tetrahedron* **1968**, *24*, 403-413.

27. Woodward, R. B.; Cava, M. P.; Ollis, W. D.; Hunger, A.; Daeniker, H. U.; Schenker, K. *Tetrahedron* **1963**, *19*, 247-288.

28. Williams, J. R.; Unger, L. R. *Chem. Comm.* **1970**, 1605-1606.

29. Bailey, P. D. *J. Chem. Res* **1987**, 202.

30. Bailey, P. D.; Hollinshead, S. P.; McLay, N. R.; Morgan, K. M.; Palmer, S. J.; Prince, S. N.; Reynolds, C. D.; Wood, S. D. *J. Chem. Soc. Perkin Trans. 1* **1993**, 431-439.

31. van Maarseveena, J. H.; Muldersa, S. J. E.; Abena, R. W. M.; Kruseb, C. G.; Scheeren, H. W. *Tetrahedron* **1995**, *51*, 4841-4852.

32. Kowalski, P.; Mokrosz, J. L. *Bull. Soc. Chim. Belg.* **1997**, *106*, 147-149.

33. Harcourt, D. N.; Taylor, N.; Waigh, R. D. *J. Chem. Soc., Perkin Trans. 1* **1978**, 722-725.

34. Harcourt, D. N.; Taylor, N.; Waigh, R. D. *J. Chem. Soc., Perkin Trans. 1* **1978**, 1330-1333.

35. Waigh, R. D. *J. Chem. Soc. Chem. Commun.* **1980**, 1164-1165.

36. Gözler, B.; Shamma, M. *J. Nat. Prod.* **1990**, *53*, 740-743.

37. Reddy, M. S.; Cook, J. M. *Tetrahedron Lett.* **1994**, *35*, 5413-5416.

38. Woodward, R. B.; Bader, F. E.; Bickel, H.; Frey, A. J.; Kierstead, R. W. *J. Am. Chem. Soc.* **1956**, *78*, 2023-2025.

39. Taylor, M. S.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2004**, *126*, 10558-10559.

40. Raheem, I. T.; Thiara, P. S.; Peterson, E. A.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2007**, *129*, 13404-13405.

41. Mergott, D. J.; Zuend, S. J.; Jacobsen, E. N. *Org. Lett.* **2008**, *10*, 745-748.

42. Klausen, R. S.; Jacobsen, E. N. *Org. Lett.* **2009**, *11*, 887-890.

43. Seayad, J.; Seayad, A. M.; List, B. *J. Am. Chem. Soc.* **2006**, *128*, 1086-1087.

44. Sewgobind, N. V.; Wanner, M. J.; Ingemann, S.; de Gelder, R.; van Maarseveen, J. H.; Hiemstra, H. *J. Org. Chem.* **2008**, *73*, 6405-6408.

45. Stadlera, R.; Kutchana, T. M.; Zenk, M. H. *Phytochemistry* **1989**, *28*, 1083-1086.

46. De-Eknamkul, W.; Suttipanta, N.; Kutchan, T. M. *Phytochemistry* **2000**, *55*, 177-181.

47. Samanani, N.; Liscombe, D. K.; Facchini, P. J. *Plant J.* **2004**, *40*, 302-313.

48. Ma, X.; Panjikar, S.; Koepke, J.; Loris, E.; Stőckigt, J. *Plant Cell* **2006**, *18*, 907-920.

49. Maresh, J. J.; Giddings, L.-A.; Friedrich, A.; Loris, E. A.; Panjikar, S.; Trout, B. L.; Stockigt, J.; Peters, B.; O'Connor, S. E. *J. Am. Chem. Soc.* **2008**, *130*, 710-723.

50. Rueffer, M.; El-Shagi, H.; Nagakura, N.; Zenk, M. H. *FEBS Lett.* **1981**, *129*, 5-9.

51. Samanani, N.; Facchini, P. J. *J. Biol. Chem.* **2002**, *277*, 33878-33883.

52. van der Heijden, R.; Jacobs, D. I.; Snoeijer, W.; Hallard, D.; Verpoorte, R. *Curr. Med. Chem.* **2004**, *11*, 607-628.

53. O'Connor, S. E.; Maresh, J. J. *Nat. Prod. Rep.* **2006**, *23*, 532-547.

54. Ishikawa, H.; Colby, D. A.; Seto, S.; Va, P.; Tam, A.; Kakei, H.; Rayl, T. J.; Hwang, I.; Boger, D. L. *J. Am. Chem. Soc.* **2009**, *131*, 4904-4916.

55. Treimer, J. F.; Zenk, M. H. *Eur. J. Biochem.* **1979**, *101*, 225-233.

56. Mizukami, H.; Nordlov, H.; Lee, S.-L.; Scott, A. I. *Biochemistry* **1979**, *18*, 3760-3763.

57. McCoy, E.; Galan, M. C.; O'Connor, S. E. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2475-2478.

58. Rueffer, M.; Zenk, M. H. *Z. Naturforsch.* **1987**, *42*, 319-332.

59. Stadler, R.; Zenk, M. H. *Liebigs Ann. Chem.* **1990**, , 555-562.

60. Loeffler, S.; Stadler, R.; Nagakura, N.; Zenk, M. H. *J. Chem. Soc. Chem. Commun.* **1987**, 1160-1162.

61. Schumacher, H.-M.; Rüffer, M.; Nagakura, N.; Zenk, M. H. *J. Med. Plant. Res* **1983**, *48*, 212-220.

62. Samanani, N.; Facchini, P. J. *Planta* **2001**, *213*, 898-906.

63. Luk, L. Y. P.; Bunn, S.; Liscombe, D. K.; Facchini, P. J.; Tanner, M. E. *Biochemistry* **2007**, *46*, 10153-10161.

64. Online Rare Codon Analysis: *http://www.doe-mbi.ucla.edu/~sumchan/caltor.html*

65. Voet, D.; Voet, J. G.; Pratt, C. W. In *Fundamentals of Biochemistry;* John Wiley & Sons: New York, 1998; , pp 69-71, 844-885.

66. Nakamura, Y.; Gojobori, T.; Ikemura, T. *Nucl. Acids Res.* **2000**, *28*, 292.

67. Murray, E. E.; Lotzer, J.; Eberle, M. *Nucl. Acids Res.* **1989**, *17*, 477-498.

68. pET System Manual, 10th Edition, Novagen, Inc.

69. BL21-CodonPlus® Competent Cells Instruction Manual, Stratagene, Inc.

70. Koshland, D. E., Jr.; Hamadani, K. *J. Biol. Chem.* **2002**, *277*, 46841-46844.

71. Meyers, A. I.; Boes, M.; Dickman, M. A. *Angew. Chem. Int. Ed.* **1984**, *23*, 458-459.

72. Kresge, A. J., Ed.; In *Magnitude of Primary Hydrogen Isotope Effects;* Cleland, W. W., O'Leary, M. H. and Northrop, D. B., Isotope Effects on Enzyme-Catalyzed Reactions; University Park Press: Baltimore, 1976; pp 37-63.

73. Klinman, J. P. In *Primary Hydrogen Isotope Effects;* Gandour, R. D., Schowen, R. L., Eds.; Transition States of Biochemical Processes; Plenum Press: New York, 1978; pp 165-200.

74. Melander, L.; Saunders Jr., W. H. In *Reaction Rates of Isotope Molecules;* John Wiley & Sons: New York, 1980; pp 91-201.

75. Kirsch, J. F. In *Secondary Kinetic Isotope Effects;* Cleland, W. W., O'Leary, M. H. and Northrop, D. B., Eds.; Isotope Effects on Enzyme-Catalyzed Reactions; University Park Press: Baltimore, 1976; pp 100-121.

76. Hogg, J. L. In *Secondary Hydrogen Isotope Effects;* Gandour, R. D., Schowen, R. L., Eds.; Transition States of Biochemical Processes; Plenum Press: New York, 1978; pp 201-224.

77. Harris, D. C. In *Quantitative Chemical Analysis;* Experimental Error; W.H. Freeman and Company: New York, 2001; pp 51-64.

78. Hansch, C.; Leo, A.; Unger, S. H.; Kim, K. H.; Nikaitani, D.; Lien, E. J. *J. Med. Chem.* **1973**, *16*, 1207-1216.

79. Zollinger, H. *Adv. Phys. Org. Chem.* **1964**, *2*, 163-199.

80. Melander, L. In *Isotope Effects on Reaction Rates;* Ronald Press: New York, 1960; pp 107-122.

81. Streitwieser, A.; Jagow, R. H.; Fahey, R. C.; Suzuki, S. *J. Am. Chem. Soc.* **1958**, *80*, 2326-2332.

82. Bethell, D.; Gold, V. *J. Chem. Soc.* **1958**, 1905-1915.

83. Garrett, B. C.; Truhlar, D. G. *J. Am. Chem. Soc.* **1980**, *102*, 2559-2570.

84. Westheimer, F. H. *Chem. Rev.* **1961**, *61*, 265-273.

85. Ilari, A.; Franceschini, S.; Bonamore, A.; Arenghi, F.; Botta, B.; Macone, A.; Pasquo, A.; Bellucci, L.; Boffi, A. *J. Biol. Chem.* **2009**, *284*, 897-904.

86. Bonamore, A.; Barba, M.; Botta, B.; Boffi, A.; Macone, A. *Molecules* **2010**, *15*, 2070-2078.

87. Timberlake, C. F. *J. Chem. Soc.* **1957**, 4987-4993.

88. Hirose, T.; Sunazuka, T.; Zhi-Ming, T.; Handa, M.; Uchida, R.; Shiomi, K.; Harigaya, Y.; Omura, S. *Heterocycles* **2000**, *53*, 777-784.

89. Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248-254.

90. Laemmli, U. K. *Nature* **1970**, *227*, 680-685.

91. Glasoe, P. K.; Long, F. A. *J. Phys. Chem.* **1960**, *64*, 188-190.

92. Ausubel, F. M.; Brent, R.; Kingston, R. E.; Moore, D. D.; Seidman, J. G.; Smith, J. A.; Struhl, K. In *Short Protocols in Molecular Biology (A Compendium of Protocols from Current Protocols in Molecular Biology;* Green Publishing Associates and John Wiley and Sons: New York, 1992.

93. Battersby, A. R.; Jones, R. C. F.; Kazlauskas, R.; Thornber, C. W.; Ruchirawat, S.; Staunton, J. *J. Chem. Soc. Perkin Trans. 1* **1981**, 2016-2029.

94. Vining, R. F.; Smythe, G. A.; Long, M. A. *J. Labelled Compd. Radiopharm.* **1981**, *18*, 1683-1692.

95. Liang, P.-H.; Ko, T.-P.; Wang, A. H.-J. *Eur. J. Biochem.* **2002**, *269*, 3339-3354.

96. Tello, M.; Kuzuyama, T.; Heide, L.; Noel, J. P.; Richard, S. B. *Cell. Mol. Life. Sci.* **2008**, *65*, 1459-1463.

97. Poulter, C. D.; Rilling, H. C. *Acc. Chem. Res.* **1978**, *11*, 307-313.

98. Dorsey, J. K.; Dorsey, J. A.; Porter, J. W. *J. Biol. Chem.* **1966**, *241*, 5353-5360.

99. Holloway, P. W.; Popják, G. *Biochem. J.* **1967**, *104*, 57-70.

100. Cornforth, J. W. *Angew. Chem. Int. Ed.* **1968**, *7*, 903-911.

101. Popják, G.; Cornforth, J. W. *Biochem. J.* **1966**, *101*, 553-568.

102. Cornforth, J. W.; Cornforth, R. H.; Popják, G.; Yengoyan, L. *J. Biol. Chem.* **1966**, *241*, 3970-3987.

103. Poulter, C. D.; Satterwhite, D. M.; Rilling, H. C. *J. Am. Chem. Soc.* **1976**, *98*, 3376-3377.

104. Okamoto, Y.; Inukai, T.; Brown, H. C. *J. Am. Chem. Soc.* **1958**, *80*, 4969-4972.

105. McBee, E. T.; Pierce, O. R.; Smith, D. D. *J. Am. Chem. Soc.* **1954**, *76*, 3725-3728.

106. Poulter, C. D.; Argyle, J. C.; Mash, E. A. *J. Am. Chem. Soc.* **1977**, *99*, 957-959.

107. Poulter, C. D.; Argyle, J. C.; Mash, E. A. *J. Biol. Chem.* **1978**, *253*, 7227-7233.

108. Mash, E. A.; Gurria, G. M.; Poulter, C. D. *J. Am. Chem. Soc.* **1981**, *103*, 3927-3929.

109. Gassman, P. G.; Fentiman, A. F. *J. Am. Chem. Soc.* **1970**, *92*, 2549-2551.

110. Poulter, C. D.; Mash, E. A.; Argyle, J. C.; Muscio, O. J.; Rilling, H. C. *J. Am. Chem. Soc.* **1979**, *101*, 6761-6763.

111. Tarshis, L. C.; Yan, M.; Poulter, C. D.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 10871-10877.

112. Tarshis, L. C.; Proteau, P.; Kellogg, B. A.; Sacchettini, J. C.; Poulter, C. D. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 15018-15023.

113. Huang, C.-C.; Hightower, K. E.; Fierke, C. A. *Biochemistry* **2000**, *39*, 2593-2602.

114. Cohen, L. H.; Pieterman, E.; van Leeuwen, R. E. W.; Overhand, M.; Burm, B. E. A.; van der Marel, G. A.; van Boom, J. H. *Biochem. Pharmacol.* **2000**, *60*, 1061-1068.

115. Dolence, J. M.; Poulter, C. D. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 5008-5011.

116. Richard, J. P.; Jencks, W. P. *J. Am. Chem. Soc.* **1984**, *106*, 1383-1396.

117. Jencks, W. P. *Chem. Soc. Rev.* **1981**, *10*, 345-375.

118. Weller, V. A.; Distefano, M. D. *J. Am. Chem. Soc.* **1998**, *120*, 7975-7976.

119. Hegazi, M. F.; Borchardt, R. T.; Schowen, R. L. *J. Am. Chem. Soc.* **1979**, *101*, 4359-4365.

120. Strickland, C. L.; Windsor, W. T.; Syto, R.; Wang, L.; Bond, R.; Wu, Z.; Schwartz, J.; Le, H. V.; Beese, L. S.; Weber, P. C. *Biochemistry* **1998**, *37*, 16601-16611.

121. Mu, Y. Q.; Omer, C. A.; Gibbs, R. A. *J. Am. Chem. Soc.* **1996**, *118*, 1817-1823.

122. Hightower, K. E.; Huang, C.-C.; Casey, P. J.; Fierke, C. A. *Biochemistry* **1998**, *37*, 15555-15562.

123. Hemmi, H.; Shibuya, K.; Takahashi, Y.; Nakayama, T.; Nishino, T. *J. Biol. Chem.* **2004**, *279*, 50197-50203.

124. Yazaki, K.; Kunihisa, M.; Fujisaki, T.; Sato, F. *J. Biol. Chem.* **2002**, *277*, 6240-6246.

125. Kuzuyama, T.; Noel, J. P.; Richard, S. B. *Nature* **2005**, *435*, 983-987.

126. Metzger, U.; Schall, C.; Zocher, G.; Unsöld, I.; Stec, E.; Li, S.-M.; Heide, L.; Stehle, T. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 14309-14314.

127. Schultz, A. W.; Lewis, C. A.; Luzung, M. R.; Baran, P. S.; Moore, B. S. *J. Nat. Prod.* **2010**, *73*, 373-377.

128. Pojer, F.; Wemakor, E.; Kammerer, B.; Chen, H.; Walsh, C. T.; Li, S.-M.; Heide, L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 2316-2321.

129. Pojer, F.; Li, S.-M.; Heide, L. *Microbiology* **2002**, *148*, 3901-3911.

130. Steffan, N.; Unsöld, I. A.; Li, S.-M. *ChemBioChem* **2007**, *8*, 1298-1307.

131. Unsöld, I. A.; Li, S.-M. *Microbiology* **2005**, *151*, 1499-1505.

132. Cress, W. A.; Chayet, L. T.; Rilling, H. C. *J. Biol. Chem.* **1981**, *256*, 10917-10923.

133. Unsöld, I. A.; Li, S.-M. *ChemBioChem* **2006**, *7*, 158-164.

134. Rementeria, A.; López-Molina, N.; Ludwig, A.; Vivanco, A. B.; Bikandi, J.; Pontón, J.; Garaizar, J. *Rev. Iberoam. Micol.* **2005**, *22*, 1-23.

135. Kremer, A.; Westrich, L.; Li, S.-M. *Microbiology* **2007**, *153*, 3409-3416.

136. Ding, Y.; Williams, R. M.; Sherman, D. H. *J. Biol. Chem.* **2008**, *283*, 16068-16076.

137. Floss, H. G. *Tetrahedron* **1976**, *32*, 873-912.

138. Weber, J. T.; O'Connor, M.-F.; Hayataka, K.; Colson, N.; Medora, R.; Russo, E. B.; Parker, K. K. *J. Nat. Prod.* **1997**, *60*, 651-653.

139. Floss, H. G. *J. Nat. Prod.* **2006**, *69*, 158-169.

140. Lakhdar, S.; Westermaier, M.; Terrier, F.; Goumont, R.; Boubaker, T.; Ofial, A. R.; Mayr, H. *J. Org. Chem.* **2006**, *71*, 9088-9095.

141. Gebler, J. C.; Poulter, C. D. *Arch. Biochem. Biophys.* **1992**, *296*, 308-313.

142. Tsai, H. F.; Wang, H.; Gebler, J. C.; Poulter, C. D.; Schardl, C. L. *Bioc. Biophys. Res. Comm* **1995**, *216*, 119-125.

143. Gupta, N. K.; Ohtsuka, E.; Sgaramella, V.; Buchi, H.; Kumar, A.; Weber, H.; Khorana, H. G. *Proc. Natl. Acad. Sci. U. S. A.* **1968**, *60*, 1338-1344.

144. Petrie, C. R.; Reed, M. W.; Adams, A. D.; Meyer, R. B. J. *Bioconjugate Chem.* **1992**, *3*, 85-87.

145. Stemmer, W. P. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 10747-10751.

146. Young, L.; Dong, Q. *Nucl. Acids Res.* **2004**, *32*, e59.

147. Hoover, D. M.; Lubkowski, J. *Nucl. Acids Res.* **2002**, *30*, e43.

148. Stemmer, W. P. C.; Crameri, A.; Ha, K. D.; Brennan, T. M.; Heyneker, H. L. *Gene* **1995**, *164*, 49-53.

149. Gao, X.; Yo, P.; Keith, A.; Ragan, T. J.; Harris, T. K. *Nucl. Acids Res.* **2003**, *31*, e143.

150. Carr, P. A.; Park, J. S.; Lee, Y.-J.; Yu, T.; Zhang, S.; Jacobson, J. M. *Nucl. Acids Res.* **2004**, *32*, e162.

151. Rosetta(DE3) pLysS Competent Cells User Manual. Novagen, Inc.

152. Lee, S.-L.; Floss, H. G.; Heinstein, P. *Arch. Biochem. Biophys.* **1976**, *177*, 84-94.

153. Midelfort, C. F.; Rose, I. A. *J. Biol. Chem.* **1976**, *251*, 5881-5887.

154. Cohn, M.; Hu, A. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 200-203.

155. Cohn, M.; Hu, A. *J. Am. Chem. Soc.* **1980**, *102*, 913-916.

156. Leonard, N. J.; Frihart, C. R. *J. Am. Chem. Soc.* **1974**, *96*, 5894-5903.

157. Vani, P. V. S. N.; Chida, A. S.; Srinivasan, R.; Chandrasekharam, M.; Singh, A. K. *Syn. Comm.* **2001**, *31*, 219-224.

158. Cane, D. E.; Iyengar, R.; Shiao, M.-S. *J. Am. Chem. Soc.* **1981**, *103*, 914-931.

159. Tiers, G. V. D. *J. Am. Chem. Soc.* **1957**, *79*, 5585.

160. Cho, Y. K.; Matsunaga, T. O.; Kenyon, G. L.; Bertagnolli, B. L.; Cook, P. F. *Biochemistry* **1988**, *27*, 3320-3325.

161. Lowe, G.; Potter, B. V. L. *J. Chem. Soc. Chem. Commun.* **1984**, 877-879.

162. Marschner, T. M.; Reynolds, M. A.; Oppenheimer, N. J.; Kenyon, G. L. *J. Chem. Soc. Chem. Commun.* **1983**, 1289-1290.

163. Potter, B. V. L. *J. Labelled Compd. Radiopharm.* **1989**, *27*, 955-963.

164. Laskovics, F. M.; Poulter, C. D. *Biochemistry* **1981**, *20*, 1893-1901.

165. Yamamitsu, T.; Ohta, S.; Suga, T. *J. Chem. Soc. Perkin Trans. 1* **1989**, 1811-1814.

166. Saito, I.; Sugiyama, H.; Yamamoto, A.; Muramatsu, S.; Matsuura, T. *J. Am. Chem. Soc.* **1984**, *106*, 4286-4287.

167. Croteau, R. B.; Shaskus, J. J.; Renstrom, B.; Felton, N. M.; Cane, D. E.; Saito, A.; Chang, C. *Biochemistry* **1985**, *24*, 7077-7085.

168. Luk, L. Y. P.; Tanner, M. E. *J. Am. Chem. Soc.* **2009**, *131*, 13932-13933.

169. Wierenga, R. K. *FEBS Lett.* **2001**, *492*, 193-198.

170. Edward, J. T.; Wong, S. C. *Can. J. Chem.* **1977**, *55*, 2492-2494.

171. Mayr, H.; Kempf, B.; Ofial, A. R. *Acc. Chem. Res.* **2003**, *36*, 66-77.

172. Steffan, N.; Unsöld, I. A.; Li, S.-M *ChemBioChem* **2007**, *8*, 1298-1307.

173. Thulasiram, H. V.; Phan, R. M.; Rivera, S. B.; Poulter, C. D. *J. Org. Chem.* **2006**, *71*, 1739-1741.

174. Poulter, C. D.; Satterwhite, D. M. *Biochemistry* **1977**, *16*, 5470-5478.

175. Webb, M. R. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 4884-4887.

176. Saito, I.; Sugiyama, H.; Yamamoto, A.; Muramatsu, S.; Matsuura, T. *J. Am. Chem. Soc.* **1984**, *106*, 4286-4287.

Appendix

NMR Spectra of Purified Products

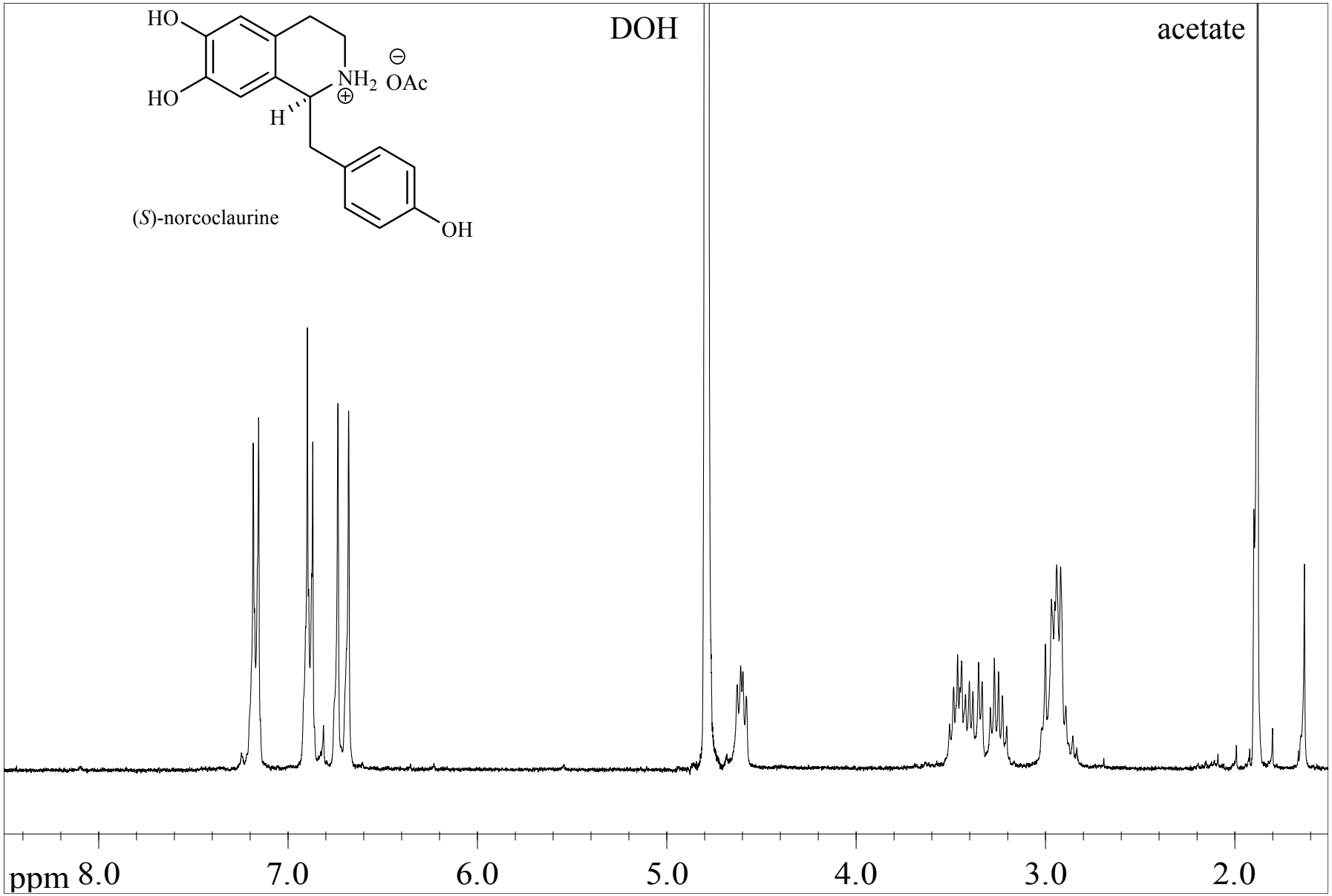


Figure A.1 ^{1}H NMR spectrum of racemized norcoclaurine

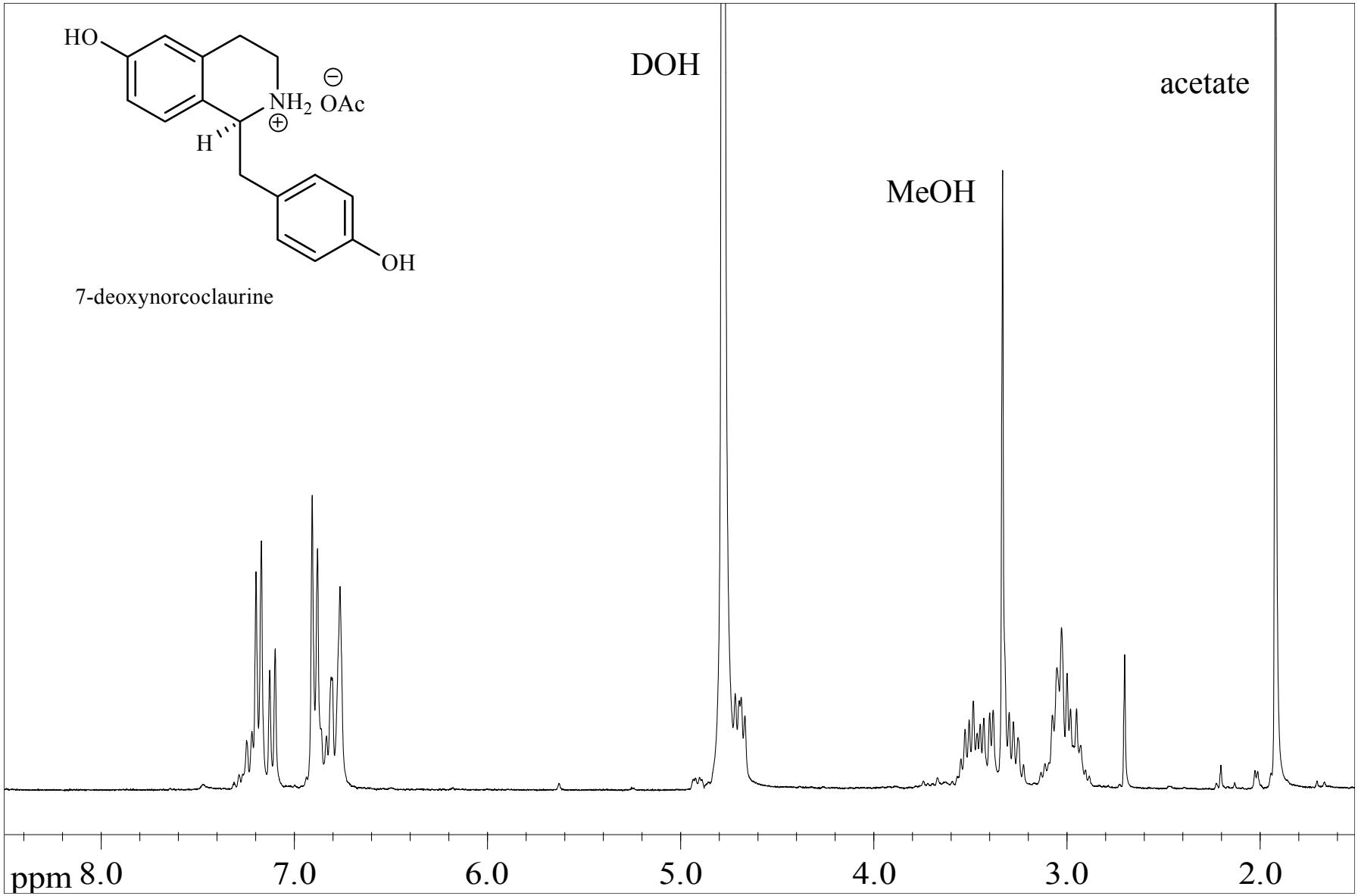


Figure A.2 ^{1}H NMR spectrum of 7-deoxynorcoclaurine

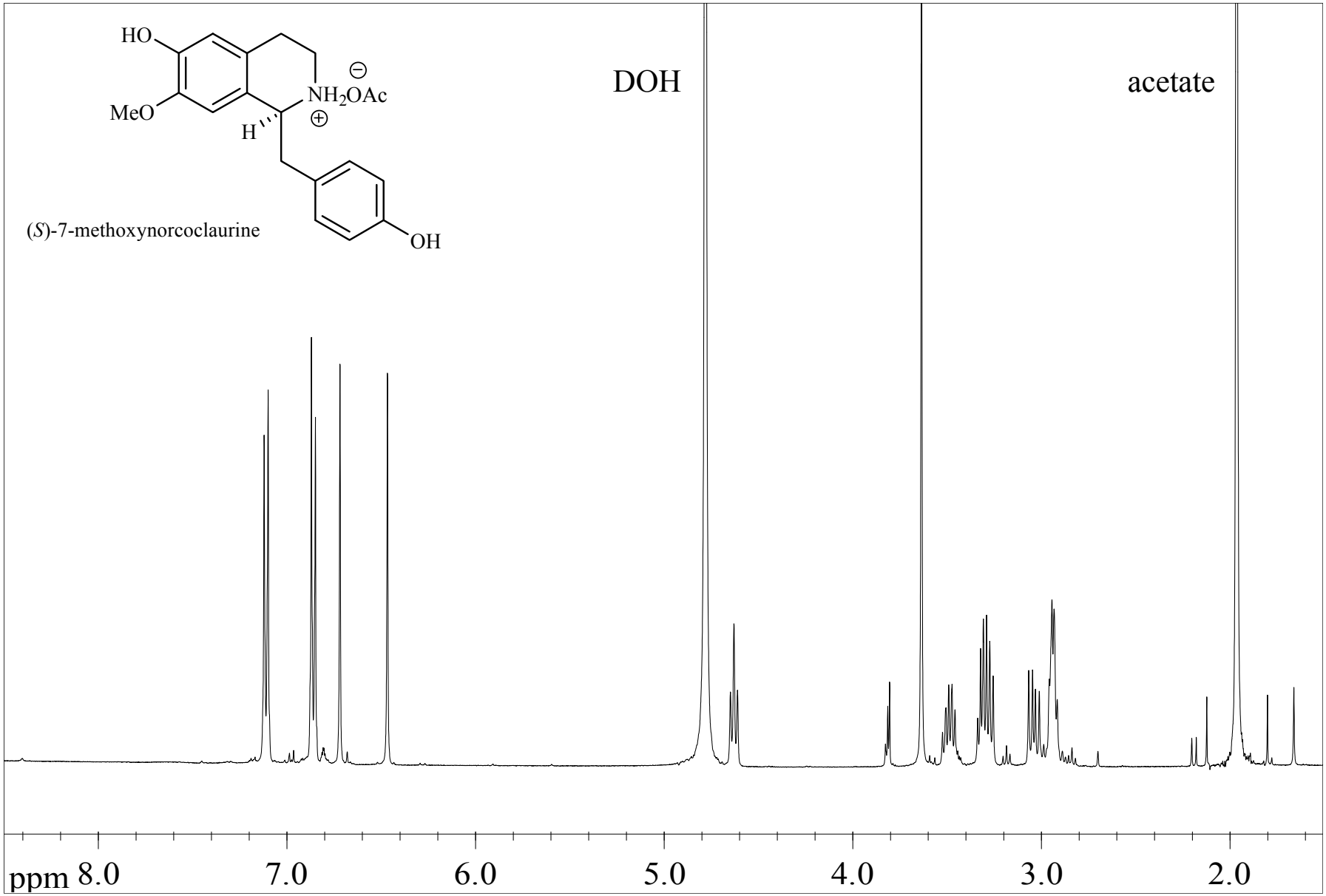


Figure A.3 ^{1}H NMR spectrum of 7-methoxynorcoclaurine

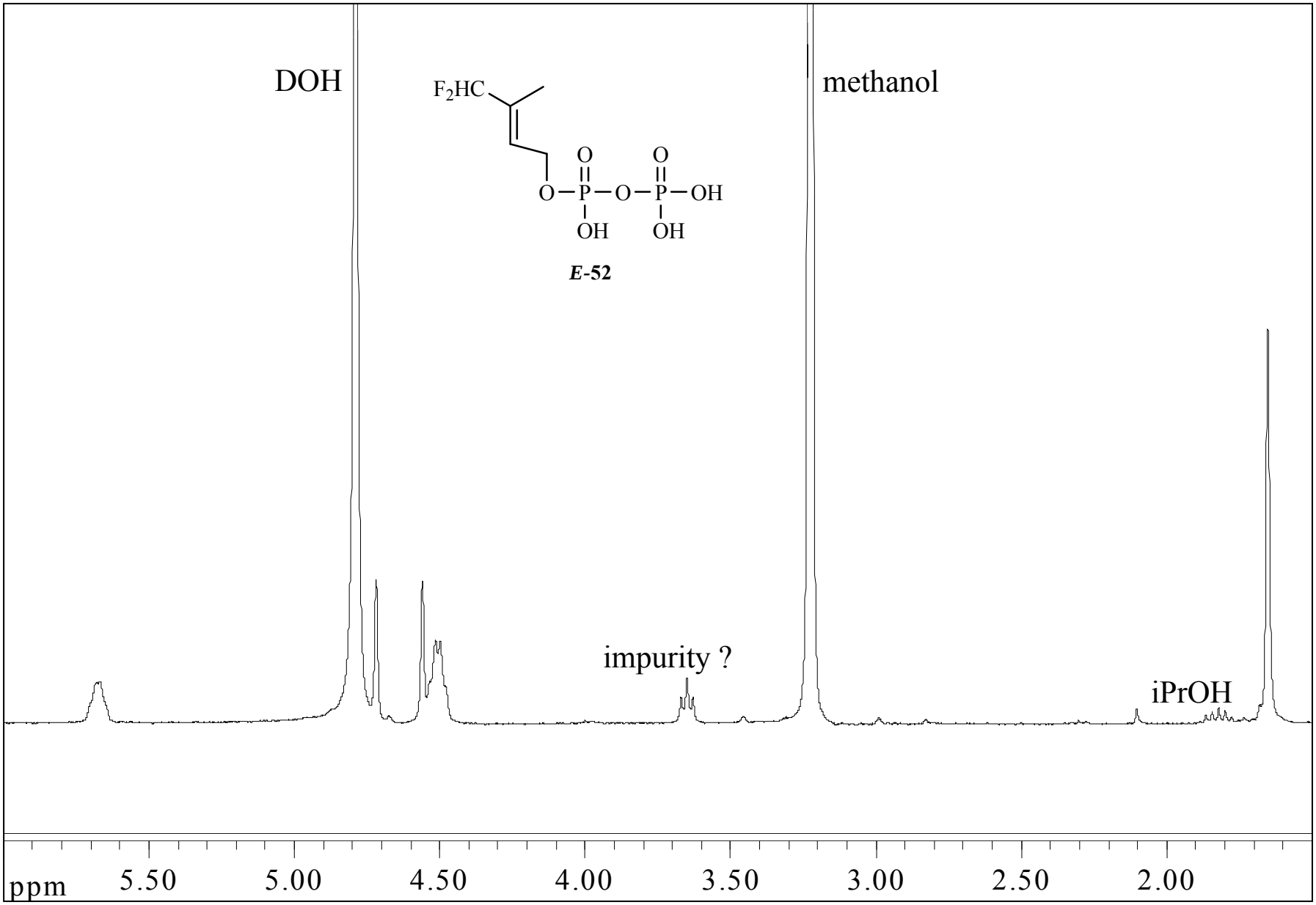


Figure A.4 ^{1}H NMR spectrum of *E*-3-(fluoromethyl)-2-butenyl diphosphate (*E*-52)

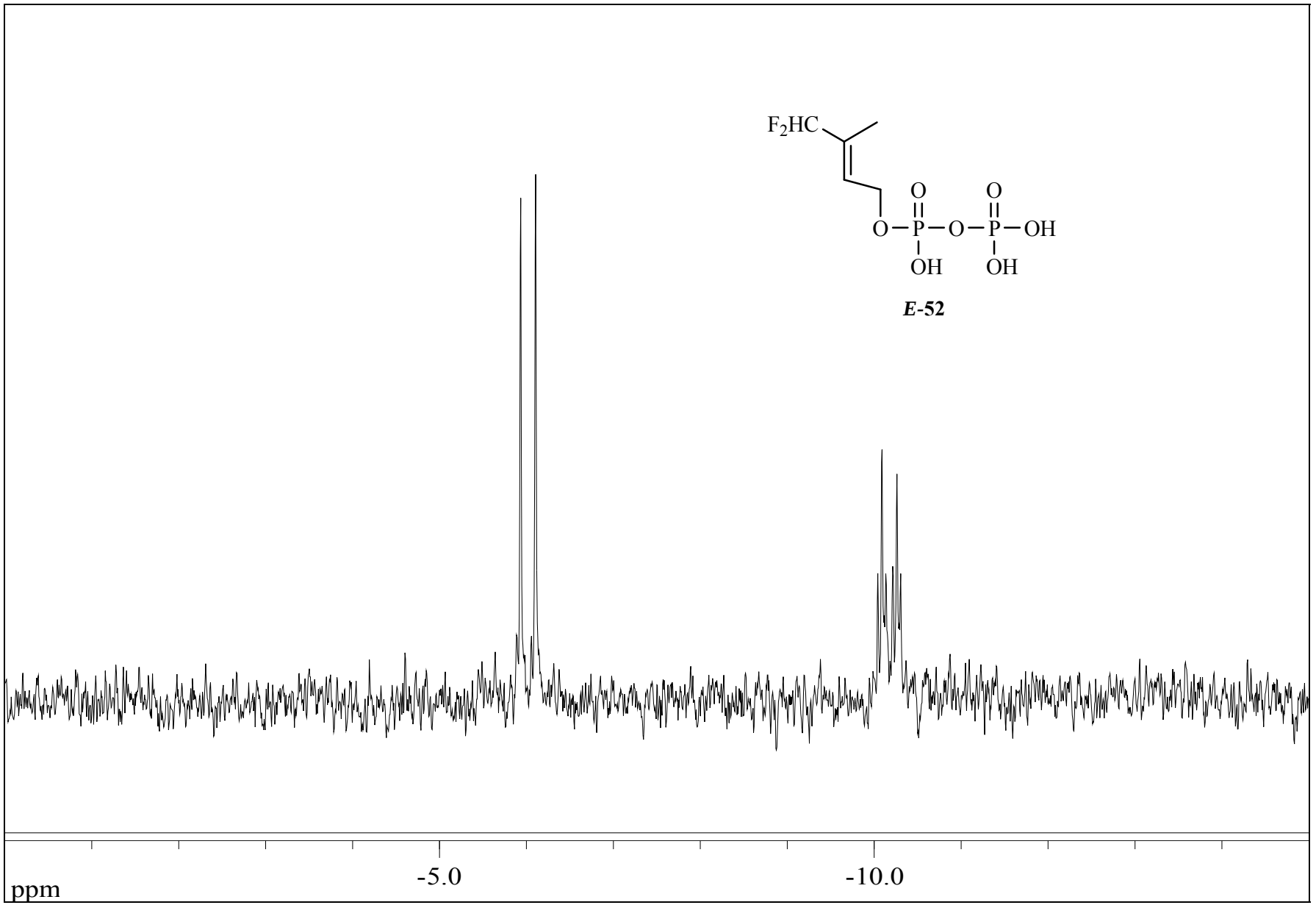


Figure A.5 Proton-coupled ^{31}P NMR spectrum of *E*-3-(fluoromethyl)-2-butenyl diphosphate (*E*-52)

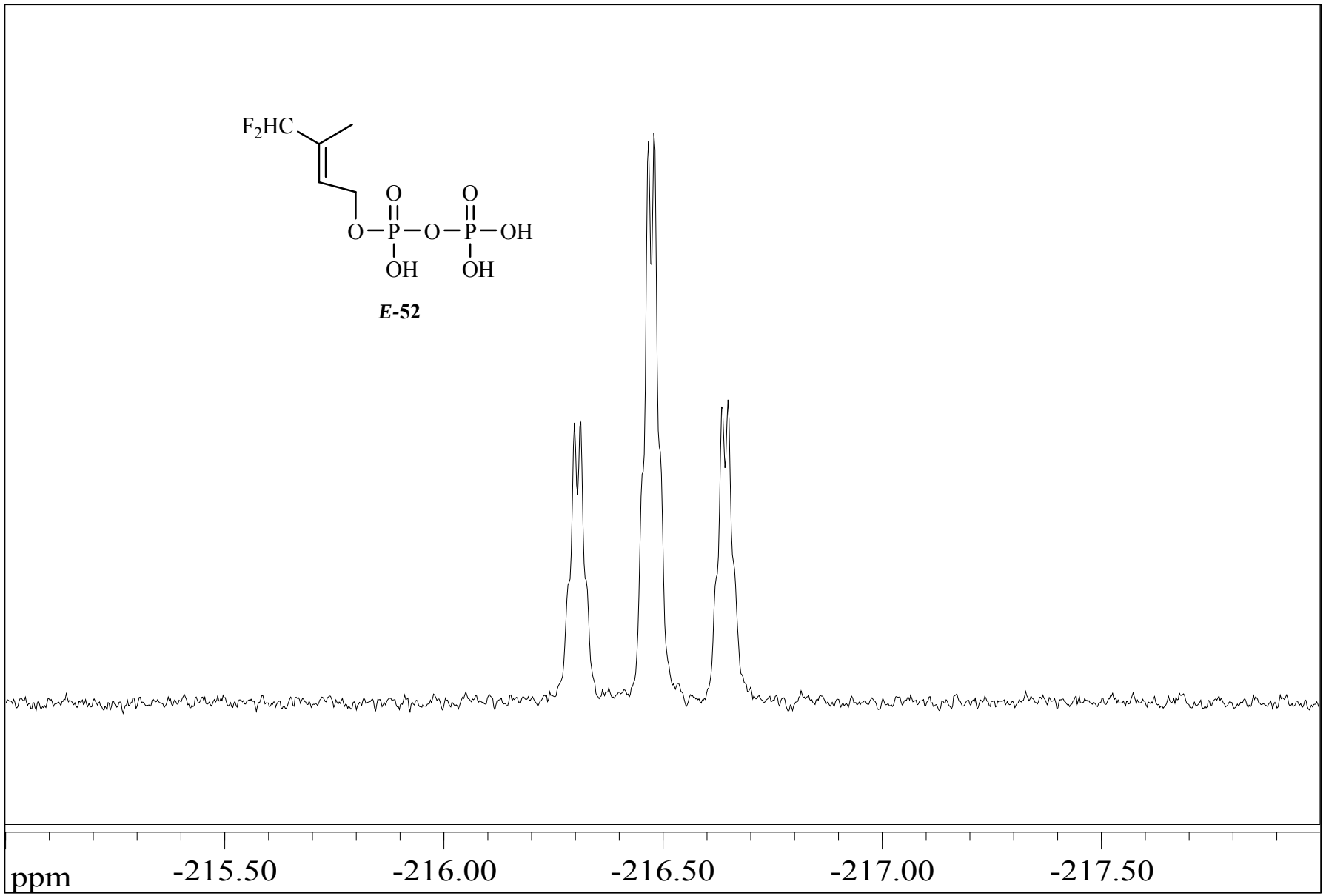


Figure A.6 ^{19}F NMR spectrum of *E*-3-(fluoromethyl)-2-butenyl diphosphate (*E*-52)

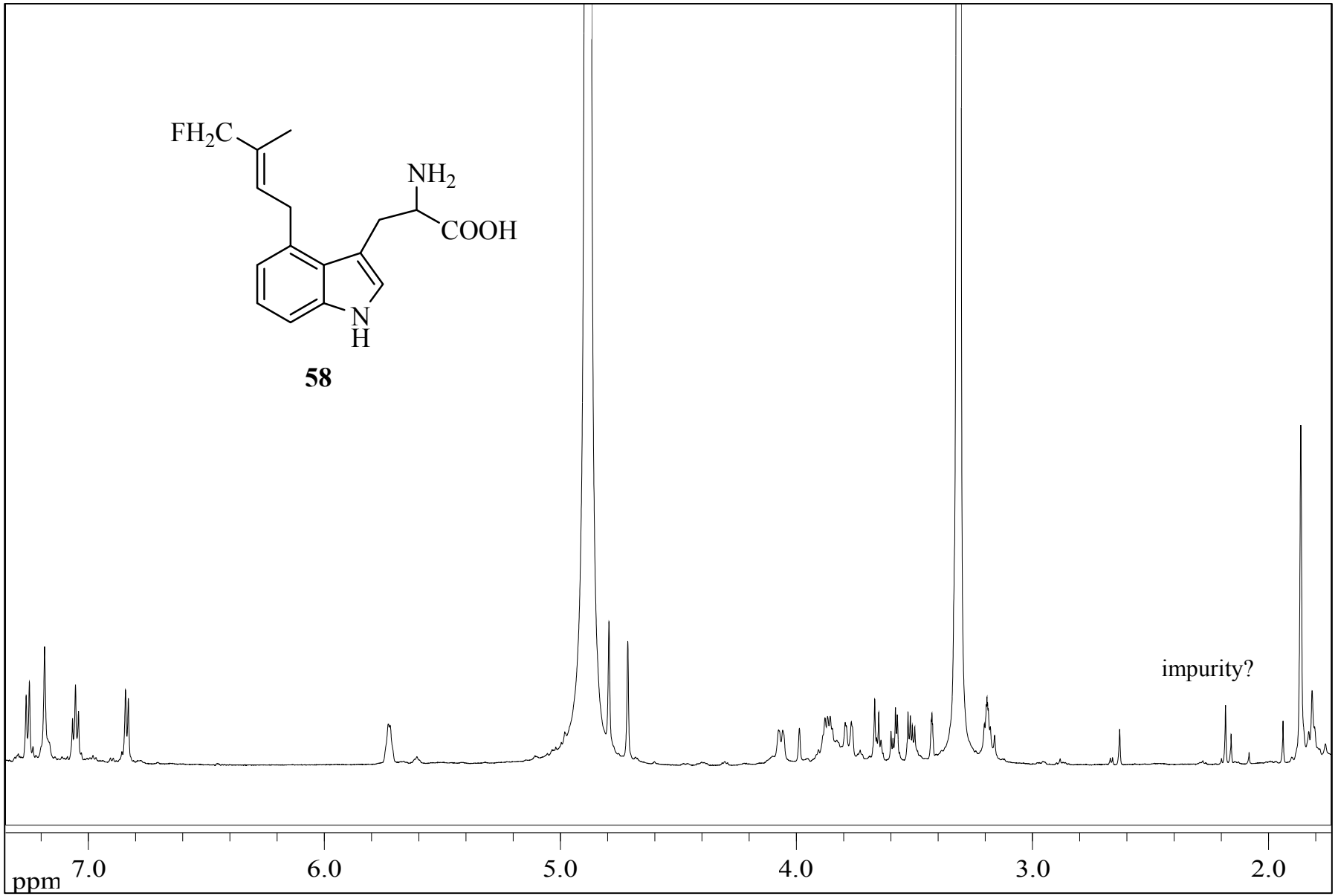


Figure A.7 ^{1}H NMR spectrum of fluorinated dimethylallyltryptophan (58) isolated from the DMATS reaction using fluorinated DMAPP analog (*E*-52)

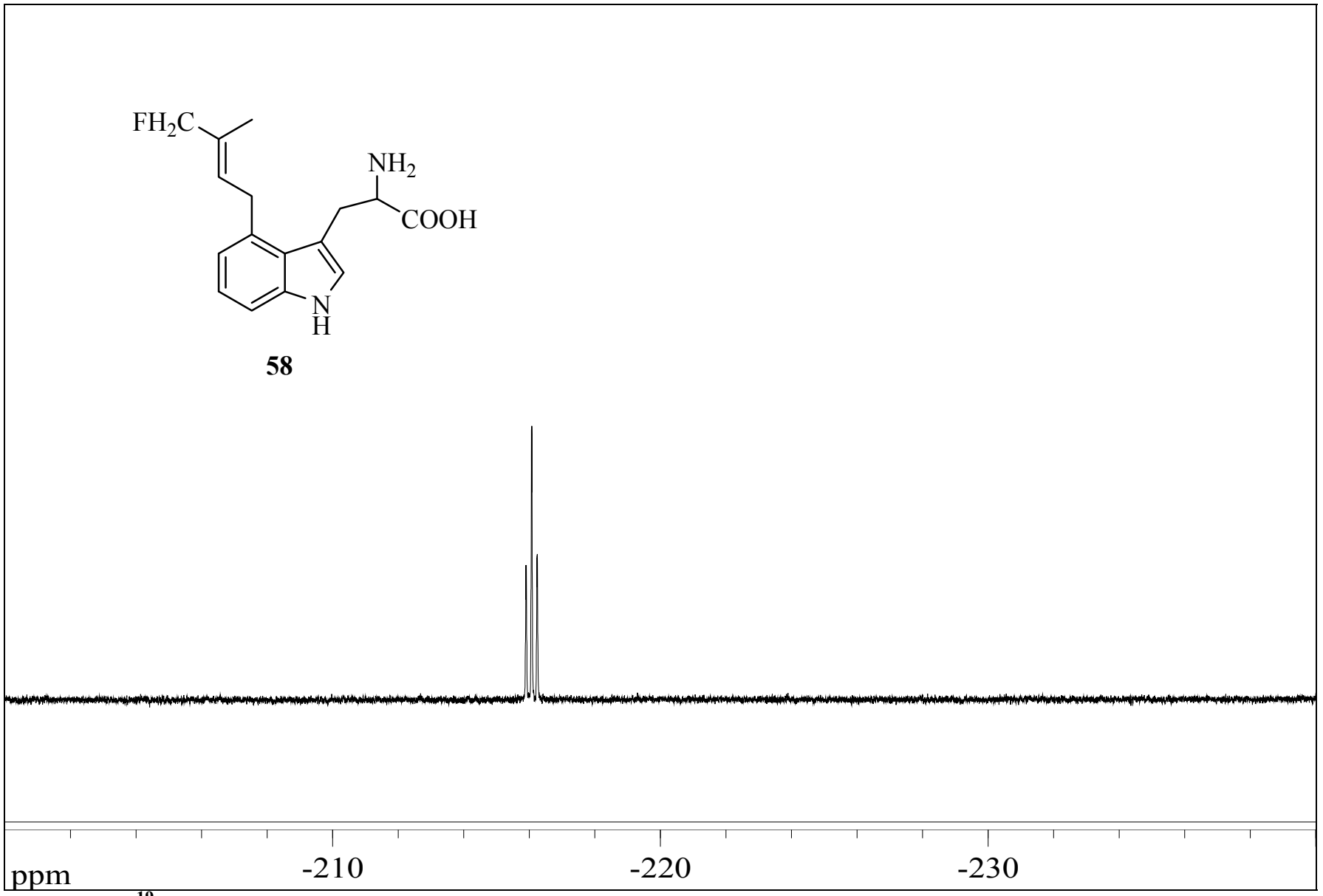


Figure A.8 ^{19}F NMR spectrum of fluorinated dimethylallyltryptophan (58) isolated from the DMATS reaction using fluorinated DMAPP analog (*E*-52)