Molecular cloning and expression of equine CYP1A2 in *Escherichia coli*

A thesis submitted in fulfillment of the requirements for the degree of

Master of Science

(Biochemistry)

at

Rhodes University

by

Lithalethu Mkabayi

March 2016

ABSTRACT

Information regarding drug metabolism in veterinary species, especially horses, remains fragmented and incomplete. This information is essential for detection of metabolites of potential performance-enhancing substances in horseracing and for veterinary drug development. Equine liver microsomes have been used to study metabolism of a limited number of drugs, but these provide little information about individual drug metabolizing enzymes. Recombinant CYP enzyme systems are commonly used to determine contribution of individual CYP to metabolism of specific drugs. A limited number of recombinant equine CYPs have been expressed in insect cells and mammalian cell lines. However, there are no reports of recombinant equine CYP1A2 enzyme. In this study, equine CYP1A2 was identified, codon-optimized, cloned and expressed in *E. coli* BL21 cells. Multiple sequence alignments of equine CYP1A2 revealed an amino acid sequence identity of 83.69% to its human homolog which has previously been expressed in *E. coli*. The enzyme was expressed using both auto-induction and IPTG induction. Expressed equine CYP1A2 had a size of about 55 kDa, and was insoluble after cell lysis. Sarkosyl-solubilized CYP1A2 was purified using nickel affinity chromatography and gel filtration. For activity reconstitution, yeast NADPH-cytochrome P450 reductase was first expressed in *E. coli* BL21 cells and exhibited activity of 0.13 U/ml. Activity assay with Glo-P450 CYP1A2 assay kit indicated that CYP1A2 was inactive. Despite numerous attempts to obtain the activity, the CYP1A2 remained inactive. Although expression of equine CYP1A2 in *E. coli* produced non-catalytically active enzyme, this study could be used as the first step in an effort to fully develop a recombinant equine CYP1A2 system.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and appreciation to the following people, without whom this work would not have been possible:

To my supervisor and mentor, Dr. Brendan Wilhelmi, your encouragement, support and endless patience gave me strength especially when things seemed challenging.

To members of lab 412, for their support and friendship. Special thanks to Roxanne Morrison and Maleke Nkululeko Lesero for their prior contributions to this research.

To the members of the Department of Microbiology and Biochemistry for the support.

The financial assistance from National Research Foundation (NRF) toward this research.

To my family and friends for their endless love and support throughout this research.

DECLARATION

I hereby declare that this is original work completed by myself, the undersigned, and is submitted for the degree of Master of Science of Rhodes University.

.....

Lithalethu Mkabayi

March 2016

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

A	Absorbance
AhR	Aryl hydrocarbon receptor
ALA	δ-aminolevulinic acid
ANF	α-naphthoflavone
BCA	Bicinchoninic acid
bp	Base pair
BSA	Bovine serum albumin
BLAST	Basic local alignment search tool
CYP	Cytochrome P450
CPR	NADPH-cytochrome P450 reductase
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
FAD	Flavin adenine dinucleotide
Fe	Iron/heme
FMN	Flavin mononucleotide
HPLC	High pressure liquid chromatography
h	Hour
IPTG	Isopropyl-1-thio-D-galacto-pyranoside
kDA	Kilo Dalton
LA	Luria agar
LB	Luria broth
NADPH	Nicotinamide adenine dinucleotide phosphate reduced tetrasodium salt
NCBI	National Center for Biotechnology Information

PAHs	Polycyclic aromatic hydrocarbons
PCR	Polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sec	Second
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
TCDD	2,3,7,8-tetrachlorodibenzo-*p*-dioxin
U	Units
V	Volts
v/v	Volume per volume
w/v	Weight per volume
XRE	Xenobiotic response element
×	Times
× *g*	Relative centrifugal field

LIST OF FIGURES

CHAPTER 1: LITERATURE REVIEW

1.1 Introduction

There has been a continuous increase of drugs used in veterinary practice; this is especially true for horses, which often receive extensive medical care. A detailed understanding of biotransformation of these drugs is important for both therapeutic and regulatory purposes. The biotransformation of several drugs has been well explored in different mammals such as humans and many laboratory animal species. However, there is limited information on horse drug metabolism. The lack of knowledge regarding equine drug metabolism presents a challenge for medicine and especially for regulatory laboratories. The current trend in equine medicine involves use of drugs that are effective in other species, while regulatory laboratories are often dependent on information obtained from different species to detect metabolites of potential performance-enhancing substances. Extrapolation of information regarding pharmacokinetics and pharmacodynamics from other species is not always accurate and there is evidence of significant differences in expression levels of drug metabolizing enzymes, substrate turnover, substrate ranges and enzyme kinetics between species (Guengerich, 1997; Flink-Gremmels, 2008).

It is challenging to conduct *in vivo* drug metabolizing studies in horses, this is because of the cost associated with maintenance of the studies, size of the animal, difficulties in obtaining research animals and ethical considerations. An alternative approach is to use *in vitro* systems which allow for determination of possible routes of metabolism. A few studies have reported the use of equine liver microsomes for drug metabolism (Chauret *et al.*, 1997; Latritz *et al.*, 2000; Nebbia *et al.*, 2003; Darwish *et al.*, 2010). However, studies with equine liver microsomes provide little

information pertaining to individual drug metabolizing enzymes. For better understanding of horse drug metabolism, it is necessary to isolate and characterize individual drug metabolizing enzymes, which can assist in determining which enzymes are responsible for metabolism of specific drugs. Recombinant enzyme systems are usually used for their ability to provide information on metabolism of drugs by individual enzymes. This information could be used to develop ways for identification of metabolites of potential performance-enhancing substances in horseracing.

In humans several drug metabolizing enzymes, called cytochrome P450s (CYPs), have been well characterized, and only six have been shown to metabolize the majority of clinically used drugs. These include CYPs from the families CYP1, CYP2 and CYP3. Human CYPs have been well studied using recombinant cytochrome P450 enzymes expressed in systems such as bacteria, yeast cells, mammalian and insect cells (Crespi *et al.*, 1997; Gonzalez and Korzekwa, 1995; Clarke, 1998; Chen *et a*l., 1996). There are limited reports on recombinant expression of equine cytochrome P450s (DiMaio Knych and Stanley, 2008; DiMaio Knych *et al.*, 2009; DiMaio Knych *et al.*, 2010; Peters *et al.*, 2013; Schmitz *et al.*, 2014; Dettwiler *et al.*, 2014).

CYP1A2 plays a major role in drug metabolism in humans. Studies with equine liver microsomes suggest the presence of CYP1A2 in horses, based on metabolic activity with selective substrates (Darwish *et al.*, 2010). However, there are no reports on recombinant expression of equine P450 1A2 enzyme. This study aims to develop a recombinant equine CYP1A2 system which could potentially be used for determination of metabolites in drug metabolism.

1.2 Cytochrome P450s

Cytochrome P450 (CYPs) represent a super family of heme enzymes involved in metabolism of several endogenous and exogenous substrates. These enzymes are found in all organisms, from bacterial to mammalian species, where they are primarily located in the smooth endoplasmic reticulum of hepatocytes. The name cytochrome P450 is based on the observation that these enzymes absorb strongly at 450 nm when reduced with carbon monoxide (Omura and Sata, 1964). CYPs are classified into families and sub-families based on the amino acid sequence identity. Classification is in accordance with a standard nomenclature system established by Nelson *et al.*, (1996). The abbreviation CYP is followed by an Arabic number which represents the family (e.g. CYP1). A family is made up of CYPs with more than 40% sequence identity. Subfamilies are created when CYPs share more than 55% sequence identity and are designated by a letter (e.g. CYP1A). Individual enzymes within a family are characterized by a number (e.g. CYP1A2); the number refers to the order of discovery. Individual CYPs belonging to a subfamily should differ by 3% from other members of the same subfamily.

CYPs from different families usually have less than 20% sequence identity, yet appear to have a common structural fold which is highly conserved. The CYP tertiary structure is composed of a helix-rich domain and a beta sheet domain which contains anti-parallel beta strands (Meunier *et al.*, 2004). The conserved structural features include a four-helix bundle, namely helices D, E, I and L and J and K. Helices K and I contain a highly conserved Glu-X-X-Arg motif and threonine residue, respectively. The Glu-X-X-Arg motif is probably the key element in structural stabilization. The heme prosthetic group (iron-protoporphyrin) is located in helix l. The heme iron is coordinated to a thiolate of absolutely conserved cysteine residue. This cysteine and several flanking residues are highly conserved across mammalian CYPs. Another important

structural element is the hydrophobic substrate access channel which is formed by beta sheets 1 and 2. The most variable regions in CYPs structure are those involved in controlling redox partner binding, substrate recognition and binding. These structural differences in substrate recognition and binding contribute to variability of CYP substrates.

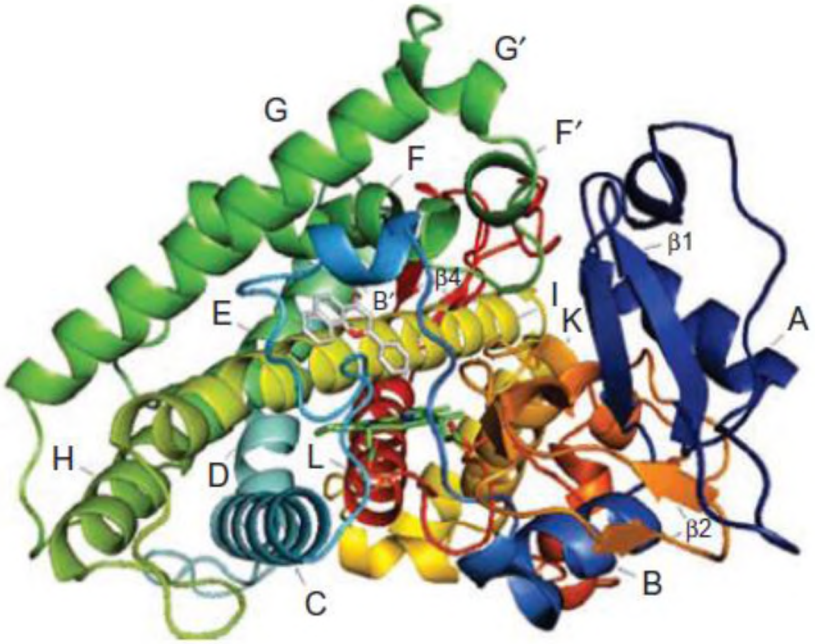


Figure 1.1: Structure of human CYP1A2 in complex with α-naphthoflavone (ANF). The twelve helices A-L are shown in green, blue and red, and anti-parallel beta sheets (ß1, ß2, ß3 and ß4) in similar colors (taken from Zhou *et al*., 2010).

CYPs catalyze oxidation reactions and require a redox partner to facilitate transfer of electrons from electron donor, mostly NADPH to CYP's heme prosthetic group. In CYP-catalyzed reactions, electron transfer is facilitated by the membrane-bound multidomain flavoprotein NADPH-cytochrome P450 reductase (CPR). CPR contains two prosthetic groups, namely flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN), which are located in different

domains (Hannemann *et al.*, 2007). These are key intermediate cofactors between NADPH (a two electron donor) and the CYP heme prosthetic group (a one electron acceptor). The FAD cofactor accepts two electrons from NADPH and transfers them one at a time to FMN cofactor, which in turn donates them to the CYP heme prosthetic group. CYP utilizes these electrons for incorporation of one atom of molecular oxygen into the substrate. The electron transfer pathway between CPR and heme is maintained by several amino acids, including the absolutely conserved cystein residue. A schematic representation of electron-transfer chain between CPR and CYP is shown in figure 1.2.

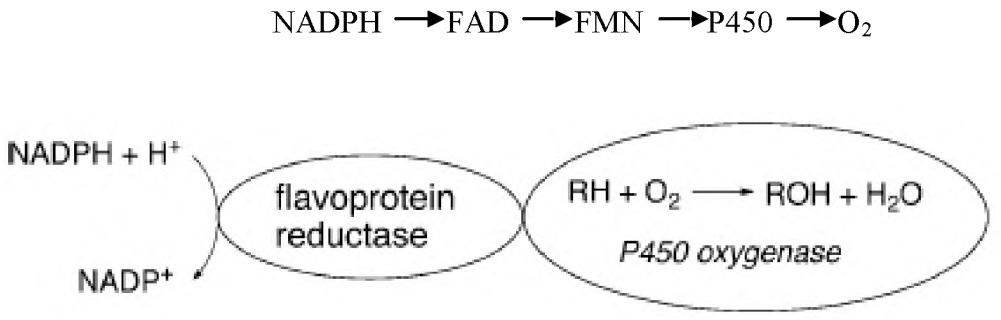


Figure 1.2: Schematic representation of the electron-transfer chain between reductase and cytochrome P450 enzymes (adapted from Meunier *et al.*, 2004).

The cytochrome P450s play a key role in both metabolic and biosynthetic reactions. Particular attention has been given to the metabolic reactions due to the functional outcomes of drug metabolism. The oxidative reactions catalyzed by these enzymes involve the introduction of an oxygen atom to the substrate, usually a saturated hydrocarbon, and reduction of a second oxygen atom to form a water molecule. The electrons required are donated by NADPH via the NADPH-cytochrome P450 reductase in an electron-transfer chain. CYPs are able to catalyze these reactions because of their ability to alter the oxidative state of the iron atom of the heme

prosthetic group. The general catalytic mechanism involves a number of controlled steps. Although other mechanisms of action have been proposed, the mechanism outlined in figure 1.3 is the widely accepted catalytic cycle.

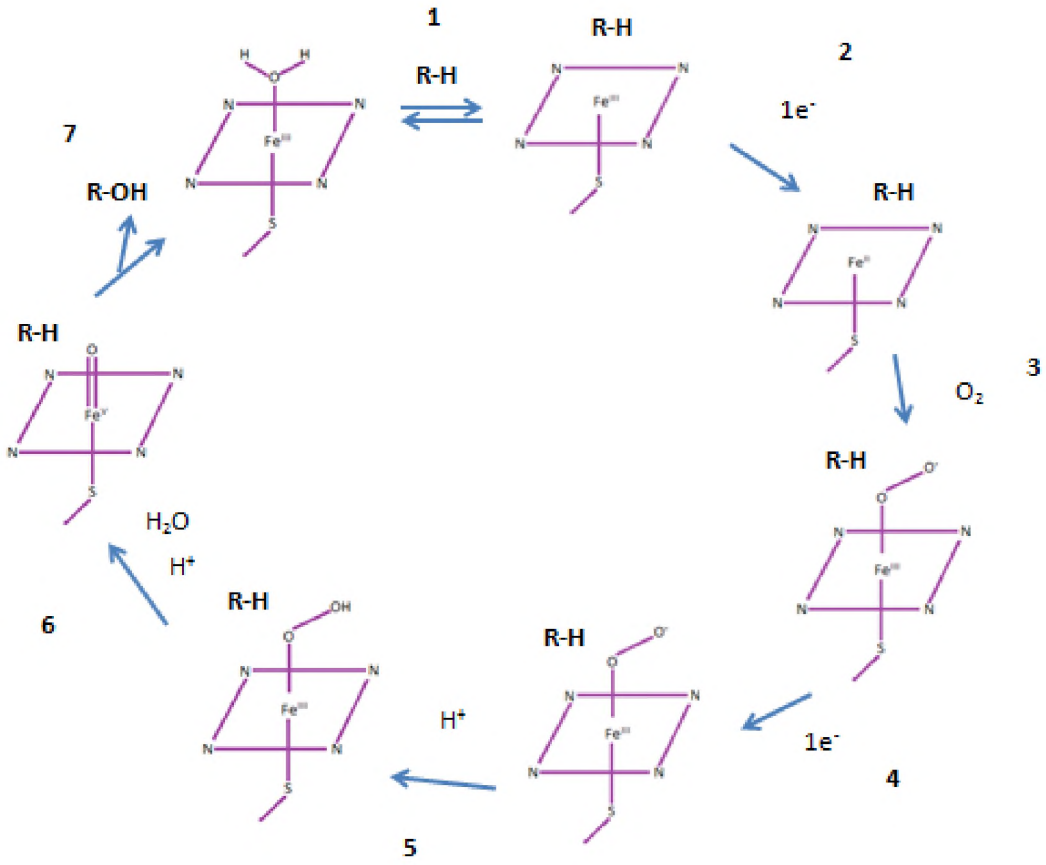


Figure 1.3: The general catalytic cycle of cytochrome P450 enzymes. 1: Binding of the substrate (R-H). 2: Reduction of heme iron by NADPH-P450 reductase. 3: Binding of oxygen and formation of dioxygen complex of CYP. 4: Second reduction of heme iron. 5: first protonation (generation of nucleophilic Iron(III)-hydroperoxo species). 6: Second protonation and formation of an electrophilic iron-oxo intermediate. 7: Product (R-OH) release and return of enzyme to its original state (adapted from Meunier *et al.*, 2004).

1.3 Cytochrome P450 1A2

1.3.1 CYP1A subfamily

Among the CYP enzymes, families 1–3 constitute the majority of CYPs in mammals. The subfamily CYP1A contains two enzymes, namely CYP1A1 and CYP1A2, involved in metabolism of polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls, nitrogenous heterocyclic and aromatic amines (Nebert *et al.*, 2004; Nebert and Dalton, 2006). CYP1A1 and CYP1A2 share more than 70% amino acid sequence (Kimura *et al.*, 1986; Omiecinski *et al.*, 1990) and show overlapping substrate specificities. Human CYP1A2 is one of the major enzymes expressed in the liver and contributes greatly to drug metabolism, as it metabolizes about 20% of clinically used drugs. Human CYP1A2 specifically metabolizes certain substrates, which have been used as markers of its metabolic activity. These include phenacetin *O*-deethylation, ethoxyrosorufin *O*-deethylation and caffeine 3-demethylation.

In the horse, the presence of this enzyme has been suggested based on the studies with selective substrates (Chauret *et al.*, 1997; Latritz *et al.*, 2000; Nebbia *et al.*, 2003; Darwish *et al.*, 2010). Latritz *et al.*, (2000) reported metabolism of ethoxyrosorufin by hepatic and pulmonary microsomes in horses. Phenacetin *O*-deethylation, another catalytic marker of CYP1A2, was observed in equine liver microsomes (Chauret *et al.*, 1997). Tyden *et a*l., (2014) also examined gene and protein expression of enzymes belonging to the CYP1A subfamily in equine intestine and liver. The study demonstrated high levels of gene expression for enzymes from the CYP1A subfamily in small intestines. The presence of CYP1A subfamily enzymes was confirmed in the liver and intestine by western blot analysis. These studies suggest that equine CYP1A2 is involved in metabolic activities that are similar to its human homolog.

1.3.2 Substrate specificity of CYP1A2

The general feature of CYP1A2 substrates is a planar ring, which enables substrate binding to the planar active site of the enzyme. The well-known CYP1A2 substrates are phenacetin, caffeine, theophylline and ethoxyresorufin, and therefore the activity of this enzyme in *in vitro* studies is commonly determined by phenacetin *O*-deethylation, caffeine 3-demethylation and ethoxyroserufin *O*-deethylation (Waxman and Chang, 2006). For *in vivo* studies, melatonin 6-hydroxylation, tarcine 1,2,3,4-tetrahydro-9-aminoacridine, 1-hydroxylation and tizanidine oxidation are used as the most common probes (Momo *et al.*, 2004; Karjalaine *et al.*, 2007). A broad spectrum of clinical drugs are metabolized preferentially by this enzyme. The contribution of CYP1A2 to metabolism of these drugs is estimated to be greater than 30%. These include clozapine, aminopyrine, olanzapine, pimobendan, tizanidine, dacarbazine, duloxetine, phenacetin, theophylline, melatonin, leflunomide, riluzole and tacrine (Zhou *et al.*, 2009).

CYP1A2 is also involved in metabolism of drugs that are substrates of other CYP enzymes, where it plays a minor role, metabolizing less than 20%. These include drugs such as fluvoxamine, acetaminophen, antipyrine, paraxanthine, promazine, naproxen, clomipromine and R-warafin (Zhou *et al.*, 2009). In addition, CYP1A2 is involved in metabolism of natural compounds including flavonoids, aristolochic acids, coumarin, estragole and paenol. These natural compounds are present in several herbal medicines, food flavors and detergents. Table 1 summarizes drugs and natural compounds known to be CYP1A2 substrates.

Table 1: Substrates of CYP1A2 and their metabolic pathways

Substrate	**Metabolic pathway**	**Reference**
Acetaminophen	3-hydroxylation	(Tassaneeyakul *et al.*, 1993)
Aminopyrine	*N*-demethylation	(Niwa *et al.*, 1999)
Caffeine	*N*-demethylation	(Ryu *et al.*, 2007)
Clozapine	*N*-demethylation	(Bertilsson *et al.*, 1994)
Clomipramine	*N*-demethylation	(Wu *et al.*, 1998)
Duloxetine	4-, 5- and 6-hydroxylation	(Lobo *et al.*, 2008)
Flutamide	2-hydroxylation	(Goda *et al.*, 2006)
Leflunomide	N-O bond cleavage	(Kalgutkar *et al.*, 2003)
Naproxen	*O*-demethylation	(Miners *et al.*, 1996)
Olanzapine	4-*N*-demethylation	(Ring *et al.*, 1996)
Paraxanthine	8-hydroxylation	(Tassaneeyakul *et al.*, 1992)
Phenacetin	*O*-deethylation	(Tassaneeyakul *et al.*, 1993)
Promazine	*N*-demethylation	(Wojcikowski *et al.*, 2003)
Riluzole	*N*-hydroxylation	(Sanderink *et al.*, 1997)
Tacrine	1-, 2- and 7-hydroxylation	(Spaldin *et al.*, 1994)
Theophylline	*N*-demethylation	(Ha *et al.*, 1995)

1.3.3 Induction and inhibition of CYP1A2

CYP1A2 is susceptible to induction and inhibition by certaian xenobiotic compounds. The inducible expression of this enzyme is controlled by the aryl hydrocarbon receptor (AhR) (Nebert *et al.*, 2004). Activation of AhR is triggered by several compounds such as 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), polycyclic aromatic hydrocarbons (PAHs), for example, benzo(a)pyrene, 3-methylcholanthrene and β-naphthoflavone. Upon ligand binding (PAHs) in the cytoplasm, AhR translocates to the nucleus where it dimerizes with an AhR nuclear translocater. This heterodimer binds to a xenobiotic response element (XRE) located in the upstream region of CYP1A2 gene promoter. The interaction between XRE and heterodimer leads to upregulation of CYP1A2 transcription.

Several compounds have been reported to be mechanism-based inhibitors of CYP1A2. These include furafylline, zeleuton, rofecoxib, carbamazepine and 5-hydroxythiabendazole (Zhou *et al.*, 2010). Mechanism-based inhibitors lead to formation of metabolic intermediates which bind irreversibly to the enzyme by forming covalent bonds with amino acids in the active cavity (Zhou *et al.*, 2010). Furafylline demonstrates mechanism-based inhibition and is commonly used as a selective inhibitor. Selective and competitive inhibitors include α-naphthoflavone and fluvoxamine. A remarkable feature of inhibitors is a planar ring and small volume. Staskal *et al.*, (2005) also reported the inhibition of human and rat CYP1A2 activity by dioxins such as TCDD in a concentration-dependent manner.

1.3.4 Structural features of CYP1A2

Data on mammalian CYP structures was obtained from crystal structure, site directed mutagenesis and homology models. There are as yet no reports on the structure of equine CYP1A2; however, the human analogue has been determined using crystallographic methods (Sansen *et al.*, 2007). Human CYP1A2 shares 83.69% amino acid sequence identity with the equine enzyme. Given the high sequence identity and conservative nature of mammalian CYPs, the human structure can serve as a template for gaining insights into structural features unique to this enzyme.

Before determination of the first crystal structure of CYP1A2 bound to α- naphthoflavone by Sansen *et al.*, (2007), studies depended on methods such as ligand-based models, site-directed mutagenesis and homology modeling using rabbit CYP2C5 as a template (Kim and Guengrich, 2004; Lewis *et al.*, 2003), which identified critical residues and the enzyme active site. As observed with other mammalian CYP structures, CYP1A2 has 12 alpha-helices (A-L) and 4

beta-sheets (1-4) (figure 1.4). The conserved secondary structural elements include regions binding a heme prosthetic group and the CPR.

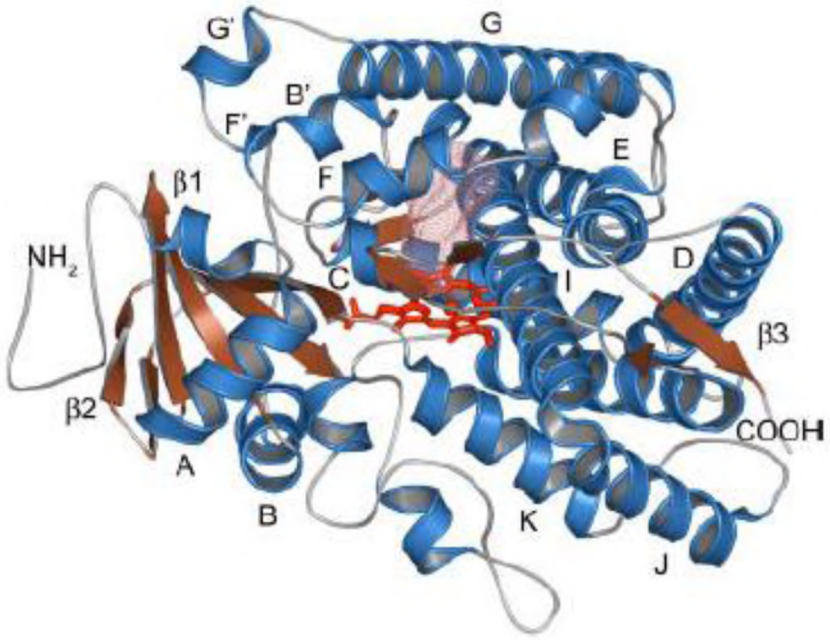


Figure 1.4: Three-dimensional structure of human CYP1A2 in complex with α-naphthoflavone (ANF). The alpha-helices are colored blue, designated A-L and beta-strands are brown, designated 1-4. The heme prosthetic group is red, and the substrate binding cavity is demonstrated as a red mesh surface (taken from Sansen *et al*., 2007).

The other distinct structural features include lengths and loop regions between the conserved secondary elements. The disruption of helix F is also a remarkable feature; this is as a result of losing an alpha-helical hydrogen bonding pattern at residues Val220 and Lys221, which causes unwinding of the helical turn in the middle of helix F (Sansen *et al*., 2007). CYP structures from families 2 and 3 contain compact helix F compared to the CYP1A2. This enzyme also displays an additional beta 3'-sheet between helix H and I and a small alpha-helix (K'') located close to the surface.

1.3.5 Active site and key amino acids of CYP1A2

The substrate binding cavity has been described as narrow and flat and is well adapted to bind planar compounds typical of CYP1A2 substrates, which include caffeine, phenacetin, clozapine, tacrine and theophylline and the inhibitor α- naphthoflavone. The CYP1A2 active site cavity is larger than that of CYP2A6, which is also narrow and flat, and binds small planer compounds such as coumarin, nicotine and naphthalone (Yona *et al*., 2005). However, CYP1A2 contains an active site that is smaller than most CYPs, this includes CYP2D6 (Rowland *et al*., 2006), CYP2C9 (William *et al*., 2003) and CYP3A4 (Yona *et al*., 2004).

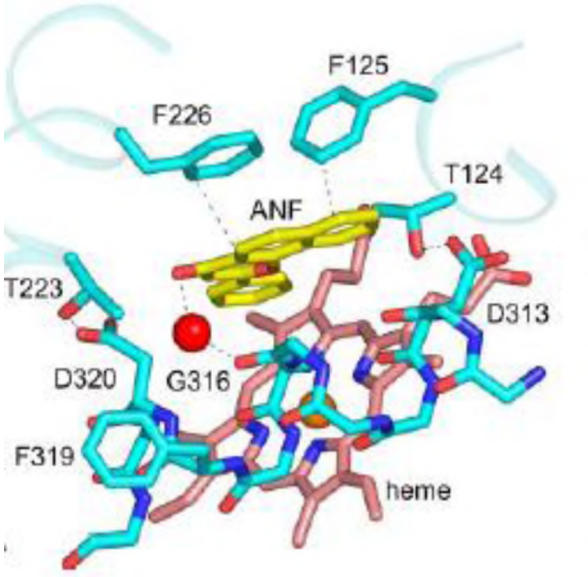


Figure 1.5: The binding of α- naphthoflavone in human CYP1A2 active site. The amino acid residues constituting the active site cavity are represented in blue sticks. The heme prosthetic group and inhibitor α- naphthoflavone are represented in pink and yellow sticks, respectively. The water molecule is shown as a red sphere (taken from Sansen *et al*., 2007).

The key amino acid residues in the active site cavity are on helices F and I, which are in proximity to the heme prosthetic group. The residues that contribute to the lining of the active

site include the side chain of Phe226 of helix F, peptide bonds of Gly316 – Ala317 and Asp320 – Thr321 (figure 1.5). Site-directed mutagenesis and docking methods have identified these residues for substrate binding. In the homology model of CYP1A2 derived from rabbit CYP2C5, Lewis *et al.*, (2003) suggested that Phe226 and Thr501 serve as substrate contacts via hydrogen bonding. In addition, substitution of Asp320 and Phe220 led to reduced catalytic activity which indicates the importance of these amino acids in forming a compact active site (Yun *et al.*, 2000). Sansen *et al.*, (2007) has described 22 residues lining the active site cavity, 7 of which are semi-conserved across human family 1. This might account for overlapping substrate specificities within this group of enzymes.

Elucidation of the crystal structure of CYP1A2 has provided important information on the mechanism for binding of substrates and inhibitors, and the key amino acid residues involved. The structure provides insights into how the active site cavity is well adapted for binding compounds with unique structural properties. This assists in understanding the substrate specificity of the enzymes in family 1. Following the recent completion of the horse genome sequence, the human CYP1A2 crystal structure can serve as a means of providing a rational platform for exploring horse CYP1A2 three-dimensional structure in *in silico* combined with *in vitro* studies. The information generated from these studies may help in understanding the major differences between human and horse CYP1A2. Several aspects relating to conservation of key residues, substrate specificity and metabolism could then be elucidated using these studies.

1.4 Cytochrome P450 expression systems

In mammals, there is a need for better understanding of the molecular mechanism of drug metabolic pathways. This is because of the danger posed by toxic products, unwanted reactions, and regulatory laboratories require information for detection of performance-enhancing substances in numerous sporting events. It is often difficult to perform *in vivo* studies because of difficulties in obtaining research animals, and ethical constraints. *In vitro* experimental approaches are usually based on animal hepatic microsomal CYP-containing systems and recombinant mammalian CYP enzymes expressed in several host systems. Studies based on hepatic microsomes unfortunately offer little information on individual CYP. Recombinant systems provide an alternative to obtain specific knowledge regarding the metabolism of compounds by individual CYP. They also offer means of assessing CYP activity in the absence of other CYP enzymes.

Several expression systems have been used. These include bacterial cells, insect cells, yeast and mammalian cells. Expression of CYPs established in mammalian cell lines is usually achieved by transfection of mammalian recombinant plasmids into cell lines such as COS-1 (Cloutier *et al.*, 1997; Crespi *et al.*, 1997). The advantage of this system is the presence of necessary proteins such as reductase, which are able to sustain CYP activity (Zuber *et al.*, 1988). They also provide an environment similar to native physiological conditions of mammalian CYPs. However, expression of recombinant proteins in this system also presents challenges such as low levels of expression and low yields.

Yeast is one of the eukaryotic expression systems used for production of CYPs. A zebrafish CYP3C1 was heterologously expressed in yeast *Saccharomyces cerevisiae* and showed the typical CYP absorption peak at 450 nm (Corley-Smith, 2006). In a study by Urban *et al.*, (1990),

a mouse CYP1A1 was co-expressed with human NADPH-cytochrome P450 reductase in yeast. This system allowed maximal CYP1A1 activity. Another eukaryotic expression system is baculovirus-insect cells. High-level expression of human and rat CYP2E1 in a baculovirus expression system has been reported (Chen *et al.*, 1996). Both enzymes were able to metabolize CYP2E1 substrates (chlorzoxazone, *p*-nitrophenol and acetaminophen). A catalytically active human CYP1A1 has also been expressed in a baculovirus system (Buter *et al.*, 1995). Both insect and yeast provide a eukaryotic environment which is normally required for proper folding and also offer an environment conducive to the membrane integration of CYP enzymes in endoplasmic reticulum. However, each of these has limitations for expressing recombinant CYPs such as low yields and low growth rate.

Most recombinant CYPs have been produced using bacterial expression systems, with the *E. coli* system favoured. CYPs from almost all major families have been expressed using this expression system, including CYP1A2 (Messina *et al.*, 2008), 2C9 (Hao *et al.*, 2007), 2D6 and 3A4 (Pan *et al.*, 2011). Bacterial systems have several desirable features, namely the ability to rapidly express very high levels of functional proteins, simplicity in development and maintenance and the absence of native CYPs, which offers a better background for activity studies.

The major disadvantage is that bacteria are prokaryotic and lack the eukaryotic posttranslational modifications and proper signal for CYP membrane integration. In some cases, the use of a bacterial system has resulted in insoluble enzymes which can form inclusion bodies. The most common strategy to resolve this has been to modify the NH_2-terminus. This is usually achieved by replacing the NH_2-terminus with the sequence (MALLLAVF), which is from CYP17A, the first mammalian CYP enzyme to be expressed in *E. coli* (Barnes *et al.*, 1991). Another strategy involves use of amphipathic peptide to replace NH_2-terminus hydrophobic sequence (Waterman

et al., 1995). The removal of NH_2-terminus hydrophobic signal anchor sequence from CYP enzymes might lead to a soluble and active enzyme. In other cases, truncated CYPs are found to be insoluble and non-functional. It has been reported that co-expression of CYPs with molecular chaperones which are known to assist in folding of bacterial proteins leads to catalytically active CYPs (Ahn *et al.*, 2004). Pritchard *et al.*, (1997) employed an alternative approach, where the NH_2-terminus was fused to a bacterial signal peptide. The signal sequence targets CYPs to the membrane and cleavage of the sequence results in production of full-length CYPs. This approach yielded unmodified functional proteins. Although bacteria do not have the necessary endogenous reductase proteins, the activity of recombinant CYPs can be reconstituted with a purified reductase system.

1.4.1 Equine recombinant CYPs

Most equine drug metabolism studies have been conducted using an *in vivo* approach (Scarth *et al.*, 2011). The drawbacks of *in vivo* work are ethical considerations, high cost, precautionary measures in sample handling, recovery in sample extraction and presence of interferences. An alternative is the use of *in vitro* methods which usually involve equine liver microsomes and lung preparations. The use of the two has advantages and limitations, such as availability of equine tissue and administration of CYP-inducing compounds before tissue preparation. However, after sequencing of the horse genome, a number of CYPs have been annotated. More recently, a limited number of CYPs have been recombinantly expressed. These include CYP2D50, CYP2C92, CYP2B6 and CYP3A members (DiMaio Knych and Stanley, 2008; DiMaio Knych *et al.*, 2009; Peters *et al.*, 2013; DiMaio Knych *et al.*, 2010; Schmitz *et al.*, 2014; Dettwiler *et al.*, 2014).

A baculovirus-insect expression system was used for expression of CYP2D50, CYP2C92 and CYP3A members (DiMaio Knych and Stanley, 2008; DiMaio Knych *et al.*, 2009; DiMaio Knych *et al.*, 2010). In other reports, CYP2B6 and CYP3A members were heterologously expressed in mammalian cells, namely, V79 Chinese hamster lung fibroblasts (Peters *et al.*, 2013; Schmitz *et al.*, 2014). All the recombinant CYPs showed activity towards their respective selective substrates and their metabolic capacity was compared with human orthologs. To date, there are no reports on expression of equine CYPs in a bacterial expression system. There are also no reports on heterologous expression of CYP1A members currently available. Tyden *et al.*, (2014) examined the CYP1A gene and protein expression levels in horse intestine and liver, and observed high levels of expression. This is the only recent publication on equine CYP1A, and functional characterization of recombinant CYPs from this family has not been reported. Availability of recombinant CYP1A members will enhance understanding of mechanism of drug metabolism in horses. It can also serve as a useful tool in closing the knowledge gap between human and horse drug metabolism.

1.4.2 Recombinant CYP1A2 in other mammals

Several publications have described the cloning, expression, purification or characterization of mammalian CYP1A2, including rat, mouse, cat, pig, monkey and human (Guengerich, 1997; Tanaka *et al.*, 2006; Messina *et al.*, 2008; Narimatsu *et al.*, 2005). A variety of expression systems have been utilized. Fisher *et al.*, (1992) reported the first heterologous expression of human CYP1A2 in *E. coli*. The high expression level was achieved through modification of the NH_2-terminal region using the sequence from CYP17A, which is believed to promote membrane insertion and correct folding. However, Kim *et al.*, (2008) reported the first human CYP1A2 to be functionally expressed in *E. coli* without modification of the NH_2-terminal region. The

metabolic capacity of the enzyme was compared to that of NH_2-terminally modified and showed very similar catalytic activity towards a number of typical CYP1A2 substrates.

Extensive research has been conducted using modified recombinant CYP1A2; this includes studies on structure-activity relationships (Parikh *et al.*, 1999; Yun *et al.*, 2000) and the crystal structure (Sansen *et al.*, 2007). Although it is believed that changes in the NH_2-terminal region do not affect catalytic capacity, use of unmodified enzyme might further confirm these claims. This approach also offers an easy method for construction of expression plasmid. After the functional expression human CYP1A2 without modification at the NH_2-terminal region in *E. coli*, investigations regarding equine CYP1A members can be facilitated.

1.5 Problem statement

Most of the current knowledge on drug metabolism is based on studies conducted in humans, and information about horse drug metabolism is limited because of the difficulties associated with *in vivo* studies. However, this does not reduce the importance of understanding drug metabolism in horses from a regulatory perspective. CYP1A2 is one of the major CYPs in the liver and metabolizes about 20% of clinically used drugs, but little is known about the activity of this enzyme in equines. Studies with equine liver microsomes have suggested the presence of CYP1A2 based on metabolic activity with selective substrates. However, there are no reports on recombinant expression of equine CYP1A2 enzyme. This study aims to express a recombinant equine CYP1A2 in an *E. coli* system, and to characterize this enzyme. An active equine CYP1A2 would potentially be useful in metabolic studies of certain drugs used for doping in horseracing.

1.6 Aim and objectives

To develop a recombinant equine CYP1A2 enzyme system.

To achieve this aim, the following objectives were set:

Identify equine CYP1A2 protein using bioinformatics approach

Ligate equine CYP1A2 gene into pET-22b(+) expression vector

Express equine CYP1A2 enzyme in *E. coli* BL21 cells

Purify equine CYP1A2 enzyme using nickel affinity chromatography

Express cytochrome P450 reductase in *E. coli* BL21 cells

Assess the CYP1A2 activity using P450-Glo assay kit

CHAPTER 2: BIOINFORMATIC IDENTIFICATION AND CLONING OF CYP1A2 GENE INTO EXPRESSION VECTOR, pET-22b(+)

2.1 Introduction

CYP enzymes belonging to humans and laboratory animals have been thoroughly characterized, but knowledge of equine CYPs is still limited. Characterization of equine CYPs is essential for better understanding of horse drug metabolism. The completion of the horse genome sequence is a significant development for CYP research. It opens possibilities for faster progress for the characterization of these enzymes. At present, most of the annotated CYP-coding genes are generated using *in silico* prediction methods. The experimental data to describe functional importance of these enzymes remain fragmented and incomplete.

A limited number of studies of equine CYP genes have been conducted. A comparison of human and horse sequence for the CYP3A subfamily gene cluster and genomic organization of the subfamily was reported by Schmtz *et al.*, (2010). It was demonstrated that horse genome has six highly similar CYP3A genes and two pseudogenes. This differs from the human CYP3A gene subfamily, which has four known functional genes and two pseudogens. Recently, equine CYP2B6 and five other potential genes of the 2B subfamily were identified from the horse genome sequence (Peters *et al.*, 2013). The CYP2B6 was recombinantly expressed in V79 Chinese hamster fibroblasts. The equine CYP2B6 *N*-demethylated ketamine to norketamine and generated metabolites of norketamine. Identification and characterization of more equine orthologs will provide information for better understanding of horse drug metabolism.

In silico predictions provide a starting point to investigate functional features of CYPs from horse genome, especially when experimental data are not available. These methods serve as tools

for comprehensive comparison of DNA and protein sequences. Sequence alignment analysis for CYPs is commonly used for prediction of defined structural features (Peterson and Deisenhofer, 1995; Sirim *et al.*, 2010). These assist in identifying conserved secondary structural elements (alpha helices and beta sheets) present in all CYP structures. For function prediction, these methods rely on identifying conservation patterns and similarities between CYP of unknown function and a well-characterized CYP. For example, human CYP1A2 structure was reported by Sansen *et al.*, (2007) and function of this enzyme is well known. Thus, sequence comparison between human and horse CYP1A2 could provide useful information on structure and function of horse CYP1A2. However, these methods can only provide reasonable predictions. Experimental data are always required for confirmation.

DiMaio Knych and Stanley, (2008) reported the first cloning and functional characterization of equine CYP, namely CYP2D50 in insect cells. The experimental approach traditionally includes isolation of total RNA from liver tissue by using polymerase chain reaction. The cDNA for equine CYP is usually obtained by reverse transcription with primers designed against predicted CYP sequence from horse genome. The cDNA is then cloned into plasmid vector for sequencing and construction of expression construct; however, this approach may pose challenges. In cases where genes are cloned for protein expression, production of recombinant proteins in *E. coli* can be subject to biased codon usage (Burgess-Brown *et al.*, 2008). The difference in codon bias between *E. coli* and the gene to be expressed may result in amino acid substitution and low levels of expression (Kane *et al.*, 1995; Calderone *et al.*, 1996).

Improvements in technology have produced an alternative technique, namely gene synthesis. The technique has been refined over the years and proved to be useful in well-established research fields as it offers essential benefits over traditional cloning approaches. Gene synthesis allows for

establishment of a cost-effective, fast and efficient alternative. With gene synthesis, a sequence for the gene to be cloned can be synthesized and codon optimized for bacterial expression. Zhou *et al.*, (2004) reported that codon optimization enhanced the expression level of recombinant *Plasmodium falciparum* protein (FALVAC-1) by threefold, and improved the *E. coli* growth after induction with IPTG.

In the present study, equine CYP1A2 sequence was retrieved from NCBI database and analyzed with multiple alignments. Virtual cloning (pDRAW32) was used to design the expression construct. The CYP1A2 sequence was then codon optimized for bacterial expression and synthesized by GenScript. The synthesized gene was inserted into a cloning plasmid, pUC57 for long term-storage. The CYP1A2 gene was subcloned into an expression plasmd, pET-22b(+). Ligation was confirmed using restriction double digest and sequencing.

2.2 Material and methods

2.2.1 Chemicals and reagents

Plasmid DNA and gel extraction kits were purchased from BioFlux (Bioer Technology Co., Ltd.). Reagents for preparation of Luria agar (LA) and Luria broth (LB) media were purchased from Merck. *E. coli* strains JM109 and BL21 (DE3) cells were purchased from Invitrogen. Restriction enzymes and T4 DNA ligase were obtained from Thermo Scientific. Buffer O and DNA ladder were purchased from Inqaba Biotechnical Industries (Pty, Ltd). The plasmid pET-22b(+) was purchased from Novagen.

2.2.2 Identification and synthesis of CYP1A2 gene into cloning vector pUC57

The equine CYP1A2 protein sequence was obtained from the National Center for Biotechnology Information (NCBI) database (accession number: XP_001493936). The CYP1A2 protein

sequences of mammals that have been characterized were also obtained from NCBI database. The selected mammals were human (*Homo sapiens*), rat (*Rattus norvegicus*), pig (*Sus scrofa*) and mouse (*Mus musculus*). Multiple alignments of protein sequences were conducted using Clustal omega program (Sievers *et al.*, 2011). The program was used to identify sequence identities and conservation of amino acid residues. Multiple sequence alignments were further analyzed using MEME Suite program for motif identification (Bailey *et al.*, 2009).

Virtual cloning was used to assist in designing the expression construct prior to synthesis. The sequence of expression vector pET-22b(+) was obtained from the Novagen website and loaded into pDRAW32. Equine CYP1A2 DNA sequence was obtained from the NCBI database and loaded into pDRAW32. *Xho*I and *Nde*I were selected as restriction enzymes for double digestion. The absence of restriction enzymes within the CYP1A2 sequence was confirmed. The orientation of His-tag was checked using the open reading frame of CYP1A2, which is available under viewing options on pDRAW32. The sequence was then sent to GenScript where it was codon-optimized and synthesized into cloning vector pUC57.

2.2.3 Plasmid amplification

Competent *E. coli* strains JM109 and BL21 (DE3) cells (Invitrogen) were prepared according to the method described by Hanahan (1983) with modifications. A single colony of *E. coli* cells was used to inoculate LB medium (5 ml) and incubated overnight at 37°C with shaking (150 rpm). The overnight culture was inoculated into super optimal broth (SOB) media (20 g/L tryptone, 5 g/L yeast extract, 0.584 g/L NaCl, 0.186 g/L KCl, 2.034 g/L $MgCl_2$, 2.464 g/L $MgSO_4$), and incubated at 37°C with shaking at 150 rpm until an OD_{600} of 0.6 was reached. The cells were placed on ice for 30 minutes and centrifuged (3000 × *g*, 10 minutes, 4°C). Cells were re-suspended in 80 ml ice cold TB buffer (10 mM HEPES pH 6.7, 15 mM $CaCl_2$, 250 mM KCl

and 55 mM $MnCl_2$) and placed on ice for 10 minutes. Cells were collected by centrifugation (3000 × *g*, 10 minutes, 4°C) and re-suspended in 20 ml TB buffer. Dimethyl sulfoxide (DMSO) was added to the suspension to a final concentration of 7% and the mixture was placed on ice for 10 minutes. The mixture was divided into aliquots (50 µl) and stored at -80°C prior to use.

The cloning vector pUC57 containing the insert (CYP1A2-like gene) was transformed into competent *E. coli* JM109 cells using the heat-shock transformation method (Froger and Hall, 2007). A stock of competent *E. coli* JM109 cells (50 µl) was thawed on ice. The plasmid (2 µl) was mixed with competent cells and the mixture kept on ice for 30 minutes, followed by heat shock at 42 minutes for 30 seconds and cooling on ice for 2 minutes. Super optimal broth with catabolite repression (SOC) media (250 µl) (20 g/L tryptone, 5 g/L yeast extract, 0.584 g/L NaCl, 0.186 g/L KCl, 2.034 g/L $MgCl_2$, 2.464 g/L $MgSO_4$) was added and the mixture was incubated with shaking at 37°C for 1 hour. The mixture was plated onto LA plates (Appendix 1) containing 100 µg/ml of ampicillin, and incubated at 37°C overnight. Single transformation colonies were picked and inoculated into a 5 ml LB culture containing 100 µg/ml of ampicillin and incubated on a Labcon shaker (160 rpm) at 37°C overnight.

2.2.4 Plasmid extraction

The plasmid was extracted from the overnight culture using BioFlux BioSpin plasmid DNA extraction kit according to the manufacturer's instructions (Bioer Technology Co., Ltd.) (Appendix 2) and plasmid DNA was quantified using a Nanodrop 2000 spectrophotometer.

2.2.5 Restriction digest

The plasmid DNA was double digested using *XhoI* and *NdeI* restriction enzymes. The reaction mixture contained 1 µl of each restriction enzyme, 7 µl of plasmid DNA and 1 µl of reaction

buffer (buffer O) and was incubated at 37°C overnight. Double digest products were analyzed by 0.8% (w/v) agarose gel (0.5 µg/ml ethidium bromide) in 1X TAE buffer (0.04 M Tris-HCl, 1 mM EDTA, pH 8.0 and 0.021 mM glacial acetic acid); the gel was run at 100 V for 1 hour and visualized under UV light with a low radiation UV source.

2.2.6 Gel excision

The CYP1A2 gene fragment was extracted from 0.8% (w/v) agarose gel (0.5 µg/ml ethidium bromide) using BioFlux Biospin Gel Extraction Kid. The extraction was carried out as per the manufacturer's instructions (Bioer Technology Co., Ltd.) (Appendix 3) and the DNA concentration was determined using a Nanodrop 2000 spectrophotometer.

2.2.7 Ligation of insert into pET-22b (+)

The expression plasmid pET-22b(+) was amplified and extracted (as described in section 2.2.4), double digested (as described in section 2.2.5 with modifications in volumes of components, 0.5 µl *NdeI*, 1 µl *XhoI* and 7.5 µl plasmid DNA) and extracted from the gel (as described in section 2.2.6). The ligation was carried out using 1: 3 plasmid: insert ratio and 50 ng of plasmid DNA. The following formula was used to calculate the amount of the insert required:

$$\text{Insert} = \frac{50 \text{ ng x size of insert (kb)}}{\text{size of plasmid (kb)}} \text{x } \frac{3}{1}$$

The reaction mixture contained 14.6 ng/µl of insert (CYP1A2), 4.8 ng/µl plasmid DNA, 2 µl reaction buffer, 1 µl T4 DNA ligase and 4 µl distilled water. The mixture was incubated at 4°C overnight. The ligation mixture was transformed into competent *E. coli* JM109 cells using the heat-shock transformation method as described in section 2.2.3. Transformed competent JM109 cells were plated on LA plates containing 100 µg/ml ampicillin and incubated at 37°C overnight. Single colonies were picked and inoculated into LB containing 100 µg/ml ampicillin and

incubated at 37°C overnight. The plasmid DNA was extracted using the method in section 2.2.4. The ligation products were analyzed on 0.8% agarose gel.

2.2.8 Confirmation of the insert

The restriction digest analysis was performed to confirm the ligation of the insert into the expression vector pET-22(b+) by using the restriction enzymes *Xho*I and *Nde*I, as described in section 2.2.5. The pET-22(b+) with CYP1A2 insert was sent for sequencing (Inqaba Biotechnical Industries, Pty, Ltd; South Africa) to further confirm the ligation. The DNA sequence obtained from Inqaba Biotechnical Industries was used in translated BLAST (blastx) to confirm the equine CYP1A2 protein sequence.

2.3 Results and discussion

2.3.1 Bioinformatic identification and multiple alignments

Several equine CYP enzymes, including CYP1A2 have been annotated and made available after the completion of the horse genome sequence. Availability of equine CYP information and sequences of other mammalian CYP1A2 which have been heterologously expressed provide significant comparative analyses. The amino acid sequence of equine othorlog to human CYP1A2 was retrieved as described in section 2.2.1. Multiple alignment among horse, human, rat, mouse and pig CYP1A2 was conducted using the clustal omega program (www.clustal.org/omega/; Sievers *et al.*, 2011). Sequence alignments revealed high amino acid identity to human CYP1A2 (83.69% identity), pig (81.17% identity), rat (75.24% identity) and mouse (75.05% identity). As seen in figure 2.1, several amino acid residues are highly conserved across these mammalian species, as shown by the asterisk (*). It was observed that conserved residues occurred in blocks. Generally, reoccurrence of blocks of highly conserved residues is

indicative of the presence of structurally and functionally important regions. It has been reported that the structure of human CYP1A2 contains twelve alpha-helices and four beta-strand sheets which are usually observed in most CYP structures (Sansen *et al.*, 2007). The presence of these conserved residues and high sequence identity therefore indicates the structural similarities among CYP1A2 of horse, human, pig, rat and mouse. These alignment results provide essential information for development of equine CYP1A2 expression system.

```
rat         MAFSQYISL-APELLLATAIFCLVFWVLRGTRTQVPKGLKSPPGPWGLPFIGHMLTLGKN
mouse       MAFSQYISL-APELLLATAIFCLVFWMVRASRTQVPKGLKNPPGPWGLPFIGHMLTVGKN
pig         MTLFQPSLFSATELLLASAIFCLVFWVVRTWQPQVPKGLKSPPGPWGWPLLGHVLTLGKS
horse       MMLSQLSPFSATELLLASTIFCLVFWVVRAWQPQIPKGLKSPPGPWGWPFLGHVLTLGKN
human       MALSQSVPFSATELLLASAIFCLVFWVLKGLRPRVPKGLKSPPEPWGWPLLGHVLTLGKN
            * : *   : * *****::*******:::  : ::*****.** *** *:::**:**:**.

rat         PHLSLTKLSQQYGDVLQIRIGSTPVVVLSGLNTIKQALVKQGDDFKGRPDLYSFTLITNG
mouse       PHLSLTRLSQQYGDVLQIRIGSTPVVVLSGLNTIKQALVRQGDDFKGRPDLYSFTLITNG
pig         PHLALARLSQCYGDVLQIRIGCTPVLVLSGLDTIRQALVRQGDDFKGRPNLYSFTLVTDG
horse       PHLALSRLSQRYGDVMQIRIGSTPVLVLSGLDTIRQALVRQGDDFKGRPDLYSFTLITNG
human       PHLALSRMSQRYGDVLQIRIGSTPVLVLSRLDTIRQALVRQGDDFKGRPDLYTSTLITDG
            ***:*:::** ****:*****.***:*** *:**:****:**********:**: **:*:*

rat         KSMTFNPDSGPVWAARRRLAQDALKSFSIASDPTSVSSCYLEEHVSKEANHLISKFQKLM
mouse       KSMTFNPDSGPVWAARRRLAQDALKSFSIASDPTSASSCYLEEHVSKEANHLVSKLQKAM
pig         QSMTFNPDSGPVWAARRRLAQKALNTFSIASDPASSSSCYLEDHVSKEAECLLGKFQELM
horse       QSMTFNPDSGPVWAARRRLAQNALNTFSIASDPASMSSCYLEEHVSKEAEALLSRLQKLM
human       QSLTFSTDSGPVWAARRRLAQNALNTFSIASDPASSSSCYLEEHVSKEAKALISRLQELM
            :*:**. ***************.**::*******:* ******:******: *:.::*: *

rat         AEVGHFEPVNQVVESVANVIGAMCFGKNFPRKSEEMLNLVKSSKDFVENVTSGNAVDFFP
mouse       AEVGHFEPVSQVVESVANVIGAMCFGKNFPRKSEEMLNIVNNSKDFVENVTSGNAVDFFP
pig         AGPGHFDPYDHIVVSVGRVIGAMCFGKGFPQSSEEMFSIVKNSHEFVETASSGNPVDFFP
horse       SVAGRFDPSSQVVASVANVIGAMCFGQHFPHSSEEMISLLRSSHEFVQTASSGNPVDFFP
human       AGPGHFDPYNQVVVSVANVIGAMCFGQHFPESSDEMLSLVKNTHEFVETASSGNPLDFFP
            :  *:*:* .::* **..********: **..*:**:.::..:::**:..:*** :****

rat         VLRYLPNPALKRFKNFNDNFVLFLQKTVQEHYQDFNKNSIQDITGALFKHSE-NYKDNGG
mouse       VLRYLPNPALKRFKTFNDNFVLFLQKTVQEHYQDFNKNSIQDITSALFKHSE-NYKDNGG
pig         ILRYLPSPTLQRFKSFNQRLLQFLRKMVQERYRDFDKNCIQDITSALFKHSEENSSTSGG
horse       ILRYLPNPPLQRFKSFNQRFLRFLQKIIQEHYRDFDKNSIQDITGALFKHREKSSRASGV
human       ILRYLPNPALQRFKAFNQRFLWFLQKTVQEHYQDFDKNSVRDITGALFKHSKKGPRASGN
            :*****.* *:*** **:.:: **:* :**:*:**:**.::***.***** :     .*

rat         LIPQEKIVNIVNDIFGAGFETVTTAIFWSILLLVTEPKVQRKIHEELDTVIGRDRQPRLS
mouse       LIPEEKIVNIVNDIFGAGFDTVTTAITWSILLLVTWPNVQRKIHEELDTVVGRDRQPRLS
pig         LISQEKTINLVNDIFGAGFDTITTAISWSLLYLVTYPEIQKKIQKELDTVIGRARRPRLS
horse       LIPQEKIINIINDIFGAGFDTVTTAITWSLTYLVTNPKIQRKIQEELDTVVGRARQPRLS
human       LIPQEKIVNLVNDIFGAGFDTVTTAISWSLMYLVTKPEIQRKIQKELDTVIGRERRPRLS
            ** :** :*::*********:*:**** **:  *** *::*:**::*****:** *:****

rat         DRPQLPYLEAFILEIYRYTSFVPFTIPHSTTRDTSLNGFHIPKECCIFINQWQVNHDEKQ
mouse       DRPQLPYLEAFILEIYRYTSFVPFTIPHSTTRDTSLNGFHIPKERCIYINQWQVNHDEKQ
pig         DRPQLPYMEAFILELFRHTSFVPFTIPHSTTRDTTLNGFYIPKERHVLVNQWQVNHDPKL
horse       DRPQLPYMEAFILETFRHSSFVPFTIPHSTVRDTTLNGFYIPKERCVFINQWHVNHDEEL
```

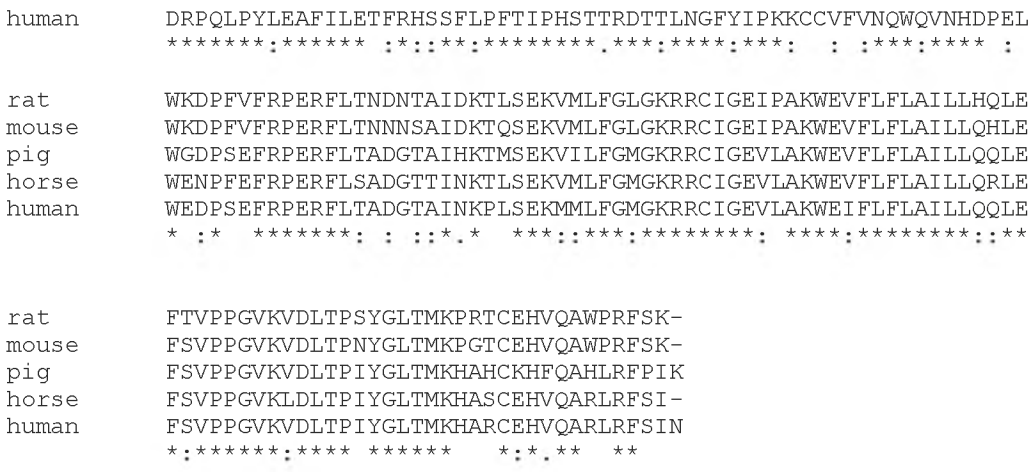


Figure 2.1: Amino acid sequence alignment of CYP1A2 from horse and human, pig, mouse and rat. Multiple alignment was carried out by Clustal omega to determine amino acid identity and conservation among these mammals. An asterisk (*) shows a position which has a fully conserved amino acid residue. A colon (:) indicates conservation between groups of strongly similar properties, while a period (.) indicates conservation between groups of weakly similar properties.

The multiple sequence alignments were analyzed using MEME Suite (www.meme-suit.org; Bailey *et al*., 2009) for determination of sequence motifs. The result in figure 2.2 shows several conserved motifs across the entire protein sequence. Another observation from this analysis is low E-values for all motif sequences, which indicates that the results are statistically significant. High conservation of these sequences suggests that CYP1A2 of selected species might have common catalytic function. It also indicates the presence of significant secondary structural elements in CYP1A2 of horse, human, pig, rat and mouse.

Figure 2.2: The logo of motif discovered by MEME Suite in the CYP1A2 protein sequence from five mammalian species. The conserved sequence motifs are shown on the logo and E-values are displayed on the right hand side.

2.3.2 Subcloning

Prior to synthesis of CYP1A2 gene into pUC57 cloning vector, virtual cloning procedure was employed to assist in designing expression construct using pDRAW32. The program automatically identifies the restriction enzyme sites of a plasmid. As shown in figure 2.3, the two restriction enzymes were found in appropriate positions; the CYP1A2 insert and His-tag were also in the correct orientation. The selected restriction enzymes were not found within the sequence of CYP1A2. Virtual cloning can only be used as an initial step to verify the designed expression construct before carrying out experimental analysis.

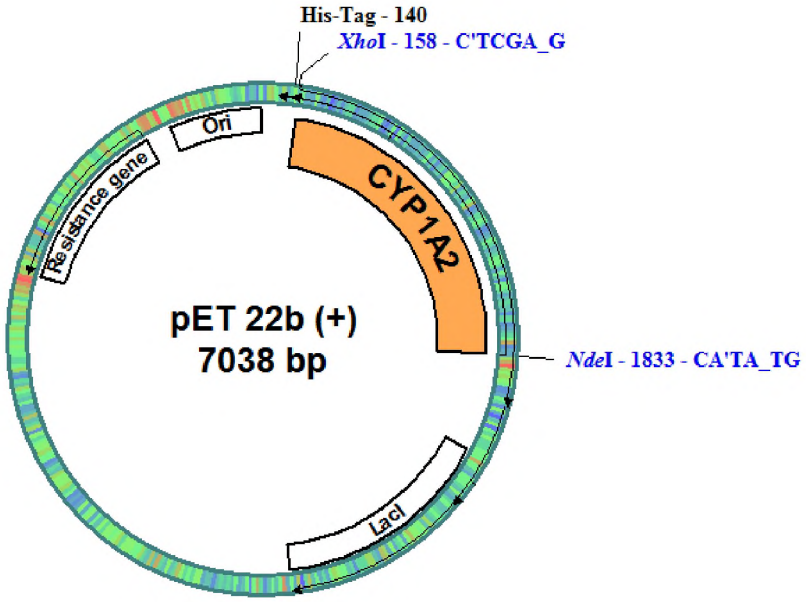


Figure 2.3: The plasmid map of pET-22b (+) plasmid with the CYP1A2 gene. The features highlighted on the map include the ampicillin resistance gene, *Xho*I and *Nde*I restriction enzymes, His-Tag and CYP1A2 gene.

The plasmid pUC57 with CYP1A2 insert was amplified as described in section 2.2.3. The plasmid was digested using restriction enzymes *Xho*I and *Nde*I as described in section 2.2.5. The recognition sequence of the two restriction enzymes is found within the multiple cloning site of pUC57 vector and flanks the CYP1A2 DNA fragment. The resulting DNA fragments were separated by agarose gel electrophoresis (figure 2.4). Two DNA bands representative of plasmid pUC57 and CYP1A2 were observed on lane 2. The sizes of these bands are 2 710 bp and 1 545 bp and are in agreement with sizes of pUC57 and CYP1A2 gene respectively. These results confirmed the presence of the CYP1A2 gene in pUC57.

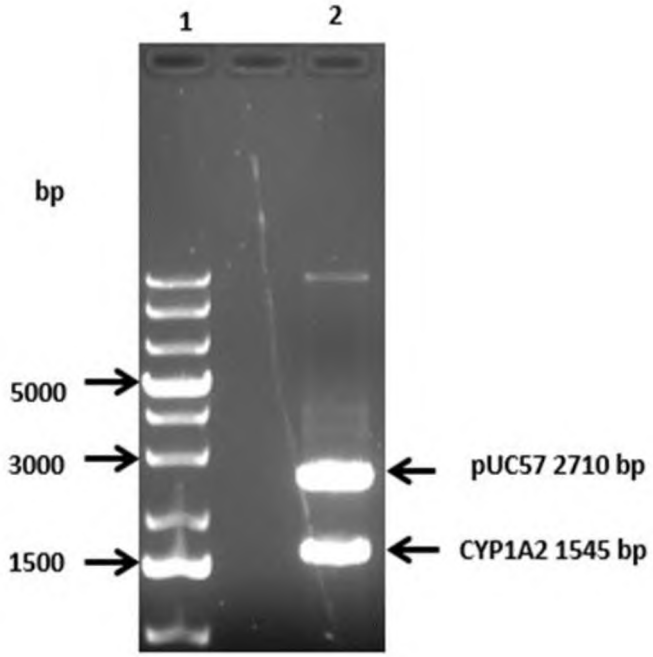


Figure 2.4: Agarose gel electrophoresis of double digestion of PUC57. The plasmid was digested with two restriction enzymes, *Xho*I and *Nde*I and analyzed by agarose gel electrophoresis using 0.8 % (w/v) gel in 1 x TAE buffer. Lane 1 represents the DNA ladder and lane 2 is the double digest products, pUC57 (2 710 bp) and CYP1A2 gene (1 545 bp).

Following the digestion of pUC57, the DNA fragment corresponding to CYP1A2 was extracted from the agarose gel and cleaned as described in section 2.2.6. The expression plasmid pET-22b(+) was subjected to double digestion as described in section 2.2.7 and extracted from the agarose gel. The CYP1A2 DNA fragment was ligated into pET-22b(+) according to the method described in section 2.2.7. Positive clones were confirmed by restriction double digestion with *Xho*I and *Nde*I enzymes.

Figure 2.5 shows the restriction double digest profile pET-22b(+) containing the CYP1A2 insert. Lane 2 -7 in figure 2.5, represent the products obtained after double digests, whereas lane 8 represents the negative control (empty pET-22b(+)). Lanes 2-7 show two bands which

correspond to the sizes of pET-22b(+) and CYP1A2 gene, 5 493 bp and 1 545 bp respectively. The results showed the expected digest pattern and no incomplete digestion or sample contamination was observed. The absence of insert band in negative control and appearance of insert only in positive clones indicated the success of ligation. However, double digest only confirmed ligation but not the integrity and orientation of the gene. It is for this reason that sequencing of positive clones was performed.

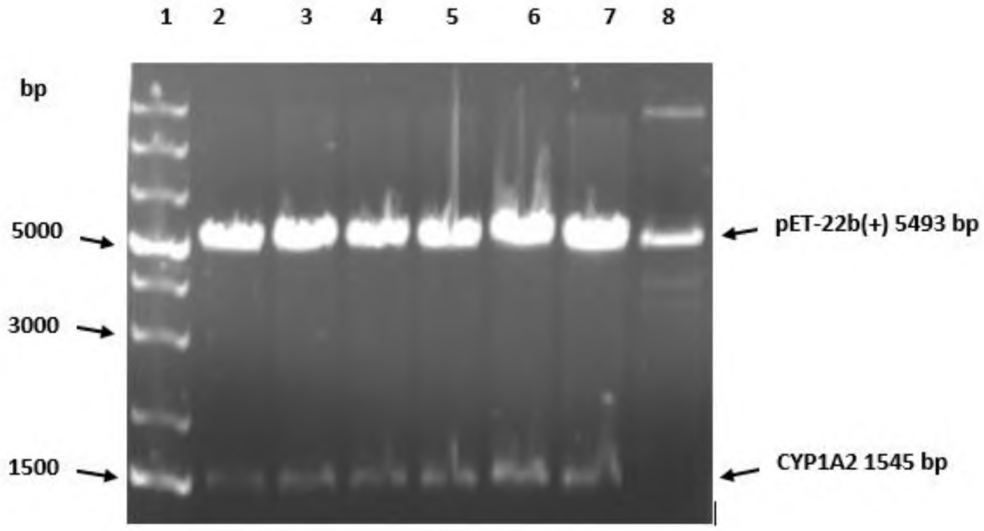


Figure 2.5: Confirmation of ligation of CYP1A2 gene by double digestion. The pET-22b(+) plasmid was digested with two restriction enzymes, *Xho*I and *Nde*I and analyzed by agarose gel electrophoresis using 0.8 % (w/v) gel in 1 x TAE buffer. Lane 1 represents the DNA ladder. Lane 2 – 7 represent pET-22b(+) plasmid extracted from six different colonies and digested with *Xho*I and *Nde*I restriction enzymes. The two digestion product are pET-22(+) (5 493 bp) and CYP1A2 (1 545 bp). Lane 8 represents the negative control, emplty pET-22b(+) plasmid digested with *Xho*I and *Nde*I.

2.3.3 Sequencing

Although ligation yielded satisfactory results, there was still a need to verify the outcome by sequencing the gene. This was done as outlined in section 2.2.8, and the results obtained are shown in figure 2.6. The sequence analysis was conducted by using FinchTV program. DNA sequence was used in translated BLAST (blastx) to retrieve equine CYP1A2-like protein sequence. The amino acid sequence showed 100% identity to the reference sequence found on BLAST. The results demonstrate that the insert was full length CYP1A2 gene ligated in the correct orientation. Thus, the expression plasmid with desired gene was produced for production of the recombinant equine CYP1A2 protein.

```
Query  1    MMLSQLSPFSATELLLASTIFCLVFWVVRAWQPQIPKGLKSPPGPWGWPFLGHVLTLGKN  180
            MMLSQLSPFSATELLLASTIFCLVFWVVRAWQPQIPKGLKSPPGPWGWPFLGHVLTLGKN
Sbjct  1    MMLSQLSPFSATELLLASTIFCLVFWVVRAWQPQIPKGLKSPPGPWGWPFLGHVLTLGKN  60

Query  181  PHLALSRLSQRYGDVMQIRIGSTPVLVLSGLDTIRQALVRQGDDFKGRPDLYSFTLITNG  360
            PHLALSRLSQRYGDVMQIRIGSTPVLVLSGLDTIRQALVRQGDDFKGRPDLYSFTLITNG
Sbjct  61   PHLALSRLSQRYGDVMQIRIGSTPVLVLSGLDTIRQALVRQGDDFKGRPDLYSFTLITNG  120

Query  361  QSMIFNPDSGPVWaarrrlaqnalnTFSIASDPASMSSCYLEEHVSKEAEALLSRLQKLM  540
            QSMIFNPDSGPVWAARRRLAQNALNTFSIASDPASMSSCYLEEHVSKEAEALLSRLQKLM
Sbjct  121  QSMIFNPDSGPVWAARRRLAQNALNTFSIASDPASMSSCYLEEHVSKEAEALLSRLQKLM  180

Query  541  SVAGRFDPSSQVVASVANVIGAMCFGQHFPHSSEEMISLLRSSHEFVQIASSGNPVDFFP  720
            SVAGRFDPSSQVVASVANVIGAMCFGQHFPHSSEEMISLLRSSHEFVQIASSGNPVDFFP
Sbjct  181  SVAGRFDPSSQVVASVANVIGAMCFGQHFPHSSEEMISLLRSSHEFVQIASSGNPVDFFP  240

Query  721  ILRYLPNPPLQRFKSFNQRFLRFLQKIIQEHYRDFDKNSIQDITGALF  885
            ILRYLPNPPLQRFKSFNQRFLRFLQKIIQEHYRDFDKNSIQDITGALF
Sbjct  241  ILRYLPNPPLQRFKSFNQRFLRFLQKIIQEHYRDFDKNSIQDITGALF  296
```

Figure 2.6: Confirmation of ligation of CYP1A2 gene by sequencing pET-22b(+) plasmid. The nucleotide sequence was used to search for the protein sequence of CYP1A2.

2.4 Conclusion

The equine CYP1A2 was found to have high sequence identity (83.69%) with the human homolog. Several amino acid residues which are believed to be significant for structural elements were found to be highly conserved across mammalian species that were selected. Due to high sequence identity and conservation, there was a desire for cloning and expression as this has never been explored for this protein. The CYP1A2 gene was synthesized into a cloning vector pUC57 as proven by restriction double digest profile. Ligation of the gene into an expression vector pET-22b(+) involved a number of steps and was successfully performed. Positive clones were confirmed by using both double digest and sequencing. Based on these results, it can be concluded that CYP1A2 gene was identified, synthesized and ligated into plasmid pET-22b(+) for production of equine recombinant CYP1A2 protein.

CHAPTER 3: HETEROLOGOUS EXPRESSION AND PURIFICATION OF EQUINE CYP1A2

3.1 Introduction

Information on heterologous expression and activity of enzymes from the CYP1A subfamily in veterinary species, especially horses, is currently limited. Knowledge of these enzymes is important as it can provide information that will aid in understanding equine drug metabolism. Furthermore, this information could be used to develop efficient ways of identifying metabolites of potential performance-enhancing drugs.

Recombinant systems have commonly been used to study CYP enzymes, particularly in humans. Bacterial expression systems have frequently been used for characterization of CYPs because of many advantages which include simplicity, low cost and rapid production of recombinant proteins (Gillam, 1998). For example, the *E. coli* expression pET vector system under the control of T7 promoter is one of the best systems as the recombinant protein can represent 50% of the total cell protein when expression is successful (Baneyx, 1999; Graumann and Premstaller, 2006). It also includes a six histidine residue (His-tag) which can be located on either the C-terminus or N-terminus of the target protein; this facilitates purification with immobilized nickel ions. In addition, the pET system contains the gene 10 5' leader, which promotes highly efficient translation. These elements make pET an ideal system for production of recombinant proteins.

High level production of recombinant proteins, especially large and membrane proteins, has limitations; these include expression of non-active insoluble protein in inclusion bodies. A number of strategies that facilitate generation of active and soluble recombinant CYP proteins have been developed. These involve introduction of amino acid modifications at the N-terminus,

truncation of the hydrophobic portion of the N-terminus, which is reported to associate with membranes and expression of CYP in the presence of molecular chaperones known to assist in folding of *E. coli* protein (Yun *et al.*, 2006; Sandhu *et al.*, 1994; Inoue *et al.*, 2000). Alternatively, CYPs can be targeted to bacterial membranes by means of N-terminus fusion to a bacterial signal peptide (Pritchard *et al.*, 1997). Use of lower temperatures (20-28°C) decreases the rate of expression to allow correct folding and incorporation of the heme. Many studies make use of these strategies and CYPs have been shown to maintain their functional activities.

Human CYP1A2 has been functionally expressed in bacteria without the N-terminus modification (Kim *et al.*, 2008). The full length mammalian CYPs expressed in this way is usually associated with the bacterial membranes. CYPs are often obtained by subcellular fractionation and the membrane fraction containing targeted protein is solubilized using low concentrations of nonionic detergents such as Triton X-100, zwitterionic detergents (CHAPS) or a combination of both. It has been reported that even if CYP is truncated, only about 47-57% of the protein is found in the cytosolic fraction and use of the solubilizing detergents increases this percentage (Mast *et al.*, 2004). Purification of solubilized material is normally carried out with chromatographic techniques and recovery of CYP protein from these steps is generally modest for characterization.

One of the important goals for bacterial expression of CYPs is to characterize these enzymes by determining their enzymatic properties. However, bacteria lack the necessary NADPH cytochrome P450 reductase (CPR) to support CYP activities, although a soluble NADPH-flavodoxin reductase from *E. coli* has been reported to support low levels of CYP activities (Jenkins and Waterman, 1994). Reconstitution of CYP activity is achieved by addition of CPR

purified from other animal sources. Other studies co-expressed CYP and CPR, in which case purification was not necessary (Iwata *et al.*, 1998).

The probe substrates are used to assess the metabolic activities with high pressure liquid chromatography (HPLC) or radiochemical-based methods. However, improvements in technology have provided an alternative luminescent method with improved sensitivity and speed for measurement of CYP activities. These P450-Glo assays provide luminometric substrates which are derivatives of D-luciferin and are converted into the active form (D-luciferin) by CYP enzymes to produce luminescence (Cali *et al.*, 2006). The amount of luminescence produced is directly proportional to CYP activity.

This study intended to develop a recombinant equine CYP1A2 system. To accomplish this, equine CYP1A2 was over-expressed in *E. coli*, solubilized, purified using nickel affinity chromatography and tested for activity using a P450-Glo assay kit. NADPH-cytochrome P450 reductase was also over-expressed, tested for activity and used for CYP1A2 activity reconstitution. A second equine enzyme CYP3A89 which has been expressed in *E coli* in an earlier study and shown to be catalytically active was used as a control for this study.

3.2 Methods and Materials

3.2.1 Chemicals and reagents

The δ-aminolevulinic acid, isopropyl-1-thio-$_D$-galacto-pyranoside (IPTG), Cytochrome *c* reductase assay reagents, Triton X-100, sarkosyl and SDS PAGE analysis reagents were purchased from Sigma-Aldrich. Lysozyme and protease inhibitor cocktail were purchased from Roche. The Pierce BCA protein assay kit was purchased from Thermo-Scientific. P450-Glo

CYP1A2 and CYP3A4 assay kits were purchased from Anatech, Tween 20 was purchased from Merck.

3.2.2 Expression of equine CYP1A2

3.2.2.1 Auto-induction

The pET-22b(+) plasmid containing CYP1A2 gene was transformed into *E. Coli* BL21 competent cells and plated on LA plates with 100 μg/ml ampicillin and incubated at 37°C overnight. Single transformation colonies were picked and inoculated into LB culture containing 100 μg/ml of ampicillin and incubated on a Labcon shaker (160 rpm) at 37°C overnight. This culture (10 ml) was used to inoculate 1 L of auto-induction media (1% tryptone, 0.5% yeast extract, 50 mM Na_2HPO_4, 50 mM KH_2PO_4, 25 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.5% glycerol, 0.05% glucose, 0.2% lactose) (Sturdier, 2005). The culture was dispensed into ten 500 ml conical flasks and incubated on a Labcon shaker (150 rpm) at 20°C for 36 hours. CYP1A2 heme precursors, $FeCl_3$ (1 mM) and δ-aminolevulinic acid (δ-ALA) (1 mM) were added after 4 hours of incubation.

3.2.2.2 IPTG

The pET-22b(+) plasmid containing CYP1A2 gene was transformed and inoculated as per section 3.2.2.1. The overnight culture was used to inoculate 100 ml of LB containing 100 μg/ml of ampicillin and incubated on Labcon shaker (180 rpm) at 37°C untill an OD_{600} of 0.6 – 0.7 was reached. Following the addition of IPTG (1 mM), $FeCl_3$ (1 mM) and δ-aminolevulinic acid (δ-ALA) (1 mM), cultures were incubated on a Labcon shaker (180 rpm) at 37°C for 4 hours.

3.2.2.3 Induction studies

Induction studies were conducted to confirm the over-expression of CYP1A2 protein and to determine the optimum incubation times. This was done on both CYP1A2 culture and negative control (culture that had been transformed with empty pET-22b(+)). For auto-induction, 1 ml samples were collected at 4 hour intervals over 36 hours. The OD_{600} of each sample was measured and cells were pelleted by centrifugation at 12 000 × *g* in a microfuge for 1 minute. The pellet was re-suspended in SDS loading buffer, boiled for 5 minutes and stored at – 20°C until used. The re-suspension volume was calculated according to the formula volume (ml) = $OD_{600}/6$. For the IPTG induction, the samples were treated as described for auto-induction. However, 1 ml samples were collected at 1 hour intervals over 4 hours.

3.2.3 Solubilization of recombinant CYP1A2 protein

The CYP1A2 protein was found in the insoluble fraction after cell lysis. In an attempt to solubilize CYP1A2, Triton X-100, Tween 20 and sarkosyl were used.

3.2.3.1 Triton X-100 and Tween 20

Following the expression of CYP1A2 protein (section 3.2.2), the culture was prepared for solubilization according to the method described by Fisher *et al.*, (1992), with modifications. The cells were chilled on ice for 30 minutes, harvested by centrifugation at 5 000 × *g* for 15 minutes in an Avanti® J-E centrifuge and a JA-14 rotor (Beckman Coulter). The pelleted cells were washed by re-suspending in one-fifth the volume of 10 mM potassium phosphate buffer (pH 7.4), containing 0.15 M NaCl and centrifuged at 5 000 × *g* for 15 minutes at 4°C. The supernate was discarded and the pellet was weighed. The cells were suspended with a volume of TSE buffer (pH 7.5, 75 mM Tris-HCl, 250 mM sucrose and 0.25 EDTA) two-fold the weight of cells, divided into 500 µl aliquots and stored at – 80°C until used. To solubilize the CYP1A2,

the *E. coli* cells were thawed, lyzed with lysozyme (1 mg/ml) at 37°C for 1 hour and sonicated for 30 seconds, 5 cycles at 60% output. The sonicated cells were diluted with 2 volumes of TSE buffer and centrifuged at 3 000 × *g* for 15 minutes at 4°C to remove unbroken cells. The supernate was centrifuged at 100 000 × *g* for 60 minutes at 4°C in Optima L-90K ultracentrifuge (Beckman-Coulter) and the pellet was suspended in TE buffer (pH 7.5, 50 mM Tris-HCl and 0.5 mM EDTA), divided into 500 µl aliquots and stored at – 80°C. The samples were thawed and 10% Triton X-100 or Tween 20 added to a final concentration of 1.5% and then incubated at 4°C for 1 hour stirring. The mixture was then centrifuged at 100 000 × *g* for 60 minutes at 4°C and the supernate was kept for protein characterization.

3.2.3.2 Sarkosyl solubilization

The cells were chilled on ice for 30 minutes, harvested by centrifugation at 5 000 × *g* for 15 minutes in an Avanti® J-E centrifuge and a JA-14 rotor (Beckman Coulter). Pelleted cells were washed by re-suspending in 100 mM potassium phosphate buffer (pH 7.4) (1 g cells/20 ml buffer) and centrifuged at 5 000 × *g* for 15 minutes at 4°C. The washed pellet was re-suspended in 100 mM potassium phosphate buffer (pH 7.4), lysozyme (1 mg/ml) and a protease inhibitor cocktail tablet added. Lysis was carried out at 37°C for 1 hour. The sample was frozen at – 80°C for a minimum of 4 hours and thawed at 4°C for solubilization. After thawing, the sample was centrifuged at 5 000 × *g* for 15 minutes at 4°C and the pellet was re-suspended in 100 mM potassium phosphate buffer (pH 7.4). A 20% solution of sarkosyl (*N*-laurylsarcosine) was added to a final concentration of 0.5% and incubated on ice for 1 hour. The sample was then centrifuged at 12 000 × *g* for 25 minutes at 4°C and the supernate was stored -20°C for purification.

3.2.4 Purification of recombinant CYP1A2 protein

3.2.4.1 Ni-NTA Superflow Cartridge

The solubilized sample was first filtered using 0.22 μm sterile filters. Purification was carried out at 4°C using nickel affinity chromatography. The 5 ml Ni-NTA Superflow Cartridge (QIAGEN®) was first equilibrated with 10 volumes of binding buffer NPI-10 (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8). The cleared lysate (20 ml) was applied to the cartridge using the same flow rate (5 ml/min) as in the previous step. The cartridge was washed using wash buffer NPI-20 (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8) to remove non-specifically bound proteins. His-tagged CYP1A2 was eluted into 5 ml fractions with elution buffer NPI-250 (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imadazole, pH 8). The 5 ml fractions were kept for desalting and characterization.

3.2.4.2 Desalting

The elution samples (5 ml fractions) with the protein of interest were concentrated to 2.5 ml by Vivaspin2 centrifugal concentrators MWCO 10 000 (Sartorius Stedim Biotech). Prior to desalting of the sample, PD-10 desalting columns (Sephadex G-25) (GE Healthcare) were equilibrated with 12 ml of 100 mM potassium phosphate buffer (pH 7.4). Sample (2.5 ml) was loaded and allowed to enter the packed bed completely. Fractions of 750 μl were eluted with 100 mM potassium phosphate buffer (pH 7.4). All steps were performed using the gravity protocol. The elution fractions were stored for SDS-PAGE analysis, protein concentration determination and enzyme activity.

3.2.5 Characterization of recombinant CYP1A2

3.2.5.1 Protein concentration

Protein concentration was determined using the bicinchoninic acid method (Smith *et al.*, 1985). The method is dependent on reduction of Cu^{+2} to Cu^{+1} by proteins in an alkaline medium, which shows a linear trend with increasing protein concentration over a broad working range (20-2 000 µg/ml). A standard curve was constructed using bovine serum albumin (BSA). Protein standards between 25–2 000 µg/ml were prepared and mixed with an assay working reagent as per manufacturer's instructions. The BSA standards and unknown samples were incubated at 37°C for 30 minutes. The absorbance values were measured at 562 nm using SynergyMx microplate reader (BioTek®).

3.2.5.2 SDS-PAGE analysis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was conducted using a mini PROTEAN® tetra system (BIO-RAD). The protein samples were analyzed with 10% resolving gel and 4% stacking gel using the procedure described by Laemmli (1970). Protein samples were mixed 1:1 with SDS loading buffer and incubated at 100°C for 10 minutes on a digital dry bath heating block (BIO-RAD) before loading. PageRuler™ prestained protein ladder (size range, 10–180 kDa; Thermo Scientific) was used as a molecular weight marker for protein size estimation. The electrophoresis was carried out at 100 V and bands on a gel were visualized by staining and de-staining using a method described by Fairbanks *et al.*, (1971). The procedure involves use of Coomassie Blue R-250, isopropanol, acetic acid and a series of heating and cooling steps in four different solutions. All the solutions used for SDS-PAGE analysis are listed in Appendix 4.

3.2.5.3 Activity determination

3.2.5.3 (a) NADPH-cytochrome P450 reductase activity

NADPH-cytochrome P450 reductase (CPR) is required for CYP activity. CPR was expressed and its activity evaluated before proceeding to CYP activity assay. Yeast CPR was expressed as described in section 3.2.2.2, the cells were lyzed with lyzosyme (1 mg/ml) and supernatant was used for activity assay. CPR activity was measured using a cytochrome *c* reductase assay with NADPH as an electron donor and cytochrome *c* (from equine heart) as electron acceptor. The assay was carried out at 30°C in a 96-well microtiter plate, the reaction constituents were 100 mM potassium phosphate buffer (pH 7.4), 0.45 mg/ml cytochrome *c*, 1 mM KCN and 15 µg/ml CPR protein in a total volume of 200 µl. The reaction was initiated by addition of NADPH to a final concentration of 0.12 mM. Increase in absorbance at 550 nm was monitored for 15 minutes and readings were taken at 1 minute intervals. The CPR activity was calculated using the equation:

$$\text{Units/ml} = \frac{\Delta A550/\min x\ dil\ x\ rxn\ vol}{21.1\ x\ enzvol}$$

Dil represents the dilution factor of original enzyme sample, enzvol is volume of enzyme sample, rxn vol is reaction volume and 21.1 $mM^{-1}.cm^{-1}$ is the extinction coefficient (ε^{mM}) for reduced cytochrome *c*.

3.2.5.3 (b) CYP assay

The enzymatic activity of CYP1A2 was determined using a P450-Glo CYP1A2 assay kit. P450-Glo Assay was performed according to the manufacturer's protocol. Briefly, the 2× CYP reaction mixture made up of CYP preparation (15 µg/ml), CPR preparation (15 µg/ml), 200 µM

luciferin-ME and 200 mM potassium buffer (pH 7.5) was pre-warmed at 37°C for 10 minutes. A 2× NADPH regenerating system containing 2.6 mM $NADP^+$, 6.6 mM glucose-6-phosphate, 6.6 mM $MgCl_2$ and 0.8 U/ml glucose-6-phosphate dehydrogenase was also pre-warmed at 37°C. The reaction was initiated by addition of a 2× NADPH regenerating system (25 µl) to the CYP reaction mixture to make a final volume of 50 µl in a white 96-well plate (Greiner Bio-one). The reaction mixture was incubated at 37°C for 30 minutes. Following the incubation period, 50 µl of luciferin detection reagent was added and the plate was kept at room temperature for 20 minutes before measuring luminescence using a SynergyMx microplate reader (BioTek®). For calculation of the amount of D-luciferin produced, a D-luciferin standard curve was constructed between 0.016 and 2 µM. The negative control reaction mixture contained a lysate from cells that were transformed with empty pET-22b(+).

3.2.6 Western blot

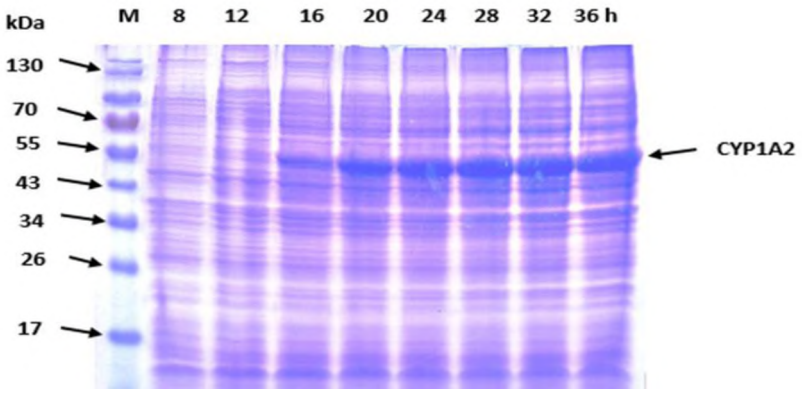
CYP protein was purified and analyzed on SDS-PAGE gel (section 3.2.5.2). The proteins were transferred onto a nitrocellulose membrane in a transfer buffer consisting of 13 mM Tris-HCl, 100 mM glycine and 20% (v/v) methanol using transblot at 100 V for 1 hour. All the materials used for transfer were pre-immersed in transfer buffer for 15 minutes. The proteins on the nitrocellulose membrane were visualized by staining with Ponceau and subsequently destained in deionised water and Tris-buffered saline-Tween (TBST; 50 mM Tris, pH 7.5, 150 mM NaCl and 1 % Tween). The membrane was kept in 2% blocking solution in TBST for 1 hour and incubated with His-probe rabbit polyclonal (1:1 000 v/v) overnight at 4°C. The membrane was washed three times in TBST for 10 minutes and incubated with anti-rabbit IgG conjugated to horseradish peroxide (HRP; 1:1 000 v/v) for 1 hour at room temperature. The membrane was washed as

previously described. Visualization was conducted with a BM chemiluminescence western blotting kit using ChemiDoc^{TM}XRS+ (Bio Rad).

3.3 Results and discussion

3.3.1 Expression of recombinant CYP1A2

The equine CYP1A2 was heterologously expressed in a bacterial system using auto-induction and IPTG-induced expression. Figure 3.1(a) shows the auto-induction SDS-PAGE analysis gel which demonstrates the over-expression of CYP1A2 over a period of 36 hours. The band representing over-expressed protein matches the predicted CYP1A2 size (≈55 kDa), which corresponds to unmodified human CYP1A2 expressed by Kim *et al.*, (2008). In a study by Tyden *et al.*, (2014), immune techniques were used for analysis of protein expression and cellular localization of CYPs in equine intestine and liver. They reported CYP sizes of 56 -58 kDa, and our results (figure 3.1(a)) show a protein of similar size. Figure 3.1(b) shows the negative control culture (empty pET22) which was subjected to the same culture conditions. No band corresponding to CYP1A2 was observed over 36 hours. The results obtained suggest that the auto-induction method can be used for expression of this enzyme.



A

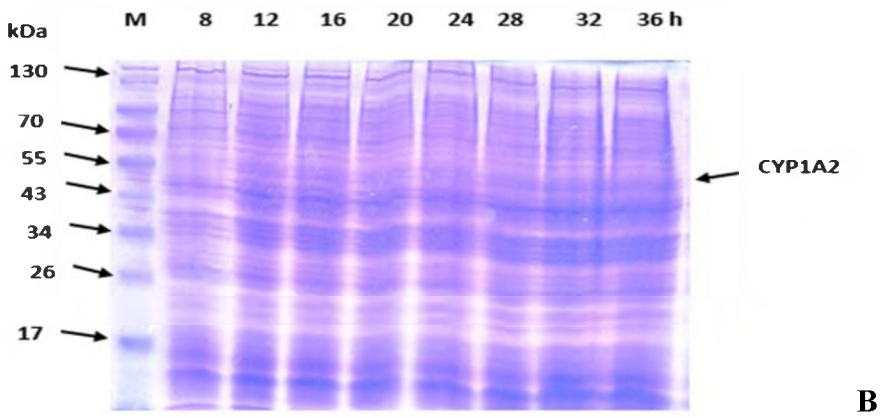


Figure 3.1: SDS-PAGE analysis of expression of CYP1A2. (**a**) Shows the SDS-PAGE gel of the auto-induction culture. Lane M: Protein molecular marker. Lanes labeled 8–36 h are samples collected from culture expressing CYP1A2 over 36 hours (4 hour intervals). (**b**) Shows the SDS-PAGE gel of negative control, a culture that has pET-22b(+) plasmid lacking the CYP1A2 gene. Lane M: protein molecular marker. Lanes labeled 8 – 36 h are samples collected from culture over 36 hours (4 hour intervals).

IPTG-induced expression, which is the most common method was also used for expression of equine CYP1A2. This approach has a few advantages, which include production of proteins over a very short time, and IPTG is not metabolized by cells, therefore remains constant throughout induction. Figure 3.2 show the SDS-PAGE analysis gel for IPTG-induced expression. Lane 1 represents the un-induced sample, lane 6 is an empty pET-22b(+), both with no protein observed at 55 kDa. Lane 2-5 shows induction samples over 4 hours with bands corresponding to CYP1A2. There was no significant difference in expression level between IPTG-induced expression and auto-induction, with both methods yielding expression levels that were satisfactory. The results indicate that the pET system under the control of T7 promoter is efficient for production of recombinant CYP1A2.

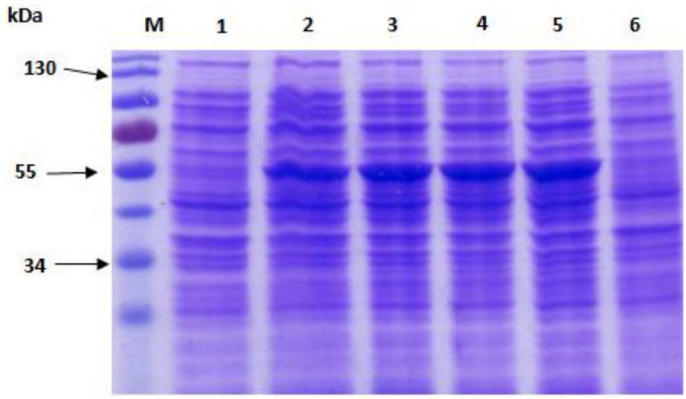


Figure 3.2: SDS-PAGE analysis of IPTG induced expression of CYP1A2. Lane M: Protein molecular marker. Lane 1: uninduced sample. Lane 2-5: Samples collected at 1 hour interval after IPTG addition. Lane 6: Negative control (empty expression vector) collected after 4 hours of induction.

3.3.2 Solubilization of CYP1A2

Although *E. coli* expression systems are commonly used, the expression of mammalian CYP enzymes is often hampered by difficulties such as the production of insoluble non-functional protein. Common strategies for production of functional and soluble protein include truncation of the hydrophobic region at the N-terminus, modification of N-terminal sequence and use of lower temperatures (20-28°C) for expression. A full-length functional human CYP1A2 has been expressed in *E. coli* without conventional modification of the N-terminal sequence (Kim *et al.*, 2008), and was therefore attempted in this study. However, the equine CYP1A2 was found to be insoluble and various strategies were used in attempts to solubilize the protein.

The first approach employed was modification of expression conditions including temperatures, concentration of IPTG and heme precursors, $FeCl_3$ and δ-ALA. One of the causes of insolubility of protein is formation of unfolded protein aggregates known as inclusion bodies. Slowing the rate of expression by reducing incubation temperatures can lead to proper folding which in turn results in soluble protein (Rosano and Ceccarelli, 2014). For auto-induction, an expression

temperature of 20°C was used and a post-induction temperature of 28°C was selected for IPTG-induced expression. However, temperature reduction did not yield the desired results as the expressed protein remained insoluble. Concentration of IPTG and heme precursors, $FeCl_3$ and δ-ALA were reduced from 1 mM to 0.5 mM, 0.2 mM and 0.1 mM at different post induction temperatures. The protein remained in the insoluble fraction. To continue with purification and characterization of the protein, different solubilizing detergents were then investigated.

3.3.3 Solubilization with Triton X-100 and Tween 20

Triton X-100 and Tween 20 are mild non-ionic detergents with hydrophilic head groups. They usually disrupt protein-lipid and lipid-lipid interactions which assist in solubilizing proteins. These detergents were used only up to 1.5%, because higher concentrations interfere with binding of His-tagged protein to the nickel affinity resin. Figure 3.3 shows SDS-PAGE analysis gel for solubilization of CYP1A2 using Tween 20. Lane 4 and 5 represent the supernatant material after addition of 1.5% of Tween 20, while lane 3 shows the pellet. Our results indicated that most of the CYP1A2 protein remained in the insoluble fraction, while only a small quantity (lanes 4 and 5) was solubilized. A low concentration (0.1%) of Triton X-100 has been reported to solubilize a human CYP1A2 from bacterial membranes (Fisher *et al.*, 1992). A second non-ionic detergent, Triton X-100, was then used to a final concentration of 1.5%. Use of Triton X-100 showed no difference, as large amount of protein was still found in the insoluble fraction. Failure to obtain soluble protein with non-ionic detergents suggests that the expressed protein accumulates in inclusion bodies. Non-ionic detergents are most commonly used for isolating membrane proteins due to their ability to disrupt lipid-lipid and protein-lipid interactions but not protein-protein associations. Ionic detergents are known to completely solubilize proteins in insoluble aggregates, and therefore sarkosyl was investigated.

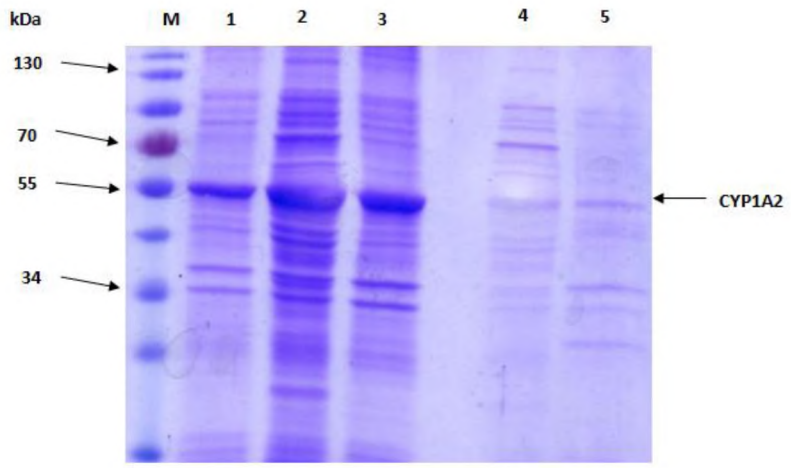


Figure 3.3: SDS-PAGE analysis of CYP1A2 protein solubilized with 1.5% Tween 20. Bacterial cells were lysed using lyzosome and sonication. Lane M: Protein molecular marker. Lane 1: Pellet after centrifugation at low speed. Lane 2: Pellet before addition of 1.5% Tween 20. Lane 3: Pellet after addition of Tween 20. Lane 4-5: Supernatant after addition of 1.5% Tween 20.

3.3.4 Solubilization using sarkosyl

Low concentrations of sarkosyl are widely used for solubilization of proteins found in inclusion bodies after bacterial expression (Frankel *et al.*, 1991; Tao *et al.*, 2010). As previously observed, the expression yielded an intractably insoluble protein which was observed in the pellet. Sarkosyl was used to a final concentration of 0.5% to solubilize the protein. Figure 3.4 shows CYP1A2 protein (55 kDa) in the soluble fraction (lane 4) after addition of buffer containing sarkosyl. The protein was not observed in the supernatant fraction with no sarkosyl added (lane 3) and in negative control samples (lane 1 and 2). The protein was previously shown to be insoluble after addition of non-ionic detergents, 1.5% Triton X-100 or Tween 20. Therefore sarkosyl was found to be an effective ionic detergent which allows large amounts of CYP1A2 to be solubilized (figure 3.4).

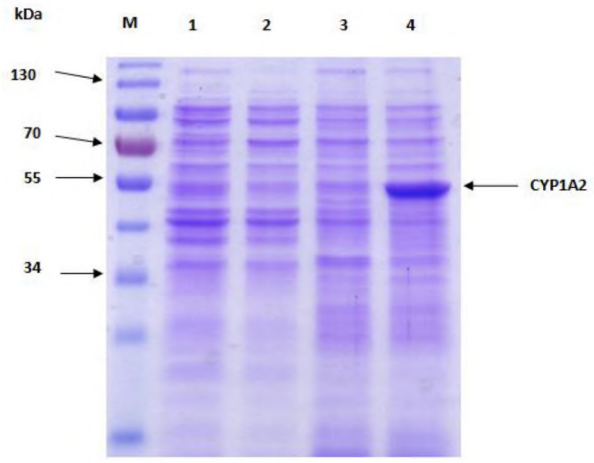


Figure 3.4: SDS-PAGE analysis of CYP1A2 protein solubilized with 0.5% sarkosyl. Bacterial cells were lysed using lyzosome, sonication and freeze-thawing. After centrifugion at low speed, the pellet was resuspended in a buffer containing 0.5% sarkosyl. Lane M: Protein molecular marker. Lane 1: Supernate of negative control prior addition of sarkosyl. Lane 2: Supernate of samples with CYP1A2 prior addition of sarkosyl. Lane 3: Supernate of negative control after addition of sarkosyl. Lane 4: Supernatant of samples with CYP1A2 after addition of sarkosyl.

3.3.5 Purification of CYP1A2

3.3.5.1 Nickel affinity chromatography

The sarkosyl-solubilized protein was purified by nickel affinity chromatography and analyzed by SDS-PAGE analysis as described in section 3.2.5.2. The recombinant CYP1A2 enzyme could be purified with this method, and a protein band of approximately 55 kDa was observed (figure 3.5). As seen in figure 3.5, lane 1 shows the crude material and lane 2 represents the flow-through. Unfortunately, a significance amount of protein is observed in the flow-through, indicating that most of the protein did not bind to the column. A small amount of protein was also observed in the wash sample (lane 3). The protein of interest was then eluted with 250 mM of imidazole, and the samples are shown in lanes 4, 5 and 6 of figure 3.5. Lane 5 shows a pure band which corresponds to the size of CYP1A2 protein. The band observed shows small amount

of CYP1A2 compared to that in the flow-through. This observation suggests that the majority of the protein was lost.

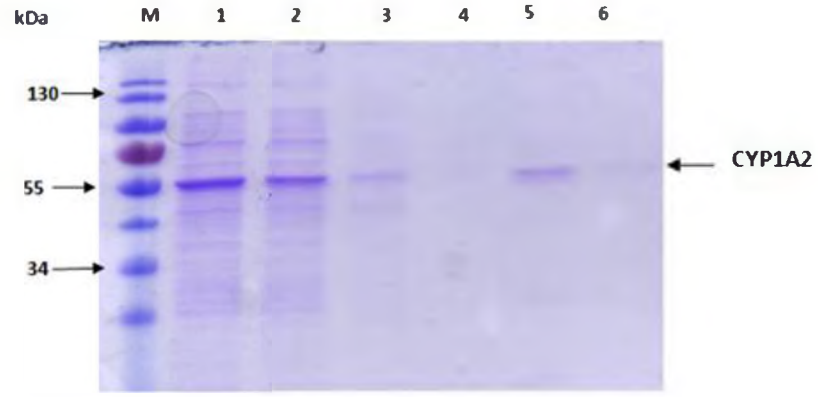


Figure 3.5: SDS-PAGE analysis of His-Tag purification of CYP1A2 protein solubilized with 0.5% sarkosyl. Lane M: Protein molecular marker. Lane 1: cleared lysate. Lane 2: Flow-through protein. Lane 3: Wash sample. Lane 4 -6: 5 ml elution samples.

The manufacturer's protocol reports that more than 0.3% of sarkosyl may interfere with binding of His-tagged protein to the column. To minimize the amount of protein found in flow-through, dialysis over 16 hours was used to reduce the concentration of sarkosyl before purification, but the results remained the same. This observation suggests that the protein was degraded or the His-tag is not fully accessible (partially exposed) to the nickel ions due to nature of protein folding. No reports are available on His-tag purification of CYP1A2 protein and positioning of the tag when the protein is folded. It could be that solubilization was incomplete and the protein observed in flow-through might be small protein aggregates. In an attempt to prevent elution in the wash step, the imidazole concentration in the wash buffer was reduced from 50 to 35 and 20 mM. However, no difference was observed after decreasing imidazole concentration. The amount of protein eluted was used for characterization studies.

3.3.5.2 Western blot analysis of CYP1A2

The purified CYP1A2 from nickel affinity (figure 3.5, lane 5) was further confirmed by western blot analysis. The western blot is based on detection of His-tag residues on the target protein by His-probe rabbit polyclonal antibody. A western blot gel image is shown in figure 3.6. Lane 1 represents purified CYP1A2 and lane 2 is a negative control sample. A protein was observed which was consistent with the size of CYP1A2 (55 kDa). No protein bands were observed for the negative control, as was expected. This confirmed that the purified protein is recombinant His-tagged CYP1A2 protein. The purified CYP1A2 was then concentrated, desalted as described in section 3.2.4.2, and used for activity assessment.

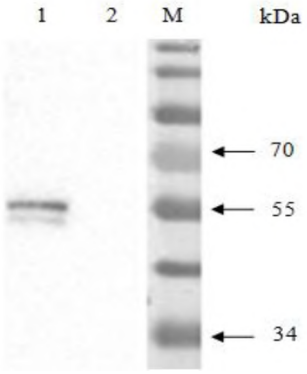


Figure 3.6: Western blot gel image of His-tagged CYP1A2 protein. Lane M represents the protein molecular marker. Lane 1: Purified CYP1A2 protein. Lane 2: Negative control (sample of empty pET22 plasmid).

3.3.6 NADPH-cytochrome P450 reductase activity assay

CPR transfers electrons from NADPH to the heme of CYP, and its activity was assessed by monitoring reduction of cytochrome *c* in the presence of substrate NADPH and CPR preparation. Figure 3.7 illustrates a progress curve over 15 minutes. The negative control reaction was monitored over the same period and the absorbance values obtained were subtracted from the

reaction values. The linear part of the progress curve was used to calculate the activity as described in section 3.2.5.3 (a). CPR activity was 0.13 U/ml, where one unit reduces 1 µmole of oxidized cytochrome c in the presence of 100 µM NADPH per minute. The results obtained show that CPR was active and should support CYP mediated reactions.

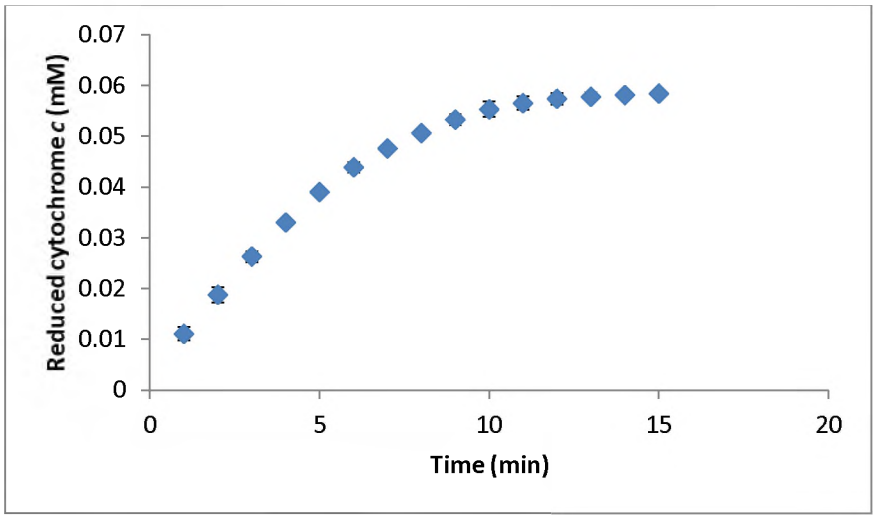


Figure 3.7: Determination of CPR activity by monitoring cytochrome c reduction at 550 nm. The curve was plotted using reduced cytochrome *c* concentration (difference between negative control reaction and reaction with substrate) and time (minutes).

3.3.7 CYP1A2 activity assay

The concentration of D-luciferin produced by CYPs in the P450-Glo assay was determined by comparing luminescence from the CYP reaction to luminescence from the D-luciferin standard curve (figure 3.8). The luminescence from all samples was less than that of the highest standard (2 µM D-luciferin).

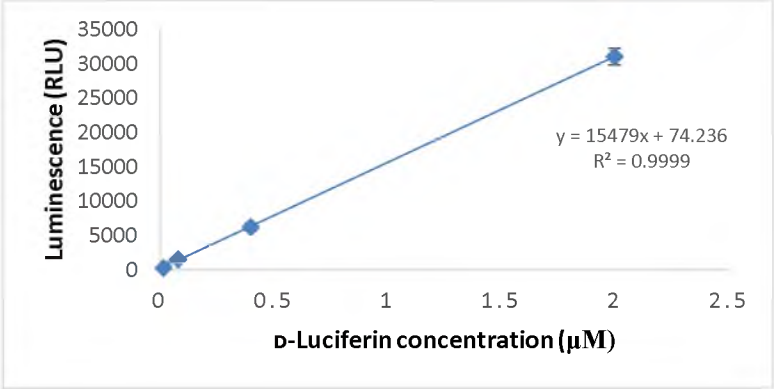


Figure 3.8: A D-Luciferin standard curve performed in parallel with CYP activity assay. A standard curve was constructed using a concentration range from 0.016 to 2.0 µM of D-luciferin. Standard deviation was observed (from triplicates).

The CYP1A2 activity assays were performed using solubilized protein and the luminescence signal produced was converted to D-luciferin using the standard curve (figure 3.8). The results indicated that solubilized CYP1A2 was not active towards the selective substrate luciferin-ME. To investigate factors that might be responsible for the effects on activity, several troubleshooting methods were performed. The ratio of CYP to CPR was investigated by increasing the amount of each component (1:1 to 1:6), but the results remained the same. It has been shown that establishing optimal incubation conditions is important for proper reconstitution of recombinant CYP (DiMaio Knych *et al.*, 2009). A range of protein concentrations from 10 to 80 µg/ml proved also to be ineffective. Incubation conditions such as temperature (25, 28, 30, 35 and 38°C) and incubation time (15, 20, 30, 40, 60, 70 and 90 minutes) were also studied. The ability of the NADPH-generating system to produce NADPH was assessed and found to be functional. The NADPH-generating system was substituted with NADPH, which was used as a substrate in the CPR assay but did not affect activity.

Despite numerous attempts to obtain activity, the CYP1A2 remained inactive. Possible causes of this situation include production of insoluble CYP1A2 protein, denaturation by sarkosyl treatment and failure to fold correctly. To attain high level expression of the equine CYP1A2, the pET system was selected, but led to the appearance of target protein in the insoluble fraction. Insoluble protein aggregates usually form because of high-level expression and as a result are incorrectly or partially folded (Ventura and Villaverde, 2006). Partially or extensively misfolded proteins are often devoid of biological activity.

When this study commenced, a full-length human CYP1A2 had been expressed in *E. coli* where it localized in the membranes (Kim *et al.*, 2008). In our study, the failure to solubilize CYP1A2 using non-ionic detergents suggests that the protein was not localized in the membrane but expressed as insoluble protein aggregates. This is because non-ionic detergents such Triton X-100 are usually used for membrane protein solubilization. They are also used in preparation of inclusion bodies (insoluble protein aggregates) to remove impurities such as membrane content, but do not solubilize insoluble protein aggregates (Khow and Suntrarachun, 2012).

To purify and characterize the CYP1A2 protein, the insoluble aggregates were solubilized with 0.5% sarkosyl. Sarkosyl is an ionic detergent but binds less strongly to proteins and dissociates more easily. It has been reported that diluting sarkosyl-solubilized protein to about 0.01% leads to dissociation of the detergent from protein and the protein refolds (Burgess, 2009). Chromatographic methods can also achieve the removal of detergent after solubilization. In the case of CYP1A2, no biological activity was retained after such treatment. This might indicate that the protein was partially or extensively misfolded. Protein activity requires folding into precise native conformation. However, it is always challenging to find the right conditions under which the solubilized protein can be refolded efficiently.

Expression of CYP1A2 as insoluble aggregates seems to be the main reason for lack of activity. Attempts to prevent formation of insoluble aggregates by modifying expression conditions did not yield the required results. One of the possible future solutions is the use of a different expression system. Several CYPs have been functionally expressed in *E. coli* using pCWOri vector which has the tac promoter (Zelasko *et al.*, 2013). Expression of CYPs using the pET system has been accomplished by co-expressing with molecular chaperones that assist in protein folding or using bacterial signal peptides (Ichinose and Wariishi, 2012; Pritchard *et al.*, 1997). This suggests that pCWOri vector may be used for expression of equine CYPs in *E. coli*. Additional elements which assist in protein folding or localization might be required when using pET, although this system leads to high-level production of protein. These will be further discussed in chapter 4.

3.3.8 Use of equine CYP3A89 as a positive control

Due to difficulties in solubility, purification and lack of activity on the expressed equine CYP1A2, it was essential to evaluate the chosen methods in order to gain deeper understanding of the target protein. This was performed using CYP3A89, which had shown activity in an earlier study in our laboratory, and had the potentially be used as a control. This enzyme was expressed using the IPTG method described in section 3.2.2.2, and analyzed using a SDS-PAGE gel shown in figure 3.6. The findings indicated a band (≈55 kDa) corresponding to the size of recombinant CYP3A89. These results are consistent with the size of equine CYP3A members expressed in baculovirus and mammalian systems (DiMaio Knych *et al.*, 2010; Dettwiler *et al.*, 2014). The absence of the same band in negative control samples (lanes 1and 6) indicated that the expressed protein was equine CYP3A89. High levels of protein expression were achieved and these were comparable to the results obtained for CYP1A2. To continue with purification

and activity determination, bacterial cells were lysed with lysozyme and sonication but the protein of interest was again found in the insoluble fraction. The protein was then solubilized with 0.5% sarkosyl as described for CYP1A2 and purified.

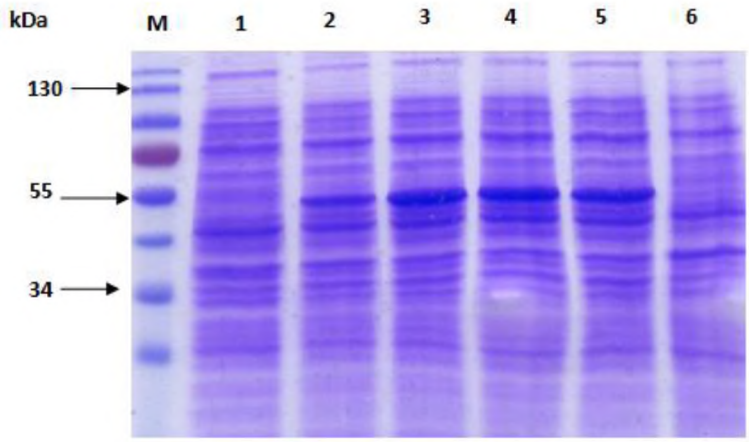


Figure 3.9: SDS-PAGE analysis of IPTG induced expression of CYP3A89. Lane M: Protein molecular marker. Lane 1: uninduced sample. Lane 2-5: Samples collected at I hour intervals after IPTG addition. Lane 6: Negative control (empty expression vector).

3.3.9 Purification of recombinant CYP3A89

Figure 3.10 shows the SDS-PAGE analysis of purification procedure. The flow-through sample (lane 2) contained part of the targeted protein with several impurities. A low concentration of CYP3A89 was found in the wash sample (lane 3). Considerable amounts of protein, especially in elution fraction 2 (lane 5) were seen after elution. This was the major difference between the purification of CYP1A2 and CYP3A89. These findings suggest that the low concentration of sarkosyl used is compatible with Ni-NTA superflow cartridge and the inability of CYP1A2 to bind efficiently to the cartridge may be due to factors such as protein degradation. The purification of CYP3A89 was found to be satisfactory. The purified CYP389 sample in figure 3.10, lane 5, was concentrated and desalted as described in section 3.2.4.2.

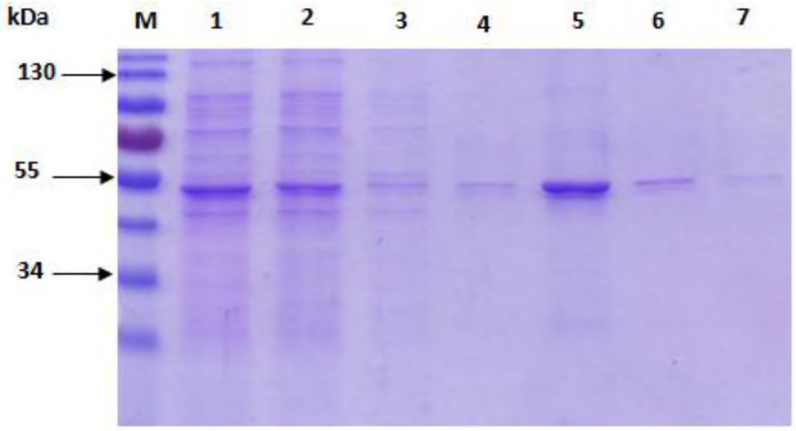


Figure 3.10: SDS-PAGE analysis of His-Tag purification of CYP3A89 protein solubilized with 0.5% sarkosyl. Lane M: Protein molecular marker. Lane 1: cleared lysate. Lane 2: Flow-through protein. Lane 3: Wash sample. Lanes 4 -7: 5 ml elution samples.

3.3.10 Desalting purified CYP3A89

Desalting was conducted mainly for two purposes: to remove high concentrations of imidazole from purified sample (figure 3.10, lane 5), and to replace the elution buffer (250 mM imidazole) with 100 mM potassium phosphate assay buffer (pH 7.4). The desalting was performed as described in section 3.2.4.2 and several fractions were obtained and analyzed with SDS-PAGE gel (figure 3.11). The findings indicate that the majority of protein eluted in fractions 2 and 3 (lanes 2 and 3). Some impurities at lower concentration were also observed, possibly histidine-rich proteins derived from the expressing bacterial host. It has been reported that host proteins with histidine-rich clusters can bind non-specifically to nickel ions (Muller, 1998). The fractions containing the protein in question (lanes 2 and 3) were used for activity and protein concentration determination.

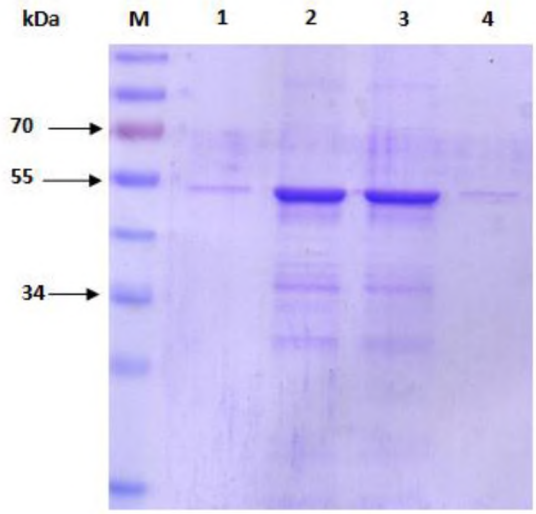


Figure 3.11: SDS-PAGE analysis of CYP3A89 protein eluted from desalting column. Elution fractions from nickel affinity chromatography were concentrated and loaded onto a desalting column. Lane M: Protein molecular marker. Lanes 1-4 contain elution samples with the protein.

3.3.11 Western blot analysis of CYP3A89

Purified CYP3A89 from nickel affinity was confirmed by western blot analysis. A western blot gel image is shown in figure 3.12. A protein band was observed which was consistent with the size of CYP3A89 (kDa). However, there was a contamination band at approximately 100 kDa. This band might be a result of histidine rich proteins from the expressing bacterial host which were not successfully removed with chromatographic methods. A few non-specific bands were also present in the desalting gel (figure 3.11), but at lower concentration (demonstrated by band intensity). Samples taken from bacterial culture with empty pET-22b(+) (lane 2) had no band corresponding to CYP3A89. These results confirmed that the purified protein was recombinant His-tagged CYP3A89.

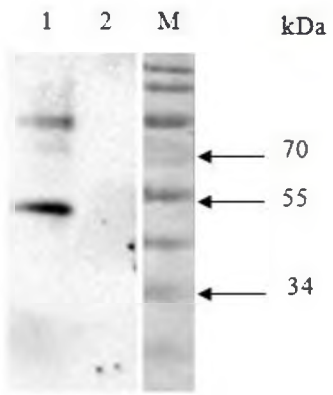


Figure 3.12: Western blot gel image of His-tagged CYP3A89 protein. Lane labeled M represents the protein molecular marker. Lane 1: Purified CYP3A89 protein. Lane 2: negative control (sample of empty pET-22b(+) plasmid).

3.3.12 CYP3A89 activity assay

The CYP3A89 which was expressed, solubilized and purified was assayed for activity. The initial incubation conditions were the same as described for CYP1A2 in section 3.2.5.3 (b) except for the selective substrate, which was luciferin-IPA (3 μM). The P450-Glo assay results indicated that there was no activity detected for CYP3A89. There was no apparent difference between the negative control sample and samples containing CYP3A89 enzyme (crude and purified). This pattern was very similar to that observed for CYP1A2 enzyme. Similar troubleshooting strategies were employed, and there was no difference in enzyme activity. All attempts at producing the desired results were unsuccessful.

From the results, it is clear that both enzymes are inactive, and initial attempts to resolve this did not yield positive results. The protein insolubility from over expression studies seems to be the major cause for production of non-functional protein. Insoluble protein is usually found in inclusion bodies formed by non-functional misfolded protein aggregates. It is also observed that solubilization and purification of these proteins from insoluble aggregates did not result in proper

refolding. It seemed unlikely that sarkosyl could have also contributed by denaturing protein as it is categorized as a mild ionic detergent. Other proteins have been shown to maintain functional activity after such treatment (Frankel *et al.*, 1991).

3.4 Conclusion

In the present study, equine CYP1A2 was heterologously expressed, purified and characterized. The expression of this CYP1A2 in *E. coli* was achieved by using both auto-induction and IPTG induction and confirmed by SDS-PAGE analysis. The high expression levels resulted in formation of insoluble protein aggregates. To prevent formation of insoluble aggregates, expression conditions were modified but the protein remained insoluble. Non-ionic detergents (Triton X-100 and Tween 20) were therefore used to solubilize the protein after cell lysis, but CYP1A2 remained in the insoluble fraction. Addition of 0.5% sarkosyl did solubilized the protein and the subsequent purification step was conducted with nickel affinity chromatography. Even though sarkosyl solubilized the protein, binding to the nickel column was inefficient as a large amount was found in the flow-through. The small amount of CYP1A2 that did bind was eluted with 250 mM imidazole. Yeast NADPH-cytochrome P450 reductase (CPR) was expressed in *E. coli* and its activity assessed before reconstitution of CYP1A2 activity. Cytochrome *c* reduction assay showed that CPR had an activity of 0.13 U/ml. Activity assessment using a P450-Glo CYP1A2 assay kit on CYP1A2 demonstrated that the protein was not functional.

To investigate factors responsible for production of non-functional CYP1A2, a second equine enzyme, CYP3A89, was expressed in *E. coli* and used as a control. This protein was also insoluble and 0.5% sarkosyl was used for solubilization. CYP3A89 had a better purification profile than CYP1A2, as the protein bound efficiently to the nickel column. However, there was no activity detected by the P450-Glo CYP3A4 assay kit despite several attempts to produce an

active protein. Production of insoluble protein aggregates is perhaps the major challenge hampering activity studies of equine CYP1A2 and CYP3A89. Under the experimental conditions of this study, the expressed equine CYP1A2 and CYP3A89 were inactive. Further investigation on high-expression levels of active equine CYP in *E. coli* is required.

CHAPTER 4: GENERAL DISCUSSION, CONCLUSION AND FUTURE WORK

4.1 General discussion and conclusion

In vitro systems provide useful tools for studying equine drug metabolism from both regulatory and therapeutic perspectives. These systems include liver microsomes, hepatocytes, lung preparations and recombinant CYP enzyme systems. Recombinant CYP enzyme systems are a valuable tool as they provide faster and more efficient drug metabolism studies. These include mammalian, baculovirus-insect and bacterial expression systems, the latter being particularly useful for expressing high levels of recombinant protein. Limited equine CYPs have been recombinantly expressed in mammalian and baculovirus-insect systems; these are CYP2D50, CYP2C92, CYP2B6 and CYP3A members (DiMaio Knych and Stanley, 2008; DiMaio Knych *et al.*, 2009; Peters *et al.*, 2013; DiMaio Knych *et al.*, 2010; Schmitz *et al.*, 2014; Dettwiler *et al.*, 2014). However, there are no reports on recombinant expression of equine CYP1A2 enzyme. In the present study, for the first time, equine CYP1A2 protein was bioinformatically identified, cloned into an expression vector and expressed using a bacterial expression system.

Human CYP1A2 has previously been expressed in an *E. coli* system and used for analysis of catalytic activities (Fisher *et al.*, 1992; Sandhu *et al.*, 1994). For those studies, functional expression was achieved because of modifications which included replacement of the native N-terminal region with sequence of bovine CYP17A1. However, Kim *et al.*, (2008) reported functional expression of human CYP1A2 without conventional N-terminal modification in *E. coli*. The catalytic activities of this enzyme were compared to modified CYP1A2, and showed very similar activities toward most substrates.

For our study, the initial step was the retrieval of the sequence for equine CYP1A2 homolog using the human CYP1A2 sequence as a template. From multiple sequence alignment, it was found that the equine CYP1A2 shared substantial sequence identity with CYP1A2 protein from human, pig, mouse and rat. A high sequence identity (83.69%) with human CYP1A2 and the presence of conserved motifs across the entire protein sequence suggest that the equine CYP1A2 might be a functional protein responsible for activities similar to that of a human homolog. This indicated that a catalytically active equine CYP1A2 could potentially be expressed in a bacterial expression system without conventional N-terminal modification, as done by Kim *et al.*, (2008) for human CYP1A2.

The *E. coli* expression system is widely used for production of recombinant proteins. The T7 expression system was selected as it often leads to high expression levels of recombinant protein (Studier and Moffatt, 1986). For equine CYP1A2 expression, the gene sequence was first synthesized and codon optimized for bacterial expression. The codon optimization was necessary because the codon bias may be applicable in *E. coli* for expression of recombinant CYPs (Yun et *al.*, 2006; Gustafsson *et al.*, 2004). Following synthesis, the gene was subcloned from pUC57 vector to expression vector pET22-b(+). The results from restriction double digest and sequencing confirmed the presence of CYP1A2 gene in pET22-b(+). The protein was produced using auto-induction and IPTG-induced expression. Both methods achieved high expression levels as the band representing the protein of interest was clearly visible and in relatively large amounts compared to host proteins. These results suggest that high-level expression of equine CYPs in *E. coli* could be achieved without modification of the N-terminal region.

Despite the high expression levels, the target protein was not detected in the soluble fraction but found only in the insoluble fraction after cell lysis. This was an indication that the protein was

possibly accumulating as insoluble misfolded aggregates (inclusion bodies) or bound to bacterial membrane. Formation of inclusion bodies is common when high levels of protein are expressed in *E. coli*, the protein is usually non-functional and requires solubilization, refolding and purification for recovery of functionally active protein (Kane and Hartly, 1988; Vallejo and Rinas, 2004). In a study by Mast *et al.*, (2004) a full length human CYP46A1 was expressed in bacteria and localized exclusively to the bacterial membrane. In an attempt to prevent the possible formation of inclusion bodies, expression temperatures were decreased to 28°C and 20°C but the protein remained in the insoluble fraction. Low temperatures slow down the rate of expression which can enhance proper folding of the protein. The second attempt was decreasing concentration of IPTG and heme precursors, $FeCl_3$ and δ-ALA, but the CYP1A2 remained inactive. These strategies are common, especially when T7 expression is used because expression levels are generally high upon induction (Studier and Moffatt, 1986).

CYPs have the N-terminal hydrophobic signal anchor which associates with the membranes. Therefore, the possibility of localization of recombinant CYP1A2 to the bacterial membranes was worth exploring. The general procedure for obtaining CYP from bacterial membranes includes subcellular fractionation, solubilization and purification. In this study, solubilization was carried out using 1.5% of non-ionic detergents (Tween 20 and Triton X-100). Both detergents yielded the same result; the majority of protein was again detected in the insoluble fraction as assessed by SDS-PAGE analysis (figure 3.3). The concentration of the detergents was not increased because higher concentrations interfere with binding of His-tagged proteins to the nickel ions. The possible explanation for the inability of these detergents to solubilize the protein is that non-ionic detergents disrupt only lipid-lipid and lipid-protein interactions but not protein-protein interactions. Because of this they are often used to wash inclusion bodies before

solubilization (Khow and Suntrarachun, 2012). This might also be an indication that the target protein was not properly inserted into bacterial membrane, but had formed insoluble aggregates.

To solubilize the highly insoluble protein, an ionic detergent (0.5% sarkosyl) was used after cell lysis. Sarkosyl is an effective ionic detergent and allows more protein to refold (Burgess, 2009). The target protein was detected in the soluble fraction after such treatment, as assessed by SDS-PAGE analysis (figure 3.4). The bulk of the expressed protein was solubilized, suggesting that sarkosyl is an efficient ionic detergent for solubilizing CYP1A2. Sarkosyl treatment also had the effect of solubilizing host proteins, and therefore a purification step was necessary before continuing with CYP1A2 characterization. Purification of the solubilized protein was conducted using nickel affinity chromatography and it was observed that most of the protein of interest was in the flow-through, indicating poor binding (figure 3.5). Tao *et al.*, (2010) reported successful solubiliziation of His-tagged protein with sarkosyl, but purification in the presence of detergent was challenging. It has been reported that concentrations of ionic detergents greater than 0.3% interfere with proper binding of protein to nickel ions. However, a relatively small quantity of CYP1A2 did bind to the column and was eluted with 250 mM imidazole and used for characterization.

The first step for characterization of purified equine CYP1A2 requires determination of the activity, however CYP1A2 activity is dependent on NADPH-cytochrome P450 reductase. A yeast NADPH-cytochrome P450 reductase was expressed in *E. coli* BL21 and crude material from cell lysis was used for activity determination with the cytochrome *c* reduction assay. NADPH-cytochrome P450 reductase had an activity of 0.13 U/ml, which is demonstrated by the progress curve in figure 3.7. This indicated that the reductase was able to transfer electrons to

cytochrome *c* and should support CYP activity. This reductase was previously found to support equine CYPs in our laboratory.

The CYP1A2 enzymatic activity was then measured with P450-Glo CYP1A2 assay kit, however the solubilized protein was not active towards luciferin-ME. The protein remained inactive despite removal of sarkosyl, which may have allowed for refolding as reported by Burgess (2009). Modification of incubation conditions such as CYP1A2 concentration, temperature, time, reductase ratio to CYP1A2 were tested, however the protein remained inactive. It is possible that CYP1A2 remained non-functional because of generation of insoluble aggregates and partial or extensive misfolding of sarkosyl-solubilized protein.

Because the CYP1A2 was inactive, a second equine CYP, namely CYP3A89 was selected as a positive control in an attempt to validate the method used. Equine CYP3A89 had been expressed in *E. coli* and showed activity in earlier studies in our laboratory. The expression was induced by IPTG and the high level of protein produced was comparable to CYP1A2 expression (figure 3.9). Unfortunately, CYP3A89 was also insoluble as the protein remained in the pellet following cell lysis. The results suggest that the selected bacterial expression system leads to generation of insoluble CYP recombinant enzymes.

To purify CYP3A89 using an affinity column, sarkosyl (0.5%) was again used for solubilization. Although some CYP3A89 was detected in the flow-through, the majority of the enzyme was eluted with 250 mM imidazole. Solubilized and purified CYP3A89 was used for activity determination. Unfortunately CYP3A89 was also inactive and could not be used as a positive control. However, a pronounced difference was observed in binding of CYP3A89 to nickel affinity column compared to CYP1A2 purification. Binding of CYP3A89 in the presence of

sarkosyl suggests that this detergent might not be the major influence on binding of CYP1A2 to a nickel affinity column. Factors that might influence CYP1A2 binding include degradation or His-tag positioning. However, it seemed unlikely that CYP1A2 could be degraded because the cell lysate had been prepared in a buffer containing protease inhibitors. The sequencing results also confirmed that His-tag sequence was in-frame and should be translated along with CYP1A2. It is therefore unlikely that His-tag is not fused to the target protein. It could be that His-tag on the C-terminus is incorporated into the structure of CYP1A2 and not exposed to nickel ions. Using pET28-expressing vector which has His-tag on the N-terminus could be an alternative to resolve His-tag positioning.

It was observed that production of insoluble protein aggregates occurred for CYP1A2 and CYP3A89 expression and both enzymes were inactive. The only difference between this study and an earlier study where CYP3A89 was active was the batch of *E. coli* BL21 cells used. However, it seemed unlikely that differences in cell batchess could be the problem because the strain is the same. The strength of T7 promoter might cause production of insoluble protein, as it has been reported that high expression levels of recombinant proteins often lead to formation of inclusion bodies. Failure to obtain soluble CYP1A2 and CYP3A89 by decreasing expression temperature and IPTG concentration led to use of solubilizing detergent. However, Sarkosyl treatment of inclusion bodies resulted in protein with no biological activity; this might be an indication that suitable refolding conditions were not found. Finding the suitable *in vitro* conditions for protein refolding after such treatment is a challenge. Strategies to prevent formation of insoluble CYP aggregates during expression are required for production of functional equine CYP1A2.

4.2 Future work

The present study is the first work involving expression of equine CYP1A2 in a bacterial expression system. It can be considered as part of an ongoing effort intended to provide an *in vitro* method for studying horse drug metabolism by a developing recombinant CYP system. The findings of this study revealed that a full-length CYP1A2 protein can be expressed in *E. coli* without modifying the protein sequence, most especially, the N-terminal region. However, more work remains to be done in order to express a functionally active CYP1A2. Production of soluble recombinant CYP1A2 protein remains a major challenge that requires further investigation.

The pET system used led to high expression levels of the target protein, but this was accompanied by formation of inclusion bodies. The T7 promoter often leads to over production of protein because of high affinity and specific binding of T7 polymerase to the promoter (Studier and Moffatt, 1986). The high-level expression and formation of insoluble protein aggregates also hinders the ability of *E. coli* to fold the protein to its native conformation. Co-expression of CYP protein with a molecular chaperone which assists in protein folding tends to improve correct folding of CYPs and proper incorporation of the heme. For example, human CYP3A7 was co-expressed with a molecular chaperone, namely GroEL (Inoue *et al.*, 2000), and this system yielded a functional protein.

GroEL is an *E. coli* chaperonin which is known to bind partially folded or misfolded proteins, preventing formation of protein aggregates (Corrales and Fersht, 1996). The bound protein undergoes conformational changes towards correct folding. Yan *et al.*, (2012) reported co-expression of human interferon-gamma protein with GroEL in *E. coli*. Over-expression of this protein was previously found to form inclusion bodies. Co-expressing with GroEL produced soluble and active interferon-gamma. Ahn *et al.*, (2004) co-expressed human CYP1A2 with a

human chaperone HDJ-1 in *E. coli*. It was shown that HDJ-1 was involved in proper folding and membrane integration of CYP1A2. Co-expression increased expression levels by 3.4-fold and catalytic activity by 5.5-fold. These studies demonstrated the importance of molecular chaperones in preventing formation of aggregates and production of catalytically active recombinant proteins. Thus, the insoluble CYP1A2 aggregates observed in this study could be prevented by co-expressing with molecular chaperone. This might produce an active equine CYP1A2 in *E. coli*.

Most of the expressed CYPs are localized in the bacterial membranes, and targeting CYP protein to the membranes leads to proper folding. Pritchard *et al*., (1997) have successfully introduced a bacterial signal sequence to target mammalian CYPs to the bacterial membranes. Human CYP3A4, CYP2A6 and CYP2E1 were fused to two different leader sequences, *omp*A and *pel*B. CYPs expressed with *omp*A showed an increase in membrane content and were catalytically active. These leader sequences were also used for production of properly folded recombinant human growth hormone (hGH) (Sockolosky and Szoka, 2013). Over-expression of recombinant hGH using a pET expression system reportedly leads to formation of inclusion bodies. Sockolosky and Szoka, (2013) showed that hGH was directed to the *E. coli* periplasm, was soluble and functionally active. Secretion of proteins into periplasm is also advantageous because isolation of the protein can be easily achieved by selective disruption of the *E. coli* outer membrane. This could be another potential alternative for expression of active equine CYP1A2 in *E. coli* and overcoming accumulation of insoluble aggregates.

Modification of the N-terminus is a common strategy for functional expression of mammalian CYPs in bacterial expression systems. The expressed CYP can be directed to the bacterial membrane by replacing the native N-terminal portion with MALLLAVFL sequence previously

established by Barnes *et al.*, (1991). The addition of this sequence in human CYP1A2 increased expression in the membrane from less than 2 nmol protein to 225 nmol protein (Sandhu *et al.*, 1994). CYP enzymes expressed using this approach include human CYP3A43, CYP2B6 and CYP2D6 (Domaski *et al.*, 2001; Hanna *et al.*, 2000; Pan *et al.*, 2011). This strategy could be used when the equine CYP1A2 is expressed using the pCWOri vector, which is commonly used for functional CYP expression and has not been reported for formation of insoluble aggregates.

Alternatively, the hydrophobic N-terminal region could be truncated in order to localize the protein in the cytosol. Truncation of the N-terminal region has been used for expression of mammalian CYPs. In the expression of CYP46A1, Mast *et al.*, (2004) reported that only about 47-57% of truncated protein was found in the cytosol in low ionic buffer and the rest remained membrane-bound. However, in the expression of rat cholesterol 7a-hydroxylase, about 85% of truncated protein was found in the cytosol and remained catalytically active (Li and Chiang, 1991). These strategies are valuable in achieving the maximal yield of catalytically active recombinant cytochrome P450 enzymes.

This study demonstrated that a full-length equine CYP1A2 could be expressed using the pET system under the control of T7 promoter. Although the recombinant equine CYP1A2 was found to be insoluble and non-functional, this work is the first step towards characterizing members of the equine CYP1A subfamily by developing a recombinant CYP enzyme system in *E. coli*. To further investigate the production of functionally active equine CYP1A2, alternative strategies discussed above need to be explored. It is to be hoped that the development of a recombinant CYP enzyme system will provide a cost-effective method for determining metabolites of illegal performance-enhancing drugs in horse racing.

REFERENCES

Ahn, T., Yang, S. and Yun, C. (2004). High-level expression of human cytochrome P450 1A2 by co-expression with human molecular chaperone HDJ-1(Hsp40). Protein Expression and Purification, 36: 48-52.

Banexyn, F. (1999). Recombinant protein expression in *Escherichia coli*. Current Opinion in Biotechnology, 10: 411-421.

Bertilsson, L., Carrillo, J. A., Dahl, M. L., Llerena, A., Alm, C., Bondesson, U., Lindstrom, L., Rodriguez de la Rubia, L., Ramos, S. and Benitez, J. (1994). Clozapine disposition covaries with CYP1A2 activity determined by a caffeine test. British Journal of Clinical Pharmacology, 38: 471-473.

Bialey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. and Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Research, 37: W202-W208.

Barnes, H. J., Arlotto, M. P., and Waterman, M. R. (1991). Expression and enzymatic activity of recombinant cytochrome P45017a-hydroxylase in *Escherichia coli*. Proceedings of the National Academy of Science, 88: 5597-5601.

Burgess, R. R. (2009). Refolding solubilized inclusion body proteins. Methods in Enzymology, 463: 259-282.

Burgess-Brown, N. A., Sharma, S., Sobott, F., Loenarz, C., Oppermann, U. and Gileadi, O. (2008). Codon optimization can improve expression of human genes in *Escherichia coli*: A multi-gene study. Protein Expression and Purification, 59: 94-102.

Buters, J. T., Shou, M., Hardwick, J. P., Korzekwa, K. R. And Gonzalez, F. J. (1995). cDNA-directed expression of human cytochrome P450 CYP1A1 using baculovirus. Purification, dependency on NADPH-P450 oxidoreductase, and reconstitution of catalytic properties without purification. Drug Metabolism and Disposition, 23: 696-701.

Calderone, T.L., Stevens, R. D. and Oas, T. G. (1996). High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli*. Journal of Molecular Biology, 262: 407-412.

Cali, J. J., Ma, D., Sobol, M., Simpson, D. J., Frackman, S., Good, T. D., Daily, W. J. and Liu, D. (2006). Luminogenic cytochrome P450 assays. Expert Opinion on Drug Metabolism and Toxicology, 2: 629-645.

Chauret, N., Gauthier, A., Martin, J. and Nicoll-Griffith, D. A. (1997). In vitro comparison of cytochrome P450-mediated metabolic activities in human, dog, cat, and horse. Drug Metabolism and Disposition, 25: 1130-1136.

Chen, W., Peter, R. M., McArdle, S., Thummel, K. E., Sigle, R. O. and Nelson S. D. (1996). Baculovirus Expression and Purification of Human and Rat Cytochrome P450 2E11. Archives of Biochemistry and Biophysics, 335: 123-130.

Clarke, S. E. (1998). In vitro assessment of human cytochrome P450. Xenobiotica, 28: 1167-1202.

Cloutier, M., Fleury, A., Courtemanche, J., Ducharme, L., Mason, J. I. and Lehoux J. G. (1997). Characterization ofthe adrenal cytochrome P450C17 in the hamster, a small animal model for the study of adrenal dehydroepiandrosterone biosynthesis. DNA Cell Biology, 16: 357-368.

Corley-Smith, G. E., Su, H. T., Wang-Buhler, J. L., Tseng, H. P., Hu, C-H., Hoang, T., Chung, WG., and Buhler, D.R. (2006). CYP3C1, the first member of a new cytochrome P450 subfamily found in zebrafish *(Danio rerio)*. Biochemical and Biophysical Research Communications, 240: 1039-1046.

Corrales, F. J. and Fersht, A. R. (1996). Toward a mechanism of GroEL-GroES chaperone activity: An ATPase-gated –pulse folding and annealing cage. Proceedings of the National Academy of Sciences, 93: 4509-4512.

Crespi, C. L., Penman, B. W., Steime, D. T., Smith, T., Yang, C. S., and Sutter T. R. (1997). Development of a human lymphoblastoid cell line constitutively expressing human CYP1B1 cDNA: substrate specificity with model substrates and promutagens. Mutagenesis, 12: 83-9.

Darwish, W. S., Ikenaka, Y., Eldaly, E. A., Ohno, M., Sakamoto, K. Q., Fujita, S. and Ishizuka, M. (2010). Cytochrome P450 1A dependent activities in deer, cattle and horses. Journal of Veterinary Medical Science, 72: 561-566.

Dettwiler, R., Schmitz, A. L. Plattet, P., Zielinski, J. and Mevissen, M. (2014). Heterologous expression of equine CYP3A94 and investigation of a tunable system to regulate co-expressed NADPH P450 oxidoreductase levels. PLoS ONE, 9: 1-20.

DiMaio Knych, H. K. and Stanley, S. D. (2008). Complementary DNA cloning, functional expression and characterization of a novel cytochrome P450, CYP2D50, from equine liver. Biochemical Pharmacology, 76: 904-911.

DiMaio Knych, H. K., McKemie, D. S. and Stanley, S. D. (2010). Molecular cloning, expression, and initial characterization of members of the CYP3A family in horses. Drug Metabolism and Disposition, 38: 1820-1827.

DiMaio Knych, H. K., DeStefano Shields, C., Buckpitt, A. R. and Stanley, S. D., (2009). Equine cytochrome P450 2C92: cDNA cloning, expression and initial characterization. Archives of Biochemistry and Biophysics, 485: 49-55.

Domanski, T. L., Finta, C., Halpert, J. R., Zaphiropoulos, P. G. (2001). cDNA cloning and initial characterization of CYP3A43, a novel human cytochrome P450. Molecular Pharmacology, 59: 386-392.

Fairbanks, G., Steak, T. L., and Wallach, D. F. H. (1971). Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. Biochemistry, 10: 2606-2671.

Fink-Gremmels, J. (2008). Implications of hepatic cytochrome P450-related biotransformation processes in veterinary sciences. European Journal of Pharmacology, 585: 502-509.

Fisher, C. W., Caudle, D. L., Martin-Wixtrom, C., Quattrochi, L. C., Tukey, R. H., Waterman, M. R. and Estabrook, R.W. (1992). High-level expression of functional human cytochrome P450 1A2 in *Escherichia coli*, FASEB Journal, 6: 759-764.

Frankel, S., Sohn, R. and Leinwand, L. (1991). The use of sarkosyl in generating soluble protein after bacterial expression. Proceedings of the National Academy of Science, 88: 1192-1196.

Froger, A. and Hall, J. E. (2007). Transformation of plasmid DNA into *E. coli* using the heat shock method. Journal of Visualized Experiments, 6: 253.

Gillam, E. M. J. (1998). Human cytochrome P450 enzymes expressed in bacteria: reagent to probe molecular interactions in toxicology. Clinical and Experimental Pharmacology and Physiocolgy, 25: 877-886.

Goda, R., Nagai, D., Akiyama, Y., Nishikawa, K., Ikemoto, I., Aizawa, Y., Nagata, K. and Yamazoe, Y. (2006). Detection of a new *N*-oxidized metabolite of flutamide, *N*-[4-nitro-3 (trifluoromethyl)phenyl]hydroxylamine, in human liver microsomes and urine of prostate cancer patients. Drug Metabolism and Disposition, 34: 828-835.

Gonzalez, F.J. and Korzekwa, K.R. (1995). Cytochrome P450 expression systems. Annual Reviews of Pharmacology and Toxicology, 35: 369-390.

Graumann, K. and Premstaller, A. (2006). Manufacturing of recombinant therapeutic proteins in microbial systems. Biotechnology Journal, 1: 164-186.

Guengerich, F. P. (1997). Comparisons of catalytic selectivity of cytochrome P450 subfamily enzymes from different species. Chemico-Biological Interactions, 106: 161-182.

Gustafsson, C., Govindarajan, S. and Minshull, J. (2004). Codon bias and heterologous protein expression. Trends in Biotechnology, 22: 346-353.

Ha, H. R., Chen, J., Freiburghaus, A. U. and Follath, F. (1995). Metabolism of theophylline by cDNA-expressed human cytochromes P-450. British Journal of Clinical Pharmacology, 39: 321-326.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmid. Journal of Molecular Biology, 166: 557-580.

Hanna, I. H., Reed, J., Guengerich, F. P. and Hollenberg, P. F. (2000). Expression of human cytochrome P450 2B6 in *Escherichia coli*: characterization of catalytic activity and expression levels in human liver. Archives of Biochemistry and Biophysics, 376: 206-216.

Hannemann, F., Bichet, A., Ewen, K. M. and Bernhardt, R. (2007). Cytochrome P450 systems – biological variations of electron transport chains. Biochimica et Biophysica Acta, 1770: 330-344.

Hao, D. C., Zhu, P. H., Yang, S. L. and Yang, L. (2007). Enhanced production of human Cytochrome P450 2C9 by *Escherichia coli* BL21 (DE3) pLysS through the novel use of grey relational analysis and Plackett–Burman design. World Journal of Microbiology and Biotechnology, 23: 71-78.

Ichinose, H. and Wariishi, H. (2012). Heterologous expression and mechanistic investigation of fungal cytochrome P450 (CYP5150A2): involvement of alternative redox partners. Archives of Biochemistry and Biophysics, 518: 8-15.

Inoue, E., Takahashi, Y., Imai, Y. and Kamataki, T. (2000). Development of bacterial expression system with high yield of CYP3A7 a human fetus-speciWc form of cytochrome P450. Biochemical and Biophysical Research Communications, 269: 623-627.

Iwata, H., Fujita, K., Kushida, H., Suzuki, A., Konno, Y., Nakamura, K., Fujino, A. and Kamataki, T. (1998). High catalytic activity of human cytochrome P450 co-expressed with human NADPH-cytochrome P450 reductase in *Escherichia coli*. Biochemical Pharmacology, 55: 1315-1325.

Jenkins, C. M. and Waterman, M. R. (1994). Flavodoxin and NADPH-Flavodoxin from *Escherichia coli* support bovine cytochrome P450c17 hydroxylase activities. Journal of Biological Chemistry, 269: 27401-27408.

Kane, J.F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*, Current Opinion in Biotechnology, 6: 494-500.

Kane, J. F. and Hartly, D. L. (1988). Formation of recombinant protein inclusion bodies in *Escherichia coli*. Trends in Biotechnology, 6: 95-101.

Kalgutkar, A. S., Nguyen, H. T., Vaz, A. D., Doan, A., Dalvie, D. K., McLeod, D. G. and Murray, J. C. (2003). *In vitro* metabolism studies on the isoxazole ring scission in the anti-inflammatory agent lefluonomide to its active α-cyanoenol metabolite A771726: mechanistic similarities with the cytochrome P450-catalyzed dehydration of aldoximes. Drug Metabolism and Disposition, 31: 1240-1250.

Karjalainen, M. J., Neuvonen, P. J. and Backman, J. T. (2007). Tolfenamic acid is a potent CYP1A2 inhibitor *in vitro* but does not interact *in vivo*: correction for protein binding is needed for data interpretation. European Journal of Clinical Pharmacology, 63: 829-836.

Khow, O. and Suntrarachun, S. (2012). Strategies for production of active eukaryotic proteins in bacterial expression system. Asian Pacific Journal of Tropical Biomedicine, 2: 159-162.

Kim, D. and Guengerich, F. P. (2004). Selection of human cytochrome P450 1A2 mutants with enhanced catalytic activity for heterocyclic amine *N*-hydroxylation. Biochemistry, 43: 981-988.

Kim, D. H., Kim, K. H., Isin, E. M., Guengerich, F. P., Chae, H. Z., Ahn, T. and Yun, C. H. (2008). Heterologous expression and characterization of wild-type human cytochrome P450 1A2 without conventional N-terminal modification in *Escherichia coli*. Protein Expression and Purification, 57: 188-200.

Kimura, S., Gonzalez, J. and Nebert, W. (1986). Tissue-specific expression of the mouse dioxin-inducible P(1)450 and P(3)450 genes: differential transcriptional activation and mRNA stability in the liver and extrahepatic tissues. Molecular and Cellular Biology, 6:1471-1477.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227: 680-685.

Lakrtiz, J., Winde,r B. S., Noorouz-Zadeh, J., Huang, T. L., Buckpitt, A. R. and Hammock, B. D. (2000). Hepatic and pulmonary enzyme activities in horses. Amercian Journal of Veterinary Research, 61(2): 152-157.

Lewis, D. F., Lake, B. G., Dickins, M., Ueng, Y. F. and Goldfarb, P. S. (2003). Homology modelling of human CYP1A2 based on the CYP2C5 crystallographic template structure. Xenobiotica, 33: 239-254.

Li, Y. C. Chiang, J. Y. (1991). Expression of catalytically active cholesterol 7a-hydroxylase cytochrome P450 in *Escherichia coli*. Journal of Biochemical Chemistry, 266: 19186-19191.

Lobo, E. D., Bergstrom, R. F., Reddy, S., Quinlan, T., Chappell, J., Hong, Q., Ring, B. and Knadler, M. P. (2008). *In vitro* and *in vivo* evaluations of cytochrome P450 1A2 interactions with duloxetine. Clinical Pharmacokinetics, 47: 191-202.

Mast, N., Andersson, U., Nakayama, K., Bjorkhem, I. and Pikuvela, I. A. (2004). Expression of human cytochrome P450 46A1 in *Escherichia coli*: effects of N- and C-terminal modifications. Archives of Biochemistry and Biophysics, 428: 99-108.

Messina, A., Chirulli, V., Gervasi, P. G. and Longo, V. (2008). Purification, molecular cloning, heterologous expression and characterization of pig CYP1A2. Xenobiotica, 38: 1453-1470.

Meunier, B., Visser, S. P., and Shaik, S. (2004). Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. Chemical Reviews, 104: 3947-3980.

Miners, J. O., Coulter, S., Tukey, R. H., Veronese, M. E. and Birkett, D. J. (1996). Cytochromes P450, 1A2, and 2C9 are responsible for the human hepatic *O*-demethylation of *R*- and *S* naproxen. Biochemical Pharmacology, 51: 1003-1008.

Momo, K., Doki, K., Hosono, H., Homma, M. and Kohda, Y. (2004). Drug interaction of tizanidine and fluvoxamine. Clinical Pharmacology and Therapeutics, 76: 509-510.

Muller, K. M., Arndt, K. M., Bauer, K. and Pluckthun, A. (1998). Tandem immobilized metal-ion affinity chromatography/immunoaffinity purification of his- tagged proteins- evaluation of two anti-his-tag monoclonal antibodies. Analytical Biochemistry, 259: 54-61.

Narimatsu, S., Oda, M., Hichiya, H., Isobe, T., Asaoka, K., Hanioka, N., Yamano, S., Shinoda, S. and Yamamoto, S (2005). Molecular cloning and functional analysis of cytochrome P450 1A2 from Japanese monkey liver: comparison with marmoset cytochrome P450 1A2. Chemico-Biological Interactions, 152: 1-12.

Nebbia, C., Dacasto, M., Rossetto Giaccherino A., Giuliano Albo, A. and Carletti M. (2003). Comparative expression of liver cytochrome P450-dependent monooxygenases in the horse and in other agricultural and laboratory species. Veterinary Journal, 165: 53-64.

Nebert D.W., Dalton T. P., Okey, A. B. and Gonzalez, F.J. (2004). Role of aryl hydrocarbon receptor-mediated induction of the CYPl enzymes in environmental toxicity and cancer. Journal of BioIogical Chemistry, 279: 23847-23850.

Nebert, D.W. and Dalton, T.P. (2006). The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. Nature Reviews Cancer, 6: 947-960.

Nelson, D. R., Koymans, L., Kamataki, T., Stegeman, J. J., Feyreisen, R. and Waxman, D. J. (1996). P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics, 6: 1-42.

Niwa, T., Sato, R., Yabusaki, Y., Ishibashi, F. and Katagiri, M. (1999). Contribution of human hepatic cytochrome P450s and steroidogenic CYP17 to the *N*-demethylation of aminopyrine. Xenobiotica, 29: 187-193.

Omiecinski, J., Redlich, A. and Costa, P. (1990). Induction and developmental expression of cytochrome P4501A1 messenger RNA in rat and human tissues: detection by the polymerase chain reaction. Cancer Research, 50: 4315-4321.

Omura, T. and Sato, R. (1964). The carbon monoxide-binding pigment of liver microsomes. Journal of Biological Chemistry, 239: 2370-2378.

Pan,Y., Abd-Rashid, B. A., Ismail, Z., Ismail, R., Mak, J. W. And Ong, C. E. (2011). Heterologous Expression of Human Cytochromes P450 2D6 and CYP3A4 in *Escherichia coli* and Their Functional Characterization. Protein Journal, 30: 581-591.

Parikh, A., Josephy, P. D. and Guengerich, F. P. (1999). Selection and Characterization of Human Cytochrome P450 1A2 Mutants with Altered Catalytic Properties. Biochemistry, 38: 5283-5289.

Peters, L. M., Demmel, S., Pusch, G., Buters, J. T. M., Thormann, W., Zielinski, J., Leeb, T., Mevissen, M. and Schmitz, A. (2013). Equine cytochrome P450 2B6 — Genomic identification, expression and functional characterization with ketamine. Toxicology and Applied Pharmacology, 266: 101-108.

Peterson, J. A. and Deisenhofer, J. (1995). Structure and function of cytochrome P450: comparative analysis of three crystal structures. Structure, 2: 41-62.

Pritchard, M. P, Ossetian, R., Li, D. N., Henderson, C. J., Burchell, B., Wolf, C. R. and Friedberg, T. (1997). A general strategy for the expression of recombinant human cytochrome

P450s in *Escherichia coli* using bacterial signal peptides: expression of CYP3A4, CYP2A6, and CYP2E1. Archives of. Biochemistry and Biophysics, 345: 342-354.

Ring, B. J., Catlow, J., Lindsay, T. J., Gillespie, T., Roskos, L. K., Cerimele, B. J., Swanson, S. P., Hamman, M. A. and Wrighton, S. A. (1996). Identification of the human cytochromes P450 responsible for the *in vitro* formation of the major oxidative metabolites of the antipsychotic agent olanzapine. Journal of Pharmacology and Experimental Therapeutics, 276: 658-666.

Rosano, G. L. and Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. Frontiers in Microbiology, 5: 172.

Rowland, P., Blaney, F. E., Smyth, M. G., Jones, J. J., Leydon, V. R. and Oxbrow, A. K. (2006). Crystal structure of human cytochrome P450 2D6. Journal of Biological Chemistry, 281: 7614-7622.

Ryu, J. Y., Song, I. S., Sunwoo, Y. E., Shon, J. H., Liu, K. H., Cha, I. J. and Shin, J. G. (2007). Development of the "Inje cocktail" for highthroughput evaluation of five human cytochrome P450 isoforms *in vivo*. Clinical Pharmacology and Therapeutics, 82: 531-540.

Sanderink, G. J., Bournique, B., Stevens, J., Petry, M. and Martinet, M. (1997). Involvement of human CYP1A isoenzymes in the metabolism and drug interactions of riluzole *in vitro*. Journal of Pharmacology and Experimental Therapeutics, 282: 1465-1472.

Sandhu, P., Guo, Z., Baba, T., Martin, M. V., Turkey, R. H. and Guengerich, F. P. (1994). Expression of modified human cytochrome P450 1A2 in *Escherichia coli*: stabilization, purification, spectral characterization, and catalytic activities of the enzyme. Archives of Biochemistry and Biophysics, 309: 168-177.

Sansen, S., Yano, J. K., Reynald, R. L., Schoch, G. A., Griffin, K. J. and Stout, C. D. (2007). Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2. Journal of Biological Chemistry, 282: 14348-14355.

Scarth, J. P., Teale, P and Kuuranne, T. (2011). Drug metabolism in the horse: a review. Drug Testing and Analysis, 3: 19-53.

Schmitz, A., Demmel, S., Peters, L. M., Leeb, T., Mevissen, M and Haase, B. (2010). Comparative human-horse sequence analysis of the CYP3A subfamily gene cluster. Animal Genetics, 41: 72-79.

Schmitz, A., Zielinski, J., Dick, B., Mevissen, M. (2014). *In vitro* metabolism of testosterone in the horse liver and involvement of equine CYPs 3A89, 3A94 and 3A95. Journal of veterinary Pharmacology and Therapeutics, 37: 338-347.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWillian, H., Remmert, M., Soding, J., Thompson, J. and Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology, 7: 539.

Sirim, D., Widmann, M., Wagner, F. and Pleiss, J. (2010). Prediction and analysis of modular structure of cytochrome monooxygenases. BMC Structural Biology, 10: 34.

Sockolosky, J. T. and Szoka, F. C. (2013). Periplasmic production via the pET expression system of soluble, bioactive human growth hormone. Protein Expression and Purification, 87: 129-135.

Spaldin, V., Madden, S., Pool, W. F., Woolf, T. F. and Park, B. K. (1994). The effect of enzyme inhibition on the metabolism and activation of tacrine by human liver microsomes. British Journal of Clinical Pharmacology, 38: 15-22.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification, 41: 207-234.

Studier, F. W. and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Biological Chemistry, 189: 113-130.

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985). Measurement of protein using bicinchoninic acid. Analytical Biochemistry, 150: 76-85.

Tanaka, N., Miyasho, T., Shinkyo, R., Sakaki, T. and Yokota H. (2006). cDNA cloning and characterization of feline CYP1A1 and CYP1A2. Life Science, 79: 2463-2473.

Tao, H., Lui, W., Simmons, B. N., Harris, H. K., Cox, C. and Massiah, M. A. (2010). Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. BioTechniques, 48: 61-64.

Tassaneeyakul, W., Birkett, D. J., Veronese, M. E., McManus, M. E., Tukey, R. H., Quattrochi, L. C., Gelbion, H. V. and Miners, J. O. (1993). Specificity of substrate and inhibitor probes for human cytochromes P450 1A1 and 1A2. Journal of Pharmacology and Experimental Therapeutics, 265: 401-407.

Tassaneeyakul, W., Mohamed, Z., Birkett, D. J., McManus, M. E., Veronese, M. E., Tukey, R. H., Quattrochi, L. C., Gonzalez, F. J. and Miner, J. O. (1992). Caffeine as a probe for human cytochromes P450: validation using cDNA expression, immunoinhibition and microsomal kinetic and inhibitor techniques. Pharmacogenetics, 2: 173-183.

Tydén, E., Tjälve, H. and Larsson, P. (2014). Gene and protein expression and cellular localization of cytochrome P450 enzymes of the1A, 2A, 2C, 2D and 2E subfamilies in equine intestine and liver. Acta Veterinaria Scandinavica, 56: 1-8.

Urban, P., Cullin, C.and Pompon, D. (1990). Maximizing the expression of mammalian cytochrome P-450 monooxygenase activities in yeast cells. Biochemistry, 72: 463-472.

Vallejo, L. F. and Rina, V. (2004). Strategy for recovery of active protein through refolding of bacterial inclusion body protein. Microbial Cell Factories, 3: 2-12.

Ventura, S. and Villaverde, A. (2006). Protein quality in bacterial inclusion bodies. Trends in Biotechnology. 24: 179-185.

Waterman, M. R., Jenkins, C. M. and Pikuleva, I. (1995). Genetically engineered bacterial cells and applications. Toxicology letters, 83: 807-813.

Waxman, D. J. and Chang, T. K. (2006). Use of 7-ethoxycoumarin to monitor multiple enzymes in the human CYP1, CYP2, and CYP3 families. Methods in Molecular Biology, 320: 153-156

Williams, P. A., Cosme, J., Ward, A., Angove, H. C., Matak Vinkovic, D.and Jhoti, H. (2003). Crystal structure of human cytochrome P4502C9 with bound warfarin. Nature, 424: 464-468.

Wojcikowski, J., Pichard-Garcia, L., Maurel, P. and Daniel, W. A. (2003). Contribution of human cytochrome P-450 isoforms to the metabolism of the simplest phenothiazine neuroleptic promazine. British Journal of Pharmacology, 138: 1465-1474.

Wu, Z.L., Bartleson, C.J., Ham, A.J. and Guengerich, F. P. (2006). Heterologous expression, purification, and properties of human cytochrome P450 27C1. Archives of Biochemistry and Biophysics, 445: 138-146.

Wu, Z. L., Huang, S. L., Ou-Yang, D. S., Xu, Z. H., Xie, H. G. and Zhou, H. H. (1998). Clomipramine *N*-demethylation metabolism in human liver microsomes. Zhongguo Yao Li Xue Bao, 19: 433-436.

Yan, X., Hu, S., Guan, Y. and Yao, S. (2012). Coexpression of chaperonin GroEL/GroES markedly enhanced soluble and functional expression of recombinant human interferon-gamma in *Escherichia coli*. Applied Microbiology and Biotechnology, 93: 1065-1074.

Yano, J. K., Hsu, M. H., Griffin, K. J., Stout, C. D. and Johnson, E. F. (2005). Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen. Nature Structural and Molecular Biology, 12: 822-823.

Yano, J. K., Wester, M. R., Schoch, G. A., Griffin, K. J., Stout, C. D. and Johnson, E. F. (2004). The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-A resolution. Journal of Biological Chemistry, 279: 38091-38094.

Yun, C. H. Miller, G. P. Guengerich, F. P. (2000). Rate-determining steps in phenacetin oxidations by human cytochrome P450 1A2 and selected mutants. Biochemistry, 39: 11319-11329.

Yun, C.H., Yim, S.K., Kim, D.H. and Ahn, T. (2006). Functional expression of human cytochrome P450 enzymes in *Escherichia coli*, Current Drug Metabolism, 7: 411-429.

Zelasko, S., Palaria, A. and Das, A. (2013). Optimizations to achieve high level expression of cytochrome P450 proteins using *Escherichia coli* expression systems. Protein Expression and Purification, 92: 77-87.

Zhou, S., Wang, B.,Yang, L-P. And Liu, J-P. (2010). Structure, function, regulation and polymorphismand the clinical significance of human cytochrome P450 1A2. Drug Metabolism Reviews, 42(2): 268-354.

Zhou, S., Yang, L-P., Zhou, Z-W., Liu, Y-H., and Chan, E. (2009). Insights into the Substrate Specificity, Inhibitors, Regulation, and Polymorphisms and the Clinical Impact of Human Cytochrome P450 1A2. The AAPS Journal, 11 (3): 481-494.

Zhou, Z., Schnake, P., Xiao, L. and Lal, A. A. (2004). Enhanced expression of recombinant malaria candidate vaccine in *Escherichia coli* by codon optimization. Protein Expression and Purification, 34: 87-94.

Zuber, M. X., Mason, J. I., Simpson, E. R. and Waterman M. R., (1988). Simultaneous transfection of COS-1 cells with mitochondrial and microsomal steroid hydroxylases: Incorporation of a steroidogenic pathway into nonsteroidogenic cells. Proceedings of the National Academy of Science, 85: 699-703.

APPENDICES

Appendix 1: Media

Table A1

Luria broth	10 g/L Tryptone 5 g/L Yeast extract 10 g/L NaCl
Luria agar	10 g/L Tryptone 5 g/L Yeast extract 10 g/L NaCl 15 g/L Agar
SOB	20 g/L Tryptone 5 g/L Yeast extract 0.584 g/L NaCl 0.186 g/L KCl 2.034 g/L $MgCl_2$ 2.464 g/L $MgSO_4$
SOB	20 g/L Tryptone 5 g/L Yeast extract 0.584 g/L NaCl 0.186 g/L KCl 2.034 g/L $MgCl_2$ 2.464 g/L $MgSO_4$ 20 ml/L Glucose
Auto-induction **ZY media** **NPS** **5052**	 10 g/L Tryptone 5 g/L Yeast extract 66.07 g/L $(NH_4)_2SO_4$ 136.09 g/L KH_2PO_4 141.96 g/L Na_2HPO_4 25% Glycerol 2.5% Glucose 10% α-Lactose

Appendix 2: Biospin Plasmid DNA Extraction Kit

1. Add 1.5 ml cultured bacteria to 1.5 micro centrifuge tube.
2. Centrifuge at 10 000 × *g* for 30 seconds, and discard the supernatant
3. Resuspend pelleted bacterial cells in 250 µl Resuspension Buffer and No cell clumps should be visisble after resuspension of the pellets.
4. Add 250 µl Lysis Buffer and gently invert the tube 4-6 times to mix (do not vortex, as this will result in shearing of genomic DNA. Do not allow this step for more than 5 minutes).
5. Add 350 µl Neutralization Buffer and gently invert the tube 4-6 times to mix (the solution should become cloudy and no local precipitate should be visible).
6. Centrifuge for 10 minutes at 13 000 × *g* until a compact white pellet form.
7. Apply the supernatant to the Spin column and centrifuge for 60 seconds at 6 000 × *g*. discard the flow-through.
8. Add 650 µl Wash Buffer to the Spin column and Centrifuge for 60 seconds at 12 000 × *g*. discard the flow-through.
9. Repeat step 8 once.
10. Centrifuge for an additional 1 minute at 12 000 × *g* and transfer the Spin column at sterile 1.5 ml micro centrifuge tube (recommend to centrifuge according to this step; otherwise, there will be residual liquid in the column)
11. Add 50 µl Elution Buffer to the Spin column and let it stand for 1 minute at room temperature.
12. Centrifuge for 1 minute at 12 000 × *g*. The Buffer in the micro centrifuge tube conatins the plasmid DNA.

13. The purified plasmid DNA can be used directly for kinds of downstream molecular biological experiments. Store at -20°C if not used immediately.

Appendix 3: Biospin Gel Extraction Kit

1. Excise the DNA fragment from the agarose gel within a clean, sharp scalpel (minimize the size of the slice by removing extra agarose).
2. Weigh the gel slice and add 3 volumes of Extraction Buffer to 1 volume of gel slice (100 mg=100 μl).
3. Incubate at 50°C until the gel melts in a heating block and vortex the tube every 2-3 minutes during the incubation (usually, it is 10 minutes. If the color of the mixture is purple add 10 μl of 3 M sodium acetate (pH 5.0), and mix. The color will return to yellow).
4. Optional: Add 1 volume of isopropanol to 1 volume of the gel and mix.
5. Apply the sample to Spin column, centrifuge for 60 seconds at 6 000 × *g*. discard the flow-through (if the sample volume is more than 750 μl, simply load and spin again).
6. Add 500 μl Extraction Buffer to Spin column, centrifuge for 60 seconds at 12 000 × *g*. Discard the flow-through.
7. Add 750 μl Wash Buffer to Spin column, centrifuge for 60 seconds at 12 000 × *g*. discard the flow-through (if the DNA will be used for salt sensitive applications, let the spin column stand 2-3 minutes after addition of wash Buffer, before centrifuging).
8. Centrifuge for additional 1 minute at 12 000 × *g* and transfer the Spin column to a sterile 1.5 ml micro centrifuge tube (recommend to centrifuge according to this step; otherwise, there will be residual liquid in the column).

9. Add 50 µl Elution Buffer to the Spin column and let it stand for 1 minute at room temperature.
10. Centrifuge for 1 minute at 12 000 × *g*. The buffer in the micro centrifuge tube contains the DNA (the extracted DNA can be used directly for kinds of downstream molecular biological experiments. Store at -20°C if not used immediately).

Appendix 4: SDS-PAGE solutions

Table A2

SDS-PAGE loading buffer	1.6 ml dH2O 0.5 ml 0.5 M Tris-HCl (pH 6.8) 0.8 ml 50% glycerol 0.8 ml 10% SDS 0.1 ml acetic acid
SDS-PAGE running buffer	25 mM Tris 192 mM Glycine 0.1% SDS
Fairbanks solution A	0.05% Coomassie Blue 25% Isopropanol 10% Acetic acid
Fairbanks solution B	0.005% Coomassie Blue 10% Isopropanol 10% Acetic acid
Fairbanks solution C	0.002% Coomassie Blue 10% Acetic acid
Fairbanks solution D	10% Acetic acid