Theses and Dissertations

Summer 2013

A. Genetic characterization of the caffeine C-8 oxidation pathway in Pseudomonas Sp. CBB1 B. Validation of caffeine dehydrogenase as a suitable enzyme for a rapid caffeine diagnostic test

Sujit Kumar Mohanty
University of Iowa

www.manaraa.com

A. GENETIC CHARACTERIZATION OF THE CAFFEINE C-8 OXIDATION

PATHWAY IN *PSEUDOMONAS* SP. CBB1

B. VALIDATION OF CAFFEINE DEHYDROGENASE AS A SUITABLE ENZYME

FOR A RAPID CAFFEINE DIAGNOSTIC TEST

by

Sujit Kumar Mohanty

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Chemical and Biochemical Engineering
in the Graduate College of
The University of Iowa

August 2013

Thesis Supervisor: Professor Mani Subramanian

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

PH.D. THESIS

_____

This is to certify that the Ph.D. thesis of

Sujit Kumar Mohanty

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Chemical and Biochemical Engineering at the August 2013
graduation.

Thesis Committee: _____

Mani Subramanian, Thesis Supervisor

_____

David Murhammer

_____

Tonya Peeples

_____

Eric Nuxoll

_____

David Weiss

To all my professors, and teachers whose blessings, guidance and constant encouragement have been my beacon and paved my path to success.

Education is not the learning of facts, but the training of the mind to think.

-Albert Einstein

ACKNOWLEDGMENTS

The journey, which started in 2002 as a student of Biotechnology Engineering from MIT School of biotechnology, Berhampur University and ending with a doctoral degree in Chemical and Biochemical Engineering at The University of Iowa in 2013, has been through many twists and turns, ups and downs. It is obvious that in a span of these eleven years of student life many people have contributed their time, effort and blessings to help me sail safely and confidently through the struggling path. I am sincerely thankful to all who played a role in my life during those wonderful and memorable school years, no matter how little we interacted. Your presence and contributions, even in the tiniest amount, have helped me in ultimately finishing a work of this magnitude. It is my privilege to thank each and every one of you here.

First of all, I would like to express my deep gratitude to my Ph.D. thesis advisor and mentor, Professor Mani Subramanian, for his tireless efforts in guiding me through graduate school, modeling for me the code of conduct of a scientist, and training me in independent research. He reassured me with his unconditional faith in me from the very commencement of my academic journey at the University of Iowa. His commitment and generosity in molding even the naïve fresher into an outstanding engineer has been a blessing and a valuable lesson for me. Let it suffice to say that a guru's place is next to none in a student's life and it is to him that ultimately belongs the glory and credit of all his student's work. I would also like to thank my thesis committee members Dr. Peeples, Dr. Murhammer, Dr. Nuxoll, and Dr. Weiss for their timely reviews and suggestions which helped me in shaping up this work to its final form. As a teaching assistant with Dr. Peeples, I acquired hands-on experience in teaching methodologies, which I believe will be extremely helpful to me in achieving my academic goals,

My sincere thanks to all my seniors and co-workers in Dr. Subramanian's lab. Dr. Chi-li Yu, Dr. Shuvendu Das, Dr. Michael Louie, Dr. Ryan Summers, Dr. Sridhar Gopishetty, Dr. James Glenn, all of you have been instrumental in extending crucial technical help in a timely manner in my day-to-day research work. Dr. Yu was always around to help me in protein purification, characterization, gene cloning and all other microbial analysis techniques. Dr. Louie and Dr. Summers were always prompt in reciprocating to discussions concerned with this research of any related advanced molecular bio-techniques. Dr. Das and Dr. Gopishetty were active in providing inputs in terms of organic/analytical chemistry and structural analysis, including, but not limited to any technical questions pertaining to high-end instruments like LC-MS, NMR, etc. I will always be thankful to my previous co-workers, Bhavita Bhatt and Kaitlin Louie, for their diligent support in organizing and maintaining the daily chores of the lab which further facilitated me in conducting experiments without interruptions. And last, but not the least, Dr. Todor Genkov, whose charming discussions kept my spirits high and my mind productive through the toil of many workaday routines.

I would like extend my sincere thanks to Dr. Santhana Velupillai, and Prof. Amnon Kohen and his student Eric Kohen in the Chemistry Department, for helping me with various experiments which were conducted in their respective laboratories. Our department secretaries ---- Linda Wheatley, Natalie Potter, Kay Spidle, Wendy Meyers ---- all deserve my deepest gratitude for their warm help within all student-related activities and their promptness in helping me in processing orders. Being an international student, it would have been impossible to even be functional in graduate school without their unwavering efforts in undergirding the university machinery. Last, but not the least, I would like to thank my parents, my teachers, and all my friends who have always extended their constant encouragement towards academic excellence. I will always be thankful to all of them and obliged, for their contribution in my academic success.

ABSTRACT

Pseudomonas sp. CBB1 degraded caffeine via C-8 oxidation. Previously, a novel quinone-dependent caffeine dehydrogenase (Cdh) was shown to catalyze the oxidation of caffeine to 1,3,7-trimethyluric acid (TMU). Initial metabolite analysis using resting cells and partially purified extract of CBB1 identified transient accumulation of 1,3,7-trimethyl-5-hydroxyisourate (TM-HIU), and 3,6,8-trimethylallantoin (TMA). TMA structure was confirmed; chiral analysis revealed that it was racemic. Further, a time-course reaction showed nine-fold higher enrichment of one enantiomer of TMA than the other, which gradually racemized in three hours. Thus, it was proposed that TMU was converted to TM-HIU and enantiomeric TMA.

A 43-kDa NADH-dependent TMU mononxygenase (TmuM) was purified and shown to convert TMU to unstable TM-HIU. The enzyme belonged to a new family of FAD-dependent monooxygenases. The enzyme was specific for methyluric acid with no activity on uric acid. Homology modeling of TmuM revealed a larger, more hydrophobic active site compared to analogous uricase in the uric acid pathway.

Genes encoding heterotrimeric Cdh (*cdhA,B,C*) and TmuM (*tmuM*) were located on a 25.2-kb fragment in the CBB1 genome. Gene cluster analysis relative to a similar cluster in uric acid degrading organisms identified five more putative genes, namely *tmuH*, *tmuD*, *orf1*, orf2, and *orf3,* of the C-8 oxidation pathway. First three genes (*tmuH*, *tmuD*, and *orf1*) are proposed to encode for TM-HIU hydrolase, TM-OHCU decarboxylase, and trimethylallantoinase, respectively, which are responsible for converstion of TM-HIU to TM-OHCU, TM-OHCU to proposed S-(+)-TMA, and S-(+)-TMA to TMAA. Two more genes (*orf2* and *orf3*) present downstream of *orf1* are proposed to encode for a YlbA like hydrolase, and a ArgE like deacetylase, respectively, which are proposed to be involved in degradation of TMAA to glyoxylate, di- and

vi

monomethylurea. This is the first report of (a) TMA structure (b) TMU monooxygenase and TM-HIU (hydroxylation product of TMU), and (c) complete delineation of C-8 oxidation pathway by a combination of enzymology and gene cluster analysis.

Excessive consumption of caffeine in various forms has created a need for a rapid diagnostic test, especially for nursing mothers and infants. Cdh was hypothesized to be suitable for this test. Sensitivity of the test was shown to be 1 ppm. A colorimetric test with partially purified Cdh was optimized to detect caffeine within a minute in drugs, nursing mother's milk, and to differentiate decaffeinated beverages.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

xvi

LIST OF ABBREVIATIONS

Abbreviations

BLAST   Basic Local Alignment Search Tool

Cdh   Caffeine Dehydrogenase

FAD   Flavin Adenine Dinucleotide

HPLC   High Performance Liquid Chromatography

HRESIMS   High Resolution Electro-Spray Ionization Mass Spectroscopy

INT   Iodonitrotetrazolium Chloride

LC-ESI-MS   Liquid Chromatography Electro-Spray Ionization Mass Spectroscopy

MSA   Multiple Sequence Alignment

NBT   Nitro Blue Tetrazolium

NMR   Nuclear Magnetic Resonance.

ORF   Open Reading Frame

PCR   Polymerase Chain Reaction

SDS-PAGE   Sodium Dodecyl Sulfate- Polyacrylamide Gel Electrophoresis

TMA   3,6,8- Trimethylallantoin

TMAA   Trimethylallantoic Acid

TM-HIU   1,3,7-Trimethyl-5-Hydroxyisourate

TM-OHCU   3,6,8- Trimethyl-2-Oxo-4-Hydroxy-4-Carboxy-5-Ureidoimidazoline

TMU   1,3,7-Trimethyluric Acid

TmuM   Trimethyluric Acid Monooxygenase

GRAVY   Grand Average of Hydrophathicity

CHAPTER 1

INTRODUCTION

Caffeine: History, Natural Occurrence and Significance

History

Caffeine (1,3,7–trimethylxanthine) is one of the most popular and commercially important plant derived purine alkaloids. Caffeine is a key component of widely consumed beverages, including coffee, tea, soft drinks, energy drinks. [Parliament, 2000]. It is a key additive to at least 150 other products that are available currently in the market (Table D1). Caffeine and related methylxanthines (Fig. 1.1) have been the subject of extensive research in recent years [Weinberg, 2001] primarily because of two reasons: their wide spread occurrence in nature [Ashihara *et al*., 2001] and their long-standing history of human consumption, possibly since the Stone Age [Escohotado *et al*., 1999]. Although, the origin of caffeine in human food (possibly somewhere from central Africa) and its way into other parts of the world is still is not certain [Weinberg, 2001], its use as a mood and performance enhancing drink in form of coffee (Africa, Europe), tea (Asia) or guarana (Amazon Basin) has been well known for centuries. In the early 1820s, a German chemist named Friedrich Ferdinand Runge isolated caffeine from the coffee plant for the first time and coined the name *Kaffein* (German for caffeine). Since then caffeine and other methylxanthines have been identified and isolated from various plant sources, studied extensively for their physiological effects, and used in applications ranging from foods and drugs to cosmetics [Parliament, 2000].

s

Methylxanthines   **Methyluric acids**
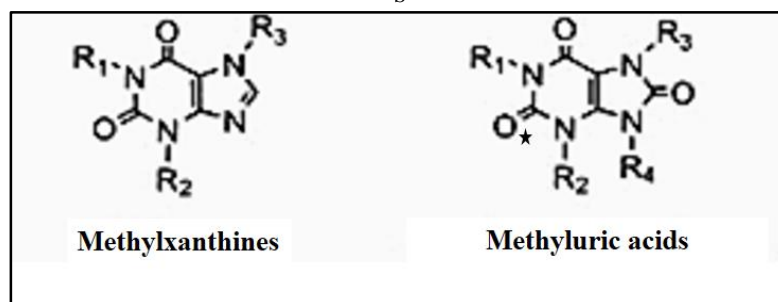
Compounds	Trivial Name	R_1	R_2	R_3	R_4	$O*$
Methylxanthines						
Xanthine		H	H	H	-	-
1-Methylxanthine		CH_3	H	H	-	-
3-Methylxanthine		H	CH_3	H	-	-
7-Methylxanthine		H	H	CH_3	-	-
1,3-Dimethylxanthine	Theophylline	CH_3	CH_3	H	-	-
1,7-Dimethylxanthine	Paraxanthine	CH_3	H	CH_3	-	-
3,7-Dimethylxanthine	Theobromine	H	CH_3	CH_3	-	-
1,3,7-Trimethylxanthine	Caffeine	CH_3	CH_3	CH_3	-	-
Methyluric acids						
Uric acid		H	H	H	H	-
1,3,7-Trimethyluric acid		CH_3	CH_3	CH_3	H	-
1,3,7,9-Trimethyluric acid	Theacrine	CH_3	CH_3	CH_3	CH_3	-
O(2),1,9-Trimethyluric acid	Liberine	CH_3	H	H	CH_3	CH_3
O(2),1,7,9-Tetramethyluric acid	Methylliberine	CH_3	H	CH_3	CH_3	CH_3

Figure 1.1. Structure of purine alkaloids based on xanthine and uric acid skeletons.

Discovery of caffeine as a psychoactive drug has led to its utilization in unnatural (synthesized) food products, beverages such as soft drinks, energy drinks, etc., and in the formulation of pharmaceuticals. However, recent toxicological studies have raised a

growing concern over its unrestricted use and excessive consumption [McGee *et al*., 1980; Mrvos *et al*., 1989; Garriott *et al*., 1985; Eteng *et al*., 1997].

Natural Occurrence

Caffeine is a member of the purine alkaloid group of metabolites widely found in the leaves and fruits of many types of plants [Ashihara *et al*., 2001]. In nature, caffeine and related purine alkaloids are produced by at least 13 different orders of plant kingdom (nearly 100 species), including *Coffea arabica* (coffee), *Camellia sinensis* (tea), *Ilex paraguariensis* (yerba mate), *Paullinia cupana* (guarana), *Cola nitida* (Cola), and *Theobroma cacao* (cacao, or cocoa) [Ashihara *et al*., 1999; Ashihara *et al*., 2001; Ashihara *et al*., 2008]. Among these purine-alkaloid producing plants, most studies have been carried out with species belonging to the genera *Camellia* and *Coffea* (Table 1.1) [Anaya *et al*., 2006]. In plants, caffeine is synthesized mostly in the young leaves and fruits during the growth phase and accumulates to highest concentration by maturation. Caffeine content in plants varies significantly depending on the species and plant organ (Table 1.1). In the same plant, different organs may accumulate different levels of caffeine and other purine alkaloids [Ashihara *et al*., 2008]. By virtue of its toxic effects, caffeine is believed to be playing a crucial role in plants in terms of protection and reproduction, as explained by two separate theories. The "chemical defense theory" suggests that caffeine in young leaves, fruits and flower buds defends the soft tissues from certain pests and insects by acting on their nervous system [Hollingsworth *et al*., 2002]. The "allelopathic theory" suggests that caffeine present in seed coats/falling

leaves gets released into the soil, where itinhibits the germination of competing seeds of other plants [Suzuki *et al*., 1987; Ashihara *et al*., 2008].

Table 1.1. Caffeine content in plants known to produce purine alkaloids *.

Latin Name	Common Name	Organ	Caffeine Content
Coffea arabica	Arabica coffee	Mature, ripened beans	1.0%
Coffea canephora	Robusta coffee	Mature, ripened beans	1.7%
Coffea kianjavatensis	Mascarocoffea	Mature, ripened beans	0.7%
Coffea dewevrei		Mature, ripened beans	1.2%
Coffea liberica		Mature, ripened beans	1.4%
Coffea eugeniodes		Mature, ripened beans	0.4%
Coffea salvatrix		Mature, ripened beans	0.7%
Coffea racemosa		Mature, ripened beans	0.8%
Camellia sinensis	Tea	Young Leaves	2.8%
Camellia assamica	Assam tea	Young Leaves	2.4%
Camellia taliensis		Young Leaves	2.5%
Paulinia cupana	Guaraná	Cotyledons	4.3%
Theobroma cacao.	Cacao	Cotyledons	2.5%
Citrus maxima	Pomelo, Chinese grapefruit	Anthers	0.9%
Ilex paraguariensis	Yerba mate	Leaves	0.9%

*Source: Ashihara, H., Sano, H., & Crozier, A. (2008) Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochemistry* **69**(4): 841-856.

Caffeine's natural role in pest and weed control can be potentially applied in agriculture, horticulture and gardening [Nathanson *et al*., 1984]. In fact, this application of caffeine has already gained scientific attention in recent years with the generation of several transgenic plants [Kim *et al*., 2006; Ashihara *et al*., 2008]. On the other hand, during the course of evolution, microorganisms that thrive in caffeine-enriched soil in the vicinity of caffeine producing plants have also developed tolerance towards this toxic molecule. Numerous studies have reported characterization of several soil microbes, mostly *Pseudomonas*, which are capable of degrading caffeine [Yu *et al*., 2008; Yu *et al*., 2009; Gopishetty *et al*., 2011; Summers *et al*., 2012; Mohanty *et al*., 2012]. Given the extensive consumption of caffeine by human [Gilbert *et al*., 1984; Barone *et al*., 1996], caffeine-metabolizing enzymes are present in humans and higher organisms as well. Hepatic cytochrome P450s (CYP 1A and 2E1) are involved in caffeine degradation and detoxification [Arnaud *et al*., 2011].

Significance of Caffeine and Other Purine Alkaloids

Caffeine containing products are listed in Table D1. Thus, it is not surprising that caffeine is "the most widely used mood-altering drug in the world" [Juliano *et al*., 2005]. Caffeine and other purine alkaloids (Fig. 1.1), often referred to as xanthine derivatives, are also active ingredients of pharmaceutical preparations. Caffeine is often used as a neurological, cardiac, and respiratory stimulant [Gokulakrishnan *et al*., 2005; Dash *et al*., 2006], as well as an analgesic enhancer in cold, cough, headache, and asthma medicines [Daly *et al*., 2007; Stavric *et al*., 1988b]. Theophylline, a potent bronchodilator, is used to treat neonatal apnea, control asthma, and relieve bronchial spasms [Mazzafera *et al*.,

2002; Stavric *et al*., 1988a]. Theobromine is used primarily as a diuretic, myocardial stimulant, or vasodilator [Stavric *et al*., 1988c]. 7-Methylxanthine has been shown to improve the quality of sclera collagen, and has been suggested for treatment and/or prevention of axial myopia, glaucoma, and macular degeneration [Trier *et al*., 1999]. Uric acid and methyluric acids are used in obesity pharmaceuticals, skin cosmetics, and anti-dandruff preparations [Madyastha and Sridhar, 1999]. Theacrine (Fig. 1.1) exhibits sedative, analgesic, and anti-inflammatory properties [Wang *et al*., 2010]. Also, 1-methylxanthine, 1-methyluric acid, and uric acid have been investigated for their antioxidant properties [Lee *et al*., 2000].

Purine alkaloids (Fig. 1.1) are also used extensively in cosmetic and beauty care products. A mixture of 15 $g \cdot L^{-1}$ caffeine, 0.07 $g \cdot L^{-1}$ paraxanthine, 1.4 $g \cdot L^{-1}$ theobromine, and 0.6 $g \cdot L^{-1}$ 7-methylxanthine known as Sveltam™ is currently being marketed as a natural skincare and slimming agent [Libragen *et al*., 2011]. This methylxanthine blend inhibits glyceraldehyde-3-phosphate dehydrogenase activity and accumulation of lipids while stimulating production of cAMP, glycerol, and free fatty acids from adipocytes. Synthetic xanthine derivatives are also finding increasing application in pharmaceutical and beauty care products. For example, Doxofylline (7-(1,3-dioxolan-2-ylmethyl)-1,3-dimethylxanthine, market name Ansimar™), Enprophylline (3-propylxanthine) and Pentoxifylline (1-(5-oxohexyl)-3,7-dimethylxanthine) are used for bronchodilation and treatment of peripheral and cerebrovascular disease, respectively [Sankar *et al*., 2008; Ward *et al*., 1987]. Other xanthine derivatives and 8-oxo derivatives of these synthetic xanthine derivatives have been used as diuretics, pulmonary and cardiac stimulants

[Zavialov *et al*., 2004], antioxidants and anti-inflammatory [Bhat and Madyastha, 2001],
and male erectile dysfunction treatment [Wang *et al*., 2002].

Given this extensive use of caffeine, it is not surprising that it is detected in the
environment. It is regarded as toxic to some aquatic organisms (like African clawed
frog) when present in unnaturally high levels (>270 ng/L or 270 ppb) in water bodies
which serve as a sink to waste-water from coffee processing units or urban areas [Bruton
et al., 2010]. Caffeine is also regarded as an anthropogenic marker of surface and ground
water contamination due to urbanization [Seiler *et al*., 1999; Buerge *et al*., 2003;
Daneshvar *et al*., 2012]. Thus, caffeine has a role in human life and the environment in
the form of food, feed, drug, and pollutant. Hence, it should not come as a surprise that
caffeine has gained huge public attention in recent years and several books have been
completely dedicated only to this compound [Feldman, 1976; Weinberg, 2001, Juliano,
2005; Spiller, 2010, Marin, 2011]. The prominent presence of caffeine and other purine
alkaloids has benefited microbes that have evolved to live on this molecule. Below is a
review of key pathways involved in biosynthesis and degradation of caffeine and other
purine alkaloids.

Caffeine Metabolism: Biosynthetic and Biodegradation Pathways

Biosynthesis of Caffeine

Caffeine and other purine alkaloids are secondary metabolites and thus, they are
synthesized from corresponding primary metabolites like purine nucleotides which are
either produced by de novo biosynthesis or by salvage pathways [Ashihara and Crozier,
1999]. Purine nucleotides are the precursors of nucleic acids, energy molecules such as

ATP and GTP, components of major co-enzymes, activated intermediates in a number of biosynthetic pathways, and metabolic regulators [Stryer, 1995]. Thus, the complete biosynthetic pathway of caffeine can be separated into two segments: biosynthesis of purine nucleotides to form the end product xanthosine (Fig. 1.2A), and biosynthesis of caffeine from xanthosine (Fig. 1.2B).
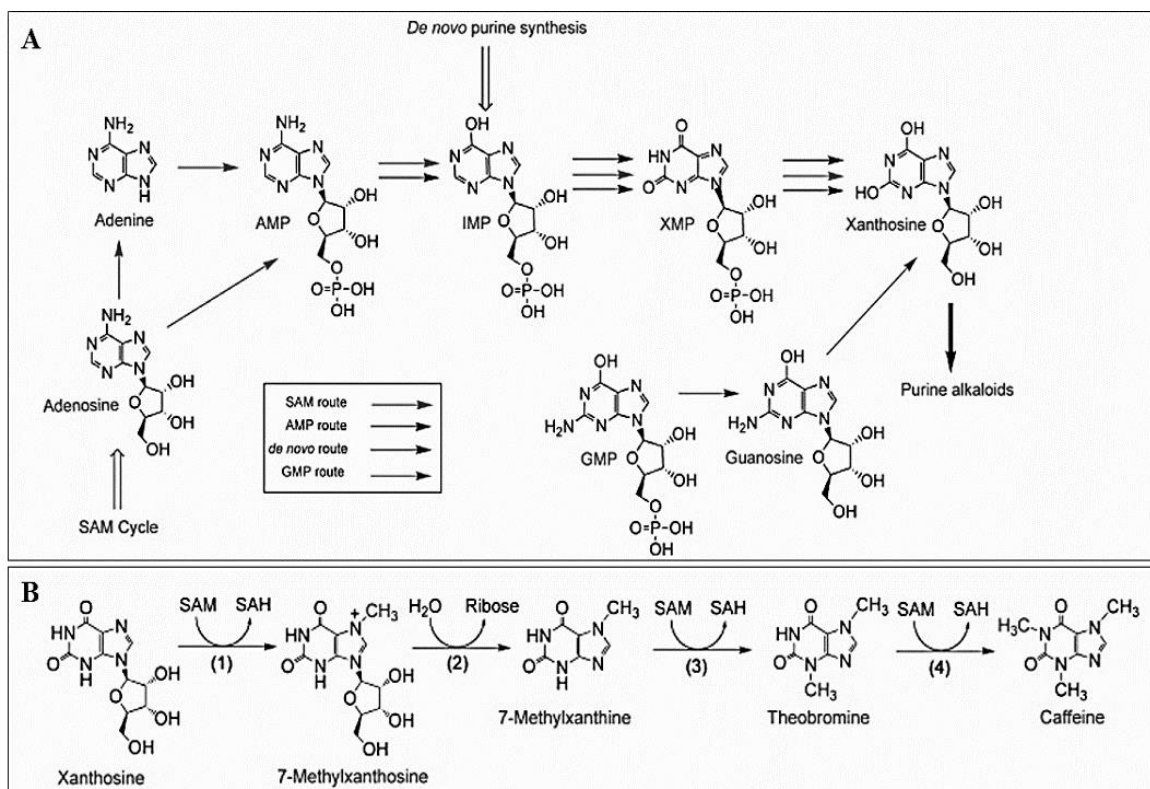
Figure 1.2. Pathways for biosynthesis of caffeine. [A] Xanthosine is produced from four routes: via adenosine from the SAM cycle (SAM route); via IMP from *de novo* purine synthesis (*de novo* route); via adenine (AMP route) and via guanine (GMP route). [B]. Caffeine-biosynthesis from xanthosine. Enzymes 1-4, see text for names and details. SAM, S-adenosyl-l-methionine; SAH, S-adenosyl-l-homocysteine.

Source: Figure adopted from Ashihara, H., Sano, H., & Crozier, A. (2008) Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochemistry* **69**(4): 841-856.

In the first segment (Fig 1.2A), xanthosine, the precursor of caffeine and other purine alkaloid, is supplied by at least four different pathways:

i. _De novo_ purine biosynthetic route: Purine nucleotides are synthesized _de novo_ from non-purine precursors by assimilating parts of several small molecules, namely CO_2, 10-formyltetrahydrofolate, 5-phosphoribosyl-1-pyrophosphate, and amino acids like glycine, glutamine, and aspartate [Ashihara _et al_., 2008]. The final product of this route, inosine 5'-monophasphate (IMP) is synthesized from 5-phosphoribosyl-1-pyrophosphate molecule by assimilation of these small molecules at specific sites: CO_2 (C-6), 10-formyltetrahydrofolate (C-2 and C-8), glycine (C-4, C-5, and _N_-7), aspartate (_N_-1), and glutamine (_N_-3 and _N_-9) [Stryer, 1995]. IMP is then converted to xanthosine via xanthosine monophosphate (XMP) by IMP dehydrogenase (EC1.1.1.205) and 5'-nucleotidase (EC 3.1.3.5) [Ashihara _et al_., 2008].

ii. Adenosine monophosphate (AMP) route: A portion of xanthosine via IMP is also produced from AMP. In this route, AMP, which is produced from previously mentioned _de novo_ or salvation routes, is converted to IMP by an AMP deaminase (EC 3.5.4.6), which is then converted to xanthosine via XMP [Koshiishi _et al_., 2001; Ashihara _et al_., 2008].

iii. _S_-adenosyl-L-methionine (SAM) cycle route: A portion of AMP is also produced from the SAM cycle. In this route, adenosine, produced _via_ SAM cycle, either gets directly converted to AMP by 5'- phosphorylation or adenine, which then gets converted to AMP (Fig. 1.2A). AMP

produced via this route follows the same AMP $\rightarrow$ IMP $\rightarrow$ XMP $\rightarrow$ xanthosine route.

iv.   Guanosine monophosphate (GMP) route: Xanthosine is also produced from GMP via a pathway independent of IMP (Fig. 1.2A). In this route, GMP produced in the cell gets converted to guanosine by 5'-nucleotidase followed by guanosine deaminase (EC 3.5.4.15) activity to form xanthosine [Ashihara *et al*., 2008]. The first committed step in biosynthesis of caffeine and other methylxanthines is the conversion of AMP, IMP, XMP, and GMP to xanthosine.

The second segment (Fig. 1.2B), which is the four-step conversion of xanthosine to caffeine, consists of three methylation steps (steps 1, 3 and 4 catalyzed by SAM-dependent *N*-methyltransferases) and one nucleosidase reaction. In step-1 of this pathway (Fig. 1.2A), xanthosine is converted to 7-methylxanthosine by 7-methylxanthosine synthase (xanthosine 7*N*-methyltransferase, EC 2.1.1.158). In the next step, 7-methylxanthosine is hydrolyzed by a nucelosidase to form 7-methylxantine. The last two steps of this pathway are by caffeine synthase (EC 2.1.1.160), which converts 7-methylxanthine to caffeine via theobromine (Fig. 1.2B) [Ashihara *et al*., 2008].

Biodegradation of Caffeine in Plants and Higher Organisms

In plants, caffeine is produced in actively growing parts like leaves, buds, fruits, etc., and simultaneously but slowly gets degraded by two known pathways, namely, *N*-demethylation and C-8 oxidation [Ashihara *et al*., 2008]. The *N*-demthylation is the

predominant pathway for caffeine degradation (Fig. 1.3A); however, C-8 oxidation is

observed in kucha and certain species of *Coffea* (Fig. 1.3B).
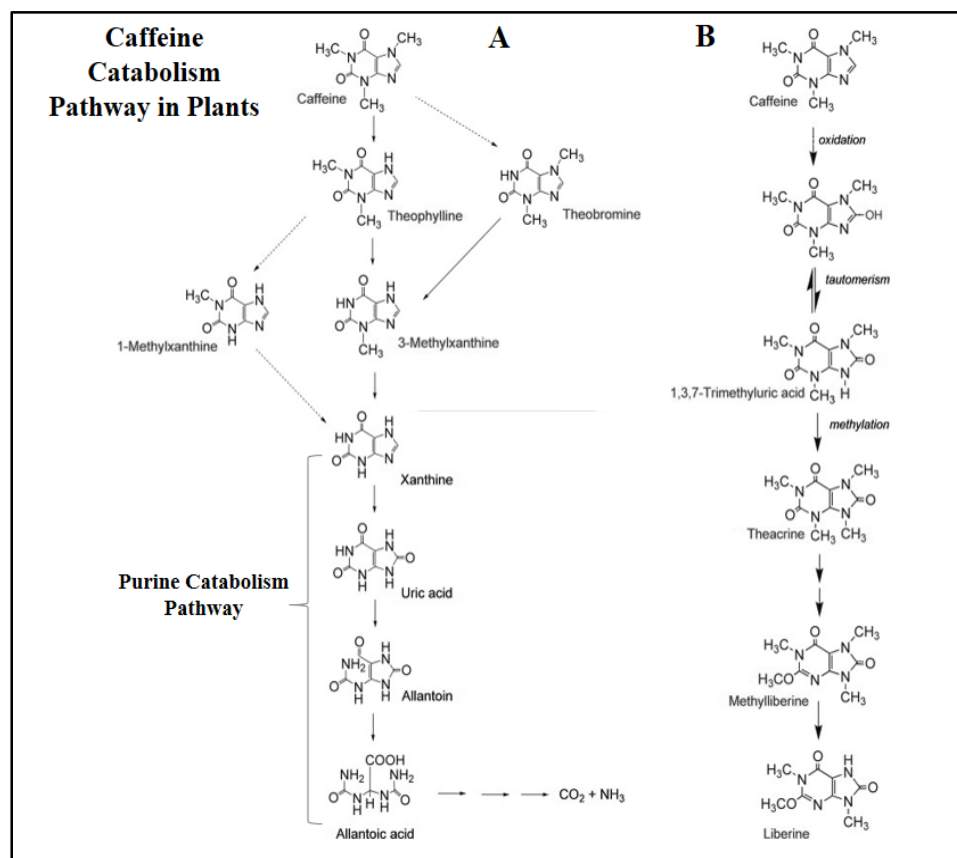
Figure 1.3. Caffeine catabolic pathways. [A] Caffeine is mainly catabolized to xanthine
via theophylline and 3-methylxanthine. Xanthine is further degraded to CO_2
and NH_3 by the conventional oxidative purine catabolic pathway. Dotted
arrows represent minor routes. [B] Alternative pathway for conversion of
caffeine to methyluric acids in kucha and some *Coffea* species.

Source: Figure adopted from Ashihara, H., Sano, H., & Crozier, A. (2008) Caffeine and
related purine alkaloids: Biosynthesis, catabolism, function and genetic
engineering. *Phytochemistry* **69**(4): 841-856.

In the *N*-demethylation pathway, caffeine is first demethylated at the *N*-7 position

to form theophylline. This is the rate-limiting step in degradation of caffeine, which

causes it to accumulate in plant organs. Theophylline then undergoes two sequential *N*-demethylation reactions at *N*-1 and *N*-3 positions to produce xanthine via 3 methylxanthine (Fig. 1.3A). Xanthine is further oxidized to uric acid and enters the conventional purine catabolism pathway to produce allantoin and allantoic acid, and ultimately to CO_2 and NH_3. Alternately, caffeine in some species of kucha and certain species of *Coffea* undergoes a direct oxidation at the C-8 position to produce 1,3,7-trimethyluric acid (TMU), which undergoes methylation at *N*-9 position to produce theacrine (Fig. 1.3B). Subsequently, theacrine gets converted to liberine via methylliberine [Petermann *et al*., 1983].

While autotrophic organisms like plants are capable of utilizing products derived from caffeine (purine, glyoxylate and ammonia), birds, mammals, primates, humans, and many reptiles degrade purines only partially and excrete urate or allantoin as end product [Vogels *et al*., 1976]. Humans metabolize caffeine by hepatic cytochrome P450 enzymes including CYP1A1, CYP1A2, CYP2E1 and CYP 2A6 [Campbell *et al*., 1987; Berthou *et al*., 1991; Tassaneeyakul *et al*., 1994]. These enzymes are known to catalyze both *N*-demethylation and oxidation of caffeine and its metabolites to produce more than 25 different metabolites including paraxanthine (80%), theobromine (10%), theophylline (4%) and others (6%) [Carrillo *et al*., 2000]. Thus, in humans, caffeine degradation occurs primarily through two *N*-demethylation steps (*N*-3 and *N*-7), to produce paraxanthine and 1-methylxanthine, respectively [Lelo *et al*., 1986]. These methylxanthines are further oxidized to form corresponding methyluric acids, with subsequent excretion into urine.

Biodegradation of Caffeine in Microbes

Caffeine is a ubiquitous natural product and enters the environment via various caffeine-containing plant parts/products. This means not only animals, but also microbes are exposed to this compound, and must have developed biochemical processes to detoxify or degrade or even consume it as a source of carbon and nitrogen. However, unlike higher organisms, caffeine metabolism in microbes did not receive much attention until the 1970s [Gopishetty *et al*., 2011]. Within the microbial kingdom, although biochemistry and enzymology of caffeine degradation in bacterial systems is well studied [Dash *et al*., 2006; Yu *et al*., 2008, Yu *et al*., 2009; Gopishetty *et al*., 2011; Summers *et al*., 2011; Summers *et al*., 2012], little is known about caffeine utilization by yeast and fungi. Caffeine metabolism in fungi like *Aspergillus niger* [Ina *et al*., 1971] and *Penicillium roqueforti* [Schwimmer *et al*., 1971] occurs *via* a similar *N*-demethylation pathway as found in higher organisms, where *N*-7 demethylation of caffeine yields theophylline [Hakil *et al*., 1998]. In yeast, caffeine is degraded in a similar way as in humans, via P450 enzymes that *N*-demethylate caffeine [Sauer *et al*., 1982]. The enzymology of fungal metabolism of caffeine lacks detailed characterization and is wide open for further investigation.

Caffeine degradation is well studied in bacterial systems, primarily from the genus *Pseudomonas*, and others like *Serratia*, *Klebsiella* and *Rhodococcus* species. Unlike higher organisms, bacteria utilize caffeine as a sole source of carbon, nitrogen and energy for growth [Woolfolk *et al*., 1975; Mazzafera *et al*., 1996; Madyastha *et al*., 1999; Dash *et al*., 2006; Yu *et al*., 2008; Yu *et al*., 2009; Summers *et al*., 2012; Mohanty *et al*., 2012]. Bacteria initiate caffeine degradation in either of the two pathways, namely *N*-

demethylation (plucking each of the three methyl groups from caffeine to form xanthine) and C-8 oxidation (oxidizing caffeine at C-8 position to form TMU). Xanthine formed by *N*-demethylation of caffeine is then oxidized to uric acid. Uric acid or TMU formed via *N*-demethylation or C-8 oxidation undergoes ring-opening reactions similar to the well-known purine degradative pathways [Vogels *et al*., 1976]. Formaldehyde formed from *N*-demethylation of caffeine is further oxidized to CO_2 via formaldehyde and formate dehydrogenases [Summers *et al*., 2011].

Bacterial Caffeine Degradation via *N*-demethylation Pathway

In the *N*-demethylation pathway (Fig. 1.4), caffeine is first demethylated predominantly at the *N*-1 position to produce theobromine (major product). *N*-3 is also demethylated resulting in paraxanthine (minor product) (Fig. 1.4A) [Yu *et al*., 2009]. These dimethylxanthines are further demethylated to 7-methylxanthine and then to xanthine [Yu *et al*., 2009]. Theophylline, a major caffeine metabolite in plants and fungi, has not been reported as a product of bacterial caffeine degradation [Yu *et al*., 2009]. Yu *et al*. [2009] showed that a caffeine degrading soil bacterium *Pseudomonas putida* CBB5 was able to consume a wide variety of methylxanthines including theophylline. In CBB5, theophylline was *N*-demethylated to 3-methylxantine or 1-methylxanthine and then to xanthine (Fig. 1.4B) [Yu *et al*., 2009]. Xanthine then enters the well-characterized xanthine degradation pathway [Vogels *et al*., 1976; Werner *et al*., 2009] for complete mineralization.
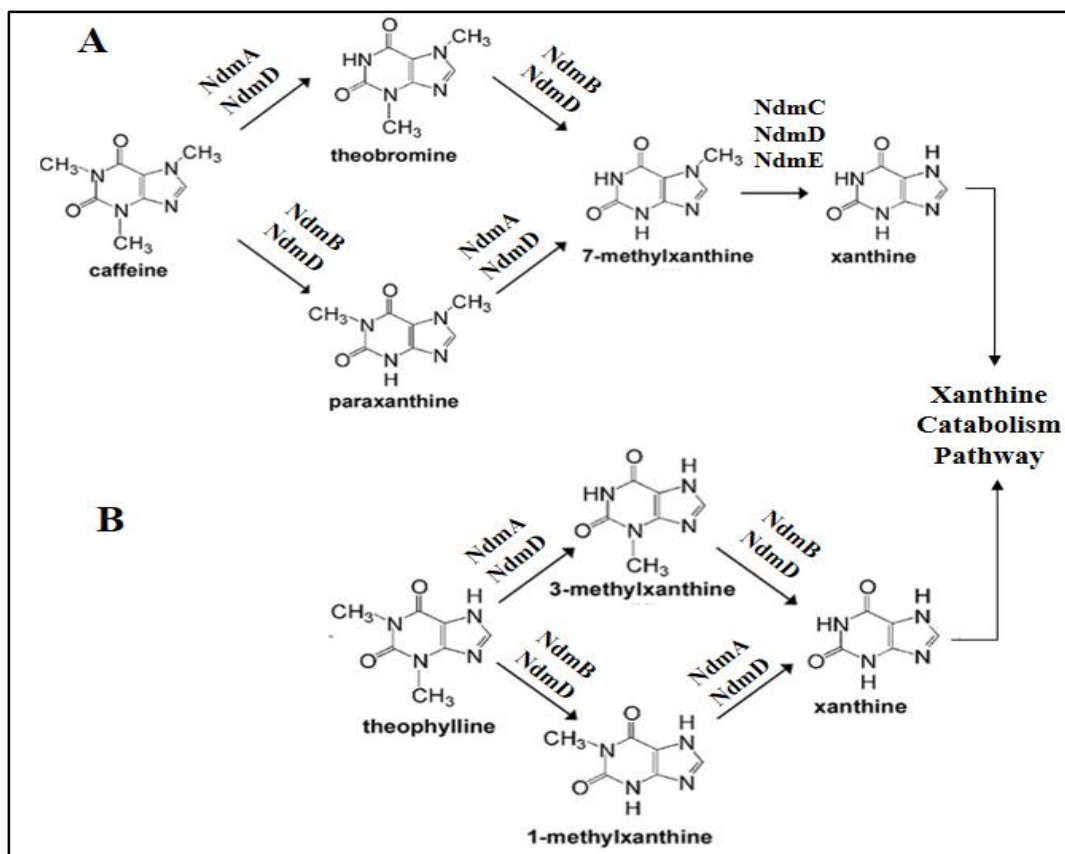
Figure 1.4. Caffeine degradation via *N*-demethylation. Caffeine [A] and theophylline
[B] are degraded via two separate pathways in *Pseudomonas putida* CBB5.

Source: Summers, R. M., Louie, T. M., Yu, C. L., Gakhar, L., Louie, K. C., &
Subramanian, M. (2012). Novel, highly specific N-demethylases enable bacteria
to live on caffeine and related purine alkaloids. *Journal of Bacteriology*. **194**:
2041-49.

Enzymology of *N*-demethylation of caffeine in bacteria was not known until

recently when two separate caffeine *N*-demethylase enzymes, namely NdmA and NdmB

from *Pseudomonas* CBB5 were isolated and characterized [Summers *et al*., 2012]. The

first report showed that a single broad-specific demethylase converted caffeine to

xanthine [Woolfolk *et al*., 1975; Summers *et al*., 2011]; however, due to the inability to

resolve the demthylase further, the question of distinct monooxygenases catalyzing

position-specific *N*-demethylations was left open [Summers *et al*., 2011; Asano *et al*., 1994; Dash *et al*., 2007]. Successful cloning of the *N*-demethylases resulted in characterization of two highly specific novel *N*-demethylases (NdmA and NdmB) from *Pseudomonas putida* CBB5 [Summers *et al*., 2012]. These enzymes were characterized as Rieske non-heme iron monooxygenases that catalyzed *N*-1 and *N*-3-specific N-demethylation of caffeine and related di- and monomethylxanthines, respectively, to produce 7-methylxanthine (Fig. 1.4). It was also reported that both the enzymes (NdmA and NdmB) were dependent on a third redox-center dense Rieske reductase called NdmD, for electron transfer from NADH. While, the NdmA-NdmD pair catalyzed *N*-1 demethylation of caffeine, theophylline, paraxanthine, and 1-methylxanthine to theobromine, 3-methylxanthine, 7-methylxanthine, and xanthine, respectively, NdmB-NdmD pair catalyzed *N*-3 demethylation of theobromine, 3-methylxanthine, caffeine, and theophylline to 7-methylxanthine, xanthine, paraxanthine, and 1-methylxanthine, respectively [Summers *et al*., 2012]. A third N-demethylase, NdmC, was also partially characterized that specifically catalyzed *N*-7 demethylation of 7-methylxantine to xanthine. NdmC was not active on caffeine to produce theophylline; this explained the distinct degradation pathways for the two compounds in CBB5 [Summers *et al*., 2012; Summers *et al*., 2013].

Bacterial Caffeine Degradation via C-8 Oxidation Pathway

In contrast to the *N*-demethylation pathway, there are very few reports in the literature for the C-8 oxidation pathway. Previously, metabolic studies with animals have reported the presence of this pathway in higher organisms where caffeine is oxidized to

TMU (Fig. 1.5) by hepatic cytochrome P450s. Unlike bacteria, higher organisms do not

have the capability to completely degrade caffeine and thus, excrete the intermediates

like TMU, 3,6,8-trimethylallantoin (TMA) [Arnaud *et al*., 1986] and allantoin. In 1998,

Madyastha *et al*. reported for the first time the presence of a caffeine C-8 oxidation

pathway and some related metabolites (TMU, TMA, glyoxylate, urea) in a mixed culture

consortium of *Klebsiella* and *Rhodocuccus* sp. [Madyastha *et al*., 1998]. This study did

not elaborate which culture did which reaction or the enzymes involved in the pathway.
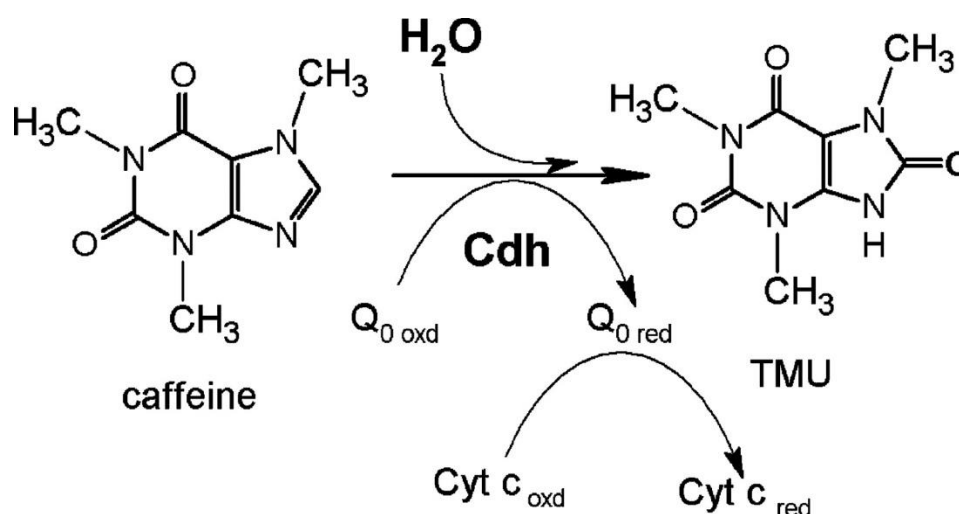
Figure 1.5. Caffeine to 1,3,7-trimethyluric acid (TMU) by caffeine dehydrogenase (Cdh)
in CBB1.

Source: Yu, C. L., Kale, Y., Gopishetty, S., Louie, T. M., & Subramanian, M. (2008). A
novel caffeine dehydrogenase in Pseudomonas sp. strain CBB1 oxidizes caffeine
to trimethyluric acid. *Journal of Bacteriology* **190**(2): 772-776.

Although caffeine is trimethyl xanthine, it is not a substrate for xanthine oxidase

or xanthine dehydrogenase [Yu *et al*., 2008]. Recently, several groups reported enzymes

involved in C-8 oxidation of caffeine; these are summarized in Table 1.2.

Table 1.2. Comparison of different caffeine and xanthine-oxidizing enzymes*.

Properties	Caffeine Oxidase	Caffeine Oxidase	Caffeine Dehydrogenase	Xanthine Oxidase/ Dehydrogenase	Aldehyde Oxidase
Organism	*Klebsiella* sp. *Rhodococcus* sp.	*Alcaligenes* sp.	*Pseudomonas* sp. CBB1	*Pseudomonas putida* L CBB5	*Thermomicr obium roseum*
M.W. (kDa)	85	65	158	290	132
Subunit M.W. (kDa)	85	65	α: 90 β: 34 γ: 20	α: 85 β: 40 γ: 20	α: 88 β: 39 γ: 18
Subunit Structure	Monomer	Monomer	Hetero Monomer, $\alpha\beta\gamma$	Hetero Dimer, $\alpha_2\beta_2\gamma_2$	Hetero Monomer, $\alpha\beta\gamma$
Electron Acceptors	Cytocrome. *c*, O_2	DCPIP, O_2	Coenzyme Q_0	O_2, NAD(P)	O_2
Substrate	Caffeine	Caffeine	Caffeine	Xanthine	Aldehydes
Product	TMU	N/A	TMU	Uric acid	Carboxylic acids

*Source: Adapted from Gopishetty, S. R., Louie, T. M., Yu, C. L., & Subramanian, M. V. (2011). Microbial degradation of caffeine, methylxanthines, and its biotechnological applications, p 44–67. *In* Thatoi HN, Mishra BB (ed), Microbial biotechnology methods and applications. Narosa Publishing House Pvt, Ltd, New Delhi, India.

In 1999, Madyastha and colleagues purified a broad substrate specific 85-kDa, flavin-containing caffeine oxidase from the mixed culture consortium, which catalyzed C-8 oxidation of caffeine to TMU [Madyastha *et al*., 1999]. However, stoichiometry of the reaction with respect to TMU and H_2O_2 was not established. Contrary to the nomenclature of the enzyme, oxygen was found to be poor electron acceptor for this enzyme. Further, decrease in the enzyme activity was observed when cytochrome *c* or

dichlorophenol indophenol (DCPIP) was used as electron acceptors in place of O_2. This brings into question the actual reaction catalyzed by this poorly characterized enzyme. Another 65-kDa caffeine oxidase was purified from *Alcaligenes* sp. CF8 [Mohapatra *et al*., 2006] with caffeine as the best substrate, DCPIP as the preferred electron acceptor *in vitro*. Oxygen conversion to H_2O_2 was also implicated with this enzyme. However, formation and stoichiometry of TMU and H_2O_2 from caffeine was not confirmed.

Recently, a new 158-kDa heterotrimeric caffeine dehydrogenase (Cdh) was purified from a caffeine-consuming *Pseudomonas* sp. strain CBB1 [Yu *et al*., 2008]. This enzyme consisted of three non-identical subunits, α, β, and γ of molecular masses 90, 40, and 19-kDa, respectively. This enzyme was structurally and catalytically different from the previously known caffeine oxidases, but similar to several bacterial xanthine dehydrogenases (Table 1.2) [Schultz *et al*., 2001; Hille *et al*., 2006]. Cdh oxidized caffeine to TMU stoichiometrically by incorporating oxygen from H_2O (Fig. 1.5). Cdh was not NAD(P)$^+$-dependent; coenzyme Q_0 was the preferred electron acceptor *in vitro*. Although, Cdh was structurally similar to xanthine dehydrogenases, it was highly specific for caffeine with a K_m of 3.7 ± 0.9 μM [Yu *et al*., 2008] with no activity towards structurally related xanthine, monomethylxantines and theophylline. Preference for theobromine as a substrate is 50 times less than caffeine [Yu *et al*., 2008]. However, fate of TMU was not elucidated.

Xanthine Degradation Pathway: A Model to Understand Caffeine C-8 oxidation Pathway

Unlike the *N*-demethylation pathway, caffeine degradation via the C-8 oxidation pathway in microbes is poorly characterized except the first oxidation step [Yu *et al*.,

2008]. Enzymology of C-8 oxidation pathway beyond Cdh, and the fate of TMU are unknown. However, it can be hypothesized from the first committed step of caffeine C-8 oxidation pathway that it follows a path similar to the xanthine degradation pathway (Fig. 1.6) via uric acid. Also, TMA, the trimethyl form of allantoin, is also a suggested metabolite of C-8 oxidation pathway [Madyastha *et al*., 1998]. This makes the xanthine degradation pathway an important reference for the present study.
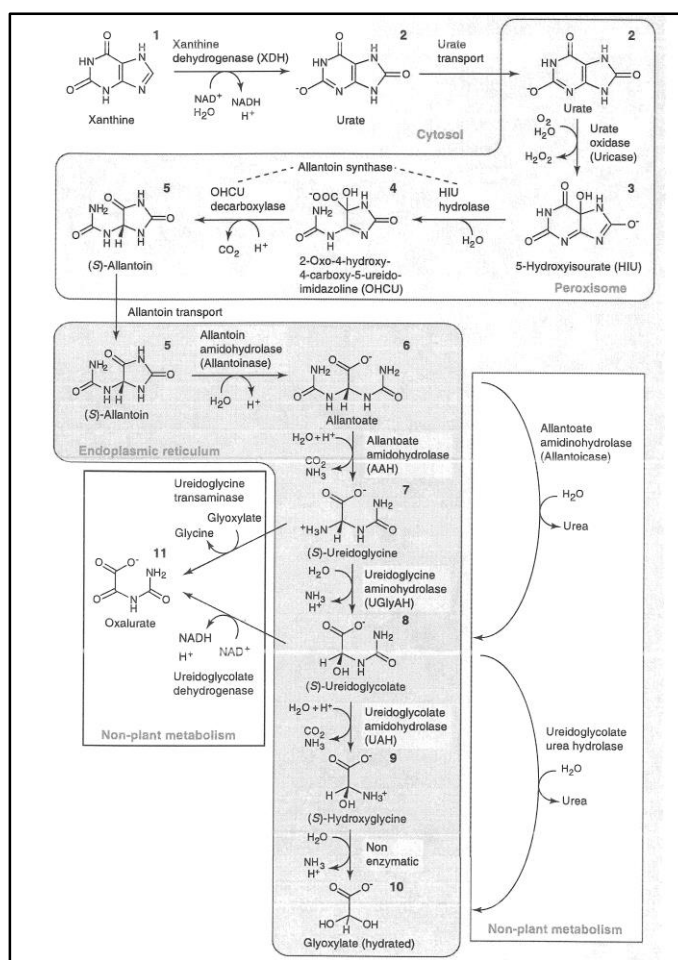
Figure 1.6. Xanthine degradation in plants and microorganisms. Humans and mammals excrete uric acid and allantoin respectively as end products.

Source: Werner A. K. & Witte, C. P. (2011). The biochemistry of nitrogen mobilization: purine ring catabolism. *Trends in Plant Science*. **16**: 381–387.

Xanthine is not fully degraded in animals, whereas plants and microbes completely mineralize it to assimilate all the nitrogen [Vogel *et al*., 1976]. In the typical xanthine degradation pathway (Fig. 1.6), xanthine is oxidized to uric acid by a xanthine dehydrogenase or oxidase. This reaction usually occurs in the cytosol of the plant cell [Werner *et al*., 2001]. Urate is then oxidized to 5 hydroxyisourate (HIU) (Fig. 1.6), which is then hydrolysed to 2-oxo-4-hydroxy-4-carboxy-5-ureido imidazoline (OHCU) by a hydrolase. OHCU gets decarboxylated enzymatically to allantoin; this reaction can also proceed non-enzymatically, but at a slower rate [Ramazzina *et al*., 2006]. These reactions occur in the peroxisomes of plant cells as both the intermediates are highly unstable and may interfere in other cellular processes [Werner *et al*., 2011]. Spontaneous degradation of HIU via OHCU results in racemic allantoin while enzymatic conversion is stereospecific to *S*-(+)-allantoin [Ramazzina *et al*., 2006].

In mammals, allantoin is the terminal metabolite of this pathway and gets excreated. In plants and microorganisms, allantoin is further degraded by the second ring-opening reaction catalyzed by amidohydrolase, to yield allantoate [Bongaerts *et al*., 1976]. This then undergoes further degradation to CO_2 and NH_3 via glyoxylate and urea, by a series of hydrolysis reactions (Fig. 1.6). These reactions occur in the endoplasmic reticulum of plant cells without the production of urea [Werner *et al*., 2011]. Alternatively, intermediate metabolites like (*S*)-ureidoglycine and (*S*)-ureidoglycolate are known to yield oxalurate and glycine in microbial systems. Glyoxylate and urea (or glycine and oxalurate) thus produced via this pathway enter the central metabolism of the cell and get further mineralized or get incorporated into anabolic reactions [Werner *et al*.,

2011]. Interestingly, in the caffeine degradation studies by a mixed culture consortium, Madyastha (1998) reported the presence of glyoxylate and urea in the media along with TMU and TMA. This implies TMU might be undergoing a similar set of hydrolytic and ring-opening reactions to yield TMA, glyoxylate and urea.

Microbial Perspectives of Caffeine Degradation

Caffeine is one of the most important psychoactive natural products with an ever growing market and popularity. A thorough understanding of caffeine metabolism including detailed enzymology and genetic analysis, especially in bacterial systems, can lead to several biotechnological applications [Roussos *et al*., 1995 and Dash *et al*., 2006]. A few of these, grouped under their individual market sectors, are briefly discussed here.

Industrial Biotechnology: Production of Chemicals and Pharmaceuticals

Synthesis of caffeine and its derivatives is a multistep process and hence it is expensive [Shamim *et al*., 1989; Strappaghetti *et al*., 2001; Wang *et al*., 2002; Zavialov *et al*., 2004]. In addition, selective alkylation of each nitrogen atom on the xanthine ring is difficult [Gopishetty *et al*., 2011]. In such cases, enzymatic processes are advantageous as they are highly specific. Thus, a thorough understanding of enzymes and genes involved in caffeine *N*-demethylation could lead to novel bioprocesses to produce commercially important xanthine-derivatives. TMA, another important molecule with applications in cosmetics and medicines treating burned skins, is also not commercially available. This molecule could be produced biologically from caffeine by CBB1.

Food Industry: Bio-decaffeination of Coffee and Tea

Decaffeinated products are also growing on an annual basis. Decaffeination of coffee or tea is usually done either by solvent extraction, or by water diffusion, or by supercritical CO_2 extraction [Feldman *et al*., 1976; Udayasankar *et al*., 1986; Coulate *et al*., 2009]. In the solvent extraction method, the moisture content of the sample is increased to about 40% followed by extraction of caffeine with methylene chloride at between 50-120 °C in a counter-current system. Decaffeinated residue is dried by steam stripping [Coulate *et al*., 2009]. In the water diffusion method, water containing 15% solids other than caffeine is used to extract caffeine from the coffee beans. Approximately 98% caffeine extraction is achieved in the decaffeinated products, which are then subjected to air- drying [Berry *et al*., 1943; Feldman *et al*., 1976]. In the third method, supercritical CO_2 selectively extracts caffeine from coffee or tea at elevated pressures (120 to 180 atm) [Udayasankar *et al*., 1986]. Although, this process is comparatively safer, it requires large input of energy and thus, cost of operation is high [Mohapatra *et al*., 2006]. Moreover, such chemical based methods use environmentally undesirable compounds.

An alternative to these chemical methods is "green chemistry" in which either whole cells or enzymes can be used for decaffeination [Gokulakrishan *et al*., 2005; Dash *et al*., 2006; Gopishettey *et al*., 2011]. "Green" methods are usually environmentally friendly and generate less waste. Such methods are sustainable, specific and generally considered safe for pharmaceutical or food production. Such operations are usually under mild conditions. However, the economics of such biological processes are unknown. Caffeine degrading organisms and enzymes could find potential use in bio-

decaffeination. It should be possible to use transgenic *Saccharomyces cerevisiae*, considered a safe industrial organism in bio-decaffeination process development [Gopishetty *et al*., 2011].

Health Industry: Sensing Caffeine in Food, Beverages and Body Fluids

Owing to its possible toxic effects, caffeine is highly restricted to certain sections of the population, such as pregnant women, nursing mothers, infants, elderly people and kids. As a result, various decaffeinated versions of caffeine containing products have gained popularity. Caffeinated and decaffeinated drinks are not differentiable by look or taste. Hence, there might be a need for a rapid, specific and portable caffeine detection kit for beverages and nursing mother's milk [Gopishetty *et al*., 2011]. The enzymes of caffeine metabolism in CBB1 and CBB5 might serve this purpose.

Environment: Bioremediation and Waste-Water Treatment

Given the extensive use of caffeine from beverages to pharmaceuticals, the compound ends up in soil via cultivation of crops, manufacturing plants and human activities. Caffeine is considered as an anthropogenic waste and marker due to its heavy consumption, especially in urban areas. The waste pulp from global consumption of coffee (139 million bags of 60 kg each or 120,000 tonnes of caffeine/year) still contains on an average, 1% caffeine, which makes them unsuitable for use either as animal feed or as crop fertilizers [Brand *et al*., 2000; Mazzafera *et al*., 2002; Gopishettey *et al*., 2011]. Caffeine is well known for its toxicity towards microorganisms, germinating seeds [Ashihara *et al*., 2008], and aquatic organisms [Bruton *et al*., 2010]. Therefore, care must

be taken to remove caffeine from industrial and human waste products before they can enter into soil and aquatic environment [Ogunseitan *et al*., 1996; Bruton *et al*., 2010]. Bioremediation of these waste pulps and liquid effluents is a potential solution [Ogunseitan *et al*., 1996; Gokulakrishnan *et al*., 2005] to reduce caffeine release into the aquatic environment [Ogunseitan *et al*., 1996]. Caffeine-degrading microbes can find use here as well.

Fundamental Science, Metabolic Engineering and Synthetic Biology

Detailed study of caffeine metabolic pathways in all systems, especially microorganisms, will lead to a thorough understanding of carbon and nitrogen cycles and its contribution to the ecosystem. The discovery of novel enzymes, genes, gene-clusters, regulatory elements, and metabolic engineering of pathways, will have a larger impact on understanding the fate of a number of natural products and xenobiotics [Quandt *et al*., 2013]. For example, metabolic engineering of *E. coli* "addicted" to caffeine resulted in the discovery of a glutathione-*S*-transferase-dependent oxygenase involved in 7-N-demethylation [Quandt *et al*., 2013]. This is the first report of such an oxygenase, which enabled complete enzymological resolution of this pathway.

Previous Research

Caffeine degrading microorganisms were isolated by enrichment from soil samples obtained from the Oakdale campus of the University of Iowa. For enrichment, M9 mineral salts medium [Sambrook *et al*., 1989] containing 2.5 g·L^{-1} caffeine was used as the sole source of carbon and nitrogen [Yu *et al*., 2009]. Cultures were incubated at 29

°C with rotary shaking at 200 rpm. Periodically, aliquots of culture media were removed to monitor growth by measuring the optical density at 600 nm (OD_{600}); caffeine utilization was measured by HPLC. Complete consumption of caffeine ($2.5 \text{ g} \cdot \text{L}^{-1}$) from the medium was observed after about 192 hours of incubation. The initial soil isolate was further enriched by three subcultures in M9-caffeine medium; the third subculture was serially diluted and plated on M9-caffeine agar. The plates were incubated at 29 °C, and morphologically different colonies were isolated and screened for the ability to degrade caffeine.

Four morphologically distinct colonies were obtained from the above procedure, and named as CBB1, CBB3, CBB5, and CES. The isolates were submitted to MIDI Inc. (Newark, DE) and identified as *Pseudomonas* based on 16S rRNA phylogenetic and fatty acid ester analysis (FATE). Phylogenetic analysis identified *P. putida* strain AJ as the closet relative to strain CBB5 (GenBank accession no. AY391278, 99% identity). Fatty acid methyl ester analysis identified CBB5 as *Pseudomonas putida* biotype B with a similarity index of 0.81. The remaining three strains were characterized as *Pseudomonas* sp. CBB1, *Pseudomonas nitroreducens* CBB3, and *Pseudomonas* sp. CES.

Strain CBB1 was shown to convert caffeine to TMU using a novel caffeine dehydrogenase [Yu *et al*., 2008]. Strain CBB5 was found to degrade caffeine via *N*-demethylation pathway [Yu *et al*., 2009]. The mechanism of caffeine degradation in strains CBB3 and CES remains to be elucidated. Strain CBB1 was chosen for the current study.

<u>Objectives</u>

Bacterial degradation of caffeine via C-8 oxidation is poorly characterized. Little is known about the enzymology of this pathway, except few reports of incomplete characterization of caffeine oxidase. Difficulty in studying this pathway beyond caffeine oxidase can be attributed to poor stability of intermediate metabolites and unavailability of standards for structural identification. Caffeine dehydrogenase, a highly caffeine specific enzyme, was isolated from strain CBB1 and characterized [Yu *et al*., 2008]. This enzyme catalyzed the first step of the C-8 oxidation pathway, caffeine to TMU. However, fate of TMU in CBB1 was not known. It was previously known that TMU transiently accumulated in the growth media of CBB1 (unpublished results). This prompted us to further investigate the C-8 oxidation pathway in CBB1 beyond TMU. Based on the xanthine degradation pathway (Fig. 1.6), it was hypothesized that TMU undergoes a three-step conversion to TMA. The initial enzyme active on TMU was hypothesized to be a hydroxylase, which incorporates an OH-group at 5' position to form TM-HIU. An enzyme similar to HIU hydrolase was proposed to convert this metabolite to TM-OHCU. The last step was hypothesized as conversion of TM-OHCU to TMA by another enzyme similar to OHCU decarboxylase. The following four research objectives were developed:

1. Identify the metabolites of the C-8 oxidation pathway up to TMA.

2. Purify the TMU degrading enzyme from CBB1, clone the gene and express the protein in *E. coli* for detailed characterization.

3. Determine the sequence of the caffeine gene-cluster in CBB1 and perform bioinformatics analysis to elucidate the entire C-8 oxidation pathway.

4. Evaluate Cdh as a suitable candidate to develop a rapid diagnostic test for caffeine and validate the test by use of coffee, soft drinks, mother's milk, and pharmaceuticals.

Under the first task, growth media, resting cell suspensions, and cell extract-assays were used for metabolite analysis. HPLC, LC-MS, and chiral-HPLC were used as the analytical methods to identify metabolites of caffeine down to TMA. NMR was used to determine the structure of TMA. The second task focused on the purification of the TMU degrading enzyme in CBB1, followed by N-terminal sequencing, gene isolation, cloning and expression of his-tagged protein in *E. coli* and characterization of the enzyme. Under the third task, all the genes present on the caffeine gene cluster were sequenced and subjected to bioinformatics and cluster analysis for function assignment. Completion of these tasks resulted in delineation of the C-8 oxidation pathway up to TMA along with isolation and characterization of a novel TMU-specific enzyme. New genes and metabolites of this pathway were identified in CBB1 and deposited in the GenBank. In the final task, previously purified Cdh enzyme was fully validated for development of a rapid semi-quantitative and qualitative diagnostic test for caffeine-content in coffee, soft drinks, pharmaceuticals, and nursing mother's milk.

CHAPTER 2

MATERIALS AND METHODS

Chemicals

Caffeine, 1,3,7-trimethyluric acid (TMU), 3,7-dimethyluric acid, 1,3-dimethyluric acid, 1-methyl uric acid, uric acid, allantoin, flavin adenine dinucleotide (FAD) disodium salt hydrate, DNase I from bovine pancreas, and glucose oxidase from *Aspergillus niger* Type II were purchased from Sigma-Aldrich (St. Louis, MO). Yeast extract and agar were obtained from Becton, Dickinson and Company (Sparks, MD), and yeast nitrogen base without amino acids and without ammonium sulfate (YNB) was purchased from ForMedium (Norfolk, United Kingdom). High-pressure liquid chromatography (HPLC)-grade methanol chloroform, hexane, and ethanol were obtained from Sigma-Aldrich.

β-Nicotinamide adenine dinucleotide (NADH) reduced form disodium salt, β-Nicotinamide adenine dinucleotide phosphate (NADPH) reduced form disodium salt, isopropyl-β-D-thiogalactopyranoside (IPTG), Tris Base and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) were obtained from RPI Corp. (Mt. Prospect, IL). Agarose was purchased from Invitrogen (Carlsbad, CA). Taq DNA polymerase, restriction enzymes *NdeI*, and *NotI*, 10x Taq polymerase buffer, 10 mM dNTP mixture, NEB buffers, Antarctic Phosphatase enzyme and buffer, 10x BSA, and pUC19 and pET32 plasmids were obtained from New England Biolabs (Ipswich, MA). Pfu polymerase and buffer were purchased from Stratgene (Santa Clara, CA). Ethidium bromide was obtained from Bio-Rad (Hercules, CA). Dithiothreitol (DTT), glacial acetic acid, and ethylenediaminetetraacetic acid (EDTA) were purchased from Fischer

Scientific (Waltham, MA). Nursing mother's milk sample (frozen) was obtained from Mother's Milk Bank of The University of Iowa. All caffeine containing products were purchased from local stores. Unless otherwise stated, all other chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

Bacterial Strains and Plasmids.

Bacterial strains and plasmids used in this study are listed in Table A1. *Escherichia coli* strains DH5α and BL21(DE3) were used for sub-cloning and gene expression studies, respectively. Competent *E. coli* EPI300™-T1R and *E.coli* StrataClone were included in CopyControl™ HTP Fosmid Library Production Kit (Epicentre, Madison, Wisconsin) and StrataClone Blunt PCR Cloning Kit (Agilent Technologies, La Jolla,CA) were used for strain CBB1 fosmid library construction and cloning of PCR products, respectively.

Media and Growth Conditions

Growth Conditions of Strain CBB1 Cells

Previously, strain CBB1 was isolated from soil samples by enriching the M9 mineral salts growth medium [Sambrook *et al.*, 1989] with 2.5 g caffeine L^{-1} as the sole source of carbon and nitrogen. Once a pure culture of strain CBB1 was obtained, further cell growth for caffeine degradation by resting cell assays and enzyme purification experiments were carried out using M9 mineral salts medium containing 2.5 g L^{-1} caffeine and 4 g L^{-1} YNB. A working volume of 20% was used for growth of cultures in Erlenmeyer Flasks. Cultures were incubated at 30°C with rotary shaking at 200 r.p.m. Cell-growth was monitored by measuring the optical density at 600 nm (OD_{600}).

Growth of *Escherichia coli* Cells

Escherichia coli strains were grown at 37°C in Luria-Bertani medium (7) or

Terrific Broth medium (26), except where indicated. For growth of recombinant *E. coli*

strains, antibiotics were added to the following final concentrations: ampicillin, 100 µg

mL^{-1} and chloramphenicol, 12.5 µg mL^{-1}. LB media were solidified with 1.5% Bacto

agar (Difco Laboratories). Electrocompetent *Escherichia coli* strains JM109 (Yanisch-

Perron *et al.*, 1985; ATCC, Manassas, VA), BL21(DE3) (Novagen, San Diego, CA),

EPI300^{TM}-T1R (Epicentre), and DH5α (Life Technologies, Gaithersburg, Md.) were

rescued following electroporation (described below) in SOC medium containing 2%

(w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 20 mM glucose, 10

mM MgCl$_2$, and 10 mM MgSO$_4$. Rescued cells were grown in LB medium containing 10

g·L^{-1} tryptone, 5 g·L^{-1} yeast extract, and 5 g·L^{-1} NaCl at 37 °C. LB agar plates were

prepared by adding 14 g·L^{-1} agar. Ampicillin was added to LB broth and LB agar to a

final concentration of 100 µg mL^{-1} (LBA broth and LBA agar, respectively). Blue/white

screening of transformants was carried out on LBA agar containing 80 µg mL^{-1} X-gal and

0.5 mM IPTG (LBAXI agar).

Assays to Track Caffeine Biodegradation and Metabolite Formation

Analysis of Growth-Media

Degradation of caffeine and TMU by strain CBB1 was investigated by adding

these compounds to M9 medium as growth substrates. A seed culture of CBB1 was first

grown as described above to an OD$_{600}$ of 1.5. After reaching the desired OD$_{600}$, the cell

culture was harvested by centrifugation (13,800 × *g* for 10 min at 4 °C) and washed twice with chilled 50 mM KPi buffer (pH 7.5). The cell pellet was then resuspended in M9 medium and inoculated into 50 mL YNB-free or YNB-supplemented M9 medium containing a test compound to an OD_{600} of approximately 0.01-0.02. All cultures were incubated at 29 °C with rotary shaking at 180 rpm. Growth of strain CBB1 on the test compounds was monitored by periodically measuring the OD_{600} of each culture. Concomitant degradation of the test compounds and accumulation of metabolites in the media (spent media) were monitored by thin layer chromatography (TLC), high pressure liquid chromatography (HPLC) and electrospray Ionized mass spectrometry (MS). Identities of the metabolites were established by comparing the metabolites' retention times (R_{ts}), absorption spectra, and electrospray ionization (ESI) mass spectra with those of authentic standards.

Caffeine Biotransformation by CBB1 Resting Cells

Degradation of caffeine and TMU by strain CBB1 was also demonstrated by performing resting cell assays. CBB1 grown in YNB-supplemented M9 mineral salts medium containing caffeine was harvested by centrifugation (13,800 × *g* for 10 min at 4°C) when the OD_{600} reached 1.5 to 2.0. The cells were washed once with 50 mM potassium phosphate (KPi) buffer (pH 7.5) and resuspended in 50 mM KP_i buffer to a final OD_{600} of 6. Initial concentration of substrate (Caffeine or TMU) in the resting cell assays varied from 1 mM to 10 mM depending on the experiment. The resting cell suspensions were incubated at 29°C with shaking at 200 r.p.m. Aliquots were removed from the cell suspensions periodically to monitor the degradation of test compounds and

the formation of metabolites by reverse-phase HPLC and/or chiral-HPLC connected with an ESI-MS as described below in the analytical methods section.

Enzyme Activity Assays

Caffeine-Oxidation Activity Assay

A spectrophotometric activity assay was developed to detect caffeine dehydrigenase activity based on caffeine-specific reduction of Nitroblue tetrazolium (NBT). A typical 1-mL reaction mixture contained an appropriate amount of enzyme (usually 1-10 µg), 0.5 mM caffeine, and 0.5 mM NBT in 50 mM KPi buffer (pH 7.5). Caffeine dehydrogenase activity was determined by monitoring with a UV-Visible spectrophotometer (Shimadzu UV-2450), the increase in absorbance at 566 nm (extinction coefficient of 15,500 M^{-1} cm^{-1}) due to formazan production. One unit of enzyme activity was defined as reduction of 1 nmol NBT per min under the defined reaction conditions (25°C and pH 7.5). The activity staining procedure and the spectrophotometric assay facilitated monitoring the purification of the caffeine dehydrogenase.

TMU-Oxidation Activity Assay

A typical TMU oxidation activity assay was carried out at 30°C in 1-mL reaction mixture containing 0.5mM NADH, 0.5mM TMU, and appropriate amounts of crude cell extract, fractions from FPLC, or purified protein, in 50 mM KPi buffer (pH 7.5). The reaction was initiated by the addition of TMU. NADH oxidation was monitored at 340 nm using a UV-Visible spectrophotometer (Shimadzu UV-2450). An extinction

coefficient of 6,200 M^{-1} cm^{-1} ($\varepsilon^{340\,nm}$ NADH = 6.2 mL $\mu mol^{-1}.cm^{-1}$) for reduced minus oxidized NADH was used in calculating the enzyme activity. To measure NADH consumption using a Tecan 96 well micro plate reader, 50 µL of master mix (0.2 mM TMU, and 0.2 mM NADH) was added to 100 µL of enzyme fraction collected from FPLC. One unit of activity was defined as the amount of protein required to oxidize 1 µmol of NADH per min under defined reaction conditions (30°C and pH 7.5). The TMU oxidation activity was also determined by measuring the disappearance of trimethyluric acid from the reaction mixture by HPLC analysis. Periodically, 100 µL from the reaction mixture was mixed with an equal volume of acetonitrile for quantifying trimethyluric acid.

Protein Purification Procedures

All purification procedures were performed at 4°C using an automated fast protein liquid chromatography system (AKTA Design, Amersham Pharmacia Biotech). Chromatography columns and column resins were from Amersham Biosciences, Piscataway, NJ. Amicon Ultra-15 (30K) ultrafiltration membrane (Millipore, Bedford, MA) was used for concentration and buffer exchange of the protein fractions containing caffeine oxidation or TMU oxidation activity.

Cell Harvesting by Centrifugation

Strain CBB1 grown in M9 mineral salts medium with 2.5 g caffeine L^{-1} and 4 g L^{-1} YNB at 30°C with shaking at 200 r.p.m. were harvested at late log phase by centrifugation (13,800 × g for 10 min at 4°C) as described by Yu *et al.,* 2008. Cell

pellets were washed twice with 50 mM potassium phosphate (KPi) buffer (pH 7.5) and

frozen at -80°C prior to lysis and enzyme purification [Scopes, 1994].

Cell Disruption and Cell Extract Preparation

About 8 g frozen cells were thawed and suspended in 30 mL 50mM KPi buffer

(pH 7.5) and DNase I was added to a final concentration of 10 µg mL^{-1}. The cell

suspension was broken by passing through a chilled French Press twice at 138 MPa.

Unbroken cells and debris were removed from the lysate by centrifugation (20,400 × g

for 20 min at 4°C) and the supernatant was designated the crude cell extract (CCE).

Partial Purification of Caffeine Dehydrogenase from CBB1

A 4.0 M solution of ammonium sulfate was added drop-wise to CCE to obtain a

final ammonium sulfate concentration of 0.8 M with constant stirring. After incubation

with shaking on ice for 1 h, the mixture was centrifuged at 20,400 × g for 20 min at 4°C

to remove precipitated proteins. The supernatant was loaded onto a 40-mL (bed volume)

phenyl Sepharose high-performance column (Amersham) preequilibrated with 50 mM

KPi buffer (pH 7.5) containing 0.8 M ammonium sulfate. Unbound proteins were

washed from the column with 40 mL 50 mM KPi buffer containing 0.8 M ammonium

sulfate. Bound proteins were eluted with a 200-mL reverse gradient of ammonium

sulfate (0.8 to 0 M in KD buffer) at a rate of 1 mL min^{-1}. The column was then washed

with 40 mL 50 mM KPi buffer (pH 7.5), followed by 40 mL water at the same flow rate

to elute hydrophobic proteins tightly bound to the column. Fractions containing caffeine

dehydrogenase were identified by the spectrophotometric enzyme activity assay as

described above. Caffeine dehydrogenase was eluted from the column during the final wash step with water.

Purification of TMU Monooxygenase from CBB1

The crude cell extract (CCE) obtained by the method described above from CBB1 was loaded onto a 100-mL (bed volume) Q-Sepharose column (GE Healthcare) pre-equilibrated with 50 mM KPi buffer (pH 7.5). After washing the unbound proteins with the same buffer at 1 mL min^{-1}, bound proteins were eluted with a linear gradient from 0 to 0.5 M KCl in the same buffer and at the same flow rate. Fractions exhibiting TMU oxidation activity eluted at 0.05 to 0.125 M KCl. The active fractions were pooled, concentrated and exchanged into 50 mM KPi buffer (pH 7.5) by ultrafiltration. This concentrated active fraction was loaded onto a 10-mL (bed volume) Blue Sepharose column (Capto Blue, GE healthcare) pre-equilibrated with 50 mM KPi buffer (pH 7.5). After washing unbound proteins from the column with 120 mL of the buffer, bound proteins were eluted with 1 M KCl at a rate of 1 mL min^{-1}. Fractions containing TMU oxidation activity were pooled, concentrated, and exchanged into 50 mM KPi buffer (pH 7.5).

A 4 M ammonium sulfate solution was added to the enzyme fraction from Blue Sepharose to a final concentration of 0.5 M. After 1 hour on ice, the precipitate was removed by centrifugation at 16,000 g for 20 min and the supernatant was loaded onto a 1-mL (bed volume) Phenyl Sepharose High-Performance column (GE healthcare) pre-equilibrated with 50 mM KPi buffer containing 0.5 M ammonium sulfate. Unbound proteins were eluted from the column with the same buffer containing 0.5 M ammonium

sulfate. Bound proteins were eluted with a reverse linear gradient of ammonium sulfate (0.5 M to 0 M in KPi buffer) at a flow rate of 1 mL min^{-1}. Fractions exhibiting TMU oxidation activity were pooled, concentrated and exchanged into 50 mM KPi buffer (pH 7.5) containing 0.1 M NaCl. This preparation was loaded onto an 80-mL (bed volume) Sephacryl S-300 HR column (Amersham) pre-equilibrated with 0.1 M NaCl in 50 mM KPi buffer. Proteins were eluted from the column with the same buffer at a flow rate of 1 mL min^{-1}. Fractions containing TMU oxidation activity were combined, concentrated, and exchanged into 50 mM KPi buffer (pH 7.5). The N-terminal amino acid sequence of the major protein band corresponding to molecular weight of 43 kDa in the purified protein fraction containing TMU oxidation activity, was determined at the Protein Facility, Iowa State University, Ames, Iowa, USA (see PAGE Gel Electrophoresis and protein blot techniques described below for details about sample preparation for N-terminal sequencing of a purified protein).

Gel Electrophoresis Methods for DNA and Protein Analysis

Agarose Gel Electrophoresis for DNA Analysis

Agarose gel electrophoresis was used to fractionate DNA fragments of different sizes and to estimate the DNA concentration in samples. A 0.5x concentration of Tris-Acetate-EDTA (TAE) buffer, pH 8.0, consisting of 20 mM Tris Base, 10 mM glacial acetic acid, and 5 mM EDTA, was used as running buffer during electrophoresis of agarose gels. DNA loading buffer (6 × concentration) was prepared with 1.7 mM bromophenol blue, 6 mM EDTA, 30% (v/v) glycerol, and 60 mM Tris-HCl. Agarose gels were prepared by adding 0.21 or 0.36 g agarose (0.7% or 1.2%, v/v, respectively) to

30 mL 0.5x TAE buffer. After adding of 0.5 μL ethidium bromide, the mixture was heated until the agarose was completely melted (about 45 s in a microwave on HIGH setting) and poured into the gel mold. Gels were allowed to cure at least 30 min at 4°C prior to electrophoresis. After the gel had set and the well comb was removed, 0.5 to 5 μL DNA sample were added to 0.5x TAE plus 1 μL 6x DNA loading buffer to a final volume of 6 μL and loaded into the gel wells. Gels were electrophoresed at 100 V of constant voltage, for 30-40 min in 0.5x TAE buffer and imaged with UV illumination at 302 nm in a Gel Doc XR (BioRad) equipped with a CCD camera and Quantity One software (BioRad).

<div align="center">Measuring DNA Concentration by Electrophoresis Method</div>

Concentration of DNA was estimated by comparing the intensity of DNA bands in an agarose gel under UV illumination as described above. A 1-kb DNA ladder stock (Invitrogen) was diluted in 0.5x TAE buffer and 6x DNA loading buffer (100:167:733 v/v/v) so that the 1.6-kb band was at a concentration of 10 ng μL^{-1}. Both the ladder (1-5 μL) and an appropriate amount of DNA sample were loaded onto an agarose gel and electrophoresed as described above. The concentration of DNA was then estimated by comparing the intensity of the DNA band with that of the 1.6-kb band in the DNA ladder and calculated according to following equation (Equation. 2.1)

$$[\text{DNA}] = \left(\frac{V_S}{V_L}\right) \times 10 \left(\frac{\text{ng}}{\mu\text{l}}\right) \tag{2.1}$$

where [DNA] is the concentration of DNA (ng μl^{-1}) in the sample and V_S and V_L are the volume of sample and ladder, respectively, which contain bands displaying similar intensity under UV imaging.

PAGE Gel Electrophoresis for Protein Analysis

Purity of the protein in the active fractions from chromatographic steps were determined using native and denaturing PAGE. For native PAGE analysis, fractions were mixed 1:2 (v/v) with Native Sample Buffer (Bio-Rad), loaded onto 4-15% Tris-HCl gels (Bio-Rad), and electrophoresed for 60 min at a constant 100 V. For denaturing PAGE analysis, fractions were mixed with 4X LDS loading buffer (Invitrogen) and 1 M DTT (70:25:5, v/v) and loaded onto 10% Bis-Tris gels. Gels were subjected to a constant 200 V for 80 min. Upon completion of PAGE, gels were washed three times in 100 mL deionized water and stained with GelCode Blue Safe Protein Stain (ThermoFisher Scientific).

Protein Blot Technique for Protein N-terminal Sequencing

Protein from the gel was transferred to a 0.2 μm Sequi-blot PVDF membrane (Bio-Rad) by electroblotting [Burnette *et al*, 1981] prior to submitting for N-terminal amino acid sequencing. Proteins were separated on 10% Bis-Tris gels as described above. Following SDS-PAGE, the gel was placed in a western blotting assembly next to a PVDF membrane and subjected to a constant 30 V for 1 h in NuPAGE Transfer Buffer (Invitrogen). Following transfer, the membrane was removed, washed twice for 1 min with H_2O, and stained for 30 min with GelCode Blue Safe Protein Stain. The membrane was destained by washing for 15 s in 100% methanol, rinsed 3 times in H_2O, and air dried overnight. Blots were stored at -20°C until they were submitted for N-terminal amino acid sequencing by Edman degradation method [Edman *et al*, 1950], a widely used

technique for sequencing of amino acids in a peptide chain. Since not more than 30 amino acids can be cleaved and identified, the method provides only the N-terminal sequence and not the whole sequence of the protein. Currently this technique is automated and the service is provided by the protein sequencing facility of The Iowa State University.

<div align="center">Characterization of TMU Monooxygenase (TmuM)</div>

<div align="center">Molecular Mass Estimation of TmuM</div>

The molecular weight of the enzyme subunit was estimated by sodium dodecyl sulfate polyacrylamide electrophoresis (SDS-PAGE) on a 10% Bis-Tris gels (Invitrogen, Carlsbad, CA) with Precision Plus Protein^{TM} molecular weight standards (Bio-Rad Laboratories, Hercules, CA). The native molecular weight of TmuM was determined by gel filtration chromatography using an 80-mL (V_c, geometrical volume) Sephacryl S-300 HR column (Amersham) equilibrated with 0.1 M NaCl in 50 mM KPi buffer at 1 mL·min^{-1}. The void volume (V_o) of the column was determined by measuring the elution volume (V_e) of a 1 mg·mL^{-1} solution of blue dextran 2000 (GE Healthcare). The column was calibrated with ferritin (440-kDa), catalase (232-kDa), adolase (158-kDa), and conalbumin (75-kDa). The V_e of each standard protein was measured, from which the respective K_{av} value was calculated according to the following equation (Equation 2.2).

$$K_{av} = \frac{V_e - V_o}{V_c - V_o},\qquad\qquad(2.2)$$

where K_{av} is the partition coefficient related to the size of a molecule, V_e is the elution volume of the protein, V_o is the void volume of the column, and V_c is the geometrical volume of the column. A standard curve of K_{av} values against the

logarithmic molecular masses of the standard proteins was then used to determine the native molecular mass of TmuM.

Determination of Optimum pH and Temperature of TmuM

To determine pH optimum, his-tagged TmuM enzyme activity was measured as described above at various pH values within the range of 3.0 to 6.5 by using 100 mM acetic acid/sodium acetate buffer, and within the range 7.0 to 11.0 by using 100 mM Tris-base buffer. The temperature optimum for his-tagged TmuM activity in 50 mM KPi (pH 7.5) buffer was determined by incubating the enzyme reactions at various temperatures between 25°C and 70°C and monitoring NADH oxidation at 340 nm as described above.

Determination of Kinetic Parameters of the TmuM

Apparent kinetic parameters of TmuM were determined by measuring the initial rate of TMU-dependent NADH oxidation (v_o) in 50 mM KPi buffer (pH 7.5) at 30°C. TmuM substrate specificity was tested by replacing TMU with 3,7-dimethyluric acid, 1,3-dimethyluric acid, 1-methyluric acid or uric acid. The concentrations ([S]) of various methyluric acid substrates used in these experiments were varied from 0.5 μM to 0.5 mM at a fixed concentration of 50 μM of NADH. Substrate was pre-incubated in KPi buffer at 30°C for 5 min before addition of NADH to the reaction mixture. The reaction was initiated by addition of the enzyme (1.5 μg) to the pre-incubated reaction mixture, followed by rapid manual mixing. The activity was then measured as described above by monitoring the absorbance at 340 nm as a function of time. The apparent kinetic

parameters were determined from Michaelis-Menten plots of v_o against $[S]$ fitted with the Michaelis-Menten equation, as expressed in Equation 2.3.

$$v_o = \frac{k_{cat}[E_T][S]}{K_M + [S]},$$ (2.3)

where $[E_T]$ is the concentration of enzyme in the reaction. All the experiments were performed in triplicate and the data were analyzed by using GraphPad Prism 5.04 software (GraphPad software).

Determination of Oxygen Dependency TmuM Catalyzed Reaction

Oxygen consumption by TmuM was determined in a closed reaction vessel equipped with a Clark-type oxygen electrode (Digital model 10, Rank brothers) as described previously [Summers *et al*, 2011]. The electrode was calibrated using glucose oxidase (Sigma) and glucose for consumption of oxygen. Oxygen dependent TMU oxidation was conducted at 30°C in 1-mL air-saturated reaction mixture containing 50 mM KPi buffer (pH 7.5) with 0.1mM TMU, 0.2 mM NADH and 50 µg of TmuM. The reaction mixture was equilibrated for 2 min with constant stirring and then initiated by adding the enzyme. Periodically, a 100 µL aliquot was withdrawn from the reaction and immediately mixed with an equal volume of acetonitrile to stop the enzyme reaction. TMU consumption during the reaction was analyzed by HPLC.

Determination of Flavin Content of TmuM

To determine the bound flavin content of TmuM, the enzyme was heat denatured at 70°C for 30 min and then centrifuged (20,400 *g* for 20 min at 4°C) to remove the protein precipitate. Supernatant was subjected to reverse-phase HPLC analysis with a

Shimadzu LC-20AT HPLC system equipped with a Hypersil BDS C18 column (4.6 by 125 mm) using an isocratic mobile phase methanol-water-acetic acid (20:80:0.5) with a flow rate of 0.5 mL min^{-1}. The detected peak was compared with the retention times of standard FAD and FMN. Stoichiometry of Enzyme:FAD was determined by measuring the absorbance at 450 nm of FAD in 50 mM KPi (pH 7.5) containing 6 M urea as previously described. [O'Leary *et al*, 2009].

Homology Modeling of TmuM

A BLASTP search with TmuM and HpxO [O'Leary *et al*, 2009] as queries against the Protein Data Bank, identified a flavin-containing hydroxylase involved in pyocyanin biosynthesis (PhzS) from *Pseudomonas aeruginosa* [Greenhagen *et al*, 2008] as the best match with 30% sequence identity to both proteins. PhzS belongs to the family of flavin-dependent aromatic hydroxylases of which the best characterized is *p*-hydroxybenzonate hydroxylase (PHBH). Homology models for TmuM and HpxO were generated using the program MODELLER 9.10 [Fiser *et al*, 2003] and PhzS (PDB ID: 3C96) as the template. The homology models (ten for each protein) were superimposed on PhzS and PHBH (PDB ID: 1PBE) using PyMOL 1.5 (Schrödinger, LLC), and the FAD- and substrate-binding cavity were identified. The ligand-bound structure of VOIDOO [Kleywegt *et al*, 1994] was used to calculate the solvent accessible volumes of the active site cavity for each of the models and then averaged for each protein.

Analytical Methods

DNA and Protein Analysis

Purity of DNA in a given sample was determined by calculating the ratio between the absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) obtained using a UV-Visible spectrophotometer (Shimadzu UV-2450). Further, the concentration of DNA in the sample was calculated using the formula [Davis *et al*, 1980] (Equation 2.4).

$$\text{Concentration } (\mu g \text{ mL}^{-1}) = A_{260} \times \text{dilution factor} \times 50. \qquad (2.4)$$

Protein concentration was determined by the Bradford method [Bradford *et al*, 1976] using bovine serum albumin as the standard with a dye reagent purchased from Bio-Rad. The N-terminal amino acid sequence of TmuM was determined by Edman degradation [Edman *et al*, 1950] with a Model 494 Procise protein/peptide sequencer (Perkin Elmer Applied Biosystems, MA) at the Protein Facility, Iowa State University.

Isotopic Oxygen ($^{18}O_2$) Incorporation Experiment

An $^{18}O_2$ (Cambridge Isotope Laboratories, Inc., Andover, MA) incorporation experiment was conducted in a 1 mL (total volume) reaction mixture containing, 0.5 mM TMU, 0.5 mM NADH and 10 μL purified recombinant TMU-oxidizing enzyme. The reaction mixture was transferred into a modified 0.5 L round bottom flask with two side arms and a quartz cuvette fused at the bottom. The enzyme was pipetted into one of the side arms. After evacuating air in the flask, argon was introduced. The enzyme in the side arm was tilted into the reaction mixture. NADH consumption was monitored at 340 nm at 30°C to confirm that oxidation did not occur under anaerobic conditions. $^{18}O_2$ was then introduced into the vessel to initiate the reaction. The enzyme activity was monitored by the decrease in absorbance at 340 nm. After 30 min, an aliquot of the reaction mixture was mixed with an equal volume of acetonitrile to precipitate the

protein. The supernatant was then analyzed by an HPLC system equipped with Shimadzu LCMS-2010EV single-stage quadruple mass analyzer. Parallel control reaction was carried out with $^{16}O_2$ (air) and the product was analyzed in a similar manner.

LC-MS Analysis of Time Course of Product Formation

The time course of the TMU-oxidizing enzyme-catalyzed reaction was conducted with the recombinant enzyme. Aliquots sampled from the reaction mixture were mixed with equal volume of acetonitrile to stop the reaction. Identification and quantification of trimethyluric acid, reaction intermediates, and end-product were conducted with a Shimadzu LC-20AT HPLC system equipped with Shimadzu LCMS-2010EV single-stage quadrupole mass analyzer and a Hypersil BDS C18 column (4.6 by 125 mm). Methanol-water-acetic acid (20:80:0.5, v/v/v) was used as an isocratic mobile phase with a flow rate of 0.5 mL min^{-1}. Substrates and products resolved by the C18 column first passed through the photodiode array detector, during which UV-visible absorption spectra were recorded. Then the molecules were ionized by electrospray ionization (ESI) in positive ion mode. The following compounds corresponding to their $(M+1)^+$ ions were monitored: caffeine, $m/z = 195$; trimethyluric acid (TMU), $m/z = 211$; 1,3,7-trimethyl-5-hydroxyisourate (TM-HIU), $m/z = 227$; 3,6,8- trimethyl-2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (TM-OHCU), $m/z = 245$; 3,6,8-trimethylallantoin (TMA), $m/z = 201$.

Separation and Identification of Trimethylallantoin (TMA)

A 100-mL reaction, containing 0.5 mM TMU, 0.5 mM NADH and 1.5 mg of recombinant enzyme in 50 mM KPi buffer (pH 7.5) was incubated at 30°C for 1 h and

TMU disappearance was monitored by the HPLC method described above. At the end,

the reaction mixture was concentrated to 5 mL using Thermo Savant AES1010

Automatic Environmental Speedvac® System (Thermo Electron, USA) at room

temperature. The reaction mixture was then loaded onto prep-scale Silica Gel 60 F254

plates of 1-mm thinkness (EM Science, Gibbstown, NJ) and developed with chloroform-

methanol, (85:15, v/v) as the solvent system. Upon the development of a TLC plate,

resolved compounds were visualized by exposing the plates to UV-light at 254 nm and

the starting point, solvent front (the level the solvent reached when the plate was removed

from the developing tank), and all the spots observed on the plate were marked using a

graphite pencil. Retention Factor (R_f) of each spot was calculated using the following

equation (Equation 2.5).

$$R_f = \frac{distance\ of\ the\ spot\ from\ the\ starting\ point}{distance\ of\ the\ solvent\ front\ from\ the\ starting\ point} \qquad 2.5$$

where R_f is a number between 0 and 1. TMU ($R_f = 0.59$) and the reaction product

($R_f = 0.51$) were subsequently extracted from preparative scale TLC plate with methanol

and subjected to LC-MS analysis (as described above). Nuclear Magnetic Resonance

(NMR), high-resolution electro-spray ionization mass spectrometry (HRESIMS) and

chiral-HPLC analyses were performed on the reaction product as described below.

Nuclear Magnetic Resonance for TMA Structure Elucidation

^{1}H and ^{13}C NMR spectra of TMA were obtained from a Bruker AMX-600 high

field spectrometer equipped with an IBM Aspect-2000 processor (Bruker Instruments,

Billerica, Massachusetts). Two mg of purified product (TMA) was dissolved in 50 µL

CDCl$_3$ and transferred into a 5-mm high-quality NMR tube (Goss Scientific Instruments, Great Baddow, Essex, UK) to acquire the spectra. Solvent CDCl$_3$ served as an internal standard. All NMR spectra were obtained in chloroform-*d* with chemical shifts expressed in parts per million (δ) and coupling constants (*J*) in Hz. Correlation Spectroscopy (COSY), Heteronuclear Multiple Bond Coherence (HMBC) and Heteronuclear Multiple Quantum Coherence (HMQC) experiments were carried out using the same Bruker AMX-600 high field spectrometer. HRESIMS was done using a Fisons VG-ZAB-HF reversed-geometry (BE configuration, where B is a magnetic sector and E is an electrostatic analyzer) mass spectrometer (VG Analytical, Inc.) at the High Resolution Mass Spectrometry Facility at The University of Iowa.

Chiral HPLC Analysis of Purified Racemic TMA

Separation and identification of the two enantiomers of TMA from the racemic mixture was performed by normal phase Chiral HPLC-ESI-MS analysis. In a typical Chiral HPLC analysis, 2 μL of a sample containing 2 mg/mL preparative TLC purified racemic TMA dissolved in methanol was injected into a Chiralcelpak IA column, Chiral Technologies, Exton, Pa. (250 mm by 4.6 mm i.d., 5 micron) with a Shimadzu LC-20AT HPLC system equipped with Shimadzu LCMS-2010EV single-stage quadrupole mass analyzer. Ethanol-Hexane (50:50, v/v) was used as an isocratic mobile phase with a flow rate of 0.5 mL min^{-1}. Substrates and products resolved by the Chiralpak IA column first passed through the photodiode array detector, during which UV-visible absorption spectra were recorded at 220 nm. Then the molecules were ionized by electrospray

ionization (ESI) in positive ion mode. The following compounds corresponding to their $(M+1)^+$ ions were monitored: 3,6,8-trimethylallantoin (TMA), m/z =201.

Basic Molecular Biology Methods

Purification of DNA from Agarose gel

DNA bands amplified by PCR [Sambrook *et al*, 1989] were purified from agarose gels using the buffers and materials provided in QIAquick Gel Extraction Kit (Qiagen, Duesseldorf, Germany). Briefly, after separating DNA by agarose gel electrophoresis, the section containing the desired band was excised from the gel with a scalpel. The gel slices were dissolved by incubation for 10 min at 50°C in three gel volumes QG buffer. One gel volume isopropanol was added to the dissolved gel and DNA mixture to provide the proper binding conditions. Up to 400 mg agarose was loaded onto a QIAquick spin column and centrifuged at 17,900×g for one min to bind the DNA to the silica membrane in the provided spin columns. Residual agarose was washed from the columns and bound DNA was then washed with 0.5 mL QG buffer. The column was washed with 0.75 mL PE buffer and DNA was eluted from the column with 50 µL warm (55°C) water.

Purification of Amplified DNA from PCR Mix

PCR amplified products were separated from the PCR reaction mixture containing DNA template, primers, polymerase, and buffers with the QIAquick PCR Purification Kit (Qiagen). All buffers and materials used were provided in the kit. Five volumes PBI buffer were added to one volume of the PCR sample, and the PCR product was bound to a silica membrane-containing spin column by centrifugation at 17,900×g for one min.

After washing the column with 0.75 mL PE buffer, purified PCR product was eluted from the column with 50 µL warm (55°C) water.

Purification of Plasmid DNA from *Escherichia coli*

Plasmids in *E. coli* cells were extracted and purified with the GeneJET Plasmid Miniprep Kit (Fermentas, Glen Burnie, MD). All buffers and materials used were provided in the kit. Cells grown in LB medium were centrifuged at $17,900 \times g$ briefly and the supernatant was discarded. After resuspending the pellet in 0.25 mL resuspension solution, cells were lysed by addition of 0.25 mL SDS and alkaline lysis solution. Following lysis, 0.35 mL neutralization solution was added to precipitate cellular components and SDS was added to provide proper binding conditions in the mixture. The resulting precipitate was separated from the mixture by centrifugation at $17,900 \times g$ for ten min, and the plasmid DNA in the supernatant was bound to a silica membrane in the spin columns by centrifugation for one min. Bound DNA was washed twice with 0.5 mL wash solution and then eluted with 50 µL warm (55°C) water.

Polymerase Chain Reaction (PCR) and Cloning Methods

PCR reactions were carried out with an Eppendorf Mastercycler ep gradient S thermocycler (Eppendorf, Hamburg, Germany) in 0.2 mL PCR tubes (RPI). A typical 25 µL PCR reaction mixture contained 0.2 µM forward primer, 0.2 µM reverse primer, 0.2 mM dNTPs, 2.5 µL 10x Taq Buffer, 0.5 µL Taq polymerase, and an appropriate amount of template (2-50 µg DNA or 0.5 µL overnight culture).

A typical PCR program contained about 30-35 cycles, where each cycle consisted of three sequential steps: denaturation, annealing and extension. An initialization or heating step was first performed in order to activate the polymerase (this step also served to lyse cells during colony/culture PCR, as described below). The temperature was then cycled through multiple melting, annealing, and extension times, followed by a final extension time of 5-10 min. Reactions were held at 4°C or ice for short-term storage. Long-term storage of PCR reactions, including storage overnight, was at -20°C. Negative controls of the PCR reactions were performed by excluding the forward primer, reverse primer, or template in order to confirm that the PCR products observed in the reaction were due to specific binding of primers to the template. Following PCR, production of amplified DNA was observed by separating between 0.5 to 5 µL reactions *via* agarose gel electrophoresis as described above. For a comprehensive list of all PCR primers and their sequences see Table A2.

Colony PCR was used as an exclusive method to screen transformants for insertion of the desired DNA sequence into the plasmid when colonies or cultures of single colonies were subjected to PCR. A sterile toothpick was used to scrape part of a single colony off of the plate and either directly added to the PCR reaction or inoculated into 4 mL LB with antibiotic. After overnight incubation at 37°C with 250 rpm shaking, 0.5 µL culture was then added to the PCR reaction. Cultures and colonies were screened with PCR, in which cells were lysed and Taq polymerase activated by initial heating to 95°C for 3 min. Reactions were then cycled 35 times at 95°C for 30 s, 60°C for 30 s, and 72°C for 1 min kb^{-1}, followed by a final extension at 72°C for 10 min.

PCR Amplification of the Partial *tmuM* Gene from CBB1 Genome

Forward degenerate primer tmuM-degF1 (Table A2) was designed based on the lowest degeneracy of amino acids (KIGADVT) in the N-terminal sequence of the purified protein (Fig. 4.3 and 4.4). A BLASTP search of the N-terminal amino acid sequence showed significant homology with five proteins (GenBank: ACF60813, EAQ44903, CAJ95547), including a FAD-binding monooxygenase (GenBank: EDZ47364) and salicylate 1-monooxygenase (GenBank: AAZ62959). A ClustalW2 [Larkin *et al.*, 2007] alignment of these five homologous proteins identified a conserved region near the 3′ end of this multiple sequence alignment (Fig. 4.4). Reverse degenerate primer tmuM-degR1 was designed from this conserved region (Table A2).

A PCR reaction using tmuM-degF1 and tmuM-degR1was conducted using *PfuUltra* DNA polymerase HF (Agilent Technologies, La Jolla, CA) with fosmid DNA pMVS848 as template. The PCR conditions were: 95°C for 30 s, 62°C for 30 s, and 72°C for 1 min, for 30 cycles, and a final extension at 72°C for 10 min. The amplified PCR product was cloned into vector pSC-B-amp/kan (Agilent Technologies, La Jolla, CA) and three clones were chosen for DNA sequencing. Specific primers tmuM-up1, tmuM-up2, tmuM-down1, and tmuM-down2 (Table A2) designed from the DNA sequence of this amplified PCR product were used to sequence the flanking region of the PCR product with fosmid DNA pMVS848 as template.

PCR Amplification of Complete *tmuM* Gene and Cloning

Primers tmuM-F-NdeI and tmuM-R-NotI (Table A2) were used with the fosmid DNA pMVS848 as the template to amplify the entire coding gene sequence of TMU-

oxidizing enzyme. The reverse primer was also designed by removing the wild-type stop codon so that the histidine-tag contained in the pET32a vector would be fused to the expressed protein. PCRs were carried out with *PfuUltra* DNA polymerase. The amplification product was cloned into pSC-B-amp/kan to generate pSC-tmuM. The restriction fragment containing *tmuM* gene obtained from *Nde*I-*Not*I digestion of pSC-tmuM was ligated into the dephosphorylated vector, pET32a (Novagen) previously digested by *Nde*I and *Not*I. DNA sequencing confirmed that no mutation was introduced in the gene encoding the TMU-oxidizing enzyme in plasmid pET32a. The resulting plasmid pMVS848a (Table A1) was electroporated into *Escherichia coli* BL21(DE3) cells as described below. Functional expression of recombinant His-tagged TMU-oxidizing enzyme in *E. coli* BL21(DE3) and the two-step purification using metal affinity and ion exchange chromatography are described below.

Procedures for Transformation of Competent *Escherichia coli* Cells

Transformation by Electroporation

Plasmids were introduced into electrocompetent cells of *E. coli* JM109 and *E. coli* BL21(DE3), by electroporation with a GenePulser Xcell electroporation unit (Bio-Rad). Approximately 0.5-3 µL plasmid stock was added to 40 µL JM109 cells in a 2-mm electroporation cuvette (Bio-Rad) and subjected to 2.5 kV, 200 Ω, and 25 µF for 4-5 ms. Alternatively, plasmid stock was added to 20 µL BL21(DE3) cells in a 1-mm electroporation cuvette (Bio-Rad) and subjected to 1.8 kV, 200 Ω, and 25 µF for 4-5 ms. Transformant cells were rescued by aseptically adding in 1 mL SOC medium to the transformation reaction mixture and incubating at 37°C and 250 r.p.m. rotary shaking for

1 h. At the end of the 1 hour incubation period, two LB plates containing appropriate antibiotic were plated with 20 and 200 µL of the transformation reaction mixture. Plates were incubated overnight at 37°C.

Chemical Transformation

Plasmids were introduced into *E. coli* DH5α , One shot and Strataclone SoloPack competent cells by chemical transformation method. Approximately, 1 µL of cloning reaction mixture containing 5 to 10 ng of DNA was added to 25 µL competent cells and incubated on ice for 20-30 minutes followed by a 30-45 s heat-shock at 42°C. The transformation reaction mixture was then immediately placed on ice promptly after the heat-shock period. A 250 µL of pre-warmed LB media or SOC medium was then aseptically added to the transformation reaction mixture and incubated at 37°C for 1 h at 225 r.p.m. At the end of the 1 h incubation period, two LB plates containing appropriate antibiotic were plated with 20 and 200 µL of the transformation reaction mixture. Plates were incubated overnight at 37°C.

Heterologous Expression and Purification of Recombinant TmuM

For recombinant His-tagged trimethyluric acid monooxygenase (TmuM) expression, *E. coli* BL21(DE3) cells carrying pMVS848a cells were grown at 37°C in 1 liter of Terrific Broth (Life Technologies, Rockville, MD) containing ampicillin (100 µg mL^{-1}) using a 2.8 liter Fernbach Flask with shaking at 250 r.p.m. When the culture reached mid-exponential growth phase (turbidity at 600 nm = 0.6 to 0.8), the temperature was reduced to 25°C and IPTG was added to a final concentration of 150 µM. After 4 h,

cells were harvested by centrifugation. Cell paste (10 g) was stored at -80°C. About 10 g frozen cells were thawed and suspended in 30 mL 50 mM KPi buffer (pH 7.5) with DNaseI at10 µg mL^{-1}. The cell-suspension was broken by French press twice as described above. Unbroken cells and debris were removed from the lysate by centrifugation (20,400 x g for 20 min at 4°C).

The clear supernatant was loaded onto a 1-mL (bed volume) immobilized Ni-Sepharose affinity chromatography (GE life sciences, USA) column pre-equilibrated with 20 mM KPi buffer containing 500mM NaCl. After washing unbound proteins with 50 mL of equilibrating buffer, bound proteins were eluted with 50 mL linear gradient of imidazole (0 to 500 mM) in 25mM KPi buffer containing 500 mM NaCl at a flow rate of 1 mL min^{-1}. Fractions containing TmuM activity were combined, concentrated, and exchanged into 50 mM KPi buffer (pH 7.5) via ultrafiltration. The concentrated active fraction was loaded onto a 5-mL (bed volume) Q-Sepharose column (GE Healthcare) pre-equilibrated with 50mM KPi buffer (pH 7.5). After washing the unbound proteins with 10 mL of 50 mM KPi buffer, bound proteins were eluted with 50 mL linear gradient of KCl (0 to1 M) in 50 mM KPi buffer at a rate 1 mL min^{-1}. Fractions exhibiting TMU oxidation activity were pooled, concentrated and exchanged into 50 mM KPi buffer (pH 7.5).

CHAPTER 3

METABOLISM OF CAFFEINE BY *PSEUDOMONAS* SP. STRAIN

CBB1 VIA C-8 OXIDATION PATHWAY

Introduction

In microbes, caffeine is known to be utilized / metabolized via two distinct

pathways namely *N*-demethylation and C-8 oxidation. Although caffeine degradation via

these two pathways in microbial system has been reported previously [Woolfolk *et al.*,

1975; Mazzafera *et al.,* 1996; Madyastha *et al.*, 1999; Dash *et al.*, 2006; Yu *et al.*, 2008;

Yu *et al.*, 2009; Gopishetty *et al.*, 2011; Summers *et al.*, 2011; Summers *et al.*, 2012],

little is known about the details of the genes, the enzymes and the fate of intermediate

metabolites involved in these pathways. Nevertheless, these reports provide two major

indications. Firstly, they highlighted the fact that these two pathways are possibly the

most predominant natural processes for microbial caffeine degradation. Secondly, these

studies indicate that these two pathways might be playing a pivotal role in the overall bio-

geochemical cycle of carbon and nitrogen in the biosphere. Thus, not only from

biochemistry and microbiology perspective but also form an environmental and ecology

standpoint, detailed genetic and enzymatic characterization of these two pathways are

important.

In order to further investigate the details of these two pathways, our research

group isolated from soil, four bacterial strains (namely *Pseudomonas* sp. strain CBB1,

Pseudomonas nitroreducens CBB3, *Pseudomonas putida* CBB5 and CES) capable of

growth on caffeine as sole source of carbon and nitrogen. It is worth mentioning here

that caffeine is one of the key metabolites produced by at least 13 different orders of plant kingdom (nearly 100 species) [Ashihara *et al.*, 1999; Ashihara *et al.*, 2001; Ashihara *et al.*, 2008], which eventually gets released into the environment. Thus, it is not surprising that bacteria capable of growth on caffeine are prevalent in soil. Two of the four isolates (*Pseudomonas* sp. strain CBB1 and *Pseudomonas putida* CBB5) were chosen for detailed investigation of caffeine degradation pathway. Preliminary experiments on these two organisms suggested that they carry out caffeine degradation via the above two caffeine biodegradation pathways. While, the *N*-demethylation was found to be the pathway of caffeine degradation in *Pseudomonas putida* CBB5, *Pseudomonas* sp. CBB1was found to utilize the C-8 oxidation pathway.

Detailed enzymological and genetic characterization of the *N*-demethylation pathway in *Pseudomonas putida* CBB5 has been accomplished by Dr. Ryan Summers [Yu *et al.*, 2009, Summers *et al.*, 2011; Summers *et al.*, 2012]. In this pathway, caffeine undergoes sequential *N*-demethylation reactions to form theobromine (3,7-dimethylxanthine) and paraxanthine (1,7-dimenthylxanthine). These N-demethylations are catalyzed by two highly specific *N*-demethylases (1-specific NdmA and 3-specific NdmB). The dimethylxanthines are further N-demethylated to 7-methylxanthine. This reaction is catalyzed by NdmC, which remains to be fully characterized (see Fig. 1.4 in Chapter 1). Xanthine then undergoes C-8 oxidation to form uric acid, and is further mineralized by the well-known xanthine degradation pathway [Werner *et al.*, 2001]. In contrast, in the C-8 oxidation pathway, caffeine undergoes direct oxidation to form 1,3,7-trimethyluric acid (TMU) (see Fig. 1.5 of Chapter 1) [Madyastha *et al.*, 1998; Madyastha *et al.*, 1999; Dash *et al.*, 2006; Mohapatra *et al.*, 2006; Yu *et al.*, 2008]. Little is known

about further metabolism of TMU. The current study aims to characterize in detail the C-8 oxidation pathway in *Pseudomonas* sp. strain CBB1.

Results

Biotransformation of Caffeine to TMU by CBB1 Resting Cells

Pseudomonas sp. strain CBB1 cells were grown, harvested, thoroughly washed, and resuspended to high cell density suspensions to carry out resting cell assays as described in Chapter 2. Biotransformation of caffeine and TMU by CBB1 resting cells were carried out at various resting cell densities. In these assays, density of CBB1 resting cells suspended in 50 mM potassium phosphate (KPi) buffer (pH 7.5) was measured at 600 nm (OD_{600}), and concentration of caffeine or TMU was measured at 273 nm and 289 nm, respectively. The above experiments were carried out at CBB1 cell density $O.D._{600}$ of 2.5 (Fig. 3.1A) and 5 (Fig. 3.1B) while keeping the initial substrate (caffeine or TMU) concentration at 1 mM. Caffeine and TMU disappearance was followed as a function of time in order to track intermediates of the C-8 oxidation pathway. Metabolites were separated and identified using HPLC as described in Chapter 2.

Caffeine and TMU were actively consumed by CBB1 when the resting cell assays were carried out at both cell densities (OD_{600} of 2.5 and 5). While CBB1 resting cells at an OD_{600} of 2.5 required almost 4 h to completely consume 1 mM caffeine (Fig. 3.1A), cells suspensions with two times higher optical density ($OD_{600} = 5$) completely consumed the same amount of caffeine in about half the previous time, i.e. within 2.5 hours (Fig. 3.1B). In contrast, TMU consumption took nearly the same amount of time by CBB1 at both cell densities. (Fig. 3.1C and D). While caffeine transformation by CBB1 resting

cells was directly proportional to the optical density of CBB1 cells, as expected, TMU transformation showed no dependency on CBB1 cell density.
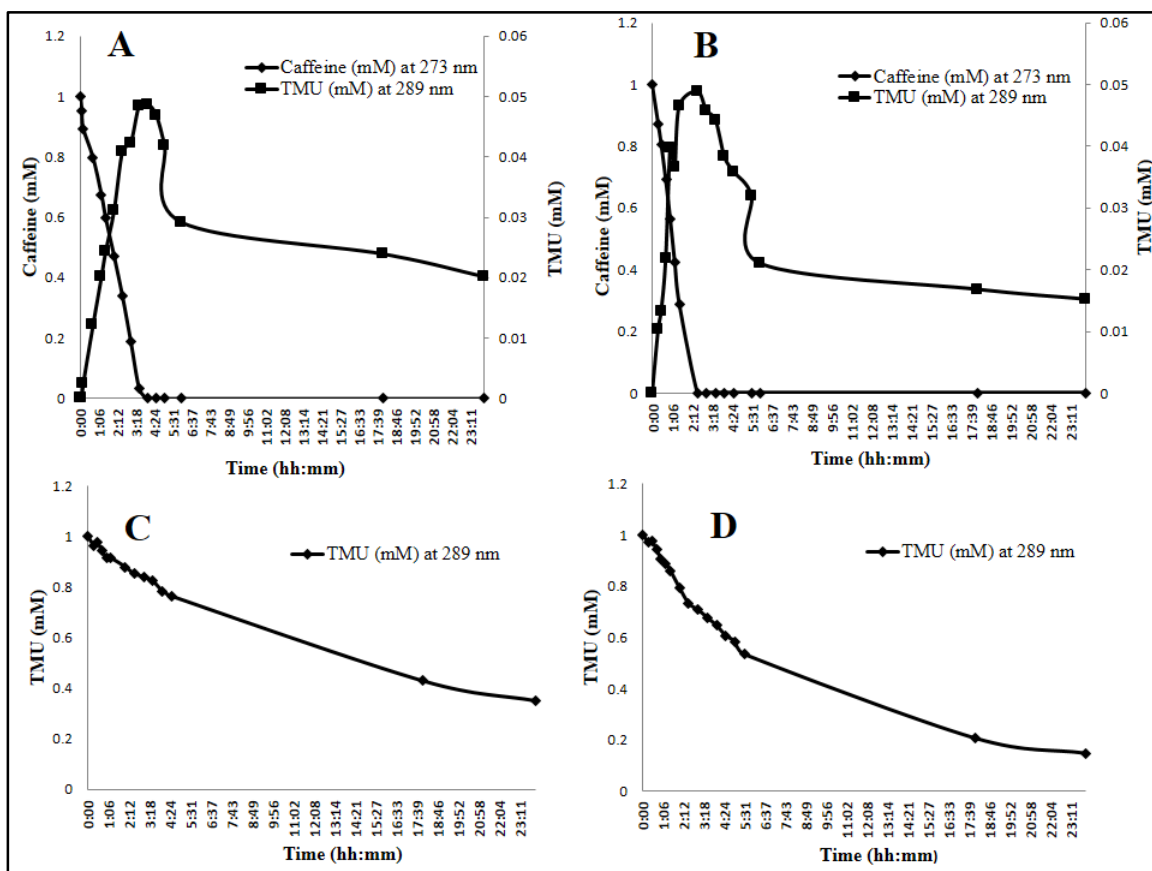
Figure 3.1. Determination of rate of substrate consumption and metabolite formation by CBB1 resting cells. HPLC analyses of caffeine (♦) consumption and concomitant TMU (■) formation at cell densities at OD_{600} of 2.5 (A) and 5 (B). The concentration of caffeine and TMU are shown in primary y-axis and secondary y-axis, respectively. HPLC analyses of TMU (♦) degradation by CBB1 resting cells at various cell densities: OD_{600} of 2.5 (C) and 5 (D).

In the above experiments, consumption of caffeine was accompanied by concomitant formation of TMU (Fig. 3.1A and B). As shown in Fig. 3.1A, resting cell assay with a cell density of OD_{600} 2.5, caffeine was consumed at a rate of 0.264 mM hr^{-1} and concomitant accumulation of TMU occurred at a rate of 0.016 mM hr^{-1}. TMU

accumulation in this reaction reached the maximum limit in about 3.5 to 4 hour and then gradually disappeared. However, the rate of disappearance or degradation of TMU was much slower than the rate at which it accumulated (Fig. 3.1A). Additionally, TMU concentration during accumulation phase linearly increased with concomitant degradation of caffeine and reached a maximum value exactly at the same time when all the caffeine was completely consumed. TMU started to degrade afterwards, but at a slower rate. The same phenomenon was observed with the resting cell assay with higher cell density (i.e., OD_{600} 5) where caffeine was consumed at a rate of 0.475 mM hr^{-1} and TMU accumulated at a rate of 0.044 mM hr^{-1} and reached the maximum at about 2.5 hr. These results further support the fact that TMU is indeed the direct oxidation product of caffeine in the C-8 oxidation pathway in CBB1 although they are degraded at different rates *in vivo*.

Stoichiometric Production of TMU from Caffeine by CBB1 Resting Cells

In the previous resting cell assays it was observed that although caffeine consumption by CBB1 resting cells was accompanied by concomitant production of TMU, the ratio of caffeine consumed to TMU produced or accumulated was not 1:1. For example, in the resting cell assay with cell density at OD_{600} equal to 5 with 1 mM caffeine (Fig 3.1B), the substrate was completely consumed in 2.5 hours, but only about 0.048 mM of TMU accumulated. This could be due to faster metabolism of TMU. In order to test this hypothesis and show stoichiometric formation of TMU from caffeine, resting cell assay was carried out at 10 mM of caffeine concentration (i.e., at an initial concentration of ten times the previous experiments) and with slightly higher cell density (i.e., OD_{600} 6 instead of OD_{600} 5 to). This condition was set up in order to avoid any

substrate inhibition (Fig. 3.2A). Under these conditions, it was observed that 10 mM caffeine was completely consumed in about 4 hours with concomitant accumulation of about 8.73 mM of TMU (Fig. 3.2A).
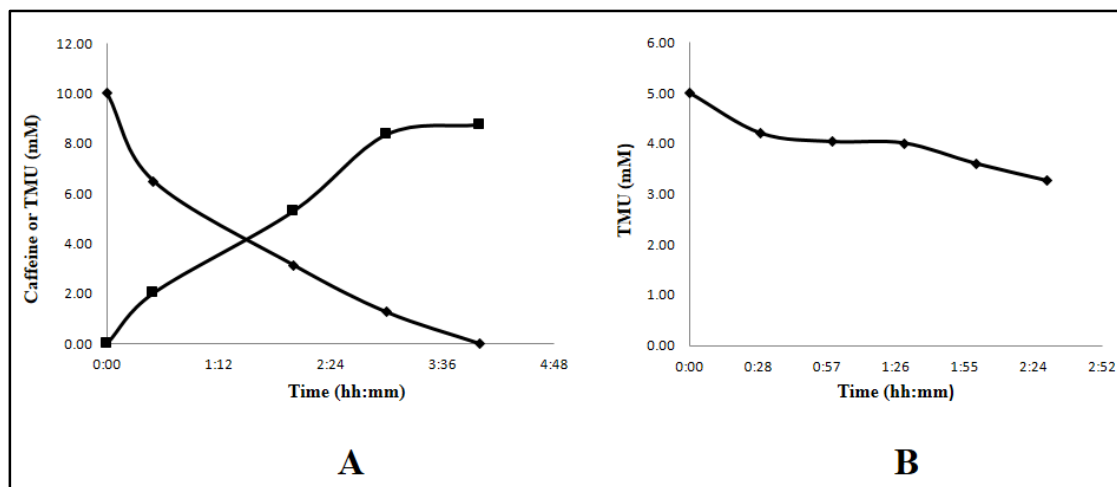
Figure 3.2. CBB1 resting cells assays. (A) Stoichiometric conversion of caffeine (♦) to TMU (■) at cell density at $OD_{600} = 6$. (B) TMU degradation at cell density at $OD_{600} = 6$ and TMU concentration of 5 mM.

These results affirmed stoichiometric conversion of caffeine to TMU by CBB1 resting cells. In these experiments, it was observed that a significant portion (about 87.3 % of 10 mM) of total TMU produced still remained in the assay buffer without getting metabolized. However, the concentration of accumulated TMU was slightly lower (about 12.7 % lower) than the expected 10 mM. This lower value could be attributed to partial degradation. Nevertheless, this resting cell condition successfully demonstrated stoichiometric conversion of caffeine to TMU conversion.

Profiling CBB1 Growth Media for Identification of C-8 Oxidation Pathway Metabolites

Pseudomonas sp. CBB1 was grown on caffeine as sole source of carbon and nitrogen, as described in Chapter 2. As this growth media contains high levels of caffeine (2.5 g caffeine L^{-1} or 12.87 mM), it was hypothesized that during growth, C-8 oxidation pathway intermediates will either accumulate or be transiently produced in the media. It is worth noting here that a previous study on caffeine degradation by a mixed culture consortium of *Klebsiella* and *Rodococcus* [Madyastha *et al*., 1998] was the first and the only report which suggested the possible existence of the C-8 oxidation pathway. This study also suggested that apart from TMU, two other potential metabolites (possibly TMA and TMAA) are also produced. However, this work was solely based growth associated metabolite formation; no biochemical or genetic evidence was provided in order to further support the suggested pathway. Nevertheless, this work served as a starting point to develop the hypothesis for the current study. Thus, our first effort was to develop a suitable analytical method to detect the presence of the proposed metabolites (TMU, TMA and TMAA) in the CBB1 spent growth media containing caffeine.

TMU is commercially available and has distinct λ_{max} at 273 nm and 289 nm. The retention times with a C18 column with a given mobile phase as described in Chapter 2 was about 19.3 minutes. This method was useful for detection and quantification of TMU (Fig. 3.3). Standards of TMA and TMAA are not available. Also, there is no described synthetic method for these compounds. Another analytical problem with these compounds is that they do not have a prominent UV maximum absorption like TMU; hence cannot be detected using a PDA detector (Fig 3.3).
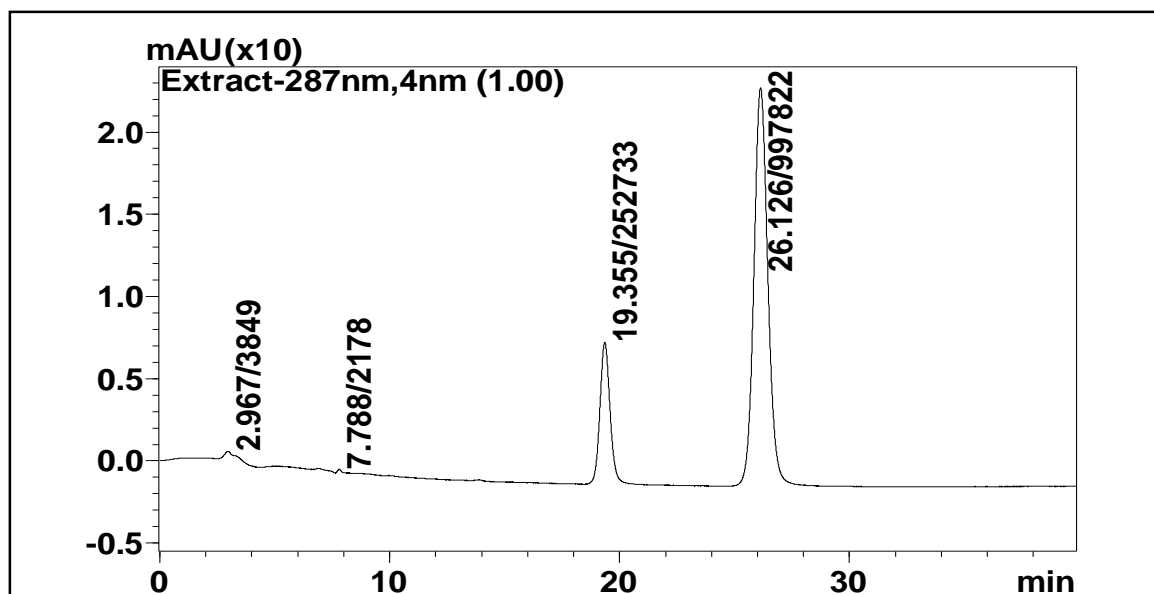
Figure 3.3 High performance liquid chromatography (HPLC) elution profiles of
metabolites in the spent medium of CBB1 in caffeine. Retention times of
26.13 and 19.35 mins corresponds to caffeine and TMU, respectively.

Analysis of the spent media of CBB1 cells grown in caffeine detected only two

peaks at retention times 26.13 and 19.63 min corresponding to caffeine and TMU,

respectively. Thus, none of the downstream metabolites of TMU like the hypothesized

TMA or TMAA could be detected by this method. In order to detect these molecules, an

alternative technique of tracking the mass-to-charge ratio of the associated ions under

defined ionization conditions was developed.

Identifying C-8 Oxidation Pathway Metabolites Using LC-MS

A Shimadzu LC-20AT reverse phase HPLC system equipped with Shimadzu

LCMS-2010EV single-stage quadrupole mass analyzer was used developed to track the

downstream metabolites of C-8 pathway in the CBB1 spent media. The ESI-LC/MS

analysis data were monitored in the positive-ion mode at m/z = 195, 211, 227, 245 or 201,

based on the expected products formed from caffeine (molecular weight [MW], 194) *via* TMU (MW, 210), which were hypothesized to be TM-HIU (MW, 226), TM-OHCU (MW, 244), and TMA (MW, 200), respectively. Caffeine and TMU were also monitored using the selected ion monitoring (SIM) of the ESI-LC/MS analysis data as $(M+1)^+$ ions at *m/z* = 195, 211 and 201 (Fig. 3.4), respectively in the spent media. In addition, analysis of the spent media revealed an ion with *m/z* value as expected for TMA. This supported the hypothesis that TMA is one of the downstream metabolites of caffeine and TMU.
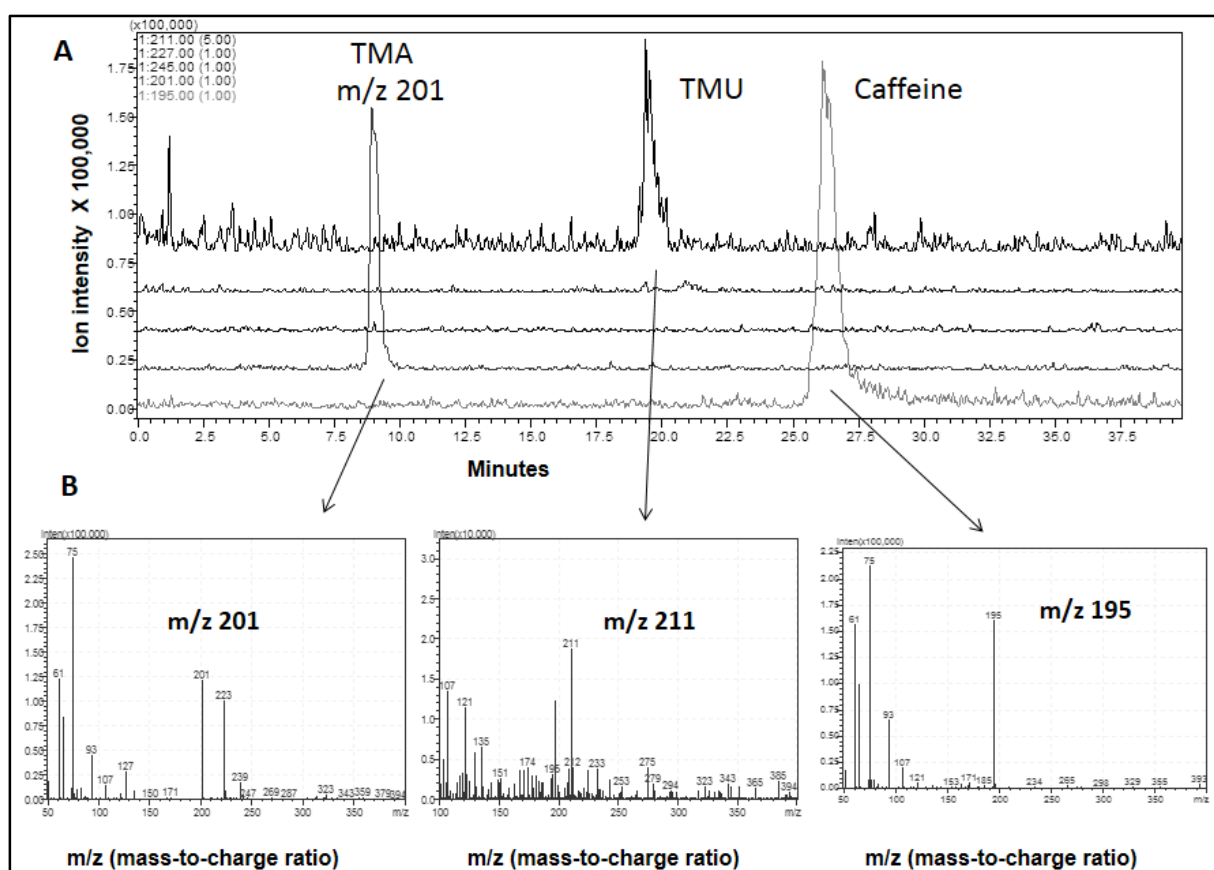
Figure 3.4 Elution profile and selected-ion monitoring (SIM) of the metabolites produced in the spent media of CBB1 grown in caffeine. (A) Detection using a HPLC-ESI-MS system in the positive-ion mode. (B) Extracted mass spectrum of various ion peaks showing enrichment of *m/z* 201, 211 and 195.

SIM analysis of the spent media of CBB1 grown on caffeine revealed the presence of TMA, but none of the other hypothesized downstream metabolites like TM-HIU or TM-OHCU could be detected. This could be attributed to several reasons such as lack of ionization under the current ESI/MS method. Also, these metabolites could be unstable like their nonmethylated counterparts, e.g., namely HIU and OHCU [Ramazzina *et al.*, 2006; Werner *et al.*, 2011]. The second reason seems to be more convincing because there might be both enzymatic and spontaneous conversion of TM-HIU and TM-OHCU to TMA. In order to minimize the conversion of TM-HIU and TM-OCHU to TMA, the following experiment was designed.

Time course of TMU Degradation by Crude Enzyme Preparation from CBB1 to Track C-8 Oxidation Pathway Metabolites Using LC-MS

In order to favor the accumulation of TM-HIU and TM-OHCU during caffeine degradation via C-8 oxidation pathway, it is necessary to slow down the rate of enzymatic and non-enzymatic degradation of these compounds in the reaction mixture. To reduce the rate of enzymatic degradation the corresponding downstream enzyme(s) present in CBB1 has to be removed. To reduce the rate of spontaneous degradation of these metabolites the reaction may be carried out at ambient temperature, and samples frozen prior to analysis. For this, TMU degradation activity was partially purified from CBB1 by using DEAE ion–exchange column (as described in "protein purification" section in Chapter 2). This preparation was used to carry out transformation of 0.2 mM TMU in presence of 0.2 mM NADH at 30°C in 50 mM potassium phosphate (KPi) buffer (pH 7.5). A negative control involved boiling the partially purified CBB1 cell extract for

10 minutes at 100°C. Samples were taken from the reaction mixture at various time

points (5, 15, 30 min), and kept on ice. The enzymatic reaction was stopped by adding

equal volume of chilled acetonitrile, and subjected to analysis by LC-MS (Fig. 3.5A-D).
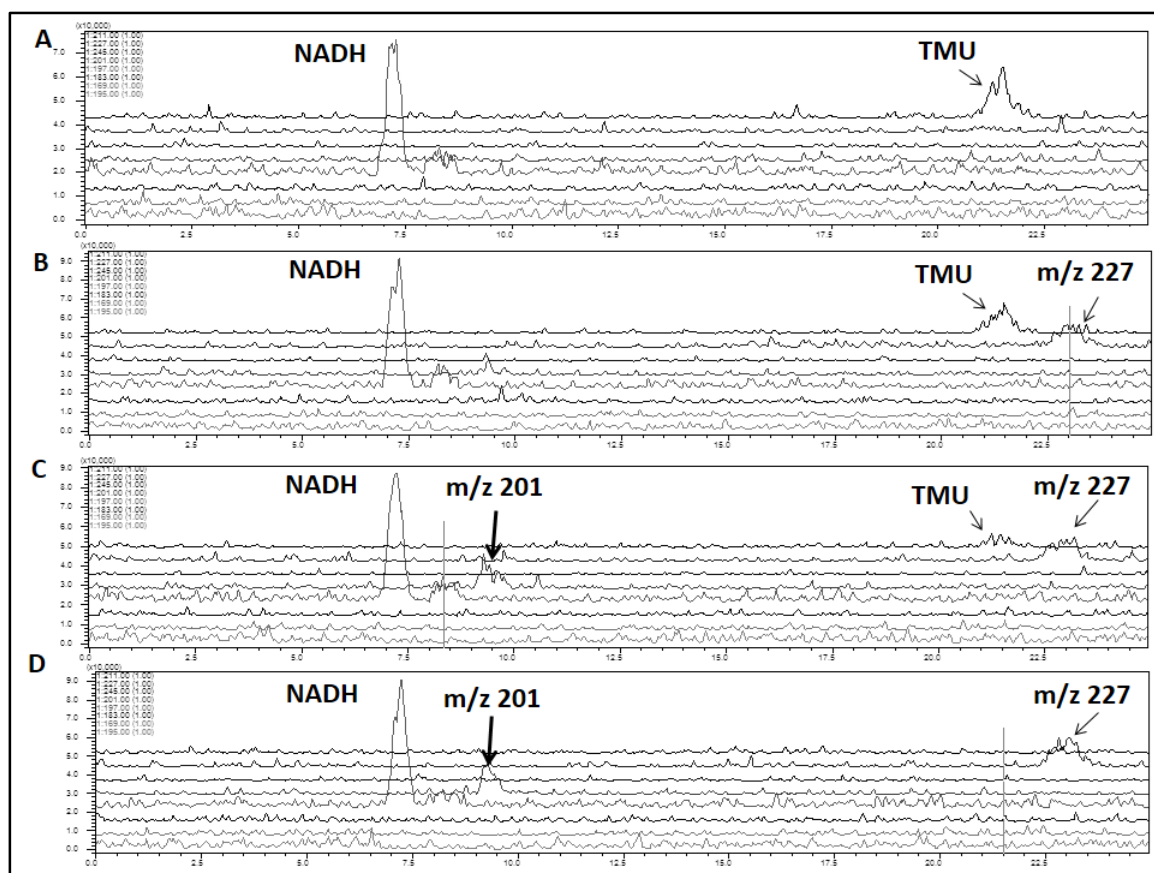
Figure 3.5 Selected-ion monitoring (SIM) of the time course of TMU degradation by
partially purified enzyme from CBB1 in the positive-ion mode. (A) 5-Minute
sample of the control reaction with boiled cell extract. (B) 5-minute (C) 15-
minute and (D) 30-minute samples of the enzymatic reaction. Peaks at *m/z*
227 and 201 show metabolite formation.

LC-MS analysis of time course of TMU degradation via tracking the

corresponding $(M+1)^+$ ions by SIM method showed that 0.2 mM TMU was completely

consumed in 30 minutes (Fig. 3.5D). Two new ion peaks were observed (Fig 3.5B-D)

during the course of the reaction at *m/z* corresponding to TM-HIU (*m/z* = 227) and TMA (*m/z* = 201). No TMU degradation was observed with the boiled control reaction even after 5 min of incubation (Fig 3.5A). However, no ion peak corresponding to *m/z* = 245, which would correspond to TM-OHCU (MW 244) was detected. Nevertheless, these results suggest that TM-HIU and TMA are the potential downstream metabolites of TMU. Any additional information of the pathway has to come from time course of metabolite analysis using LC-MS.

Production of Pure TMA and Analysis of Purity

The results obtained from previous experiments suggest that TMA and TM-HIU are possible downstream metabolites of caffeine and TMU. These results also demonstrate that TM-HIU is unstable and transiently formed during the degradation of caffeine to TMA. In contrast, TMA accumulates in the reaction mixture and is relatively stable under normal physiological conditions. It was important to isolate and purify TMA since this compound is not available commercially. The isolation method for this compound using the preparative TLC is described in Chapter 2, page 45. The quality and purity of isolated TMA was analyzed by LC-MS. When 2 µL of 0 5 mM purified TMA was analyzed by LC-MS in the positive ion mode, only a single ion peak was observed in the total ion spectrum (Fig. 3.6A). When the data were extracted from the same spectrum in selective ion mode for $(M+1)^+$ ions, it was observed that only a single ion peak had eluted exactly at the same retention time to the total ion peak (Fig. 3.6B) of *m/z* =201 (Fig. 3.6C). This highly purified sample of TMA was further analyzed for physical and chemical properties, and structural confirmation.
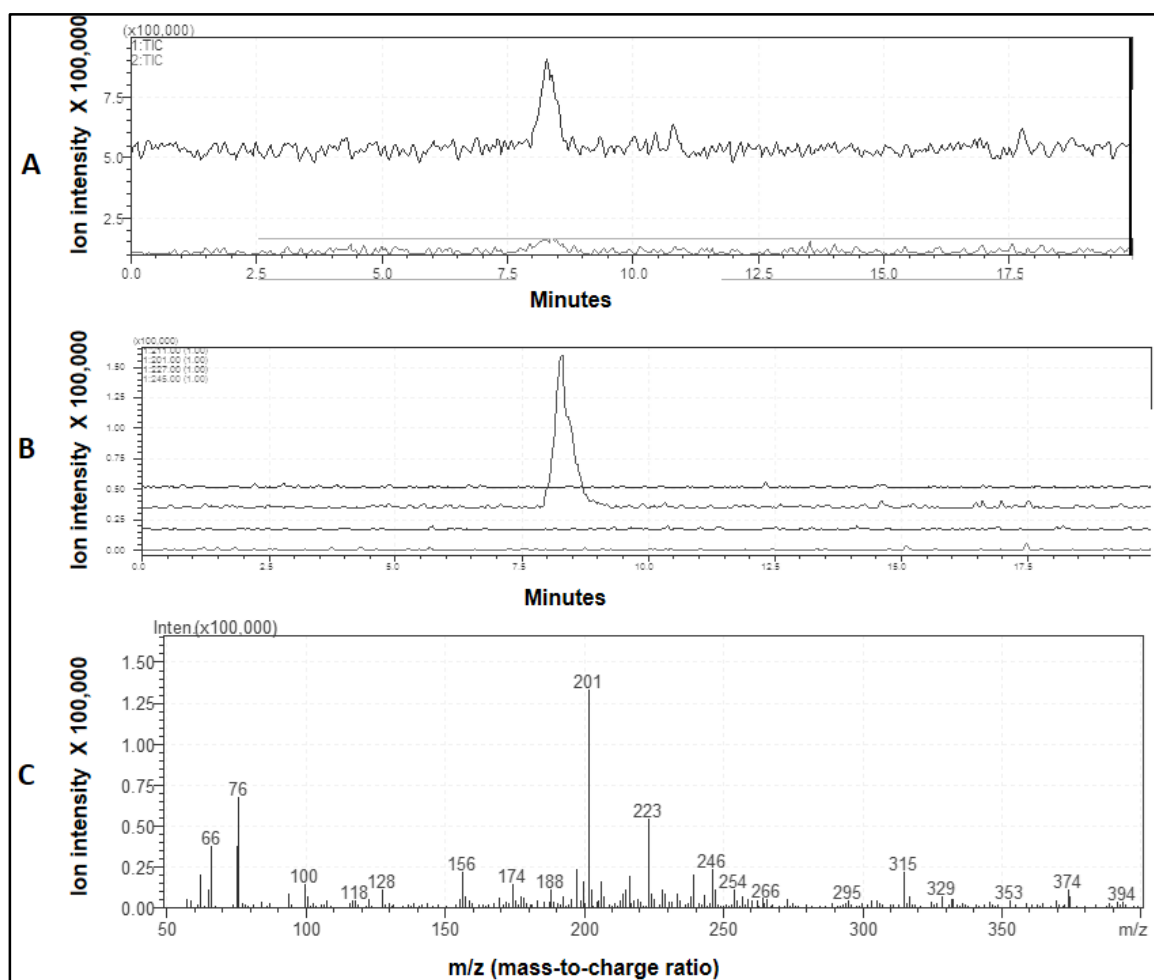
Figure 3.6. LC-MS analysis of purified TMA. (A) Total-ion monitoring and (B)
　　　　　Selected-ion monitoring (SIM) of the sample (C) Extracted mass spectrums
　　　　　in the positive ion mode of various ion peaks showing enrichment of the
　　　　　particular ion at *m/z* =201.

UV-visible and HPLC-MS Analysis of Purified TMA (Metabolite IV)

Determination of the UV-Visible spectrum of the prep-TLC purified sample of

metabolite IV was performed by using a UV-Visible spectrometer (Shimadzu UV-2450).

The absorption maximum was observed at 215 nm with a shoulder from 220 to 250 nm

and a plateau between 220 to 230 nm (inset of Fig. 3.7A). These lower absorption maxima of metabolite IV compared to caffeine or TMU can be explained using the structure of TMA in which the six-carbon ring structure is ruptured with loss of one carbon atom. The UV-Visible absorption spectrum further supports our hypothesis of Metabolite IV as TMA. This result explains why Metabolite IV was previously undetectable by HPLC analysis although it was getting accumulated in the spent media and was detectable in LC-MS.
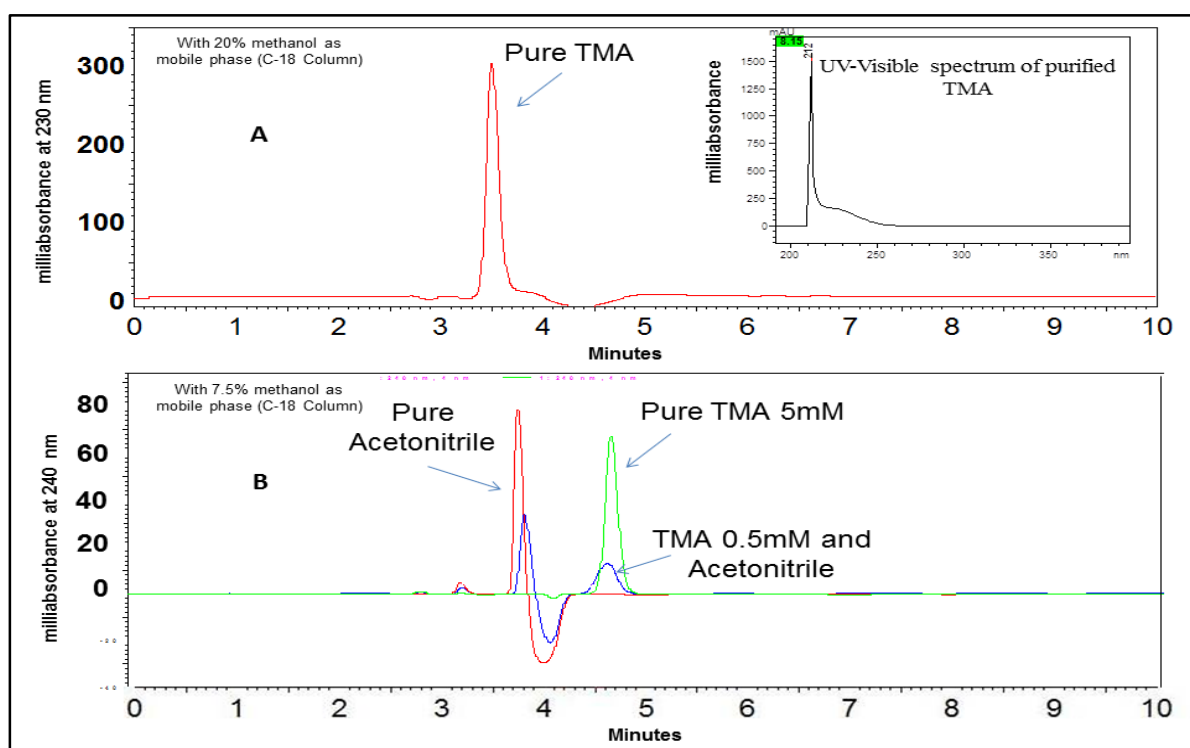
Figure 3.7. HPLC analysis of purified TMA. (A) Analysis using a C18 column in reverse phase (20% methanol as mobile phase) and detection at 230 nm. (Inset) UV-Visible spectrum of pure TMA peak. (B) Resolution of acetonitrile peak from pure TMA peak by changing the methanol concentration in mobile phase and detected at 240 nm

Development of a Standard Curve for TMA Using HPLC

Previously, all the metabolites were detected using a PDA detector at their respective λ_{max} (273 and 289 nm for TMU and caffeine, respectively), which were well above 250 nm. On the other hand, the absorption spectrum of metabolite IV (hence forth called TMA) had no absorption above 250 nm. It is worth mentioning here that no commercial sample of TMA is available and this is the first report of the UV-visible spectrum of this compound. Further, when 2 µL of 0 5 mM purified TMA was analyzed using HPLC using the same method previously used (20% methanol) and detected at 230 nm, the retention time (3.5 minutes) matched with that of acetonitrile (Fig. 3.7A). In order to separate acetonitrile and TMA, the HPLC analysis was changed to 7.5% methanol in the mobile phase. Under these conditions, acetonitrile peak eluted at 3.7 min and TMA at 4.7 min (Fig. 3.7B; detection at 240 nm). This result was reconfirmed by doing a similar analysis with 10 times lower concentration of metabolite TMA (0.5 mM). Thus, these results provided first level evidence that Metabolite IV is TMA.

The standard curve of the purified TMA is shown in Fig. 3.8. There is excellent correlation between the concentration and peak area, further confirming the purity of the compound. Given that the compound is not available commercially, this standard curve will serve as a reference for future studies on the C-8 oxidation pathway of CBB1.
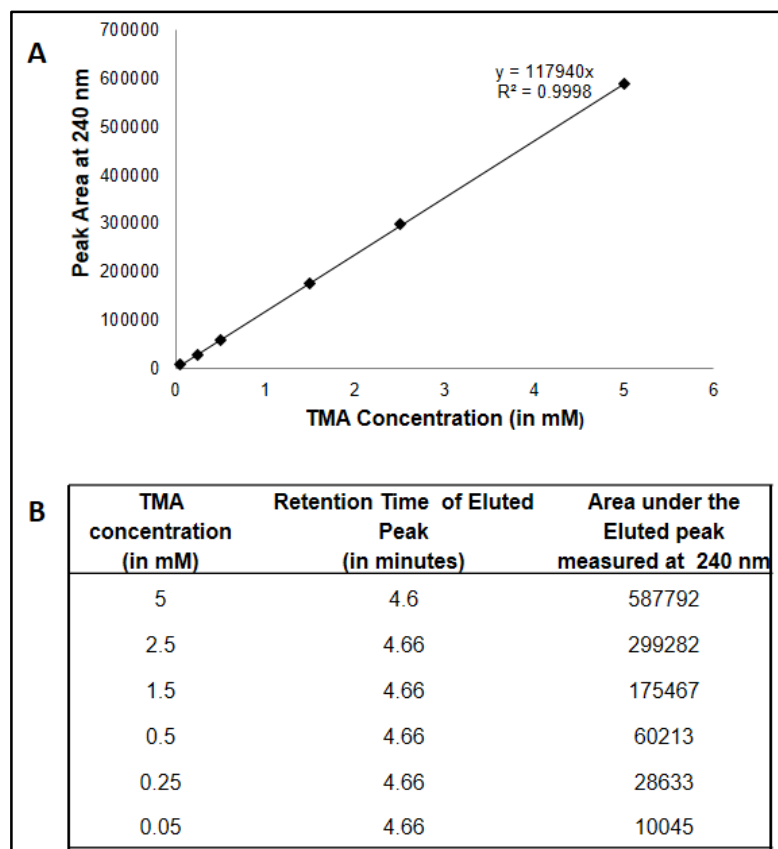
Figure 3.8. Standard curve of purified TMA. (A)Standard curve prepared using HPLC equipped with a C18 column and mobile phase containing 7.5 % methanol and detected at 240 nm. (B) Data used for preparing the standard curve.

TMA Structure Elucidation by HRESIMS and ^{1}H and ^{13}C NMR

High resolution electro-spray ionization mass spectroscopy (HRESIMS) was conducted at the mass spectroscopy facility of The University of Iowa to calculate the exact molecular weight of TMA and estimate the chemical composition. HRESIMS analysis of pure TMA (Fig. 3.9) showed that M^+ was $m/z = 200.0907$, indicating an empirical formula of $C_7H_{12}N_4O_3$ (calculated value, 200.0909). Thus, based on the results obtained from HRESIMS it is inferred that the molecule contains two hydrogen atoms more and one carbon atom less than TMU. This indicates that a hydrolysis accompanied

by decarboxylation at specific molecular positions of TMU can result in formation of TMA.
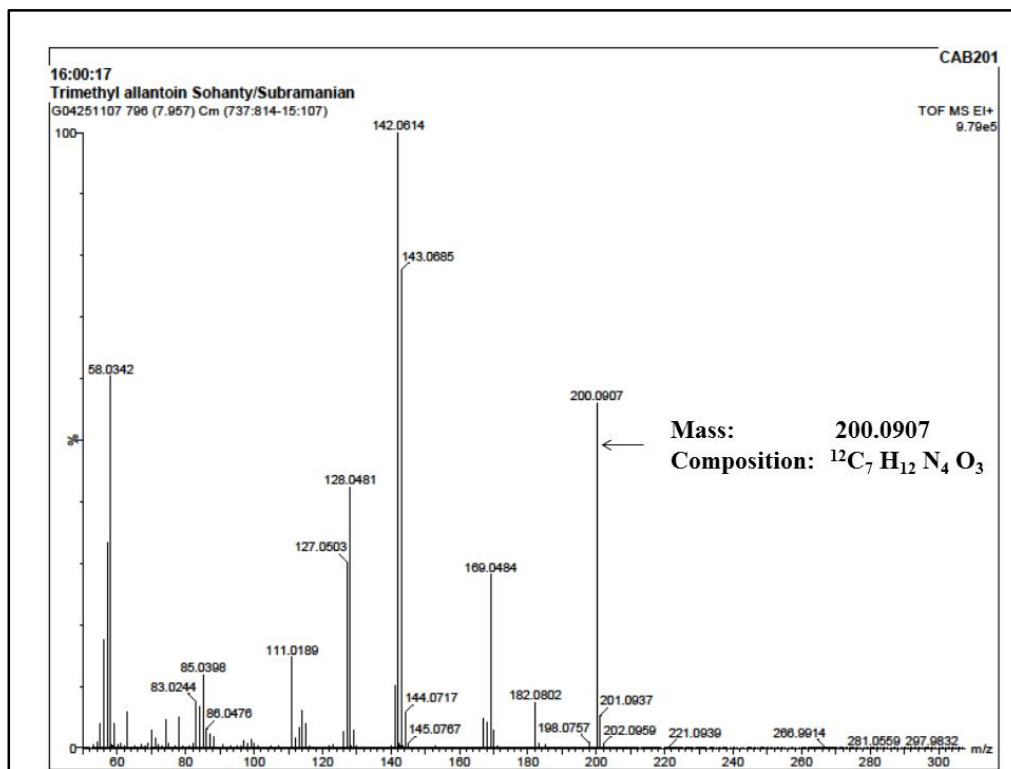
Figure 3.9. Mass spectrum of TMA from high resolution electro-spray ionization mass spectroscopy (HRESIMS).

In order to confirm structure of purified TMA, ^{1}H NMR and ^{13}C NMR techniques were employed as described in chapter 2. The ^{1}H NMR spectrum of metabolite IV contained all signals of TMU, in addition to two extra signals (Table 3.1). These included signals for a new secondary amine proton at 5.77 ppm for H-8 (singlet, broad) and a doublet at 5.19 ppm (J = 7.77Hz) for H-5 coupled to the H-1 signal of the secondary amine group. The ^{13}C NMR spectra of metabolite IV showed one less signal

than TMU but contained all three methyl group signals at 40.7 ppm (C-9), 26.8 ppm (C-10), and 24.9 ppm (C-11) (Table 3.1).

Table 3.1. ^{1}H and ^{13} C NMR data of TMA

H	Chemical shift (ppm)	Peak type	J coupling (Hz)	C
H-1 (1H)	6.51	Doublet	7.77	C-2 (CO)
H-5 (1H)	5.19	Doublet	7.77	C-4 (CO)
H-8 (1H)	5.77	Singlet, broad		C-5 (CO)
H-9 (3H)	2.98	Singlet		C-7 (CO)
H-10 (3H)	2.88	Singlet		C-9 (CH$_3$)
H-11 (3H)	2.69	Doublet	4.6	C-11 (CH$_3$)

From these data it is clear that loss of the carbon atom in metabolite IV was not due to *N*-demethylation but occurred *via* ring rupture. The ^{13}C NMR spectrum of metabolite IV also showed two carbonyl group signals at 156.1 ppm (C-4), which is farther downfield than the original TMU signal. In addition, the spectrum contained signal for a new secondary carbon atom at 65.93 ppm (C-5). Signal assignments and connectivities were confirmed by using COSY, HMBC, and HMQC spectral editing. COSY analysis correlated H-8 with H-11 and correlated H-1 with H-5. HMBC showed 3-bond correlations between H-11 and C-7, i.e., H-9 between C-2 and C-4, H-10 between C-5 and C-7 and H-5 between C-2 and C-10. Based on all of these data, the structure of purified TMA was confirmed (Fig. 3.10).
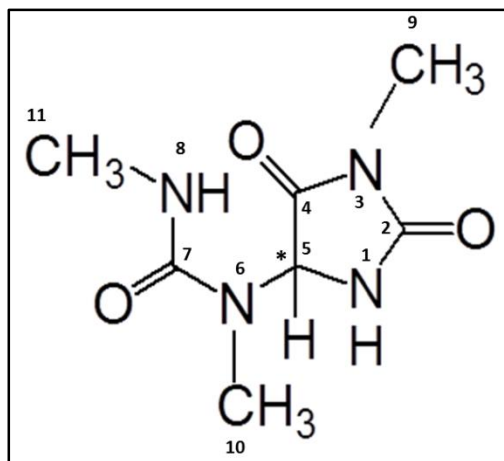
Figure 3.10. Structure of TMA showing a chiral center (*) at C-5 carbon atom.

Chiral Resolution of the Purified (Racemic) TMA into Enantiomers

When a purified sample of TMA was injected into a reverse phase C18 column

(20% methanol mobile phase running at 0.5 mL min^{-1} HPLC method), it eluted as a

single peak detectable at 230 nm with a retention time of 3.5 minute (Fig. 3.7A). When

the same injected sample after elution from HPLC at a slower flow rate (0.2 mL min^{-1})

was injected into ESI-MS, a single total ion and selected ion peak with $m/z = 201$ was

detected at 8.5 min (Fig. 3.6). These results established the purity of TMA with a

molecular weight of 200. Upon further investigation of the structure of TMA obtained

from NMR analysis (Fig. 3.10), it was found that this molecule contains a chiral center at

C-5 position. Thus, this molecule could exist in either of its enantiomeric forms. In

order to confirm this, the compound was subjected to Chiral HPLC-MS analysis as

described in Chapter-2. The pure sample of TMA which eluted as single peak in a C18

reverse phase HPLC (Fig. 3.7) resolved into its individual enantiomers on a chiral

column and eluted as two separate peaks TMA-A and TMA-B at retention time 10.7 and

12.9 minutes respectively (Fig. 3.11a and 3.11c). Further, it was observed that the UV-Visible spectrum of both of these resolved peaks were the same as that of the unresolved purified TMA (data not shown). It was also observed that the resolved enantiomers, when analyzed with ESI-MS (Fig. 3.11b) produced ions of equal mass-to-charge ratio (*m/z* =201) (Fig. 3.11d), which is same as that of unresolved TMA. From these results, it is inferred that the purified sample of TMA is a mixture of both the enantiomers (S and R).
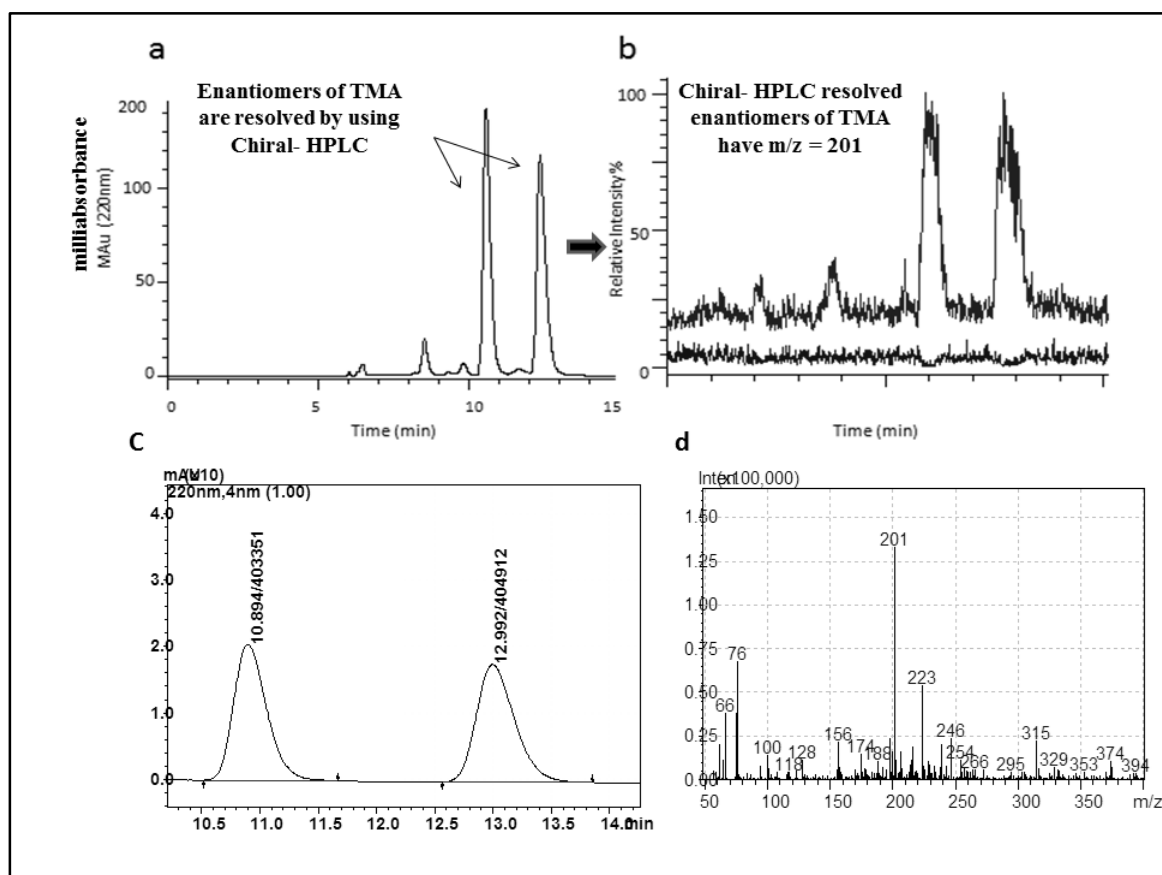
Figure 3.11. Chiral analysis of purified TMA. (a) Analysis with HPLC equipped with ChiralPak IA column and subsequent (b) ESI-MS analysis of the resolved peaks after elution. (c) Enlarged view of portion of chiral-HPLC chromatogram showing retention time and peak area of the resolved peaks analyzed at 220 nm. (d) Extracted mass spectrums of selected ion peaks in positive ion mode showing enrichment of a particular ion with *m/z* =201.

Furthermore, the results showed that both the eluted peaks, TMA-A and TMA-B, had nearly equivent peak area and slightly different peak shapes (Fig. 3.11c). Thus, the purified TMA is a racemic product of both enantiomers (TMA-A and TMA-B) in equimolar quantities.

Standard Curve for Both the Enantiomers of TMA Using Chiral-HPLC

Concentration of purified TMA (in mM)	Retention Time of Eluted Peak TMA-A (in minutes)	Area under the Eluted peak TMA-A measured at 220 nm	Retention Time of Eluted Peak TMA-B (in minutes)	Area under the Eluted peak TMA-B measured at 220 nm	Concentration of each enantiomer of TMA (TMA-A and TMA-B)* (in mM)
5	10.7	649697	12.9	647518	2.5
3	10.7	380213	12.9	378264	1.5
2.5	10.7	333435	12.9	330177	1.25
1.5	10.7	193061	12.9	190430	0.75
1	10.7	123902	12.9	123244	0.5
0.5	10.7	57919	12.9	55960	0.25
0.25	10.7	28642	12.9	26793	0.125
0.1	10.7	13171	12.9	13105	0.05
0.05	10.7	6659	12.9	6544	0.025

*Assuming racemic TMA is an equimolar mixture of both enantiomers of TMA.

Figure 3.12. Standard curves generated by analyzing purified TMA with a HPLC equipped with Chiralpak IA column. Concentration, retention time and peak area of both the enantiomers of TMA are shown in a tabular form (top) and in graphical form (bottom).

TMA is a novel compound that is not commercially available. There is no synthesis method described in the literature. At present, this molecule can be produced only via growth of CBB1 on caffeine and isolation/purification of the metabolite. Further, this is the first chiral resolution of the metabolite (Fig 3.11). The standard curve (for detailed methods, see Chapter 2) shown in Fig. 3.12 is an excellent fit for both enantiomers. This will be useful in further studies on the nature of TMA enantiomer that is formed during enzymatic and non-enzymatic conversion of TM-HIU.

Enrichment of One of the Enantiomers of TMA and its Gradual Racemization

One of the key questions is whether the TMA formed from TMU is chiral. Although TMA obtained from the partially purified enzymatic reaction is racemic, the racemization could have occurred during the isolation procedures. Biological reactions are stereospecific, and thus the prevailing hypothesis is that the enzymatic product is chiral. In order to investigate the chirality of TMA, a CBB1 resting cell assay was done with cell density at an OD_{600} of 6 and initial TMU concentration at 5 mM (Fig 3.2B). Samples were taken at various time points and clear supernatant was immediately analyzed using a chiral-HPLC under the established protocol (see Chapter 2). As shown in Fig. 3.13, at 30 minute time point, the ratio of TMA-A (i.e. peak area under peak A) to TMA-B (i.e. peak area under peak B) was 93:7. As the reaction progressed, the ratio of TMA-A to TMA-B dropped to 77:23 by 2.5 h (Fig. 3.13).
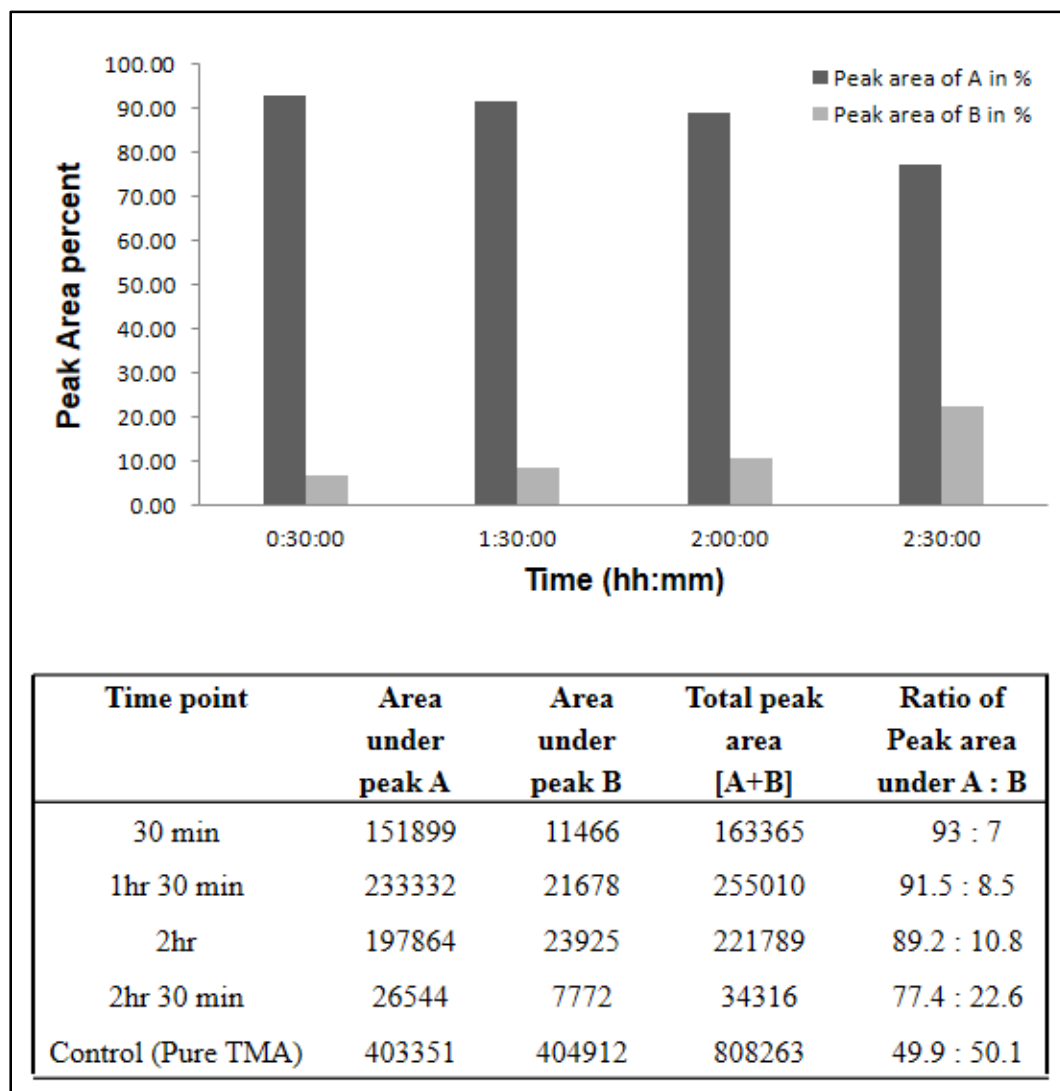
Figure 3.13. Analysis of TMU to TMA conversion using resting cells of CBB1 at various time points. TMA was resolved in a chiral column as described in Chapter 2. The raw data are shown in the table at the bottom.

These results indicate that TMA-A is the actual metabolite formed from TMU. With time, there is racemization to TMA-B (Fig. 3.13). It is clear that the TMA obtained as Metabolite IV above is a product of racemization due to isolation and purification procedures. Thus the TMA produced in the C-8 oxidation pathway is stereospecific. The

absolute stereochemistry and optical rotation of the enantiomers remains to be established.

Discussion

C-8 Oxidation Pathway in *Pseudomonas* sp. CBB1: Caffeine to TMA

Soil bacterium *Pseudomonas* sp. strain CBB1 has the ability to grow on caffeine ($C_8H_{10}N_4O_2$: a nitrogen rich compound) as the sole source of carbon and nitrogen. A novel caffeine dehydrogenase (cdh) was purified from CBB1 and shown to oxidize caffeine at C-8 position to form TMU [Yu *et al.*, 2008]. However, the fate of TMU was not established. It was hypothesized that C-8 pathway may be similar to the xanthine degradation pathway where xanthine gets converted to uric acid *via* oxidation at C-8 position by xanthine oxidase/dehydrogenase (see Fig. 1.6 of Chapter 1) [Werner *et al.*, 2011]. Uric acid is then hydroxylated at the 5' position to form 5-hydroxyisourate (HIU) by uric acid oxidase. HIU is unstable under normal physiological conditions with a half-life of 30 minutes and gets converted to OHCU [Kahn *et al.*, 1998] and then to racemic allantoin. These reactions are also catalyzed by HIU hydrolase and a decarboxylase. However, the product of the enzymatic reactions is S-(+)-allantoin).

Our hypothesis was that in the C-8 oxidation pathway, TMU, generated from caffeine is hydoxylated at the C-5 position to form TM-HIU, which is the trimethyl form of HIU (Fig. 3.14). Further, it was hypothesized that TM-HIU is unstable under normal physiological conditions and thus breaks down spontaneously into TM-OHCU and then to racemic TMA, the trimethyl forms of OHCU and allantoin, respectively (Fig.3.14). It

was also hypothesized that TMA formed via the enzymatic conversion of TMU would be stereospecific, likely (S-(+)-TMA) (Fig. 3.14).
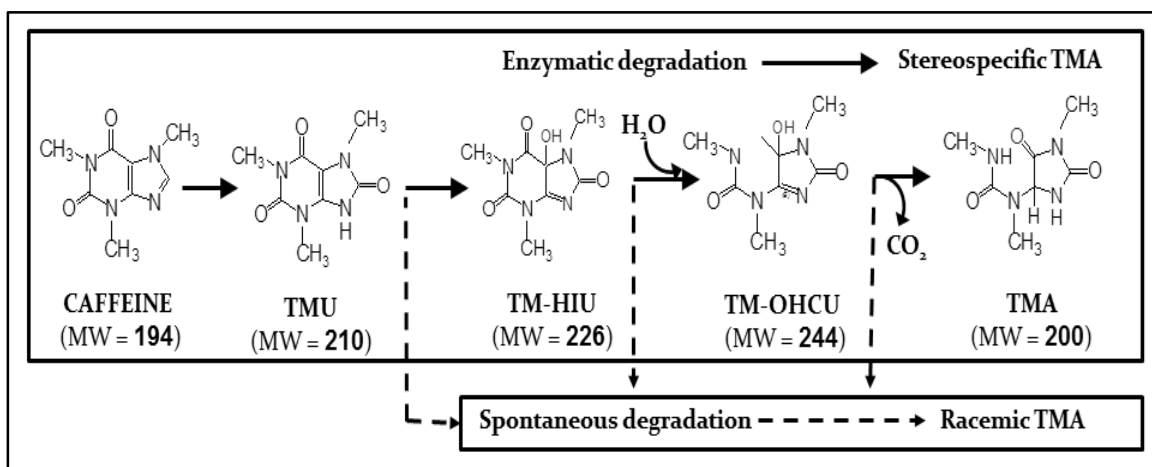
Figure 3.14. Diagrammatic representation of the C-8 oxidation pathway showing conversion of caffeine to TMA *via* TMU. TM-HIU and TM-OHCU. Non-enzymatic conversion is shown at the bottom.

Initial analysis of caffeine degradation by CBB1 cells both in spent media and resting cell assays indisputably showed that TMU is the first metabolite of C-8 oxidation. Upon further analysis of the spent media by LC-MS, it was observed that a new metabolite (designated Metabolite IV) is getting accumulated along with TMU. Later, this metabolite was isolated and identified as TMA by structural and molecular characterization methods. These results supported the initial hypothesis that TMA is one of the downstream metabolites of caffeine catabolism *via* C-8 oxidation pathway in CBB1 (Fig. 3.14). Molecular formula of TMA revealed that it contains two hydrogen (H) atoms more and one carbon (C) atom less than TMU. This information when combined with the corresponding TMA structural data further revealed that the loss of carbon atom from TMU during formation of TMA is not due to *N*-demethylation but due

to the ring rupture at a specific position (C1-C6). This could be by hydrolysis followed by a decarboxylation step (Fig. 3.14). This information further supports our hypothesis that TMU generated from caffeine undergoes hydroxylation at C-5 position to form TM-HIU followed by hydrolysis/ring opening at C1-C6 to form TM-OHCU. It is this molecule that undergoes decarboxylation to form TMA (in total net addition of 2 H atoms and liberation of 1 C atom) (Fig. 3.14).

TM-HIU: A New Metabolite of C-8 Oxidation Pathway

Two of the proposed intermediates in the working hypothesis namely TM-HIU and TM-OHCU could not be identified. It was assumed that these metabolites are transiently formed during conversion of TMU to TMA. When transformation of TMU was carried out with a crude enzyme preparation from CBB1 and metabolites were subsequently analyzed by LC-MS, transient formation of a metabolite (designated Metabolite II) was observed. This had a mass-to-charge ratio expected for TM-HIU in positive ion mode. The reason Metabolite II (possibly TM-HIU) was observed only in this experiment could be due to better control of reaction time with the partially purified enzyme. In addition to that, the samples collected from the enzymatic reaction were immediately stored on ice, which significantly reduced the rate of spontaneous degradation of TM-HIU to TMA. This experiment further supported the hypothesis that TM-HIU is formed from TMU possibly by a hydroxylation reaction (Fig. 3.14).

However, TM-OHCU, the hypothesized product of TM-HIU could not be detected by any of the reaction schemes employed. This could be due to much faster decomposition to TMA, similar to that observed for HIU [Ramazzina *et al.,* 2006]. One

way to detect TM-OHCU is by trapping at low temperatures or characterization of the reaction in vivo by methods such as NMR.

TMA: A Well Characterized Metabolite of C-8 Oxidation Pathway

TMA isolated and purified from the spent media was found to be racemic based on chiral-HPLC analysis. The enantiomers of TMA were successfully resolved and using a ChiralPak IA column. The resolved enantiomers of TMA had same UV-visible spectrum and almost equal peak area at the detection wavelength of 220 nm. This was in agreement with the hypothesis that the non-enzymatic product would be racemic. In a separate experiment of TMU degradation with CBB1 resting cells in which spontaneous generation of TMA was minimized, it was possible to demonstrate substantial enrichment of one of the enantiomers at 93:7 (Fig. 3.13). Over a period of 2.5 hours, it was possible to track the partial loss of the enriched enantiomer (Fig. 3.13). These results provided first level evidence that TMA derived enzymatically was racemic. Taken together, the results support the hypothesis for the C-8 oxidation pathway in CBB1 shown in Fig. 3.14.

Summary and Conclusion

Pseudomonas sp. strain CBB1 is capable of growth on caffeine as sole source of carbon and nitrogen. CBB1 uses C-8 oxidation pathway where caffeine gets oxidized at the C-8 position to form TMU. TM-HIU was identified as the next metabolite, which spontaneously degraded to TMA. For the first time, TMA was isolated, purified and fully characterized with respect to structure and molecular properties. TM-HIU was found to be unstable and difficult to isolate. Nevertheless, its formation was established

by LC-MS. TM-OHCU, the downstream metabolite of TM-HIU was also not detectable due to its instability. Further, TMA formed spontaneously was shown to be racemic, in contrast to the highly enriched enantiomer formed enzymatically.

These results supported the hypothesis, which was initially developed based on the xanthine degradation pathway via uric acid. The next goal was to fully characterize the enzyme involved in the formation of TM-HIU and to isolate the genes involved in the caffeine C-8 oxidation pathway.

CHAPTER 4

PURIFICATION AND CHARACTERIZATION OF A NOVEL

TRIMETHYLURIC ACID MONOOXYGENASE

Introduction

In the previous chapter (Chapter 3), using a partially purified cell extract of

CBB1, TM-HIU formation as a transient product was confirmed by LC-MS. Formation

of TM-HIU in a caffeine utilizing mixed culture was also proposed [Madysatha *et al.,*

1998]. However, little is known about the enzymology of TM-HIU formation and further

degradation of caffeine via C-8 oxidation pathway in CBB1. A detailed characterization

of TMU degradation will help (i) in explaining the formation of TMA via TM-HIU and

(ii) in identifying the true product of TMU conversion. In order to accomplish this goal,

the enzyme was purified. This chapter describes the first detailed purification and

biochemical characterization of a (tri)methyluric acid specific enzyme responsible for

metabolism of TMU in *Pseudomonas* sp. strain CBB1. The native enzyme purified from

CBB1 was heterologously expressed in *E. coli* as a His-tagged protein, which was used

for biochemical characterization. The enzyme was found to be a FAD-containing

NAD(P)H-dependent enzyme, which oxidized TMU at the C-5 position to produce TM-

HIU. As expected, uric acid was not a substrate for this enzyme. Further, the

involvement of TM-HIU in the pathway was confirmed and its spontaneous degradation

to TMA was explained.

<center>Results</center>

<center>TMU Oxidation Activity in CBB1 Cell Extracts</center>

TMU accumulation in CBB1 resting cell assays and crude cell extract (CCE) is
described in Chapter 3. Further studies with CCE showed that TMU degradation was
incomplete even after 24 hrs of incubation. Upon addition of 0.5 mM NADH to the
reaction mixture complete degradation of 0.5 mM TMU was observed in 15 minutes.
(Fig. 4.1A). The cell extracts showed very little background NADH oxidation activity
(Fig. 4.1B). Thus, it was concluded that TMU degrading enzyme was NADH-dependent
and designated as TMU oxidoreductase.
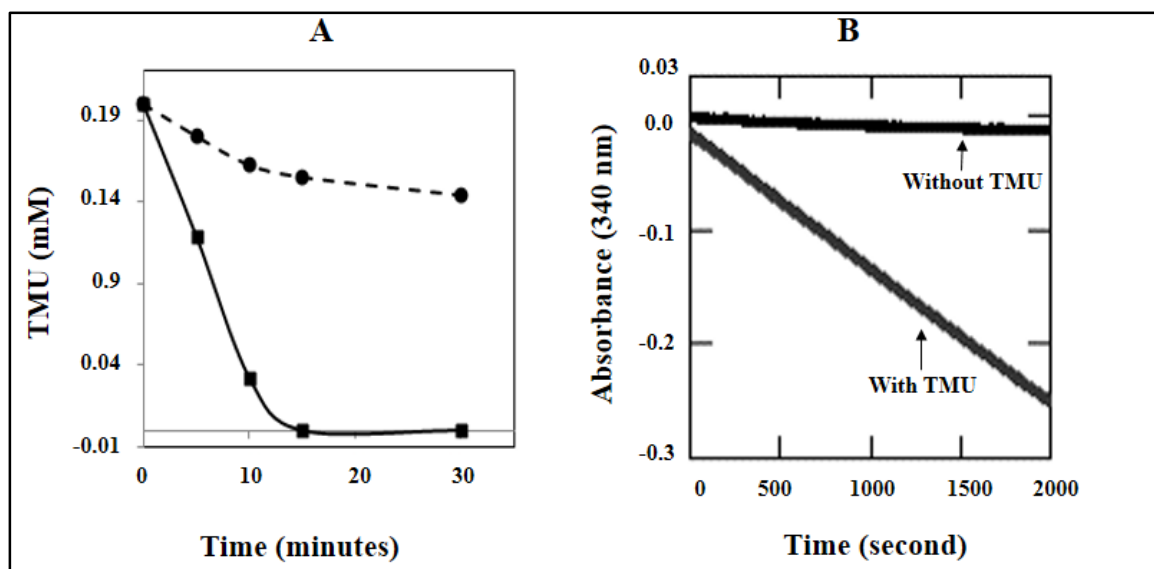
Figure 4.1. TMU oxidation by crude cell extracts (CCE) of *Pseudomonas* sp. CBB1. [A]
Monitoring TMU degradation by CCE with (solid) and without NADH
(dash) using HPLC. [B] UV-Visible spectrophotometric method to
determine kinetics of TMU dependent NADH oxidation at 340 nm.

Based on TMU-specific oxidation of NADH, a spectrophotometric assay was
developed using a UV/visible spectrophotometer (Shimadzu UV-2450), as described in

Chapter 2, to monitor TMU oxidoreductase activity during purification (Fig. 4.1B). In a typical 1-mL reaction, 0.5 mM TMU was completely consumed in 33 minutes in presence of 0.5 mM NADH by 100 µL of enzyme fraction (Fig. 4.1B).

Purification of TMU-Oxidoreductase from CBB1

Based on NADH dependent TMU oxidation, one unit of activity was defined as the amount of protein required to reduce 1 µmol of NADH per min under the defined conditions. In order to purify the TMU oxidoreductase, CCE from CBB1 grown on YNB-supplemented M9-caffeine medium was used as the source of the enzyme (Table 4.1). The four-step procedure shown in Table 4.1 resulted in 55-fold purification of the enzyme relative to the activity in cell extracts. The specific activity of the purified enzyme was 4128 mU per mg protein (Table 4.1).

Table 4.1. Purification of trimethyluric acid oxidoreductase from *Pseudomonas* sp. strain CBB1.

Purification step	Total protein (mg)[a]	Total activity (mU)[b]	Specific activity (mU/mg)	Purification (fold)	Yield (%)
Crude cell extract	903	66806	74	1	100
Q-Sepharose	83	33703	407	5	50
Butyl-Sepharose	20.3	24916	1228	16	37
Hydroxyapatite	0.97	3478	3577	48	5
Phenyl-Sepharose HP	0.03	116.4	4128	55	0.17

[a]Determined by using the Bradford assay (Bio-Rad) with bovine serum albumin as standard.

[b]One unit of activity was defined as the amount of protein required to oxidize 1 µmol of NADH per minutes under the reaction conditions as described in the Materials and Methods (Chapter 2).

The purified enzyme was dark yellow in color, suggesting the presence of a tightly bound cofactor. However, this purification scheme resulted in a low recovery, about 0.2% (Table 4.1). This might be due to loss of enzyme activity and protein precipitation observed during the four-step purification process. The yield loss may also be due to high salt content of the buffers used. Thus, every attempt was made to reduce the residence time of the enzymatic fractions in high-salt buffers by promptly exchanging to 50 mM KPi. As a result, enzyme activity was stable for at least 24 hours at 4°C when stored in 50 mM KPi buffer and indefinitely at -80°C. Freshly purified TMU oxidoreductase was subjected to preliminary biochemical characterization (SDS-PAGE gel analysis, native molecular weight determination by Gel filtration, and N-terminal sequencing) without storage in order to minimize any loss and or oligomerization.

Preliminary Biochemical Characterization of TMU-Oxidoreductase

Freshly purified TMU-oxidizing enzyme from Phenyl Sepharose was loaded onto a Sephacryl S-300 gel filtration chromatography column for molecular mass estimation. The enzyme eluted as a single peak with a minor shoulder towards the end (Fig. 4.2). The native molecular weight of the enzyme was estimated to be about ~45-kDa based upon the elution volume (*Ve*) compared with the *Ve* of various protein standards of known sizes as described in Chapter 2. All fractions of TMU-oxidoreductase (including CCE, purified samples from each of the columns (Table 4.1), and Sephacryl S-300) were subjected to PAGE under denaturing conditions (Fig. 4.3).
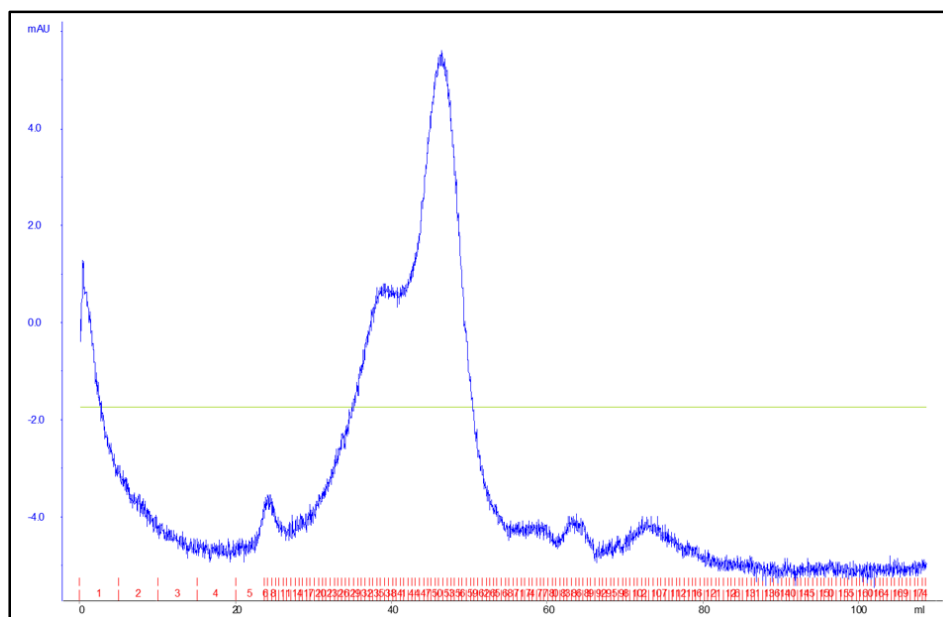
Figure 4.2. Elution profile of purified TMU oxidoreductase on Sephacryl S-300 showing absorbance at 280 nm A_{280} (mAU, left axis).

Purified enzyme fractions from Phenyl Sepharose and Sephacryl S-300 columns resolved into one major band with an apparent molecular weight of ~43-kDa, with a faint minor band at a higher molecular weight (Lane 4 and 5 of Fig. 4.3). The major band was assigned as TMU oxidoreductase and the faint minor band as a co-eluted impurity. This also explains the presence of the shoulder on the peak, which eluted from the gel filtration column (Fig. 4.2). This suggests TMU oxidoreductase is monomeric. Purification of enzymes to homogeneity from native organisms using multiple steps of conventional protein purification techniques is laborious and often low-yielding. Hence, it is desirable to clone the corresponding gene and purify the protein in much larger quantities.
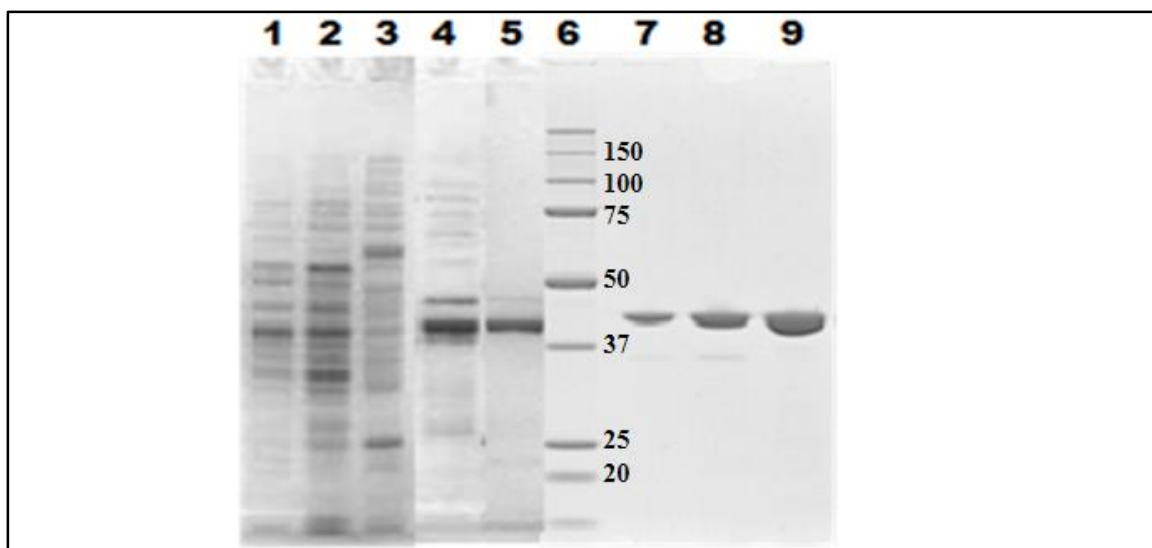
Figure 4.3. SDS-PAGE analysis of purified TMU oxidoreductase from CBB1, Lane 1-5, and the purified recombinant TMU oxidoreductase (TmuM-His$_6$), Lane 7-9. Molecular masses of markers (in kDa) are shown on Lane 6. Lane 1, Crude cell extract; Lane 2, Q-Sepharose; Lane 3, Blue-Sepharose (8 μg of protein); Lane 4, Phenyl-Sepharose HP (8 μg of protein); Lane 5, Sephacryl-S300 (8 μg of protein); Lane 6, molecular weight standards; Lane 7-9, 4, 8 and 12 μg of purified recombinant TmuM-His$_6$, respectively. Proteins were stained with Coomassie brilliant blue R250.

Cloning can be accomplished by determining the amino-terminal (N-terminal) sequence of the purified protein, based on which degenerate primers can be designed (forward primer). Reverse primer can be designed based on homology of this protein to other proteins in the GenBank (Fig. 4.4). Based on this, partial gene sequence can be obtained, which will help in designing specific forward and reverse to get full-length gene sequence of the native organism (Table A2 of Appendix A). The sequence of first 30 amino acids (after methionine) of the N-terminus of TMU oxidoreductase was determined by the Edman degradation method [Edman *et al.*, 1950] (described in Chapter-2) as "SRPLRVTIIGAGIGGLSAAVALRKIGADVT".

```
TmuM        --------------------MSRPLRVTIIGAGIGGLSAAVALRKIGADVTVVERAPE 38
ACF60813.1  -------------------------MKAIVIGAGIGGLSAAVALKQSGIDCDVYEAVKE 34
EDZ47364.1  --------------------MELKGLKITVIGAGIGGLTAALALRRQGAAVTVLEQAEA 39
EAQ44903.1  MSET------------QQAAMTLRHLNITIIGAGIGGLAVALVLRRKGAQVTVLEQSEA 47
AAZ62959.1  MPTAIETRSADGAPPTGRKQADNMTSLKIGINGAGIGGLAAAIALRKLGMDVTVYEQAAR 60
CAJ95547.1  MP---------------------LKIGINGAGIGGLAAAIALRKLGMDAIVYEQAPR 36
TmuM        LRAAGAGICMWPNGAQALHALGIANPLEMVSPILHRVCYRDQH-GRVIREMSIDKLTELV 97
ACF60813.1  IKPVGAAISVWPNGVKCMAHLGMGDIMETFGGPLRRMAYRDFRSGENMTQFSLAPLIERT 94
EDZ47364.1  ISEVGAGLQITPNGVAVLKALGLADDLAWCSQRARAVVLRGHRRGNEVLRLDLDEYAAG- 98
EAQ44903.1  ISEVGAGLQVTPNGVAVLTALELNDALAWSSQRAKAVVLREHRQGREVLRLDLDQYAAD- 106
AAZ62959.1  FARVGADINLTPNAVRALDGLGVGEALRETAARPTHRISRTWDTGEETSRLPMSDEAEQR 120
CAJ95547.1  FARVGADINLTPNAVRVLDGLGIGDALRETAARPSHRISRTWDTGEETSRLPMQKDAERR 96
TmuM        G-QRPFPLARSDLQAALLSRLD--PALVRLGGACVSVEQDANGVRAVLDDGTEIASDLLV 154
ACF60813.1  G-SRPCPVSRAELQREMLDYWG--RDSVQFGKRVTRCEEDADGVTVWFTDGSSASGDLLI 151
EDZ47364.1  --LQYYFVHRSDLVGILAGAARREGVQVRLLQKVERVEPGPQ-PVVHLGNGAQCGGDLVI 155
EAQ44903.1  --LRYYFVHRADLIKILADAAREAGVQVRLLQKVERVETGPK-PMVHLANGAQCGGDLVI 163
AAZ62959.1  YGAPQLTMHRADLMTALEGALP--AANVKLGHKAVAIERNDNGTTVRFADGGEDKVDVLV 178
CAJ95547.1  YGAPQLTMHRADLMTALEAAVP--AECVRLGHKAVAIEPHADSATLRFASGAEEERVDVLV 154
TmuM        GADGIRSVVRNHVTGGTDRLRYH-YTTWLGLVSFG----LNLTPPGTFTFHVQDSKRVGL 209
ACF60813.1  AADGSHSALRPWVLGFTPQRRYAGYVNWNGLVEID----EALAPGDQWTTFVGEGKRVSL 207
EDZ47364.1  GADGLHSKTRAALNGAD-KPVFTGQVAWRATVPNH-----LNLPPEAQLF-MGPGRHLVA 208
EAQ44903.1  GADGLHSKARQALNEVS-QPFFTRQVAWRAIVPNH-----SKLPAEAQVF-MGPGRHLVA 216
AAZ62959.1  GADGIHSVVRTALFGQE-SPIFTGVVAYRAVVPAERLAGVPNLNAFTKWWGTDPTSQIVT 237
CAJ95547.1  GADGIHSTVRTALFGQE-SPIFTGVVAYRAVVPAERLAGVPNLGAFTKWWGPDAATQIVT 213
TmuM        LNVGDDRLYFFFDA-VPSGEANP----DGVRAELRHHFDGWCSEVTTLVEALDEAKTNRL 264
ACF60813.1  MPVSAGRFYFFFDVPLPAGLAEDR---DTLRADLSRYFAGWAPPVQKLIAALDPQTTNRI 264
EDZ47364.1  YPLRDGSLVNLVAVQERRAWAEEGWNLQDDPANLRAAFAGFGGSAAALLEAVEEVLLWGL 268
EAQ44903.1  YPLRDGSMINLVAAQERRAWAEEGWHQPDDPENMRRAFAGFGGEAQQLLAQVQDCALWGL 276
AAZ62959.1  FPLNRGRDIFIFATVAQESWRNESWTTPGRVEDLRSAYAGFHPEARALLDACDDVLISAL 297
CAJ95547.1  FPLNRGRDIFIFATVAQEAWRHESWTTPGRVQDLRSAYAGFHPEARALLDACDEVLISAL 273
TmuM        PVHDLDPLASFVNGRIVLIGDAAHATTPTLGQGGALAMEDSLVLARHLAEST-----DYG 319
ACF60813.1  EIHDIEPFSRLVRGRVALLGDAGHSTTPDIGQGGCAAMEDAVVLGAVFRQTR-----DIA 319
EDZ47364.1  FRH--PVAARWQGGNSAILGDAAHPTLPFMAQGANLALEDAWVLARALQEA-----DSAG 321
EAQ44903.1  FRH--PVANQWHRGAVALLGDAAHPTLPFMAQGGNLALEDAWVLGEALEAVP----DDIP 330
AAZ62959.1  YVR--DPLPTWSDGHVTLMGDACHPMMPFMAQGAGMAIEDGVVLARCLADAAQDGYAAVP 355
CAJ95547.1  YVR--DPLPAWCAGPVTLMGDACHPMMPFMAQGAGMAIEDGVVLARCLADAAPGGAAGVP 331
TmuM        SALASYDNERLMRTRQVVLASRARTAATLG-IDNTSAQTWQKQLTDDASQDFLEQLVDIH 378
ACF60813.1  AALCEYEAQRCDRVRDLVLKARKRCDITHG-KDMQLTEAWYQELREETGERIINGMCDTI 378
EDZ47364.1  AGLALYQSRRLERARKVVQAASRNAWKYHLRAPLSWPAHQVLKLGGRVAPDKMVRQFDWI 381
EAQ44903.1  SALARYQLLRRERAVRVVDAASRNAWKYHLRAPLNWPAHQLLKLGGYLAPQRVVSQFDWI 390
AAZ62959.1  SALARYQRARHERTSRIQIGSRSNAW--------------LKEGG---------NADWV 391
CAJ95547.1  AALARYQAARHERTSRIQIGSRGNAW--------------LKEGG---------NADWV 367
TmuM        RAGPLAAWPHDRQEGVAT 396
ACF60813.1  LSGPLG------------ 384
EDZ47364.1  YRHDVTA----------- 388
EAQ44903.1  YRHDVTQQN--------- 399
AAZ62959.1  YDYDAWNVALG------- 402
CAJ95547.1  YDYDAWNVALG------- 378
```

Figure 4.4. Multiple sequence alignment of (1) TMU oxidoreductase from CBB1 (30
amino acids of N-terminal protein sequence determined from Edman
degradation are in bold and italic), (2) ACF60813.1: HpxO [*Klebsiella
pneumoniae*], (3) EDZ47364.1: FAD-binding monooxygenase
[*Rhodobacterales bacterium Y4I*], (4) EAQ44903.1: salicylate hydroxylase
[*Roseobacter* sp. MED193], (5) AAZ62959.1: salicylate 1-monooxygenase
[*Ralstonia eutropha* JMP134], and (6) CAJ95547.1: salicylate hydroxylase
[*Ralstonia eutropha* H16]. Consensus sequences used for degenerate primers
are highlighted in dark grey.

A BLASTP [Altschul *et al.*, 1997] search of the GenBank [Benson *et al.*, 2008] database using the above N-terminal amino acid sequence as query showed significant homology (about 70% identity) with HpxO from *Klebsiella pneumoniae* (GenBank accession no. ACF60813), FAD-binding monooxygenase from *Rhodobacterales bacterium* Y4I (GenBank accession no. EDZ47364), and salicylate 1-monooxygenase from *Rastonia eutropa* JMP134 (GenBank accession no. AAZ62959) (Fig. 4.4).

Cloning, Expression and Purification of TMU-Oxidoreductase

PCR reaction using primers designated tmuM-degF1 and tmuM-degR1 successfully amplified a single, 900 bp PCR product from the fosmid DNA pMVS848 (Table A1 of Appendix A). A BLASTX search using the DNA sequence of this 900-bp product showed that this amplified partial gene has significant homology to FAD-binding monooxygenase family of proteins (Fig. 4.4). Further sequencing of the flanking region of this partial gene revealed a complete open reading frame (ORF) with 1191 nucleotides. The gene was designated as *tmuM* and it encoded a 396-amino-acid protein with a theoretical M_r of 42,619 and pI of 6.12. N-terminal sequence derived from the *tmuM* gene matched the experimentally determined N terminus of the purified TMU-oxidizing enzyme from CBB1. The amplified gene was cloned into the pET32a plasmid with His$_6$-tag at C-terminal and expressed in soluble form in *E. coli* BL21(DE3). The over-expressed protein was purified to electrophoretic homogeneity using immobilizied-metal affinity chromatography and Q-Sepharose chromatography (Fig. 4.3 Lane 7-9). The specific activity of the recombinant TMU oxidoreductase was 9,700 mU per mg, which is

twice that of the purified wild-type enzyme. Given the higher activity of the recombinant

enzyme, this preparation was used for further characterization.

Biochemical Characterization of Recombinant TMU Oxidoreductase

The recombinant TMU oxidoreductase was similar to the wild-type enzyme

purified from CBB1 in terms of native (~45-kDa) and sub-unit molecular weights (~43-

kDa), and UV-Visible absorption spectrum (Figs. 4.3 and 4.5). Both the native and

recombinant enzymes were monomeric with UV-Visible absorption spectrum maxima at

271 nm, 380 nm, and 456 nm, with a shoulder at 485 nm (Fig. 4.5), which is

characteristic of flavoproteins [Nakamura *et al.*, 1963; Megerle *et al.*, 2011].
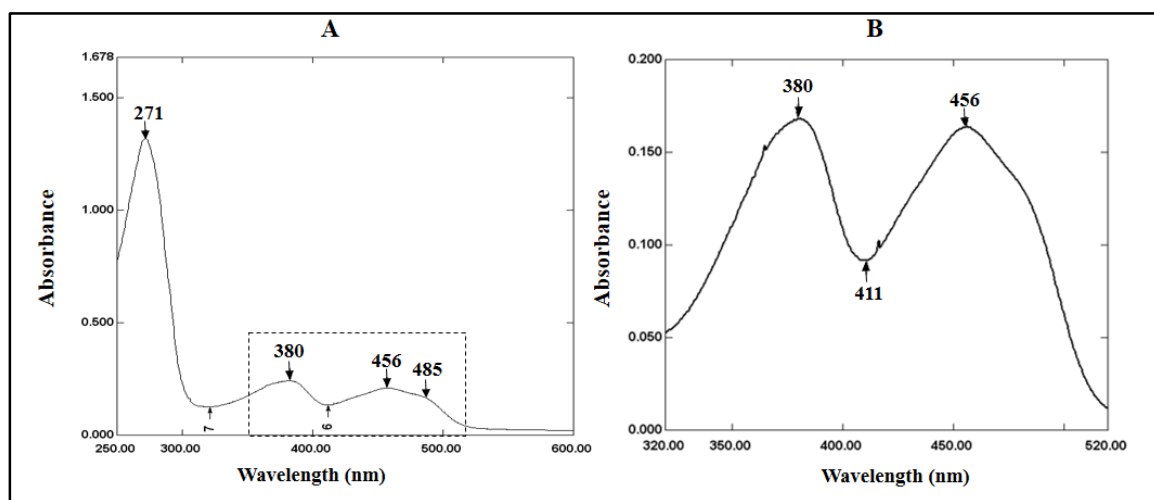
Figure 4.5. UV/visible spectrum of purified TMU-oxidoreductase (native and
recombinant). [A] Full spectrum exhibited three peaks at 271, 380, 456, and
485 nm, characteristic of flavoproteins. [B] Magnification of the spectrum
from 300-500 nm.

HPLC analysis of the cofactor released from the recombinant enzyme upon heat

denaturation was identified as FAD with enzyme:FAD stoichiometry 1:1 (Fig. 4.6).

Previously, it was shown that the N-terminal amino acid sequence of TMU-oxidizing

enzyme showed homology with various FAD dependent monooxygenases including

HpxO (Fig. 4.4). Based on the NADH-dependent activity, color, UV-visible spectrum,

HPLC-based cofactor analysis, and homology of TMU degrading enzyme, it was

concluded that this enzyme is a FAD-dependent TMU oxidoreductase. However, further

characterization is needed to determine whether this enzyme is a monooxygenase.
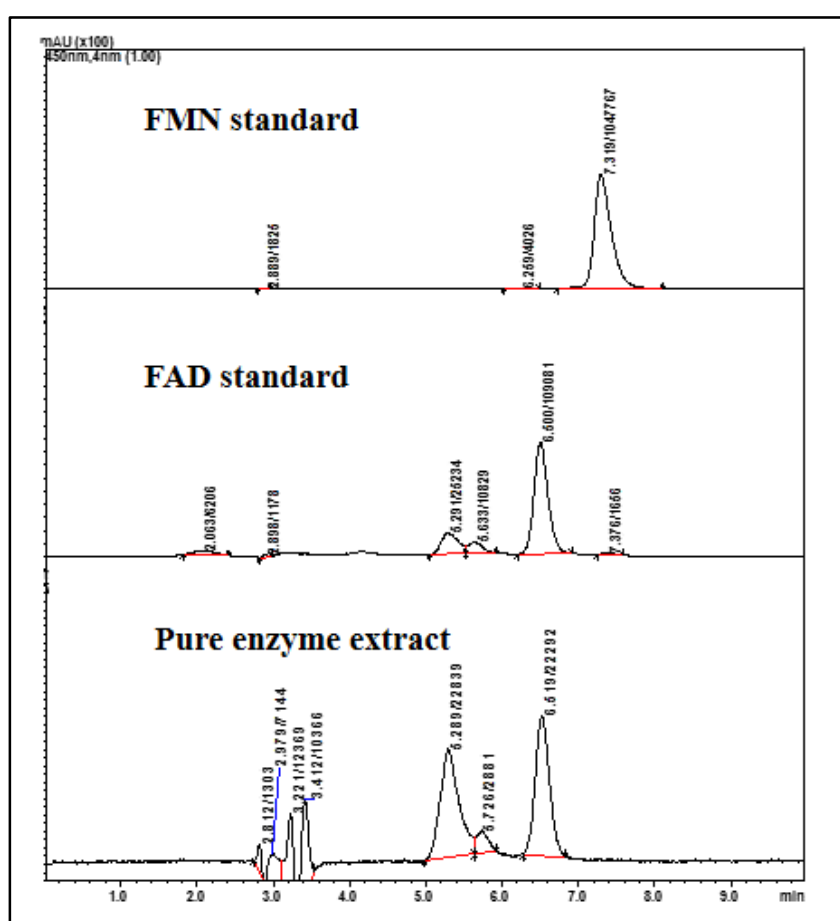
Figure 4.6. HPLC analysis of cofactor released from purified TMU oxidoreductase.

Optimum pH and Temperature of TMU Oxidoreductase

Activity of the enzyme was tested in the pH range 7.0 to 11.0; maximum activity was at 9.0 (Fig. 4.9A). Activity increased at a rapid rate from pH 7.0 (12.5% of highest activity) up to 9.0. Beyond pH 9, the enzyme activity decreased at a much slower rate with almost 70% activity even at pH 11.0.
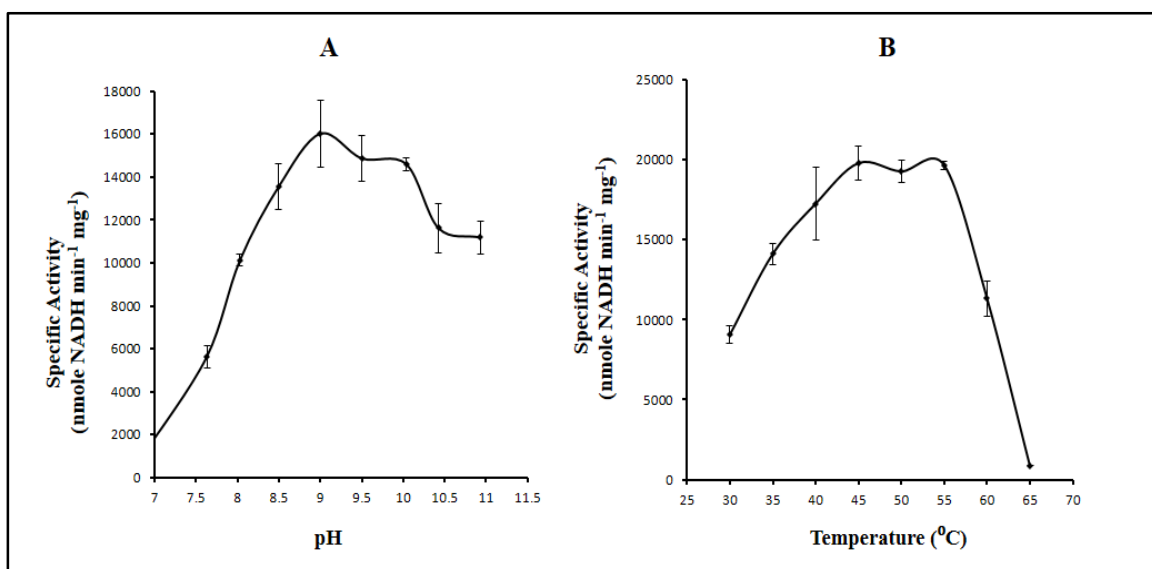
Figure 4.7. Relative activity of TMU oxidation by TMU oxidoreductase at different pH [A] and temperatures [B]. Error bars show standard deviations calculated from three independent assays.

Similarly, TMU oxidoreductase was tested for activity in the range 25°C to 70°C. Maximum TMU oxidation activity was observed at 45°C to 55°C (Fig. 4.7B). Activity steadily increased from 25°C (~45% of highest activity) to 45°C and plummeted beyond 55°C to almost zero at 65°C.

Oxygen Requirement of TMU Oxidation Reaction and Characterization of TMU

Oxidoreductase as a Monooxygenase

When TMU oxidation reaction was carried out in a chamber pre-flushed with N_2

gas, no TMU degradation or NADH-oxidation was observed (experimental set-up in

Chapter-2: Isotopic oxygen incorporation experiment). Upon introduction of $^{18}O_2$ into

the reaction chamber, activity was restored immediately, and was monitored

spectrometrically (Fig. 4.8A). Complete TMU conversion was observed in 15 min, as in

previous cases (Fig. 4.1A). Thus, it was concluded that TMU oxidation reaction by TMU

oxidoreductase is oxygen dependent.
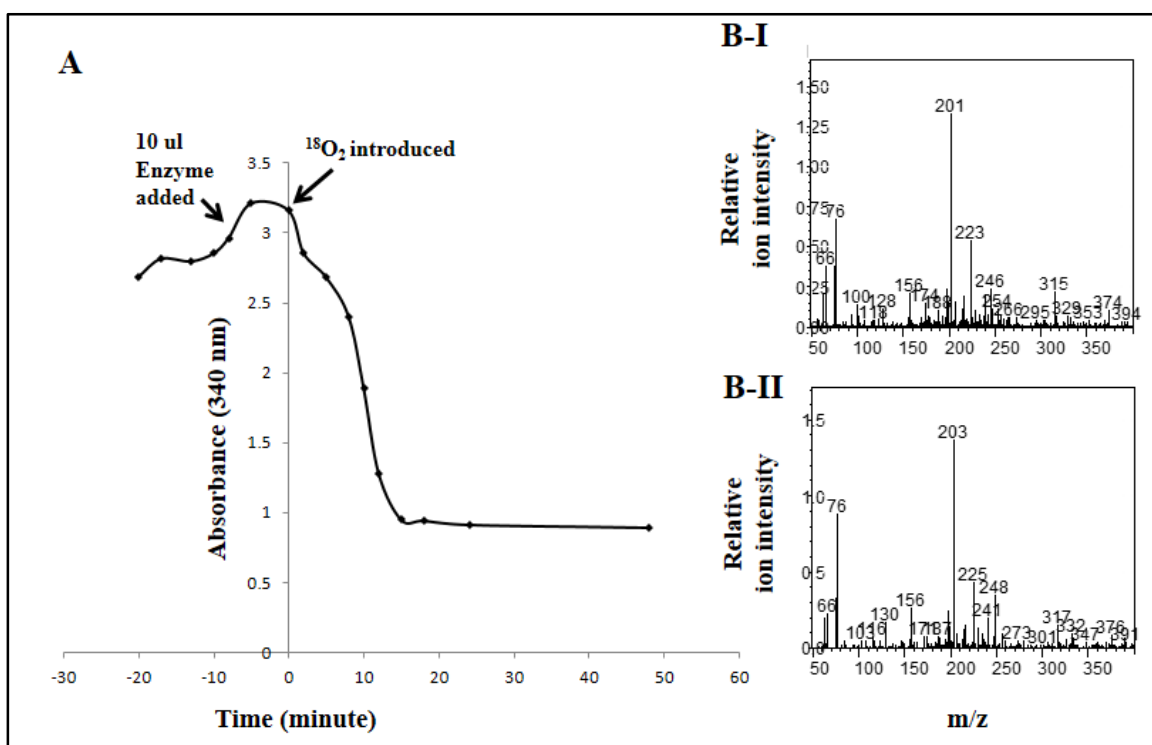
Figure 4.8. TMU oxidation reaction carried out in oxygen-controlled environment to incorporate isotopic oxygen ($^{18}O_2$) into the product (TMA). [A] Monitoring TMU degradation without and in the presence of $^{18}O_2$ spectrometrically. [B] ESI-LC/MS analysis of reaction product showing [I] ^{16}O-incorporated TMA and [II] ^{18}O-incorporated TMA.

Analysis of the TMU oxidation reaction using ESI-LC/MS confirmed that TMA is the stable end product of this reaction. TMA produced in the presence of atmospheric O_2 had an *m/z* of 201, which agreed with the $(M+1)^+$ ion of the compound (Fig. 4.8B-I). TMA produced in the presence of $^{18}O_2$ had an *m/z* of 203 indicating incorporation of one atom of isotopic oxygen (Fig. 4.8B-II). The incorporated oxygen atom comes from molecular oxygen (O_2). Based on these results, the TMU oxidoreductase was reclassified as TMU monooxygenase (TmuM).

Stoichiometric Analysis of TMU Oxidation Reaction

The stoichiometry of TMU oxidation by TmuM was established in two separate experiments, i.e., oxygen measurement and NADH consumption. In the first experiement, the stoichiometry of oxygen in the reaction was demonstrated by monitoring the oxygen consumption in a Clark-type oxygen electrode and simultaneously quantifying TMU disappearance by HPLC. It was observed that after 200, 300, and 500 seconds of reaction, 74.6, 190.5, and 275.4 nmole of O_2 was consumed while 78.5, 185, and 274 nmole of TMU disappeared, respectively (Fig. 4.9A). This indicates that approximately 1 molecule of O_2 was consumed per molecule of TMU oxidized. In the second experiment NADH to TMU stoichiometry was established by monitoring NADH consumption spectrometrically with simultaneous quantification of TMU disappearance by HPLC. In this case, after 200, 300, and 500 seconds of reaction, 20.5, 41.8, and 34.4 nmoles of NADH was consumed while 21.7, 43.3, and 37.8 nmoles of TMU disappeared, respectively (Fig. 4.9B). This established the NADH: TMU stoichiometry as 1:1. Overall TMU: NADH: O_2 stoichiometry of TmuM was established as 1:1:1.
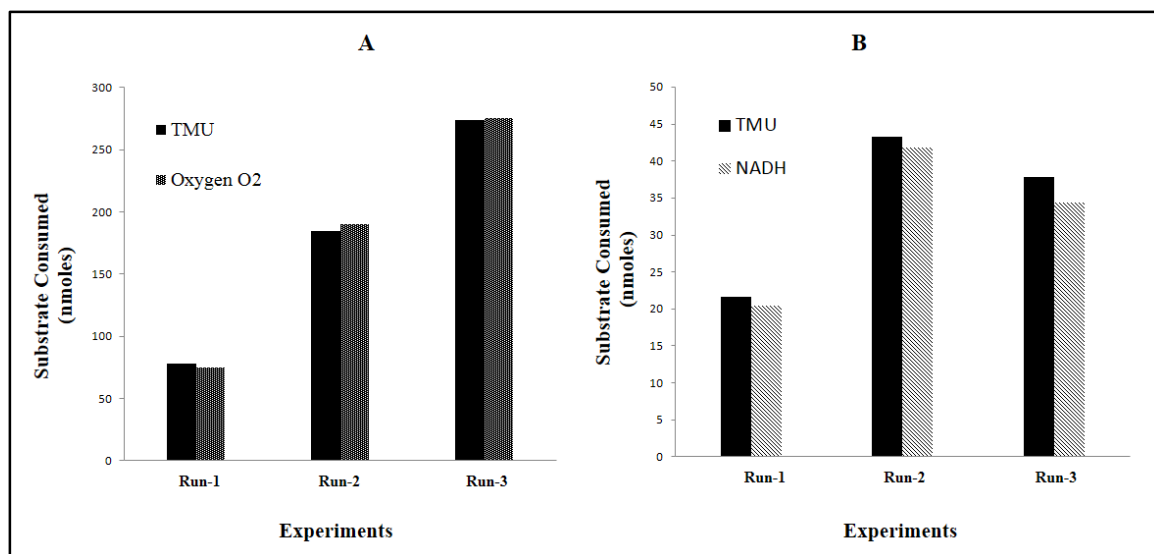
Figure 4.9. Stoichiometric oxidation of TmuM reaction. (A) Consumption of oxygen and disappearance of TMU and (B) NADH oxidation vs. TMU disappearance. The bars represent average of three separate experiments. Run-1 was 200 seconds, Run-2 was 300 seconds, and Run-3, 500 seconds. The amount of TmuM was the same in both experiments (except Run-3 of Set-B, which was lower).

Enzyme Kinetics and Substrate Specificity of TmuM

The apparent kinetic parameters (K_M and k_{cat} values) of the recombinant TmuM were determined using various tri-, di-, monomethyluric acids and uric acid as substrates over a range of 0.5 μM to 500 μM (Table 4.2). The apparent K_M for TMU and k_{cat} of the recombinant TmuM in 50 mM KPi (pH 7.5) at 30°C were 10.2 ± 2.2 μM and 448.9 ± 21.7 min^{-1}, respectively (Table 4.2, ± represents S.D., n=3). 1,3-Dimethyluric acid had an apparent K_M and k_{cat} of 126.5 ± 29.3 μM and 185.0 ± 16.4 min^{-1} respectively (Table 4.2). In contrast, the K_M values of 3,7-dimethyluric acid and 1-methyluric acid were much lower (1.3 ± 0.6 μM and 1.2 ± 0.5 μM respectively).

Table 4.2. Kinetic parameters and substrate specificity of TmuM-His$_6$.

Substrate	Average ± SD[a]		
	k_{cat} (min^{-1})	K_m (µM)	k_{cat}/K_m (min^{-1} µM^{-1})
Trimethyluric acid	448.9 ± 21.7	10.2 ± 2.2	44.1 ± 2.1
1,3-dimethyluric acid	185.0 ± 16.4	126.5 ± 29.3	1.5 ± 0.1
3,7-dimethyluric acid	118.2 ± 8.4	1.3 ± 0.6	89.4 ± 6.3
1-methyluric acid	29.1 ± 2.4	1.2 ± 0.5	24.5 ± 2.0
Uric acid	NA		NA

[a] Average and standard deviations were derived from three independent assays. Initial reaction rates were determined by following substrate-dependent NADH oxidation at 340 nm. NA, no activity.

Initial assays with wild type TmuM and recombinant TmuM-His$_6$ suggested that the enzyme was active on TMU and other methyluric acids, but not on uric acid (data not shown). The catalytic activity k_{cat} (turnover number) of TmuM-His$_6$ for TMU was found to be 2.5, 3.8, and 15.4 times higher than that of 1,3-, 3,7- and 1-methyluric acids, respectively (Table 4.2). These results suggest that TMU is the preferred substrate for TmuM, followed by 1,3-dimethyluric acid, 3,7-dimethyluric acid and 1-methyluric acid. In contrast, the catalytic efficiency k_{cat}/K_M for 3,7-dimethyluric acid was 2 times higher than that of TMU due to an order of magnitude lower K_M. 1-Methyluric acid was the least preferred substrate, although it's K_M was only an order of magnitude lower than that of TMU (Table 4.2). No activity was observed with uric acid, which makes TMU functionally different from other known FAD-dependent uricase like HpxO.

Time course of TMU Oxidation by TmuM to Identify the True Reaction Product

Time course of TMU oxidation and product formation catalyzed by recombinant TmuM-His$_6$ was monitored by LC-MS (Fig. 4.10 and 4.8B-I). Initial HPLC analysis showed that TMU was completely consumed while NADH was converted to NAD$^+$ (Fig 4.1). In Chapter 3 of this document, rigorous metabolite analysis (including time course) of TMU degradation by CBB1 resting cells and crude cell extracts revealed the presence of substrate TMU, metabolite II, and metabolite IV (Fig. 3.4 and 3.5) with (M+1)$^+$ ions at $m/z = 211$, $m/z = 227$ (Fig. 4.10A) and $m/z = 201$ (Fig. 4.8B-I), respectively. The observed molecular masses of 226 (with $m/z = 227$) for metabolite II and 200 (with $m/z = 201$) for metabolite IV matched with that of the TM-HIU- and TMA, respectively. Corresponding non-methylated metabolites have been reported for the uric acid pathway [de la Riva *et al.*, 2008; Pope *et al.*, 2009; Ramazzina *et al.*, 2006].

In the product analysis of the TmuM catalyzed reaction, metabolite II at $m/z = 227$ could be detected initially, but it disappeared after 10 min (Fig. 4.10B). In contrast, TMA at $m/z = 201$ continued to accumulate until the end of the reaction (Fig. 4.10B). Based on these results, and the observation of 5-hydroxyuric acid in the uric acid degradation pathway [Kahn *et al.*, 1997], metabolite II (TM-HIU) was proposed as the true product of TMU monooxygenase reaction. Metabolite II (TM-HIU) was unstable under the reaction condition and spontaneously converted to TMA (possibly via metabolite III), as is the case in the uric acid degradation pathway [de la Riva *et al.*, 2008; Pope *et al.*, 2009: Ramazzina *et al.*, 2006]. Metabolite III (TM-OHCU) could not be detected in the enzymatic reaction as well.
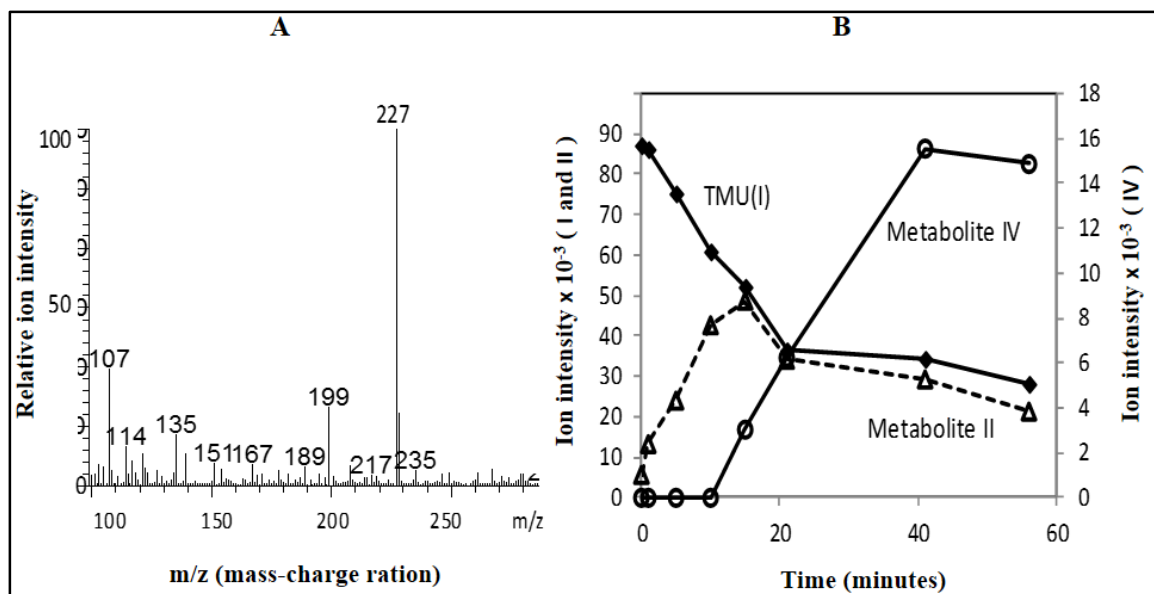
Figure 4.10. Identification of product of TMU oxidation in CBB1. [A] Mass spectra of metabolite II formed from TMU by recombinant TmuM-His$_6$. [B] Time course of SIM/LCMS analysis of TMU oxidation and product formation by TmuM-His$_6$. Ion intensities of metabolite I and II are shown in the primary y-axis. Ion intensity of metabolite IV is shown in secondary y-axis. TMU (I) (■), TM-HIU (II)(Δ) and TMA (IV) (○).

Discussion

TmuM: A Novel Trimethyluric Acid Monooxygenase from CBB1

In Chapter 3, based on the metabolite analysis of CBB1 growth media, resting cell, and crude cell extract, TMU conversion to 3,6,8-trimethylallantoin (TMA) via 1,3,7-trimethyl-5-hydroxyisourate (TM-HIU) and 3,6,8- trimethyl-2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (TM-OHCU) (see Fig. 3.14 of Chapter 3) was proposed. TM-HIU was identified as a possible metabolite based on *m/z* data obtained from LC-MS analysis. TMA was purified and structure fully established by various analytical methods. The thrust of subsequent work was to fully characterize TMU oxidoreductase and confirm

hydroxylation of TMU at the C-5 position to produce TM-HIU, as hypothesized
previously (Fig. 3.14).

TMU oxidoreductase was found to be a NADH-dependent flavoprotein (Fig. 4.1).
This monomeric (Fig. 4.2), single subunit protein (Fig. 4.3) had 1 molecule of bound
FAD as cofactor (Fig. 4.5 and 4.6). The enzyme was also found to be oxygen dependent
(Fig. 4.8). LC-MS analysis of the TMU oxidation reaction by the purified enzyme in
presence of oxygen and NADH, yielded a product of m/z 227 (hypothesized to be TM-
HIU, Fig. 4.10). This product was unstable. Likewise, 5-HIU formed in the uric acid
pathway was also unstable [Kahn *et al.*, 1997; Ramazzina *et al.*, 2006]. Hence it was
concluded that TM-HIU is the product of this enzyme. Since the reaction required
molecular oxygen, the enzyme was reclassified as trimethyluric acid monooxygenase
(TmuM) (Fig. 4.11).
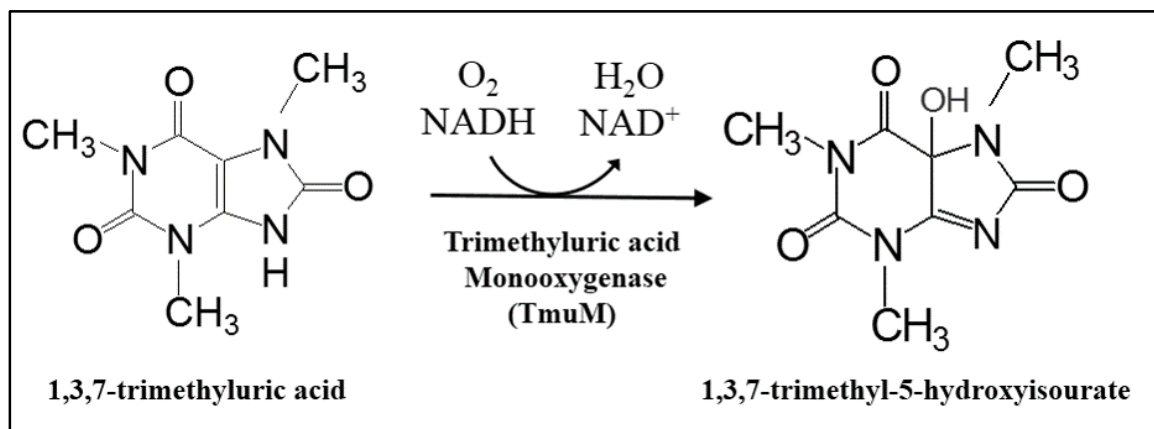
Figure 4.11. TMU oxidation to TM-HIU by a novel FAD-containing NAD(P)H-
dependent trimethyluric acid monooxygenases (TmuM) from *Pseudomonas*
sp. strain CBB1.

TM-HIU is the True Product TMU Oxidation in CBB1

A time course of TmuM catalyzed TMU oxidation, analyzed by ESI-MS in the positive ion mode, revealed gradual disappearance of TMU, transient accumulation of an unstable metabolite (with MW similar to TM-HIU), and continuous accumulation of TMA (Fig. 4.10B). Also, in the isotopic oxygen incorporation experiment one atom of molecular oxygen got incorporated into TMU, possibly at the C-5 position forming TM-HIU, which was detected in TMA (Fig. 4.8). The incorporated oxygen atom in TMU is retained in TMA during the ring opening and subsequent decarboxylation reaction (TM-OHCU to TMA) (Fig. 4.12). All these data agree with the above mentioned reaction scheme proposed for the TmuM reaction (Fig. 4.11).
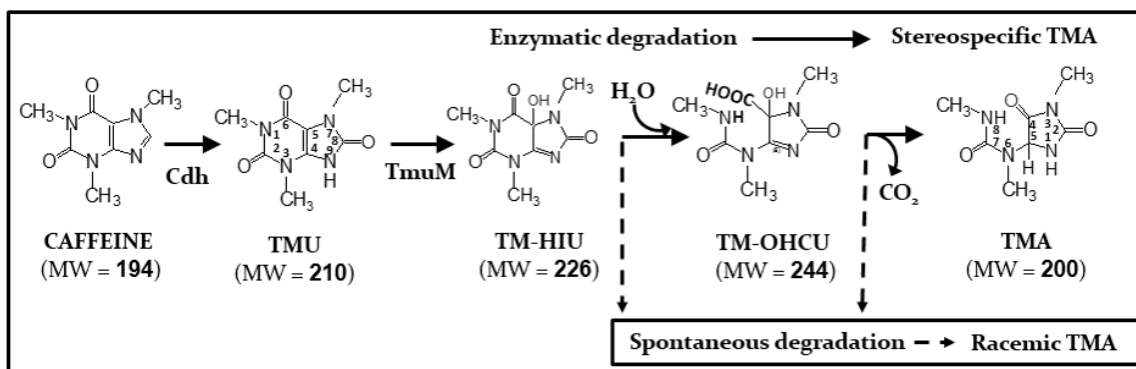
Figure 4.12. Diagrammatic representation of the C-8 oxidation pathway showing enzymatic conversion of caffeine to TMU by Cdh, and TMU to TM-HIU by TmuM, and enzymatic and spontaneous conversion TM-HIU to TMA *via* TM-OHCU.

In the ^{18}O-incorporation experiment, isotopic labeling of TM-HIU was not detected. This is because it was an end-point reaction (Fig. 4.8 B-I and B-II) which

resulted in spontaneous ^{18}O-TM-HIU decomposition to ^{18}O-TMA (Fig. 4.10). Nevertheless, the retention of isotopic oxygen in TMA (Fig. 4.8B-I) conclusively established TM-HIU as the true product of TmuM (Fig. 4.11).

Rationale for TM-HIU Decomposition to TMA via TM-OHCU

In the hypothesis for C-8 oxidation pathway (Fig. 4.12), TM-HIU degradation to TMA was proposed to be via TM-OHCU. This hypothesis was based on the uric acid degradation reaction [Ramazzina *et al.*, 2006] where oxidation of uric acid by urate oxidase results in transient accumulation of 5-hydroxyisourate (HIU) [O'Leary *et al.*, 2009]. HIU formed from uric acid by the traditional urate oxidase (not a monooxygenase) from soybean root nodules, was also shown to be unstable, with a half-life of less than 30 min at neutral pH [Kahn *et al.*, 1997; Kahn *et al.*, 1998; Kahn *et al*., 1997]. This unstable compound spontaneously hydrolyzed to 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (OHCU). OHCU has been reported to be unstable under normal physiological conditions as well, and gets further decarboxylated spontaneously to racemic allantoin (stable end product of urate oxidase reaction) [Ramazzina *et al.*, 2006]. The instability of TM-HIU can be rationalized in a similar way, given the structural similarity to HIU.

TmuM Belongs to a New Family of FAD-Dependent Monooxygenases

The gene encoding TmuM (*tmuM*) is 1191 nt long. Sequence analysis of *tmuM* (Fig. 4.4) revealed that this enzyme belongs to the family of FAD containing aromatic-ring hydroxylases which includes FAD-dependent urate oxidase (HpxO, GenBank:

ACF60813) from *Klebsilella pnemoniae* [O'Leary *et al.*, 2009], a flavin dependent

hydroxylase (PhzS, GenBank: AAG07605) from *Pseudomonas aeruginosa* [Greenhagen

et al., 2008], salicylate 1-monooxygenase (GenBank: AAZ62959), and other

uncharacterized/putative FAD-binding monoyxgenases (GenBank: EAQ44903,

CAJ95547, EDZ47364). A multiple sequence alignment of these homologous proteins

revealed that primary sequence of TmuM contains conserved residues that include an

FAD binding domain and NADH-binding domain (Fig. 4.4). Plus, a phylogenetic

analysis of these seven homologous flavoprotein monooxygenases (Fig. 4.13) as

expected, revealed TmuM and HpxO clustered together in a clade, distinct from the other

five flavin monooxgenases. TmuM and HpxO appear diverging from a common

ancestor. However, it is worth noting that *hpxO* (the gene encoding HpxO) had 38%

sequence identity with *tmuM*,
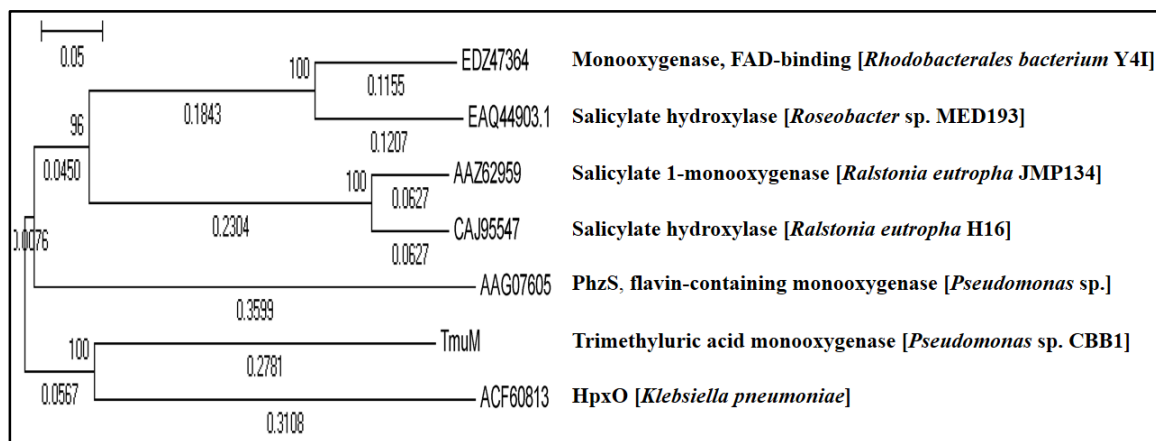
Figure 4.13. Phylogenetic analysis of TmuM from *Psedomonas* sp. CBB1 with six other
homologous proteins including flavin-containing monooxygenases and
salicylate hydroxylases (see text for details). TmuM and HpxO (ACF60813)
form a single clade, distinct from other five homologous proteins

There are two types of enzymes that catalyze reactions similar to TmuM. The first one is the most prevalent urate oxidase (EC 1.7.3.3), which does not require any cofactor [Kahn and Tipton, 1997]. The co-product of this reaction, as per the enzyme nomenclature, is hydrogen peroxide. The second type of enzyme is NADH-dependent, FAD-containing urate oxidase. This enzyme follows a catalytic mechanism similar to flavin containing aromatic-ring hydroxylases [Harayama *et al.*, 1992]. HpxO from *Kelbsiella pneumonia* [O'Leary *et al.*, 2009], and TmuM are examples of the second type. TmuM is closely related to HpxO in terms of sequence similarity (Fig. 4.4), phylogeny (Fig. 4.13), cofactor content (FAD), NADH-dependence and likely, the catalytic mechanism. However, the enzymatic conversion of urate to HIU catalyzed by HpxO, although was shown to be oxygen dependent [O'Leary *et al.*, 2009], has not been characterized as a monooxygenase. It may be mis-assigned as urate oxidase; hydrogen peroxide production was not reported as one of the products of this reaction. Thus, based on this evidence and the similarity of HpxO catalyzed reaction to TmuM, HpxO is likely a uric acid monooxygenase. Both TmuM and HpxO belong to a unique class of (methyl)urate monooxygenases.

TmuM is Not an Uricase but a Methyluric Acid Specific Enzyme

Although, TmuM and HpxO belonged to the same family of FAD-containing NADH-dependent monooxygenases, TmuM is distinct from HpxO in terms of substrate they act upon. While HpxO was identified as an uricase, TmuM has no activity with uric acid (Table 4.2). Kinetic studies of TmuM with various tri-, di-, and monomethyluric acids suggested that TmuM has the highest catalytic activity for trimethyluric acid,

compared to di-, and monomethyluric acids (Table 4.2). A comparison of the active site

cavities of TmuM and HpxO based on homology models (Fig. 4.14) reveals a much

bigger (244 ± 45 Å^3) cavity for TmuM compared to the cavity for HpxO (121 ± 10 Å^3).

Also, the grand average of hydrophathicity (GRAVY) value based on the Kyte and

Doolittle hydropathy scale [Kyte *et al.*, 1982] for the residues lining the putative active

site (Table in Fig. 4.14) reveals a much more hydrophobic cavity for TmuM (0.673) (Fig.

4.14IIA) compared to HpxO (-0.218) (Fig. 4.14IIB).
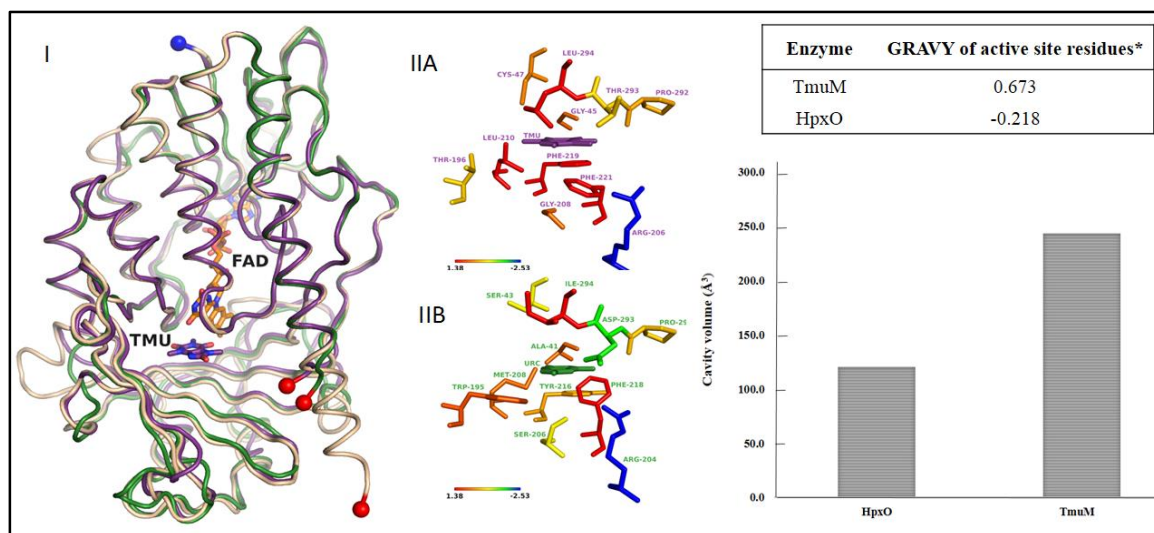
Figure 4.14. Homology models for TmuM and HpxO. (I) The best homology models for TmuM (purple) and HpxO (green) are superimposed with the modeling template PhzS (PDB ID: 3C96, wheat). Blue and red spheres indicate the N- and C-terminal regions of the protein that were either observed in the X-ray structure or could be modeled reliably. FAD (orange sticks) is modeled into place based on PHBH (1PBE) with the isoalloxazine ring in the "in" position as would be expected in the presence of the ligand. TMU (purple sticks) is modeled in the ligand-binding pocket based on the ligand position in PHBH (1PBE). (II) Active site cavities for TmuM and HpxO homology models. Residues lining the modeled active site for TmuM (A) and HpxO (B) are shown as sticks colored by hydropathy values based on the Kyte-Doolittle scale (24) with the most hydrophobic in red and the least hydrophobic in blue. TmuM has a bigger, more hydrophobic active site cavity (trimethyluric acid modeled in in purple) than HpxO (uric acid modeled is in green).

These results are consistent with the enzyme data (Table 2) and provide an initial explanation of the specificity of TmuM to the larger TMU and 3,7-dimethyluric acid.

Summary and Conclusion

A novel caffeine dehydrogenase catalyzes the first step of caffeine C-8 oxidation pathway in *Pseudomonas* sp. CBB1. TMU has been shown to be transiently formed in the crude cell extract reactions along with TM-HIU and TMA. A novel FAD-containing NADH-dependent trimethyluric acid monooxygenase (TmuM) was purified from CBB1. TmuM is a monomeric, single subunit protein with an apparent molecular weight of ~43-kDa and 1 molecule of FAD/subunit. Optimum enzyme activity was observed at 45°C and at a pH of 9.0. TmuM activity was found to be oxygen dependent. TmuM oxidized TMU with an equimolar consumption of NADH and molecular oxygen. TmuM was found to be highly specific for methyluric acids with the highest activity with TMU. No activity was observed with uric acid. Homology model of TmuM indicated that the active site of TmuM is larger and more hydrophobic compared to a homologous FAD-containing urate oxidase/monooxygenase (HpxO).

One atom of molecular oxygen got incorporated into TM-HIU, which was shown to be retained in TMA. Plus, a time course of TMU oxidation reaction with purified TmuM demonstrated the transient formation of TM-HIU, which decomposed to TMA. This confirmed TM-HIU as the product of TMU oxidation by TmuM enzyme. Based on these results, a reaction scheme was developed for the TmuM catalyzed TMU oxidation to TM-HIU. The previously developed scheme for the C-8 oxidation pathway was refined by adding TmuM as catalyzing the second reaction of this pathway. This work

further supported the previous proposal of spontaneous conversion of TM-HIU to TMA via TM-OHCU. Overall, caffeine degradation pathway via C-8 oxidation in CBB1 has been elucidated via TMU, TM-HIU. This compound undergoes spontaneous conversion to racemic TMA *in vitro*. *In vivo*, TM-HIU is converted to S-(+)-TMA via TM-OHCU.

In order to further delineate the enzymatic conversion of TM-HIU to TMA, corresponding enzymes need to be characterized. However, this task is impeded by lack of commercial availability TM-HIU, TM-OHCU and TMA. Another approach is to take a bioinformatics approach to study the genes and gene-cluster of these reactions.

CHAPTER 5

GENE CLUSTER ANALYSIS OF CBB1 CAFFEINE C-8 OXIDATION
PATHWAY

Introduction

The work so far has shown that *Pseudomonas* sp. CBB1 degrades caffeine to
TMU, which gets converted to TM-HIU by TmuM. It has also been shown that TM-
HIU, being unstable, spontaneously undergoes degradation to racemic TMA. Our
hypothesis is, in CBB1, that TM-HIU undergoes a two-step enzymatic conversion (a ring
opening reaction followed by decarboxylation) to *S*-(+)-TMA via TM-OHCU. The bulk
of the earlier work was on full characterization of the conversion of TMU to 5-hydroxy-
TMU (TM-HIU) catalyzed by a novel flavoprotein TMU monooxygenase (Chapter 4). It
has been difficult to characterize the reactions from TMU to *S*-(+)-TMU due to the
instability of the intermediates and the lack of availability of standards.

Hence a gene-cluster analysis of the C-8 oxidation pathway was undertaken. In
this work, a 25.2-kb caffeine gene-cluster, consisting of 23 open reading frames (ORF),
was sequenced from the CBB1 genome. Homology based gene identification and
multiple sequence alignment [Chenna *et al.*, 2003] based function assignment was
performed on the gene-cluster. Genes encoding Cdh and TmuM, which were previously
sequenced, are present this cluster. Five new genes were identified which might encode
enzymes involved in the degradation of TM-HIU, TM-OHCU, TMA, and TMAA (See
Appendix E for full names) to di- and monomethyl urea and glyoxylate (in that
sequence). Further, through gene cluster analysis of this 25.2-kb genome fragment, with

hpx gene cluster(s) (responsible for hypoxanthine →xanthine →uric acid pathway) of other microbial systems supplied more information about organization of these genes of the C-8 oxidation pathway. Based on this, the complete C-8 oxidation pathway in CBB1 from caffeine to glyoxylate was elucidated.

<center>Results</center>

<center>Previous Work: Sequencing a Part of Caffeine Gene Cluster in CBB1</center>

Previously, caffeine dehydrogenase (Cdh) was purified from CBB1 and N-terminal sequence of the protein was determined [Yu *et al.*, 2008]. Using degenerate PCR primers designed from the N-terminal protein sequences of the large (90-kDa) and medium (32-kDa) subunits of Cdh, a 2.4-kb gene (*cdhA*) was amplified from the CBB1 genome [Yu *et al.*, 2008; Louie *et al.*, ASM 2009 Poster Q-271]. Using sequences derived from *cdhA* as probe to screen a genomic library of CBB1, a clone (with fosmid pMVS848) containing *cdhA* was identified. Further sequencing of pMVS848 using specific primers derived from *cdhA* revealed 2 ORFs (namely *cdhB* of 1.0-kb and *cdhC* of 0.5-kb) immediately downstream (Fig. 5.2). It was proposed that *cdhA, cdhB* and *cdhC* corresponds to the large (90-kDa), medium (35-kDa) and small subunits (20-kDa) of the Cdh enzyme. Further, upon sequencing the flanking DNA of Cdh genes, a total of 19.02-kb containing 15 ORFs, including *cdhA,B,C* and *orf1,2,and 3* genes (Fig. 5.2, Table C1 of Appendix C: gene-8-23), were obtained.

A BLASTX analysis of the three Cdh genes revealed that protein products encoded by *cdhABC* displayed significant homologies to the molybdopterin-, FAD-, and [2Fe-2S]-binding subunits, respectively, (Fig. 5.1) of various members of the

molybdenum hydroxylase family such as xanthine dehydrogenases, aerobic carbon monoxide dehydrogenases, and aldehyde dehydrogenases (Fig. 5.2). A multiple sequence alignment of amino acid sequences from each of the subunit genes *cdhA,B,C* with their corresponding homologs revealed that while CdhA contains conserved Molybdenum-cofactor-interacting-motifs, CdhB and CdhC contains a FAD-binding motif and [2Fe-2S] cluster motifs, respectively.

```
A
        Apical ligand                 Active site loop
           motif       Motif I            motif          Motif II                     Motif IV          Motif V
           226      254    258        369    375       507  509                    669       680        745
CdhA     GTQIIN   RLGGGFG-AKF      SGSY GWGQ--PQ     CDMGQGVTN       CdhA     MGTIVNPMIVDGCLH     GGMKGVGGGG
Pp_QorL  ATQMPS   DVGGGFG-QKA      IGAYRGVGF--TA     ASSGQGHET       Pp_QorL  CGTMINPMIVEGCIR     GGIKGMGSA
Hp_CutL  TSQAPH   DIGGGFG-NKV      GVAYRCSFRV-TE     ITQGQGHQT       Hp_CutL  CGTRINPMIIEGCIH     IGAKGVASP
Oc_CodL  TFQAPH   DIGGGFG-NKV      GVAYRCSFRV-TE     KSQGQGHET       Oc_CodL  CGTRINPMIIEGCVH     IGAKGVGSP
Hs_Xdh   STQNTM   RMGGGFGGKET      NTAFRGFGG--PQ     TEMGQGLHT       Hs_Xdh   VGSSLNPAIDIGQVE     YASKAVGPP
Bt_Xdh   STQNAM   RMGGGFGGKET      NTAFRGFGG--PQ     TEMGQGLHT       Bt_Xdh   VGSSLNPAIDIGQVE     YASKAVGPP
Rc_XdhB  SSQHPS   RMGGGFGGKES      NTAFRGFGG--PQ     TEMGQGLHA       Rc_XdhB  ACASLNPALDIGCIE     FRSKAVGPP
Dd_Mod   KSVAIH   TTGGTFG--YK      GAAFRGYGA--PE     EDHGQGADA       Dd_Mod   IGKICNKLVVDGQIY     FGASGVGMP

B
         24                          45    100                      125
CdhB     HSDDVAIL AGGQS LMPLLNFRM....  SVAHIQIKTRG TLGGN LCNAHPASEM
Pp_QorM  D-PDARII AGGQS LLPLLAFRS....  HVAHQAVRNRG TLGGS LAHADAGAEM
Hp_CutM  LGSDAKLL AGGHS LLPMMKLRE....  LIADPQVRNRG TIGGD IAHGDPGNDH
Oc_CoxM  LGEDARPL AGGHS LIPIMKTRE....  LIADPQIRYMG TIGGN AANGDPGNDM
Rc_XdhA  AHPEATLI AGGTD VSLWVTKAP....  RFASEQVRQVA TIGGN IANGSPIGDG

C
         42   47   50                         62          101       139
CdhC     GCEHGVCGACTIQFDGEP----------ARSCLM........ LQCGFCTPG.....NLCRCTGY
Pp_QorS  GCEQGVCGSCTILIDGAP----------MRSCLT........ LQCGFCTAG.....NLCRCTGY
Dd_Mod   GCGKGQCGACTVILDGKV---------VRACII........ AQCGFCTPG.....NICRCTGY
Hp_CutS  GCETSHCGACTVDIDGRS---------VKSCTH........ LQCGFCTPG.....NLCRCTGY
Oc_CoxS  GCDTSHCGACTVDLDGMS---------VKSCTM........ LQCGYCTPG.....NLCRCTGY
Hs_Xdh   GCGEGGCGACTVMLSKYDRLQNKIVHFSANACLA........ SQCGFCTPG.....NLCRCTGY
Bt_Xdh   GCGEGGCGACTVMLSKYDRLQDKIIHFSANACLA........ SQCGFCTPG.....NLCRCTGY
Rc_XdhA  GCNEGDCGACTVMIRDAAG------SRAVNACLM........ SQCGFCTPG.....NLCRCTGY
```

Figure 5.1. Multiple Sequence Alignment of Cdh subunits with Molybdenum (Mo)-containing hydroxylases with crystal structures are used in the alignment: Pp_Qor (*P. putida* 86 Quinoline 2-oxidoreductase), Hp_Cut (*H. pseudoflava* CO dehydrogenase), Oc_Cod (*O. carboxidovorans* CO dehydrogenase), Rc_Xdh (*R. capsulatus* xanthine dehydrogenase), Dd_Mod (*D. desulfuricans* aldehyde oxidoreductase), Hs_Xdh (human xanthine oxidoreductase), Bt_Xdh (bovine xanthine dehydrogenase). (A) Conserved Mo-cofactor-interacting motifs, (B) FAD-binding motif, (C) [2Fe-2S] cluster motifs identified in CdhA, CdhB, and CdhC respectively. Numbering corresponds to the sequence of Cdh.

Sequencing and Analysis of the Complete Caffeine Gene Cluster from CBB1

In Chapter 4, the *tmuM* gene (1.1-kb) was isolated from fosmid pMVS848 as described in the section entitled "Cloning, Expression and Purification of TMU-Oxidoreductase". Subsequently, flaking DNA from both ends of the *tmuM* gene were sequenced by obtaining specific primers derived from the *tmuM* gene sequence. This iterative process ultimately led to sequencing of a total of 6.18-kb DNA from the fosmid pMVS848 with identification of eight new ORFs, including the *tmuM* gene. Combined with previously sequenced 19.02-kb, a total of 25.2-kb was sequenced, which constituted the total length of the CBB1 g-DNA fragment present in pMVS848. This 25.2-kb g-DNA sequence of CBB1 thus contains (15 + 8) 23 ORFs, including two partial ORFs at both ends of the fragment (Fig. 5.2, Table C1 of Appendix C).

A BLASTX analysis of these 23 ORFs revealed that at least six of them are involved in caffeine C-8 oxidation pathway in CBB1 (tabulated data in Fig. 5.2). A detailed list of BLASTX results of all the 23 ORFs are in Appendix C. Thus, along with *cdhA,B,C,* and *tmuM*, five other ORFs were found associated with C-8 oxidation pathway namely, *tmuH*, *tmuD*, *orf1*, *orf2* and *orf3*. These genes had significant homology with hydroxyisourate (HIU) hydrolases, OHCU decarboxylases, allantoinases, *ylbA*, and acetylornithine deacetylases, respectively (Fig. 5.2). Sequences of these genes with putative assigned functions were deposited in the GenBank with accession numbers *cdhA*: ADH15879; *cdhB*: ADH15880; *cdhC*: ADH15881, *tmuM*: JQ743481; *tmuH*: JQ743482; and, *tmuD*: JQ743483.
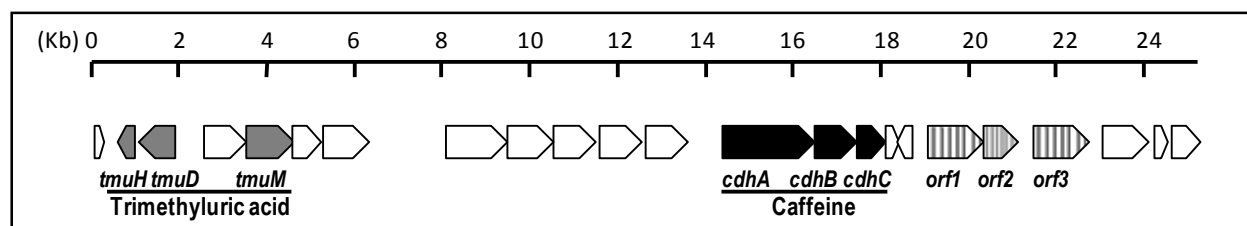
Gene	No.of codons	Proposed / confirmed fucntion	Homologous protein : % Identity[a]			
			NR Database	%	SwissProt Database	%
Trimethyluric acid oxidation						
tmuH	114	TM-HIU hydrolase	Hydroxyisourate hydrolase	50	5-hydroxyisourate hydrolase 2	44
tmuD	291	TM-OHCU decarboxylase	OHCU decarboxylase	39	OHCU decarboxylase	37
tmuM	385	TMU monooxygenase	FAD-binding monooxygenase	38	Zeaxanthin epoxidase	24
Caffeine oxidation						
cdhA	791	Cdh molybdopterin binding subunit	Xanthine dehydrogenase molybdopterin binding protein	49	Carbon monoxide dehydrogenase large chain	32
cdhB	297	Cdh FAD-binding subunit	Alcohol dehydrogenase medium subunit	39	Carbon monoxide dehydrogenase medium chain	31
cdhC	167	Cdh [2Fe-2S]-binding subunit	aldehyde oxidase small subunit	59	Carbon monoxide dehydrogenase small chain	47
Trimethylallantoin and downstream metabolism						
Orf1	466	Trimethylallantoinase	Allantoinase	41	Allantoinase	42
Orf2	266	TMAA degrading enzyme	Allantoin catabolism protein	60	YlbA (encoded by *ylbA*)	21
Orf3	429	TMAA degrading enzyme	Acetylornithine deacetylase ArgE	58	Acetylornithine deacetylase	24

[a] % identity was determined by aligning the gene product of each *orf* with the homologous protein using ClustalW2 [Larkin *et al.*, 2007]

Figure 5.2. Physical map of the genes for caffeine degradation in a 25.2-kb gene cluster in *Pseudomonas* sp. CBB1. Genes are denoted by arrows and show the extents and directions of transcription. The two genetic modules are denoted by the substrate upon which the gene-products act: caffeine, black; trimethyluric acid, dark grey. Sequence similarity and function assignment of the ORFs-encoded proteins are indicated in table.

Multiple Sequence Alignment of *tmuH* and *tmuD* and Function Assignment

A BLASTX analysis of *tmuH* gene exhibited 50% and 44% identity with putative HIU hydrolases from *Rhizobium leguminosarum bv. Trifolii* and *Sinorhizobium meliloti*, respectively (Fig. 5.2, Table C1). Similarly, BLASTX of gene *tmuD* revealed 39% and 37% identity with putative OHCU decarboxylases from *Starkeya novella* and *Haloferax volcanii* (Fig. 5.2, Table C1), respectively. Upon further investigation, it was found that *hpxT* and *hpxQ* from the *hpx* gene cluster of *Klebsiella pneumoniae* [de la Riva *et al.*, 2008] had 29% and 19% sequence identity with *tmuH* and *tmuD,* respectively. However, as previously mentioned, *hpxO* had 38% identity with *tmuM*. In order to further investigate the function of these genes (or gene products), sequences of similar proteins with known function were obtained from these database search and aligned using ClustalW2 [Larkin *et al.*, 2007]. Conserved amino acids were visualized in accordance with the ESPript [Gouet, 1999] equivalence measurement.

A BLASTP analysis of *tmuH* and *hpxT* revealed that both the gene products belong to transthyretin-like protein (TLP) super family [French *et al.*, 2011]. According to NCBI Conserved domain database [Marchler-Bauer *et al.*, 2005], the TLP super family includes transthyretin (a highly conserved vertebrate protein that transports thyroid hormones and retinol) and 5-hydroxyisourate hydrolase (HIU hydrolase). Proteins of this family are usually homotetrameric with each subunit consisting of eight beta-strands arranged in two sheets and a short alpha-helix. The central channel of the tetramer contains two independent binding sites, each located between a pair of subunits. Thus, HIU hydrolases from various sources (mammals to bacteria) were aligned with *tmuH*

(Fig. 5.3) in order to identify conserved domains and residues within them either

unaltered or altered with physicochemically-equivalent residues.
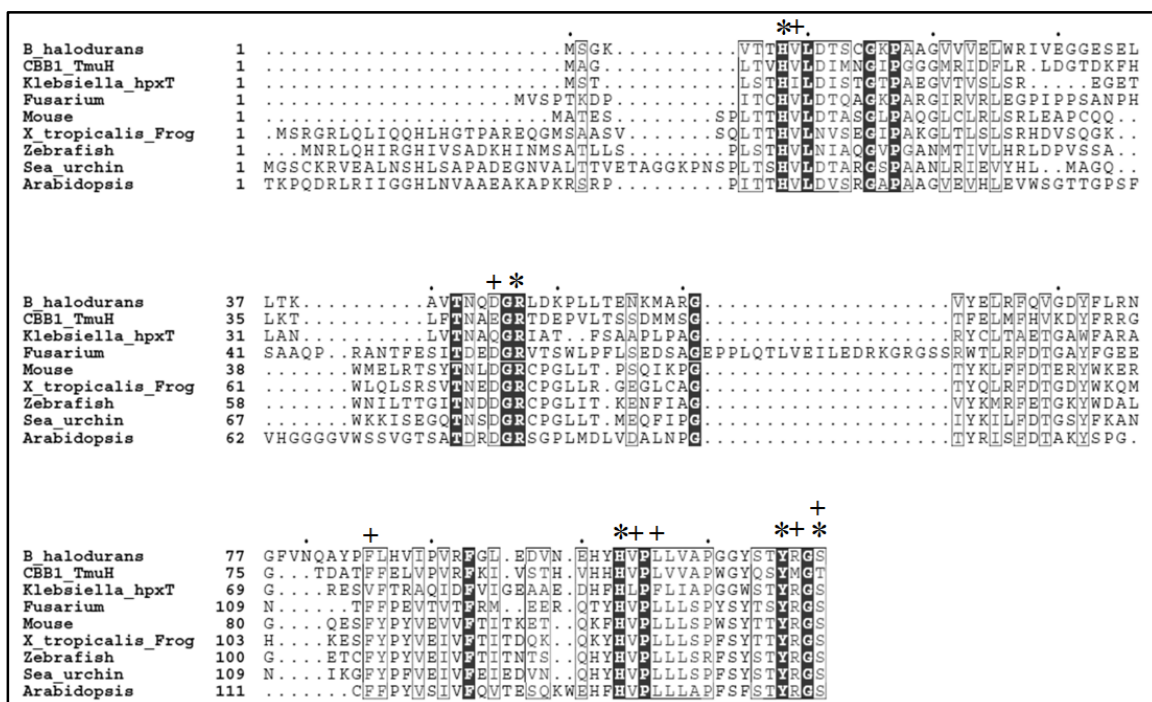
Figure 5.3. Multiple sequences alignment of HIU hydrolases from various organisms: bacteria (*Bacillus halodurans*, *Pseudomonas* sp. CBB1, *Klebsiella pneumonia)*, fungi (*Fusarium oxysporum*), plant (*Arabidopsis thaliana*), amphibian (*Xenopus tropicalis* or Western clawed frog), fish (*Danio rerio* or Zebra fish), marine animal (*Sea urchin*), and mammal (*Mus musculus* or Mouse). Amino acid conservation visualized in accordance with the ESPript [Gouet *et al.*, 1999] equivalence measurement: invariant residues (dark background), physicochemically equivalent residues (boxed). (*) residues lining the active site, (+) residues within the conserved domain (box) with single change in physicochemically equivalent residues either in *tmuH* or *hpxT* or both.

A multiple sequence alignment (MSA) of HIU hydrolases from various sources

(Fig. 5.3) identified several conserved domains within these proteins from bacteria to

mammals which are also present in *tmuM*. Upon close inspection, it was observed that

these conserved domains contains active site residues which were mostly unaltered in *tmuM* or altered with a physicochemically equivalent amino acid. Moreover, it was also found that *tmuM* aligned well with *hpxT* (both bacterial HIU hydrolases) with very few gaps or alterations in the conserved domains. This provides evidence that tmuM encodes for a putative TM-HIU hydrolase and any alterations in residues in the active site or conserved domains between *tmuM* and *hpxT* can be attributed to the difference in substrates (TM-HIU vs HIU).

Similarly, a MSA of OHCU decarboxylases from various biological sources (Fig. 5.4) also demonstrated equivalent results. As shown in Fig. 5.4, *tmuD* perfectly aligned with other OHCU decarboxylases with presence of most of the conserved domains, very few gaps and one of the unaltered active site residues. Other active site residues were altered but with physicochemically equivalent residues, which can again be attributed to substrate specificity (TM-OHCU vs OHCU). This suggests that *tmuD* possibly encodes for a TM-OHCU decarboxylase. However, it is worth mentioning that *tmuD* is a fused protein with a portion homologous to HIU hydrolase. Currently, we do not have any explanation for this extra fused HIU hydrolase gene, which is present side by side with the *tmuH* on the same genome. More detailed bioinformatics and functional genomic analysis might explain the origin of gene fusion (*tmuD* with a copy of *tmuH*) and their relation with the stand-alone *tmuH*. However, the sequence organization indicates that tmuD is a novel TM-OHCU decarboxylase, probably with a fused (TM)-HIU hydrolase.

```
Mouse              1  ..........MDMVKVNSMDFGEFVDVFGNIVEKCPLIAAAVWSQRPFSGLEDLENHFFAFIDALPRSGQE
Zebrafish          1  ..........MDINVVNALAYEDFVKLFGNVVEKCPLISAAIWSYRPFKDLADIEARISEFIHSLPDSGKE
X_tropicalis_Frog  1  ..........MDLNTINSMSYEEFLDIFGNIIERCPIITAAIWSQFPFASVTELENSVYDFIESLPLTGKE
Sea_urchin         1  .....MAEKFTIEAINQQEYEEFIENFENIVEHGILIAGAVWSHRPFRSFDHLHRCFCDFMDSLPESGKQ
B_halodurans       1  MMGMAVSTKLSIDEVNMLEKEDFVTKIGPIFEHSPWVAERAWAHRPFTSAENMYECMLEKVYEADKRLQL
CBB1_TmuD          1  ...............MTLREFVDVFGGIFEHAPWVAERACLSRPFASVYEMHRKMMDVVRRCDESEKV
Klebsiella_hpxQ    1  .........MIALSQFNSLSKDEAAGLLAPCVAIPAWG.ETLVSLRPFASRHALLQTAREAMANWGEDELN
Arabidopsis        1  ............MAMEIGEDEWKVCCG.....SSEFAKQMSTSGPLTSQEAIYTARDIWPNQVNVTDWL


Mouse             62  GILRCHPDLAGRDLQQGTLTAESQREQSQ.AGLTSLDTDDRLRLQQLNAQYRERFGFPFVLAARLSDR.A
Zebrafish         62  GILRCHPDLAGRDLQSGTLTPESQEEQSQ.AGMTTLDSAEIVHMYRLNSEYKERFGFPFVICARLNNK.A
X_tropicalis_Frog 62  GILRCHPDLAGRDLMRGTLTDESQTEQAQ.AGLTALTPKERETLNLLNSQYKAKFGFPFVICAKMSDK.N
Sea_urchin        66  SILRCHPNLAGKLARQGKLTSESEQEQAS.AGLSSLTDEQYNEIHKNNDIYRKKFSFPFVICARENKI.A
B_halodurans      71  ALLRAHPDLGTR....LEISETSQSEQQR.AGLSQLTEEEFAVFAELNKCYVDTFRFPFIMAVRGQTK.N
CBB1_TmuD         54  GLLRGHPELGGREARAGEMTPDSSSEQSR.LGLDRLSEAEFSEMTELNRLYDEKFGFPCMICLRLHVKRD
Klebsiella_hpxQ   62  AALSAHPRIGEK..PTGCQAHAALSRQEQ.SSVDSENERLAQALREGNARYEARFGWVFLIRAKGRSG.E
Arabidopsis       53  EAFSAHPQIGNTPSPSINSDFARRSVSEQSTAFATTSASALQELAEWNVLYKKKFGFIFIICASGRTH.A


Mouse            130  TVPRELARRLQCQPESELRTALGEVKKISHLRLTDLLGAHSHSARVELP.....
Zebrafish        130  DIVRQLSERLKNRRTAELECLRLHSIVLSDIQTKL..........
X_tropicalis_Frog130  KIMRELASRLQNEQSQELQIGIAEVKKICHLRVNDLFLNVKLPTKL.....
Sea_urchin       134  AILQGLQTRIQNARELELQKGIEEVKKISYYRLLDMIQEKTLSPKKPKL.....
B_halodurans     135  SIKEQMRKRLVNDEEQERKTALREVAKIAKFRLADLVVMGSRS...........
CBB1_TmuD        123  TVVAEHRRRLLNDRQTEIENCIEQVGQLTERRLRERVRESQRTRRAISTHVLDS
Klebsiella_hpxQ  128  EMLQALTRRLQHTADEEVAEALAQLREITMLRLEGVIGE...............
Arabidopsis      122  EMLHALKERYENRPIVELEIAAMEQMKITELRMAKLFSDKAKVISETDSSSS..
```

Figure 5.4. Multiple sequences alignment of OHCU decarboxylases from various organisms: bacteria (*Bacillus halodurans*, *Pseudomonas* sp. CBB1, *Klebsiella pneumonia)*, plant (*Arabidopsis thaliana*), amphibian (*Xenopus tropicalis* or Western clawed frog), fish (*Danio rerio* or Zebra fish), marine animal (*Sea urchin*), and mammal (*Mus musculus* or Mouse). Amino acid conservation visualized in accordance with the ESPript [Gouet *et al.*, 1999] equivalence measurement: invariant residues (dark background), physicochemically equivalent residues (boxed), (*) residues lining the active site, (+) residues within the conserved domain (box) with single change in physicochemically equivalent residues either in *tmuH* or *hpxT* or both.

Analysis of the Caffeine Gene Cluster in CBB1 with Hpx Gene Cluster

Comparison of the organization of *tmuH* and *tmuD* in the 25.2-kb gene cluster

(Fig. 5.5) with the corresponding (characterized) genes *hpxO*, *hpxT*, and *hpxQ* of the

related uric acid pathway in the *hpx* gene cluster from *Klebsiella pneumoniae* [de la Riva

et al., 2008] and 22.872-kb purine utilization gene cluster from *Klebsiella oxytoca* [Pope

et al., 2009] reveals some interesting features. Organization of *tmuM*, *tmuH*, and *tmuD* in

the caffeine gene cluster of CBB1 (Fig. 5.5) is similar to that of corresponding *hpxO*,

hpxT and *hpxQ* on the *hpx* gene clusters in *Klebsiella*, in terms of orientation, direction of

transcription and gene location. Genes, *tmuH* and *tmuD* are next to each other with same

direction of transcription and separated from the oppositely transcribing gene *tmuM* by a

single 1188 nt long ORF (Fig. 5.2). This is same as the organization of corresponding

hpxT, and *hpxQ* genes and their separation from the oppositely transcribing *hpxO* by

hpxP (Fig. 5.5) [de la Riva *et al.*, 2008; Pope *et al.*, 2009].
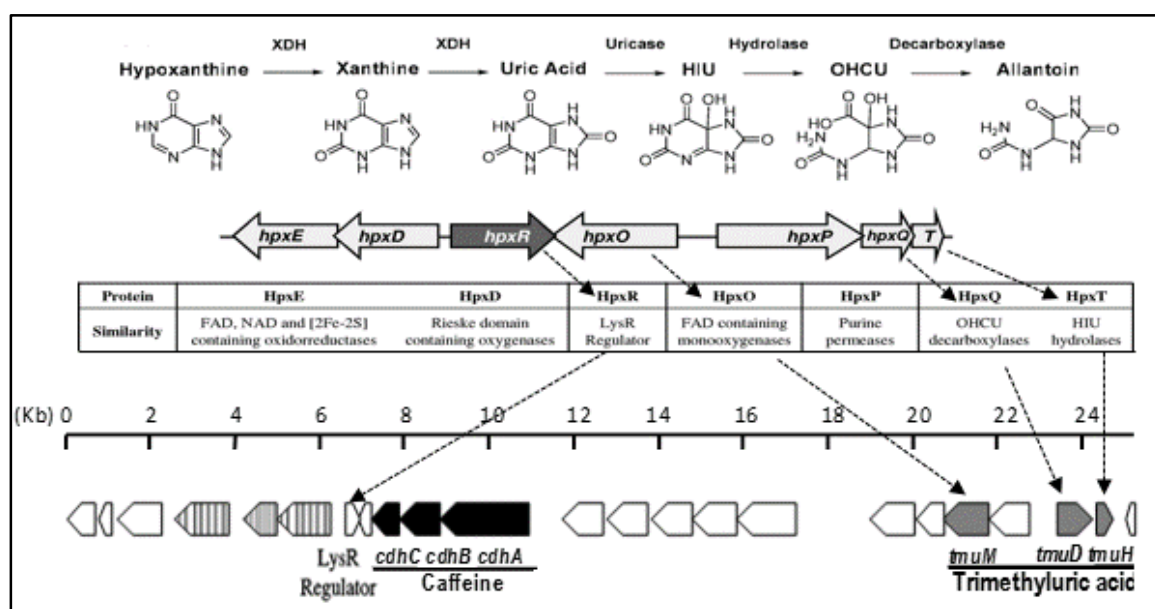
Figure 5.5. Analysis of 25.2-kb caffeine gene-cluster in CBB1 (bottom) vs *hpx* gene
        cluster (top) in *Klebsiella pneumoniae* [de la Riva *et al.*, 2008].

Further, in the *Klebsiella oxytoca* gene cluster, *hpxB* is located about 8-kb

downstream of *hpxO* and codes for an allantoinase (data not shown) [Pope *et al.*, 2009].

In the caffeine gene cluster (Fig. 5.2), *orf1* with homology to allantoinases, along with

other genes (*orf2* and *orf3*) are 14-kb downstream of *tmuM*. It is tempting to propose that

these genes are involved in further downstream metabolism of S-(+)-TMA; but this needs further substantiation. The larger separation of *orf1* from *tmuM,* can be attributed to the recruitment of 3.765-kb long *cdhABC* genes, encoding for the heterotrimetic Cdh (Fig. 5.2).

Discussion

Sequence Analysis for Co-factor Assignment to Cdh Subunits

Previously, Yu *et al*. [2008] established by sequence homology that *cdhA*, the large (α 90-kDa) subunit of Cdh, contained the molybdopterin cofactor (Fig. 5.6). Based on the UV-Visible absorbance spectrum of Cdh, it was also suggested that the other two sub-units (medium β 40-, and small γ 20-kDa) of Cdh contained FAD and iron-sulfur clusters as cofactors (Fig. 5.6). However, the assignment of the cofactors to the β and γ could not be made due to unavailability of total sequence of these subunits [Yu *et al.*, 2008]. With availability of complete Cdh gene sequence (*cdhA,B,C*), homology-based analysis of the complete gene as well as individual subunits was possible, which led to further characterization of this enzyme (Fig. 5.1). Total sequence analysis suggests that members of this family of proteins include other heterotrimeric enzymes, like carbon monoxide dehydrogenase, xanthine dehydrogenase/oxidase, and aldehyde oxidase (Figs. 5.1 and 5.2).
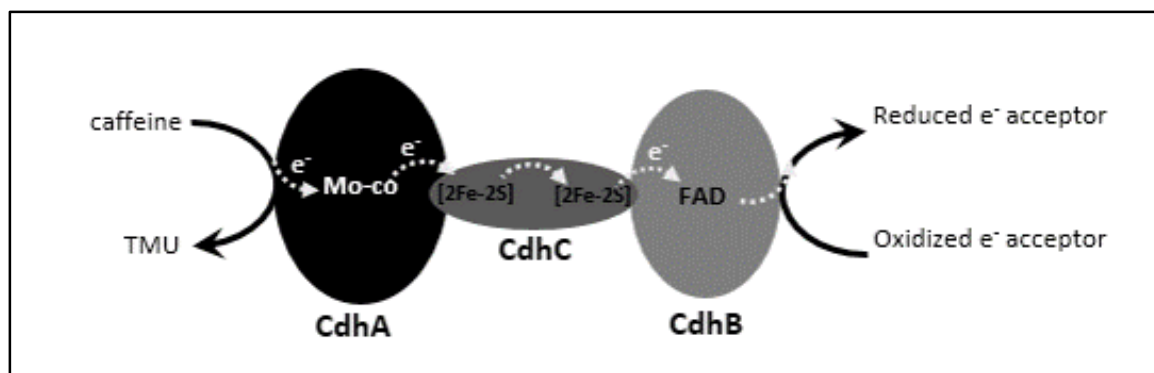
Figure 5.6. Organization of Cdh subunits and proposed electron transfer during the oxidation of caffeine to TMU by Cdh.

Source: Adopted from Louie *et al*., 2009.

As expected, this family of proteins is associated with cofactors like molybdopterin, flavin adenine dinucleotide (FAD), and [2Fe-2S] clusters [Brondino *et al.*, 2006; Hille *et al.*, 2006]. Sequence analysis of individual subunits (*cdhB* and *cdhC*) revealed that *cdhB* had significant homology with FAD binding medium subunits of similar enzymes like alcohol dehydrogenase and carbon monoxide dehydrogenase. Plus, N-terminal protein sequence deduced from *cdhB* also matched the experimentally determined N-terminal protein sequence of the medium (β, 40-kDa) subunit of Cdh purified from CBB1. Additionally, *cdhC* was homologous to the corresponding iron-sulfur containing small subunits of all the heterotrimeric enzymes. This suggests that the β (40-kDa) subunit of Cdh contains the FAD cofactor and the the γ subunit contains the iron–sulfur cluster. Cdh is similar to the known NAD^+-dependent xanthine dehydrogenases in terms of catalysis, sub-unit structure (αβγ) and cofactor content.

Elucidation of the Complete C-8 Oxidation Pathway Using Homology Analysis of
Caffeine Gene Cluster in CBB1

The first reaction of the proposed C-8 oxidation pathway in *Pseudomonas sp.*
strain CBB1 is the hydrolytic oxidation of caffeine to TMU, catalyzed by Cdh [Yu *et al.*,
2008]. Earlier, it was hypothesized that in CBB1, TMU undergoes a three-step enzymatic
degradation process to produce S-(+)-TMA (Chapter 3). In Chapter 4, the second
reaction of the C-8 oxidation pathway, i.e., hydroxylation of TMU to TM-HIU by TMU
monooxygenases, was reported. However, further enzymatic degradation of TM-HIU
and the fate of TMA in CBB1 was unknown. Thus, in order to characterize the
enzymatic degradation of metabolites downstream of TM-HIU, gene cluster analysis of
the 25.2-kb genomic DNA segment of CBB1 was employed, which harbored almost all
the genes of the pathway (Fig. 5.2). Although bioinformatics is not a regularly employed
method for pathway/enzyme characterization, it is highly helpful in certain special cases
(like the current work) where there is a (i) lack of availability of the metabolites like TM-
HIU, TM-OCHU and racemic and S-(+) TMA proposed in this pathway (Fig. 5.7), and
(ii) lack of published synthetic methods for these compounds.

Homology and gene cluster-based analysis is a valid tool for elucidation of the
pathway [Kaneko *et al.*, 1996; Blattner *et al.*, 1997]. It has been employed extensively in
recent years for elucidation of uric acid degradation pathway in various microbial strains,
including *Bacillus subtillis* [Schultz *et al.*, 2001], *Klebsiella pneumoniae* [de la Riva *et
al.*, 2008; Guzman *et al.*, 2011] and *Klebsiella oxytoca* [Pope *et al.*, 2009]. For example,
HIU hydrolase (MuraH) and OHCU decarboxylase (MuraD) from mouse were recently
identified by phylogenetic comparison of genomes and sequence analysis [Ramazzina *et*

al., 2006; Ramazzina *et al.,* 2008]. This led to cloning of these genes in *E.coli* and the establishment of HIU conversion to S-(+)-allantoin via OHCU by ^{13}C-NMR [Ramazzina *et al.*, 2006]. Subsequently, analysis of the *hpx* gene cluster (Fig. 5.5) in *Klebsiella pneumoniae* [de la Riva *et al.*, 2008] led to the identification of seven genes involved in the oxidation of hypoxanthine to allantoin via uric acid, HIU and OHCU. To our knowledge, the above is the first report on three-step enzymatic conversion of uric acid to S-(+)-allantoin in bacteria based on gene cluster analysis. This led to identification of three relevant genes: *hpxO* encoding urate oxidase (HpxO) (Fig. 4.4 and 4.13 of Chapter-4), *hpxT* (Fig. 5.3) encoding HIU hydrolase (Fig. 5.2) and *hpxQ* (Fig. 5.4) encoding OHCU decarboxylase (Fig. 5.2). Later, this led to the functional expression of *hpxO*, *hpxT* and *hpxQ* in *E.coli* and crystal structure elucidation of KpHIUH (*hpxT*) [French *et al.*, 2011] and KpOHCU decarboxylase (*hpxQ*) [French *et al.*, 2010]. Even though the substrates of KpHIUH and KpOHCU are unstable, the catalytic activities of these enzymes were confirmed by observation of substrates *in situ*.

A similar gene homology (Fig. 5.2), MSA (Fig. 5.3 and 5.4) and cluster based analysis (Fig. 5.5) of the 25.2-kb genomic DNA segment of CBB1 composed of 21 full length genes, revealed the presence of *tmuH* and *tmuD* upstream of *tmuM*, and three other genes (*orf1*,*orf2* and *orf3*) downstream of *cdhABC*. Based on these bioinformatics studies, it is proposed that while *tmuH* encodes for a putative TM-HIU hydrolase which catalyzes the N1-C6 hydrolysis and ring opening reaction of TM-HIU to TM-OHCU (Fig. 5.7), *tmuD* encodes for a putative TM-OHCU decarboxylase which catalyzes the stereo-selective decarboxylation of TM-OHCU to S-(+)-TMA in CBB1 (Fig. 5.7). In

addition, these outcomes indicate that TM-OHCU and S-(+)-TMA are possible

intermediates of caffeine degradation via C-8 oxidation (Fig. 5.7).
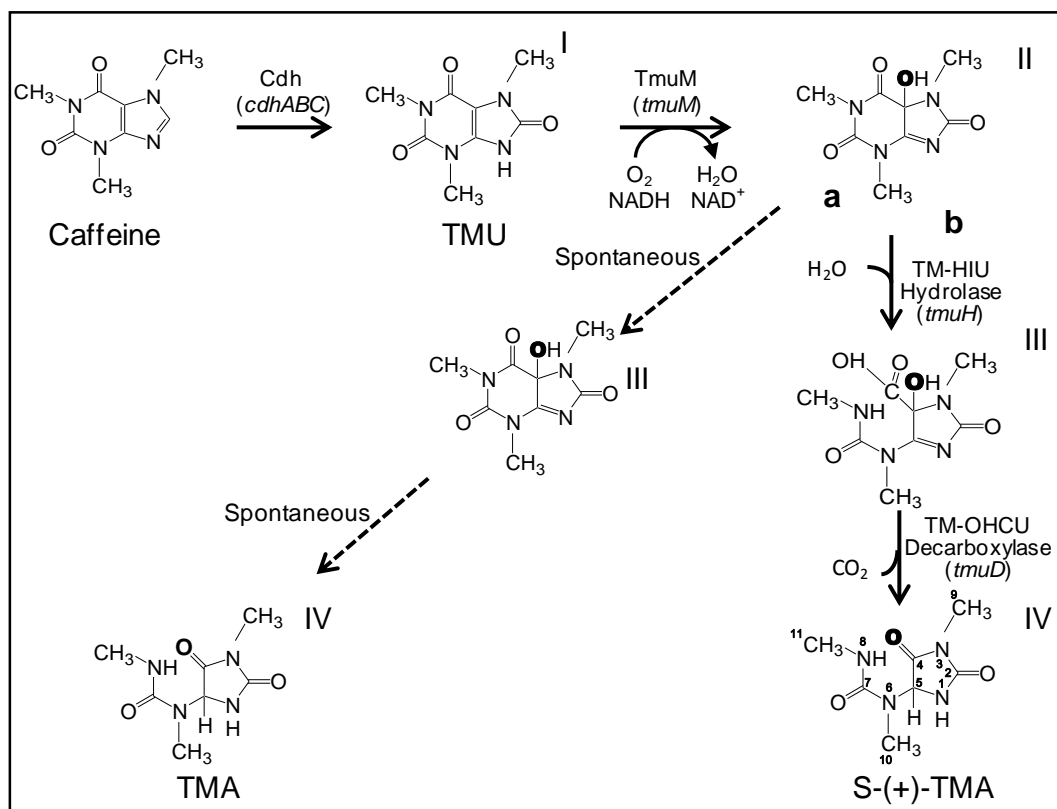
Figure 5.7. Proposed caffeine degradation pathway in *Pseudomonas* sp. CBB1. Caffeine dehydrogenase (Cdh) oxidizes caffeine to TMU (I), which is oxidized to TM-HIU (II) by trimethyluric acid monooxygenase (TmuM). (a) Spontaneous conversion of TM-HIU to racemic TMA through the TM-OHCU (III) intermediate. (b) Proposed pathway for the enzymatic conversion of TM-HIU to S-(+)-TMA. Gene representation is in italics. Metabolite: I: 1,3,7-trimethyluric acid (TMU); II: 1,3,7-trimethyl-5-hydroxyisourate (TM-HIU); III: 3,6,8- trimethyl-2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (TM-OHCU); IV: 3,6,8- trimethylallantoin (TMA).

In Chapter 4, it was shown that a specific enantiomer of TMA is formed which

gradually racemizes with time (Chapter 3, Fig. 3.13). Even though this isomer has not

been stereo-chemically characterized, it supports the formation of a chiral TMA (Chapter

3, Fig. 3.13). Other strong evidence for the proposed S-(+)-TMA is that only S-(+)-allantoin seems to be formed in cells [Vogel *et al.*, 1976] capable of utilizing uric acid. These organisms rely on the enzymatic conversion of HIU to S-(+)-allantoin via OCHU [Ramazzina *et al.*, 2006]. Our contention is that the putative TM-HIU hydrolase and a putative TM-OHCU decarboxylase in CBB1 are specific for methyl-substituents in the respective substrates (Fig. 5.7). This is based on the preference of both Cdh and TmuM for trimethylated substrates with no activity on non-methylated xanthine and uric acid. Hence, in the proposed conversion of TMU to S-(+)-TMA, all enzymes are designated with preference for (tri)methylated substrates (Fig. 5.7). However, any clarity on the biochemistry of formation and metabolism of S-(+)-TMA will come from the functional expression of *tmuH* and *tmuD* along with chemical characterization of S-(+)-TMA.

Delineation of C-8 Oxidation Pathway beyond S-(+) TMA in CBB1

Organisms capable of utilizing uric acid as a source of carbon and nitrogen possess allantoinase and allantoin racemase [Guzman *et al.*, 2011; Ramazzina *et al.*, 2008]. Recently, Guzmán *et al.* [2011] reported the presence of another gene cluster in *Klebsiella pneumoniae that* harbors genes responsible for further degradation of allantoin. In particular, *hpxA* and *hpxB* were identified that encode allantoin racemase and allantoinase, respectively. This lead to the crystal structure determination of heterologously expressed *hpxA* gene-product KpHpxA [French *et al.*, 2011]. Our initial analysis of the 25.2-kb genomic DNA sequence, downstream of Cdh genes, has identified three ORFs *orf1*, *orf2*, and *orf3* showing homology to allantoinase, *ylbA* and acetylornithine deacetylase, respectively (Fig. 5.2). Based on the homology of these

three ORFs and cluster analysis with *hpx* gene cluster of *Klebsiella oxytoca* [Pope *et al.*, 2009], we propose that *orf1* encodes for trimethylallantoinase, which catalyzes a hydrolytic ring opening reaction with S-(+)-TMA to produce trimetylallantoic acid (TMAA) (Fig. 5.8).
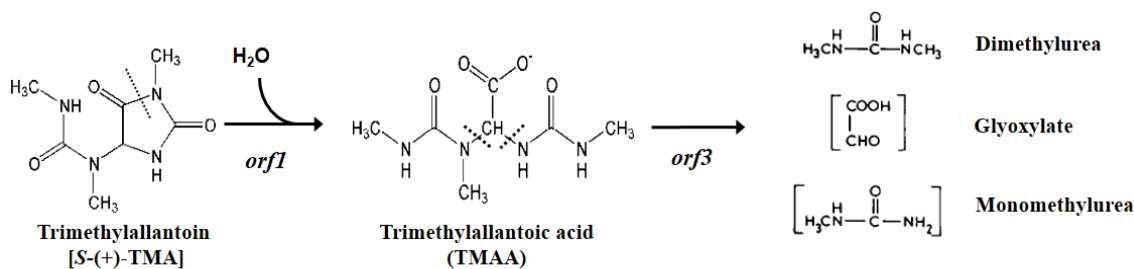
Figure 5.8.   Proposed function for *orf1* and *orf3* gene products based on homology and cluster analysis. While, *orf1* encodes for a putative trimethylallantoinase, *orf3* is proposed to encode for an enzyme like acetylornithine deacetylase which cleave C-N (non-peptide) bonds of trimethylallantoic acid (TMAA),

Furthermore, *orf2* and *orf3* shows similarity to YlbA (encoded by *ylbA*) and ArgE, respectively. YlbA is identified as a candidate *S*-Ureidoglycine Aminohydrolase [Fabio *et al.,* 2010], an enzyme which converts *S*-Ureidoglycine to *S*-Ureidoglycolate in the xanthine degradation pathway in plants (Fig. 1.6 Chapter 1). ArgE is acetylornithine deacetylase, an enzyme that hydrolyses carbon-nitrogen bonds other than peptide bonds, specifically in linear amides in the urea cycle, and amino group metabolism. Based on this, we hypothesize that *orf2* and *orf3* are likely to encode for enzymes which break non-peptide C-N bonds of TMAA (Fig. 5.8) to generate glyoxylate, di- and monomethylurea. Previously, Madyastha and Sridhar had reported transient accumulation of di- and monomethylurea, and glyoxylate in the media during caffeine consumption by a mixed culture consortium [Madyastha *et al.*, 1998].

Summary and Conclusion

A 25.2-kb caffeine gene cluster has been sequenced from the CBB1 genomic library. 23 ORFs were found on the gene cluster, out of which six genes, namely *cdhA,B,C*, *tmuM*, *tmuH*, and *tmuD*, are proposed to be encoding enzymes of C-8 oxidation pathway in CBB1. Sequence homology of *cdhA,B,C* suggests that CdhA contains a Molybdopterin cofactor, while CdhB and CdhC contains FAD and the iron–sulfur cluster, respectively. Sequence homology, multiple sequence analysis and gene-cluster analysis of *tmuM*, *tmuH*, and *tmuD* suggest that *tmuH* encodes for a TM-HIU hydrolase, which converts TM-HIU to TM-OHCU, and *tmuD* encodes for a TM-OHCU decarboxylase, which converts TM-OHCU to S-(+)-TMA. Gene cluster analysis has led to clarity and further refinement of the C-8 oxidation pathway, which now includes two new enzymes (TM-HIU hydrolase and TM-OHCU decarboxylase) on the two-step degradation of TM-HIU to S-(+)-TMA via TM-OHCU. Furthermore, based on sequence homology, *orf1* is proposed to be encoding for a TM-allantoinase and *orf2* and *orf3* proposed to encode for enzymes that might be involved in TMAA catabolism. In summary, a new pathway has been proposed for degradation of caffeine via C-8 oxidation to S-(+)-TMA and further metabolism of TMAA to di- and monomethyl urea and glyoxylate. The first two steps, i.e., conversion of caffeine to TMU by Cdh and TMU to TM-HIU by TmuM are based on detailed enzymology and gene analysis. Subsequent degradation of TM-HIU to S-(+)-TMA is based on gene cluster analysis, and comparison to uric acid degradation, which is based on gene-cluster analysis as well. To our knowledge, this is the first report of biochemical and genetic analysis of caffeine degradation via C-8 oxidation.

CHAPTER 6

DEVELOPMENT OF A RAPID COLORIMETRIC ASSAY FOR DETECTION OF CAFFEINE IN BEVERAGES AND BODY FLUIDS USING CAFFEINE DEHYDROGENASE FROM *PSEUDOMONAS* SP. CBB1

Introduction

Caffeine Consumption and Human Health

Human consumption of caffeine dates back to as far as the Paleolithic period primarily from its natural sources such as coffee (harvested for beans), tea (harvested for leaves), cocoa (harvested for seeds), guarana (harvested for fruit), kola nuts, and yerba mate [Weinberg, 2001]. In 1820, caffeine, a bitter white crystalline alkaloid, was isolated and identified as the chemical entity in coffee which is responsible for its stimulating effects. Since then, caffeine has been synthesized in its purest form and widely used as an artificial stimulant in various pharmaceuticals, beverages like soft drinks, energy drinks, and other food products [Durrant *et al.*, 2002; McCusker *et al.*, 2006; Heckman *et al.*, 2010] as listed in Table D.1 of Appendix D. In fact, caffeine is one of the most widely consumed psychoactive substance/drug in the world with an estimated global consumption of 1.2×10^8 kg per year [https://en.wikipedia.org/wiki/Caffeine] and in a range of 80 to 400 mg per person per day [Gokulakrishnan *et al.*, 2005]. According to the report published by International Coffee Organization (ICO) in September 2012, the annual global consumption of coffee was 139 million bags (of 60 kg each), and the demand is growing rapidly each year [http://dev.ico.org/documents/cmr-0912-e.pdf].

A low (< 200 mg/day/adult) to moderate (200-400 mg/day/adult) consumption of caffeine is regarded as beneficial in terms of increasing alertness and overcoming fatigue. In medicine and pharmaceutics, caffeine is used as an analgesic, antipyretic, diuretic, cardiac and respiratory stimulant, as well as in certain types of painkillers to increase the effectiveness [Spiller, 2010]. Research shows that beneficial effects of caffeine include reduced risk to diseases like Parkinson, Alzheimer, Type 2 diabetics, Gallstone, Asthma, and certain types of cancers like oral, esophageal and pharyngeal [Parliament, 2000]. In contrast to that, various studies in recent years have reported adverse health effects of excess caffeine consumption (400-700 mg/day/adult), e.g., changes in sleep pattern, increased blood pressure, palpitations, anxiety, irritability, nausea, restlessness, etc. [Eteng *et al.*, 1997; Marin *et al.*, 2011, Schmidt *et al.*, 2008]. Furthermore, caffeine is an addictive substance. Excessive and long term caffeine consumption leads to severe health condition like insomania, auditory hallucinations, loss of appetite, abdominal pain and heart burn due to excess acid, vitamin deficiency, increased risk of coronary heart disease, and several types of cancer [van Dam *et al.*, 2008]. Moreover, in a few rare instances, sudden but lethal doses of caffeine have also resulted in human death [McGee *et al.*, 1980; Garriott *et al.*, 1985; Mrvos *et al.*, 1989]. This brings into question the safety of this widely consumed drug [Marin, 2011].

FDA Restictions and Need for a Rapid Caffeine Diagnostic Test

Unlike other psychoactive substances consumed worldwide, caffeine is legal and generally unregulated with a classification of "Multiple Purpose Generally Recognized as Safe Food Substance" by the U.S. Food and Drug Administration (FDA) [Under 21 Code

of Federal Regulations section 182.1180]. However, the FDA does not consider it as a nutrient although it is a natural product. The FDA considers concentrations of caffeine above 200 parts per million (ppm), equivalent to 0.02 percent in food and beverages, as unsafe [21 Code of Federal Regulation section 182.1180(b)] and recommends no more than 400 milligrams of caffeine per day for adults [U.S. Food and Drug Administration 2013]. In recent years there is an increasing trend of adding caffeine to various food and other consumable products that is of concern. These include instant energy/alert boosters or faster weight loss supplements, etc. On last count, at least 150 such items are available [http://www.energyfiend.com/caffeine-in-candy], ranging from usual caffeinated soft drinks and chocolates, to the newly developed caffeinated chewing gums, chap sticks, bathing soaps and tattoos (Table D1). Shockingly, most of these products are neither FDA approved, nor restricted to children, and often contain 150-250 mg of caffeine per serving (Table D1) and considered unsafe, especially when consumed without restriction. This is a disturbing trend keeping in view the possible negative effects of excessive caffeine consumption on human health [Marin 2011], particularly teenagers.

Recently, the FDA has also shown growing concern over the popular trend of caffeine-infused food products and beverages (Table D.1), particularly the ones which are widely consumed by children and teenagers like energy drinks and chewing gums [Reissig *et al.*, 2009; Tomson *et al.*, 2013; McGee *et al.*, 1980; Marin, 2011; Johannes *et al.*, 2010; U.S. Food and Drug Administration 2013]. Under current FDA guidelines, any manufactured food or beverage product should not be supplemented with caffeine above 200 ppm (0.02 %) [U.S. Food and Drug Administration 2013] and if caffeine is added, it has to be declared clearly on the content label. There are instances of drug manufacturers

who have recalled the mislabeled bottles, which might lead to unintentional excessive intake of caffeine [U.S. Food and Drug Administration 2011]. Thus, keeping in mind the negative health effects of caffeine over consumption, and the FDA's restrictions over caffeine intake, the safety of this natural product is under scrutiny. Given this situation, there is a clear need for developing a rapid caffeine diagnostic test, easily usable by consumers, to detect caffeine levels in food and beverages which are consumed on a daily basis. Such a test may also be useful for quality control of pharmaceuticals formulated with caffeine.

In a recent survey on "Caffeine intake by the U.S. population" conducted in September 2009 (revised in August 2010) by The FDA, Laszlo P. Somogyi did an extensive analysis of daily caffeine intake for the subpopulations of children (aged 2-13), male and female teenagers (ages 14-21), adults (over age 22) and women of childbearing (age 16-45) [Somogyi, 2010]. Data and information used in this study was gathered from relevant sources like National Nutrient Database for Standard Reference (NDB), National Health and Nutrition Examination Survey (NHANES), NPD Groups Food Consumption surveys, The National Coffee Association, The Tea Association of the USA, the American Beverage Association, US trade statistics, comprehensive review of scientific publications, trade association data, the Internet and other industry sources. One of interesting outcomes reported in this survey is that a significant portion (about 95-97%) of caffeine intake comes from beverage sources and not from solid food supplies [Somogyi, 2010]. Thus, a rapid caffeine diagnostic test for beverages will be of value to consumers. Additionally, the caffeine content of instantly prepared beverages like coffee and tea is not constant (usually varies with the amount of coffee/tea powder and method

of preparation). Also, decaffeinated coffee, tea and soft drinks that are available in market are not readily distinguishable from their respective caffeinated counter parts either by visual or taste or aroma. Here again, especially in situations where these drinks are served in non-labeled containers, a quick and sensitive diagnostic test may be of value, preferably in a dip-stick test format.

Excessive caffeine is also moderately restricted for lactating mother and highly restricted for pregnant women, owing to its possible teratogenic effects on the growing fetus and infant [Eteng *et al.*, 1997, Rossignol *et al.*, 1990]. Caffeine has the ability to cross the blood brain barrier and placenta; thus, any caffeinated drink consumed by pregnant women can directly enter the central compartment of the fetus/infant, and can impact brain development [Eteng *et al.*, 1997]. If a nursing mother consumes 100 mg of caffeine, then within 1 hr, about 1-1.5 % (1 mg of caffeine) shows up in 1 L of mother's milk or saliva (body fluid) [http://kellymom.com/bf/can-i-breastfeed/lifestyle/caffeine/, Tyrala *et al.,* 1979]. In the previous FDA report by Somogyi, women of childbearing (age 16-45) were found to be consuming on an average 115.6 mg of caffeine per day, which is in the above range. Hence a pregnant woman/nursing mother would benefit from knowing the caffeine content in drinks or in the milk prior to nursing a child, to make sure that it is below the restricted limit. Plus, such a subgroup of population, who are under highly restricted caffeine diet, will definitely find such a test helpful in safeguarding against any unintentional caffeine intake.

Thus, at the consumer level, there is a clear need for a rapid caffeine diagnostic test. The test should preferably score caffeine content within a minute, be portable (preferably as a dip-stick) and easy to read (preferably by a visible color change). The

test should be sensitive enough to detect caffeine at 1-10 ppm levels, which is the expected range in nursing mother's milk consuming 100-700 mg of caffeinated drink/day. The test should also readily indicate or warn when a higher limit of caffeine is present in beverages, which as per FDA is about 400 mg/day/adult [U.S. Food and Drug Administration 2013].

Desired Attributes of a Rapid Caffeine Diagnostic Test

Based on the above identified need, the following attributes are essential for a caffeine diagnostic test:

1. *In-situ* (on-the-spot) detection of caffeine: Detection of caffeine by High-Performance Liquid Chromatography (HPLC) is very accurate and routinely performed. However, this method is tedious and time consuming from sample preparation to final analysis. Detection methods are not colorimetric. This is not portable or suitable for "at home tests". Alternately, a suitable test has to be as simple and light weight like a litmus paper, or pH paper (may be a dip-stick type test) and portable to test at home or at large.

2. Rapid detection of caffeine: The dip-stick type assay has to be rapid with ability to detect approximate caffeine levels in about one minute. This will enable a coffee drinker or nursing mother to make quick decisions based on the result.

3. Easy to read and semi-quantitative or qualitative: The expectations of the end user of such a test is not accurate detection of caffeine concentration, but to know the range, to make quick decisions. Examples of such diagnostic tests in the market are digital read out of approximate blood sugar levels for diabetics

and qualitative Yes/No pregnancy test. The caffeine diagnostic test needs to be semi-quantitative to readily indicate the range of caffeine concentration. The test must differentiate decaffeinated drinks from caffeinated, and also indicate the level of caffeine in for example, nursing mother's milk. Different levels of caffeine must be visual or easily discernable like a pH indicator strip. Caffeine level could also be coded as Safe-Moderate-Unsafe or High-Moderate-Low.

4. <u>Sensitivity of the test</u>: Based on the FDA recommendation for safe caffeine levels for pregnant woman or infants nursed by mothers, the caffeine diagnostic test should be able to detect caffeine at ~20 ppm in milk or saliva. If the test is colorimetric instead of numerical, the color should readily differentiate various levels of caffeine especially < 20 ppm.

5. <u>Cost of the test and product safety</u>: The test should be reasonably priced at few cents/test. The kit itself should not use any toxic substance. The test strips should have a shelf life of at least a few months, and be stable to normal day to day temperatures.

Caffeine Dehydrogenase: Enzyme Suitable for Caffeine Diagnostic Test Development

Enzyme-based diagnostic tests usually carry all of the above attributes for a rapid diagnostic test. In the previous chapters, caffeine degradation via C-8 oxidation by *Pseudomonas* sp. strain CBB1 was covered in detail. The first enzyme of this pathway, caffeine dehydrogenase (Cdh, EC number 1.17.5.2 and systematic name caffeine:ubiquinone oxidoreductase), is highly suitable for development of a caffeine diagnostic test. The enzyme converts caffeine to TMU and in the process, reduces

quantitatively an electron acceptor (Fig. 6.1) such as nitro blue tetrazolium (NBT, a tetrazolium salt) [Yu *et al.,* 2008]. Details of this reaction and role of Cdh in the C-8 oxidation pathway is discussed in detail in Chapter 1. Cdh assay using NBT is described in Materials and Methods (chapter 2).
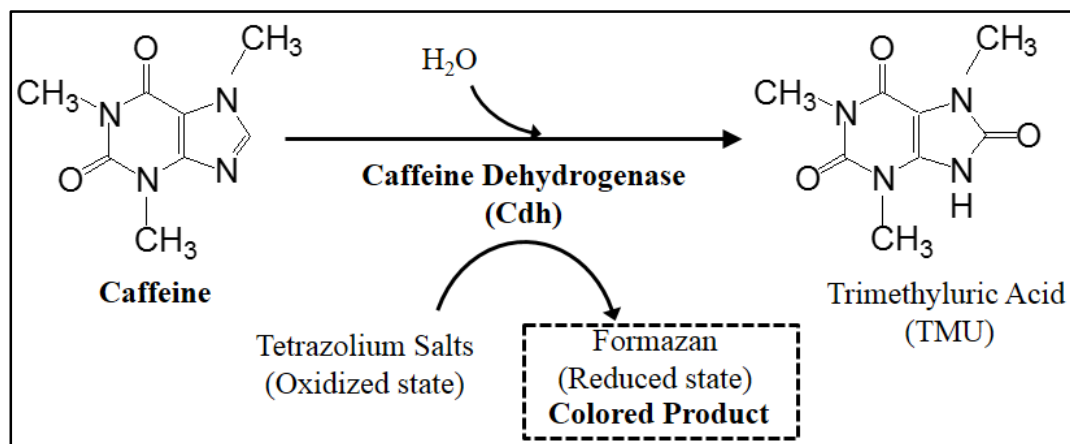
Figure 6.1. Caffeine oxidation to TMU by caffeine dehydrogenase (Cdh) with concomitant reduction of tetrazolium dyes to formazans. Formazans are intensely colored products.

Caffeine dehydrogenase (Cdh) is able to reduce a number of tetrazolium dyes to colored formazan in the presence of caffeine (Fig 6.1). For example, if NBT is used in the above reaction, the color changes from light yellow to dark blue. This reaction was chosen as the basis for caffeine diagnostic test using partially purified Cdh for the following reasons:

i. Specificity for caffeine as substrate: Cdh is highly specific for caffeine with a K_m of 3.7 ± 0.9 µM [Yu *et al.,* 2008] with no activity towards structurally related xanthine or theophylline. Preference for theobromine as a substrate is 50 times less than caffeine [Yu *et al.,* 2008].

ii. <u>Stability of diagnostic test and Cdh at higher temperature</u>: Cdh-based diagnostic test must be suitable to measure caffeine in hot beverages. Mesophilic enzymes are usually unstable or inactive at temperatures above 50°C. In contrast, Cdh is not only active at temperatures above 50°C but the activity linearly increased from 25°C to 66°C [Yu *et al.,* 2008]. Thus the hot beverage could be sampled in small volume and quickly tested without waiting to cool.

iii. <u>Rapid colorimetric results under ambient conditions</u>: Cdh reduction of NBT in the presence of caffeine is very rapid (color development in less than a minute) at pH 7 and temperature 25°C. The reaction is faster at higher temperatures. This makes Cdh highly suitable for rapid diagnostic test development [Yu *et al.,* 2008].

<u>Development of Caffeine Diagnostic Test Using Cdh</u>

Purification of Caffeine Dehydrogenase from CBB1

A detailed three step purification of Cdh enzyme from *Pseudomonas* sp. strain CBB1 is described by Yu *et al*. [2008]. In this work, Cdh was purified for detailed characterization; hence, high purity was needed. However, for the current application, partially purified Cdh was deemed economical and suitable due to lack of interference from other proteins in the preparation. Cdh from crude extract was obtained by a one-step purification from a Phenyl Sepharose column and had a specific activity of 210 Units (U) per mg of protein (Table 6.1). Enzyme recovery was 48.6 % and the preparation had no non-specific reduction of NBT or other interfering components. This partially

purified preparation of Cdh was used for further assay optimization and development of

the caffeine diagnostic test.

Table 6.1. Purification of caffeine dehydrogenase from *Pseudomonas* sp. strain CBB1.

Purification step	Total protein (mg)[a]	Total activity (U)[b]	Specific activity (U/mg)	Purification (fold)	Yield (%)
Crude cell extract	468.0	8200	18	1	100
Phenyl-Sepharose HP	19	3990	210	11.7	48.6

[a]Determined by using the Bradford assay (Bio-Rad) with bovine serum albumin as standard. [b]One unit of activity (U) was defined as the amount of protein required to reduce 1 nmol NBT/min under the assay conditions.

Result I: Testing Various Tetrazolium Dyes for Visual Suitability

Cdh generates a striking purple color with NBT in the presence of caffeine.

However, several commercially available (Sigma chemical company) tetrazolium dyes

were screened to meet the following criteria:

1. Rapid reduction in presence of Cdh and caffeine to give a bright colored
   product relative to background. NBT has a light yellowish background.

2. No color formation in the absence of caffeine

3. Should be reasonably priced and non-toxic.

As shown in Table 6.2, six different tetrazolium dyes were tested at 0.5 mM for

reduction by Cdh in presence of caffeine (0.5 mM) in a total assay volume of 200 μL

potassium phosphate buffer at pH 7.5. All the dyes were purchased from Sigma

Chemical Company. As shown in Table 6.2, all dyes were readily reduced by Cdh in the presence of caffeine except TB and TV. Of the four dyes that showed activity, MTT and TNBT showed strong background color (Table 6.2) which made it difficult to distinguish between reaction and no reaction at lower concentrations of caffeine (data not shown). On the other hand, INT had no background color at all and produced shades of light pink to dark red depending on the caffeine concentration. As mentioned before, NBT has a light yellow color, but upon reduction produced shades of dark blue to dark purple which was quite distinguishable from the background color. Thus, INT was found to be most suitable for further diagnostic test development. It also was the lowest priced that worked dye. NBT was chosen as a back-up for INT. Assuming a 1:1 stoichiometry between caffeine and tetrazolium dye in the reaction (Fig. 6.1), the dye concentration was fixed at 0.5 mM. This ensured that the dye is not limiting in the assay.

Table 6.2. Screening of various tetrazolium dyes for suitability for developing the Cdh-based caffeine diagnostic test.

Tetrazolium dyes	Reduction with Cdh and Caffeine[a]		Background color	Retail price from Sigma (in $/ mg)
Iodonitrotetrazolium Chloride (INT) ✔	Yes		Dark Red	No	0.140
Tetrazolium Blue Chloride (TB)	No		_	No	0.033
Thiazolyl Blue Tetrazolium Bromide (MTT)	Yes		Dark Blue	Yes	0.200
Tetrazolium Violet (TV)	No		_	No	0.055
Nitro Blue Tetrazolium (NBT)	Yes		Purple	Yes	0.400
Tetranitro blue Tetrazolium (TNBT)	Yes		Black	Yes	0.680

[a]Assays were performed as described in Materials and Methods (Chapter-2) with 8-10 U (40-50 µg) of Cdh/assay.

Result II: Optimization of Enzyme Load for Color Development with a Threshold of One

Minute

Given the application envisaged, it is important that the Cdh caffeine diagnostic

test produce visible color within one minute, at the lowest detection limit of caffeine. In

applications like detection of caffeine in mother's milk, about 5 ppm of caffeine is

deemed as the safe upper-limit in breast milk, assuming the mother consumes 500 mg of

caffeine (which is 100 mg higher than the safe upper-limit set by FDA for adults/day

[U.S. Food and Drug Administration 2013]) and about 1-1.5% of it shows up in breast

milk after an hour [http://kellymom.com/bf/can-i-breastfeed/lifestyle/caffeine/, Tyrala

1979]. Thus, for this set of experiments, a fixed concentration of caffeine (5 ppm) was

used with varying enzyme load (Fig. 6.2), and the corresponding time required for

maximum color development was noted (inset table in Fig. 6.2).

Enzyme Units (U)	Time required for maximum color development (min)
3.99	4.52 ± 0.09
9.975	2.42 ± 0.07
19.95	1.22 ± 0.05
39.9	0.63 ± 0.06

20 U enzyme gives maximum color in 1.2 min with 5 ppm caffeine
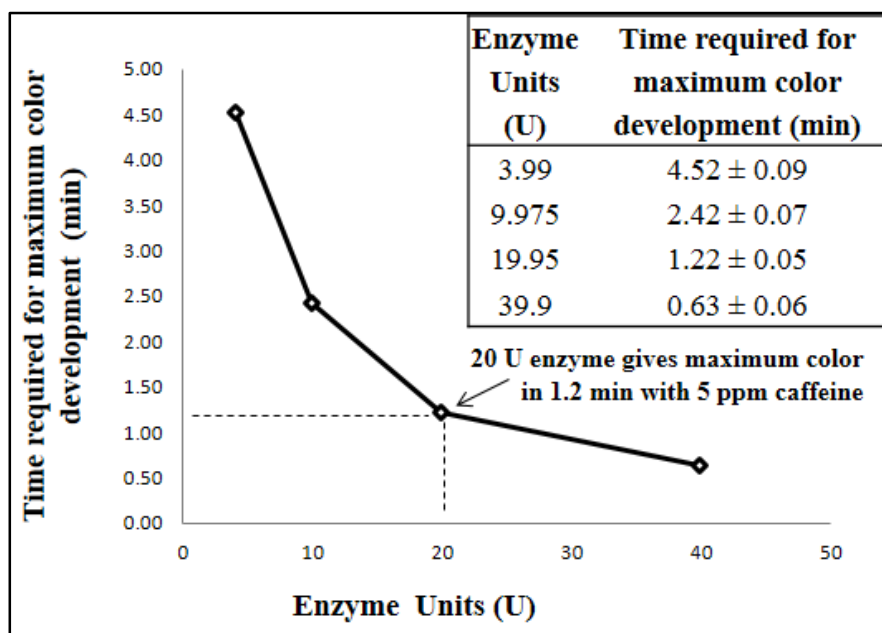
Figure 6.2. Cdh assay at varying enzyme levels, at fixed caffeine concentration (5 ppm).
Inset table shows the numerical data obtained from the experiment.

As expected, increased enzyme load decreased the time required for color development exponentially (Fig. 6.2). Based on this, for maximum color development within a threshold of one minute, the minimum enzyme load was determined to be 20 U. While this enzyme load is adequate for further validation of the test in microtiter plate, the enzyme load may have to be higher to compensate for (i) loss of activity due to immobilization of Cdh in the test strip (ii) shelf life of the enzyme and the dye. These parameters will have to be further optimized during the development of 'at home' format of the diagnostic test.

Caffeine concentration (ppm)	Time required for maximum color development (min)
1	0.22
2.5	0.58
5	1.17
10	1.66
20	3.50
25	4.32
50	8.18
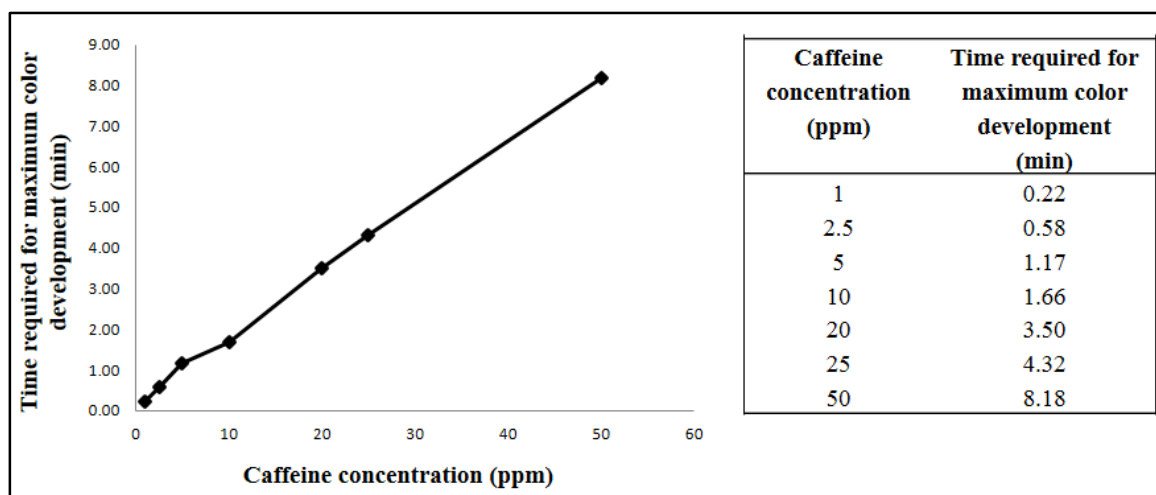
Figure 6.3. Color development with fixed Cdh load (20 U) vs. increasing caffeine concentration (0-50 ppm). Inset Table shows the numerical data obtained from the experiment.

The color development in a diagnostic kit with the threshold of one minute or less, must be independent of caffeine concentration in the test sample. This will demand higher enzyme load than 20 U. The results above shows the relationship between maximum color development time vs. caffeine concentration (Fig. 6.3). At minimum

enzyme load of 20 U, longer time is required for maximum color development with caffeine concentration higher than 5 ppm. Thus, in the diagnostic kit, the enzyme load must be much higher than 20 U, in order to achieve the threshold time of one minute for maximum color development irrespective of caffeine concentration in the test sample. Taking into consideration the threshold of one minute for color development and independent of caffeine concentration, the actual enzyme load will be re-optimized during the development of 'at home' format of the diagnostic test.

Result III: First Level Attributes for Cdh-Based Caffeine Diagnostic Test

The first level key attributes established for further validation of the test with real world samples (caffeinated and caffeine-containing beverages, nursing mother's milk and Pharmaceuticals) are summarized below.

1. Minimum Enzyme Loading: 20 U Cdh

2. Tetrazolium Dye: Iodonitrotetrazolium Chloride (INT)

3. Dye Concentration: 0.5 mM = 97.1 ppm

4. Color to detect caffeine: Light pink (at 1 ppm caffeine) to readily differentiated shades of red at 5, 20, 100, >>100 ppm)

5. Threshold for caffeine detection: 1 minute

6. Background color: None

The above attributes are optimized for the current 96-well micro plate assay with a total reaction volume of 200 μL. Further optimization will be needed for the development of the "at home" diagnostic test-strip. For example, the enzyme and dye load will have to be increased in the final immobilized formulation on the strip to compensate for the:

a. Shelf-life activity loss

b. Activity variation due to pH of the test fluid

c. Activity loss due to immobilization and formulation

d. Activity loss due to presence of any inhibitory components in the sample.

Result IV: Validation of the Color Development at 1 ppm Caffeine (Lowest Limit)

In mother's milk, a safe limit of caffeine is between 0-4 ppm, assuming that the mother consumes 0-400 mg of caffeine (which is safe limit set by FDA for adults/day [U.S. Food and Drug Administration 2013]) with about 1-1.5% of it partitioning into breast milk in one hour [http://kellymom.com/bf/can-i-breastfeed/lifestyle/caffeine/, Tyrala 1979]. Thus, the lowest detection limit for a diagnostic test should be ~1 ppm (on a ppm scale) in a threshold time of 1 minute, in order to designate safe/unsafe for caffeine in such fluids. The color development for Cdh-based test is shown below (Fig. 6.4).
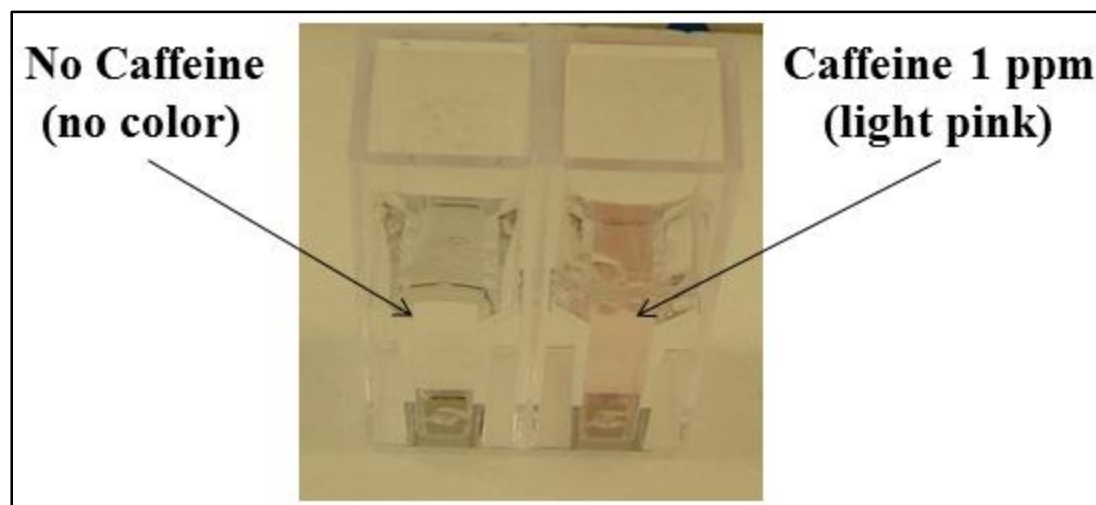
Figure 6.4. Color development with Cdh-caffeine diagnostic test at 1 ppm caffeine. The control reaction on the left has no caffeine, but only INT. Color development time was one minute.

It is clear from the figure that the light pink color is distinguishable within the one minute threshold time. There is no background in the control where caffeine is absent. These results establish the fact that the Cdh-based caffeine diagnostic test can detect caffeine at a concentration as low as 1 ppm within the threshold time.

<u>Validation of the Caffeine Diagnostic Test</u>

Validation of the Cdh-caffeine diagnostic test with "real world samples" is very critical. Real world samples are different in that they are finished products in complex formulations. Conditions may not be as optimized when these samples are used for caffeine test due to interfering components. Also, pH of the test solution may not be optimum for the enzyme. Background color of the test samples will also be an issue; for example the color of coffee, colas and milk. For Cdh-based caffeine diagnostic test validation, three different categories of commercial samples containing caffeine were

chosen. These include (i) pharmaceuticals formulated with caffeine, (ii) caffeinated and caffeine-free beverages and (iii) milk samples (spiked with caffeine).
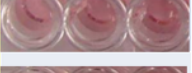
Validation Test I: Semi-Quantitative Estimation of Caffeine in Pharmaceuticals

The first validation for the Cdh-based caffeine diagnostic test was done with pharmaceuticals containing caffeine. Interestingly, caffeine is added in the final formulation of a number of drugs [Durrant *et al.,* 2002]. Caffeine is not an active ingredient in these drugs, but is present along with other components, including formulating agents. Thus, testing the caffeinated pharmaceuticals would be an excellent validation of the diagnostic test. Caffeine concentrations in the drug formulations are already in the label. We further confirmed this by HPLC as described in the Materials and Methods (Chapter II). The caffeine levels in these drugs can be tested for validation of the Cdh-based diagnostic test.

For this validation test six commercially available caffeine containing pharmaceuticals were chosen namely, Diurex (a diuretic), Menstrual Relief, Vanquish, Midol, Excedrin (pain relievers), and Vivarin (a stimulant). Caffeine amount in these drugs as reported by the manufacturer is shown in Table 6.3. Prior to testing of the drugs, one tablet from each of the listed pharmaceutical was powdered separately, dissolved in deionized water, and diluted appropriately to bring the caffeine concentration range to 5-50 ppm (see column 3 of Table 6.3). Correspondingly, a reference color chart was prepared by spiking known concentrations of pure caffeine purchased from Sigma Chemicals, and developing the shades of red color using the Cdh-diagnostic test in the range 0-60 ppm (bottom of Table 6.3). Caffeine concentration in the diluted samples

were determined in triplicate, using HPLC (as described in Materials and Methods, Chapter II). Concurrently, caffeine in these samples were measured using the diagnostic test in triplicate (last column in Table 6.3) and compared with the color reference chart to obtain caffeine concentration.

Table 6.3. Validation of Cdh-caffeine diagnostic test for semi-quantitative estimation of caffeine in six pharmaceutical preparations.

Name of the Drug	Manufacturer specified caffeine amount/ tablet	Caffeine after dilution (in ppm)	As determined by (in ppm)		Color developed in triplicate
			HPLC	Caffeine diagnostic test	
Diurex (Diuretics)	50 mg	25	25.05 ± 0.1	25 ± 5	
Menstrual Relief (Pain Relievers)	60 mg	30	29.67 ± 0.29	30 ± 5	
Vanquish (Pain Relievers)	33 mg	33	31.18 ± 0.64	30 ± 5	
Midol (Pain Relievers)	60 mg	6	5.675 ± 0.19	5	
Excedrin (Pain Relievers)	65 mg	6.5	5.65 ± 0.36	5-10	
Vivarin (Stimulants)	200 mg	20	19.39 ± 0.66	20 ± 5	
Color Reference Chart (caffeine concentration in ppm)		0 5 10 20 30 40 50			

[a] ± represents standard deviations obtained from three independent HPLC measurements.

Reference color chart for caffeine concentration from 0-50 ppm is shown at the bottom of the table.

As shown in Table 6.3, the caffeine concentration readings obtained from the caffeine diagnostic test closely matches the HPLC results. The Cdh diagnostic test gave credible results in the range 5 - 60 ppm. The components in the drug formulations did

not interfere with caffeine-dependent color development. Also, color development was within the threshold time of one minute. These results show that the Cdh-based caffeine diagnostic test works very well with commercial drugs formulated with caffeine. All the test parameters established earlier were also confirmed by this test. It is tempting to propose that the Cdh-based diagnostic test could be useful in pharmaceutical formulation laboratories to establish the caffeine concentration for labeling. It will be faster, easier, simpler and cheaper compared to the laborious, complex and expensive HPLC procedures.

Validation Test II: Rapid Determination of the Presence of Caffeine in Commercial Beverages (Coffee and Soft-Drinks)

The second validation of the Cdh-based caffeine diagnostic test was performed with caffeine containing beverages like coffee, soft drinks, and their decaffeinated counter parts. For this test, two sets of caffeine containing beverages were selected

Set-1: Regular coffee of Dunkin Donut brand, half-decaffeinated coffee from Folgers and Decaffeinated coffee from Dunkin Donut, with caffeine concentration of about 600 ppm, 330 ppm, and 20 ppm, respectively (Table 6.4).

Set-2: Regular Coca Cola with caffeine concentration of 96.4 ppm and caffeine free cola (zero caffeine); Regular Pepsi with caffeine concentration of 107 ppm and caffeine-free Pepsi; and regular Mountain Dew with caffeine concentration of 152 ppm and caffeine-free diet mountain dew (Table 6.4).

Coffee extract used in the above test (Table 6.4) was obtained by boiling 10 grams of coffee powder for 10 minutes with 200 mL of deionized water and filtering

using a filter paper. Caffeine content in the extract was determined by HPLC. Caffeine

content in soft drinks listed in Table 6.4 is based on the label in the can. The caffeine

content of the soft drinks were also determined by HPLC. The caffeine content on the

label and that determined by HPLC were close, but not exact.

Table 6.4. Cdh-based caffeine diagnostic test validation using various caffeinated and
non-caffeinated beverages.

Type of Beverage	Manufacturer specified caffeine content (ppm) [a]	As determined by HPLC (ppm) [b]	Cdh-caffeine Diagnostic Test	
Coffee			**Blank**	**Assay**
Dunkin Donut (Regular)	NA	601.6 ± 21.4		
Folgers (50% Decaffeinated)	NA	331.6 ± 5.9		
Dunkin Donut (Decaffeinated)	NA	20.4 ± 1.3		
Soft Drinks (Caffeine free *vs* Regular) [¥]			**Caffeine Free**	**Regular**
Coca-cola	96.4	83.28 ± 4.2		
Pepsi	107	96.14 ± 5.6		
Mountain Dew	152	142 ± 9.4		

[a] Powdered coffee either from Dunkin Donut or Folgers do not have manufacturer
specified caffeine content in the label. [b] Caffeine values reported for coffee is based on
brewing methods described below followed by analysis with HPLC. Caffeine content
determination by HPLC was done in triplicate and reported as average ± standard
deviations. [¥] For soft drinks, caffeine content is based on manufacturers' specifications
and HPLC analysis.

Beverages listed in Table 6.4 were analyzed for caffeine using the Cdh-based

diagnostic test. The objective here was qualitative; to validate the diagnostic test to

differentiate between presence, absence and low level of caffeine in decaffeinated coffee. The fact that many of these beverages have a background color is also a challenge; the color formation with caffeine should still be readily differentiable relative to background. The results with the coffee types were very clear. There was intense color produced with caffeinated and 50% decaffeinated coffee (Table 6.4). In contrast, there was very light color with decaffeinated coffee (~20 ppm caffeine). The background color of the coffee extracts did not interfere with Cdh-based diagnostic test in terms of clearly and instantly differentiating caffeinated coffee from decaffeinated. Interestingly, the color is slightly lighter in 50% decaffeinated coffee (330 ppm caffeine) relative to regular (600 ppm caffeine).

Likewise, the diagnostic test instantly and readily differentiated the caffeinated beverages from caffeine-free beverages. These beverages, which also contain other ingredients in the form of natural (milk and/or sugar) and artificial chemicals, coloring agents, etc., did not interfere with the test. This test further affirmed that the Cdh-based caffeine diagnostic test is indeed robust and can be used to quickly differentiate caffeinated drinks from decaffeinated or non-caffeinated counterparts. This test could be very useful to avoid mix-ups in restaurants and other outlets where all these beverages are served extensively in non-labelled cups.

Validation Test III: Rapid Determination of Unsafe Levels of Caffeine in Nursing Mother's Milk

The importance of caffeine restriction for pregnant woman and infants was well documented in the beginning of this chapter. Possible teratogenic effects of excess

caffeine on infants was also documented. To the best of our knowledge, there is no rapid diagnostic test for caffeine that is commercially available for testing nursing mother's milk. Hence we tested the validity of the present Cdh-based test to detect caffeine in various milk, including nursing mother's milk. Milk is a complex mixture of proteins, fats, sugars and other ingredients, including calcium and vitamin D. In addition, nursing mother's milk has immunoglobulins, metal ions, cofactors, hormones and other biochemicals. It was of great interest to see if any of these components in milk would interfere with the caffeine detection using the Cdh-based test. The result of the Cdh-based diagnostic test for caffeine detection in various milk samples is shown in Fig. 6.5. For this specific test, both INT and NBT was used, assuming that there might be non-specific reduction of these dyes and one of them would still be able to function in milk. Also, since milk does not have caffeine, all the tests were done with the substance spiked into the fluid.
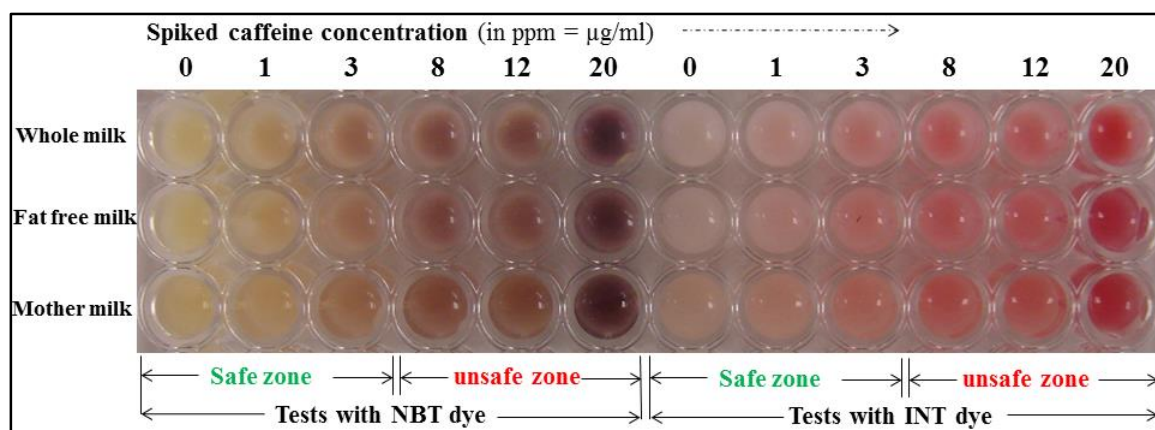
Figure 6.5. Cdh-based caffeine diagnostic test with milk samples containing spiked caffeine as noted. Commercially available whole and fat free milk was used for the test. Nursing mother's milk was obtained from a nearby laboratory. Where indicated, caffeine was spiked to the indicated amounts. Left panel is NBT-based test and right, INT-based. The duration of the test was one minute.

As shown in Fig. 6.5, the Cdh-based test scored very well in all the tests. Both dyes were able to differentiate safe (3-8 ppm) and unsafe levels of caffeine in all milk samples, especially mother's milk. The complex components in the milk did not interfere with the diagnostic test and the rapidity of color development. Clearly, the test is suitable for a nursing mother to check the milk using the Cdh-based test, and instantly determine if the caffeine level is safe for nursing the infant. There is light background color with the dye at "0" caffeine level; this might be due to very low-level reduction of the dye by milk components (Fig 6.5). Nevertheless, this does not interfere with the speed or the ability of the diagnostic test to determine safe and unsafe levels of caffeine in milk.

Discussion

Comparison of the New Cdh-based Caffeine Diagnostic Test with Other Known Caffeine Detection Methods

Caffeine is a widely popular psychoactive drug and has been a significant component of beverages and other food items in an ever increasing trend [Karen *et al.*, 2002]. However, negative effects of excess caffeine is well documented. FDA recommends no more than 400 mg of caffeine/day/adult and also considers caffeine concentration of above 200 ppm in food and beverages as unsafe [U.S. Food and Drug Administration 2013]. There is growing concern in FDA about introduction of caffeinated products, including chewing gums and their unrestricted use especially among children [Tomson *et al.*, 2013]. In numerous publications and public information portals like FDA websites, etc., negative effect of caffeine is reviewed and drawing increased attention [Garriott *et al.*, 1985; Eteng *et al.*, 1997]. This has resulted in several

attempts to create a rapid caffeine detection kit, which are listed in Table 6.5. None of these are available commercially, probably due to the merits (or demerits) of these tests, which is compared in Table 6.5.

Table 6.5. Evaluation of various caffeine detection methods along with currently developed Cdh-based caffeine diagnostic test, for suitability for various applications.

ASSAY METHODS	ASSAY ATTRIBUTES				
	In-situ	Rapid (< 1 Min)	Sensitivity	Easy to Read	Commercial Availability
Bacterial-cell based	NO (DO probe)	NO (> 3 Min.)	POOR [100 ppm]	YES (Digital Output)	NO
Lipid Polymer Membrane based	NO (Electrode)	YES (< 30 sec.)	Data Not Available	YES (Digital Output)	NO
Fluorescent Dye based	YES (Test strips)	NO (> 20 Min.)	POOR [100 ppm]	NO (UV light, TLC)	NO
Enzyme based	YES (Test strips)	Unknown	Data Not Available	YES (Colorimetric)	NO
Antibody based (D+caf)	YES (Test strips)	YES (< 60 Sec)	POOR [124 ppm] Limited Use¥	YES (Colorimetric)	Discontinued
✔ **Cdh Enzyme based Test (current work)**	**YES**	**YES** (< 60 Sec)	**YES** [< 1 ppm]	**YES** (Colorimetric)	**Work in Progress**

¥ Does not work with milk and sugared drinks. Test has to be done before addition of milk and sugar to coffee.

As shown in Table 6.5, almost half-a-dozen diagnostic tests were found in literature for detecting caffeine ranging from using live bacterial whole-cells [Sarath Babu *et al.*, 2006], enzyme/antibody-based methods [Scherl *et al.*, 1997 US patent # 5610072A, Akyilmaz *et al.*, 2010] to other non-biological methods like lipid/polymer membranes [Shen *et al.*, 2008] or fluorescent dyes [Luisier *et al.*, 2012; Rochat *et al.*,

2011]. The enzymatic method is based on xanthine oxidase [Scherl *et al.*, 1997 US patent # 5610072A] or alkaline phosphatase [Akyilmaz *et al.*, 2010]. The antibody method is based on caffeine monoclonal antibodies [Franco *et al.*, 2010, Graham *et al.*, 2012 US patent # 8137984B2]. As shown in Table 6.5, none of these tests meet the desired attributes of "at home or portable caffeine diagnostic test". Thus, it is not surprising that none of the existing methods are commercially viable for rapid caffeine detection.

The bacterial immobilized cell-based assay is an electrode based amperometric test in which a Dissolved Oxygen (DO) probe measures oxygen depletion as a function of caffeine metabolism [Sarath Babu *et al.*, 2006]. This method was not found to be suitable as it could not meet any of the required criteria like an *in-situ* method (bulky electronic instrumentation), or rapid (response time of more than 3 minutes), or required sensitivity (>100 ppm). Similarly, the lipid/polymer membrane based method (also known as electronic tongue) is also electrode based amperometric test [Shen *et al.*, 2008], which claimed to be rapid (response time less than half a minute), but has similar drawbacks as the bacterial cell-based method. Thus, this test was also not commercially suitable. In 2011 and 2012, two methods [Luisier *et al.*, 2012; Rochat *et al.*, 2011] were described based on binding of fluorescent dyes to caffeine on a thin-layer chromatography (TLC) plate and visualizing using UV light. Although these methods are simple compared to bacterial cell-based method, the test failed to meet the other criteria like rapid detection (more than 20 minutes for color) and sensitivity (cannot detect <100 ppm).

Enzyme and antibody-based methods are known for high specificity for substrates and ligands. Both xanthine oxidase-based method and anti-caffeine MAb-method (Table

6.5) were colorimetric tests that met most of the required attributes, except sensitivity. While the sensitivity of the anti-caffeine MAb-based method was found to be about 124 ppm (D+caf manual), the sensitivity of enzyme-based method was not reported. In addition, the other notable drawbacks of these methods are (i) the enzyme-based method used xanthine oxidase (which is not a caffeine specific enzyme) and (ii) the antibody-based method (which is the only method that had resulted in development of a commercial caffeine test strip called D+caf) was unable to measure caffeine in milk and sugared beverages. Thus, neither of these methods found commercial success in the form of a caffeine diagnostic test.

In contrast, the presently developed Cdh-based caffeine diagnostic test meets all the requirements for various applications. Cdh has high specificity for caffeine [Yu *et al.*, 2008], unlike xanthine oxidase. The sensitivity for caffeine detection by Cdh is 1 ppm, and is highly suitable for detection of caffeine in nursing mother's milk. This test can, within one minute, diagnose the safe and unsafe levels of caffeine in mother's milk to make a decision with respect to feeding the infant (Fig. 6.5). The complex biological components, including metal ions and immunoglobulins, do not interfere with the Cdh-based diagnostic test (Fig. 6.5). The diagnostic test is also applicable for hot or cold beverages to detect the caffeine level and differentiate decaffeinated products (Table 6.4). Again, this will be very useful for pregnant women, where FDA limits safe level of caffeine to 400 ppm/day. The Cdh-based caffeine diagnostic test is rapid (takes less than a minute to develop color). The test successfully detected caffeine in the formulations of several drugs; the caffeine level matched that determined by HPLC (Table 6.3). Again, the formulation components in the drug did not interfere with the test. The Cdh-based

diagnostic test is robust and versatile in terms of applications for various purposes. It can be used for qualitative tests (Yes/No caffeine levels, Safe/Unsafe caffeine levels) and for quality control tests for caffeine in drug formulations. Complex and expensive HPLC procedures are not needed on a routine basis.

The Cdh-diagnostic test needs further development to make it a commercial kit. These include, but not limited to, the following:

1. Development of a dip stick method by immobilizing Cdh and INT dye on a test strip.

2. Determination of shelf-life of Cdh and dye at various temperatures.

3. Formulation of Cdh for appropriate activity levels due to

   a. Shelf-life activity loss,

   b. pH of test samples,

   c. Presence of inhibitory compounds in test samples.

4. Other quality control issues to optimize the diagnostic assay for quick visual analysis (< 1 min) with various caffeinated beverages, especially nursing mother's milk.

5. Production of a reliable reference color-chart to facilitate the semi-quantitative caffeine determination (1-100 ppm)

6. Potential extension of this diagnostic test to other applications like environmental sampling (required sensitivity 1 ppb).

Recently, caffeine has been designated as a marker of ground water pollution due to urbanization [Seiler *et al.*, 1999; Buerge *et al.*, 2003; Dhaneswar *et al.*, 2012]. Pollution of ground and surface water by man-made chemicals like DTT, Atrazine and other pesticides have always been a serious issue for the last few decades mostly due to their excessive use in farming and their gradual leaching from farm lands to nearby water bodies. In recent years, studies have shown that not only farmlands, but also big cities can contribute to surface and ground water pollution. For example, highly populated urban areas generate a lot of polluted water, which are either wastewater from daily human activities or water seeping from landfills [Buerge *et al.*, 2003]. The polluted water contains high levels of unnatural chemicals such as pharmaceuticals [Seiler *et al.*, 1999], which are typical to polluted water originating from urban sources and are highly toxic to marine life of nearby water bodies even in minute quantities. Interestingly, it was found that caffeine is also found in these polluted waters coming from urban areas [Dhaneswar *et al.*, 2012].

Researchers currently believe that if a quick test can detect the presence of unnatural levels of caffeine in water bodies connected to urban areas, then pollution of these water bodies can be checked actively and in a preventive way before any destruction to ecosystem [Buerge *et al.*, 2003; Dhaneswar *et al.*, 2012]. However, the level of caffeine in these polluted water samples is at parts-per-billion level (ppb). Our currently developed Cdh-based caffeine detection test is about a 100-fold less sensitive (detection to 1 ppm level). Another future application of Cdh-based caffeine diagnostic

test is to make it more sensitive by amplification of caffeine detection signal to ppb. This will open up a whole new environmental application for pollution control.

<div align="center">Summary and Conclusion</div>

FDA characterizes caffeine as a psychoactive drug. Excessive consumption of this substance is not encouraged; for specific age or health groups like infants, seniors, pregnant-women and nursing mothers, FDA recommends restriction of caffeine consumption. This led us to develop a rapid diagnostic test for caffeine detection that is also portable. This will enable rapid testing of caffeine levels in various beverages by the age or health groups with caffeine-restricted diet, to make proper decisions. This test will be particularly useful for pregnant women and nursing mothers who are severely caffeine restricted because its presence even in minute amounts in blood or milk might result in possible teratogenic effects on growing fetus or on the breast feeding infant.

The presently developed caffeine diagnostic test is based on colorimetry, using caffeine dehydrogenase (Cdh) and a tetrazolium dye (INT). Partially purified Cdh from *Pseudomonas* sp. strain CBB1 was used for the test, which has suitable properties like thermostability and high specificity for caffeine. Of the various tetrazolium dyes tested, INT was the best with no background color and readily recognizable intense red color at various caffeine concentrations. The test was developed with the following attributes: enzyme load (20 U Cdh), dye load (0.5 mM INT), time for color development (≤ 1 minute), color range to detect caffeine (no color in absence of caffeine, light pink at 1 ppm to shades of red at increasing levels of caffeine with color saturation at > 100 ppm). The test was also validated with the following commercial products containing caffeine:

pharmaceuticals (semi-quantitative), beverages (coffee and soft drinks), and milk (commercial and nursing mother's milk spiked with caffeine).

There are several caffeine diagnostic methods described in literature such as bacterial cell based, membrane based, fluorescent dye based, enzyme based, and antibody based (D+caf). None of these methods meet all the desired attributes required for a suitable caffeine diagnostic test, e.g., being portable, rapid, and sensitive at ppm level. The presently developed Cdh based diagnostic test fits all the required attributes. Our future plan is to develop this into a dip stick test for general public use and industrial use, such as in a pharmaceutical quality control laboratory. If the sensitivity of the test can be increased about 100-fold then the application can be extended to detect ppb levels of caffeine in environment water samples. Caffeine is recognized as a marker for human pollution activities due to its extensive use and consequent release into the environment.

CHAPTER 7

SUMMARY OF COMPLETED RESEARCH AND SUGGESTIONS

FOR FUTURE WORK

Soil bacterium *Pseudomonas* sp. CBB1 is capable of growing on caffeine as a sole source of carbon and nitrogen. CBB1 degrades caffeine via the C-8 oxidation pathway, where caffeine gets converted to TMU. Previously, a novel caffeine dehydrogenase was shown to catalyze this conversion. However, little was known about the cofactor content of this heterotrimeric enzyme or the fate of TMU. In the present work, based on sequence homology of *cdhA,B,C* (Cdh genes located on 25.2-kb caffeine gene cluster from CBB1) CdhA was identified as carrying a Molybdopterin cofactor, while CdhB and CdhC as FAD and iron–sulfur cluster-containing, respectively.

Secondly, a new FAD-containing NADH-dependent trimethyluric acid monooxygenase (TmuM) was purified from CBB1, which oxidized TMU to TM-HIU. TM-HIU was found to be unstable and spontaneously degraded to TMA. Subsequently, TMA was isolated, purified and fully characterized with respect to structure and molecular properties. The compound formed spontaneously was racemic. In contrast, a time-course reaction showed that TMA formed from TM-HIU was 93% enriched with one of the enantiomers (proposed to be S-(+)-TMA based on uric acid metabolism), which racemized over time. This is the first report of the enzymology of the caffeine C-8 oxidation pathway beyond TMU, to TM-HIU and then to S-(+)-TMA via TM-OHCU. TM-HIU was identified as a transient product while TM-OHCU (unstable) was proposed based on similarity to uric acid pathway. None of the metabolites downstream of TMU

are available commercially. The present work has opened a very interesting area for future research in terms of stabilizing TM-HIU by derivatization and further structural characterization. Likewise for TM-OHCU and S-(+)-TMA.

Biochemical characterization of TmuM suggested that it is a single subunit protein with an apparent molecular weight of ~43-kDa, and contains 1 molecule of FAD bound non-covalently/subunit. TmuM was found to be oxygen dependent, and highly specific for methyluric acids; the enzyme had no activity on uric acid. An homology model of TmuM indicated that the active site of TmuM is larger and more hydrophobic than to the analogous FAD-containing urate oxidase/monooxygenase (HpxO). Future work on the crystal structure of TmuM can offer further insight into the catalytic mechanism and specificity of this enzyme, compared to HpxO [Hicks *et al.,* 2013].

To further delineate the enzymatic conversion of TM-HIU to TMA and beyond, a 25.2-kb caffeine gene cluster was sequenced from a CBB1 genomic library. Three new genes, *tmuH*, *tmuD*, and *orf1* were identified as encoding TM-HIU hydrolase (converts TM-HIU to TM-OHCU), TM-OHCU decarboxylase (converts TM-OHCU to S-(+)-TMA) and trimethylallantoinase (converts S-(+)-TMA to TMAA), respectively. Further, gene cluster analysis revealed *orf2* and *orf3*, which are proposed to encode hydrolase and deacetylase type enzymes involved in TMAA catabolism. Gene cluster analysis helped in refining the entire C-8 oxidation pathway in CBB1. Any clarity on the biochemistry of formation and metabolism of S-(+)-TMA will come from functional expression of these putative genes along with chemical characterization of S-(+)-TMA. This future work can lead to:

    I.   Confirmation of functional assignments of *tmuH, D and orf1-3*

II. Biochemical characterization of the above enzymes,

III. Structural and mechanistic studies of the gene products, and its differentiation from the uric acid pathway enzymes, HIU hydrolase [French *et al.*, 2010] and OHCU decarboxylase [French *et al.,* 2011].

IV. Identification of metabolites and the reaction sequence by NMR.

One of the genes missing in the 25.2 Kb caffeine gene cluster is trimethylallantoin racemase. This is another interesting extension of the present research. Comparative geneomic and proteomic analysis of CBB1 with the organisms that elaborate the uric acid pathway can elucidate details about these missing genes/enzymes. Identification of the terminal metabolites of the C-8 oxidation pathway namely, glyoxylate and di-and monomethylurea is also a rich area of future research (possibly via *in situ* NMR).

Finally, Cdh was found to be ideally suited for development of a dip- stick caffeine diagnostic test. The enzyme suitability was validated with respect to attributes such as speed of caffeine diagnosis (<1 min), dye choice (INT), sensitivity (1 ppm), semi-quantitation of caffeine in drug formulations, differentiation of decaffeinated beverages and detection of caffeine in nursing mother's milk. This work needs to be extended to develop a commercial test strip with Cdh and INT. Complete validation of the diagnostic strip with respect to shelf-life, formulations, loss in activity due to pH variation, inhibitory compounds, etc., must be done before commercialization. If the sensitivity of the test can be increased about 100-fold then the application can be extended to detect ppb levels of caffeine in environment water samples. This would be a valuable tool for environmental applications since caffeine is a marker for human pollution activity.

APPENDIX A

STRAINS, PLASMIDS AND PRIMERS USED IN THIS STUDY

Table A1. A complete list of all the strains and plasmids used in this study.

Bacterial strains	Relevant characteristics	Source or reference(s)
E. coli		
DH5α	Δ(*lacZYA-argF*)*U169 hsdR17 relA1 supE44 endA1 recA1 thi gyrA96* φ80d*lacZ*ΔM15	Life technologies, Gaithersburg, Md.
BL21(DE3)	F- *ompT hsdS*B (rB-mB-) *gal dcm* (DE3)	Invitrogen
EPI300™-T1R		Epicentre.
Pseudomonas sp. Strain CBB1	Caffeine-degrading strain	Yu *et al.,* 2008

Plasmids	Relevant characteristics	Source or reference(s)
pMVS848a	Ap^r, *tmuMO* (encoding the trimethyluric acid monooxygenase) under the control of the T7 promoter of pET32a	This study
pET32a	Ap^r, T7 expression vector	Novagen.
pSC-B-amp/kan	Cloning vector in StrataClone Blunt PCR Cloning Kit	Agilent Technologies
pMVS848	Cm^r, pCC2FOS containing *cdhABC* gene encoding caffeine dehydrogenase and *tmuMO* encoding the trimethyluric acid monooxygenase from *Pseudomonas sp. CBB1*	This study
pCC2FOS™	Vector in CopyControl™ HTP Fosmid Library Production Kit	Epicentre.

Ap^r, ampicillin resistance; Cm^r, Chloramphenicol resistance.

Table A2. A complete list of all the primers used in this study and their sequences.
Sequences are listed from 5′ to 3′ end of each primer.

Primer Name	DNA sequence (5′-N$_x$-3′) (restriction enzyme sites are underlined)	Comment
tmuMO-degF1	5′- AARATHGGNGCNGAYGTNAC-3′	Degenerate forward and reverse primers for *tmuM* studies
tmuMO-degF2	5′- ATHGGNGCNGGNATHGGNGGN-3′	
tmuMO-degR1	5′- NGGRTGNGCNGCRTCNCC-3′	
tmuMO-degR2	5′- NCCRTCNGCNGCDATNARRTA-3′	
tmuMO-up1	5′- CGACATTTCGCGTATCACACG -3′	Specific forward and reverse primers for sequencing flanking sequences of *tmuM*
tmuMO-up2	5′- CAGTCGGTCAGTTCCTCCGGT -3′	
tmuMO-down1	5′- TCGCGCCTCGATCCGGCCCTG -3′	
tmuMO-down2	5′- TCCTTTGGCTTGAACCTGACC -3′	
tmuMO-F-Nde I	5′-GTGCCG<u>CATATG</u>AGCCGTCCG-3′	Specific forward and reverse primers used for *tmuM* cloning
tmuMO-R-Not I	5′- ACCAT<u>GCGGCCGC</u>GAAGCCAG -3′	
Cdh-AF1	5′-ATGTTCGCSGAYATHAAYAARGG-3′	Degenerate forward and reverse primers for *cdhA* studies
Cdh-AF2	5′-ATGTTTGCWGAYATHAAYAARGG-3′	
Cdh-AR1	5′-CCYTTRTTDATRTCSGCGAACAT-3′	
Cdh-AR2	5′-CCYTTRTTDATRTCSGCGAACAT-3′	
ACF	5′- GGCGAGAACTTGAATGTC-3′	Primers used to screen the CBB1 genomic library
ACR	5′-TCCGGCAGCTCGATGATT-3′	

APPENDIX B

DNA SEQUENCE OF CAFFEINE GENE CLUSTER IN CBB1

DNA Sequence

The complete DNA sequence of the caffeine gene cluster (25.2-kb) containing the

genes responsible for caffeine transformation in *Pseudomonas* sp. strain CBB1. A

physical map of this gDNA segment of strain CBB1 containing caffeine gene cluster is

illustrated in Figure 5.2 of Chapter-5. For better visualization, start and stop codons have

been **bolded**, and the corresponding open reading frames are *italicized*.

Table B1. Predicted genes using the GeneMark Program of DNASTAR software.

Gene	Strand	Left End	Right End	Gene Length (bp)
orf-1	+	138	329	192
orf-2 (tmuH)	-	468	812	345
orf-3 (tmuD)	-	812	1687	876
orf-4	+	2421	3341	921
orf-5 (tmuM)	+	3364	4554	1191
orf-6	+	4551	5228	678
orf-7	+	5225	6358	1134
orf-8	+	7903	9474	1572
orf-9	+	9471	10562	1092
orf-10	+	10572	11495	924
orf-11	+	11533	12612	1080
orf-12	+	12768	13577	810
orf-13 (cdhA)	+	14359	16734	2376
orf-14 (cdhB)	+	16725	17618	894
orf-15 (cdhC)	+	17623	18126	504
orf-16	+	18123	18572	450
orf-17	-	18559	18879	321
orf-18 (orf1)	+	19308	20708	1401
orf-19 (orf2)	+	20736	21536	801
orf-20 (orf3)	+	21910	23199	1290
orf-21	+	23347	24378	1032
orf-22	+	24409	24663	255
orf-23	+	24729	>25232	504

GACGATCATGTGACGTTTTCTTCGAACTCAACGTCCAGCCGCGCGCGTCAAA
GTGCCGGATCGGCGCATGGCGCATTCGCGTCCAAGGCAAGGCGCAAGGCAG
AGCCCAGCAATGATGCCCATATCTGGCGCCGCCG**ATG***CCGGGTGTCGTCTCGA*
CGCTCGCCGTCGCGGCCGGCCAGCCCATCAAGGCCGGCGACGTGCTGCTGTCGA
TCGAGGCGATGAAGATGGAAACCGCACTGCACGCCGAGCGCGACGGCACGATCG
CCGAAGTGCTGGTGCGCGCCGGCGACCAGATCGACGCCAAGGACCTGCTGGTCG
*TCTACGGC***TGA**GGCCGATGTTGGCTAAGCTGGGCCAAAGTTACGCGATCGGC
ACAGCATCAATCAAACAATCGCCTGAAGAGATTACCTGGACGACTTTTGGCC
GCCAGCCCGTCACCGCGAAACGCTTCGTGTAGGCAAAGCAAATCA**TCA***AGTC*
CCCATGTAGCTTTGATATCCCCACGGTGCCACAACTAGGGGTACGTGATGGTGTAC
ATGGGTACTGACAATCTTGAACCGCACTGGCACGAGTTCGAAGAACGTGGCGTCG
GTGCCACCGCGGCGAAAATAATCCTTTACATGGAACATCAACTCGAATGTTCCTGA
CATCATATCTGATGAGGTGAGAACCGGTTCGTCGGTCCGACCCTCGGCATTGGTG
AAGAGTGTCTTCAAGTGAAACTTATCTGTGCCGTCCAGGCGCAAGAAATCGATCCG
CATGCCGCCGCCCGGGATACCATTCATTATGTCGAGTACGTGTACGGTCAAACCA
*GC***CATTA***ACGGACATCCTCGCGTGTTGGCCCGGAAGCGAGCGCCGGAATGCCG*
CCCTTGTAAACTGAAAAGGCATAGCGCCCTACAATAAGTGGGATATGGTAATTCTG
GTAGGGATCATCGACCTCGAACAACACGGGTATCACGTCAATAAACGTGCTGGGA
GTGCCGTGCTCGGCGAGATATGCTCCAGCCTGCAGGATCATCTGGTAACGGCCCG
GCCGCATTTCCTCGGCGTCAAGCAGTTGGGCGCGGCCGACTTCATCTGTCGTCAC
CCTTGCGACTGTCGTGTATCCGCCACCTTCGAACCGCTGCAATTCGATCACCATTC
CCTGCGCACCACCGGCATTGCCGCTGTCGAGGACATGGGTCGAGATTGCGCGGC
GTGTCCTCTGGCTCTCCCTAACCCGTTCGCGCAACCGGCGCTCCGTAATCTGCCC
GACCTGTTCGATGCAATTTTCGATCTCGGTCTGACGGTCGTTCAACAGACGCCGCC
TGTGTTCGGCAACAACGGTGTCGCGCTTTACGTGAAGCCGCAAGCATATCATGCAA
GGGAAGCCAAATTTCTCGTCGTAAAGACGGTTAAGCTCCGTCATCTCCGAGAATTC
TGCCTCGCTGAGGCGATCGAGGCCGAGCCGGCTTTGTTCTGAACTAGAATCGGGT
GTCATTTCACCAGCACGCGCCTCGCGGCCGCCGAGTTCTGGATGCCCCGCAGA
AGGCCCACCTTTTCACTTTCGTCGCATCGTCGCACCACATCCATCATCTTTCGATG
CATTTCATACACCGATGCGAACGGACGGCTAAGACAAGCCCGCTCAGCCACCCAG
*GGAGCATGCTCGAAAATGCCTCCGAATACATCAACAAATTCGCGCAAAGT***CAT**TC
GATTGACCTGGTCAATATCCACGGTTCACCTCATTCGCACATACTAGGATGCC
GGCTTCCATCCGGAGATCGGCGGTCAAGATCTGTAATTATAAGTATTAAATTA
TACGATTAAAAAATAACGAATAATTATTGCACAAATATTTAGCATATAATCA
TTAATTTATCTTAATAAAATACCAGTTATTGCATGCTAATTGTGATTGGGTAG
TAGATATTCCATCAAGCAAAGGTCAATCGATATGCTCGCACTGTTGTCGATCC
CGTGTGATCTGAGGTCCTACCGCTTGCCGGGTTCGTCTCGTCGCCGTCGCTCA
CACATCCGATGATGTTGGCTGCTTCAGCCCAGGTTCCGGAGACCTGGCCAGC
GCAAAGGGGAATGCTGGCTACTGTTTCAATTTTGGGCCTCTCTCCATCGTGGC
GCTGGCGCCTTCACCAAACCATACATGAGGAACGCCAATAGTCAAACCTGTA
TGATAGCTATTGACACGCTGAGCCCTGTATGACAGCTATTCCATGCTAACGAA
GGGGCACGCGAGGAGGGATATGGCCTTTTCGCAATAGCGAGAAACGCGTCG
ATCGCGACTATCTGATAGTGGGGCATTTGGTCGGTCGCGGGAAGGGTCACGC
CATGCGTCCCTGAACATGCGAAGTTCACAGGCCGGGGTGCTTTTAGACGTGA
ACAACGTGAAAGCATTGTTTCAACAAAGAACAGGTTGAGCTGCGAGGGA**AT**
G*GGTCAAATGAATGGGAACCATCCGCAAAAGGACAGGAATTATCGCGGGTACGGC*

CGTCACGCTCCCCGCGTCTTTTGGCCCGGAGGCGCTCGGGTAGCAGTGAACATC
CAGGTCAACTACGAAGAGGGGGCCGAGTATTCATACGCTCGCGATGGTCGCAACG
AGGGTGCCGGTGAATTCCTTGGCTCTCTGTCTGGCATAGCCGATCGGTCGGTAGA
GTCGGCGTTCGAATATGGGAGCCGAGCGGGCTTCTGGCGCTTGGCGCGGTTGCT
TGATGAATACAGGATTCAGGCAACAGTAAATGCCTGTGCAGTGGCGGTCGAGCGC
AACTTGGAGGTTGGCGACTACCTGCGCGAGGCCAAGCATGAGATTGCCTGCCACG
GGTACCGCTTTGAGGAAGTTTGGAACCTGACGCGGGAGGAGGAACGCCAGCGGA
TACACGCCGCCATCGAGTCGCTGACCAAAACCTGCGGCGAACGGCCGGTTGGCT
GGATATCGCGGTTGATGTCCTCTGACAACACCCGCGAACTGTTGGTGGAGGAGGG
TGGATTCCTATACGACTCGGACGCCCTTAATGACGATCTGCCCTATTTCGTTGATGT
TTCCAACAAGCCACATTTGGTCGTGCCGCTGTCGTTCACCTACAATGACGGCCGCT
TTATCATGGGCGGCTGCGACGATCCGGCGGCTTTCGCTCAATATTGTTGTGGTGC
GCTGGACGAACTTCGGCGCGAAGGCATTACGCATCCGAAAATGATGACCGTCAGC
CTGCATCCGCGGATCATTGGACAGGCAGGAAGGATTGTGGCGTTGCGGCGGTTCA
TTGAGCATGCGTCGGAAGCCGGAGACGTATGGTTCGCGCGGCGTTCTGACATCGC
*GCAGTGGTGGATTGATCATCATAAGGAATTTGCATCA***TGA***AATCTCAGAAGGAGT*
*GCCGGTT***ATG***AGCCGTCCGCTTCGTGTAACGATAATTGGCGCAGGAATTGGCGG*
CCTCAGCGCCGCCGTTGCGCTACGCAAGATCGGCGCTGACGTCACCGTCGTTGA
GCGGGCGCCTGAACTGCGGGCTGCTGGTGCAGGAATATGCATGTGGCCGAATGG
AGCGCAGGCATTACATGCGCTTGGCATTGCTAATCCCCTCGAGATGGTCAGCCCG
ATATTGCATCGGGTCTGCTATCGCGACCAGCACGGTCGTGTGATACGCGAAATGT
CCATCGACAAACTCACTGAGCTTGTCGGGCAGCGTCCGTTTCCGCTTGCGCGGTC
CGATCTGCAGGCAGCATTACTGTCGCGCCTCGATCCGGCCCTGGTGCGACTGGG
GGGCGCCTGCGTGAGCGTAGAGCAGGATGCAAACGGTGTCCGCGCTGTCTTGGA
TGATGGCACCGAGATCGCTTCCGACCTGTTGGTGGGCGCGGATGGTATCCGCTCA
GTAGTCCGAAACCACGTTACCGGAGGAACTGACCGACTGCGCTATCATTATACGAC
TTGGCTCGGTCTCGTCTCCTTTGGCTTGAACCTGACCCCGCCCGGTACCTTCACTT
TCCATGTTCAGGATTCCAAACGCGTGGGTCTGCTAAACGTGGGAGACGACAGGCT
CTACTTTTTCTTCGACGCCGTTCCGAGCGGCGAGGCCAATCCCGATGGCGTGCGG
GCAGAACTTCGGCATCATTTTGATGGATGGTGTTCTGAAGTAACAACTCTTGTGGA
AGCCCTCGACGAGGCGAAAACGAACAGGCTGCCCGTGCATGACCTCGACCCTTTA
GCGTCATTCGTGAACGGACGCATTGTTCTGATTGGCGATGCCGCGCACGCCACGA
CGCCAACCTTGGGGCAAGGCGGTGCGTTGGCTATGGAAGACTCACTCGTACTGGC
TCGCCATCTAGCCGAAAGCACCGATTATGGGAGCGCGCTCGCCAGTTACGATAAC
GAGCGTTTAATGCGAACACGCCAGGTTGTTCTTGCATCACGTGCCCGTACGGCGG
CGACGCTCGGAATTGACAACACCTCGGCACAGACTTGGCAGAAGCAGTTGACCGA
CGACGCCAGCCAAGATTTTCTGGAGCAACTCGTGGATATCCATCGCGCTGGCCCG
*CTGGCAGCCTGGCCGCATGATCGCCAAGAGGGCGTAGCAAC***ATGA***CTCATTCGT*
TGATTGTATCTGTCGACGTTATAACCCAGCACCGGCCGGCGTTTGAAGAATGGTTG
CGATCTTCTTTCCTTGCCGCGACGGGAGTGCCCGGCGTAGCCGCAGATCCGGTGT
GTTTCCGAGGCATGCGCGTCGAGACGCGTTACCGTACCTATCAGCCAACTCCATC
TTACATTGTAATTTTCCAACTCGAGGGTGATCCGGCTGAAGTGACTGGTGCCAAGG
CTTTTTCCGACTGGTGGACCAAAGGCGTCGCAGCGCGGTTCGAATGGGTTGAAAA
GCAAAGTTGGGTCGTCGCTGAGTTAGTCTCCGGCCCGCCGCCGCCTTTCGACTAC
TCCCGTGTGCTGTTTACCCAGGTCGATCTCAGACCAGCGCATCAGAGTGGTTGGG
CCCGCTGGTATGACGAGGTTCATATTCCCCAATCACGGCTTGTCCCAGGAATGTTC

CGTGCGGAAAACCGACGCTTTGCGGCGGTCGAGATCGCCACGGCACGTTGGCAT
TGTTCGGCCAAACCGTCTTTCATACATTTGGTGCCTATCGAGGATAGTGCCGATAT
CGTGCAGGGAGCTGCTACGCCGGAATTTCTTGCCTTGATCGCAGATACCCAAGCA
*CAGTGGGCCGGCGCCCTTGAAAGCGCTGCGTCGACACTCTGCGAGCGGTTTCG**A**
TG*AGCCGCTCCGGCAAACTCATCGTTGTCGGAGACGTTGTCCCAACCCGAGCACT*
GCCAAGCGACTGGCACCCGTTCGGCGAGGCGGACTTCGTGTTTGGTGACCTTGA
GGTCCCGCTAACTGAACGCGGACTGCGCTGGGACAAGCCGGTCTCCTACCGCGC
GGCTCCAGACCGAGCTCGCGAACTGGCTGGTGTCGGCTTCGGCCTGGTAAGCCT
AGCCAACAATCAAGCAATGAACTATGGCCGCACCGGCCTAGTCGATACGATCACT
GCTCTTGATCAAGCCGGGGTGCCATTTATTGGTGCTGGGCCTGACCTGGAGGCAG
CGCTGAAACCTCACATCACCGAGTTGGGCGGTAAGAGGATCGCCTTTATCGGGCT
CACCTGCGTAGCCCCACGGCACTGGGACGCGGGTGCAGATCGCAGCGGCTTGGC
TGTGCTCCGACCTCGCCAAGCGGCCGAAATCGACCCCGCCTGGGCTGCTGAGGA
ACCGGGTGTTGCTCCGACCGTGCATACCTGGCTCGATGACGATGCACTCGCTCGT
GCCGCGGCTGCTGTTCAACGGGCGCTGGAGCACGCAGATCTGGTGGTGGCGGCC
GTCCATTGGGGCGTTGGGTCCAGCTATCAAATCACGGGCTACCAGCGTCAGCTTG
GCCATGCGCTGCTCAATGCCGGGGCGGCTCTTGTGCTTGGCTCGCACCCGCCTC
CTCTGCAGGGGTTCGAACGAACGAATGCTGGGCTGATATCCTACAGTCTGGGCAC
CTTCATCCGCCAACAGCCGCAAGAGGGAGTTGGACTGGCGCTCAGCGCGGCATAT
TCTAGGATGCCGCGAGAATCGGCAATTCTTGAACTGACTGTCAAAGGCGGCCGGT
TCAGTGAGGCGCACGTTCACCCTGCGGTACTGGACGACGACGGAATTGCTCGCCA
GGTCAGCGGCGAACGGGCACGACGCATAATCCGAACCATCCTTGATCGTTCTACA
GGGCTAATTCCTAACCCGGCACAAGGGGAAGAGCCGGAAACCTTGACCATCCATC
TGGACAATGAGTTCGCCACCGGATCGGATCGACCATCGGATGGAAGTATGGTGCC
*ATCCGGACCGGAGGTCGCCGCGCGCCAAGGTCCCCGA**TAG**CGCTGTCAACAAT*
TGGTCTGATCGACCATCGCTCGTCAATGACCGGCTACCGGAGGGGCGAGTCG
GAGGCAGAGGCTCTAAAAGTTGAGTTCGGCGGGGCTACCTGCAGCTTAAGAG
TGATCGAAATACCGCTCGTCTTCGATACGAATGATCGCCTCAGCGGTATGCTT
GCAGGAGCCGCAGGCAGCTACCTCAGCTTGTCGCGACGCGCGACAGCGCCAG
GGTTGGCGGCCTTCGACCGTTTTCGAAGCTGGCGTAAAAGTGCCTGCTATCGT
GGATAGTAGCCATGAGATTTCCCATGGATTGAAGTCGCTCGCCACCGCCCTTG
GTGTATAGGTGCTGGCAAGCTGAACAGTGGGCACTGGTCGCCGGCAGCGCTT
CGCATCGGTTTGTAGTGTCAATCCCACCGGCCCGCGTCATTGTCTACGACGCC
CTCTTTACGATTGGCGGTGGGCGCTGACGTCTTCACCTGTTCCGCATACCAAA
GGCAGGCTCATTGCATTTGTTCGTTCTCGGGGATCGCGAGGCAATTGGCGGTG
ATACAGCGAAGGTTCGGTCCATTGTCAGCCGGCGGCCTTTGGATGGTGTAAG
AGCCGCTGTCCGCTTGATGTATTATGCTCCATCAGACCGCCCCTGGGCGCGCC
CGACGTAGAGAGGACGTGAGGTAGTCTCTCCTGATGGCGCCGAGGATTTTAG
AGTTCGAGGAAGGCTGCAAACTGGGTATCGCGACCACCTAACGAAAGGTGG
GTCTTCGAGCTGGATTCGTTTCTTCATCGTCTGGCGCCAGCGTTCGATCTTGCC
CTACGTCTTAGGGTGACTGGCATCCGCAATGAGCCTGAAATCATTCGTCTCTC
TATGGACCCTGAATTTGTTCATTGCCGTTAAGCCGGGCGGCGCCTTTAGGTCA
ATTGCCTATATGATAGCTATTGACACAAGGCTAGCTATATGAAAGCTTGCTAA
AAACAAAAGAACAAAAAAAGGGAACTGGAATGGGGAAAATGCAGGGCATGT
TTCTGGACGTGACGGCTACCGCCAATCACCAACCGTGAGCCTTCAACGATGT
GTCGGAGAAGGACGCAGTTGTATCAAGATCGTACTTGGTTGGCTGTGTTATGT

CGGCAGAGGCCGACCGGACGGCAGCAAAGAGCTTCACAAGTCAATTTGATGT
TCATGCCTGTAAATAGTCAGGCGTGGGCTATGTAAGACTGAAGATGAAGGTC
ACTTGTTTATTGAACAAGTGTCATGACTTCAGATTTTGCATATGTAAACGCGC
AACACACATGCATGCGTGTTGCGGGGCATTTTCCCGTATCGAGACTAACAGC
TTGGCCGAGCCGGCATTCTTGCCCGGAATGCTGCTGAAGACCTCGCGGCGAT
CAACCGCTATTTCGTTGAATATGCATGATAGTCGTCCAATGAAGGATCGGCTG
GTGACAGAGAGATAGGCCCATCTGAGTCTCCCGCAATTCGTGATTGCTAAAG
CCGGATGATGGACTTTCCGGCTTCATCGAATTGCCAGACATCAACCTCAAATA
GGCGGGTAA**ATG***ACCAGCAACCCAGCAGTACAAGTCGGCGAGGCGCCCAGGAA*
TGGAGCGCCGCTGCTTTCCCTTTCTGGAATTACCAAGAGATATCCCACTGTAGTCG
CGAACGAAAACATCGATCTTGACGTAAGCGCCGGCGAAATCCACGCGGTACTTGG
CGAAAACGGCGCCGGCAAATCAACGTTGATGAAGATCATCTTCGGTGTCGCCCAA
CCGGATGCCGGCACCATTTATTGGCAAGGTCATCCCGTGCGTATCGATAGCCCGG
CTCATGCCCGGGAGTTGGGCATCAGCATGGTGTTTCAACACTTCTCCTTATTCGAA
GCGATCACCGTAGCTGAAAACATCTCGCTGACCGTCCCAGGTACGCTCAAGGAAC
TGTCTCGCAGGATCAGGGAAACGGGACAGGAGTTTGGCCTCAAGGTCGAGCCTG
CGGCGCTGGTTTGCAATCTTTCCGTTGGCCAACGTCAAAGGGTCGAGATCATCCG
ATGCCTATTGCAGGAACCGAAACTCCTCATTTTGGATGAGCCGACATCGGTGTTAC
CTCCACCGAACGTCGAGCAACTCTTCGAGACATTGCGACGACTTGTCGCCAAAGG
CATGTCGATCCTCTACATCAGCCACAAACTTGAAGAAATACGTTCGCTCTGTCACTC
GGCGACGATATTGCGCCAGGGTCGTGTTTCTGGACATGTGATCCCTGCTGAAACA
TCAAGCCATGCGCTCGCTTCGCTGATGATTGGCCGGGATTTGCCGAAGACTGCTC
ATCCGCCGGCTCGACAGGATGGTGCGGTAAGGCTTCGTGTCGCGAATCTCTCCTC
GCACGACCCGGATCCGCACATTGTCAATATTGATGACCTGAGCTTGGAGGTGCGC
TCAGGCGAAATCCTCGGAATTGCGGGCGTTTCCGGAAACGGACAGCACATGCTTT
CAAAGATGTTGTCCGGCGAAGAGGTTCTCGCCCCATCCGAAAAGCATCGCATCGA
GCTCATGGGTACGGCTGTCGCACACGAAAACGCGGGCCGCCGCCGCAACCTTGG
GCTTTCATTCGTCCCGGAAGAACGGCTTGGCCGTGGTGCCGTCCCTCCCTTCAGC
TTGGCTGATAATTCGCTGCTTACGGGCCATCGAATGGGTATGACCGGTCGCGGGA
TGGTGCGCAATCAGCCGCGGGACGCGTTTACGGACCAATGCATCGCCGACATGAA
TGTCGTGTGCGCCGGCAGCAAGTCGAACGCGGCCAGCCTCTCTGGAGGAAATCTA
CAGAAGTTCATTGTTGGCCGTGAGATTATGATGGCGCCCAAGGTACTGGTGATTGC
GCAGCCAACTTGGGGGATTGATGTCGGCGCAGCAGCGATGGTCCGTCAGAAACTC
GTGGATTTGCGCGATACGGGAACCGCTATTCTGCTGATCTCGGAAGAGCTCGAGG
AACTGTTCGAGATTTCCGACCGCATCGTGGTGATGTTTAAGGGCAGGGCATCGCC
TGCCCTTGATGCGCGCGCGACGGACGCCGAAACAATCGGACGCATGATGATTGG
*CGACATCGCTGCTCCCCTGTTCAGCGAGGCTGCACA**ATGA***ATATCAGACAGCCGT*
TCGTCCTGGTGCCGCGGCAAAGGCAATCGCGCATTGCTTTAATTGCGGTGCCTCT
TGTAACCATTTCGCTGACTGTCATCGCGGCATTTTCCTGTTCGCTGGACTGGGTG
CCGACCCTGGCATTGTGCTCTATACGTTCTTTATCGAACCGCTTTCATCCGCTTACA
ACCTCGGAGAGGTGCTGATCAAAGCAAGCCCGCTGATCCTGATCGCGCAGGGACT
TGCAATAAGCTTTCGCGCAAAGGTTTGGAACATCGGCGCGGAAGGACAGCTCATT
ATCGGCGCGATCTGCGCCAGCCTCATTCCGATTTACTGGTCCCATAGCGACTCAAT
GCTGATGCTACCCGGGATGATCCTCATTGGGCGGCGGCCGGCATGGTCTGGGC
TGGCATCGCGGCGCTGCTGCGTACTCGCTTCAACGCCAGCGAGGTGATCGTTACG
CTCATGCTGACAGAGATTGCCCGGCAGCTTCTCTACTTTCTGGTCACGGGGCCGC

TTCGCGACCCGATGGGATACAACTTCCCACAATCGGTGCTATTTCCCTCGGCCGC
GCTCTACCCGACGTTCGGTGGCGTCGGCGTGAGGGCCAACCTTTCTATTTTCATCA
CGCTTGCGGTCACAATCATCTGCTGGGTGTTCGTCTCGAAGAGCTTCGCGAGCTT
CAAGTTGCTGGTTGGCGGAACAACGCCTGCCGCCGCAAGGTACGCCGGCTTCCT
GCCAAAACAGGCGGTCTGGATATCGCTTCTGATCGGAGGTGCTGTCGCGGGGCTT
GCAGGCGTTGGAGAGATCGCCGGACCTATCGGAAAGCTGCAACGAATCATATCCC
CCGGCTACGGATTTGCGGCGATCATCGTCGCCTTCCTCGGAGGATTGAACCCCAT
AGGAATCGTCTTCGCCGGAATGCTCATGGCCGTTGTCTATGTTGGTGGCGACAATT
CTCTCGTGACGGCAAACATTCCGGCTTCCGCATCGGTCGTGTTCCAGGGATTGCT
ACTGGTGTTCTACCTCGCGACTGCCGTGTTCGCGCGGTACGAAATCCGGCGTGCG
*CCGCCCCCGTTAATCCCCGCG***TAA***AGGACCTTT***ATG****GATGGATGCGCTCACTTTT*
GTTGTGGCCGGGGCATTTGCGATGGCCACTCCCTTGATGCTCGCAGCGTTGGGTG
AACTGGTCACCGAGAGATCGGGCGTTCTCAATCTGAGTGTGGAAGGGATGATGGC
GATCGCTGCCGCCATCGCGTTTATGACAACGCATGATACGGGATCGTTCCTGCTG
GGCTTCGCTGCTGGTGGCGTTTCAGGACTCCTGCTGTCGGTCGTGTTCGCGCTCC
TGGTTCTCGTTTTCCTTGCCAATCAGGTCGTAGCAGGGCTCGCAGTGGGCATCCTT
GGCCTCGGATTGTCCGCGCTACTTGCTCGATCATATGAAGGCATGACGATCTTACC
AATGGGAAAGATCACCATACCGATCCTCTCCGATCTGCCACTCGTCGGTCCAATCC
TCTTTCGTCAAGACGTTATGGTGTATCTGGCAATCGGTTCAGCCATCACACTCGCTT
GGTTCTTGTTCAAAACGAAGCGGGGCCTTGTGTTGCGGGTCGTCGGGGAGAATCC
ACAGGCTGCTCATTCGATAGGTCTTCGCCCTATCCGGGTCCGCTTTATCGCTGTTC
TCTTCGGTGGCCTGATGGCTGGTCTGGGTGGTGCGTACGTCTCCATCGTCCTGAC
GCCACTCTGGTCGCAAGGAATGATTGCAGGTCGTGGCTGGATCGCCGTGGCGCT
CGTCGTATTTGCCTGTTGGAGGCCAATCCGCCTCACGGTGGGGGCCTATCTGTTC
GGTGTCGTCTTGCTTCTCGATCTGGCGATCCAGGCCCTCGGGTTAAGCATACCGT
CGGTCATCCTGACCTGCATGCCTTACGCATTGACGATAGTGGTGCTGACCATCGTG
TCCTCGGACGCGTCTCGCATCCGTCTCAACGCACCGGTCTCCCTTGGAGAAAACT
*TCCGGCCGGCAAAC***TAG****CTGGTCGGATTCAAGCAACAAGAATGGAGTTGGGA*
AC***ATG****AAGATCTGGGTACGCACTTTATTGGGAGCAACCGCGCTGGGCTTGCTGCT*
TTCGGGGCCTGCTGCCTCGGACCCTTTCAAGATCGGCTACATAATCCCGAACGCA
ATTGGCGAAGCGGGGTGGGACCATGAACTTGAACGCGGCCGCCAGGCGATCGTC
GATCACTTCGGCGACAAGGTCCAGGTCAATGCCGTCAACGGTGTCGGCGAAGGG
CCTGATGCCACGCGCGTCATGAACAGGATGGCAGCTGACGGTACTAACATGCTCA
TTCTTGGAGCATTCGGACTCATGAATGACGGACTGGTGCTCGCGCGACGCTATCC
AGACCTTAAGGTTCTGCATTATGGCGGTTACGTCAACGAACCAAACTTCGCGACGC
TTGCACTCCGTCACTACGAGGCATCCTATCTTTGCGGCATGGCCGCGGGCATGGC
CGCCAAAACGGGCAATCTCGGCGTTGTAGCCGCATTCCCTTTGCCGGAAGTCCTA
AGCATCATGAACGCCTATGTACTTGGCGCACAGAGCGTGAACCCCGACATTAAGC
CGGTAAAGGTCGTGTGGCTCAACTCCTGGTTCGATCCTGCAAAAGAAAAGCGGC
TGCTGAATCGCTCGCATCCCAGGGTGCGGAAGTGCTGTATTCGCTGTTCCCCGGA
ACTCCTTCGACCGTGTCGACAGCCGAGCAACTTGGTGTCTACGTAACAGTTACGTT
GTCCGACAACACAATGTTTGCACCGACCAAGCATCTTTGCTCCGCACAAGCTGAAT
TCGGCCCCGCTCTGATCCGGAAGATCCAGGACGCTATCGACGGGAAGTTCGTTGG
AGACGACACCTTTTCAGGCGTGAAGGATGGCTCGATGGGCATTGCTGGGCTCAGT
CCCGACCTTAGCGAAGAGCAAAAGAAGCAGATCATGGCCAGGCTCGAGGAGATGC
GAAACGGCAGCTTCCAGCCATTCAGAGGGCCTATTATGTCAAACACCGGGCAGGA

AATTGCCGCCGAAGGTACAAATCTCGACGACACTGCAATCAAGTCGATGAAATTTC
*TAGTCAAAGGCATAGACACCACGCTTCCCAAC***TAA***AGACGAACCTTGTCTCGGA*
TTGCCAGACGGCTCGCGAGGGCCGCTTGGCAGTCCACGAAACCTCGTCGAAC
GAATGGGATCGGACAGTCTGGAAGACTAAGGCTGTCTCCTTGCTGGTACTCA
CATCATCGAAGCAACTTCCACGGGAGGGGAAA**ATG***AACTTCACGCCAAAATCC*
TGTCTGAATGGACTGAAGAAATTCTTGATTCGAGCCATGTTGCTCGTTGCGGCCAG
CCACGGCCTGTCAACTTCCGCATTGGCGCAGACCTTTCTGCCAATCAACTGCAGC
TTCATTTCGGCGGCGGGTATCGTTTTGGGGGAAATGGTGCTGAAACAACGACCCG
AACGATGCTAGAATTCCAGCATTTCAGCGCTTACAAATATGGCGATCTCTTTTTCTT
CGCCGACGTCTATCGGGATCATCAGTGGGACGGGACTTCAAACCGAGTCAACTTC
TATGGCGAGGGTTATGCACATCTGAGCGCGAAGTCGCTGGGAGGCATTCCATTCG
GAGAGGAATCGTTCCTGGCAGATATCGGGCCGGGTATTGGCTTCAACGCCGGTCA
GGATTTCTTGGTTGCCATTTACGGTGCGAGAGCCAGCTTCAAAGTACCGGGTTTTA
GCGTGTTGACCTTTGGGGTCTATGCCTATGACAATCAGATAGACCCGTATGGTCGT
GACCTAGATACGACCTATCAGGCCACTCTCGTTTGGGACATTCCGTTCGCAATAGG
CGATCAGAAGTTCTCAACGGGAGGGGTTCTCGACCTGATCGGCAGCCAGGGTGTG
GGAGTAGAGGAGCAAATCTATTTTCAGCCGGAGATCAGGTGGGACGTAGCCAATG
CGCTCGGCCAGAAAGCCGGCAGTTTCGATCTGGGTTTGAGATACGTCCACTTCAA
CAACAAATACGGGGTCGATGGCGTGGATGAGAATGCGATGTCCGTGTTCGTCTCC
*AAGAAATTC***TAA***TGTATCAGGGACTTGCCGCGGCCTTGTGACGAGGCCGCGGT*
ATGTGCAAACAGCTCGAAGCGGCTGGGCAAAACAAAATCCGGTTTGATCGAC
GGGCATTGGCTCGATGATGCGTGCCACCGCACGCGACTCAAGCCGAGCTCGT
CCGTCTTGTCACTTCTGTCACGCTCGATCGCGCACGGCCGATCTCAGAGCGTA
ACTTCGTTGAGCTATTTCAACAACGAGGTTGAGCACCACATTTCGGTATGACG
ATCCGATCGCACTTTAGAAAACTTAGGGGCGCAGCGCGTCCGCCATCGTCAT
ATCACAAGGAACCTCTTCCCTCCGAAGGGTAAAGCACGGCGCGGGTGCCTGT
ATGATAGCTATTGACACGGGGATAGCTGTATGAAAGCTTTCCTGAGAACAAT
AATAAAAGGGAACATAGAGATTGGACCGTAAGACAGTTCGCCCCATTGGGA
ATCGTTGTCGAGGCCAGTATTCACTGGTGAGGTAGGATTCGGACGCGTCGAA
GAAAAGACGTCGGTCGGATTGCCTGGCAGTTTGGCTTCAGCGGCAGTTGACG
CTCGCGTCAGTATCGCGAAGACCCCGCCACAAGAACCGAAGGTTTGCAAGCG
CCAGCAACTTCGGCAACTCGAAAAGAGAGCATTGCCATGTGGCTGGTTGAGA
ATTTGGTTCCCTGAGGATTTAGAAGTCCAACGGGTGCACTGCGGACGGTGTC
CGGATCGACACTCTTTCCGGCAAATGGTTGCGCGACAAGCTGTGTAGAACTG
TGGAGGTTTGC**ATG***TTCGCGGATATAAACAAAGGCGACGCCTTCGGTACCTGGG*
TCGGAAAAGCGTACCGCGTCGCGAGGATGCGGATATCCTTGCAGGGCGAGCCG
AGTACATTGCCGATATCAAGCTTCCTGGCATGCTTGAGGCCGCGTTCCTGCGTAGT
CCGTTTGCTCATGCTAGGATCGTCTCGATAGACGTCTCACAAGCCCTGGCTCTGCC
TGGTGTGTACGACGTCATGGTGGGCGCGGATATTCCGGACTATGTCAAACCGCTT
CCCCTCATGATTACCTACCAGAACCATCGCGAAACCCCGACTTCCCCGCTGGCCC
GAGACATCGTTCGCTATGCCGGTGAACCGGTCGCTGTGGTCGCCGCGATAAATCG
CTATGTCGCAGAAGATGCACTGGAGTTGATCGTCGTCAAATATGAGGAATTGCCTG
TCGTCGCTTCCATCGATGCTTCGCTTGCTGTTGACGGGCCACGTCTCTATGAGGGA
TGGCCTGACAATGTCGTGGCCAAGGTCAGCTCGGAAATTGGCGATGTGGACGCG
GCCATGGCATCTGCCGACCTCGTCTTTGAGGAGCGGTTCGAAATCCAACGCTGCC
ATCCGGCTCCGCTCGAGACGCGCGGCTTCATCGCGCAATGGGACTTCAAGGGCG

AGAACTTGAATGTCTGGAATGGTACGCAGATCATCAATCAGTGCCGTGATTTCATG
TCGGAAGTGCTGGATATTCCGGCCAGCAAGATACGCATCAGATCGCCCCGTTTGG
GTGGAGGCTTCGGTGCGAAATTTCATTTCTACGTCGAAGAACCCGCGATCGTGCTC
CTTGCCAAGCGGGTCAAGGCACCGGTGCGCTGGATAGAAGACCGGCTGGAGGCT
TTTTCGGCGACAGTTCACGCCCGCGAGCAAGTCATCGACGTCAAACTCTGTGCCAT
GAACGACGGGCGCATTACCGGCATTGTCGCGGACATAAAAGGCGATCTCGGCGC
GAGCCATCATACAATGTCGATGGGACCGGTCTGGCTGACTTCGGTGATGATGACG
GGTGTCTATTTGATCCCGAACGCACGTTCCGTCGCAAAGGCGATTGTAACCAACAA
GCCGCCGTCAGGTTCCTATCGCGGCTGGGGACAGCCTCAGGCAAACTTTGCAGTC
GAACGGATGGTGGATCTTTTGGCGCACAAACTCCAGCTTGATCCGGCTGCGGTGC
GGCGGATCAATTACGTGCCTGAAGCCAGAATGCCTTACACCGGTCTTGCTCATACG
TTCGACAGCGGCCGGTATGAGGTATTGCATGACAGAGCACTGAAGACGTTCGGCT
ACGAGGCATGGTTGGAGCGCCAGGCAGCAGCGCAGGCCCAAGGACGGCGTATTG
GGATCGGAATGTCGTTTTACGCCGAGGTTAGCGCCCATGGCCCATCGCGTTTTCTT
AATTATGTTGGTGGACGGCAAGGTGGTTACGACATTGCACGTATTCGGATGGATAC
GACCGGTGATGTGTATGTCTACACAGGCCTCTGCGACATGGGTCAGGGCGTGACC
AATAGCTTGGCGCAAATAGCTGCCGATGCGTTAGGTTTGAATCCTGATGACGTGAC
GGTAATGACGGGAGATACGGCCCTTAACCCCTACACCGGCTGGGGCACAGGGGC
TAGCCGTTCGATCACTATCGGCGGTCCGGCTGTCATGCGCGCGGCGACGCGTTTG
CGGGAAAAGATACTCTCTATTGCTCGTCATTGGCTTCAGGCCGATCCGGATACGTT
AGTTCTGGCAAATCGCGGCGTCATGGTGCGCGACGATCCGGGTCGTTATGTATCT
TTCGCTTCCATCGGCAGGGCTGCCTATTGCCAAATCATCGAGCTGCCGGAGGATG
TAGAGCCGGGGCTCGAGGCGGTAGGAGTTTTCGACACAGTGCAGTTAGCTTGGCC
ATACGGCATGAACCTGGTTGCCGTTGAAGTCGATGAAGATACCGGCGCGGTGTCG
TTCCTGGACTGCATGCTGGTTCACGACATGGGAACGATCGTCAATCCGATGATCGT
GGATGGCCAACTTCACGGCGGAATTGCTCAAGGCATCGCACAGGCTCTCTATGAA
GAGTTGCGCTATGACGAGAACGGTCAACTAGGGACAGGCTCCTTTGCGGACTTCC
TGATGCCTACAGCCAGCGAAATACCAAACATGCGTTTTGACCATATGGTGACCGAG
TCACCGCTCATACCGGGGGGGGATGAAGGGAGTCGGAGAAGGAGGCACTATCGGT
ACTCCGGCCGCAGTAGTGAACGCTATCGAGAATGCACTACGCCCAATCACGAACT
CGAAATTGAACCGAACTCCGGTTACCCCAGATCGCATCTTGACCGCCATTTCTGCC
*GGAGC**ATG**CGCA**TGA**AACCTACCGCGTTTGACTATATTCGGCCGACCAGCTTGC*
CAGAAGCGCTCGCAATCCTGGCCGAACATTCCGACGATGTCGCTATACTTGCTGG
CGGGCAAAGCCTGATGCCGCTGCTCAACTTCCGAATGTCCCGTCCGGCACTGGTA
CTCGATATCAATGACATTTCCGAGCTGCAGCAGGTTCGATGCGAAAACGATACCTT
GTATGTCGGGTCGATGGTCCGGCACTGTCGTGTTGAACAGGAAGAAATCTTCCGC
TCGACGATACCGTTGATGTCCGAGGCGATGACTTCGGTCGCGCACATTCAGATCA
AGACGCGAGGTACTTTGGGTGGCAATCTTTGCAATGCGCACCCGGCGTCGGAAAT
GCCTGCGGTGATAACTGCATTAGGCGCCTCAATGGTCTGCAAGTCCGAGAAACGG
GGGGAGCGCGTTCTTACGCCCGAAGAGTTCTTCGAAGGCGCCTTGCAAAACGGTC
TTCAAAGCGATGAACTGCTTTGCGAAATCCGCATTCCCGTACCATCCCAATACGTT
GGATGGGCGTTTGAGGAAGTCGCGCGGCGGCATGGCGATTTCGCACAATGCGGC
GCCGCAGTTCTGATCGGCGCTGAGGATAGGAAGATCGATTATGCGCGCATAGCCC
TTTGCAGCATTGGCGAAACGCCGATCCGTTTCCATGCTCTGGAGCAATGGCTTATC
GGCAGGCCGGTCGGAAATGATCTGCCTGCAGATGTAAAGCTCCATTGCCGCGAAA
TCCTCGACGTTGCGGAGGACTCGACGATGACTGCCGAGAACAGGGCGAAGCTTG

CCTCCGCCGTGACTTCCAGAGCCATCGCCAGGGCCGCAGACCGGATTGTTCATCT
*CGATGTAAAGAGAGGT***TGA***AGTTA***ATG****AGCAGTCACGTAATTTCTCTTACCGTCAA*
CGGCCAAGCAATTGAACGAAAAGTGGACAGCAGAACTTTGCTGGCGGATTTCCTG
CGCGACGAGTTGCGCCTCACCGGGACCCATGTCGGTTGCGAACATGGCGTGTGT
GGCGCATGCACAATCCAGTTTGATGGGGAACCGGCACGCAGTTGTTTGATGCTTG
CGGTTCAGGCCGAAGGCCATTCGATCCGGACTGTCGAAGCCTTGGCCGTCGATG
GCTGTTTGGGAGCGCTCCAGCAGGCATTCCATGAAAAACACGGTCTTCAGTGCGG
GTTTTGTACGCCTGGCCTGCTAATGACACTGGACTACGCGCTGACAGCGGATCTTC
ACATTGACTTTTCAAGCGACAAGGAAATTCGAGAGCTCATTTCCGGAAACCTTTGTC
GTTGCACGGGCTACCAGAACATAATCAACGCAATAAAGTCCGTTTCTCCTACGACT
*GAAATTGCAAAGAGTGAGGAACTTGT***ATGA***AGTTCGAAGGAAATTACGACATAGAC*
GCGGAACCAAAAAAAGTATGGGAATTGCTCAACAATCCTGAGGTGCTTCATAAGAG
CATTCCTGGATGCGAAGAAGTTATCCAGACTTCTCCTGTGAAGTATGAAGCGGTTG
TGACGTTGAAGATCGGGCCAATCAAGGCGCGCTTCCAGGGTCACGCAGAAATGAC
AGAGCAGCAGCCACCAAATCAGTGTGTAATCATATTCCAAGGTTCAGGTGGCATTG
CGGGCATGGCACGCGGCGAAGCCAAGGTTCTGCTTACACCGAATGACCGAGGGA
CTGAGTTAACCTATATTGCGGATGTCGTGATCGGTGGAAAAATTGCACAAATAGGA
TCTAAGCTGATCGAGGGAACTGCTCGTAAAATCGCTGACCAGTTTTTTAGCAATTTC
*GTCAATTATATCAGCGTCAATGTCCCG***TAA***ATACACAATTGAAGGTGGCTATAT*
GCTCAGTTGAGTTCATGCGCAGTCTAATCAGGTGGCACTGCATCTGAAACAA
CTCCGAAATTGCATTTCGTTGGTCTACGGCCAAGCCATCGATCGCGGCCAGCT
CTGTCTCGATGCGATCAACGCCGGGAGCGACGCCGACGATCAGTCGCGTGCC
AGCCTCGGTCGGCGCAACGCTTCGGGTGGTGCGTGTCAGTAGGCGAACCCCG
AGGCGTTCTTCAGCAGGCGGATTGTGTGGCTCAAGGCGGATTGTGAGACGCC
GAGCCGCGCTGCAAGCAATCATAAACCTATCATAATGATGTTGACAGCCTGT
AGAACAGGGGTCTTGATTGCATCGCAGGGTCGAAGAGATCTTGATCCTGATA
ATCGGTGGGAACTGGCATGAACGCTACGACAAGGGTCTCCAAGGCGGGGCAT
CGATCAATCGCAGTTCATGCTGGTGCCAACCGTTCTGCTCGCCGGCACCGGG
GGAGCGGTGCGTTCGTTCAGGGCAGTAAGACCATTTCAGTAAAGCGACGAAC
AACATTGTTCCGCTGAAATTGATAATGCCGTATCCGAGGCGATCAATAAGTA
TCGAAAGGAACAAGGCGGGATCGGGGTGTAATCGTGCACTCCGGGCCGGCAT
TGGTCGGCGATCGCGCCGGCTTTGAAAGTCGTCGAGTCAGCCCGAAATACAG
*TGTATCCGGGCAAAGCCTAGTTAAGGAACCATCC***ATG***GATCTTTTTATCAAGAA*
TGTGAGGGTCATCACCGAAAACGGCAGTTTCGAGGGCGGCGTGGCGGTCGACAA
AGGTGTAATCGCGCAGATCACCGGTGCGAACGGACATGTCGACGCGGATGAAACC
ATCGATGCGGAAGGATTGTATCTTCTGCCCGGCATCATCGACGCGCATTCCCATTT
CAGCCATCCTGGTCGCGAGTTTGAGGGGTTCGAGGCCGGTACGCGGGCGTCCGC
GGCAGGTGGCGTGACAACAGCGATCGACATGCCGCTCGCCGACATTCCGCCTAT
GGCCACGGCGAAGGCGCACGCCGCCAAATATGAAATGGTGAAAGATCTTTGTGTG
GCCGACTACGCATTCTGGGGCGGCCTAATCGATGACAATCTCGAACATCTCGATG
AACAAAATGAACTGGGGGTCGTAGCCTACAAGGCGTTCATGCGCAGGTCCAGCAG
CTACCCCAAGGTCGATGACGGACAGCTATTCACCGGCCTGCAGAAGACGGCGAAG
TTTGGCAACATTGTTGGCGTGCATGCGGAAAACGACACGCTGATCGAACATTTGGA
AAAACGGTTGAAGGCGGAGGGCCGTCACGATCGCATGGCTTGGAACGATTCGCAT
CCGATCGAATCGGAGCTTGAAGCCATCAATCGGGCGATCCTTCTGGCAAAGGCTG
CCGCAAGCAGTTTGTATATAGTCCATATCAGTTGGGCCGGGGGCGTCGACACGGT

GCGTGAAGCCGCTCATCGGGGCCAGGATGTTTATGCCGAAGTCTGTGCCCACCAT
CTGTGCTTTGACCTGGACGACTATCTACGCGTCGGGCCTCGCTTACGCTCCGCGC
CGCCGATCCGGTCGCGGCCGGAGGTGGAGGAACTTTGGGACCGCGTGCTGCGC
GGCAAGGTTGACGTGATCGCGTCGGATCACTCGCCGTTTCTCCCGGACTTGTACA
AGGCCGGCGAAAACGATGTCTGGCAGGGCACCGGCGGCATTACGGGCATCCAGT
CGATCCTGCCGGCGATTATCACGGAGGGCGTGCACATGCGCGGCCTCTCCTGGG
AACTCCTGGTAAAAATGATGTCGTCCAACCCTGCGCGCATTTTTGGCATTTATCCG
CGCAAGGGAGCGATCGTACCGGGAGCGGACGCAGATTTCGCGCTGGTTGACCCC
CAAAAGGTCTGGACCCTTCAAACGGAAGATCTTTTCTACCGCTACAAGGAAAGCCC
GTACGTCGGCAAAGAATTCTGCGGTCGTGTCGAACGCACTATTGTTCGCGGCAAG
ACAGTTTTCCTGGAAGGGAAAATAACTGCCGACAAAGGGTATGGTCAGCTGGTTCG
ACGAGAAAGCGCCACTCACGGGCGACCGCTTCATAAACACCAGCAATCGGTTCGA
*CATTTG**TGA**GAGGTTCTCGTGGATAGAAATTCATCC**TTG**GTTGCTGGCGGCCT*
TCCTGATCAAAGACAGCTTCTTTCCAGCAGGGCAGTCTTTACGCCAGCCTATGCCA
TTATTCCGGGCTCCGTCCTGAGGGACATCGTTTCAAGTGCCCTGCCGGAATGGCG
CGACACGCGGCTTTGGGTGTTGGCGCGGCCGCTTTCAGGTTTCGCTGAAACCTTT
TCGCAATATTACATGGAAGTGTCGCCCGGCGGCGGCAGTGAAGATCCGGAGCCC
GATCCGCACGCGCAAAGCGCAATCTTCGTGACTGCTGGAACGATGACGATCACCA
CCAACGGCGTTCAACATCGCATGGTGAAGGGCTCATTCGCTTTTATACCCGCCGG
CCAGGTGTGGCAGATCATCAATCCGTCAGCAGAGATCTGCCGGTTTCAATGGATC
CGCAAGAAGTTCGAGCCGGCTGGCAAACTGGGAACGCCACCGGCGATCTTCACC
CATGAAGATAAAGCTGCCGGCGTAAGCATGTGCGATGAGCCACCCTTGTGGGGCT
CCGTCCGGTTCATCGACCCAAGCGATCTACGCTACGACATGCACATCAACATCGTC
TGGTTTGAAGGCGGTGCGACGATCCCCTTCATGGAAACGCATGTGATGGAGCATG
GCATCTACATATTGGAAGGCAAAGCGGTTTACCGGCTCAACAGCGATTGGGTGGA
AGTCGAGGCCGGCGACTTTCTCTGGCTTCGTGCCTTTTGCCCACAGGCTCTTTATG
CCGGCGGTCCTGACGGACTTCGCTATCTCCTGTACAAAGACGTGAACCGCCACAT
*GAAACTGCTA**TGA**CCATCGTAACAGTCTGACCCTGTCAGCCAAGGATTCTGAA*
CCAAATACTACGGCATGGTCAGTTTGTACGTGAGTCAAATATAAGCTCAGTTT
CTTATTATCACCTGCGATCTCATTGTCCGGCCGGTTCACCTAAACTACGGCTG
GCTCCCTGGTAGGCGGAAGCTTGTCTCGCCAAAACGCGCTGATCGGCCGACT
GGCCGATCGCGAGCAATGCCGCTGGGCAAGCAGACTTTTGGCCCGATCAGGT
TCTGTCTCGATTGTCGCCGCAGAGCCGGCTGCCGATCTACGCGCGCGGGCTCC
AGACATCGAGTTGATAGGTGTCTGCCATACGGCCGGTGTGCGAGCCGACAAG
TAACGAGACGCTTCGTTG**ATG**ATCGAAATGACTGGGTGTCAATTTTAGAATTGA
CTAAACTAATGGACTGTAGGAAAAAAGTTCATCTGAAGGAAACTTTCCGCCTGACA
CAAAAATTTGGAGTGGTACGATGAACGGCTCATCAGCGGTCGGTCTTGAATCGACT
TTGAATGTCATTGAACGTTTCATTTCTTTTAACACCGAAAGCTCTCGCTCCAACCTC
GAACTCGTCGATTGGGTTGAGGGCAGCGTCAAAGGACCGCGAGTCAGCTGCTTCA
GGACGTTCAACGCAGAGAAAAGCAAGGCGACGCTTTGCATAACGGTAGGCCCTCC
GGTGGACGGCGGCGTAGTCCTTTCGGGGCATATCGACGTCGTCCCGGTCGAAGG
ACAGGCATGGAGCAGTGATCCGTATTCGTTGCGTCAGGCTGACGGGCGGCTCTAT
GGCCGCGGAACCTGCGACATGAAGGCCTTCGGTGCCATTGCGGTGGCACACATT
CCCTATTTCCTCGCATCCAACATCGCCAAGCCGGTTCACATAGTTCTGAGCTATGA
TGAAGAAGTGACGAGCCGAGGGTCGATTGCGGTAATCGAGGAATTCGGAAGATCG
TTGCCCAAGCCTTCCTCGGTCATCGTCGGCGAGCCGACGATGATGAAGGTCGCAG

ATACCCACAAGAGCATCTCGACCTATTACACCACAGTGACGGGCCACGAAGCACAT
TCCTCGAAACCGAACATCGGGGCAAATGCAATCATGGGTGCCTGCGAATTGGTCG
ATGAACTGTACAGGTTCGCCGCCCGGCTTGAAAAGAGACAGACCCCAGCGGAGG
GCTGGAGCCCGCCTATTCGACCATTAACGTCGGCACGATTTCCGGCGGCACGGCC
CGAAACATACTGGCTCGTGAATGTAAGTTCCACTGGGAGTTCCGTGGCTTGCCCG
GAATACCGCAGGACCTCGCCGTGCGGCATCTGGAGCGCTATGCCAGCGCCAAGG
TACTTCCAAGACTGCGCCGGTACACTGACGACGCCACCATATCCACTATTGTCGGC
ACAGAAGTGCCGGGCCTTGTGCGCGAGCCAGGTTCGATGGCGGAGACGCTGGCG
CTGAAACTCAGCGAGTCCAATGAGACGGTCGCGATGCCGTTCGCCACCGAGGCTG
GTCAGTTCCAAGTGCGCGGATATCCAACTGTGCTGTGCGGTCCGGGCAATGTGGA
CCAAGCGCATCAGCCAGATGAATACATCGCCGTGGATCAGGTCGAATCTTGCATG
*CGCTTCATGCGACGATTGGGCGAGTATTTGAAGGGA***TAA***CAAGCCAAACGTGCT*
TGAATCGAGTGTCAACCTGTATCATAGTATGAACCTATATGATAGTATGAAG
ATGCCTAGCTAACATGAGGTGAAATCGTGCATGCGTTGGGTGTGGTATTGAC
*AGGGGATATCTGCCCGACACGCGGTCTT***GTG***GCTGATGCCGCAGACGTAAAAC*
AGACGTTCCGGCTCATTGAAAACGCAGATTTTGCCCTCGGCAATTTTGAGATGCCG
CTCACGGATCAAGTTAGAGCAGTTGAGAAGCTGCTAAACATTAAGGCTGATCCGTC
CATCGCTCCTACACTTTCGGAGCTTGGGCTCGACATGGTGACGGTAGCCAACAAT
CACGCGGTCGATTACGGCTGGGAGGGGCTTTCCCAGACCATCGACCTGTTGCGGT
CGCAGGGTCTGACTGTCATAGGCGCCGGCTCAAACATCGCTGAGGCTACGGCGC
CGGCGATCGCAACGGTGGCAGGCAGGCAAATTGGAGTGCTTGCCTATTCGTGCCT
CCTGCCGACCGGCATGTCGGCAGGGATCAATCGACCTGGCATCGCTCCCATCCAC
GTGCAGACATCATACGAAATCGATCCCTACTATCAGATGGAGGAGCCTGGCGACC
TGTCTGCAGTGCGCGCCAGAACGCAAGCCAGGGAGACCGACGTGGAGGCGGCG
GTTGCGGCAATCCGCGCGCTCCGCAACAAGTGCGACACGCTCATCGTCACAATTC
ATTGGGGTTTCGGATCGGGAGAGGAGCTCGCCGAATATCAGCTACCGCTGGCGCA
AAGGCTAATTGACGCCGGCGCAGACATTATCCATGGGCACCATCCGCATGCAGTT
CACCCCATCGGCTTCTATCGCGGCAAGCCTATTTTCTTTGGTTTGGGCACCCTGGT
CGGCCAGCAGGTTTTTCTCGATGCGCCGCCGGCCGTGCGGGCGCTCTGGGCAGA
AATGTCGCCGGACGGATATGTCGCAACCATTTCGATTTCGCAATCCAACGTTCTCG
GGATCGAGTTAACCCCGACTACCCTGAATTCTGATCGCCTTCCCTTAATCGCAAAG
GACAATGATTTCGATCGGATACATCACCGTCTGGTGAGCCTGTCGGCGCCGTATG
GAGCCAATATCGAATTGGACGGAACGGTCCTGCGTGCCAGTGCCGCCGCACTGG
*ATTGCCAGAGAAAT***TGA***ATTTTCCACAACAAAAAAAGGGGAATAGCC***ATG***GAA*
AAGAAAACGACAACTATCTGCCTCGATGCGGAGATGACGCGCCGCTTGTTTACTG
GCGGTGCTGCAGCTTTCTTCGCGGCGCTTGGCCTCGGGGTAGGGTCCGAAGCCA
GTGCGCAGGAAAAGGTCATCGCCATCAGCTTTCCGAACTCCTCGACAATCGGTGC
AGTGATCACCACCCTCAATCAGGCAAAGCGCGCGGCAAGGAACTCGGCTTTAATG
*TCATCGTCGATGATCCGGGCAGCGACC***TGA***ACAAGCAGATCAACACGATCAAG*
*ACCTGGATCCAACAGAAGGTCGCCGTTATCGTCTGTGTCGTC***TTG***CAGCCACA*
GGTCTTTGTGGCGATCGCAAAACAAGCACGCGATGCCGGCATCGCGTGGATAACC
TATGGCGAGAAGATCGAGAATCAGAACGCCACGGTGGGGTATGCTCAATATGACG
ACGGTCGCCGCCTAGGCGAGTATGCTGGCCAGTGGATCACCGAGAATCTCGGCG
GCGCCGCCAAGGTCGCGATCCTCGGCTATGAAAGGGCTCATGGGGACAACTGC
GCGGCAATGGCATTAAGGACGGCCTTAAGGCACACGCTCCCAATGTAGAGATCGT
CGCTGAGCAAGACGCGATCACTCCGACGGAAGGCCTGAACGTCACCCGCACGAT

CCTCCAGGCGCATCCGGACGTGAACGTCATCCTCGGTGTTGAGGATCCAGCGACA
GAAGGCGCCTACAAGGCTTGGATTGCCGCCGGCAAGGATTCCAAGGATCCCAAG
GCCTTCATCGGCGGCATGGACGGAACGCCCCCGGCGCTTAAGCTGCTGCGCGAG
GGGG

Protein Sequences Obtained from Translation of ORFs

Open reading frames in the 25.2-kb region were translated to their corresponding

peptide sequences using DNASTAR software and are listed below ("aa" represents

amino acids).

orf-1 **(63 aa) Strand: + From: 138 To: 329**
MPGVVSTLAVAAGQPIKAGDVLLSIEAMKMETALHAERDGTIAEVLVRAGDQID
AKDLLVVYG

orf-2 **(114 aa) Strand: - From: 468 To: 812**
MAGLTVHVLDIMNGIPGGGMRIDFLRLDGTDKFHLKTLFTNAEGRTDEPVLTSS
DMMSGTFELMFHVKDYFRRGGTDATFFELVPVRFKIVSTHVHHHVPLVVAPWG
YQSYMGT

orf-3 **(291 aa) Strand: - From: 812 To: 1687**
MTLREFVDVFGGIFEHAPWVAERACLSRPFASVYEMHRKMMDVVRRCDESEKV
GLLRGHPELGGREARAGEMTPDSSSEQSRLGLDRLSEAEFSEMTELNRLYDEKFG
FPCMICLRLHVKRDTVVAEHRRRLLNDRQTEIENCIEQVGQITERRLRERVRESQ
RTRRAISTHVLDSGNAGGAQGMVIELQRFEGGGYTTVARVTTDEVGRAQLLDAE
EMRPGRYQMILQAGAYLAEHGTPSTFIDVIPVLFEVDDPYQNYHIPLIVGRYAFS
VYKGGIPALASGPTREDVR

orf-4 **(306 aa) Strand: +   From: 2421 To: 3341**
MGQMNGNHPQKDRNYRGYGRHAPRVFWPGGARVAVNIQVNYEEGAEYSYAR
DGRNEGAGEFLGSLSGIADRSVESAFEYGSRAGFWRLARLLDEYRIQATVNACA
VAVERNLEVGDYLREAKHEIACHGYRFEEVWNLTREEERQRIHAAIESLTKTCGE
RPVGWISRLMSSDNTRELLVEEGGFLYDSDALNDDLPYFVDVSNKPHLVVPLSFT
YNDGRFIMGGCDDPAAFAQYCCGALDELRREGITHPKMMTVSLHPRIIGQAGRI
VALRRFIEHASEAGDVWFARRSDIAQWWIDHHKEFAS

orf-5 **(396 aa) Strand: + From: 3364 To: 4554**
MSRPLRVTIIGAGIGGLSAAVALRKIGADVTVVERAPELRAAGAGICMWPNGAQ
ALHALGIANPLEMVSPILHRVCYRDQHGRVIREMSIDKLTELVGQRPFPLARSDL
QAALLSRLDPALVRLGGACVSVEQDANGVRAVLDDGTEIASDLLVGADGIRSVV
RNHVTGGTDRLRYHYTTWLGLVSFGLNLTPPGTFTFHVQDSKRVGLLNVGDDR

LYFFFDAVPSGEANPDGVRAELRHHFDGWCSEVTTLVEALDEAKTNRLPVHDLD
PLASFVNGRIVLIGDAAHATTPTLGQGGALAMEDSLVLARHLAESTDYGSALAS
YDNERLMRTRQVVLASRARTAATLGIDNTSAQTWQKQLTDDASQDFLEQLVDI
HRAGPLAAWPHDRQEGVAT

orf-6 (225 aa) Strand: + From: 4551 To: 5228
MTHSLIVSVDVITQHRPAFEEWLRSSFLAATGVPGVAADPVCFRGMRVETRYRT
YQPTPSYIVIFQLEGDPAEVTGAKAFSDWWTKGVAARFEWVEKQSWVVAELVS
GPPPPFDYSRVLFTQVDLRPAHQSGWARWYDEVHIPQSRLVPGMFRAENRRFAA
VEIATARWHCSAKPSFIHLVPIEDSADIVQGAATPEFLALIADTQAQWAGALESA
ASTLCERFR

orf-7 (377 aa) Strand: + From: 5225 To: 6358
MSRSGKLIVVGDVVPTRALPSDWHPFGEADFVFGDLEVPLTERGLRWDKPVSYR
AAPDRARELAGVGFGLVSLANNQAMNYGRTGLVDTITALDQAGVPFIGAGPDLE
AALKPHITELGGKRIAFIGLTCVAPRHWDAGADRSGLAVLRPRQAAEIDPAWAA
EEPGVAPTVHTWLDDDALARAAAAVQRALEHADLVVAAVHWGVGSSYQITGY
QRQLGHALLNAGAALVLGSHPPPLQGFERTNAGLISYSLGTFIRQQPQEGVGLAL
SAAYSRMPRESAILELTVKGGRFSEAHVHPAVLDDDGIARQVSGERARRIIRTILD
RSTGLIPNPAQGEEPETLTIHLDNEFATGSDRPSDGSMVPSGPEVAARQGPR

orf-8 (523 aa) Strand: + From: 7903 To: 9474
MTSNPAVQVGEAPRNGAPLLSLSGITKRYPTVVANENIDLDVSAGEIHAVLGENG
AGKSTLMKIIFGVAQPDAGTIYWQGHPVRIDSPAHARELGISMVFQHFSLFEAITV
AENISLTVPGTLKELSRRIRETGQEFGLKVEPAALVCNLSVGQRQRVEIIRCLLQEP
KLLILDEPTSVLPPPNVEQLFETLRRLVAKGMSILYISHKLEEIRSLCHSATILRQG
RVSGHVIPAETSSHALASLMIGRDLPKTAHPPARQDGAVRLRVANLSSHDPDPHI
VNIDDLSLEVRSGEILGIAGVSGNGQHMLSKMLSGEEVLAPSEKHRIELMGTAVA
HENAGRRRNLGLSFVPEERLGRGAVPPFSLADNSLLTGHRMGMTGRGMVRNQP
RDAFTDQCIADMNVVCAGSKSNAASLSGGNLQKFIVGREIMMAPKVLVIAQPTW
GIDVGAAAMVRQKLVDLRDTGTAILLISEELEELFEISDRIVVMFKGRASPALDAR
ATDAETIGRMMIGDIAAPLFSEAAQ

orf-9 (363 aa) Strand: + From: 9471 To: 10562
MNIRQPFVLVPRQRQSRIALIAVPLVTISLTVIAAFFLFAGLGADPGIVLYTFFIEPL
SSAYNLGEVLIKASPLILIAQGLAISFRAKVWNIGAEGQLIIGAICASLIPIYWSHSD
SMLMLPGMILIGAAAGMVWAGIAALLRTRFNASEVIVTLMLTEIARQLLYFLVT
GPLRDPMGYNFPQSVLFPSAALYPTFGGVGVRANLSIFITLAVTIICWVFVSKSFA
SFKLLVGGTTPAAARYAGFLPKQAVWISLLIGGAVAGLAGVGEIAGPIGKLQRIIS
PGYGFAAIIVAFLGGLNPIGIVFAGMLMAVVYVGGDNSLVTANIPASASVVFQGL
LLVFYLATAVFARYEIRRAPPPLIPA

orf-10 (307 aa) Strand: + From: 10572 To: 11495
MMDALTFVVAGAFAMATPLMLAALGELVTERSGVLNLSVEGMMAIAAAIAFM
TTHDTGSFLLGFAAGGVSGLLLSVVFALLVLVFLANQVVAGLAVGILGLGLSALL
ARSYEGMTILPMGKITIPILSDLPLVGPILFRQDVMVYLAIGSAITLAWFLFKTKRG

LVLRVVGENPQAAHSIGLRPIRVRFIAVLFGGLMAGLGGAYVSIVLTPLWSQGMI
AGRGWIAVALVVFACWRPIRLTVGAYLFGVVLLLDLAIQALGLSIPSVILTCMPY
ALTIVVLTIVSSDASRIRLNAPVSLGENFRPAN

orf-11 (359 aa) Strand: + From: 11533 To: 12612
MKIWVRTLLGATALGLLLSGPAASDPFKIGYIIPNAIGEAGWDHELERGRQAIVD
HFGDKVQVNAVNGVGEGPDATRVMNRMAADGTNMLILGAFGLMNDGLVLAR
RYPDLKVLHYGGYVNEPNFATLALRHYEASYLCGMAAGMAAKTGNLGVVAAF
PLPEVLSIMNAYVLGAQSVNPDIKPVKVVWLNSWFDPAKEKAAAESLASQGAE
VLYSLFPGTPSTVSTAEQLGVYVTVTLSDNTMFAPTKHLCSAQAEFGPALIRKIQ
DAIDGKFVGDDTFSGVKDGSMGIAGLSPDLSEEQKKQIMARLEEMRNGSFQPFR
GPIMSNTGQEIAAEGTNLDDTAIKSMKFLVKGIDTTLPN

orf-12 (269 aa) Strand: + From: 12768 To: 13577
MNFTPKSCLNGLKKFLIRAMLLVAASHGLSTSALAQTFSANQLQLHFGGGYRFG
GNGAETTTRTMLEFQHFSAYKYGDLFFFADVYRDHQWDGTSNRVNFYGEGYA
HLSAKSLGGIPFGEESFLADIGPGIGFNAGQDFLVAIYGARASFKVPGFSVLTFGV
YAYDNQIDPYGRDLDTTYQATLVWDIPFAIGDQKFSTGGVLDLIGSQGVGVEEQI
YFQPEIRWDVANALGQKAGSFDLGLRYVHFNNKYGVDGVDENAMSVFVSKKF

orf-13 (791 aa) Strand: + From: 14359 To: 16734
MFADINKGDAFGTWVGKSVPRREDADILAGRAEYIADIKLPGMLEAAFLRSPFA
HARIVSIDVSQALALPGVYDVMVGADIPDYVKPLPLMITYQNHRETPTSPLARDI
VRYAGEPVAVVAAINRYVAEDALELIVVKYEELPVVASIDASLAVDGPRLYEGW
PDNVVAKVSSEIGDVDAAMASADLVFEERFEIQRCHPAPLETRGFIAQWDFKGE
NLNVWNGTQIINQCRDFMSEVLDIPASKIRIRSPRLGGGFGAKFHFYVEEPAIVLL
AKRVKAPVRWIEDRLEAFSATVHAREQVIDVKLCAMNDGRITGIVADIKGDLGA
SHHTMSMGPVWLTSVMMTGVYLIPNARSVAKAIVTNKPPSGSYRGWGQPQANF
AVERMVDLLAHKLQLDPAAVRRINYVPEARMPYTGLAHTFDSGRYEVLHDRAL
KTFGYEAWLERQAAAQAQGRRIGIGMSFYAEVSAHGPSRFLNYVGGRQGGYDI
ARIRMDTTGDVYVYTGLCDMGQGVTNSLAQIAADALGLNPDDVTVMTGDTAL
NPYTGWGTGASRSITIGGPAVMRAATRLREKILSIARHWLQADPDTLVLANRGV
MVRDDPGRYVSFASIGRAAYCQIIELPEDVEPGLEAVGVFDTVQLAWPYGMNLV
AVEVDEDTGAVSFLDCMLVHDMGTIVNPMIVDGQLHGGIAQGIAQALYEELRY
DENGQLGTGSFADFLMPTASEIPNMRFDHMVTESPLIPGGMKGVGEGGTIGTPA
AVVNAIENALRPITNSKLNRTPVTPDRILTAISAGACA

orf-14 (297 aa) Strand: + From: 16725 To: 17618
MRMKPTAFDYIRPTSLPEALAILAEHSDDVAILAGGQSLMPLLNFRMSRPALVLD
INDISELQQVRCENDTLYVGSMVRHCRVEQEEIFRSTIPLMSEAMTSVAHIQIKTR
GTLGGNLCNAHPASEMPAVITALGASMVCKSEKRGERVLTPEEFFEGALQNGLQ
SDELLCEIRIPVPSQYVGWAFEEVARRHGDFAQCGAAVLIGAEDRKIDYARIALC
SIGETPIRFHALEQWLIGRPVGNDLPADVKLHCREILDVAEDSTMTAENRAKLAS
AVTSRAIARAADRIVHLDVKRG

orf-15 (167 aa) Strand: + From: 17623 To: 18126

MSSHVISLTVNGQAIERKVDSRTLLADFLRDELRLTGTHVGCEHGVCGACTIQFD
GEPARSCLMLAVQAEGHSIRTVEALAVDGCLGALQQAFHEKHGLQCGFCTPGLL
MTLDYALTADLHIDFSSDKEIRELISGNLCRCTGYQNIINAIKSVSPTTEIAKSEELV

orf-16 (149 aa) Strand: + From: 18123 To: 18572
MKFEGNYDIDAEPKKVWELLNNPEVLHKSIPGCEEVIQTSPVKYEAVVTLKIGPI
KARFQGHAEMTEQQPPNQCVIIFQGSGGIAGMARGEAKVLLTPNDRGTELTYIA
DVVIGGKIAQIGSKLIEGTARKIADQFFSNFVNYISVNVP

orf-17 (106 aa) Strand – From: 18559 To: 18879
MIACSAARRLTIRLEPHNPPAEERLGVRLLTRTTRSVAPTEAGTRLIVGVAPGVDR
IETELAAIDGLAVDQRNAISELFQMQCHLIRLRMNSTEHIATFNCVFTGH

orf-18 (466 aa) Strand: + From: 19308 To: 20708
MDLFIKNVRVITENGSFEGGVAVDKGVIAQITGANGHVDADETIDAEGLYLLPGII
DAHSHFSHPGREFEGFEAGTRASAAGGVTTAIDMPLADIPPMATAKAHAAKYEM
VKDLCVADYAFWGGLIDDNLEHLDEQNELGVVAYKAFMRRSSSYPKVDDGQLF
TGLQKTAKFGNIVGVHAENDTLIEHLEKRLKAEGRHDRMAWNDSHPIESELEAI
NRAILLAKAAASSLYIVHISWAGGVDTVREAAHRGQDVYAEVCAHHLCFDLDD
YLRVGPRLRSAPPIRSRPEVEELWDRVLRGKVDVIASDHSPFLPDLYKAGENDV
WQGTGGITGIQSILPAIITEGVHMRGLSWELLVKMMSSNPARIFGIYPRKGAIVPG
ADADFALVDPQKVWTLQTEDLFYRYKESPYVGKEFCGRVERTIVRGKTVFLEGK
ITADKGYGQLVRRESATHGRPLHKHQQSVRHL

orf-19 Strand: 266 aa Strand: + From: 20736 To: 21536
LVAGGLPDQRQLLSSRAVFTPAYAIIPGSVLRDIVSSALPEWRDTRLWVLARPLS
GFAETFSQYYMEVSPGGGSEDPEPDPHAQSAIFVTAGTMTITTNGVQHRMVKGS
FAFIPAGQVWQIINPSAEICRFQWIRKKFEPAGKLGTPPAIFTHEDKAAGVSMCDE
PPLWGSVRFIDPSDLRYDMHINIVWFEGGATIPFMETHVMEHGIYILEGKAVYRL
NSDWVEVEAGDFLWLRAFCPQALYAGGPDGLRYLLYKDVNRHMKLL

orf-20 (429 aa) Strand: + From: 21910 To: 23199
MIENDWVSILELTKLMDCRKKSSSEGNFPPDTKIWSGTMNGSSAVGLESTLNVIE
RFISFNTESSRSNLELVDWVEGSVKGPRVSCFRTFNAEKSKATLCITVGPPVDGG
VVLSGHIDVVPVEGQAWSSDPYSLRQADGRLYGRGTCDMKAFGAIAVAHIPYFL
ASNIAKPVHIVLSYDEEVTSRGSIAVIEEFGRSLPKPSSVIVGEPTMMKVADTHKSI
STYYTTVTGHEAHSSKPNIGANAIMGACELVDELYRFAARLEKETDPSGGLEPAY
STINVGTISGGTARNILARECKFHWEFRGLPGIPQDLAVRHLERYASAKVLPRLRR
YTDDATISTIVGTEVPGLVREPGSMAETLALKLSESNETVAMPFATEAGQFQVRG
YPTVLCGPGNVDQAHQPDEYIAVDQVESCMRFMRRLGEYLKG

orf-21 (343 aa) Strand: + From: 23347 To: 24378
VADAADVKQTFRLIENADFALGNFEMPLTDQVRAVEKLLNIKADPSIAPTLSELG
LDMVTVANNHAVDYGWEGLSQTIDLLRSQGLTVIGAGSNIAEATAPAIATVAGR
QIGVLAYSCLLPTGMSAGINRPGIAPIHVQTSYEIDPYYQMEEPGDLSAVRARTQ
ARETDVEAAVAAIRALRNKCDTLIVTIHWGFGSGEELAEYQLPLAQRLIDAGADII

HGHHPHAVHPIGFYRGKPIFFGLGTLVGQQVFLDAPPAVRALWAEMSPDGYVAT
ISISQSNVLGIELTPTTLNSDRLPLIAKDNDFDRIHHRLVSLSAPYGANIELDGTVLR
ASAAALDCQRN

orf-22 (84 aa) Strand: + From: 24409 To: 24663
MEKKTTTICLDAEMTRRLFTGGAAAFFAALGLGVGSEASAQEKVIAISFPNSSTIG
AVITTLNQAKRAARNSALMSSSMIRAAT

orf-23 (168 aa) Strand: + From: 24729 To: 25232
LQPQVFVAIAKQARDAGIAWITYGEKIENQNATVGYAQYDDGRRLGEYAGQWI
TENLGGAAKVAILGYEKGSWGQLRGNGIKDGLKAHAPNVEIVAEQDAITPTEGL
NVTRTILQAHPDVNVILGVEDPATEGAYKAWIAAGKDSKDPKAFIGGMDGTPPA
LKLLREG

APPENDIX C

SEQUENCE SIMILARITY OF ALL THE OPEN READING FRAMES

IDENTIFED ON CAFFEINE GENE CLUSTER OF CBB1

Table C1. A complete list of all the 23 open reading frames (ORFs), including two incomplete ORFs (*orf-1* and *orf-23*) of the caffeine gene cluster present on the 25.2-kb genomic DNA fragment of *Pseudomonas* sp. strain CBB1.

Gene no.	No. of codon	Homologous protein	Database*	GenBank accession no.: Organism name	Identity (%)¥
Gene-1	108 / 1151	pyruvate carboxylase	NR	CCE97797.1 *Sinorhizobium fredii*	87
		Glutaconyl-CoA decarboxylase subunit gamma	SwissProt	Q9ZAA7.1 *Acidaminococcus fermentans* DSM 20731	46
Gene-2 *tmuH*	114	hydroxyisourate hydrolase	NR	YP_002976942.1 *Rhizobium leguminosarum*	48
		5-hydroxyisourate hydrolase 2	SwissProt	Q92UG5.2 *Sinorhizobium melilot* 1021	45
Gene-3 *tmuD*	291	OHCU decarboxylase	NR	YP_003695277.1 *Starkeya novella* DSM 506	41
		OHCU decarboxylase	SwissProt	D4GPU8.1 *Haloferax volcanii* DS2	40
Gene-4	396	Chitin deacetylase	NR	ZP_08883006.1 *Saccharopolyspora spinosa*	56
		Chitin deacetylase1	SwissProt	O13842.1 *Schizosaccharomyces pombe*	34
Gene-5 *tmuM*	385	FAD-binding monooxygenases	NR	YP_003741647.1 *Erwinia billingiae* Eb661	39
		Zeaxanthin epoxidase	SwissProt	Q40412.1 *Nicotiana plumbaginifolia*	32

Table C1 continued

Gene	No. of codon	Homologous protein	Database	GenBank accession no.: Organism name	Identity (%)
Gene-6	255	transcriptional regulator (MerR)	NR	NP_824803.1 *Streptomyces avermitilis*	29
		TBCC domain-containing protein	SwissProt	Q5FVR8.1 *Rattus norvegicus*	31
Gene-7	377	capsule synthesis protein, capA	NR	YP_003320833.1 *Sphaerobacter thermophilus*	36
		PGA biosynthesis protein CapA	SwissProt	P96738.1 *Bacillus subtilis*	29
Gene-8	523	ABC transporter, fused ATPase	NR	ZP_05782548.1 *Citreicella* sp. SE45	60
		Ribose import ATP-binding protein RbsA 1	SwissProt	Q1AXG5.1 *Rubrobacter xylanophilus* DSM 9941	37
Gene-9	363	ABC transporter, inner membrane subunit	NR	YP_559078.1 *Burkholderia xenovorans* LB400	52
		Uncharacterized ABC transporter permease (yufP)	SwissProt	O05254.1 *Bacillus subtilis*	
Gene-10	307	ABC transporter permease (sugar)	NR	YP_472019.1 *Rhizobium etli* CFN 42	55
		Uncharacterized ABC transporter permease (yufQ)	SwissProt	O05255.1 *Bacillus subtilis*	41
Gene-11	359	basic membrane lipoprotein Purine-binding protein	NR	ZP_05781700.1 *Citreicella* sp. SE45	41
			SwissProt	Q2YKI6.1 *Brucella melitensis*	33

Table C1 continued

Gene	No. of codon	Homologous protein	Database	GenBank accession no.: Organism name	Identity (%)
Gene-12	269	nucleoside-specific channel-forming protein Tsx	NR	YP_003023750.1 *Geobacter* sp. M21	27
		Uncharacterized protein C11orf52	SwissProt	Q96A22.2 *Homo sapiens*	32
Gene-13 *cdhA*	791	Xanthine dehydrogenase molybdopterin binding subunit	NR	YP_003395893.1 *Conexibacter woesei* DSM 14684	50
		Carbon monoxide dehydrogenase large chain	SwissProt	P19919.2 *Oligotropha carboxidovorans* OM5	35
Gene-14 *cdhB*	297	Alcohol dehydrogenase medium subunit	NR	ADV16272.1 *Alicycliphilus denitrificans*	41
		CO dehydrogenase medium chain	SwissProt	P19914.2 *H. pseudoflava*	32
Gene-15 *cdhC*	167	aldehyde oxidase small subunit	NR	YP_002521823.1 *Thermomicrobium roseum*	66
		CO dehydrogenase small chain	SwissProt	P19915.2 *H. pseudoflava*	51
Gene-16	149	CO dehydrogenase subunit G (CoxG)	NR	YP_004305439.1 *Polymorphum gilvum*	50
		Tripartite motif-containing protein	SwissProt	Q5BIM1.1 *Bos taurus* (cattle)	27
Gene-17	106	transcriptional regulator (lysR)	NR	YP_004730098.1 *Salmonella bongori*	72
		Histone deacetylase 4	SwissProt	Q613L4.1 *Caenorhabditis briggsae*	30

Table C1 continued

Gene	No. of codon	Homologous protein	Database	GenBank accession no.: Organism name	Identity (%)
Gene-18 *orf1*	466	allantoinase	NR	ZP_09354491.1 *Bacillus smithii* 7-3-47FAA	42
		Allantoinase	SwissProt	Q1AS71.1 *Rubrobacter xylanophilus*	45
Gene-19 *orf2*	266	allantoin catabolism protein	NR	EFY89705.1 *Metarhizium acridum*	61
		YlbA (encoded by *ylbA*)	SwissProt	P75713.1 *Escherichia coli* K-12	26
Gene-20 *orf3*	429	acetylornithine deacetylase ArgE	NR	YP_001832172.1 *Beijerinckia indica*	60
		Acetylornithine deacetylase	SwissProt	Q9CLT9.2 *Pasteurella multocida*	60
Gene-21	343	capsule synthesis protein, capA	NR	YP_003320833.1 *Sphaerobacter thermophilus*	36
		PGA biosynthesis protein CapA	SwissProt	P96738.1 *Bacillus subtilis*	29
Gene-22	84	alkyl hydro-peroxide reductase	NR	YP_004602138.1 *[Cellvibrio] gilvus*	38
		7-cyano-7-deazaguanine synthase	SwissProt	Q2FS65.1 *Methanospirillum hungatei* JF-1	37
Gene-23	168	sugar ABC transporter protein	NR	YP_004223511.1 *Microbacterium testaceum*	39
		D-ribose-binding periplasmic protein	SwissProt	P0A2C5.1 *Salmonella typhimurium*	37

¥ % identity was determined by aligning the gene product of each *orf* with the homologous protein using the ClustalW2 [Larkin *et al.*, 2007].

*Database used in BlastP search.

APPENDIX D

CAFFEINE CONTENT OF COMMERCIALLY AVAILABLE

PRODUCTS

Table D1. A comprehensive list of all the known food, beverage and other consumables containing caffeine*.

Product	Serving size	Caffeine consumed per serving (mg)	Caffeine consumed per unit volume (mg/oz.)	(ppm)
Coffee				
Generic Brewed (Espresso)	1 oz.	40 (range: 30-90)	40	1352
Generic Brewed (Regular)	8 oz.	133 (range: 102-200)	16.6	562
Generic Brewed (Instant)	8 oz.	93 (range: 27-173)	11.6	393
Generic (Decaffeinated)	8 oz.	5 (range: 3-12)	0.6	21
Starbucks (Espresso) doppio	2 oz.	150	75	2536
Starbucks (Espresso) solo	1 oz.	75	75	2536
Starbucks Brewed (Grande)	16 oz.	320	20	676
Starbucks Latte (Grande)	16 oz.	150	9.4	318
Starbucks (Espresso Decaf.)	1 oz.	4	4	135
Einstein Bros. (Espresso)	1 oz.	75	75	2536
Einstein Bros. (Regular)	16 oz.	300	18.7	634
Dunkin Donuts (Regular)	16 oz.	206	13	440

Table D1 continued

Product	Serving size	Caffeine consumed per serving (mg)	Caffeine consumed per unit volume	
			(mg/oz.)	(ppm)
Tea				
Generic Brewed tea	8 oz.	53 (range: 40-120)	6.625	224
Black tea 1 Minute brew	5 oz.	21 - 33	4.2 - 6.6	142 - 223
Black tea 3 Minute brew	5 oz.	35 - 46	7 - 9.2	237 - 311
Black tea 5 Minute brew	5 oz.	39 - 50	7.8 - 10	264 - 338
Canned iced tea	5 oz.	22 - 36	4.4 – 7.2	149 - 243
Iced tea	12 oz.	67 - 76	5.6 – 6.3	190 - 214
Instant tea	5 oz.	22-36	4.4 – 7.2	149 - 243
Generic Brewed (Grande)	8 oz.	53 (range: 40 - 120)	6.6	224
Starbucks Tazo Chai Latte	16 oz.	100	6.25	211
Snapple, Lemon (Diet Version)	16 oz.	42	2.6	89
Snapple, Peach (Diet Version)	16 oz.	42	2.6	89
Snapple, Plain Unsweetened	16 oz.	18	1.1	38
Snapple, Kiwi Teawi	16 oz.	10	0.6	21
Arizona iced tea, black	16 oz.	32	2	68
Arizona iced tea, green	16 oz.	15	0.9375	32
Nestea	12 oz.	26	2.2	73
Loose-leaf tea (Black)	N.A.	25-110	N.A.	N.A.

Table D1 continued

Product	Serving size	Caffeine consumed per serving (mg)	Caffeine consumed per unit volume (mg/oz.)	(ppm)
Soft Drinks				
Coca cola (Regular)	12 oz.	35	2.9	98.6
Coca cola (Diet coke)	12 oz.	47	3.9	132
Coca cola lime (Diet)	12 oz.	47	3.9	132
Coke Red (Reg. and Diet)	12 oz.	54	4.5	152
Coke Black	12 oz.	69	5.75	194
Coke Black Cherry Vanilla	12 oz.	35	2.9	98.6
Pepsi (Regular)	12 oz.	38	3.2	107
Pepsi light (Diet)	12 oz.	36	3	101
Pepsi Lime	12 oz.	38	3.2	107
Pepsi Vanilla	12 oz.	37	3	101
Pepsi Twist	12 oz.	38	3.2	107
Pepsi One	12 oz.	54	4.5	152
Mountain Dew (Reg. and Diet)	12 oz.	54	4.5	152
Mountain Dew MDX	12 oz.	71	5.9	200
Dr. Peeper (Regular)	12 oz.	42	3.5	118
Dr. Peeper (Diet)	12 oz.	44	3.6	124
Mr. Pibb (Regular)	12 oz.	41	3.4	116

Table D1 continued

Product	Serving size	Caffeine consumed per serving (mg)	Caffeine consumed per unit volume	
			(mg/oz.)	(ppm)
Mr. Pibb (Diet, Zero, and Xtra)	12 oz.	41	3.4	116
Tab, Shasta Cola (Reg. & Diet)	12 oz.	45	3.75	127
Royal Crown Cola (Reg./Diet)	12 oz.	36	3	101
Sunkist Orange Soda	12 oz.	41	3.4	116
Jolt Cola	12 oz.	72	6	203
Vault	12 oz.	71	5.9	200
Alcoholic Beverages[¥]				
MillerCoors Sparks	16 oz.	214	13.4	452
Anheuser-Busch	10 oz.	55	5.5	186
Energy Drinks				
5-hour Energy (Extra Strength)	2 oz.	242	121	4091
5-hour Energy (Regular)	2 oz.	215	107	3635
5-hour Energy (Decaf)	2 oz.	6	3	101
Spike shooter	8.4 oz.	300	36	1208
Cocaine	8.4 oz.	288	34	1159
Monster Energy	16 oz.	160	10	338
Full Throttle	16 oz.	144	9	304
Rip it	8 oz.	100	12.5	423

Table D1 continued

Product	Serving size	Caffeine consumed per serving (mg)	Caffeine consumed per unit volume	
			(mg/oz.)	(ppm)
SoBe No Fear	8 oz.	83	10.4	351
SoBe Adrenaline Rush	8.3 oz.	79	9.5	322
Red Bull (Reg. and Sugar-free)	8.3 oz.	80	9.6	326
Rockstar Energy	8 oz.	80	10	338
Amp	8.4 oz.	75	8.9	302
Tab Energy	10.5 oz.	95	9	306
Glaceau Vitamin Water	20 oz.	50	2.5	84.5
Over-the Counter Drugs Diuretics				
Aqua Ban	1 tab.	200	200 mg / tablet	
Permathene Water Off	1 tab.	200	200 mg / tablet	
Pre-Mens Forte	1 tab.	100	100 mg / tablet	
Diurex	2 tabs	100	50 mg / tablet	
Weight control aids				
Zantrex-3	2 tab	1,223	611.5 mg / tablet	
Dexatrim	1 tab	200	200 mg / tablet	
Dietrac	1 tab	200	200 mg / tablet	
Prolamine	1 tab	140	140 mg / tablet	
Appredrine	1 tab	100	100 mg / tablet	

Table D1 continued

Product	Serving size	Caffeine consumed per serving (mg)	Caffeine consumed per unit volume (mg/oz.)	(ppm)
Stimulants				
No Doz (max)	1 tab.	200	200 mg / tablet	
Caffedrine capsules	1 tab.	200	200 mg / tablet	
Vivarin	1 tab.	200	200 mg / tablet	
Cold/Allergy Remedies				
Dristan / Drisdan A-F	1 tab.	30	30 mg / tablet	
Triaminicin	1 tab.	30	30 mg / tablet	
Coryban-D	2 tab.	32	16 mg / tablet	
Neo-Synephrine	2 tab.	32	16 mg / tablet	
Pain Relievers				
Menstrual Relief	1 tab.	60	60 mg / tablet	
Anacin	2 tab.	64	32 mg / tablet	
Excedrin	2 tab.	130	65 mg / tablet	
Midol	2 tab.	64	32 mg / tablet	
Vanquish	2 tab.	66	33 mg / tablet	
Cope	1 tab.	32	32 mg / tablet	
Other Caffeine Products Containing Commonly Available to Children. Like Breakfasts/ Chocolates/ Snack Bars				
Pit Bull Energy Bar	1 oz.	165	165	5579

Table D1 continued

Product	Serving size	Caffeine consumed per serving (mg)	Caffeine consumed per unit volume (mg/oz.) (ppm)	
Snickers Charged	1.83 oz.	60	32.8	1108
Hershey's Dark Chocolate Bar	1.45 oz.	31	21.8	737
Hot Cocoa mix	8 oz.	3-10	0.4 – 1.2	13 - 42
SumSeeds (Energized seeds)	3.5 oz.	120	34	1159
Oatmeal (Morning spark)	1 packet	60	60 mg. / meal	
Breath Fresheners/Gums[¥]				
Jolt Caffeinated Gum	1 stick	33	33 mg / stick of gum	
Penguin's Mints (all 3 types)	3 mint	21	7 mg / tablet of mint	
Foosh Energy Mints	1 mint	100	100 mg / mint tablet	
Perky Jerky	1 oz.	150	150	5072
Wrigley's Alert Energy Gum	8 piece.	320	40 mg / piece of gum	
Other Commonly Use Products				
Shower Shock Bathing Soap	12 wash	2400	200 mg / wash	
Spazz Stick lip balm	1 stick.	250	250 mg / stick	
Spot on Energy Patches/Tattoo	1 Patch.	20	20 mg / Patch	

*A complete list of 150 items containing caffeine is available on the Web at the following address http://www.energyfiend.com/caffeine-in-candy. Other sources of information are Barone, 1996; Carillo, 2000; Juliano, 2005; McCusker, 2006; Reissig, 2009; Heckman, 2010; Gopishetty, 2011; Summers, 2011 and other Web sites like FDA and www.cspinet.org/new/cafchart.htm.

[¥] These products are either withdrawn from market or currently under FDA scrutiny for potential health safety related issues.

APPENDIX E

PUBLICATIONS AND PRESENTATIONS

Publications

Mohanty, S. K., Yu, C. L., Das, S., Louie, T. M., Gakhar, L., & Subramanian, M. (2012). Delineation of the Caffeine C-8 Oxidation Pathway in Pseudomonas sp. Strain CBB1 via Characterization of a New Trimethyluric Acid Monooxygenase and Genes Involved in Trimethyluric Acid Metabolism. *Journal of Bacteriology* **194**(15): 3872-3882.

Mohanty, S. K., Yu, C. L., Gopishetty, S., & Subramanian, M. (2013). Validation of Caffeine Dehydrogenase as a Suitable Enzyme for Development of a Rapid Caffeine Diagnostic Test. *Chemical Research in Toxicology* (manuscript submitted).

GeneBank Contributions (Nucleotide)

1. JQ743481.1, *Pseudomonas* sp. CBB1 TmuM (*tmuM*) gene, complete cds.
2. JQ743482.1, *Pseudomonas* sp. CBB1 TmuH (*tmuH*) gene, complete cds.
3. JQ743483.1. *Pseudomonas* sp. CBB1 TmuD (*tmuD*) gene, complete cds.

Presentations

Mohanty, S. K., Yu, C. L., & Subramanian, M. (2012). Development of a Rapid Colorimetric Assay for Detection of Caffeine in Beverages and Body Fluids Using Caffeine Dehydrogenase from *Pseudomonas* sp. CBB1. University of Iowa 21th Center for Biocatalysis and Bioprocessing Conference, Iowa City, IA, USA.

Mohanty, S. K., Yu, C. L., Louie, T. M., & Subramanian, M. (2012). Development of a Rapid Colorimetric Assay for Detection of Caffeine in Beverages and Body Fluids Using Caffeine Dehydrogenase from *Pseudomonas* sp. CBB1. 112th General Meeting of American Society for Microbiology, San Francisco, CA, USA. (**ASM Student Award**).

Mohanty, S. K., Yu, C. L., Louie, & Subramanian, M. (2011). New Pathway for Caffeine Degradation in *Pseudomonas* sp. CBB1: Novel Enzymes, Genes and Metabolites. . University of Iowa College of Engineering Research Open House, Iowa City, IA, USA.

Mohanty, S. K., Yu, C. L., Louie, & Subramanian, M. (2011). New Pathway for Caffeine Degradation in *Pseudomonas* sp. CBB1: Novel Enzymes, Genes and Metabolites. University of Iowa 20th Center for Biocatalysis and Bioprocessing Conference, Iowa City, IA, USA. (**Usha Balakrishnan Outstanding Student Poster Award**)

Yu, C. L., Mohanty, S. K., & Subramanian, M. (2011). Purification, Characterization, and Molecular Cloning of a NAD(P)H-Dependent Trimethyluric Acid (TMU) Oxidoreductase from Caffeine-Degrading Strain *Pseudomonas* sp. CBB1. 111th General Meeting of American Society for Microbiology, New Orleans, LA, USA.

Mohanty, S. K., Yu, C. L., Louie, T. M., & Subramanian, M. (2010). Purification of a Novel NAD(P)H-Dependant Trimethyluric Acid (TMU) Oxidoreductase Responsible for TMU Degradation in *Pseudomonas* sp. CBB1. University of Iowa College of Engineering Research Open House, Iowa City, IA, USA.

Yu, C. L., Mohanty, S. K., & Subramanian, M. (2010). Purification and Characterization of a NAD(P)H Dependent Trimethyluric Acid (TMU) Oxidoreductase from Caffeine-Degrading Strain *Pseudomonas* sp. CBB1. 19th Center for Biocatalysis and Bioprocessing Conference, Iowa, IA.

Yu, C. L., Mohanty, S. K., Louie, T. M., & Subramanian, M. (2009). A Novel NAD(P)H Dependent Trimethyluric Acid (TMU) Oxidoreductase is Responsible for TMU Degradation in *Pseudomonas* sp. CBB1. 18th Center for Biocatalysis and Bioprocessing Conference, Iowa City, IA.

REFERENCES

Akyilmaz, E., & Turemis, M. (2010) An inhibition type alkaline phosphatase biosensor for amperometric determination of caffeine. *Electrochimica Acta.* **55**: 5195-5199.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* **25**(17): 3389-3402.

Anaya, A. L., Cruz-Ortega, R., & Waller, G. R. (2006). Metabolism and ecology of purine alkaloids. *Frontiers in Bioscience* **11**: 2354-2370.

Arnaud, M. J. (2011). Pharmacokinetics and metabolism of natural methylxanthines in animal and man, p 33–91. *In* Fredholm BB (ed), Handbook of experimental pharmacology, vol 200. Springer-Verlag, Berlin, Germany.

Asano, Y., Komeda, T., & Yamada, H. (1994). Enzymes involved in theobromine production from caffeine by *Pseudomonas putida* No. 352. *Bioscience, Biotechnology, and Biochemistry* **58**(12): 2303-2304.

Ashihara, H., & Crozier, A. (1999). Biosynthesis and metabolism of caffeine and related purine alkaloids in plants. *Advances in Botanical Research* **30**: 117-205.

Ashihara, H., & Crozier, A. (2001). Caffeine: a well-known but little mentioned compound in plant science. *TRENDS in Plant Science* **6**(9): 407-13.

Ashihara, H., & Suzuki, T. (2004). Distribution and biosynthesis of caffeine in plants. *Frontiers in Bioscience* **9**: 1864-1876.

Ashihara, H., Monteiro, A. M., Moritz, T., Gillies, F. M., & Crozier, A. (1996). Catabolism of caffeine and related purine alkaloids in leaves of *Coffea arabica* L. *Planta* **198**: 334-339.

Ashihara, H., Sano, H., & Crozier, A. (2008) Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochemistry* **69**(4): 841-856.

Barone, J. J., & Roberts, H. R. (1996). Caffeine consumption. *Food and Chemical Toxicology* **34**(1): 119-129.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2008). GenBank. *Nucleic Acids Research* **36**: D25-D30.

Berthou, F., Flinois, J. P., Ratanasavanh, D., Beaune, P., Riche, C., & Guillouzo, A. (1991). Evidence for the involvement of several cytochromes P-450 in the first steps of caffeine metabolism by human liver microsomes. *Drug Metabolism and Disposition* **19**(3): 561-567.

Berthou, F., Guillois, B., Riche, C., Dreano, Y., Jacqz-Aigrain, E., Beaune, P. H. (1992). Interspecies variations in caffeine metabolism related to cytochrome P4501A enzymes. *Xenobiotica* **22**(6): 671-680.

Bhat, V. B., & Madyastha, K. M. (2001). Antioxidant and radical scavenging properties of 8-oxo derivatives of xanthine drugs pentoxifylline and lisofylline. *Biochemical and Biophysical Research Communications* **288**: 1212-1217.

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., & Mayhew, G. F. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**(5331): 1453-1462.

Blecher, R., & Lingens, F. (1977). The metabolism of caffeine by a *Pseudomonas putida* strain. *Hoppe-Seyler's Zeitschrift für Physiologishe Chemie* **358**: 807-817.

Bongaerts, G. P. A., & Vogels, G. D. (1976). Uric Acid Degradation by *Bacillus fastidiosus* Strains. *Journal of. Bacteriology* **125**: 689-697.

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

Brand, D., Pandey, A., Roussos, S., & Soccol, C. R. (2000). Biological detoxification of coffee husk by filamentous fungi using a solid state fermentation system. *Enzyme and Microbial Technology* **27**(1-2): 127-133.

Brondino, C., Romão, M. J., Moura, I., & Moura, J. J. G. (2006). Molybdenum and tungsten enzymes: the xanthine oxidase family. *Current Opinions in Chemical Biology* **10**: 109 –114.

Bruton, T., Alboloushi, A., de la Garza, B., Kim, B. O., & Halden, R. U. (2010) Fate of Caffeine in the Environment and Ecotoxicological Considerations, p 257-273. *In* Halden, R. U.(ed), Contaminants of Emerging Concern in the Environment: Ecological and Human Health Considerations. *ACS Symposium Series*. Vol. 1048, ACS Publications, USA.

Buerge, I. J., Poiger, T., Muller, M. D., & Buser, H. R. (2003). Caffeine, an anthropogenic marker for wastewater contamination of surface waters. *Environmental Science & Technology* **37**(4): 691-700.

Burnette, W. N. (1981). "Western blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Analytical Biochemistry* **112**(2): 195-203.

Campbell, M. E., Grant, D. M., Inaba, T., & Kalow, W. (1987). Biotransformation of caffeine, paraxanthine, theophylline, and theobromine by polycyclic aromatic hydrocarbon-inducible cytochrome (s) P-450 in human liver microsomes. *Drug Metabolism and Disposition* **15**(2): 237-249.

Carrillo, J. A., & Benitez, J. (2000). Clinically significant pharmacokinetic interactions between dietary caffeine and medications. *Clinical Pharmacokinetics* **39**(2): 127-153.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., & Thompson, J. D. (2003). Multiple sequence alignment with the clustal series of programs. *Nucleic Acids Research* **31**(13): 3497-3500.

Daly, J. W. (2007). Caffeine analogs: Biomedical impact. *Cellular and Molecular Life Sciences* **64**(16): 2153-2169.

Daneshvar, A., Aboulfadl, K., Viglino, L., Broséus, R., Sauvé, S., Madoux-Humery, A. S., Weyhenmeyer, G. A., & Prévost, M. (2012) Evaluation Pharmaceuticals and caffeine as indicators of fecal contamination in drinking water sources of the Greater Montreal region. *Chemosphere* **88**(1): 131-139.

Dash, S. S., & Gummadi, S. N. (2006). Catabolic pathways and biotechnological applications of microbial caffeine degradation. *Biotechnology Letters* **28**: 1993-2002.

Dash, S. S., & Gummadi, S. N. (2007). Degradation kinetics of caffeine and related methylxanthines by induced cells of *Pseudomonas* sp. *Current Microbiology* **55**: 56-60.

Dash, S. S., & Gummadi, S. N. (2008). Inducible nature of the enzymes involved in catabolism of caffeine and related methylxanthines. *Journal of Basic Microbiology* **48**: 227-233.

Davis, R. W., Botstein, D., Roth, J. R., (1980). Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

de la Riva, L., Badia, J., Aguilar, J., Bender, R. A., & Baldoma, L. (2008). The *hpx* genetic system for hypoxanthine assimilation as a nitrogen source in *Klebsiella pneumoniae*: Gene organization and transcriptional regulation. *Journal of Bacteriology* **190**(24): 7892-7903.

Durrant, K. L. (2002) Known and hidden sources of caffeine in drug, food, and natural products. *Journal of the American Pharmacist Association.* **42**(4): 625-637.

Edman, P. (1950). Method for determination of the amino acid sequence in peptides. *Acta Chemica Scandinavica* **4**: 283-293.

Escohotado, A. & Symington, K. (1999). A Brief History of Drugs: From the Stone Age to the Stoned Age. Park Street Press. ISBN 0-89281-826-3.

Eteng, M. U., Eyong, E. U., Akpanyung, E. O., Agaiang, M. A., & Aremu, C. Y. (1997). Recent advances in caffeine and theobromine toxicities: a review. *Plant Foods for Human Nutrition.* **51**: 231–243.

Feldman, J. R., & Katz, S. M. (1976). Caffeine. In *Encyclopedia of Chemical Processing and Design: Blowers to Calcination.* Edited by J. J. McKetta. Marcel Dekker, New York, NY, p. 424-440.

Fiser, A., & Sali, A. (2003). Modeller: generation and refinement of homology-based protein structure models. *Methods in Enzymology* **374**: 461– 491.

Franco E. J., Sonneson, G. J., DeLegge, T. J., Hofstetter, H., Horn, J. R., Hofstetter, O. (2010) Production and characterization of a genetically engineered anti-caffeine camelid antibody and its use in immunoaffinity chromatography. *Journal of Chromatography B*. **878**: 177-186.

French, J. B., & Ealick, S. E. (2010). Structural and mechanistic studies on *Klebsiella pneumoniae* 2-oxo-4-hydroxy-4-carboxy-5- ureidoimidazoline decarboxylase. Journal of Biological Chemistry **285**: 35446 –35454.

French, J. B., & Ealick, S. E. (2011). Structural and kinetic insights into the mechanism of 5-hydroxyisourate hydrolase from *Klebsiella pneumoniae*. *Acta Crystallographica*. D **67**: 671– 677.

French, J. B., Neau, D. B., & Ealick, S. E. (2011). Characterization of the structure and function of *Klebsiella pneumoniae* allantoin racemase. Journal of Molecular Biology **410**: 447– 460.

Garriott, J. C., Simmons, L. M., Poklis, A., & Mackell, M. A. (1985). Five cases of fatal overdose from caffeine containing "look-alike" drugs. *Journal of Analytical Toxicology* **9**(3): 141-143.

Gilbert, R. M. (1984). Caffeine consumption. Progress in Clinical and Biological Research **158**:185–213.

Gokulakrishnan, S., Chandraraj, K., & Gummadi, S. N. (2005). Microbial and enzymatic methods for the removal of caffeine. *Enzyme and Microbial Technology* **37**: 225-232.

Gopishetty, S. R., Louie, T. M., Yu, C. L., & Subramanian, M. V. (2011). Microbial degradation of caffeine, methylxanthines, and its biotechnological applications, p 44–67. *In* Thatoi HN, Mishra BB (ed), Microbial biotechnology methods and applications. Narosa Publishing House Pvt, Ltd, New Delhi, India.

Gouet, P., Courcelle, E., Stuart, D. I., & Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305-308.

Graham, A. Q., Hathaway, C., & Geisberg, M. S. (2012) Caffeine detection via internally referenced two part assay. U.S. patent No.: 8,137,984 B2.

Greenhagen, B. T., Shi, K., Robinson, H., Gamage, S., Bera, A. K., Ladner, J. E., & Parsons, J. F. (2008). Crystal structure of the pyocyanin biosynthetic protein PhzS. *Biochemistry* **47**: 5281–5289.

Guzmán, K., Badia, J., Giménez, R., Aguilar, J., & Baldoma L. (2011). Transcriptional regulation of the gene cluster encoding allantoinase and guanine deaminase in *Klebsiella pneumoniae*. *Journal of Bacteriology*. **193**: 2197–2207.

Hakil, M., Denis, S., Viniegra-Gonzalez, G., & Augur, C. (1998). Degradation and product analysis of caffeine and related dimethylxanthines by filamentous fungi. *Enzyme and Microbial Technology* **22**: 355-359.

Harayama, S., Kok, M., Neidle, E. L. (1992). Functional and evolutionary relationships among diverse oxygenases. *Annual Review of Microbiology* **46**:565–601.

Heckman, M. A., Weil, J., & De Mejia, E. G. (2010). Caffeine (1, 3, 7-trimethylxanthine) in foods: A comprehensive review on consumption, functionality, safety, and regulatory matters. *Journal of Food Science* **75**(3): R77-R87.

Hille, R. (2006). Structure and function of xanthine oxidoreductase. *European Journal of Inorganic Chemistry* **2006**: 1913–1962.

Hicks, K. A., O'Leary, S. E., Begley, T. P., & Ealick, S. E. (2013). Structural and Mechanistic Studies of HpxO, a Novel Flavin Adenine Dinucelotide-Dependent Urate Oxidase from *Klebsiella pneumoniae*. *Biochemisty* **52**: 477-487.

Hollingsworth, R. G., Armstrong, J. W., & Campbell, E. (2002). Pest control: Caffeine as a repellent for slugs and snails. *Nature* **417**(6892): 915-916.

Johannes, L. (2010). Can a caffeine-packed plant give a boost? *The Wall Street Journal*. U.S. edition (March 2, 2010): p D3.

Juliano, L. M., & Griffiths, R. R. (2005). Caffeine. In *Substance Abuse: A Comprehensive Textbook*, 4th ed., Edited by J. H. Lowinson, P. Ruiz & R. B. Millman. Lippincott Williams & Wilkins, Philadelphia, PA.

Kahn, K., & Tipton, P. A. (1997). Kinetic mechanism and cofactor content of soybean root nodule urate oxidase. *Biochemistry* **36**: 4731– 4738.

Kahn, K., & Tipton, P. A. (1998). Spectroscopic characterization of intermediates in the urate oxidase reaction. *Biochemistry* **37**: 11651–11659.

Kahn, K., Serfozo, P., & Tipton, P. A. (1997). Identification of the true product of urate oxidase reaction. *Journal of American Chemical Society* **119**: 5435–5442.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., & Sasamoto, S. (1996). Sequence analysis of the genome of the unicellular *Cyanobacterium synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Research* **3**(3): 109-136.

Kim, Y.S., Uefuji, H., Ogita, S., & Sano, H. (2006). Transgenic tobacco plants producing caffeine: a potential new strategy for insect pest control. *Transgenic Research*. **15**: 667-672.

Kleywegt, G. J., Jones, T. A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallographica*. *D* **50**: 178–185.

Koshiishi, C., Kato, A., Yama, S., Crozier, A., & Ashihara, H. (2001). A new caffeine biosynthetic pathway in tea leaves: utilization of adenosine released from the *S*-adenosyl-L-methionine cycle. *FEBS letters* **499**: 50-54.

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*. **157**: 105–132.

Larkin, M. A., et al. (2007). ClustalW and ClustalX version 2. *Bioinformatics* **23**:2947–8.

Lee, C. (2000). Antioxidant ability of caffeine and its metabolites based on the study of oxygen radical absorbing capacity and inhibition of LDL peroxidation. *Clinica Chimica Acta* **295**: 141-154.

Lelo, A., Miners, J. O., Robson, R. A., & Birkett, D. J. (1986). Quantitative assessment of caffeine partial clearances in man. *British Journal of Clinical Pharmacology* **22**: 183-86.

Libragen (2011). Sveltam: Green Innovation for Organic Slimming Cosmetics. p. 6. http://www.libragen.com/files/Leaflet-Sveltam.pdf, Accessed 23/06/2011.

Louie, M. T. M., Kale, Y., Yu, C. L., & Subramanian, M. (2009). Cloning and sequence analysis of a gene cluster in *Pseudomonas* sp. CBB1 involved in caffeine catabolism. *General meeting of American Society for Microbiology.* poster number Q-271.

Luisier, N., Ruggi, A., Steinmann, S. N., Favre, L., Gaeng, N., Corminboeuf, C., & Severin, K. (2012) A ratiometric fluorescence sensor for caffeine. *Organic and Biomolecular Chemistry.* **10**(37): 7441-7636.

Madyastha, K. M., & Sridhar, G. R. (1998). A novel pathway for the metabolism of caffeine by a mixed culture consortium. *Biochemical and Biophysical Research Communications* **249**: 178-181.

Madyastha, K. M., & Sridhar, G. R. (1999). Highly efficient C-8 oxidation of substituted xanthines with substitution at the 1-, 3-, and 7-positions using biocatalysts. *Journal of the Chemical Society, Perkin Transactions 1* **1999**(6): 677-680.

Madyastha, K. M., Sridhar, G. R., Vadiraja, B. B., & Madhavi, Y. S. (1999). Purification and partial characterization of caffeine oxidase- A novel enzyme from a mixed culture consortium. *Biochemical and Biophysical Research Communications* **263**: 460-464.

Marchler-Bauer, A., Anderson, J. B., Cherukuri, P. F., DeWeese-Scott, C., Geer, L. Y., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., & Ke, Z. (2005). CDD: A conserved domain database for protein classification. *Nucleic Acids Research* **33**(Suppl 1): D192-D196.

Marin, N. (2011) Caffeine: The Side Effects and Possible Effects It Could Be Having on Your Body. Webster's Digital Services, USA. ISBN: 1270790544

Mazzafera, P. (2002). Degradation of caffeine by microorganisms and potential use of decaffeinated coffee husk and pulp in animal feeding. *Scientia Agricola* **59**(4): 815-821.

Mazzafera, P. (2004). Catabolism of caffeine in plants and microorganisms. *Frontiers in Bioscience* **9**: 1348-1359.

Mazzafera, P., Olsson, O., & Sandberg, G. (1996). Degradation of caffeine and related methylxanthines by *Serratia marcescens* isolated from soil under coffee cultivation. *Microbial Ecology* **31**: 199-207.

McCusker, R. R., Goldberger, B. A., & Cone, E. J. (2006). Caffeine content of energy drinks, carbonated sodas, and other beverages. *Journal of Analytical Toxicology* **30**: 112-114.

McGee, M. B. (1980) Caffeine poisoning in a 19-year-old female. *Journal of Forensic Science*. **25**(1): 29-32.

Megerle, U., Wenninger, M., Kutta, R., Lechner, R., König, B., Dick, B., & Riedle, E. (2011) Unraveling the flavin-catalyzed photooxidation of benzylic alcohol with transient absorption spectroscopy from sub-pico- to microseconds. *Physical Chemistry Chemical Physics*. **13**: 8869-8880.

Mohanty, S. K., Yu, C. L., Das, S., Louie, T. M., Gakhar, L., & Subramanian, M. (2012) Delineation of the caffeine C-8 oxidation pathway in *Pseudomonas* sp. strain CBB1 via characterization of a new trimethyl acid monooxygenase and genes involved in trimethyluric acid metabolism. *Journal of Bacteriology*. **194** (15): 3872-3882.

Mohapatra, B. R., Harris, N., Nordin, R., & Mazumder, A. (2006). Purification and characterization of a novel caffeine oxidase from *Alcaligenes* species. *Journal of Biotechnology* **125**: 319-327.

Mrvos, R. M., Reilly, P. E., Dean, B. S., & Krenzelok E. P. (1989) Massive caffeine ingestion resulting in death**.** *Veterinary and Human Toxicology*. **31**(6): 571-572.

Nakamura, S., Nakamura, T., & Ogura, Y. (1963) Absorption spectrum of flavin mononucleotide semiquinone. *The Journal of Biochemistry*. **53**(2): 143-146.

Nathanson, J. A. (1984). Caffeine and related methylxanthines: Possible naturally occurring pesticides. *Science* **226**(4671): 184-187.

O'Leary, S.E., Hicks, K.A., Ealick, S.E., & Begley, T.P. (2009). Biochemical characterization of the HpxO enzyme from *Klebsiella pneumonia*, a novel FAD-dependent urate oxidase. *Biochemistry* **48**: 3033-3035.

Ogunseitan, O. A. (1996). Removal of caffeine in sewage by *Pseudomonas putida*: Implications for water pollution index. *World Journal of Microbiology and Biotechnology* **12**: 251-256.

Parliment, T. H., Ho, C. T., & Schieberle, P. (2000) Caffeinated Beverages: Health Benefits, Physiological Effects, and Chemistry. American Chemical Society symposium series ISSN 0097-6156; 754.

pET systems manual. (2005). 11^(th) Edition, Novagen, EMD Biosciences, Germany.

Petermann, J. B., & Baumann, T. W. (1983). Metabolic relations between methylxanthines and methyluric acids in *Coffea* L. *Plant Physiology* **73**: 961-964.

Pope, S. D., Chen, L., & Stewart, V. (2009). Purine utilization by *Klebsiella oxytoca* M5al: genes for ring-oxidizing and ring opening enzymes. *Journal of Bacteriology* **191**(3): 1006-1017.

Quandt, E. M., Hammerling, M. J., Summers, R. M., Otoupal, P. B., Slater, B., Alnahhas, R. N., Dasgupta, A., Bachman, J. L., Subramanian, M. V., & Barrick, J. E. (2013). Decaffeination and measurement of caffeine content by addicted *Escherichia coli* with a refactored N-demethylation operon from Pseudomonas putida CBB5. ACS Synthetic Biology 2(6): 301-307.

Ramazzina, I., Cendron, L., Folli, C., Berni, R., Monteverdi. D., Zanotti, G., & Percudani, R. (2008). Logical identification of an allantoinase analog (*puuE*) recruited from polysaccharide deacetylases. *The Journal of Biological Chemistry* **283**: 23295–23304.

Ramazzina, I., Folli, C., Secchi, A., Berni, R., & Percudani, R. (2006). Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nature Chemical Biology* **2**: 144-148.

Reissig, C. J., Strain, E. C., & Griffiths, R. R. (2009). Caffeinated energy drinks- A growing problem. *Drug and Alcohol Dependence* **99**: 1-10.

Rochat, S., Steinmann, S. N., Corminboeuf, C., & Severin, K. (2011) Fluorescence sensing of caffeine in water with polysulfonated pyrenes. *Chemical Communications*. 47: 10584-10586.

Rossignol, A. M., & Bonnlander, H. (1990) Caffeine- containing beverages, total fluid consumption, and premenstrual Syndrome. American Journal of Public Health **80**(9): 1106-1110.

Roussos, S., de los Angeles Aquiáhuatl, M., del Refugio Trejo-Hernández, M., Gaime Perraud, I., Favela, E., Ramakrishna, M., Raimbault, M., & Viniegra- González, G. (1995) Biotechnological management of coffee pulp-isolation, screening, characterization, selection of caffeine- degrading fungi and natural microflora present in coffee pulp and husk. **42**(5): 756-762.

Sambrook J., Fritsch E. F., & Maniatis T. (1989). Molecular cloning: a laboratory manual, 3rd ed. Cold Springs Harbor Laboratory, Cold Spring Harbor, N.Y.

Sankar, J., Lodha, R., & Kabra, S. K. (2008). Doxofylline: The next generation methylxanthine. *Indian Journal of Pediatrics* **75**(3): 251-254.

Sarath Babu, V. R., Patra, S., Karanth, N. G., Kumar, M. A., & Thakur, M. S. (2007) *Analytica Chimica Acta*. **582**: 329-334.

Sauer, M., Kappeli, O., & Fiechter, A. (1982). Comparison of the cytochrome P450-containing monooxygenases originating from two different yeasts. In *Cytochrome P-450,*

Biochemistry, Biophysics, and Environmental Implications pp. 453-457. Edited by E. Hietanen, M. Laitinen & O. Hanninen. Amsterdam.

Scherl, M., Scherl, J., & Fiedler, S. B. (1997) Detection of caffeine in beverages. U.S. patent No.: 5,610,072 A.

Schmidt, A. & Karlson-Stiber, C. (2008) Caffeine poisoning and lactate rise: an over looked toxic effect? *Acta Anaesthesiologica Scandinavica*. 52: 1012-1014.

Schultz, A. C., Nygaard, P., & Saxild, H. H. (2001). Functional analysis of 14 genes that constitute the purine catabolic pathway in *Bacillus subtilis* and evidence for a novel regulon controlled by the PucR transcription activator. *Journal of Bacteriology*. **183**: 3293–3302.

Schwimmer, S., & Kurtzman, R. H. (1971). Caffeine metabolism by *Penicillium roqueforti*. *Archives of Biochemistry and Biophysics* **147**(1): 109-113.

Scopes, R. K. (1994). Protein Purification: Principles and Practice. 3rd ed. Springer-Verlag New York.

Seiler, R. L., Zaugg, S. D., Thomas, J. M., & Howcroft, D. L. (1999). Caffeine and pharmaceuticals as indicators of waste water contamination in wells. *Ground Water* **37**(3): 405-410.

Shamim, M. T., Ukena, D., Padgett, W. L., & Daly, J. W. (1989). Effects of 8-phenyl and 8-cycloalkyl substituents on the activity of mono-, di, and trisubstituted alkylxanthines with substitution at the 1-, 3-, and 7-positions. *Journal of Medicinal Chemistry* **32**(6): 1231-1237.

Shen, H., Habara, M., & Toko, K. (2008) Development of caffeine detection using taste sensor with lipid/polymer membranes. *Sensors and Materials*. 20(4): 171-178.

Somogyi, L. P. (2010) Caffeine intake by the U.S. population. The Food and Drug Administration Oakridge National Laboratory, subcontract number: 70000073494 (http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofFoods/CFSAN/CFS ANFOIAElectronicReadingRoom/UCM333191.pdf). Accessed on 05/28/2013.

Spiller G. A. (2010) Caffeine. CRC press, USA. ISBN: 1420050133.

Stavric, B. (1988a). Methylxanthines: Toxicity to humans. 1. Theophylline. *Food and Chemical Toxicology* **26**(6): 541-565.

Stavric, B. (1988b). Methylxanthines: Toxicity to humans. 2. Caffeine. *Food and Chemical Toxicology* **26**(7): 645-662.

Stavric, B. (1988c). Methylxanthines: Toxicity to humans. 3. Theobromine, paraxanthine and the combined effects of methylxanthines. *Food and Chemical Toxicology* **26**(8): 725-733.

Strappaghetti, G., Corsano, S., Barbaro, R., Giannaccini, G., & Betti, L. (2001). Structure-activity relationships in a series of 8-substituted xanthines as A1-adenosine receptor antagonists. *Bioorganic & Medicinal Chemistry* **9**(3): 575-583.

Stryer L. (1995). Biochemistry, 4th ed. WH Freeman and Co., New York, NY.

Summers, R. M., Louie, T. M., Yu, C. L., & Subramanian, M. (2011). Characterization of a broad-specificity non-haem iron *N*-demethylase from *Pseudomonas putida* CBB5 capable of utilizing several purine alkaloids as sole carbon and nitrogen source. *Microbiology* **157**(2): 583-592.

Summers, R. M. (2011). "Metabolism, enzymology, and genetic characterization of caffeine degradation by pseudomonas putida CBB5." dissertation, University of Iowa. http://ir.uiowa.edu/etd/3389.

Summers, R. M.**,** Louie, T. M., Yu, C. L., Gakhar, L., Louie, K. C., & Subramanian, M. (2012). Novel, highly specific *N*-demethylases enable bacteria to live on caffeine and related purine alkaloids. *Journal of Bacteriology*. **194**: 2041–2049.

Suzuki, T., & Waller, G. R. (1987). Allelopathy due to purine alkaloids in tea seeds during germination. *Plant and Soil* **98**: 131-136.

Tassaneeyakul, W., Birkett, D. J., McManus, M. E., Tassaneeyakul, W., Veronese, M. E., Andersson, T., Tukey, R. H., & Miners, J. O. (1994). Caffeine metabolism by human hepatic cytochromes P450: Contributions of 1A2, 2E1 and 3A isoforms. *Biochemical Pharmacology* **47**(10): 1767-1776.

Tomson, B. (2013). Caffeinated snacks draw FDA scrutiny. *The Wall Street Journal*. U.S. edition (May 1, 2013): p B8.

Trier, K., Olsen, E. B., Kobayashi, T., Ribel-Madsen, S. M. (1999). Biochemical and ultrastructural changes in rabbit sclera after treatment with 7-methylxanthine, theobromine, acetazolamide, or L-ornithine. *British Journal of Ophthalmology* **83**(12): 1370-1375.

Tyrala, E. E., & Dodson, W. E. (1979) Caffeine secretion into breast milk. *Archives of Disease in Childhood.* **54**: 787-800.

U.S. Food and Drug Administration (2011) Qualitest Pharmaceuticals Issues, Nationwide Retail Level Recall of Four Lots of Butalbital, Acetaminophen, and Caffeine Tablets, USP 50mg/325mg/40mg and Four Lots of Hydrocodone Bitartrate and Acetaminophen Tablets, USP 7.5mg/500mg. (*http://www.fda.gov/safety/recalls/ucm260837.htm*) Accessed on 05/29/2013.

U.S. Food and Drug Administration. Medicines in my home: caffeine and your body. (*http://www.fda.gov/downloads/Drugs/ResourcesForYou/Consumers/BuyingUsingMedicineSafely/UnderstandingOver-the-CounterMedicines/UCM205286.pdf*). Accessed on 05/28/2013

Udayasankar, K., Manohar, B., & Chokkalingam, A. (1986). A note on supercritical carbon dioxide decaffeination of coffee. *Journal of Food Science and Technology* **23**(6): 326–328.

van Dam, R.M. (2008). Coffee consumption and risk of type 2 diabetes, cardiovascular diseases, and cancer. *Applied Physiology, Nutrition, and Metabolism* **33**: 1269-1283.

Vogels, G. D., & C. Van der Drift. (1976). Degradation of purines and pyrimidines by microorganisms. *Bacteriology Reviews* **40**: 403–468.

Wang, Y., Chackalamannil, S., Hu, Z., Boyle, C. D., Lankin, C. M., Xia, Y., Xu, R., Asberom, T., Pissarnitski, D., & Stamford, A. W. (2002). Design and synthesis of xanthine analogues as potent and selective PDE5 inhibitors. *Bioorganic & Medicinal Chemistry Letters* **12**(21): 3149-3152.

Wang, Y., Yang, X., Zheng, X., Li, J., Ye, C., & Song, X. (2010). Theacrine, a purine alkaloid with anti-inflammatory and analgesic activities. *Fitoterapia* **81**: 627-631.

Ward, A., & Clissold, S. P. (1987). Pentoxifylline. A review of its pharmacodynamic and pharmacokinetic properties, and its therapeutic efficacy. *Drugs* **34**(1): 50-97.

Weinberg, B. A., & Bealer, B. K. (2001). The world of caffeine: the science and culture of the world's most popular drug. Routledge Publication, New York ISBN 415-92722-6.

Werner A. K. & Witte, C. P. (2011). The biochemistry of nitrogen mobilization: purine ring catabolism. *Trends in Plant Science*. **16**: 381–387.

Werner, A.K., Romeis, T., & Witte, C. (2009). Ureide catabolism in *Arabidopsis thaliana* and *Escherichia coli*. *Nature Chemical Biology* **6**: 19-21.

White, P. A., & Rasmussen, J. B. (1998). The genotoxic hazards of domestic wastes in surface waters. *Mutation Research* **410**: 223-236.

Woolfolk, C. A. (1975). Metabolism of *N*-methylpurines by a *Pseudomonas putida* strain isolated by enrichment on caffeine as the sole source of carbon and nitrogen. *Journal of Bacteriology* **123**(3): 1088-1106.

Yu, C. L., Kale, Y., Gopishetty, S., Louie, T. M., & Subramanian, M. (2008). A novel caffeine dehydrogenase in *Pseudomonas* sp. strain CBB1 oxidizes caffeine to trimethyluric acid. *Journal of Bacteriology* **190**(2): 772-776.

Yu, C. L., Louie, T. M., Summers, R., Kale, Y., Gopishetty, S., & Subramanian, M. (2009). Two distinct pathways for metabolism of theophylline and caffeine are coexpressed in *Pseudomonas putida* CBB5. *Journal of Bacteriology* **191**(14): 4624-4632.

Zavialov, I. A., Dahanukar, V. H., Nguyen, H., Orr, C., & Andrews, D. R. (2004). New and practical method for synthesis of 1-and 1, 3-substituted xanthines. *Organic Letters* **6**(13): 2237-2240.