DEVELOPMENT OF EFFICIENT *E. coli* HOST STRAIN FOR EXPRESSION OF FUNCTIONALLY ACTIVE HUMAN CYTOCHROME P450 2A6 BY ENDOGENOUS INCREASED BIOSYNTHESIS OF 5-AMINOLEVULINIC ACID (ALA)

M. Sc. Thesis
2012

Submitted to
CENTRAL DEPARTMENT OF BIOTECHNOLOGY
Tribhuvan University
Kirtipur, Kathmandu, Nepal

Sujina Mali
Roll No. : BT 021/066
T.U. Regd. No. : 5-2-33-632-2005

Development of efficient *E. coli* host strain for expression of functionally active human cytochrome P450 2A6 by endogenous increased biosynthesis of 5-Aminolevulinic acid (ALA)
Sujina Mali

DEVELOPMENT OF EFFICIENT *E. coli* HOST STRAIN FOR EXPRESSION OF FUNCTIONALLY ACTIVE HUMAN CYTOCHROME P450 2A6 BY ENDOGENOUS INCREASED BIOSYNTHESIS OF 5-AMINOLEVULINIC ACID (ALA)

M.Sc. Thesis
2012

Submitted to
Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

Sujina Mali
Roll No. : BT 021/066
T.U. Regd. No. : 5-2-33-632-2005

Supervisors

Dr. Tribikram Bhattarai
Head of Department
Central Department of
Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

Dr. Pramod Aryal
Chief Scientific Officer
Supra Inventica Pvt. Ltd.
Presently, Director
Capsule Inc. Pvt. Ltd.
Kathmandu, Nepal

Central Department of Biotechnology
Tribhuvan University

Kirtipur, Kathmandu, Nepal
Tel. No. : +977 1 4336221

Date: - 10th June, 2012

Certificate of Evaluation

This is to certify that this thesis entitled **"Development of efficient *E. coli* host strain for expression of functionally active human cytochrome P450 2A6 by endogenous increased biosynthesis of 5-Aminolevulinic acid (ALA)"** presented to evaluation committee by Ms. **Sujina Mali** is found satisfactory for the partial fulfillment of Master of Science in Biotechnology.

……………………………………………
Dr. Tribikram Bhattarai
Supervisor
Head of Department
Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

……………………………………………
Dr. Jyoti Maharjan
External Examiner
Senior Scientific Officer
Nepal Academy of Science and Technology (NAST)
Khumaltar, Lalitpur, Nepal

……………………………………………
Dr. Pramod Aryal
Supervisor
Chief Scientific Officer
Supra Inventica Pvt. Ltd
Presently, Director, Capsule Inc. Pvt. Ltd
Kathmandu

……………………………………………
Dr. Deepak Pant
Internal Examiner
Lecturer
Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

Acknowledgement

I would like to express my profound gratitude to my supervisor **Dr. Tribikram Bhattarai**, Head of Department, Central Department of Biotechnology, Tribhuvan University, for his precious supervision, constant support, and valuable guidance that was pivotal for the completion of my thesis works.

I am eternally grateful to my supervisor **Dr. Pramod Aryal**, Chief Scientific Officer, Supra Inventica, Pvt. Ltd., who had been the most important person during the entire period of this study. He had made it all possible from "culture to construct". I will always remain indebted for his unquantifiable help and support. He is a fabulous supervisor and it was an honor to work with him.

I would like to express my deepest gratitude to Dr. F.P. Guengerich, Professor and Interim Chair, Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, TN, USA for providing bacterial strains and plasmid. I must thank Dr. Elizabeth Gillam, Professor, School of Chemistry and Molecular Biosciences, The University of Queensland, Australia for providing pCW 2A6: hNPR plasmid sequence information and her papers.

I owe special thanks to Research and Development officers of Supra Inventica Pvt. Ltd., namely Rumika Maharjan, Suman Lama, Samidha Shah and Ayesha Dhungel for their unconditional help and endless support throughout this study period. I would like to additionally thank Rumika Maharjan, Ayesha Dhungel and Damodar Parajuli for being part in preparing grant as co-investigators that signified importance of the present research nationally, too. I would like to thank Ministry of Science and Technology (MoST), Government of Nepal for partially funding this study and promoting the innovative scientific works.

I would like to extend my huge gratitude to Mr. Hari Bhakta Sharma (Executive Director) and all the members of Deurali-Janta Pharmaceuticals Pvt. Ltd. for providing fundamental financial support to carry out my research study at Supra Inventica Pvt. Ltd. Similarly, I would like to acknowledge Dr. Deepak Khadka, Research and Development-Chief Scientific Officer, Deurali-Janta Pharmaceuticals Pvt. Ltd., who had helped to improve my laboratory skills.

I am indebted to Damodar Parajuli, Kalpana Bakey, Keshav Raj Budda and Sushil Khanal for their incredible support and cooperation during entire research period without them this work would not have been possible. My special thanks go to Dev Mahato for his technical assistance and constant support. I would like to thank all the interns, students

and entire team of Supra Inventica Pvt. Ltd. for creating such a warm and helpful environment during the research period.

I am deeply indebted to Dr. Lakshmaiah Sreerama, Professor and Chair, Department of Chemistry, Fulbright Schlolar-2010, St. Cloud State University, MN, USA for his support, care and guidance. Similarly, I thank Dr. Huanting Liu, Academic, School of Chemistry, University of St Andrews, UK for sharing his expertise for successful site directed mutagenesis PCR in this study. I gratefully acknowledge Dr. Thomas Elliott, Professor, Department of Microbiology, Immunology and Cell Biology, School of Medicine, West Virginia University, WV, USA for his generous response and help to create *Nde*I deleted pUC19 plasmid by Fill-in method.

It is my immense pleasure to acknowledge Professor Dr. Mohan Kharel, Professor Dr. Ganga Prasad Kharel, Dr. Sampoornanand Jha, Dr. Rajani Malla, Dr. Krishna Das Manandhar, Dr. Deepak Pant, Dr. Giri Raj Tripathi, Mrs. Jarina Joshi, Mrs. Smita Shrestha, Mr. Bal Hari Poudel, Late. Mr. Janardan Pandey, Mr. Shirsir Sharma, Mr. Rajendra Bahadur Shahi and all the faculty members and staffs of Central Department of Biotechnology for their love, consistent support and encouragement. I particularly thank Mohan Shrestha, Laboratory assistant, who had generously helped me whenever needed.

I owe special thanks and regards to my dear friends Jeni, Dinesh, Jagat, Govind, Om, Ram, Dinita, Bal Krishna, Bishnu, Padam, Bhim, Roshan, Krishna, Megh, Raju, Buddha and Saroj. They have been helpful, supportive and inspiring.

I am grateful to my loving parents Sudarshan Mali and Ganga Laxmi Mali who always tried to make me feel a lucky and privileged person and without their unconditional love and support this study would not have been possible. Thank you for always being there for me. I would also like to thank my dear brother Sagar Mali for his valuable contribution and assistance to prepare the final manuscript. Finally, I need to thank all the people who had directly or indirectly helped me during the entire study period.

Glossary Acronyms

µ	micro
ALA	5-Aminolevulinic acid
ATP	Adenosine Tri Phosphate
BLAST	Basic Local Alignment Search Tool
bp	base pairs
cat	chloramphenicol acetyltransferase gene
CYP	Cytochrome P450
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
dsDNA	double stranded DNA
DTT	Dithiothreitol
EDTA	Ethylene Diamine Tetra Acetic Acid
GLB	Gel Loading Buffer
GluTR	Glutamyl tRNA Reductase
hNPR	human NADPH-P450 Reductase
IDT	Integrated DNA Technologies
IPTG	Isopropyl Thiogalactopyranoside
kb	kilo base pairs
L	Liter
LB	Luria Bertani
m	milli
M	Molar
min	minute
NADPH	Nicotinamide Adenosine Diphosphate (H)

NCBI	National Center for Biotechnology Information
NEB	New England’s Biolabs
OD_{600}	Optical Density at 600 nm
orf	open reading frame
ori	origin of replication
PALP	Pyridoxal Phosphate
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
rbs	ribosome binding site
RNase	Ribonuclease
rpm	revolutions per minute
RT	Room Temperature
SOC	Super Optimal Broth With Catabolite
TB	Terrific Broth medium
T_m	Melting Temperature
Tris	Tris-(hydroxymethyl)-aminmethane
U	Units
UV	Ultraviolet
v/v	volume per volume
w/v	weight per volume

CONTENTS

Chapters **Page no.**

List of Figures and Tables

Abstract

Cytochrome P450s are large family of heme-containing monooxygenases catalyzing NADPH dependent hydroxylation and one of the most known versatile biocatalysts having central importance in pharmaceutical industries for toxicological, mutagenesis and drug-drug interactions studies. Since cytochrome P450s are commonly found in liver microsomes, robust heterologous P450 expression system such as *E. coli* has been practiced. However, expression of active P450 (holo-form) is compromised mainly due to limited availability of free heme since heme biosynthetic pathway is tightly regulated and intrinsic heme proteins have higher affinity for intracellular heme. Moreover, *E. coli* does not have effective heme transport mechanism and expressions of active P450s have been carried with exogenous supplementation of expensive ALA, a rate limiting precursors in heme biosynthesis. However, higher concentration of ALA is toxic and mutagenic in presence of $Fe^{2+/3+}$ ions making its use limiting. In the present study attempt has been made in increasing endogenous pool of ALA for facilitating free heme availability in supporting ample P450 expression by cloning and expression of mutated *hem*A gene, the rate limiting enzyme in ALA biosynthesis, as the mutant HemA protein is insensitive to heme mediated proteolytic cleavage. Plasmid pACYC184 (cloning vector) has been modified by site directed mutagenesis PCR to create *Nde*I site (CATATG) overlapping *cat* ATG initiation codon in developing expression vector pAM101. This vector was used for sub-cloning of stabilized *hem*A gene under regulation of *cat* promoter giving plasmid pAM102 since chloramphenicol supports heterologous expression of proteins. Chloramphenicol resistant *E. coli* HME5 *tna<>cat* harboring bicistronic plasmid carrying L240C/N297Q double mutant P450 2A6 and human NADPH-P450 reductase in pCW” was transformed with pAM102. The bacteria thus developed, *E. coli* AM102/2A6, was able to support expression of catalytically active P450 without exogenous supplementation of ALA because functionally active P450 2A6 L240C/N297Q metabolizes indole into non-toxic indoxyl products and relatively high survival rate of this host strain at toxic level of indole (5mM) clearly suggested expression of functional cytochrome P450 2A6 double mutant. This indicated that cloned mutated *hem*A gene co-expressed stabilized HemA protein and increased ALA biosynthesis thus supporting augmented cellular heme concentration in facilitating expression of functionally active cytochrome P450 2A6. This system could be further enhanced by stable integration in chromosome by reamplifying the construct designed to be cloned in pUC*ΔNde*I developed in the present study. This construct can also swap engineered protein. This simple and effective strategy can be utilized in expression of P450s and hemeproteins.

Keywords: P450s; heme pathway; ALA; hemeproteins; glutamyl tRNA reductase; *hem*A; chromosomal engineering

CHAPTER ONE

INTRODUCTION

1.1 Background

Cytochrome P450s (CYPs) are large family of heme-containing monooxygenases found in all domains of life ranging from lower archaebacteria to mammals (Danielson PB, 2002). Mostly, P450s catalyze NADPH dependent hydroxylation of substrates and is known as nature's most versatile biocatalyst due to its ability to catalyze diversified reactions utilizing innumerate substrates (reviewed in Gillam EMJ, 2008). Cytochrome P450s have central important role in toxicology to study drug metabolism and clearance of xenobiotics (Gillam EMJ, 2007; 2008). Human cytochrome P450s accounts for nearly 70-80% of all so called drug metabolizing enzyme phase I dependent metabolism of clinically relevant drugs (Hasler JA, 1999). In addition, cytochrome P450s play critical role in drug-drug interaction (Badyal and Dadhich, 2001) and metabolic profiling of drug candidates for bioactivation and generation of potentially mutagenic compounds (Aryal *et al.*, 1999; 2000; Oda *et al.*, 2001) which are investigated during drug development stages.

Human cytochrome P450s are mostly found in liver microsomes apart from other tissues (Gillam EMJ, 2007). The limited availability of human or animal tissue samples and also due to animal right issues to use these samples, study of structure-function relationships and functional characteristics of human cytochrome P450s are often compromised. But the vast applications of human cytochrome P450s in pharmaceutical and other industries for study of metabolic profiling, mutagenicity, toxicity studies of innumerate chemical compounds had necessitated the requirement for efficient expression of functionally active human cytochrome P450s in easy and renewal systems such as bacteria. Despite the fact that cytochrome P450s are also found in many bacterial systems, *E. coli* (Parikh *et al.,* 1997; Harnastai *et al.*, 2006) and *Salmonella typhimurium* (Aryal *et al.,* 1999) lacks its own P450s which makes it suitable host for expression of functional cytochrome P450s for enzyme based assays in absence of potential noise from native background of the host.

In *E. coli* and *S. typhimurium*, the expression of membrane bound active eukaryotic P450s are compromised due to several factors. The presence of N-terminus hydrophobic residues forming the membrane-spanning segments of cytochrome P450s, the proline rich clusters in the hinge between the membrane anchoring and the catalytically active domain of the enzyme, requirement of intact heme group in the structure for proper enzyme folding, and the pivotal auxiliary NADPH reductase binding domain for efficient electron transfer have been limiting factors for heterologous expression and functional assay (Harnastai *et al.*, 2006 and references there in). To overcome these limitations, common strategies have been employed to alleviate expression of functional cytochrome P450s in *E. coli* which include N-terminal sequence modification (Fisher *et al.*, 1992; Pernecky *et al.*, 1993) and use of lowered temperature (20-28^{0}C) during cytochrome P450s expression (Barnes *et al.*, 1991; Fisher *et al.*, 1992) for proper membrane folding, and co-expression of human NADPH reductase (Fisher *et al.*, 1992; Parikh *et al.,* 1997; Kranendonk *et al.,* 1999; Aryal *et al.,* 2000) for efficient electron transfer.

However, in these *E. coli* and *S. typhimurium* based expression systems, the heterologous expression of functionally active cytochrome P450s are still compromised due to limited availability of free heme that can be incorporated for functional protein (Harnastai *et al.*, 2006). In *E. coli* and *S. typhimurium*, heme biosynthetic pathway is under strict regulatory control as free heme and its intermediate metabolites are known to be potentially toxic to cells in higher concentration (Nakahigashi *et al.*, 1991). In addition, under normal physiological conditions, the newly synthesized heterologous hemeproteins having lower affinity for heme than that of intrinsic cellular proteins such as cytochrome C (Springer and Sligar, 1987) are often only expressed as apo-proteins (inactive) due to competitiveness with available free heme in the cytoplasm (Chudaev and Usanov, 1997) thus requiring higher amount of heme in the cell for expression of functional cytochrome P450.

Enterobacteriaceae including *E. coli* K12 derivative or *S. typhimurium* L2 derivative strains are commonly employed for heterologous expression of cytochrome P450s (Fisher *et al.*, 1992; Parikh *et al.,* 1997; Kranendonk *et al.,* 1999) and mutagenicity assays (Josephy *et al.,* 1998; Aryal *et al.,* 1999, Kranendonk *et al.,* 1999; Aryal *et al.,* 2000, Oda *et al.,* 2001) but they are impermeable to free heme (Jacobs *et al.*, 1972) that limits use of heme directly but exogenous supplementation of ALA-a rate limiting precursor for heme (Kang *et al.*, 2011), has been standardized as an integral medium constituent for heterologous expression (Guengerich *et al.,* 1996; Gillam *et al.,* 1999; 2000; Nakamura *et al.,* 2001; Kim *et al.,* 2002; Guengerich *et al.,* 2004; Kim *et al.,* 2005; Wu *et al.,* 2005) of properly folded active form of cytochrome P450s. However, in some cases exogenous ALA supplementation might not support high level expression of fully active cytochrome P450s if maintenance of constant cellular heme biosynthesis levels is hindered and does not coincide with cytochrome P450s expression levels (Harnastai *et al.*, 2006). This can potentially explain disparity in expression observed among different isoforms which make cellular metabolism study complicated due to differences in the ratio of NADPH-

P450 reductase and P450 (Parikh *et al.,* 1997; Oda *et al.,* 2001) that might not replicate microsomal ratio making it prudent to reconstitute the system for metabolic study (Guengerich *et al.,* 1996; Parikh *et al.,* 1997; Nakamura *et al.,* 2001) costing time and resources. In addition, supplementation of expensive ALA precursor may account up to 60% costs of P450s expression and ALA itself is known to be potentially toxic and mutagenic compound in the presence of Fe $^{2+/3+}$ ions (Onuki *et al.*, 2002) thus its supplementation at high concentration is not feasible to perform cytochrome P450s mediated mutagenicity assays as it might give false positive results.

Thus, endogenous supplementation of ALA has been envisaged for sufficient amount of heme biosynthesis for supporting heterologous expression of hemeproteins. To this effect, cloning and expression of *hem*A gene has been in forefront to provide cellular ALA for subsequent heme biosynthesis (Harnastai *et al.,* 2006; Kang *et al.*, 2011). This gene is known to encode the rate limiting enzyme glutamyl tRNA reductase in ALA biosynthesis (Chen *et al.,* 1994) through C5 pathway (Philipp-Dornston and Doss, 1973) that is found in *E. coli*. However, *hem*A gene is known to be transcriptionally regulated (Nakayashiki *et al.*, 1995; Choi *et al.,* 1996; Schobert and Jahn, 2002) and Glutamyl tRNA reductase (GluTR or HemA) protein is also under feedback inhibition by heme (Verderber *et al.*, 1997). HemA protein is mainly regulated by conditional protein stability in response to cellular heme concentration. Two ATP dependent Lon/ClpAP proteases mediate heme regulated proteolytic degradation of HemA protein thus this protein is more stable and abundant in heme limited cells while its stability significantly reduces in normally growing cells (Wang *et al.*, 1999a).

Overcoming these regulatory mechanisms are critical for endogenous supplementation of ALA for augmented heme biosynthesis to supplement for heterologous expression of non-intrinsic hemeproteins. Moreover, heme regulated proteolysis of HemA protein can be overcome by introducing two lysine residues close to N terminus of the protein (change from ^{1}MTLLALGI^{8} to ^{1}MT$\underline{\text{KK}}$LLALGI^{10}), denoted as HemA(KK), which stabilizes HemA protein (Wang *et al.*, 1999b). This mutant *hem*A gene has been co-expressed for expression of stabilized HemA(KK) protein and it was reported that exogenous supplementation of ALA was not required for high-level expressions of functionally active four human microsomal cytochrome P450s (CYP1A1, CYP1A2, CYP21 and CYP17s) and a mitochondrial CYP11A1 (Harnastai *et al.*, 2006). This indicated that this mutant could sufficiently catalyze ALA biosynthesis which then subsequently acts as the precursor for heme biosynthesis to be integrated in heterologously expressed non-intrinsic hemeproteins.

1.2 Current Studies

In the present study, co-expression of *E. coli hem*A(KK) gene has been envisaged for endogenous supplementation of ALA for functional expression of human cytochrome P450 2A6 double mutant L240C/N297Q. This P450 2A6 double mutant carries the gain of function mutation for indole metabolism (Gillam *et al.,* 2000) and is cloned in plasmid pCW"2A6:hNPR (Nakamura *et al.*, 2001). The cytochrome P450 2A6 isoform has been cloned with its cognate redox partner human NADPH-P450 reductase (hNPR) cDNA as the bicistronic construct under double *tac* promoter (Gillam *et al.*, 1999) in vector pCW"2A6:hNPR (Parikh *et al.*, 1997).

It has been validated that the catalytically active P450 2A6 metabolizes indole (Gillam *et al.*, 2000), a toxic substrate to *E. coli* cells at concentration of 5mM, to non-toxic indoxyl derivatives (Nakamura *et al.,* 2001). Normally, in *E. coli*, indole is produced during tryptophan degradation due to action of tryptophanase (*tna*) (Burns *et al.,* 1962). However, in the homologous recombination efficient host strain, *E. coli* HME5 (Yu *et al.,* 2000) *tna* gene has been replaced with *cat* gene, *E. coli* HME5 *tna<>cat* (Wu *et al.*, 2005) that does not produce indole endogenously, making it suitable host for screening of functional P450 2A6 activity without noise and can support maintenance of constant indole levels for screening. Furthermore, swapping of *cat* gene renders the strain resistant to chloramphenicol which is also known to facilitate heterologous protein expression through induction of cold shock response (Kusano *et al.*, 1999; Kagawa *et al.*, 2003).

In this study, pAM101 expression vector has been constructed by site directed mutagenesis of pACYC184 plasmid to create *Nde*I site overlapping *cat* ATG initiation codon. The mutant *hem*A gene was then cloned in pAM101 at place of *cat* gene to be expressed under *cat* promoter giving plasmid pAM102. The cloned gene can be co-expressed along with cytochrome P450, hNPR due to compatibility of the plasmids, pCW"2A6 L240/N297Q and pAM102, for effective screening of indole metabolism by *E. coli* HME5 *tna<>cat* strain harboring both these plasmids.

Compared to *E. coli* HME5 *tna<>cat* strain harboring pCW"2A6:hNPR and pAM101 plasmid (without cloned *hem*A gene; *E. coli* AM101), the *E. coli* HME5 *tna<>cat* strain harboring pCW"2A6: hNPR and pAM102 plasmid (with clone *hem*A gene; *E. coli* AM102) had higher survival rate in toxic levels of indole (5mM). In addition, ALA supplementation in *E. coli* AM101 culture had higher survival rate than *E. coli* AM102 indicating that *E. coli* AM102 expressed HemA protein that supported on production of additional ALA that is possibly lower than exogenously supplemented 1 mM but sufficient to express functionally active cytochrome P450 2A6 L240C/N297Q and reductase that metabolized indole for decreasing toxicity. This indicated that the expression of functionally active cytochrome P450 2A6 L240C/N297Q was possible because cloned stabilized *hem*A gene expressed stable HemA protein that provided ALA endogenously required for heme biosynthesis which eventually could be integrated in the expressed cytochrome P450 2A6 L240C/N297Q for retaining its functional activity. However, the newly developed host strain *E. coli* AM102/2A6 is applicable for mutagenicity studies too where exogenous ALA supplementation is not suitable due to its potential toxicity at high concentration (Nakahigashi *et al.*, 1991).

Furthermore, pAM102 could be used for generation of random mutant library of HemA protein employing random mutagenesis (Zhao *et al.,* 1998; Nakamura *et al.,* 2001; Kim *et al.,* 2002, Miyazaki *et al.,* 2002; Kim *et al.,* 2005) that could be screened on the basis of indole metabolism as been done for cytochrome P450 2A6 (Nakamura *et al.,* 2001; Kim *et al.,* 2005). The mutant library upon being cloned in newly developed pUCΔ*Nde*I in which 5'- and 3'-homologous regions flanking *hem*A gene construct can be cloned and this region can be amplified by reamplification primer designed to generate linear cassette for homologous recombination for easy integration in host strain chromosomal *hem*A gene locus through recombination as described for development of the host strain (Wu *et al.,* 2005) since the parent strain *E. coli* HME5 strain is compatible for homologous recombination (Yu *et al.,* 2001). This would allow pAM101 to be used for cloning of *hem*L gene for co-expression (Kang *et al.,* 2011) since HemA and HemL protein are known to make complex (Moser *et al.,* 2001) for efficient transfer of unstable GSA generated by HemA to give ALA by the action of HemL (Luer *et al.,* 2005) and efficiently produce ALA (Luer *et al.,* 2005).

1.3 Objectives

1.3.1 General objective

- To develop the host strain for heterologous expression of functionally active human cytochrome P450 2A6 by endogenous supplementation of ALA through stabilized *hem*A gene co-expression

1.3.2 Specific objectives

- To construct pAM101 plasmid by creating *Nde*I site (CATATG) overlapping *cat* ATG initiation codon in pACYC184 vector by site directed mutagenesis PCR.
- To amplify *hem*A(KK) gene from genomic DNA of *E. coli* HME5 *tna*<> *cat* strain using specifically designed primers.

- To develop pAM102 by sub-cloning stabilized *hem*A(KK) gene under *cat* promoter in pAM101.
- To transform *E. coli* HME5 *tna*<> *cat* harboring pCW"2A6: hNPR with pAM102 or pAM101.
- To determine endogenous supplementation of ALA due to clone HemA(KK) for expression of catalytically active P450 2A6 (L240C/N297Q) that would metabolize indole to reduce the toxicity and support cell growth by comparing the growth rate ($0D_{600}$) in presence of toxic levels of indole (5mM).

1.4 Rationale and Scope

Exogenous ALA supplementation for cytochrome P450s expression is costly as ALA is highly expensive. To overcome the dependence of exogenous ALA supplementation a new *E. coli* host strain is developed, which would co-express mutant *hem*A gene under regulation of *cat* promoter from a compatible plasmid that can harbor along with another plasmid expressing human cytochrome P450 2A6 and hNPR in same bacterial strain. This mutant *hem*A gene coding for HemA(KK) protein is insensitive to feedback regulation by heme, because incorporation of two lysine residues (KK) between second and third amino acids in mutant protein is presumed to hinder essential conformational change that would be necessary for recognition by Lon and ClpAP energy dependent proteases for heme regulated proteolysis. Thus co-expression of stabilized HemA(KK) would endogenously supplement ALA and subsequently heme in supporting high level expression of catalytically active cytochrome P450 2A6 making its expression possible also in the absence of exogenous 5-ALA. This host strain can be used not only for expression of functionally active cytochrome P450 2A6 but also for efficient expression of different cytochrome P450s isoforms.

In addition, overproduction of ALA can support production of tetrapyrrole containing biological molecules such as vitamin B_{12}, heme-containing enzyme (siroheme, cytochrome), porphyrins and also ALA itself that has many therapeutic potential and used in tumor-localizing and photodynamic therapy for various cancers.

CHAPTER TWO

LITERATURE REVIEW

2.1 Cytochrome P450s

Cytochrome P450s (CYP or P450, E.C. 1.14.14.1) are large family of heme-containing monooxygenases found in all domains of life ranging from lower archaebacteria to mammals (Danielson PB, 2002). P450s have molecular masses of appoximately 50 kDa and are usually localized in membrane anchored by a hydrophobic N-terminus region (VonWachenfeldt and Johnson, 1995). Cytochrome P450s are named on the basis of spectral absorbance peak at around 450 nm (generally at 448 nm) due to reduced heme iron bound to carbon monoxide. Only functionally active and properly folded cytochrome P450s show this characterictic Soret peak whereas misfolded and heme unbound P450s absorbs maximally at 420 nm (Omura and Sato, 1964). The principal reaction catalyzed by most cytochrome P450s include monooxygeneses activity, transfering single oxygen atom into a substrate. This reaction can be generalized as:

$$RH + O_2 + NADPH + H^+ \rightarrow ROH + H_2O + NADP^+$$

In this reaction, RH represents a substrate (alkanes, fatty acids, drugs, steroids, carcinogens, etc.) which gets hydroxylated (ROH) in the presence of NADPH, to become pharmacologically active compound (Strobel *et al.,* 1995; Nakamura *et al.,* 2001). Since cytochrome P450s cannot directly receive the electron from NAD(P)H it requires flavoprotein, NADPH-P450 reductase, that transfers electron to cytochrome P450 but some of the isoforms receive electron from cytochrome b5, also. The electron generated from oxidation of NAD(P)H is transferred to FMN/FAD containing flavoprotein which then directly transfers the electron to cytochrome P450 or through cytochrome b5 (Schenkman and Jansson, 2003).

It has been now elucidated that individual cytochrome P450s are capable of catalyzing vast array of molecular reactions such as hydroxylation, dehalogenation, epoxidation, N, S- and O-dealkylation, N-oxidation, sulfoxidation and a number of other reactions (Nelson *et al.,* 1993). Thus, cytochrome P450s are amongst the most versatile biocatalyst found in nature (Gillam EMJ, 2007). They are primarily involved in oxidation of innumerate exogenous and endogenous chemicals (Fujita and Kamataki, 2002).

In general, cytochrome P450 mediated metabolism makes hydrophobic chemicals to water soluble forms in order to facilitate their clearance (Rosic *et al.*, 2009). Human cytochrome P450s are well recognized for their ability to detoxify xenobiotic compounds. However, some procarcinogenic and promutagenic compounds also undergo bioactivation by different isoforms of human cytochrome P450s producing carcinogenically active metabolites (Shimada and Nakamura, 1987; Aryal *et al.*, 1999; Aryal *et al.*, 2000; Oda *et al.*, 2001). Thus, cytochrome P450s are very important enzyme systems in pharmaceutical industries for drug development to study the potential mutagenic effect of the compound or its metabolites, drug-drug interaction of parent compound and its metabolites. Other major functions of cytochrome P450s include synthesis of growth regulators, sterols and molecules involved in cell signalling (Capdevila *et al.,* 2005; Harnastai *et al.,* 2006). In general use of cytochrome P450s can be summarized as illustrated in figure 2.1.

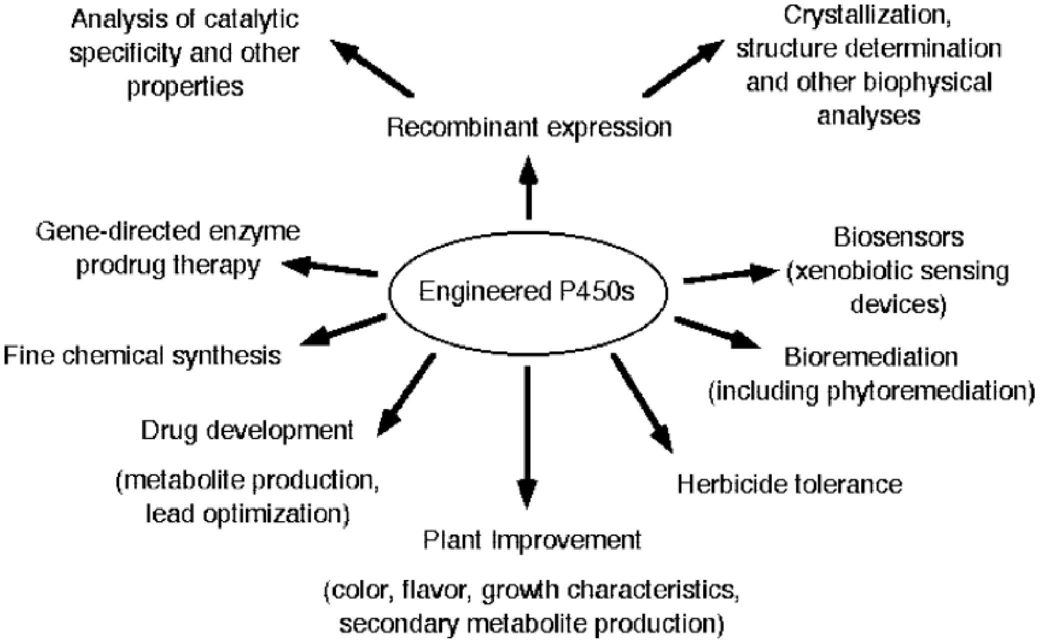


Figure 2.1. Schematic illustration of various applications of cytochrome P450s (Adopted from Gillam EMJ, 2008).

2.1.1 Role of Cytochrome P450s

2.1.1.1 Drug Metabolism

Cytochrome P450s are primarily involved in drug metabolism and accounts for nearly 70-80% of all so called drug metabolizing phase I enzyme dependent metabolism of clinically relevant drugs (Hasler JA, 1999). While some drugs become catalytically active only after biotransformation reaction catalyzed by cytochrome P450s, most of the drugs gets deactivated after cytochrome P450 mediated oxidation reaction which helps in clearance of drugs from body (Nebert and Russell, 2002; McFadyen *et al.,* 2004; Rosic *et al.*, 2009).

2.1.1.2 Drug Interaction

In clinical pharmacology, when patient are under multidrug therapy, drug-drug interaction is important factor to be considered to avoid adverse drug metabolism. Many drugs are known to increase or decrease the metabolic activity of different cytochrome P450 isozymes as a result of induction or inhibition of P450 enzymes (Badyal and Dadhich, 2001; Wu *et al.,* 2012). This has direct impact on cellular drug concentration as one drug might inhibit the P450 isozyme that is responsible for second drug clearance thus accumulating the second drug within body to toxic levels. Hence, study of enzymatic activity of cytochrome P450 isozymes in response to different drug concentration is pivotal to understand drug-drug interactions and to minimize side effects of drugs (Orr *et al.,* 2012).

2.1.1.3 Genotoxic screening/ bioactivation of procarcinogens

Promutagenic and procarcinogenic compounds are suggested to undergo cytochrome P450 mediated bioactivation and exhibit their detrimental biological effects. Environmental pollutants, smoke condensates of proteinous food products, river and industrial wastes have been reported to contain such potent procarcinogenic compounds (Reifferscheid *et al.,* 1991; Oda *et al.,* 1995; Sugimura T*,* 1977). The in-vitro assays for screening of cytochrome P450 mediated transformation of potential carcinogenic metabolites would be helpful to identify genotoxins and prevent induction of pro-oncogenes. In Ames test, *umu* assay or short term bacterial genotoxicity assay, rat liver microsome or S9 fraction of phenobarbitol induced rat liver have been used, but these assays have been modified by expressing human/rat cytochrome P450s along with reductase in these assays (Ames *et al.*, 1973; Oda *et al.*, 1985; Aryal *et al.*, 1999; 2000; Oda *et al.*, 2001).

2.1.1.4 Production of Industrially significant molecules

The ability of cytochrome P450 isoforms to selectively metabolize vast array of biomolecules makes it suitable enzymatic system for production of industrially significant molecules that are not feasible or difficult to synthesize by chemical process. Apart from its central importance in clinical toxicological studies, this enzyme also has potential to generate putative drug metabolites for pharmaceutical industries (Gillam EMJ*,* 2008).

Furthermore, engineering of cytochrome P450s enzymes by directed evolution is envisaged to significantly improve catalytic properties to expand its applications ranging from fine chemical synthesis (Nakamura *et al.*, 2001), drug development (Wu *et al.*, 2005), gene therapy and bioremediation to xenobiotic-detecting biosensors (Nakamura *et al.,* 2001; Bistolas *et al.,* 2005; Gillam EMJ*,* 2008).

2.1.2 Nomenclature of Cytochrome P450s

The first systemic nomenclature of the complex and diverse group of cytochrome P450 super family enzymes has been provided by Nebert *et al.,* 1989. These enzymes are highly polymorphic and their polymorphic characterization is being continued and updated periodically (Nebert *et al.,* 1991; Nelson *et al.,* 1993, 1996). The fully annotated comprehensive cytochrome P450 nomenclature including recent changes can be retrieved from David Nelson's homepage for cytochrome P450 nomenclature: http://drnelson.uthsc.edu/human.P450.table.html. Similarly, the list of human CYP allele nomenclature can be found in Karolinska Institute's official website http://www.cypalleles.ki.se/.

2.1.2.1 Nomenclature of Human CYP families

The human cytochrome P450 super family has been divided into number of families and subfamilies according to the percentage of their amino acids sequence similarity. Members of same cytochrome P450 family (designated as CYP1 or CYP2, etc.) must share at least 40% amino acid identity among the same family (between CYP1 members with different sub-family), and in the same subfamily (denoted as 2A or 2B or 2C etc.) protein should have at least 55% amino acid identity between individual members within the same sub-family such as members of CYP 2A (Levy RH, 1995). Thus, those cytochrome P450s which fall in the same family but within different sub-families (such as CYP1A and CYP1B) there should be at least 40% amino acids similarity and those individual isoforms of same sub-family (such as CYP1A1 and CYP1A2) there should be at least 50% similarities between the members of same sub-family.

In general, cytochrome P450 enzymes are abbreviated as CYP or P450 followed by an Arabic numeral indicating gene family, a capital letter indicating subfamily and again another numeral for assigning individual gene (Nelson *et al.,* 1996) as described in illustrative example below. Thus, this system of nomenclature is solely based on genetic information and has no functional implication. Human cytochrome P450s has been grouped among 18 CYP families and 43 subfamilies (Nelson DR, 2003) and list of some important human cytochrome P450s and their metabolic functions are listed in Table 2.1.

For example: CYP 2A6
CYP = cytochrome P450
2 = genetic family
A = genetic sub-family
6 = specific gene/ polypeptide

Table 2.1: List of some of important human cytochrome P450s with its members and essential metabolic functions.

Family	Function	Members	Names
CYP1	drug and steroid (especially estrogen) metabolism	3 subfamilies, 3 genes, 1 pseudogene	CYP1A1, CYP1A2, CYP1B1
CYP2	drug and steroid metabolism	13 subfamilies, 16 genes, 16 pseudogenes	CYP2A6, CYP2A13, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP2E1, CYP2F1, CYP2J2, CYP2R1, CYP2S1, CYP2W1
CYP3	drug and steroid (including testosterone) metabolism	1 subfamily, 4 genes, 2 pseudogenes	CYP3A4, CYP3A5, CYP3A7, CYP3A43
CYP4	arachidonic acid or fatty acid metabolism	6 subfamilies, 12 genes, 10 pseudogenes	CYP4A11, CYP4A22, CYP4B1, CYP4F2

2.2 Heterologous expression of Cytochrome P450

Due to vast application of cytochrome P450 and its significant role in drug metabolism, toxicity and in carcinogenicity/mutagenicity studies, this enzyme system has been of primary interest for heterologous expression in pharmaceutical industries. In addition, use of human liver microsome (where CYP is abundantly found) in drug metabolism studies has always demanded for expression of functional enzymes in easily reproducible system. Thus, for elucidation of primary structure and study of functional properties of different CYP isoforms, an essential prerequisite had been the development of an efficient system for heterologous expression of catalytically active P450s enzyme.

Unlike other bacteria (*Bacillus* sp., *Pseudomonas* sp. etc) and eukaryotes, *Escherichia coli* lacks its own cytochrome P450s making it excellent host for functional assay of cytochrome P450s in the absence of native background (Harnastai *et al.*, 2006). However, expression of membrane bound active eukaryotic cytochrome P450s are compromised in *E. coli* due to several factors (Barnes *et al.,* 1991; Fisher *et al.,* 1992; Pernecky *et al.,* 1993; Inoue *et al.,* 2000). The presence of hydrophobic residues forming the membrane-spanning segment of cytochrome P450s, the proline rich clusters in the hinge between the membrane anchor and the catalytic part of the protein, high frequency of positively charged amino acids (Arg) and rare codons collectively account for relatively low CYP expression (Harnastai *et al.,* 2006).

Common strategies employed to alleviate amount of functional cytochrome P450 levels in *E. coli* includes N-terminal sequence modification (second codon GCT) or N-terminal truncation of hydrophobic region of cytochromes (Fisher *et al.,* 1992; Pernecky *et al.,* 1993) and use of lowered temperature (20-28^{0}C) during cytochrome P450 expression (Barnes *et al.,* 1991; Fisher *et al.,* 1992). Moreover, to assist proper folding of cytochrome P450 in *E. coli*, efforts had been made to co-express molecular chaperon (GroEL) along with cytochrome P450 (Inoue *et al.,* 2000). In addition, cytochrome P450s require their cognate redox partner for efficient transfer of electron derived from oxidation of NAD(P)H for its monooxygenase activity and *E. coli* lacks this flavoprotein to transfer electron to eukaryotic cytochrome P450s. This has been achieved through several strategies which are as described below.

2.2.1 Cytochrome P450 and NADPH-P450 reductase

The normal functioning of cytochrome P450s require a cognate flavoprotein redox partner (NADPH-P450 reductase) and supplementation of NADPH or NADH redox cofactors (Roberts *et al.,* 1999). Three different strategies have been used to co-express NADPH-P450 reductase for efficient transfer of electron to heme and augment catalytic efficiency of expressed cytochrome P450s. These include co-expression of NADPH-P450 reductase through two plasmids system (Kranendonk *et al.,* 1999), expression as fusion protein with cytochrome P450 (Fisher *et al.,* 1992) and a bicistronic construct (Figure 2.2, Parikh *et al.,* 1997). The expression of NADPH-P450 reductase through two plasmid system may limit expression of other proteins such as expression of cytochrome *b5* that may be required for some cytochrome P450s or other phase II drug metabolizing enzymes while expression as fusion proteins with cytochrome P450s may limit compatibility with all cytochrome P450 isoforms. Thus, expression as bicistronic construct consisting human cytochrome P450 as the first cistron and human NADPH-P450 reductase (hNPR) as the second cistron can support cloning of other flavoproteins if they are found to be compatible and also allow expression of other proteins in another compatible plasmids for co-transfection and co-expression studies (Oda *et al.*, 2001).

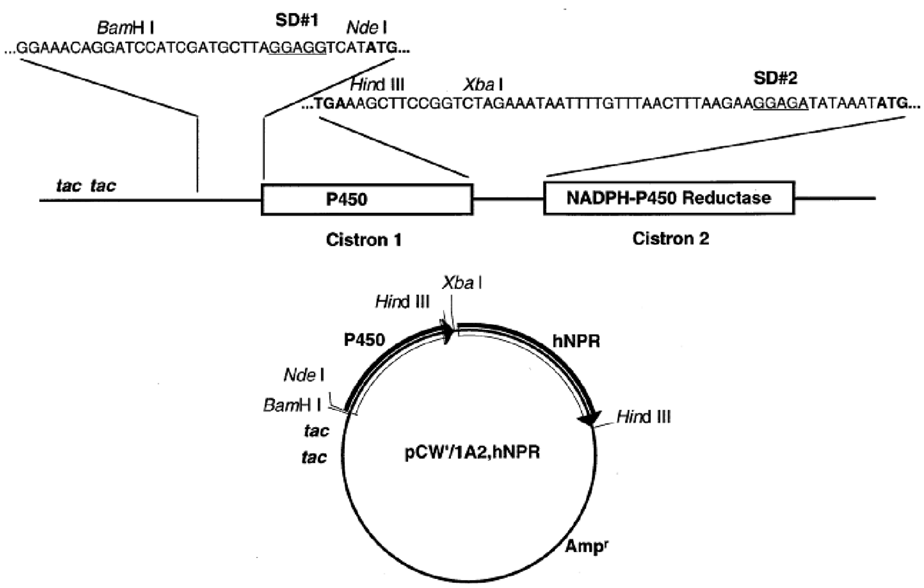


Figure 2.2: Bicistronic construct consisting of human P450 followed by human NADPH-P450 reductase cDNA under double tac promoter cloned in pCW vector (Parikh *et al.*, 1997) for expression of P450 and reductase that has been used in present study. SD, Shine-Dolgarno box.

Initially, seven cytochrome P450 isoforms had been expressed as bicistronic construct with human NADPH system. Cytochrome P450 is expressed as the first cistron followed by short sequences containing ribosome binding site for the second cistron containing NADPH-P450 reductase under double *tac* promoter. In this construct induction of double *tac* promoter by IPTG transcribes single mRNA that translates two discrete proteins (Parikh *et al.,* 1997) with independent ribosome binding sites. This bicistronic construct has been successfully used for expression of other proteins through compatible plasmids (Aryal *et al.*, 1999; 2000) and had been reported as better option than two plasmid system of expression of redox partner and cytochrome P450. Similar bicistronic construct with cytochrome P450 2A6 followed by NADPH-P450 reductase (Gillam *et al.,* 1999) that carries L240C/N297Q double mutant (Nakamura *et al.*, 2001) has been used in the present study in transforming two compatible plasmids for co-expressing genes from both the plasmids.

2.2.2 Cytochrome P450 and heme cofactor

However, in *E. coli* based expression systems, the heterologous expression of functionally active cytochrome P450s and other hemeproteins are still compromised due to limitation of free heme availability (Harnastai *et al.*, 2006). In *E. coli*, heme biosynthetic pathway is under strict regulatory control as free heme and its intermediate metabolites are potentially toxic compounds when present in excess amount in cell (Nakahigashi *et al.,* 1991). In these bacterial systems heterologously expressed hemeproteins are synthesized as apo-proteins and become catalytically active (holo-proteins) only when reconstituted with heme (Harnastai *et al.,* 2006). Thus, under normal growth conditions the hemeproteins with lower affinity for heme than that of intrinsic cellular hemeproteins (cytochrome C) are often expressed predominantly as inactive apo-proteins (Springer and Sligar, 1987; Chudaev and Usanov, 1997). This necessitated development of strategies to supplement heme or its precursor molecules either through exogenous addition or endogenous overproduction.

Since *E. coli* K-12 strain derivatives are impermeable for free heme thus heme limitation for efficient cytochrome P450s expression have been circumvented by exogenous supplementation of 5-Aminolevulinic acid (ALA), a rate limiting precursor for heme. However, this approach is not suitable for mutagenicity assay as high concentration of ALA is itself mutagenic in presence of $Fe^{2+/3+}$ ions (Onuki *et al.,* 2002). Furthermore, 5-ALA supplementation may not support high level expression of active cytochrome P450s when heme concentration could not be maintained at constant levels and does not coincide with hemeprotein synthesis. In addition, 5-ALA is an expensive precursor that accounts for nearly 60% of the total expense of cytochrome P450s expression in *E. coli* (Harnastai *et al.,* 2006). Thus in the present study, endogenous supplementation of heme precursors by augmented heme biosynthetic pathway within the cell has been envisaged for efficient expression of functionally active cytochrome P450s in the absence of exogenous 5-ALA.

2.3 Heme biosynthesis

In biosynthesis of tetrapyrrole containing biological molecules ALA is the first committed precursor (Schobert and Jahn, 2002). The biosynthetic pathway branch out from uroporphyrinogen III, as illustrated in figure 2.3, to form respective porphinoids like corrinoids (vitamin B12), siroheme or coenzyme F_{430} and heme according to biological requirements in different organisms (Dailey HA, 1990; Jordan PM, 1991; Beale SI, 1996). The formation of heme from ALA requires seven enzymatic reactions, including assembly of eight ALA molecules into cyclic tetrapyrrole, side chains modification, and chelation of reduced iron molecule. All of these reactions are tightly regulated (Nakahigashi *et al.*, 1991) and remarkably conserved among all organisms (Verderber *et al.*, 1997; Schobert and Jahn, 2002).

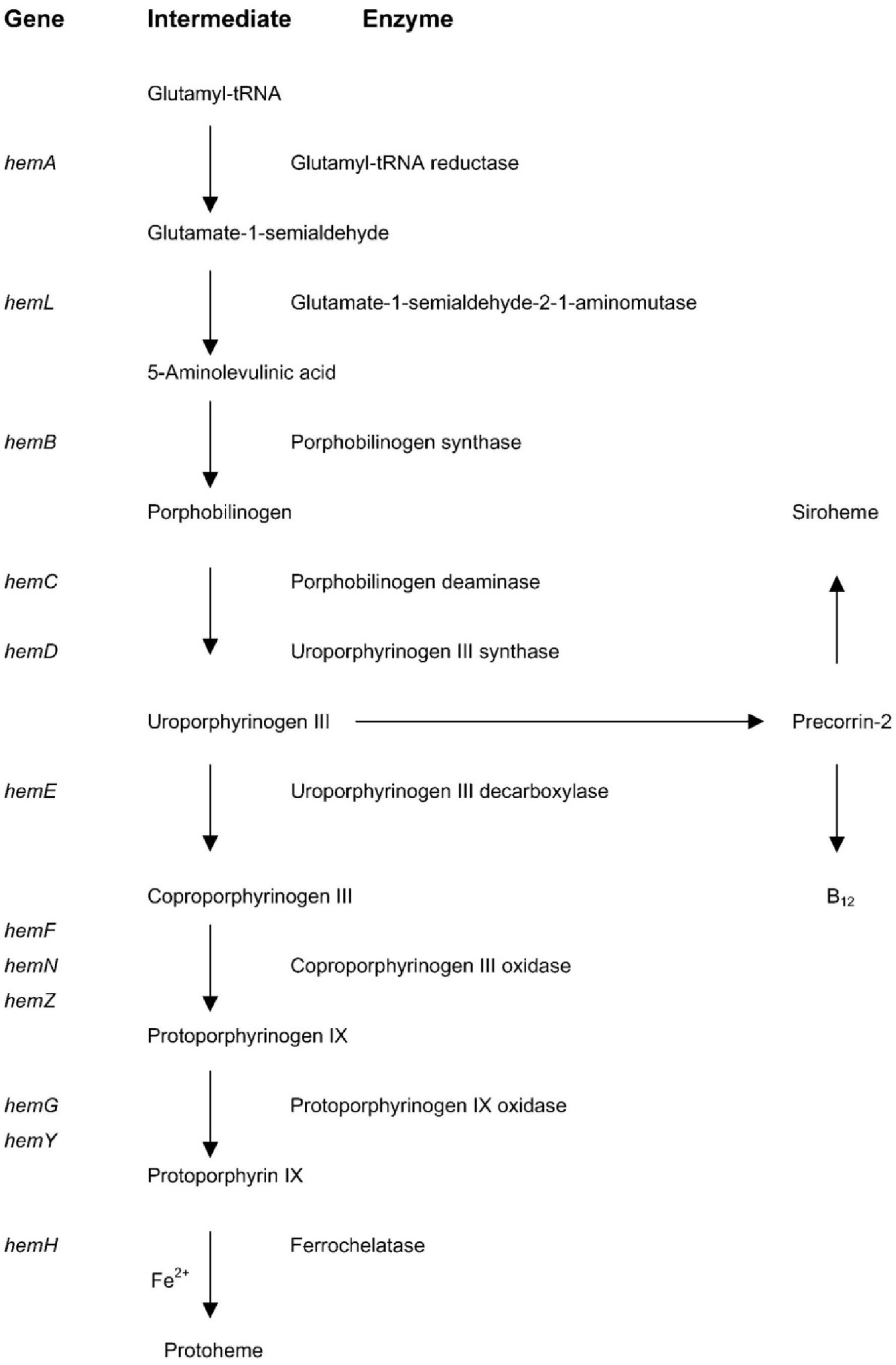


Figure 2.3: Tetrapyrroles biosynthesis pathway in bacteria. The main biosynthetic pathway branches out from central intermediate precursor Uroporphyrinogen III to various tetrapyrrole biomolecules including protoheme, vitamin B_{12} and siroheme (Adopted from Schobert and Jahn, 2002).

These tetrapyrrole containing molecules have vast applications in biological systems as they are found in chlorophyll involved in carbon assimilation and carbohydrate biosynthesis, cytochrome C in respiration, cytochrome P450 in drug metabolism, toxicity, drug discovery and drug development, siroheme for sulfur transfer in cysteine biosynthesis, corrin ring of vitamin B12 molecule thus having diverse implications in medical, agricultural and biotechnological applications (reviewed in Schobert and Jahn, 2002).

2.3.1 Regulation of Heme Biosynthesis in *E. coli*

In *E. coli*, the demand of heme is modulated depending upon its metabolism pattern, aerobic or anaerobic. Requirement of heme significantly increases during aerobic respiration due to requirement of heme containing respiratory enzymes for efficient electron transfer whereas its requirement is reduced when bacteria are metabolizing anaerobically (Doss and Philipp-Dornston, 1971; Jacobs *et al.*, 1972). As metabolic state appears to determine heme concentration in the cell it clearly indicates that the cell could have potential regulatory mechanism for heme biosynthesis regulation based on the cellular requirement. This can be substantiated because when heme binding proteins such as cytochrome, hemoglobin or catalases were heterologously overexpressed in *E. coli*, heme level was found to increase upto 20 fold (Verderber *et al.*, 1997). This clearly suggested that heme most probably regulates its own biosynthesis according to its cellular demand (Woodard and Dailey, 1995; Schobert and Jahn, 2002).

Two regulatory steps in heme biosynthetic pathway have been identified. One is during biosynthesis of main precursor molecule, ALA, as demonstrated by ALA feeding experiments where more heme and porphyrins were found to be accumulated upon exogenous supplementation of ALA (Jacobs *et al.*, 1972; Philipp-Dornston and Doss, 1973) and other rate limiting step was found to be formation of protoporphyrinogen IX from coproporphyrinogen III (Tait GH, 1969; Xu *et al.*, 1992). To overcome these limitations in overexpression of cytochrome or other tetrapyrrole containing molecules ALA has been supplemented in culture media (Wu *et al.*, 2005). However, cost associated, optimal P450 expression can be limited and attempts have been made to increase ALA biosynthesis inside the cell than exogenous supplementation (Harnastai *et al.*, 2006) for sufficient and optimal heme biosynthesis.

2.3.2 ALA Biosynthesis

In living organisms ALA is known to be biosynthesized from two different pathways, either C4 (Shemin) or C5 pathway (Figure 2.4). In alpha proteobacteria, yeast and mammalian cells ALA is biosynthesized through C4 pathway directly by condensation of glycine and succinyl-CoA (C4 compound) in a single enzymatic reaction catalyzed by 5-Aminolevulinate synthase (Avissar *et al.*, 1989; Suzuki *et al.*, 1997; Ferreira GC, 1999).

In higher plants, algae and many bacteria including *E. coli* biosynthesis of ALA is through C5 pathway (Jahn *et al.*, 1992; Avissar and Moberg, 1995). In C5 pathway ALA is biosynthesized from glutamate, a five carbon compound, involving three enzymatic reactions. In first enzymatic step GTR synthetase converts glutamate to glutamyl-tRNA (GTR) (McNicholas *et al.*, 1997) which is followed by second enzymatic reaction in which glutamyl-tRNA reductase (GTR reductase; HemA) in the presence of NADPH converts GTR to glutamate 1-semialdehyde (GSA) (Avissar *et al.*, 1989; Verkamp *et al.*, 1992) and subsequently in the final enzymatic step GSA aminotransferase (HemL) transaminates GSA to ALA (Ilag and Jahn, 1992; Jahn *et al.*, 1992; Smith *et al.*, 1992).

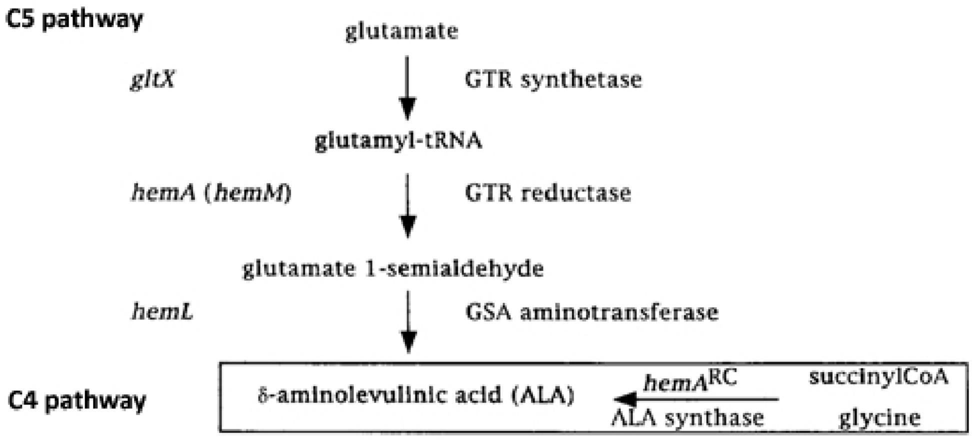


Figure 2.4: Schematic illustration of ALA biosynthesis through C5 and C4 pathways (Adopted from Verderber *et al.*, 1997).

In different studies it has been demonstrated that over expression of homologous or heterologous glutamyl-tRNA reductase (encoded by *hem*A gene) in *E. coli* overproduced ALA indicating that glutamyl-tRNA reductase could be rate limiting enzyme for ALA biosynthesis that subsequently limits heme biosynthesis (Chen *et al.*, 1994; Verderber *et al.*, 1997). Qi and coworkers (2011) further validated this hypothesis by separately over expressing genes involved in C5 pathway, *gltX*, *hem*A and *hem*L and found that increased ALA accumulation was observed in *E. coli* strain only when *hem*A gene was over expressed (Kang *et al.*, 2011). This result suggested that over production of heme and other tetrapyrrole containing biological molecules would require improved ALA biosynthesis for which up-regulation of expression of *hem*A gene and the activity of HemA protein would be pivotal that potentially requires deregulation of *hem*A gene transcription and protein stability.

2.3.2.1 Transcriptional Regulation of *hemA* gene

As mentioned above heme regulates its own biosynthesis and ALA being the main precursor for heme, *hem*A gene might be regulated. It has been found that both transcriptional and post transcriptional regulation is observed for *hem*A gene expression and its protein function, respectively (Schobert and Jahn, 2002).

In *E. coli*, an operon consisting *hemA* gene followed by *prfA* gene encoding protein release factor 1, and *prmC* (*hemK*) gene encoding N5-glutamine methyltransferase has been reported (Figure 2.5). Verkamp and Chelm (1989) had identified two rho-dependent promoters in *hemA* gene containing operon. Proximal promoter (P_1) contributes more than 90% transcription of *hemA,* along with two other genes in *hemA-prfA-prmC* operon as single transcriptional unit, whereas distal promoter (P_2) transcribes only *hemA* gene and is only 10-15% active (Nakayashiki *et al*., 1995; Schobert and Jahn, 2002).

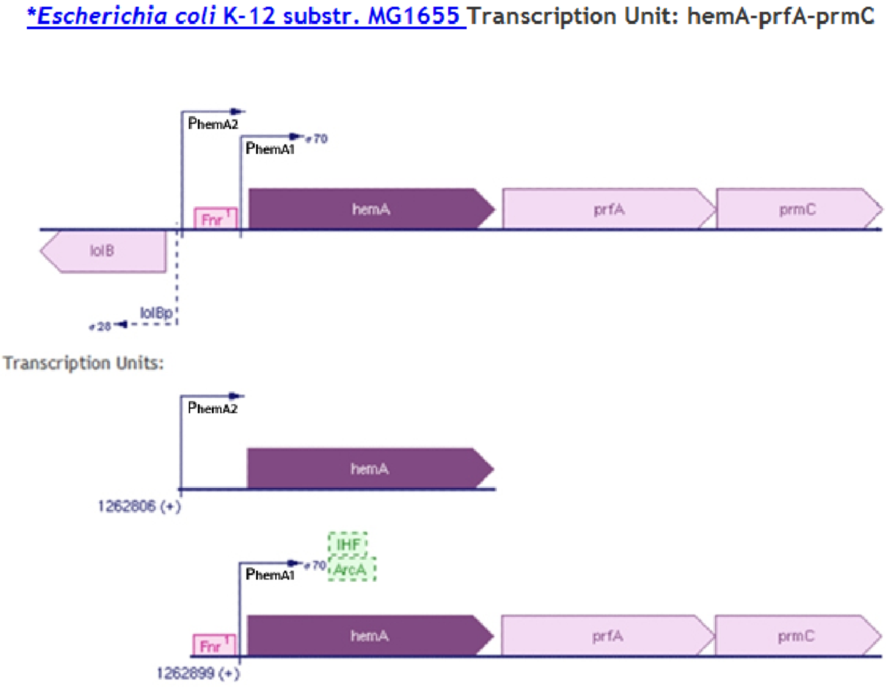


Figure 2.5: *E. coli hem*A-*prf*A-*prm*C operon showing two promoters locus, corresponding transcriptional units and transcription regulators. (Derived from Ecocyc: http://ecocyc.org)

Several experiments have been conducted to study the transcriptional regulation along with the effects of transcriptional activators and repressors in *hem*A gene expression. The different studies have suggested the roles of *fnr*, *arc*A gene products and high level of heme for regulation of *hem*A gene expression in cell (Darie and Gunsalus, 1994). The upstream region of proximal promoter (P_{hemA1}) contains Fnr repressor binding single consensus half-site (TTGAT; Figure 2.6). This indicated possible effect of oxygen concentration regulating *hem*A transcription and it has been found that *hem*A gene expression is repressed under anaerobic condition triggered by Fnr modulated repression of gene expression involved in aerobic metabolism (Melville and Gunsalus, 1996).

Figure 2.6: Fnr gene product binding consensus sequence found in proximal promoter (P_{hemA1}) of *hem*A operon, also indicating transcription initiation site, TATA binding site (-10 region) and -35 consensus region (Derived from Ecocyc: http://ecocyc.org).

However, another transcription regulator, *arcA* gene product, activates *hemA* transcription during both aerobic and anaerobic growth. Integration Host Factor (IHF) is also known to be required for positive control of *hemA* gene expression (Figure 2.5) (Darie and Gunsalus, 1994; Melville and Gunsalus, 1996).

In the studies to understand regulation of *hemA* gene transcription in *E. coli*, *lac* operon fused to *hem*A promoter was constructed and integrated in chromosome of ALA auxotroph *E. coli* to express as single copy (Darie and Gunsalus, 1994; Choi *et al.*, 1996; McNicholas *et al.*, 1997). For normal growth, this *hemA* mutant strain requires exogenous supplementation of ALA and faces serious heme limitation in the media without ALA. Under aerobic conditions, 4-6 fold increase in induction of *hem*A-*lac*Z fusion was observed during ALA starvation as indicated by cellular β-galactosidase activity. However, the induction was restored to normal level after addition of ALA (Choi *et al.*, 1996) indicating that *hemA* gene is transcriptionally controlled by ALA concentration in the cell. Furthermore, the transcription of *hem*A gene was also found to be elevated by 1.5 fold during anaerobic condition, too, compared to that of aerobic growth during heme starvation (McNicholas *et al.*, 1997). Thus, these results clearly indicated that *hem*A gene is only moderately regulated at transcriptional level due to heme starvation and provides strong indication towards post transcriptional control of glutamyl-tRNA reductase, HemA protein (Wang *et al.*, 1999b).

2.3.2.2 Conditional protein stability

Suggestions that HemA protein might go posttranslational regulation lead to various studies which have postulated that glutamyl tRNA reductase (HemA) is under feedback regulation by heme. Despite the moderate level of transcriptional control of *hemA* in both heme limiting and abundant conditions, increase in abundance of HemA protein under heme limiting condition indicated importance of post transcriptional regulation. HemA activity was substantially elevated upto 10- to 25-fold when different heme mutant *Salmonella typhimurium* was imposed to severe heme starvation (Wang *et al.*, 1997). Furthermore, pulse-labeling and immunoprecipitation experiments showed the increase in half-life of HemA protein from 20 minutes in unrestricted cells to >300 minutes in heme-limited cells that clearly indicated that the mechanism of regulation involved could be protein stabilization that is dependent upon heme. Thus, it was inferred that the HemA protein is vulnerable to its stability which is greatly reduced when heme gets accumulated in the cell and this conditional protein stability could play the major role during regulation of heme biosynthetic pathway (Wang *et al.*, 1999a).

In one of the studies, in vitro addition of heme had no direct inhibitory effect on enzymatic activity of *E. coli* HemA protein (Jahn *et al.,* 1991), however, the glutamyl tRNA reductase activity is greatly reduced upon addition of heme to cell extracts of *E. coli* (Javor and Febre, 1992) indicating that most probably heme alone does not alter the enzymatic activity of HemA but could facilitate in destablizing by inducing other regulatory mechanisms to be activated within the cellular component. This clearly provided suggestive evidence for requirement of additional factors, apart from heme, for regulation of HemA activity and possibly *E. coli* HemA could have similar function as that of observed with *S. typhimurium* HemA.

2.3.2.3 Heme regulated proteolysis of HemA

Earlier evidence of involvement of other factors were first postulated by Elliott and coworkers who first identified the involvement of two ATP dependent proteases, Lon and ClpP, in *S. typhimurium* HemA turnover rate. HemA protein was found to be completely stabilized in *lon* and *clpP* double mutant *E. coli*. However, either of these single mutant, by itself, only contributed small amount for stabilization of HemA. This clearly indicated the collective activities of both of these energy dependent proteases, Lon and ClpP, are important for destablizing the HemA protein (Wang *et al.,* 1999a). Furthermore, the proteolytic cleavage of HemA protein being facilitated by heme has been reviewed, and illustrated in Figure 2.7 (Schobert and Jahn, 2002) where N-terminus region of HemA protein is prone to these proteases mediated cleavage.

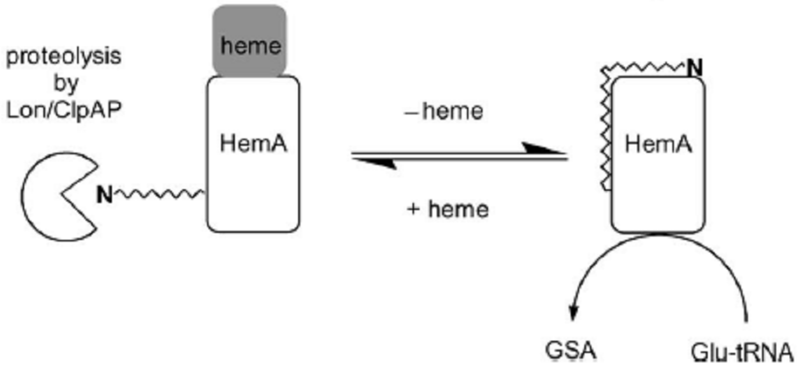


Figure 2.7: Heme regulated proteolytic cleavage of HemA protein by two energy dependent proteases Lon/ClpAP (Adopted from Schobert and Jahn, 2002).

2.3.2.4 Stabilized HemA(KK) protein against feedback regulation

There are considerable evidences to support the role of N and C termini of protein for stability or susceptibility towards proteolysis. Based on similar hypothesis, hybrid protein constituting N terminal 18 amino acids of HemA fused with *lacZ* gene, $HemA_{1\text{-}18}$-LacZ, had been constructed to determine the role of N-terminal amino acids for heme regulated proteolysis. Interestingly, $HemA_{1\text{-}18}$-LacZ underwent proteolytic cleavage by same two proteases in unrestricted cells and the hybrid protein was found to be stabilized in *lon* and *clpP* double mutant strain, as observed in case of wild type HemA. These results confirmed that N-terminal part of HemA posses the degradation tag which is recognized by proteases to start its progressive degradation (Wang *et al.,* 1999a).

Since Lon and ClpAP proteases specifically recognize 18 N-terminal amino acids of HemA or the subset of them, several independent site-directed mutations within this sequence were created to alter important residues thereby interfering with degradation process. The mutations include introduction of positively or negatively charged amino acids, addition of epitope tag or deletion of few amino acids. Fortunately, one of mutant HemA(KK) with two lysine residues (AAGAAG) inserted between second and third amino acids (change from ^{1}MTLLALGI^{8} to ^{1}MT$\underline{KK}$LLALGI^{10}) was abundant and completely stabilized against Lon and ClpAP proteases. Addition of positively charged amino acid might have interfered with initial binding of proteases or prevent subsequent conformational change of HemA (Wang *et al.,* 1999b), most probably after binding of heme which is required for initiation of protease attack.

Although the initial investigations on heme regulated proteolysis of HemA protein were carried out using *S. typhimurium* HemA protein, but, the glutamyl tRNA reductase activity that was found to be greatly reduced upon addition of heme to cell extracts of *E. coli* (Javor and Febre, 1992) clearly indicated that *E. coli* and *S. typhimurium* HemA proteins would undergo similar fate. This was evident from recent works on introduction of two lysine residues close to N-terminus of *E. coli* HemA which rendered protection and insensitivity towards heme regulated proteolysis that efficiently produced ALA and heme, irrespective of cellular heme concentration (Harnastai *et al.*, 2006; Kang *et al.*, 2011).

2.3.2.5 Glutamyl tRNA reductase

Avissar and Beale (1989) first identified enzymatic basis for ALA auxotrophy in *hemA* mutant *E. coli*. This mutant strain was able to convert charged glutamyl tRNA to glutamate-1-semialdehyde and subsequently formed ALA (Figure 2.8) only when glutamyl tRNA reductase was exogenously supplemented. This confirmed the important role of *hemA* gene product in ALA biosynthesis. The involvement of glutamyl $tRNA^{Glu}$ as precursor molecule for biosynthesis of tetrapyrrole containing molecules also signified potential role of charged tRNA in metabolite biosynthetic pathway apart from universal dogma of protein biosynthesis in inserting respective amino acids in elongating peptide chain.

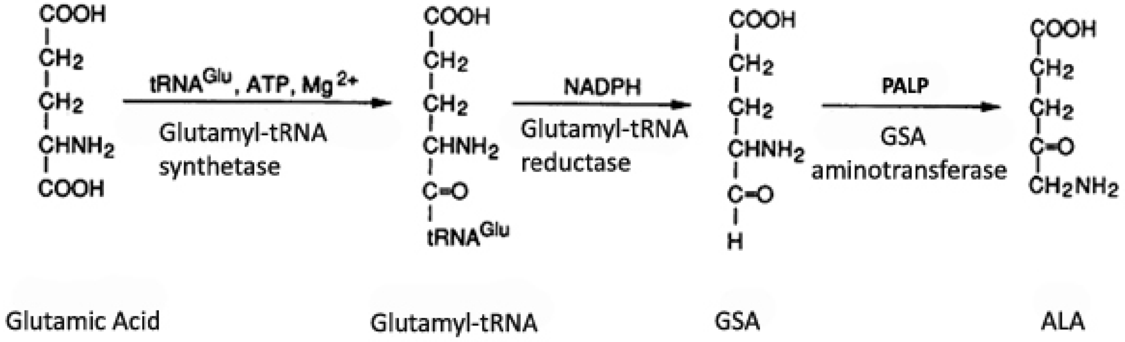


Figure 2.8: C5 pathway of ALA biosynthesis in *E. coli*. Chemical structure of intermediates, cofactors and enzymes involved in the reactions are illustrated (Adopted from Avissar and Beale, 1989).

Verkamp and Chelm (1989) reported that the open reading frame of *E. coli* glutamyl tRNA reductase (GluTR or HemA protein) encoded by *hemA* gene consists of 418 amino acids with a calculated molecular mass of 46 kDa (Figure 2.9). However, two proteins with distinct glutamyl tRNA reductase activities having molecular weights of 85 kDa (GluTR 85) and 45 kDa (GluTR 45), respectively have been reported in *E. coli* (Jahn *et al.*, 1991). Later, it was elucidated that Glu-tRNA reductase can exist in multiple multimeric forms, but dimeric structure (GluTR 85) represents functionally active enzyme (Moser *et al.*, 2001).

(A) *E. coli hemA* gene information derived from NCBI

Gene symbol	*hem*A
Gene description	glutamyl tRNA reductase (GTR reductase, GluTR, HemA)
Organism	*Escherichia coli* str. K-12 substr. MG 1655
Lineage	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; *Escherichia*
Also known as	ECK1198; gtrA; JW1201; b1210
Annotation	NC_000913.2 (1262937…1264193)
Length	1257 bp
Gene ID	945777
GI	16129173
ECOCYC	EG10427
EcoGene	EG10427
Protein	NP_415728.1
UniProtKB	P0A6X1, http://www.uniprot.org/uniprot/P0A6X1
Amino acid	418
Molecular weight	46.3 kDa (calculated based on amino acid sequence in uniprot P0A6X1)

(B) Deduced amino acids sequence of Glutamyl-tRNA reductase

```
        10         20         30         40         50         60
MTLLALGINH KTAPVSLRER VSFSPDKLDQ ALDSLLAQPM VQGGVVLSTC NRTELYLSVE

        70         80         90        100        110        120
EQDNLQEALI RWLCDYHNLN EEDLRKSLYW HQDNDAVSHL MRVASGLDSL VLGEPQILGQ

       130        140        150        160        170        180
VKKAFADSQK GHMKASELER MFQKSFSVAK RVRTETDIGA SAVSVAFAAC TLARQIFESL

       190        200        210        220        230        240
STVTVLLVGA GETIELVARH LREHKVQKMI IANRTRERAQ ILADEVGAEV IALSDIDERL

       250        260        270        280        290        300
READIIISST ASPLPIIGKG MVERALKSRR NQPMLLVDIA VPRDVEPEVG KLANAYLYSV

       310        320        330        340        350        360
DDLQSIISHN LAQRKAAAVE AETIVAQETS EFMAWLRAQS ASETIREYRS QAEQVRDELT

       370        380        390        400        410
AKALAALEQG GDAQAIMQDL AWKLTNRLIH APTKSLQQAA RDGDNERLNI LRDSLGLE
```

Figure 2.9: *E. coli hemA g*ene information (A) and deduced amino acid sequence (B).

2.3.2.6 Crystal Structure of GluTR

The first molecular structure of glutamyl-tRNA reductase (GluTR) was determined from hyperthermophilic Archaea *Methanopyrus kandleri*. The crystal structure of GluTR revealed V-shaped dimeric protein consisting of three distinct domains arranged along a curved 'spinal' α-helix as shown in figure 2.10. N-terminal catalytic domain interacts with charged glutamyl $tRNA^{Glu}$, NADPH binding domain supplies essential cofactor and C-terminal domain is responsible for dimerization (Moser *et al.*, 2001).

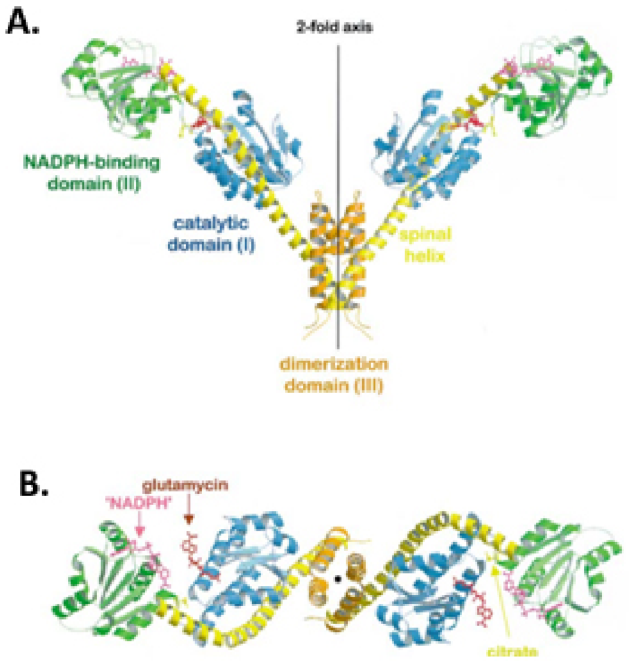


Figure 2.10: Ribbon model of V-shaped dimeric glutamyl tRNA reductase viewed (A) perpendicular to and (B) along the 2-fold axis. Each monomer consist of three distinct domain (I) N-terminal catalytic domain, (II) NADPH-binding domain and (III) C-terminal dimerization domain connected by α-helix (Adopted from Schauer *et al.*, 2002).

N-terminal catalytic domain (I) has been reported to constitute residues 1-145 with two sub domains. It is predicted to carry a small βαββααβ-motif that resembles the common βαββαβ structure and this is followed by 3 α-helices aligned with the spinal helix to form a 4-helix bundle. Second, NADPH binding domain II (residues 149-285) that is connected with domain I by a short linker has a conserved nucleotide-binding fold (Carugo and Argos, 1997) composed of central six-stranded parallel β-sheet, βαβ motifs and a glycine-rich loop. Domain II is followed by spinal α-helix consisting of 18 α-helical turns with a visible kink introduced by Pro305. The C-terminal part dimerizes with neighboring monomeric structure and together forms six-helix bundle (Moser *et al.*, 2001; Schubert *et al.*, 2002). The elucidation of X-ray crystallographic structure combined with biochemical and biophysical studies helped to understand the detail catalytic mechanism of GluTR and this was found to be similar to that of *E. coli* GluTR in further investigations (Moser *et al.*, 1999; 2001; Schauer *et al.*, 2002).

2.3.2.7 Catalytic mechanism of GluTR (HemA)

To determine the essential structural features of charged glutamyl-tRNA required for GluTR recognition, Jahn and coworkers used various glutamate analogues, uncharged tRNAs and a chemically synthesized compound mimicking the intact 3′ end of tRNA as the substrate for GluTR inhibition assay. It was found that glutamycin (Figure 2.11) inhibited GluTR activity, suggesting importance of intact chemical bond between glutamate and its cognate tRNA for efficient substrate binding and enzyme turnover (Moser *et al.*, 1999).

Glutamycin

Glutamyl-tRNA

Figure 2.11: Structure of GluTR inhibitor glutamycin compared with the 3′ end of native substrate glutamyl-tRNA (Adopted from Moser *et al.*, 1999).

Chemical modifications of *M. kandleri* GluTR enzyme and site directed mutagenesis studies helped to postulate its catalytic mechanism (Figure 2.12). A single cysteine residue was found to be important for catalytically active enzyme. Biochemical trapping and autoradiography visualization revealed that the reactive sulfhydryl group of Cys-48 attacks the aminoacyl linkage of glutamyl-tRNA^{Glu} generating an enzyme-bound thioester intermediate and frees tRNA^{Glu} during the reaction. This potential thioester intermediate then gets reduced by direct hydride transfer from NADPH to form GSA and releases the enzyme for subsequent reactions. However, in the absence of NADPH, enzyme bound glutamate moiety is released due to GluTR esterase activity (Moser *et al.*, 1999). The determination of high-resolution crystal structure of GluTR in complex with substrate like inhibitor glutamycin (Figure 2.11) has provided further evidence to validate the biochemical data and catalytic mechanism of GluTR (Schubert *et al.*, 2002).

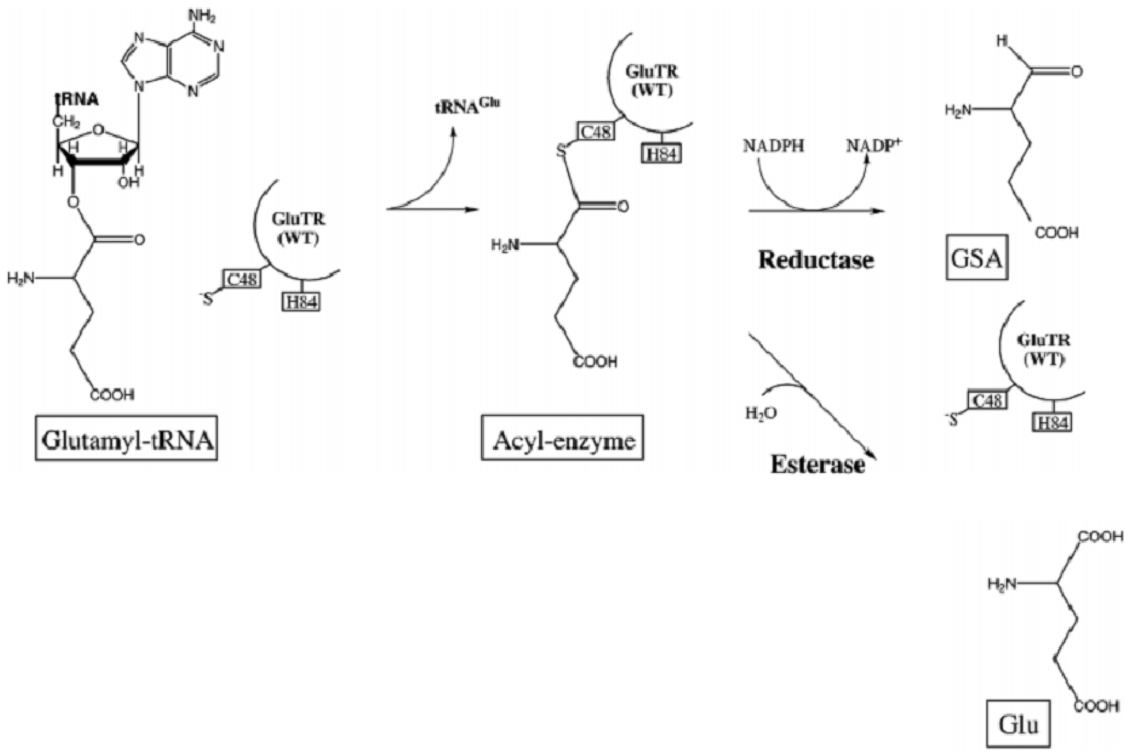


Figure 2.12: Schematic diagram representing enzymatic mechanism of *M. kandleri* GluTR. The catalytic activity of Cys-48 is facilitated by His-84 (Adopted from Moser *et al.*, 1999).

The GluTR catalytic mechanism of *E. coli* was found to be similar to that of previously well characterized *M. kandleri* GluTR in similar experimental investigations. To determine the nucleophilically active cysteine residue of GluTR found in *E. coli*, all cysteines (Cys-50, Cys-74 and Cys-170) were individually replaced with serine by site directed mutagenesis. Only GluTR C50S mutant lost its catalytic activity while other two mutants fully retained their catalytic activities suggesting that Cys-50 acts as the active-site nucleophile. Furthermore, when nearly 40 known GluTR genes from different strains were aligned together, Cys-50 was the only cysteine residue found to be conserved, thus confirming the importance of this nucleophile amino acid in GluTR catalytic activity (Schauer *et al.*, 2002).

2.3.2.8 Metabolic Channeling of GSA through GluTR and GSA-AM in ALA Biosynthesis

ALA biosynthesis requires cooperative actions of *hemA* gene product GluTR and *hemL* gene product GSA 2,1-aminomutase (GSA-AM). The reaction product of GluTR, glutamate-1-semialdehyde (GSA), is subsequently converted to ALA by aminotransferase activity catalyzed by GSA-AM. It had been proposed that the highly reactive aldehyde species, GSA, is protected from exposure to external environment by direct metabolic channeling of GSA to active site of GSA-AM. In order to perform this reaction, the formation of tight complex between the two enzymes has been hypothesized on the basis of docking model generated from solved crystal structure of GluTR of *M. kandleri* and GSA-AM from *Synechococcus* (Moser *et al.*, 2001). The V-shaped homodimer of GluTR forms complex with dimeric GSA-AM (Figure 2.13) allowing direct metabolic transfer of labile GSA without being exposed to environment for efficient ALA formation (Luer *et al.*, 2005).

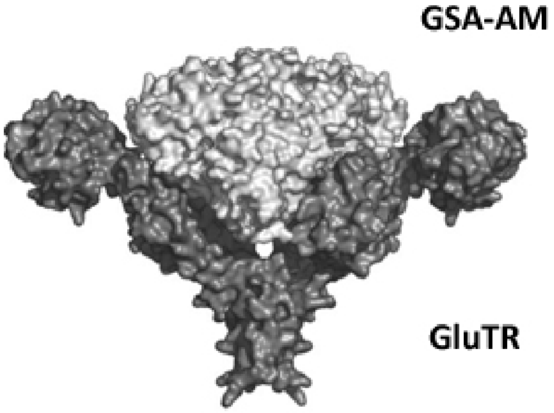


Figure 2.13: Docking model complex of *E. coli* GluTR (dark gray) and GSA-AM (light gray) (Adopted from Luer *et al.*, 2005).

The existence of this *in silico* model of GluTR-GSA-AM complex was further verified by co-immunoprecipitation experiments and gel permeation chromatography. In co-immunoprecipitation test, rabit anti-GluTR and rabit anti-GSA-AM antibodies were used to detect the physical interaction between GluTR and GSA-AM of *E. coli* in cell-free extracts or between purified recombinant enzymes (Luer *et al.*, 2005). The comigration pattern clearly indicated that the enzymes are in complex form.

2.4 Production of 5-Aminolevulinic acid (ALA)

The expression of stabilized *E. coli* HemA(KK) protein can potentially be used for ALA production. Recently, Qi and his coworkers has developed an engineered *E. coli* for efficient production of 5-ALA from glucose (Kang *et al.*, 2011). The feedback regulation of heme on HemA protein was circumvented by expressing stabilized form of *S. arizona* HemA(KK) protein. To enhance the ALA production, *hem*L was coexpressed with *hem*A(KK) as bicistronic construct and the intracellular ALA was exported out by overexpression of *rht*A gene encoding threonine/homoserine exporter with broad substrate specificity. With this new strategy of ALA production through C5 pathway, 4.13g/L ALA have been produced form glucose in modified minimal medium without adding any other co-substrate or inhibitors. Further metabolic engineering and optimization of fermentation process could improve cost effective ALA production.

Apart from role in tetrapyrrole ring containing bio-molecules such as vitamin B_{12}, heme-containing enzyme (siroheme, cytochrome), porphyrins, etc (Sasaki *et al.*, 2002), ALA has attracted much attention in recent years due to its significant implications in tumor-localizing and photodynamic therapy for various cancers (Bhowmick and Girotti, 2010; Mikolajewska *et al.*, 2010). In addition, ALA can also be used as biodegradable herbicide and insecticide due to its nontoxicity to crops, animals and humans (Edwards *et al.*, 1984; Sasaki *et al.*, 2002). The potential applications of ALA have been briefly summarized in Table 2.2.

Table 2.2: Potential applications of ALA in agriculture, medicine and biotechnology. (Sasaki *et al.*, 2002)

Field	Application	References
Agriculture	Biodegradable herbicide	Rebeiz *et al.*, 1984
	Biodegradable insecticide	Rebeiz *et al.*, 1988
	Herbicide accelerator	Sasaki *et al.*, 1998
	Growth promoting factor and yield enhancement	Tanaka *et al.*, 1992; Hotta and Watanabe, 1999
	Salt tolerance	Kuramochi *et al.*, 1997
	Cold temperature tolerance	Hotta and Watanabe, 1999
	Quality improvement of vegetables	Hotta and Watanabe, 1999
	Color intensifying effects	Hotta and Watanabe, 1999
	Growth and green-color maintenance of grass	Hotta and Watanabe, 1999
Medicine	Diagnosis of heavy-metal poisoning	Takeya *et al.*, 1997
	Diagnosis of porphyria	Takeya *et al.*, 1997
	Cancer treatment	Sasaki *et al.*, 1998
	Diagnosis of brain tumor	Kaneko *et al.*, 1998
	Treatment of rheumatoid arthritis	Trauner and Hassen, 1994
	Restore hair growth and prevent hair loss	Ishino *et al.*, 1999
	Treatment of mycosis	Peter *et al.*, 1994
	Cosmetic and dermatological applications	Uhlmann *et al.*, 1996
Biotechnology	Porphyrin production	Miyachi *et al.*, 1998
	Vitamin B12 production	Vladimir *et al.*, 1997
	Plant cell culture	Masuda *et al.*, 1994
	Animal cell culture	Fujita *et al.*, 1997
	Genetically treated bacteria culture	Nakayashiki and Inokuchi, 1996
	Heme-containing enzyme production	Imai *et al.*, 1993

2.5 Cytochrome P450 2A6 and indole metabolism

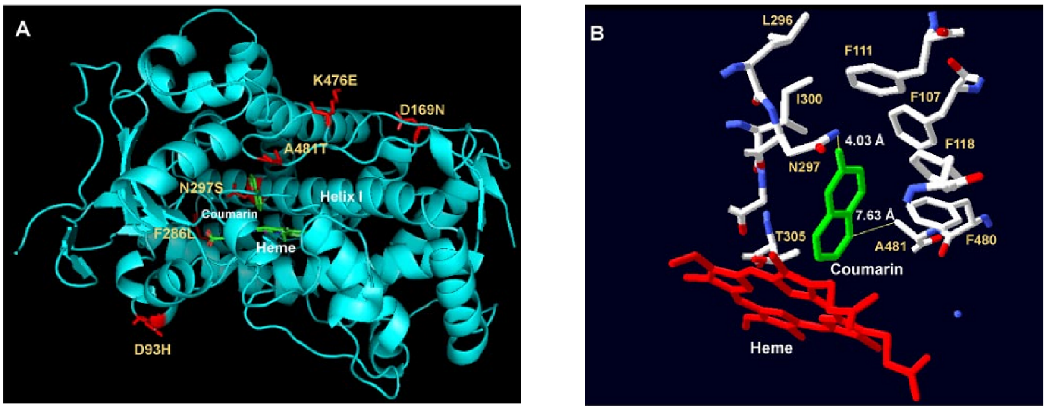


Figure 2.14: X-ray molecular structure of cytochrome P450 2A6 in complex with substrate coumarin (A) in ribbon model and (B) expanded region of active site with heme and coumarin (Adopted from Kim *et al.*, 2005).

Cytochrome P450 2A6 is an important human P450 isoform (Figure 2.14) that is involved in the metabolism of various xenobiotic substances including coumarin, indole, nicotine, aflatoxin and carcinogenic nitrosamines (Kim *et al.*, 2005). The spectral and catalytic characterization of P450 2A6 was performed by cloning its cDNA (obtained from F. J. Gonzalez, National Cancer Institute, Bethesda, MD) in expression vector which was further modified at N-terminus (second amino acid change: Ala instead of Leu) to facilitate its expression in *E. coli* as reported by Soucek, 1999. Furthermore, the construction of bicistronic pCW/2A6: hNPR plasmid, to express P450 2A6 along with its cognate redox partner NADPH-P450 reductase under double *tac* promoter enabled P450 activity to be reconstituted within bacterial cells (Gillam *et al.*, 1999).

In this particular system, the heterologous expression of cytochrome P450 2A6 resulted in development of blue pigmentation in bacterial culture and this blue pigment had been found to be indigo as the pigment extracted from the cultures was found to comigrate with standard indigo sample on Thin layer chromatography (TLC). Further, UV-visible spectroscopy and mass spectrometric analysis helped to characterize the pigment as indigo. It was then elucidated that cytochrome P450 2A6 catalyzes oxidation of indole (the product of tryptophan degradation by tryptophanase A in *E. coli*) to isatin and indoxyl products, which in turn reacted to produce indigoids including indigo (blue) as major product and indirubin (pink) as minor constituents. The major reactions in P450-catalyzed pigment formation (Figure: 2.15) was postulated by Guengerich and his group (Gillam *et al.*, 2000).

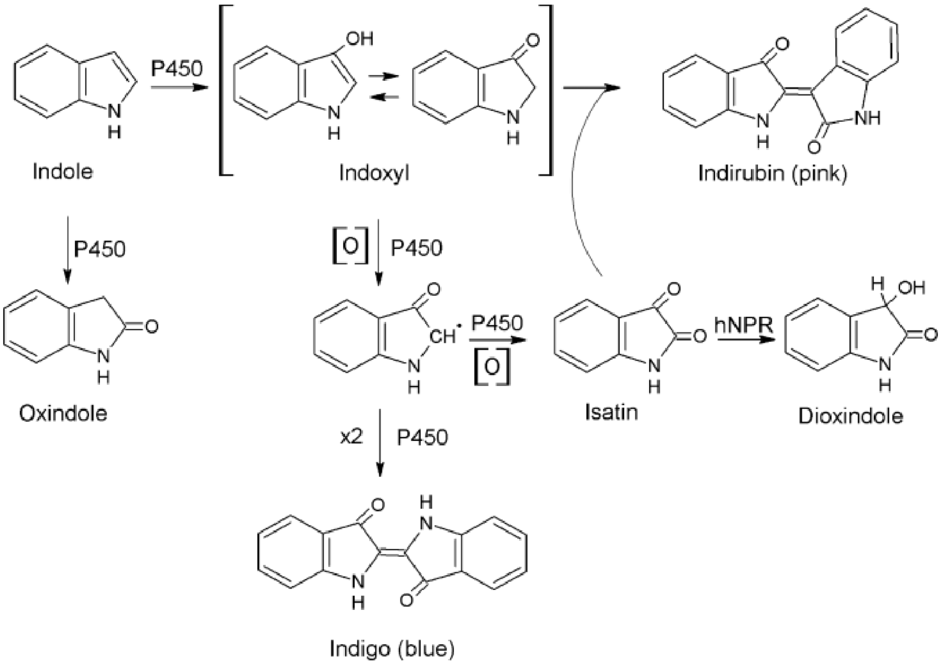
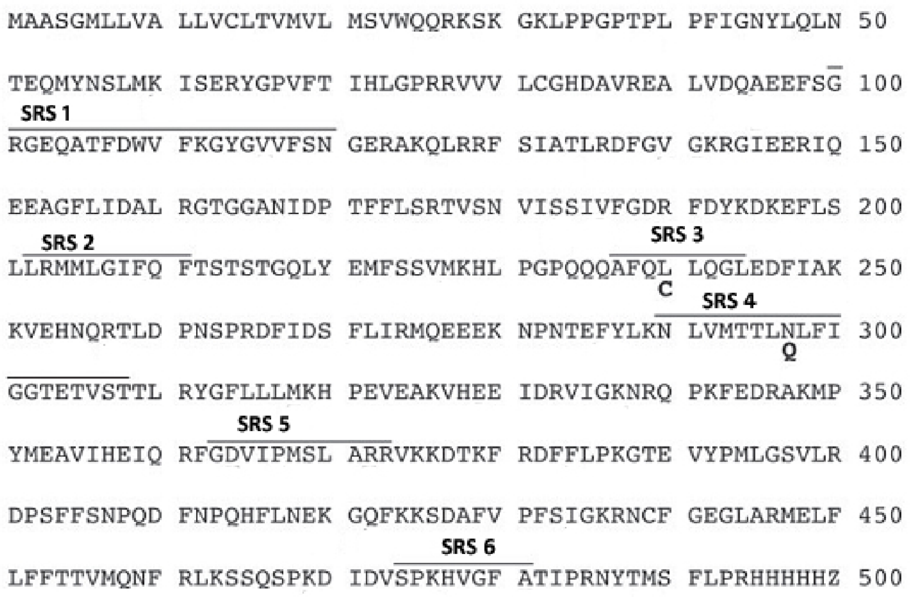


Figure 2.15: Schematic presentation of cytochrome P450 2A6 mediated metabolism of indole to indoxyl products and various diindole derivatives (Gillam *et al.*, 2000; Rosic NN, 2009).

Thus, with an efficient system where HemA protein is stabilized and expressed it can be further used for industrially valuable product formation including indigoids using cytochrome P450s. The parent bicostronic construct consisting cytochrome P450 2A6 and hNPR has been used in the present study to assess applicability of stabilized *hemA* gene co-expression strategy for endogenous supplementation of heme for functional cytochrome P450s expression.

2.5.1 Random mutagenesis of Cytochrome P450 2A6

Cytochrome P450 2A6 has broad spectrum catalytic selectivity and had potential to generate new chemical entities of commercial interest (Guengerich FP, 2002). In order to further enhance its catalytic efficiency and for the study of structure-function relationships, random mutagenesis techniques had been applied by Guengerich and his group. Putative six substrate recognition sequences (SRS) of cytochrome P450 2A6 (Figure 2.16) that were identified on the basis of Gotoh (1992) methodology were selected for generation of mutant library targeting individual SRS by site-saturation mutagenesis (Nakamura *et al.,* 2001). The functional P450 2A6 mutants were screened on the basis of colorimetric assay using indole oxidation product, indigo.



```
MAASGMLLVA LLVCLTVMVL MSVWQQRKSK GKLPPGPTPL PFIGNYLQLN 50

TEQMYNSLMK ISERYGPVFT IHLGPRRVVV LCGHDAVREA LVDQAEEFSG 100
 SRS 1
RGEQATFDWV FKGYGVVFSN GERAKQLRRF SIATLRDFGV GKRGIEERIQ 150

EEAGFLIDAL RGTGGANIDP TFFLSRTVSN VISSIVFGDR FDYKDKEFLS 200
   SRS 2                                   SRS 3
LLRMMLGIFQ FTSTSTGQLY EMFSSVMKHL PGPQQQAFQL LQGLEDFIAK 250
                                          C  SRS 4
KVEHNQRTLD PNSPRDFIDS FLIRMQEEEK NPNTEFYLKN LVMTTLNLFI 300
                                                  Q
GGTETVSTTL RYGFLLLMKH PEVEAKVHEE IDRVIGKNRQ PKFEDRAKMP 350
                 SRS 5
YMEAVIHEIQ RFGDVIPMSL ARRVKKDTKF RDFFLPKGTE VYPMLGSVLR 400

DPSFFSNPQD FNPQHFLNEK GQFKKSDAFV PFSIGKRNCF GEGLARMELF 450
                                SRS 6
LFFTTVMQNF RLKSSQSPKD IDVSPKHVGF ATIPRNYTMS FLPRHHHHHZ 500
```

Figure 2.16: Amino acid sequences of cytochrome P450 2A6. The six SRS regions are indicated by lines above residues and change in amino acids are in bold face (L240C/N297Q).

Among the six SRS libraries, SRS 3 and SRS 4 produced blue colored colonies that were at least as well as wild-type P450 2A6 and yielded more blue colonies upon replating. Furthermore, SRS 3 and SRS 4 mutants were recombined by staggered extension process (StEP) PCR (Zhao *et al.*, 1998) which created P450 2A6 double mutant (L240C/N297Q). This double mutant was found to be efficient in consistent blue color formation than wild type P450 2A6. Moreover, L240C/N297Q mutant has ability to catalyze wide range of different substituted indole derivatives to form variety of colored products (Figure 2.17) that is not shown by any other isolated mutants (Nakamura *et al.*, 2001). This double mutant has been used in the present study.

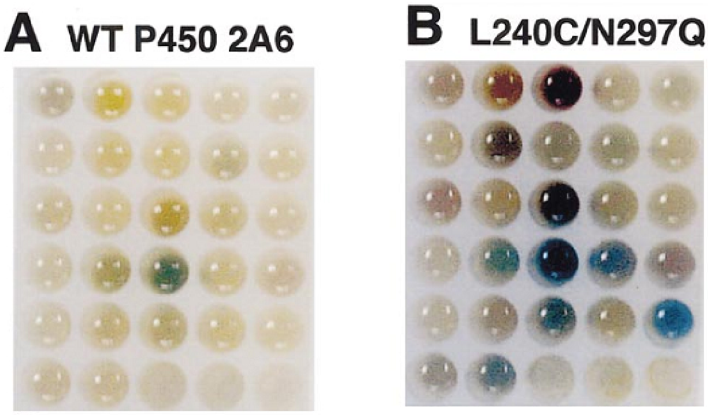


Figure 2.17: Production of indigoid dyes from different indole derivatives in *E. coli* expressing WT P450 2A6 (A) and L240C/N297Q (B) (Courtesy: Nakamura *et al.*, 2001).

2.5.2 Indigoid molecules as protein kinase inhibitors

Protein kinases are considered to be an attractive therapeutic target for treatment of number of diseased conditions like GSK-3 for cancer, type II diabetes, stroke, Alzheimer's disease and other neurodegenerative diseases (Wu *et al.*, 2005). Traditional Chinese medicine Danggui Longhui Wan, used to treat chronic diseases including leukemia (Tang and Eisenbrand, 1992), contains indirubin as one of the active constituents of it. Indirubin is known to act as inhibitor of both glycogen synthase kinase-3 (GSK-3) (Leclerc *et al.*, 2001) and cyclin-dependent kinases (Hoessel *et al.*, 1999) which could be possible mechanism of this medication to exhibit its efficacy.

As already mentioned, cytochrome P450 2A6 catalyzes indole to diindole derivatives, indigo and indirubin (Nakamura *et al.*, 2001) and these indigoids are thought to act as potent agonist of aryl hydrocarbon (Ah) receptor (Adachi *et al.*, 2001). So, due to the potential implication of indigoid derivatives in treatment of numerous diseases, many efforts have been made for synthesis of potent GSK-3 and cyclin-CDK inhibitors. The poor solubility, efficacy and specificity have been primary concern in developing indigoids as potential therapeutic agents thus efforts have been in developing indigoids having desired properties. But still its production is hindered due to ample availability of highly unstable precursor isatins and indoxyls (Polychronopoulos *et al.*, 2004).

Guengerich *et al.* (2004) attempted to generate new protein kinase inhibitors through combinatorial chemistry approach in generating library of indigoids from mutant library of cytochrome P450 2A6 generated by random mutagenesis approach to convert different substituted indole to indigo and indirubin derivatives. In one of the studies, P450 2A6 double mutant (L240C/N297Q) was found to be most versatile in generating diverse indigoid dyes when co-incubated with substituted indole (Nakamura *et al.*, 2001). Using this clone, the culture extracts harboring each indigoid derivative were then screened for inhibition of human cyclin-dependent kinases (CDK5/p25 and CDK1/cyclinB) and glycogen synthase kinase-3 (GSK-3α/β). Among them some of the substituted indirubins contained upto 10 fold more activity than un-substituted indirubin thus attributing the inhibition to substituted moieties of indole (Guengerich *et al.*, 2004).

However, when each potent indigoid were isolated and characterized, it revealed that this bacterial expression system is not much capable of supporting synthesis of wide range of disubstituted indigo and indirubin molecules because of the endogenous production of 1-H indole. In normal bacteria, the tryptophanase (*tna*A) gene product degrade tryptophan to produce 1-H indole (Figure 2.18) (Burns and DeMoss, 1962), which will subsequently compete with added indole to form monosubstituted indigo and indirubin, thus decreasing the product diversity and efficiency (Guengerich *et al.*, 2004; Wu *et al.*, 2005). In order to have specific reactions between different substituted indoles, and generate combinatorial library of different indigoids it is prudent to remove indole production in the cell.

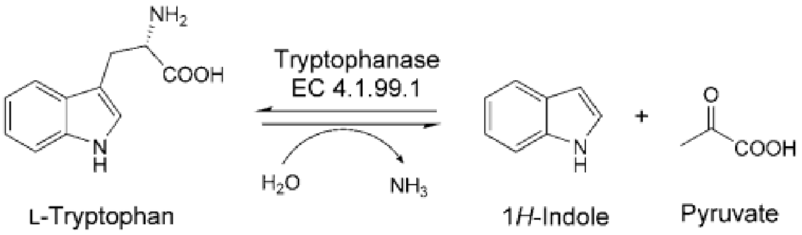


Figure 2.18: Tryptophan degradation catalyzed by *Tryptophanase* producing 1*H*-Indole, ammonia and pyruvate (Adopted from Wu *et al.*, 2005).

2.6 Development of *E. coli* HME5 *tna<>cat* strain

Guengerich and coworkers then developed a host strain *E. coli* HME5 *tna<>cat*, to express P450 2A6 mutants which can support production of library of different disubstituted indigoids (Wu *et al.*, 2005). For this, the native tryptophanase (*tna*) gene of *E. coli* HME5 was replaced by chloramphenicol acetyltransferase (*cat*) gene. (The development of *E. coli* HME5, homologous recombination proficient bacteria is being discussed in following chapters). In tryptophanase deleted mutant *E. coli* strain, the absence of endogenously produced unsubstituted indole allows the bacteria to produce disubstituted indigoids when cultured with two different substituted indoles. This can give all possible combination of substituted indigoids which can then be screened as potential protein kinase inhibitors. Interestingly, disubstituted indigoids with enhanced potency towards inhibition of GSK and CDK/cyclin were isolated (Wu *et al.*, 2005). In addition, this biochemical approach can be used as a paradigm for use of enzyme mutagenesis coupled with combinatorial chemistry for synthesis of libraries of compounds having pharmacological properties of interest (Guengerich *et al.*, 2004; Wu *et al.*, 2005).

2.6.1 Homologous recombination efficient *E.coli* HME5 strain

Normally, *E. coli* do not readily get transformed and recombine with electroporated linear DNA. This lower recombination efficiency is attributed to cellular RecBCD exonuclease activity which rapidly degrades linear DNA (Benzinger *et al.*, 1975). In order to develop homologous recombination efficient *E. coli* in both recA+ or recA- backgrounds, Court and his coworkers integrated λ prophage genes into *E. coli* chromosome (Figure 2.19; Yu *et al.*, 2000). The expression of three λ prophage recombination genes *exo, bet* and *gam,* under temperature sensitive cI-repressor, significantly increases the recombination efficiency. Based on λ prophage molecular mechanism as described for prophage the cloned prophage genes products, Exonuclease (Exo) would degrade linear dsDNA from 5'ends and generate 3'-overhangs (Little JW, 1967), while Beta then would bind to these single stranded DNA protecting it from nuclease attack (Carter and Radding, 1971) and Gam would inhibit host RecBCD exonuclease activity (Murphy KC, 1991; Karu *et al.*, 1975). This flanking 3'-overhangs would then be integrated in chromosomal DNA with homology as part of homologous recombination of cellular dsDNA break repair mechanism. Due to combined effort of these genes, the developed host strain *E. coli* HME5 can proficiently recombine with DNA homologies as short as 30-50 bp at the ends of linear DNA (Yu *et al.*, 2000).

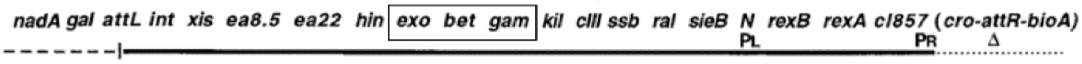


Figure 2.19: The recombination genes (*exo*, *bet* and *gam*) of defective λ prophage on *E. coli* HME5 chromosome. Genes of λ prophage are shown in solid line, and genes of host are shown in broken line. P_L and P_L represent early left and right promoters. The dotted line indicates deletion of *cro-bio*A genes of λ prophage (Courtesy: Yu *et al.*, 2000).

However, λ prophage also consists of lytic gene, but the deletion of *cro* to *bio*A genes (as shown in Figure 2.19) in this defective prophage genome integrated in the chromosomal DNA removes the lytic genes which could have otherwise caused cell lysis within 7 minutes of normal prophage induction (Weisberg and Gallant, 1967). However, some of remaining gene products is also potentially lethal but they require continuous induction for 60 minutes (Kourilsky and Knapp, 1974; Greer H, 1975). Thus, P_L operon consisting of *exo, bet* and *gam* genes, can be selectively induced by shifting the culture from 32^0C to 42^0C for 15 minutes without causing cell lysis and at the same time making cells recombinogenic and ready to take up linear DNA without its destruction (Yu *et al.*, 2000).

The basic strategy of gene replacement technique in *E. coli* HME5 strain by homologous recombination is illustrated in figure 2.20. First the gene to be recombined is amplified by PCR with primers containing at least 20 nucleotides homologous region at 3' end (arrowheads) and 5' end with 30-50 nucleotides (shaded rectangles) identical to sequences of the target DNA (Figure 2.20 A). Thus amplified double stranded DNA is flanked by 30-50 bp homologies present at the target site where the gene is to be recombined. Then the host strain *E. coli* HME5 carrying the target chromosomal region for DNA swapping is induced for Exo, Beta and Gam function by shifting cultures from 37^{0}C to 42^{0}C for 15 minutes. Then it is mixed with amplified cassette and electroporated (Figure 2.20 B). After electroporation, recombination occurs between the homologous sequences on the linear DNA cassette and the target DNA, swapping the target DNA segment with the amplified cassette (Figure 2.20 C).

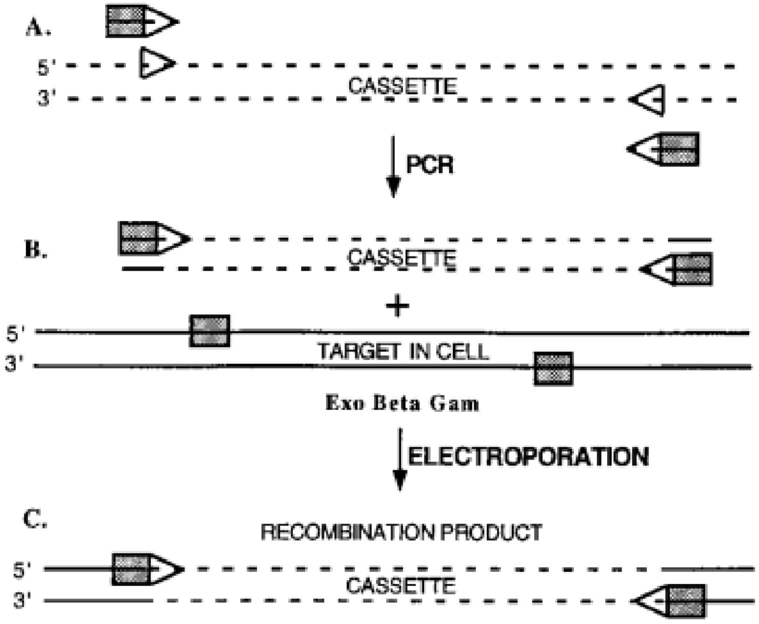


Figure 2.20: General scheme of gene replacement techniques in *E. coli* HME5 expressing λ prophage *exo, bet* and *gam* genes (Adopted from Yu *et al.*, 2000).

2.6.2 Plasmid pACYC184 vector and *cat* gene

Plasmid pACYC184 is *E. coli* based cloning vector originally developed by Chang and Cohen (1978). It is low copy number vector with p15A origin of replication (Selzer *et al.*, 1983; Rose RE, 1988) which can co-exist with plasmids of ColE1 origin compatibility group such as pBR322, pCW and pUC19. It contains two selection markers: chloramphenicol and tetracycline resistance genes. In this plasmid, the chloramphenicol acetyltransferase (*cat*) gene was derived from *Tn*9 (Alton and Vapnek, 1979) and tetracycline resistance gene was obtained from pSC101 plasmid (Bernardi *et al.*, 1984).

2.6.2.1 Promoter of *cat* gene

The detailed study of *cat* gene induction and its promoter sequences have been performed with chloramphenicol acetyltransferase '*cat*-86' gene from gram positive bacteria (*Bacillus subtilis*). The key regulatory mechanism involved in *cat*-86 gene induction has been elucidated to be translational attenuation during *cat* gene expression (Ambulos *et al.*, 1988; Lovett PS, 1990). The study of *cat* promoter sequences revealed the presence of two overlapping but functionally distinct domains (Figure 2.21). Domain A consists of a ribosome-binding site (RBS-2: GAAAGGA) and an open reading frame encoding for small leader peptide of nine amino acids that ends with TGA codon in the upstream inverted-repeat sequence within domain B. Domain B contains two 14 bp inverted repeat sequences that are separated by 12 bp and which spans *cat*-86 ribosome binding site (RBS-3: AGGAGG). When the leader peptide is being translated these inverted repeats tend to form a stable secondary structure sequestering *cat*-86 RBS (Figure 2.21) thus blocking its access to ribosome and severely affecting *cat* gene translation (Rogers and Lovett, 1990).

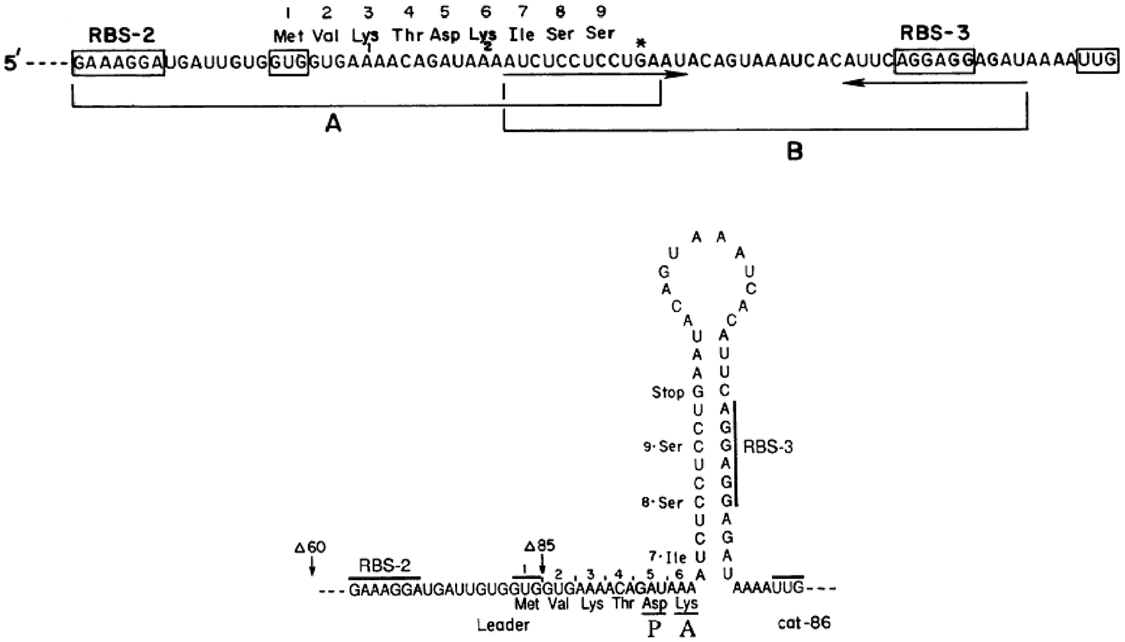


Figure 2.21: The regulatory region of *cat* gene with two domains A and B that is essential for inducible expression of *cat* gene (upper panel) and the stable secondary structure formed by *cat*-86 transcripts (lower panel) sequestering *cat*-86 RBS (Adopted from Rogers and Lovett, 1990).

When chloramphenicol is added, it inhibits protein synthesis and stalls ribosome at precise position in the leader peptide in *cat*-86 mRNA. This then destabilizes the stemloop structure that normally sequesters *cat*-86 RBS (Alexieva *et al.*, 1988) thus facilitating translation of *cat* gene (Duvall and Lovett, 1986) due to changed secondary mRNA structure where RBS is accessible for incoming ribosome. This confirms the importance of stable secondary structure for induction of *cat* gene with chloramphenicol. However, equivalent regulatory sequences are not present in upstream regions of Tn9 *cat* gene which clearly indicates the constitutive rather than inducible expression of *cat* gene present in pACYC184 plasmid (Alton and Vapnek, 1979).

2.6.3 Effect of chloramphenicol on P450 2A6 expression

In the study by Wu *et al.*, (2005) the homologous recombination efficient strain *E. coli* HME5 that was originally developed by Court and coworkers (Yu *et al.*, 2000) was used as host strain to develop bacteria that is devoid of tryptophanase A activity. The constitutive *cat* gene, with its native promoter and terminator sequences, was amplified from pACYC184 vector and integrated into chromosomal DNA replacing tryptophanase A gene with 70 bases of homology. This allowed selection of *E. coli* HME5 *tna<>cat* with chloramphenicol resistance marker and the strain was also tested for absence of endogenous indole production by biochemical test and at genomic level to confirm that tryptophanase A has been deleted and replaced with *cat* gene (Wu *et al.*, 2005).

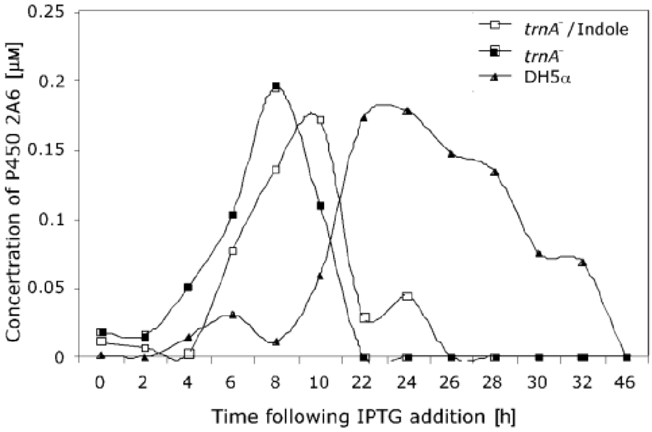


Figure 2.22: Kinetics of expression levels of cytochrome P450 2A6 L240C/N297Q in *E. coli* DH5α and tryptophanase negative *E. coli* HME5 strains (Courtesy: Wu *et al.*, 2005).

The comparison of cytochrome P450 2A6 L240C/N297Q levels in *E. coli* HME5 *tna<>cat* and *E. coli* DH5α strains clearly indicated that there was a positive effect on integration of *cat* gene at *tna*A locus. This could be either due to effect of inhibition of indole formation or effect of basal level of *cat* gene induction on cytochrome P450 2A6 expression (Figure 2.22). The strain has been observed to express highest level of cytochrome P450 2A6 at around 8 hours compared to *E. coli* DH5α F'IQ that produces around at 22 hours, and even in presence of indole *E. coli* HME5 *tnaA<>cat* at around 10 hours (Wu *et al.,* 2005). It is reported that *cat* gene is also known to assist in heterologous protein expression by induction of cold stress response (Kusano *et al.*, 1999; Kagawa *et al.*, 2003). Thus, potentially addition of chloramphenicol in the culture medium could have had positive effect in protein expression and detrimental effect of chloramphenicol could have been circumvented due to cloned *cat* gene at chromosome that degraded chloramphenicol and made the strain resistant. Similar positive effect of chloramphenicol was also observed with different cytochrome P450s expression levels in the study by Harnastai *et al.*, (2006). The correlation between cytochrome P450s expression levels and addition of chloramphenicol can also be explained due to fact that partial inhibition of protein synthesis by chloramphenicol increases the pool of charged glutamyl-tRNA that could potentially increase ALA and ultimately heme biosynthesis (Nakayashiki *et al.*, 1995). Increased pool of heme could facilitate expression of catalytically active cytochrome P450s (holo-protein). This tryptophanaseA deleted mutant host bacteria has been used for present study.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Chemicals and Restriction Enzymes

The sources of chemicals, restriction enzymes, DNA polymerase and other DNA modifying enzymes employed in this study are from following sources: 5-Aminolevulinic acid (ALA), Ribonuclease A and IPTG were purchased from Sigma-Aldrich, USA. Restriction enzymes (*Eco*RI-HF^{TM}, *Nde*I, *Dpn*I and *Hind*III-HF^{TM}), Phusion High-Fidelity (HF) DNA polymerase and T_4 DNA ligase were from New England Biolabs Inc. USA. DNA amplifying and Site directed mutagenesis primers were purchased from Integrated DNA technologies (IDT), USA. Other chemicals and reagents were obtained from commercial suppliers, Himedia and Merck.

3.2 Bacterial Strains and Plasmids

Escherichia coli HME5 *tna<>cat* (Yu *et al.*, 2000 was generous gift from Dr. F. P. Guengerich, Director, Center in Molecular Toxicology and Interim Chair, Department of Biochemistry, Vanderbilt University School of Medicine, TN, USA) was used to isolate chromosomal DNA for *hem*A gene amplification. *E. coli* K12 ER2420/pACYC184 (Chang *et al.*, 1978) and pUC19 plasmid was obtained from New England Biolab Inc. USA. Plasmid pAM101 carrying *Nde*I site (CATATG) under *cat* promoter overlapping ATG start codon of chloramphenicol acetyl transferase (*cat)* gene was developed by site directed mutagenesis PCR of pACYC184 vector. This newly developed expression vector (pAM101) was used for cloning of *hem*A gene and its expression under the regulation of *cat* promoter. The bacterial strain *E. coli* HME5 *tna<>cat* harboring expression plasmid pCW"2A6 L240C/N297Q:hNPR (Parikh *et al.*, 1997; Gillam *et al.*, 1999; Nakamura *et al.*, 2001) was also a generous gift from Dr. Guengerich. *E. coli* DH5α and *E. coli* HME5 *tna<>cat* (Wu *et al.,* 2005) were used as host strains for transformation. The details of development of host bacterial strain and construction of P450 2A6 expression plasmid (pCW"2A6:hNPR) has been discussed in literature review sections. The bacterial strains and plasmids employed in the study are listed in Table 3.1.

Table 3.1: Strains and plasmids used in this study.

Strains and plasmids	Relevant properties	Source
Strains		
E. coli DH5α	F^{-}*endA*1 glnV44 *thi*$^{-1}$*recA*1 *rel*A1 *gyr*A96 *deo RnupG* ϕ80 *dlacZ* ΔM15Δ(*lacZYA-arg*F)U169,*hsd*R17(*rK*$^{-}$ *mK*$^{+}$),*λ*$^{-}$	Lab stock
E. coli HME5 *tna<>cat*	W3110 Δ*lacU169 λcl857* Δ(*cro-bioA*)	Wu *et al.*,2005
E. coli HME5 *tna<>cat*, pCW"2A6: hNPR	*E. coli* HME5 harboring pCW2A6 (L240C/N297Q): hNPR	Wu *et al.*, 2005
E. coli K12 ER2420/pACYC184	*E. coli* K12 harboring pACYC184	NEB
E.coli AM101	*E. coli* DH5α harboring pAM101	This work
E.coli AM102	*E. coli* DH5α harboring pAM102	This work
E. coli AM101/2A6	*E. coli* HME5 *tnaA<>cat* harboring pCW"2A6: hNPR and pAM101	This work
E. coli AM102/2A6	*E. coli* HME5 *tnaA<>cat* harboring pCW"2A6: hNPR and pAM102	This work
Plasmids		
pACYC184	Tet^{R}, Cm^{R}- low copy vector	NEB
pUC19	Amp	NEB
pCW"2A6: hNPR	Amp^{R}, pCW2A6 (L240C/N297Q):hNPR	Nakamura *et al.*, 2001
pAM101	pACYC184::NdeI	This work
pAM102	pAM101 containing *hemA*(KK)	This work
pUC19ΔNdeI	pUC19 with deleted *Nde*I site	This work

3.3 Construction of expression vector pAM101

The plasmid pAM101 was constructed by site directed mutagenesis of vector pACYC184 to create *Nde*I site (CATATG) overlapping ATG start codon of *cat* gene. Thus constructed expression vector pAM101 will facilitate one step cloning of foreign gene under *cat* promoter and the cloned gene will be constitutively expressed at basal level in absence of chloramphenicol and higher level of gene expression can be induced with addition of chloramphenicol.

To construct pAM101, pACYC184 plasmid was amplified using specifically designed site directed mutagenesis primers (SSMX06403SDF and SSMX06403SDR) to create *Nde*I site overlapping *cat* initiation codon. After PCR reaction, the methylated plasmid template DNA was digested with *Dpn*I and remaining DNA was used to transform into *E. coli* DH5α competent cells. The transformants harboring pACYC184 plasmid with inserted *Nde*I site were screened and confirmed by restriction mapping and designated as plasmid pAM101. The preparation and composition of all the reagents and solutions have been described in Appendix 4.

3.3.1 Isolation of pACYC184 plasmid

Plasmid pACYC184 was recovered from *E. coli* K12 ER2420/pACYC184 strain by Alkaline lysis method and purified using HiPurA Silica kit (HiMedia) following manufacturer's protocol.

3.3.1.1 Bacterial culture and cell harvest

Briefly, *E. coli* K12 ER2420/pACYC184 strain was revived by streaking on Luria Bertani (LB)/tetracycline (5 µg/ml) agar plate and incubated overnight at 37^0C. A single isolated colony was inoculated in 3 ml of LB/tetracycline (10 µg/ml) medium and incubated at 37°C for 18 hours at 200 rpm. Then, 1.5 ml of the overnight culture was transferred to sterilized eppendorf tube and centrifuged at 5,000 rpm for 5 minutes at 4°C. The supernatant was poured off and remaining culture (1.5 ml) was added to same tube with cell pellet and again centrifuged at 5,000 rpm for 5 minutes at 4°C. Then supernatant was removed as much as possible without disturbing the cell pellet and the cells were re-suspended in 500 µl of ice cold STE solution and again centrifuged at 5,000 rpm for 5 minutes at 4°C.

3.3.1.2 Cell lysis: Alkaline lysis method

The cell pellet after washing with STE solution was re-suspended in 200 µl of alkaline lysis solution I and mixed well by pipetting. To this solution, 200 µl of freshly prepared alkaline lysis solution II was added and mixed by gentle inversion of tube for three times. The bacterial solution becomes clear as the cell lysis occurs. To this mixture 200 µl of ice cold alkaline lysis solution III was added and mixed by inverting the tube gently for several times. During this step, a cloudy precipitate becomes visible due to precipitation of proteins. The mixture was then centrifuged at 13,000 rpm for 10 minutes at 4°C. Then supernatant solution consisting plasmid DNA was transferred to a fresh eppendorf tube.

3.3.1.3 Plasmid DNA purification by HiPurA Silica kit

To the supernatant containing plasmid DNA, 100 µl of chaotropic salt solution (provided by manufacturer) and 5 µl of glass powder suspension (GPS) was added and left at room temperature for 10 minutes with intermittent mixing. The solution was centrifuged at 10,000 rpm for 1 minute at 4°C. Then supernatant was removed without disturbing the pellet. The pellet was re-suspended by pipetting after addition of 500 µl of diluted wash solution (provided by manufacturer). Then it was centrifuged at 10,000 rpm for 1 minute at 4°C. Then the supernatant was again discarded and the pellet was washed two more times with 100 µl of diluted wash solution and centrifuged as above. Then supernatant was removed and the pellet was allowed to air dry for 5- 10 minutes at room temperature. The pellet was dissolved in 20 µl of MilliQ water by pipetting and incubated at 55^0C in a water bath for 10 minutes. The solution was mixed in every 2-3 minutes to properly extract DNA from silica gel. Then it was centrifuged at 10,000 rpm for 1 minute at room temperature and the supernatant solution containing plasmid DNA was transferred to a fresh eppendorf tube without touching the pellet and stored at -20^0C until use.

3.3.2 Gel Quantification of plasmid DNA

The quantification of isolated plasmid DNA was done by gel electrophoresis of 5 µl of plasmid DNA along with 1kb NEB DNA ladder in 1% agarose gel. The DNA concentration of supercoiled plasmid DNA was determined by comparing the band intensity of sample DNA with known concentration of standard DNA ladder bands.

Briefly, the agarose gel was prepared by boiling 0.2 gm of agarose dissolved in 20 ml of 0.5X TAE buffer. The gel was casted and allowed to solidify. Then, it was placed in gel tank containing 0.5X TAE buffer. Plasmid DNA sample (1µl of 6X NEB DNA loading dye + 5µl of Plasmid DNA sample) was loaded in the well and also 6 µl of 1X NEB 1kb DNA ladder was loaded as the marker in another well. The gel was run for 2 hours at 50 V in Mini Run GE-100, *S.* Korea. After completion of electrophoresis the gel was stained for few minutes in ethidium bromide solution (10 µg/ml) and visualized under Wise Doc WGD 30, China and photographed after properly washing in MilliQ water.

3.3.3 Site directed mutagenesis

3.3.3.1 Primer designing

The site directed mutagenesis primers were designed according to general guidelines provided by Liu and Naismith (2008) and using the nucleotide sequences of pACYC184 deposited in GenBank (Accession number: X06403.1) and information provided by NEB-technical references (Appendix 6). The basic properties of mutagenesis primers are presented in Appendix 1.1.

The protocols developed by Liu and Naismith (2008) describes the novel strategy of site directed mutagenesis that preserved the simple one step procedure of the Stratagene QuikChange^{TM} site-directed mutagenesis but with modified primer designing technique that enhances efficiency to incorporate desired mutation. Each primer set contains primer-primer non-overlapping sequences at their 3′ end while primer-primer complementary (overlapping) sequences at the 5′ end as shown in figure 3.1. The desired mutation (addition/deletion) is usually placed at the centre of overlapping region. The non-overlapping sequences are extended than priming region because the melting temperature of non-overlapping sequences ($T_{m\ no}$) is 5 to 10^{0}C greater than melting temperature of primer-primer complementary sequences ($T_{m\ pp}$). This new primer designing scheme facilitates the primer annealing with template by eliminating primer-primer dimerization as observed in Stratagene QuikChange^{TM} protocol and even promote exponential increase in PCR product by enabling primers to use the PCR product as the template which is not feasible with QuikChange mutagenesis protocol (Liu and Naismith, 2008).

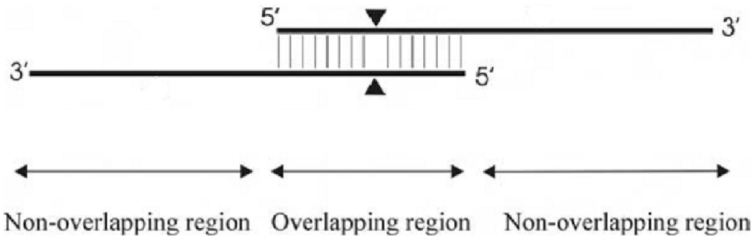


Figure 3.1: The scheme of primer designing used to create *Nde*I site (triangle) in plasmid pACYC184 showing priming-priming (overlapping) and non-overlapping regions.

3.3.3.2 Site directed mutagenesis primer

SSMX06403SDF: 5′ **TTCTC CATATG AGCT TCCT** TAG CTCCTG AAA ATC TCGATAAC 3′

SSMX06403SDR: 5′ **AGGAAGCT CATATG GAGAA** AAAAATCACTGGATATACCACCG TTG 3′

Key: Bold nucleotides (**ATGA): 5′** primer-primer complementary region
Other nucleotides (TAGC): 3′ non-overlapping region
Highlighted nucleotides (**AGTC**): Change in nucleotide

Underlined Nucleotide (**CATATG**): *Nde*I restriction site

3.3.3.3 Site directed PCR

The reaction mixture and PCR conditions for site directed mutagenesis were optimized with slight modification of protocols provided by Liu and Naismith (2008). The unique Phusion High-Fidelity DNA polymerase (NEB) was applied for error-proof DNA amplification process. This enzyme offers robust performance with high speed and accuracy, having an error rate 50-fold lower than that of *Taq* DNA polymerase and 6-fold lower than that of *Pyrococcus furiosus* DNA polymerase.

The melting temperatures of primer-primer overlapping and non-overlapping regions were calculated from IDT: Oligo analyzer 3.1. For first 20 PCR cycles, $T_{m\ no}$– 5^0C was used as annealing temperature and another 4 cycles were run at $T_{m\ pp}$ –5^0C with additional one cycle at 50^0C to increase the PCR efficiency. The reaction mixture and PCR conditions has been presented in Appendix 1.1.

After completion of PCR reactions, 5 µl of PCR product was subjected to 1% agarose gel electrophoresis along with 1 kb NEB DNA ladder in separate well and PCR amplification was confirmed by observing linear plasmid DNA band of 4245 bp. Then 3 µl of *Dpn*I (6U) was directly added to the remaining PCR product (45 µl) and incubated overnight at 37°C to digest methylated template DNA. Then after purification of PCR product that effectively removes cleaved methylated DNA, enzymes and buffer used, PCR amplified plasmid DNA was concentrated by ethanol precipitation.

Briefly, 135 µl of 100% ethanol (3 volume of PCR product) was added to the solution and incubated at -40^0C for 30 minutes. The mixture was centrifuged at 13,000 rpm for 15 minutes at 4^0C and the supernatant was discarded. The precipitate containing DNA was washed with 1 ml of 70% ethanol without disturbing the pellet and again centrifuged at 13,000 rpm for 2 minutes at 4°C. The supernatant was discarded and tube was left for 5-10 minutes at room temperature to air dry DNA sample. Then the precipitated DNA was re-suspended in 10 µl of MilliQ water and used for transformation of *E. coli* DH5α competent cells.

3.3.4 Preparation of *E. coli* DH5α competent cells

E. coli DH5α competent cells were prepared by Rubidium chloride (RuCl) method. First, the bacterial strain was revived by streaking on LB agar plate and incubated overnight at 37^0C. After incubation, a single isolated colony was inoculated in 2 ml of LB medium and incubated at 37^0C for 18 hours at 200 rpm. Then 20 µl of overnight culture was 100 times diluted in 2 ml of SOC medium and incubated at 37^0C with shaking at 200 rpm for an hour or until OD_{600} reaches 0.4-0.6. The bacterial cells were pelleted by centrifuging at 10,000 rpm for 5 minutes at 4^0C. Then the cells were re-suspended into 800 µl of ice cold RF1 solution (Note: add 0.4 volume of RF1 to original volume of culture) and incubated in ice bath for 15 minutes. Again the cells were collected by centrifuging at 10,000 rpm for 5 minutes at 4^0C. The supernatant was discarded and cell pellet was re-suspended in 80 µl of ice cold RF2 solution (Note: add 0.04 volume of RF2 to original volume of culture) and incubated in ice bath for 15 minutes. Thus prepared *E. coli* DH5α competent cells were ready for transformation or stored at -40^0C until use.

3.3.5 Transformation

Purified *Dpn*I digested and ethanol precipitated PCR amplified plasmid DNA sample was used for transformation of *E. coli* DH5α competent cells by heat shock method. Briefly, 10 µl of plasmid DNA was added to 80 µl of *E. coli* DH5α competent cells, mixed well and placed in ice bath for 30 minutes. Then it was subjected to heat shock by keeping in water bath at 42^{0}C for 2 minutes and immediately transferred to ice bath and kept for 10 minutes. Then 1 ml of SOC medium was added to the transformed competent cells and incubated at 37^{0}C for 90 minutes keeping the tube perpendicular and slightly shaking (100 rpm) to stabilize the plasmids in transformed cells. After incubation, 100 µl of transformed cells were directly plated into LB/tetracycline (5 µg/ml) agar plates whereas remaining cells were centrifuged at 10,000 rpm for 10 minutes at 4^{0}C and cells were re-suspended in 100 µl of sterilized ice cold phosphate buffer (pH 7.0) and spread in LB/tetracycline (5 µg/ml) agar plates.

Both plates were incubated at 37^{0}C for overnight. Then few of the isolated tetracycline resistant transformed colonies were selected to inoculate 3 ml of LB/tetracycline (10 µg/ml) medium and incubated as described previously.

3.3.6 Confirmation of Transformants

3.3.6.1 Plasmid Extraction

Plasmid isolation from the transformed colonies was performed as described previously by Alkaline lysis method. During plasmid extraction, 30 µl of RNase solution (1 mg/ml) was also added after addition of 200 µl of lysis solution II. The isolated plasmid DNA was purified using chilled phenol-chloroform method and recovered after ethanol precipitation as described below.

A. Plasmid DNA purification: Chilled phenol-chloroform method

To the supernatant containing plasmid DNA, equal volume of chilled phenol: chloroform (1:1 v:v) mixture was added and vigorously vortexed for three times. Then the mixture was centrifuged at 13,000 rpm for 2 minutes at 4^{0}C and aqueous solution was collected in fresh eppendorf tube. Again, to this solution equal volume of chloroform was added and centrifuged at 13,000 rpm for 2 minutes at 4^{0}C. Then aqueous layer containing DNA was carefully collected in fresh eppendorf tube avoiding carryover of lower organic phase.

B. Recovery of plasmid DNA

To the aqueous solution (500 µl), 50 µl (1/10th volume of aqueous solution) of 3 M Chilled sodium acetate (pH 5.2) and equal volume of isopropanol (550 µl) was added and mixed well by inverting the tubes several times. Then the mixture was centrifuged at 13,000 rpm for 10 minutes at 4^0C. The supernatant was discarded and DNA pellet was washed with 1 ml of 70% ethanol without disturbing pellet and again centrifuged at 13,000 rpm for 2 minutes at 4^0C. Then the supernatant was immediately removed as much as possible and the tube was left at room temperature to air dry for 5-10 minutes. Finally, plasmid DNA was dissolved in 50 µl of MilliQ water and stored at -20^0C until use.

3.3.6.2 Restriction digestion

Thus recovered plasmid DNA was digested with restriction enzymes to verify the creation of *Nde*I site in pACYC184 vector by site directed mutagenesis PCR. In this experiment, pACYC184 vector was used as control plasmid. The different restriction digestion mixtures were prepared as following.

Restriction digestion with *Nde*I

Plasmid DNA	5 µl
*Nde*I (3U/ µl)	2 µl
NEB buffer 4 (10X)	2 µl
MilliQ water	11 µl
Total	20 µl

Restriction digestion with *Nde*I and HindIII

Plasmid DNA	5 µl
*Nde*I (3U/ µl)	2 µl
HindIII-HF™ (5U/ µl)	1 µl
NEB buffer 4 (10X)	2 µl
MilliQ water	10 µl
Total	20 µl

Digestion of pACYC184 with *Eco*RI-HF^{TM}

Plasmid pACYC184	5 µl
*Eco*RI-HF^{TM} (3U/ µl)	2 µl
NEB buffer 4 (10X)	2 µl
MilliQ water	11 µl
Total	20 µl

All the restriction digestion mixtures were incubated at 37^{0}C for 2 hours and run in 1% agarose gel. After gel electrophoresis the band pattern was observed and the presence of *Nde*I site at correct position was verified. Then the cell which harbor the correct plasmid was streaked in LB/tetracycline (5 µg/ml) agar plate. The single isolated colony was then sub-cultured in 8 ml of LB/tet (10 µg/ml) medium and 2 ml of the overnight culture was used to isolate plasmid DNA as described previously by Alkaline lysis and chilled phenol-chloroform method. The plasmid was re-confirmed by *Nde*I digestion and thus obtained plasmid was designated as plasmid pAM101. The remaining culture (6 ml) harboring pAM101 was designated as *E. coli* AM101 and was mixed with 4 ml of 40% glycerol (15% glycerol final concentration) and dispensed 2 ml aliquots in each 5 sterilized cryo vials and stored at -40^{0}C after proper labeling.

3.3.6.3 Functional assay of plasmid pAM101

In order to assess the integrity of *cat* gene in newly constructed plasmid pAM101 with inserted *Nde*I site under *cat* promoter, *E. coli* DH5α harboring pAM101, *E. coli* K12 ER2420/pACYC184 (positive control) and *E. coli* DH5α (negative control) were streaked in LB/chloramphenicol (25 µg/ml) agar plates and incubated at 37^{0}C for 24 hours. The strain that survived in plate containing chloramphenicol that indicated the creation of *Nde*I site without disrupting *cat* gene expression was isolated and used for further works.

3.4 Construction of expression plasmid pAM102

The chromosomal DNA isolated from *E. coli* HME5 *tna<>cat* strain was used as template DNA for *hem*A gene amplification using specifically designed primers SSMPOA6X1F01 and SSMPOA6X1R01. The plasmid pAM102 was constructed by sub-cloning *hem*A gene into the plasmid vector pAM101 previously digested by *NdeI* and *Eco*RI. The cloning of *hem*A gene in pAM101 under *cat* promoter would disrupt chloramphenicol resistance marker but for high level induction of *hem*A gene addition of chloramphenicol would be necessary. The host strain *E. coli* HME5 *tna<>cat* (Wu *et al.*, 2005) has chloramphenicol resistance gene integrated into its chromosomal DNA which makes the strain resistant to chloramphenicol and making it suitable host strain for expression of *hem*A gene cloned under *cat* promoter.

PCR amplified *hem*A gene was completely digested with *Eco*RI and partially digested with *Nde*I due to presence of internal *Nde*I site in the coding region of *hem*A gene. After gel extraction of correct DNA insert (1.3 kb), the DNA was ligated in plasmid vector pAM101 previously digested with *EcoRI* and *NdeI* and gel purified. The detail of construction of plasmid pAM102 is described and the protocols employed are described below. The compositions of reagents are presented in Appendix 4.

3.4.1 Genomic DNA isolation

3.4.1.1 Preparation of cell

E. coli HME5 *tna<>cat* strain was revived and streaked on LB/chloramphenicol (10 µg/ml) plate and incubated overnight at 37^{0}C. Then, a single isolated colony was inoculated in 2 ml of LB medium and incubated at 37°C for 18 hours at 200 rpm. From this overnight culture 1.5 ml of the culture was transferred to sterilize eppendorf tube and centrifuged at 5,000 rpm for 5 minutes at 4°C. Immediately, supernatant was discarded by aspiration and remaining overnight culture (0.5 ml) was also added to same tube containing the cell pellet and centrifuged again at 5,000 rpm for 5 minutes at 4°C. Then, supernatant was removed as much as possible without disturbing the cell pellet.

3.4.1.2 Cell Lysis

The cell pellet was re-suspended in 450 µl of TE1 buffer by gentle pipetting. The solution was split into two fresh sterilized eppendorf tubes by transferring 225 µl of above suspension to each tube. To each tube 180 µl of lysozyme (1 mg/ml) was added. Both tubes were incubated at 37°C for 30 minutes gently mixing the solution by inverting the tube every 5 minutes for proper cell lysis. Then 45 µl of STEP solution was added in both tubes and incubated in ice for 45 minutes or until the solution became clear due to cell lysis, with gentle inversion in between the ice incubation period. Equal volume of chilled phenol (450 µl) was added and mixed by vortexing. The mixture was centrifuged at 13,000 rpm for 10 minutes. The upper aqueous layer containing DNA was transferred to fresh sterilized eppendorf tube without carryover of lower organic phase. Again, equal volume of chilled phenol: chloroform: isoamyl alcohol (25:24:1) was added to above aqueous solution, mixed by vortexing and centrifuged at 13,000 rpm for 10 minutes at 4°C. After collection of aqueous layer in a fresh tube, equal volume of chloroform was added and vortexed. The mixture was then centrifuged at 13,000 rpm for 2 minutes and the aqueous phase was collected in a fresh eppendorf tube.

3.4.1.3 Genomic DNA recovery

To the aqueous solution (450 µl) containing genomic DNA, 100 µl of 3M chilled sodium acetate (pH 5.2) and double volume of 95% ethanol (1,100 µl) was added and incubated at -20°C for 30 minutes. Then the mixture was centrifuged at 13,000 rpm for 20 minutes at 4°C. The supernatant was poured off and pellet was washed with 250 µl of 70% ethanol without disturbing the pellet. Then the solution was centrifuged at 13,000 rpm for 10 minutes at 4°C. After draining the supernatant, remaining ethanol was removed by keeping the tube open in room temperature for 5- 10 minutes. Care was taken not to over dry the DNA pellet. The genomic DNA was re-suspended in 100 µl of autoclaved MilliQ water or TE buffer (pH 8.0) and stored at -20^{0}C until use.

3.4.2 Genomic DNA Quantification

The genomic DNA quantification and purity assessment was done by UV-visible Spectrophotometer. The absorbance of the sample was measured at 260 nm and 280 nm wavelengths. The concentration of genomic DNA was calculated using the following formula

DNA concentration (µg/ml) = A_{260} X 50 X Dilution factor

A_{260} of 1 corresponds to approximately 50 µg/ml of DNA .

DNA purity = A_{260} / A_{280}

Where, A_{260} and A_{280} are Absorbance of DNA or RNA at wavelength 260 nm and protein at wavelength 280 nm, respectively.

Pure DNA sample has A_{260} / A_{280} ratio equals to 1.8, the ratio above 1.8 indicates RNA contamination whereas the ratio below 1.8 suggests protein contamination in the DNA sample.

3.4.3 Amplification of *hem*A gene

3.4.3.1 Primer designing

The *hem*A gene amplification primers were designed manually using different web based tools discussed below and using the nucleotide sequence deposited in the GenBank for *E. coli* K-12 MG1655 strain (NC_000913.2). The gene was amplified from its coding region with inserted *Nde*I (CATATG) site overlapping the *hem*A ATG initiation codon and introducing two lysine residues (AAA AAA) between second and third amino acids in the forward primer to overcome the heme regulated proteolysis. Reverse primer was designed downstream of stop codon of *hem*A gene that would amplify the gene few bases downstream at 3'-end of the coding sequences to meet the general criteria's of primer designing regarding length, melting temperature (T_m), GC content, secondary structure, self and hetero dimer, etc. BLAST- search (non-redundant database in NCBI) was performed for each primer to ensure that the designed primers would not prime to non-specific regions. The basic information of designed primers is presented in Appendix 1.2.

Bioinformatics Tools used for Primer Designing

a. NCBI for whole genome and genome map of *E. coli*
   http://www.ncbi.nlm.nih.gov/genome
   http://www.ncbi.nlm.nih.gov/guide/genomes-maps/

b. Integrated DNA technologies; Oligo analyzer 3.1 for Tm, hetero/homo dimer
   http://eu.idtdna.com/analyzer/applications/oligoanalyzer/default.aspx

c. The mfold Web server- DNA folding form for folding structure of primer DNA
   http://mfold.rna.albany.edu/?q=mfold/dna-folding-form

d. NCBI- Standard Nucleotide BLAST
   http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?PAGE=Nucleotides

e. New Englands Biolabs webtool: NEB cutter V2.0 for restriction map analysis of amplicon
   http://tools.neb.com/NEBcutter2/cutshow.php?name=35a4abff-

3.4.3.2 *hem*A gene amplification primers

SSMPOA6X1F01: CTAC**GAATTC**TCCAAA **CATatg** acc ***AAA AAA*** ct c tta gca ctc gg
*Eco*RI *Nde*I
SSMPOA6X1R01: AA ACC ACT **GAATTC** CTG ATG **TCTAga** tgt act gct act cc
*Eco*RI *Xba*I

3.4.3.3 Primer Dilution

Primers were purchased from Integrated DNA technologies (IDT) in lyophilized form. The primers were provided in standard desalting purification process and used without any further purification. The freezer and working stock solutions of primers were prepared as described below.

First, the lyophilized primer containing vials were briefly centrifuged for few minutes before opening the vial for first time to avoid loss of the DNA pellet. Then 1 mM stock solution was prepared by adding a volume of TE buffer (pH 8.0) equal to the number of nano moles of DNA present in the tube. For example, 77.4 nMoles of oligo was dissolved in 77.4 µl of TE buffer (pH 8.0) to make a stock 1 mM solution. Similarly, 50 µM stock solutions were prepared by adding 5 µl from 1 mM stock solution in 95 µl of TE buffer (pH 8.0). Finally 10 µM working stock solution was prepared by adding 10 µl of 50 µM stock solution in 40 µl of MilliQ water. For long term storage, 1 mM and 50 µM stock solutions were stored at -40^{0}C and working solution of 10 µM was stored at -20^{0}C.

3.4.3.4 PCR amplification of *hem*A gene

The chromosomal DNA sample isolated from *E. coli* HME5 *tna<>cat* strain was used as the template DNA for PCR amplification of *hem*A gene using specifically designed forward (SSMPOA6X1F01) and reverse (SSMPOA6X1R01) primers. The DNA sample was quantified spectrophotometrically and diluted in MilliQ water to prepare template DNA with concentration of 10-20 ng/µl. PCR amplification was performed with Phusion High-Fidelity (HF) DNA polymerase, NEB. The general protocol provided by NEB technical reference was followed to prepare the reaction mixtures and to determine PCR conditions and are presented in Appendix 1.2.

The two step reaction conditions was used for *hem*A gene amplification. For the first 5 PCR cycles, the annealing temperature targeted to only priming region (3'-end nucleotide complimentary to template DNA) was used whereas for rest 25 PCR cycles, annealing temperature (T_m– 5^{0}C) was determined regarding the whole primer sequence. This strategy increases overall PCR efficiency and yield of PCR product. PCR was performed using Thermal cycler: Bioer-XP technology, China. Then 5 µl of PCR product was subjected to 1% agarose gel electrophoresis along with 1 kb NEB DNA ladder in a separate well as described previously.

3.4.4 PCR product purification

After confirmation of correct amplicon size (1.3 kb DNA fragment) from gel electrophoresis, PCR product was purified by chilled phenol-chloroform method to remove Phusion DNA polymerase and buffers that might interfere with subsequent restriction digestion and ligation steps.

Briefly, 155 µl of MilliQ water was added to the remaining PCR product (45µl) to increase the final volume to 200 µl. Then equal volume of chilled phenol: chloroform (1:1 v/v) solution (200 µl) was added, briefly vortexed and centrifuged at 13,000 rpm for 2 minutes at 4^0C. The aqueous layer was carefully collected in a fresh eppendorf tube avoiding carry off of lower organic phase. Equal volume of chloroform (200 µl) was added to this aqueous solution and centrifuged at 13,000 rpm for 2 minutes at 4^0C. Again the aqueous layer containing DNA was collected in a fresh tube. To this solution, 25 µl of 3M Chilled sodium acetate (pH 5.2) (1/10th volume of aqueous solution) and equal volume of isopropanol (250 µl) was added and incubated at -20^0C for 30 minutes. Then the DNA was recovered by centrifuging the mixture at 13,000 rpm for 15 minutes at 4^0C and the pellet was washed with 1 ml of 70% ethanol without disturbing the pellet. The mixture was then centrifuged at 13,000 rpm for 2 minutes at 4^0C and supernatant was removed and allowed to air dry for 5-10 minutes until alcohol was evaporated. Then the precipitated DNA was re-suspended in 20 µl of MilliQ water and used for restriction digestion.

3.4.5 Restriction digestion of *hem*A gene

The cloning of *hem*A gene digested with *Nde*I and *Eco*RI into pAM101 vector will ensure correct positioning of *hem*A gene under *cat* promoter overlapping ATG initiation codon of *hem*A gene instead of chloramphenicol acetyl transferase start codon. When *hem*A gene was analysed by NEB Cutter tool (Appendix 3), *Nde*I site was found to be present in internal coding sequence (385 bp from *hem*A start codon). Thus, *hem*A gene was completely digested with *Eco*RI followed by partial digestion with *Nde*I.

3.4.5.1 Complete digestion with *Eco*RI

DNA solution	20 µl
*Eco*RI-HF™ (5 U/ul)	1.5 µl
NEB buffer 4 (10X)	3 µl
MilliQ water	5.5 µl
Total	30 µl

The reaction mixture was incubated at 37^{0}C for 4 hours and proceeded for partial digestion with *Nde*I as described.

3.4.5.2 Partial digestion with *Nde*I

Once the digestion with *Eco*RI was completed, 63 µl of MilliQ water and 7 µl of 10X NEB buffer 4 was added to make the final volume to 100 µl. This tube was labeled as "A" and placed on ice. All the works from now for serial dilution was done on ice bath. Then 4 more eppendorf tubes were taken and labeled as B, C, D and E. From tube "A", 20 µl solutions was transferred to 3 tubes "B", "C" and "D" each and 10 µl in tube "E". After this aliquoting 30 µl of reaction mixture was left on tube "A" and all the tubes were kept on ice. Only to tube labeled "A", 1 µl of *Nde*I (20 U/µl) was added and mixed well. This solution was then serially diluted to rest of the tubes by transferring 10 µl from tube "A" to tube "B", 10 µl from tube "B" to tube "C" and so on up to tube "E". Every time the solution was mixed well before transferring to the next tube and now each tubes contained 20 µl of reaction mixture were placed on ice. Then all tubes were incubated at 37^{0}C for 4 minutes and immediately placed back in ice bath. The restriction digestion was stopped by adding 60 µl of chilled phenol: chloroform (1:1 v:v) solution in each tube and then pooled out in one tube. The volume of aqueous solution was increased upto 250 µl by adding 150 µl MilliQ water and the mixture was centrifuged at 13,000 rpm for 2 minutes at 4^{0}C. The aqueous layer was collected in fresh tube and mixed with equal volume of chloroform and again centrifuged at 13,000 rpm for 2 minutes at 4^{0}C. Then the upper aqueous layer containing DNA was transferred to the fresh tube and DNA was recovered by isopropanol precipitation method.

To the collected solution, 25 µl of 3M chilled sodium acetate (pH 5.2) and 270 µl of isopropanol was added and mixed well by inverting tubes several times. DNA was recovered by centrifuging the mixture at 13,000 rpm for 15 minutes at 4^{0}C and the pellet was washed with 1 ml of chilled 70% ethanol. The mixture was then centrifuged at 13,000 rpm for 2 minutes at 4^{0}C and supernatant was removed and allowed to air dry for 5-10 minutes until alcohol gets evaporated. Then the precipitated DNA was re-suspended in 15 µl of MilliQ water and subjected to 1% low melting agarose gel electrophoresis to separate the desired band of 1.3 kb from completely digested insert DNA fragments.

3.4.6 1% Low melting Agarose Gel electrophoresis

For gel electrophoresis and to recover proper size digested DNA containing *hem*A gene, 1% low melting agarose gel was prepared by boiling 0.2 gm of low melting agarose (HiMedia) dissolved in 20 ml 0.5X TAE buffer. Once the gel was cooled to about 50^0C-60^0C, 1 µl of EtBr solution (5 mg/ml) was added and gel was casted. Then it was placed in gel tank containing 0.5X TAE buffer. The restriction digested DNA solution pooled to 15 µl that consisted insert DNA sample was mixed with 10X gel loading buffer and was loaded in a well. In another well 6 µl of 1X NEB 1kb DNA ladder was loaded and run for nearly 2 hours at 50V. The gel was visualized under Wise DOC and correct size insert (1.3 kb) DNA band was cut with clean knife leaving other fragments in the gel for gel extraction.

3.4.7 DNA extraction from low-melting gel

DNA recovery efficiency from low melting gel is greater than from high melting gel. Thus, multiple DNA bands of partially digested insert DNA sample was separated using low melting agarose gel as above and the DNA extraction protocols employed were as described below.

Briefly, the weight of gel was measured and approximately 1.5 volume of TE buffer (pH 8.0) was added (For example, for 100 mg of agarose gel, 150 µl of TE buffer, pH 8.0 was added). The gel was melted by boiling the gel at 65^0C for 5-10 minutes. Then, equal volume of phenol (250 µl) was added to the solution, briefly vortexed and centrifuged at 13,000 rpm for 10 minutes at 4^0C. The aqueous layer was carefully collected in fresh eppendorf tube avoiding carry off of lower organic phase and DNA was recovered by alcohol precipitation. The solution was mixed with 25 µl of 3M chilled sodium acetate (pH 5.2), 1 µl of glycogen (2 mg/ml) and 280 µl isopropanol solution. The mixture was incubated at room temperature for few minutes and centrifuged at 13,000 rpm for 15 minutes at 4^0C. The pellet was washed with 1 ml of chilled 70% ethanol and again centrifuged at 13,000 rpm for 2 minutes at 4^0C. Ethanol was evaporated by allowing the tube to air dry for 5-10 minutes. Then recovered DNA was re-suspended in 10 µl of MilliQ water and directly used for ligation reaction.

3.4.8 Preparation of Vector pAM101

3.4.8.1 Restriction digestion

Plasmid pAM101 was used as expression vector to clone *hem*A gene. The verified plasmid pAM101 was digested with *Nde*I and *Eco*RI and the reaction mixture was prepared as following.

Plasmid DNA	34 µl (250 ng)
*Nde*I (5U/µl)	1 µl
*Eco*RI-HF™ (5U/µl)	1 µl
NEB buffer 4 (10X)	4 µl
Total	40 µl

The reaction mixture was incubated in water bath set at 37^{0}C for 4 hours. Then the digested linear plasmid DNA was subjected to electrophoresis in 1% agarose gel as described previously. The vector DNA sample corresponding to 4.0 kb was recovered and purified from agarose gel using HiPurA™ Agarose gel DNA purification spin kit (HiMedia, Mumbai, India) following manufacturer's instructions as described below.

3.4.8.2 DNA extraction from gel

The band corresponding to ~ 4 kb DNA fragment was cut with clean knife and the gel was put into eppendorf tube and the weight of the gel was measured. Then Gel binding buffer (HG) was added three times the weight (v/w) and incubated at 50^{0}C for 5-10 minutes until all the gel gets properly dissolved. For example: 300 µl of gel binding buffer was added to 100 mg gel slice. One gel volume (290 µl) of isopropanol was added to the solution, and mixed well and 700 µl of this mixture was loaded into the HiElute Miniprep Spin column. It was incubated at room temperature for 2 minutes and then centrifuged at 13,000 rpm for 1 minute. The flow through was saved in separate eppendorf tube and the remaining sample was loaded in the same spin column and again centrifuged at 13,000 rpm for 1 minute. The column was washed with 700 µl of diluted wash solution and centrifuged at 13,000 rpm for 1 minute. The flow through was discarded and the column was placed in same tube and again it was centrifuged at 13,000 rpm for 1 minute to remove excess ethanol. The column was transferred then to fresh collection tube and loaded with 50 µl of elution buffer. The column was left at room temperature for 1 minute and then centrifuged at 13,000 rpm for 1 minute. The elute solution containing linear vector DNA was stored at -20^{0}C until use.

3.4.9 Sub-cloning of *hem*A in plasmid pAM101

The *hemA* gene recovered after complete digestion with *Eco*RI and partial digestion with *Nde*I was sub-cloned in plasmid pAM101, previously digested with same restriction enzymes, employing T4 DNA ligase (NEB). For the ligation reaction, vector: insert molar ratio was kept within range of 1:3 to 1:10 according to known quantity of vector and insert DNA. The concentration of vector and insert DNA was determined by gel quantification against the known amount of standard DNA ladder. The amount of 4 kb vector DNA required to be added to 25 ng of 1.3 kb insert (*hemA*) DNA with molar ratio of 1:5 was calculated using following formula.

$$\text{ng of vector to be added} = \frac{\text{ng of insert X kb size of vector}}{\text{kb size of insert}} \text{ X Molar ratio of vector: insert}$$

Ligation reaction mixture

*hem*A (insert)	10 µl	(25 ng)
pAM101 (vector)	5 µl	(15 ng)
T4 DNA ligase (1U/µl)	3 µl	
10X T4 DNA ligase buffer	2 µl	
Total	20 µl	

The ligation mixture was incubated at 16^{0}C for overnight and the whole ligation mixture was used for transformation of *E. coli* DH5α competent cells by heat shock method as described previously.

3.4.10 Confirmation of Transformants

3.4.10.1 Plasmid Extraction

The newly transformed *E. coli* DH5α cells with ligation mixture was plated in LB/tetracycline (5 µg/ml) agar plate. The cloning of *hem*A gene in pAM101 vector would disrupt chloramphenicol resistance gene so making the bacteria sensitive to chloramphenicol. The tetracycline resistant transformant colonies were also streaked in LB/chloramphenicol (25 µg/ml) agar plate to confirm that the resistance to tetracycline was not of pAM101 but of pAM102. Then, the tetracycline resistant and chloramphenicol sensitive colonies were sub-cultured in 2 ml of LB/tet (10 µg/ml) medium and incubated at 37^{0}C with shaking at 200 rpm. The plasmid was isolated as described previously by Alkaline lysis method, purified by chilled phenol-chloroform mixture and recovered after ethanol precipitation.

3.4.10.2 Restriction digestion

Thus recovered plasmid DNA was digested with restriction enzymes to verify sub-cloning of *hem*A gene in plasmid pAM101. The different restriction digestion mixtures were prepared as following.

Restriction digestion with *Eco*RI-HF^{TM}

Plasmid	8 µl (40 ng)
*Eco*RI-HF^{TM} (5U/ µl)	1 µl
NEB buffer 4 (10X)	2 µl
MilliQ water	9 µl
Total	20 µl

Restriction digestion with *Nde*I and *Eco*RI-HF^{TM}

Plasmid	8 µl (40 ng)
*Eco*RI-HF^{TM} (5U/µl)	1 µl
*Nde*I (5U/µl)	1 µl
NEB buffer 4 (10X)	2 µl
MilliQ water	8 µl
Total	20 µl

All the restriction digestion mixtures were incubated at 37^{0}C for 2 hours and electrophoresed in 1% agarose gel as described previously. The positive transformant harboring *hem*A gene sub-cloned in pAM101 vector was confirmed by restriction mapping. Then the positive transformant was streaked in LB/tetracycline (5 µg/ml) agar plate and a single isolated colony was re-cultured in 8 ml of LB/tet (10 µg/ml) medium. From this overnight culture, 2 ml was used to isolate plasmid DNA as described previously and restriction digested with *Eco*RI to re-confirm the sub-cloning of *hem*A gene in pAM101. Thus confirmed plasmid DNA was designated as plasmid pAM102. The remaining culture (6 ml) was mixed with 4 ml of 40% glycerol and dispensed 2 ml aliquots in each 5 sterilized cryo vials and stored at -40^{0}C after proper labeling.

3.5 Development of *E. coli* AM102/2A6 and AM101/2A6

E. coli HME5 tna<>*cat* strain harboring plasmid pCW"2A6 L240C/N297Q:hNPR was revived and streaked on LB/ampicillin (50 µg/ml) agar plate and incubated overnight at 37^{0}C. Then, a single isolated colony was inoculated in 2 ml of LB/amp (50 µg/ml) medium and incubated at 37°C for 18 hours at 200 rpm. The bacterial culture was 100 fold diluted in SOC/ampicillin (25 µg/ml) medium to prepare competent cells as described previously by Rubidium Chloride method. Then, the competent cells of *E. coli* HME5 *tna<>cat* harboring pCW''2A6 L240C/N297Q:hNPR were transformed with 5 µl (~50 ng) of plasmid pAM102 and pAM101, respectively by heat shock method (as described) and the resulting ampicillin (50 µg/ml) and tetracycline resistant (5 µg/ml) colonies were selected. The strain *E. coli* HME5 *tna <>cat* harboring pCW''2A6 L240C/N297Q:hNPR and pAM102 was designated as *E. coli* AM102/2A6 and used for heterologous expression of functionally cytochrome P450 2A6 supported by *hem*A gene co-expression. *E. coli* HME5 *tna<>cat* strain harboring pCW''2A6 L240C/N297Q:hNPR and pAM101 (without *hem*A gene insert) was designated as *E. coli* AM101/2A6 and used as control strain for the study.

3.6 Heterologous expression of cytochrome P450 2A6 in newly established *E. coli* AM 102/2A6 strain through *hem*A gene coexpression

The newly established host strain *E. coli* AM102/2A6 harboring two compatible plasmids pCW''2A6 L240C/N297Q: hNPR and pAM102 was used for heterologous expression of functionally active cytochrome P450 2A6 without exogenous supplementation of ALA. In this host strain, ALA precursor required for expression of active form (holo-protein) of cytochrome P450 2A6, would be endogenously supplemented by expression of stabilized *hem*A gene cloned under *cat* promoter in plasmid pAM102. The stabilized *hem*A gene is constitutively expressed at basal level in absence of chloramphenicol whereas the gene can be induced to greater extent with addition of chloramphenicol. In this study, the strain *E. coli* AM101/2A6 harboring pCW''2A6 L240C/N297Q:hNPR and pAM101 (without *hem*A) was used as control strain. The protocols employed for heterologous expression of cytochrome P450 2A6 has been described below.

A single isolated colony of *E. coli* AM102/2A6 was grown overnight in LB/ampicillin (50 µg/ml), tetracycline (10 µg/ml) and chloramphenicol (25 µg/ml) medium at 37^0C with shaking (200 rpm). Then the overnight culture was 100 fold diluted in Terrific Broth (TB) medium supplemented with additives (2.0 g of Peptone/l, 0.4% (v/v) glycerol, 1 mM thiamine, 25 µg/ml ampicillin, 5 µg/ml tetracycline and 0.025% (v/v) of a mixture of trace elements) and incubated at 37^0C with shaking at 200 rpm. After 4 hours of incubation, *hem*A gene and cytochrome P450 2A6 gene were induced with addition of 25 µg/ml chloramphenicol and 1 mM IPTG, respectively. At the same time, the culture was sub-divided into two cultures. In one set of culture, indole (in DMSO) was added to final concentration of 5 mM and in another set of culture, only DMSO was added as control. Then the cultures were incubated for 44 hours at 28^0C with shaking (200 rpm).

The lower incubation temperature (28^0C instead of 37^0C) facilitates cytochrome P450 2A6 expression which metabolizes indole to non-toxic indoxyl products.

In this study, control strain *E. coli* AM101/2A6 was cultured in presence and absence of 1 mM ALA in TB medium as described previously for *E. coli* AM102/2A6. The culture of *E. coli* AM101/2A6 in presence of ALA (1 mM) acts as positive control while culture of *E. coli* AM101/2A6 in absence of ALA acts as negative control for this study.

In order to study the comparative effect of duration of *hem*A gene induction on expression level of functional cytochrome P450 2A6, another set of replica cultures were established with addition of 25 µg/ml chloramphenicol in TB medium along with the other additives during the culture initiation. To these set of cultures, only 1 mM IPTG and 5 mM indole (or DMSO in control) was added after 4 hours of culture incubation.

The applicability of the developed host strain *E. coli* AM102/2A6 to support heterologous expression of functionally active cytochrome P450 2A6 was assessed indirectly by measuring the bacterial growth rate. From the culture 100 µl of sample was aliquoted from the culture at different time intervals (0 hr, 1 hr, 2 hr, 4 hr, 24 hr and 44 hr) and optical density (OD_{600}) was measured with the help of UV-visible spectrophotometer against 0.1 M phosphate buffer (pH 7.0) as blank. The expression of catalytically active cytochrome P450 2A6 metabolizes indole present at toxic level (5mM) to non-toxic indoxyl products and decreases its cellular toxicity which can be monitored by determining the relative bacterial survival rate. After incubation of the culture in presence of toxic levels of indole (5 mM) the percentage of bacteria that survived the indole toxicity can be calculated using following formula.

$$\text{Relative bacterial survival rate} = \frac{OD_{600}\text{ (48 hrs) in presence of 5 mM indole}}{OD_{600}\text{ (4 hrs) before addition of indole}} \text{ X } 100\%$$

3.7 Chromosomal engineering

Endogenous ALA production can be further enhanced by chromosomal engineering of *E. coli* HME5 *tna<>cat* strain. The stable integration of stabilized *hem*A gene into chromosomal DNA of host strain in place of native *hem*A gene can be performed for which a cassette with 5′ homologous region, stabilized *hem*A gene and 3′ homologous region in a cloning vector has been designed. Primers has been designed as described for *hem*A gene amplification primers using different web based online softwares for PCR amplification of 5′ and 3′ homologous region of *hem*A gene. The restriction enzyme sites were introduced through forward and reverse primers for subsequent cloning steps. The designed primer set are as following and the details of oligonucleotide properties has been presented in Appendix 1.3 and 1.4.

3.7.1 Primer set for PCR amplification of 5’ homologous region of *hem*A

SSMP61320F01: CCTACTG ***AAGCTT*** gtaa acctaaaatc cactg
*Hind*III
SSMP61320R01: GATA ***GGATC CATATG*** tgc ggg aaa taa tac c
*Bam*HI *Nde*I

3.7.2 Primer set for PCR amplification of 3’ homologous region of *hem*A

SSMPOA710F01: CGT GGT GTG C ***TCTAGA*** ca ttt tca ttt ttt aca gg
*Xba*I
SSMPOA710R01: TGT CGG CTA ***GAATTC*** a gt tac gtt cgt cat cag
*EcoR*I

3.7.3 Reamplification primers for amplification of linear cassette containing *hem*A gene flanked by homologous region in 5’- and 3’-end to chromosomal *hem*A gene locus

For chromosomal integration of stabilized *hem*A gene, *hem*A gene flanked by 5′ and 3′ homologous region needs to be amplified. Thus reamplification primers were designed which does not contain any extra added nucleotide or restriction sites and could specifically amplify *hem*A gene flanked with upto 100 bp homology in both 5′ and 3′ region. The basic oligonulceotide properties have been presented in Appendix 1.5.

Reamplification primer set for Homologous recombination

SSMPOA6X1F03: cgtccg cta tcg tct atg ttc aag ttg tc

SSMPOA6X1R03: gta att gct gtt cca gtt gct cgc ttt ttt c

3.7.4 PCR amplification of 5’ and 3’ homologous region of *hem*A gene

Using *E. coli* HME5 *tna<>cat* chromosomal DNA as the template 5′ and 3′ homologous region of *hem*A gene was amplified. The reaction mixtures were prepared as described for *hem*A gene amplification but respective primer set was used to successfully amplify the DNA fragment. The PCR conditions (Appendices 1.2) were determined according to melting temperature calculated for oligos by using Oligoanalyzer webtool. After PCR amplification, 5 µl of PCR product was checked by 1% agarose gel electrophoresis.

3.7.5 Amplification of *hem*A gene locus with reamplification primers

E. coli HME5 *tna<>cat* chromosomal DNA was amplified using remplification primers (SSMPOA6X1F03 and SSMPOA6X1R03) to ascertain the designed reamplification primer specifically amplifies *hem*A gene flanked by 5′ and 3′ homologous region. The PCR conditions were same as that for *hem*A gene amplification. The PCR product was visualized after 1% agarose gel electrophoresis.

3.7.6 Construction of *Nde*I deleted pUC19 vector

To simplify the sub-cloning steps to construct the cassette containing 5′ homologous region, stabilized *hem*A and 3′ homologous region a cloning vector was developed by deleting *Nde*I site from pUC19 plasmid by Fill-in method. In this technique, the vector was first cut with *Nde*I and linear plasmid was subjected to 1% low melting agarose gel electrophoresis. The linear plasmid DNA with 5′ and 3′ overhangs was gel extracted and filled- in and blunted with Phusion DNA polymerase (or Klenow fragment) due to its 5′ to 3′ polymerase activity. The resulting blunt ended plasmid was ligated using T4 DNA polymerase and transformed into *E. coli* DH5α competent cells as described before. The ampicillin resistant transformants were used for plasmid extraction by alkaline lysis method. The deletion of *Nde*I site was verified by restriction mapping.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Construction of expression vector pAM101

Expression vector pAM101 was developed so that it can coexist with plasmids of ColE1 origin such as pBR322, pCW, pUC since this carries p15A origin of replication as that of pACYC184 (Cohen *et al.,* 1978) and support in expression of genes from both the plasmids. Creation of *Nde*I site overlapping ATG start codon of *cat* gene would make it feasible to clone gene (Parikh *et al.,* 1997) at this locus which would be under the regulation of *cat* promoter. Chloramphenicol has been found to support heterologous protein expression (Kusano *et al.*, 1999; Kagawa *et al.*, 2003), and low amount of chloramphenicol is not cytotoxic since its mode of action is much more bacteriostatic than bacteriocidal (Hahn *et al.*, 1995).

4.1.1 Isolation of pACYC184 plasmid and gel quantification

Plasmid pACYC184 was isolated from *E. coli* K12 ER2420/pACYC184 using HiPurA Silica kit. To determine the plasmid DNA concentration, 5 µl of supercoiled plasmid DNA sample and 1 kb NEB DNA ladder was subjected to 1% agarose gel electrophoresis and visualized after ethidium bromide staining as shown in figure 4.1. Two bands corresponding to nicked and supercoiled plasmid DNA sample was observed and the plasmid DNA concentration was determined to be 40 ng/µl by comparing its band intensity with standard known concentration of NEB DNA ladder. Thus quantified pACYC184 was used as template for site directed mutagenesis PCR.

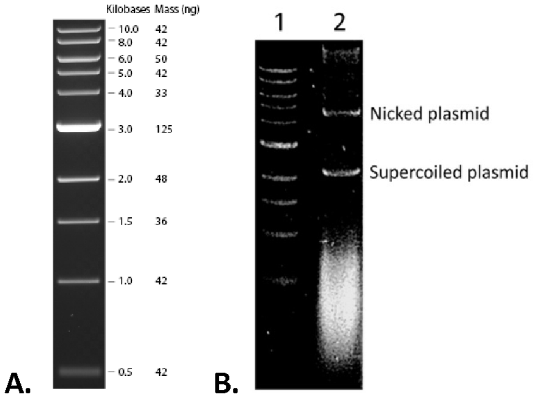
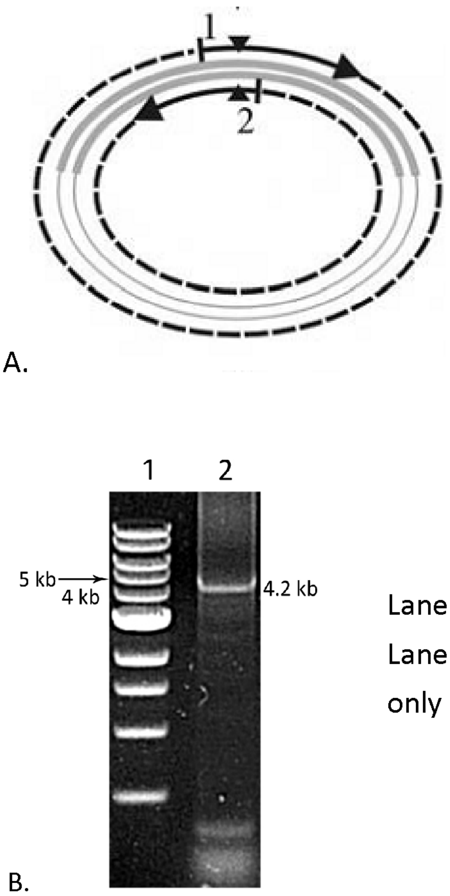


Figure 4.1: A. 1 kb NEB DNA ladder with standard known DNA concentrations. B. Gel quantification of supercoiled pACYC184 plasmid DNA along with 1 kb NEB DNA ladder in 1% agarose gel.

4.1.2 Site directed Mutagenesis PCR

Site directed mutagenesis PCR to create *Nde*I site in pACYC184 vector was performed by using specifically designed mutagenic primers (SSMX06403SDF and SSMX06403SDR) according to protocols of Liu and Naismith (2008). The different PCR conditions were used for successful amplification of plasmid DNA and optimized PCR condition has been presented in materials and method sections. Slight modification to the method of Liu and Naismith (2008) by increasing both concentration of template DNA (80 ng) and PCR amplification cycles (20 cycles instead of 12 cycles) helped to increase the overall PCR efficiency. The successful amplification of plasmid DNA was confirmed by observing 4.2 kb linear DNA fragment (Figure 4.2 B) after loading 5 µl of PCR product in 1% agarose gel. Site directed mutagenesis PCR using non-phosphorylated mutagenic primers introduces two nicks at both strands of amplified plasmid DNA (Figure 4.2 A, Liu and Naismith 2008) thus upon agarose gel electrophoresis only single 4.2 kb DNA band corresponding to nicked plasmid (Figure 4.2.B) was observed contrary to nicked and supercoiled DNA observed for template plasmid (Figure 4.1).



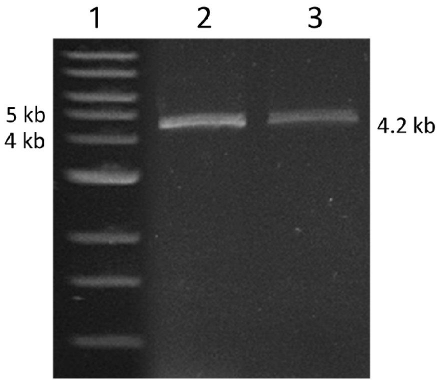
Lane 1: 1 kb NEB DNA ladder

Lane 2: Site directed mutagenesis PCR product (4.2kb), only nicked plasmid DNA is observed.

Figure 4.2: A. PCR amplified plasmid with two nicks (Lui and Naismith, 2008). B. Site directed mutagenesis PCR product of pACYC184 vector

4.1.3 Confirmation of creation of *Nde*I site in plasmid pAM101 through restriction mapping

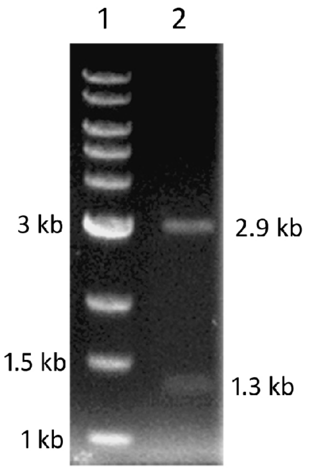
The PCR product was then treated with *Dpn*I to cleave methylated cellular DNA (template) so that only mutant plasmid amplified from PCR could be selectively transformed. During *Dpn*I treatment, PCR product would not be digested as it does not contain methylation at CpG region. Since the PCR product was *Dpn*I treated, it can be assumed that template DNA has been successfully cleaved as it consists of methylated CpG region because it was extracted from cellular matrix and only PCR amplified non-methylated DNA whose nick has not yet been ligated has been purified. The purified DNA after *Dpn*I treatment was used for transforming *E. coli* DH5α competent cells and plated in tetracycline plates. Four transformants from overnight incubated plates were taken for plasmid preparation. Plasmid DNA was extracted by Alkaline lysis method and confirmed by restriction mapping. A single linear DNA band of 4.2 kb was observed when plasmid was digested with *Nde*I restriction enzyme corresponding to 4.2 kb DNA band of *Eco*RI digested pACYC184 vector (Figure 4.3) indicating that this plasmid from the transformant has single restriction site recognized by *Nde*I restriction enzyme and is of similar molecular weight to pACYC184 thus suggesting that *Nde*I restriction site has been successfully created in pACYC184.



Lane 1: 1 kb NEB DNA ladder
Lane 2: pAM101 digested with *Nde*I
Lane 3: pACYC184 digested with *Eco*RI

Figure 4.3: Restriction digestion of plasmid DNA to confirm presence of *Nde*I site

Furthermore, the plasmid DNA was double digested with *Nde*I and *Hind*III and two DNA fragments (2.9 kb and 1.3 kb) were observed (Figure 4.4). This further confirms creation of *Nde*I site in pACYC184 by amplification with mutagenic primers since pACYC184 does not carry *Nde*I recognition site but *Hind*III recognition site. *Hind*III recognition site is at 1.5 kb which is around 1.3 kb from *cat* start codon, thus creation of *Nde*I recognition site at this start codon site would give band of 1.3 kb and 2.9 kb upon double digestion and cumulative size of both bands correspond to size of pACYC184 cleaved by *Hind*III confirming creation of *Nde*I site at ATG start codon of *cat* gene.Thus confirmed plasmid was designated as pAM101 and *E. coli* DH5α harboring pAM101 plasmid as *E. coli* AM101.



Lane 1: 1 kb NEB DNA ladder
Lane 2: pAM101 digested with *Nde*I and *Hind*III

Figure 4.4: Restriction digestion of plasmid DNA to confirm orientation of *Nde*I site

4.1.4 Functional assay of plasmid pAM101

In order to confirm that the insertion of *Nde*I site has not disrupted coding region of *cat* gene and the promoter activity, *E. coli* AM101, *E. coli* K12 ER2420/pACYC184 (positive control) and *E. coli* DH5α (negative control) were streaked in LB/chloramphenicol (25 μg/ml) agar plates and incubated at 37^{0}C for 24 hrs. After incubation, *E. coli* AM101 and *E. coli* K12 ER2420/pACYC184 were able to survive on LB/Cm^R plate but no growth was seen with *E. coli* DH5α. This clearly indicated that the presence of *Nde*I site in plasmid pAM101 does not affect the expression of chloramphenicol resistance marker gene and proved that the promoter is intact along with coding sequence as can be deduced from sequence alignment of primer regions and is illustrated in figure 4.5 (B) below.

RBS
A) pACYC184 cagtgatttt tttctccat t tt ag**cttcct** tagctcctga aaatctcgat
B) pAM101 cagtgatttt tttctccat A tG ag**cttcct** tagctcctga aaatctcgat
*Nde*I recognition site (catatg)

Figure 4.5: The *cat* gene promoter region showing translation initiation site (cat corresponds to ATG start codon) and relative position of ribosome binding site (RBS) (**cttcct** corresponding to AGGAAG) in pACYC184 vector and two point mutation sites in pAM101 are highlighted (Note: cat gene is in reverse orientation in pACYC184 vector)

4.2 Construction of expression vector pAM102

Since *hem*A gene product catalyze rate limiting enzyme in ALA biosynthesis (Chen *et al.,* 1994; Verderber *et al.,* 1997; Kang *et al.*, 2011), expression plasmid pAM102 was developed because it can coexist with plasmid pCW"2A6 L240C/N297Q:hNPR, and possibly express HemA protein to support higher amount of cytochrome P450 isoforms. Similar co-expression system for cytochrome P450 has been developed (Harnastai *et al.*, 2006) or cytochrome P450 and reductase from bicistronic construct (Parikh *et al.,* 1997) has been co-expressed with bacterial *O*-acetyltransferase (Aryal *et al.*, 1999; 2000; Oda *et al.,* 2001). In addition, this plasmid can be utilized in generation of random mutant library (Cadwell and Joyce, 1992; Zhao *et al.,* 1998; Nakamura *et al.,* 2001; Murakami *et al.,* 2002) of *hem*A gene that can be screened on the basis of indole metabolism by cytochrome P450 2A6 (Gillam *et al.,* 1999; 2000) expressed from pCW"2A6 L240C/N297Q:hNPR (Nakamura *et al.,* 2001; Guengerich *et al.,* 2004; Wu *et al.,* 2005). Furthermore, the mutant screened can be then inserted in host strain *E. coli* HME5 *tnaA<>cat* (Wu *et al.,* 2005) chromosomal DNA at *hem*A gene locus by homologous recombination (Yu *et al.,* 2000).

4.2.1 Primer designing for *hem*A gene amplification

When *S. typhimurium* and *E. coli* HemA proteins were aligned together, 94% identity on amino acid level were observed (Figure 4.6) and with conserved 18 amino acids in N-terminus of the both proteins. This provided strong evidence for similar heme regulatory feedback inhibition of *E. coli* HemA as that has been reported for *S. typhimurium* HemA protein (Jones and Elliott, 2010) which is sensitive to two proteases Lon and ClpAP mediated cleavage depending upon the cellular heme concentration (Wang *et al.,* 1999a). The recent works on introduction of two lysine residues close to N-terminus of *E. coli* HemA which has shown to provide protection against this proteolytic cleavage and insensitive towards heme regulated proteolysis irrespective of cellular heme concentration (Harnastai *et al.*, 2006; Biedendieck *et al.*, 2010; Kang *et al.*, 2011).

Thus, *hem*A gene amplification forward primer spanning N-terminus of the protein was designed using different web based bioinformatics tools to introduce two lysine encoding codons (AAA AAA; Appendix 2) between second and third amino acid residues in the amplified HemA protein to make it insensitive towards heme regulated proteolysis. Forward primer and reverse primer downstream of the stop codon of HemA protein was also designed using similar bioinformatics tools as described below.

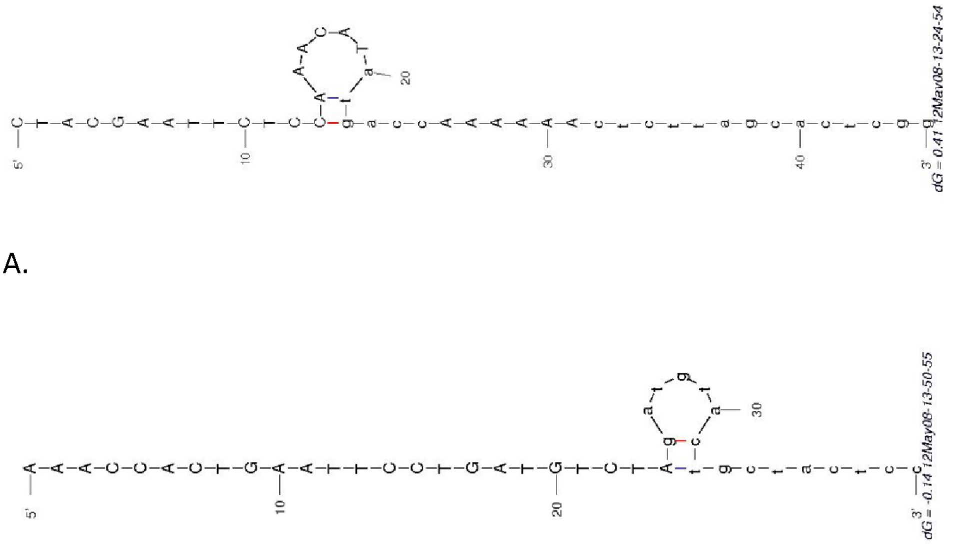
Score = 786 bits (2029), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 394/418 (94%), Positives = 407/418 (97%), Gaps = 0/418 (0%)

```
           1                                                           60
E.coli     MTLLALGINHKTAPVSLRERVSFSPDKLDQALDSLLAQPMVQGGVVLSTCNRTELYLSVE
           MTLLALGINHKTAPVSLRERV+FSPD LDQALDSLLAQPMVQGGVVLSTCNRTELYLSVE
S.typh     MTLLALGINHKTAPVSLRERVTFSPDTLDQALDSLLAQPMVQGGVVLSTCNRTELYLSVE


           61                                                         120
E.coli     EQDNLQEALIRWLCDYHNLNEEDLRKSLYWHQDNDAVSHLMRVASGLDSLVLGEPQILGQ
           EQDNLQEALIRWLCDYHNLNE+DLR SLYWHQDNDAVSHLMRVASGLDSLVLGEPQILGQ
S.typh     EQDNLQEALIRWLCDYHNLNEDDLRNSLYWHQDNDAVSHLMRVASGLDSLVLGEPQILGQ


           121                                                        180
E.coli     VKKAFADSQKGHMKASELERMFQKSFSVAKRVRTETDIGASAVSVAFAACTLARQIFESL
           VKKAFADSQKGH+ AS L RMFQKSFSVAKRVRTETDIGASAVSVAFAACTLARQIFESL
S.typh     VKKAFADSQKGHLNASALRRMFQKSFSVAKRVRTETDIGASAVSVAFAACTLARQIFESL


           181                                                        240
E.coli     STVTVLLVGAGETIELVARHLREHKVQKMIIANRTRERAQILADEVGAEVIALSDIDERL
           STVTVLLVGAGETIELVARHLREHKVQKMIIANRTRERAQ LADEVGAEVI+LSDID RL
S.typh     STVTVLLVGAGETIELVARHLREHKVQKMIIANRTRERAQALADEVGAEVISLSDIDARL


           241                                                        300
E.coli     READIIISSTASPLPIIGKGMVERALKSRRNQPMLLVDIAVPRDVEPEVGKLANAYLYSV
           ++ADIIISSTASPLPIIGKGMVERALKSRRNQPMLLVDIAVPRDVEPEVGKLANAYLYSV
S.typh     QDADIIISSTASPLPIIGKGMVERALKSRRNQPMLLVDIAVPRDVEPEVGKLANAYLYSV


           301                                                        360
E.coli     DDLQSIISHNLAQRKAAAVEAETIVAQETSEFMAWLRAQSASETIREYRSQAEQVRDELT
           DDLQSIISHNLAQR+AAAVEAETIV QE SEFMAWLRAQ ASETIREYRSQ+EQ+RDELT
S.typh     DDLQSIISHNLAQRQAAAVEAETIVEQEASEFMAWLRAQGASETIREYRSQSEQIRDELT


           361                                                      418
E.coli     AKALAALEQGGDAQAIMQDLAWKLTNRLIHAPTKSLQQAARDGDNERLNILRDSLGLE
            KAL+AL+QGGDAQAI+QDLAWKLTNRLIHAPTKSLQQAARDGD+ERLNILRDSLGLE
S.typh     TKALSALQQGGDAQAILQDLAWKLTNRLIHAPTKSLQQAARDGDDERLNILRDSLGLE
```

Figure 4.6: Sequence alignments of GluTR amino acid of *E. coli* MG1655 (UniProtKB/Swiss-Prot: P0A6X1.1) and *Salmonella typhimurium* (UniProtKB/Swiss-Prot: POA1Q6). The amino acid sequences were retrieved from and the alignments were made using the NCBI website.

Using different freely available web based accesses *hem*A gene amplification primers were designed to amplify genomic *hem*A gene. *E. coli* HME5 *tnaA<>cat* strain is more efficient than other *E. coli* strains in expression of cytochrome P450 (Wu *et al.,* 2005), and is also efficient in homologous recombination and is *E. coli* K12 derivative (Yu *et al.,* 2000) thus the primer was designed to amplify its *hem*A gene for cloning in pAM101.

Using Oligo analyzer 3.1, the self complimentary and other parameters were determined. Then the DNA folding structure was checked using mfold Web server- DNA folding form and was found to be as illustrated in figure 4.7 with Tm of the structures (30^0C for forward, and 39^0C for reverse) which is way below Tm of the primers (63.6^0 C for forward and 63.2^0 C for reverse primer) thus having the 3'-end flanking open and making it compatible to make complimentary hybridization with the template DNA. Similarly, heterodimer structure was also analyzed and found to have structure having ΔG = -8.51 kcal/mole. Thus, these sequences with the complementary region of the primer at 3'-end to genomic *hem*A gene DNA sequence and exhibiting 100% match with *E. coli* genomic *hem*A gene during the NCBI BLAST search was selected as the amplification primers. They were designated as SSMPOA6X1F01 for forward primer in which at the 5'-end had restriction sites created for recognition by restriction enzymes *Eco*RI followed by *Nde*I overlapping ATG start codon of *hem*A gene and SSMPOA6X1R01 for reverse primer in which at 5'-end had restriction sites created for recognition by restriction enzymes *Eco*RI followed by *Xba*I.



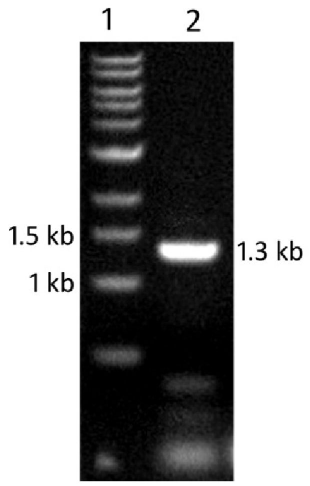
A.

B.

Figure 4.7: Secondary structure formed by forward (SSMPOA6X1F01) primer (A) and reverse (SSMPOA6X1R01) primer (B). The structures were generated using mfold-DNA folding form (http://mfold.rna.albany.edu/?q=mfold/DNA-Folding-Form).

4.2.2 PCR amplification of *hem*A gene

Genomic DNA was extracted from *E. coli* HME5 *tna<>cat* strain and its concentration and purity were determined spectrophotometrically. The concentration of genomic DNA was calculated to be 720 µg/ml with purity of 1.67 (A_{260}/A_{280}). Using specifically designed primers (SSMPOA6X1F01 and SSMPOA6X1R01) mutant *hem*A gene which is insensitive to heme mediated regulation (Jones and Elliott, 2010) was amplified from the isolated genomic DNA using High Fidelity Phusion DNA polymerase that has greater accuracy and efficiency than Taq DNA polymerase. The PCR product gave a clear band corresponding to around 1.3 kb DNA fragment on comparing with 1 kb NEB DNA ladder in another well (Figure 4.8). This DNA band was similar to nucleotide size of *hem*A gene as calculated from genome sequence of *E.* coli MG1655 strain (Appendix 5). Taking the size of the amplified DNA and the specificity of both primers as confirmed from NCBI BLAST results, it was assumed that *hem*A gene has been successfully amplified.



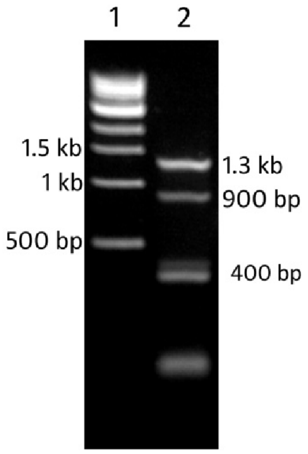
Lane 1: 1 kb NEB DNA ladder

Lane 2: *hem*A gene PCR product (1.3 kb)

Figure 4.8: PCR product of *hem*A gene in 1% agarose gel

4.2.3 Restriction digestion of *hem*A gene

The PCR product that gave 1.3 kb DNA fragment size and presumed to be *hem*A gene was purified by phenol:chloroform method. After purification, the amplified DNA was completely digested with *Eco*RI and followed by partial digestion with *Nde*I for subsequent cloning in pAM101 digested with *Eco*RI and *Nde*I. In order to purify corresponding size DNA digested with restriction enzymes, the restriction digested PCR sample was run in 1% low melting agarose gel (Figure 4.9) to separate DNA fragments because *hem*A gene has *Nde*I restriction site in its coding sequence (Gene ID: NC_000913.2) apart from the newly inserted during PCR amplification that generates multiple bands. Multiple DNA bands observed in figure 4.9 corresponded to single digestion with *Nde*I that is close to 1.3 kb, and DNA bands of around 900 bp and 400 bp corresponded to double digestion of PCR product that would yield 385 bp fragment between two *Nde*I recognition sites, and around 900 bp between second *Nde*I and *Eco*RI recognition sites. This would give cumulative length of PCR product to be 1.3 kb that corresponds to *hem*A gene. This clearly indicated that the amplified PCR product is potentially *hem*A gene. The lower band below 200 bp is supposed to be that of hybridized primers. Thus the correct DNA band size (~1.3 Kb) was cut with clean knife and proceeded for gel extraction.



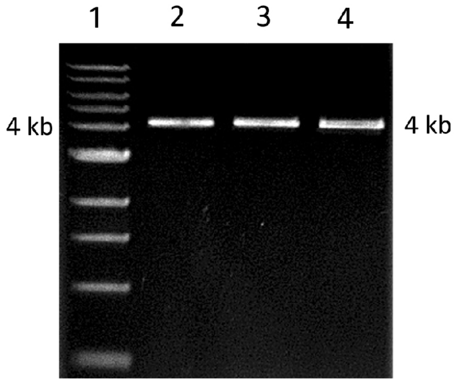
Lane 1: 1 kb NEB DNA ladder

Lane 2: *hem*A gene digested with *Nde*I and *Eco*RI

Figure 4.9: Partial restriction digested *hem*A gene in 1% Low melting agarose gel

4.2.4 Restriction digestion of pAM101 with *Nde*I and *Eco*RI

The expression vector pAM101 was prepared for sub-cloning of *hem*A gene by double digestion with *Nde*I and *Eco*RI and whole reaction mixture was subjected to gel electrophoresis in 1% agarose gel. The linearized vector corresponding to 4 kb DNA fragment (Figure 4.10) was subjected to gel purification.



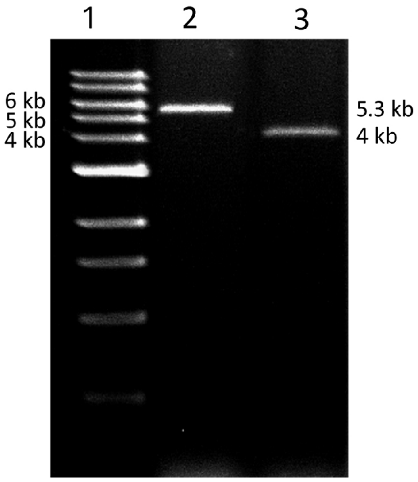
Lane 1: 1 kb NEB DNA ladder

Lane 2-4: pAM101 digested with *Nde*I and *Eco*RI

Figure 4.10: Restriction digestion of pAM101 with *Nde*I and *Eco*RI

4.2.5 Construction of expression vector pAM102 and restriction digestion to confirm Positive transformants harboring plasmid pAM102

Both PCR amplified *hem*A gene fragment and pAM101 digested with *EcoR*I and *Nde*I, were electrophoresed and their respective quantities were empirically calculated by visualization in 1% agarose gel comparing with known quantity of 1 kb NEB DNA marker. With the formula for subcloning as described in materials and methods, respective DNA was subjected to ligation. After ligation, the reaction mixture was used to transform *E. coli* DH5α competent cells by heat shock method as described. Sub-cloning of *hem*A gene (1.3 kb) in pAM101 vector (4 kb) was confirmed by restriction map analysis of isolated plasmids from transformants.



Lane 1: 1 kb NEB DNA ladder
Lane 2: pAM102 digested with *Eco*RI
Lane 3: pAM102 digested with *Nde*I and *Eco*RI

Figure 4.11: Restriction digestion of plasmid to confirm sub-cloning of *hem*A gene in pAM101

Plasmid DNA was isolated from positive clones surviving in tetracycline plate and was confirmed by restriction digestion for insertion of *hem*A gene in pAM101. The linearized plasmid showed 5.3 kb DNA band with *EcoR*I digestion corresponding to cumulative bp of pAM101 (4 kb) and PCR amplified *hem*A gene (1.3 kb). Furthermore, the 4 kb DNA band with *EcoR*I and *Nde*I double digestion corresponded to pAM101 size (Figure 4.11) and *hem*A gene that was cleaved at multiple sites was not detected in the present gel. This indicated that *hem*A gene was successfully sub-cloned in plasmid pAM101 and this newly developed plasmid was designated as plasmid pAM102 and *E. coli* DH5α harboring this plasmid as *E. coli* AM102. The overall strategy of development of pAM102 is illustrated in figure 4.12.

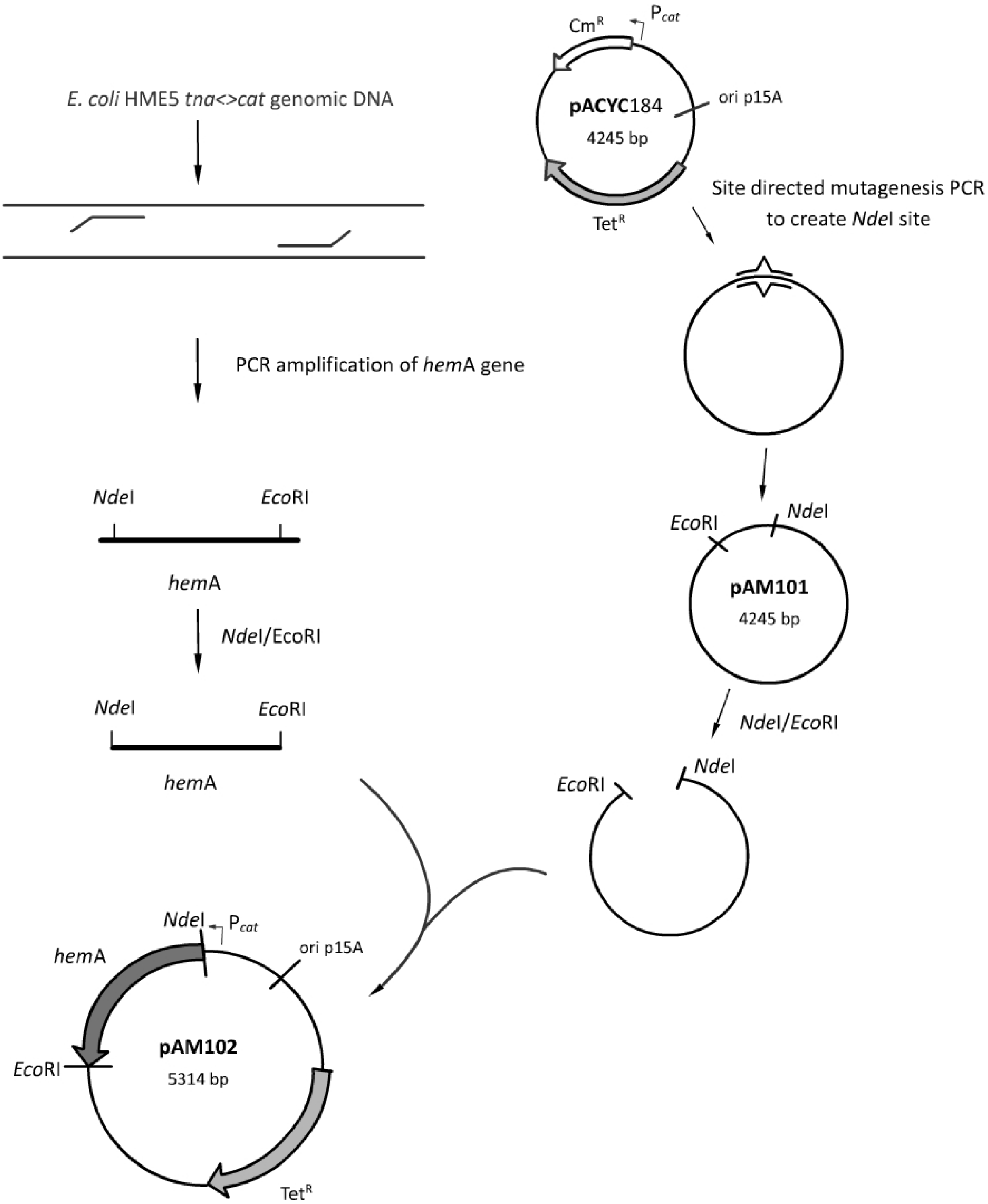


Figure 4.12: Schematic diagram of sub-cloning *hem*A gene in plasmid pAM101 to construct the plasmid pAM102

4.3 Development of *E. coli* AM102/2A6 and *E. coli* AM101/2A6

E. coli HME5 *tna<>cat* strain does not express indole endogenously (Wu *et al.,* 2005) making it efficient host to study indole cytotoxicity because this strain harboring pCW”2A6 L240C/N297Q:hNPR is known to metabolize substituted indoles efficiently by cloned cytochrome P450 2A6 L240C/N297Q double mutant (Nakamura *et al.,* 2001). The competent cell of *E. coli* HME5 *tna<>cat* harboring pCW”2A6 L240C/N297Q:hNPR plasmid was prepared and transformed with newly developed pAM102 plasmid which carry mutated *hem*A gene that is supposed to express stabilized HemA protein, and resulting ampicillin/tetracycline resistant transformants carrying pCW”2A6 L240C/N297Q:hNPR and pAM102 were designated as *E. coli* AM102/2A6. Similar sequential transformation of pCW”P450:hNPR expressing cytochrome P450 and NADPH reductase (Kranendonk *et al.,* 1999) or different P450 isoforms along with reductase and pOA101 (Aryal *et al.,* 1999) or pOA102 (Aryal *et al.,* 2000; Oda *et al.,* 2001) plasmids which are derivatives of pBR322 and pACYC184 have been done for expression of genes cloned in respective plasmids. Similarly, *E. coli* AM101/2A6 strain was established by introducing the null vector without *hem*A gene (pAM101) and pCW”2A6 L240C/N297Q:hNPR that is used as control strain in the present study. The significance of developing this *E. coli* AM102/2A6 host strain is illustrated schematically in the figure 4.13.

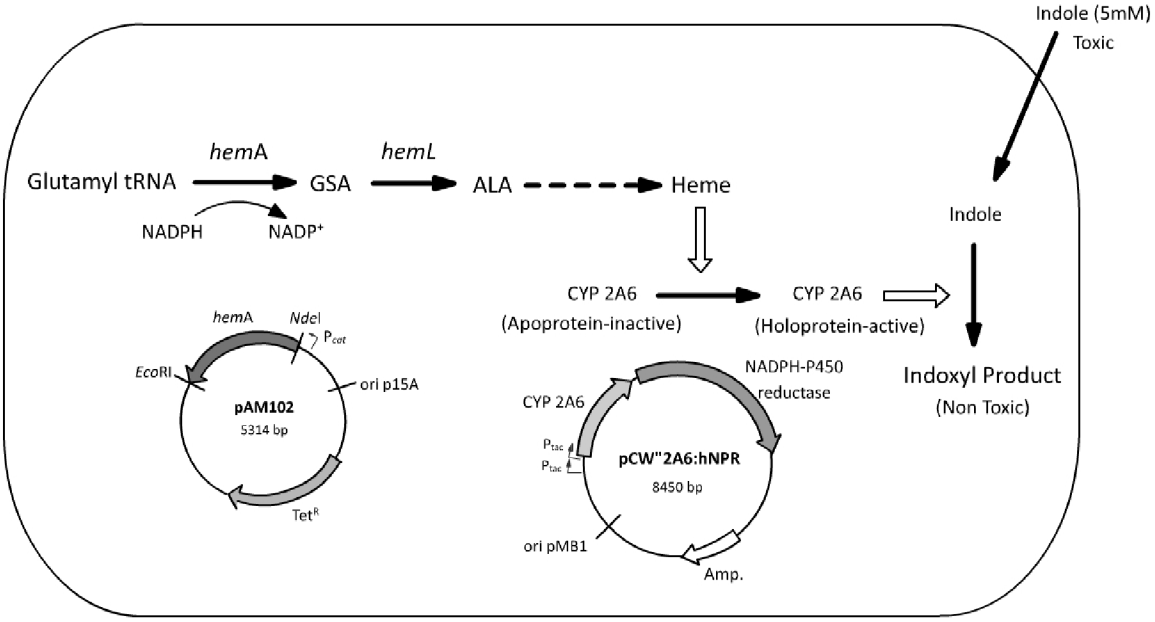


Figure 4.13: Schematic diagram of newly developed host strain *E. coli* AM102/2A6 harboring two compatible plasmids, pCW”2A6 L240C/N297Q:hNPR and pAM102, which can support heterologous expression of functionally active cytochrome P450 2A6 by endogenous supplementation of ALA through *hem*A gene co-expression.

4.4 Functional activity of sub-cloned genes expressed in newly developed *E. coli* AM102/2A6 and *E. coli* AM101/2A6

It is presumed that the newly developed *E. coli* AM102/2A6 strain upon induction of chloramphenicol acetyl transferase (*cat*) promoter with chloramphenicol would induce expression of stabilized *hem*A gene that is cloned under *cat* promoter in pAM102. Additional expression of stabilized *hem*A gene from plasmid could potentially support expression of functional cytochrome P450 2A6 double mutant along with human NADPH-P450 reductase for functional metabolism of supplemented indole without addition of ALA in the expression media. The hypothesis of the present work has been that the feedback insensitive stabilized *hem*A gene product (Jones and Elliott, 2010) in association with chromosomal *hem*L gene product catalyzes conversion of charged glutamyl tRNA to augment endogenous production of ALA (Moser *et al.*, 2001; Kang *et al.,* 2011) and subsequently increase heme biosynthesis inside the cells (Woodard and Dailey, 1995). This would presumably support heterologous expression of non-intrinsic heme proteins such as cytochrome P450 in *E. coli* as that has been reported (Harnastai *et al.,* 2006) since ALA is the concomitant precursor of heme biosynthesis (Jahn *et al.,* 1992). The increased ALA biosynthesis would increase cellular heme concentration and making it available for non-intrinsic heme proteins in the cell. Thus, formation of cytochrome P450 holo-protein (active form) would be augmented from the apo-protein (non-active form) expressed from pCW"2A6:hNPR upon induction by IPTG (Parikh *et al.,* 1997; Aryal *et al.,* 1999; 2000; Gillam *et al.,* 2000; Nakamura *et al.,* 2001; Wu *et al.,* 2005) mostly due to increased heme biosynthesis and making availability of increased amount of heme that can be then integrated in cytochrome P450 2A6. The bacteria thus developed is assumed to be able to perform cytochrome P450 mediated metabolism as illustrated in the figure 4.13, with increased level of holo-protein and expression of its cognate redox partner NADPH-P450 reductase cloned as second cistron downstream of cloned cytochrome P450 cDNA (Parikh *et al.,* 1997).

4.4.1 Study of effect of chloramphenicol in bacterial growth rate on newly developed bacterial strains

Host strain *E. coli* HME5 *tna*A<>cat harboring pAM102 and pCW"2A6 L240C/N297Q: hNPR (*E. coli* AM102/2A6) was cultured in the absence of 1mM ALA, and as the positive control, *E. coli* HME5 *tna*A<>cat harboring pAM101 and pCW"2A6 L240C/N297Q:hNPR (AM101/2A6 strain) was sub-cultured in the presence of 1mM ALA and as the negative control *E. coli* AM101/2A6 was sub-cultured without exogenous supplementation of ALA. Bacterial survival was monitored for resistance to chloramphenicol at 25 µg/ml and there was not substantial cytotoxicity from chloramphenicol (Figure 4.14) to any of the cultures since host bacteria *E. coli* HME5 *tna*A<>*cat* has chloramphenicol resistant gene incorporated in *tna*A locus at its chromosome (Wu *et al.,* 2005) despite *hem*A gene being cloned under *cat* promoter at locus of *cat* gene of plasmid pACYC184 that disrupts *cat* gene. This clearly indicated that even use of chloramphenicol for induction of *hem*A gene would not have substantial cytotoxicity effect in the newly constructed host strain *E. coli* AM102/2A6.

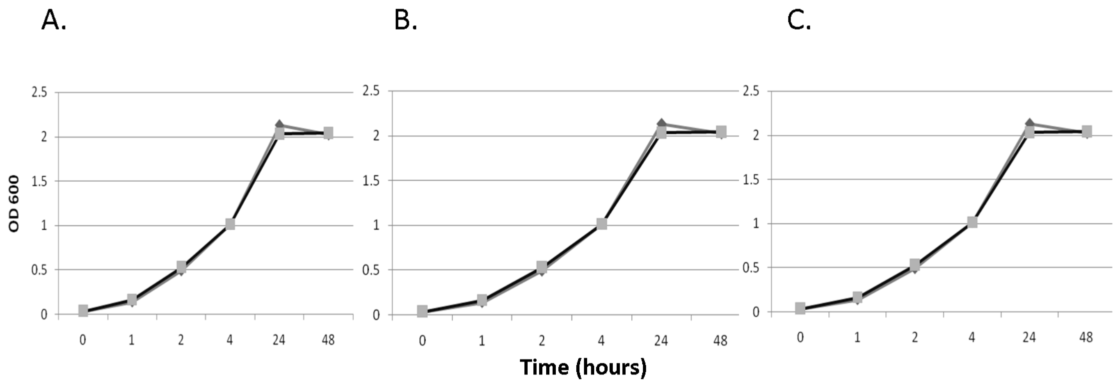


Figure 4.14: Comparative study of growth characteristics of (A) *E. coli* AM102/2A6, (B) *E. coli* AM101/2A6 and (C) *E. coli* AM101/2A6 with ALA (1mM). Supplementation of 25 µg/ml of chloramphenicol in the bacterial culture at 0 hour (◆) or at 4 hour of incubation (■). After 4 hours of incubation IPTG for induction of cytochrome P450 2A6 and hNPR, and DMSO (in subsequent indole would be dissolved in DMSO) were added in all cultures.

When host bacteria *E. coli* AM102/2A6 was cultured in presence or absence of 25 µg/ml chloramphenicol during first 4 hours of culture initiation, the bacterial growth rate was shown to be slightly retarded in culture with chloramphenicol as evident in figure 4.14 A. But, in case of control strain *E. coli* AM101/2A6 (with or without ALA) no such difference in growth rate was observed in presence or absence of chloramphenicol (Figure 4.14 B and C). Compared to control strain *E. coli* AM101, the slight retardation in the bacterial growth rate of host bacteria *E. coli* AM102/2A6 in presence of chloramphenicol can be explained due to the differences in copy number of *cat* gene present in the bacteria. The host strain consists of only one copy of *cat* gene that is present in chromosomal DNA but has been disrupted in pAM102 plasmid during sub-cloning of *hem*A gene under *cat* promoter whereas the control strain *E. coli* AM101/2A6 has two copies of *cat* gene in plasmid pAM101 and chromosomal DNA that could have accounted for efficient neutralization of chloramphenicol that would otherwise hinder bacterial growth due to inhibition of protein biosynthesis (Hahn *et al.*, 1995).

4.4.2 Heterologous expression of functionally active cytochrome P450 2A6 in newly established *E. coli* AM102/2A6 strain through *hem*A gene coexpression

The host strain *E. coli* AM102/2A6 was studied for its ability to support expression of functionally active cytochrome P450 2A6 by augmented ALA biosynthesis through co-expression of stabilized *hem*A gene. The host strain was sub-cultured in TB medium supplemented with trace elements and additives (Appendix 4; Parikh *et al.,* 1997). In this expression system, *hemA* gene, cloned under *cat* promoter in pAM102, was induced with addition of 25 µg/ml of chloramphenicol since the bacteria can survive in presence of chloramphenicol, too, while cytochrome P450 2A6 with its auxiliary redox partner, cloned under double *tac* promoter in pCW''2A6 L240C/N297Q:hNPR, was induced with addition of 1mM IPTG as described in material and methods section.

Expression of functionally active cytochrome P450 2A6 L240C/N297Q without exogenous supplementation of ALA was monitored by observing bacterial survival rate of the host strain *E. coli* AM102/2A6 that was exposed to toxic level of indole (5mM final concentration) (Nakamura *et al.,* 2001; Zhang *et al.*, 2009). After induction of cytochrome P450 2A6 and addition of 5mM indole, the cultures were further incubated for 44 hours and bacterial growth rate was determined by UV-visible spectrophotometer measuring OD_{600} at different time intervals as shown in figure 4.15.

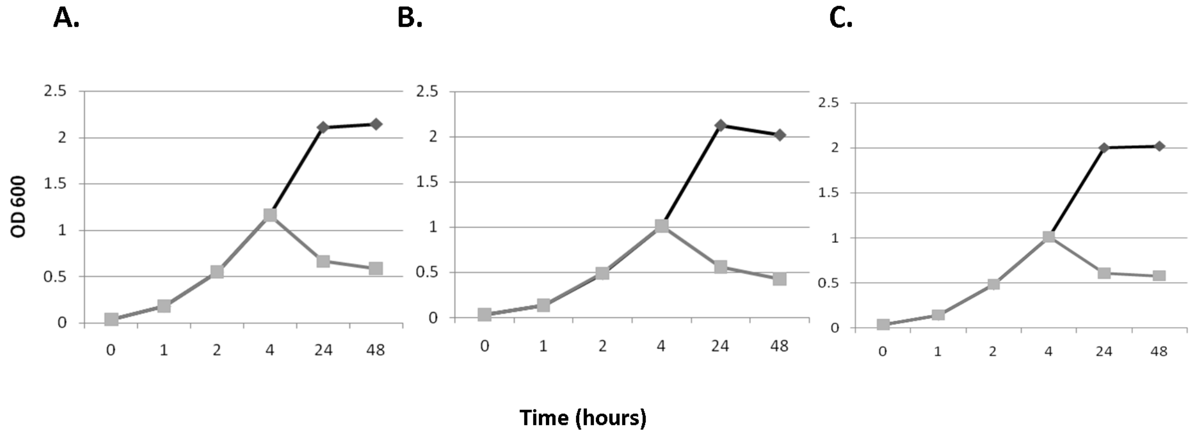


Figure 4.15: Comparative study of growth characteristics of (A) *E. coli* AM102/2A6, (B) *E. coli* AM101/2A6 and (C) *E. coli* AM101/2A6 with ALA (1mM). The overnight culture was 1:100 times diluted in TB expression medium with 25 µg/ml chloramphenicol. After 4 hours, IPTG was added and culture was sub-divided into two culture flask supplemented with 5mM Indole (■) and DMSO only (◆).

Compared with DMSO only, all the cultures exhibited cytotoxicity when 5 mM indole was added that was evident from the decrease in optical density at OD_{600} as shown in fig

4.15. As illustrated in the figure (4.15 C) when ALA was supplemented the *E. coli* AM101/2A6 (*E. coli* HME5 *tna*A<>cat harboring pAM101 lacking *hem*A gene and pCW''2A6 L240C/N297Q: hNPR) culture had higher survival rate than the culture of same bacteria without supplementation of ALA when cultured in presence of 5 mM indole (Figure 4.15 B). To further analyze the data, Relative bacterial survival rate (%) was also calculated by considering the bacterial population (OD_{600} at 4 hrs) present during the time of induction of cytochrome P450 2A6 by IPTG and after 44 hours of culture in presence of 5 mM indole (OD_{600} at 48 hrs) as shown in figure 4.16. The effect of *hem*A gene induction period for cytochrome P450 2A6 expression levels was also studied.

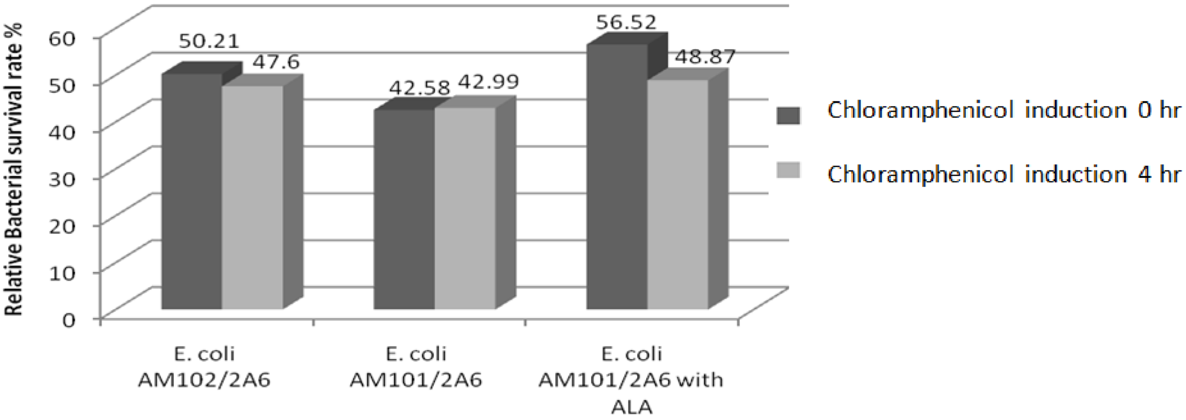


Figure 4.16: Study of effect of duration of *hem*A gene induction for functional CYP 2A6 expression levels.

The graph demonstrates the effect of duration of *hem*A gene induction on CYP 2A6 expression levels. The induction of *hem*A gene by addition of chloramphenicol from culture initiation (0 hour) helped to accumulate more ALA precursor (Jones and Elliott, 2010) which could have been eventually consumed to convert newly expressed CYP 2A6 to active holo-protein (Harnastai *et al.*, 2006) that subsequently metabolized more indole (5 mM) (Gillam *et al.,* 1999; Nakamura *et al.,* 2001; Wu *et al.,* 2005) thus decreasing its toxicity and increasing the bacterial survival rate to 50.21% for *E. coli* AM102/2A6 which is in close proximity to 56.52% for ALA supplemented *E. coli* AM101/2A6. This can be taken as an evident that functional cytochrome P450 could be expressed in *E. coli* by expression of stabilized *hem*A gene under *cat* promoter if the strain is resistant to chloramphenicol. This system could be better than the IPTG inducible *hem*A gene (Harnastai *et al.,* 2006) because the inducer would compete for expression of cytochrome P450, NADPH-P450 reductase and HemA.

Similarly, the control strain *E. coli* AM101/2A6 (without stabilized *hem*A) supplemented with 1 mM ALA also exhibited higher bacterial survival rate 56.52% when chloramphenicol was added from culture initiation compared to 48.87% when added after 4 hours. This indicated that chloramphenicol inhibited protein synthesis and has accumulated glutamyl tRNA pool which potentially increased intracellular ALA (Nakayashiki *et al.*, 1995) and hence increased the expression of functionally active holo-form of cytochrome P450. This can be presumed because in negative control with *E. coli* AM101/2A6 that did not have either exogenous ALA or *hem*A gene cloned in the plasmid, the survival rates did not differ much in both the conditions where chloramphenicol was added either at 0 or 4 hours after incubation of the cells indicating that cytochrome P450 expression was limited without ALA and indole metabolism could not occur thus exhibiting cytotoxicity.

As explained above that accumulation of charged tRNA including glutamyl tRNA occurs when cells are exposed to chloramphenicol (Nakayashiki *et al.*, 1995). It can be assumed that this phenomenon would increase the availability of precursors for HemA protein in ALA biosynthesis through C5 pathway and thus increasing free heme availability. This would reduce competitiveness with intrinsic heme protein and support subsequently in increasing level of functional CYP 2A6 expressed. In addition, chloramphenicol, a protein synthesis inhibitor (Hahn *et al.*, 1995), is known to enhance heterologous expression of P450s and related proteins by induction of cold-shock response in *E. coli* (Kunsano *et al.*, 1999; Kagawa *et al.,* 2003) and higher expression levels of different P450 isoforms were observed with addition of chloramphenicol (Harnastai *et al.*, 2006). This further supports present results that relatively more bacterial cell survived in 5mM indole in the culture where chloramphenicol was added from culture initiation.

However, compared to the newly developed host strain *E. coli* AM102/2A6 the positive control strain *E. coli* AM101/2A6 with exogenous ALA supplementation showed relatively higher bacterial survival rate (56.52%) than that of *hem*A gene co-expression strategy. This clearly indicated that the amount of ALA produced by just cloning of *hem*A gene only is not efficient since it is reported that HemA protein dimerizes with HemL protein in ALA biosynthesis (Moser *et al.*, 2001). HemA converts glutamyl tRNA to GSA (Avissar *et al.*, 1989; Verkamp *et al.*, 1992) which is then channelized into HemL groove finally giving ALA (Ilag and Jahn, 1992; Jahn *et al.*, 1992; Smith *et al.*, 1992). In addition, GSA is known to be unstable compound so it is presumed that only expression of *hem*A gene was not sufficient for biosynthesis of enough ALA that could support higher expression of cytochrome P450. Co-expression of *hem*A and *hem*L gene as bicistronic construct have been attempted for overproduction of ALA by Kang and his group (2011) and higher production of ALA has been observed. In the study *hem*A and *hem*L gene coexpression had synergistic effect on ALA production. Thus supporting our hypothesis that *hem*L gene coexpression would further enhance ALA production and for efficient heterologous expression of cytochrome P450s.

4.5 Chromosomal engineering

Despite the fact that *E. coli* strains are incompatible for homologous recombination without further manipulation for chromosomal engineering, the host strain *E. coli* HME5 *tna<>cat* (Wu *et al.,* 2005) developed using *E. coli* HME5 originally developed by Court and coworkers (Yu *et al.,* 2000) is homologous recombination proficient bacteria. The bacteria has integrated λ prophage genes into its chromosomal DNA which can be selectively induced and recombinogenic genes can be expressed to make the bacteria competent to recombine electroporated linear foreign DNA fragments into chromosomal locus flanked by homologous region of the locus. These competent bacteria has unusual capacity to uptake DNA fragments and recombine the electroporated DNA fragments with even 30-50 bp of homologous region flanking the DNA segment to be recombined.

For the development of robust host strain for heterologous expression of functionally active cytochrome P450s by increasing ALA level by increasing half life of HemA protein by stable integration of stabilized *hem*A gene that is insensitive to heme regulated proteolysis into chromosomal DNA by swapping native *hem*A gene can be envisaged as that have been done to swap *cat* gene with chromosomal tryptophanase A gene (Wu *et al.,* 2005). Integration of mutated *hem*A gene would not require co-transfection of pAM102 for expression of *hem*A gene for augmented ALA biosynthesis for cytochrome P450 as has been performed in this study and this would reduces the burden of plasmid based expression to the cell (Flores *et al.*, 2004), and also would overcome limitation of low copy number plasmid pAM102. In addition, if further amount of ALA would be required then co-transfection of pAM102 would support in higher expression of HemA protein.

Furthermore, dimeric HemA protein is known to make complex structure with HemL (Moser *et al.*, 2001) that not only protects unstable GSA but properly channelizes to give ALA (Luer *et al.*, 2005). Thus cloning of *hem*L gene becomes more prudent and with two plasmids expression system in the developed bacteria *E. coli* AM102/2A6, the cell would be put in additional burden if *hem*L is also be cloned in plasmid, thus chromosomal integration of stabilized mutant *hem*A would relieve the use of pAM102 expression and this would support cloning of *hem*L gene under the *cat* promoter in pAM101 for promoting proportionate ratio of HemA/HemL complex by increasing of half life of HemA to match the amount of HemL protein which would be possibly required to synchronize the expression level of HemA protein that is sensitive to proteolytic cleavage and thus support sufficient ALA formation.

E. coli HME5 *tna<>cat* strain is efficient in homologous recombination of linear DNA that is flanked by 50 bp homologous DNA sequence at both 5'- and 3'-ends to the locus where it is to be integrated in chromosome (Yu *et al.,* 2000; Wu *et al.,* 2005). The stabilized *hem*A gene amplified with nearly 100 bp homology region upstream and downstream of native *hem*A gene could be electroporated in competent *E. coli* HME5 *tna<>cat* strain because recombination efficiency mediated by Exo, Beta and Gam (Murphy, 1991 and references there in) of λ prophage genes (Yu *et al.,* 2000). The combined actions of all three gene products help to recombine stabilized *hem*A gene in place of native *hem*A gene. Thus developed host strain with stabilized *hem*A gene under native regulatory promoter is thought to circumvent exogenous requirement of ALA supplementation for heterologous cytochrome expression. In addition, at the same time this strategy would rescue the host bacteria from metabolic burden due to plasmid based gene expression system (Nakayashiki *et al.*, 1995). Thus the gene cassette which has mutated *hem*A gene flanked by homologous regions at both 5'- and 3'-ends of native *hem*A gene should be constructed as illustrated.

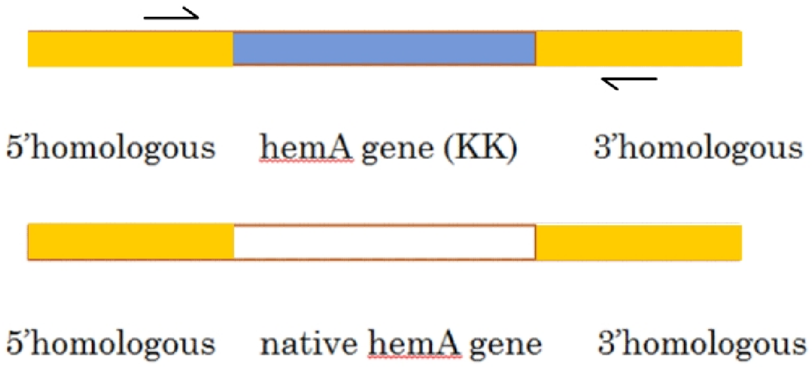


Figure 4.17: Schematic representation of DNA cassette carrying mutated *hem*A gene flanked by 5'- and 3'-homologous region of *hem*A gene for chromosomal swapping.

Initially, it was customary to investigate whether our strategy of co-expressing mutated *hem*A gene (cloned in pAM101) along with pCW"2A6:hNPR L240C/N297Q double mutant would support indole metabolism. As observed in bacterial survival data it is now clear that expression of HemA protein would supplement ALA for additional heme biosynthesis to have functional cytochrome P450, it is assumed that the strain has to be further developed for higher efficiency.

In order to develop this host strain that has mutated *hem*A gene cloned in chromosome at native locus, strategies and experimental designs have been carried. According to the strategy some experiments to develop construct that has mutated *hem*A gene flanked by both 5′- and 3′- regions homologous regions of native chromosomal *hem*A gene have already been done.

Single swapping could be feasible by amplifying *hem*A gene with mutation inserted and homologous flanking regions designed in primers as that has been done to amplify *cat* gene (Wu *et al.,* 2005). However, generation of random mutant library (Nakamura *et al.,* 2001; Kim *et al.*, 2005) cannot be screened for gain of function without promoter being intact. Thus, cloning has to be done in a vector containing either native *hem*A gene promoter amplified along with the mutant *hem*A gene or under another constitutive/inducible promoter. Furthermore, if native *hem*A promoter is intact then the gain of function mutant cannot be distinguished whether the mutation was in promoter that expressed higher amount of gene or the mutation was in coding sequence that change the functional activity of protein. In addition, if the three dimensional study has to be carried out then the screening of random mutant library has to be created using only at coding sequence hence make it necessary to clone in a vector with the promoter. Thus *hem*A was cloned in pAM101 to develop pAM102.

Plasmid pAM102 vector is limited with restriction enzyme recognition sites to integrate flanking homologous regions development of new construction vector is required. Alternatively expression vector pUC19 has been chosen due to its multiple cloning sites. 5′ homologous insert has designed in such a way that *Hind*III site at 5′-end and *Nde*I followed by *Bam*HI site at 3′end (see below) created for efficient cloning to develop the construct. Cloning of *Hind*III and *Bam*HI digested 5′ homologous region in pUC19 vector digested by these restriction enzymes would create *Nde*I site. This *Nde*I site can be used to clone *hem*A gene digested by *Nde*I and *Eco*RI (pUC19 vector has *Eco*RI site downstream to *Bam*HI site) so that the initiation ATG codon of mutated *hem*A would be intact at same position as in native chromosomal form whose expression would be under the control of native promoter. Then to this construct 3′-homologous region would be cloned by digesting with *Xba*I and *Eco*RI making complete cassette. To amplify linear DNA without any sequences of restriction sites by reamplifying this cassette from the primers designed internal sequences of 5′- and 3′-homologous regions with at least 100 nucleotide homologous region flanking the mutated *hem*A gene.

However, pUC19 vector has internal *Nde*I restriction recognition site at nucleotide position 183 in the plasmid (Appendix 7). As mentioned above after cloning of 5′-homologous region then there would be two *Nde*I sites created mandating partial digestion for cloning of *hem*A gene. In addition, when mutant library has to be generated by modified random insertion and deletion mutagenesis (Murakami *et al.,* 2002) two *Nde*I sites would make it difficult to get the desired coding sequence. Furthermore, substantial mutant library would be lost during partial digestion. This may limit screening of potential gain of function mutants. As this construct has to be cloned in pUC19 whose *Nde*I site has been deleted and new pUC19ΔNdeI plasmid was developed whose *Nde*I restriction recognition site has been removed to avoid partial digestion due to PCR amplified 5′-homologous has *Nde*I recognition site to clone *hem*A gene in the construct.

However, the linear DNA required for homologous recombination retrieved from the construct for cloned in pUC19ΔNdeI plasmid by restriction digestion (or amplified from forward primer of 5′-homologous region and reverse 3′-homologous region) for DNA swapping would be having restriction enzyme recognition sites at 5′- and 3′- ends making it difficult to make homologous recombination (Yu *et al.,* 2000), linear DNA flanked with homologous regions of at least 100 bp would be needed.

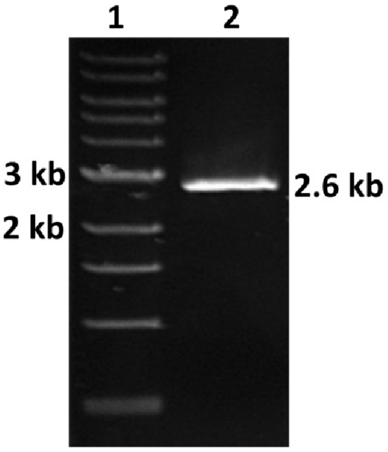
Thus, subsequent amplification of mutant *hem*A gene library would be easier from cloning into pUC19ΔNdeI plasmid than using primers with flanking region. This is because such primers would have long size, high Tm and possibility of secondary structures between primers that would make amplification complicated and would be difficult to use the primer for cloning. Thus linear reamplification primer set was designed without any restriction sites that would amplify *hem*A gene with homologous region at both 5′- and 3′- region of *hem*A gene (Figure 4.18).

Figure 4.18 PCR amplification of *hem*A gene containing cassette for homologous recombination without any restriction enzyme sites or non priming sequences at both 5′- and 3′-ends

4.5.1 Construction of pUCΔNdeI vector

In order to construct the cassette containing stabilized *hem*A gene flanked by 5′- and 3′- homologous regions of *hem*A gene, a special cloning vector was developed with *Nde*I restriction recognition site of pUC19 deleted by Fill-in method. Since, *hem*A gene contains internal *Nde*I site at its coding region and for cloning of *hem*A gene in this cassette *Nde*I site has been also created at ATG start codon of this gene. So, to avoid the multiple partial digestions in subsequent cloning steps native *Nde*I restriction site of pUC19 plasmid has been deleted by Fill-in method.

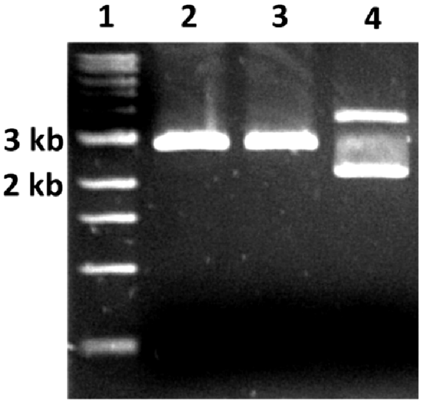
Initially, pUC19 plasmid was digested with *Nde*I restriction enzyme and the digested plasmid was run in 1% agarose gel that gave single band of 2.6 kb DNA corresponding to linear pUC19 plasmid (Figure 4.19) and confirming that there is only one *Nde*I recognition site in the plasmid.



Lane 1: 1kb NEB DNA ladder
Lane 2: pUC19 digested with *Nde*I

Figure 4.19: pUC19 plasmid digested with *Nde*I

The *Nde*I digested pUC19 vector was extracted from gel and the recessive ends were filled-in by the action of Phusion DNA polymerase. Then the blunt ended vector was ligated and transformed into *E. coli* DH5α cells. The plasmid was extracted and digested with different restriction enzymes to confirm the absence of *Nde*I site in pUC19ΔNdeI vector. The plasmid DNA when digested with *Nde*I produced two DNA fragments corresponding to nicked and supercoiled DNA (Figure 4.20). While digestion with *Hind*III only or in combination with *Nde*I produced only single linear band confirming deletion of *Nde*I recognition site from the plasmid.

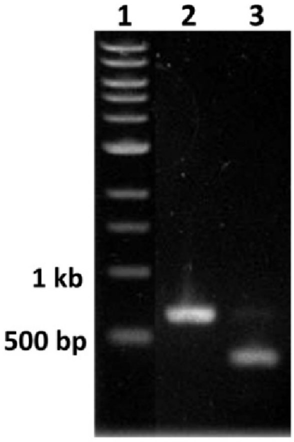


Lane 1: 1kb NEB DNA ladder
Lane 2: pUCΔNdeI digested with *Hind*III (2.6 kb)
Lane 3: pUCΔNdeI digested with *Hind*III and *Nde*I (2.6 kb)
Land 4: pUCΔNdeI digested with *Nde*I (uncut plasmid)

Figure 4.20: Confirmation of pUCΔNdeI plasmid with restriction digestion

4.5.2 PCR amplification of 5'- and 3'- Homologous region of *hem*A gene

Genomic DNA of *E. coli* HME5 *tna<>cat* strain was used for PCR amplification of 5′- and 3′- homologous region of *hem*A gene using specifically designed primers. PCR amplification produced of 5′ and 3′ homologous region of *hem*A gene amplicon with 665 bp and 395 bp DNA band, respectively on 1% Agarose gel electrophoresis (Figure 4.21) as expected.



Lane 1: 1kb DNA ladder
Lane 2: 5′ homologous region of *hem*A (665 bp)
Lane 3: 3′ homologous region of *hem*A (395 bp)

Figure 4.21: PCR product of 5′ and 3′ homologous region of *hem*A

4.5.3 Construct containing 5'homologous region, stabilized *hem*A gene and 3' homologous region in pUC∆NdeI vector

The upstream and downstream regions of *hem*A gene have been successfully amplified and *hem*A gene has been amplified and cloned in pAM102. In this newly constructed pUCΔNdeI, first 5′ homologous region would be sub-cloned between *Bam*HI and *Hind*III restriction sites as described above. During cloning of 5-′ homologous region *Nde*I site would also be created because it has been inserted at 3′-end of the reverse primer. After this sub-cloning then the newly developed plasmid containing 5′-end of *hem*A would be digested with *Nde*I and *Eco*RI. The amplified stabilized *hem*A would be digested with *Eco*RI and partially digested by *Nde*I as described above. Then this stabilized *hem*A gene would be cloned between *Nde*I and *Eco*RI recognition sites that has *Xba*I site preceeding *Eco*RI recognition site in *hem*A gene reverse primer. Finally, the 3′- homologous region would be cloned between *Xba*I and *Eco*RI recognition sites. Once the construct is completed then this cassette with stabilized *hem*A gene flanked by homologous regions at both 5′- and 3′-ends (Figure 4.22) would be amplified by reamplification primers that would give a linear cassette with at least 100 bp homologous regions at both ends of *hem*A gene. This linear cassette would be then electroporated in competent *E. coli*

HME5 *tna<>cat* strain that expressed Exo, Beta and Gam products (Yu *et al.,* 2000) to swap the native *hem*A gene with stabilized version of *hem*A gene as that has been done of tryptophanase A gene (Wu *et al.,* 2005).

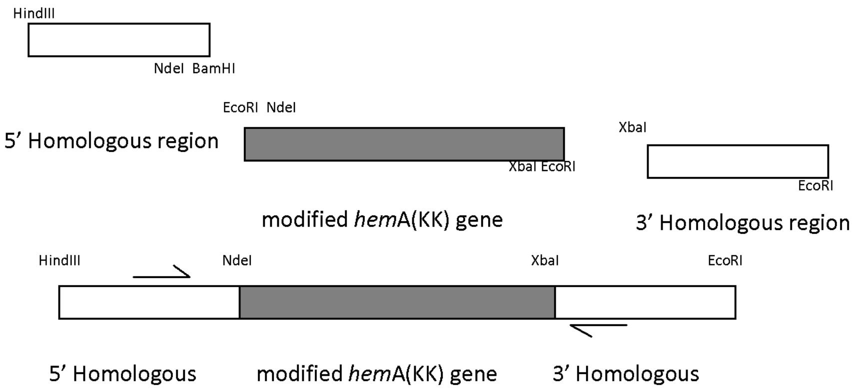


Figure 4.22: Cloning strategy to design the construct containing 5′ homologous region, modified mutant *hem*A(KK) gene and 3′ homologous region of *hem*A gene for chromosomal engineering

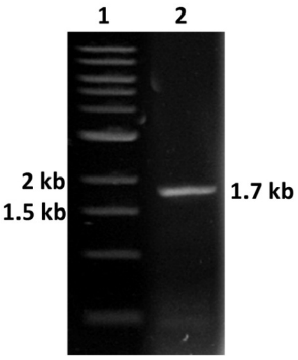
For homologous recombination linear DNA is required (Yu *et al.,* 2000) or self integrating plasmid, however, pUC19 is not self integrating plasmid thus DNA swapping has to be with linear DNA obtained either by reamplification or digested with restriction enzymes. Restriction digestion would have flanking sequences as cohesive ends and these sequences are not homologous to chromosomal DNA and for homologous recombination the DNA sequences at both 5′- and 3′- ends has to be at least 30-50 nt long to recombine in this bacteria. Alternative mechanism to get such linear DNA with 30-50 nt long at both 5′- and 3′- ends through PCR amplification and blunting and Phusion polymerase has this blunting capacity which gives PCR product without flanking hang making the amplified DNA compatible for homologous recombination. Using this Phusion polymerase and primers designed complementary to homologous sequences at least 100 bp in both forward and reverse would give amplicon that can be directly electroporated for chromosomal integration after purification of PCR product.

Although single swapping could have been feasible by amplifying *hem*A gene with mutation inserted and homologous flanking regions designed in primers as that has been done to amplify *cat* gene (Wu *et al.,* 2005), but generation of random mutant library (Nakamura *et al.,* 2001; Kim *et al.,* 2005) using pAM102 and subsequent amplification of mutant *hem*A gene library would be easier from further cloning to pUC19ΔNdeI plasmid than using primers with flanking region. This is because such primers would have long size, high Tm and possibility of secondary structures between primers that would make amplification complicated and would be difficult to use the primer for cloning. Thus linear reamplification primer set was designed without any restriction sites that would amplify *hem*A gene with homologous region at both 5'- and 3'- regions of *hem*A gene.

Furthermore, protein engineering of *hem*A gene by random or semi-random process can be performed to obtain the gain of function mutants that would augment endogenous supplementation of ALA for heme biosynthesis. The internal *Nde*I recognition site could be mutated by site directed mutagenesis without altering the coding amino acids, and random mutant library generated could be cloned in pAM102 construct by digesting with *Nde*I and *Eco*RI for effective screening by indole metabolism (Nakamura *et al.,* 2001; Kim *et al.,* 2002; 2005). Once the mutant is validated then mutated *hem*A gene could be redigested with *Nde*I and *Xba*I and cloned in the vector carrying cassette. Then this cassette could be reamplified using reamplification primer and used for chromosomal engineering. This is presumed to ultimately increase the expression of functionally active cytochrome P450s levels. After chromosomal engineering of *hem*A gene, plasmid based *hem*L gene expression could be carried out by cloning in pAM101 that would be expressed under *cat* promoter to effectively give sufficient amount of ALA to have reasonable amount of hemeproteins expressed.

4.5.4 Amplification of linear *hem*A gene containing 5'- and 3'- homologous regions

From the construct spanning about 2.35 kb linear segment consisting of 5' homologous region, stabilized *hem*A gene and 3' homologous region of *hem*A gene in pUC19ΔNdeI plasmid, the reamplification primer amplifies around 1.7 kb DNA segment. In order to confirm that the designed reamplification primers can successfully amplify the *hem*A gene with 5' and 3' homologous region, *E. coli* HME5 *tna<>cat* genomic DNA was used as template for PCR. Observation of 1.7 kb DNA fragment on agarose gel electrophoresis confirmed the successful amplification of *hem*A gene with 5' and 3' homologous region (Figure 4.23).



Lane 1: 1 kb NEB DNA ladder
Lane 2: PCR product of *hem*A gene with homologous region

Figure 4.23: *hem*A gene amplification with 5' and 3' homologous region

Conclusion

Normally, HemA protein is susceptible to heme regulated proteolytic cleavage mediated by Lon/ClpAP energy dependent proteases that regulates ALA biosynthesis which is rate limiting precursor for heme synthesis. However, introduction of two positively charged lysine residues close to N-terminus of HemA protein would render it insensitive towards retro-inhibition due to heme, most probably the positively charged amino acids would render repulsive force to positively charged Fe^{++} ion of heme thus preventing docking of it in the protein structure that makes the flanking 18 amino acids intact in the structure than otherwise which could have been protruded making it sensitive to proteolytic cleavage.

Thus, in *E. coli* the heterologous expression of cytochrome P450s is mainly compromised due to limitation in heme precursors that would be necessary for proper folding of cytochrome P450 enzymes. The newly synthesized heterologous hemeproteins have relatively lower affinity for heme than that of intrinsic cellular hemeproteins and are only expressed as apo-protein in functionally inactive form. Thus, expression of functionally active (holo-protein) cytochrome P450s are possible only with increased heme biosynthesis that has been modulated by exogenous supplementation of expensive ALA precursor. In the present study, in order to circumvent the dependence of exogenous ALA supplementation for heterologous expression of cytochrome P450s, a host strain *E. coli* AM102/2A6 was developed which co-expresses stabilized *hem*A gene (from plasmid pAM102 that has *hem*A gene cloned in under *cat* promoter) that encodes feedback insensitive glutamyl tRNA reductase [HemA(KK)], a rate limiting enzyme in the ALA biosynthetic pathway and ALA is rate limiting precursor that is critical for heme biosynthesis.

With this strategy, newly developed host strain *E. coli* AM102/2A6 was able to support heterologous expression of cytochrome P450 2A6 double mutant L240C/N297Q in functionally active form as evident from relatively higher bacterial survival rate of the host strain as compared to control strain *E. coli* AM101/2A6 (without *hem*A) when cultured in presence of toxic level of indole (5 mM). The stabilized HemA protein endogenously supplemented ALA and subsequently heme cofactor that facilitated cytochrome P450 2A6 to reconstitute into active holo-protein. The expressed catalytically active cytochrome P450 2A6 metabolized indole to non-toxic indoxyl products decreasing indole toxicity which was corroborated with higher percentage of bacterial survival rate of newly developed host strain.

The expression of stabilized *hem*A gene by inducing with chloramphenicol before the induction of cytochrome P450 2A6 by IPTG resulted in higher proportions of bacterial survival rate compared with parallel gene induction of *hem*A and cytochrome P450 2A6 after four hours. Although exact reason for this phenomenon could not be explained from present works, but it can be assumed that most probably addition of cholramphenicol might have supported higher intrinsic hemeprotein production that increased the bacterial concentration at 4 hours of incubation as observed in figure 4.16, most probably due to continuous supply of heme precursors.

Thus newly developed host strain *E. coli* AM102/2A6 can be extended for heterologous expression of different cytochrome P450 isoforms and also aid in production of tetrapyrrole containing biological molecules including vitamin B_{12}, heme-containing enzymes (siroheme, cytochrome), porphyrins etc. The expression of stabilized *hem*A gene can potentially be used for ALA production that can be further enhanced for commercial scale production with metabolic and genetic engineering strategies. The protein engineering of *hem*A gene can be envisaged for obtaining gain of function protein mutant and stable integration of mutant *hem*A gene in chromosomal DNA in homologous recombination efficient *E. coli* HME5 *tna<>cat* strain would develop the robust host strain for expression of tetrapyrrole containing biological molecules by simple and cost effective approach.

REFERENCES

Adachi J, Mori Y, Matsui S, Takigami H and Fujino J (2001) Indirubin and Indigo are potent aryl hydrocarbon receptor ligands present in human urine. *J Biol Chem*. **276**: 31475-31478

Alexieva Z, Duvall EJ, Ambulos Jr. NP, Kim UJ and Lovett PS (1988) Chloramphenicol induction of *cat-86* requires ribosome stalling at a specific site in the leader. *Proc Natl Acad Sci USA*. **85**: 3057-3061

Alton NK and Vapnek D (1979) Nucleotide sequence analysis of the chloramphenicol resistance transposon Tn*9*. *Nature*. **282**: 864-866

Ames BN, Lee FD and Durston WE (1973) An improved bacterial test system for the detection and classification of mutagens and carcinogens. *Proc Natl Acad Sci USA*. **70**: 782-786

Aryal P, Terashita T, Guengerich FP, Shimada T and Oda Y (2000) Use of genetically engineered *Salmonella typhimurium* OY1002/1A2 strain coexpressing human cytochrome P450 1A2 and NADPH-cytochrome P450 reductase and bacterial O-acetyltransferase in SOS/umu assay. *Environ Mol Mutagen.* **36**(2): 121-126

Aryal P, Yoshikawa K, Terashita T, Guengerich FP, Shimada T and Oda Y (1999) Development of a new genotoxicity test system with *Salmonella typhimurium* OY1001/1A2 expressing human P450 1A2 and NADPH-P450 reductase. *Mutat Res.* **442**: 113-120

Avissar YJ and Moberg PA (1995) The common origins of the pigments of life- early steps of chlorophyll biosynthesis. *Photosynthesis Research*. **44**: 221-242

Avissar YJ, Ormerod JG and Beale SI (1989) Distribution of delta-Aminolevulinic acid biosynthetic pathways among phototrophic bacterial groups. *Arch Microbiol.* **151**: 513–519

Badyal DK and Dadhich AP (2001) Cytochrome P450 and drug interactions. *Indian J Pharmocol.* **33**: 248-259

Barnes HJ, Arlotto MP and Waterman MR (1991) Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*. *Proc Natl Acad Sci USA.* **88**: 5597-5601

Beale S and Weinstein (1990) Tetrapyrrole Metabolism in Photosynthetic Organisms. McGrawHill, New York.

Beale SI (1996) Biosynthesis of hemes. In: *Escherichia coli* and Salmonella: cellular and molecular biology. Frederick C and Neidhardt (Eds) *ASM* (Pub Washington, DC) pp 731-748

Benzinger R, Enquist LW and Skalka A (1975) Transfection of *Escherichia coli* spheroplasts. V. Activity of recBC nuclease in rec+ and rec minus spheroplasts measured with different forms of bacteriophage DNA. *J Virol.* **15**(4): 861-871

Bernardi A and Bernardi F (1984) Complete sequence of pSC101. *Nucleic Acids Res*. **12**(24): 9415-9426

Berry A, Dodge TC, Pepsin MP and Weyler W (2002) Application of metabolic engineering to improve both the production and use of biotech indigo. *J Ind Microbiol Biotechnol.* **28**: 127–133

Bhowmick R and Girotti AW (2010) Cytoprotective induction of nitric oxide synthase in a cellular model of 5-aminolevulinic acid based photodynamic therapy. *Free Radical Biol Med.* **48**: 1296-1301

Biedendieck R, Malten M, Barg H, Bunk B, Martens JH, Deery E, Leech H, Warren MJ and Jahn D (2010) Metabolic engineering of cobalamin (vitamin B12) production in *Bacillus megaterium*. *Microb Biotechnol.* **3**(1): 24-37

Bistolas N, Wollenberger U, Jung C and Scheller FW (2005) Cytochrome P450 biosensors-A Review. *Biosens Bioelectron.* **20**: 2408-2423

Burns RO and DeMoss RD (1962) Properties of tryptophanase from *Escherichia coli*. *Biochim Biophys Acta*. **65**: 233-244

Cadwell RC and Joyce GF (1992) Randomization of genes by PCR mutagenesis. PCR *Methods Appl.* **2**(1): 28–33

Capdevila J, Holla VR and Falck JR (2005) Cytochrome P450 and the metabolism and bioactivation of arachidonic acid and eicosanoids, In: Montellano PR (Ed.), *Cytochrome P450 Structure, Mechanism and Biochemistry*. (Pub Kluwer Academic/Plenum Plublishing, New York) pp 531-551

Carter DM and Radding CM (1971) The role of exonuclease and beta protein of phage lambda in genetic recombination. II. Substrate specificity and the mode of action of lambda exonuclease. *J Biol Chem*. **246**(8): 2502-2512

Carugo O and Argos P (1997) NADP-dependent enzymes. I. Conserved stereochemistry of cofactor binding. *Protiens Struct Genet.* **28**: 10-28

Chang ACY and Cohen SN (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. *J Bacteriol.* **134**(3): 1141-1156

Chen W, Russell CS, Murooka Y and Cosloy SD (1994) 5-Aminolevulinic acid synthesis in *Escherichia coli* requires expression of *hemA*. *J Bacteriol.* **176**: 2743-2746

Choi P, Wang L, Archer CD and Elliott T (1996) Transcription of the glutamyl-tRNA reductase (*hem*A) gene in *Salmonella typhimurium* and *Escherichia coli*: role of the *hem*A P1 promoter and the *arc*A gene product. *J Bacteriol.* **178**(3): 638-646

Chudaev MV and Usanov SA (1997) Expression of functionally active cytochrome b_5 in *Escherichia coli*: isolation, purification, and use of the immobilized recombinant heme protien for affinity chromatography of electron transfer proteins. *Biochem.* **62**: 401-411

Dailey HA (ed.) (1990) Biosynthesis of heme and chlorophylls. Mc Graw-Hill Publishing Co., Hightstown, NJ. USA.

Danielson PB (2002) The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. *Curr Drug Metab*. **3**(6): 561-597

Darie S and Gunsalus RP (1994) Effect of heme and oxygen availability on *hemA* gene expression in *Escherichia coli*: role of the *fnr, arcA*, and *himA* gene products. *J Bacteriol*. **176**: 5270-5276

Doss MO and Philipp-Dornston WK (1971) Porphyrin and heme biosynthesis from endogenous and exogenous δ-aminolevulinic acid in *Escherichia coli, Pseudomonas aeruginosa* and *Achromobacter metalcaligenes*. *Hoppe-Seylers's Z Physiol Chem*. **352**: 725-733

Duvall EJ and Lovett PS (1986) Chloramphenicol induces translation of the mRNA for a chloramphenicol resistance gene in *Bacillus subtilis*. *Proc Natl Acad Sci USA.* **83**: 3939-3943

Edwards S, Jackson D, Reynoldson J and Shanley B (1984) Neuropharmacology of delta-aminolevulinic acid II. Effect of chronic administration in mice. *Neurosci Lett.* **50**:169-173

Elizabeth JR and Lovett PS (1990) Erythromycin Induces Expression of the Chloramphenicol Acetyltransferase Gene *cat-86. J Bacteriol.* **172**(8): 4694-4695

Ellis HM, Yu D, DiTizio T and Court DL (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc Natl Acad Sci USA*. **98**(12): 6742-6746

Ferreira GC (1999) 5-Aminolevulinate synthase and mammalian heme biosynthesis. In: Ferreira GC, Moura JJG and Franco R ed. *Iron Metabolism: Inorganic biochemistry and regulatory mechanisms* (Pub Wiley-VCH) pp. 15-34

Fisher CW, Caudle DL, Martin-Wixtrom C, Quattrochi LC, Tukey RH, Waterman MR and Estabrook RW (1992) High-level expression of functional human cytochrome P450 1a2 in *Escherichia coli*. *FASEB J.* **6**: 759-764

Fisher CW, Shet MS, Caudle DL, Martin-Wixtrom and Estabrook RW (1992) High-level expression in *Escherichia coli* of enzymatically active fusion proteins containing the domains of mammalian cytochromes P450 and NADPH-P450 reductase flavoprotein. *Proc Natl Acad Sci USA*. **89**: 10817-10821

Flores S, de Anda-Herrera R, Gosset G and Bolivar FG (2004) Growth-rate recovery of *Escherichia coli* cultures carrying a multicopy plasmid, by engineering of the pentose-phosphate pathway. *Biotechnol Bioeng*. **87**(4): 485-494

Fujita K and Kamataki T (2002) Genetically engineered bacterial cells co-expressing human cytochrome P450 with NADPH cytochrome P450 reductase: prediction of metabolism and toxicity of drugs in humans. *Drug Metab Pharmacokinet.* **17**: 1-22

Fujita M, Tanaka T and Hotta Y (1997) Additive for animal cell culture. *Jpn Kokai Tokkyo Koho JP* 2598726

Gillam EMJ (2008) Engineering cytochrome p450 enzymes. *Chem Res Toxicol.* **21**(1): 220-231

Gillam EMJ, Aguinaldo AM, Notley LM, Kim D, Mundkowski RG, Volkov A, Arnold FH, Soucek P, DeVoss JJ and Guengerich FP (1999) Formation of indigo by recombinant mammalian cytochrome P450. *Biochem Biophys Res Commun.* **265**(2): 469-472

Gillam EMJ, Notley LM, Cai H, Voss J and Guengerich FP (2000) Oxidation of Indole by Cytochrome P450 Enzymes. *Biochem.* **39**: 13817-13824

Gotoh O (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and nucleotide sequences. *J Biol Chem.* **276**(1): 83-90

Grant WE, Hooper C, MacRobert AJ, Speight PM and Bown SG (1993) Photodynamic therapy of oral cancer photosentitisation with systemic aminolevulinic acid. *Lancet.* **342**: 147-148

Greer H (1975) The *kil* gene of bacteriophage lambda. *Virology*. **66**(2): 589-604

Guengerich FP (2002) Cytochrome P450 enzymes in the generation of commercial products. *Nat Rev Drug Discovery.* **1**: 359-366

Guengerich FP, Martin MV, Guo Z and Chun YJ (1996) Purification of recombinant human cytochrome P450 enzymes expressed in bacteria. *Methods Enzymol*. **272**: 35-44

Guengerich FP, Sorrells JL, Schmitt S, Krauser J, Aryal P and Meijer L (2004) Generation of New Protein Kinase Inhibitors Utilizing Cytochrome P450 Mutant Enzymes for Indigoid Synthesis. *J Med Chem*. **47**: 3236-3241

Guengerich FP, Wheeler JB, Chun YJ, Kim D, Shimada T, Aryal P, Oda Y and Gillam EMJ (2002) Use of heterologously-expressed cytochrome P450 and glutathione transferase enzymes in toxicity assays. *Toxicol.* **181–182**: 261–264

Hahn FE, Wisseman CL and Hopps HE (1995) Mode of action of chloramphenicol. III. Action of chloramphenicol on bacterial energy metabolism. *J Bacteriol.* **69**(2): 215-223

Harnastai IN, Gilep AA and Usanov SA (2006) The development of an efficient system for heterologous expression of cytochrome P450s in *Escherichia coli* using *hem*A gene co-expression. *Protein Expr Purif*. **46**(1): 47-55

Hasler JA (1999) Pharmacogenetics of cytochromes P450. *Mol Aspects Med*. **20**(1-2): 12-24, 25-137

Hoessel R, Leclerc S, Endicott JA, Nobel ME, Lawrie A, Tunnah P, Leost M, Damiens E, Marie D, Marko D, Niederberger E, Tang W, Eisenbrand G and Meijer L (1999) Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. *Nat Cell Biol*. **1**(1): 60-67

Hotta Y and Watnabe K (1999) Plant growth-regulating activities of 5-aminolevulinic acid. *Syokubutu-no-Kagaku-Tyousetu (Chemical regulation of plants)*. **34**: 85-96

Ilag LL and Jahn D (1992) Activity and spectroscopic properties of the *Escherichia coli* glutamate 1-semialdehyde aminotransferase and the putative active site mutant K265R. *Biochem*. **31**: 7143-7151

Imai T, Globerman JM, Gertner JM, Kagawa N and Waterman MR (1993) Expression and purification of functional human 17 α-hydroxylase /17, 20-lyase (P450c17) in *Escherichia coli. J Biol Chem.* **268**: 199681-199689

Inoue E, Takahashi Y, Imai Y and Kamataki T (2000) Development of bacterial expression system with high yield of CYP3A7, a human fetus-specific form of cytochrome P450. *Biochem Biophys Res Commun*. **269**(2): 623-627

Ishino A, Magara A, Tajima M, Tuji Y, Takahashi K, Tanaka T and Hotta Y (1999) External medicine for head hair. *Jpn Kakai Tokkyo Koho: Toku Kai Hei* 11-116446

Jacobs NJ, Jacobs JM and Morgan HEJ (1972) Comparative effect of oxygen and nitrate on protoporphyrin and heme synthesis from delta-aminolevulinic acid in bacterial cultures. *J Bacteriol.* **112**: 1444-1445

Jahn D, Michelsen U and Söll D (1991) Two glutamyl-tRNA reductase activities in *Escherichia coli. J Biol Chem.* **266**(4): 2542–2548

Jahn D, Verkamp E and Söll D (1992) Glutamyl- transfer RNA: a precursor of heme and chlorophyll biosynthesis. *Trends Biochem Sci.* **17**: 215-218

Javor GT and Febre EF (1992) Enzymatic basis of thiol-stimulated secretion of porphyrins by *Escherichia coli. J Bacteriol.* **174**(3): 1072-1075

Jones AM and Elliott T (2010) A purifed mutant HemA protein from *Salmonella enterica* serovar Typhimurium lacks bound heme and is defective for heme-mediated regulation in vivo. *FEMS Microbiol Lett.* **307:** 41–47

Jordan PM (1991) Biosynthesis of tetrapyrroles. In: Neuberger A and van Deenen LLM (ed) *New comprehensive biochemistry*. Vol. 19 (Elsevier, Amsterdam).

Josephy PD, Evans DH, Parikh A and Guengerich FP (1998) Metabolic activation of aromatic amine mutagen by simultaneous expression of human P450 1A2, NADPH-Cytochrome P450 reductase, and *N*-acetyltransferase in *Escherichia coli. Chem Res Toxicology.* **11**: 70-74

Kagawa N, Cao Q and Kusano K (2003) Expression of human aromatase (CYP19) in *Escherichia coli* by N-terminal replacement and induction of cold stress response. *Steroids.* **68**: 205-209

Kamiyama H, Hotta Y, Tanaka T, Nishizawa Y and Sasaki K (2000) Production of 5-aminolevulinc acid by a mutant strain of a photosynthetic bacteria. *Seibutu-Kougaku.* **78**:48-55

Kaneko S, Aoki T, Nanato H, Miyoshi N, Houki S and Fukuda Y (1998) Intraoperative photodynamibc diagnosis of human glioma using 5-ALA induced protoporphyrin IX. *Iwamizawa-sir-itu Sougou Byounin-shi.* **24**: 71-79

Kang Z, Wang Y, Gu P, Wang Q and Qi Q (2011) Engineering *Escherichia coli* for efficient production of 5-aminolevulinic acid from glucose. *Metab Eng.* **13**(5): 492-498

Karu AE, Sakaki Y, Echols H and Linn S (1975) The gamma protein specified by bacteriophage lambda. Structure and inhibitory activity for the recBC enzyme of *Escherichia coli. J Biol Chem.* **250**(18): 7377-7387

Kim D and Guengerich FP (2002) Random Mutagenesis by Whole-Plasmid PCR Amplification. *Methods Mol Biol.* **192**: 241-45

Kim D, Wu ZL and Guengerich FP (2005) Analysis of Coumarin 7-Hydroxylation Activity of Cytochrome P450 using Random Mutagenesis. *J Biol Chem.* **280**(48): 40319-40327

Knockaert M, Greengard P and Meijer L (2002) Pharmacological inhibitors of cyclin-dependent kinases. *Trends Pharmacol Sci*. **23**: 417-425

Kourilsky P and Knapp A (1974) Lysogenization by bacteriophage lambda. III. Multiplicity dependent phenomena occuring upon infection by lambda. *Biochimie.* **56**(11-12): 1517-1523

Kranendonk M, Carreira F, Theisen P, Laires A, Fisher CW, Rueff J, Estabrook RW and Vermeulen NP (1999) *Escherichia coli* mtc, a human NAPH P450 reductase competent mutagenicity tester strain for the expression of human cytochrome P450 isoforms 1A1, 1A2, 2A6, 3A4, or 3A5: catalytic activities and mutagenicity studies. *Mutat Res*. **441**: 73-83

Kuramochi H, Konnai M, Tanaka T and Hotta Y (1997) Method for improving plant salt tolerance. *US Patent No*. 5 661 111

Kusano K, Waterman MR, Sakaguchi M, Omura T and Kagawa N (1999) Protein synthesis inhibitors and ethanol selectively enhance heterologous expression of P450s and related proteins in *Escherichia coli*. *Arch Biochem Biophys*. **367**(1): 129-136

Leclerc S, Garnier M, Hoessel R, Marko D, Bibb JA, Snyder GL, Greengard P, Biemat J, Wu YZ, Mandelkow, Eisenbrand G and Meijer L (2001) Indirubins inhibit glycogen synthase kinase-3 beta and CDK5/p25, two protein kinases involved in abnormal tau phosphorylatin in Alzheimer's disease. A property common to most cyclin-dependent kinase inhibitors. *J Biol Chem*. **276**(1): 251-260

Levy RH (1995) Cytochrome P450 isozymes and antiepileptic drug interactions. *Epilepsia*. **36** Suppl 5: 8S-S13

Li P, Tong L, Liu K and Wang Y (2009) Indole degrading of ammonia oxidizing bacteria isolated from swine wastewater treatment system. *Water Sci Technol*. **59**(12): 2405-2410

Little JW (1967) An exonuclease induced by bacteriophage lambda. II. Nature of the enzymatic reaction. *J Biol Chem*. **242**(4): 679-686

Liu H and Naismith JH (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol*. **8**: 91

Lovett PS (1990) Translational Attenuation as the Regulator of Inducible *cat* Genes. *J Bacteriol.* **172**(1): 1-6

Luer C, Schauer S, Mobius K, Schulze J, Schubert WD, Heinz DW, Jahn D and Moser J (2005) Complex formation between glutamyl-tRNA reductase and glutamate-1-semialdehyde 2,1-aminomutase in *Escherichia coli* during the initial reactions of porphyrin biosynthesis. *J Biol Chem*. **280**(19): 18568-18572

Malke H and Ferretti JJ (1985) Expression in *Escherichia coli* of streptococcal plasmid-determined erythromycin resistance directed by the *cat* gene promoter of pACYC 184. *J Basic Microbiol.* **25**(6): 393-400

Masuda T, Tanaka R, Shioi Y, Takamiya K, Kannangara CG and Tsuji H (1994) Mechanism of benzyl-adenine-induced stimulation of the synthesis of 5-aminolevulinic acid in greening cucumber cotyledons: benzyladenine increases level of plasmid tRNAGlu. *Plant Cell Physiol*. **35**: 183-188

McFadyen MCE, Melvin WT and Murray GI (2004) Cytochrome P450 enzymes: Novel options for cancer therapeutics. *Mol Cancer Ther.* **3**: 363-371

McNicholas PM, Javor G, Darie S and Gunsalus RP (1997) Expression of the heme biosynthetic pathway genes *hemCD*, *hemH*, *hemM*, and *hemA* of *Escherichia coli*. *FEMS Microbiol Lett.* **146**: 143-148

Melville SB and Gunsalus RP (1996) Isolation of an oxygen-sensitive FNR protein of *Escherichia coli*: interaction at activator and repressor sites of FNR-controlled genes. *Proc Natl Acad Sci USA.* **93**(3): 1226-1231

Mikolajewska P, Donnelly RF, Garland MJ, Morrow DI, Singh TR, Iani V, Moan J and Juzeniene A (2010) Microneedle pre-treatment of human skin improves 5-aminolevulininc acid (ALA)- and 5-aminolevulinic acid methylester (MAL)-induced PpIX production for topical photodynamic therapy without increase in pain or erythema. *Pharm Res.* **27**:2213-2220

Miyachi N, Tanaka T, Nishikawa S, Takeya H and Hotta Y (1998) Preparation and chemical properties of 5-aminolevulinic acid and its derivatives. *Porphyrins.* **7**: 342-347

Miyazaki K and Misa T (2002) Creating Random Mutagenesis Libraries Using Megaprimer PCR of Whole Plasmid. *BioTechniques.* **33**: 1033-1038

Moser J, Lorenz S, Hubschwerlen C, Rompf A and Jahn D (1999) *Methanopyrus kandleri* glutamyl-tRNA reductase. *J Biol Chem.* **274**(43): 30679-30685

Moser J, Schubert WD, Beier V, Bringemeier I, Jahn D and Heinz DW (2001) V-shaped structure of glutamyl-tRNA reductase, the first enzyme of tRNA-dependent tetrapyrrole biosynthesis. *EMBO J.* **20**(23): 6583-6590

Murakami H, Hohsaka T and Sisido M (2000) Random insertion and deletion mutagenesis for construction of protein library containing non natural amino acids. *Nucleic Acids Symposium Series* **44**: 69-70

Murphy KC (1991) Lambda Gam protein inhibits the helicase and chi-stimulated recombination activities of *Escherichia coli* RecBCD enzyme. *J Bacteriol.* **173**(18): 5808-5821

Nakahigashi K, Nishimura K, Miyamoto K and Inokunchi H (1991) Photosensitivity of a protoporphyrin-accumulating, light-sensitive mutant (visA) of *Escherichia coli* K-12 *Proc Natl Acad Sci USA.* **88**: 10520-10524

Nakamura K, Martin MV and Guengerich FP (2001) Random Mutagenesis of Human Cytochrome P450 2A6 and Screening with Indole Oxidation Products. *Arch Biochem Biophys.* **395**: 25-31

Nakayashiki T and Inokuchi H (1996) Control of the availability of exogenous 5-aminolevulinic acid in *Escherichia coli*. *Genes Genet Syst.* **71**: 237-241

Nakayashiki T, Nishimura K and Inokuchi H (1995) Cloning and sequencing of a previously unidentified gene that is involved in the biosynthesis of heme in *Escherichia coli*. *Gene.* **153**(1): 67-70

Nakayashiki T, Nishimura K, Tanaka R and Inokuchi H (1995) Partial inhibition of protein synthesis accelerates the synthesis of porphyrin in heme-deficient mutants of *Escherichia coli*. *Mol Gen Genet*. **249**(2): 139-146

Nebert DW, Nelson DR, Adesnik M, Coon MJ, Estabrook RW, Gonzalez FJ, Guengerich FP, Gunsalus IC, Johnson EF and Kemper B (1989) The P450 superfamily: update listing of all genes and recommended nomenclature. *DNA Cell Biol*. **10**: 1-14

Nebert DW, Nelson DR, Coon MJ, Estabrook RW, Feyereisen R, Fujii-Kuriyama Y, Gonzalez FJ, Guengerich FP, Gunsalus IC and Johnson EF (1991) The P450 superfamily: update on new sequences, gene mapping and recommended nomenclature. *DNA Cell Biol*. **10**:1-14

Nebert DW and Russell DW (2002) Clinical importance of the cytochromes P450. *Lancet*. **360**: 1155-1162

Neidle EL and Kaplan S (1993) Expression of the *Rhodobactoer sphaeroides hemA* and *hem*T genes, encoding two 5-aminolevulinc acid synthesis isozymes. *J Bacteriol.* **175**: 2292-2303

Nelson DR, Kamataki T, Waxman DJ, Guengerich FP, Estabrook RW, Feyereisen R, Gonzalez FJ, Coon MJ, Gunsalus IC and Gotoh O (1993) The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trival names of enzymes and nomenclature. *DNA Cell Biol.* **12**: 1-51

Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC and Nebert DW (1996) P450 supefamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics*. **6**: 1-42

Nicholas P, Ambulos JR, Elizabeth JR, Zlatka A and Lovett PS (1988) Induction of *cat-86* by Chloramphenicol and Amino Acid Starvation in Relaxed Mutants of *Bacillus subtilis. J Bacteriol***.** **170**(12): 6542-5646

Nishikawa Y, Watanabe K, Tanaka T, Miyachi N, Hotta Y and Murooka Y (1999) *Rhodobacter sphaeroides* mutant which accumulate 5-aminolevulinic acid under aerobic and dark conditions. *J Biosci Bioeng*. **87**(6):798-804

Oda Y, Aryal P, Terashita T, Gillam EM, Guengerich FP and Shimada T (2001) Metabolic activation of heterocyclic amines and other procarcinogens in *Salmonella typhimurium* umu tester strains expressing human cytochrome P4501A1, 1A2, 1B1, 2C9, 2D6, 2E1, and 3A4 and human NADPH-P450 reductase and bacterial O-acetyltransferase. *Mutat Res.* **492**(1-2): 81-90

Oda Y, Nakamura S, Oki I, Kato T and Shinagawa H (1985) Evaluation of a new test system (umu-test) for the detection of environmental mutagens and carcinogens. *Mutat Res*. **47:** 219-229

Ohtsubo H and Ohtsubo E (1978) Nucleotide sequence of an insertion element, IS1. *Proc Natl Acad Sci USA*. **75**(2): 615-619

Omura T and Sato R (1964) The carbon monoxide-binding pigment of the liver microsomes. I . Evidence for its hemoprotien nature. *J Biol Chem*. **239**: 2370-2378

Onuki J, Rech CM, Medeiros MH, De AUG and Mascio PDi (2002) Genotoxicity of 5-aminolevulinic and 4,5-dioxovaleric acids in the *Salmonella*/microsuspension mutagenicity assay and sos chromotest. *Enivoron Mol Mutagen*. **40**: 63-70

Orr ST, Ripp SL, Ballard TE, Henderson JL, Scott DO, Obach RS, Sun H and Kalgutkar AS (2012) Mechanism-Based Inactivation (MBI) of Cytochrome P450 Enzymes: Structure-Activity Relationships and Discovery Strategies To Mitigate Drug-Drug Interaction Risks. *J Med Chem.* In press.

Parikh A, Gillam EMJ and Guengerich FP (1997) Drug metabolism by *Escherichia coli* expressing human cytochromes P450. *Nat Biotechnol*. **15**: 784-88

Patterson TA , Costantino N, Dasgupta S and Court DL (1993) Improved bacterial hosts for regulated expression of genes from lambda pL plasmid vectors. *Gene*. **132**: 83-87

Pernecky SJ, Larson JR, Philpot RM and Coon MJ (1993) Expression of truncated forms of liver microsomal P450 cytochromes 2b4 and 2e1 in *Escherichia coli*: influence of nh2-terminal region on localization in cytosol and membranes. *Proc Natl Acad Sci.* **90**: 2651-2655

Peter W, Fink-Puches R, Cerroni L and Kerl H (1994) Photodynamic therapy for mycosis fungicides after topical photosensitization with 5-aminolevulinic acid. *J Am Acad Dermatology.* **31**: 678-680

Philipp-Dornston WK and Doss MO (1973) Comparision of porphyrin and heme in various heterotrophic bacteria. *Enzyme.* **16**: 57-64

Polychronopoulos P, Magiatis P, Skaltsounis AL, Myrianthopoulos V, Mikros E, Tarricone A, Musacchio A, Roe SM, Pearl L, Leost M, Greengard P and Meijer L (2004) Structural basis for the synthesis of indirubins as potent and selective inhibitors of glycogen synthase kinase-3 and cyclin-dependent kinases. *J Med Chem*. **47**(4): 935-946

Prentki P, Karch F, Iida S and Meyer J (1981) The plasmid cloning vector pBR325 contains a 482 base-pair-long inverted duplication. *Gene*. **14**(4): 289-299

Rebeiz CA, Juvik JA and Rebeiz CC (1988) Porphyric insecticides. I. Concept and phenomenology. *Pestic Biochem Physiol*. **30**: 11-27

Rebeiz CA, Montazer-Zouhool A, Hopen H and Wu SM (1984) Photodynamic herbicides. I Concepts and phenomenology. *Enzyme Microb Technol.* **6**: 390-401

Reifferscheid G, Heil J, Oda Y and Zahn RK (1991) A microplate version of the SOS/umu-test for rapid detection of genotoxines and genotoxic potentials as environmental samples. *Mutat Res*. **253**: 215-222

Roberts GC (1999) The power of evolution: accessing the synthetic potential of P450s. *Chem Biol*. **6**(10): R269-R272

Rogers EJ and Lovett PS (1990) Erythromycin induces expression of the chloramphenicol acetyltransferase gene *cat*-86. *J Bacteriol*. **172**(8): 4694-4695

Rose RE (1988) The nucleotide sequence of pACYC184. *Nucleic Acids Res*. **16**(1): 355

Rosic NN (2009) Versatile capacity of shuffled cytochrome P450s for dye production. *Appl Microbiol Biotechnol*. **82**(2): 203-210

Sambrook J and Russell DW (2001) Molecular Cloning- A Laboratory Manual Third Edition Volume 1-3 Cold Spring Harbor Laboratory Press, Cold Spring Habor, New York.

Sasaki K, Ikeda S, Nishizawa Y and Hayashi M (1987) Production of 5-aminolevulinc acid by photosynthesis bacteria. *J Ferment Technol.* **65**: 511-515

Sasaki K, Ikeda S, Nishizawa Y and Hayashi M (1990) Production of a herbicide, 5-aminolevulinc acid. By *Rhodobacter sphaeroides* using the effluent waste from an anaerobic digestor. *Appl Microbiol Biotechnol.* **32**: 727-731

Sasaki K, Tanaka T and Nagai S (1998) Use of photosynthetic bacteria for the production of SCP and chemicals from organic wastes. In Bioconversion of waste materials to industrial products, second edition. Martin AM eds (Blackie Academic and professionals) pp. 247-291

Sasaki K, Watanabe M and Tanaka T (2002) Biosynthesis, biotechnological production and applications of 5-aminolevulinic acid. *Appl Microbiol Biotechnol.* **58**(1): 23-29

Schauer S, Chaturvedi S, Randau L, Moser J, Kitabatake M, Lorenz S, Verkamp E, Schubert WD, Nakayashiki T, Murai M, Wall K, Thomann HU, Heinz DW, Inokuchi H, Soll D and Jahn D (2002) *Escherichia coli* glutamyl-tRNA reductase. Trapping the thioester intermediate. *J Biol Chem.* **277**(50): 48657-48663

Schenkman JB and Jansson I (2003) The many roles of cytochrome b5. *Pharmacol Ther.* **97**(2): 139-152

Schobert M and Jahn D (2002) Regulation of Heme Biosynthesis in Non-Phototrophic Bacteria. *J Mol Microbiol Biotechnol.* **4**(3): 287–294

Schubert WD, Moser J, Schauer S, Heinz DW and Jahn D (2002) Structure and function of glutamyl-tRNA reductase, the first enzyme of tetrapyrrole biosynthesis in plants and prokaryotes. *Photosynth Res.* **74**(2): 205-215

Shimada T and Nakamura S (1987) Cytochrome P450-mediated activation of procarcinogens and promutagens to DNA- damaging products by measuring expression of *umu* gene in *Salmonella typhimurium* TA1535/pSK1002. *Biochem Pharmacol.* **36**: 1979-1987

Smith MA, Kannangara CG and Grimm B (1992) Glutamate 1-semialdehyde aminotransferase: anomalous enantiomeric reaction and enzyme mechanism. *Biochem.* **31**(45): 11249-11254

Springer BA and Sligar SG (1987) High-level expression of sperm whale myoglobin in *Escherichia coli. Proc Natl Acad Sci USA.* **84**: 8961-8965

Strobel HW, Hodgson AV and Shen S (1995) NADPH cytochrome P450 reductase and its structural and functional domains. In: Ortiz de Montellano PR (ed) *Cytochrome P450: structure, mechanism and biochemistry.* (Pub *Plenum,* New York) pp. 225-244

Sugimura T (1997) Overview fo carcinogenic heterocyclic amines. *Mutat Res.* **376**: 211-219

Suzuki JY, Bollivar DW and Bauer CE (1997) Genetic analysis of chlorophyll biosynthesis. *Annu Rev Genet.* **31**: 61–89

Tait GH (1969) Coproporphyrinogenase activity in extracts from *Rhodopseudomonas sphaeroides. Biochem Biophys Res Commum*. **37**: 116-122

Takeya H, Tanaka T Hotta T and Sasaki K (1997) Production methods and applications of 5-aminolevulinic acid. *Porphyrins*. **6**: 127-135

Tanaka T, Takahashi K, Hotta T, Takeuchi Y and Konnai M (1992) Promotive effects of 5-aminolevulinic acid on yield of several crops. In: Proceedings of the 19th annual meeting of plant growth regulators Society of America, San Francisco. Plant growth regulator Society of America, Washington DC: 237-241

Tanaka T, Watanabe K, Nishikawa S, Hotta Y, Sasaki K, Murooka Y and Nagai S (1994) Selection of a high 5-aminolevulinic acid producing *Rhodobacter sphaeroides* mutant which is insensitive to yeast extract. *Seibutu-Kougaku.* **72**: 461-467

Tang W and Eisenbrand G (1992) Chinese Drugs of Plant Origin: Chemistry, Pharmacology and Use in Traditional and Modern Medicine. (Pub. *Springer-Verlag.* Heidlberg)

Trauner K and Hassen T (1994) Photodynamic therapy for the destruction of the synovium in the treatment of rheumatoid arthritis and the inflammatory arthritides. *US Patent No.* 5 368 841

Uhlmann B, Mann T, Gers-BarlagH and Sauermann G (1996) Cosmetic or dermatological preparation comprising delta-aminolevulinic acid content as an active ingredient. *US Patent No*. 5 520 905

Valdimir YB, Demain Al and Zaitseva NI (1997) The crucial contribution of starved resting cells to the elucidation of the pathway of vitamin B_{12} biosynthesis. *Crit Rev Biotechnol.* **17**: 21-37

Verderber E, Lucast LJ, Vandehy JA, Cozart P, Etter JB and Best EA (1997) Role of the *hemA* Gene Product and d-Aminolevulinic Acid in Regulation of *Escherichia coli* Heme Synthesis. *J Bacteriol.* **179**(14): 4583-4590

Verkamp E and Chelm BK (1989) Isolation, nucleotide sequence, and preliminary characterization of the *Escherichia coli* K-12 *hem*A gene. *J Bacteriol.* **171**(9): 4728-4735

Verkamp E, Jahn M, Jahn D, Kumar AM and Soll D (1992) Glutamyl–tRNA reductase from *Escherichia coli* and *Synechocystis* 6803. Gene structure and expression. *J Biol Chem.* **267**: 8275-8280

VonWachenfeldt C and Johnson EF (1995) Structures of eukaryotic cytochrome P450 enzymes. In: Oritz de Montellano PR (ed) *Cytochrome P450: structure, mechanism and biochemistry* (Pub Plenum, New York) pp. 183-233

Wang L, Elliott M and Elliott T (1999a) Conditional stability of the HemA protein (glutamyl-tRNA reductase) regulates heme biosynthesis in *Salmonella typhimurium. J Bacteriol.* **181**: 1211–1219

Wang L, Wilson S and Elliott T (1999b) A mutant HemA protein with positive charge close to the N-terminus is stabilized against heme-regulated proteolysis in *Salmonella typhimurium. J Bacteriol.* **181**: 6033–6041

Wang LY, Brown L, Elliott M and Elliott T (1997) Regulation of heme biosynthesis in *Salmonella typhimurium*: activity of glutamyl-tRNA reductase (HemA) is greatly elevated during heme limitation by a mechanism which increases abundance of the protein. *J Bacteriol.* **179**(9): 2907-2914

Weisberg RA and Gallant JA (1967) Dual function of the lambda prophage repressor. *J Mol Biol.* **25**(3): 537-544

Woodard SI and Dailey HA (1995) Regulation of Heme Biosynthesis in *Escherichia coli*. *Arch Biochem Biophys.* **316**(1): 110-115

Wu JJ, Ai CZ, Liu Y, Zhang YY, Jiang M, Lv AP and Yang L (2012) Interactions between Phytochemicals from Traditional Chinese Medicines and Human Cytochrome P450 Enzymes. *Curr Drug Metab.* In press.

Wu ZL, Aryal P, Lozach O, Meijer L and Guengerich FP (2005) Biosynthesis of New Indigoid Inhibitors of Protein Kinases Using Recombinant Cytochrome P450 2A6. *Chem Biodivers.* **2**(1): 51–65

Xu K and Elliott T (1994) Cloning, DNA sequence, and complementation analysis of the *Salmonella typhimurium hem*N gene encoding a putative oxygen-independent coproporphyrinogen III oxidase. *J Bacteriol.* **176**: 3196-3203

Xu K, Delling J and Elliott T (1992) The genes required for heme synthesis in *Salmonella typhimurium* include those encoding alternative functions for aerobic and anaerobic coproporphyrinogen oxidation. *J Bacteriol.* **174**(12): 3953-3963

Yanofsky C (2007) RNA-based regulation of genes of tryptophan synthesis and degradation in bacteria. *RNA.* **13**(8): 1141-1154

Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG and Court DL (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA.* **97**(11): 5978-5983

Zhang ZG, Liu Y, Guengerich FP, Matse JH, Chen J and Wu ZL (2009) Identification of amino acid residues involved in 4-chloroindole 3-hydroxylation by cytochrome P450 2A6 using screening of random libraries. *J Biotechnol.* **139**(1): 12-18

Zhao H, Giver L, Shao Z, Affholter JA and Arnold FH (1998) Molecular evolution by staggered extension process (StEP) *in vitro* recombination. *Nat Biotechnol.* **16**: 258-261

APPENDICES

Appendix 1. Basic information of oligonucleotides

Appendix 1.1 Site directed mutagenesis primer

SSMX06403SDF: **TTCTC CATATG AGCT TCCT** TAG CTCCTG AAA ATC TCGATAAC
SSMX06403SDR: **AGGAAGCT CATATG GAGAA** AAAAATCACTGGATATACCACCG TTG

Key: Bold nucleotides (**ATGA**)**:** 5′ primer-primer complementary region
Other nucleotides (TAGC): 3′ non-overlapping region
Highlighted nucleotides (**AGTC**): Change in nucleotide
Underlined Nucleotide (**CATATG**): *Nde*I restriction site

S.No	**Specifications**		**Forward Primer (SSMX06403SDF)**	**Reverse Primer (SSMX06403SDR)**
OligoAnalyzer 3.1 (Integrated DNA Technologies, IDT)				
1.	Primer-primer overlapping sequence	$T_{m(pp)}$	50^{0}C	50^{0}C
		GC content	42 %	42 %
		Length	19 bp	19 bp
2.	Non-overlapping sequence	T_m	52^{0}C	55^{0}C
		GC content	39%	38 %
		Length	23 bp	26 bp
3.	Secondary structure (Hairpin)	T_m	54^{0}C	43^{0}C
		ΔG	-3.9 kcal/mole	-3.01 kcal/mole
4.	Self-Dimer	ΔG	-7.91 kcal/mole	-7.82 kcal/mole
5.	Hetero dimer		ΔG = -33.68 kcal/mole	

PCR reaction mixture: site directed PCR

Components		Concentration
Template plasmid DNA (pACYC 184)	2 µl	80ng
2mM each dNTPs	5 µl	200µM
10µM SSMX06403SDF	2.5 µl	50pmoles (1µM)
10µM SSMX06403SDR	2.5 µl	50pmoles (1µM)
5 X Reaction Buffer	10 µl	1X
100% DMSO	4 µl	4%
0.5U/ µl DNA polymerase	2.5 µl	1.25U/50µl
MilliQ water	To 50µl	

PCR condition

	Cycles	Temp	Time
Initial Denaturation	1 Cycle	95°C	5 min
Denaturation	20 Cycles	95°C	1 min
Annealing (T_{no}-5^0C)		52°C	1 min
Extension		72°C	5 min (4245 bp)
Denaturation	4 Cycle	95°C	1 min
Annealing (T_{pp}-5^0C)		45°C	1 min
Extension		72°C	10 min
Denaturation	1 cycle	95°C	1 min
Annealing		50°C	1 min
Extension		72°C	20 min
Hold		4^0C	∞

Appendix 1.2: Primers used for *hem*A gene amplification

SSMPOA6X1F01: CTAC**GAATTC**TCCAAA **CATatg** acc ***AAA AAA*** ct c tta gca ctc gg
*Eco*RI *Nde*I

SSMPOA6X1R01: AA ACC ACT **GAATTC** CTG ATG **TCTAga** tgt act gct act cc
*Eco*RI *Xba*I

S.No	**Specifications**		**Forward Primer (SSMPOA6X1F01)**	**Reverse Primer (SSMPOA6X1R01)**
OligoAnalyzer 3.1 (Integrated DNA Technologies, IDT)				
1.	Annealing sequence	T_m	56.3^{0}C	45.2^{0}C
		GC content	55 %	50 %
		Length	20 bp	16 bp
2.	Whole Primer sequence	T_m	63.6^{0}C	63.2^{0}C
		GC content	40%	42.5 %
		Length	45 bp	40 bp
3.	Secondary structure (Hairpin)	T_m	19.4^{0}C	28.8^{0}C
		ΔG	0.35 kcal/mole	-0.26 kcal/mole
4.	Self-Dimer	ΔG	-8.51 kcal/mole	-8.51 kcal/mole
5.	Hetero dimer		ΔG = -8.51 kcal/mole	
6.	Amplicon size		1315 bp	

PCR reaction mixture: *hem*A gene amplification

Components		Final concentration
Template: Genomic DNA of *E. coli* HME 5	1.5 µl	25 ng
2mM dNTPs each	3.2 µl	250 µM dNTPs of each
10µM Forward primer (SSMPOA6X1F01)	2.5 µl	1 µM
10µM Reverse primer (SSMPOA6X1R01)	2.5 µl	1 µM
5 X Reaction Buffer	5 µl	1X
0.5U/ µl DNA polymerase	1 µl	0.5U
Nuclease free water	Upto 25µl	

PCR conditions: *hem*A gene amplification

	Cycles	Temp	Time
Initial Denaturation	1 Cycle	95^0C	5 min
Denaturation	5 Cycles	95^0C	30 sec
Annealing		52^0C	30 sec
Extension (1 min/kb)		72^0C	1 min 38 sec
Denaturation	25 Cycles	95^0C	30 sec
Annealing		59^0C	30 sec
Extension		72^0C	1 min 38 sec
Final Extension	1 Cycle	72^0C	5 minutes
Hold	1 Cycle	4^0C	∞

Appendix 1.3 Primer set for PCR amplification of 5’ homologous region of *hem*A

SSMP61320F01: CCTACTG **AAGCTT** gtaa acctaaaatc cactg
*Hind*III
SSMP61320R01: GATA ***GGATC CATATG*** tgc ggg aaa taa tac c
*Bam*HI *Nde*I

S.No	Specifications		Forward Primer (SSMP61320F01)	Reverse Primer (SSMP61320R01)
OligoAnalyzer 3.1 (Integrated DNA Technologies, IDT)				
1.	Annealing sequence	$T_{m(pp)}$	45.5^{0}C	45.5^{0}C
		GC content	37 %	44 %
		Length	19 bp	16 bp
2.	T_m	T_m	58.8^{0}C	58.8^{0}C
		GC content	41%	42 %
		Length	32 bp	31 bp
3.	Secondary structure (Hairpin)	T_m	14.4^{0}C	27^{0}C
		ΔG	0.71 kcal/mole	-0.13 kcal/mole
4.	Self-Dimer	ΔG	-10.23 kcal/mole	-10.76 kcal/mole
5.	Hetero dimer		ΔG = -6.12 kcal/mole	

Appendix 1.4 Primer set for PCR amplification of 3' homologous region of *hem*A

SSMPOA710F01: CGT GGT GTG C ***TCTAGA*** ca ttt tca ttt ttt aca gg
*Xba*I
SSMPOA710R01: TGT CGG CTA ***GAATTC*** a gt tac gtt cgt cat cag
*Eco*RI

S.No	**Specifications**		**Forward Primer (SSMPOA710F01)**	**Reverse Primer (SSMPOA710R01)**
OligoAnalyzer 3.1 (Integrated DNA Technologies, IDT)				
1.	Annealing sequence	$T_{m(pp)}$	42.2^{0}C	47.3^{0}C
		GC content	26 %	47 %
		Length	19 bp	17 bp
2.	T_m	T_m	60.7^{0}C	62.1^{0}C
		GC content	40%	45 %
		Length	35 bp	33 bp
3.	Secondary structure (Hairpin)	T_m	21.8^{0}C	20.6^{0}C
		ΔG	0.21 kcal/mole	0.34 kcal/mole
4.	Self-Dimer	ΔG	-7.31 kcal/mole	-8.51 kcal/mole
5.	Hetero dimer		ΔG = -5.73 kcal/mole	

PCR reaction mixture

Components		Final concentration
Template: Genomic DNA of *E. coli* HME 5	1.5 µl	25 ng
2mM dNTPs each	3.2 µl	250 µM dNTPs of each
10µM Forward primer	2.5 µl	1 µM
10µM Reverse primer	2.5 µl	1 µM
5 X Reaction Buffer	5 µl	1X
0.5U/ µl DNA polymerase	1 µl	0.5U
Nuclease free water	Upto 25µl	

PCR reaction condition

	Cycles	Temp	Time
Initial Denaturation	1 Cycle	95°C	5 min
Denaturation	5 Cycles	95°C	30 sec
Annealing		52^{0}C	30 sec
Extension		72°C	50 sec
Denaturation	25 Cycles	95°C	30 sec
Annealing		59^{0}C	30 sec
Extension		72°C	50 sec
Final Extension	1 Cycle	72°C	5 minutes
Hold	1 Cycle	4°C	--

Appendix 1.5 Reamplification primer set for Homologous recombination

SSMPOA6X1F03: cgtccg cta tcg tct atg ttc aag ttg tc

SSMPOA6X1R03: gta att gct gtt cca gtt gct cgc ttt ttt c

S.No	Specifications		Forward Primer (SSMPOA6X1F03)	Reverse Primer (SSMPOA6X1R03)
OligoAnalyzer 3.1 (Integrated DNA Technologies, IDT)				
1.	T_m	T_m	60.3^{0}C	60.5^{0}C
		GC content	48 %	42 %
		Length	29 bp	31 bp
2.	Secondary structure (Hairpin)	T_m	31.1^{0}C	28.5^{0}C
		ΔG	-0.45 kcal/mole	-0.3 kcal/mole
3.	Self-Dimer	ΔG	-3.9 kcal/mole	-5.36 kcal/mole
4.	Hetero dimer		ΔG = -3.9 kcal/mole	

Appendix 2: 20 Amino acids, their single-letter data-base codes, and their corresponding DNA codons

Amino Acid	SLC	DNA codons
Isoleucine	I	ATT, ATC, ATA
Leucine	L	CTT, CTC, CTA, CTG, TTA, TTG
Valine	V	GTT, GTC, GTA, GTG
Phenylalanine	F	TTT, TTC
Methionine	M	ATG
Cysteine	C	TGT, TGC
Alanine	A	GCT, GCC, GCA, GCG
Glycine	G	GGT, GGC, GGA, GGG
Proline	P	CCT, CCC, CCA, CCG
Threonine	T	ACT, ACC, ACA, ACG
Serine	S	TCT, TCC, TCA, TCG, AGT, AGC
Tyrosine	Y	TAT, TAC
Tryptophan	W	TGG
Glutamine	Q	CAA, CAG
Asparagine	N	AAT, AAC
Histidine	H	CAT, CAC
Glutamic acid	E	GAA, GAG
Aspartic acid	D	GAT, GAC
Lysine	**K**	**AAA, AAG**
Arginine	R	CGT, CGC, CGA, CGG, AGA, AGG
Stop codons	Stop	TAA, TAG, TGA

APPENDIX 3: Analysis of Restriction site present in *hem*A gene sequence by using NEBcutter V2.0

(http://tools.neb.com/NEBcutter2/cutshow.php?name=35a4abff-)

Appendix 4 Preparation of Stock solutions, Buffers and Reagents commonly used in Molecular Biology

Tris-Cl (1 M, pH 7.5 and 8.0)

Tris-Cl buffer was prepared by adding 12.11 g of Tris base in 80 ml of MilliQ water and pH was adjusted to 7.5 or 8.0 by adding concentrated HCl. Then final volume was maintained to 100 ml. The solution was sterilized by autoclaving and stored at 4^0C.

EDTA (0.5 M, pH 8.0)

18.61 g of disodium EDTA.$2H_20$ was added to 80 ml of MilliQ water, stirred vigorously on magnetic stirrer and pH was adjusted to 8.0 with NaOH. The final volume of the solution was made upto 100 ml and sterilized by autoclaving and stored at 4^0C.

Tris EDTA (TE) buffers

pH 7.5	*pH 8.0*
10 mM Tris-Cl (pH 7.5)	10 mM Tris-Cl (pH 8.0)
1 mM EDTA (pH 8.0)	1 mM EDTA (pH 8.0)

Potassium Phosphate buffer (0.1 M, pH 7.0)

1 M K_2HPO_4	61.5 ml
1 M KH_2PO_4	38.5 ml
MilliQ water	upto 1000 ml

Preparation of solutions for chromosomal DNA isolation

TE1 solution

50 mM Tris-Cl (pH 7.5)
50 mM EDTA (pH 8.0)

STEP solution

50 mM Tris-Cl (pH 7.5)
0.2 mM EDTA (pH 8.0)
0.5% SDS
1 mg/ml Proteinase K

Lysozyme (1 mg/ml)

1 mg of lysozyme was dissolved in 10mM Tris-Cl (pH 8.0) with final volume of 1 ml. The solution was prepared immediately before use.

Proteinase K (20 mg/ml)

The lyophilized proteinase powder was dissolved at a concentration of 20 mg/ml in sterilized 50 mM Tris (pH 8.0), 1.5 mM calcium acetate. The solution was stored at -20^{0}C.

Sodium Acetate (3 M, pH 5.2)

40.83 g of sodium acetate.3H_20 was dissolved in 80 ml of MilliQ water and pH was adjusted to 5.2 with glacial acetic acid and final volume was made upto 100 ml. The solution was sterilized by autoclaving and stored at 4^{0}C.

Preparation of solutions for Plasmid DNA extraction by Alkanline-lysis method

Alkaline Lysis Solution I

50 mM glucose
25 mM Tris-Cl (pH 8.0)
10 mM EDTA (pH 8.0)
The solution was sterilized by autoclaving and stored at 4^{0}C.

Alkaline Lysis Solution II (Freshly prepared)

0.2 N NaOH
1% (w/v) SDS

Alkaline Lysis Solution III

5 M Potassium acetate	60.0 ml
Glacial acetic acid	11.5 ml
H20	28.5 ml

STE

10 mM Tris-Cl (pH 8.0)
0.1 M NaCl
1 mM EDTA (pH 8.0)

Preparation of solutions for Transformation

RF1

30 mM KOAc (Potassium acetate)
100 mM RbCl
10 mM $CaCl_2$
50 mM $MnCl_2$
15% Glycerol
pH of the solution was adjusted to 5.8 with acetic acid and filter sterilized (0.22 µm) and stored in dark bottle at 4^0C.

RF2

10 mM MOPS
75 mM $CaCl_2$
10 mM RbCl
15% Glycerol
pH of the solution was adjusted to 6.5 with KOH and and filter sterilized (0.22 µm) and stored in dark bottle at 4^0C.

Preparation of solutions for Agarose gel electrophoresis

10 X TAE Buffer

Tris base	4.84 g
Glacial acetic acid	1.142 ml
0.5 M EDTA (pH 8.0)	2 ml
MilliQ water	upto 100 ml

The working solution (0.5 X) of TAE buffer was prepared by diluting 17.5 ml of 10 X TAE stock solutions with MilliQ water to make the final volume upto 350 ml.

10X Gel loading buffer

Bromophenol blue	10.5 mg
0.5 M EDTA (pH 8.0)	2 ml
98% Glycerol	2.5 ml
MilliQ water	upto 5 ml

Ethidium Bromide (10 mg/ml)

100 mg of ethidium bromide was weighed and dissolved in 10 ml of MilliQ water. The solution was protected from sun light by wrapping with aluminum foil and stored at room temperature.

Preparation of Media and Additives

IPTG (1 M)

238 mg of IPTG powder was dissolved in 1 ml of MilliQ water, filter sterilized (0.22 µm) and stored at -20^{0}C.

5-Aminolevulinic acid (1 M)

167.59 mg of 5- Aminolevulinic acid was dissolved in 1 ml of MilliQ water, filter sterilized (0.22 µm) and stored at -20^{0}C.

Indole (5 M)

585.75 mg of indole was dissolved in 1 ml of DMSO and stored at 4^{0}C.

Thiamine (1 M)

337.27 mg of thiamine was dissolved in 1 ml of MilliQ water, filter sterilized (0.22 µm) and stored at 4^{0}C.

Trace elements

$FeCl_3 \cdot 6H_20$	27.0 g
$ZnCl_2 \cdot 4H_20$	2.0 g
$CoCl_2 \cdot 6H_20$	2.0 g
$Na_2MoO_4 \cdot 2H_20$	2.0 g
$CaCl_2 \cdot 2H_20$	1.0 g
$CuCl_2$	1.0 g
H_3B0_3	0.5 g
30% HCl	100 ml

The mixture was dissolved in MilliQ water and final volume was adjusted to 1000 ml.

Antibiotics

	Stock Solution (-20^0C)	Working solution	
		LB broth	LB/Agar plate
Ampicillin	50 mg/ml in H_20 (filter sterilized)	50 µg/ml	50 µg/ml
Tetracycline	10 mg/ml in 70% (v/v) ethanol	10 µg/ml	5 µg/ml
Chloramphenicol	25 mg/ml in 70% (v/v) ethanol	25 µg/ml	10 µg/ml

Luria Bertani (LB) medium

2.5 g of LB medium was weighed and dissolved in 100 ml MilliQ water and sterilized by autoclaving. Once the medium is cooled (25^0C), appropriate antibiotics were supplemented if required and stored at 4^0C.

LB/Agar plate

LB medium	2.5 g
Agar	1.5 g

The volume was adjusted to 100 ml with MilliQ water and sterilized by autoclaving. Once the medium is cooled (60^0C), then appropriate antibiotics were added to the desired concentrations (as mentioned above) and the plates were stored at 4^0C.

SOC medium

Tryptone	20 g
Yeast extract	5 g
Sodium chloride	0.5 g
Potassium chloride	0.186 g

pH was adjusted to 7.0, final volume was made upto 1000 ml and autoclaved. Then 10 ml of separately autoclaved 20% glucose and 1 M $MgCl_2$ was added respectively.

Terrific Broth (TB) expression medium

Terrific Broth	4.93 g
Peptone	0.2 g
40% Glycerol	1 ml
Trace elements	25 µl
MilliQ water	100 ml

The medium was sterilized by autoclaving and once cooled (25^{0}C), appropriate antibiotics and additives were added to desired concentrations as mentioned in materials and method sections.

Appendix 5: PCR amplified *hem*A gene sequence (GenBank: NC_000913.2)

0001 ctac***gaattc*** tccaaa***cata tg***accaaaaa actcttagca ctcggtatca accataaaac
0061 ggcacctgta tcgctgcgag aacgtgtatc gttttcgccg gataagctcg atcaggcgct
0121 ttgacagcct gcttgcgcag ccgatggtgc agggcggcgt ggtgctgtcg acgtgcaacc
0181 gcacggaact ttatcttagc gttgaagagc aggacaacct gcaagaggcg ttaatccgct
0241 ggctttgcga ttatcacaat cttaatgaag aagatctgcg taaaagcctc tactggcatc
0301 aggataacga cgcggttagc catttaatgc gtgttgccag cggcctggat tcactggttc
0361 tgggggagcc gcagatcctc ggtcaggtta aaaaagcgtt tgccgattcg caaaaaggtc
0421 atatgaaggc cagcgaactg gaacgcatgt tccagaaatc tttctctgtc gcgaaacgcg
0481 ttcgcactga aacagatatc ggtgccagcg ctgtgtctgt cgcttttgcg gcttgtacgc
0541 tggcgcggca gatctttgaa tcgctctcta cggtcacagt gttgctggta ggcgcgggcg
0601 aaactatcga gctggtggcg cgtcatctgc gcgaacacaa agtacagaag atgattatcg
0661 ccaaccgcac tcgcgaacgt gcccaaattc tggcagatga agtcggcgcg gaagtgattg
0721 ccctgagtga tatcgacgaa cgtctgcgcg aagccgatat catcatcagt tccaccgcca
0781 gcccgttacc gattatcggg aaaggcatgg tggagcgcgc attaaaaagc cgtcgcaacc
0841 aaccaatgct gttggtggat attgccgttc cgcgcgatgt tgagccggaa gttggcaaac
0901 tggcgaatgc ttatctttat agcgttgatg atctgcaaag catcatttcg cacaacctgg
0961 cgcagcgtaa agccgcagcg gttgaggcgg aaactattgt cgctcaggaa accagcgaat
1021 ttatggcgtg gctgcgagca caaagcgcca gcgaaaccat tcgcgagtat cgcagccagg
1081 cagagcaagt tcgcgatgag ttaaccgcca aagcgttagc ggcccttgag cagggcggcg
1141 acgcgcaagc cattatgcag gatctggcat ggaaactgac taaccgcttg atccatgcgc
1201 caacgaaatc acttcaacag gccgcccgtg acggggataa cgaacgcctg atattctgcg
1261 cgacagcctc gggctggagt agcagtaca***t ctaga***catca g***gaattc***agt ggttt

Appendix 6: pACYC184 plasmid information from NEB

pACYC184

4,245 base pairs
GenBank Accession #: X06403

pACYC184 is available as a transformant of ER2420 (#E4152S) at no charge when shipped with an order or for the cost of shipping if ordered separately.

There are no restriction sites for the following enzymes: AarI (x), AatII, Acc65I, AflII, AflIII, AleI, ApaI, ApaLI, AscI, AsiSI, AvrII, BaeI, BbvCI, BciVI, BglII, BlpI, BmgBI, BsaI, BsaXI, BsiWI, BspQI, BsrGI, BssHII, BstEII, BstXI, CspCI, DraIII, FseI, HpaI, I-CeuI, I-SceI, KpnI, MfeI, MluI, NdeI, NotI, NsiI, PI-PspI, PI-SceI, PacI, PaeR7I, PciI, PmeI, PmlI, PsiI, PspOMI, PspXI, PstI, PvuI, RsrII, SacI, SanDI (x), SapI, SbfI, SexAI, SfiI, SmaI, SnaBI, SpeI, SrfI (x), StuI, SwaI, TliI, TspMI, XcmI, XhoI, XmaI, ZraI.

(x) = enzyme not available from NEB

pACYC184 is an *E. coli* plasmid cloning vector containing the p15A origin of replication (1-4). This allows pACYC184 to coexist in cells with plasmids of the ColE1 compatibility group (e.g., pBR322, pUC19). It is a low copy number vector, at about 15 copies per cell (5), but can be amplified with spectinomycin. Chloramphenicol cannot be used for amplification due to the presence of the *cat* gene.

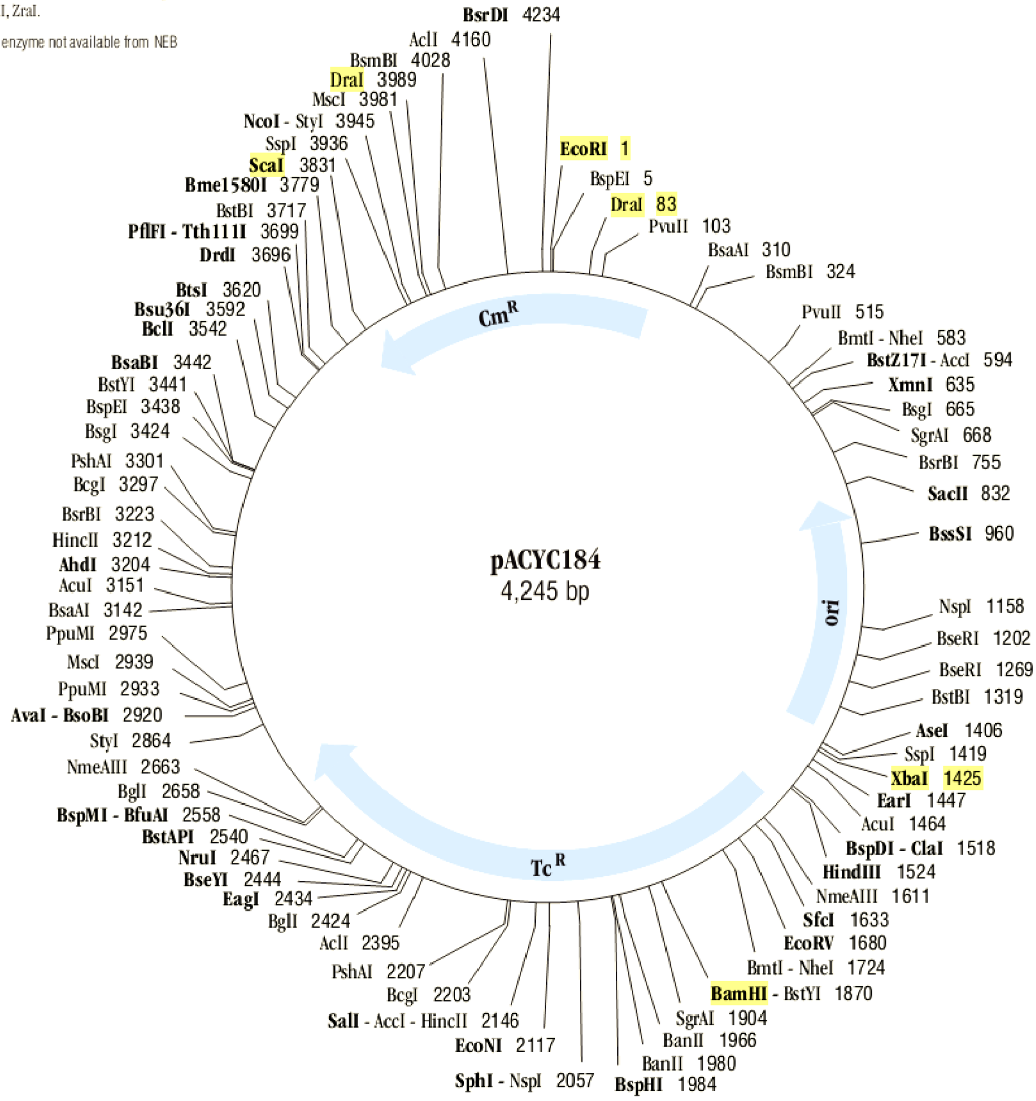
Enzymes with unique restriction sites are shown in **bold** type and enzymes with two restriction sites are shown in regular type. Location of sites of all NEB restriction enzymes can be found on the NEB web site (choose Technical Reference > DNA Sequences and Maps). Restriction site coordinates refer to the position of the 5´-most base on the top strand in each recognition sequence.

Open reading frame (ORF) coordinates are in the form "translational start – translational stop"; numbers refer to positions on the top (clockwise) strand, regardless of the direction of transcription and include the start and stop codons.

Origin of replication coordinates include the region from the -35 promoter sequence of the RNAII transcript to the RNA/DNA switch point.

Feature	**Coordinates**	**Source**
cat (Cm^R)	219-3805 (ccw)	*Tn9*
Tc^R	1581-2771	pSC101
origin	1393-848	p15A

ori = origin of replication
Cm = chloramphenicol, Tc = tetracycline
(ccw) = counterclockwise



References

1. Chang, A.C.Y., and Cohen, S.N. (1978) *J. Bacteriol.* 134, 1141–1156.
2. Rose, R.E. (1988) *Nucleic Acids Res.* 16, 355.
3. Mok, Y.K., Clark, D.R., Kam, K.M., and Shaw, P.C. (1991) *Nucleic Acids Res.* 19, 2321-2323.
4. Selzer, G., Som, T., Itoh, T., and Tomizawa, J (1983) *Cell* 32, 119-129.
5. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Appendix 7: pUC19 plasmid information from NEB

pUC19

GenBank Accession #: L09137

Feature	Coordinates	Source
lacZα	469-146	–
origin	1455-867	pMB1 (mutant)
bla (Ap^R)	2486-1626	*Tn3*

ori = origin of replication
Ap = ampicillin

There are no restriction sites for the following enzymes: AarI(x), AfeI, AflII, AgeI, AleI, ApaI, AscI, AsiSI, AvrII, BaeI, BbsI, BbvCI, BclI, BglII, BlpI, BmgBI, BmtI, Bpu10I, BsaAI, BsaBI, BseRI, BsgI, BsiWI, BsmFI, BsmI, BspDI, BspEI, BsrGI, BssHII, BstBI, BstEII, BstXI, BstZ17I, Bsu36I, BtgI, BtgZI, ClaI, CspCI, DraIII, EagI, EcoNI, EcoRV, FseI, FspAI(x), HpaI, I-CeuI, I-SceI, MfeI, MluI, MscI, NaeI, NcoI, NgoMIV, NheI, NotI, NruI, NsiI, PI-PspI, PI-SceI, PacI, PaeR7I, PflFI, PflMI, PmeI, PmlI, PpuMI, PshAI, PsiI, PspOMI, PspXI, RsrII, SacII, SanDI(x), SexAI, SfiI, SgrAI, SnaBI, SpeI, SrfI(x), StuI, StyI, SwaI, TliI, Tth111I, XcmI, XhoI

(x) = enzyme not available from NEB

pUC19 is a small, high-copy number *E. coli* plasmid cloning vector containing portions of pBR322 and M13mp19 (1). It contains the pMB1 origin of replication from pBR322, but it lacks the *rop* gene and carries a point mutation in the RNAII transcript (G 2975 in pBR322 to A 1308 in pUC19; 2). These changes together result in a temperature-dependent copy number of about 75 per cell at 37°C and >200 per cell at 42°C (2, 3). The multiple cloning site (MCS) is in frame with the *lacZα* gene, allowing screening for insertions using α-complementation.

pUC18 is identical to pUC19 except that the MCS region (nt 397-454) is inverted.

pNEB193 is also identical to pUC19 except for the addition of several restriction endonuclease sites to the MCS. Its total length is 2713 bp.

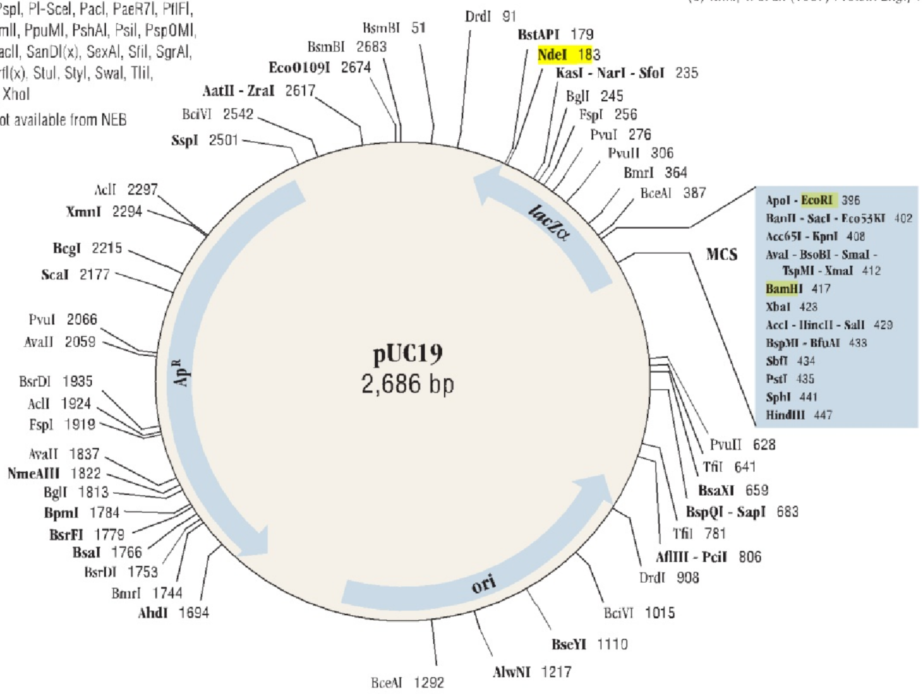
Enzymes with unique restriction sites are shown in **bold** type, and enzymes with two restriction sites are shown in regular type Location of sites of all NEB restriction enzymes can be found on the NEB web site (choose Technical Reference > DNA Sequences and Maps). Restriction site coordinates refer to the position of th 5´-most base on the top strand in each recognition sequence.

Open reading frame (ORF) coordinates are in the form "translational start – translational stop"; numbers refer to positions on the top (clockwise) strand, regardless of the directi of transcription and include the start and stop codons.

Origin of replication coordinates include the region from the -35 promoter sequence of the RNAII transcript to the RNA/DNA switc point. *bla* (Ap^R) gene coordinates include the signal sequence.

References

(1) Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene*, 33, 103–119.
(2) Lin-Chao, S., Chen, W.-T. and Wong, T.-T. (1992) *Mol. Microbiol.*, 6, 3385–3393
(3) Miki, T. et al. (1987) *Protein Eng.*, 1, 327–332.



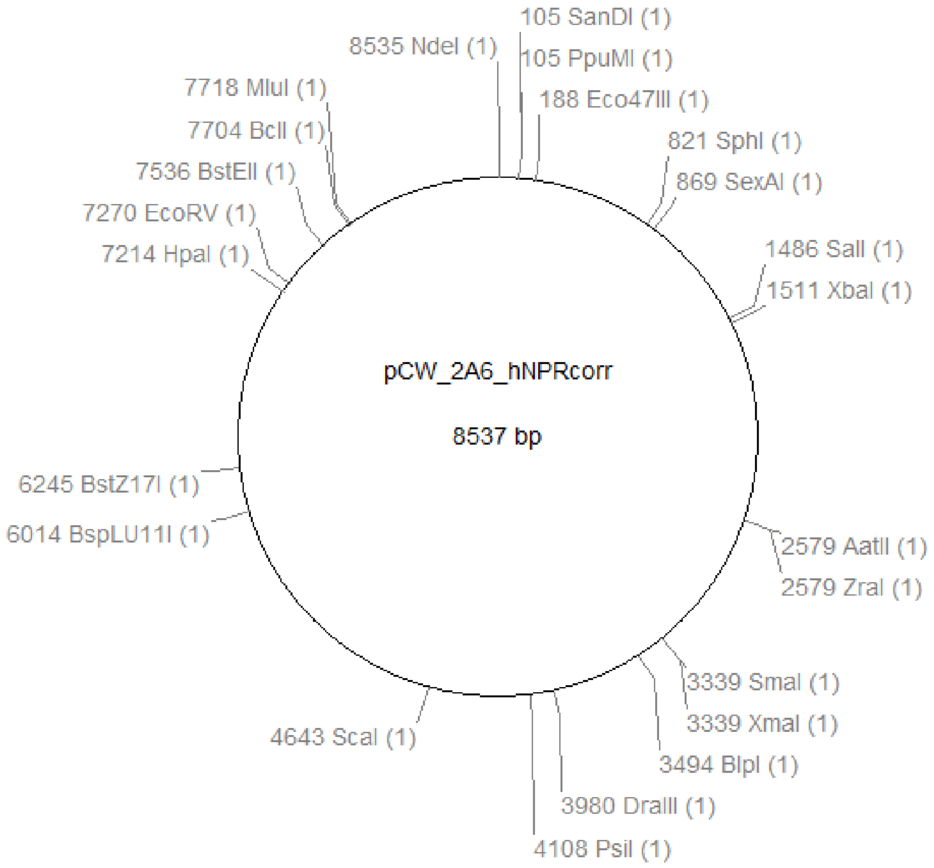
pUC19 MCS

```
       SacI       SmaI       XbaI       SbfI
   EcoRI       KpnI     BamHI       SalI PstI  SphI  HindIII
agtgAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGcgtaatcatggtcat
       .         .         .         .         .         .         .
      400       410       420       430       440       450       460

...S  N  S  S  P  V  R  P  D  E  L  T  S  R  C  A  H  L  S  P  T  I  M  T  M

                                       <------------ lacZα translational start
```

pNEB193 MCS

```
       SacI       SmaI   BssHII       PacI                          SbfI
   EcoRI       KpnI       AscI     BamHI          XbaI  SalI   PmeI    PstI  SphI  HindIII
agtgAATTCGAGCTCGGTACCCGGGGGCGCGCCGGATCCTTAATTAAGTCTAGAGTCGACTGTTTAAACCTGCAGGCATGCAAGCTTGGcgtaatcatggtcat
       .         .         .         .         .         .         .         .         .         .
      400       410       420       430       440       450       460       470       480       490

...S  N  S  S  P  V  R  P  R  A  P  D  K  I  L  D  L  T  S  Q  K  F  R  C  A  H  L  S  P  T  I  M  T  M

                                                                         <------------ lacZα translational start
```

Appendix 8: pCW''2A6: hNPR plasmid harboring human cytochrome P450 2A6 and human NADPH-P450 reductase as bicistronic construct under double tac promoter. Restriction map and nucleotide sequences (Courtsey: Gillam EMJ).



atggctgcttcaggaatgttattagtggccttgctggtctgcctgactgtgatggtcttgatgtctgtttggcagcagaggaagagcaaggg
gaagctgcctccgggacccaccccattgcccttcattggaaactacctgcagctgaacacagagcagatgtacaactccctcatgaagatc
agtgagcgctatggccccgtgttcaccattcacttggggccccggcgggtcgtggtgctgtgtggacatgatgccgtcagggaggctctggt
ggaccaggctgaggagttcagcgggcgaggcgagcaagccaccttcgactgggtcttcaaaggctatggcgtggtattcagcaacgggg
agcgcgccaagcagctccggcgcttctccatcgccaccctgcgggacttcggggtgggcaagcgaggcatcgaggagcgcatccaggag
gaggcgggcttcctcatcgacgccctccggggcactggcggcgccaatatcgatcccaccttcttcctgagccgcacagtctccaatgtcat
cagctccattgtctttggggaccgctttgactataaggacaaagagttcctgtcactgttgcgcatgatgctaggaatcttccagttcacgtc
aacctccacggggcagctctatgagatgttctcttcggtgatgaaacacctgccaggaccacagcaacaggcctttcagttgctgcaaggg
ctggaggacttcatagccaagaaggtggagcacaaccagcgcacgctggatcccaattccccacgggacttcattgactcctttctcatcc
gcatgcaggaggaggagaagaaccccaacacggagttctacttgaaaaacctggtgatgaccacgttgaacctcttcattgggggcaccg
agaccgtcagcaccaccctgcgctatggcttcttgctgctcatgaagcacccagaggtggaggccaaggtccatgaggagattgacagag
tgatcggcaagaaccggcagcccaagtttgaggaccgggccaagatgccctacatggaggcagtgatccacgagatccaaagatttgga
gacgtgatccccatgagtttggcccgcagagtcaaaaaggacaccaagtttcgggatttcttcctccctaagggcaccgaagtgtaccct a
tgctgggctctgtgctgagagaccccagtttcttctccaacccccaggacttcaatccccagcacttcctgaatgagaaggggcagtttaag
aagagtgatgcttttgtgcccttttccatcggaaagcggaactgtttcggagaaggcctggccagaatggagctctttctcttcttcaccacc
gtcatgcagaacttccgcctcaagtcctcccagtcacctaaggacattgacgtgtcccccaaacacgtgggctttgccacgatcccacgaa
actacaccatgagcttcctgccccgctgaGTCGACCTGCAGCCCAAGCTTCCGGTCTAGAAATAATTTTGTTTAACTT
TAAGAAGGAGATATAAATatggctgactcccacgtggacaccagctccaccgtgtccgaggcggtggccgaagaagtatctctttt
cagcatgacggacatgattctgttttcgctcatcgtgggtctcctaacctactggttcctcttcagaaagaaaaaagaagaagtccccgagt
tcaccaaaattcagacattgacctcctctgtcagagagagcagctttgtggaaaagatgaagaaaacggggaggaacatcatcgtgttct

acggctcccagacggggactgcagaggagtttgccaaccgcctgtccaaggacgcccaccgctacgggatgcgaggcatgtcagcggac
cctgaggagtatgacctggccgacctgagcagcctgccagagatcgacaacgccctggtggttttctgcatggccacctacggtgaggga
gaccccaccgacaatgcccaggacttctacgactggctgcaggagacagacgtggatctctctggggtcaagttcgcggtgtttggtcttg
ggaacaagacctacgagcacttcaatgccatgggcaagtacgtggacaagcggctggagcagctcggcgcccagcgcatctttgagctg
gggttgggcgacgacgatgggaacttggaggaggacttcatcacctggcgagagcagttctggccggccgtgtgtgaacactttggggtg
gaagccactggcgaggagtccagcattcgccagtacgagcttgtggtccacaccgacatagatgcggccaaggtgtacatgggggagat
gggccggctgaagagctacgagaaccagaagccccccttgatgccaagaatccgttcctggctgcagtcaccaccaaccggaagctga
accagggaaccgagcgccacctcatgcacctggaattggacatctcggactccaaaatcaggtatgaatctggggaccacgtggctgtgt
acccagccaacgactctgctctcgtcaaccagctgggcaaaatcctgggtgccgacctggacgtcgtcatgtccctgaacaacctggatga
ggagtccaacaagaagcacccattcccgtgccctacgtcctaccgcacggccctcacctactacctggacatcaccaacccgccgcgtacc
aacgtgctgtacgagctggcgcagtacgcctcggagccctcggagcaggagctgctgcgcaagatggcctcctcctccggcgagggcaag
gagctgtacctgagctgggtggtggaggcccggaggcacatcctggccatcctgcaggactgcccgtccctgcggcccccccatcgaccacc
tgtgtgagctgctgccgcgcctgcaggcccgctactactccatcgcctcatcctccaaggtccaccccaactctgtgcacatctgtgcggtgg
ttgtggagtacgagaccaaggccggccgcatcaacaagggcgtggccaccaactggctgcgggccaaggagcctgccggggagaacgg
cggccgtgcgctggtgcccatgttcgtgcgcaagtcccagttccgcctgcccttcaaggccaccacgcctgtcatcatggtgggccccggca
ccggggtggcacccttcataggcttcatccaggagcgggcctggctgcgacagcagggcaaggaggtgggggagacgctgctgtactac
ggctgccgccgctcggatgaggactacctgtaccgggaggacgtggcgcagttccacagggacggtgcgctcacccagctcaacgtggcc
ttctcccgggagcagtcccacaaggtctacgtccagcacctgctaaagcaagaccgagagcacctgtggaagttgatcgaaggcggtgcc
cacatctacgtctgtggggatgcacggaacatggccagggatgtgcagaacaccttctacgacatcgtggctgagctcggggccatggag
cacgcgcaggcggtggactacatcaagaaactgatgaccaagggccgctactccctggacgtgtggagctagaAGCTTATCGATGA
TAAGCTGTCAAACATGAGCAGATCTGAGCCCGCCTAATGAGCGGGCTTTTTTTTCAGATCTGCTTGAAGA
CGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGCGTCAAAG
CAACCATAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC
GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGC
TTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCC
CAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGA
CGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGC
TATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAA
ATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCG
GAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAA
TGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTG
CGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTG
GGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAA
GAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGG
GCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAA
AAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTG
CGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGG
ATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACA
CCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCC
CGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCG
GCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGG
GGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAAC
GAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTC
ATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAA
TCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAG
GATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCG
GTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGA
TACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACA
TACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGA
CTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCA
GCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC
CCGAAGGGAGAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGA

GCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGAT
TTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCT
GGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACC
GCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGA
AGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATATGGTG
CACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTG
GGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCA
TCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGA
AACGCGCGAGGCAGAACGCCATCAAAAATAATTCGCGTCTGGCCTTCCTGTAGCCAGCTTTCATCAACAT
TAAATGTGAGCGAGTAACAACCCGTCGGATTCTCCGTGGGAACAAACGGCGGATTGACCGTAATGGGAT
AGGTTACGTTGGTGTAGATGGGCGCATCGTAACCGTGCATCTGCCAGTTTGAGGGGACGACGACAGTAT
CGGCCTCAGGAAGATCGCACTCCAGCCAGCTTTCCGGCACCGCTTCTGGTGCCGGAAACCAGGCAAAGC
GCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCA
GCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACG
TTGTAAAACGACGGCCAGTGAATCCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTC
ACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA
CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATG
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTG
AGACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGG
TTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGT
ATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCG
CCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGG
TTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTG
AGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCG
ATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAA
TACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCAC
AGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATT
GTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTT
GATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCA
ACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCG
CCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACG
GTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAA
TTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCTGGCACG
ACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCAC
ACAGGAAACAGGATCGATCCATCGATGAGCTTACTCCCCATCCCCCTGTTGACAATTAATCATCGGCTCGT
ATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGGATCAGCTTACTCCCCATCCCCC
TGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAAC
AGGATCCATCGATGCTTAGGAGGTCAT

PUBLICATIONS

- Sujina Mali, Rumika Maharjan, Tribikram Bhattarai, F.P. Guengerich and Pramod Aryal (2012) Development of efficient *E. coli* host strain for expression of cytochrome P450s by endogenous increased biosynthesis of 5-Aminolevulinic acid (ALA) abstract submitted to Nepal Academy of Science and Techonology (NAST) for The Sixth National Conference on Science and Technology: June 6-8, 2012, Khumaltar, Lalitpur, Nepal.

- Damodar Parajuli, Sujina Mali, Tribikram Bhattarai, Ayesha Dhungel, Rumika Maharjan and Pramod Aryal (2011) Protein and chemical engineering for efficient host development on production of biologically important molecule, Abstract, Organized by Biotechnology Unit, Nepal Agricultural Research Council (NARC), Khumaltar, Lalitpur, Nepal.

- Sujina Mali, Rumika Maharjan and Pramod Aryal (2011) Protein Engineering and Chromosomal Engineering: A Novel Strategy for Directed Evolution. Transcript, Volume I, Annual Magazine, Central Department of Biotechnology, Tribhuvan University.

Grant

- Grant from Ministry of Science and Technology (MoST), Government of Nepal, for Research project “Deregulation and protein engineering of *hem*A gene: Implication in Pharmaceutical, Industrial and Agricultural Biotechnology” (2011) as one of the co-investigator with Dr. Pramod Aryal as principal investigator.