MECHANISTIC STUDIES ON TRYPTOPHAN 2,3-DIOXYGENASE, AND CHARACTERIZATION OF GENES INVOLVED IN VITAMIN B_6 AND PYRIMIDINE CATABOLISM

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Tathagata Mukherjee

August 2009

MECHANISTIC STUDIES ON TRYPTOPHAN 2,3-DIOXYGENASE, AND CHARACTERIZATION OF GENES INVOLVED IN VITAMIN B_6 AND PYRIMIDINE CATABOLISM

Tathagata Mukherjee, Ph. D.

Cornell University 2009

Tryptophan 2,3-dioxygenase catalyzes the first and the rate limiting step in the kynurenine pathway of tryptophan degradation to form N-formyl kynurenine. Mechanistic and kinetic studies were done on this enzyme and the results were shown to favor the Criegee rearrangement as the most likely pathway for the catalytic transformation.

There are two characterized catabolic routes for vitamin B_6 degradation. Even though the enzymes involved in the degradation of vitamin B_6 have been characterized, no information on the genes encoding for these enzymes were available for a long time. Recently, 4 out of the 7 genes involved in Vitamin B_6 catabolism were identified from *Mesorhizobium loti*, while the other 3 genes remained elusive. Bioinformatics search revealed three candidates which could function as an oxidoreductase (mlr6793), a decarboxylase (mlr6791) and a hydrolase (mlr6787). These were cloned and heterologously overexpressed in *Escherichia coli*. Each protein was purified to homogeneity and *in vitro* biochemical studies were carried out to rigorously characterize mlr6793, mlr6791 and mlr6787 gene-products to be 4-pyridoxic acid dehydrogenase, 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase and 2-(acetamidomethylene)succinate hydrolase respectively.

Apart from the well known reductive and oxidative catabolic pathway of uracil (or thymine) another novel pyrimidine degradation pathway was very recently discovered in *Escherichia coli* K12. The genes participating in this pathway were identified but no

information on the enzymology and intermediates involved in this pathway were known. Our work characterized the intermediates of the pathway and identified three enzymes in this novel pathway that were able to convert uracil to 3-oxopropionate. The first enzyme in this pathway is RutA which does a novel flavin dependant hydrolysis of uracil. With the crystal structure of this protein and model chemistry we were able to hypothesize a plausible mechanism for the reaction catalyzed by RutA.

BIOGRAPHICAL SKETCH

Tathagata was born on a sunny afternoon on 25th June in the year 1981. He grew up in the industrial township of Burnpur where he received his education from St. Vincent's High and Technical School. His father was a Chemical Engineer and was instrumental in motivating him towards science. His mother was always there to answer any queries that he might have. Through the explanations from his mother and motivation from his father very early on he had decided that he wanted to take up Science for higher studies. He loved Biology and at one point of time had considered to become a physician. Fate, however had other plans. After completing his high school he took admission at St. Stephen's College, Delhi where he studied Chemistry. The choice of Chemistry as a subject was made primarily because the college did not offer Biology and he hated Physics. There he was very ably guided by Prof. Eshwaran, who not only made him love Organic Chemistry but also taught him Biochemistry, which was never a part of the curriculum. Serendipitously he was amongst brilliant likeminded students with whom he would often discuss new scientific discoveries in the filed of Biochemistry. He then went on to do Masters in Chemistry from the prestigious Indian Institute of Technology, Delhi, where he did a project in Biochemistry under Prof. Gupta, who encouraged him to pursue a PhD degree. He then travelled half the world and came to Cornell University to work in the field of Bioorganic Chemistry with Prof. Tadhg Begley, who taught him not only to enjoy research but absolutely love it. Half way through his PhD degree he was introduced to Paulami, the love of his life, whom he married in 2008. Paulami's constant encouragement and support in the final stages pushed him hard to complete the PhD requirements in time to graduate in the summer of 2009. Both of them will be moving to Maryland where Tathagata will be pursuing post doctoral research with Dr. Clifton Barry III at NIH doing research in the field of tuberculosis.

To

Ma, Baba, Sankha &

my beloved wife Paulami

ACKNOWLEDGMENTS

I take this opportunity to express my heartfelt gratitude towards all those who, with their sincere help and cooperation, have been instrumental in my getting a PhD degree.

Tadhg, I could never thank you enough. You have been a great mentor. You have always believed in me and have motivated me to do better. At times I felt you were more excited by my results than I was, I guess that was your way of passing it on to me. Thank you for your encouragement and support. Thanks to my committee members, Steve for the suggestions and insightful discussions and Peng for providing me help whenever I needed.

Keri, AmyH, Kristin, Grant, Tim, Jeremiah, AmyG, Jennie, Dave, Abhishek, Kalyan, Sean, Amrita, Debashree, Sameh, Ahmed, Dinuka, Udami, Tomoshige thank you all for training me in various ways and making my stay in the lab very memorable. All the suggestions and insightful comments helped me troubleshoot my experiments and taught me more than any amount of reading would have ever done. Katie and Yang for being such excellent collaborators and teaching me crystallography. Cynthia, thank you for all the mutants and teaching me molecular biology. Ivan thanks for helping me out with the designing of the complex NMR experiments. Sharon, for taking care of so many things for me and helping me with all the paper works. You had answers to all the questions, nothing seemed difficult when you were around. Brendon thanks for providing a clean environment for work, Denise thanks for calling immediately whenever a shipment arrived and June for taking care of all the last minute orders. I would like to thank all my friends who made my staying in Ithaca very enjoyable. I would like to thank the Department of Chemistry and Chemical Biology for its resources and the lovely environment for nurturing science, the ISSO for taking such great care of international students and Cornell University for providing such a congenial atmosphere for education and the Graduate School for all the help.

I would like to thank all my teachers for preparing me for this day, my parents for their teaching, motivation and inspiration, my brother for his constant encouragement and my wife Paulami for being my best friend. Paulami, your selfless support during the time when most needed cannot be acknowledged in mere words. Thanks for being such a patient listener. You held my hands through the highs and lows of my life.

Last but not the least I would like to thank God, without whose grace this work would not have seen the light of the day.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ATCC	American Type Culture Collection
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI-FTMS	Electrospray Ionization – Fourier Transform Mass Spectrometry
FAD	Flavin Adenine Dinucleotide
FMN	Flavin Adenine Mononucleotide
HPLC	High Performance Liquid Chromatography
HMBC	Heteronuclear Multiple Bond Coherence
HMPDdc	3-hydroxy-2-methyl-4,5-dicarboxylate decarboxylase
IDO	Indoleamine-2,3-dioxygenase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KFA	N-formylkynurenine formamidase
LB	Luria-Bertani Broth
LCMS	Liquid chromatography Mass Spectrometry
MWCO	Molecular Weight Cut Off
NAD	β-Nicotinamide Adenine Dinucleotide
NADP	β-Nicotinamide Adenine Dinucleotide Phosphate
NADH	β-Nicotinamide Adenine Dinucleotide, reduced form
Ni-NTA	Nickel-nitrilotriacetic acid
NMDA	N-methyl-D-aspartate
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Enhancement Spectroscopy
MHPCO	2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase
MHPC	2-methyl-3-hydroxypyridine-5-carboxylic acid
MHPD	2-methyl-3-hydroxypyridine-4,5-dicarboxylic acid
PCR	Polymerase Chain Reaction
PLP	Pyridoxal 5′-phosphate
PSI-BLAST	Position Specific Iterative Basic Local Alignment Search Tool

SDS-PAGE	Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis
SelMet	Selenomethionine
TDO	Tryptophan-2,3-dioxygenase
TFA	Trifluoroacetic acid

CHAPTER 1

Introduction

1.1 Tryptophan 2,3-dioxygenase

The amino acid tryptophan, first discovered in 1901 by F.G. Hopkins (*1*), is an essential amino acid which is not only required for protein synthesis but also is a precursor in important metabolic pathways like the biosynthesis of nicotinamide cofactors (NAD^+, NADP) in eukaryotes (*2, 3*) and neurotransmitters like serotonin and melatonin in mammals (*4*). *figure 1.1* shows the 'kynurenine pathway' (shown in blue) that catabolizes 99% of the dietary tryptophan, not assimilated into proteins, to either NAD^+ (deficiency of which causes pellagra in humans (*5, 6*)) or to CO_2, NH_3 and H_2O, while the alternative 'serotonin pathway' (shown in red) converts the rest to serotonin and melatonin (*4*). Quinolinic acid, which is the precursor to the biosynthesis of NAD also serves as an agonist for the N-methyl-D-aspertate (NMDA) receptors, while kynurenic acid (formed from kynurenine via kynurenine aminotranferase) serves as the antagonist for the NMDA receptors (*7*). Thus the flux of metabolites via the 'kynurenine pathway' has implication in the pathology of certain neurological conditions. Tryptophan-2,3-dioxygenase, TDO, catalyzes the first and the rate limiting step in the 'kynurenine pathway' of tryptophan degradation to form N-formyl kynurenine (NFK) (*8*). The mechanism by which TDO catalyzes the oxidative ring opening of the indole ring of tryptophan is poorly understood.

As outlined in the *figure 1.2* the heme iron catalyzes the formation of the tryptophan hydroperoxide (**13**) most likely by a single electron transfer mechanism. From this intermediate, three mechanistic routes to NFK (**2**) can be postulated. According to mechanism A (red), addition of water to the endocylic imine would form (**14**) which can rearrange to give NFK. In mechanism B (in pink) rearrangement of the hydroperoxide

(**13**) to a dioxetane intermediate (**16**) followed by a retro [2+2] cycloaddition can form NFK (*9, 10*).

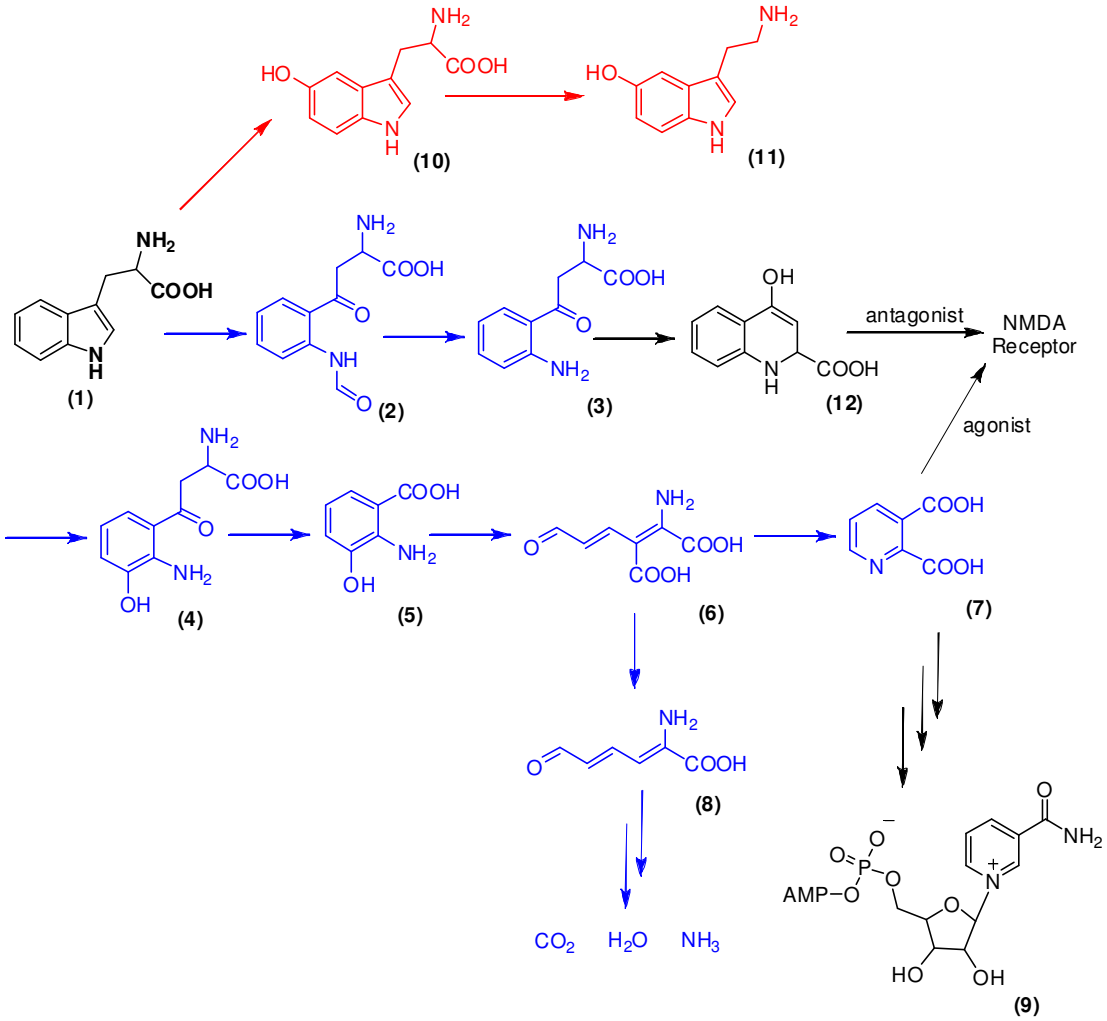


Figure 1.1: The tryptophan catabolic pathway. The 'serotonin pathway' is shown in red and the 'kynurenine pathway' is shown in blue. (**1**) tryptophan; (**2**) N-formylkynurenine; (**3**) Kynurenine; (**4**) 3-hydroxykynurenine; (**5**) 3-hydroxyanthranilate; (**6**) 2-amino-3-carboxy muconate semialdehyde; (**7**) quinolinic acid; (**8**) 2-aminomuconate semialdehyde; (**9**) nicotinamide adenine dinucleotide (NAD^+); (**10**) 5-hydroxytryptophan; (**11**) serotonin and (**12**) kynurenic acid.

By mechanism C (blue) (analogous to the mechanism of extradiol dioxygenases (*11*)), a Criegee rearrangement would give the cationic intermediate (**17**) which could then undergo a nucleophilic attack by the iron-bound oxygen anion to give (**18**) followed by its collapse to form NFK. Recombinant TDO from *Ralstonia metallidurans* was overexpressed in *E. coli* and purified. Efforts were made to elucidate the mechanism of this reaction, as detailed out in chapter 2 and 3.

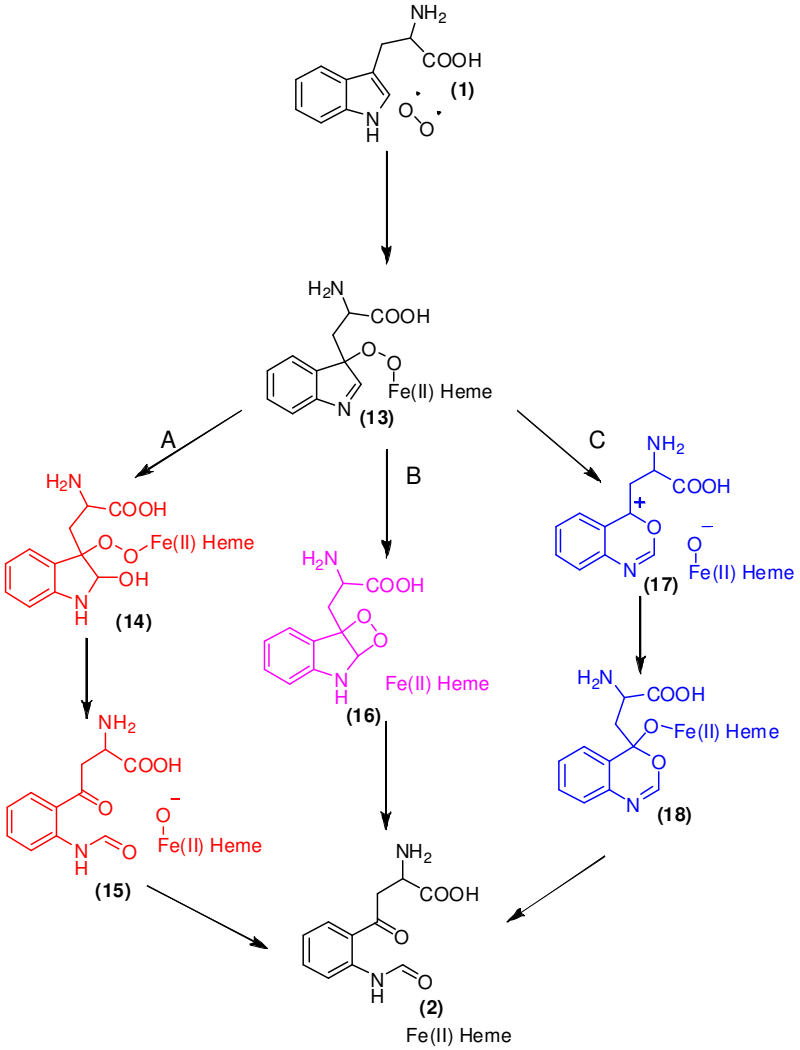


Figure 1.2: Three palusible mechanistic routes which the ternary complex (**13**) could possibly take to produce NFK.

1.2 PLP catabolic pathway

The catabolic pathway of vitamin B_6 (**19**) is probably the best studied amongst all other cofactors. In animals (including humans) 4-pyridoxic acid (**22**) is the primary catabolic product of vitamin B_6, found in urine (*12*). It is formed by oxidation of pyridoxal (**20**) by a relatively non-specific FAD-dependent aldehyde oxidase. The catalytically active form of vitamin B_6, pyridoxal-5'-phosphate, PLP (**35**), does not undergo similar oxidation to form (**22**) (*13, 14*). The various forms of vitamin B_6 (**19**, **20**, **33**) and their respective phosphate esters (**34, 35**)) are readily enzymatically interconvertible (*15*). Further degradation of (**22**) is unlikely in humans as subjects administered with doses of (**22**), were found to excrete it quantitatively (*16*). Estimation of (**22**) in urine serves as a nutritional marker, as lower than normal amounts of (**22**) is indicative of vitamin B_6 deficiency (*17*).

The catabolic pathway in microorganisms does not stop at (**22**). Selective culture techniques were used by Snell et al. to identify ten different bacterial strains and 1 yeast strain, able to grow on one or more forms of vitamin B_6 as the sole source of carbon and nitrogen (*18*). The catabolic intermediates have been isolated and characterized from the growth medium (*18-20*), and two catabolic pathways have been proposed as shown in *figure 1.3*. Both these pathways are inducible and the enzymes catalyzing these steps have also been characterized.

In pathway A *figure 1.3*, observed in *Pseudomonas* MA-1, pyridoxal (**20**) is produced from either pyridoxamine (**33**) by a transamination reaction with pyruvate catalyzed by pyridoxamine pyruvate transaminase (*21-24*), or from pyridoxine (**19**) by an oxidation reaction catalyzed by the FAD dependent pyridoxine 4-oxidase (*25*). Pyridoxal (**20**) is then oxidized to form the lactone (**21**) by the NAD dependent pyridoxal dehydrogenase (*26*) which is then hydrolyzed to form (**22**) by 4-pyridoxolactonase (*26, 27*). FAD dependent 4-pyridoxic acid dehydrogenase catalyzes the two electron oxidation of (**22**) to

(a)

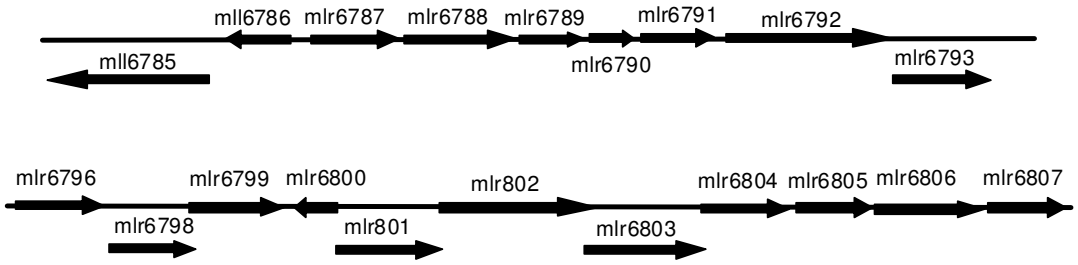
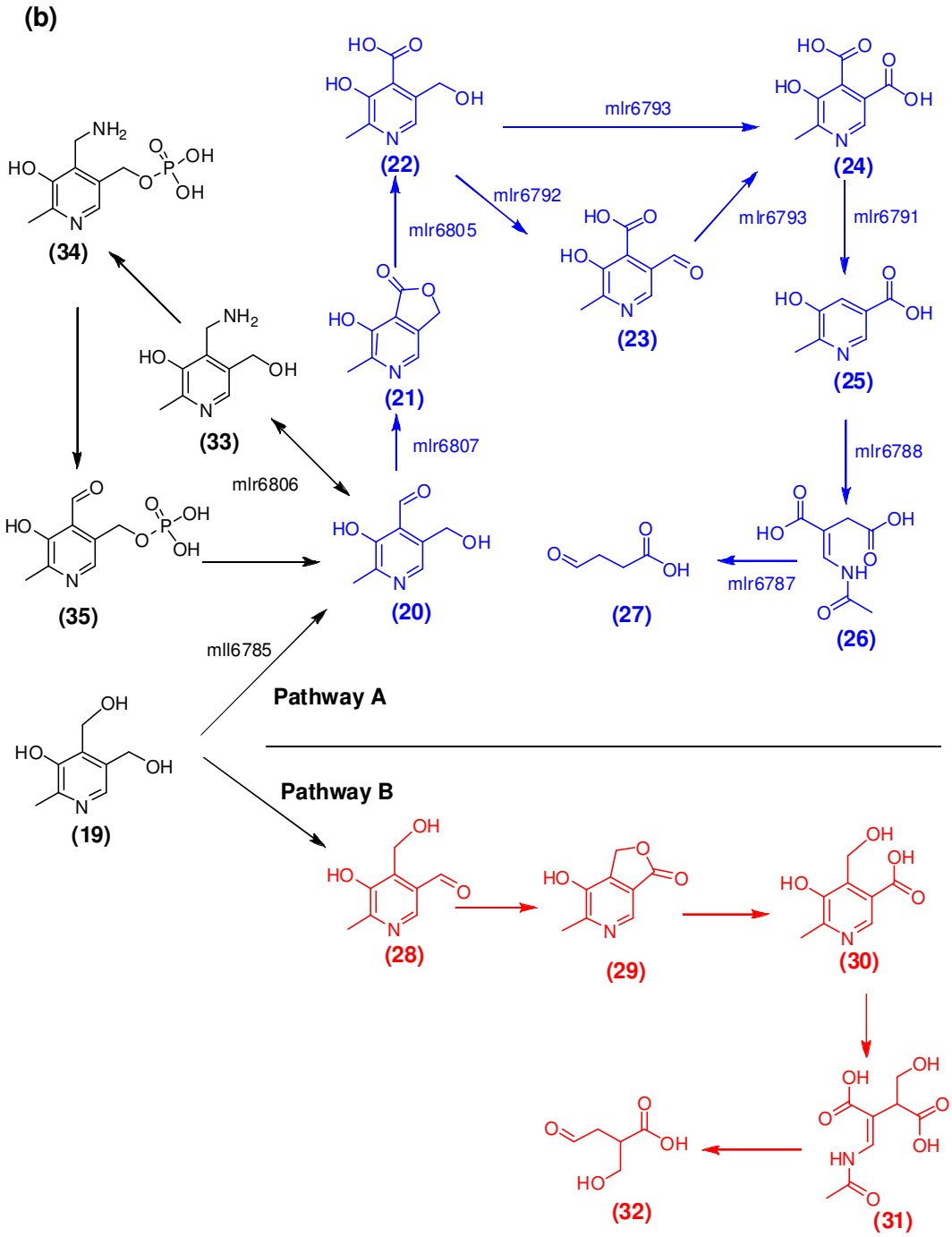


Figure 1.3: (a) The gene organization in *Mesorhizobium Loti* MAFF303099. The genes that have been identified to participate in PLP degradation are shown in catabolic pathway A. **mlr6806** - Pyridoxamine pyruvate transaminase, **mll6785** - pyridoxine 4-dehydrogenase, **mlr6807** - pyridoxal dehydrogenase, **mlr6805** - 4-pyridoxolactonase, **mlr6793** - 4-pyridoxic acid dehydrogenase, **mlr6791**- 2-methyl-3-hydroxypyridine-4,5-dicarboxylic acid decarboxylase, **mlr6788** - 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase, **mlr6787** - 2-(N-acetamidomethylene) succinic acid hydrolase.

(b) The vitamin B_6 catabolic pathways. The pathway A is observed in *Pseudomonas sp.* MA-1 while the pathway B is observed in *Pseudomonas* IA and in *Arthrobacter* Cr-7. A minor variation of pathway A is seen in *Mesorhizobium loti* MAFF303099. (**19**) pyridoxine; (**20**) pyridoxal; (**21**) 4-pyridoxolactone; (**22**) 4-pyridoxic acid; (**23**) 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylate; (**24**) 3-hydroxy-2-methylpyridine-4,5-dicarboxylate; (**25**) 3-hydroxy-2-methylpyridine-5-carboxylate; (**26**) 2-(acetamidomethylene) succinate; (**27**) succinic semialdehyde; (**28**) isopyridoxal; (**29**) 5-pyridoxolactone; (**30**) 3-hydroxy-4-hydroxymethyl-2-methylpyridine-5-carboxylate; (**31**) 2-(acetamidomethylene)-3-(hydroxymethyl) succinate; (**32**) 2-(hydroxymethyl)-4-oxobutanoate; (**33**) pyridoxamine; (**34**) pyridoxamine-5-phosphate; (**35**) pyridoxal-5-phosphate.

Figure 1.3 (continued)



form the aldehyde (**23**) (*28, 29*), which then undergoes a two electron oxidation by the NAD dependent 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylic acid dehydrogenase to form the diacid (**24**) (*28, 30*). The diacid (**24**) undergoes a decarboxylation at C4 by a Mn^{2+} dependent 2-methyl-3-hydroxypyridine-4,5-dicarboxylic acid decarboxylase to form (**25**) (*31*), which undergoes a unique ring opening oxygenation reaction catalyzed by the FAD dependent 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase, MHPCO, to produce the first acyclic compound (**26**) of the pathway, (*32-34*). (**26**) is hydrolyzed to form succinic semialdehyde (**27**) , acetate, ammonia and carbon dioxide by 2-(N-acetamidomethylene) succinic acid hydrolase (*35, 36*).

In pathway B *figure 1.3*, observed in *Pseudomonas* IA and *Arthrobacter* Cr-7, pyridoxine (**19**) is oxidized to the aldehyde at C5 (**28**) by an FAD-dependent pyridoxine-5-dehydrogenase (*25, 37*), which further undergoes oxidation by the NAD dependent isopyridoxal dehydrogenase to produce the lactone (**29**) (*30*), which then undergoes hydrolysis catalyzed by 5-pyridoxolactonase to produce 5–pyridoxic acid (**30**) (*27*). FAD dependent 5-pyridoxic acid oxygenase then catalyzes the ring opening reaction of (**30**) to form (**31**) (*38, 39*), which is hydrolyzed by 2-hydroxymethyl-(N-acetamidomethylene) succinc acid hydrolase to form 2-(hydroxymethyl)-4-oxobutanoate (**32**), acetate, ammonia and carbon dioxide (*35, 36*).

Even though the participating catabolic enzymes were known and the various intermediates were characterized, the genes involved have only recently been identified. Yagi and coworkers identified five genes in *Mesorhizobium Loti* MAFF303099 involved in pyridoxine (**19**) catabolism. These genes encoded pyridoxamine pyruvate transaminase (mlr6806) (*40*), pyridoxine-4-oxidase (mlr6785) (*41*), 4-pyridoxolactonase (mlr6805) (*42*), pyridoxal-4-dehydrogenase (mlr6807) (*43*) and the 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (mlr6788) (*44*). Efforts directed towards the identification and characterization of 4-pyridoxic acid dehydrogenase (mlr6793) (*45*), 3-hydroxy-2-

methylpyridine-4,5-dicarboxylate decarboxylase (mlr6791) (*46*) and 2-(N-acetamidomethylene) succinic acid hydrolase (mlr6787) (*47*) have been detailed out in chapter 4, 5 and 6. Chapter 7 deals with the studies directed towards elucidation of the mchanism for the MHPCO reaction.

All the genes involved in the vitamin B_6 degradation have been identified in *Mesorhizobium loti* MAFF303099. Apart from gene identification in *Mesorhizobium loti,* genes encoding for pyridoxal 4-dehydrogenase (*48*) and pyridoxine 4-oxidase (*49*), (*50*) have been identified and characterized in *Microbacterium luteolum* YK-1. The gene coding for pyridoxal 4-dehydrogenase has also been identified in *Aureobacterium luteolum (51)*. The remaining PLP catabolic genes have not yet been identified.

1.3 The rut pathway

Apart from the well characterized reductive(*52*) and oxidative pathways (*53*) of degradation of uracil (**36**) (and thymine) a new pathway was discovered in *Escherichia coli* K12 very recently (*54, 55*). The b1012 operon, one of the largest cluster of genes of unassigned function in *E.coli,* is highly expressed under the control of the transcriptional activator nitrogen regulatory protein C (NtrC) thus suggesting in its participation in catabolism of a nitrogen source. Mutants of each of the genes present in the operon did not grow when it was subjected to uracil (**36**) (or thymine) as the sole source of nitrogen at 25 °C (not at 37 °C). The parental strain could use them at 25 °C (not at 37 °C), thereby asserting the participation of these genes in the pathway. The end products of the pathway were characterized to be 3-hydroxy propionate (**37**), carbon dioxide and two

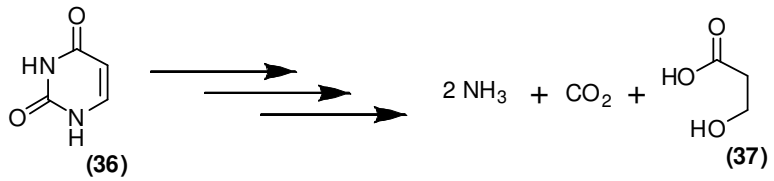


Figure 1.4: A novel pathway of uracil (**36**) degradation in *E.coli* K12.

molecules of ammonia (*54*), *figure 1.4*.

The first enzyme in the pathway is the gene product of rutA(b1012) which was annotated as a putative FMN dependent monooxygenase (*56*). This enzyme along with a flavin reductase was seen to convert uracil (or thymine) to another product which was shown to be the substrate for rutB(b1011), the next enzyme in the pathway. The intermediates of this pathway were not characterized and no information on the enzymology was known. Efforts directed towards the elucidation of the pathway intermediates and the enzymology required for their transformation is described in chapter 8. The studies towards understanding the reaction mechanism for the gene product of rutA is reported in chapter 9.

REFERENCES

(1) Curzon, G. (1986) Hopkins and the Discovery of Tryptophan. *Progress in Tryptophan Research*, Cardiff, Wales, United Kingdom: Walter de Gruyter & Co.

(2) Magni, G., Amici, A., Emanuelli, M., Raffaelli, N., and Ruggieri, S. (1999) Enzymology of NAD+ synthesis. *Adv. Enzymol. Relat. Areas Mol. Biol. 73*, 135-182.

(3) White, H. B., III. (1982) *Biosynthetic and salvage pathways of pyridine nucleotide coenzymes.* Pyridine Nucleotide Coenzymes. *Nucleotide Coenzymes*, 225-248.

(4) Brown, R. R. (1994) Tryptophan metabolism: A review. *L-Tryptophan: Curr. Prospects Med. Drug Saf.*, 17-30.

(5) Hegyi, J., Schwartz, R.A., and Hegyi, V. (2004) Pellegra: Dermatitis, dementia and diarrhea. *International Journal of Dermatology 43*, 1-5.

(6) Kresge, N., Simoni, R. D., and Hill, R. L. (2004) Blacktongue and Nicotinic Acid Metabolism: Philip Handler. *J. Biol. Chem. 279*, e8-e9.

(7) Stone, T. W., and Darlington, L. G. (2002) Endogenous kynurenines as targets for drug discovery and development. *Nat Rev Drug Discov 1*, 609-20.

(8) Tanaka, T., and Knox, E. (1959) The Nature and Mechanism of the Tryptophan Pyrrolase (Peroxidase-Oxidase) Reaction of Pseudomonas and of Rat Liver. *J. Biol. Chem. 234*, 1162-1170.

(9) Leeds, J. M., Brown, P. J., McGeehan, G. M., Brown, F. K., and Wiseman, J. S. (1993) Isotope effects and alternative substrate reactivities for tryptophan 2,3-dioxygenase. *J Biol Chem 268*, 17781-6.

(10) Terentis, A. C., Thomas, S. R., Takikawa, O., Littlejohn, T. K., Truscott, R. J., Armstrong, R. S., Yeh, S. R., and Stocker, R. (2002) The heme environment of

recombinant human indoleamine 2,3-dioxygenase. Structural properties and substrate-ligand interactions. *J Biol Chem 277*, 15788-94.

(11) Bugg, T., and Winfield, C. (1998) Enzymic cleavage of aromatic rings: mechanistic aspects of the catechol dioxygenases and later enzymes of bacterial oxidative cleavage pathways. *Natural Product Reports 15*, 513.

(12) Huff, J. W., and Perlzweig, W. A. (1944) An Oxidative Metabolite Of Pyridoxine In Human Urine. *Science 100*, 15.

(13) Schwartz, R., and Kjeldgaard, N. O. (1951) The enzymic oxidation of pyridoxal by liver aldehyde oxidase. *Biochem J 48*, 333-7.

(14) Rabinwitz, J. C., and Snell, E. E. (1949) Vitamin B6 group; urinary excretion of pyridoxal, pyridoxamine, pyridoxine, and 4-pyridoxic acid in human subjects. *Proc Soc Exp Biol Med 70*, 235-40.

(15) Snell, E. E., and Haskell, B. E. (1971) *The Metabolism of Vitamin B6, In: Comprehensive Biochemistry*, Vol. 21, Elsevier/North Holland, New York.

(16) Reddy, S. K., Reynolds, M. S., and Price, J. M. (1958) The determination of 4-pyridoxic acid in human urine. *J Biol Chem 233*, 691-6.

(17) Baysal, A., Johnson, B. A., and Linkswiler, H. (1966) Vitamin B 6 depletion in man: blood vitamin B 6, plasma pyridoxal-phosphate, serum cholesterol, serum transminases and urinary vitamin B 6 and 4-pyridoxic acid. *J Nutr 89*, 19-23.

(18) Rodwell, V. W., Volcani, B. E., Ikawa, M., and Snell, E. E. (1958) Bacterial oxidation of vitamin B6. I. Isopyridoxal and 5-pyridoxic acid. *J Biol Chem 233*, 1548-1554.

(19) Burg, R. W., Rodwell, V. W., and Snell, E. E. (1960) Bacterial oxidation of vitamin B6. II. Metabolites of pyridoxamine. *J Biol Chem 235*, 1164-1169.

(20) Ikawa, M., Rodwell, V. W., and Snell, E. E. (1958) Bacterial oxidation of vitamin B6. II. Structure of 260 compound. *J Biol Chem 233*, 1555-1559.

(21) Ayling, J. E., Dunathan, H. C., and Snell, E. E. (1968) Stereochemistry of transamination catalyzed by pyridoxamine--pyruvate transaminase. *Biochemistry 7*, 4537-42.

(22) Ayling, J. E., and Snell, E. E. (1968) Relation of structure to activity of pyridoxal analogs as substrates for pyridoxamine pyruvate transaminase. *Biochemistry 7*, 1626-36.

(23) Ayling, J. E., and Snell, E. E. (1968) Mechanism of action of pyridoxamine pyruvate transaminase. *Biochemistry FIELD Full Journal Title:Biochemistry 7*, 1616-25.

(24) Kolb, H., Cole, R. D., and Snell, E. E. (1968) Molecular weight and subunit structure of pyridoxamine pyruvate transaminase. *Biochemistry FIELD Full Journal Title:Biochemistry 7*, 2946-54.

(25) Sundaram, T. K., and Snell, E. E. (1969) The bacterial oxidation of vitamin B6. V. The enzymatic formation of pyridoxal and isopyridoxal from pyridoxine. *J Biol Chem 244*, 2577-2584.

(26) Burg, R. W., and Snell, E. E. (1969) The bacterial oxidation of vitamin B6. VI. Pyridoxal dehydrogenase and 4-pyridoxolactonase. *J Biol Chem 244*, 2585-2589.

(27) Jong, Y. J., and Snell, E. E. (1986) Enzymes of vitamin B6 degradation. Purification and properties of 4- and 5-pyridoxolactonases. *J Biol Chem 261*, 15112-15114.

(28) Burg, R. W., Rodwell, V. W., and Snell, E. (1960) Bacterial oxidation of vitamin B6. Metabolites of pyridoxamine. *J. Biol. Chem. 235*, 1164-9.

(29) Yagi, T., Kishore, G. M., and Snell, E. E. (1983) The bacterial oxidation of vitamin B6. 4-Pyridoxic acid dehydrogenase: a membrane-bound enzyme from Pseudomonas MA-1. *J Biol Chem 258*, 9419-9425.

(30) Lee, Y. C., Nelson, M. J., and Snell, E. E. (1986) Enzymes of vitamin B6 degradation. Purification and properties of isopyridoxal dehydrogenase and 5-

formyl-3-hydroxy-2-methylpyridine-4-carboxylic-acid dehydrogenase. *J Biol Chem 261*, 15106-15111.

(31) Snell, E. E., Smucker, A. A., Ringelmann, E., and Lynen, F. (1964) [Bacterial Oxidation Of Vitamin B6. Iv. Enzymatic Decarboxylation Of 2-Methyl-3-Hydroxypyridine-4,5-Dicarboxylic Acid.]. *Biochem Z 341*, 109-119.

(32) Kishore, G. M., and Snell, E. E. (1981) Kinetic investigations on a flavoprotein oxygenase, 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *J Biol Chem 256*, 4228-4233.

(33) Kishore, G. M., and Snell, E. E. (1981) Interaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase with FAD, substrates, and analogues. Spectral and fluorescence investigations. *J Biol Chem 256*, 4234-4240.

(34) Sparrow, L. G., Ho, P. P., Sundaram, T. K., Zach, D., Nyns, E. J., and Snell, E. E. (1969) The bacterial oxidation of vitamin B6. VII. Purification, properties, and mechanism of action of an oxygenase which cleaves the 3-hydroxypyridine ring. *J Biol Chem 244*, 2590-2600.

(35) Nyns, E. J., Zach, D., and Snell, E. E. (1969) The bacterial oxidation of vitamin B6. 8. Enzymatic breakdown of alpha-(N-acetylaminomethylene) succinic acid. *J Biol Chem 244*, 2601-2605.

(36) Huynh, M. S., and Snell, E. E. (1985) Enzymes of vitamin B6 degradation. Purification and properties of two N-acetylamidohydrolases. *J Biol Chem 260*, 2379-2383.

(37) Jong, Y. J., Nelson, M. J., and Snell, E. E. (1986) Enzymes of vitamin B6 degradation. Purification and properties of pyridoxine 5'-dehydrogenase (oxidase). *J Biol Chem 261*, 15102-15105.

(38) Sparrow, L. G., Ho, P. P. K., Sundaram, T. K., Zach, D., Nyns, E. J., and Snell, E. E. (1969) Bacterial oxidation of vitamin B6. VII. Purification, properties, and

mechanism of action of an oxygenase which cleaves the 3-hydroxypyridine ring. *J. Biol. Chem. FIELD Full Journal Title:Journal of Biological Chemistry 244*, 2590-600.

(39) Nelson, M. J., and Snell, E. E. (1986) Enzymes of vitamin B6 degradation. Purification and properties of 5-pyridoxic-acid oxygenase from Arthrobacter sp. *J Biol Chem 261*, 15115-15120.

(40) Yoshikane, Y., Yokochi, N., Ohnishi, K., Hayashi, H., and Yagi, T. (2006) Molecular cloning, expression and characterization of pyridoxamine-pyruvate aminotransferase. *Biochem J 396*, 499-507.

(41) Yuan, B., Yoshikane, Y., Yokochi, N., Ohnishi, K., and Yagi, T. (2004) The nitrogen-fixing symbiotic bacterium Mesorhizobium loti has and expresses the gene encoding pyridoxine 4-oxidase involved in the degradation of vitamin B6. *FEMS Microbiol Lett 234*, 225-230.

(42) Funami, J., Yoshikane, Y., Kobayashi, H., Yokochi, N., Yuan, B., Iwasaki, K., Ohnishi, K., and Yagi, T. (2005) 4-Pyridoxolactonase from a symbiotic nitrogen-fixing bacterium Mesorhizobium loti: cloning, expression, and characterization. *Biochim Biophys Acta 1753*, 234-239.

(43) Yokochi, N., Nishimura, S., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2006) Identification of a new tetrameric pyridoxal 4-dehydrogenase as the second enzyme in the degradation pathway for pyridoxine in a nitrogen-fixing symbiotic bacterium, Mesorhizobium loti. *Arch Biochem Biophys 452*, 1-8.

(44) Yuan, B., Yokochi, N., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2006) Molecular cloning, identification and characterization of 2-methyl-3-hydroxypyridine-5-carboxylic-acid-dioxygenase-coding gene from the nitrogen-fixing symbiotic bacterium Mesorhizobium loti. *J Biosci Bioeng 102*, 504-510.

(45) Mukherjee, T., Kinsland, C., and Begley, T. P. (2007) PLP catabolism: Identification of the 4-pyridoxic acid dehydrogenase gene in Mesorhizobium loti MAFF303099. *Bioorg Chem 35*, 458-464.

(46) Mukherjee, T., McCulloch, K. M., Ealick, S. E., and Begley, T. P. (2007) Gene Identification and Structural Characterization of the Pyridoxal 5'-Phosphate Degradative Protein 3-Hydroxy-2-methylpyridine-4,5-dicarboxylate Decarboxylase from Mesorhizobium loti MAFF303099. *Biochemistry 46*, 13606 - 13615.

(47) Mukherjee, T., Hilmey, D. G., and Begley, T. P. (2008) PLP catabolism: identification of the 2-(Acetamidomethylene)succinate hydrolase gene in Mesorhizobium loti MAFF303099. *Biochemistry 47*, 6233-41.

(48) Yokochi, N., Yoshikane, Y., Trongpanich, Y., Ohnishi, K., and Yagi, T. (2004) Molecular Cloning, Expression, and Properties of an Unusual Aldo-Keto Reductase Family Enzyme, Pyridoxal 4-Dehydrogenase, That Catalyzes Irreversible Oxidation of Pyridoxal. *J. Biol. Chem. 279*, 37377-37384.

(49) Yoshikane, Y., Yokochi, N., Ohnishi, K., and Yagi, T. (2004) Coenzyme precursor-assisted cooperative overexpression of an active pyridoxine 4-oxidase from Microbacterium luteolum. *Protein Expression Purif. 34*, 243-248.

(50) Kaneda, Y., Ohnishi, K., and Yagi, T. (2002) Purification, molecular cloning, and characterization of pyridoxine 4-oxidase from Microbacterium luteolum. *Biosci., Biotechnol., Biochem. 66*, 1022-1031.

(51) Trongpanich, Y., Abe, K., Kaneda, Y., Morita, T., and Yagi, T. (2002) Purification and characterization of pyridoxal 4-dehydrogenase from Aureobacterium luteolum. *Biosci., Biotechnol., Biochem. 66*, 543-548.

(52) Vogels, G. D., and Van der Drift, C. (1976) Degradation of purines and pyrimidines by microorganisms. *Bacteriol Rev 40*, 403-68.

(53) Soong, C. L., Ogawa, J., Sakuradani, E., and Shimizu, S. (2002) Barbiturase, a novel zinc-containing amidohydrolase involved in oxidative pyrimidine metabolism. *J Biol Chem 277*, 7051-8.

(54) Loh, K. D., Gyaneshwar, P., Markenscoff Papadimitriou, E., Fong, R., Kim, K. S., Parales, R., Zhou, Z., Inwood, W., and Kustu, S. (2006) A previously undescribed pathway for pyrimidine catabolism. *Proc Natl Acad Sci U S A 103*, 5114-9.

(55) Osterman, A. (2006) A hidden metabolic pathway exposed. *Proc Natl Acad Sci U S A 103*, 5637-8.

(56) http://theseed.uchicago.edu/FIG/index.cgi.

CHAPTER 2

Mechanistic studies on Tryptophan 2,3-dioxygenase

2.1 Introduction

The mechanism by which tryptophan 2,3-dioxygenase (TDO) catalyzes the oxidative ring opening of the indole ring of tryptophan (**1**) to produce N-formylkynurenine (NFK) (**2**) is still poorly understood. This reaction is significantly different form the well-studied oxidative aromatic ring opening reaction catalyzed by the non-heme dioxygenases. For the latter enzymes, the substrate and molecular oxygen are coordinated to the active site iron and this coordination is a key part of the catalytic mechanism (*1*). Such coordination is not possible for TDO because the heme has only one vacant coordination site, which is occupied by molecular oxygen during the reaction, thus preventing tryptophan (**1**) to be coordinated to the heme iron in TDO.

The catalytically active form of TDO is the ferrous state of heme (*2*). From the works of Ishimura *et al*., (*3*) it has been established, based on kinetic and spectroscopic evidences, that tryptophan (**1**) first binds with the iron (II) heme containing enzyme inducing a conformational change which activates the heme to react with oxygen, thus forming an intermediary heme-tryptophan-oxygen ternary complex.

As outlined in the *figure 2.1* the heme iron catalyzes the formation of the tryptophan hydroperoxide (**3**) most likely by a single electron transfer mechanism (*4*). From this intermediate, three mechanistic routes to NFK (**2**) can be postulated. According to mechanism A (red), water can add to the endocylic imine to form (**4**) which can rearrange to give NFK (2). In mechanism B (in pink) rearrangement of the hydroperoxide (**3**) to a dioxetane intermediate (**6**) followed by a retro [2+2] cycloaddition can form NFK (**2**). Although there is no precedent of dioxetane formation in dioxygenase-catalyzed ring opening reactions of aromatic compounds, this is the mechanism proposed in the literature (*5, 6*). In mechanism C (blue) (analogous to the

mechanism of extradiol dioxygenases (*7*)), a Criegee rearrangement would give the cationic intermediate (**7**) which could then undergo a nucleophilic attack by the iron-bound oxygen anion to give (**8**) which could collapse to form NFK (**2**).

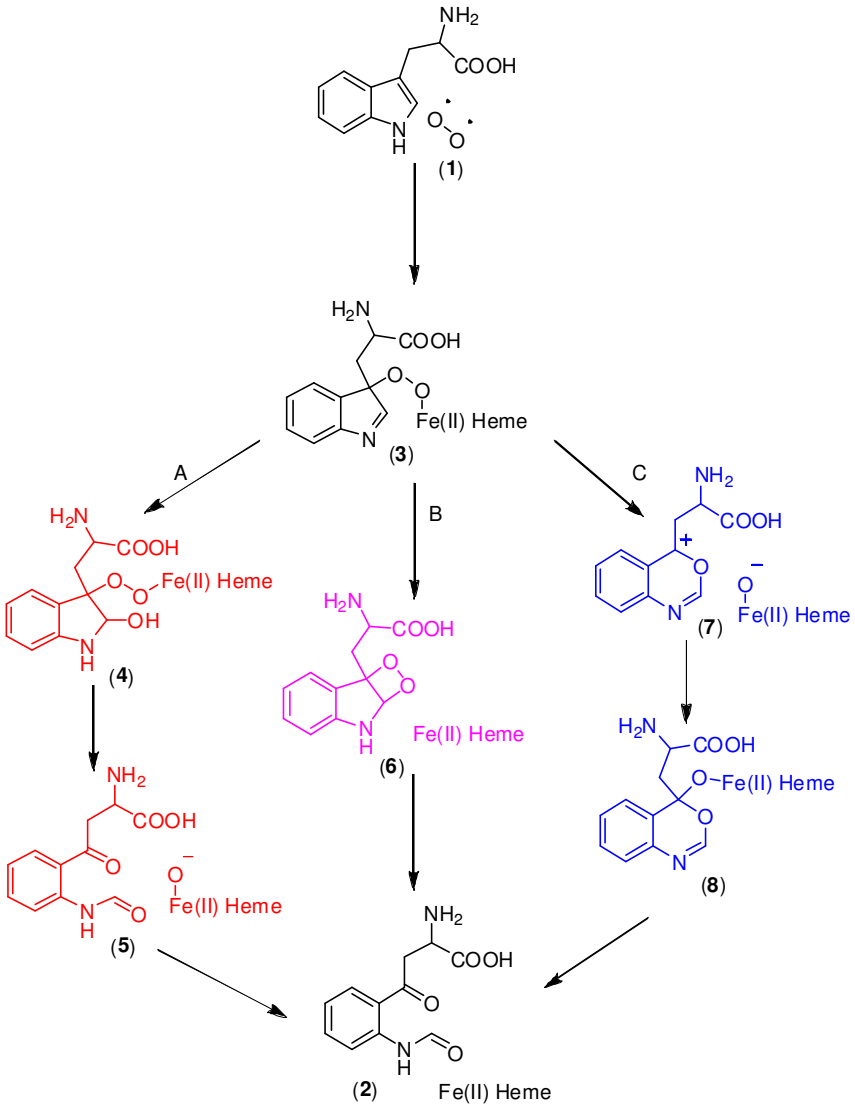


Figure 2.1: Three plausible mechanistic routes which the ternary complex (**3**) could possibly take to produce NFK (**2**).

It should be possible to differentiate between three mechanisms by running the reaction in presence of labeled oxygen. Mechanism A predicts that the oxygen of the formyl group of NFK (**2**) is derived from solvent and the ketone oxygen is derived from molecular oxygen. Mechanism B predicts that the formyl and the ketone oxygen atoms are both derived from molecular oxygen. Mechanism C predicts that the formyl oxygen is derived from molecular oxygen and that the ketone oxygen may be derived from solvent if the lifetime of intermediate (**7**) is sufficient for iron-bound oxygen to exchange with water. Such exchange has been observed for catechol dioxygenase when the corresponding intermediate was stabilized by an extra oxygen substituent on the catechol ring (*8*).

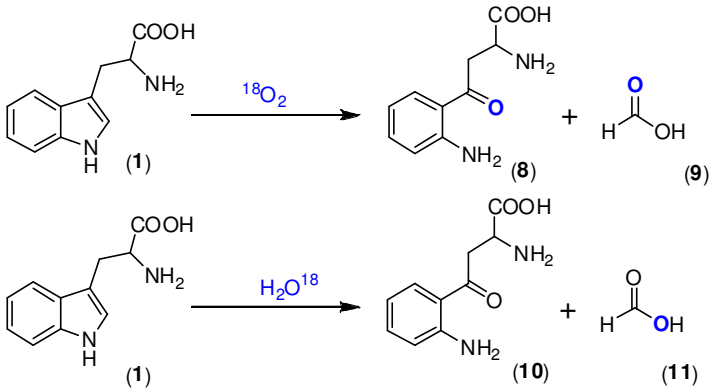


Figure 2.2: Incorporation of labeled oxygen in kynurenine (**8**, **10**) and formic acid (**9**, **11**) as observed by Hayaishi et al (*9*) in 1957, when the TDO catalyzed oxidation of tryptophan (**1**), was performed in presence of labeled water and oxygen.

Hayashi et al (*9*) reported the results of an oxygen incorporation study in 1957. The experiments, though providing solid data against mechanism A, if repeated today would be prohibitively expensive, specifically with the usage of 75 ml of H_2O^{18} reaction mixtures. Current techniques would significantly speed up the determination of the ^{18}O-enrichment of the product, which was originally accomplished by its crystallization and

subsequent pyrolysis. Moreover ^{18}O incorporation into NFK (**2**), the product of the reaction catalyzed by TDO, was never directly demonstrated. His experiments showed the incorporation of oxygen in kynurenine (**8**, **10**) and formic acid (**9**, **11**), the hydrolyzed products of NFK (**2**), *figure 2.2*.

When the reaction was performed in presence of $^{18}O_2$, 84% enrichment of oxygen in formate (**9**) and 82% enrichment in kynurenine (**8**) was observed, while in presence of $^{18}H_2O$, 94% enrichment was observed for formate (**11**) and 22% for kynurenine (**10**) (*9*). Though the data were mutually consistent with oxygen gas being the sole source of both the oxygen atoms in NFK (**2**), we felt that this important experiment must be repeated using better analytical tools with purified highly active and stable enzyme. If the reaction went via the Criegee intermediate (mechanism C) then the percentage of product, if any, with oxygen incorporation from labeled water would be in trace amounts. Detection of such trace amounts of labeled product would not be possible under the experimental conditions employed by Hayaishi et al (*9*). We looked into the solvent incorporation into NFK (**2**) by ESI-MS analysis. The results were confirmed by re-running the reaction using $^{18}O_2$. Controls were run in order to determine the extent to which the carbonyl oxygen of NFK (**2**) would undergo exchange with buffer.

Although there is no precedent of dioxetane formation in dioxygenase-catalyzed ring opening reactions of aromatic compounds, this is the mechanism proposed in the literature for TDO (*5*). The formation of dioxetane (**14**) is seen in case of firefly luciferase enzyme. Hydroperoxide intermediate (**13**) possess a high degree of nucleophilicity. The intermediate hydroperoxide (**13**) is formed next to the carbonyl group in case of adenylated D-luciferin (*figure 2.3*), which then undergoes a nucleophilic attack to form the dioxetane intermediate (**14**) (*1*). Decomposition of the dioxetane intermediate (**14**) is followed by light emission.

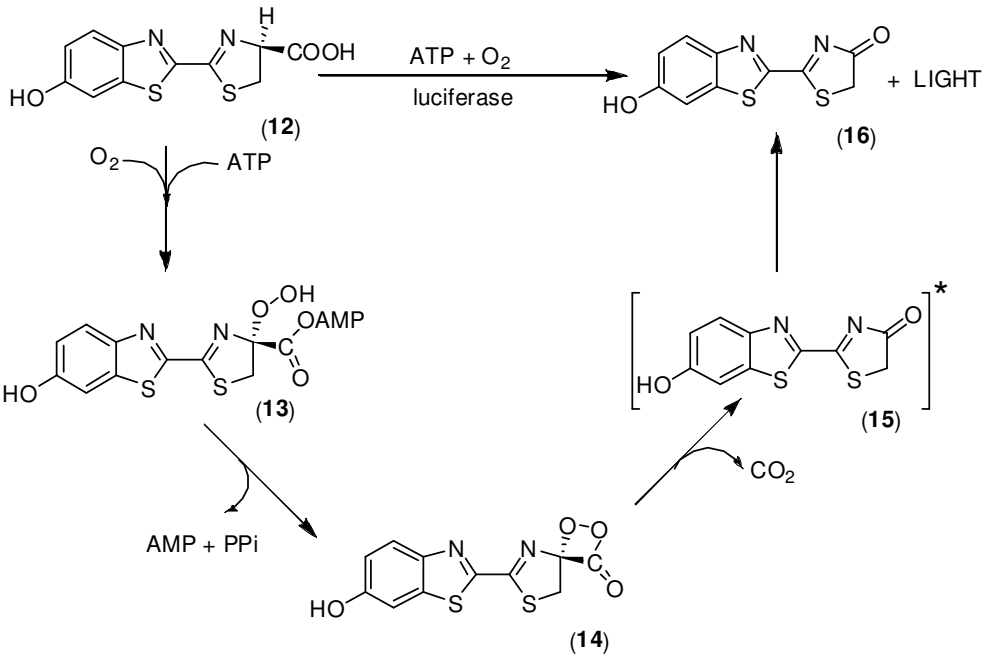


Figure 2.3: Mechanism of oxidation of D-luciferin by luciferase

In the enzymatic reaction of tryptophan 2,3-dioxygenase (TDO), no light emission has been detected, which would be expected to accompany the decomposition of dioxetane provided that the dioxetane (**6**) formed underwent a retro [2+2] cycloaddition to produce NFK (**2**) in the excited state, which would releases energy in the form of light on its decomposition to product. Efforts were directed towards detection of light emission during the course of the reaction, if any, by stopped flow experiments.

2.2 Experimental

2.2.1. Protein Overexpression, Purification and Activity Assays

2.2.1.1. **Ralstonia metallidurans** ***Tryptophan 2,3-dioxygenase:*** The protein was overexpressed and purified by the method described previously (*10*) with some modifications. While culturing the cells prior to induction 5-aminolevulinic acid hydrochloride (final concentration of 20 mg/L), which is a heme precursor, was added

when the culture reached an A_{590} of 0.3. IPTG was added to a final concentration of 0.8 mM when A_{590} was 0.6 and the culture was incubated for an additional 4-5 hours at 37°C. The cells were then harvested by centrifugation at 13,000 xg for 8 min at 4°C. The cell pellets were then frozen in liquid nitrogen and stored at -80°C. The purification procedure of TDO was same as previously described (*10*). The protein was buffer exchanged into 1 mM ammonium acetate buffer at pH 7.0. DTT, the reducing agent which is usually added to the desalting buffer was not added. The purified protein (appears as a single band on SDS-PAGE, *figure 2.4(a)*) was immediately lyophilized. Care was taken to degas the buffer and then purge it with argon to remove maximum possible dissolved air. The choice of ammonium acetate buffer was made as it is a volatile buffer and ideally suited ESI-MS.

The activity of the green colored protein (absorbance at 405nm, *figure 2.4(b)*) was determined by detection of the formation of NFK (**2**) [λ_{max} = 321nm] from L-tryptophan (**1**) by a UV-Visible spectrophotometer. The reaction mixture (500 μL) consisted of 800 μM tryptophan (**1**), 200 μM ascorbic acid and 6 μM RmTDO in 100 mM ammonium acetate buffer at pH 7.0. UV-visible spectra in the region 200-800 nm were taken at one minute time intervals. The data is shown in *figure 2.4(c)*.

The protein was lyophilized in 1 ml aliquots and the green lyophilized protein was also found to be active, as shown in *figure 2.4(d)*. The lyophilized protein was dissolved in appropriate volume of 100 mM ammonium acetate buffer at pH 7.0 such that the final concentration of the protein, determined by Bradford assay, was 150 μM. Exactly the same reaction condition as mentioned above was used for the activity assay for the lyophilized protein. The activity of the enzyme was assayed by monitoring the production of NFK (2) (characteristic λ_{max} at 321 nm) over time.

After allowing the reaction to go to completion, the enzyme was removed using Amicon Ultra-4 MWCO 5000 (Millipore) centrifugal membrane filter. A white powder

was obtained on lyophilization of the filtrate. It was characterized by ^{1}H NMR and ESI-MS and was identified as NFK (2). (^{1}H NMR ,300 MHz, CD_3OD, ppm): 8.56 (d, 1H), 8.44 (s, 1H), 8.06 (d, 1H), 7.63 (t, 1H), 7.28 (t, 1H), 4.05 (m, 1H), 3.82 (m, 1H) and 3.63 (m, 1H). ESI-MS (positive ion mode), m/z = 237

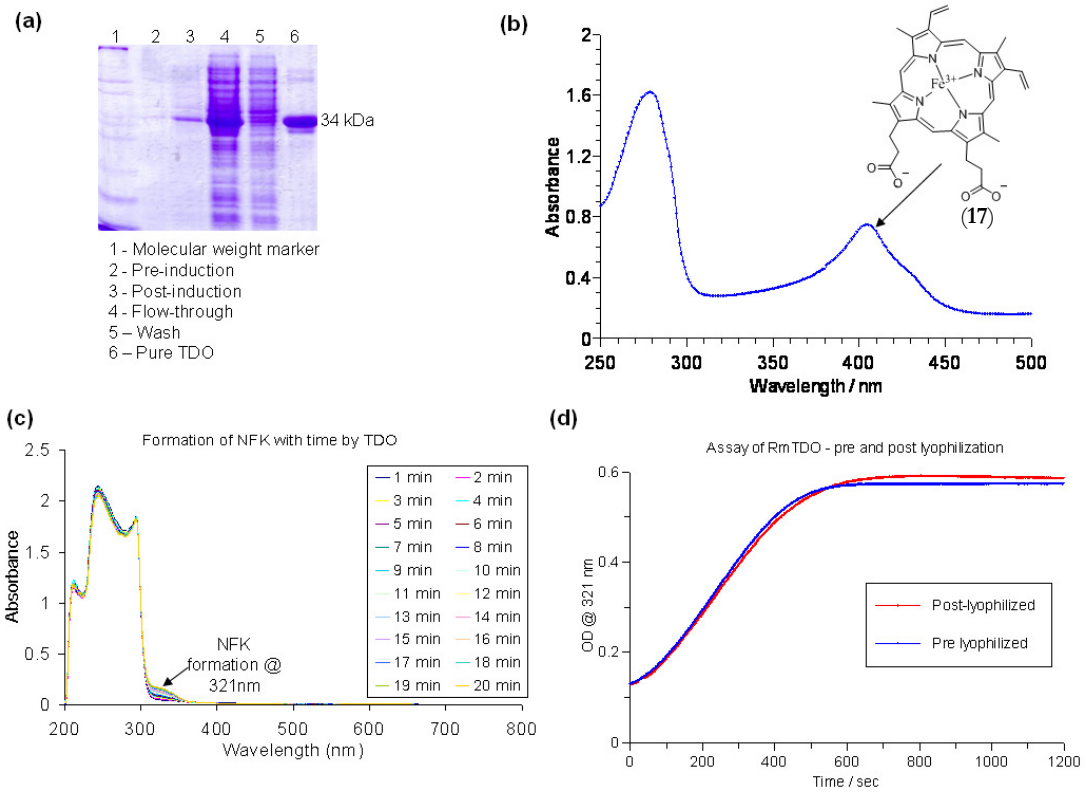


Figure 2.4: (a) SDS-PAGE analysis showing various stages of purification of recombinant RmTDO. (b) The UV-Visible spectrum of purified RmTDO showing the Soret at 405 nm, which corresponds to the Fe(III)-bound heme (**17**). (c) UV-visible spectra, showing the formation of NFK (2) over time from tryptophan (**1**) by RmTDO. NFK (**2**) has a characteristic λ_{max} at 321 nm. (d) The activity of RmTDO, before and after lyophilization, as observed by monitoring the absorbance at 321 nm, which corresponds to the formation of NFK (**2**).

2.2.1.2. **Ralstonia metallidurans *Kynurenine formamidase:*** The protein was overexpressed and purified using the protocol described before (*10*). The protein was found to be 95% pure by SDS-PAGE analysis. A coupled assay was performed to test the activity of the pure protein(*figure 2.5 (a)*). Instead of adding NFK (**2**) as a substrate to kynurenine formamidase (KFA), tryptophan (**1**) was used as the substrate. The reaction mixture (500 μL) consisted of 30 mM tryptophan (**1**), 20 μM of RmTDO, 20 μM of $ZnCl_2$ and 85 μM RmKFA in 100 mM Tris buffer at pH 8.0. The solution was vortexed and 20 μl reaction mixture was aliquoted from time to time and added to 480 μl of Tris buffer and an UV-vis scan was performed on the sample (*figure 2.5(b)*). Accumulation of NFK (**2**) was not observed under the experimental conditions, due to the much greater rate of reaction of KFA compared to that of TDO (*10*). Kynurenine (**10**) has a λ_{max} at 365 nm. On plotting the absorbance at 356 nm with time we get the rate of formation of kynurenine (**10**) from TDO by the coupled assay (*figure 2.5(c)*).

A sample of 1 ml of this protein was lyophilized. The activity of the lyophilized protein, white powder, was determined by detection of the disappearance NFK (**2**) (λ_{max}= 321 nm) or formation of kynurenine (**10**) (λ_{max} = 365 nm) by a UV-Visible spectrophotometer. The disappearance of the NFK (**2**) peak (*figure 2.5 (d)*) on addition of KFA showed that the protein was active.

After allowing the reaction to go to completion, the enzyme was removed using Amicon Ultra-4 MWCO 5000 (Millipore) centrifugal membrane filter. A white powder was obtained on lyophilization of the filtrate. It was characterized by ^{1}H NMR and ESI-MS and was identified as kynurenine. (^{1}H NMR, 300 MHz, CD_3OD, ppm): 7.75 (d, 1H), 7.26 (t, 1H), 6.75 (d, 1H), 6.6 (t, 1H), 4.00 (m, 1H), 3.72 (m, 1H) and 3.51(m, 1H). ESI-MS (positive ion mode) 209 Da.

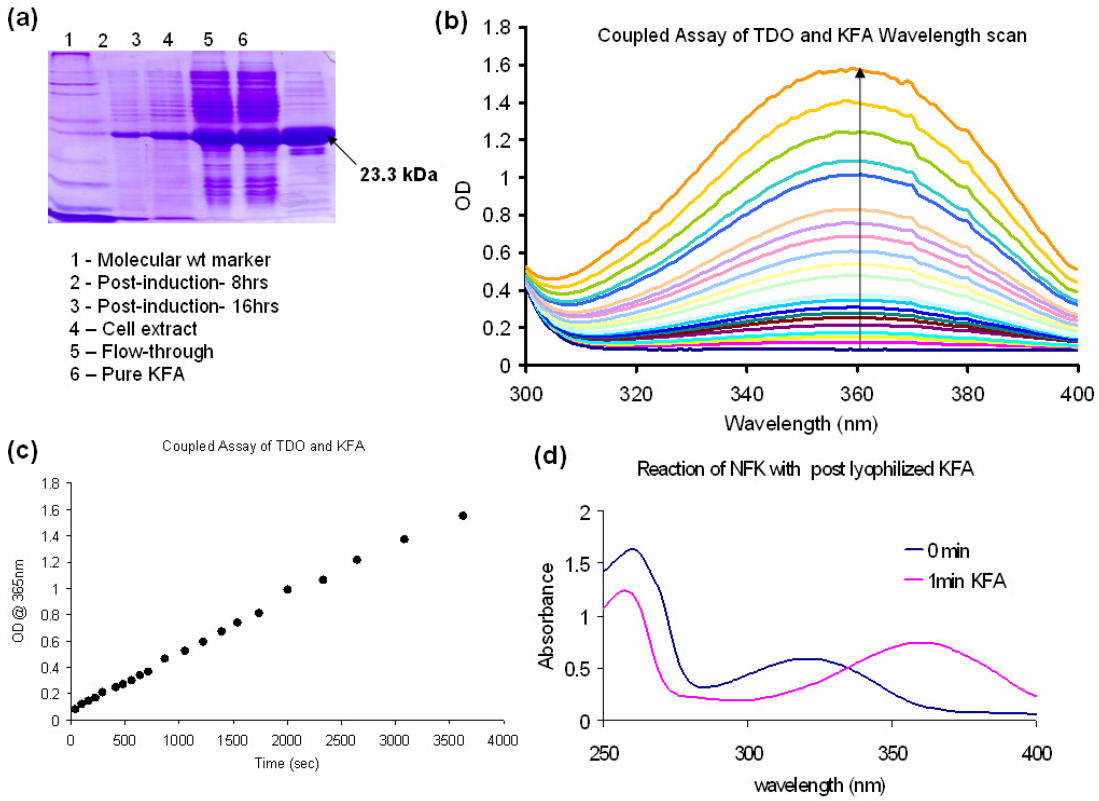


Figure 2.5: (a) SDS-PAGE analysis showing various stages of purification of recombinant RmKFA. (b) Coupled assay of RmTDO and RmKFA. The UV-Visible spectra showing the formation of kynurenine (**10**) (λ_{max} = 365 nm) from tryptophan (**1**) over time. (c) Absorbance at 365 nm from (b) is plotted over time, showing the formation of kynurenine (**10**) from tryptophan (**1**) by the coupled reaction of RmTDO and RmKFA. (d) The activity of lyophilized KFA. The spectra above show the disappearance of the peak at 321 nm (corresponding to λ_{max} NFK (**2**)) and appearance of peak at 365 nm (corresponding to λ_{max} of kynurenine (**10**)) at pH 8.0 and wavelength scan at 1200 nm/sec was done between 400 nm-200 nm.

2.2.1.3. **Homo sapiens** ***Indoleamine 2,3-dioxygenase (IDO):*** Wild type human IDO cDNA was obtained from Prof. Hiroshi Sugimoto, Japan. This plasmid contained a gene for ampicillin resistance and also coded for a 6x-His tag with a thrombin cleavage site. The plasmid was transformed into BL-21(DE-3) *E. coli* cells. While culturing the cells prior to induction 5-aminolevulinic acid hydrochloride (final concentration of 20 mg/L), which is a heme precursor, was added when the culture reached an A_{590} of 0.3. IPTG was added to a final concentration of 0.5 mM when A_{590} was 0.6 and the culture was incubated at 37°C for 12 hours (*11*). The cells were then harvested by centrifugation at 13000 xg for 10 min at 4°C. The cell pellets were frozen in liquid nitrogen and stored at -80°C. The purification procedure described for TDO was used for IDO, with one modification. Instead of washing with 100 mL wash buffer 200 ml of it was used to get a more pure protein preparation. SDS-PAGE of HuIDO showed the presence of 95% pure protein, *figure 2.6(a)*. The protein purified out with the iron bound to the heme in its active +II oxidation state (**18**) (*12*) as evident from the UV-visible spectrum is shown in *figure 2.6(c)*.

The assay conditions for IDO were different from those of TDO. The assay mixture (500 µl total volume) consisted of 400 µM tryptophan (**1**), 20 mM ascorbic acid, 0.2 mg/ml catalase, 10 µM methylene blue and10.5 µM HuIDO. All solutions were made in 50 mM phosphate buffer at pH 6.5. The pH of the ascorbic acid stock solution (200 mM) was found to be 3.0 and was subsequently adjusted to 7.0 with 2M NaOH. Care was taken not to exceed pH 7.0. The stock solution of catalase (4 mg/ml, 2800 units/ mg of solid) from *Bovine* liver was made fresh in 50 mM phosphate buffer at pH 6.5. The activity was measured by monitoring the absorbance at 321 nm, characteristic for the production of NFK (**2**). Both the freshly purified protein as well as a lyophilized sample showed the same activity, *figure 2.6 (b)*.

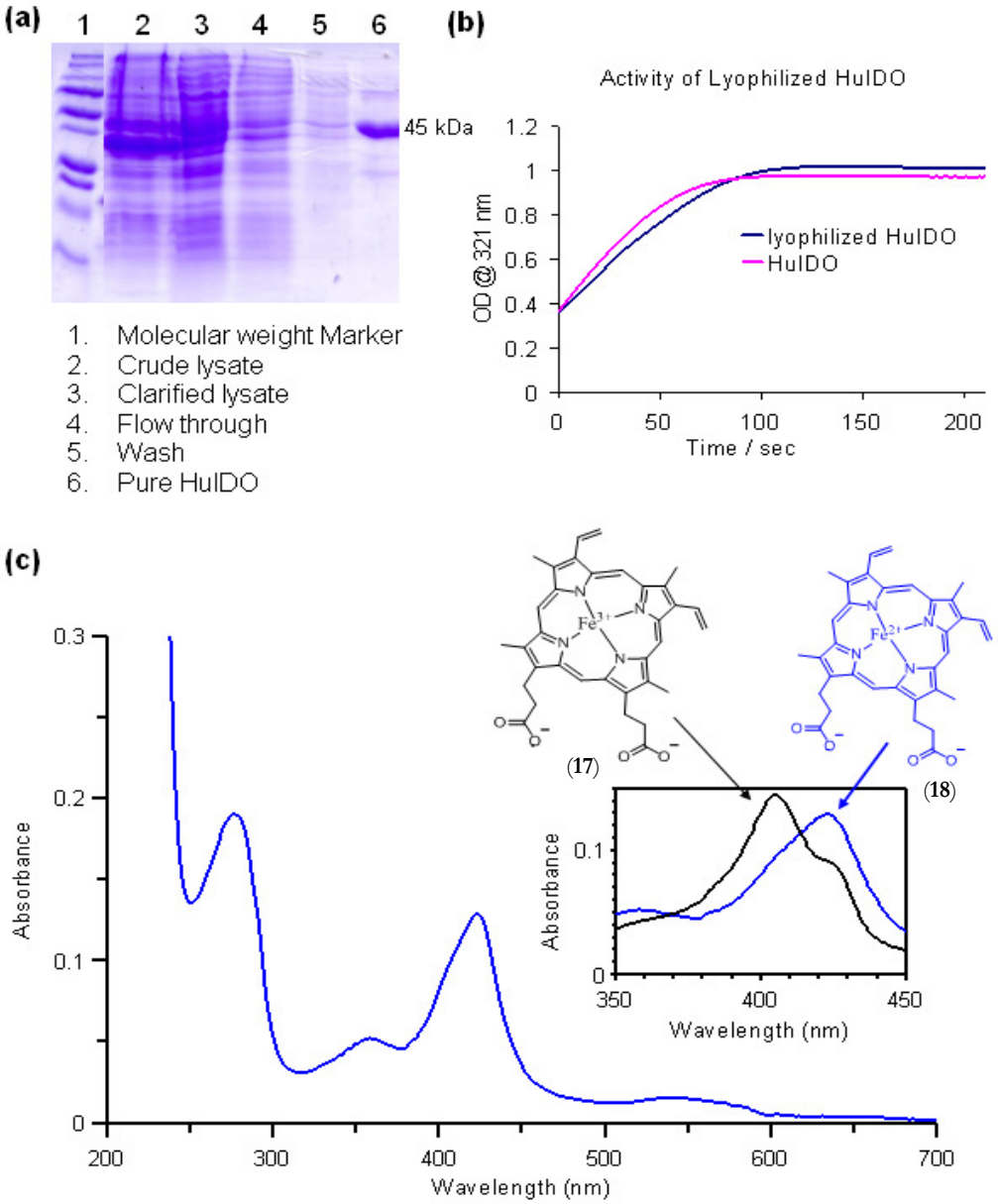


Figure 2.6: (a) SDS-PAGE analysis showing various stages of purification of recombinant HuIDO. (b) The activity of HuIDO, before and after lyophilization, as observed by monitoring the absorbance at 321 nm, which corresponds to the formation of NFK (**2**). (c) UV-Visible spectrum of HuIDO. It purified out as the active Fe(II) form (**18**), as evidenced from the Soret at 425 nm. On keeping it exposed to oxygen it slowly converts to the oxidized Fe(III) state (**17**) (inset).

2.2.2. Synthesis

6-hydroxytryptophan (**23**) was prepared following literature procedure (*13*) with some modifications *figure 2.7*.

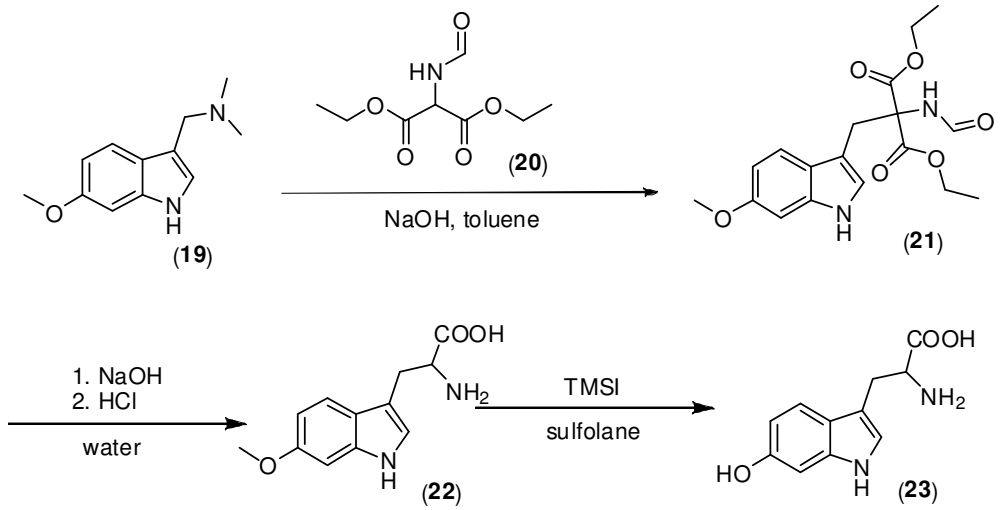


Figure 2.7: The synthetic scheme for the preparation of 6-hydroxytryptophan (**23**) from 6-methoxygramine (**19**).

2.2.2.1 Ethyl-β-(6-methoxyindolyl-3)-α-carbethoxy-α-formamidopropionate (21). A mixture of 1.09 g (5.3 mmol) 6-methoxygramine (**19**), 1.36 g (6.7 mmol) diethylformamidomalonate (**20**) and 120 mg (2.9 mmol) powdered NaOH and 70 ml toluene was heated under reflux for 22 h, while a vigorous stream of argon was bubbled through. After running the reaction for 22 hours the hot suspension was filtered and extracted with water. On shaking with water three layers were seen, the top being the toluene layer, the bottom being the water and the middle which looked like an emulsion was white in color. The aqueous layer was separated and then the organic layer was extracted with 1 M HCl, 2 M NaOH and water respectively. The middle layer which looked like an emulsion remained after each extraction. The toluene layer was filtered in order to remove the air bubbles in the emulsion as they would crack at the surface of the

filter paper. On filtering a white residue was obtained and clear toluene was the filtrate. On evaporating the toluene yellow grease was left behind, which, in accordance with the literature was not the product. TLC of the white solid that had crashed out during extraction was different from that of the substrate spot. NMR spectra of this compound revealed it to be the desired product. ^{1}H NMR (300 MHz, CD_3OD) δ 8.05 (s, 1H), 7.27 (d, *J* = 8.7, 1H), 6.87 (s, 1H), 6.84 (d, *J* = 2.2, 1H), 6.66 (dd, *J* = 8.7, 2.3, 1H), 4.24 – 4.13 (m, 4H), 3.79 (s, 4H), 3.31 (dt, *J* = 3.2, 1.6, 3H), 1.24 (d, *J* = 7.1, 6H).

2.2.2.2 6-methoxytryptophan (22). A mixture of 663 mg (17 mmol) NaOH, 500 mg (1.4 mmol) of the methoxyformamidoester (**21**) and 15 ml water was refluxed for 3 h. After neutralization with 2 M HCl, the aminomalonic acid precipitated. After further addition of 700 µl (1.4 mmol) 2 M HCl refluxing was continued for 30 min and the solid went completely into solution with the evolution of CO_2. The hot solution was filtered and the pH of the yellow filtrate was adjusted to 6.8 (isoelectric point of tryptophan (1)) with 1 M ammonia. On evaporating water the amino acid precipitated. The off white solid was collected, washed with cold ethanol and dried under high vacuum overnight. The entire dried sample was not soluble either in water or methanol. The NMR solution was centrifuged before loading into the tube. NMR spectra in D_2O (both ^{1}H as well as ^{13}C NMR) showed that all the peaks corresponded to 6-methoxytryptophan (**22**) as reported in literature. ^{1}H NMR (300 MHz, D_2O) δ 7.60 (d, *J* = 8.7, 1H), 7.20 (s, 1H), 7.10 (d, *J* = 2.2, 1H), 6.87 (dd, *J* = 8.7, 2.3, 1H), 4.03 (dd, *J* = 8.0, 4.8, 1H), 3.88 (s, 3H), 3.44 (dd, *J* = 15.1, 4.6, 1H), 3.26 (dd, *J* = 15.4, 8.1, 1H).

2.2.2.3. 6-hydroxytryptophan (23). The reaction was carried out in a 5 ml screw cap glass vial, which was taped with a black tape on the outside, under argon. 100 mg (427 µmol) of 6-methoxytryptophan (**22**) were suspended in 2 ml sulfolane. Then 72 µl (512

μmol) triethylamine and 700 μl (5.2 mmol) trimethylsilyliodide were added and the red solution was stirred and heated for 2 h at 50°C. Determining the extent of reaction by TLC was difficult as sulfolane caused the silica plate to streak. The intermediate trimethylsilylethers formed during the reaction were cleaved and excess trimethylsilyliodide was decomposed by addition of 2 ml methanol. After quenching the reaction with methanol the pH of the organic solution could not be determined by the pH meter, since the volume was very low. pH paper could not be used as the solvent was organic. So 2 drops of 1M ammonium hydroxide was added to the solution and loaded onto the column (10×50 mm), filled with cation-exchange resin Dowex-50W (H^+ form). The column was washed first with 6 ml methanol-water (1:1, v/v), then with 10 ml water and eluted with 0.5 M ammonia at a flow-rate of 0.5 ml/min. 2ml fractions were collected using a fraction collector and all the fractions were tested for UV activity. Those fractions which were UV active were spotted on a TLC plate along with 6-methoxytryptophan (**22**) as a control. All of them showed a new spot and migrated with the same R_f. On collecting the UV active fractions and putting it on high vacuum overnight, a grey colored solid (30 mg) was left behind. NMR of this gray solid showed no H^1 NMR peak corresponding to the methoxy. This preparation resulted in 90-95% pure product. Following the literature procedure (*13*) this sample was passed through Dowex-1 anion exchange column to further purify the 6-hydroxytryptophan (**23**). The resins were washed with 1 M acetic acid (20 times the column volume) before loading on to the column (10×50 mm). The compound was eluted with 1 M acetic acid at a flow-rate of 0.5 ml/min and 2 ml fractions were collected using a fraction collector. The fractions which were UV active were accumulated and were subjected to a high vacuum overnight. A grey film was left behind [~15 mg]. NMR of the sample after anion exchange revealed it to be 6-hydroxytryprophan (**23**). NMR (300 MHz, D_2O) δ 7.50 (d, J = 8.7, 1H), 7.14 (s, 1H), 6.93 (d, J = 2.0, 1H), 6.74 (dd, J = 8.6, 2.2, 1H), 4.23 (dd, J = 7.4, 5.2, 1H), 3.42 (dd, J =

15.3, 5.2, 1H), 3.31 (dd, *J* = 15.4, 7.4, 1H). Both 6-methoxytryptophan (**22**) as well as 6-hydroxytryptophan (**23**) was a racemic mixture of D and L isomers.

2.3. Results

2.3.1. Reaction of RmTDO with tryptophan (1) in H_2O^{18}. A solution of L-tryptophan (**1**) (2 mg/ml) in 1 mM ammonium acetate buffer at pH 7.0 was prepared and 1 ml of the solution was lyophilized. The lyophilized sample was dissolved in 1 g of H_2O^{18}. 10 mg of lyophilized TDO was added to the resulting solution and the reaction mixture was briefly vortexed. All of enzyme did not dissolve. 150 µl of the reaction mixture was removed at 20 minutes and filtered through a Microcon YM-10(MWCO 10 kDa) cellulose filter by centrifugation for 20 minutes. The filtrate was then mixed with equal volume of anhydrous methyl alcohol (such that the solvent system for ESI-MS was 1:1 methanol: water) and frozen immediately by immersion in an acetone-dry ice mixture. ESI-MS of the samples showed the presence of a peak at m/z = 237 which corresponded with the calculated mass of NFK (**2**) (*figure 2.8*).

MS^MS of this peak indicated the presence of de-aminated compound (**27**) at m/z = 220 and another fragmentation peak (**26**) at m/z = 202 nm, probably produced by rearrangement with loss of water (*figure 2.9*). When the reaction was carried out in H_2O^{18} only the peak (m/z = 237) corresponding to NFK (**2**) with no ^{18}O was observed, suggesting that no oxygen incorporation in NFK (**2**) from H_2O^{18} water occurred, as no peak at m/z = 239 (single ^{18}O incorporation (**25**)) was seen.

To eliminate the possibility of oxygen exchange from the solvent, pure NFK (**2**) (synthesized enzymatically and purified by HPLC) was incubated for 100 minutes (more than twice the reaction time) in H_2O^{18}. No $^{16}O/^{18}O$ exchange was observed (*figure 2.10*).

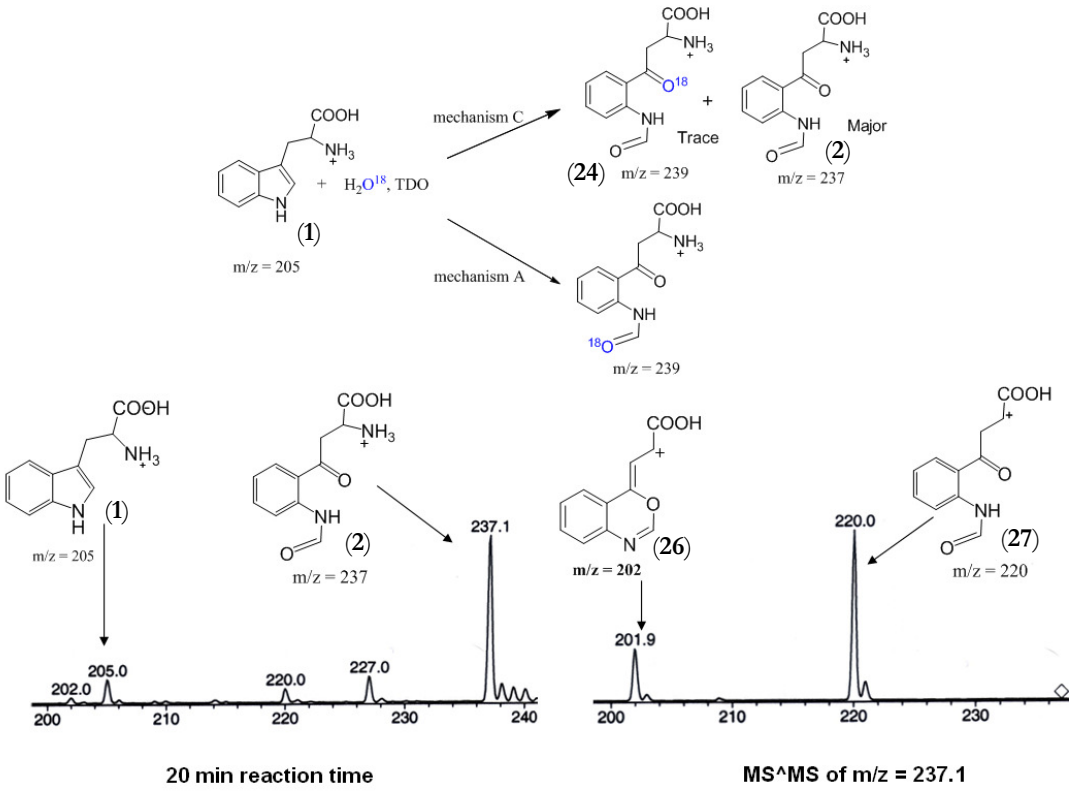


Figure 2.8: The scheme on the top shows the expected incorporation pattern of ^{18}O from labeled water via mechanism A and C. Bottom left: Positive mode ESI-MS analysis of the reaction mixture after 20 minutes. Bottom right: Fragmentation pattern of species at m/z = 237 is shown in the spectrum on the right.

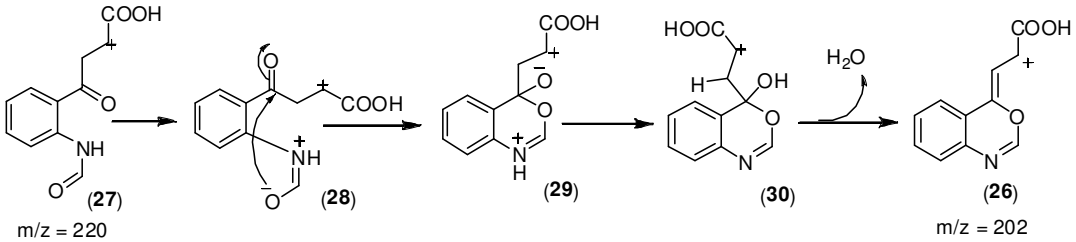


Figure 2.9: A plausible mechanism for the fragmentation of the species with m/z = 220

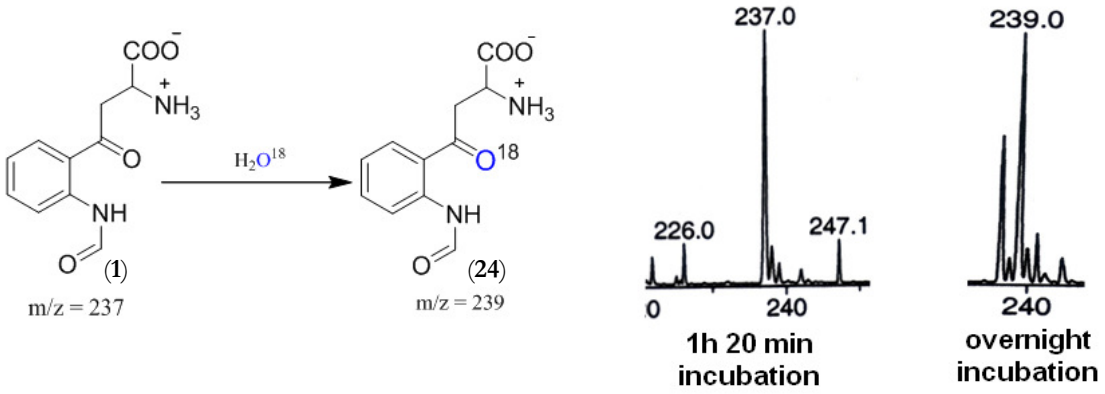


Figure 2.10: Exchange of the ketonic oxygen from NFK (**2**) over time. The mass spectra on the right show that no exchange occurs after 80 minutes of incubation of NFK (**2**) in H_2O^{18} water. However, exchange is seen after overnight incubation (MS analysis on the right).

Only one oxygen exchange was observed on overnight incubation of pure NFK (**2**) in ^{18}O water as indicated by the presence of a peak at m/z = 239 (**24**). This is probably the ketonic oxygen as it is more labile than the amide oxygen. Further experiments were performed to confirm this and the results are discussed later on. Thus oxygen atoms in NFK (**2**) do not come from water during the course of the enzymatic reaction.

2.3.2. Reaction of RmTDO with tryptophan (1) in $^{18}O_2$ gas. 1.5 ml of 1 mM ammonium acetate buffer at pH 7.0 was degassed by alternately stirring it under vacuum and purging it with argon for 2 hours. $^{18}O_2$ gas was then passed through the buffer for 20 minutes. The resulting $^{18}O_2$ gas containing buffer was transferred to a sealed 10 ml pear shaped flask, pre-purged with argon, containing 6 mg of tryptophan (**1**) and 7.2 mg of lyophilized TDO. The reaction mixture was vortexed briefly to dissolve the tryptophan (**1**) and stirred at room temperature. After 20 minutes, 100 µl aliquot was taken and

frozen in acetone-dry ice mixture. Before taking ESI-MS of the sample it was diluted 4 times (as the enzyme was not filtered off, to save time) and mixed with equal volume of anhydrous methyl alcohol. ESI-MS of the sample (*figure 2.11*) showed the presence of a peak at m/z = 241 (**31**) along with a peak at m/z = 237 (**2**). MS^MS of the species with peak at m/z = 241 (**31**) produced fragments at m/z = 194 (**32**), 204 (**33**) and 224 (**34**). The species at m/z = 224 (**34**) fragment probably corresponds to the de-amination of the molecule with the peak of m/z = 241 (**31**), the species with m/z peak of 204 (**33**) could be due to a rearrangement with loss of water (*figure 2.10*).

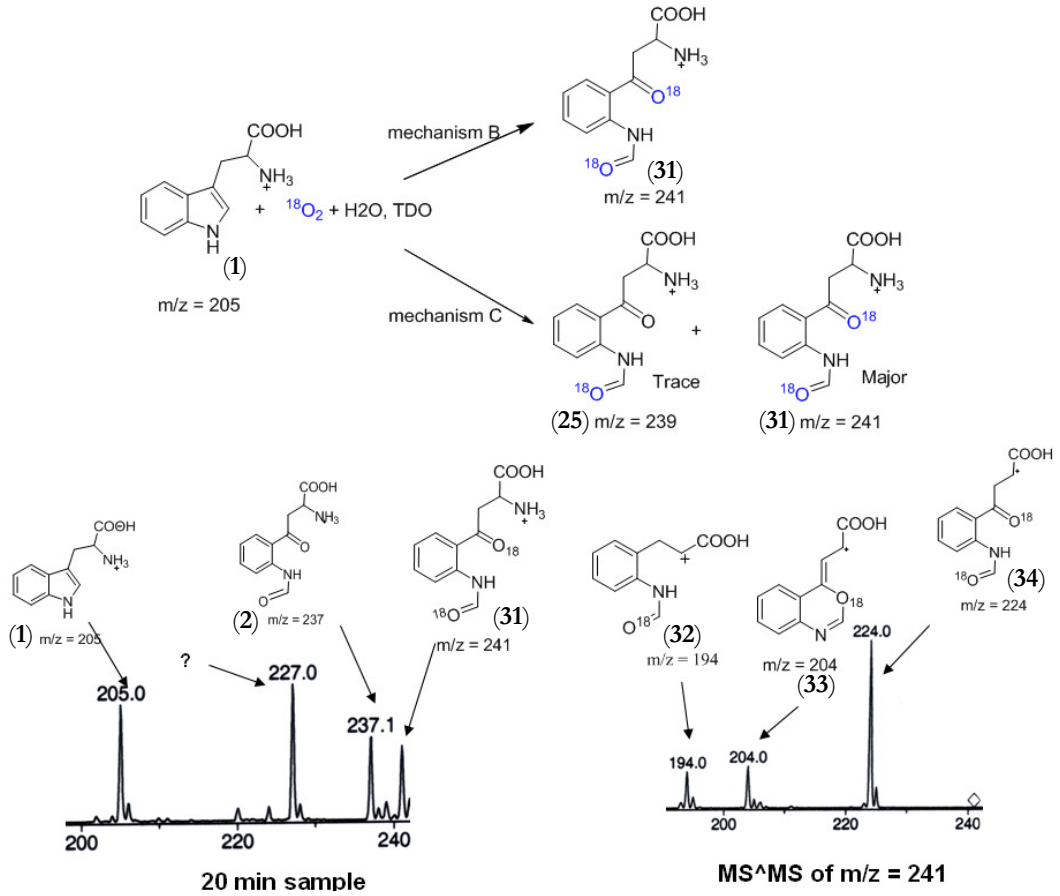


Figure 2.11: The scheme on the top shows the expected incorporation pattern of labeled oxygen gas mechanism B and C. Bottom left: Positive mode ESI-MS analysis of the reaction mixture after 20 minutes. Bottom right: Fragmentation analysis of the species at m/z = 241 is shown in the spectrum on the right.

The presence of a peak at m/z = 241 (**31**) on carrying out the reaction in presence of $^{18}O_2$ gas provides strong evidence that both the oxygen atoms in NFK (**2**) come from oxygen gas dissolved in water, thus it safely rules out the mechanistic route A for the formation of NFK (**2**) from tryptophan (**1**). There is however a peak at m/z =237 (**2**) the presence of which can be attributed to incomplete removal of $^{16}O_2$ from the buffer prior to the reaction and/or because of diffusion of $^{16}O_2$ from air.

On letting the reaction run in presence of $^{18}O_2$ for 4 hours and then carrying out an ESI-MS, a peak at m/z = 239 (**25**) was observed (*figure 2.12(a, b)*) along with peaks at m/z = 237 (**2**) and 241 (**31**). The peak at m/z = 239 (**25**), could be due to a single ^{18}O incorporation into the unlabelled NFK (2) produced as a result $^{18}O/^{16}O$ exchange of either the ketone or the formyl oxygen with H_2O^{16}. The species at m/z = 239 could be either (**24**) (if the formyl oxygen exchanges) or (**25**) (if the ketone oxygen exchanges) and they would react with KFA to give different products (*figure 2.12(c)*). The species (**24**) would produce labeled kynurenine (m/z = 211) (**8**) and (**25**) would produce unlabelled kynurenine (m/z = 209) (**10**).

On the other hand the peak at m/z = 241 (**31**), which is due to two ^{18}O incorporation in the NFK (**2**), will generate an ^{18}O labeled kynurenine (m/z = 211) (**8**) on treatment with KFA. So if the m/z = 239 species is (**24**) then the prediction would be that we get only m/z = 211 (**8**), otherwise we will get both m/z = 209 (**10**) as well as 211 (**8**). On treatment of the 4 hour reaction mixture with KFA, carried out under $^{18}O_2$ gas saturated buffer condition, ESI-MS showed the presence of peaks at m/z = 209 (**10**) as well as 211 (**8**) (*figure 2.13*) which could only be derived from (**25**) as (**24**) would give rise to the only peak at m/z = 211 (**8**) and no peak at m/z = 209 (**10**). MS^MS of the species with peaks at m/z = 209 (**10**) and 211 (**8**) showed the presence of de-amidated peaks at m/z = 192 (**36**) and 194 (**37**) respectively. The evidence of (**25**), being the species

contributing to m/z = 239 is a proof in itself that the ketonic oxygen is the labile one and not the amide one.

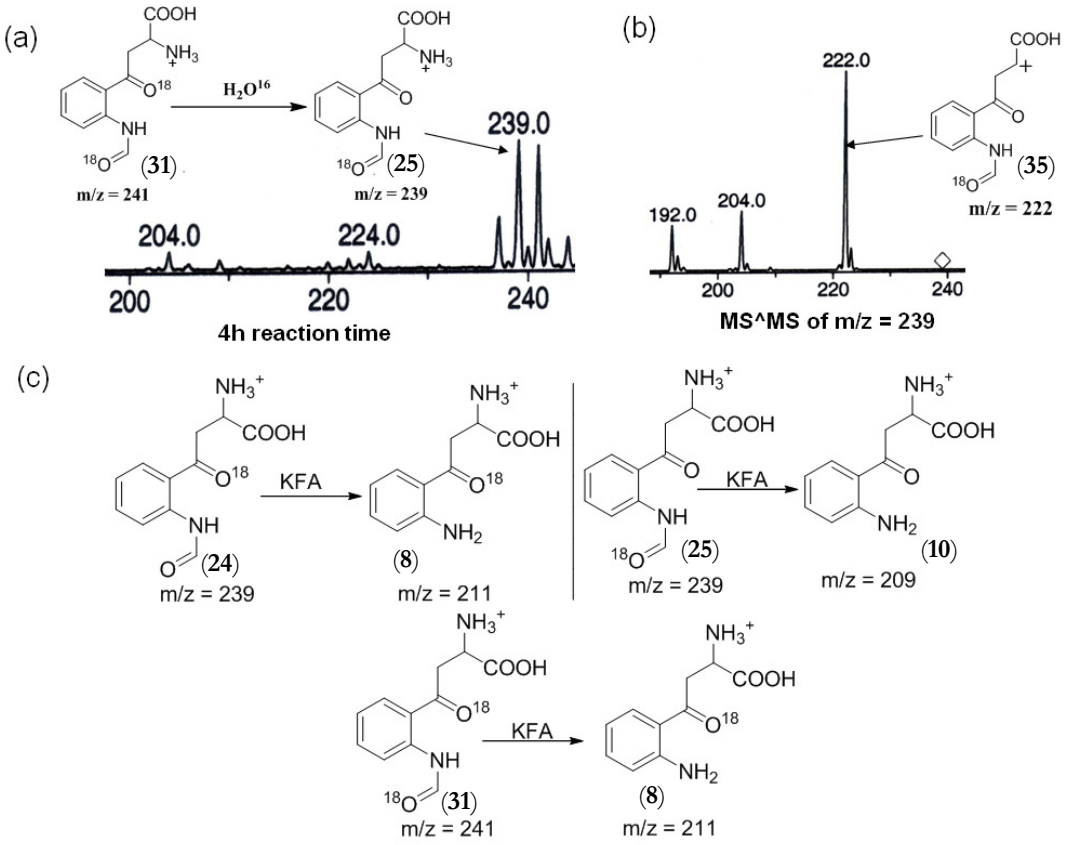


Figure 2.12: (a) ESI-MS analysis of the reaction mixture (tryptophan (**1**) and RmTDO in $^{18}O_2$ saturated buffer) after 4 hours. (b) Fragmentation analysis the species at m/z = 239 (**25**) (c) Various outcome of the species with m/z = 239 (**25**) & 241 (**31**) on treatment with KFA is shown. The scheme at the bottom shows the product that would be formed upon reaction of the species at m/z = 241 (**31**) with KFA.

2.3.3. Reaction of RmTDO with 6-methoxytryptophan (22) and 6-hydroxytryptophan (23). Both 6-methoxytryptophan (**22**) as well as 6-hydroxytryptophan (**23**) was tested as potential substrates for RmTDO. UV-visible scan was performed over time on an enzymatic reaction mixture (500 μL) containing 2 mM 6-methoxytryptophan (**22**), 18 μM RmTDO and 500 μM L-ascorbic acid in 100 mM

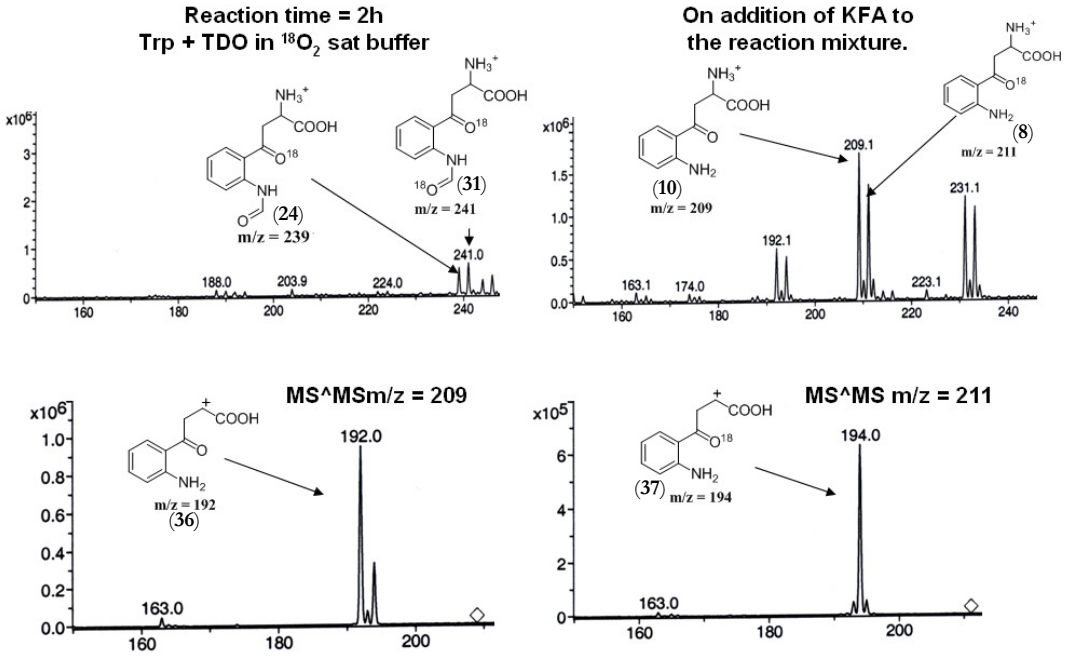


Figure 2.13: Top left: ESI-MS analysis of the two hour reaction of tryptophan (**1**) with RmTDO in $^{18}O_2$ saturated buffer. Top right: ESI-MS analysis of the reaction mixture after addition of KFA to the reaction mixture at the left. Bottom: Fragmentation analysis of the species with peaks at m/z = 211 (**8**) and m/z = 209 (**10**).

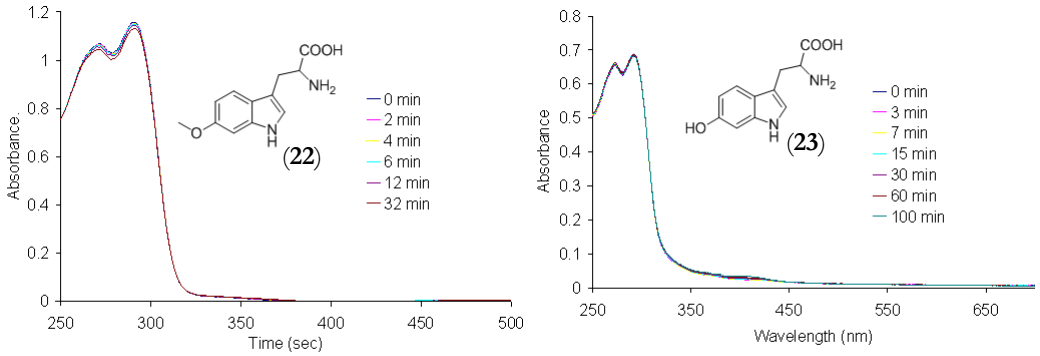


Figure 2.14: UV-Visible analysis of the RmTDO catalyzed oxygenation of 6-methoxytryptophan (**22**) (left) and 6-hydroxytryprophan (**23**) (right) over time.

phosphate buffer at pH 7.0. An exactly identical assay was set up with 6-hydroxytryptophan (**23**). The UV-visible spectrum of the reaction mixture, in both cases, in the region 250-700 nm did not change over the course of the reaction (*figure 2.14*) thereby showing no turnover. Neither 6-methoxytryptophan (**22**) nor 6-hydroxytryptophan (**23**) was a substrate for RmTDO. Both of them were assayed as potential inhibitors, but neither of them inhibited the rate of formation of NFK (**2**) over time to an appreciable extent.

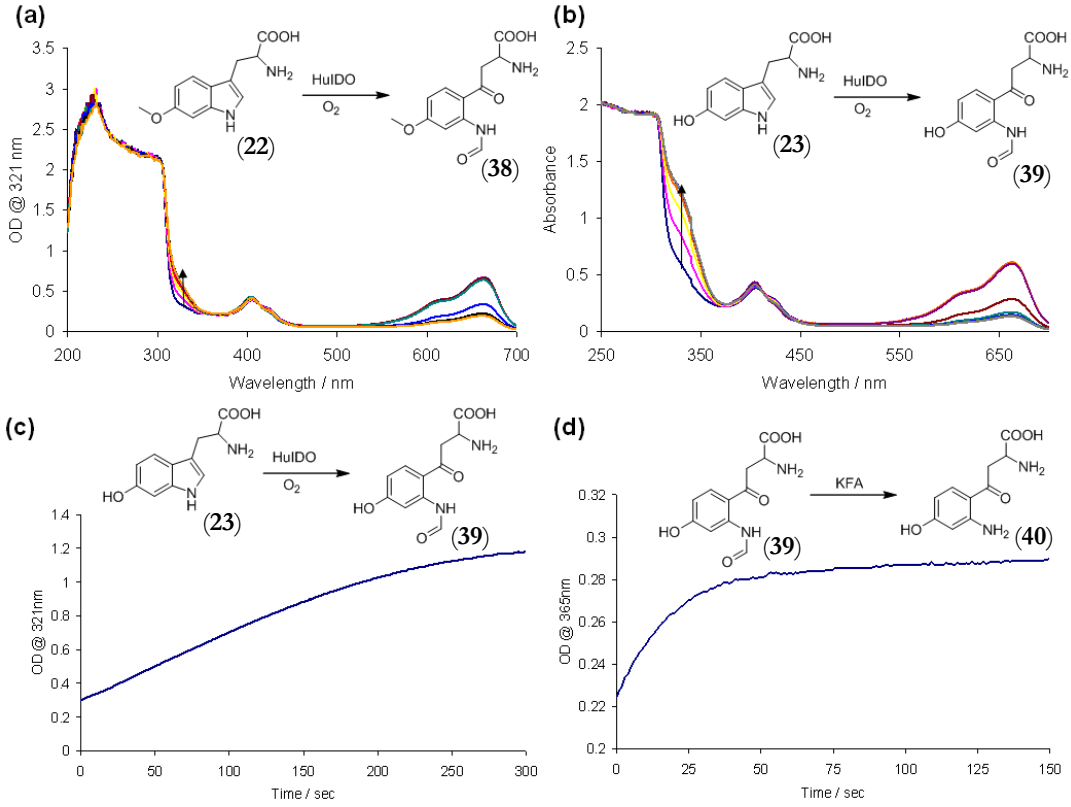


Figure 2.15: UV-Visible analysis of the reaction mixture containing (a) 6-methoxytryptophan (**22**) and HuIDO and (b) 6-hydroxytryptophan (**23**) over time. Assay of HuIDO with 6-hydroxytryptophan (**23**) as a substrate. (c) The formation of 6-hydroxy-NFK (**39**) was observed by monitoring the absorbance at 321 nm. (d) Formation of 6-hydroxykynurenine (**40**) over time upon addition of KFA to the previous reaction mixture.

2.3.4. Reaction of HuIDO with 6-methoxytryptophan (22). The reaction mixture (500 µl total volume) consisted of 400 µM 6-methoxytryptophan (**22**), 20 mM neutralized ascorbic acid solution, 0.2 mg/ml catalase solution, 10 µM methylene blue and 10.5 µM HuIDO in50 mM phosphate buffer at pH 6.5. The UV-visible spectrum between 200 and 700 nm for the reaction mixture at various time points was monitored. It was observed that there was an increase in absorbance at 321 nm, the characteristic absorbance maxima of NFK (**2**), *figure 2.15(a)*.

A decrease in absorbance at the absorbance maximum of methylene blue (675 nm) was observed, which was not investigated further. It has been reported by Sono (*14*) that methylene blue acts an electron mediator from ascorbic acid to the iron bound to the heme, thereby reducing any Fe(III) back to the active Fe(II) form.

2.3.5. Reaction of HuIDO with 6-hydroxytryptophan (23). The reaction mixture (500 µl total volume) consisted of 400 µM 6-hydroxytryptophan (**23**), 20 mM neutralized ascorbic acid solution, 0.2 mg/ml catalase solution, 10 µM methylene blue and 10.5 µM HuIDO in50 mM phosphate buffer at pH 6.5. UV-visible spectrum between 200 and 700 nm for the reaction mixture at various time points was monitored. It was observed that there was an increase in absorbance at 321 nm, the characteristic absorbance maximum for NFK (**2**), *figure 2.15 (b)*. The formation of 6-hydroxy-NFK (**39**) was quantified by the time dependent increase in absorbance at 321 nm wavelength (*figure 2.15(c)*).

An indirect proof for the production of 6-hydroxy-NFK (**39**) was obtained from the following experiment. 50 µl of KFA (unknown concentration), was added to the above reaction mixture (6-hydroxytryptophan (**23**) with HuIDO), and the absorbance at 365 nm, characteristic of kynurenine (**10**), was monitored. This absorbance was found to

increase after addition of KFA, suggesting that KFA catalyzed the formation of 6-hydroxykynurenine (**40**) from 6-hydroxy-NFK (**39**) (*figure 2.15(d)*).

A 1 ml sample of HuIDO was lyophilized. The green lyophilized powder was dissolved in 500 µl of 50 mM phosphate buffer at pH 6.5. Bradford assay indicated the protein solution to be 219 µM in strength. This protein sample was assayed for activity by monitoring the production of 6-hydroxy-NFK (**39**) over time at 321 nm wavelength. The assay mixture (500 µl total volume) consisted of 400 µM 6-methoxytryptophan (**23**), 20 mM neutralized ascorbic acid solution, 0.2 mg/ml catalase solution, 10 µM methylene blue and 10.5 µM HuIDO in50 mM phosphate buffer at pH 6.5. It was found that the protein was as active as before.

2.3.6. Reaction of HuIDO with 6-hydroxytryptophan (23) in H_2O^{18} 100 µl of 10 mM 6-hydroxytryptophan (**23**), 100 µl of 200 mM ascorbic acid neutralized with 2 M ammonium hydroxide, 50 µl of 4 mg/ml catalase solution and 2 µl of 5 mM methylene blue were mixed together and lyophilized. 1 ml of 62.98 µM HuIDO was also lyophilized separately. All solutions were made in 20 mM ammonium acetate buffer at pH 7.0. 1 g of H_2O^{18} was added to the green colored lyophilized protein and the protein solution was added to the lyophilized reaction mixture. The reaction mixture was briefly vortexed and after three minutes 400 µl of it was filtered through YM-10(MWCO 10kDa) centrifugal filter for 15 minutes at 30,000 xg. 100 µl of the filtrate was mixed with 100 µl of 95% methanol and the sample was sprayed in the mass spectrometer. Had there been any incorporation of ^{18}O from solvent then we would have expected a peak at m/z = 255 due to the incorporation of a single ^{18}O atom. The mass spectrum for the reaction is shown in *figure 2.16 (a)*.

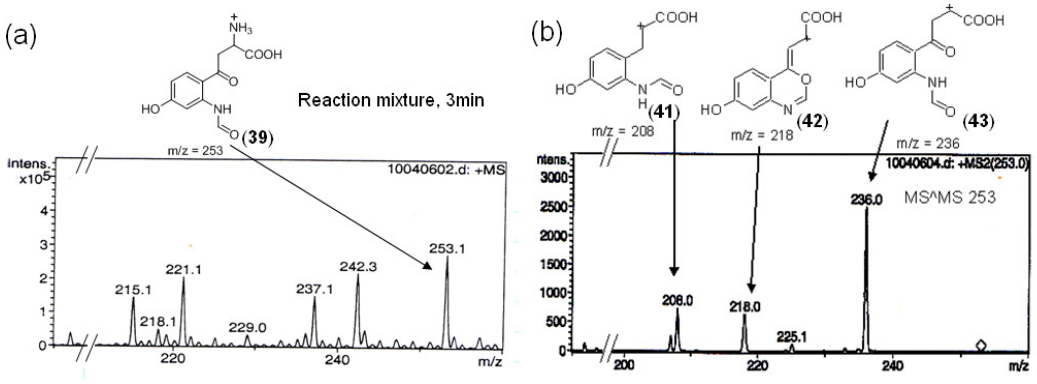


Figure 2.16 (a) ESI-MS analysis of the reaction mixture containing HuIDO and 6-hydroxytryptophan (**23**) in H_2O^{18} water. (b) Fragmentation analysis of the species with peaks at m/z = 253 (**36**).

There was no peak at m/z = 255, indicating that there was no incorporation of ^{18}O from the solvent. Fragmentation of the species at m/z = 253 matched with the expected fragmentation pattern for 6-hydroxy-NFK (**39**), there is a de-aminated peak at m/z = 236 (**43**), a peak at m/z = 218 (**42**) which could be explained by a rearrangement with a loss of water and a peak at m/z = 208 (**41**) which can be explained by loss of carbon monoxide, thereby proving the validity of the reaction, *figure 2.16 (b)*. It was not possible to isolate and fragment the species appearing at m/z = 255.

2.3.7 Reaction of TDO with tryptophan (1) – light emission experiment. TDO was purified previously and lyophilized. Green lyophilized TDO powder was added to 500 µl of 50 mM phosphate buffer at pH 7.0 till a saturated enzyme solution was produced. The resulting solution was allowed to sit on ice for half an hour. A small amount of protein precipitated. The solution was centrifuged for 10 minutes at 30,000 xg. The supernatant was taken and the amount of dissolved enzyme was calculated by the Bradford assay (*15*). The enzyme solution was found to be 943 µM in strength. Maximum concentration of

intermediate is build up at single turnover concentrations. Two kinds of substrate solutions were prepared: Solution (a) :1 ml solution containing 900 μM tryptophan (**1**) and 225 μM ascorbic acid (prepared by adding 45 μl of 20 mM tryptophan (**1**) and 11.3 μl of 20 mM ascorbic acid to 944 μl of 50 mM phosphate buffer at pH 7.0) and Solution (b): 1 ml solution containing 900 μM tryptophan (**1**), 20 mM ascorbic acid, 0.2 mg/ml of catalase and 10 μM of methylene blue (prepared by adding 45 μl of 20 mM tryptophan (**1**), 100 μl of 200 mM ascorbic acid that was netralized by NaOH prior to addition, 50 μl of catalase (2800 units/mg of solid from *Bovine* liver) of strength 4 mg/ml, 2 μl of 5 mM methylene blue to 803 μl of 50 mM phosphate buffer at pH 7.0).

Solution (a) is the solution that has been used for TDO assays through out and solution (b) was used for the assay of IDO. Both these substrate solution was used for the reaction to see any light emission during the course of the reaction.

In the stopped flow instrument [KinTek cooperation, Model SF-2004.5], 500 μl of the substrate solution was loaded in one of the syringe and 500 μl of the enzyme solution was loaded into the other. Care was taken to remove any air bubble from the syringe. Filters were removed from the photomultiplier tube, such that light emitted at any wavelength would get detected and the photomultiplier tube voltage setting was kept at its maximum, 1000V, so that the instrument gives maximum sensitivity for any light emission.

The enzyme and the substrate was mixed in the reaction chamber and the emission of light during the course of the reaction, if any, as a result of the decay of the dioxetane intermediate, was monitored over a time of 2 seconds, for reaction of solution (a), (b) as well as buffer (as a control) with the enzyme, *figure 2.17(a)*.

2.3.8. Reaction of luciferase with luciferrin – light emission experiment. An assay for detection of ATP using luciferase is reported in literature (*16*). Luciferase from

Photinus pyralis (firefly) was obtained as a lyophilized powder from Sigma. The activity of the protein was reported to be 18500000 light units/mg of protein by the manufacturer. D-luciferin (**12**) as the sodium salt was obtained from the same vendor. 25 mM Tris at pH 7.5 was used as the buffer for this assay instead of 25 mM glycine glycine buffer as mentioned in the literature. Tris was chosen as a replacement for glycine glycine buffer because the lyophilized sample of the protein has Tris in it. A 200 µl solution containing 40 µl of 80% glycerol, 2 µl of 200 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and 158 µl of 25 mM Tris buffer at pH 7.5 (to a final concentration of 20% and 20 mM respectively) was used to dissolving 1 mg of the protein in order to make a stock solution of 5 mg/ml. TCEP was added because luciferase is a sulfhydryl containing enzyme and requires two sulfhydryl group for catalytic activity. 5 mg of luciferin (**12**), was dissolved in 500 µl of 25 mM Tris buffer at pH 7.5; the final concentration of the stock solution being 33 mM. A 100 mM solution of ATP and $MgSO_4$ was made using the same buffer.

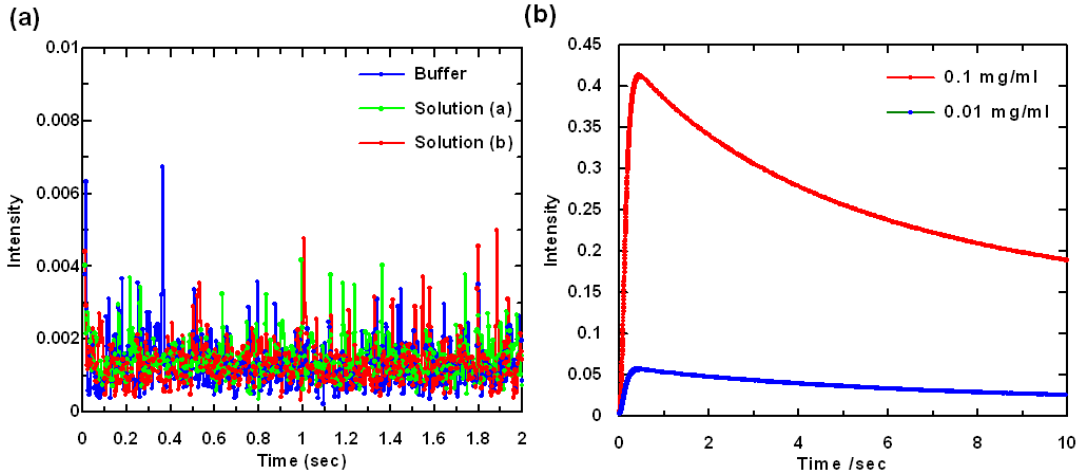


Figure 2.17: (a) Reaction of TDO with tryptophan (1). (b) Reaction of luciferase with D-luciferin (**12**). Light emission was monitored with the PMT voltage setting at the maximum in the stopped flow apparatus.

1 ml of the substrate solution was prepared by mixing 3.03 μl of 33 mM luciferin (**12**), 1 μl of 100 mM ATP, 10 μl of 100 mM $MgSO_4$ and 986.7 μl of 25 mM Tris buffer at pH 7.5, such that the final concentration of luciferin (**12**), ATP and $MgSO_4$ was 100 μM, 100 μM and 1 mM respectively. The 5 mg/ml stock solution of the protein was diluted to 500 μL of 0.1 mg/ml of protein by dissolving 10 μL of protein stock solution in 490 μL of buffer. This solution was further diluted 10 times to yield a protein solution of 0.01 mg/ml strength. The protein solution and the substrate solution were loaded into two syringes of the stopped flow instrument. Care was taken to remove any air bubble from the syringe. Filters were removed from the photomultiplier tube, such that light emitted at any wavelength would get detected, even though the emission is known to occur at 615 nm, (*16*) in order to keep the settings consistent with that of the TDO experiment. The emission of light over 10 seconds due to the reaction was monitored with the photomultiplier tube voltage setting at 500 V and 600 V for 0.1 mg/ml and 0.01 mg/ml of protein solution respectively. The result is shown in *figure 2.17(b)*.

2.4. Discussion

2.4.1. ^{18}O labeling studies. The results obtained from the ^{18}O labeling studies done on RmTDO showed that the $^{18}O_2$ experiment and the H_2O^{18} experiments are mutually consistent with one another and both the oxygen atoms incorporated into NFK (**2**) by RmTDO come from atmospheric oxygen. This agrees with the results of Hayaishi (*9*). Our experiment improves on that method, in that a significantly smaller quantity of H_2O^{18} was used and the time-consuming crystallization was avoided. This experiment also showed the presence of labeled oxygen in NFK (**2**) as well in kynurenine (**10**).

It does not however distinguish between the mechanistic routes B and C (*figure 2.1*). It cannot be excluded that the rate of exchange of the oxygen bound to Fe (as shown in mechanism C) with solvent, might be slower than the recombination reaction,

resulting in the presence of both the labeled oxygen atoms in NFK (**2**). There has been no evidence for dioxetane (mechanism B) playing a role in any of the characterized dioxygenase catalyzed ring opening reactions of aromatic compounds. The retro [2+2] cycloaddition could occur by emission of light as seen in the case of the enzyme luciferase, in firefly. No such evidence is reported in case of mechanism B. Sono in his review (*4*) points out that the high endothermicity of the formation of dioxetane and the high exothermicity of its decomposition makes this pathway the less likely one.

If the reaction proceeds *via* mechanism C then, if by chemical means we could slow down the rate of recombination then it might be possible to observe exchange of the ketonic oxygen with the solvent enzymatically. This could be done by stabilizing the intermediate carbocation (**7**) formed. Such an effect would be expected to be maximal for tryptophan (**1**) derivatives possessing an electron donating group at the C-4 and/or C-6 positions. Indoleamine 2,3-dioxygenase (IDO) catalyzes the same step in the degradation of tryptophan (1). TDO is localized in the liver, while IDO is ubiquitously present in the body. From the studies on tryptophan (1) analogues done in IDO it was seen that the enzyme was more tolerable to substituents at the C-6 position. It was thus decided to synthesize 6-methoxytryptophan (**22**) and 6-hydroxytryptophan (**23**) (instead of the analogues with substitution at the C-4 position) and use them as substrate, with the hope that they would be taken up by the enzyme, RmTDO, and methoxy or the hydroxyl group being electron donors would stabilize the intermediate carbocation such that the rate of the recombination would be slow enough to allow the exchange of the oxygen bound to iron with the solvent.

Neither 6-methoxytryptophan (**22**) nor 6-hydroxytryptophan (**23**), that was synthesized were substrates for RmTDO.

IDO accepts a wide variety of substrates like D-tryptophan, tryptamine, serotonin, as well as 5-hydroxytryptophan, 5-methoxytryptophan, 5-methyltryptophan,

6-methyltryptophan etc (*4*). The catalytic mechanism of both IDO and TDO are believed to be similar, proceeding via the hydroperoxide intermediate followed by either a dioxetane intermediate or undergo a Criegee rearrangement (*1*), neither of which have been established. Thus, IDO offered a more flexible system for mechanistic studies than is available with RmTDO.

Both 6-methoxytryptophan (**22**) as well as 6-hydroxytryptophan (**23**) was found to be substrate for HuIDO. Hydroxy group, being a better electron donating group than the methoxy group was thought to be a better stabilizer of the intermediate carbocation (**7**), in case the reaction went via the Criegge mechanism. Moreover, it is less bulky than methoxy group and thus could be a better substrate for the enzyme. So, 6-hydroxytryptophan (**23**) was used as a substrate for the reaction of the HuIDO in H_2O^{18}.

As evident from the results, no incorporation of ^{18}O from solvent into 6-hydroxy-NFK (**39**) was seen. We would have been able to see the incorporation of ^{18}O from solvent, provided that:

a) Criegge rearrangement was the catalytic route the enzyme took for the production of 6-hydroxy-NFK (**39**), via the carbocation intermediate.
b) The stabilization of a carbocation intermediate in the presence of the electron donating hydroxyl group decreases the rate of the recombination reaction sufficiently to allow the exchange of the oxygen bound to the heme with the solvent. This would lead to the incorporation the ^{18}O from the solvent into the product.

Based on the results of this experiment, it was not possible to differentiate between mechanism B and C using oxygen labeling. Even though there was no ^{18}O incorporation in 6-hydroxy-NFK (**39**), we still could not rule out the mechanistic route C. The assumption (b) might not have been a very valid one. Even if the intermediate carbocation were stabilized by a hydroxyl group at position C-6 of tryptophan (**1**), the

stabilization might not be sufficient to slow the rate of recombination, thereby allowing the exchange of the oxygen bound to the heme with H_2O^{18}. Moreover at neutral pH, the exchange of the oxygen bound to the heme with water will be very slow. Likelihood of the exchange could be enhanced at a lower pH. However the labiality of the ketonic oxygen in NFK (**2**) would also increase under these conditions, and this would pose a problem in accurate determination of whether the oxygen incorporated in the ketonic position of NFK (**2**) is due to exchange of the oxygen bound to heme or due to non-enzymatic exchange of the labile oxygen with the solvent. If the two rates mentioned above are sufficiently different and if the reaction indeed went via the Criegee intermediate (mechanism C) then we would be able to observe ^{18}O incorporation in 6-hydroxy-NFK (**39**).

2.4.2. Light emission studies. For the reaction of tryptophan (**1**) with TDO the final concentration of both enzyme and substrate was 450 μl in the reaction chamber. As seen in *figure 2.17*, no light emission occurs during the course of the reaction. If there were any detectable light emission we would have expected it to be above the noise level (buffer in this case).

Arguments can easily be made against the efficiency of the instrument and blaming it for not being able to detect low quanta of light emission. No information stating the lower limit of the photomultiplier tube was available from the manual of the instrument. Since firefly-luciferase enzyme is known to emit light, an assay for light emission by luciferase was performed in this instrument to test the latter's ability for detection of low amounts of photon.

For the reaction of luciferin (**12**) with luciferase, there was a rapid flash of light followed by slow decay over a long period of time. This proves beyond doubt that light emitted can be detected by the photomultiplier tube in the stopped flow. A difficulty with

using very low concentration of protein is that the enzyme gets deactivated on standing. A very low dilution of 0.1 ng/ml of protein solution was made and immediately used for the experiment. The photomultiplier tube setting was at its maximum, 1000V. The result of light emission from 0.1 ng/ml of enzyme concentration is superposed in the same graph obtained for TDO (*figure 2.18*).

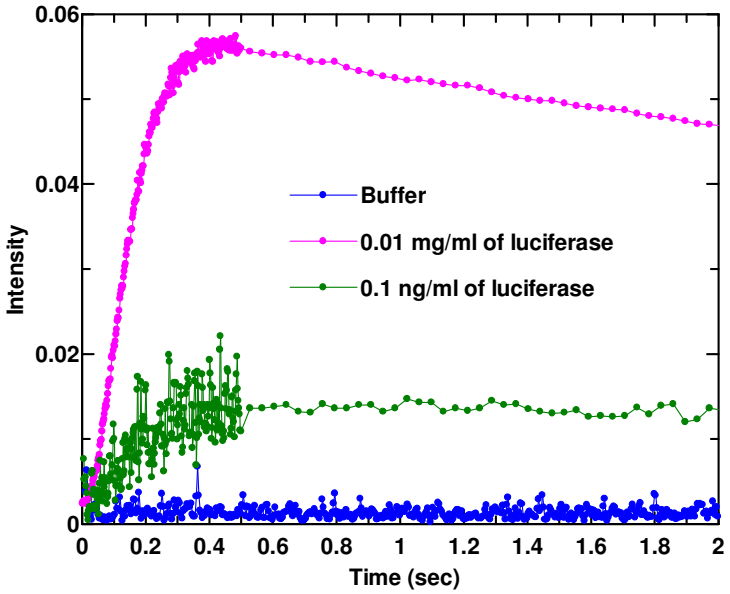


Figure 2.18: The reaction of TDO is compared with 0.1 ng/ ml of luciferase. Light emission was monitored with the PMT voltage setting at the maximum in the stopped flow apparatus.

The flow cell volume is 10 µl. Equal volumes of substrate and enzyme enter the reaction cuvette. Thus, the enzyme concentration [0.1 ng/ml] is further diluted to 0.05 ng/ml. 10 µl of the reaction mixture contains 5×10^{-13} g of protein i.e. 4.17×10^{-18} mol [MW: 120 kDa], which corresponds to 2.5×10^{6} molecules of protein. Light emission could thus be detected from such a low number of protein molecules. In contrast to the luciferase concentration the amount of TDO used was 900 µM, which corresponds to

2.7×10^{-15} molecules of protein. Assuming that the entire TDO preparation was active the concentration of TDO (determined by Bradford assay) is 10^9 times more than that of luciferase. The quantum efficiency of light emission for luciferase is reported to be very close to 1 (*7*). Under the assumption that the efficiency of detection of photon emission from the two enzymes are same/comparable the experimental results suggest that the quantum efficiency of the decay of dioxetane (**6**) in TDO must be less than 10^{-9} when it could be said that the detection of light was not possible due to the limitation of the photomultiplier tube.

This experiment does not suggest that dioxetane (**6**) is not the intermediate. It only proves that light is not emitted, or if emitted then not detectable, during the course of the reaction. Dioxetane (**6**) could very well be formed as an intermediate and then decompose to NFK (**2**) without emission of light. The emission of light is only possible if the dioxetane (**6**) undergoes a retro [2+2] cycloaddition and produces NFK (**2**) in the excited state and which then decays to the product in the ground state with emission of light. Moreover the decomposition of the dioxetane (**6**) may not be concerted (retro [2+2] cycloaddition) and could very well happen in a sequential manner. Moreover, if it did occur in the concerted manner then it could very well be possible that the decay to the ground state is non-radiative.

This experiment did not nail any mechanism but showed that the photomultiplier tube in the stopped flow instrument could be used for detection of very low amount of light. This also paved a new method for detection of bioluminescence.

REFERENCE:

(1) Bugg, T. D. H. (2003) Dioxygenase enzymes: catalytic mechanisms and chemical models. *Tetrahedron 59*, 7075.

(2) Tanaka, T., and Knox, W. E. (1959) The nature and mechanism of the tryptophan pyrrolase (peroxidase-oxidase) reaction of Pseudomonas and of rat liver. *J Biol Chem 234*, 1162-70.

(3) Ishimura, Y., Nozaki, M., and Hayaishi, O. (1970) The oxygenated form of L-tryptophan 2,3-dioxygenase as reaction intermediate. *J Biol Chem 245*, 3593-602.

(4) Sono, M., Roach, M. P., Coulter, E. D., and Dawson, J. H. (1996) Heme-Containing Oxygenases. *Chem Rev 96*, 2841-2888.

(5) Terentis, A. C., Thomas, S. R., Takikawa, O., Littlejohn, T. K., Truscott, R. J., Armstrong, R. S., Yeh, S. R., and Stocker, R. (2002) The heme environment of recombinant human indoleamine 2,3-dioxygenase. Structural properties and substrate-ligand interactions. *J Biol Chem 277*, 15788-94.

(6) Leeds, J. M., Brown, P. J., McGeehan, G. M., Brown, F. K., and Wiseman, J. S. (1993) Isotope effects and alternative substrate reactivities for tryptophan 2,3-dioxygenase. *J Biol Chem 268*, 17781-6.

(7) Bugg, T., and Winfield, C. (1998) Enzymic cleavage of aromatic rings: mechanistic aspects of the catechol dioxygenases and later enzymes of bacterial oxidative cleavage pathways. *Natural Product Reports 15*, 513.

(8) Mayer, R. J., and Que, L., Jr. (1984) 18O studies of pyrogallol cleavage by catechol 1,2-dioxygenase. *J Biol Chem 259*, 13056-60.

(9) Hayaishi, O., Rothberg, S., Mehler, A. H., and Saito, Y. (1957) Studies on oxygenases; enzymatic formation of kynurenine from tryptophan. *J Biol Chem 229*, 889-96.

(10) Kurnasov, O., Goral, V., Colabroy, K., Gerdes, S., Anantha, S., Osterman, A., and Begley, T. P. (2003) NAD biosynthesis: identification of the tryptophan to quinolinate pathway in bacteria. *Chem Biol 10*, 1195-204.

(11) Oda, S., Sugimoto, H., Yoshida, T., and Shiro, Y. (2006) Crystallization and preliminary crystallographic studies of human indoleamine 2,3-dioxygenase. *Acta Crystallogr Sect F Struct Biol Cryst Commun 62*, 221-3.

(12) Littlejohn, T. K., Takikawa, O., Skylas, D., Jamie, J. F., Walker, M. J., and Truscott, R. J. (2000) Expression and purification of recombinant human indoleamine 2, 3-dioxygenase. *Protein Expr Purif 19*, 22-9.

(13) Wickern, B. v., Müller, B., Simat, T., and Steinhart, H. (1997) Determination of [gamma]-radiation induced proudcts in aqueous solutions of tryptophan and synthesis of 4-, 6- and 7-hydroxytryptophan. *Journal of Chromatography A 786*, 57.

(14) Sono, M. (1989) The roles of superoxide anion and methylene blue in the reductive activation of indoleamine 2,3-dioxygenase by ascorbic acid or by xanthine oxidase-hypoxanthine. *J Biol Chem 264*, 1616-22.

(15) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem 72*, 248-254.

(16) DeLuca, M. a. M., W. D. (1978) Purification and Properties of Firefly Luciferase. *Methods in Enzymology* ***LVII***, 3.

CHAPTER 3

Kinetic studies on mutants of Tryptophan 2,3-dioxygenase*

3.1 Introduction

The crystal structure of tryptophan 2,3-dioxygenase (TDO) from *Ralstonia metallidurans* (PDB: 2NOX) with cofactor heme bound in the active site was solved by Dr. Yang Zhang (Ealick group). The crystal structure contains four tetramers per asymmetric unit. Each active site resides at the interface of two monomers, comprised mostly by one monomer and the N-terminus of the other monomer *figure 3.1*. Based on the active site information from the crystal structure, modeling as well as mutagenesis studies were carried out to investigate the enzyme mechanism.

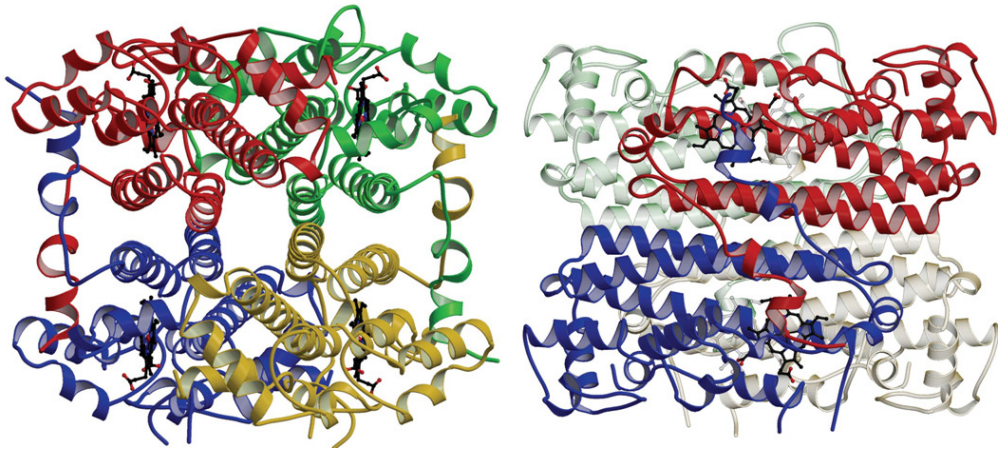


Figure 3.1: Crystal structure of RmTDO. Left: The four tetramers with heme bound at the active site. Right: The active sites at the interface of the two monomers.

3.2 Experimental

3.2.1. Molecular Cloning. Standard methods were used for DNA manipulations (*1, 2*). Plasmid DNA was purified with a Qiagen Miniprep kit. *E. coli* strain MachI (Invitrogen)

* Reproduced in parts with permission from Yang Zhang, Seong A. Kang, Tathagata Mukherjee, Shridhar Bale, Brian R. Crane, Tadhg P. Begley and Steven E. Ealick. (2007). *Biochemistry,* **46**, 145-155. © [2007] American Chemical Society.
The protein was crystallized by Seong, the X-ray crystallographic and analysis were done by Yang and modeling studies were done by Shridhar.

was used as a recipient for transformations during plasmid construction and for plasmid propagation and storage. Site-directed mutagenesis was performed on pRmTDO.PXHta by a standard PCR protocol using *Pfu*Turbo DNA polymerase per the manufacturer's instructions (Invitrogen) and *Dpn*I (New England Biolabs) to digest the methylated parental DNA prior to transformation. For each mutant, a third primer was designed to screen for the presence of the mutant by colony PCR with an appropriate vector-specific primer. Only clones that produced a PCR product were sequenced. In every case, the mutagenesis primer pair consisted of the primer whose sequence is in *table 2.1* and its reverse complement.[†]

Table 3.1: Primer sequences used for mutagenesis and screening.

Mutant	Mutagenesis Primer (top strand)/Screening Primer
Y130F	GATGACGCCGCCCGAGTTCAGCGCGATGCGCCCATACC TGACGCCGCCCGAGTTCAGC
F68A	GATCACAATGAGATGCTCGCAATCGTCCAACACCAGACCACC GATCACAATGAGATGCTCGC
T271A	GGTGCCGCCCGCGCCGCGCTTGAATCCGATTACCCGCTCGACTGTGG CTGACGCCTTCGGTGCCGCCCGC
R134A	GAGTACTCGGCCATGGCTCCATACCTGGGGGCATCGTCC GCCCCCAGGTATGGAGC

3.2.2. Activity of TDO Mutants: Substrate solutions of varying concentrations were prepared by dissolving the required amount of L-tryptophan and ascorbic acid (reducing agent to keep the enzyme in the Fe(II) oxidation state) in molar ratio of 4:1 in 100 mM phosphate buffer at pH 7.0. 50 μL of freshly purified enzyme (strength varying from 100

[†] The molecular cloning was done by Dr. Cynthia Kinsland

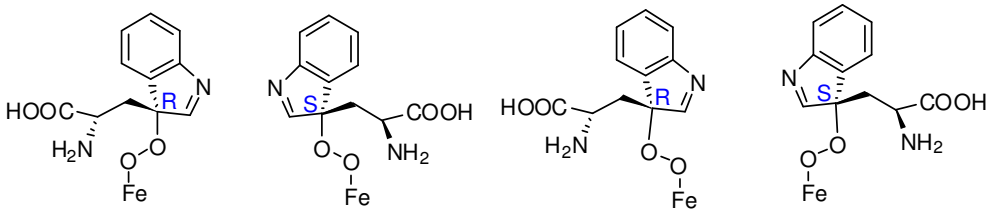
to 200 μM depending on the preparation of enzyme) was added to 450 μL of the substrate solution (always ensuring that the steady state condition is not violated) and the apparent catalytic activity of all the mutants and wild type was determined by measuring the rate of product (N-formyl kynurenine) formation at 321 nm (ε = 3152 $M^{-1}cm^{-1}$ for NFK). The K_M for tryptophan and k_{cat} for the active mutants and wild type were determined by fitting the rate of product formation against tryptophan concentration to the Michaelis-Menten equation using Grafit 5.0.11 (Erithacus Software, Horley, UK).

3.2.3. Activity with substrate analogues of tryptophan. 20 mM solutions of the tryptophan analogues were prepared. 5 μL of 20 mM solution was added to a mixture of 25 μL of 100 μM RmTDO, 5 μL of 5 mM ascorbic acid and 465 μL of 100 mM phosphate buffer at pH 7.0. All solutions were made in 100 mM phosphate buffer at pH 7.0. A solution of L-tryptophan was assayed in order to test the activity of the enzyme. Because of poor solubility of tryptamine; 20 μL of 5 mM solution was used in the assay. The final strength of each substrate was 200 μM, enzyme 5 μM and ascorbic acid 50 μM. For reference solution everything other than the substrate was added. UV-visible spectrum of this reaction mixture was taken at various time intervals.

***3.2.4. Modeling of the Tryptophan Hydroperoxide into the Active Site*†.** The modeling of the putative hydroperoxide intermediate (**1, 2, 3,** and **4**) (*figure 3.2*) into the active site of TDO was carried out using the program Macromodel, version 7.2 (*3*). Modeling simulations were preformed in which the Cβ-Cγ torsion was either constrained (*anti* periplanar or *syn* coplanar) or allowed to freely rotate. The base shell of atoms included all residues within 20.0 Å of the heme iron and was used as the starting model for energy minimization. The intermediate was built into the active site, followed by

† The modeling studies were performed by Dr. Shridhar Bale

removal of water molecules and hydrogen atoms were added appropriately. The starting structure was subjected to mixed Monte Carlo MCMM/lowmode conformational search steps allowing residues in the 4 Å shell around the active site to move. This included the key residues His72, Tyr43 and Phe68. The conformational search allowed torsional rotation about the Fe-O1, O1-O2, O2-Cγ, Cγ-Cβ, Cβ-Cα and Cα-C bonds (O1 and O_2 are the two incorporated oxygen atoms). The generated structures were energy minimized to a gradient of 0.01 kJ/molA *in vacuo* using the AMBER* force field (*4, 5*), a distance dependant electrostatics further attenuated by a factor of 4 (4r electrostatics) and a TNCG minimization technique (*6*). The AMBER* parameters for the heme were adapted from D. A. Giammona (*7*). To test the effect of solvation on the minimization, the conformational search was repeated with a minimization that uses the GB/SA (*8*) solvation model. This search yielded a global minimum similar to the one from the *in vacuo* search.



(1) (2) (3) (4)

Figure 3.2: The hydroperoxide compounds that were modeled into the active site of *R. metallidurans* TDO.

3.3. Results

3.3.1. Modeling of hydroperoxide intermediate. Although the substrate L-tryptophan was present in the crystallization condition, there was no corresponding electron density in any of the 16 subunits. Because attempts to obtain experimental data for a substrate or

product complex have so far been unsuccessful, we prepared a model of the hydroperoxy intermediate into the predicted active site (*table 3.2*).

Table 3.2: The energies and torsion angles from the modeling

Putative intermediate	Energy (kJ/mol)	Torsion angle
1	-171.03	-173.6
2	-228.06	163.0
3	-172.15	13.8
4	-202.58	-18.9
1 - no torsional constraints	-221.39	-51.0
2 - no torsional constraints	-234.82	90.6

The global energy minimum structure for the hydroperoxy intermediate showed two residues with significant conformational change: Tyr43* and His72. The rest of the heme binding site residues are essentially the same as the unliganded TDO structure. The tryptophan ring of the intermediate stacks perpendicularly to the side chain of Phe68. While the minimum energy structure corresponded to a Cβ-Cγ torsion angle in between *anti* periplanar (*figure 3.3 (a), (d)*) and *syn* coplanar (*figure 3.3 (b), (c)*), the constrained simulations showed a clear preference for the S-*anti* periplanar structure (**2**) over structures (**1**, **3** and **4**). In this model the substrate forms good hydrogen bonds between its carboxylate group and Tyr130, Arg134 and the amide group and side chain of Thr271, and hydrogen bonds between its amino group and Thr271 and a propionate of the heme group (*figure 3.2(a)*).

3.3.2. Activities of Mutants and Substrate Analogs.

Several mutants of TDO was made and tested for activity. The steady state kinetic parameters of the active ones are listed in *table 3.3*. Various substrate analogues (D-Tryptophan, tryptamine, indole propionic acid and indole) were used to test the

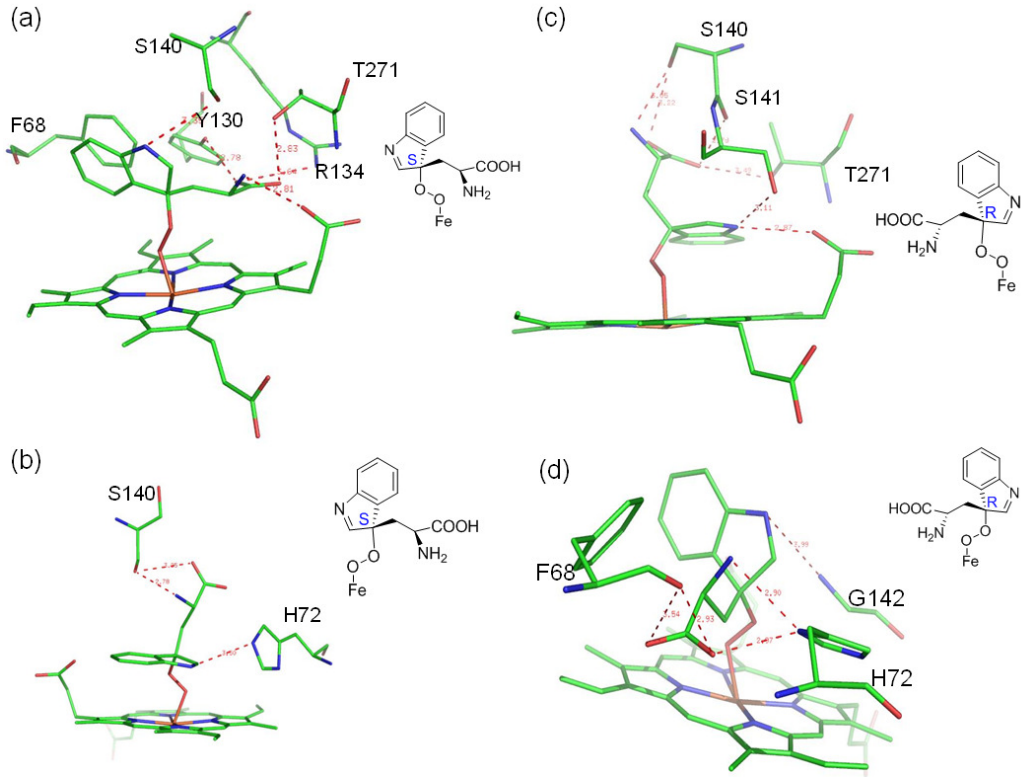


Figure 3.3: Model of (a) S-*anti* periplanar hydroperoxide (b) S-*syn* coplanar hydroperoxide (c) R-*syn* coplanar hydroperoxide and (d) R-*anti* periplanar hydroperoxide in the active site of RmTDO.

Table 3.3: Steady state kinetic parameters RmTDO and its mutants.

Mutant	**k_{cat} (s^{-1})**	**K_M (μM)**	**k_{cat}/K_M ($\mu M^{-1}s^{-1}$)**	**Relative specificity**
Native	0.01 (-)	385 (42)	3.06×10^{-5}	1
H72A	0.13 (0.01)	5927 (725)	2.24×10^{-5}	0.73
R134A	Inactive			
T271A	Inactive			
Y130F	0.20(0.01)	438 (54)	45.8×10^{-5}	15
F68A	Inactive			
F68C	Inactive			
F68D	Inactive			
F68E	Inactive			
F68H	0.05 (-)	389 (30)	11.8×10^{-5}	4
F68Y	0.005 (-)	609 (74)	0.1×10^{-5}	0.03

activity for TDO. TDO did not accept any of the analogues as alternative substrates. The F68A, F68C, F68D, F68E, R134A and T271A mutants of TDO were found to be inactive under the assay conditions. Activity of F68Y mutant was found to be 0.03 times that of the wild type enzyme, while the F68H was found to be 4 times more than the native enzyme (*figure 3.4(a)*). H72A mutant was found to be 0.73 times less active than the wild type, however, the Y130F mutants showed a 15-fold increase in specific activity compared to wild type TDO using L-tryptophan as the substrate (*figure 3.4(b)*).

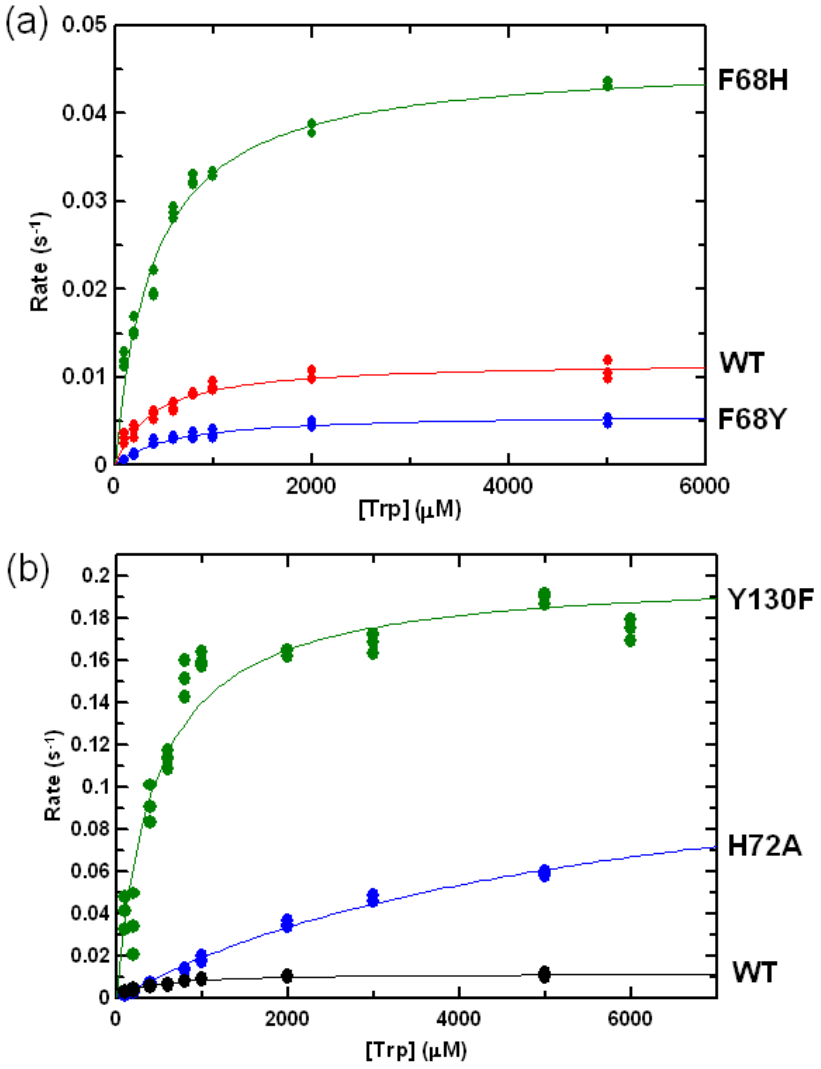


Figure 3.4: Steady state kinetic parameters of (a) F68H, F68Y and (b) Y130F, H72A mutants of RmTDO.

3.4. Discussion

The Criegee rearrangement requires an *anti* periplanar arrangement of the O-O bond and the C2-C3 bond of the indole (i.e., **1** or **2**) while the dioxetane mechanism requires a *syn* coplanar configuration (i.e., **3** or **4**). These four putative intermediates have been modeled into the active site. The resulting energies and torsion angles are given in *table 3. 2*. The modeling suggests that the binding of the R-peroxides (**1** and **3**) is less favorable than the binding of the S-peroxides (**2** and **4**) and that binding of the *anti* periplanar conformation (**2**) is favored over binding of the *syn* coplanar conformation (**4**). While these results provide some support for the Criegee mechanism, this support is tentative: when the torsional angle constraint is removed, both the R and the S peroxides bind tightly with the oxygen located halfway between the *syn* and the *anti* conformations.

The structures generated by modeling the S-peroxides (**2** and **4**) into the active site are shown in *figure 3.3(a).* For the *anti* hydroperoxide (**2**) the carboxylate is anchored at the active site by hydrogen bonding interactions with Arg134, Tyr130 and the hydroxyl and amide NH of Thr271. The tryptophan amino group is hydrogen bonded to the hydroxyl of Thr271 and to the heme propionic acid side chain. The indole ring is stacked against Phe68. Overall, there is good structural complementarity between the intermediate and the active site functionality with only one possible exception: a basic residue for the deprotonation/reprotonation of the indole nitrogen is not apparent.

For the *syn* hydroperoxide (**4**), there are no interactions between the carboxylate and the enzyme, the tryptophan amino group is hydrogen bonded to the carbonyl group of Ser140 and the indole is stacked over the heme. A basic residue for the deprotonation/reprotonation of the indole nitrogen is also not apparent in this model. Overall, there is poor structural complementarity between this intermediate and the active site functionality.

From the alignment of RmTDO with all other TDO's whose sequence are available, it was seen that the residues H-72 is fairly conserved, occasionally replaced by serine in some species. It is present at the axial side of the heme, close to the iron centre and was thought to be the potential base which could abstract the indole NH of tryptophan to trigger the reaction. On mutating it to alanine no large loss of activity was seen, as was expected had it been the base abstracting the indole proton.

The oxidation of substrate analogs, as well as the catalytic properties of active site mutants, were examined to experimentally differentiate between the *syn* and *anti* conformations of the bound (S)-tryptophan hydroperoxide. The results are shown in *table 3.3*.

The model for the binding of the *anti* hydroperoxide (**2**) predicts specific strong interactions between the enzyme and the amino and carboxylate groups of the substrate as well as high selectivity for the L-isomer of tryptophan over the D-isomer. In contrast, the model for the binding of the *syn* hydroperoxide (**4**) predicts weak or no interactions between the enzyme and the amino and carboxylate groups of the substrate as well as low selectivity for the L-isomer of tryptophan. This suggests that that D-tryptophan (**5**), tryptamine (**6**), indole propionic acid (**7**), and indole (**8**) will be good substrates for the enzyme if the oxidation proceeds via the dioxetane mechanism and poor substrates if the oxidation proceeds via the Criegee mechanism. Our observation that none of these analogs are substrates for TDO (*table 3.3*) supports the model for the binding of the (S)-tryptophan hydroperoxide in the *anti* periplanar conformation and suggests that the oxidation occurs by the Criegee mechanism. The results of active site mutagenesis experiments are generally in agreement with this analysis. In the model for the binding of the *anti* hydroperoxide **2**, Phe68 forms the indole-binding pocket, Arg134 forms hydrogen bonds to the substrate carboxylate group and the hydroxyl of Thr271 forms

hydrogen bonds to the substrate amino and carboxylate groups. As predicted, the F68A, R134A and T271A mutants are inactive.

Phe68 could be involved in hydrophobic interactions with the indole ring of tryptophan. Phe68 when mutated to non-aromatic residues like alanine, cystine, aspartate and glutamate, showed no activity. However the activity was restored on mutating it to aromatic residues. F68Y mutant showed much lower specific activity, while the F68H mutant has 4 times higher specific activity than native enzyme. The histidine group could interact favorably with the tryptophan group positioning it in a better catalytic orientation, while the tyrosine group interferes with the stabilizing hydrophobic interactions of phenylalanine which results in the lowering and raising of the specific activity of the mutant (with respect to the wild type) respectively. This explanation is consistent with the size of the mutants. Histidine is smaller than the phenyl group and hence do not interfere with the spatial orientation of the active site residues, where as tyrosine being bulkier than phenyl group because of the presence of the hydroxyl moiety might interfere with the conformation of the active site hence lowering the activity of the mutant

The model suggests that the hydroxyl of Tyr130 also forms a hydrogen bond to the substrate carboxylate. The mutagenesis experiments do not support this interaction because the Y130F mutant shows a 15-fold increase in catalytic activity rather than the expected decrease. Y130 has been mutated by nature to phenylalanine in 39% of the sequences available. This mutation could interact with the F-68 which in turn interacts with the indole ring of tryptophan thereby positioning the intermediate better for catalysis. However no solid evidence to this fact is available currently to offer an explanation for higher turnover of the enzyme. All mutants bound the heme cofactor suggesting that loss of activity was not due to protein misfolding. The mutants were also analyzed using the model for the binding of the *syn* hydroperoxide. The large loss of

catalytic activity observed for the F68A, R134A and T271A mutants is not consistent with this model.

In conclusion, the modeling of the tryptophan hydroperoxide into the active site of the enzyme coupled with substrate analog and active site mutagenesis studies are consistent with a Criegee mechanism for the indole oxidative ring opening reaction. Additional biochemical experiments are in progress to further elucidate the mechanistic details of this important oxidation.

.

REFERENCES

(1) Ausubel, F. M., and Brent, F. (1987) in *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.

(2) Sambrook, J., Fritsch, G. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Guide*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

(3) Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., and Still, W. C. (1990) MacroModel - an Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. *J. Comput. Chem. 11*, 460-467.

(4) Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc. 106*, 765-784.

(5) Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem. 7*, 230-252.

(6) Ponder, J. W., and Richards, F. M. (1987) An Efficient Newton-like Method for Molecular Mechanics Energy Minimization of Large Molecules. *J. Comput. Chem. 8*, 1016-1024.

(7) Giammona, D. A. (1984) *Univ. California, Davis, Davis, California*, 151 pp.

(8) Qiu, D., Shenkin, P. S., Hollinger, F. P., and Still, W. C. (1997) The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii. *J. Phys. Chem. A 101*, 3005-3014.

CHAPTER 4

Identification of 4-pyridoxic acid dehydrogenase gene in *Mesorhizobium loti* MAFF303099*

4.1 Introduction

Only a few bacteria have been identified that are capable of surviving on vitamin B_6 as a sole source of carbon and nitrogen (*1*) and two catabolic routes for vitamin B_6 degradation have been identified (*2*), *figure 1.3*. Each of the catabolic intermediates have been isolated, characterized and the enzymes which produce them have been identified over the past few decades (*2*). At the start of this project four genes in *Mesorhizobium loti* MAFF303099 involved in PLP catabolism had been identified. The genes that were identified encoded pyridoxine-4-oxidase (mlr6785) (*3*), 4-pyridoxolactonase (mlr6805) (*4*), pyridoxal-4-dehydrogenase (mlr6807) (*5*) and 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (mlr6788) (*6*). The genes that have been identified to date are clustered on the chromosome. Inspection of the nearby genes revealed two putative dehydrogenases mlr6793 and mlr6792. The former, mlr6793 was annotated as 3-hydroxybutyryl-CoA dehydrogenase (*7*). Both these genes looked to be the likely candidate for the oxidation of 4-pyridoxic acid (**1**) to 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**), *figure 4.1*. In this chapter, we report the cloning and overexpression of mlr6793 and demonstrate that the purified gene product has a dual dehydrogenase activity catalyzing the 4-electron oxidation of 4-pyridoxic acid (**1**) to 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**).

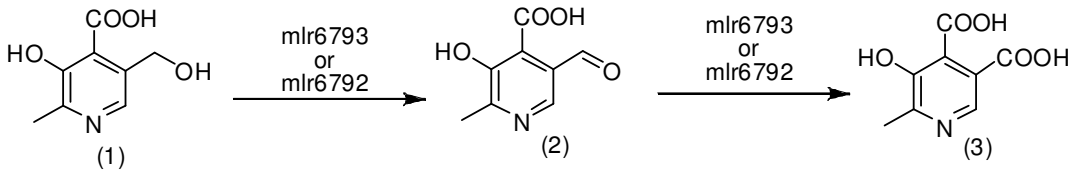


Figure 4.1 Reaction which could be catalyzed by mlr6793 or mlr6792.

* Reproduced in parts with permission from Mukherjee, T., Kinsland, C. K., and Begley, T.P. (2007). *Bioorganic Chemistry,* **35**, 458-464. © [2007] Elseiver. Molecular cloning was done by Dr. Kinsland.

4.2 Experimental

4.2.1. Molecular Cloning.[†] Standard methods were used for DNA manipulations (*8, 9*). Plasmid DNA was purified with the Qiagen Miniprep kit and DNA fragments were purified from agarose gel with the Nucleospin Purification kit. *E. coli* strain MachI was used as a recipient for transformations during plasmid construction and for plasmid propagation. Phusion DNA polymerase was used for PCR following the manufacturer's recommendations. The pENTR-TEV-D-TOPO and the Gateway system were used following the manufacturer's instructions with slight modifications.

4.2.2. Cloning of* M. loti *mlr6793: The *M. loti* mlr6793 gene was amplified from genomic DNA by PCR with the following primer pair: 5'-CAC CAT GAT CCG AAA TAT CGC CAT CAT C-3' and 5'-CTA CTC CCG GCC TTC CAG GAT GCG TC-3'. The PCR product was purified and used in a topoisomerase mediated reaction with pENTR-TEV-D-TOPO essentially following the manufacturer's instructions. Clones were screened by PCR and verified by sequencing. A correct clone was used in an LR recombination reaction with the plasmid pDESTF1, which is a Gateway adapted vector based on the pET-system. The plasmid pDESTF1 encodes an N-terminal 6xHis tag and is under the control of the T7lac promoter. Clones were screened by restriction digestion. A correct clone was named pMl5337.XF1.

4.2.3. Overexpression and purification. The plasmid pMl5337.XF1 was used to transform *Escherichia coli* BL21(DE3). A starter culture was prepared by growing a single colony of transformed cells in 10 ml of LB media containing 100 μg/ml of ampicillin at 37°C with overnight agitation. 1 liter LB medium (20 g/L), containing 100 μg/ml of ampicillin, was inoculated with this starter culture. The cells were grown at 37°C with

† The molecular cloning was done by Dr. Cynthia Kinsland

agitation until the culture reached an OD_{590} of 0.6 at which point they were induced by adding IPTG to a final concentration of 0.8 mM, the temperature was lowered to 15°C and the cells were allowed to grow for a further 12 hours. The cells were then harvested by centrifugation at 10000 xg for 8 min at 4°C.

Cells from 1 liter of culture were re-suspended in 20 ml of binding buffer (50 mM NaH_2PO_4, 150 mM NaCl, 10 mM imidazole, pH 7.7) and approximately 2 mg of lysozyme was added. The cells were then lysed by sonication (Misonix Sonicator 3000, pulse 'on' time 1.0 sec, pulse 'off' time 1.0 sec, output level 0.8, 30 cycles) 5 times on ice. The cell debris was removed by centrifugation at 30,000 xg for 40 minutes at 4°C. The clarified supernatant was loaded on to a 2 ml Ni-NTA-affinity column pre-equilibrated with binding buffer kept at 4°C. The Ni-NTA-affinity column was then washed with 100 ml wash buffer (50 mM NaH_2PO_4, 150 mM NaCl, 20 mM imidazole, pH 7.7). The protein was eluted from the column with elution buffer (50 mM NaH_2PO_4, 150 mM NaCl, 200 mM imidazole, pH 7.7) at 4°C. The fractions containing protein were pooled and concentrated using YM-10 Amicon Ultracentrifugal filters at 5000 xg to a final volume of 500 μl. The concentrated sample was desalted into 100 mM phosphate buffer at pH 8.0 using an Econo-Pac 10DG disposable chromatography column. The yield of the purified protein was 6 mg/liter.

4.2.4. Enzymatic assay based on NADH production. The reaction mixture consisted of: enzyme (5 μM), NAD (1 mM) and 4-pyridoxic acid (**1**) (50 μM) in 100 mM potassium phosphate buffer at pH 8.0. The enzymatic activity was monitored by measuring the increase in absorbance at 340 nm due to the production of NADH. A control reaction without NAD showed negligible background activity, *figure 4.2*.

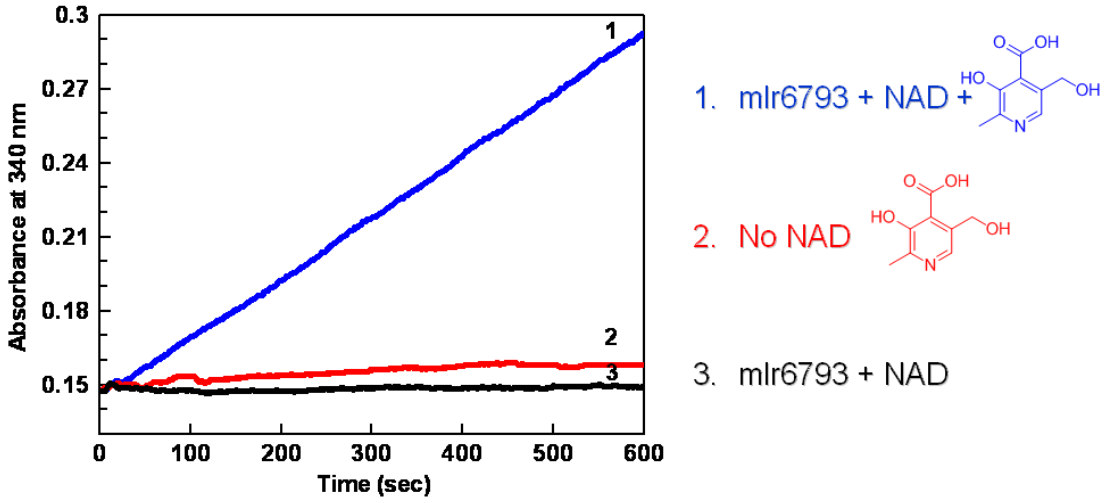


Figure 4.2: Assay for mlr6793 gene product. NADH is only formed (monitored by absorbance change at 340 nm, λmax for NADH) when all the components of the reaction mixture are present, as mentioned in section 4.2.4.

4.2.5. HPLC analysis. HPLC analysis of the enzymatic reaction mixture was performed on a Hewlett-Packard 1100 instrument using a Supelcosil LC-18-T (15 cm X 4.6 mm, 3.0 μM) column. Solution A contained water with 0.1% TFA, and solution B contained methanol with 0.1% TFA. The following linear gradient mixing solution A with solution B was used: 100% solution A for 0 to 2 min, 100% to 80% solution A from 2-12 min, 80% to 0% solution A in 12-13 min, 0% to 100% solution A in 13-14 min and 100% solution A in 14-17 minutes. Flow rate was 1 ml/min, absorbance was measured at 254 nm (characteristic for NAD) and 320 nm (characteristic of 4-pyridoxic acid, **1**). Under these conditions the following compounds were readily separated (retention time in parenthesis): 4-pyridoxic acid (**1**, 9.8 min), 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**, 7.35 min), NAD (8.75 min).

4.2.6. Reaction time course. A time course was determined with a reaction mixture (1 mL) containing 17 mM NAD, 1.5 mM 4-pyridoxic acid (**1**) and 150 μM of pure enzyme. At various time points, 100 μl of the reaction mixture was quenched by addition to 100 μl of 10% TFA. This mixture was filtered through Microcon YM-10 and 100 μl of the filtrate was analyzed by HPLC.

4.2.7. Product purification. A reaction mixture (5.5 ml) containing 78 mg of NAD (20 mM), 6 mg of 4-pyridoxic acid (**1**) (6 mM) and freshly purified enzyme (final concentration of 50 μM) was incubated overnight at room temperature. It was then filtered through a YM-10 Microcon centrifugal filter at 14,000 xg for 30 minutes and the filtrate was analyzed by HPLC. The enzymatic product (retention time 7.35 minutes) was collected over multiple injections. Methanol was removed by rotary evaporation; TFA and water were removed under high vacuum overnight. The resulting white powder was characterized by NMR and ESI-MS analysis.

4.2.8. Steady state kinetic analysis. The steady state kinetic parameters for 4-pyridoxic acid dehydrogenase were determined by monitoring the absorbance at 370 nm over time. To a reaction mixture (500 μl) containing 243 nM enzyme and 400 μM of NAD, varying concentration of 4 –pyridoxic acid (**1**) was added. The rate of formation of NADH was monitored over 15 minutes at 370 nm for each concentration of the substrate. The extinction coefficient of NAD in 100 mM phosphate buffer at pH 8.0 at 370 nm was determined to be 2600 $M^{-1}cm^{-1}$. The K_M for 4-pyridoxic acid (**1**) and k_{cat} for the production of NADH by the enzyme was determined by fitting the rate of product formation versus substrate concentration using non-linear regression to the Michaelis-Menten equation using Grafit 5.0.11 (Erithacus Software Limited, Surrey, UK). In order to determine the K_M for NAD a similar experiment was done in which the reaction mixture (500 μL)

contained 243 nM enzyme and 100 μM of 4-pyridoxic acid (**1**) and the concentration of NAD was varied. All solutions were made in 100 mM phosphate buffer at pH 8.0.

4.3 Results and Discussion

4.3.1. Reaction time course. The formation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**) and the disappearance of 4-pyridoxic acid (**1**) as a function of time is shown in *figure 4.3*. Neither 320 nm nor 254 nm wavelength showed the presence of any other peaks which indicated that under the reaction conditions, the two electron oxidation of 4-pyridoxic acid (**1**) to 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylate (**2**) was not observed. Under the acidic conditions used to quench the enzymatic reactions NADH is degraded to a non-chromophoric compound at 320 nm and is therefore not detected in the HPLC-chromatogram, monitored at 320 nm.

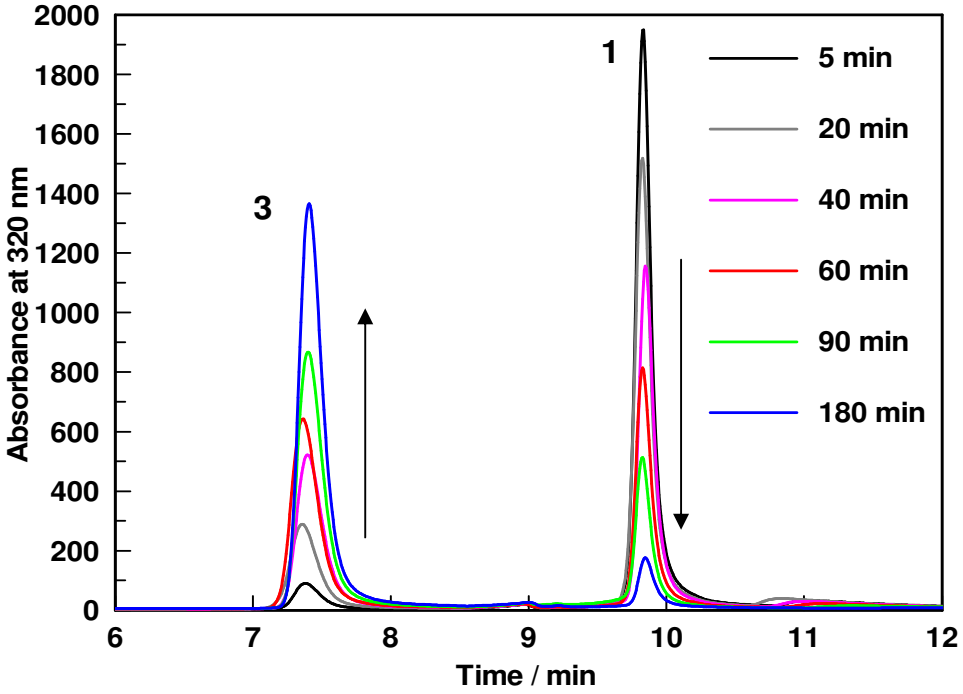


Figure 4.3: HPLC trace showing the disappearance of the reactant (retention time of 9.8 minutes) and the appearance of the product (retention time of 7.35 minutes).

4.3.2. Product characterization. The white colored compound obtained as a result of the enzymatic reaction mixture was identified as 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**) by NMR and MS analysis. ^{1}H NMR (300 MHz, D_2O) δ 2.69 (s, 3H, CH_3) and 8.42 (s, 1H, C_6-H), *figure 4.4*. Negative ion mode ESI-MS (Esquire-LC_00146 instrument, Bruker) showed a species with m/z = 196, which corresponded to the calculated mass of mono anionic 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**). Fragmentation analysis resulted in the formation of a species with m/z =152 (single decarboxylation, -44), which when isolated and fragmented further, showed the formation of the species with m/z = 108 (double decarboxylation, -88).

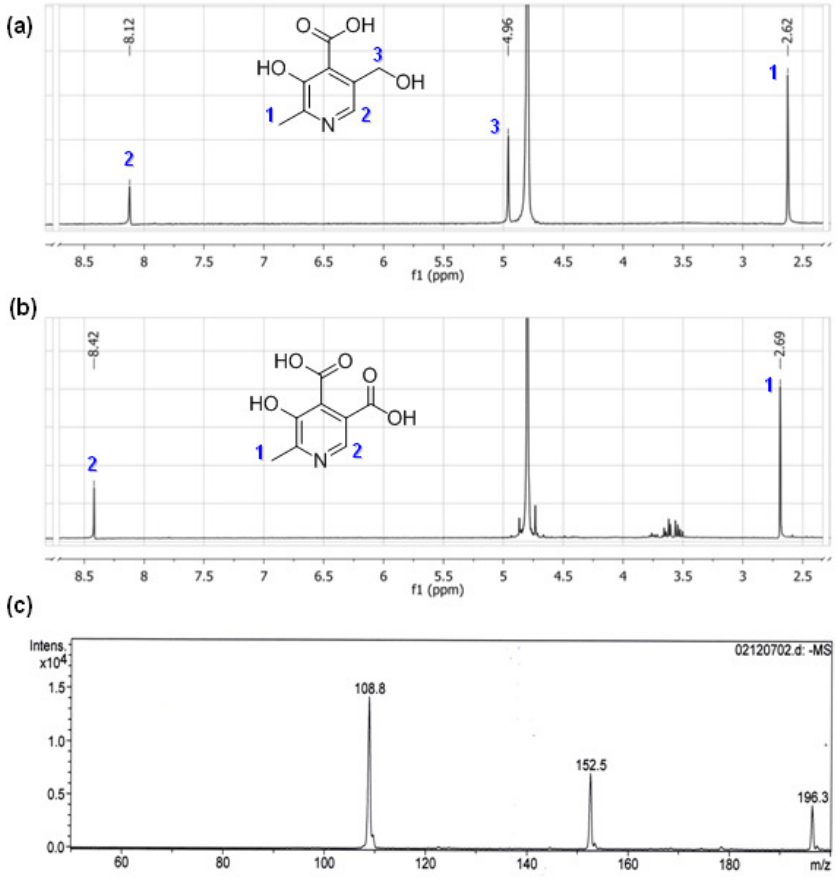


Figure 4.4: (a) ^{1}H NMR of the substrate, 4-pyridoxic acid (**1**). (b) ^{1}H NMR of the product, purified by HPLC, i.e. 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**). (c) Negative mode ESI-MS analysis of the product (**3**).

4.3.3. Steady state kinetic analysis. The steady state kinetic parameters for the enzyme were determined by monitoring the production of NADH in a UV-visible spectrophotometer. The k_{cat} for the NADH production is $0.01s^{-1}$. The K_M for 4-pyridoxic acid (**1**) (under saturating concentration of NAD) is 5.8 (± 0.9) μM. The K_M for NAD (under saturating concentration of 4-pyridoxic acid, **1**) is 6.6 (± 1.3) μM, *figure 4.5*. The absorbance for NADH production was monitored at 370 nm instead of 340 nm (λ_{max} for NADH) since both 4-pyridoxic acid (**1**) and 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**) have some absorbance at 340 nm.

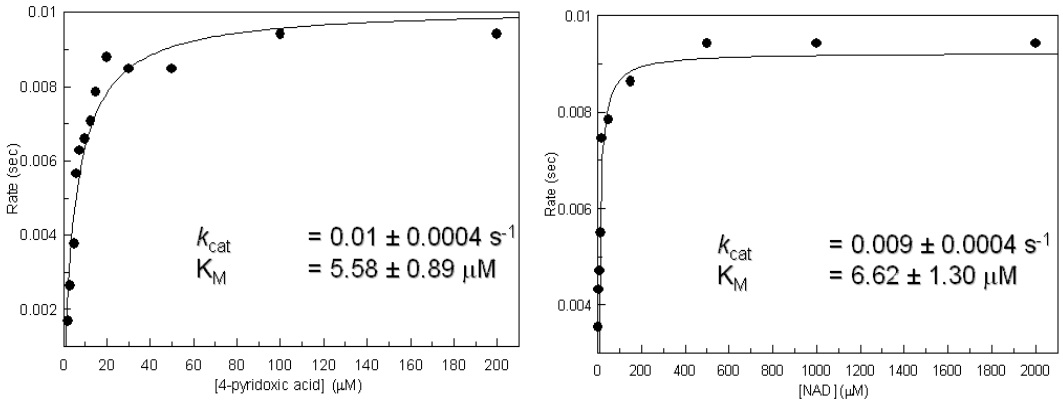


Figure 4.5: The steady state kinetic parameters of mlr6793 gene product were determined by measuring the rate of NADH formation over time for different substrate concentration. The experimental parameters have been detailed out in section 4.2.8.

4.4. Conclusion.

We have demonstrated that the mrl6793 gene product catalyzes the 4-electron oxidation of 4-pyridoxic acid (**1**) to 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**) using NAD as a cofactor. The intermediate 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylate (**2**) did not accumulate in the enzymatic reaction mixture. This implies that the enzyme catalyzes the oxidation of (**2**) to (**3**), faster than the oxidation of (**1**) to (**2**). Since we see no reduction of NAD in absence of the substrate and the reaction product has been

unambiguously identified, it is reasonable to assume that two moles of NAD are consumed for each mole of substrate oxidized. Sequence analysis indicates only one nucleotide binding site so we assume that both hydride transfer reactions are occurring at the same active site. The mrl6793 encoded pyridoxic acid dehydrogenase is different from the previously characterized enzyme from *Pseudomonas sp.* MA-1 which uses FAD as a cofactor and only catalyzes the 2-electron oxidation of 4-pyridoxic acid (**1**) to 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylate (**2**). FAD is not required by the mrl6793 encoded pyridoxic acid dehydrogenase that has been characterized in this paper.

Later work by Yagi et al. showed that the mlr6792 actually coded for 4-pyridoxic acid dehydrogenase which catalyzed the two electron oxidation of 4-pyridoxic acid (**1**) to the 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylate (**2**), which in turn undergoes a further 2-electron oxidation to form the 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**3**), catalyzed by 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylic acid dehydrogenase (mlr6793) (*10*). However, no structural evidence for the reaction products has been provided and would hopefully appear in due course of time.

4.5 Acknowledgement

We are grateful to Dr. Yasunobu Ohkawa, Director of Genebank, National Institute of Agrobiological Sciences, Ibaraki, Japan for providing us with the *M. loti* MAFF303090 strain. This research was supported by a grant from the National Institutes of Health (GM069618).

REFERENCES

(1) Rodwell, V. W., Volcani, B. E., Ikawa, M., and Snell, E. E. (1958) Bacterial oxidation of vitamin B6. I. Isopyridoxal and 5-pyridoxic acid. *J Biol Chem 233*, 1548-1554.

(2) Snell, E. E., and Haskell, B. E. (1971) *The Metabolism of Vitamin B6, In: Comprehensive Biochemistry*, Vol. 21, Elsevier/North Holland, New York.

(3) Yuan, B., Yoshikane, Y., Yokochi, N., Ohnishi, K., and Yagi, T. (2004) The nitrogen-fixing symbiotic bacterium Mesorhizobium loti has and expresses the gene encoding pyridoxine 4-oxidase involved in the degradation of vitamin B6. *FEMS Microbiol Lett 234*, 225-30.

(4) Funami, J., Yoshikane, Y., Kobayashi, H., Yokochi, N., Yuan, B., Iwasaki, K., Ohnishi, K., and Yagi, T. (2005) 4-Pyridoxolactonase from a symbiotic nitrogen-fixing bacterium Mesorhizobium loti: cloning, expression, and characterization. *Biochim Biophys Acta 1753*, 234-9.

(5) Yokochi, N., Nishimura, S., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2006) Identification of a new tetrameric pyridoxal 4-dehydrogenase as the second enzyme in the degradation pathway for pyridoxine in a nitrogen-fixing symbiotic bacterium, Mesorhizobium loti. *Arch Biochem Biophys 452*, 1-8.

(6) Yuan, B., Yokochi, N., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2006) Molecular cloning, identification and characterization of 2-methyl-3-hydroxypyridine-5-carboxylic-acid-dioxygenase-coding gene from the nitrogen-fixing symbiotic bacterium Mesorhizobium loti. *J Biosci Bioeng 102*, 504-510.

(7) http://theseed.uchicago.edu/FIG/index.cgi.

(8) Sambrook, J., Fritsch, G. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Guide*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

(9) Ausubel, F. M., and Brent, F. (1987) in *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.

(10) Ge F, Y. N., Yoshikane Y, Ohnishi K, Yagi T. (2008) Gene Identification and Characterization of the Pyridoxine Degradative Enzyme 4-Pyridoxic Acid Dehydrogenase from the Nitrogen-fixing Symbiotic Bacterium Mesorhizobium loti MAFF303099. *J Biochem. 143*, 603-609.

CHAPTER 5

Gene identification and structural characterization of the PLP degradative protein 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase from *Mesorhizobium loti* MAFF303099*

5.1 Introduction

In the PLP catabolic pathway 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**), the product of 4-pyridoxic acid dehydrogenase, undergoes a decarboxylation to form 3-hydroxy-2-methylpyridine-5-carboxylate, (**2**), *figure 5.1*. The gene coding for the enzyme producing this metabolite has not previously been discovered and is the subject of this chapter.

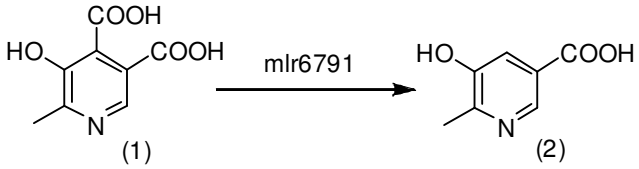


Figure 5.1: Reaction catalyzed by mlr6791 gene product.

All the genes participating in PLP catabolism that have been identified in *Mesorhizobium Loti* are not part of an operon, but are all close to each other on the *M. loti* chromosome. This suggested that other PLP catabolic genes might also be found in this region. In particular, we felt that the mlr6791gene, annotated as ribulose-5-phosphate 4-epimerase (*1*), was a likely candidate for the decarboxylase gene because the epimerase and the decarboxylase catalyzed reactions both proceed via enzyme stabilized enolate intermediates. In this chapter, the cloning and overexpression of mlr6791 is reported, the purified gene product is shown to catalyze the decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) and the structure of this enzyme has been described.

---

* Reproduced in parts with permission from Mukherjee, T., Kathryn M. McCulloch, K. M., Ealick, S. E., and Begley, T.P. (2007). *Biochemistry,* **46**, 13606-13615. © [2007] American Chemical Society
All crystallographic studies and analysis were done by Kathryn.

While the decarboxylation of hydroxy-substituted benzene rings is a common motif in biosynthesis, the mechanism of this reaction is still poorly characterized. The structural studies described here suggest that catalysis of such decarboxylations proceeds by an aldolase-like mechanism.

5.2 Experimental.

***5.2.1. Molecular Cloning*[‡].** Standard methods were used for DNA manipulations (*2,3*). Plasmid DNA was purified with the Qiagen Miniprep kit and DNA fragments were purified from agarose gel with the Nucleospin Purification kit. Escherichia coli strain Mach1 was used as a recipient for transformations during plasmid construction and for plasmid propagation. Phusion DNA polymerase was used for PCR following the manufacturer's recommendations. The pENTR-TEV-D-TOPO and the Gateway system were used following the manufacturer's instructions with slight modifications.

5.2.2. Cloning of* M. loti *mlr6791. The *M. loti* mlr6791 gene was amplified from genomic DNA by PCR with the following primer pair: 5'-CAC CAT GCG TCG GAA GGT CTT CGA AGA G-3' AND 5'-TCA GGC GAG GCC TGC TTG CCT GAG G-3'. The PCR product was purified and used in a topoisomerase mediated reaction with pENTR-TEV-D-TOPO essentially following the manufacturer's instructions. Clones were screened by PCR and verified by sequencing. In initial sequencing, no clones were found with a completely correct sequence. One clone was subjected to standard site-directed mutagenesis with the following complementary primer pair: 5'-GGA TAC GTT CGG GCA CAT ATC TGC CCG TGA CCC CGA G-3' and 5'-CTC GGG GTC ACG GGC ACA TAT GTG CCC GAA CGT ATC C-3'. Colonies from the mutagenesis were screened with the M13-reverse primer and the mutant specific primer 5'-GAT ACG TTC

‡ Molecular cloning was done by Dr. Cynthia Kinsland.

GGG CAC ATA TC-3'. PCR positive clones were sequenced and a correct clone was used in an LR recombination reaction with the plasmid pDESTF1, which is a Gateway adapted vector based on the pET-system. The plasmid pDESTF1 encodes an N-terminal 6xHis tag and is under the control of the T7/lac /promoter. Clones were screened by restriction digestion. A correct clone was named pMl5335.XF1.

5.2.3. Overexpression and Purification. The plasmid pMl5335.XF1 was used to transform *E. coli* BL21(DE3). The cells were grown in 1L LB medium, containing 100 μg/mL of ampicillin, at 37 °C with agitation until the culture reached an OD_{590} of 0.6, at which point overexpression was induced by adding IPTG to a final concentration of 0.5 mM, the temperature was lowered to 15 °C and the cells were allowed to grow for a further 12 h before being harvested. The protein was purified by Ni-affinity chromatography following the purification procedure described in section 4.2.1.

For crystallographic studies[†], the selenomethionine (SelMet) protein was prepared by transforming the plasmid pMl5335.XF1 into *E. coli* B834(DE3) cells, a strain auxotrophic for methionine. Cells were grown at 37 °C with shaking in minimal M9 media which was supplemented with 20 mg/L of all amino acids except methionine, 1X MEM vitamin mix, 0.4% glucose, 50 mg/L L-selenomethionine, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 25 mg/L $FeSO_4$, and 100 mg/L ampicillin. An identical purification procedure was followed. In both preparations, SDS-PAGE analysis showed 95% pure protein and the yield of the purified protein was 20 mg/liter. The protein concentration was measured by the Bradford assay (*4*).

5.2.4. HPLC analysis. HPLC analysis of the enzymatic reaction mixture was performed on a Hewlett-Packard 1100 instrument using a Supelcosil LC-18-T (15 cm X 4.6 mm,

[†] All crystallographic studies and analysis were done by Kathryn McCulloch

3 μM) column. Solution A contained water, solution B contained 100 mM sodium phosphate buffer at pH 6.6 and solution C contained methanol. The following linear gradient was used: 0% to 10% solution A and 100% to 90% solution B for 0 to 5 min, 10% to 48% solution A, 90% to 40% of solution B and 0% to 12% of solution C from 5-12 min, 48% to 50% solution A, 40% to 30% of solution B and 12% to 20% of solution C in 12-14 min, 50% to 30% solution A, 30% to 10% of solution B and 20% to 60% of solution C in 14-18 min, 30% to 0% solution A, 10% to 100% of solution B and 60% to 0% of solution C in 18-20 min and 0% of solution A, 100% of solution B and 0% of solution C in 20-25 minutes. The flow rate was 1 mL/min and the absorbance was measured at 254 nm (characteristic for NAD, 3-hydroxy-2-methylpyridine-5-carboxylate (**3**)) and 320 nm (characteristic of 4-pyridoxic acid (**3**), 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) and 3-hydroxy-2-methylpyridine-5-carboxylate, (**2**)). Under these conditions the following compounds were readily separated (retention time in parenthesis): 4-pyridoxic acid (**3**), (17.8 min), 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) (2.8 min), 3-hydroxy-2-methylpyridine-5-carboxylate (**2**) (5.5 min) and NAD (12.9 min).

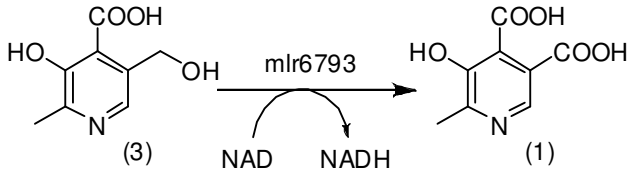


Figure 5.2: Reaction catalyzed by 4-pyridoxic acid dehydrogenase, mlr6793.

5.2.5. Enzymatic Assay. The substrate, 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) was not available commercially. A coupled reaction was set up with 4-pyridoxic acid dehydrogenase, mlr6793 whose product is compound (**1**), *figure 5.2*. The reaction mixture consisted of: 4-pyridoxic acid dehydrogenase (2 μM), 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase (HMPDdc) (1 μM), NAD (1 mM), $MnCl_2$ (10 μM) and 4-

pyridoxic acid (**3**) (100 μM) in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT. It was incubated at room temperature for 30 minutes, filtered using Microcon YM-10 to remove the enzyme, and analyzed by HPLC. The analysis was performed at 320 nm as well as 254 nm in order to see each component of the reaction mixture. A control reaction without mlr6791 gene product was set up and analyzed in an identical fashion, *figure 5.3.*Unreacted 4-pyridoxic acid (3) and NAD as well as NADH formed due to the 4-pyridoxic acid catalyzed reaction is also seen. In the control reaction, where no decarboxylase is present, one can see the product of 4-pyridoxic acid dehydrogenase, mlr6793, ie 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**). The assay clearly shows the conversion of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) to a new species that eluted at 5.5 min.

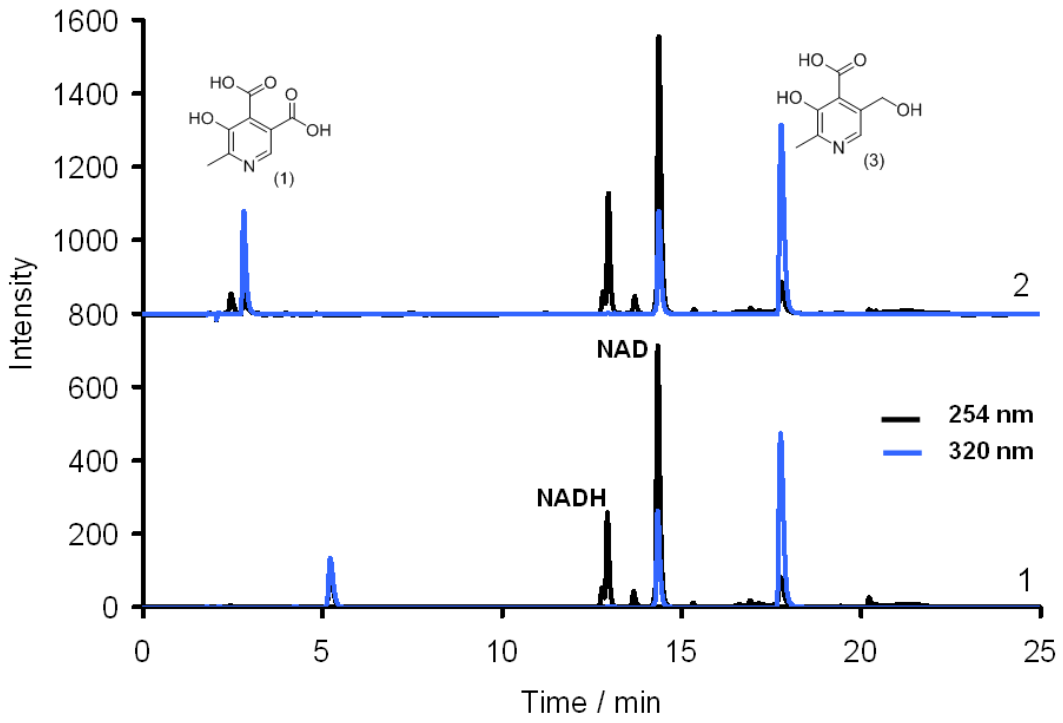


Figure 5.3: HPLC analysis of the assay mixture with (1) and without (2) the decarboxylase enzyme, mlr6791. The analysis was performed at 320 nm as well as 254 nm in order to see each component of the reaction mixture.

5.2.6. Enzymatic synthesis of the substrate. 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) was enzymatically synthesized from 4-pyridoxic acid (**3**), *figure 5.2*. The 5 mL enzymatic reaction mixture containing 12 mM NAD, 6 mM 4- pyridoxic acid (**3**) and 100 μM freshly purified 4-pyridoxic acid dehydrogenase (*5*) in 100 mM sodium phosphate buffer at pH 8.0 was incubated overnight at room temperature. The reaction mixture was subsequently concentrated by lyophilizing and redissolving in a minimum volume of water. It was then filtered through YM-10 Microcon centrifugal filter at 14,000 *g* for 30 min to remove the protein and the filtrate was purified by HPLC. 100 mM Ammonium acetate, pH 6.6, was used instead of 100 mM sodium phosphate buffer at pH 6.6 as solution B, to facilitate removal of the buffer salts from the isolated 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) during lyophilization. Compound (**1**) is a stable white solid. ^{1}H NMR (300 MHz, D_2O) δ 2.69 (s, 3H, CH_3) and 8.42 (s, 1H, C_6-H).

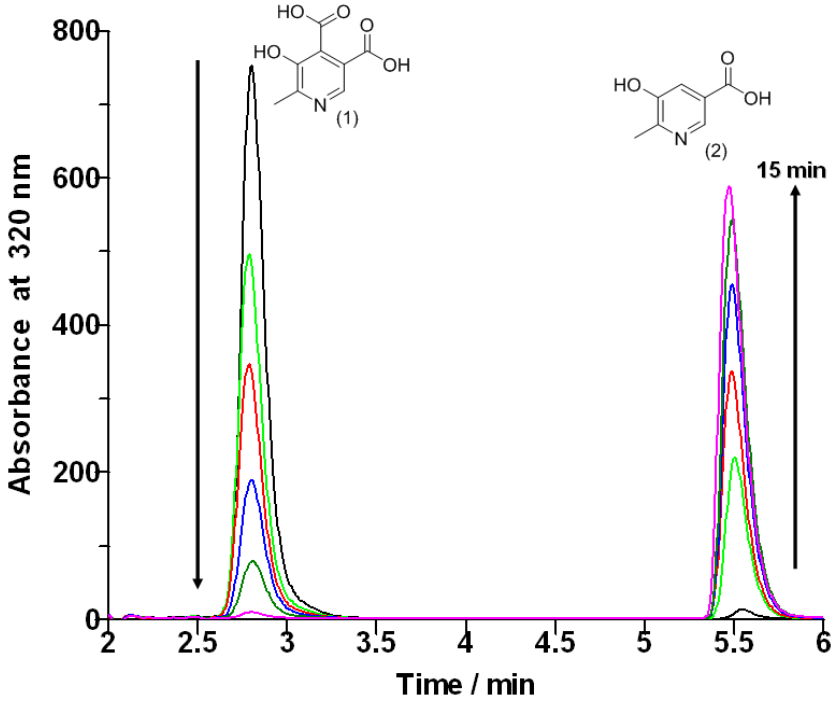


Figure 5.4: HPLC trace showing the disappearance of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) (retention time of 2.8 minutes) and the appearance of 3-hydroxy-2-methylpyridine-5-carboxylate (**2**) (retention time of 5.5 min).

5.2.7. Reaction time course. A time course was determined with a reaction mixture (1 mL) containing 1.58 mM 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**), 2 μM pure HMPDdc and 10 μM $MnCl_2$ in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT (*6*). At various time points, 100 μL of the reaction mixture was quenched by addition to 100 μL of 10% TFA. This mixture was filtered through Microcon YM-10 and 100 μL of the filtrate was analyzed by HPLC, *figure 5.4*.

5.2.8. Product purification and characterization. A reaction mixture (10.0 ml) containing 12 mM NAD, 6 mM 4-pyridoxic acid (**4**), 100 μM 4-pyridoxic acid dehydrogenase, 100 μM of HMPDdc and 1mM $MnCl_2$ in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT was incubated overnight at room temperature. It was then filtered through an YM-10 Microcon centrifugal filter at 14,000g for 30 min. The desired enzymatic product (retention time 5.5 min) was purified by HPLC over multiple injections. Methanol was removed by rotary evaporation; TFA and water were removed under high vacuum overnight. The resulting white powder was characterized by NMR and ESI-MS. ^{1}H NMR (300 MHz, D_2O) δ 2.46 (s, 3H, CH_3), 7.68 (s, 1H, C4-H) and 8.12 (s, 1H, C6-H). ESI-MS (Esquire-LC_00146 instrument, Bruker, Negative ion mode) m/z = 152, (mono anionic 3-hydroxy-2-methylpyridine-5-carboxylate (**2**)). Fragmentation analysis resulted in the formation of a species with m/z = 108 (M-44, decarboxylation of (**2**)).

5.2.9. Steady state kinetic parameters. The steady state kinetic parameters for HMPDdc were determined by monitoring the change in absorbance at 265 nm over time. To a reaction mixture (500 μL) containing 1 μM enzyme and 5 μM $MnCl_2$, varying concentrations of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) were added. The rate

of formation of 3-hydroxy-2-methylpyridine-5-carboxylate (**2**) was monitored over 3 min at 265 nm for each concentration of the substrate. The K_M and k_{cat} for the enzyme were determined by fitting the rate of product formation as a function of substrate concentration using non-linear regression to the Michaelis-Menten equation using Grafit 5.0.11 (Erithacus Software Ltd., Surray, UK). All solutions were made in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT.

5.2.10. Protein Crystallization. The hanging drop vapor diffusion method was used with 1 μL of SelMet protein solution and 1 μL of reservoir solution at 22 °C. Sparse matrix screens were used to determine initial hits (Crystal Screen and Crystal Screen 2, Hampton Research). Optimized crystallization conditions consisted of 6-9% polyethylene glycol 8000 and 100 mM Tris buffer with a pH ranging from 7.0 to 7.5. Crystals grew in a conical shape to a size of approximately 0.3 mm x 0.1 mm in roughly one week and were cryoprotected by a quick transfer into crystallization conditions with an additional 17% glycerol. SelMet-HMPDdc crystals were then flash frozen by plunging in liquid nitrogen and stored frozen until data collection. Crystals belong to the space group *I4* with unit cell parameters a = 72.0 Å and c = 90.4 Å. The unit cell contains one monomer per asymmetric unit with a solvent content of 45% and a Matthews coefficient of 2.24 $Å^3$/Da (*7*).

5.2.11. X-Ray Data Collection and Processing. A single wavelength anomalous diffraction dataset was collected on a SelMet-HMPDdc crystal at the NE-CAT 24-ID-C beamline at the Advanced Photon Source using a Quantum315 detector (Area Detector Systems Corp). The dataset was collected at the maximum f'' for selenium as determined using a fluorescence scan of the SelMet-HMPDdc crystal. The crystal diffracted to 1.9 Å

resolution and data was collected over 360° using a 1° oscillation range. The HKL2000 suite of programs was used to index, integrate, and scale the data (*8*).

5.2.12. Structure Determination, Model Building, and Refinement. The computer program hkl2map was used to determine the positions of the Se atoms using data cut off at 2.2 Å resolution to maximize the anomalous signal (*9*). Three of the possible five selenium atoms were located. The program autoSHARP was used for refinement of the heavy atom positions, phasing, calculation of residual maps, density modification, and automated model building (*10*). Automated model building built 222 of 234 residues with the correct side chain. Examination of the density modified maps allowed the manual building of residues 139-141, 209-216, and 234 using COOT (*11*). Refinement of the HMPDdc model was then performed using CNS (*12*). A metal binding site was identified and modeled as a manganese ion, based on the magnitude of the electron density, coordination geometry, and the results of a fluorescence scan performed (results not shown). Water molecules were added as refinement continued also using CNS. The HMPDdc structure was verified using PROCHECK and no residues were located in disallowed regions of the Ramachandran plot (*13*). Figures were generated using Pymol (*14*).

5.3. Results

5.3.1. Product purification and characterization. The product of the enzymatic reaction was purified as a stable white powder. It was identified as 3-hydroxy-2-methylpyridine-5-carboxylate (**2**) by NMR and ESI-MS analysis, *figure 5.5*. The presence of a singlet at 7.68 ppm in the aromatic region, corresponding to the C-4 hydrogen of (**2**), indicated a clean conversion of substrate (**1**) to product (**2**). The time course for the HMPDdc catalyzed decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**)

to 3-hydroxy-2-methylpyridine-5-carboxylate (**2**), as analyzed by HPLC is shown in *figure 5.4*.

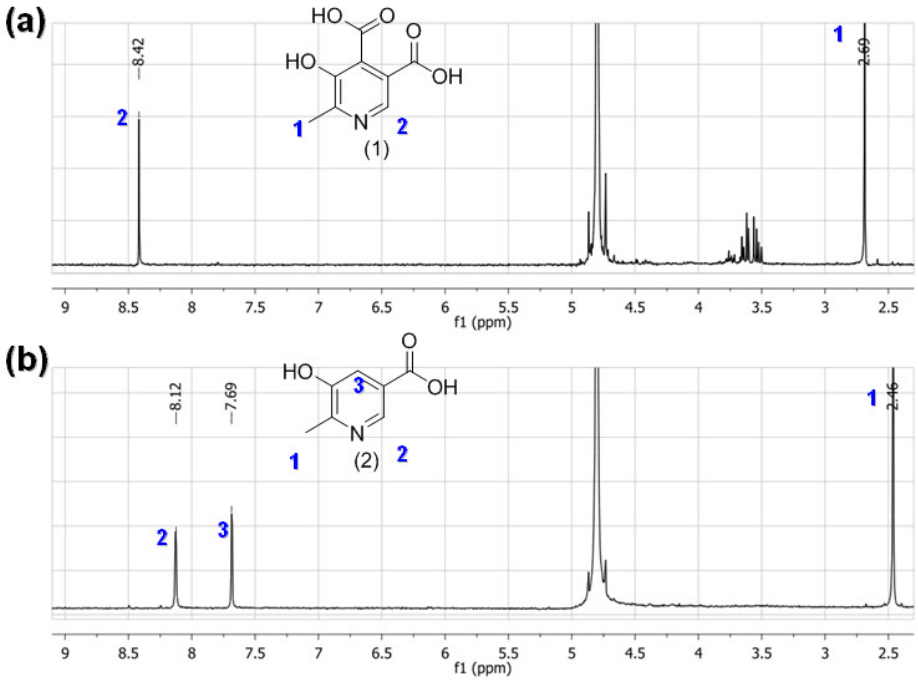


Figure 5.5: (a) NMR spectra of the substrate, 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**). (b) NMR spectra of the product, purified by HPLC, i.e. 3-hydroxy-2-methylpyridine-5-carboxylate (**2**).

5.3.2. Steady state kinetic parameters. UV visible spectra of the enzymatic reaction mixture (500 µL), containing 1.58 mM of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**), 4 µM $MnCl_2$ and 780 nM of the HMPDdc in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT, taken at various time points, *figure 5.6(a)*, showed an increase in absorbance at 265 nm. Both 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) and 3-hydroxy-2-methylpyridine-5-carboxylate (**2**) absorb at this wavelength. The extinction coefficients of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**) and 3-hydroxy-2-methylpyridine-5-carboxylate (**2**) at 265 nm in 100 mM Tris HCl at pH 8.0

containing 100 mM NaCl and 2 mM DTT were determined to be 1017 $M^{-1}cm^{-1}$ and 3623 $M^{-1}cm^{-1}$ respectively. The rate of product formation was determined using the difference in molar extinction coefficients ($\Delta\varepsilon_{265}$ = 2615 $M^{-1}cm^{-1}$). Steady state kinetic parameters were obtained from the concentration dependence of the rate of formation of 3-hydroxy-2-methylpyridine-5-carboxylate (**2**) at constant concentration of HMPDdc under saturating concentration of $MnCl_2$. The enzymatic reaction exhibited Michealis Menten kinetics with K_M and k_{cat} of 366 μM and 0.6 s^{-1} respectively. The k_{cat}/K_M for HMPDdc was determined to be 1530 $M^{-1}s^{-1}$, *figure 5.6(b).*

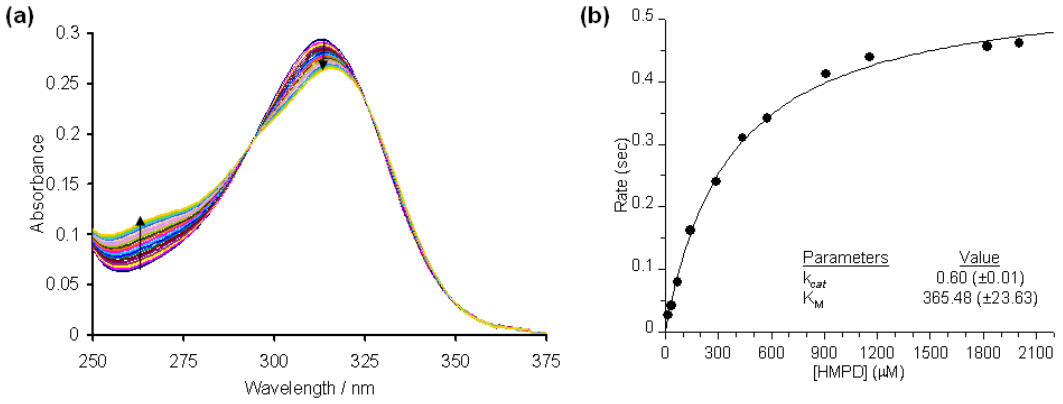


Figure 5.6: (a) UV visible spectra of the enzymatic reaction mixture (500 μL) containing 1.58 mM of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**), 4 μM $MnCl_2$ and 780 nM of the HMPDdc in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT were taken at 1 minute intervals over 20 minutes. (b) The steady state kinetic parameters for HMPDdc were determined by monitoring the absorbance at 265 nm over time.

5.3.3. Monomeric Structure of HMPDdc. The structure of HMPDdc was determined at 1.9 Å resolution using SAD phasing. All 234 residues of the protein and 3 residues from the 6X N-terminal His tag were modeled into the final structure of the monomer, as

well as 205 water molecules and a manganese ion. The monomer is composed of a single domain with an α/β/α fold, *figure 5.7*.

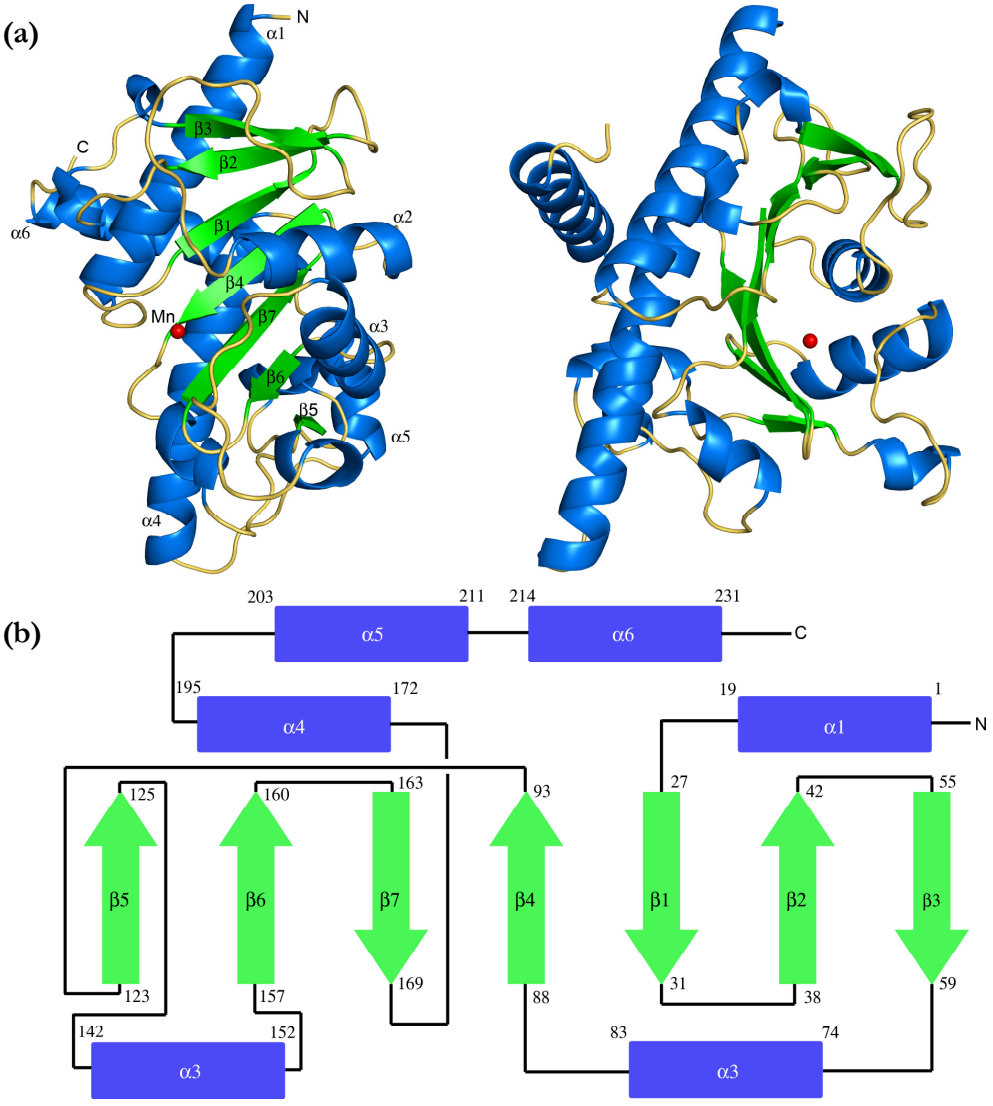


Figure 5.7: Monomeric structure of *Mt*HMPDdc. (a). Ribbon diagram of HMPDdc at two orientations 90° rotated from each other. The secondary structure is labeled and α-helices are colored in blue, β-strands are colored in green, and loop regions are yellow. The manganese ion is shown as a non-bonded red sphere. (b) Topology diagram of HMPDdc.

The central 7 stranded mixed β-sheet is mostly antiparallel with a strand order of β5↑β6↑β7↓β4↑β1↓β2↑β3↓ where only β5 and β6 run parallel to each other as seen in *figure 5.7(b)*. The β-sheet forms a half-barrel with four α-helices flanking one side of the β-sheet and two α-helices flanking the opposite side. Three of the α-helices, α1, α4 and α6, are unusually long with 19, 24 and 18 amino acids, respectively. There are also three 3_{10} helices.

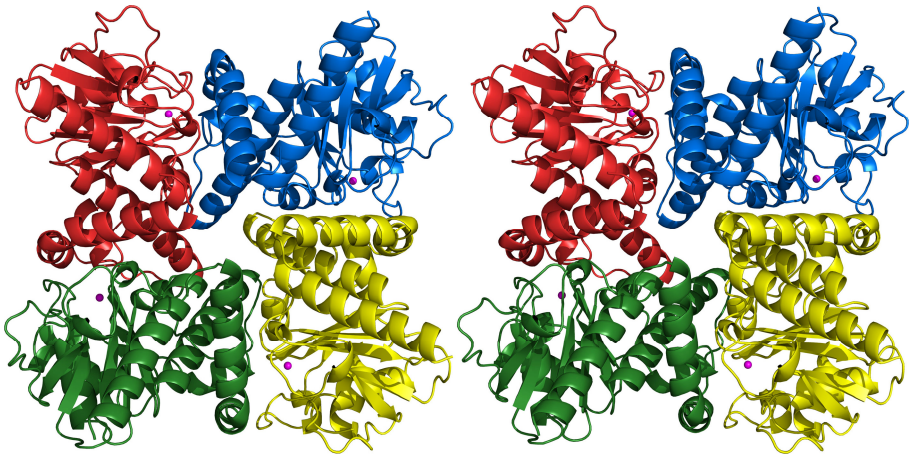


Figure 5.8: Stereoview diagram of the tetrameric structure of HMPDdc. The tetramer has been color coded by subunit. The manganese ion is shown in magenta.

5.3.4. Tetrameric Structure of HMPDdc. The quaternary structure of HMPDdc is a tetramer formed by using the fourfold crystallographic axis of the space group I4 and is shown in *figure 5.8*. The tetramer is roughly 65 Å wide, 45 Å tall and 80 Å across the diagonal. A channel with a diameter of 10 Å runs through the tetramer. α-Helices α1, α4, and α6 face the channel and it is composed mostly of hydrophilic side chains and backbone carbonyl groups. The opening is much wider on the top, near the N and C termini, and nearly closed on the bottom by the last turn of α4 and the side chain of Lys194. The subunit interface is formed by interactions between the C-terminal end of one HMPDdc subunit (α4, α5, α6, and the connecting loops) and the loops between β2

and β3, β6 and β7, and the loop between the third 3_{10} helix and α3 from a neighboring subunit. The first 3_{10} helix also contributes to the interface between subunits.

Approximately 3000 $Å^2$ of surface area is buried at the interface between subunits (*15*). The residues lining the subunit interface are mostly hydrophobic, including a patch rich in aromatic residues, including Phe25, Phe177, Phe180, and Tyr181. There are two sets of salt bridges at the interface, one between Arg128 and Glu205 and the second between Arg162 and Glu126. The arginine residues are found close to each other and there also appears to be some stacking between these side chains. The interface is also stabilized by five hydrogen bonds: Asp23 and Thr24 both hydrogen bonding to Tyr181, Ser48 to His223, and Asp105 to Asn197. Three backbone carbonyl groups form hydrogen bonds: the carbonyl group from Ile195 to the side chain of Thr106, the carbonyl group of Asp133 to the hydroxyl group of Ser202, and the carbonyl oxygen atom of Glu134 to the oxygen of the Thr208 side chain.

5.3.5. Metal Binding Site. A cleft 13 Å long, 8 Å wide, and 10 Å deep forms near the hydrophobic patch at the interface between two subunits. The cleft is found in the middle of the tetramer and lies closer to the external solvent than to the channel that runs through the tetramer. This cleft contains six histidine residues in close proximity to each other: His27, His92, His94, His163, and His177 from one subunit and His113* from the neighboring subunit.

In addition to the histidine residues, a large unexpected peak in the electron density was found in this cleft and was modeled as a manganese ion. No metal was added during purification, and treatment of the protein sample with EDTA was unsuccessful in removing the metal, showing that the metal is tightly bound in MlHMPDdc. A fluorescence scan identified the metal as a manganese ion. The metal has a tetrahedral coordination to His92, His94, His163, and Glu73, *figure 5.9*. The

bonding distance between the manganese ion and His92 is 2.2 Å, while the distances to the other residues coordinated to the metal are longer than expected with a distance of 2.7 Å for His94, His163 and Glu73. Adjacent to the bound metal is a pocket 8 Å long and 7 Å wide, which could potentially accommodate 3-hydroxy-2-methylpyridine-4,5-dicarboxylate 6 . Hydrophobic residues that protrude into and line the cleft include Phe25, Phe28, and Tyr71 and hydrophilic residues in the pocket include His113*, Gln216*, and Arg219*.

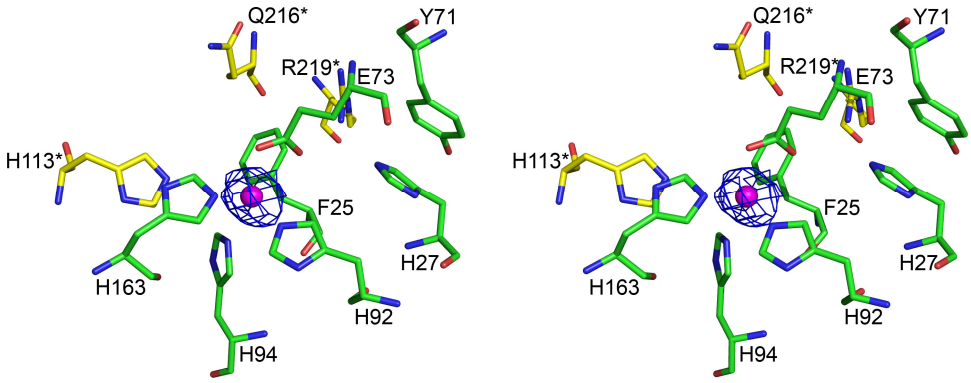


Figure 5.9: Stereoview diagram of the metal binding site. Composite omit density is shown around the manganese ion at a contour level of 1.0 σ and is shown in blue. The manganese ion is magenta. Residues shown in green are from one subunit and the residues shown in yellow are from the neighboring subunit.

5.4. Discussion

5.4.1. Comparison of HMPDdc to other proteins. An iterative BLAST (*16*) search using the non-redundant protein sequence database revealed a large cluster of proteins with sequence identities to HMPDdc ranging from 20 to 40%. Most of these proteins are annotated as class II aldolases/adducins or hypothetical proteins. Alignment of the top hits identified five strictly conserved residues and several highly conserved residues (*17, 18*). Pro47 is strictly conserved and is found at the interface between monomers as the

first residue of a 3_{10} helix in HMPDdc. Two of the residues coordinating to the manganese ion, Glu73 and His94, are also absolutely conserved. The two other histidine residues, His92 and His163, are mostly conserved among the proteins, with His92 being replaced by an arginine and His163 being replaced by asparagine in a class II aldolase from *Nocardioides sp. JS614*. His77 is strictly conserved, and is one of the six histidine residues found near the putative active site cleft at the interface. The final conserved residue, Gly164, is adjacent to His163, which coordinates to the manganese ion and is the first residue of a β-strand.

A BLAST search was then conducted on a class II aldolase/adducin-like protein from *Burkholderia sp. 383*, which had the highest sequence similarity to *Ml*HMPDdc in the initial BLAST search. Most of the sequences with significant alignments were the same class II aldolases/adducins and hypothetical proteins identified using HMPDdc as a search subject; however, at a lower similarity level several 3,4-dihydroxyphthalate 2-decarboxylase (DHPdc) sequences were identified. DHPdc is found in the phthalate catabolism pathway of gram-positive bacteria and catalyzes a decarboxylation reaction of an aromatic ring very similar to the reaction catalyzed by HMPDdc (*19, 20*). Alignment of these DHPdc sequences with the sequence of HMPDdc and the nearest class II aldolase identified several residues that are conserved (*18*). All metal coordinating residues are conserved and His77 and His27, two histidine residues found near the cleft between subunits, are also conserved. The aromatic residues near the active site, Phe25, Tyr71, His113, and Phe138, are not conserved in DHPdc but are replaced by hydrophobic residues. Gln216 in HMPDdc is found to be an asparagine residue in DHPdc and Arg219 is replaced by a threonine residue.

The *Ml*HMPDdc monomer was submitted to DALI to identify structurally related proteins (*21*). Five proteins were identified as being structurally similar to HMPDdc with Z-score greater than 10; all other proteins have Z scores less than 4. Top

hits included two L-fuculose 1-phosphate aldolases, one L-ribulose 5-phosphate epimerase, a rhamnulose 1-phosphate aldolase, and the C terminal fragment of phosphomethylpyrimidine kinase from *Pyrococcus furiosus* (*22-24*). The aldolases and epimerase belong to the superfamily of AraD-like proteins and class II aldolases. L-Fuculose 1-phosphate aldolase from *E. coli* showed the highest structural similarity with a Z score of 24.4 and an r.m.s.d. of 2.4 Å for 205 of a possible 210 residues. The sequence identity between HMPDdc and the aldolase is about 20%. The Z score for *E. coli* L-ribulose 5-phosphate epimerase was 22.5 with an r.m.s.d. of 2.4 Å for 207 of 223 residues.

The structures of the L-fuculose 1-phosphate aldolases, L-ribulose 5-phosphate epimerase, and the rhamnulose 1-phosphate aldolase superimpose well on the structure of HMPDdc, as seen in *figure 5.10*. The topology of the long β-sheet is conserved for each of the aldolases and HMPDdc; however, HMPDdc has a C-terminal α-helix not observed in the other structures. Rhamnulose 1-phosphate aldolase is the least similar to HMPDdc and has two extra β-strands and four extra α-helices. All of the enzymes shown in *figure 5.10* adopt a tetrameric oligomeric state.

Despite the strong structural similarity between HMPDdc and the class II aldolase family, there is low sequence conservation and apart from the coordination of the metal ion few of the putative active site residues are conserved. The class II aldolase superfamily members each bind a zinc ion in the active site, while HMPDdc has a manganese ion bound. L-fuculose 1-phosphate aldolase from *Bacterosides thetaiotaomicron* has the highest sequence identity at 20% and the sequence identity between HMPDdc and rhamnulose 1-phosphate aldolase from *E. coli* is only 15%. The zinc ion is found in the same cleft and coordinated by three conserved histidine residues. The fourth residue coordinated to the metal ion is a glutamate residue in HMPDdc and the three aldolases, and an aspartate residue in the epimerase. Several glycine and proline residues are

conserved, most likely playing roles in positioning structural elements. His77 is conserved, as is Ser29, which is hydrogen bonded to His 77. Arg31, Arg84, Pro85, and Asp86 are all structurally conserved and are found in turns directly exposed to solvent. Glu183 is conserved and is found within α4 and faces the channel that runs through the tetramer.

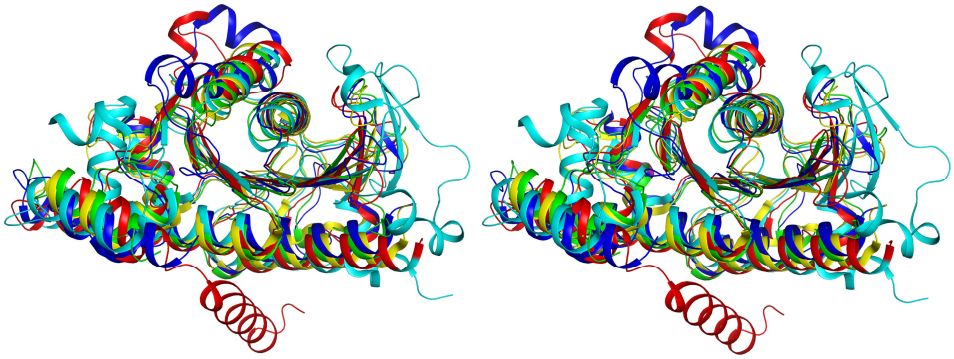


Figure 5.10: Stereoview diagram of the superposition of the top four DALI hits on HMPDdc. HMPDdc is shown in red, 1-ribulose-5-phosphate-4-epimerase is shown is blue, rhamnulose-1-phosphate aldolase is in cyan, 1-fuculose-1-phosphate aldolase from *E. coli* is green, and 1-fuculose-1-phosphate aldolase from *B. thetaiotaomicron* is colored yellow.

5.4.2. Active Site Comparison. The structure of L-fuculose 1-phosphate aldolase from *E. coli* with an inhibitor bound (PDB ID: 4FUA) was used to compare the active sites of the class II aldolase and HMPDdc *figure 5.11* (*22*). The zinc ion is located closer to the coordinating residues in the aldolase structure, only 2.0 Å from His94 and His 155 and 2.1 Å from His 92. The corresponding distances in HMPDdc to the manganese ion are 2.7 Å for His 94 and His163 and 2.2 Å for His92. With the inhibitor phosphoglycolohydroxamate bound Glu73 has been pushed out of position for coordination to the zinc ion, suggesting the same could occur upon binding of 3-

hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**). His113* is replaced by Tyr113*, which points toward the zinc ion. The adjacent residue, Met114*, is found in both structures. His27 is replaced by Asn29. Phe138 is conserved, although adopting a different side chain orientation. Gly28 still composes the floor of the binding pocket and the pocket is roughly the same size in both structures. No other residues are conserved within the binding pocket, unsurprising given that the substrates are very different.

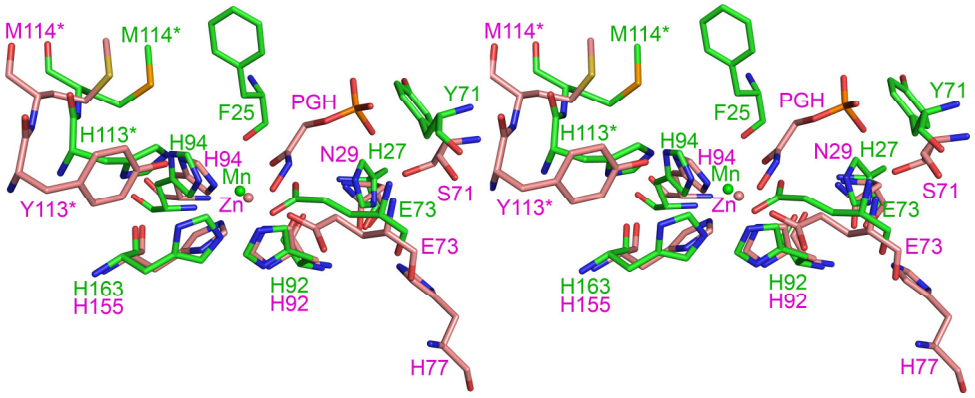


Figure 5.11: Stereoview diagram of the active sites of HMPDdc and 1-fuculose1-phosphate aldolase. HMPDdc is colored in green and 1-fuculose1-phosphate aldolase is colored in pink. The phosphoglycolohydroxamate ligand is abbreviated as PGH.

5.4.3. Structural Implications for Mechanism. Enzyme-catalyzed decarboxylations constitute a well-studied family of reactions and the role of pyridoxal phosphate, the pyruvoyl cofactor, imine formation with lysine and metal ions in the catalysis of these reactions is now well established. The general rule for the catalysis of such decarboxylations is that the enzyme provides a mechanism for the stabilization of an enolate intermediate by charge delocalization (*25*). The decarboxylation of orotidine monophopshate that occurs during pyrimidine biosynthesis is a well-known exception (*26*). The decarboxylation of hydroxy-substituted benzene rings is a common motif in polyketide biosynthesis. It is generally assumed that this reaction proceeds by initial

tautomerization of the phenol, followed by decarboxylation of the resulting keto-acid. However, this proposal has not been experimentally validated and only one structure of a hydroxybenzoic acid decarboxylase has been reported (PDB ID: 2DVX).

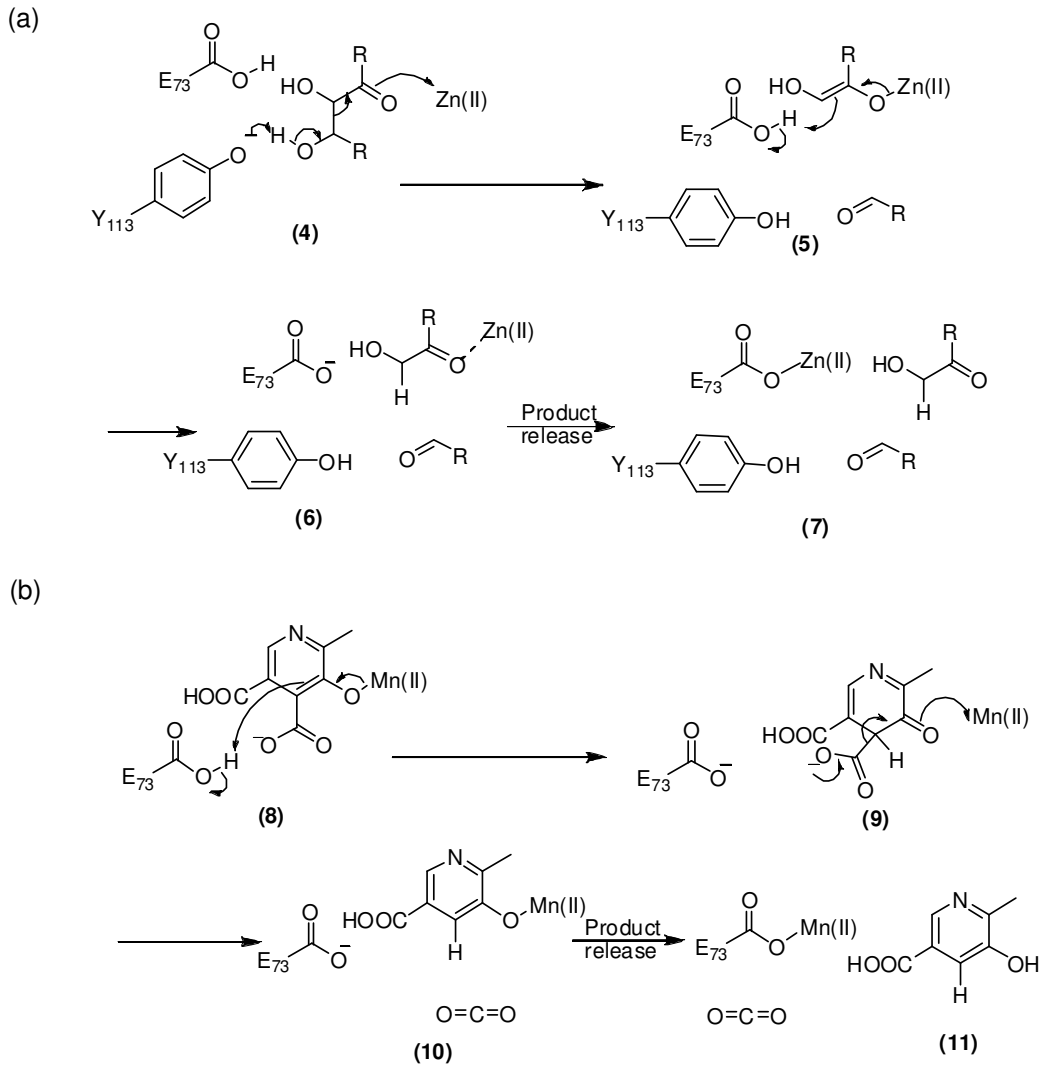


Figure 5.12: (a) The retroaldol condensation reaction catalyzed by fuculose aldolase. (b) The proposed mechanism for the decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**1**).

The Protein Data Bank contains three zinc-dependent decarboxylases: 2,6-dihydroxybenzoate decarboxylase (PDB ID: 2DVX), α-amino-β-carboxymuconate-ε-

semialdehyde decarboxylase (PDB ID: 2HBV) and a protein of unknown function similar to α-acetolactate decarboxylase (PDB ID: 1XV2) (*27*); however, none of these are structurally similar to HMPDdc. In contrast several members of the class II aldolase/adducin family were revealed by a DALI search. The structural similarity between fuculose aldolase and HMPDdc suggests that the two reactions share common mechanistic features (*28*). For the aldolase catalyzed reaction, Tyr113 initiates the retroaldol reaction by alcohol deprotonation, *figure 5.12(a)*. The resulting enolate is stabilized by the active site zinc ion. Protonation of this enolate by Glu73 followed by product release completes the reaction. In the resting state of the enzyme, Glu73 is coordinated to the zinc ion and is released from the metal upon substrate binding. Based on this proposal, we suggest an analogous mechanism for HMPDdc. In this mechanism, binding of the substrate displaces Glu73 from the manganese ion replacing it with the substrate hydroxyl. Glu73 then provides the proton for the keto-enol tautomerization to give (**9**), *figure 5.12(b)*. The decarboxylation reaction is analogous to the retroaldol reaction except that is does not need a base as the carboxylate is likely to be deprotonated under the reaction conditions. Displacement of the product from the manganese ion by Glu73 followed by product dissociation completes the reaction. The testing of this mechanistic proposal is in progress and will require additional structural and mechanistic studies.

5.5 Acknowledgement

We thank Dr. Yasunobu Ohkawa, National Institute of Agrobiological Sciences, Ibaraki, Japan for providing us with the *M. loti*/ MAFF303090 strain; the NE-CAT beamline staff for assistance during the collection and processing of the X-ray diffraction data, Dr. Cynthia Kinsland for preparing of the overexpression plasmid and Leslie Kinsland for the preparation of the manuscript.

REFERENCES

(1) http://theseed.uchicago.edu/FIG/index.cgi.

(2) Ausubel, F. M., and Brent, F. (1987) in *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.

(3) Sambrook, J., Fritsch, G. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Guide*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

(4) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem 72*, 248-254.

(5) Mukherjee, T., Kinsland, C., and Begley, T. P. (2007) PLP catabolism: Identification of the 4-Pyridoxic acid Dehydrogenase gene in Mesorhizobium loti MAFF303099. *Biorganic Chemistry in Press*.

(6) Snell, E. E., Smucker, A. A., Ringelmann, E., and Lynen, F. (1964) [Bacterial Oxidation Of Vitamin B6. Iv. Enzymatic Decarboxylation Of 2-Methyl-3-Hydroxypyridine-4,5-Dicarboxylic Acid.]. *Biochem Z 341*, 109-19.

(7) Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol. 33*, 491-7.

(8) Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307-326.

(9) Pape, T., and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for phasing with SHELX programs. *J Appl Cryst 37*, 843-844.

(10) Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2006) Automated Structure Solution With autoSHARP. *Methods Mol Biol 364*, 215-30.

(11) Cowtan, P. E. a. K. (2004) Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. D 60*, 2126-2132.

(12) Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J.,

Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D 54*, 905-21.

(13) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283-291.

(14) DeLano, W. L. (2002), DeLano Scientific, San Carlos, CA.

(15) Hasel, W., Hendrickson, T. F., and Still, W. C. (1988) A rapid approximation to the solvent-accessible surface areas of atoms. *Tetrahedron Comp. Meth. 1*, 103-16.

(16) Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389-402.

(17) Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics 15*, 305-8.

(18) Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res. 22*, 4673-80.

(19) Habe, H., Miyakoshi, M., Chung, J., Kasuga, K., Yoshida, T., Nojiri, H., and Omori, T. (2003) Phthalate catabolic gene cluster is linked to the angular dioxygenase gene in Terrabacter sp. strain DBF63. *Appl Microbiol Biotechnol 61*, 44-54.

(20) Eaton, R. W. (2001) Plasmid-encoded phthalate catabolic pathway in Arthrobacter keyseri 12B. *J Bacteriol 183*, 3689-703.

(21) Holm, L., and Sander, C. (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res. 26*, 316-319.

(22) Dreyer, M. K., and Schulz, G. E. (1996) Catalytic mechanism of the metal-dependent fuculose aldolase from Escherichia coli as derived from the structure. *J Mol Biol 259*, 458-66.

(23) Luo, Y., Samuel, J., Mosimann, S. C., Lee, J. E., Tanner, M. E., and Strynadka, N. C. (2001) The structure of L-ribulose-5-phosphate 4-epimerase: an aldolase-like platform for epimerization. *Biochemistry 40*, 14763-71.

(24) Kroemer, M., and Schulz, G. E. (2002) The structure of L-rhamnulose-1-phosphate aldolase (class II) solved by low-resolution SIR phasing and 20-fold NCS averaging. *Acta Crystallogr D Biol Crystallogr 58*, 824-32.

(25) Begley, T. P., and Ealick, S. E. (2004) Enzymatic reactions involving novel mechanisms of carbanion stabilization. *Curr Opin Chem Biol 8*, 508-15.

(26) Miller, B. G., and Wolfenden, R. (2002) Catalytic proficiency: the unusual case of OMP decarboxylase. *Annu Rev Biochem 71*, 847-85.

(27) Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235-242.

(28) Samuel, J., Luo, Y., Morgan, P. M., Strynadka, N. C., and Tanner, M. E. (2001) Catalysis and binding in L-ribulose-5-phosphate 4-epimerase: a comparison with L-fuculose-1-phosphate aldolase. *Biochemistry 40*, 14772-80.

CHAPTER 6

Identification of the 2-(acetamidomethylene)succinate hydrolase gene in *Mesorhizobium loti* MAFF303099*

6.1 Introduction

The gene encoding for 2-(acetamidomethylene)succinate hydrolase, which catalyzes the last step in the degradation of vitamin B_6 (*1*) was the last gene in the catabolic pathway that remained to be identified. The genes that have been identified to date are all clustered on the chromosome. Inspection of the nearby genes revealed a putative hydrolase (mlr6787) which is currently annotated as a 'putative DHNA-CoA thioesterase' proposed to be involved in menaquinone biosynthesis (*2*). This chapter details out the purification and characterization of the mlr6787 gene product and shows that it catalyzes the hydrolysis of 2-(acetamidomethylene) succinate (**1**) to form succinic semialdehyde (**2**), acetic acid, ammonia and carbon dioxide, *figure 6.1*. The enzyme was shown to utilize the *E* isomer of 2-(acetamidomethylene)succinate.

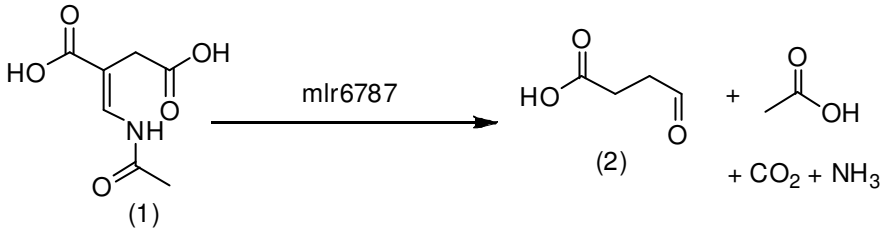


Figure 6.1: Hydrolysis reaction catalyzed by mlr6787.

6.2 Experimental Section

6.2.1. Over-expression and purification. The plasmid pMl5331.XF1 was used to transform *Escherichia coli* BL21(DE3). A starter culture was prepared by growing a single colony of transformed cells in 10 ml of LB media containing 100 μg/ml of ampicillin at

* Reproduced with permission from Mukherjee, T., Hilmey, D. G., and Begley, T.P. (2008). *Biochemistry,* **47**, 6233-4. © [2008] American Chemical Society.
The chemical synthesis was carried out by Dr. Dave Hilmey.

37 °C with overnight agitation. 1 liter LB medium (20 g/L), containing 100 μg/ml of ampicillin, was inoculated with this starter culture. The cells were grown at 37 °C with shaking at 250 rpm until the culture reached an OD_{590} of 0.6 at which point they were induced by adding IPTG to a final concentration of 0.8 mM, the temperature was lowered to 15 °C and the cells were allowed to grow for a further 12 hours. The cells were then harvested by centrifugation at 10,000g for 8 min at 4 °C.

The protein was purified by Ni-affinity chromatography following the purification procedure described in section 4.2.1. SDS-PAGE analysis showed 95% pure protein. The yield of the purified protein was 25 mg/liter. The protein concentration was measured by the Bradford assay (*3*).

6.2.2. Overexpression and purification of succinic semialdehyde dehydrogenase. The gene product of MtbH37*Rv gabD1* (Rv0234c) gene from *Mycobacterium tuberculosis* is reported to be succinic semialdehyde dehydrogenase (*4*). The MtbH37*Rv gabD1* (Rv0234c) gene in pET28b vector was overexpressed and purified in a similar manner as described above. Kanamycin was used as the antibiotic to the final concentration of 40 μg/mL. After loading the protein onto the Ni-NTA-affinity column pre-equilibrated with binding buffer kept at 4°C, the column was washed with 300 ml of wash buffer containing 50 mM NaH_2PO_4, 150 mM NaCl, 40 mM imidazole, pH 7.7. The purified protein was 80% pure by SDS-PAGE. Bradford assay (*3*) was used to measure the protein concentration and the yield of the purified protein was 2 mg/liter.

6.2.3. HPLC analysis. HPLC analysis of the enzymatic reaction mixture was performed on a Hewlett-Packard 1100 instrument using a Supelcosil LC-18-T (15 cm X 4.6 mm, 3.0 μm) column. Two different methods were used. Method A: solution A contained water, solution B contained 100 mM sodium phosphate buffer at pH 6.6 and solution C

contained methanol. The following linear gradient was used: 0% to 10% solution A and 100% to 90% solution B for 0 to 5 min, 10% to 48% solution A, 90% to 40% of solution B and 0% to 12% of solution C from 5-12 min, 48% to 50% solution A, 40% to 30% of solution B and 12% to 20% of solution C in 12-14 min, 50% to 30% solution A, 30% to 10% of solution B and 20% to 60% of solution C in 14-18 min, 30% to 0% solution A, 10% to 100% of solution B and 60% to 0% of solution C in 18-20 min and 0% of solution A, 100% of solution B and 0% of solution C in 20-25 minutes. Flow rate was maintained at 1 mL/min and the following compounds were readily separated (retention time, wavelength in parenthesis): 2-(acetamidomethylene)succinate (**1**) (2.1 min, 254 nm), 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**8**) (2.8 min, 320 nm), 3-hydroxy-2-methylpyridine-5-carboxylate (5.5 min, 320 nm), NAD (12.9 min, 254 nm), NADH (14.3 min, 254 nm), and FAD (19.3 min, 254 nm). Method B: solution A contained water with 0.1% TFA and solution B contained methanol with 0.1% TFA. The following linear gradient mixing solution A with solution B was used: 100% solution A for 0 to 2 min, 100% to 60% solution A from 2-7 min, 60% to 0% solution A in 7-10 min, 0% to 100% solution A in 10-11 min and 100% solution A in 10-15 minutes. Flow rate was 1 ml/min, and by this method the following compounds were readily separated (retention time, wavelength in parenthesis): 2-(acetamidomethylene)succinate (**1**) (4.3 min, 261 nm), DNP (**14**)(11.3 min, 261 nm) and DNP hydrazone of succinic semialdehyde (**15**) (11.65 min, 261 nm).

6.2.4. Synthesis

6.2.4.1. 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (8). The synthesis of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**8**) began with the known protocols for making N-formyl alanine (**4**) from L-alanine (**3**) (*5*) *figure 6.2(a)*, which was then converted

to 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**8**) according to literature procedures(*6*), m.p. 270-272°C; ^{1}H NMR (D_2O, 300 MHz) δ 2.60 (s, 3H), 8.21 (s, 1H); ^{13}C NMR (D_2O, 75 MHz) δ 173.7, 171.2, 155.3, 146.1, 135.7, 130.1, 110.0, 16.0.†

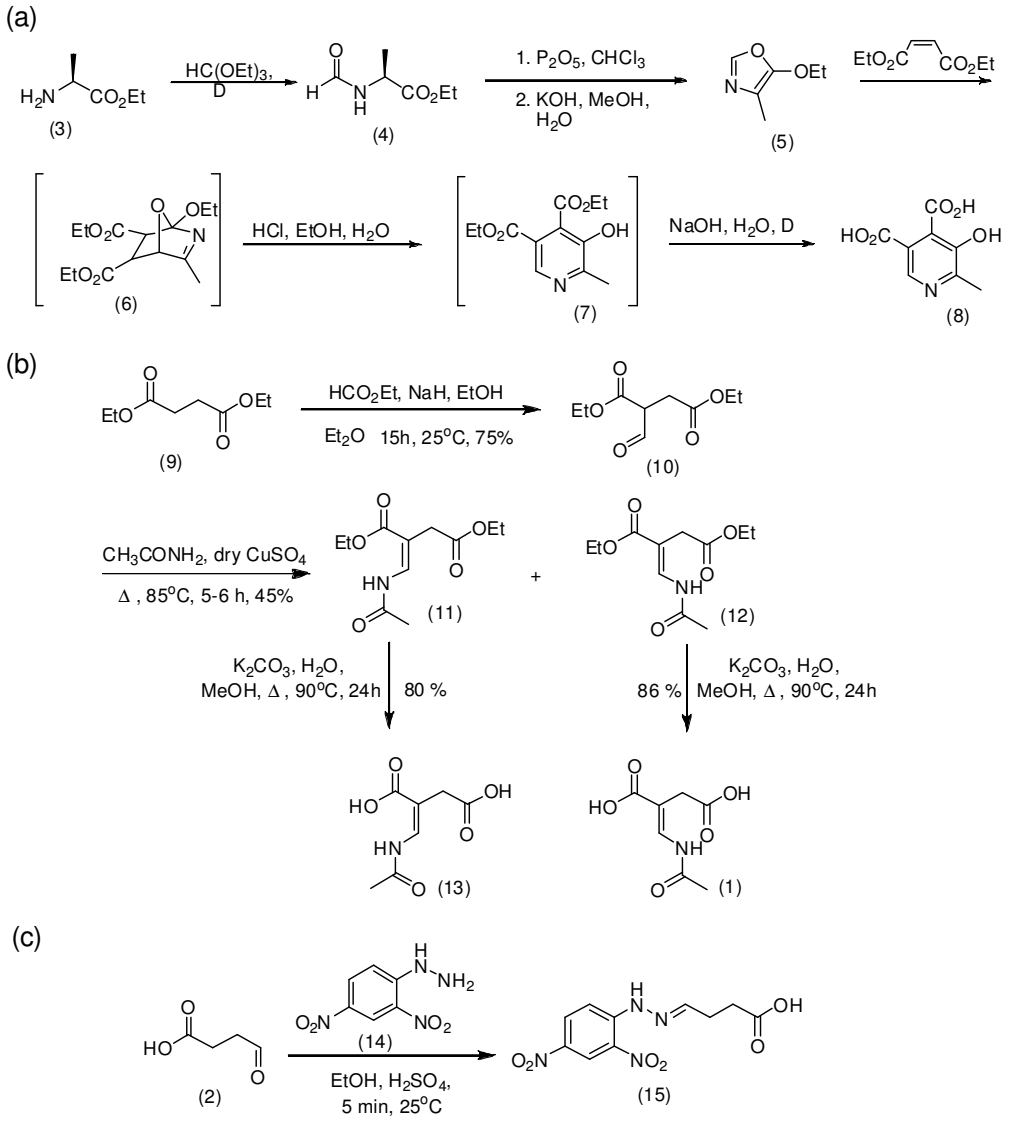


Figure 6.2: Synthetic scheme for the preparation of substrate and reference compounds (a) 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**8**); (b) *E*- and *Z* isomer of 2-(acetamidomethylene) succinate (**1** and **13**) (c) DNP hydrazone of succinic semialdehyde (**15**).

† The synthesis of (**8**) was carried out by Dr. Dave Hilmey.

6.2.4.2.. 2-(acetamidomethylene)succinic acid (1). 2-(acetamidomethylene)succinic acid (**1**) was synthesized starting from diethyl succinate (**9**) according to the scheme shown in *figure 6.2(b)*. α-Formyl diethyl succinate (**10**) was prepared from diethyl succinate (**9**) and ethyl formate in accordance with known literature procedure(*7*). Acetamidomethylene diethyl succinate (**11**, **12**) was prepared from acetamide and α-formyl diethyl succinate (**10**) from known literature procedure (*8*). After workup following the literature protocol, yellow colored oil was produced. TLC in 30% ethyl acetate and hexane showed three well separated spots with r_f values of 0.48, 0.37 and 0.15. The three compounds were isolated by flash chromatography. The spot with an r_f value of 0.48 was found to be unreacted starting material, while the other two spots corresponded to the *E* and *Z* geometric isomers (**11**, **12**)**,** both as pale yellow solids. The assignment of *E* and *Z* isomer was made based on their 2D-NOESY spectra (*figure 6.6*). The spot with an r_f value of 0.37 was assigned the *Z* form (**11**) and the one with an r_f value of 0.15 was assigned the *E* form (**12**). For the *Z*-form (**11**), ^{1}H NMR ($CDCl_3$, 600 MHz) δ 1.17 (t, *J* = 7.2 Hz, 3H), 1.19 (t, *J* = 7.2Hz, 3H), 2.07 (s, 3H), 3.10 (s, 2H), 4.05 (q, *J* = 7.2 Hz, 2H), 4.12 (q, *J* = 7.2 Hz, 2H), 7.37 (d, *J* = 11.4 Hz, 1H), 10.40 (d, *J* = 11.4 Hz, 1H) and for the *E*-form (**12**), ^{1}H NMR ($CDCl_3$, 600 MHz) δ 1.23 (t, *J* = 7.2 Hz, 3H), 1.25 (t, *J* = 7.2 Hz, 3H), 2.12 (s, 3H), 3.34 (s, 2H), 4.11 (q, *J* = 7.2 Hz, 2H), 4.17 (q, *J* = 7.2 Hz, 2H), 8.11 (d, *J* = 11.4 Hz, 1H), 8.9 (d, *J* = 11.4 Hz, 1H). The *E*-2-(acetamidomethylene) diethyl succinate (**12**) (13 mg, 0.05 mmol) was dissolved in MeOH (75 μL) and H_2O (0.5 mL) and treated with K_2CO_3 (27 mg, 0.2 mmol). After heating to 90 °C for 24 h, the solution was acidified with 4 M HCl and the solvent was evaporated *in vacuo*. The residue was dissolved in EtOH and filtered to remove salt. Evaporation of the filtrate resulted in a precipitate which proved to be the *E*-2-(acetamidomethylene)succinic acid (**1**) (8 mg, 86%) as a white solid; mp 218-220 °C ^{1}H NMR (D_2O, 300MHz) δ 2.14 (s, 3H), 3.41 (s, 2H), 8.05 (s, 1H). The *Z*-2-

(acetamidomethylene) diethyl succinate (**11**) on similar treatment as mentioned above produced *Z*-2-(acetamidomethylene)succinic acid (**27**) in 80% yield as a white solid; mp 215-218 °C ^{1}H NMR (D_2O, 300MHz) δ 2.14 (s, 3H), 3.06 (s, 2H), 7.09 (s, 1H).

6.2.4.3. DNP hydrazone of succinic semialdehyde (15). Succinic semialdehyde (**2**) (200 μL of 15% solution in water) was added to 2.2 ml of ethanol, to which 63 mg of DNP in 600 μL of concentrated sulfuric acid was added. The hydrazone formation was rapid and the orange product was filtered and dried (*figure 6.2(c)*) It was purified by flash chromatography with 5% methanol-chloroform. ^{1}H NMR (CD_3OD, 300MHz) δ 2.48 - 2.64 (m, 4H), 7.7 (t, *J* = 3.9 Hz, 1H), 7.86 (dd, *J* = 1.2, 9.6 Hz, 1H), 8.21 (ddd, *J* = 1.5, 2.7, 9.6 Hz, 1H), 8.92 (dd, *J* = 1.5, 2.7 Hz, 1H).

6.2.5. Enzymatic synthesis of the 2-(acetamidomethylene)succinate (1). The biologically relevant isomer of 2-(acetamidomethylene)succinate (**1**) was synthesized enzymatically from 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**8**). A reaction mixture (10 ml) containing 20 μM freshly purified 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase (*9*), 10 μM 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (*10*), 2 mM of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**8**), 120 mM β-mercaptoethanol, 600 μM FAD and 4 mM NADH in 100 mM sodium phosphate buffer at pH 8.0 was incubated overnight at room temperature. It was filtered using YM-10 Amicon ultracentrifugal filter at 5000 xg to remove all protein, lyophilized, redissolved in a minimum volume of 100 mM sodium phosphate buffer at pH 8.0, and then purified by HPLC using method A. It was then rotary evaporated to remove methanol and lyophilized again. The lyophilized sample was then redissolved in a minimum volume of water containing 0.01% TFA and desalted by HPLC over multiple injections using method B. Methanol was removed by rotary evaporation and TFA was removed *in vacuo*.

Trace amounts of TFA and methanol were finally removed by lyophilization, leaving a white solid. NMR of this sample was identical with the chemically synthesized *E*-2-(acetamidomethylene) succinic acid (**1**).

6.2.6. Reaction time course - depletion of substrate. A time course was determined by monitoring the disappearance of *E*-2-(acetamidomethylene) succinate (**1**) by UV-visible spectrophotometry as shown in *figure 6.3*. UV-visible scans were taken every 20 sec from 300 nm to 250 nm at the rate of 800 nm/sec. The reaction mixture (500 µL) contained 60 µM *E*-2-(acetamidomethylene) succinate (**1**) and 2 µM freshly purified enzyme in 100 mM sodium phosphate buffer at pH 8.0.

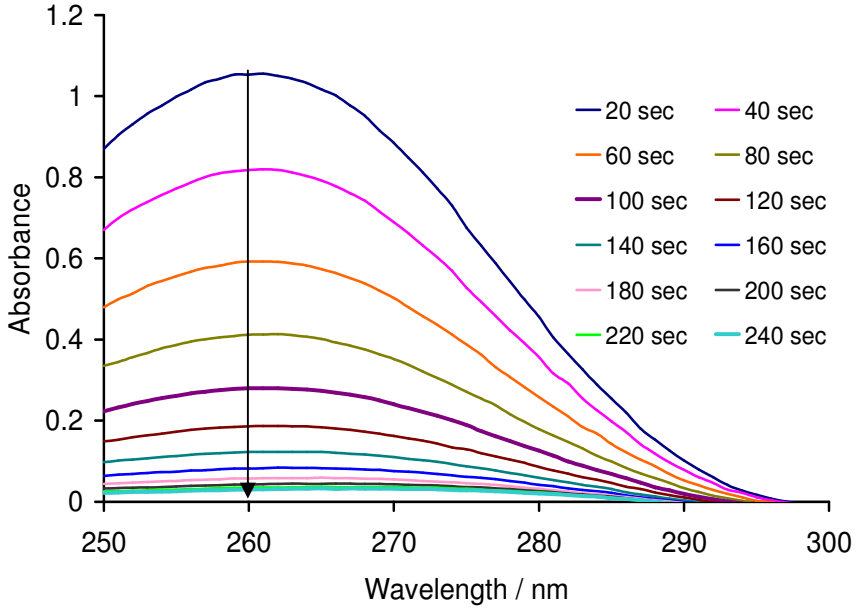


Figure 6.3: UV-visible scan of the enzymatic reaction mixture over time.

6.2.7. Reaction time course - formation of DNP hydrazone of succinic semialdehyde (15). To a 3 mM solution of *E*-2-(acetamidomethylene)succinate (**1**) (500 µL) in 100 mM sodium phosphate buffer at pH 8.0, 1 µM of freshly purified *E*-2-(acetamidomethylene)succinate hydrolase was added. At various time intervals, 40 µL of

the reaction mixture were removed and quenched by addition to 50 μL of 8 mM DNP (**14**)(prepared in 3mM HCl) and extracted with ethyl acetate (2 x 300 μL). 100 μL of the extracted sample were analyzed by HPLC, *figure 6.4*. Enzyme and substrate negative controls were run in a similar fashion in which buffer was substituted for the enzyme and the substrate, respectively. A standard of the DNP hydrazone of succinic semialdehyde (**15**) was also treated similarly.

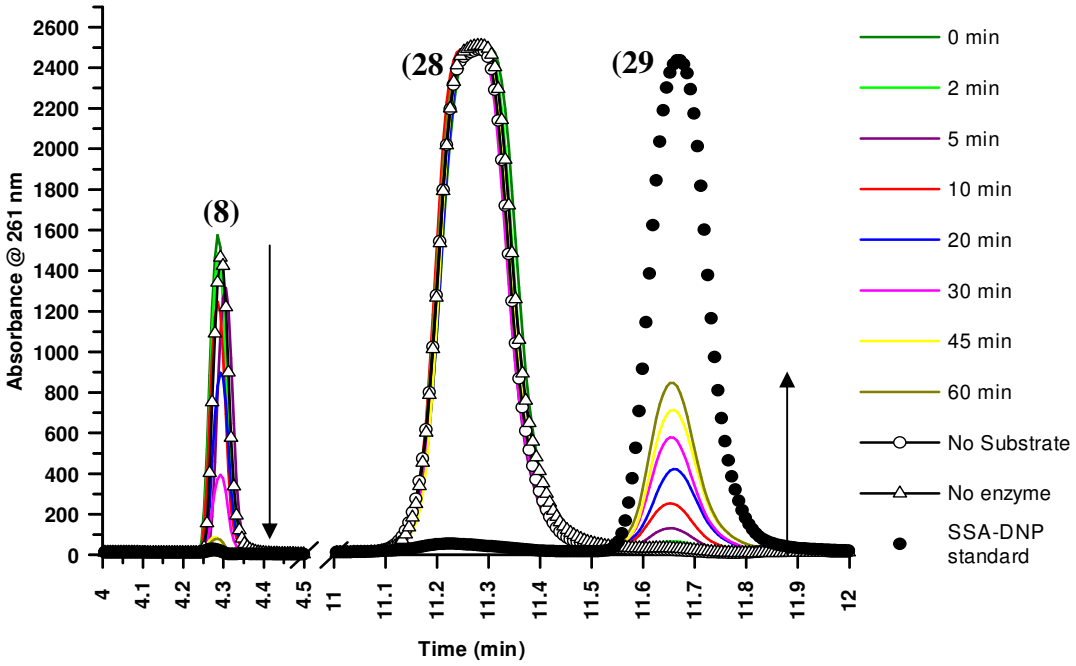


Figure 6.4: HPLC trace showing the formation of the DNP hydrazone of succinic semialdehyde (**15**) with time.

6.2.8. Assay for Ammonia. The assay for ammonia was done by monitoring the disappearance of NADPH over time in a coupled assay with α-ketoglutarate and glutamate dehydrogenase (*11*). The assay mixture (500 μL) consisted of 10 U of glutamate dehydrogenase (unit is defined as the amount of glutamate dehydrogenase that will reduce 1 μmol of α-ketoglutarate to glutamate per min at pH 8.3 at 30 ˚C), 10 mM α-

ketoglutarate, 1 mM EDTA, 250 μM NADPH, 123 μM *E*-2-(acetamidomethylene) succinate (**1**) and 2.4 μM of *E*-2-(acetamidomethylene)succinate hydrolase. The reduction of α-ketoglutarate to glutamate was not rate limiting under these assay conditions. The reaction was started by the addition of *E*-2-(acetamidomethylene) succinate hydrolase and was monitored at 340 nm. All the solutions were made in 100 mM sodium phosphate buffer at pH 8.0.

6.2.9. Assay for succinic semialdehyde. The enzymatic hydrolysis of *E*-2-(acetamidomethylene)succinate (**1**) was coupled with an assay for succinic semialdehyde dehydrogenase in order to detect the formation of succinic semialdehyde (**2**). Succinic semialdehyde dehydrogenase catalyzes the oxidation of succinic semialdehyde (**2**) to succinic acid using NAD. Succinic semialdehyde dehydrogenase was used in excess and was not rate limiting. The enzymatic reaction mixture (500 μL) consisted of 2 mM NAD, 1 mM 2-mercaptoethanol, 135 μM *E*-2-(acetamidomethylene)succinate (**1**), 1.2 μM of *E*-2-(acetamidomethylene)succinate hydrolase and 17 μM of succinic semialdehyde dehydrogenase. The *E*-2-(acetamidomethylene)succinate hydrolase was added last and the reaction was monitored at 340 nm. All the solutions were made in 100 mM sodium phosphate buffer at pH 8.0.

6.2.10. Steady state kinetic parameters. The *E*-2-(acetamidomethylene)succinate (**1**) has a λ_{max} at 261 nm and an extinction coefficient at 261 nm of 18,600 $M^{-1}cm^{-1}$(*1*). The extinction coefficient of this molecule was also determined at 280, 290 and 295 nm from the absorbance measurements at various concentration of *E*-2-(acetamidomethylene) succinate (**1**)**.** For lower concentrations of *E*-2-(acetamidomethylene)succinate (**1**) the rate of the reaction was determined from the change in absorbance at 261 nm over time. At higher concentrations of substrate the rate of change of absorbance of *E*-2-

(acetamidomethylene)succinate (**1**) was monitored at higher wavelengths (see *figure legend 6.5* for details). To a reaction mixture (500 μL) containing 250 nM enzyme, varying concentrations of *E*-2-(acetamidomethylene)succinate (**1**) were added. The rate of depletion of *E*-2-(acetamidomethylene)succinate (**1**) was monitored over 2 min. The K_M and k_{cat} for the enzyme were determined by fitting the initial rate (less than 10% conversion) of substrate depletion as a function of substrate concentration to the Michaelis-Menten equation by non-linear regression using Grafit 5.0.11 (Erithacus Software Ltd., Surrey, UK) *figure 6.5*. All solutions were made in 100 mM sodium phosphate buffer at pH 8.0.

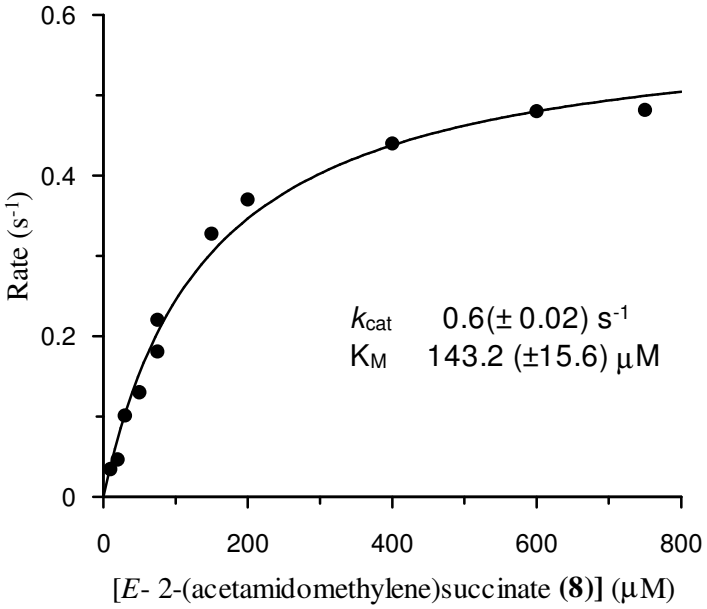


Figure 6.5: The steady state kinetic parameters for 2-(acetamidomethylene)succinate hydrolase with *E*- 2-(acetamidomethylene)succinate as a substrate were determined by monitoring the absorbance at 261 nm (for substrate concentrations of 10, 20, 30 and 50 μM), 280 nm (for substrate concentrations of 75 and 150 μM), 290 nm (for substrate concentrations of 200, 400 and 600 μM) and 295 nm (for substrate concentration of 750 μM) over time. Amount of substrate depleted was determined by dividing the change in the absorbance at a particular wavelength by the extinction coefficient at that wavelength.

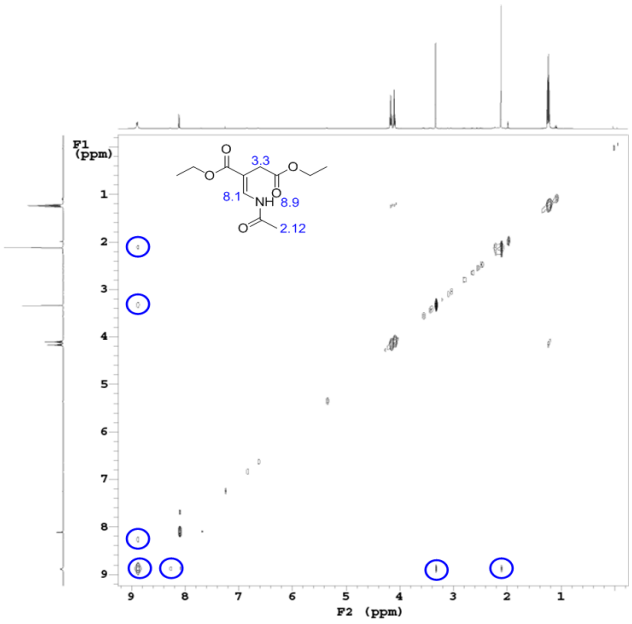
6.3. Results and Discusssion

6.3.1. Protein overexpression and purification: *E*-2-(acetamidomethylene)succinate hydrolase was overexpressed and 25 mg of protein per liter of cells were obtained at 95% purity by SDS-PAGE analysis. In contrast, even though the succinic semialdehyde dehydrogenase overexpressed well the best preparation yielded only 2 mg of protein per liter of cells at 80% purity.

6.3.2. Substrate identification: 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase catalyzes the oxidative pyridine ring opening reaction in vitamin B_6 catabolism, *figure 1.3* (*12*). This enzyme has been characterized in *Pseudomonas sp.* MA-1 (*13-18*). The product, 2-(acetamidomethylene)succinate (**1**) could exist as two geometric isomers, *E* (**1**) and *Z* (**13**). The stereochemistry of the enzymatic product had not been previously determined and the rate of isomer conversion was unknown.

During the chemical synthesis of 2-(acetamidomethylene)succinate (**1**), we observed that treatment of acetamide with α-formyl diethyl succinate resulted in two compounds which were purified by flash chromatography and characterized by two dimensional nuclear overhauser enhancement spectroscopy (NOESY) as the *E* (**1**) and *Z* (**13**) isomers of acetamidomethylene diethyl succinate, *figure 6.6*. In the case of the *E*-isomer (**1**) the amide hydrogen at 8.9 ppm has a cross peak that correlates with the allylic hydrogen at 3.3 ppm, while no cross peak is seen between the vinyl hydrogen at 8.1 ppm and the allylic hydrogen at 3.3 ppm. In case of the *Z* isomer (**13**) no cross peak is observed between the amide hydrogen at 10.4 ppm and the allylic hydrogen at 3.1 ppm while the vinyl hydrogen at 7.37 ppm has a cross peak that correlates with the allylic hydrogen at 3.1 ppm.

(a)



(b)

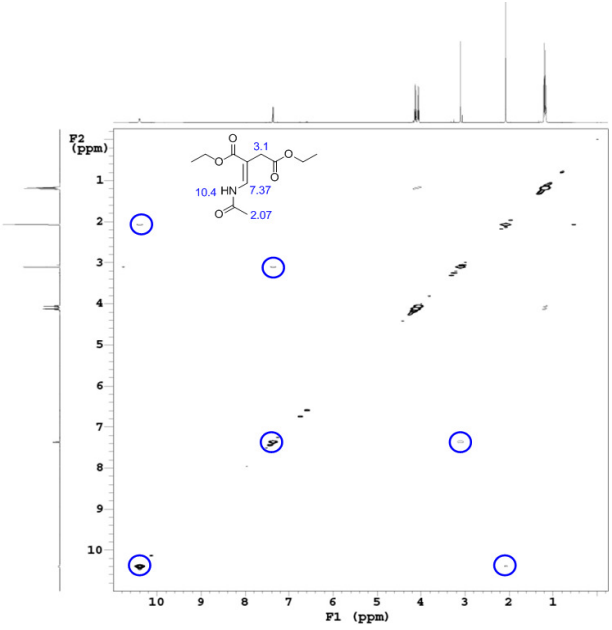


Figure 6.6: Phase sensitive 2D NOESY was acquired with a sweep width of 9.6 kHz. 200 points were collected in the indirectly detected dimension with 8 scans and 5758 points per increment. The resulting matrix was zero filled to 4096 X 4096 complex data points and Gaussian window functions were applied in both dimensions prior to Fourier transformantion. (a) 2D NOESY of *E*- 2-(acetamidomethylene)succinate (**1**). (b) 2D NOESY of *Z*- 2-(acetamidomethylene)succinate (**13**).

The biologically relevant isomer of 2-(acetamidomethylene)succinate was enzymatically synthesized starting from 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (**8**) using 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase (*9*) and 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (*10*). NMR analysis demonstrated that the enzymatic product was the *E* isomer of 2-(acetamidomethylene)succinate (**1**).

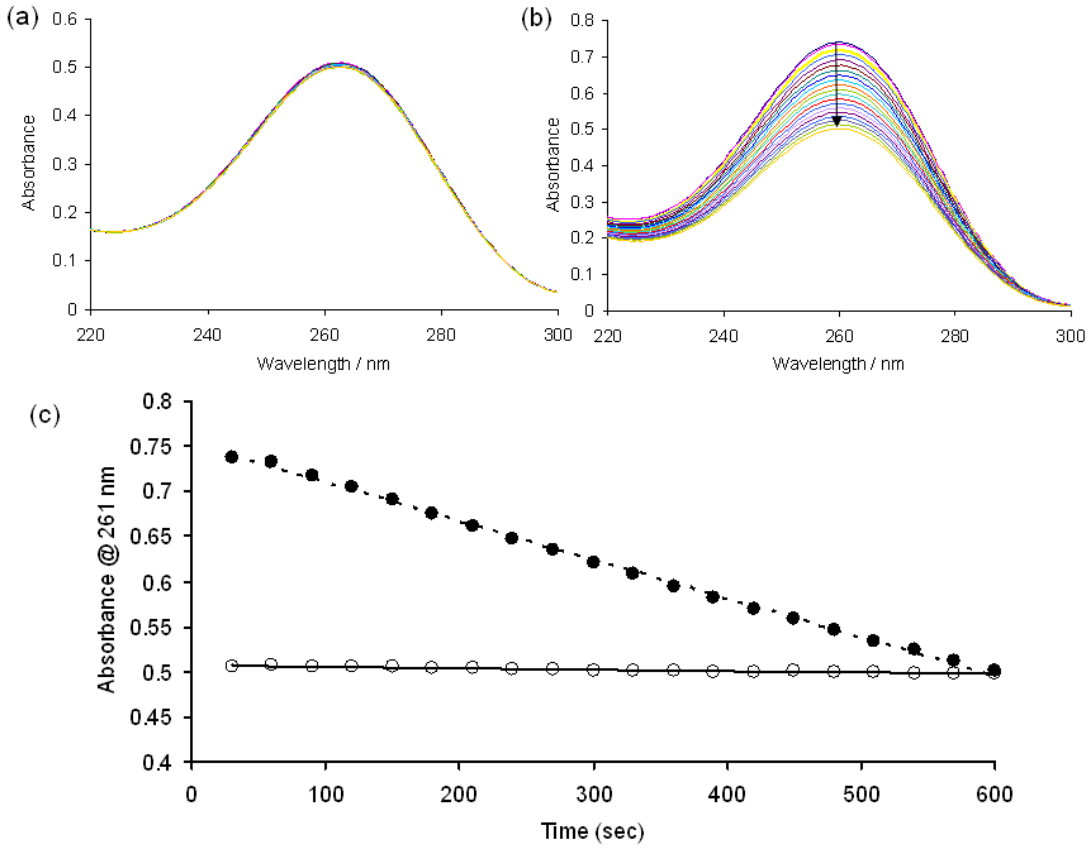


Figure 6.7: (a) Assay of the mlr6787 gene product with *Z*- 2-(acetamidomethylene) succinate (**13**). Assay mixture consisted of 75 μM of *Z*- isomer (**13**) and 250 nM freshly purified enzyme. (b) Assay of mlr6787 gene product with *E*- 2-(acetamidomethylene) succinate (**1**). Assay mixture consisted of 75 μM *E*- isomer (**1**) and 250 nM freshly purified enzyme. (c) Plot showing the decrease in absorbance at 261 nm over time for the E and Z substrate isomers. Solid line: *Z*- 2-(acetamidomethylene)succinate (**13**) Dotted line: *E*- 2-(acetamidomethylene) succinate (**1**).

Both the *E* and the *Z* isomers of 2-(acetamidomethylene)succinate (**1**, **13**) were tested as substrates for the mlr6787 gene product. The *E* - isomer (**1**) was readily hydrolyzed. In contrast, no hydrolysis was observed for the *Z*- isomer (**13**), indicating that the enzyme was specific for the hydrolysis of (**1**) *figure 6.7*. There is no detectable non-enzymatic isomerization of *E*-2-(acetamidomethylene)succinate (**1**) over a period of 15 days in 100 mM sodium phosphate buffer at pH 8.0 in D_2O kept at room temperature.

6.3.3. Product characterization: The 2,4-dinitrophenyl hydrazone of sucinnic semialdehyde was chemically synthesized and found to elute at 11.65 min by HPLC, using method B. The reaction time course shown in *figure 6.4,* indicates the depletion of the *E*-2-(acetamidomethylene)succinate (**1**) at 4.26 min and the formation of the DNP hydrazone of succinic semialdehyde (**15**) at 11.65 min over time, thereby showing that succinic semialdehyde (**2**) is the product of the enzymatic reaction. No hydrazone (**15**) was formed in the absence of either the enzyme or the substrate.

A further proof that succinic semialdehyde (**2**) is a product of the enzymatic hydrolysis of *E*-2-(acetamidomethylene)succinate (**1**) comes from the coupled assay with succinic semialdehyde dehydrogenase (*4*) in which NADH formed as a result of the coupled enzymatic reaction can be monitored at 340 nm. An increase in absorbance at 340 nm was seen on addition of succinic semialdehyde dehydrogenase to the reaction mixture containing NAD, *E*-2-(acetamidomethylene)succinate (**1**) and the *E*-2-(acetamidomethylene)succinate hydrolase. No NADH was formed in the absence of the succinic semialdehyde dehydrogenase or the *E*-2-(acetamidomethylene)succinate hydrolase *figure 6.8 (a).*

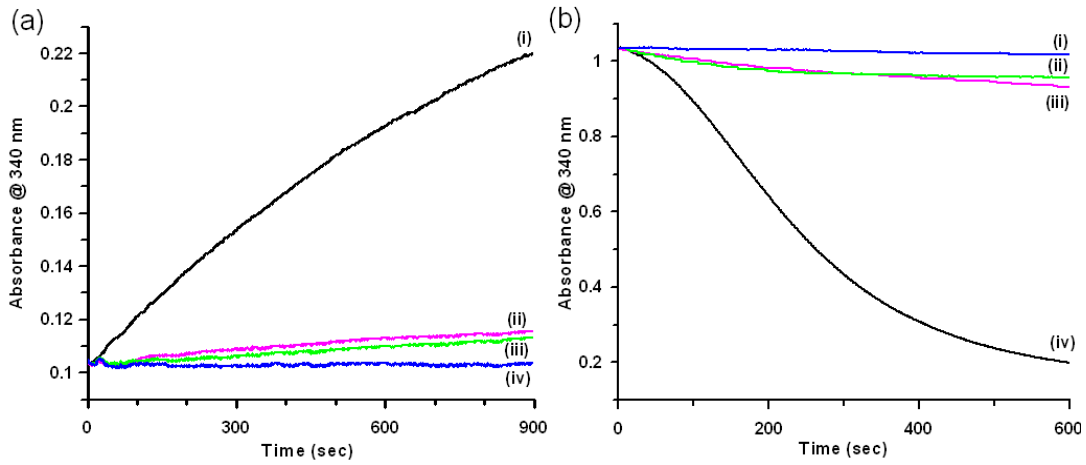


Figure 6.8: Enzymatic assays for the production of succinic semialdehyde and ammonia. (a) Trace (i) shows the formation of NADH (at 340 nm) for the enzymatic reaction mixture (500 μL) consisting of 2 mM NAD, 1 mM β-mercaptoethanol, 135 μM *E*-isomer (**1**), 1.2 μM of the hydrolase and 17 μM of succinic semialdehyde dehydrogenase. Traces (ii), (iii) and (iv) shows the formation of NADH for the reaction mixture which had all the components of trace (i) except *E*-isomer (**1**), the hydrolase and succinic semialdehyde dehydrogenase, respectively. (b) Trace (iv) shows the formation of NADH (at 340 nm) for the enzymatic reaction mixture (500 μL) consisting of 10 U of glutamate dehydrogenase, 10 mM α-ketoglutarate, 1 mM EDTA, 250 μM NADPH, 123 μM *E*-isomer (**1**) and 2.4 μM of the hydrolase. Traces (i), (ii) and (iii) shows the formation of NADH for the reaction mixture (500 μL) which had all the components of trace (iv) except glutamate dehydrogenase, *E*-isomer (**1**) and the hydrolase, respectively.

A coupled assay with glutamate dehydrogenase in the presence of α-ketoglutarate was used for the detection of ammonia (*11*). In the presence of ammonia, α-imino glutarate is formed and subsequently reduced to glutamine by glutamate dehydrogenase. This enzyme requires NADPH as the cofactor, which is oxidized to NADP during the

reaction resulting in a decrease in absorbance at 340 nm. The coupled assay with mlr6787 showed a decrease in absorbance at 340 nm. No decrease in absorbance at 340 nm was seen in the absence of either the glutamate dehydrogenase or the *E*-2-(acetamidomethylene)succinate hydrolase. This confirmed the production of ammonia during the enzymatic hydrolysis of *E*-2-(acetamidomethylene)succinate (**1**) *figure 6.8(b).*

The previously reported hydrolysis of 2-(acetamidomethylene)succinate (**1**) in *Pseudomonas MA-1* (*1*) and *Mesorhizobium Loti* also yields the same products.

6.3.4. Steady state kinetic parameters: The *E*-2-(acetamidomethylene)succinate (**1**) has a λ_{max} at 261 nm and the extinction coefficient at 261 nm has been reported as 18,600 M^{-1} cm^{-1} (*1*). A continuous assay at 261 nm at higher concentrations of the substrate was not possible as the initial absorbance was too high to obtain an accurate measurement. To avoid using a discontinuous assay, the extinction coefficients at longer wavelengths were measured. The extinction coefficients of *E*-2-(acetamidomethylene)succinate (**1**) at 280, 290 and 295 nm in 100 mM sodium phosphate buffer at pH 8.0 were determined to be 6900, 2000 and 700 $M^{-1}cm^{-1}$, respectively. Steady state kinetic parameters were obtained from the concentration dependence of the rate of depletion of *E*-2-(acetamidomethylene) succinate (**1**) at constant concentration of the hydrolase. The enzymatic reaction exhibited Michealis Menten kinetics with K_M and k_{cat} of 143 μM and 0.6 s^{-1} respectively. The k_{cat}/K_M for HMPDdc was determined to be 4196 $M^{-1}s^{-1}$, *figure 6.5*.

6.3.5. Mechanistic analysis: Two mechanisms for the production of succinic semialdehyde are shown in *figure 6.9*. In mechanism A (*1*) the amide bond of *E*-2-(acetamidomethylene)succinate (**1**) is first hydrolyzed to give acetic acid and the intermediate (**18**) which could then undergo a decarboxylation followed by deamination to generate succinic semialdehyde (**2**). The order of these two steps could also be

reversed. Mechanism B predicts a conjugate addition of water to the *E*-2-(acetamidomethylene)succinate (**1**) to generate intermediate (**23**) which could then eliminate acetamide resulting in (**24**). Facile tautomerization would give (**21**) which could then decarboxylate to produce succinic semialdehyde (**2**). Since the enzymatic hydrolysis of *E*-2-(acetamidomethylene)succinate (**1**) generates ammonia rather than acetamide and acetamide is stable under our reaction conditions, the hydrolysis of *E*-2-(acetamidomethylene)succinate (**1**) is likely to proceed via mechanism A.

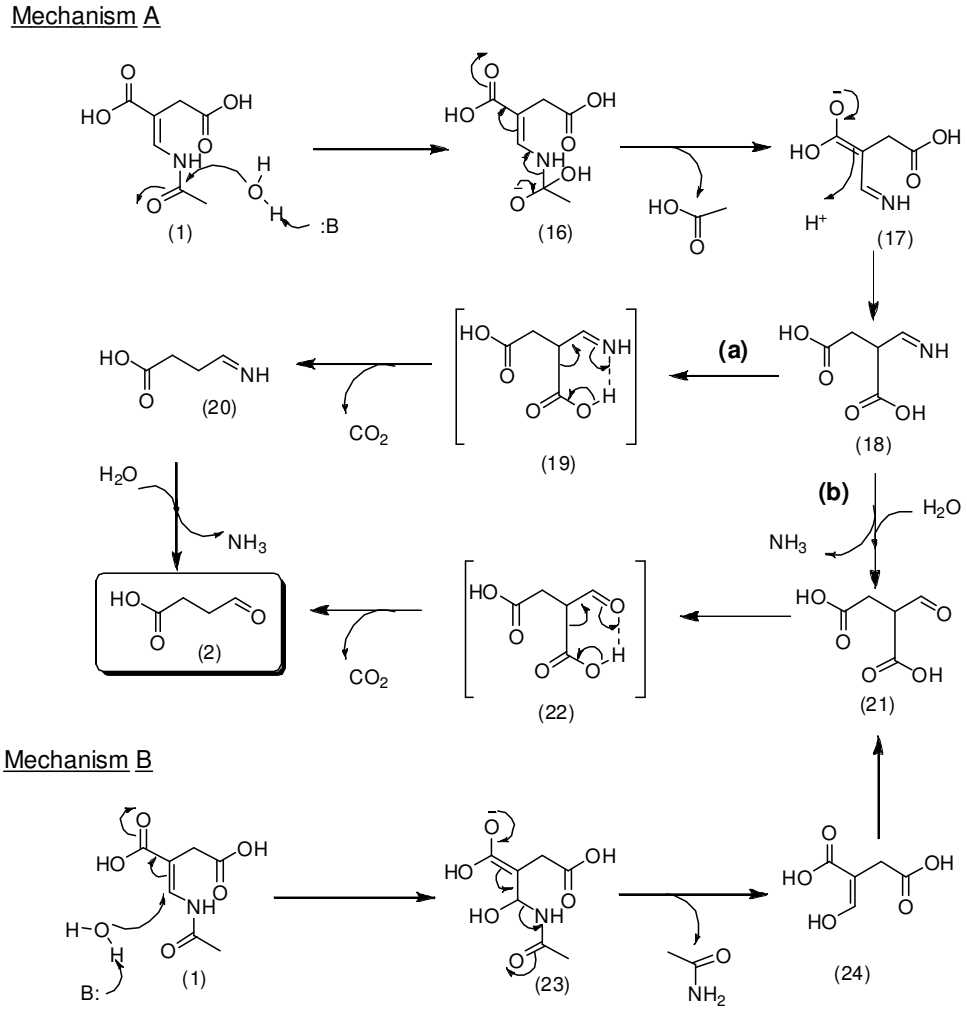


Figure 6.9: Mechanism of hydrolysis of *E*-2-(acetamidomethylene)succinate (**1**).

6.3.6. Comparative genomics analysis: *E*-2-(acetamidomethylene)succinate hydrolase catalyzes the last step in the degradative pathway of PLP, generating products that can be directly incorporated into primary metabolism. The identification of its gene completes the gene assignment of all the enzymes involved in PLP catabolism in *M. loti*. Surprisingly, when we search for an equivalent pathway in available sequenced bacterial genomes using the SEED database, we find that this pathway is unique to *M. loti*, suggesting that other PLP catabolic pathways still remain to be discovered.

6.4. Acknowledgement

We gratefully acknowledge the help from Dr. Yasunobu Ohkawa, National Institute of Agrobiological Sciences, Ibaraki, Japan for providing us with the /M. loti/ MAFF303090 strain; Dr. Cynthia Kinsland for preparation of the overexpression plasmid, Dr. Carl Nathan, Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York for providing us with the MtbH37*Rv gabD1* (Rv0234c) plasmid, Dr. Ivan Keresztes and Anthony Condo for their assistance with the NOESY experiment and Dr. Jeremiah Hanes for insightful discussions.

REFERENCES

(1) Huynh, M. S., and Snell, E. E. (1985) Enzymes of vitamin B6 degradation. Purification and properties of two N-acetylamidohydrolases. *J Biol Chem 260*, 2379-2383.

(2) http://theseed.uchicago.edu/FIG/index.cgi.

(3) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem 72*, 248-254.

(4) Tian, J., Bryk, R., Itoh, M., Suematsu, M., and Nathan, C. (2005) Variant tricarboxylic acid cycle in Mycobacterium tuberculosis: identification of alpha-ketoglutarate decarboxylase. *Proc Natl Acad Sci U S A 102*, 10670-10675.

(5) Chancellor, T., Morton, C. (1994) A facile procedure for the synthesis of N-formyl amino acid esters. *Synthesis 10*, 1023-1025.

(6) Firestone, R. A., Harris, E. E., Reuter, W. (1967) Synthesis of pyridoxine by Diels-Alder reactions with 4-methyl-5-alkoxyoxazoles. *Tetrahedron 23*, 943-955.

(7) Tonari, K., Yonemoto, K. (2002) Synthesis and antibacterial activity of a-methylene-g(b)-carboxy-g-butyrolactones. *Journal of Oleo Science 51*, 259-264.

(8) Lingens, F., Hankwitz, R. (1963) Synthesis of acetamidomethylenesuccinic acid and derivatives of dimethyl formylsuccinate. *Ann. 670*, 31-47.

(9) Mukherjee, T., McCulloch, K. M., Ealick, S. E., and Begley, T. P. (2007) Gene Identification and Structural Characterization of the Pyridoxal 5'-Phosphate Degradative Protein 3-Hydroxy-2-methylpyridine-4,5-dicarboxylate Decarboxylase from Mesorhizobium loti MAFF303099. *Biochemistry 46*, 13606 - 13615.

(10) Yuan, B., Yokochi, N., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2006) Molecular cloning, identification and characterization of 2-methyl-3-hydroxypyridine-5-

carboxylic-acid-dioxygenase-coding gene from the nitrogen-fixing symbiotic bacterium Mesorhizobium loti. *J Biosci Bioeng 102*, 504-510.

(11) Day, N., and Keillor, J. W. (1999) A continuous spectrophotometric linked enzyme assay for transglutaminase activity. *Anal Biochem 274*, 141-144.

(12) Sparrow, L. G., Ho, P. P., Sundaram, T. K., Zach, D., Nyns, E. J., and Snell, E. E. (1969) The bacterial oxidation of vitamin B6. VII. Purification, properties, and mechanism of action of an oxygenase which cleaves the 3-hydroxypyridine ring. *J Biol Chem 244*, 2590-2600.

(13) Chaiyen, P., Sucharitakul, J., Svasti, J., Entsch, B., Massey, V., and Ballou, D. P. (2004) Use of 8-substituted-FAD analogues to investigate the hydroxylation mechanism of the flavoprotein 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 43*, 3933-3943.

(14) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Thermodynamics and reduction kinetics properties of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 36*, 2612-2621.

(15) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Unusual mechanism of oxygen atom transfer and product rearrangement in the catalytic reaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 36*, 8060-8070.

(16) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Reaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid (MHPC) oxygenase with N-methyl-5-hydroxynicotinic acid: studies on the mode of binding, and protonation status of the substrate. *Biochemistry 36*, 13856-13864.

(17) Kishore, G. M., and Snell, E. E. (1981) Interaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase with FAD, substrates, and analogues. Spectral and fluorescence investigations. *J Biol Chem 256*, 4234-4240.

(18) Kishore, G. M., and Snell, E. E. (1981) Kinetic investigations on a flavoprotein oxygenase, 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *J Biol Chem 256*, 4228-4233.

CHAPTER 7

Mechanistic studies of 2-Methyl-3-hydroxypyridine-5-carboxylic acid oxygenase from *Mesorhizobium loti* MAFF303099*

7.1 Introduction

The most interesting enzyme in the PLP catabolic pathway is the 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (MHPCO) that catalyzes the ring opening of 2-methyl-3-hydroxypyridine-5-carboxylic acid (MHPC) (**1**) with the incorporation of two oxygen atoms producing *E*-2-acetaminomethylene succinate (**2**), the first acyclic intermediate on the pathway, *figure 7.1*. MHPCO (E.C. 1.14.12.4) has been classified as a flavin-dependent monooxygenase based on its sequence and biochemical similarity to other enzymes in this family. This class of enzymes utilize FAD and molecular oxygen to form a C(4a)-hydroperoxy-FAD intermediate which then donates one oxygen atom to the substrate (*1*). The pyridine ring-opening is an unusual reaction for this enzyme family because most oxidative ring-opening reactions of aromatic compounds are catalyzed by iron-dependent dioxygenases.

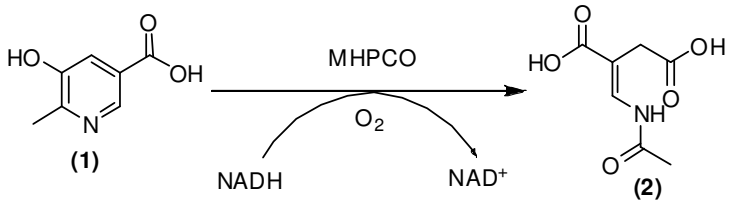


Figure 7.1: The reaction catalyzed by MHPCO (mlr6788)

MHPCO has been characterized in sufficient detail in *Pseudomonas sp.* MA-1. The binding studies (*2*) and effect of substrate and FAD analogues (*3*) have been reported, the mechanism of hydroxylation (*4*) and oxygen atom transfer and product rearrangement (*5*)

* Reproduced in parts with permission from K. M. McCulloch, T. Mukherjee, T. P. Begley and S. E. Ealick (2009). *Biochemistry,* **48** (19), pp 4139–4149. © [2009] American Chemical Society. The crystallography and its analysis was done by K. M. McCulloch.

has been investigated and the kinetic (*6*) and thermodynamic characterization (*7*) has been done. Previous work has shown the *Pseudomonas* MA-1 MHPCO to crystallize and diffract to reasonable resolution, although the structure was not reported (*8*). In this chapter the cloning and overexpression of mlr6788 is reported, along with a brief description of the X-ray crystal structures of *M. loti* MHPCO which was determined at 2.1 Å resolution[†]. The crystal structure was also obtained for the protein complex with MHPC (**1**). These structures confirm the relationship of MHPCO to other members of the flavin monooxygenase family and provide insight into the catalytic mechanism. Based on the active site information from the crystal structure, mutagenesis studies were carried out to investigate the enzyme mechanism. Substrates analogous were designed and chemical reactions were carried out to probe into the mechanism of MHPCO.

7.2 Experimental

7.2.1. Cloning of* M. loti *MHPCO. Standard methods were used for DNA manipulations (9, 10). The *M. loti* MHPCO gene was amplified from genomic DNA (purified from ATCC strain # 700743) by PCR with the following primer pair: 5'-CAC CAT GGC CAA TGT AAA CAA AAC TCC-3' and 5'-CTA CTG CGG CCA CGA GTA GAC ACG GCG CAG C-3'. The PCR product was purified and used in a topoisomerase-mediated reaction with pENTR-TEV-D-TOPO (Invitrogen) following the manufacturer's instructions. Clones were screened by PCR and verified by sequencing. A correct clone was used in an LR reaction with the plasmid pDESTF1, a Gateway-adapted vector based on the pET-system from Novagen that encodes an N-terminal 6xHis tag under the control of the T7lac promoter. Clones were again screened by restriction digestion. A correct clone was named pMl5332.XF1[‡].

[†] All crystallographic studies and analysis were done by Kathryn McCulloch.

[‡] The molecular cloning was done by Dr. Cynthia Kinsland.

7.2.2. Protein overexpression and purification. *E. coli* Tuner (DE3) was transformed first using the chaperone-encoding plasmid pGro7 (Takara Bio Inc.) and then using pMl5332.XF1. The cells were grown in 1L LB medium, containing 100 μg/mL of ampicillin at 37 °C with agitation. Chaperone proteins GroES and GroEL were induced at cell culture inoculation by addition of 2.5 g/L L-arabinose. When the culture reached an OD_{590} of 0.6, it was induced by adding IPTG to a final concentration of 0.5 mM, the temperature was lowered to 15 °C and the cells were allowed to grow for a further 12 h before being harvested. The protein was purified by Ni-affinity chromatography following the purification procedure described in section 4.2.1. 60 mM imidazole was used to wash the Ni-column in order to get rid of the chaperone protein as it stuck non specifically to the column.

For crystallographic studies[†], the selenomethionine (SelMet) protein was prepared by transforming the plasmid pMl5332.XF1 into E. coli B834(DE3) cells, a strain auxotrophic for methionine. Cells were grown at 37 °C with shaking in minimal M9 media which was supplemented with 20 mg/L of all amino acids except methionine, 1X MEM vitamin mix, 0.4% glucose, 50 mg/L L-selenomethionine, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 25 mg/L $FeSO_4$, and 100 mg/L ampicillin. After NiNTA chromatography and additional step of size exclusion chromatography was followed. In both preparations, SDS-PAGE analysis showed 95% pure protein and the yield of the purified protein was 1-2 mg/liter. The protein concentration was measured by the Bradford assay (11).

7.2.3. Synthesis

7.2.3.1. Enzymatic synthesis of substrate (1). The enzymatic synthesis and purification of 2-methyl-3-hydroxypyridine-4,5-dicarboxylic acid (MHPD) has been

[†] All crystallographic studies and analysis were done by Kathryn McCulloch.

previously described (*12*). MHPC (**1**) was enzymatically synthesized from MHPD in a 10 mL reaction mixture containing 5 mM MHPD and 100 μM of MHPD decarboxylase in 100 mM sodium phosphate buffer (pH 8.0) incubated over night at room temperature. MHPC was purified by the HPLC (*12, 13*) and upon lyophilization yielded a stable white solid. ^{1}H NMR (300 MHz, D_2O) δ 2.46 (s, 3H, CH_3), 7.68 (s, 1H, C4-H), and 8.12 (s, 1H, C6-H). The purified MHPC was dissolved in 100 mM Tris (pH 7.7) to a final concentration of 100 mM and stored frozen at -20 °C until use.

7.2.3.2. 3-hydroxy-5-(methoxycarbonyl)-1-methylpyridinium (4). 3-hydroxy-5-(methoxycarbonyl)-1-methylpyridinium (**4**) was synthesized in one step from methyl-5-hydroxynicotinate (**3**) and methyl iodide following literature procedure (*2*). Quantitative yield was seen and the NMR agreed with that reported.

7.2.3.3. 3-methoxy-5-(methoxycarbonyl)-1-methylpyridinium (6). 3-methoxy-5-(methoxycarbonyl)-1-methylpyridinium (**6**) was synthesized in a similar manner as described above from methyl-5-methoxynicotinate (**5**). ^{1}H NMR (400 MHz, D_2O) δ 9.01 (s, 1H), 8.75 (s, 1H), 8.55 (s, 1H), 4.44 (s, 3H), 4.09 (a, 3H), 4.03 (s, 3H).

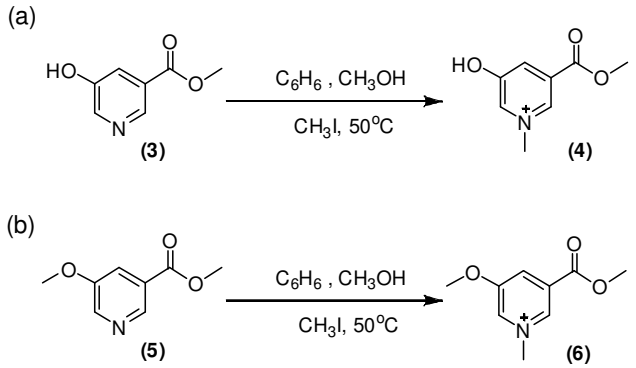


Figure 7.2: Synthetic scheme (a) 3-hydroxy-5-(methoxycarbonyl)-1-methylpyridinium (**4**) and (b) 3-methoxy-5-(methoxycarbonyl)-1-methylpyridinium (**6**)

7.2.4. Reaction with H_2O_2. Compound (**4**) (4.3 μmol) was treated with H_2O_2 (60 μmol) in presence and absence of NaOH. In the later case the H_2O_2 was premixed with the base for 30 min before its addition to (**4**). The reaction mixture was allowed to sit at 25 °C for 1 hour before analyzing by HPLC. Compound (**6**) was treated identically as (**4**).

7.2.5. Reaction with mCPBA. m-Chloroperbenzoic acid (25 mM) has sparing solubility in water so the reaction was done in 50% methanol. Compound (**4**) (25 mM) was treated with mCPBA (50 mM) in presence and absence of K_2CO_3 (50 mM). In the later case the mCPBA was premixed with the base for 30 min before adding to (**4**). The reaction mixture was allowed to sit at 25 °C for 2 hour before analyzing by HPLC. Compound (**6**) was treated in an identical fashion.

7.2.6. Activity Assay for MHPCO The recombinant His-tagged MHPCO was assayed by monitoring the disappearance of NADH absorbance at 340 nm as previously described (*13*).

7.2.7. HPLC. HPLC analysis of the enzymatic reaction mixture was performed on a Hewlett-Packard 1100 instrument using a Supelcosil LC 18 T (15 cm X 4.6 mm, 3.0 μm) column. Solution A contained water with 0.1% TFA and solution B contained acetonitrile with 0.1% TFA. The following linear gradient mixing solution A with solution B was used: 100% solution A for 0 to 2 min, 100% to 40% solution A from 2-7 min, 40% to 100% solution A in 7-10 min, 100% to 0% solution A in 10-11 min and 100% solution A in 11-15 minutes. Flow rate was 1 ml/min, and by this method the following compounds were readily separated (retention time, wavelength in parenthesis): compound (**4**) (7.7 min), N-methyl-5-hydroxynicotinate (4.6 min), compound (**6**) (7.87 min), N-methyl-5-methoxynicotinate (6.8 min) and mCPBA (11.7 min).

7.3. Results and Discussion

7.3.1. Protein overexpression and purification. The protein over expressed very well and but had problems with its solubility. Heterologus overexpression of the protein in BL21 (DE-3) cell lines grown in LB media with reduced IPTG concentration did not improve the solubility of the protein. BL-21(DE-3) cells have *lacY* gene which codes for β-galactosides permease, thus the local concentration of IPTG in the cells cannot be regulated by lowering the overall concentration of IPTG, during induction in the cell culture. To enhance correct protein folding and preventing the formation of inclusion body it was necessary to decrease the local concentration of IPTG. Tuner (DE-3) cell lines are same as Bl-21(DE-3) cell lines but do not have the *lacY* gene in them. Thus they were be used for overexpression in LB media with very low IPTG concentration (0.2 mM). No improvement on the protein solubility was observed. In order to give more time for the protein to fold properly Tuner (DE-3) cell transformed with MHPCO gene was grown in minimal media with 0.2 mM IPTG concentration, however no improvement in protein solubility was observed.

The plasmid for MHPCO was then co-expressed with each of the plasmids containing various combinations of genes encoding for a mixture of chaperone proteins, *table 7.1*.

Table 7.1: The various plasmids marketed by *Takara Bio Inc.* containing various combinations of chaperone proteins.

No.	Plasmid	Chaperone	Promoter	Inducer
1	pG-KJE8	dnaK-dnaJ-grpE groES-groEL	*araB* *Pzt1*	L-Arabinose Tetracyclin
2	pGro7	groES-groEL	*araB*	L-Arabinose
3	pKJE7	dnaK-dnaJ-grpE	*araB*	L-Arabinose
4	pG-Tf2	groES-groEL-tig	*Pzt1*	Tetracyclin
5	pTf16	Tig	*araB*	L-Arabinose

All the combinations were tried and pure soluble yellow colored protein was obtained when pMl5332.XF1 was co-expressed with pGro7 in competent Tuner (DE-3) and induced with 0.5 mM IPTG. 1mg of protein was obtained from 1 L of cells.

.

7.3.2. MHPCO activity The recombinant His-tagged MHPCO was purified using metal affinity chromatography and was found to be active; however, the His-tagged *Ml*MHPCO was found to be approximately five-fold slower than the *Ml*MHPCO purified by classical methods without a His tag (*13*). The activity of recombinant MHPCO with the His-tag cleaved by digestion with TEV protease was comparable to the activity of the His-tagged MHPCO sample.

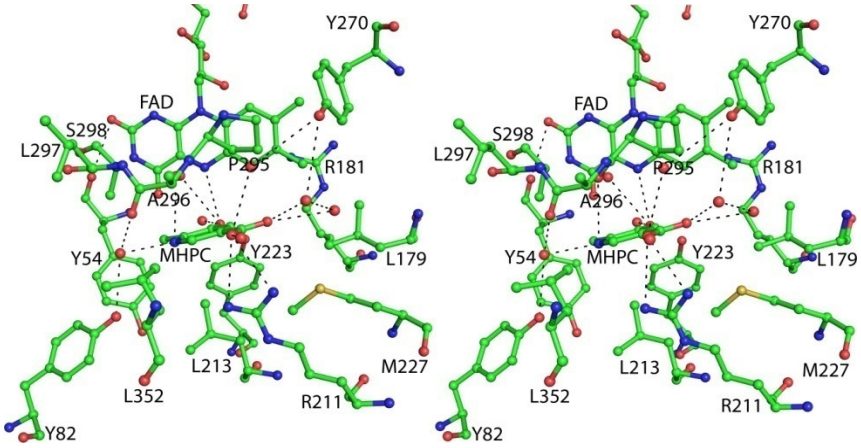


Figure 7.3: Stereoview diagram of the active site of MHPCO with FAD and MHPC bound. Water molecules are shown as nonbonded red spheres. Hydrogen bonds between MHPC, water molecules and the protein are shown as dashed lines.

7.3.3. Active Site of MHPCO. The active site of MHPCO is a large pocket lined mostly by hydrophobic and aromatic residues, *figure 7.3*. These residues include Tyr54, Tyr82, Tyr223, Phe358, Trp203, Asp 203, and Arg211. The hydroxyl groups of the tyrosine residues point into the active site, and the side chain of Arg211 is also positioned within

the active site. These residues are hydrogen bonded to the many water molecules that are found within the active site in both structures. Comparisons of the active sites of flavin aromatic monooxygenases whose structures are available show some similarities.
All structures have an absolutely conserved proline residue above the FAD isoalloxazine ring containing C7 and C8 (Pro295 in MHPCO). The other commonality is the presence of several aromatic residues at the back of the binding pocket. MHPCO has three tyrosine residues, two tryptophan residues, and two phenylalanine residues in the rear of the active site, and none form any direct contacts with the substrate MHPC (**1**). *p*-Hydroxybenzoate hydroxylase, PHBH (PDB ID 1pbe)(*14*) and 2,6-dihydroxypyridine hydroxylase (*15*) have a similar wall in the active site. p-Hydroxybenzoate, p-OHB, the substrate of PHBH, forms a hydrogen bond to a tyrosine hydroxyl group, which can deprotonate p-OHB for activation. The carboxylate group of p-OHB forms a salt bridge with an arginine side chain, helping correctly orient the substrate within the active site. The binding of MHPC within the active site of MHPCO is quite different. Unlike PHBH, MHPCO makes no direct interactions with its substrate, *figure 7.3*. The 3-hydroxyl group of MHPC forms a water-mediated hydrogen bond to Tyr223, and the N1 atom is hydrogen bonded to Tyr82 through a water molecule. The carboxyl oxygen atoms are each hydrogen bonded to two water molecules that lead into the tunnel. Arg211 is one of the few charged residues found in the active site, and it is poorly positioned to help stabilize MHPC. The surprising lack of direct interactions between MHPC (**1**) and MHPCO and the large open cavity of the active site is unusual and differentiates MHPCO from other members of this enzyme family.

7.3.4. Mechanistic implications of the MHPCO Structure. Four plausible mechanism for the catalytic transformation of MHPC (**1**) is shown in *figure 7.4*. All four mechanisms are consistent with all previous $^{18}O_2$ and $H_2{}^{18}O$ labeling experiments, substrate analogue

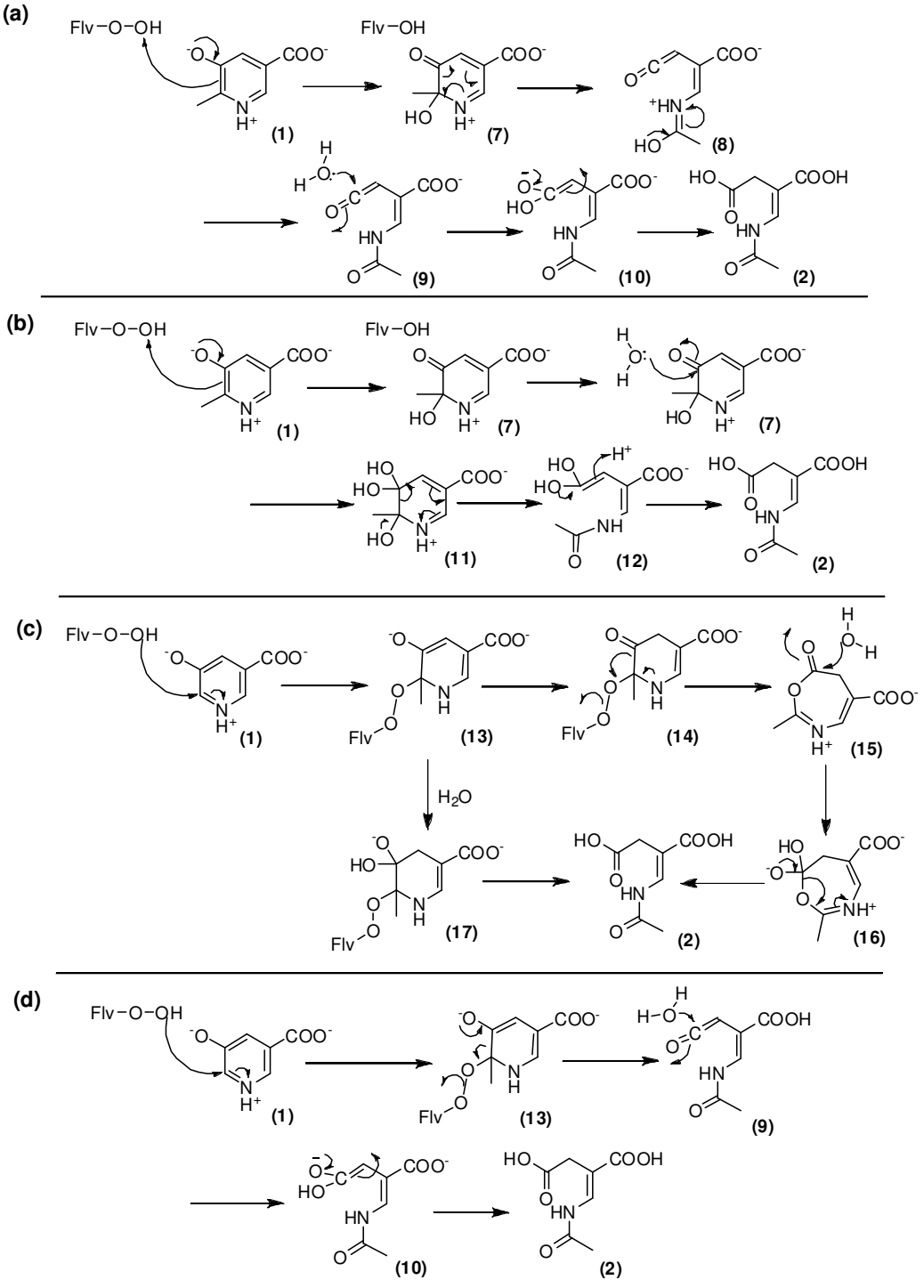


Figure 7.4: Possible mechanisms for the reaction catalyzed by MHPCO. Flavin hydroperoxide can act both as an electrophile (a) and (b) or a nucleophile (c) and (d).

studies, and kinetics data (*3, 5-7, 16*). Mechanism (b) and (c) have been proposed previously by Ballou and co workers (*2*). Although there are no experimental evidences against the nucleophilic attack of the flavin hydroperoxide Ballou and coworkers favors electrophilic attack of the flavin hydroperoxide on the pyridine ring based on favorable sequence homology of MHPCO with other aromatic flavoenzymes which have been shown to react as electrophilic flavin hydroperoxide, chemical calculations and more similarity in certain properties of MHPCO with aromatic hydroxylases than nucleophilic monooxygenase.

Alternative mechanisms for the catalysis by MHPCO for both electrophilic and nucleophilic attack by the flavin hydroperoxide is shown in (a) and (d) respectively, *figure 7.4*. According to mechanism (a), the hydroxylation of MHPC (**1**) first occurs by an electrophilic attack by the flavin hydroperoxide on C2 of pyridine (**1**) to give (**7**), which could then undergo an electrocyclic ring opening to produce ketene (**8**). Tautomerization to (**9**) followed by addition of water to the ketene would generate the product *E*-2-acetaminomethylene succinate (**2**). Such a ketene intermediate can also be traversed by the route of nucleophilic attack by the flavin hydroperoxide on the pyridine ring, *figure 7.4(d)*. The arguments against nucleophilic attack of the flavin hydroperoxide are still valid with the alternative mechanistic hypothesis, but in absence of hard experimental evidences it is difficult to rule it out with certainty. The current structure of MHPCO bound with MHPC (**1**) seems to position the substrate optimally for hydroxylation by the flavin hydroperoxide.

The structure of MHPCO complexed with the substrate MHPC (**1**) clarifies many of the details of the active site and provides further insight into the catalytic mechanism. The C2 of MHPC (**1**) is located 5.3 Å from C4a of the isoalloxazine ring. This separation, in addition to the orientation of the planes of the pyridine and the isoalloxazine rings, gives a reasonable geometry for reaction of the distal oxygen atom of the flavin

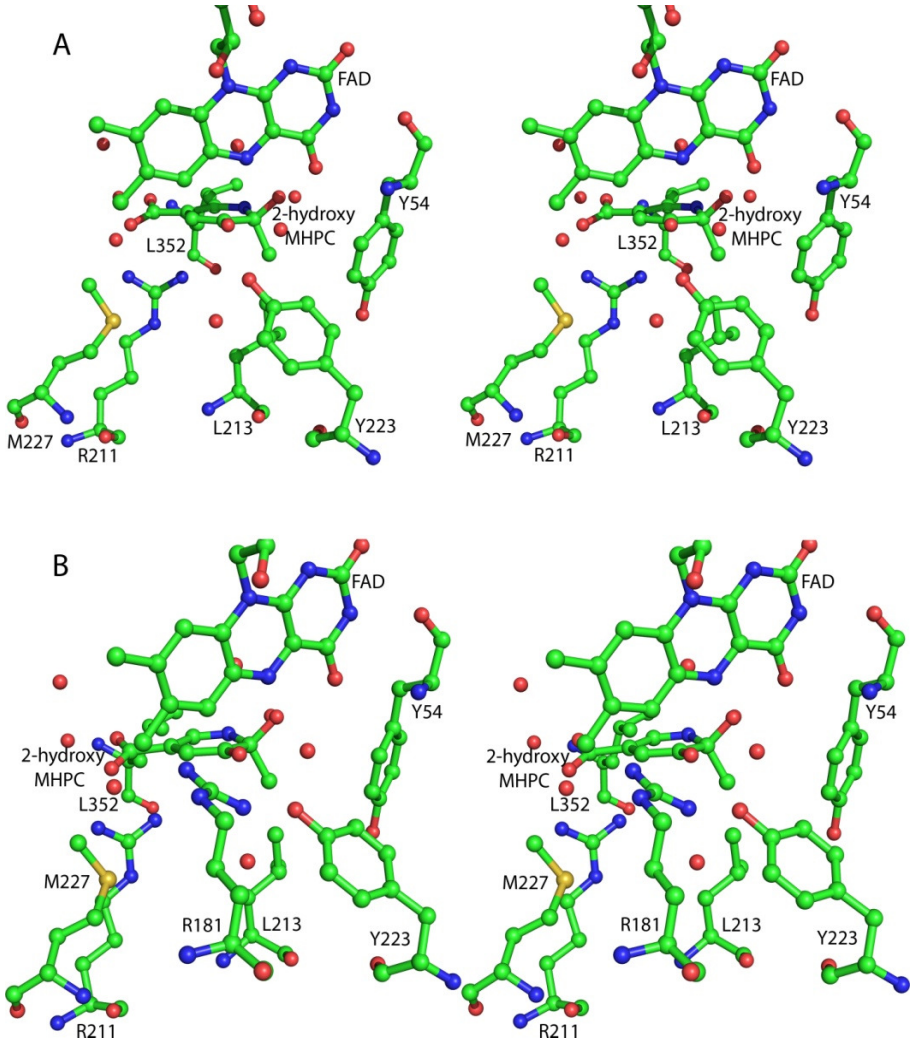


Figure 7.5: Stereoview diagrams of the active site with possible bases for the ring-opening reaction. All residues and ligands are shown in ball and stick representation and colored according to atom type. Red nonbonded spheres are water molecules. The 2-hydroxy-MHPC (**X**) was manually modeled into the active site using the binding of MHPC (**1**) as a guide. (A) Tyr223 as the active site acid. Tyr223 activates a water molecule for attack at the C3 carbonyl and could provide stabilization of the intermediate. (B) Arg181 as the active site base. Arg181 could protonate the carbonyl and stabilize the intermediate after attack of water at the carbonyl. Arg181 has adopted a different conformation than that seen in the crystal structure.

hydroperoxide with C2 of the substrate. The position of this oxygen is indicated by structurally conserved water molecules seen in both structures of MHPCO described here. Comparable substrate-C4a distances have been observed in other aromatic ring oxidizing flavin-dependent monooxygenases. Tyr223 and Tyr82 are likely catalytic residues, *figure 7.5(A)*. Tyr223 could activate the water molecule located 2.7 Å from the substrate C3 oxygen and facilitate the conversion of (**1**) to (**8**). Tyr82 is hydrogen bonded to the water molecule located 2.6 Å from the pyridine nitrogen and may facilitate the conversion of (**8**) to (**9**). Arg181 is also a possible candidate for an active site base. In the crystal structure, this residue is oriented along the isoalloxazine ring, but a conformational change of the side chain may position the guanidinium group within hydrogen bonding distance of the substrate C3 oxygen (modeled in *figure 7.5B*). However, for Arg181 to adopt this conformation some conformational changes of the active site would have to occur to avoid close contacts between the isoalloxazine ring and the arginine.

Removal of the hydroxyl group of Tyr223 (Y223F) or substitution of the guanidinium group of Arg181 with an amide group (R181Q) both render the enzyme inactive. Arg211 is located below the plane of the pyrimidine ring, and its guanidinium group is facing the C5 carboxylate of MHPC. While poorly positioned to form a salt bridge in the enzyme-substrate complex, conformational changes during the course of the reaction could increase the strength of this interaction.

7.3.5. Model chemistry. There has been no precedence of a ketene intermediate (**8**), *figure 7.4 (a, d)* in an enzyme catalyzed reaction. Also, there are no evidence for or against the electrophilic and the nucleophilic mechanism. To explore the possibility of such reactions and to investigate the mechanism we decided to do the oxygenation reaction with model compounds which are easy to handle. Two such model compounds, 3-hydroxy-5-(methoxycarbonyl)-1-methylpyridinium (**4**) and 3-methoxy-5-

(methoxycarbonyl)-1-methylpyridinium (**6**) were synthesized and their oxygenation in presence of the peracid and hydrogen peroxide were studied.

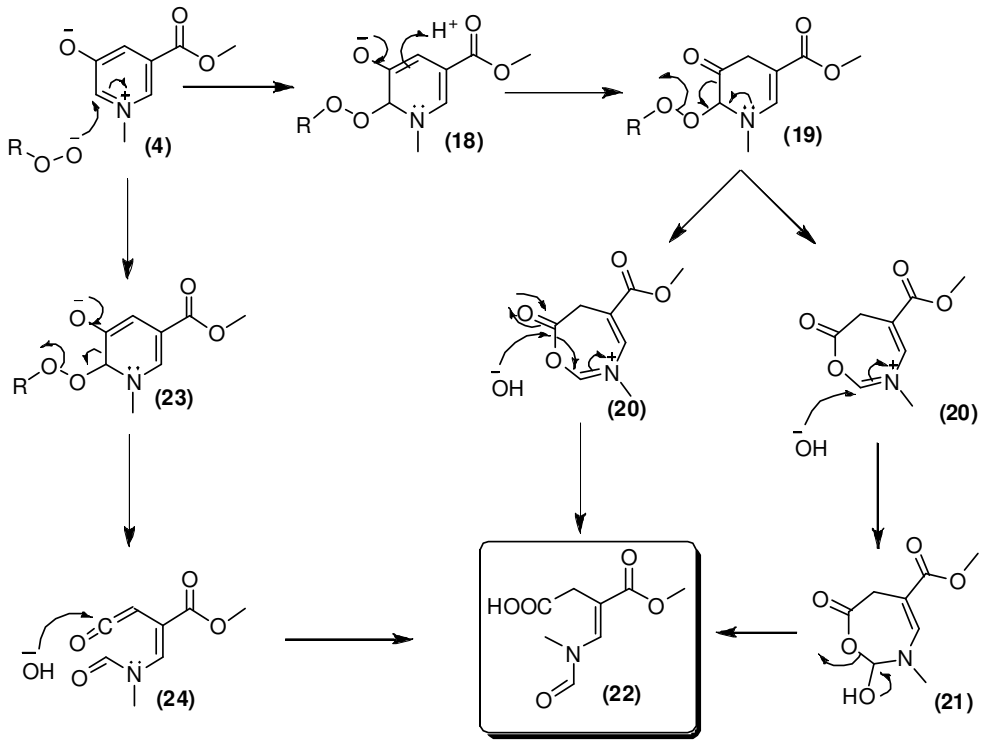


Figure 7.6: Postulated reaction scheme for oxidation of 3-hydroxy-5-(methoxycarbonyl)-1-methylpyridinium (**4**) in presence of peracids and peroxides.

7.3.6. 3-hydroxy-5-(methoxycarbonyl)-1-methylpyridinium (4). Possible reaction of 3-hydroxy-5-(methoxycarbonyl)-1-methylpyridinium (**4**) with peracid and H_2O_2 is shown in *figure 7.6*. There are two pathways in which the final ring open product (**22**) can be formed. The first one involves the ketene intermediate (**24**) and the second involves the Bayer-Villiger type of ring expansion of (**19**) to give the lactone (**20**) which can then undergo a nucleophilic attack by water to give the final ring opened product (**22**). It was treated with hydrogen peroxide as well as peracid (in presence and absence of base) at 25 °C for 1 hour and then analyzed by HPLC. No new product was seen. Reaction samples in which base was present hydrolyzed the ester (**4**) to form N-methyl-5-hydroxy-

nicotinic acid. No new species were seen on elevating the temperature to 50 °C. The fact that 3-hydroxy-5-(methoxycarbonyl)-1-methylpyridinium (**4**) did not react with peracid or peroxide was not very surprising as the presence of the hydroxide group at C3 makes C2 of the pyridine ring less electrophilic for nucleophilic attack of the peracid or the peroxide.

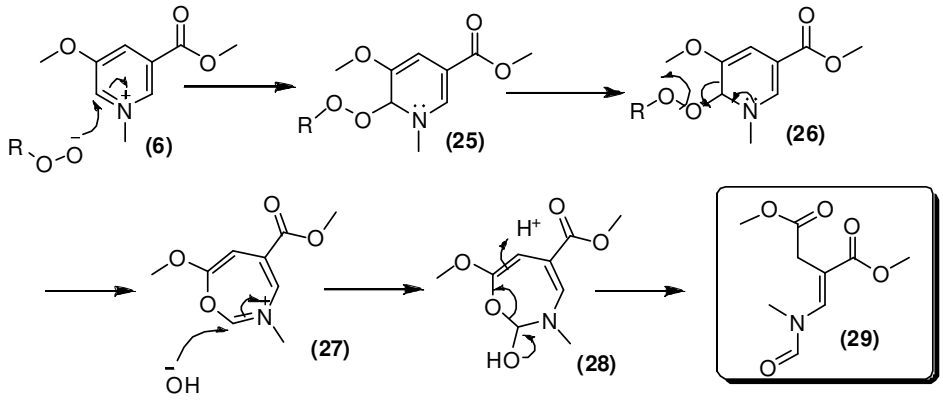


Figure 7.7: Postulated reaction scheme for oxidation of 3-methoxy-5-(methoxycarbonyl)-1-methylpyridinium (**6**) in presence of peracids and peroxides.

7.3.7. 3-methoxy-5-(methoxycarbonyl)-1-methylpyridinium (6). Possible reaction of 3-methoxy-5-(methoxycarbonyl)-1-methylpyridinium (**4**) with peracid and H_2O_2 is shown in *figure 7.7*. Unlike compound (**4**), (**6**) can form the ring opened product (**29**) in only one way. The ketene intermediate cannot be formed because of the methoxy group. It was treated with hydrogen peroxide as well as peracid as mentioned before in the presence and absence of base at 25 °C and 50 °C and analyzed by HPLC. As in the previous case the presence of base caused hydrolysis of the (**6**) to form N-methyl-5-methoxy-nicotinic acid. One hour treatment at 25 °C with H_2O_2 some of (**6**) is converted into a new species which elutes out at 10.0 min, *figure 7.8*. On addition of H_2O_2 premixed with NaOH to (**6**) the same species is seen at 10.1 min along with two other peaks at 8.4 and 8.9 min. The

peak at a 10.0 min (for only H_2O_2) and 10.1 min (for H_2O_2 premixed with NaOH) are likely to be the same since they have the same absorption spectra and the presence of base in a non buffered solvent system could change the retention time of the species. This new species is highly chromophoric with a very high extinction coefficient. Attempt to isolate this compound over multiple injections by HPLC have failed to generate sufficient amount to characterize by NMR. The tiny tail of the peak corresponding to (**6**), present at 10.0 min is all that is seen after collection of 10.0 min peak over multiple injections of a ten fold scaled up reaction mixture.

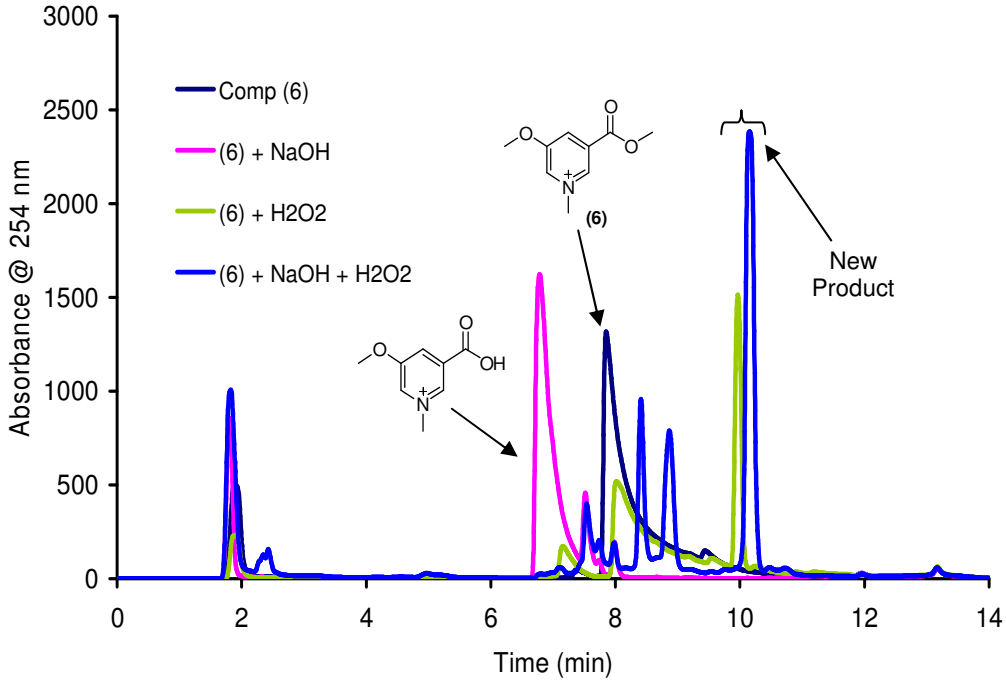


Figure 7.8: Treatment of (**6**) with H_2O_2 in presence and absence of NaOH-HPLC trace.

The same new peak at 10.0 minutes was seen when mCPBA was used in presence and absence of base. Clearly essentially the same chemistry is happening in this case as well. There is much more of the new compound with retention time of 10.0 min in case of (**6**) treated with mCPBA premixed with K_2CO_3. This peak also has similar absorption profile

as the one produced on treatment of (**6**) with H_2O_2.

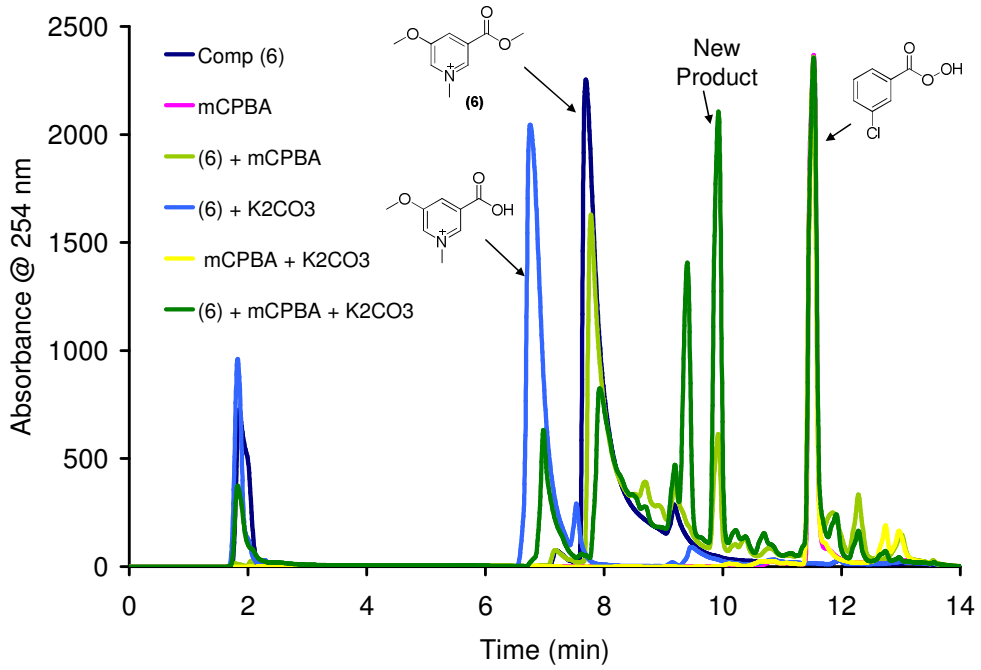


Figure 7.9: Treatment of (**6**) with mCPBA in presence and absence of K_2CO_3-HPLC trace.

This product needs to be characterized. It could likely be the ring opened product (**29**) in which case if not NMR (for the difficulty in isolation) at least ESI-MS would be able to confirm it. The high extinction coefficient lends hope that it could be the ring opened product since *E*-2-(acetamidomethylene)succinate (**2**), which is the enzymatic product of MHPCO, also has a very high extinction coefficient.

Methyl-5-methoxy nicotinate which lacks the N-methyl group does not undergo any reactions on treatment with H_2O_2 both in presence and absence of NaOH. The presence of methyl group on the nitrogen of the pyridine ring in (**4**) and (**6**) is essential to maintain the positive charge on the nitrogen which in turn makes the position C2 of the pyridine ring susceptible for electrophilic addition reaction with peracids and/or

peroxides. More over the presence of the methoxy group also make the position C2 of the pyridine ring more electrophilic in (**6**) than (**4**), which facilitate the attack of peracid as well as peroxide and do the ring opening chemistry.

7.3.8. Engineering the enzyme as a mechanistic probe. In order to distinguish between the electrophilic and the nucleophilic mechanism an effort was made to engineer the enzyme active such that one can bias the enzyme towards the formation of a desired product which can be formed by only one mechanism. When 5-hydroxynicotinate (**30**) was used as a substrate for MHPCO, it did not undergo rearomatization but instead produced the ring opened product, α-(N-formyl-N-methyl-amino-methylene) succinate (*2*). If flavin hydroperoxide acts as a nucleophile then it will form intermediate (**31**) and if it acts as an electrophile it would form the intermediate (**34**), *figure 7.10 (a, b)*. If by some mechanism the hydrogen at C2 (**31** & **34**) could be abstracted then we would have a way to distinguish between the electrophilic and the nucleophilic mechanism. When we looked into the active site of the protein we found a hydrophobic residue (Leu 213) close to C2 of the pyridine ring which when mutated to an Asp or a Glu could act as a base and assists in the abstraction of the C2 hydrogen, thereby biasing the reaction towards the rearomatization of the pyridine ring (instead of ring opening) to form 5,6 dihydroxynicotinate (**33**), *figure 7.10*. If flavin hydroperoxide acts as a nucleophile, for the ring opening to happen the C-C bond in the intermediate (**13**, **14** & **31**) must be transperiplanar to O-O bond, which excludes the possibility of the C-H bond being transperiplanar to the O-O bond. Thus, when the mutant of MHPCO is used and if the reaction goes via flavin hydroperoxide acting as a nucleophile, then the reaction will stop at the stage of the intermediate (**31**). This is because the C-H bond is not trans-periplanar with the O-O bond to facilitate its cleavage in order to release flavin hydroxide. If 5,6 dihydroxynicotinate (**33**) is seen, then it has to be via the electrophilic mechanism.

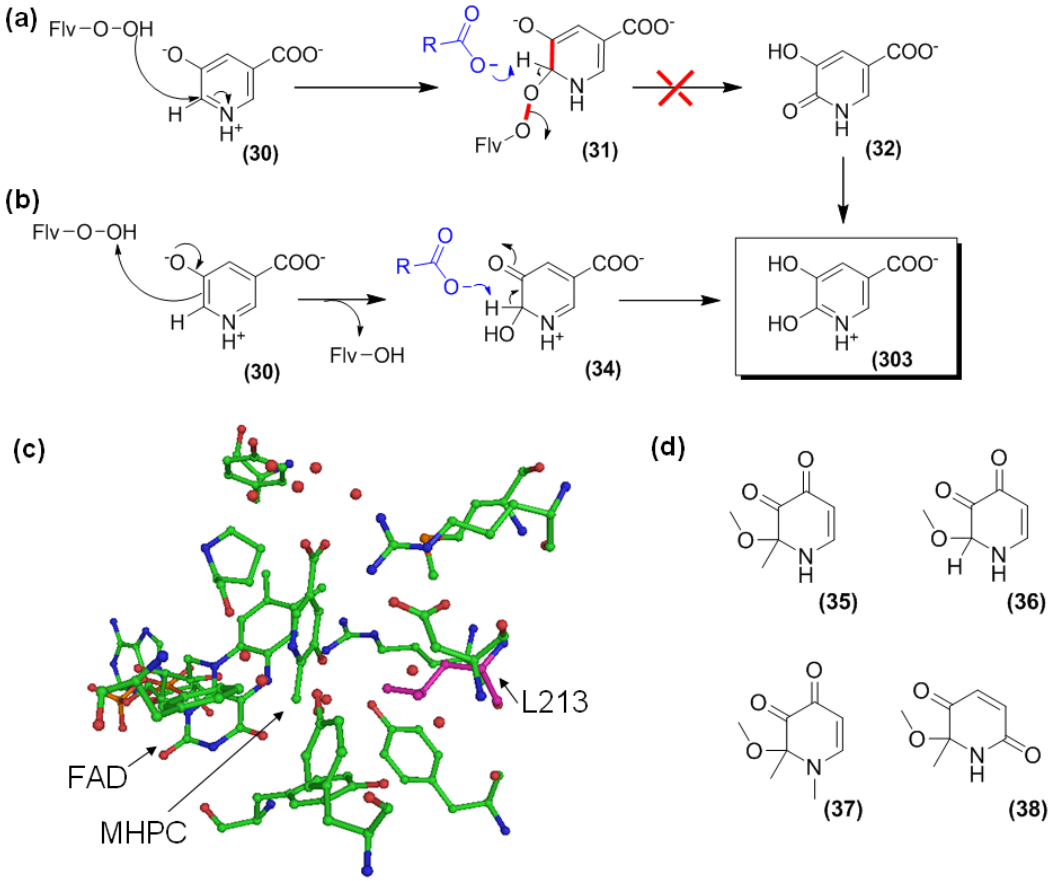


Figure 7.10: Postulated mechanism of MHPCO mutant in which an active site base is introduced with 5-hydroxynicotinate (**30**) when flavin hydroperoxide acts as (a) a nucleophile and (b) an electrophile. (c) The active site of MHPCO showing Leu 213 which could be mutated to Asp or Glu. (d) Few examples of 2-methoxy pyridine derivatives which are stable and do not undergo spontaneous rearomatization.

2-methoxy pyridine derivatives do not spontaneously rearomatize and are stable which suggests that (**33**) can be formed only via the electrophilic mechanism *figure 7.10(b)*. Leu 213 was mutated to Asp and Glu. Both the mutants did not turnover (**30**).

However L213D mutant was found to oxidize NADH in absence of the substrate. In the wild type MHPCO, the reduction of the flavin by NADH is greatly enhanced by the presence of the substrate within the active site. The L213D mutant

seems to have decoupled these two events. It can be speculated that, the presence of a highly charged, hydrophilic residue in the floor of the active site (containing only hydrophobic residues) could somehow mimic the charges seen on the natural substrate MHPC or trigger a change in the FAD binding from the IN conformation to the OPEN (or READY) conformation, where it can be reduced by NADH. The water molecules could relay these changes to the FAD in the absence of MHPC.

7.4. Conclusion

The work done towards elucidation of the mechanism of the reaction catalyzed by MHPCO is far from complete and lot more studies must be done to prove the mechanism. Chemical synthesis of compound (**29**) must be done and co-injected with the compound (**6**) and H_2O_2 reaction mixture to see if it co-elutes with the new product that was seen. If this is the case, then the nucleophilic attack of the flavin hydroperoxide on the pyridine ring is the mechanism by which the reaction proceeds. Moreover, no ketene intermediate can be formed from compound (**6**). This information is valuable to distinguish between all the plausible reaction mechanism for the very interesting and unique oxidative ring opening reaction catalyzed by MHPCO.

7.5. Acknowledgement.

We thank Dr. Yasunobu Ohkawa, National Institute of Agrobiological Sciences, Ibaraki, Japan for providing us with the *M. loti* MAFF303090 strain. The authors are grateful to the staff scientists at the APS NE-CAT beamlines for their assistance in data collection. We thank the Cornell University Protein Facility for providing the clones.

REFERENCES

(1) van Berkel, W. J., Kamerbeek, N. M., and Fraaije, M. W. (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J Biotechnol 124*, 670-89.

(2) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Reaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid (MHPC) oxygenase with N-methyl-5-hydroxynicotinic acid: studies on the mode of binding, and protonation status of the substrate. *Biochemistry 36*, 13856-13864.

(3) Kishore, G. M., and Snell, E. E. (1981) Interaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase with FAD, substrates, and analogues. Spectral and fluorescence investigations. *J Biol Chem 256*, 4234-4240.

(4) Chaiyen, P., Sucharitakul, J., Svasti, J., Entsch, B., Massey, V., and Ballou, D. P. (2004) Use of 8-substituted-FAD analogues to investigate the hydroxylation mechanism of the flavoprotein 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 43*, 3933-3943.

(5) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Unusual mechanism of oxygen atom transfer and product rearrangement in the catalytic reaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 36*, 8060-8070.

(6) Kishore, G. M., and Snell, E. E. (1981) Kinetic investigations on a flavoprotein oxygenase, 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *J Biol Chem 256*, 4228-4233.

(7) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Thermodynamics and reduction kinetics properties of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 36*, 2612-2621.

(8) Oonanant, W., Sucharitakul, J., Yuvaniyama, J., and Chaiyen, P. (2005) Crystallization and preliminary X-ray crystallographic analysis of 2-methyl-3-

hydroxypyridine-5-carboxylic acid (MHPC) oxygenase from Pseudomonas sp. MA-1. *Acta Crystallogr Sect F Struct Biol Cryst Commun 61*, 312-4.

(9) Ausubel, F. M., and Brent, F. (1987) in *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.

(10) Sambrook, J., Fritsch, G. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Guide*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

(11) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem 72*, 248-254.

(12) Mukherjee, T., McCulloch, K. M., Ealick, S. E., and Begley, T. P. (2007) Gene identification and structural characterization of the pyridoxal 5'-phosphate degradative protein 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase from mesorhizobium loti MAFF303099. *Biochemistry 46*, 13606-15.

(13) Yuan, B., Yokochi, N., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2006) Molecular cloning, identification and characterization of 2-methyl-3-hydroxypyridine-5-carboxylic-acid-dioxygenase-coding gene from the nitrogen-fixing symbiotic bacterium Mesorhizobium loti. *J Biosci Bioeng 102*, 504-510.

(14) Wierenga, R. K., de Jong, R. J., Kalk, K. H., Hol, W. G., and Drenth, J. (1979) Crystal structure of p-hydroxybenzoate hydroxylase. *J Mol Biol 131*, 55-73.

(15) Treiber, N., and Schulz, G. E. (2008) Structure of 2,6-dihydroxypyridine 3-hydroxylase from a nicotine-degrading pathway. *J Mol Biol 379*, 94-104.

(16) Kishore, G. M., and Snell, E. E. (1981) Interaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase with FAD, substrates, and analogs. Spectral and fluorescence investigations. *J. Biol. Chem. 256*, 4234-40.

CHAPTER 8

A novel pyrimidine degradation pathway in *Escherichia coli* K12*

8.1 Introduction

There are two well characterized catabolic pathways for pyrimidines. The first one is the reductive pathway (*1*), which is more prevalent, has been discovered in bacteria, archaea as well as in humans. The pyrimidine bases are first reduced to their respective dihydro derivatives (**2**), which then undergoes two hydrolysis reactions to produce ammonia, carbon dioxide and β-alanine (**4**) as end products. The other pathway is the oxidative pathway which is less prevalent and has been reported in some bacteria. The pyrimidine bases undergo oxidation to form barbituric acid (**5**) which undergoes two hydrolysis reactions to produce urea (**7**) and malonic acid (**8**) (*2*), *figure 8.1*.

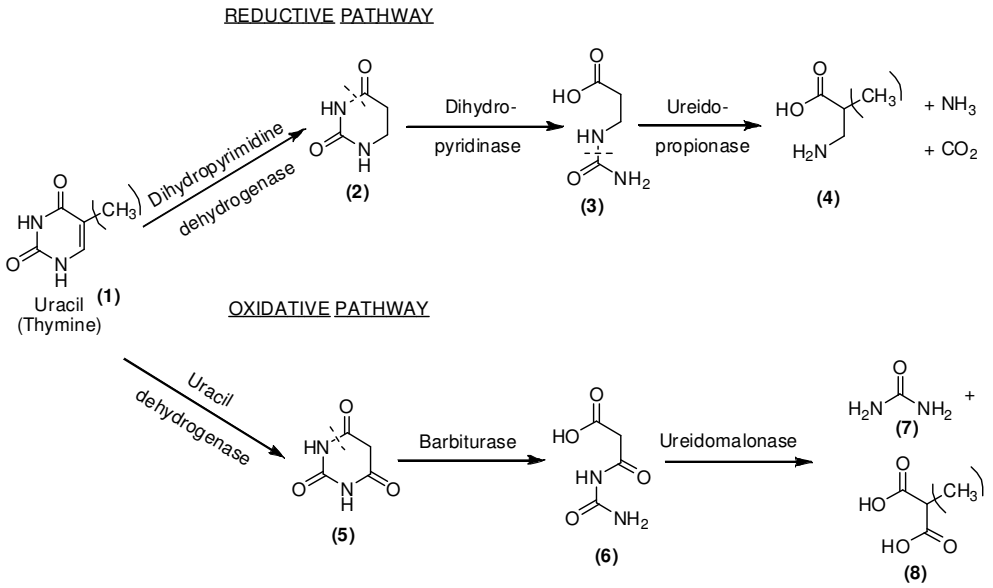


Figure 8.1: Two well known pathways of pyrimidine catabolism.

Apart from these two pathways, another new pathway of degradation of uracil (**1**) (or thymine) was recently discovered in *Escherichia coli* K12 (*3, 4*). Even though the

* Synthesis and Molecular cloning was done by Dr. Sameh Abdelwahed and Dr. Cynthia Kinsland respectively

genome of *E.coli* was sequenced a decade ago, date one-fifth of it still remained uncharacterized. The b1012 operon in *E.coli* was one of the largest clusters of uncharacterized genes which were regulated under the control of the transcriptional activator, nitrogen regulatory protein C (NtrC). Kustu and coworkers, made mutants for each of the genes in the b1012 operon and the *E.coli* mutants were grown in the presence of various nitrogen sources. It was found that the wild type grew in the presence of uracil (**1**) and thymine as the sole nitrogen source at 25 ˚C (not 37 ˚C) while none of the mutants did. This highlighted the importance of all the genes in the operon participating in a catabolic pathway of uracil (**1**) (and thymine) which ultimately produced utilizable nitrogen in the form of ammonia. Each molecule of uracil (**1**) produces two molecules of ammonia, one molecule of carbon dioxide and 3-hydroxy propionate (**9**), *figure 8.2*. The b1012 operon which contained seven genes of unassigned functions were named as *rutA-rutG* (py**r**imidine **ut**ilization). However, the enzymology and intermediates involved in this pathway were not known. In this chapter the identification and characterization of the intermediates in this new pathway are reported.

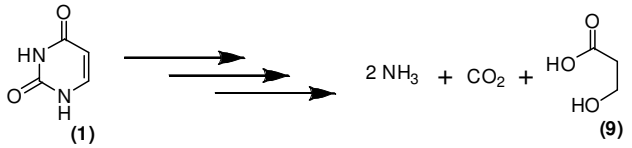


Figure 8.2: A novel pathway of uracil (**1**) degradation in *E.coli* K12 (*3*).

8.2 Experimental

8.2.1. Cloning, over-expression and purification. The pCA24N plasmid containing genes encoding for the Rut proteins (b1012, b1011, b1010, b1009, b1008 and b1007 genes) was transformed into *Escherichia coli* BL21(DE3). A starter culture was prepared by growing a single colony of transformed cells in 10 ml of LB media containing 20 μg/ml of chloramphenicol at 37 °C with overnight agitation. 1 liter LB medium (20 g/L),

containing 20 μg/ml of chloramphenicol, was inoculated with this starter culture. The cells were grown at 37 °C with shaking until the culture reached an OD_{600} of 0.6 at which point they were induced by adding IPTG to a final concentration of 0.5 mM, the temperature was lowered to 15 °C and the cells were allowed to grow for a further 12 hours. The cells were then harvested by centrifugation at 10,000g for 8 min at 4 °C.

Cells from 1 liter of culture were re-suspended in 20 ml of binding buffer (50 mM KH_2PO_4, 150 mM NaCl, 10 mM imidazole, pH 7.7) and approximately 2 mg of lysozyme was added. The cells were then lysed by sonication (Misonix Sonicator 3000, pulse 'on' time 1.0 sec, pulse 'off' time 1.0 sec, output level 0.8, 30 cycles) on ice for 5 times. The cell debris was removed by centrifugation at 39,000g for 40 minutes at 4 °C. The clarified supernatant was loaded onto a 2 mL Ni-NTA-affinity column pre-equilibrated with binding buffer kept at 4°C. The Ni-NTA-affinity column was then washed with 100 ml wash buffer (50 mM KH_2PO4, 150 mM NaCl, 20 mM imidazole, pH 7.7). The protein was eluted from the column with elution buffer (50 mM KH_2PO_4, 150 mM NaCl, 200 mM imidazole, pH 7.7) at 4°C. The fractions containing protein were pooled and concentrated using YM-10 Amicon Ultracentrifugal filters at 5000g to a final volume of 500 μl. The concentrated sample was desalted into 100 mM phosphate buffer at pH 8.0 containing 100 mM NaCl and glycerol to a final concentration of 30% using an Econo-Pac 10DG disposable chromatography column.

RutA, RutB, RutC and RutD over expressed well and soluble protein was obtained upon purification by Ni-affinity chromatography, *figure 8.3*. 25, 23, 12 and 11 mg of RutA, RutB, RutC and RutD respectively were obtained per liter of cell culture. The protein concentration was measured by the Bradford assay (*5*). RutE and RutF had good over expression patterns but very little soluble protein was obtained upon purification. RutE was subcloned into a pET28b vector and soluble protein was obtained. RutF was subcloned into pET28b vector and then had to be co-expressed with chaperon plasmid

pG-Tf2 (encoding for chaperone proteins groES-groEL-tig, marketed by the TaKaRa Bio Inc.) in order to get the soluble fraction. Kanamycin (40 mg/L) was used in the later case as the choice of antibiotic.

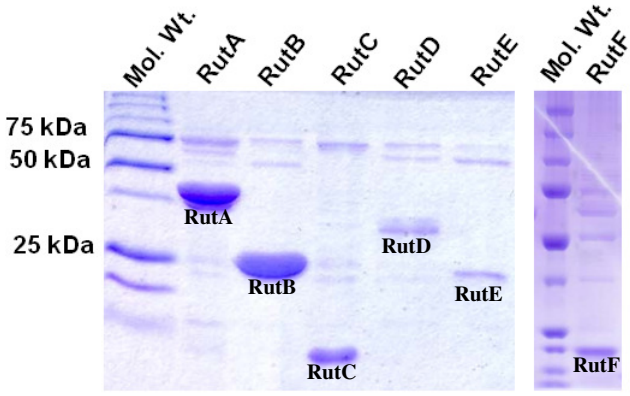


Figure 8.3: SDS-PAGE analysis showing pure proteins of the Rut pathway which were purified by Ni-affinity chromatography.

The clone for the Fre protein did not code for 6xHis tag. The *fre* gene was excised from the pET30LIC vector by digestion with *Nde*I and *Eco*RI and ligated into similarly digested pTHT (which is a pET-28 derived vector which allows attachment of a modified 6xHisTag followed by a TEV protease cleavage site onto the N-terminus of the expressed protein) to construct pFRETHT[1]. It was then over expressed and purified by the Ni-affinity chromatography according to the procedure mentioned above. 60 mg of protein was obtained per liter of cell culture.

8.2.2. HPLC analysis. HPLC analysis of the enzymatic reaction mixture was performed on a Hewlett-Packard 1100 instrument using a Supelcosil LC 18 T (15 cm X 4.6 mm, 3.0 μm) column. Two different methods were used.

Method A: Solution A contained water, solution B contained 100 mM potassium

[1] The molecular cloning was done by Dr. Cynthia Kinsland

phosphate buffer at pH 6.6 and solution C contained methanol. The following linear gradient was used. 0% to 10% solution A and 100% to 90% solution B for 0-5 min, 10% to 48% solution A, 90% to 40% of solution B and 0% to 12% of solution C from 5-12 min, 48% to 50% solution A, 40% to 30% of solution B and 12% to 20% of solution C in 12-14 min, 50% to 30% solution A, 30% to 10% of solution B and 20% to 60% of solution C in 14-18 min, 30% to 0% solution A, 10% to 100% of solution B and 60% to 0% of solution C in 18-20 min and 0% of solution A, 100% of solution B and 0% of solution C in 20-25 minutes. The flow rate was 1 ml/min and the absorbance was measured at 254 nm. Under these conditions the following compounds were readily separated (retention time in parenthesis): 3-ureidoacrylate (**16**) (2.3 min), uracil (**1**) (3.7 min), NAD (12.9 min) and NADH (14.3 min).

Method B: Solution A contained water with 0.1% TFA and solution B contained acetonitrile with 0.1% TFA. The following linear gradient mixing solution A with solution B was used: 100% solution A for 0 to 2 min, 100% to 60% solution A from 2 7 min, 60% to 0% solution A in 7-12 min, 0% A from 12-14 min, 0% to 100% solution A in 14-15 min and 100% solution A in 15-17 minutes. Flow rate was 1 ml/min and the absorbance was measured at 370 nm. The following compounds were readily separated (retention time in parenthesis): 2,4-dinitrophenylhydrazine (**18**) (11.6 min) and DNP hydrazone of malonic semialdehyde (cis/trans (**19**/**20**) 12.42 min and trans/cis (**20**/**19**) 14.0 min).

8.2.3. Synthesis

8.2.3.1. 3-hydroxyuracil (13)†. Literature procedure (*6*) was followed for the two step synthesis of 3-hydroxyuracil (**13**) from cytidine (**10**) *figure 8.4(a)*. ^{1}H NMR (300 MHz, D_2O): δ 38.33 (d, 1H), 6.76 (d, 1H), 6.34 (d, 1H), 4.8-4.0 (m, 5H).

† The synthesis was done by Dr. Sameh Abdelwahed.

8.2.3.2. 3-ureidoacrylate (16). The synthesis of 3-ureidoacrylate (**16**) began with the known protocols for making 3-oxauracil (**15**) from maleic anhydride (**14**) (*7*) *figure 8.4(b)*, which was converted to 3-ureidoacrylate (**16**) with minor modification of literature procedure (*8*). 3-Oxauracil (**15**) (1.13 g, 0.01 mol; well ground) was added to a solution of ammonia in water (20 mL) at 5 °C with constant stirring over 10 min till all of it dissolved, after which the reaction mixture was allowed to stand for another 12 hrs at 4°C instead of 25 °C as reported in the literature. There was no change in the workup procedure and quantitative yields of 3-ureidoacrylate (**16**) was obtained. This is a vast improvement on the literature procedure where the reported yield is 12%. ^{1}H NMR (DMSO, 300MHz) δ 4.73 (d, 1H), 6.80 (br s, 2H), 7.32 (dd, 1H), 9.75 (br d, 1H).

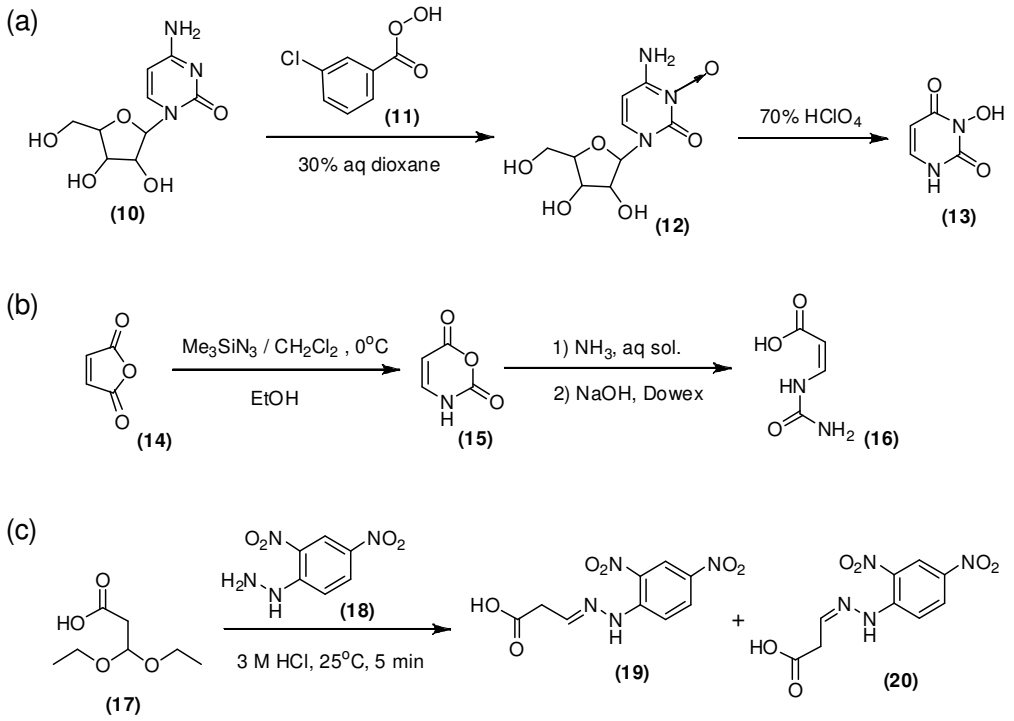


Figure 8.4: Synthetic schemes for the preparation of substrates and reference compounds (a) 3-hydroxyuracil (**13**), (b) 3-ureidoacrylate (**16**) and (e) DNP hydrazone of malonic semialdehyde (**19**/**20**).

8.2.3.3. DNP hydrazone of malonic semialdehyde (19, 20). 3,3-diethoxypropanoic acid (**17**) (10 µL) was added to 990 µL of 20 mM 2,4-dinitrophenyl hydrazine (**18**) prepared in 3 M HCl. The hydrazone formation was rapid and the orange product was extracted twice with 500 µL of ethyl acetate. The yellow colored ethyl acetate layer was washed with 1M HCl and the solvent was removed *in vacuo*. A yellow-orange powder was seen, (*figure 8.4(c)*). ^{1}H NMR (CD_3OD, 300MHz) δ 3.33 (d, 2H), 7.71 (t, 1H), 7.86 (d, 1H), 8.20 (dd, 1H), 8.91 (d, Hz, 1H).

8.2.4. Enzymatic Assays

8.2.4.1. RutA and RutF catalyzed reaction of uracil (1). A 500 µL of reaction mixture containing 200 µM of uracil (**1**), 5 µM of RutA, 5 µM of RutF, 40 µM of FMN and 2 mM of NADH solution was incubated at room temperature for 2 hours. All solutions were made in 100 mM potassium phosphate buffer at pH 8.0. The reaction mixture was filtered to remove protein using Microcon YM10 and analyzed by HPLC (method A), *figure 8.5*.

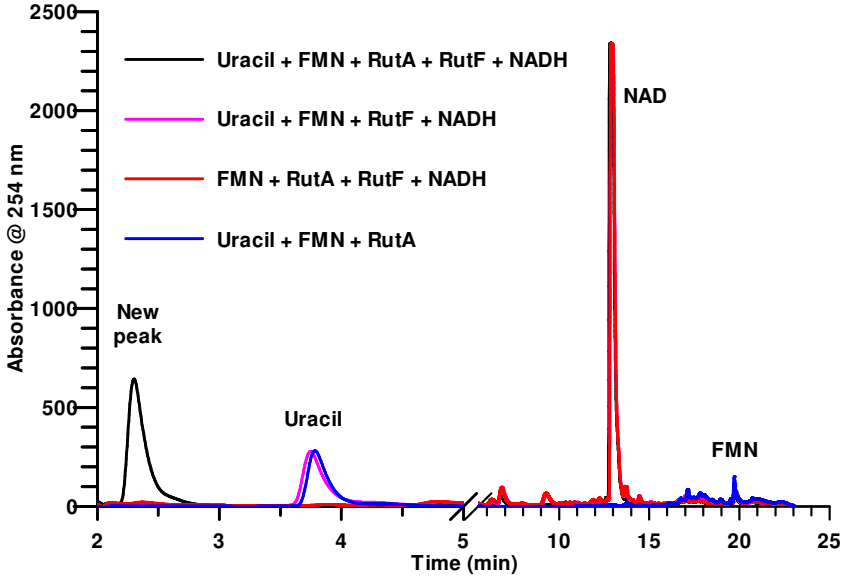


Figure 8.5: HPLC analysis showing the formation of a new product as a result of enzymatic reaction of RutA and RutF with uracil (**1**).

8.2.4.2. Assay for RutB.

8.2.4.2.1. HPLC assay. The product of RutA/F catalyzed reaction of uracil (**1**) was the substrate for RutB, which is the next enzyme in the pathway. A reaction mixture was set up as mentioned in section 8.2.4.1. The reaction mixture was divided into two equal halves. RutB was added to one of them to a final concentration of 8 μM while an equal volume of buffer was added to the other one. Both the reaction mixture as well as the control was incubated at 25°C for 30 minutes, filtered using Microcon, YM10 and analyzed by HPLC (method A). The substrate was totally consumed by RutB, *figure8.6*. The product of this reaction is non-chromophoric at 254 nm and could not be observed by the analytical method that was employed.

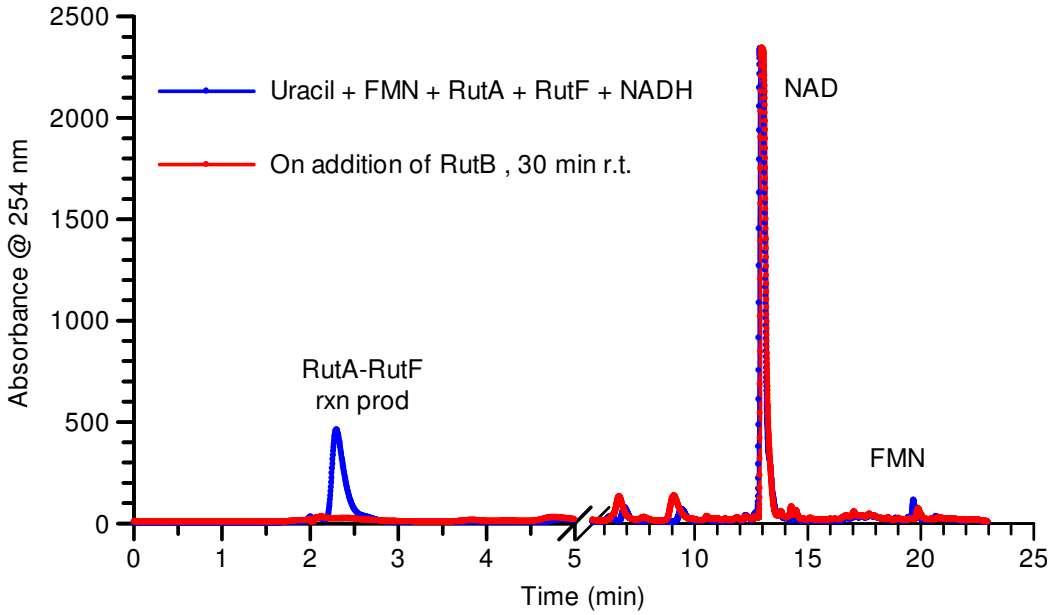


Figure 8.6: HPLC analysis showing the disappearance of the product of RutA/F catalyzed reaction of uracil (**1**) by RutB.

8.2.4.2.2. UV-vis assay – depletion of substrate. A time course was determined by monitoring the disappearance of 3-ureidoacrylate (**16**) by UV-visible spectrophotometry as shown in *figure 8.7*. UV-visible scans were taken every 100 sec from 360 nm to 200 nm

at the rate of 100 nm/sec. The reaction mixture (500 μL) contained 64 μM 3-ureidoacrylate (**16**) and 0.1 μM freshly purified enzyme in 100 mM potassium phosphate buffer at pH 8.0.

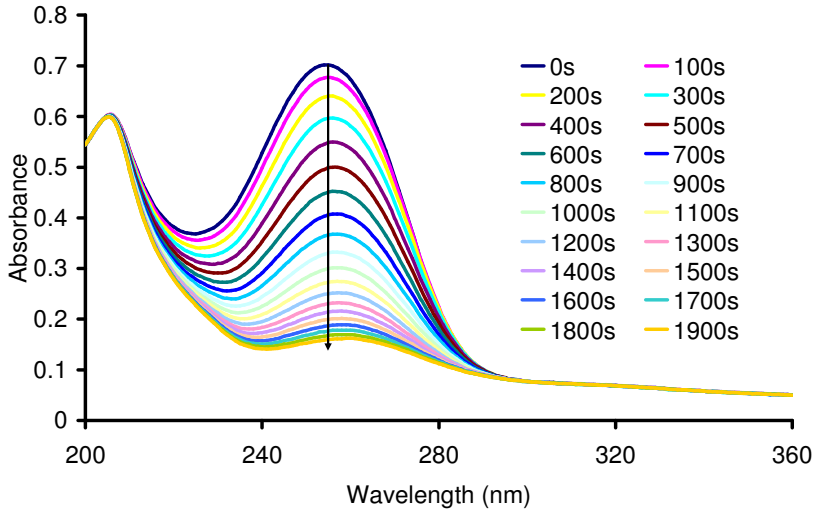


Figure 8.7: UV-visible scan of the RutB enzymatic reaction mixture over time.

8.2.4.2.3. Assay for ammonia production. The product of the RutA/F catalyzed reaction of uracil (**1**) was isolated by HPLC in 100 mM potassium phosphate buffer at pH 8.0 and the isolated sample was analyzed for the presence of ammonia. To 200 μL of the isolated 3-ureidoacrylate (**16**) (unknown concentration), 1 mM α-ketogluterate (**21**), 200 μM of NADPH and 2 U of glutamate dehydrogenase (unit is defined as the amount of glutamate dehydrogenase that will reduce 1 μmol of α-ketoglutarate (**21**) to glutamate (**22**) per min at pH 8.3 at 30 ˚C), 2 μL of 0.6 mM RutB (final concentration 2.4 μM) was added. The reduction of α-ketoglutarate (**21**) to glutamate (**22**) was not rate limiting under these assay conditions. The reaction was started by the addition of RutB and was monitored at 340 nm, characteristic for the absorbance of NADPH, *figure 8.8*. Controls were run in which one of the components of the enzymatic reaction mixture was missing. All the solutions were made in 100 mM potassium phosphate buffer at pH 8.0.

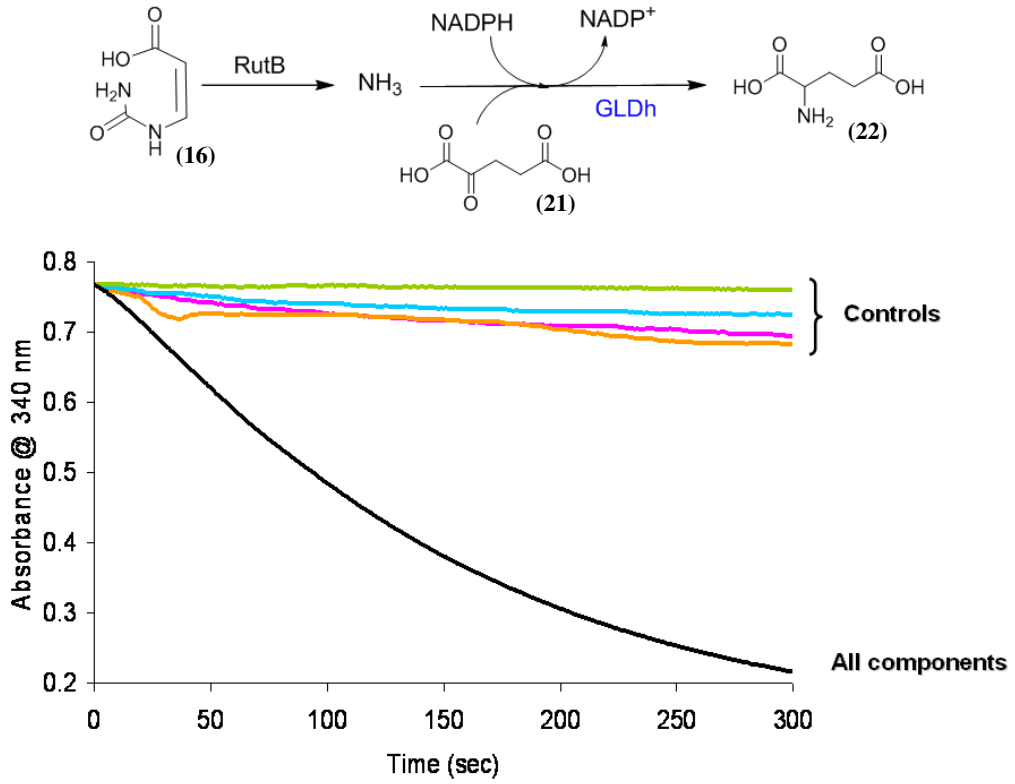


Figure 8.8: Coupled assay of RutB with glutamate dehydrogenase showing the production of ammonia as observed by the decrease in absorbance of NADPH.

8.2.4.2.4. Assay for the formation of 3-oxopropionate (33). 3-ureidoacrylate (**16**) was isolated from a scaled up reaction mixture containing uracil (**1**), RutA/F, NADH and FMN, incubated overnight at 25˚C by HPLC using method A. The product was collected in small aliquots and frozen immediately in liquid nitrogen and stored at -20˚C. This sample was used in this assay. 5 µL of 0.6 mM RutB was added to 2.5 mL of 3-ureidoacrylate (**16**) in 100 mM phosphate buffer at pH 7.0 at 25˚C. At various time points 300 µL of the reaction mixture was taken and added to 300 µL of 20 mM 2,4-dinitrophenyl hydrazine (**18**), mixed thoroughly and incubated at 25˚C for 5 minutes. It was extracted twice with ethyl acetate and the organic layer was pooled together. 100 µL of the organic layer was analyzed by HPLC using method B, *figure 8.9*.

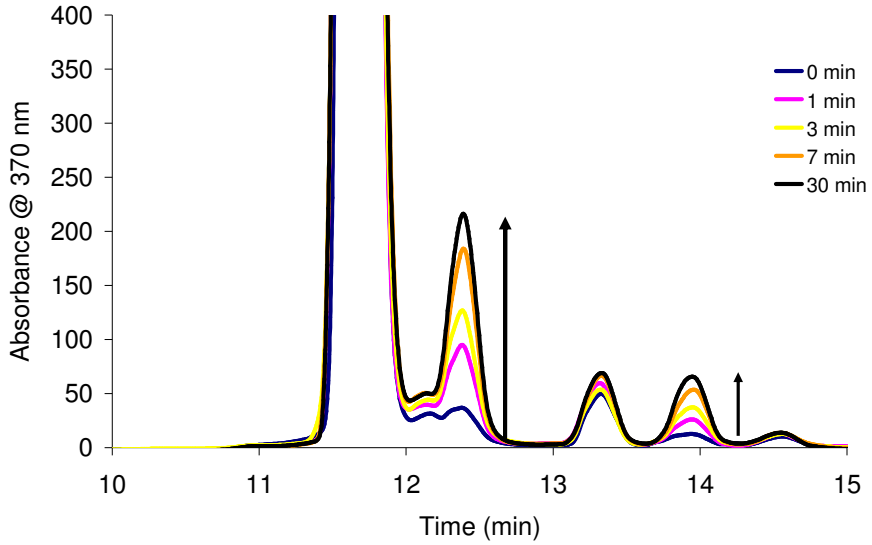


Figure 8.9: HPLC chromatogram showing the time course of the formation of 2,4 dinitrophenyl hydrazone of 3-oxopropionate (**33**) during the RutB assay.

8.2.5. Steady state kinetic parameters. The 3-ureidoacrylate (**16**) has a λ_{max} at 254 nm and the extinction coefficient at 254 nm was determined to be 12,200 $M^{-1}cm^{-1}$ in 100 mM phosphate buffer at pH 8.0. To a reaction mixture (500 µL) containing 100 nM enzyme, varying concentrations of 3-ureidoacrylate (**16**) were added. The rate of depletion of 3-ureidoacrylate (**16**) was monitored over 5 min. The K_M and k_{cat} for the enzyme were determined by fitting the initial rate (less than 10% conversion) of substrate depletion as a function of substrate concentration to the Michaelis-Menten equation by non-linear regression using Grafit 5.0.11 (Erithacus Software Ltd., Surrey, UK). All solutions were made in 100 mM potassium phosphate buffer at pH 8.0.

8.3. Results and Discusssion

8.3.1. Background information. The first enzyme involved in this pathway is RutA (b1012), which is an FMN dependent mono-oxygenase (*9*). It requires a partner protein which provides the reducing equivalent to FMN to form $FMNH_2$. The ideal candidate

to reduce flavin is RutF(b1007) which is a putative flavin reductase (*9*). However, RutF is not obtainable in pure soluble form. An alternative is the Fre protein which is a flavin reductase (*10*). Kustu and coworkers showed that RutA along with Fre catalyzed the first step in the Rut pathway. TLC assays were performed with C^{14} labeled uracil (**1**) and appearance of a new spot for the product was seen. The structure of the product was not known. They also showed that RutB(b1011) catalyzed the next step in the pathway.

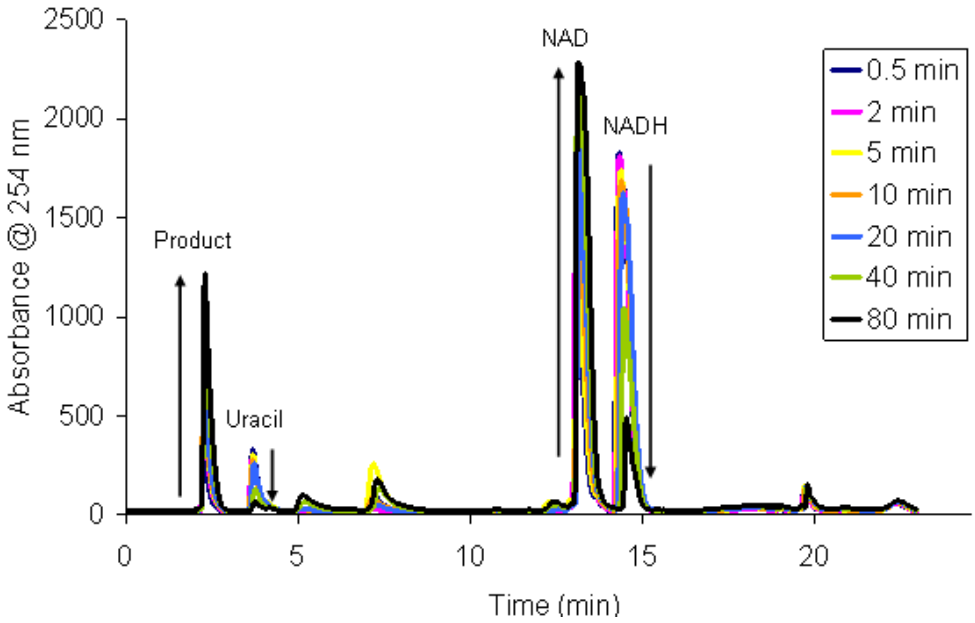


Figure 8.10: HPLC chromatogram of a time course of RutA and RutF reaction with uracil (**1**). The final concentration of uracil (**1**), protein, NADH and FMN are 400 mM, 20 mM, 2 mM and 40 mM respectively. The reaction was done in 100 mM phosphate buffer pH 8.0 at room temperature. Each time point was obtained by heat quenching a part of the reaction mixture.

8.3.2. Activity assay for RutA. RutA over expressed very well and purified out to near homogeneity (25 mg/ L of cell) by Ni-affinity chromatography. The RutF protein over expressed very well but had problems with its solubility. Efforts were made to solubilize it

by the usage of chaperone proteins, but the purity of the soluble fraction was not great. The RutF preparation was found to selectively use NADH instead of NADPH in order to reduce FMN. A coupled assay with RutF, RutA, FMN, NADH and uracil (**1**) showed product formation when each of the components of the reaction mixture was present, as shown in *figure 8.5*. No product was detected on substituting RutF with DTT and/or dithionite as alternative chemical reductants. A time course of the reaction showed the time dependant disappearance of uracil (**1**) and NADH and appearance of the product and NAD, *figure 8.10*.

8.3.3. Product isolation and characterization. All attempts to isolate and characterize the product by NMR and ESI-MS failed as the product was unstable to acid and would degrade on lyophilization. In order to characterize the product without isolation the enzymatic reaction was set up with $^{13}C^{15}N$ labeled uracil (**1**). 5 mM $^{13}C^{15}N$ labeled uracil (**1**) was incubated over night with 50 mM NADH, 200 μM FMN and 120 μM of RutA and RutF mixture, at 25 °C overnight. The final volume was 700 μL. A control reaction was set up in an identical manner with the enzymes were substituted with equal volume of desalting buffer. All solutions made in 25 mM D_2O phosphate buffer pH 8.0. The enzymes were desalted into 100 mM potassium phosphate buffer at pH 8.0 containing 100 mM NaCl and 2 mM DTT and glycerol was added to a final concentration of 50 %. The complete conversion of the substrate to the product was assessed by HPLC (data not shown). The control reaction, which is that of unreacted $^{13}C^{15}N$ labeled uracil (**1**), is shown in the top panel, *figure 8.11*. $J_{C4\text{-}C5}$ = 65.5 Hz, $J_{C5\text{-}C6}$ = 66.5 Hz, $J_{C6\text{-}N1}$ = 11.6 Hz, $J_{N1\text{-}C2}$ = 24.5/20.5 Hz, $J_{C2\text{-}N3}$ = 20.5/24.5 Hz and $J_{N3\text{-}C4}$ = 10.6 Hz. The bottom panel in *figure 8.11* shows the product ($^{13}C^{15}N$ labeled) under identical conditions. $J_{C4\text{-}C5}$ = 66.6 Hz, $J_{C5\text{-}C6}$ = 2.3 Hz, $J_{C6\text{-}N1}$ = 15.5 Hz, $J_{N1\text{-}C2}$ = 20.3/17.7 Hz, $J_{C2\text{-}N3}$ = 17.7/20.3 Hz, $J_{N3\text{-}C4}$ = 0 Hz. The ^{13}C NMR of the product, showed no coupling between N3 and C4, thereby

proving that the bond between them is broken during the enzymatic transformation.

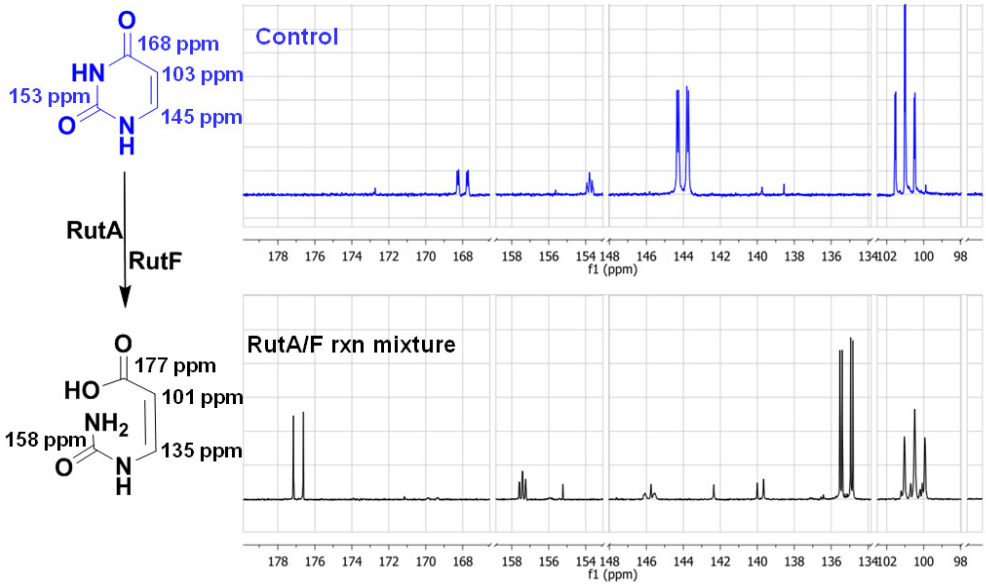


Figure 8.11: NMR (125 MHz, D_2O, C^{13}) of the reaction mixture with and without the enzyme.

HMBC of the labeled sample of uracil (**1**) (the control in section 8.3.3.) showed that the N1 has a chemical shift of about 140 ppm and the N3 has a chemical shift of about 165 ppm *figure 8.12(a)*. The literature values of chemical shifts for N1 and N3 for uracil (**1**) are 133 ppm and 159 ppm in DMSO (*11*). The experiment was so set up that it was not be possible to see the coupling of a ^{13}C which is bonded to two ^{15}Ns. Thus we do not see the coupling between C2 (153 ppm) and either N1 (140 ppm) or N3 (165 ppm). We only see the N1 (140 ppm) and C6 (145 ppm) one bond coupling, N3 (165 ppm) and C4 (168 ppm) one bond coupling and N3 (165 ppm) and C5 (103 ppm) two bond coupling. HMBC of the product only shows one bond coupling between N1 (116 ppm) and C6 (135 ppm) *figure 8.12(b)*. There is no coupling observed between N3 and C4 (177 ppm) and C5 (101 ppm). The way the experiment was set up we would not see any

coupling between the C2 (158 ppm) and with either N1 or N3. So the HMBC tells us the chemical shift of only N1.

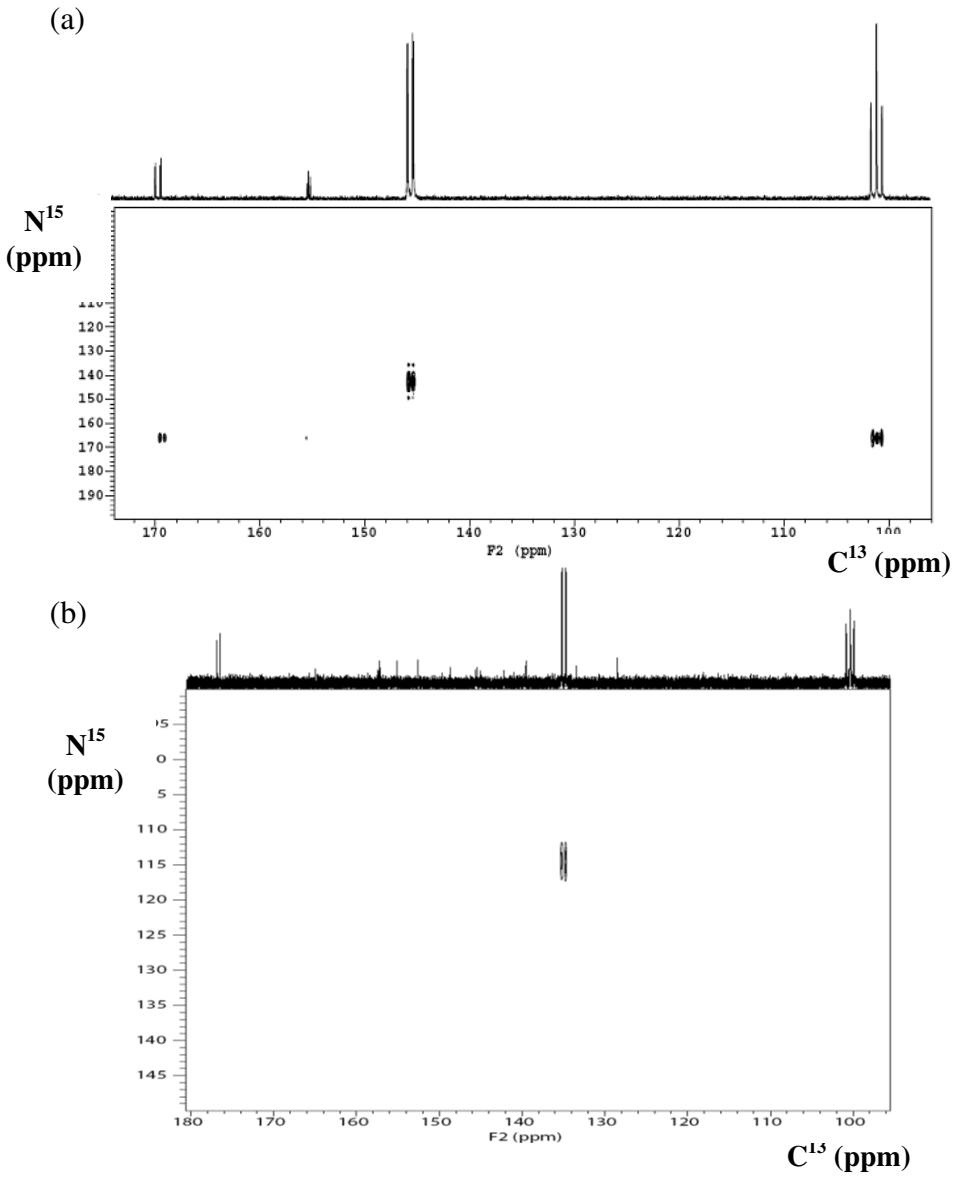


Figure 8.12: 150 MHz. (a) HMBC of the labeled sample of uracil (**1**) (the control sample) shows that the N1 has a chemical shift of about 140 ppm and the N3 has a chemical shift of about 165 ppm. (b) HMBC of the product only shows one bond coupling between N1 (113 ppm) and C6 (135 ppm). All the peaks are referenced to deuterium signal as the external reference. The accuracy is about ±1 ppm.

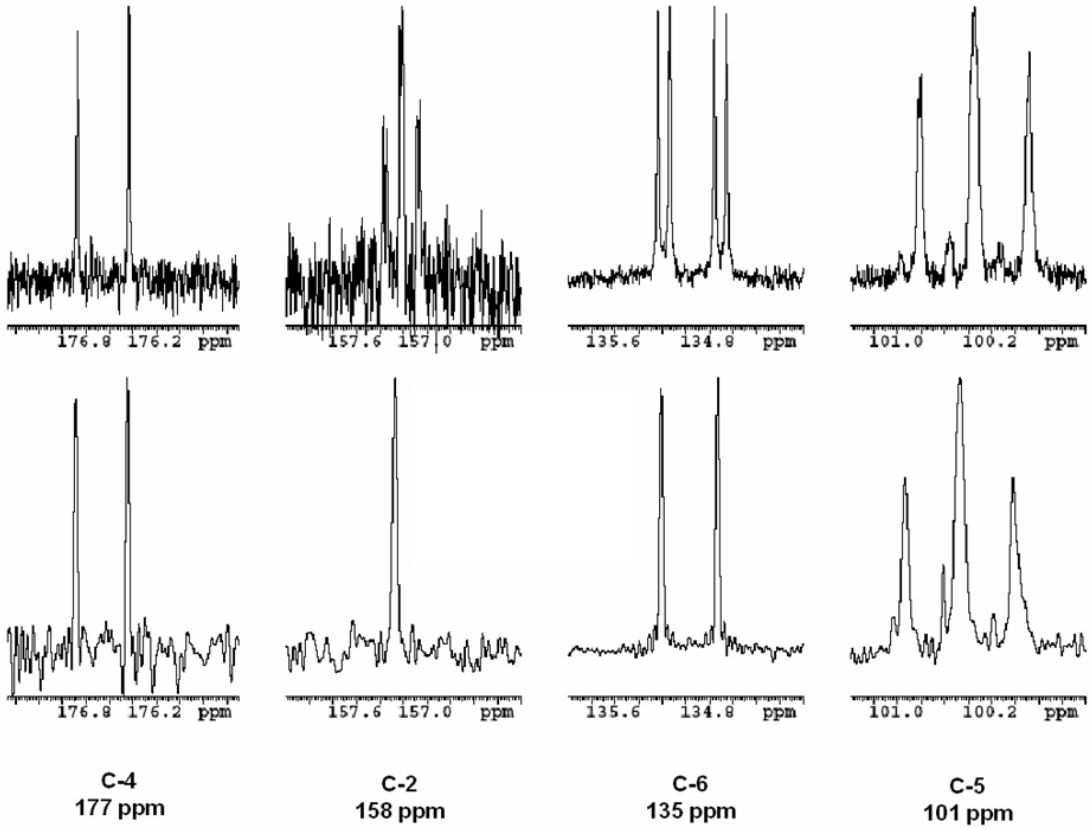


Figure 8.13: 150 MHz**.** Top row shows the ^{13}C NMR of the product of Rut A/F catalyzed reaction of uracil (**1**) (same sample as described in 8.3.3) with ^{15}N coupling, while in the bottom row the ^{13}C NMR is decoupled from the ^{15}N by broad band decoupling of the ^{15}N.

A broad band ^{15}N decoupling was done and a ^{13}C NMR of the product was acquired *figure 8.13*. The C5 (101 ppm) and the C4 (177 ppm) carbon peaks remained unaffected when we use broad band nitrogen decoupling. C5 (101 ppm) is bonded to C4 (177 ppm) and C6 (135 ppm) and so it will not have any significant coupling with any of the two nitrogens (there are however 2 bond coupling – with very low J value, which is not present in the decoupled spectra) while the C4 (177 ppm) is disconnected from the N3 as shown previously. This experiment serves as another proof of the disconnection

between the N3 and C4 (177 ppm). The C6 (135 ppm) which is a doublet of doublet (because it is coupled to C5 (101 ppm) and one nitrogen, N1) gets simplified to a doublet on introducing the broad band coupling. The C2 (158 ppm) which is not bonded to any other carbons becomes a singlet on removing the coupling from two adjacent nitrogens (N1 and N3). This experiment also confirms the connectivity of C4-C5-C6-N1-C2-N3.

To get the chemical shift of N3, an arrayed narrow band decoupling of the ^{15}N on the ^{13}C NMR of the product was performed. We can electronically decouple a nucleus by applying radio frequency at the resonance frequency of the nucleus such that a rapid transition between the two spin states of the nucleus is induced which would make it transparent to other nucleus coupling with it. The frequency at which N3 resonates in not known, hence the ^{13}C NMR was taken at various potential decoupling frequencies from 138.3 ppm to 56.3 ppm, every 4.1 ppm interval. A 90° decoupler pulse was provided for 4.8 msec with the decoupling band width of 15 ppm. The result of this experiment is shown in *figure 8.14*. Each of the ^{13}C nuceli spectra is shown and its fate at various potential decoupling frequencies of the ^{15}N. The C5 (101 ppm) remains unaffected by this experiment as it does not couple with any nitrogen. C6 (135 ppm) is coupled with the N1 (113 ppm). We see the decoupling for this nucleus when a decoupler pulse of 113 ppm was applied. From doublet of doublets it decoupled into a doublet (because it still coupled with C5 (101 ppm)). C2 (158 ppm) is coupled to N1 (113 ppm) and hence the triplet gets simplified to a doublet when a decoupler pulse of 113 ppm was applied. It is however also bonded with the N3. On continuing application of potential decoupler pulse at other frequencies we again see C5 (158 ppm) getting coupled and then decoupling at 80.8 ppm and then again coupling. It becomes a doublet (decoupled) at 80.8 ppm and thus this must be the chemical shift for N3. It will decouple only when the decoupling frequency matches with the resonance frequency of the nucleus. Thus, N3 must resonate at 80.8 ppm. This is a significant change in the chemical environment of

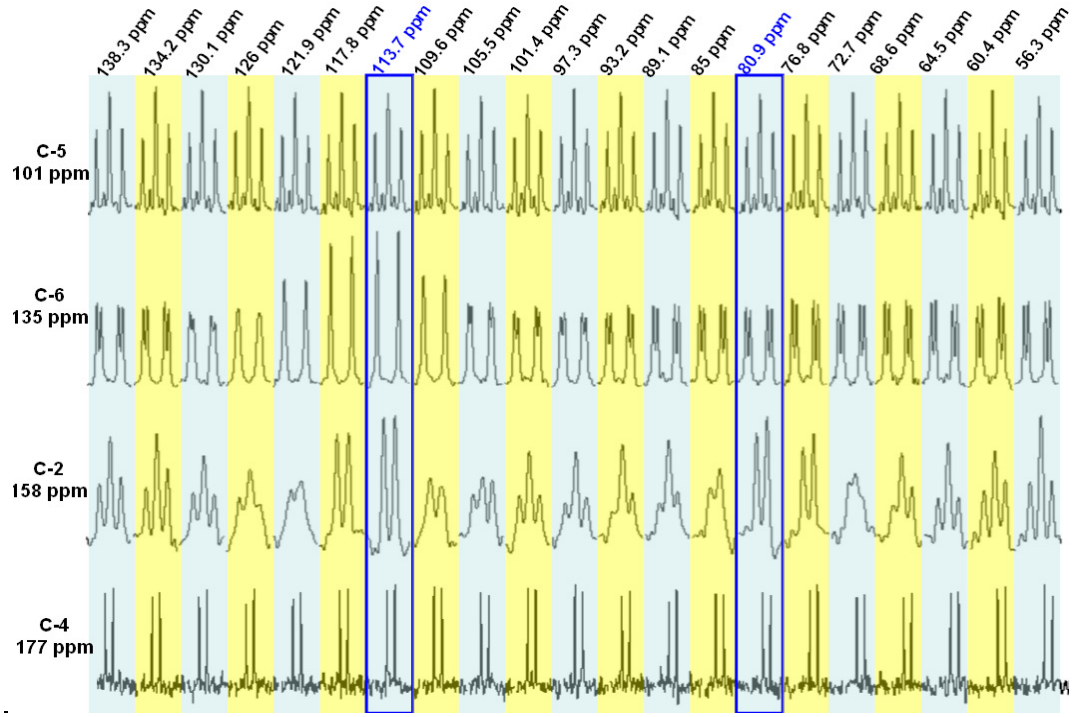


Figure 8.14: 150 MHz. An arrayed narrow band decoupling of ^{15}N on the C^{13} NMR of the product of RutA/F catalyzed reaction with uracil (**1**) (same sample as described in 8.3.3). The potential decoupling frequency is labeled on top of each column.

N3. In uracil (**1**) N3 had a chemical shift at 165 ppm and here it is nearly half of that. Moreover, N3 does not couple with C4. This is where the reaction has happened.

There are four possible structures for the product of RutA/F catalyzed reaction of uracil (**1**), as shown in *figure 8.15(a)*. Once, the reduced form of flavin reacts with molecular oxygen to form the flavin hydroperoxide (*12*), it could act as an electrophile, as in the case of p-hydroxybenzoate hydroxylase (*13*), to form 3-hydroxyuracil (**23**) by the mechanism shown in *figure 8.15 (c)*. This does not agree with J=0 for N3 and C4. Compound (**24**) can be produced by nucleophilic attack of the flavin hydroperoxide on uracil (**1**), as in the case of cyclohexanone monooxygenase (*14*), by the mechanism shown in *figure 8.15(b)*, which could then undergo hydrolysis at C4 to form 3-(3-hydroxyureido

acrylate) (**25**). Both compound (**24**) and 3-(3-hydroxyureido acrylate) (**25**) do not have an N3-C4 bond and hence no coupling (J=0). This agrees with the NMR results. The only other structure that agrees with the NMR result is 3-ureidoacrylate (**16**), but its formation by flavin dependent chemistry is unprecedented.

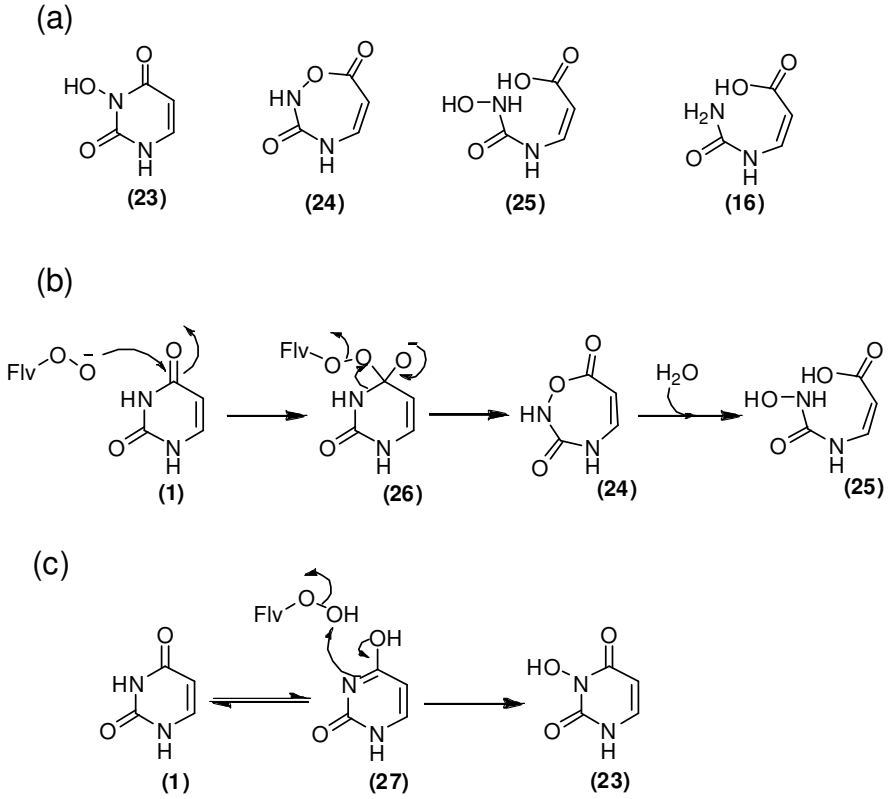


Figure 8.15: (a) Possible structures of the product of RutA/F catalyzed reaction of uracil (**1**). (b) The reaction mechanism in which flavin hydroperoxide acts as a nucleophile. (c) The reaction mechanism in which flavin hydroperoxide acts as an electrophile.

***8.3.4. Reactions with 3-hydroxyuracil* (23).** 3-hydroxyuracil (**23**) was analyzed by HPLC (method A) to see if it is the product of the RutA/F catalyzed reaction of uracil(**1**),

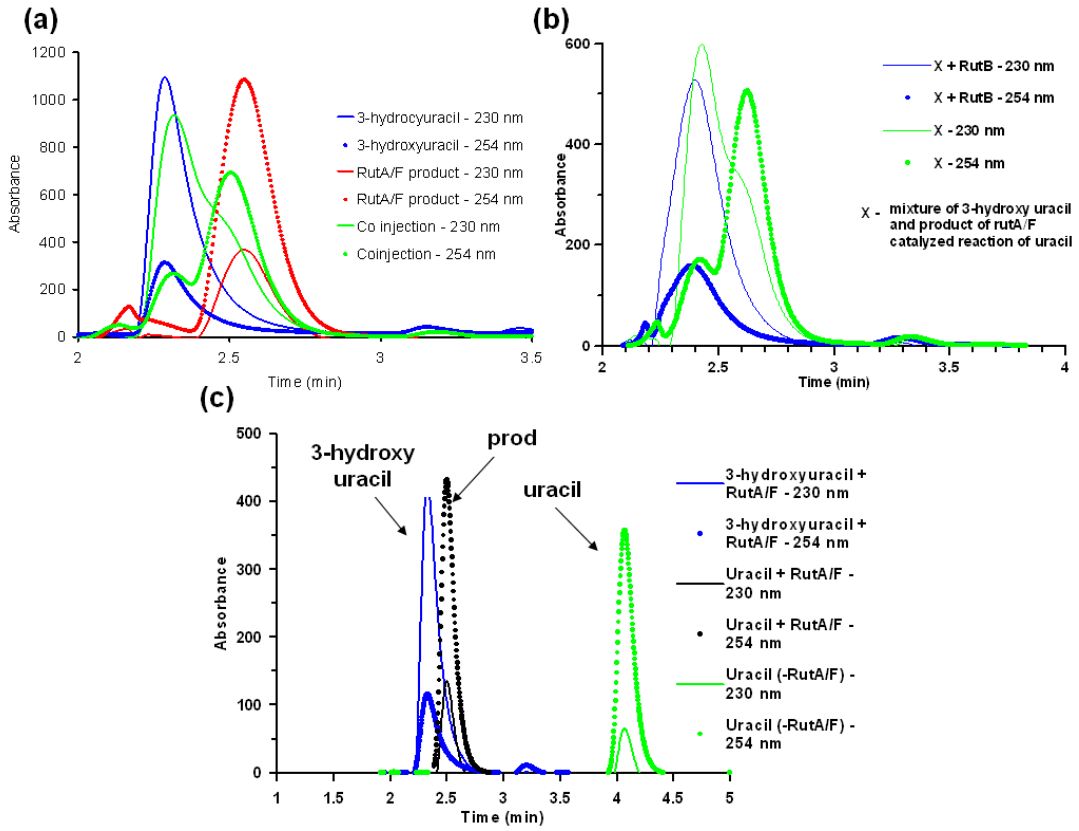


Figure 8.16: (a) HPLC chromatogram of 3-hydroxyuracil (**23**) and the product of RutA/F catalyzed reaction of uracil (**1**). (b) HPLC chromatogram of the above samples after incubation with RutB. (c) HPLC chromatogram after 2 hours of incubation of 3-hydroxyuracil (**23**) with RutA/F. The controls are also shown. The absorbance was measured at 230 nm as well as 254 nm with 5 nm as the bandwidth.

figure 8.16(a). Though it eluted with a retention time very close to that of the product, it had a totally different absorbance profile. 3-hydroxyuracil (**23**) had an absorbance maximum at 230 nm while the product of RutA/F catalyzed reaction of uracil (**1**) had an absorbance maximum of 254 nm. The product of RutA/F catalyzed reaction of uracil (**1**) was prepared by incubation of 200 μM uracil (**1**) with 10 μM of enzyme in presence of 2 mM NADH and 20 μM FMN in 100 mM potassium phosphate buffer at pH 8.0, at 25°C

for 1 hour. 3-hydroxyuracil (**23**) was untouched on treatment with RutB, while the product of RutA/F catalyzed reaction of uracil (**1**) was converted to a non-chromophoric species, *figure 8.16(b)*. The samples were prepared by incubation of RutB (10 μM) with 200 μM of 3-hydroxyuracil (**23**) and the product of RutA/F catalyzed reaction of uracil (**1**) respectively for 30 min at 25°C. 3-hydroxyuracil (**23**) was then tested to see if it was an intermediate for the reaction catalyzed by RutA/F. Incubation of 200 μM 3-hydroxyuracil (**23**) with 10 μM of RutA/F for 2 hours at 25°C, did not yield the product of RutA/F catalyzed reaction of uracil (**1**). However under identical conditions, RutA/F catalyzed the transformation of uracil (**1**) to form the enzymatic product, *figure 8.16 (c)*. Thus, 3-hydroxyuracil (**23**) is not an intermediate in the RutA/F catalyzed reaction of uracil (**1**).

8.3.5. RutA reaction in the presence of H_2O^{18} and $^{18}O_2$. Two identical solutions were made, each (300 μL) containing 180 μL of D_2O, 10 μL of 1M K_2HPO_4 pH 8.0, 10 μL of 20 mM FMN and 100 μL of 1 mM RutA/F. In one of them, 300 μL of 9.4 mM $C^{13}N^{15}$ labeled uracil (**1**) solution made in H_2O^{18} was added. This was the 'reaction mixture' which had a final concentration of 50% H_2O^{18}. The other sample was used as a 'control' equal volume of 9.4 mM $C^{13}N^{15}$ labeled uracil (**1**) solution made in H_2O^{16} was added. The cap of the eppendorf tubes were left open for atmospheric oxygen to dissolve. The reaction mixture was allowed to incubate at room temperature for 3 hours after which both the samples were analyzed by C^{13} NMR. Reaction carried out with $^{13}C^{15}N$ labeled uracil (**1**) in presence of 50% H_2O^{18} showed no isotopic shift at C4 of the product, *figure 8.17(a)*, which would have been observed, had 3-(3-hydroxyureido)acrylic acid (**25**) was the product of RutA and RutF catalyzed reaction of uracil (**1**). Thus one can rule out 3-(3-hydroxyureido acrylate) (**25**) to be the product of the enzymatic reaction.

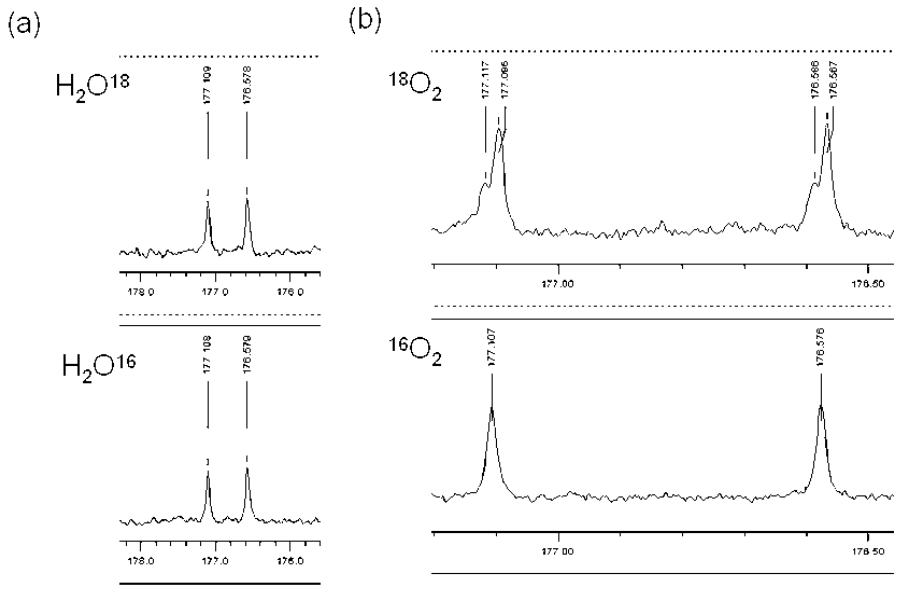


Figure 8.17: ^{13}C NMR (125 MHz) of the product of RutA/F catalyzed reaction of uracil (**1**). (a) when the reaction is done in H_2O^{18} (control reaction done in H_2O^{16}) and (b) when the reaction is done in $^{18}O_2$ (control reaction done in $^{16}O_2$).

Two 15 mL Nalgene tube containing 500 μL of $C^{13}N^{15}$ labeled uracil (**1**) dissolved in 25 mM phosphate buffer at pH 8.0 in D_2O was set up and $^{18}O_2$ was passed through one (reaction mixture) while $^{16}O_2$ was passed through the other (control) with a long needle attached to the respective cylinders. The rate of the gas flow was about 1 small bubble per second. The solutions were purged with oxygen for half hour, after which 100 μL of 1 mM RutA/F, 50 μL of 20 mM FMN and 75 mg of NADH was added in quick succession in the order mentioned. The tubes were capped tightly and vortexed briefly to dissolve the NADH. Both the reaction mixture as well as the control solution was incubated at room temperature for 2 hours after which they were analyzed by C^{13} NMR. Enzymatic reaction carried out in the presence of $^{18}O_2$ shows a clear isotopic shift at C4 for the product *figure 8.17(b)*.

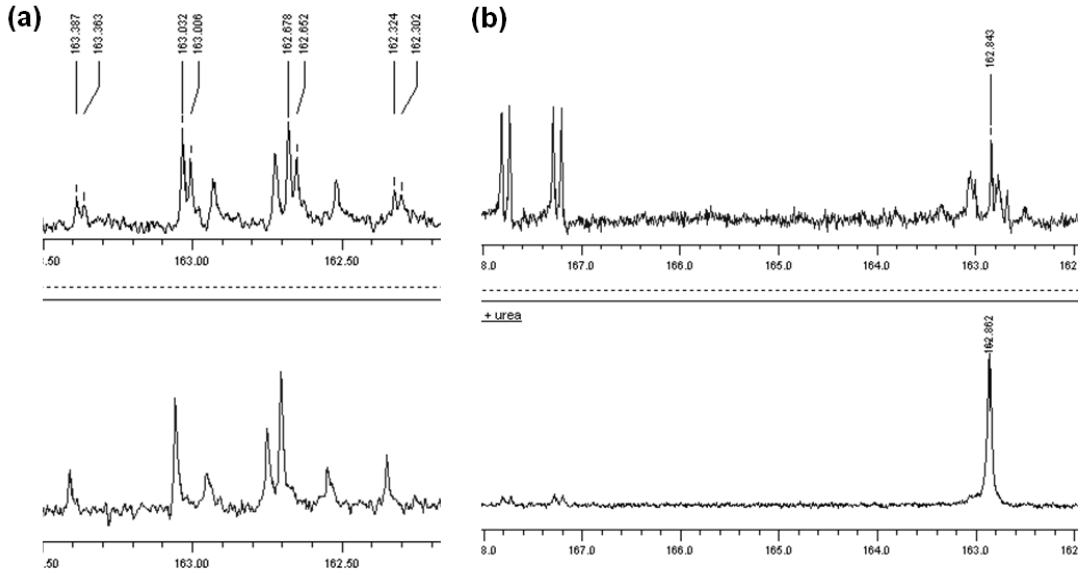


Figure 8.18 (a) NMR 125 MHz, D_2O. Top panel shows the ^{13}C NMR of treatment of the reaction mixture in 50% H_2O^{18} treated with 10% TFA. The carbonyl carbon of trifluoroacetic acid is shown here. The bottom panel shows an identical sample prepared in 100 % H_2O^{16}. Because of isotopic chemical shifts due to H_2O^{18} each of the peaks of the quartet is seen as a doublet. 22 ppb isotopic chemical shift was seen for the sample which had 50% H_2O^{18}. No such isotopic shift is seen for the sample prepared in 100% H_2O^{16}. Amongst the carbonyl peak a triplet can be seen in both the samples. (b) NMR, 500 MHz, D_2O. ^{13}C NMR of the sample shown in (a) after spiking it with urea (**7**).

This observation, however, did not rule out between compound (**24**) and 3-ureidoacrylate (**16**). The enzymatic reaction product was then treated with 10% trifluoroacetic acid, with the hope that the structure of acid degraded product could lead to the structure of the parent compound. A triplet was observed at 162.7 ppm which corresponded to the chemical shift of urea (**7**), *figure 8.18*. The triplet clearly originated from C2 of uracil (**1**) (flanked by 2 ^{15}N nitrogens) and this observation suggested that acid hydrolysis of the product of RutA and RutF catalyzed reaction of uracil (**1**) produces

urea (**7**). The product is likely to be 3-ureidoacrylate (**16**) as it is susceptible to addition elimination reaction under acidic condition to give urea (**7**). Such a reaction can also occur for compound (**24**) but in that case it would have produced hydroxyl urea, which has a different chemical shift.

8.3.6. LCMS analysis of RutA reaction product. 200 μM of uracil (**1**) was incubated with 10 μM of RutA/F in 25 mM ammonium actetate buffer at pH 7.0 in the presence of 1 mM NADH and 40 μM FMN at 25˚C for 3 hours. All solutions were made in 25 mM ammonium acetate at pH 7.0. The protein was buffer exchanged to 50 mM ammonium acetate buffer at pH 7.0. The product of the reaction was purified out by HPLC using method A in which the 100 mM potassium phosphate buffer at pH 6.6 was substituted with 25 mM ammonium acetate at pH 6.6. A preparative column was used and the product peak was collected in small aliquots by HPLC (method A, where the solution B was 25 mM ammonium acetate buffer at pH 6.6) and frozen immediately in liquid nitrogen and stored at -20˚C. A control reaction was set up in which no enzyme was added and unreacted uracil (**1**) was isolated from the reaction mixture in a similar manner. Both the samples were subjected to LC-EMS analysis. Another set of reaction mixture as well as the control was made in an identical fashion and they were directly subjected to LCI-MS analysis, without any HPLC purification. *Figure 8.19* shows the result of all the analysis. For the control reactions *figure 8.19(a, c)* the peak with m/z = 113 was isolated, corresponding to uracil (**1**), while for the reaction samples in which the product was expected *figure 8.19(b, d)* the peak with m/z =131 was isolated that corresponds to 3-ureidoacrylate (**16**). No peaks with m/z = 129 corresponding to compound (**24**) and m/z = 147 corresponding to 3-(3-hydroxyureido acrylate) (**25**) were isolable.

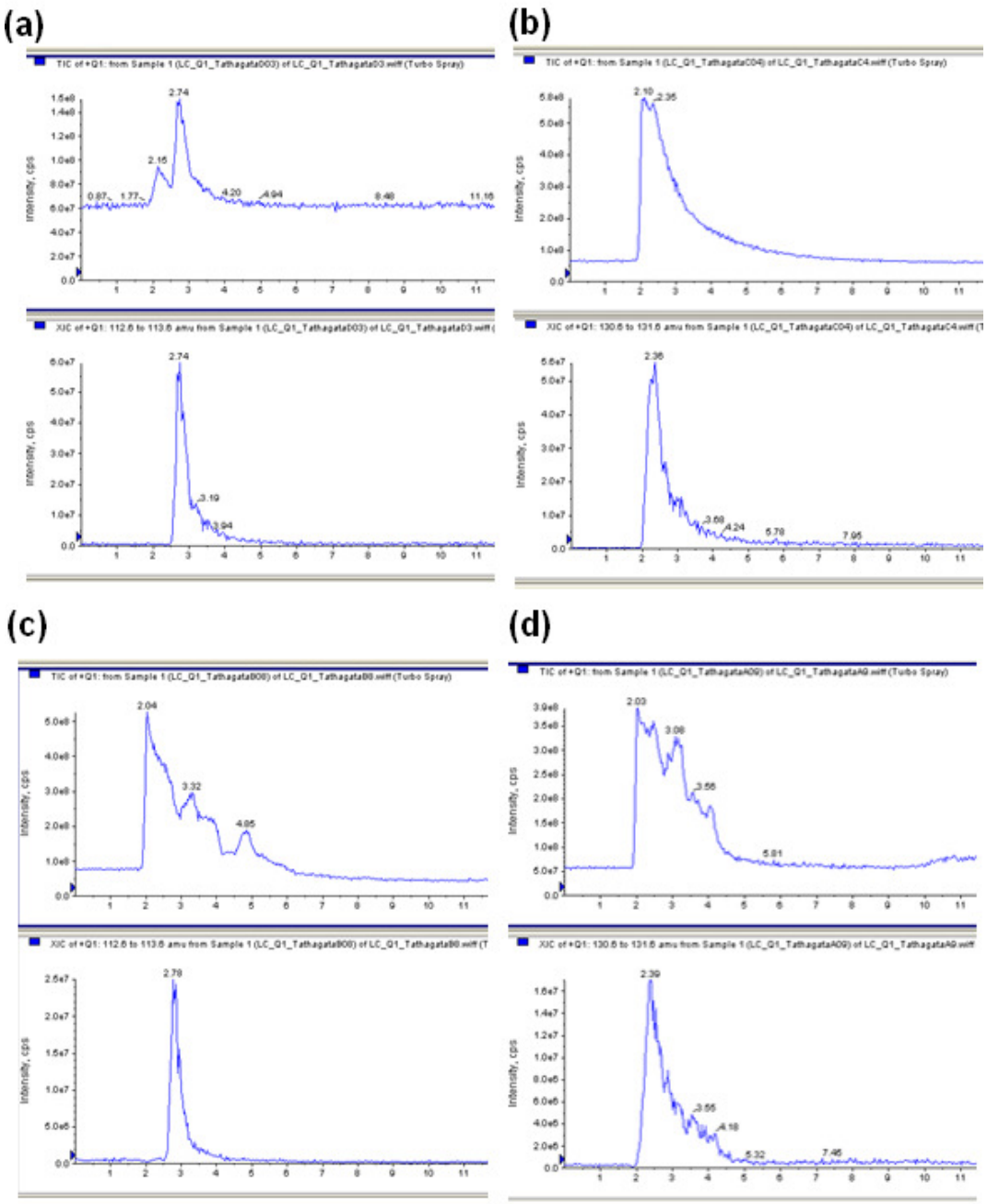


Figure 8.19: LC-EMS of (a) HPLC purified uracil (**1**), (b) HPLC purified product of RutA/F catalyzed reaction of uracil (**1**), (c) a control reaction sample with uracil (**1**), NADH and FMN, incubated at 25°C for 3 hours and (d) a reaction mixture containing uracil (**1**), NADH, FMN and RutA/F, incubated at 25°C for 3 hours. The top panel shows the chromatogram and the bottom panel shows the isolated peak.

8.3.7. Co-injection with 3-ureidoacrylate (16). The RutA/F product was prepared by incubation of 200 μM of uracil (**1**), 1 mM NADH, 20 μM of FMN and 10 μM of RutA/F at 25°C for 2 hours. All solutions were prepared in 100 mM potassium phosphate buffer at pH 8.0. The synthesized 3-ureidoacrylate (**16**) was dissolved in 100 mM potassium phosphate buffer at pH 8.0. Co- injection of synthesized sample of 3-ureidoacrylate (**16**) with the product of RutA/F catalyzed reaction of uracil (**1**) showed an increase in intensity of the product peak, *figure 8.20,* thereby ascertaining that 3-ureidoacrylate (**16**) was indeed the product of the RutA/F catalyzed reaction of uracil (**1**).

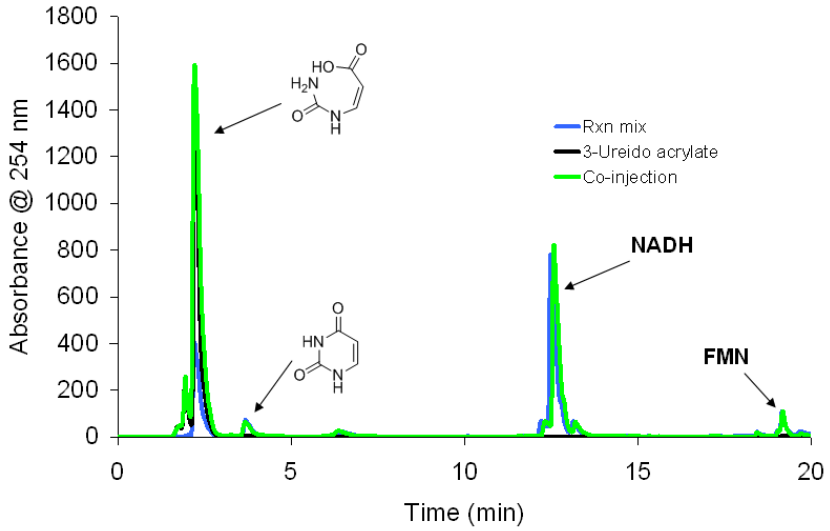


Figure 8.20: HPLC trace showing the co-migration of 3-ureidoacrylate (**16**) with the product of RutA/F catalyzed reaction of uracil (**1**).

8.3.8. Reaction of RutA/F product with RutB. The 3-ureidoacrylate (**16**) was prepared in the same manner as described in section 8.3.3. 50 μL of 0.6 mM RutB was added to it and the reaction was monitored by ^{13}C NMR (arrayed experiment) over time in 500 MHz instrument. The 3-ureidoacrylate (**16**), produced enzymatically from $^{13}C^{15}N$ labeled uracil (**1**), when treated with RutB showed the growth of a singlet at 161 ppm, over time, along with the disappearance of the peaks of 3-ureidoacrylate (**16**) *figure 8.21(a).*

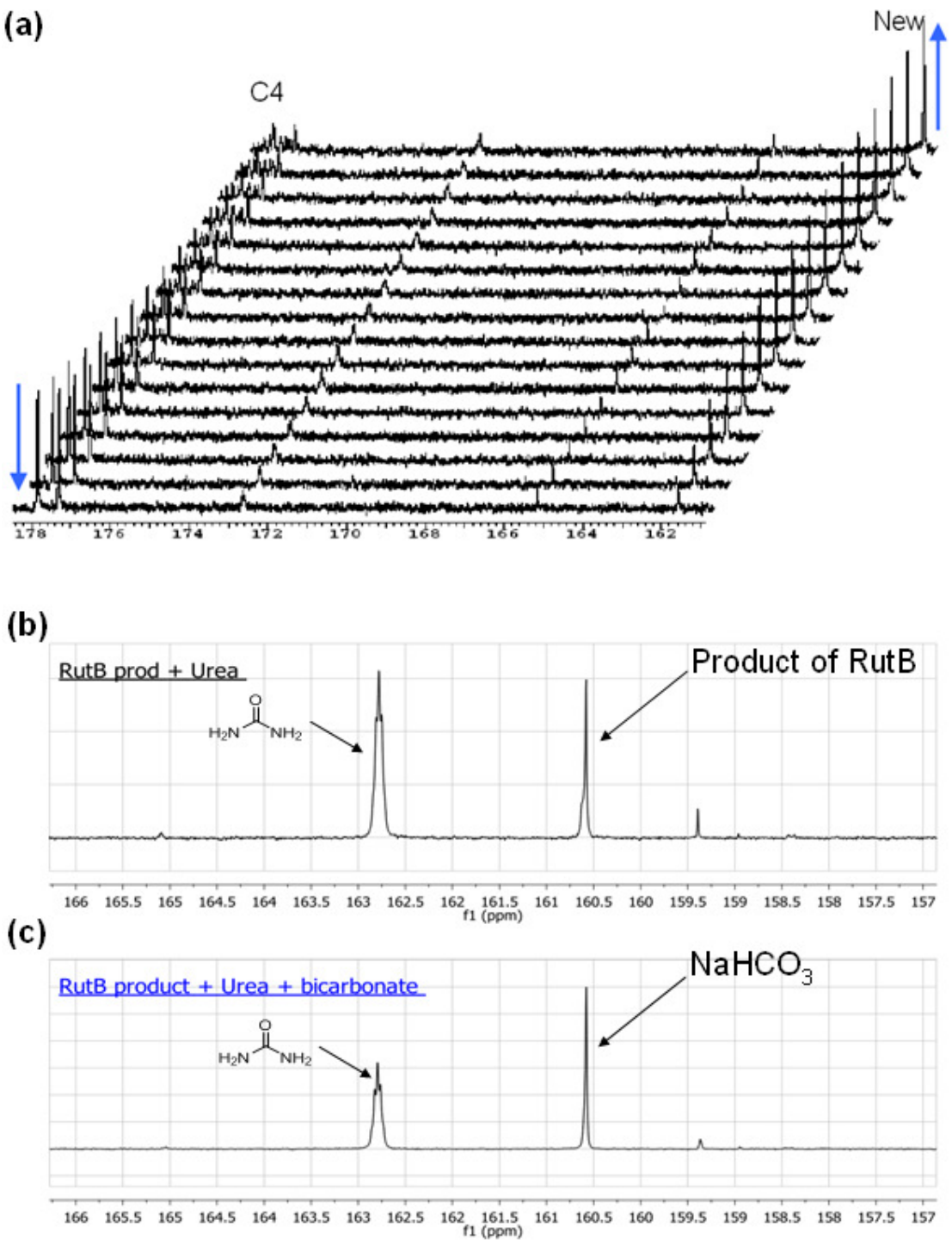


Figure 8.21: (a) Time course of the reaction mixture containing RutB and 3-ureidoacrylate (**16**) as monitored by ^{13}C NMR. 250 scans were taken at each time point. (b) On addition of urea (**7**) to the above sample (c) On addition of bicarbonate to the above sample.

Upon addition of urea (**7**) (unknown concentration) to the sample a peak at 162.5 ppm came up showing that the new peak is not urea (**7**), *figure 8.21(b)*. On addition of sodium bicarbonate to the sample the new peak spiked up, thereby showing that it is indeed due to bicarbonate formation *figure 8.21(c)*. Since, it was a singlet, which means that the ^{13}C was not bonded to either a ^{13}C or a ^{15}N, it is likely to be due to dissolved CO_2 which originates from C2 of labeled uracil (**1**). The evolution of CO_2 as a result of enzymatic reaction of RutB, suggests that RutB which has similarities with amidases (*9*), hydrolyses the amide bond between N1 and C2. The carbamate (**29**) formed would readily decarboxylate to give NH_3 and CO_2 *figure 8.25*.

8.3.9. Ammonia production by RutB. A coupled assay for ammonia with glutamate dehydrogenase in presence of α-ketogluterate (**21**) and NADPH was carried out to test the presence of ammonia in the RutB reaction sample. Ammonia liberated during the course of the reaction would react with α-ketogluterate (**21**) to form α-imminogluterate which would in turn be reduced by the enzyme glutamate dehydrogenase to glutamate (**22**). This enzyme requires NADPH as a cofactor which will get oxidized to NADP during the course of the reaction thereby showing a change in the absorbance at 340 nm, indicating the presence of ammonia in the solution. When all the reaction components (as described in 8.2.4.2.3) were present a huge decrease in the absorbance at 340 nm was seen due to production of ammonia. Such large decrease is not seen in any of the controls, *figure 8.8*.

The amount of ammonia produced per unit of uracil (**1**) was determined by two simultaneous experiments. A 500 μL reaction mixture containing 200 μM of uracil (**1**), 400 μM of NADH, 10 μM of RutA/F, 20 μM of FMN and 5 μM of RutB was incubated at 25°C for 2 hours. It was then filtered using Microcon YM10 centrifugal filters to remove the protein. 50 μL of the filtrate was assayed for ammonia by its addition to a

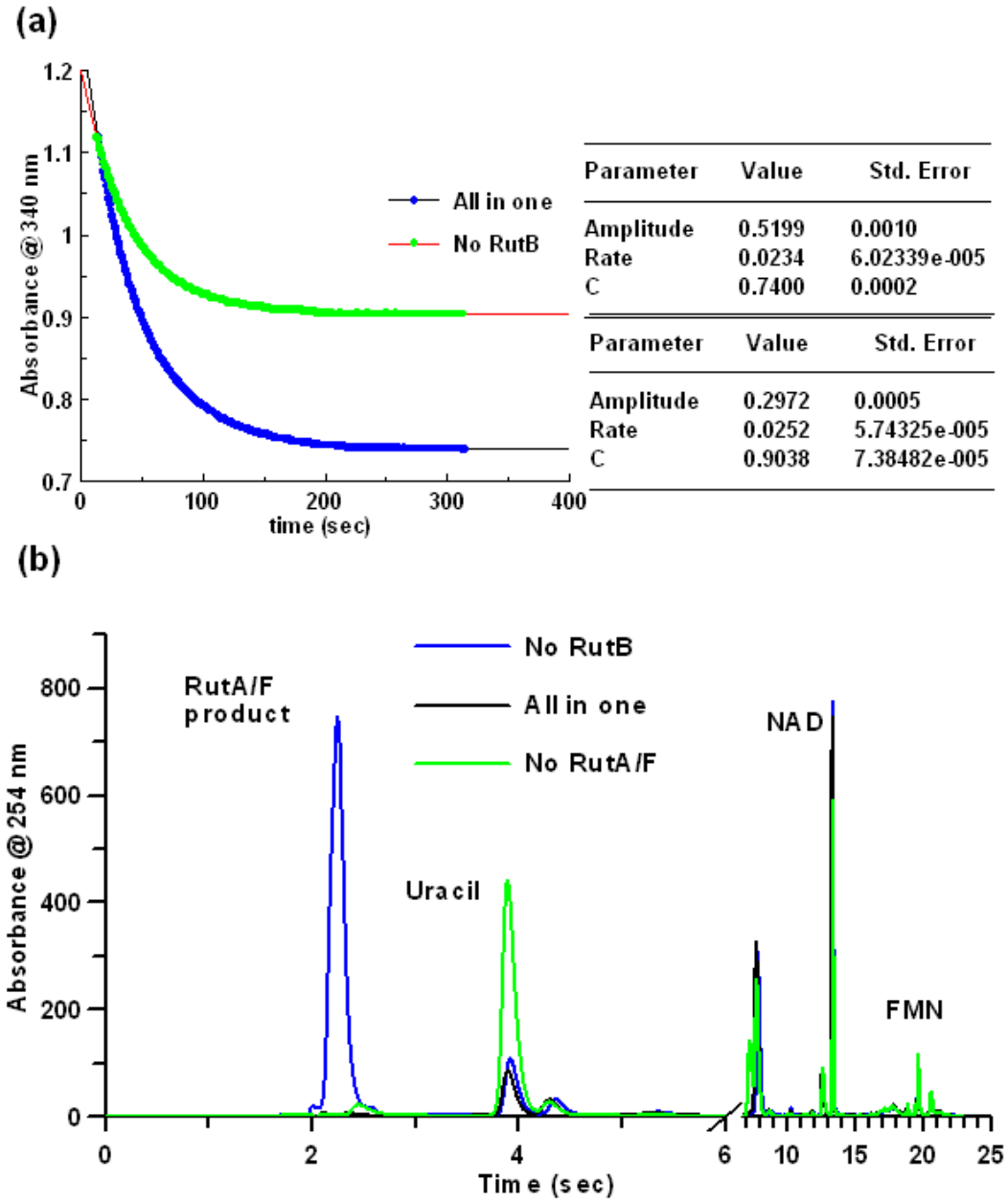


Figure 8.22: (a) UV-visible spectra showing the decrease in absorbance at 340 nm quantifying the total amount of ammonia production in the reaction of RutA/F/B and uracil (**1**) in presence of NADH and FMN. The reaction was assayed by the coupled assay mentioned in the text. (b) HPLC chromatogram of the reaction mixture and the controls of the above samples.

450 μL solution of 1 mM α-ketogluterate (**21**), 200 μM of NADPH and 2 U of glutamate dehydrogenase and monitoring the decrease in the absorbance at 340 nm over 5 min. All the solutions were made in 100 mM potassium phosphate buffer at pH 8.0. 50 μL of filtrate was analyzed by HPLC following method A. Two identical control reactions were run in parallel and treated in identical fashions. One control sample did not have RutB (this sample would give us the estimate of background ammonia in the reaction mixture by UV-vis assay) and the other did not have RutA/F (this would give us an estimate of the amount of uracil (**1**) consumed during the reaction by HPLC). The results of the ammonia assay are shown in *figure 8.22(a)*. From the difference in the amplitude of the reaction mixture and the control the amount of ammonia present in the reaction mixture was estimated to be 370 μM. The HPLC chromatogram shown in *figure 8.22(b)* showed about 80% conversion of uracil (**1**). Thus 160 μM of uracil (**1**) in presence of RutA/F/B produces 370 μM of ammonia thereby proving that RutB reaction produces 2 molecule of ammonia per molecule of 3-ureidoacrylate (**16**) produced. The second molecule of ammonia could come from the addition–elimination reaction of 3-aminoacrylate (**31**) with water, forming 3-oxoproionate (**33**) *figure 8.25*.

8.3.10. 3-oxopropionate (33) formation over time. The formation of 3-oxopropionate (**33**) was observed by derivatization of the RutB reaction mixture with 2,4- dinitrophenyl hydrazine (**18**) and analyzing by HPLC (Method B). The elution time of the hydrazone was identical with the elution time of the synthetically prepared 2,4- dinitrophenyl hydrazone of 3-oxopropionate (**19**/**20**), *figure 8.9*.

8.3.11. RutA –Fre reaction with uracil (1). Since the RutF preparation was not very pure, it was substituted with a pure preparation of Fre which is a flavin reductase (*10*). Assays were set up with Fre (in place of RutF) in an identical way as described in 8.2.4.1.

The reaction mixture was allowed to incubate for 30 min at 25°C, filtered using Microcon YM10 and analyzed by HPLC (Method A). The result was identical as seen with the impure preparation of RutF, *figure 8.23 (a)*. The product formed had the exact retention time as the chemically synthesized 3-ureidoacrylate (**16**) *figure 8.23(b)*. On treatment with RutB the product disappears, *figure 8.23 (c)*. Thus the impure sample of RutF had a flavin reductase activity.

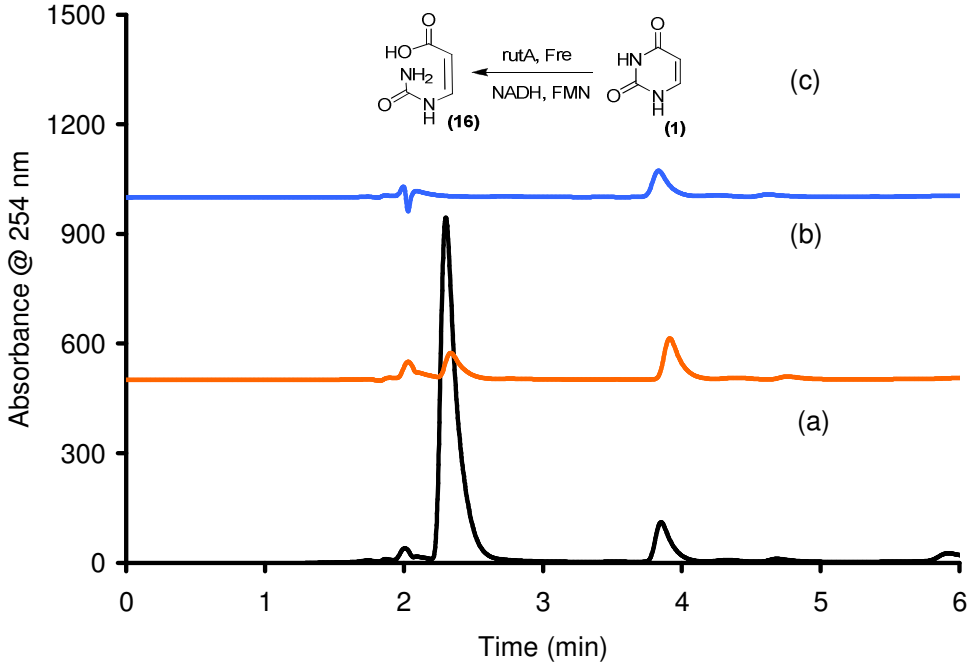


Figure 8.23: HPLC chromatogram of the assay mixture of RutA-Fre catalyzed reaction of uracil (**1**). (a) Co-injection of chemically synthesized 3-ureidoacrylate (**16**) with the reaction mixture set up in an identical manner as described in 8.2.4.1. (b) Reaction mixture showing unreacted uracil (**1**) and the 3-ureidoacrylate (**16**). (c) On treatment of the above sample with RutB.

8.3.12. Steady State Kinetic parameters of RutB. The 3-ureidoacrylate (**16**) has a λ_{max} at 254 nm and the extinction coefficient at 254 nm was determined as 12,200 M^{-1} cm^{-1} in 100 mM phosphate buffer at pH 8.0. Steady state kinetic parameters were obtained from

the concentration dependence of the rate of depletion of 3-ureidoacrylate (**16**) at constant concentration of the RutB. The enzymatic reaction exhibited Michealis Menten kinetics with K_M and k_{cat} of 28 μM and 0.4 s^{-1} respectively. The k_{cat}/K_M for RutB was determined to be 14285 $M^{-1}s^{-1}$, *figure 8.24*.

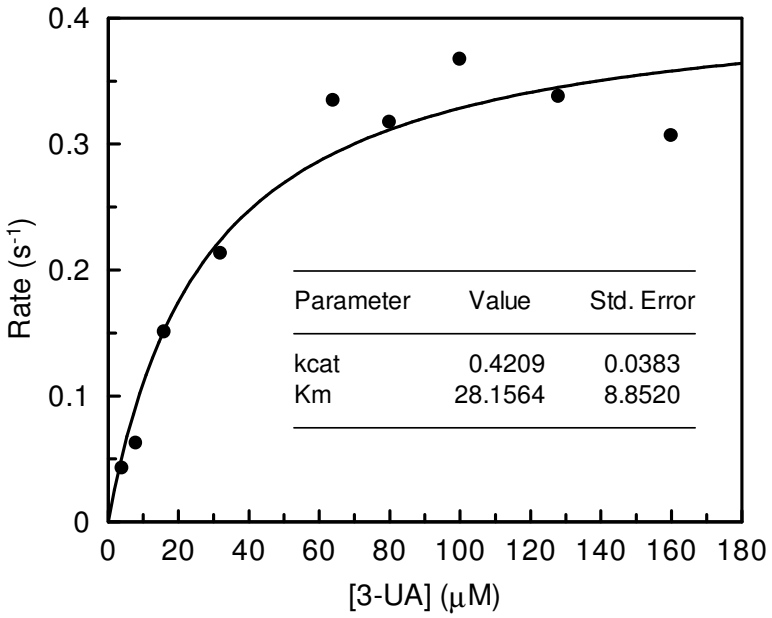


Figure 8.24: The steady state kinetic parameters for RutB with 3-ureidoacrylate (**16**) as a substrate was determined by monitoring the absorbance at 254 nm.

8.4. Conclusion:

The entire pathway is represented in the *figure 8.25* based on all the experimental evidences. In order to form 3-hydroxypropionate (**9**) an oxidoreductase was required to do the last transformation. Sequence similarities revealed that RutE, which was a putative oxioreductase could be catalyzing the last step. No 3-hydroxy propionate (**9**) was observed when RutE was added to the enzymatic reaction mixture of 3-ureidoacrylate (**16**) and RutB in the presence of both NADH as well as NADPH.

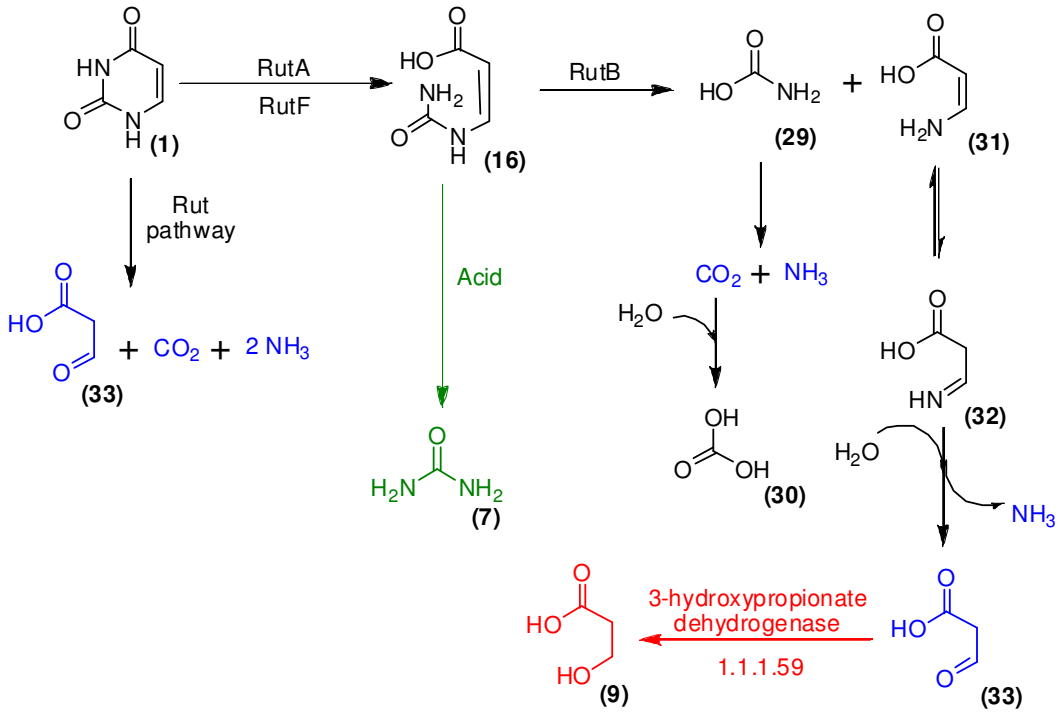


Figure 8.25: The Rut pathway of the degradation of pyrimidines.

This result makes one wonder if 3-hydroxy propionate (**9**) is indeed the last intermediate in the novel degradative pathway of uracil (**1**). *Escherichia coli* has a 3-hydroxypropionate dehydrogenase (*15*), which could very well do this transformation *in vivo*. Moreover, 3-oxopropionate (**33**) is a readily utilizable metabolite in bacteria. It can be transformed to acetyl CoA, by malonate-semialdehyde dehydrogenase and to malonyl CoA by malonic semialdehyde oxidative decarboxylase and used up for energy production via the TCA cycle as well as in fatty acid biosynthesis. It can be converted to β-alanine by 3-aminopropanoate:2-oxoglutarate aminotransferase, which is a precursor to pantothenate and CoA biosynthesis. Thus, we propose that 3-oxopropionate is the end product of the Rut pathway of degradation of uracil (**1**).

8.5. Acknowledgement

This research was supported by a grant from the National Institutes of Health (GM069618). We cordially thank Prof. Sydney Kustu and Prof. David Wemmer of the University of Berkeley, California for their collaboration in this project. We thank Dr. Ivan Keresztes, Director of the NMR facility at the Department of Chemistry and Chemical Biology, Cornell University, for all the 2D NMRs, Prof, Anderi Osterman, Department of Bioinformatics & Systems Biology, the Burnham Institute, for providing us with all the clones of the Rut genes, Prof. Luying Xun, School of Molecular Biosciences, Washington State University for providing us with the Fre clone and Dr.Sheng Zhang, Director of Proteomics and Mass Spectrometry Core Facility at Cornell University for running the LC-EMS.

REFERENCES

(1) Vogels, G. D., and Van der Drift, C. (1976) Degradation of purines and pyrimidines by microorganisms. *Bacteriol Rev 40*, 403-68.

(2) Soong, C. L., Ogawa, J., Sakuradani, E., and Shimizu, S. (2002) Barbiturase, a novel zinc-containing amidohydrolase involved in oxidative pyrimidine metabolism. *J Biol Chem 277*, 7051-8.

(3) Loh, K. D., Gyaneshwar, P., Markenscoff Papadimitriou, E., Fong, R., Kim, K. S., Parales, R., Zhou, Z., Inwood, W., and Kustu, S. (2006) A previously undescribed pathway for pyrimidine catabolism. *Proc Natl Acad Sci U S A 103*, 5114-9.

(4) Osterman, A. (2006) A hidden metabolic pathway exposed. *Proc Natl Acad Sci U S A 103*, 5637-8.

(5) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem 72*, 248-254.

(6) Ohigashi, H. K., Mikio; Sakaki, Masaharu; Koshimizu, Koichi. (1989) 3-Hydroxyuridine, an allelopathic factor of an African tree, Baillonella toxisperma. *Phytochemistry 28*, 1365-1368.

(7) Rehberg, G. M. G., Brian M. (1995) Improved synthesis of 2H-1,3(3H)-oxazine-2,6-dione (oxauracil). *Organic Preparations and Procedures International 27*, 651.

(8) Farkas, J. H., Jaroslav; Jindrova, Olga; Skoda, Jan. (1982) Reaction of 2,3-dihydro-1,3-6H-oxazine-2,6-dione with aliphatic amines and amino acids. *Collection of Czechoslovak Chemical Communications 47*, 2932-2945.

(9) http://theseed.uchicago.edu/FIG/index.cgi.

(10) Xun, L., and Sandvik, E. R. (2000) Characterization of 4-hydroxyphenylacetate 3-hydroxylase (HpaB) of Escherichia coli as a reduced flavin adenine dinucleotide-utilizing monooxygenase. *Appl Environ Microbiol 66*, 481-6.

(11) M. Chahinian, H. B. S. a. B. A. (1998) Hydration strcuture of uracil as studied by 1D and 2D heteronuclear Overhauser spectroscopy: evidence for the formation of a trihydrate in the first solvation shell. *Chem. Phys. Lett. 285*, 337-345.

(12) Massey. V. (1994) Activation of molecular oxygen by flavins and flavoproteins. *Journal of Biological Chemistry 269*, 22459-22462.

(13) Entsch, B., and van Berkel, W. J. (1995) Structure and mechanism of para-hydroxybenzoate hydroxylase. *Faseb J 9*, 476-83.

(14) Sheng, D., Ballou, D. P., and Massey, V. (2001) Mechanistic studies of cyclohexanone monooxygenase: chemical properties of intermediates involved in catalysis. *Biochemistry 40*, 11156-67.

(15) Den, H., Robinson, W. G., and Coon, M. J. (1959) Enzymatic conversion of beta-hydroxypropionate to malonic semialdehyde. *J Biol Chem 234*, 1666-71.

CHAPTER 9

Mechanistic studies on RutA from *Escherichia coli* K12*

9.1 Introduction

RutA catalyzes the first step in the degradation of uracil (**1**) (or thymine) to form 3-ureidoacrylate (**2**) with the assistance of RutF, which is a flavin reductase. This new pathway of pyrimidine degradation was recently discovered in *Escherichia coli*K12 (*1, 2*). It utilizes flavin mono nucleotide (**5**) as a cofactor, *figure 9.1*. The reduced form of the cofactor is required for the catalysis. Thus RutA requires the assistance of a partner protein which would be able to supply the reducing equivalents to FMN (**5**) in order to convert it into $FMNH_2$ (**6**), the active form of the cofactor. This partner protein is RutF, which is a flavin reductase (*3*) which oxidizes NADH (**3**) to NAD (**4**) and transfers the reducing equivalent to FMN (**5**). It over expresses well, but has poor solubility. It can also be substituted by Fre which is a flavin reductase from *E.coli* (*4*).

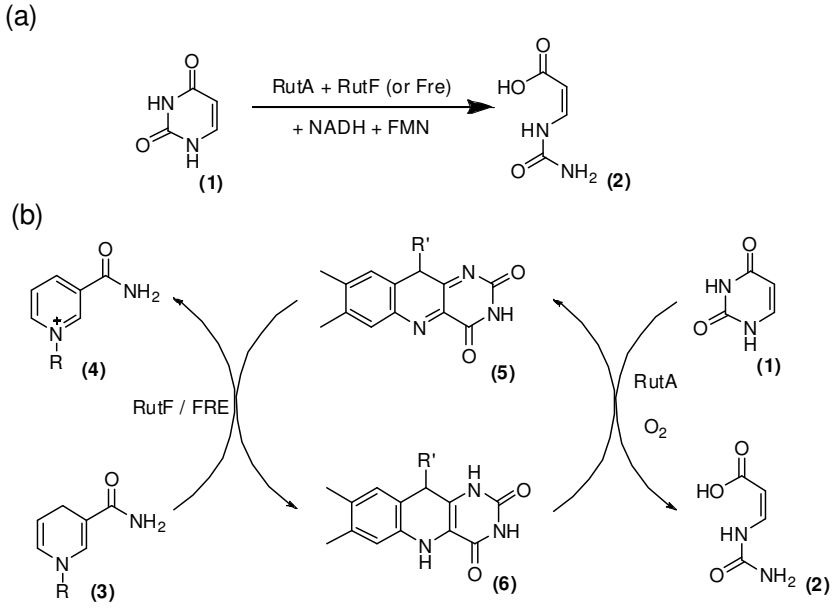


Figure 9.1: (a) Overall reaction catalyzed by RutA. (b) Two sub reactions catalyzed by RutF and RutA respectively.

* The synthesis of 3-ureidoacrylic peracid and its stability was done by Dr. Sameh Abdelwahed. The crystallographic studies and analysis were done by Dr. Yang Zhang.

The experimental details for the characterization of the product of this enzymatic reaction have been provided in the previous chapter. However, the mechanism of the formation of 3-ureidoacrylate (**2**) is not known. This chapter details the efforts towards understanding the reaction mechanism.

9.2 Experimental Section

9.2.1. Cloning over-expression and purification. The protein overexpression and purification was done following the procedure described in section 8.2.1. SDS-PAGE analysis showed over 95% pure protein for RutA and Fre. The yield of the purified RutA and Fre was 25 and 60 mg per liter of cell culture respectively. The protein concentration was measured by the Bradford assay (*5*).

9.2.2. HPLC analysis. HPLC analysis of the enzymatic reaction mixture was performed on a Hewlett-Packard 1100 instrument using a Supelcosil LC 18 T (15 cm X 4.6 mm, 3.0 μm) column. Solution A contained water, solution B contained 100 mM potassium phosphate buffer at pH 6.6 and solution C contained methanol. The following linear gradient was used. 0% to 10% solution A and 100% to 90% solution B for 0-5 min, 10% to 48% solution A, 90% to 40% of solution B and 0% to 12% of solution C from 5-12 min, 48% to 50% solution A, 40% to 30% of solution B and 12% to 20% of solution C in 12-14 min, 50% to 30% solution A, 30% to 10% of solution B and 20% to 60% of solution C in 14-18 min, 30% to 0% solution A, 10% to 100% of solution B and 60% to 0% of solution C in 18-20 min and 0% of solution A, 100% of solution B and 0% of solution C in 20-25 minutes. The flow rate was 1 ml/min and the absorbance was measured at 254 nm and 450 nm. Under these conditions the following compounds were readily separated (retention time in parenthesis): 254 nm: 3-ureidoacrylate (**16**) (2.3 min),

uracil (**1**) (3.7 min), NAD (12.9 min) and NADH (14.3 min), 450 nm: p-nitrophenol (20.7 min).

9.2.3. Synthesis[†]

9.2.3.1. (Z)-4-nitrophenyl-3-ureidoacrylate (9). The synthesis of *(Z)*-4-nitrophenyl 3-ureidoacrylate (**9**) was adapted from known literature procedures of similar compounds (*6, 7*) . To a solution of *(Z)*-3-ureidoacrylic acid (**2**) (100 mg, 0.77 mmol) and 4-dimethylaminopyridine (DMAP) (10 mg, 0.08 mmol) in anhydrous dimethylformamide (DMF) (2 mL), 4-nitrophenol (**8**) (130 mg, 0.92 mmol) was added, followed by dicylcohexylcarbodiimide (DCC) (190 mg, 0.93 mmol). The reaction mixture was stirred at room temperature overnight. N,N-dicyclohexylurea (DCU) was removed by filtration, and DMF was evaporated from the filtrate under reduced pressure. The residue, which contained some DCU that was partially soluble in DMF, was dissolved in ethyl acetate (30 mL), the insoluble DCU was removed by filtration and the solvent was removed under reduced pressure. The product was washed with water and filtered under vacuum and dried to give yellow colored 4-nitrophenyl 3-ureidoacrylate (55 mg, 30%). ^{1}H NMR (300 MHz, DMSO): δ 5.14 (d, 1H), 6.85 (br s, 1H), 7.31 (br s, 1H), 7.47 (d, 2H), 7.67 (dd, 1H), 8.32 (d, 2H), 9.80 (d, 1H). GC-MS: m/z = 251.1 ($M^{+\cdot}$) calculated for $C_{10}H_9N_3O_5$.

9.2.3.2. 3-ureidoacrylic peracid (10). The synthesis of *(Z)*-3-ureidoacrylic peracid (**10**) was adapted from known literature procedures with significant modifications. To a suspension of 4-nitrophenyl 3-ureidoacrylate (**9**) (30 mg, 0.12 mmol) in D_2O (0.5 mL) Na_2O_2 (16 mg, 0.14 mmol) was added at 4° C and the reaction mixture was stirred for 5 min. *(Z)*-3-ureidoacrylic peracid (**10**) was identified by NMR without any further

† The synthesis was done by Dr. Sameh Abdelwahed. The stability of the peracid was also determined by him.

purification which was not possible due to its instability. ^{1}H NMR (300 MHz, D_2O): σ 5.03 (d, 1H), 7.23 (d, 1H).

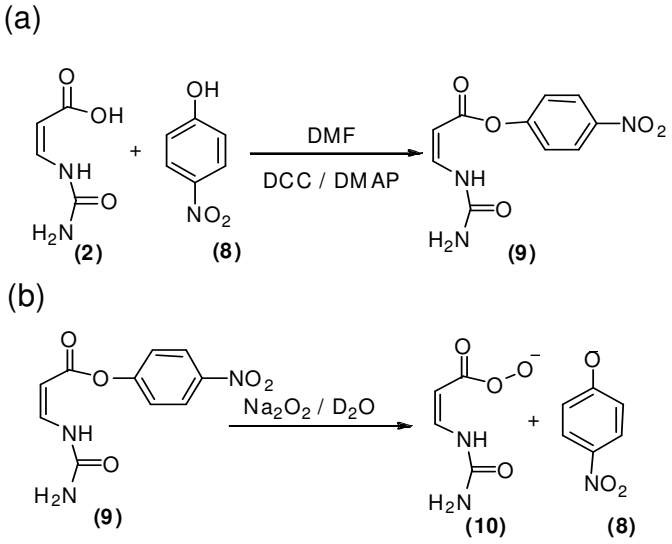


Figure 9.2: Synthetic scheme for the preparation of (a) *(Z)*-4-nitrophenyl 3-ureidoacrylate (**9**) and (b) *(Z)*-3-ureidoacrylic peracid (**10**).

9.2.4. Protein Crystallization and structure determination. † Crystals of RutA were grown from 2 mM FMN, 1.5 M ammonium sulfate, and 100 mM MES pH 6.0 using the hanging drop vapor diffusion method at 18 °C. Crystals in yellow color appeared after 1 month with a maximum size of approximately 50 μm × 50 μm × 50 μm. Prior to flash freezing in liquid nitrogen, these crystals were cryo-protected by a quick sock in a solution containing 11% glycerol, 10% ethylene glycol, 2.0 M ammonium sulfate, and 100 mM MES pH 6.0. X-ray diffraction data of RuA at 1.8 Å resolution were collected at the Advanced Photon Source NE-CAT beamline ID24-E using an ADSC Quantum 315 detector. The data were integrated and scaled using HKL2000 (*8*). Crystals of RutA belong to the space group $P3_121$ with unit cell parameters a = 87.85 Å and c = 96.37 Å.

† The crystallographic studies and analysis were done in Dr. Yang Zhang.

The unit cell contains one protomer per asymmetric unit with a solvent content of 49% and Matthews coefficient of 2.41 $Å^3$/Da. The structure of RutA was determined using automated molecular replacement program MrBUMP (*9*). The top solution was obtained using one protomer of protein Yntj (1YW1) as the search model, which has approximately 21% sequence identity to RutA. The initial model had an R factor of 50% and Rfree of 55%. The phases were gradually improved by iterative cycles of manually rebuilding in COOT (*10*) and restrained refinement in REFMAC5 (*11*) and PHENIX *(12)*. Water molecules and ligand FMN were included after the refinement of the protein molecule converged. The final structure has an R factor of 16.2% and Rfree of 19.0%.

9.3. Results

9.3.1. Structure of RutA. The structure of RutA was determined at 1.8 Å resolution using molecular replacement. The final model of the protein monomer is shown in *figure 9.3(a)* and contains residues 17 – 293 and 310 – 377 of the protein and one molecule of FMN. One protomer is found within the asymmetric unit and the biological unit dimer shown in *figure 9.3(b)* is generated using a crystallographic two-fold symmetry. RutA, which has a TIM barrel fold, shows strong structural similarity to other members of the FMN-dependent alkane sulfonate monooxygenase family. Each protomer of RutA contains a complete active site at the center of the β-barrel and the two active sites are independent of each other. FMN is tightly bound at the active site in a non-covalent fashion, *figure 9.3(c)*. The phosphate group of FMN is anchored by five hydrogen bonds from residues Arg159, Tyr160, and Ser209. The ribose moiety is additionally stabilized through a hydrogen bond to the backbone carbonyl oxygen atom of Ala206. The isoalloxazine ring stacks against the inner surface of the β-barrel and also make a hydrogen bond to the backbone amine of Ile68. The substrate binding cavity is located above the isoalloxazine ring where one side has hydrophilic residues Arg69 and Glu143

and the other side are lined with hydrophobic residues including Phe25, Trp33, Met67, Trp80, Phe224, Phe226, Leu259, and Leu347. In addition, residues 294 – 309, which lack interpretable electron density therefore are not built in the final model, presumably are flexible and can serve as a flap coving the substrate binding site as in the case of other sulfonate monooxygenase family members.

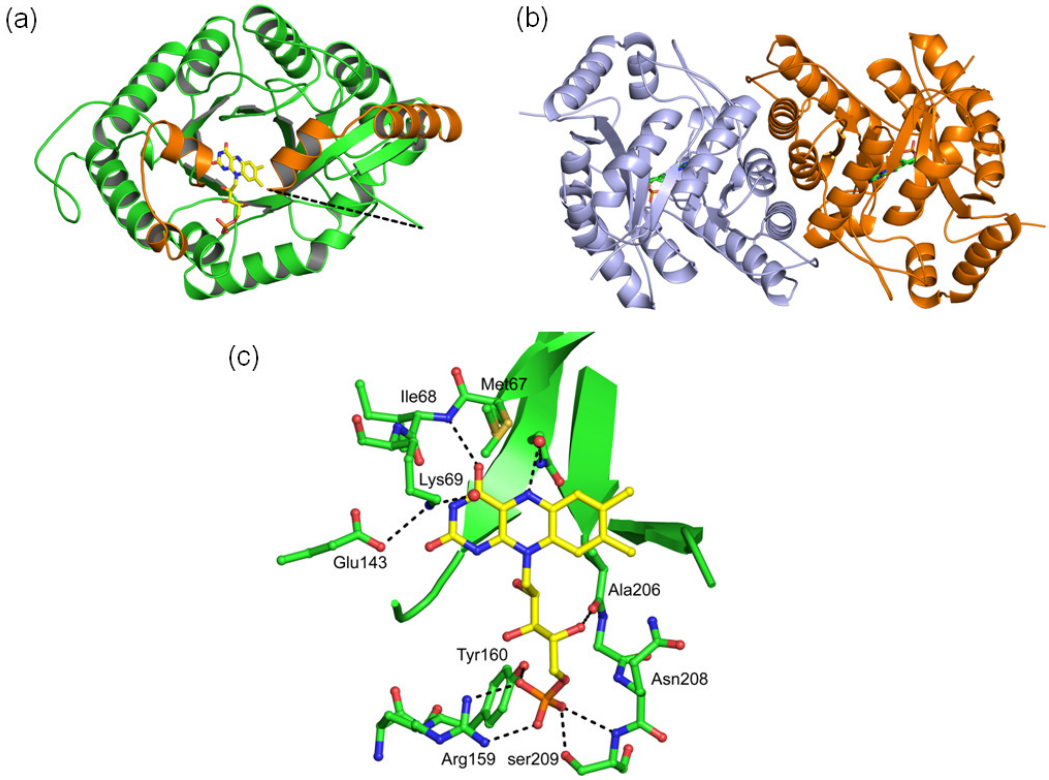


Figure 9.3 (a) Monomeric (b) dimeric and (c) active site structure of RutA

9.3.2. Stability of the 3-ureidoacrylic peracid (10). There are many procedures that describe the preparation of peracids from the corresponding acid using hydrogen peroxide in an acidic medium (*13, 14*). The *(Z)*-3-ureidoacrylic acid (**2**) was found to be susceptible to addition elimination reaction under acidic condition to produce urea (section 8.3.4). A similar reaction is likely to occur for *(Z)*-3-ureidoacrylic peracid (**10**) if it

is prepared under acidic conditions, hence basic conditions were employed for its synthesis.

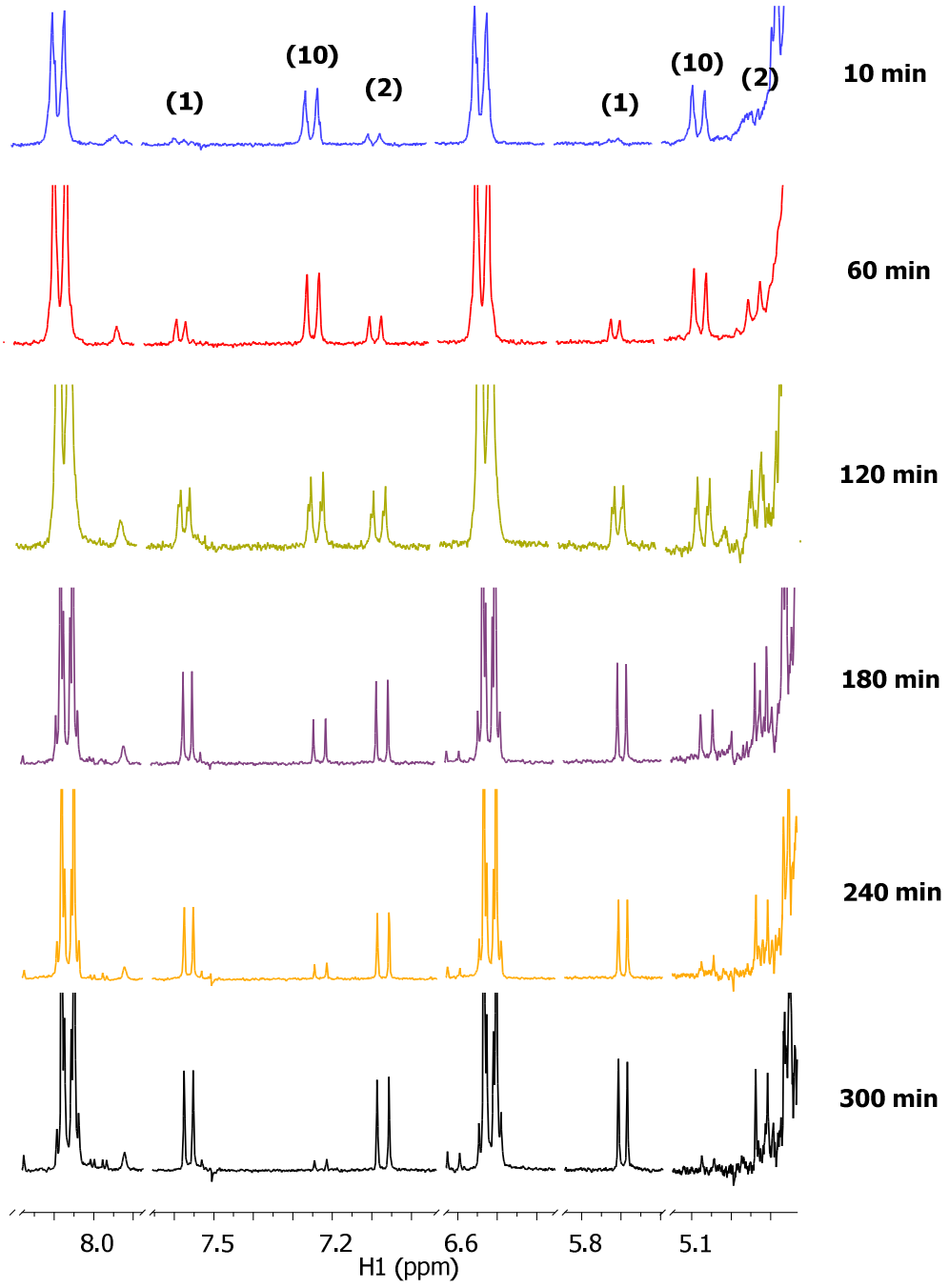


Figure 9.4: ^{1}H NMR for *(Z)*-3-ureidoacrylic peracid (**10**) over time^{1}H NMR (300 MHz, D_2O): δ 5.03 (d, J=9 Hz, 1H), 7.23 (d, J=9 Hz, 1H), decomposed to it acid: σ 4.92 (d, J=9 Hz, 1H), 7.07 (d, J=9 Hz, 1H) and uracil: δ 5.70 (d, J=7 Hz, 1H), 7.50 (d, J=7 Hz, 1H).

Unfortunately, the *(Z)*-3-ureidoacrylic peracid (**10**) is not stable for purification as it decomposed over time to form *(Z)*-3-ureidoacrylic acid (**2**) and uracil (**1**) along with some unidentified species, *figure 9.4*. The half life of the *(Z)*-3-ureidoacrylic peracid (**10**) was calculated to be 3 hrs at 25 °C, pH 10.0. The degradation of *(Z)*-3-ureidoacrylic peracid (**10**) to *(Z)*-3-ureidoacrylic acid (**2**) and uracil was confirmed by HPLC analysis of the degraded sample, *figure 9.5* as well as spiking the degraded sample with *(Z)*-3-ureidoacrylic acid (**2**) and uracil (data not shown).

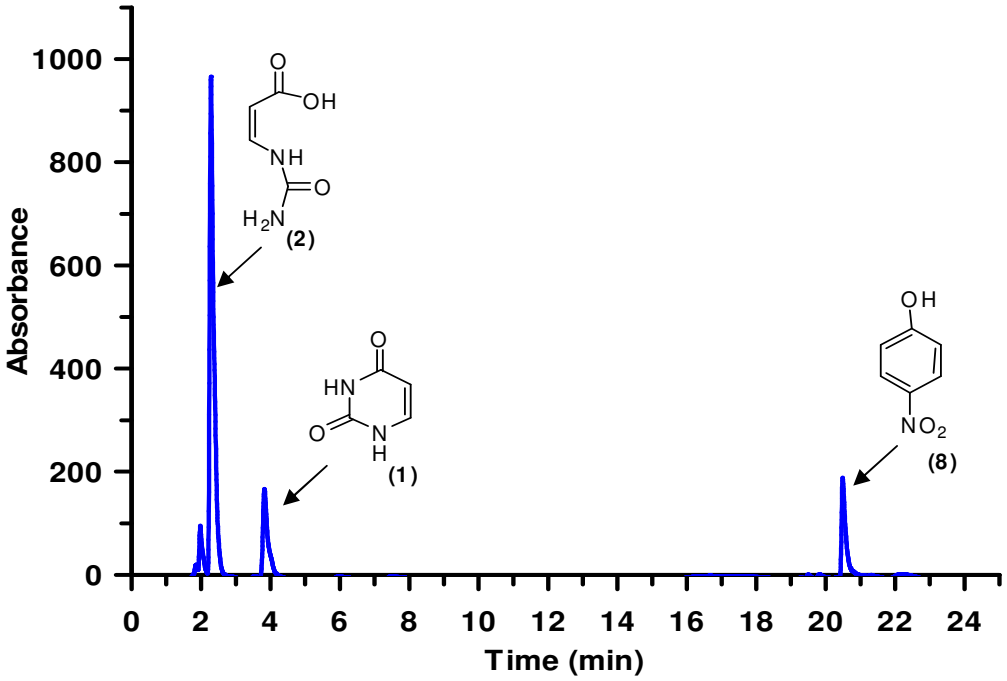


Figure 9.5: HPLC trace of the sample of *(Z)*-3-ureidoacrylic peracid (**10**) kept at 25 °C for over 5 hours

9.4. Discussion. A mechanism of the formation of *(Z)*-3-ureidoacrylic peracid (**10**) by a flavin dependent enzymatic reaction can be hypothesized *figure 9.6*. The flavin is first reduced to its active form $FMNH_2$ (**6**) by transferring of the reducing equivalent from NADH (**3**) by RutF (or Fre) which then reacts with molecular oxygen to form the flavin hydroperoxide (**11**) (*15*). The C4 of uracil (**1**) then undergoes a nucleophilic attack by the

flavin hydroperoxide (**11**) to form the tetrahedral intermediate (**12**) which then collapses to break the bond between N3 and C4 of uracil (**1**). C^{13}NMR experiments done with the product of this reaction mixture shows that the bond between N3 and C4 is broken (section 8.3.3). The lone pair on N of the isoaloxazine ring of (**13**) then kicks out the *(Z)*-3-ureidoacrylic peracid (**10**) to regenerate flavin in its oxidized form which is then ready to do the second round of catalysis. This is a novel reaction in terms of flavin dependent enzymology.

No evidence of the *(Z)*-3-ureidoacrylic peracid (**10**) was seen during the analysis of the enzymatic reactions (by NMR or HPLC). The reaction of RutA and RutF with uracil (**1**) done in the presence of H_2O^{18} did not show any C^{13} isotopic shift at C4 of *(Z)*-3-ureidoacrylic acid (**2**). The C^{13} isotopic shift was only observed when the reaction was done in the presence of $^{18}O_2$ (section 8.3.4). Washing away of the O^{18} label was not observed over a period of 4 days. The reactions were done in pH 8.0. This suggests that the *(Z)*-3-ureidoacrylic acid (**2**) was not formed as a result of a hydrolysis reaction but reduction of the *(Z)*-3-ureidoacrylic peracid (**10**).

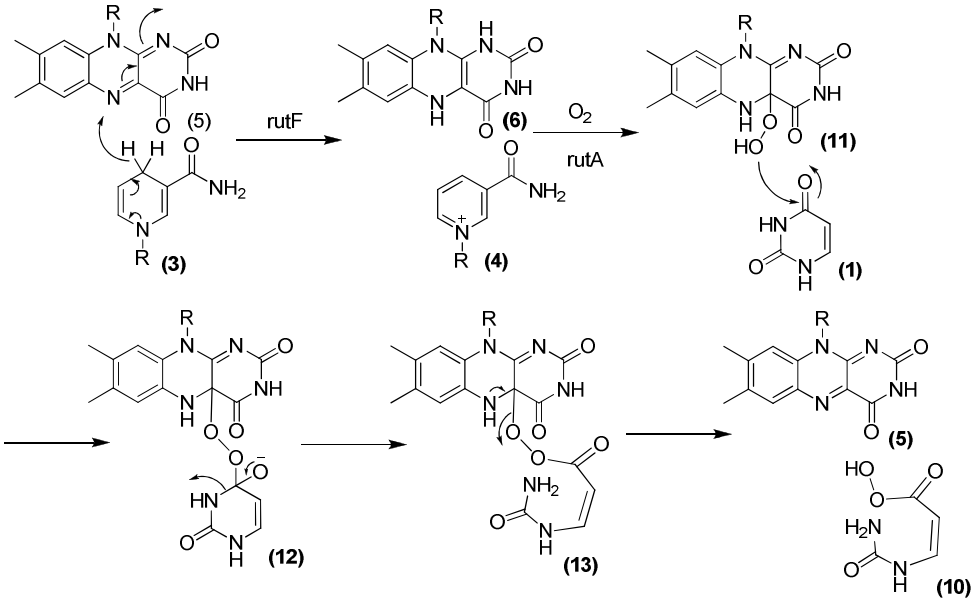


Figure 9.6: A plausible mechanism of the formation of *(Z)*-3-ureidoacrylic peracid (**10**).

The active site of RutA does not show the presence of any cysteine residues which could potentially have reduced the *(Z)*-3-ureidoacrylic peracid (**10**) to form *(Z)*-3-ureidoacrylic acid (**2**). Thus it is reasonable to assume that the RutA does not play any role in the reduction of the *(Z)*-3-ureidoacrylic peracid (**10**) to form *(Z)*-3-ureidoacrylic acid (**2**). The half life of the *(Z)*-3-ureidoacrylic peracid (**10**) is 3 hours at pH 10.0 at 25 °C. In the enzymatic reaction mixture there is an excess of NADH which in principal could reduce the *(Z)*-3-ureidoacrylic peracid (**10**). 1.5 equivalents of NADH was added to a freshly prepared sample of chemically synthesized *(Z)*-3-ureidoacrylic peracid (**10**) and within 5 minutes all the *(Z)*-3-ureidoacrylic peracid (**10**) was converted to *(Z)*-3-ureidoacrylic acid (**2**) *figure 9.7*. No peaks corresponding to uracil was observed, which suggested that the *(Z)*-3-ureidoacrylic peracid (**10**) was indeed reduced to form *(Z)*-3-ureidoacrylic acid (**2**). Had the presence of NADH catalyzed the degradation of the *(Z)*-3-ureidoacrylic peracid (**10**), we would have observed the formation of uracil (**1**) as well. This reaction was carried out at a pH>8.0.

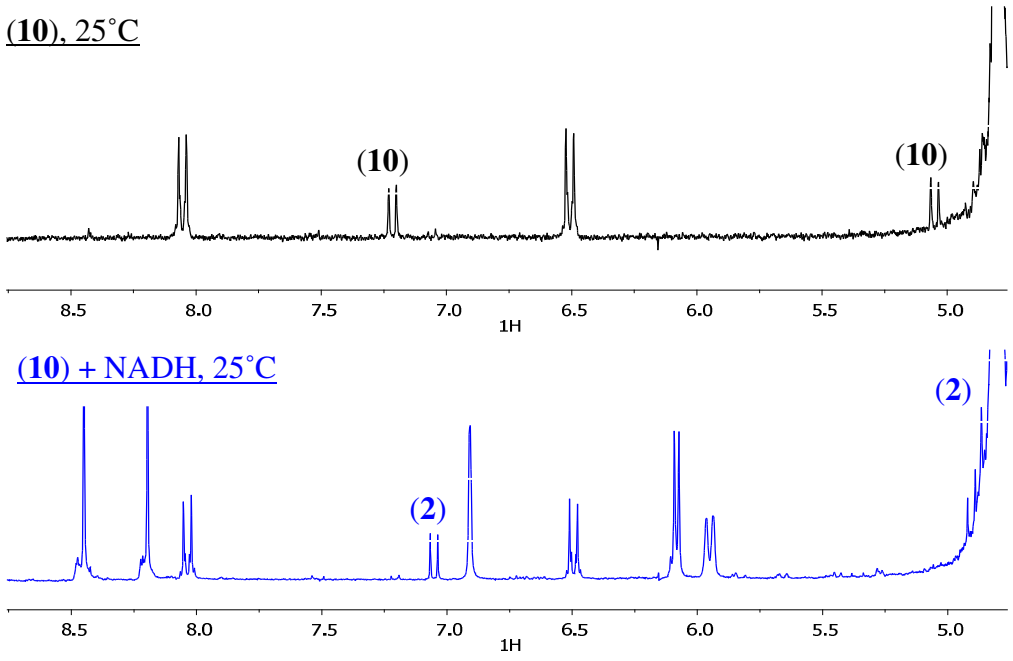


Figure 9.7: ^{1}H NMR of the peracid in presence and absence of NADH at 25 °C.

DTT was also present in the enzymatic reaction mixture which can also reduce the *(Z)*-3-ureidoacrylic peracid (**10**) to *(Z)*-3-ureidoacrylic acid (**2**), *figure 9.8*. Two equivalents of DTT was added to a freshly prepared solution of *(Z)*-3-ureidoacrylic peracid (**10**) which was instantaneously reduced to the *(Z)*-3-ureidoacrylic acid (**2**). No uracil (**1**) was seen in this case as well.

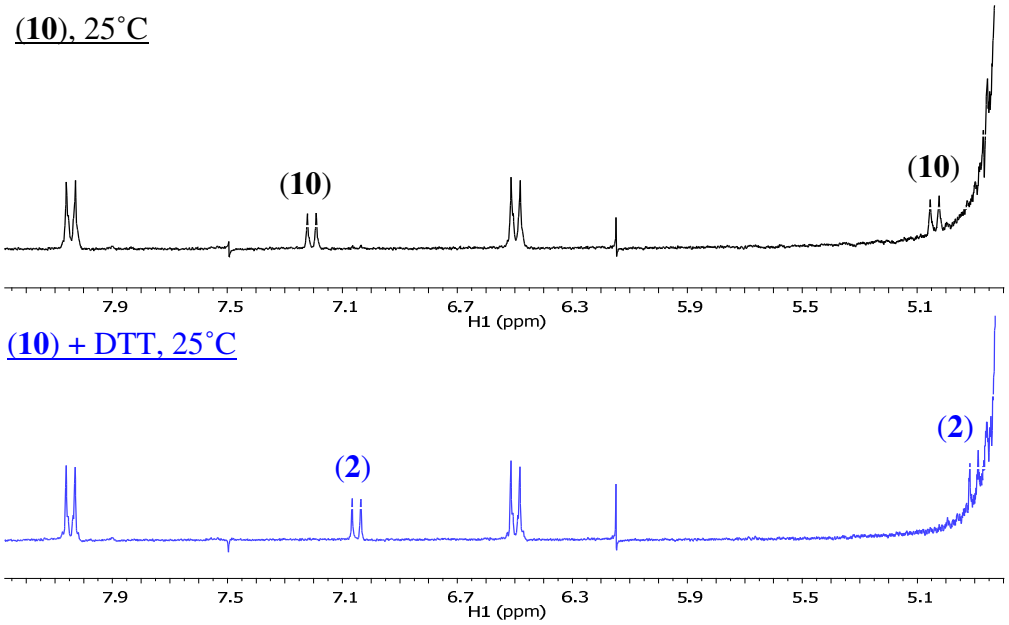


Figure 9.8: ^{1}H NMR of the peracid in presence and absence of DTT at 25 ˚C.

These results suggests that the *(Z)*-3-ureidoacrylic peracid (**10**) formed during the enzymatic reaction of RutA was spontaneously reduced to the *(Z)*-3-ureidoacrylic acid (**2**) non-enzymatically under the reaction condition.

9.5. Acknowledgement

This research was supported by a grant from the National Institutes of Health (GM069618). We thank Dr. Ivan Keresztes, Director of the NMR facility at the Department of Chemistry and Chemical Biology, Cornell University, for all the 2D NMRs.

REFERENCES

(1) Loh, K. D., Gyaneshwar, P., Markenscoff Papadimitriou, E., Fong, R., Kim, K. S., Parales, R., Zhou, Z., Inwood, W., and Kustu, S. (2006) A previously undescribed pathway for pyrimidine catabolism. *Proc Natl Acad Sci U S A 103*, 5114-9.

(2) Osterman, A. (2006) A hidden metabolic pathway exposed. *Proc Natl Acad Sci U S A 103*, 5637-8.

(3) http://theseed.uchicago.edu/FIG/index.cgi.

(4) Xun, L., and Sandvik, E. R. (2000) Characterization of 4-hydroxyphenylacetate 3-hydroxylase (HpaB) of Escherichia coli as a reduced flavin adenine dinucleotide-utilizing monooxygenase. *Appl Environ Microbiol 66*, 481-6.

(5) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem 72*, 248-254.

(6) Neises, B., and Steglich, W. (1978) Simple Method for the Esterification of Carboxylic Acids. *Angew. Chem. Int. Ed. 17*, 522-524.

(7) Campagne, J.-M., Coste, J., and Jouin, P. (1995) (1H-Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium Hexafluorophosphate- and (1H-Benzotriazol-1-yloxy)tripyrrolidinophosphonium Hexafluorophosphate-Mediated Activation of Monophosphonate Esters: Synthesis of Mixed Phosphonate Diesters, the Reactivity of the Benzotriazolyl Phosphonic Esters vs the Reactivity of the Benzotriazolyl Carboxylic Esters. *J. Org. Chem. 60*, 5214-5218.

(8) Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307-326.

(9) Keegan, R. M., and Winn, M. D. (2007) Automated search-model discovery and preparation for structure solution by molecular replacement. *Acta Crystallogr. D Biol. Crystallogr. 63*, 447-57.

(10) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr. 60*, 2126-32.

(11) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr. 53*, 240-55.

(12) Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr. 58*, 1948-54.

(13) Silbert, L. S. S., Elaine; Sworn, Daniel. (1962) Peroxides. IX. New method for the direct preparation of aromatic and aliphatic peroxy acids. *Journal of Organic Chemistry 27*, 1336-1342.

(14) Schneider, H. J. M., Walter. (1985) Selective functionalization of hydrocarbons. 6. Mechanistic and preparative studies on the regio- and stereoselective paraffin hydroxylation with peracids. *Journal of Organic Chemistry 50*, 4609-4615.

(15) Massey. V. (1994) Activation of molecular oxygen by flavins and flavoproteins. *Journal of Biological Chemistry 269*, 22459-22462.

CHAPTER 10

Summary and Outlook

10.1 Tryptophan 2,3-dioxygenase

10.1.1. ^{18}O labeling experiments. Tryptophan2,3-dioxygenase (TDO) is the first enzyme in the kynurenine pathway of tryptophan degradation which catalyzes the formation of N-formyl kynurenine (NFK) from tryptophan (*1*). We have directed our efforts towards the investigation of the mechanism of TDO by looking into the ^{18}O incorporation into the product of the enzymatic reaction. A similar experiment was done by Hayashi (*2*) but it is very costly to run an analogus experiment today and is not sensitive enough to detect low level of ^{18}O incorporation. Our experimental design is a marked improvement in terms of effciency, sensitivity and cost. The product NFK, was analyzed for ^{18}O incorporation by positive mode ESI-MS. The results of these experiments showed that there were no ^{18}O incorporation from water, both the oxygen atom in NFK came from molecular oxygen. This could happen via a Criegee rearrangement (*3*) or a dioxetane intermediate (*4, 5*). In case it is the former and if the putative carbocation generated is stable enough to allow for the exchange of oxygen bound to the iron with water before recombination, then we would be able to see it by our experimental procedure. In order to facilitate the exchange of the oxygen bound to iron with water, efforts were made to stabilize the putative carbocation by introducing an electron donating group at C-6 of tryptophan and so 6-hydroxytryptophan was synthesized as a possible candidate. It was however, not accepted as a substrate by TDO. While TDO is localized in the liver, indoleamine-2,3-dioxygenase, IDO, which substitutes TDO in the kynurenine degradation pathway of tryptophan, is present ubiquitously in the body. IDO is known to have a much higher tolerance to substrate modifications (*6, 7*). 6-hydroxytryptophan was accepted as a substrate by IDO. Positive mode ESI-MS analysis of the product(s) formed upon IDO catalyzed oxidation of 6-hydroxytryptophan, in presence of H_2O^{18}, showed no

incorporation of ^{18}O label from the solvent. This, however, was not conclusive evidence in favor of the dioxetane formation as the rate of exchange of the oxygen bound to iron with the solvent might still be much slower with respect to the rate of recombination with the putative carbocation formed during the Criegee rearrangement.

10.1.2. Dioxetane intermediate. If the reaction went via a dioxetane intermediate that decomposed by retro [2+2] cycloaddition, a light emission might be associated with it. Efforts were directed towards detection of light emission during the course of the TDO reaction, if any, by stopped flow experiments. No light emission was observed during the course of the reaction catalyzed by both TDO and IDO. Luciferase from firefly, *Photinus pyralis*, which catalyzes the oxidation of D-luciferin to oxy-luciferin is believed to go via a dioxetane intermediate and emits light(*8*). The minimum concentration of luciferase for which light could be detected under the experimental condition was 10^9 times less than the concentration of TDO used. However, this did not rule out the possibility of the dioxetane intermediate as the dioxetane intermediate, once formed, could open up in a non-concerted manner, in which case there would be no light emission.

10.1.3. Modeling and mutation studies on TDO. The crystal structure of TDO from *R. metallidurans* was solved at 2.4 Å resolution (*9*). Although the crystal structure revealed heme bound at the active site, there was no substrate bound to it. Models of the hydroperoxide intermediate were built into the active site and energy minimization revealed that the S-hydroperoxides were energetically more favored than the R-hydroperoxides. Mutants of all the relevant amino acid residues in the active site that interacted with the modeled in hydroperoxide were constructed and their activities were tested and steady state kinetic parameters were determined for the active ones. The results suggest that the S-antiperiplanar hydroperoxide could be the intermediate involved

in the reaction catalyzed by TDO. This suggests that the reaction route involving Criegee rearrangement might be the actual route of formation of NFK from tryptophan.

10.1.4. Outlook. No definitive proof for the Criegee rearrangement as the preferred route of oxidation of tryptophan has been shown. In order to see ^{18}O incorporation from the solvent the intermediate carbocation must be stabilized long enough to allow for the exchange of the oxygen bound to the iron with water before recombination. The ^{18}O labeling experiment was carried out at neutral pH, where the exchange of the oxygen bound to the heme with water will be very slow. Likelihood of the exchange could be enhanced at a lower pH. However the labiality of the ketonic oxygen in NFK would also increase under these conditions, and this would pose a problem in accurate determination of whether the oxygen incorporated in the ketonic position of NFK is due to exchange of the oxygen bound to heme or due to non-enzymatic exchange of the labile oxygen with the solvent. If the two rates mentioned above are sufficiently different and if the reaction indeed went via the Criegee intermediate then we would be able to observe ^{18}O incorporation in the product. The crystal structure, failure to detect light emission during the course of the reaction as well as the modeling studies favors the Criegee rearrangement. Thus it is worth the effort to repeat the labeling experiment at lower pH for the enzymatic reaction to see if there is any ^{18}O incorporation from water in the NFK generated. If ^{18}O incorporation is observed then repeating the same experiment with synthesized NFK in absence of the enzyme would tell us whether the ^{18}O exchange was indeed enzymatic or not.

10.2. Vitamin B_6 catabolism

10.2.1. 4-Pyridoxic acid dehydrogenase. The function of the mlr6793 gene from

Mesorhizobium loti MAFF303099 was identified (*10*). This gene encodes 4-Pyridoxic acid dehydrogenase, an enzyme involved in the catabolism of PLP (Vitamin B_6). This enzyme was overexpressed in *Escherichia coli* and characterized. 4-Pyridoxic acid dehydrogenase is a 33 kDa protein that catalyzes the four electron oxidation of 4-pyridoxic acid to 3-hydroxy-2-methylpyridine-4,5-dicarboxylate, using nicotinamide adenine dinucleotide as a cofactor. The k_{cat} for NADH production is 0.01 s^{-1}. The K_M values for 4-pyridoxic acid and NAD are 5.8 μM and 6.6 μM respectively.

10.2.2. 3 hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase. The function of the mlr6791 gene from *Mesorhizobium loti* MAFF303099 was identified (*11*). This gene encodes 3 hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase (HMPDdc), an enzyme involved in the catabolism of pyridoxal 5-phosphate (Vitamin B_6). This enzyme was overexpressed in *Escherichia coli* and characterized. HMPDdc is a 26 kDa protein that catalyzes the decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate to 3-hydroxy-2-methylpyridine-5-carboxylate. The K_M and k_{cat} were found to be 366 μM and 0.6 s^{-1} respectively. The structure of this enzyme was determined at 1.9 Å resolution using SAD phasing and belongs to the class II aldolase/adducin superfamily. While the decarboxylation of hydroxy-substituted benzene rings is a common motif in biosynthesis, the mechanism of this reaction was characterized. The structural studies suggested that catalysis of such decarboxylations proceeded by an aldolase-like mechanism.

10.2.3. E- 2-(acetamidomethylene)succinate hydrolase. The function of the mlr6787 gene from *Mesorhizobium loti* MAFF303099 was identified (*12*). This gene encodes 2-(acetamidomethylene)succinate hydrolase, an enzyme involved in the catabolism of pyridoxal 5'-phosphate (Vitamin B_6). This enzyme was overexpressed in *Escherichia coli*, purified to homogeneity and characterized. 2-(Acetamidomethylene)succinate hydrolase

catalyzes the hydrolysis of 2-(acetamidomethylene)succinate to yield succinic semialdehyde, acetic acid, carbon dioxide and ammonia. The k_{cat} and K_M for this reaction were 0.6 s^{-1} and 143 μM, respectively. The enzyme was shown to utilize the *E* isomer of 2-(acetamidomethylene)succinate.

10.2.4. 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. One of the enzymes on the PLP degradative pathway, 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase, is a flavin-dependent enzyme and catalyzes the ring opening of 2-methyl-3-hydroxypyridine-5-carboxylic acid to form *E*-2-acetaminomethylene succinate. The gene for this enzyme was cloned and the corresponding protein was overexpressed in *Escherichia coli* and purified. The crystal structure of MHPCO has been solved to 2.1 Å using SAD phasing with and without the substrate MHPC bound (*13*). These crystal structures provide insight into the reaction mechanism and suggest roles for active site residues in the catalysis of a novel oxidative ring-opening reaction. Chemical reactions mimicking the enzymatic reaction with substrate analogues were performed which through some light on the reaction mechanism of MHPCO.

10.2.5. Outlook.

10.2.5.1. Mechanistic studies on MHPCO. The mechanism of the ring cleavage reaction performed by the MHPCO is not yet established. The chemical model reactions that were done with 3-methoxy-5-(methoxycarbonyl)-1-methylpyridinium suggested that the enzymatic reaction could occur by a nucleophilic attack of the flavin hydroperoxide on the pyridine ring. The end product of this chemical transformation has not been characterized. If it is characterized to be *E*-dimethyl 2-((N-methylformamido)methylene)succinate then it is a definitive proof that the reaction does not go via a ketene intermediate as no such intermediate can be formed from the starting

material. Attempts to isolate and characterize the product of the chemical reaction of 3-methoxy-5-(methoxycarbonyl)-1-methylpyridinium with H_2O_2 failed. It is thus necessary to synthesize *E*-dimethyl 2-((N-methylformamido)methylene)succinate and see if it co-elutes with the product of the chemical reaction by HPLC. A mass spectrum of this HPLC peak corresponding to the product would also characterize the product. This study will lead to further experiments in assigning the most likely mechanism for the catalytic transformation by MHPCO.

10.2.5.2. Identification of PLP catabolic genes in other organisms. *Mesorhizobium loti* MAFF303099 is the first organism in which all the genes participating in the PLP degradative pathway have been identified. Surprisingly however, when we search for an equivalent pathway in the available sequenced bacterial genomes using the SEED database (*14*), we find that this pathway is unique to *Mesorhizobium loti* suggesting that other vitamin B_6 catabolic pathways still remain to be discovered. Apart from gene identification in *Mesorhizobium loti,* two genes encoding for pyridoxal 4-dehydrogenase (*15*) and pyridoxine 4-oxidase (*16, 17*) have been identified and characterized in *Microbacterium luteolum* YK-1. The gene coding for pyridoxal 4-dehydrogenase has also been observed in *Aureobacterium luteolum (18)*. The remaining genes encoding for the protein participating in the degradation of vitamin B_6 in these two organisms are yet to be discovered. In *Pseudomonas* IA and *Arthrobacter* Cr-7 a slightly different pathway of vitamin B_6 degradation occurs. The participating genes and their respective gene products have not been characterized. Characterization of the genes of the above three organisms would throw light upon the diversified vitamin B_6 degradative pathway.

10.2.5.3. Characterization of 5-pyridoxic acid oxygenase. MHPCO and 5-pyridoxic acid oxygenase are the only two known enzymes which catalyzes ring cleavage reaction of aromatic compounds without the use of metal cofactor (*19*). The latter enzyme is from *Pseudomonas* IA and *Arthrobacter* Cr-7 (*20*). MHPCO has been characterized significantly

(*19, 21-26*), but very little characterization of the 5-pyridoxic acid oxygenase has been done(*20*). This will also be an interesting enzyme to perform crystallographic studies on as the substrate for the two enzymes are significantly different at C4 of the pyridine ring. The carboxyl group at C4 of the pyridine ring is first removed by the decarboxylase (*11*) to produce 3-hydroxy-2-methylpyridine-5-carboxylate which is the substrate for MHPCO. The 5-pyridoxic acid oxygenase seems to tolerate substitution at C4 of the pyridine ring. Since these are the only two enzymes of its kind and accepts slightly different substrates it would be important to understand the structural similarities and their differences. The substrate tolerance of both these enzymes has not been rigorously tested. If the 5-pyridoxic acid oxygenase has a more open active site, this enzyme would be the ideal candidate to study the mechanism for the oxidative ring opening reaction by using substrate analogous.

10.2.5.4. Identification of a PLP transporter. The PLP catabolic pathway is induced under stress condition where the vitamin B_6 is the sole source of carbon and nitrogen (*27*). This is an ideal system to look for a PLP transporter in microorganisms as it is reasonable to assume that under the stress condition there will be an increase in the transcription of the transporter gene. Two ABC transporters (mlr6799 and mlr6804) are present in the neighborhood of the PLP catabolic genes. Preliminary studies show that the mlr6799 is not the ideal candidate; however it has not been shown rigorously. The other gene product has not been analyzed. A transcription profile of this organism under stress condition needs to be completed in order to identify any the suitable candidate genes which would serve as a transporter for vitamin B_6 in cells.

10.3. Pyrimidine catabolism – the rut pathway.

10.3.1. Characterization of the rut genes. The rut genes were shown to be involved in the degradation of uracil (and thymine) in *Escherichia coli* K12 (*28, 29*). No functions were

assigned to the genes. Our work showed that rutA catalyzes the ring opening reaction of uracil via a flavin dependent hydrolysis at C4 to form 3-ureidoacrylate which is subsequently taken up by rutB to form two molecules of ammonia, carbon dioxide and 3-oxopropionate. We have shown that rutF clarified lysate has the flavin reductase activity and assists rutA in the first step of pyridine degradation.

10.3.2. Mechanistic analysis. The mechanism of the reaction catalyzed by rutA has been looked into and evidences suggest that 3-ureidoacrylic peracid is the actual product of the enzymatic reaction, which is reduced by DTT and/or NADH present in the enzymatic assay mixture. The crystal structure of the enzyme has been solved at 1.9 Å resolution.

10.3.3. Outlook. The 3-ureidoacrylic peracid is not very stable and hence it is difficult to isolate and characterize. An HPLC method needs to be developed at pH = 10.0, where the peracid is stable long enough for characterization. This will provide for a direct evidence of its formation. Complete kinetic characterization of rutA needs to be done. Effort should be directed towards obtaining a pure preparation of rutF. Subcloning it into a THT vector from pCA24N vector, fusing it with SUMO protein and coexpressing it with chaperone proteins were done, but nothing yielded pure preparation of rutF. A C-terminal His-tag construct needs to be made and purified to see if soluble pure preparation of the protein is obtained. The activity of the other genes in the pathway (rutC, rutD and rutE) has not been assigned. Their characterization is necessary to complete the characterization of the entire operon. There are two steps in the pathway that are happening non-enzymatically – the reduction of the 3-ureidoacrylate peracid to 3-ureidoacrylate and the hydrolysis of 3-aminoacrylate to 3-oxo-propionate. These steps could be easily catalyzed by the other enzymes in the operon.

REFERENCES

(1) Bugg, T. D. H. (2003) Dioxygenase enzymes: catalytic mechanisms and chemical models. *Tetrahedron 59*, 7075.

(2) Hayaishi, O., Rothberg, S., Mehler, A. H., and Saito, Y. (1957) Studies on oxygenases; enzymatic formation of kynurenine from tryptophan. *J Biol Chem 229*, 889-96.

(3) Bugg, T., and Winfield, C. (1998) Enzymic cleavage of aromatic rings: mechanistic aspects of the catechol dioxygenases and later enzymes of bacterial oxidative cleavage pathways. *Natural Product Reports 15*, 513.

(4) Leeds, J. M., Brown, P. J., McGeehan, G. M., Brown, F. K., and Wiseman, J. S. (1993) Isotope effects and alternative substrate reactivities for tryptophan 2,3-dioxygenase. *J Biol Chem 268*, 17781-6.

(5) Terentis, A. C., Thomas, S. R., Takikawa, O., Littlejohn, T. K., Truscott, R. J., Armstrong, R. S., Yeh, S. R., and Stocker, R. (2002) The heme environment of recombinant human indoleamine 2,3-dioxygenase. Structural properties and substrate-ligand interactions. *J Biol Chem 277*, 15788-94.

(6) Littlejohn, T. K., Takikawa, O., Skylas, D., Jamie, J. F., Walker, M. J., and Truscott, R. J. (2000) Expression and purification of recombinant human indoleamine 2, 3-dioxygenase. *Protein Expr Purif 19*, 22-9.

(7) Sono, M., Roach, M. P., Coulter, E. D., and Dawson, J. H. (1996) Heme-Containing Oxygenases. *Chem Rev 96*, 2841-2888.

(8) DeLuca, M. a. M., W. D. (1978) Purification and Properties of Firefly Luciferase. *Methods in Enzymology* ***LVII***, 3.

(9) Zhang, Y., Kang, S. A., Mukherjee, T., Bale, S., Crane, B. R., Begley, T. P., and Ealick, S. E. (2007) Crystal structure and mechanism of tryptophan 2,3-

dioxygenase, a heme enzyme involved in tryptophan catabolism and in quinolinate biosynthesis. *Biochemistry 46*, 145-55.

(10) Mukherjee, T., Kinsland, C., and Begley, T. P. (2007) PLP catabolism: Identification of the 4-pyridoxic acid dehydrogenase gene in Mesorhizobium loti MAFF303099. *Bioorg Chem 35*, 458-464.

(11) Mukherjee, T., McCulloch, K. M., Ealick, S. E., and Begley, T. P. (2007) Gene Identification and Structural Characterization of the Pyridoxal 5'-Phosphate Degradative Protein 3-Hydroxy-2-methylpyridine-4,5-dicarboxylate Decarboxylase from Mesorhizobium loti MAFF303099. *Biochemistry 46*, 13606 - 13615.

(12) Mukherjee, T., Hilmey, D. G., and Begley, T. P. (2008) PLP catabolism: identification of the 2-(Acetamidomethylene)succinate hydrolase gene in Mesorhizobium loti MAFF303099. *Biochemistry 47*, 6233-41.

(13) McCulloch, K. M., Mukherjee, T., Ealick, S. E., and Begley, T. P. (2009) Structure of the PLP Degradative Enzyme 2-Methyl-3-hydroxypyridine-5-carboxylic Acid Oxygenase from Mesorhizobium loti MAFF303099 and Its Mechanistic Implications. *Biochemistry*.

(14) http://theseed.uchicago.edu/FIG/index.cgi.

(15) Yokochi, N., Yoshikane, Y., Trongpanich, Y., Ohnishi, K., and Yagi, T. (2004) Molecular Cloning, Expression, and Properties of an Unusual Aldo-Keto Reductase Family Enzyme, Pyridoxal 4-Dehydrogenase, That Catalyzes Irreversible Oxidation of Pyridoxal. *J. Biol. Chem. 279*, 37377-37384.

(16) Yoshikane, Y., Yokochi, N., Ohnishi, K., and Yagi, T. (2004) Coenzyme precursor-assisted cooperative overexpression of an active pyridoxine 4-oxidase from Microbacterium luteolum. *Protein Expression Purif. 34*, 243-248.

(17) Kaneda, Y., Ohnishi, K., and Yagi, T. (2002) Purification, molecular cloning, and characterization of pyridoxine 4-oxidase from Microbacterium luteolum. *Biosci., Biotechnol., Biochem. 66*, 1022-1031.

(18) Trongpanich, Y., Abe, K., Kaneda, Y., Morita, T., and Yagi, T. (2002) Purification and characterization of pyridoxal 4-dehydrogenase from Aureobacterium luteolum. *Biosci., Biotechnol., Biochem. 66*, 543-548.

(19) Chaiyen, P., Sucharitakul, J., Svasti, J., Entsch, B., Massey, V., and Ballou, D. P. (2004) Use of 8-substituted-FAD analogues to investigate the hydroxylation mechanism of the flavoprotein 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 43*, 3933-3943.

(20) Nelson, M. J., and Snell, E. E. (1986) Enzymes of vitamin B6 degradation. Purification and properties of 5-pyridoxic-acid oxygenase from Arthrobacter sp. *J Biol Chem 261*, 15115-15120.

(21) Kishore, G. M., and Snell, E. E. (1981) Kinetic investigations on a flavoprotein oxygenase, 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *J Biol Chem 256*, 4228-4233.

(22) Kishore, G. M., and Snell, E. E. (1981) Interaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase with FAD, substrates, and analogues. Spectral and fluorescence investigations. *J Biol Chem 256*, 4234-4240.

(23) Chaiyen, P., Ballou, D. P., and Massey, V. (1997) Gene cloning, sequence analysis, and expression of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Proc Natl Acad Sci U S A 94*, 7233-7238.

(24) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Reaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid (MHPC) oxygenase with N-methyl-5-hydroxynicotinic acid: studies on the mode of binding, and protonation status of the substrate. *Biochemistry 36*, 13856-13864.

(25) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Unusual mechanism of oxygen atom transfer and product rearrangement in the catalytic reaction of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 36*, 8060-8070.

(26) Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) Thermodynamics and reduction kinetics properties of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry 36*, 2612-2621.

(27) Rodwell, V. W., Volcani, B. E., Ikawa, M., and Snell, E. E. (1958) Bacterial oxidation of vitamin B6. I. Isopyridoxal and 5-pyridoxic acid. *J Biol Chem 233*, 1548-1554.

(28) Loh, K. D., Gyaneshwar, P., Markenscoff Papadimitriou, E., Fong, R., Kim, K. S., Parales, R., Zhou, Z., Inwood, W., and Kustu, S. (2006) A previously undescribed pathway for pyrimidine catabolism. *Proc Natl Acad Sci U S A 103*, 5114-9.

(29) Osterman, A. (2006) A hidden metabolic pathway exposed. *Proc Natl Acad Sci U S A 103*, 5637-8.